International Journal of Library and Information Services
Volume 9 • Issue 2 • July-December 2020

The Collection and Service Optimization of China's Academic Library ILL Based on Bipartite Matching:

A Case Study of Soochow University

Yue Ma, School of Information Management, Central China Normal University, Wuhan, China

https://orcid.org/0000-0002-0875-2405

Jingxian Han, Soochow University Library, Suzhou, China

Zhuozhuo Li, Department of Archive and E-Government, School of Social Science, Soochow University, Suzhou, China

ABSTRACT

The interlibrary loan (ILL) service is essential to a multi-campus university library. This article builds a dynamic collection optimization model based on readers' needs using ILL data. This article first examined the status quo and methods of ILL. Then, refers to some existing algorithms of bipartite matching and recommendation such as network-based inference (NBI). Based on the above analysis, this article builds a model to optimize university library collection. With the loaning data of Soochow Academic Library from 2013 as the training data, an optimizing case was performed. The 2469 interlibrary loan records including readers and book they borrow from other library in 2014 was used for testing the effects of the model. The model is able to identify potential ILL demands and improve the efficiency of book circulation.

KEYWORDS

Academic Libraries, Bipartite Matching, Collection Management, Collections Optimizing, Interlibrary, Loaning

INTRODUCTION

Interlibrary loan is viewed as a service platform that combines libraries in different geographical locations. Readers of member libraries all have access to all of the resources in the library network. Access channels are either on site or by telephone and internet application services. Self-service book borrowing and returning can be done in the nearest library, and door-to-door book service is available as well (Li, 2016).

China experienced a tide of universities merging in 1990s and a large-scale construction of university towns in the early 21st century. To optimize the resource allocation of university libraries and facilitate proximity utilization for readers, branch building, resource reorganization, multi-point service and interlibrary loan have become part of the new normal state. Not only facing university libraries, interlibrary loan also plays an important role in urban library service. In particular, with the improvement of library sharing systems (such as one card solution), the standardization of public culture services and the promotion of a nationwide reading movement in recent years, interlibrary loan has become one of the basic public culture services. Meanwhile it is also the premise of the

DOI: 10.4018/IJLIS.2020070101

This article, originally published under IGI Global's copyright on April 3, 2020 will proceed with publication as an Open Access article starting on January 20, 2021 in the gold Open Access journal, International Journal of Library and Information Services (converted to gold Open Access January 1, 2021), and will be distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

formation of a library resources and service network. In this case, university library interlibrary loan services should attach more importance to integrating dispersed teaching and research resources and services according to the subject setting of each campus. Additionally, the adjustment of collection structure and layout is also in urgent need, so as to satisfy relatively stable knowledge preferences in different campuses. Moreover, book circulation networks between central libraries and service nodes has been further promoted with the prevalence of community library, subway library and self-check. Data generated from this interlibrary circulation network provides with a dynamic collection optimization scheme. This study is to analyze the interlibrary loan data of campuses in Soochow University and to find out and adapt to the diverse needs of readers, so as to achieve an accurate prediction of readers' needs and a dynamic collection structure system. As a result, time and space costs due to collection division by campuses will be minimized, and circulation efficiency and service quality will be improved.

LITERATURE REVIEW

Interlibrary Lending Service (ILL)

In 1894 Rowell initiated U.C. Berkeley's first program of interlibrary lending, with the California State Library as its partner. Later that year Rowell expanded the invitation to a group of libraries, named NUCMC (National Union Catalog of Manuscript Collections), which can be seen as the first practice of interlibrary loan (Wikipedia). In 1994, the Reference and User Services Association (RUSA) of the American Library Association (ALA) led the development of the "Interlibrary Loan Guidelines" (ALA, 2008) which has become a de facto standard for libraries at all levels after several revisions. It states: "the needs of each reader should be satisfied as soon as possible," and "keep communication between libraries to adjust the collection in advance according to potential needs." These principles indicate a focus on convenience and friendliness of specific service. Seal points out that university academic libraries also undertake the educational mission of providing assistance to students' growth and scientific research (Seal, 2016). Sharing of open access, online resources and foreign language resources is a significant area of interlibrary loan service as well (Celikbas, 2016; McGrath, 2016). With the gradual enrichment of information type and content, content building and distribution of collection resources display more and more importance nowadays.

