Post-print of Book Professional Quality Measurement Method and Book Personalized Recommendation Service Method in Library Borrowing Scenarios

Authors: Li Shuqing, Zhuang Guangguang, Qin Jiahang, Xu Xia

Date: 2023-08-26T00:00:00+00:00

Abstract

[Purpose/Significance] Grounded in existing library borrowing records and through in-depth mining of book reading relevance, this study explores effective methods for identifying the professional quality of books in borrowing contexts and implementing corresponding personalized book recommendation services. [Method/Process] Leveraging book reading relevance, the study proposes a book relevance linkage relationship and identifies professional book resources through an iterative identification algorithm for book quality. Simultaneously, utilizing book category relevance linkage relationships, it proposes a representation method for personalized user patterns, and presents design principles and implementation methods for personalized book recommendation strategies from two perspectives: long-term interest recommendation and short-term interest instant recommendation. [Results/Conclusion] Regarding book quality identification, this method more readily identifies high-quality professional book resources, demonstrates considerable flexibility in application, and can also conduct professional book identification within constrained book collections. Regarding personalized book recommendation, the findings indicate that regardless of long-term or short-term interest recommendation methods, the average recommendation hit rate of type-II users surpasses that of type-I users. Among type-I users, the average recommendation hit rate of short-term interest recommendation methods in both the highest similarity interval (above 75%) and the lower similarity interval (15%-50%) exceeds that of long-term interest recommendation methods. This study identifies professional books through reader borrowing sequence analysis and implements corresponding personalized book recommendation methods, which is conducive to improving existing library borrowing service levels and enhancing reader satisfaction.

Full Text

Preamble

ChinaXiv Collaborative Journal
Volume 62, Issue 11, June 2018

Methods for Measuring the Professional Quality of Books and Personalized Book Recommendation Services in Borrowing Scenarios

Li Shuqing[1], Zhuang Guangguang[1], Qin Jiahang[2], Xu Xia[3]
[1]School of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210046
[2]Library, Nanjing University of Finance and Economics, Nanjing 210046
[3]School of Management, Nanjing University of Posts and Telecommunications, Nanjing 210046

Abstract

[**Purpose/Significance**] Based on existing library borrowing records and in-depth mining of book reading relevance, this paper explores effective methods for identifying the professional quality of books in borrowing scenarios and implementing corresponding personalized book recommendation services. [**Method/Process**] We propose a book relevance linking relationship based on reading relevance, combined with an iterative identification algorithm for book quality to identify professional book resources. Simultaneously, utilizing book category relevance linking relationships, we present a method for expressing reader user personalization patterns and provide design principles and implementation methods for personalized book recommendation strategies from both long-term interest recommendation and short-term instant interest recommendation perspectives. [**Result/Conclusion**] In terms of book quality identification, this method more easily identifies high-quality professional book resources with strong specialization, offers flexible applicability, and can also conduct professional book identification within limited book collections. Regarding personalized book recommendation, we find that regardless of whether using long-term or short-term interest recommendation methods, the average recommendation hit rate for the second type of users is higher than that for the first type. Among the first type of users, the short-term interest recommendation method demonstrates a higher average hit rate than the long-term interest method in both the highest similarity range (above 75%) and lower similarity ranges (15%-50%). This study identifies professional books through reader borrowing sequence analysis and implements corresponding personalized recommendation methods, which can help improve existing library borrowing service levels and enhance reader satisfaction.

Keywords: personalized recommendation; book borrowing; library services; book quality

1. Introduction

Personalized recommendation services are currently most prominently applied in the e-commerce domain, and due to their demonstrated success, they have gradually been adopted in many other fields. However, unlike e-commerce, library personalized recommendation services can leverage reader behavior information with distinct characteristics. For instance, the scale and variety of data reflecting reader user interest levels are relatively limited. Library borrowing activities require physical presence, resulting in comparatively constrained data generation. According to our analysis of borrowing statistics from Nanjing University of Finance and Economics between 2011 and 2014 (a four-year period), the average annual borrowing volume was approximately 30,000 instances, with an average of about 11 books per person per borrowing session. Within this relatively limited dataset, information analogous to product ratings from e-commerce website users—which is crucial for classic collaborative filtering recommendation methods—is absent.

Nevertheless, from another perspective, due to the inherent patterns of reading content and knowledge acquisition, the same reader often borrows the same book or the same category of books repeatedly, sometimes even forming a relatively stable borrowing trajectory. This provides additional research angles through book categories and temporal relationships. Consequently, the relevant design concepts and implementation methods require specialized investigation.

From the recommendation process perspective, achieving better service effectiveness necessitates in-depth research on two critical aspects: (1) The books recommended to current readers should avoid the Matthew effect. While general book recommendations often employ popularity metrics such as bestseller lists and borrowing volume rankings, borrowing scenarios feature diverse reader needs and frequent demands for high-quality professional resources. Therefore, designing effective methods to identify quality professional books constitutes key content for improving library personalized recommendation service workflows. (2) Effective expression of current reader user interest characteristics can be obtained by analyzing readers' existing borrowing histories, and experiments demonstrate that further combining borrowing histories with reader relationships can provide more effective expression approaches [?]. These two aspects respectively constitute the two main research components of this paper.

2. Literature Review

Library personalized recommendation services can significantly improve reader user experience and enhance the comprehensive utilization rate of existing library collections [?]. Numerous implementation methods exist, and with deepening research on Web 2.0 and media audience influence, leveraging greater audience participation in information product creation, dissemination, and sharing—and indirectly discovering information resource value and identifying quality resources through audience behavior—has become a highly promising research approach. Based on existing book borrowing records, effective book quality analysis methods and recommendation service strategies can be developed, with related research gradually increasing in recent years.

Regarding book quality measurement research, high-quality professional books can be identified through various standards, including expert evaluation, reader evaluation, or other indirect methods. For library borrowing services, the determination of high-quality professional books must comprehensively consider both content and readers. The strong professional characteristics of high-quality professional book content often result in limited attention from the general borrowing population and difficulty gaining recognition from the overall reader community. Conversely, books that receive widespread attention are often popular due to diverse reading preferences rather than high-quality professional content. Book quality and professionalism cannot be directly reflected through borrowing volume alone. Therefore, employing indirect methods that comprehensively consider both aspects can provide more comprehensive identification approaches for high-quality professional books.

Some scholars have utilized collaborative filtering methods based on user evaluation and rating to propose measurement methods for professional book quality and implement personalized book recommendation services accordingly [?]. Others have pointed out the drawbacks of using borrowing frequency and average borrowing time as traditional indicators—for instance, high borrowing frequency of a single book does not justify purchasing multiple copies, books have timeliness constraints, and manual methods like expert recommendations suffer from subjectivity and limited scalability. However, the specific book quality evaluation methods proposed only employ the book N-index method based on book categories without achieving quality assessment for individual books [?]. Alternative approaches have used metrics such as average borrowing time per loan, borrowing frequency, and whether a book is new to design evaluation methods measuring popularity and indirectly expressing quality [?].

More research utilizing borrowing records ultimately aims to serve effective recommendation method design. Some scholars employ big data resources and association rule analysis to discover reader interest patterns from borrowing records, implementing personalized library services through improved frequent pattern growth algorithms [?]. Others have addressed the cold-start problem in digital library personalized recommendation by adapting e-commerce cold-start handling methods, using improved K-medoids algorithms to cluster existing readers and books based on borrowing-return time intervals, thereby achieving personalized services for new readers and new books [?].

Notably, these methods heavily depend on effective processing of reader borrowing records and accurate understanding of related reader borrowing behavior. Current personalized book recommendation research methods include association rule methods [?], topic model methods [?], etc. Traditional book recommendation algorithms using only borrowers' historical borrowing data typically result in low recommendation accuracy. Therefore, comprehensive methods combining book similarity and borrower similarity [?], and multi-feature methods incorporating book classification [?] have achieved good results in book recommendation research.

Regarding reader borrowing behavior itself, some scholars categorize reader borrowing into four types: renewal, overdue borrowing (long-term and short-term), normal borrowing, and blind borrowing, calculating relative borrowing time and arguing that blind borrowing and long-term overdue borrowing cannot effectively reflect reader interest [?]. Others classify borrowing behavior into three types: borrowing, renewal, and reservation, believing that borrowing time correlates with reader interest and introducing time decay strategies to improve analysis methods [?].

In recent years, increasing numbers of scholars have focused on analyzing the network structure formed by borrowing information, providing a new research starting point for related book recommendation services. Some have conducted comprehensive analyses from a network structure perspective, noting that borrowing networks exhibit relatively high average clustering coefficients, small average shortest path lengths, and obvious complex network characteristics [?]. Others have utilized borrowing networks to derive readers' co-borrowing relationships, adding user personal attributes and book classifications to the network to further study borrowing connection strengths between different user categories and book categories, proposing many targeted personalized library borrowing service measures [?].