Three main perspectives are proposed in terms of interlibrary loan service efficiency. The first is from the perspective of collection circulation, measuring collection rate and other related indicators. In 1995 F.W. Lancaster put forward that interlibrary loan efficiency is measured mainly through collection structure analysis, circulation analysis and availability (Lancaster, 1995). ISO11620 "Information and literature - library performance evaluation indicators" views speed and success rate (percentage of successful interlibrary loans) as main indicators of efficiency evaluation (British Standards, 2014). Additionally, indicators like coverage rate and circulation time (Stein, 2001) are also included. As referred to in "Interlibrary Loan Guidelines" (ALA, 2011), coverage rate is the satisfaction rate of readers' interlibrary loan requests, while circulation time refers to the time by which readers' demands are satisfied, representing service speed. Some studies have examined request source and response situation on the basis of borrowing request analysis in order to better assess service status (Benn, 2009). Ochola introduces the concept of overuse to discover and optimize collection categories that do not meet reader demands (Ochola, 2002). The second is from the reader's perspective; that is, directly investigate readers' satisfaction towards interlibrary loan (Stevens, 1974). For instance, applying a field method to the service satisfaction evaluation (Van Den Hoogen & Parrott, 2012), or identifying loyal users through user portraits to provide reference for service optimization (Mangiaracina, Cocever, Chiandoni, & Arabito, 2014). A survey regarding a library network in Hong Kong found that readers were sensitive to the type of resources borrowed; there was a shortage of some types of resources, while the unpopular resources remained in deep freeze

(Sidorko, Wong, Tai, & Wong, 2006). The third is to study the overall interlibrary loan process from a macroscopic perspective, such as a system that cares about time and economic cost, the establishment of TDABC (Time-Driven Activity-Based Costing) model, impacting factor analysis concerning the service quality of interlibrary loan (Pernot, Roodhooft, & Van den Abbeele, 2007), or an integrated model explaining resource procurement flow (Lindy, Oktaviani, Willy, & Prasetio, 2015). The aim is to achieve an integration of online browsing, discovery and offline access of resources. Interlibrary loan service studies that emphasize practice put great consideration into "effective input" (Leon & Kress, 2012). The best library services are supposed to satisfy readers as much as possible without interlibrary loan service. A better interlibrary loan service is supposed to communicate needs and deliver books in advance and cooperate with recommending and purchasing systems to maximize resource utilization efficiency, eventually attaining effective utilization of library expenditure (Austin Booth & O'Brien, 2011). Reflections on interlibrary loan service emerge with the convenient and widespread use of digital resource. Although digital resources cannot completely replace paper resources, more attention should be paid to digital resources utilization instead of inefficient paper resources services (Pedersen, 2010). In this case, it is worth learning from internet product services to enhance the flexibility and intelligence of library services (*Posner, 2007*).

The study of interlibrary loan service efficiency provides references for dynamic optimization of library collections. On the one hand, the inherent logic of the interlibrary loan service development requires a mature resource construction and distribution model; the library mission requires that interlibrary loan service should supply information resources to users more accurately. The actual situation requires that libraries must make full use of limited expenditure, avoiding worthless effort and input. The existing interlibrary loan model lacks flexible communication between library branches and has difficulty applying to practical work. Therefore, it is necessary to coordinate the resource allocation of branches. On the other hand, existing empirical research on interlibrary loan service mainly focuses on simple statistical methods. It was found that resource construction of ILL service cannot meet readers' demands in regard to circulation and category, while specific countermeasures for collection optimization are insufficient. The extensive applications of big data exploit new methods for dynamic collection optimization in order to achieve accurate collection adjustment. However, in terms of the correlation between ILL efficiency and collection optimization, there is a still lack of effective ways and utilization modes to predict readers' demands and achieve dynamic adjustment of collection structures.

Based on the status quo of university libraries, collecting and analyzing ILL data, and using it to explore a method or mode to optimize the collections of multi-campus libraries, they are finally to solve the problem of "what to deliver" and enhance the "Sensitivity" and "Responsiveness" of library collection. The final purpose is to satisfy readers' real needs and to find out readers' potential needs. Positive changes include transforming the mode from traditional "passive" library collection and circulation into "active" prediction and demand guidance, along with improving circulation efficiency, reducing circulation cost and activating of collection resources.

Bipartite Matching and ILL Collection Optimization

A graph, made up of a set of points and edges, is an abstract way of expressing a set of objects and the relationships between objects (Easley & Kleinberg, 2010). A bipartite graph is a tool to express two different types of objects and their relationships, which can be seen as a mapping of realistic relation (Wasserman & Faust, 1994). For example, relationships between books and readers, goods and consumers are suitable for using a bipartite graph to present. Bipartite Matching matches the elements of the two sets according to a certain relationship. In the borrowing process, it is for readers to find desired books. In other words, a bipartite graph reveals the correlation of "finding books for readers and finding readers for books" in libraries. With the rapid change of readers' needs, library resources development plans and collection systems need to be adapted and adjusted. Li Yang discussed book distribution issues in ILL by using a logistics center location model but did not take into consideration users' need for books (Li, 2010).