Among these research methods, the bipartite network structure analysis method is relatively common [?]. In our preliminary experiments, we used reader borrowing behavior characteristics to form the basis for judging book recommendability quality, designed an iterative algorithm for measuring book recommendability quality based on the borrowing bipartite network structure, and proposed three forms of personalized book recommendation services: specific topic recommendations, modified recommendations for currently borrowed books, and new book recommendations [?]. Similar research includes using energy transfer six-step algorithms to repeatedly diffuse weights in the borrowing bipartite network structure, obtaining recommendation lists composed of books with different weights to achieve personalized book borrowing push services [?]. Further research has combined web-scraped book purchase records with university library borrowing records, using user book ratings as weights in the borrowing bipartite network graph, and comprehensively predicting preferences for the same book ratings and borrowing preferences based on differences in user rating preferences to complete personalized book recommendation services [?].

From the perspective of recommendation strategy innovation, using temporal and evolutionary information to enhance the identification of user personalization interest patterns and the expression of user interest characteristic information constitutes a promising research direction in personalized recommendation services [?]. However, research on personalized book recommendation incorporating time factors remains limited. Some scholars propose calculating users' short-term demand preferences using only the most recent semester's borrowing records and long-term demand preferences using users' entire borrowing records [?]. Our preliminary research has gradually recognized the importance of temporal information analysis, achieving some initial results, such as using weighted interest expression methods to propose weighted keyword co-occurrence temporal elements, studying personalized academic research temporal path discovery methods and their visualization interface design through keyword temporal path discovery and keyword temporal network structure expression [?]. This paper continues this exploration and strives to achieve comparative analysis and characteristic analysis of long-term interest recommendation methods and short-term interest recommendation methods.

From the practical application perspective of domestic university libraries, the popularization of personalized book recommendation services remains limited, with related web applications and services being quite inadequate. Except for a few universities such as Nanjing University, few university library systems provide such personalized book recommendation services, and even when recommendation services are provided, they are not personalized [?]. This fully demonstrates the necessity of related research. This paper begins with the analysis of reader borrowing record bipartite network structures to complete the identification of high-quality professional books and the measurement of reader user similarity, thereby exploring a new method for personalized professional book recommendation services.

3. Method for Identifying High-Quality Professional Books

3.1 Basic Conceptual Framework

Reader borrowing records can reflect readers' own reading intentions and attention levels. Professional readers typically better understand the quality of borrowed professional books, and borrowing records themselves demonstrate borrowers' recognition of related books. Generally, the more excellent the professional reader, the more likely they are to borrow higher-quality professional books. Due to professional differences, library measurement of high-quality books must not ignore the recognition level of professional readers. In continuous borrowing records, readers' attention to a series of borrowed books can be indicated, and the correlation between books can also be reflected.

However, measuring the quality of professional readers and distinguishing their excellence is not easy, as pure book borrowing records provide limited direct analysis basis. More commonly, analyzing the connections between borrowed books themselves through book relevance can measure book associations and distinguish book quality. Therefore, this paper proposes the following basic approach: first, define book reading relevance, then construct book relevance linking relationships, and finally use an iterative algorithm to identify high-quality professional books.

3.2 Book Reading Relevance

Many methods exist to obtain reader borrowing sequences, with the key being how to extract required temporal information. Different extraction methods yield different sequence generation approaches. Borrowing time and return time are the most fundamental temporal information. Typically, readers complete concentrated borrowing of multiple books in one borrowing operation and concentrated returning of multiple books in one return operation, forming a borrowing sequence with interleaved borrowing and returning orders.

Based on current analysis, we propose the following hypothesis: For all readers, if book A is frequently borrowed after book B and before book B is returned, this indicates that books A and B have certain reading relevance.

This reading relevance has multiple possible semantic interpretations—it may indicate obvious content correlation between books, extended reading expansion relationships, or other unknown reasons. However, due to high frequency of occurrence, it can largely provide an effective approach for measuring book relevance. Even when the underlying reasons cannot be determined, it offers a data resource worth deep mining. For further analysis, quantitative measurement methods for this reading relevance need to be defined.

Borrowing duration is relatively intuitive data content. However, in experiments, we found that simply using days or other borrowing duration units for measurement often introduces many adverse effects. The main reasons include vacation impacts (as shown in Table 1, obvious summer vacation impact characteristics exist) and personal reading habit influences—many readers have longer average borrowing durations, which does not indicate interest in these books and sometimes even reflects lower borrowing activity. Therefore, borrowing duration is not considered in our measurement method design.