In recent years, the application of social network methods in the field of library circulation has become more and more mature. Social network methods can roundly describe the operation status of ILL service, and find information, such as subgroup and centrality, that is not easily revealed by using original linear methods. The purpose is to make full use of reader behaviour data, providing new methods and ideas for collection optimization. In studies on ILL, C. Glover, et al. regard library locations as network nodes, building a resource network based on the consideration of free flow of library resources (Glover & Langstaff, 2006). Social network analysis has also been applied in studies on ILL circulation and construction of librarian consortiums, identifying important branches and subjects with high ILL rates (Jalalimanesh & Yaghoubi, 2013).

In the course of resource distribution, the essence is a match between readers and books as the homogeneity and correlation of readers' needs. Correspondingly, the ultimate goal of recommendation system is to allocate resources to readers in need, consistent with the essential requirements of ILL service (Pernot, et al., 2007). Besides, recommendation systems can capture the tendency of readers' need with good flexibility thanks to readers' feedback. Therefore, it is likely to achieve more efficient collection optimization by introducing a recommendation system to dynamically adjust library resources.

METHODOLOGY

At present, there are four main approaches for recommendation systems: collaborative filtering based on similarity between readers, a global ranking method, a content-based filtering method, and a network-based inference (NBI) method (Bobadilla, Ortega, Hernando, & Gutiérrez, 2013; KG & Sadasivam; Liu, 2016). Existing studies on book-loaning matches mostly practice collaborative filtering methods and global ranking methods (Fei-juan, MIAO, XU, & Bi, 2015; LAN, 2016; N.-n. Li & Zhang, 2009). This study, mainly uses network-based inference method to repeatedly improve the weight of links between readers and books, and achieves a relatively stable state by iteration, and finally describes the ultimate weight of resource allocation (Zhou, Ren, Medo, & Zhang, 2007).

The core of the NBI algorithm is to use a bipartite projection to project one class of nodes onto another class. For example, the homogenous degree of readers' book loaning behavior can be presented as a single-mode network between books, which means two books that are usually borrowed at the same time show a stronger link in the single-mode network. Distinguished from simple single-mode networks, edge weights between nodes are allocated by an iteration of common loaning relationships between readers (Zhou et al., 2007). The loaning relationship is binary in the original NBI algorithm, that is, only consider whether to borrow (0 or 1) or use specific values to describe loaning relationships (Liu, 2016; ZHANG & JIANG, 2012). The resource allocation matrixes made up of single-mode relationships reflect the degree to which one resource is recommended for another. Combined with previous loaning habits, readers' potential demands can be targeted.

For university library systems adopting authorization loaning mode, although the system control center can uniformly allocate resources, resources are dispersed in different branches. Readers are welcome to borrow books from any one of the branches in the library system. The operation mode is: reader proposes a request to certain branch (requesting library); then the request library sends a request to the branch where the book is collected (requested library); and the book is shipped to the requesting library at a regular time. Likewise, when returning, readers return books to requesting library, and then to original requested library after sorting and delivery. Logistical pressure can be reduced on the condition that delivery and allocation are done in advance.

For ILL in university libraries, readers' demands are of strong sensitivity and inclination. Students' loaning demands are often closely related to a specialty (research and teaching books), accompanied with stable and periodic demands (such as high and stable demand for novels, and a periodicallychanging demand for reference book for exams). Therefore, the focus of collection allocation is the structure readers' demands for book content.

Using existing ILL data, first create a model to calculate the allocation of books for various disciplines. Then readjust the collection structure of branches according to the inclination of readers' demands. Finally allocate books beforehand to branches most likely in need, so as to increase logistics and distribution efficiency. What is mentioned above is a theoretical collection optimization program, independent from implementation methods of distribution, location, time cost, etc. Here are the following assumptions:

H1: Readers from the same school generally share similar demands for books.
H2: Inclination of readers' demands can be extrapolated from previous loaning records.
H3: Readers with a tendency towards high demand should have priority to utilize books.

In empirical research, R language is used for data processing and calculating, while igraph and ggplot2 packages are used for data visualization. First, build a correlation matrix between readers and books, and modify the matrix according to the number of readers and books. Second, project book nodes to the bipartite graph shown in the matrix; that is, calculate the resource allocation matrix to obtain the degree of mutual recommendation among books. Third, in terms of each reader's loaning inclination, based on the resource allocation matrix, calculate the demand inclination of each reader for each kind of book. Fourth, consider the actual collection, count the average of readers' demands to obtain the resource allocation matrix between books and readers. The whole process is shown in Figure 1.