Borrowing sequence number refers to the unique identifier for each borrowing operation in all borrowing records of a user. In the user's first borrowing operation, all borrowing records are assigned borrowing sequence number 1, incrementing with each subsequent borrowing. Since readers may borrow multiple books at once, some borrowing records contain the same borrowing sequence number corresponding to multiple borrowed books. This indicator can express the borrowing order of different books.

3.3 Construction Method for Book Relevance Linking Relationships

Based on the definition of book reading relevance, we extract required book relevance links from reader borrowing sequences. As shown in Table 1, for the same reader's borrowing records, according to the increasing relationship of borrowing sequence numbers and the inclusion relationship of borrowing times, we can obtain book relevance linking relationships as shown in Figure 1 [Figure 1: see original paper].

From Figure 1, we can see that before returning book 118808, this reader subsequently borrowed four books: 153604, 49490, 160691, and 147269, thus establishing directed connections between relevant nodes. Book 153604 only had one subsequent borrowed book (49490) before being returned, therefore only one directed connection exists.

Processing all readers' borrowing records in the same manner yields a typical network structure for this book relevance linking graph, which from experimental analysis exhibits typical complex network characteristics. Notably, many generation methods exist for book relevance linking graphs—traditional methods often construct them based on co-borrowing relationships. The method proposed in this paper avoids the characteristic of traditional methods where books with high borrowing volumes often have high in-degree nodes. Instead, only books with more book reading relevance can form high in-degree or out-degree nodes, clearly providing a novel basis for measuring book quality and user similarity.

From the perspective of reader user behavior, we believe that in continuous borrowing behavior, as readers continuously understand the content of borrowed books, they are more likely to borrow more professionally relevant book resources in subsequent borrowing actions. Therefore, high-quality books measured by this method are often more professional in content, thereby providing good recommendation object resources for personalized recommendation services.

3.4 Iterative Identification Algorithm for Book Professional Quality

Iterative calculation methods in complex network structures can yield converged node information, such as PageRank methods [?]. These methods primarily operate in network structures composed of numerous nodes by assigning initial weights to each node, then using weight expansion methods based on link-out relationships through multiple iterative calculations to finally obtain stable node weights, thereby achieving effective measurement of network node quality.

However, these traditional methods do not consider characteristics of specific application domains in node weight allocation and weight diffusion strategy selection. As previously noted, book reading relevance is established on the important premise of certain frequency of occurrence—whether for individual readers or all readers, frequently occurring book relevance linking relationships better indicate reading relevance between linked book nodes. According to this design principle, necessary modifications to traditional network node iterative algorithms are required.

Based on the standard PageRank method, this paper proposes the algorithmic approach shown in Equation (1):

$$\text{WeightRPR}(Book_i) = c \sum_{DOC_{ij} \in F(Book_i)} \frac{\text{WeightRPR}(Book_{ij}) \times count_{ij}}{NF(Book_i)} + (1-c)$$

Equation (1)

WeightRPR represents the book weight based on the modified PageRank method. The set $F(Book_i)$ represents the collection of all in-link books for book $Book_i$ in the book relevance linking relationship graph, and $NF(Book_i)$ represents the number of books in this in-link collection. Three points require explanation:

(1) $Count_{i,j}$ represents the quantity of linking relationships between corresponding books $Book_i$ and $Book_j$. This coefficient amplifies the influence of each in-link book's weight, thereby reflecting that frequently occurring book relevance linking relationships better indicate reading relevance between linked book nodes.

(2) In actual calculations, $Book_i$ and $Book_j$ are not necessarily different books. On the contrary, the behavior of continuously borrowing the same book better reflects that book's quality, and retaining such linking relationships helps discover high-quality books.

(3) Equation (1) provides a basic calculation method. In practical applications, more adjustments can be made according to application environments and requirements, such as replacing books with book categories to mine the attention levels of different book categories among designated reader groups, or using this method within specific reader groups to hierarchically understand the concerns and most highly regarded high-quality professional books of different reader groups.

4. Personalized Professional Book Recommendation Method Based on Reader Borrowing Sequence Analysis

4.1 Basic Conceptual Framework

Reader borrowing sequences can also provide analytical conditions for personalized book recommendation services, enabling the formation of various methods to express user personalization patterns and related user similarity determination methods, such as directly using borrowed books as processing units. However, these methods often suffer from certain sparsity and expression precision issues.

To answer whether users who have borrowed the same books are interest-similar users, we can first address a related question: Will interest-similar users borrow the same books? Several common scenarios can be analyzed:

(1) In library borrowing systems, many books have different versions, and often the same version has multiple different IDs. Even for the same book, borrowing volume limitations prevent all interested readers from borrowing it, so readers often must choose similar books or other versions.

(2) Borrowing behavior motivations are complex, and the degree to which

borrowed books reflect reader interest is not directly corresponding. In many cases, readers exhibit tentative or incidental borrowing behavior. Even for professionally related book borrowing, understanding differences exist, and different readers have different evaluation standards for the same book, with varying abilities of borrowed books to reflect readers' own interests.

From the above analysis, directly using borrowed identical books as the basis for judging user interest similarity has inherent limitations. Based on the borrowing order existing in reader borrowing sequences, improvements can be made from two aspects:

(1) Using book categories as processing units can better balance accurate reader interest restoration and avoiding missed reports caused by overly detailed expression. The specific method can utilize classification numbers from the Chinese Library Classification (CLC) annotated for each book. Since these classification numbers have multi-level hierarchies, our actual experiments adopted a processing strategy that only retains the English letter prefix and the last two digits of the classification number—for example, retaining "F75" for "F752.68/27" and "TP39" for "TP391.13/24".

(2) Based on the aforementioned book relevance linking relationships, mapping book linking relationships to corresponding book category mapping relationships, then using book category relevance linking relationships as interest expression units, as shown in Table 2 , where higher frequency indicates greater effectiveness. This information also provides quantitative data foundations for subsequent user similarity calculations.

4.2 Similarity Calculation and Recommendation Method Design

Using all book category relevance links as vector units and frequency information as vector unit values, we can obtain interest feature vectors for each reader user. For each reader user i, we can obtain the reader user interest pattern shown in Equation (2):

$$ReaderVector_i = \{(\text{Book Category Relevance Link 1}, \text{Frequency 1}), (\text{Book Category Relevance Link 2}, \text{Freque$$

where m is the total number of book category relevance links.

Specific reader user similarity calculation methods can employ Pearson coefficients or cosine similarity coefficients, ultimately obtaining similarity between each user and other related users. Due to the large number of readers and book category relevance links, in actual calculations, the comparison scope can be limited by setting a threshold for the number of identical book category relevance links shared by reader users.

This paper first normalizes the weights of each reader user interest pattern. To avoid influence from absolute frequency numbers of individual users, we adopt a weight normalization method that divides each frequency value in the vector by the user's maximum frequency. Then, we use the cosine similarity coefficient between every two reader user interest pattern vectors to obtain final user similarity:

$$sim(ReaderVector_i, ReaderVector_j) = \frac{ReaderVector_i \cdot ReaderVector_j}{|ReaderVector_i||ReaderVector_j|}$$

In the personalized recommendation stage, for the target reader user, we first obtain the sequence of most similar other reader users. In actual calculations, similarity thresholds can be set to control the size of this sequence. We designed two personalized book recommendation service forms with different service objectives:

(1) **Long-term interest recommendation**: Based on all borrowing situations of the target user, obtain relevant borrowed book category information, then go to the most similar other reader user sequence to aggregate and obtain a recommended book list, and output the recommended book list in reverse order according to the aforementioned high-quality book identification standards. This recommendation form primarily addresses long-term interest characteristics of reader users, with recommended content having certain stability and user relevance.

(2) **Short-term instant interest recommendation**: Based on the most recent n borrowing situations of the target reader user (where n can be selected according to experimental data conditions, such as 2 or 3 times), obtain relevant borrowed book category information, then go to the most similar other reader user sequence to aggregate and obtain a recommended book list based on the most recent borrowing situation, and output the instant recommended book list in reverse order according to the aforementioned high-quality book identification standards. This recommendation form primarily addresses short-term interest characteristics of reader users, with recommended content having strong timeliness.

5. Experimental Results

5.1 Experimental Environment Preparation

We used nearly four years of book borrowing records from the Huiwen borrowing system of Nanjing University of Finance and Economics Library from January 1, 2011, to June 16, 2014, as experimental data. The effective borrowing record data volume obtained was 1,076,749 [Figure 1076749: see original paper] entries, involving 138,696 types of books, each with a unique book ID, and 42,750 readers, including school faculty and nearly four cohorts of undergraduate and graduate students.

For comparison experiments, we retained all readers' most recent borrowing content as comparative data, which was not used in the high-quality book identification experiment or user similarity calculation experiment. These borrowing records all lacked return dates, totaling 13,791 entries, accounting for 1.28% of total borrowing records.