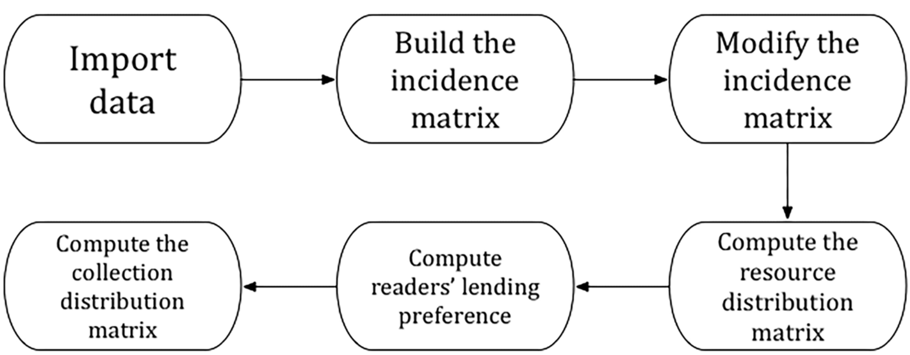
Build and Modify the Incidence Matrix

In the bipartite graph of loaning relationship, the set (nodes) of readers (schools) is given as $R = \left\{r_1, r_2, .., r_m\right\}$; the set (nodes) of books is given as $B = \left\{b_1, b_2, \ldots, b_n\right\}$, and the set of loaning relationship (edge set) between R and B is given as $L = \left\{l_{11}, l_{12}, \ldots, l_{mn}\right\}$:

$$l_{ij} = \begin{cases} k, k \ of \ b_j \ are \ borrowed \ by \ r_i \\ 0, no \ b_j \ is \ borrowed \ by \ r_i \end{cases} \tag{1}$$

This contributes to a bipartite graph, represented as the correlation matrix $L_{m \times n}$.

Figure 1. Process of building the collections optimization model



In this study, the match is between reader set R and book set B. For any reader r, the purpose of the network inference algorithm is to sort b according to the preference degree of r. The higher ranking the book *b* is the more likely it is to produce a demand for resources that should be assigned to specific r with priority. B's with higher rankings are the resource with higher demand from the r, and should be allocated to the r with priority. Once resource b is chosen by r, the algorithm can also recommend more items in set b to r. This two-way interaction allows the prediction of potential demands. In other words, books borrowed by more active readers are of higher value, while readers borrowing more valuable books are more active. By giving priority to active readers in term of delivering the valuable books, basically the efficiency of ILL work can be improved.

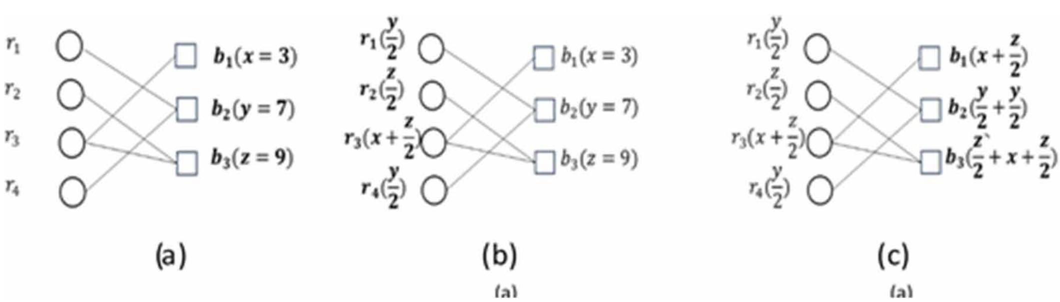
Specifically, in Figure 2, the total frequency of b_1, b_2, b_3 being borrowed are the initial resources x, y, z. Then b_1, b_2, b_3 share resources equally to related readers *r*. In this process, readers who borrow more valuable books tend to obtain more resources (which means its lending capacity for one specific kind of books raises), as shown in (2-b). Similarly, readers r share their resources equally to adjoining book nodes b_1, b_2, b_3, and books borrowed by active readers are of higher value. This improved method determines the value of x, y, z ($f\left(b_j\right)$, as shown in Figure (2-c)). The amount of resources reflect the value of b_1, b_2, b_3, providing an edge weight reference for single-mode network construction. As for the calculation process, the first step is to assign activity values of each reader according to book loaning situation:

$$f(r_i) = \sum_{j=1}^{n} \frac{l_{ij} f\left(b_j\right)}{k\left(b_j\right)} \tag{2}$$