5.2 Professional Book Quality Identification Experimental Results

According to the book relevance linking relationship graph construction method described earlier, we extracted 2,595,690 records from the experimental dataset, involving a total of 126,033 books, accounting for 90.87% of the total book collection. The main reason is that some readers borrowed books very infrequently and with long intervals, and such books were also rarely borrowed by other readers, meaning not all books could form in-link or out-link nodes for other books in the book relevance linking relationship graph.

For convenience in displaying identification effectiveness, we provide comparative displays of search results. The comparison object is the common output result based on descending borrowing volume ranking in existing library borrowing systems, as shown in Tables 3 through 5 :

Table 3 shows the comparison results for information retrieval books (CLC classification number G25). **Table 4** shows the comparison results for management principle books (CLC classification number C93). **Table 5** shows the comparison results for data mining books (with "data mining" in the title).

From Tables 3-5, several characteristics of the proposed method can be observed:

(1) In book quality identification, the method focuses on mining professional book resources with strong specialization, which is a primary advantage. In fact, among high-quality books identified by traditional descending borrowing volume methods, exam-oriented and test-preparation books often achieve high borrowing volumes in university libraries. For example, in the "information retrieval" category, the top-ranked book by borrowing volume is a case analysis book, which does not appear in the professional book catalog identified by the book relevance linking relationship graph iterative algorithm. In the "management principles" category, three books related to study notes rank second, third, and fifth by borrowing volume, but only one appears in the professional book catalog identified by our algorithm, ranking fourth.

(2) For quantitative comparison evaluation, we used the Chinese Dangdang website's positive rating indicators. Following the NDCG (Normalized Discounted Cumulative Gain) indicator's basic principle that users generally click on top search results first, we introduced discounting factors and calculated final scoring values for the top five search results:

$$discountingScore = \sum_i Score_i \cdot \frac{\log(2)}{\log(1+i)}$$

The comparison calculations for the three results from Tables 3-5 are shown in Table 6 .

We subsequently completed random testing across 20 professional domains. The average discountingScore of our algorithm was 2.7362, while the average discountingScore of the descending borrowing volume method was 2.37569174, representing an overall improvement of 20.38%. The results show the method is applicable across various professional book domains without obvious professional differences, only affected by the number of book query results—some professional books have few readers, making effective identification based on user behavior difficult.

(3) The method offers flexible applicability, enabling analysis to meet specific needs within all queried book ranges. Tables 3 and 4 show query results under dual constraints of keyword queries and CLC classification number limitations, while Tables 5 and 6 show only keyword query results. Different query strategies were adopted mainly because both book titles and classification numbers have limitations in representing book content.

5.3 Reader Similarity Experimental Results

Due to limitations in borrowing record content, not all readers share identical borrowed book categories with other readers. Therefore, in user similarity results obtained through book category relevance links, the actual number of effective readers (with similarity greater than 0%) was 23,937 [Figure 23937: see original paper], accounting for 56% of all reader users.

Experimental results indicate that user similarity values have large variation space, ranging from 100% to 0.000006%. The specific value distribution is shown in Table 7 .

The total number of matching user pairs is 422,984, with nearly 77.13% of user pairs having similarity below 10%. Among the remaining 23% of user pairs, the total number of involved reader users is 14,669, accounting for 61.3% of the actual effective reader total (23,937).

5.4 Personalized Recommendation Experimental Results

These experiments verify the effectiveness of the personalized recommendation method itself, including two parts testing long-term interest recommendation methods and short-term instant interest recommendation methods.

The test reader user objects selected for the experiments are mainly divided into two categories: (1) Users with more than 200 borrowing records, matched with similar reader users each having at least 10 borrowing records, resulting in 139 test reader users and 25,220 matching similar reader users, generating 3,505,580 user pairs. (2) Users with borrowing volumes between 100 and 200, matched with similar reader users each having at least 10 borrowing records, resulting in 1,213 test reader users and 25,220 matching similar reader users, generating 3,505,580 user pairs.

Since we retained all readers' most recent borrowing content as comparative data, we treat borrowed book categories in all user borrowing behaviors as in-links of book category relevance relationships. For long-term interest recommendation methods, the record scope is each reader user's complete borrowing records. For short-term interest recommendation methods, only the most recent 3 borrowing records of each user are used. We then observe recommended book category results based on these in-links and, according to actual users' subsequent borrowing book categories, calculate the proportion of book categories recommended by the method that appear in actual subsequent borrowing behavior—that is, the recommendation hit rate—to measure the effectiveness of the personalized recommendation method itself:

$$\text{Hit Rate} = \frac{\text{Number of users with recommended book categories appearing in subsequent borrowing}}{\text{Total number of users}}$$

Table 8 shows the specific effects of the long-term interest recommendation method for the first category of users (borrowing volume > 200). **Table 9** shows the specific effects of the short-term instant interest recommendation method for the first category. **Table 10** shows the long-term interest recommendation effects for the second category (borrowing volume 100-200). **Table 11** shows the short-term instant interest recommendation effects for the second category.

The overall comparison is shown in Figure 2 [Figure 2: see original paper].

From Figure 2, three characteristics can be observed:

(1) Regardless of long-term or short-term interest recommendation methods, the average recommendation hit rate for the second category of users is higher than for the first category. The main reason is that the total number of test users in the second category is higher (8.73 times that of the first category), thus having a greater probability of finding similar users and obtaining more effective recommendation results. However, in the similarity range of [90%, 100%], the average hit rate for the first category is higher than for the second category, indicating that for users with higher borrowing volumes, using higher user similarity for recommendations has advantages.

(2) Overall, among the first category of users, the short-term interest recommendation method's average hit rate is higher than the long-term interest method in both the highest similarity range (above 75%) and lower similarity ranges (15%-50%). Among the second category of users, the

short-term interest recommendation method's average hit rate is higher than the long-term interest method in most similarity ranges (above 20%), which also represents the range with the highest overall average hit rate. This demonstrates that the short-term interest recommendation method has obvious advantages, showing relatively high average hit rates across larger similarity ranges, fully indicating that reader users' reading interests have obvious timeliness and short-term characteristics, especially that interest changes in the recent period have significant impact on current user interests.

(3) Both long-term and short-term interest recommendation methods show a dependency relationship between recommendation hit rate and user similarity threshold range—relaxing the user similarity matching range can increase the number of recommended users but adversely affects recommendation hit rate.

For effectiveness comparison, we tested the standard collaborative filtering method on the above two categories of users for short-term instant interest recommendation. The results are shown in Table 12 .

As shown in Table 12, both methods have unsatisfactory overall recommendation hit situations, with the second category of users (with smaller borrowing volumes) showing even worse recommendation hit effects. This indicates that traditional standard recommendation methods, without considering user interest evolution trends reflected in borrowing history and limited by data volume, do not achieve ideal results with pure similarity fitting methods.

It should be noted that current experiments only target existing reader borrowing records and have not yet considered integration with high-quality book resource recommendations. We have developed an online user satisfaction testing system, inviting users to search for books of interest, and have begun user testing of actual recommendation effects in the Nanjing University of Finance and Economics Library borrowing system. This work requires time, and we will provide more detailed analysis and explanation of related research results upon completion.

6. Conclusion

As an exploration of personalized recommendation service research in the book borrowing domain, this paper strives to use existing reader borrowing records to achieve high-quality professional book resource identification and reader user similarity comparison methods, propose strategies for improving existing library personalized services, and conduct practice to initially achieve expected design goals. However, the method still has areas requiring further research and improvement. The main issue is that personalized recommendations based on reader borrowing records currently require relatively large-scale reader user borrowing records and demand that readers have relatively stable long-term borrowing habits. Conversely, for occasional borrowing reader users, it is difficult to obtain more interest features from borrowing records, though this also indicates that for these users, how to stimulate their attention to book borrowing is particularly important for improving library book borrowing services. This constitutes our next research objective.

References

[1] WU F, HU Y, WANG P. Developing a novel recommender network-based ranking mechanism for library book acquisition[J]. Electronic library, 2017, 35(1): 50-68.

[2] OUNI A, KULA R, KESSENTINI M, et al. Search-based software library recommendation using multi-objective optimization[J]. Information & software technology, 2017, 83(3): 55-75.

[3] MATHEW P, KURIAKOSE B, HEGDE V. Book recommendation system through content based and collaborative filtering method[C]//Proceedings of international conference on data mining and advanced computing. Ernakulam, India: IEEE, 2016: 47-52.

[4] Xu Bin. Research on the h-index of library books[J]. Journal of the China Society for Scientific and Technical Information, 2014(8): 892-896.

[5] Li Kechao, Liang Zhengyou. Personalized book recommendation algorithm based on multiple features[J]. Computer Engineering, 2012, 38(11): 34-37.