As shown in Equation (1), l_{ji} is the frequency of book b_j being borrowed by reader r_i, while $k\left(b_j\right)$ is the total frequency of b_j being borrowed. This step calculates the preliminary resource of readers. In the next step, resources are reallocated to book b_j, and the value of $f(b_j)$ will also be updated:

$$f'\left(b_j\right) = \frac{\sum_{i=1}^{m} l_{ji} f\left(r_i\right)}{k\left(r_i\right)} = \sum_{i=1}^{m} \frac{l_{ji}}{k\left(r_i\right)} \sum_{j=1}^{n} \frac{l_{ji} f\left(b_j\right)}{k\left(b_j\right)} \tag{3}$$

Figure 2. Network based inference algorithm



Since readers and books are mappings of actual schools and book types, the matrix needs to be modified after importing the data. In other words, it ought to be taken into consideration the deviation of the number of students at each school and the number of books in certain categories when calculating. The corrected weight is calculated as:

$$c_i = \frac{v_i}{\overline{v}} \tag{4}$$

In Formula (4), C_i is the modified weight of entity *I*; v_i is the actual number of items in entity *I*; $\overline{v}$ is the average of the total item numbers of all entities in the set. This step makes sure there will not be too huge difference in number of C_i and be in the range of 0 to 1.

For instance, for school r_i, v_i is the number of students in each school, while for book b_j, v_j is the number of collections in each category. Once the set of modified weight for R and B have been obtained, calculate the collection number of all kinds of books from each campus as:

$$l_{ij}^{'} = l_{ij} \times c_i \times c_j \tag{5}$$

The modified result is a new matrix $L' = (l_{ij}^{'})_{m\times n}$. For simplicity, later mentions of l_{ij} in this paper are all elements in the modified set ***L’*** unless otherwise stated.

Compute the Resource Distribution Matrix

Transform the bipartite graph represented by the correlation matrix into a single-mode network between books, that is, a resource allocation matrix that predict the weight (in other words, lending priority) of reader r_i borrowing another book b_z based on previous borrowing record of book b_j:

$$w_{jz} = \frac{1}{d\left(b_j\right)} \sum_{i=1}^{m} \frac{l_{iz} l_{ij}}{d\left(r_i\right)} \tag{6}$$

In Formula (6), w_{jz} . is the predictive weight of b_j to b_z; $d\left(b_j\right)$ is the measurement of books b_j, that is, the sum of column b_j in set ***L’*** (how many times b_j has been borrowed). $d\left(r_i\right)$ is the measurement of reader r_i, that is, the sum of column r_i in set ***L’*** (how many books r_i has borrowed). As a result, $W = \left(w_{jz}\right)_{n\times n}$. Basically, this step projects the 2-mode graph to one-mode by multiplyinthe matrix. This is a key way of NBI.

Compute Readers’ Lending Preference

For each r_i, the inclination of borrowing book b_j is represented by the number of resources flowing back to the reader through the iteration, which is simplified by Formula (3), as shown below:

$$f\left(b_i\right) = \sum_{j=1}^{n} w_{ij} l_{jr} \tag{7}$$

Depending on previous reading preference, the algorithm recommends books that might possibly be needed to readers.

However, books are limited and the algorithm is supposed to give priority to readers with stronger demands. Therefore, readers' inclination needs to be averaged:

$$\delta_{rj} = f_r\left(b_j\right) \times \frac{1}{\sum_{i=1}^{m} f_i\left(b_j\right)} \tag{8}$$

δ_{rj} represents readers' demands for book b_j after average processing, so actually it represents the averaged priority value. The result is a matrix $\Delta = \left(\delta_{rj}\right)_{m\times n}$, that indicates the weight of any b_j being allocated to any reader *r* among all the collections.

Compute Collections Distribution Matrix

Let set $C_c = \left\{r_1, r_2, \ldots, r_r\right\}, \left(c = 1,2,3,4,5\right)$, represents the schools in each campus. Eventually, books of different categories for each campus are represented in matrix as below:

$$A_c = \left(\delta_{rj} \times v_j\right)_{m\times n} \left(r \in C_c, b = 1,2,\ldots,m\right) \tag{9}$$

In Formula (9), v_j is the participation of all redistributed books. A_c is the allocation program of campus *c*. The collection optimization program for branches is a correlation matrix that displays the number of bibliographies of each category obtained by each college.

EMPIRICAL RESEARCH AND RESULT

Research Object

Soochow University Library was chosen as the research object. They optimized collection according to practical situations, and checked the effect of the model with testing.

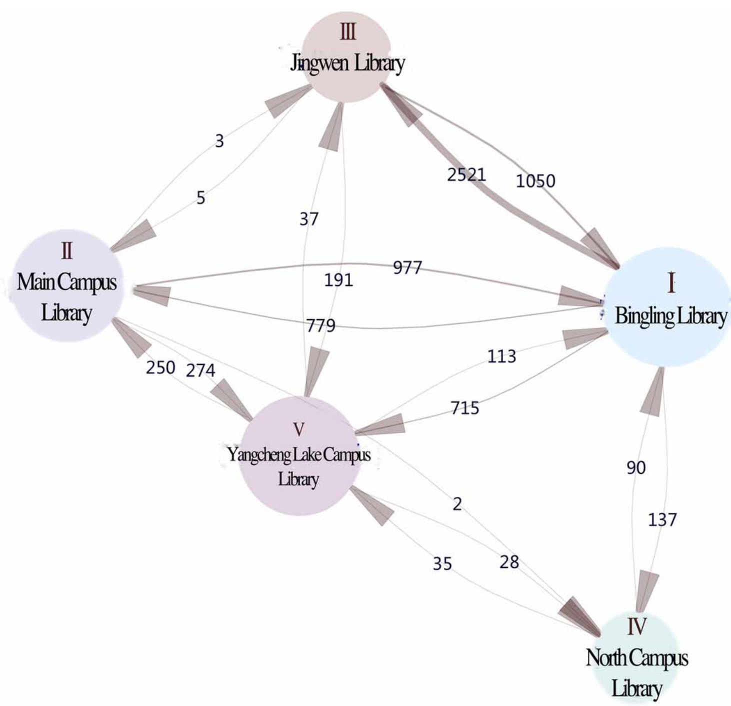
Soochow University is located in Suzhou, Jiangsu Province of China. It is a comprehensive university with 5 campuses and 29 first-order schools. Excluding C.W.Chu College and Tang Wenzhi College, there are 27 independent schools with recorded library data. Each campus has a library branch, including: I. Bingling library; II. Main campus library; III. Jingwen library; IV. North campus library; V. Yangcheng Lake campus library. The three branches: III, IV, and V are geographically closer together, but still in different campuses. Figure 3 depicts Location of Soochow academic library.

With the ILL service of Soochow University Library, readers obtain books from other branches through a branch agent (*Soochow-University-Library, 2016*), which is one of the most popular services. An igraph was used to virtualize ILL data of Soochow University Library from 2013-2016. In Figure 4, the starting point of the arrow is the collection library; the end is the authorized library and the figure is the number of books.

It can be concluded that in some branches, such as between I and III, book flow is more significant, indicating an existing collection structure that is not fully adapted to the readers' needs of that campus. Meanwhile the inflow of branch V is greater than its outflow due to the lack of collection resources themselves.

Figure 3. Location of Soochow Academic Library

Figure 4. ILL situation of Soochow University Library in 2013-2016



Data Overview

Soochow University Library has a total of 755,231 books in the 523 categories of "Chinese Library Classification" that is not including books under "Library Classification of Chinese Academy of Sciences" or other classifications. Reallocating existing collections to satisfy readers' demands that popular books be placed in branches as much as possible.

The training set included loaning data (including ILL data) from 15340 transactions at Soochow University Library in 2013. Each record consists of a requesting library, requested library, book classification number, readers' ID, etc. Some data samples are shown in Table 1.

Table 1. Soochow Academic Library inter library lending data in 2013(part)

Request Library	Reader's ID Number	Requested Library	Book Classification Number	Department of Readers
Main campus library warehouse	122*******	Bingling library warehouse	I267.1/1118	Electronics school
Bingling library warehouse	120*******	Main campus library warehouse	H316/122	School of Politics and Public Administration
Main campus library warehouse	130*******	Bingling library warehouse	H030/6	Foreign Language School

The overview of book loaning data in 2013, shows that readers' preferences for books were relatively concentrated (novels), while specific school (such as school of basic medicine) had a great demand for certain kinds of books (medical and health). This remains to be verified. The number of books under each category of the "Chinese Library Classification" was obtained through the OPAC. The number of students in each school was accessible though the yearbook of Soochow University in 2013, and some inaccessible data, was inferred from the official information release platform of Soochow University and the enrollment plan of that year.