[6] PING H. The research on personalized recommendation algorithm of library based on big data and association rules[J]. Open cybernetics & systems journal, 2015, 9(1): 2554-2558.

[7] Ling Xiaoe, Zhou Bing, Li Kechao. Research on cold start problem in digital library personalized recommendation for new readers and new books[J]. Information Studies: Theory & Application, 2014, 37(8): 100-104.

[8] Liu Dan. Deploying personalized book recommendation service using Apache Mahout[J]. New Technology of Library and Information Service, 2015, 31(10): 102-108.

[9] HE B, ZHANG H. Library personalized information recommendation of big data[C]//Proceedings of online analysis and computing science. Chongqing, China: IEEE, 2016: 289-292.

[10] Zheng Xiangyun, Chen Zhigang, Huang Rui, et al. Personalized book recommendation algorithm based on topic model[J]. Journal of Computer Applications, 2015, 35(9): 2569-2573.

[11] ZHU W. Research on user similarity calculation methods in recommender systems[D]. Chongqing: Chongqing University, 2014.

[12] Ma Jian, Du Zeyu, Li Shuqing. Library personalized book recommendation method based on multi-interest feature analysis[J]. New Technology of Library

and Information Service, 2012, 28(6): 1-8.

[13] Li Kechao, Lan Dongmei, Ling Xiaoe. Research on uncertainty book recommendation for university readers' borrowing preferences using cloud model and multi-features[J]. New Technology of Library and Information Service, 2013, 29(5): 54-58.

[14] Jing Minchang, Yu Yinghui. Collaborative book recommendation model and application based on borrowing time scoring[J]. Library and Information Service, 2012, 56(3): 117-120.

[15] Jiang Zhoufeng, E Haihong, Yang Jun. Borrowing frequency scoring model and application based on temporal context information[J]. Library and Information Service, 2014, 58(s2): 220-223.

[16] Wang Jinliang, Zhang Peng, Di Zengru, et al. Network analysis of the book borrowing system of Beijing Normal University[J]. Journal of the China Society for Scientific and Technical Information, 2009, 28(1): 137-141.

[17] Fu Linhua, Guo Jianfeng, Zhu Jianyang. Library book borrowing system and single-scale binary network model[J]. Journal of the China Society for Scientific and Technical Information, 2004, 23(5): 571-575.

[18] Yan Fei, Zhang Ming, Sun Tao, et al. Analysis of user book borrowing behavior based on network characteristics: A case study of Peking University Library[J]. Journal of the China Society for Scientific and Technical Information, 2011, 30(8): 875-882.

[19] ZHAO S, ZHAO Y, SUN F, et al. Study on single mode weighted network of library lending network[J]. Journal of residualss science & technology, 2016, 13(6): 241-246.

[20] Li Shuqing, Xu Xia, Xu Minjia. Book recommendability quality measurement method and personalized book recommendation service based on reader-book bipartite network[J]. Journal of Library Science in China, 2013, 39(3): 83-95.

[21] Yuan Husheng, Zhao Xichen. Personalized recommendation algorithm and implementation based on weighted borrowing network[J]. Library and Information Service, 2016, 60(10): 130-134.

[22] Lan Dongmei. Diversified recommendation of bipartite graphs with multiple datasets under large-volume books[J]. Information Studies: Theory & Application, 2016, 39(2): 69-72.

[23] SUN Y, YIN H, REN X. Recommendation in context-rich environment: an information network analysis approach[C]//Proceedings of the 26th international conference on World Wide Web companion. Perth, Australia: International WWW conferences steering committee, 2017: 941-945.

[24] Qiu Junping, Zhang Cong. Research on collaborative recommendation system for university library collection resources[J]. Library and Information Ser-

vice, 2013, 57(22): 132-137.

[25] Li Shuqing, Sun Ying. Personalized academic research temporal path discovery and its visualization method based on weighted keyword co-occurrence temporal elements[J]. Journal of the China Society for Scientific and Technical Information, 2014, 33(1): 55-67.

[26] PAGE L. The PageRank citation ranking: bringing order to the web[J]. Stanford Digital Libraries Working Paper, 1998, 9(1): 1-14.

Author Contributions:
Li Shuqing: Proposed research ideas, conducted paper writing, revision, and finalization.
Zhuang Guangguang: Designed experimental algorithms and conducted related data analysis.
Qin Jiahang: Responsible for original borrowing record information organization and related data processing.
Xu Xia: Responsible for evaluation and analysis of recommendation experimental results.

Note: Figure translations are in progress. See original paper for figures.