Computing Collection Allocation Matrix

After importing the training set into R and removing useless fields, the resulting dataset consists of four fields: requesting library, requested library, department of the readers and the book classification number. Then, build a data dictionary, processing all the fields to correlate readers and their schools, books and their classification numbers. This model defines books with the same classification using a book classification number consisting of the same 2-digit number after a letter, which compresses the amount of information and reduces computing pressure, avoiding over-scattered predictions (Easley & Kleinberg, 2010; Jalalimanesh & Yaghoubi, 2013).

Depending on the incidence matrix between books and readers which is described above, modifications are made for the number of students in each school and the number of book classifications. The modified result is a 529×27 matrix with correction significance of 1. The resource allocation matrix is calculated by Formula (4) to correlate schools with affiliated branches. The resource allocation matrix A_c of the five campuses is obtained from the model, and can be used to reallocate loanable books by demand strength.

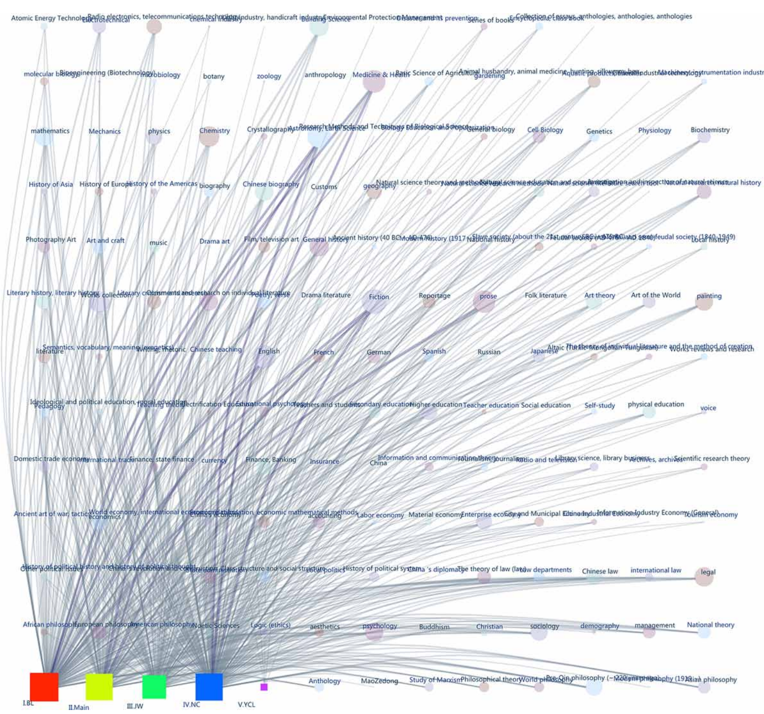
Collection Allocation Result

Five allocation matrixes are integrated into an allocation scheme for each campus. The matrix is shown in Table 2. Figure 5 is the Visualization of the allocation scheme.

Table 2. Collections distribution matrix (part)

	Anthology	Mao Zedong	Marxist Study	Philosophical Theory	World Philosophy	Pre-Qin Philosophy (~ 220 Years Ago)	Modern Philosophy (1919 ~)	Asian Philosophy
I. Bingling	8	0	7	8	7	78	0	18
II. Main	2	0	1	3	5	37	0	10
III. Jingwen	3	0	2	2	2	17	0	5
IV .North	8	1	5	8	7	49	2	14
V.YC Lake	0	0	0	0	0	0	0	0

Figure 5. Results of collections distributing



As it turns out, branch I possesses around 5,430 out of 15,340 books after reallocation. Book categories are diversified. Except for 747 novels, the number of books in fields like philosophy, education and literary criticism are all around 10 to 30, which corresponds to the subject setting of the campus. Branch II accounts for 3,253 books. Except for novels, specialized subjects include chemistry (109), law (162), mathematics (111) and architecture (65). Relevant books such as legal theory, physics and electric machinery also have their places. Branch III has 2,140 books. Based on Figure 5, the books come from widely dispersed categories, and there is no outstanding edge like other branches. Books on economics make up a certain proportion, but not a strong concentration. Branch IV has the most medical books due to the three main schools of the Medical College. Of the total 3,728 books, 366 books are medical books. While there are 234 and 317 books about earth science and essays respectively, branch IV is, in general, relatively centralized. Branch V is a bit special since it has only one school. Specialized books are abundant and lending out is not often demanded.

However, demand for novels and English learning books accounts for 234 of a total 789 books for allocation. In short, among these five branches, the hottest books are novels, essays, astronomy, earth science (including popular science) and English learning books. The specific allocation scheme of each branch has different characteristics according to the readers' discipline. It is indicated that this allocation scheme gives consideration to basic books with large demand (such as English learning, novel, essay) and specialized books (such as medicine and chemistry), while optimizing the circulation of unpopular books (such as semantics, vocabulary and basic science of agriculture). Libraries should increase purchasing for the hottest books, while emphasizing the adjustment of the existing collection for specialized books.

We can say that such an allocation matrix is a theoretical "should" plan. In practice, it can be combined with library workflow, and coordinating the collection based on the actual situation. For example, in a large-scale optimization of all the collection, more attention should be attached to books with high demand (novels), and in regard to the local optimization of the branches, special needs should be taken into consideration. To optimize branch IV, focus should be given to medical and health books that are urgently needed.

Collection Allocation Effect Evaluation

To test the effect of resource allocation, this paper imports the ILL data from 2014 as a test set; 2,469 ILL requests are included. Use the model to optimize the collection, then calculate the ratio of ILL request being satisfied within the branch. Expected satisfaction index for the demand is:

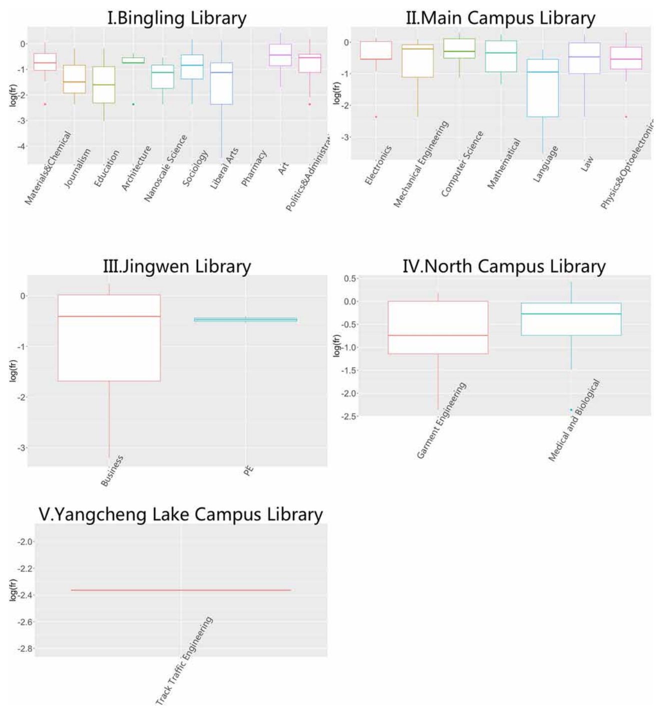
$$fr_b = \frac{\text{number of optimizated collections}}{\text{number of actual collections}} \quad (10)$$

It is worth noting that, since the optimized collection and the actual collection are measured in branches and schools respectively, collection redundancy exists when compared with actual demand. The redundancy index may be higher than 1, which indicates this branch is able to satisfy this school's demand for this kind of book. Simultaneous demands for a certain kind of book from multiple schools are not taken into consideration.

The ILL demand of all schools in the test set as fr_b, and was calculated and the distribution of fr was divided by branches. The result is displayed in a boxplot as shown in Figure 6 (names of schools have been simplified):

The length of the "line" indicates the number of total demands and the range of satisfaction rates for each school, while the "box" position indicates the school's average demand satisfaction rate. Since the fr value is logarithmic, if the data point's ordinate value is greater than 0, then $fr > 1$, meaning that demand can be satisfied by the branch collection. Taking the expected value, it can be inferred that 56.75% of the demands of all the schools can be satisfied within their affiliated branches after optimization. In other words, it shows the lending capacity of a specific school. The predictive capability of the model has been testified, but a relative surplus still exists. For instance, the School of Basic Medical and Life Science's demand satisfaction rate is much greater than other schools affiliated to the same branch. This may attribute to the strong specialization of the books it demands with "exclusivity". Branch I and II have a large number of readers from various schools. The demand satisfaction rate has a huge range, suggesting this model's prediction is unbalanced. By contrast, with limited resources, branch V can satisfy most of its requests. It can be inferred that unique needs (specialized books) and general needs (novels and other books) have different allocation effects. In other words, even if large amounts of resources are obtained in the training set, if the demand is biased towards books of a general need such as literature and English learning, the demand satisfaction rate may not be too high.

Figure 6. Expected satisfaction of each library



CONCLUSION AND DISCUSSION

After modeling the loaning data of Soochow University, the collection allocation scheme shows a certain degree of predictability in the test. In comparison with the existing ILL model, advanced resource allocation helps to decrease the pressure of ILL delivery service. A practical collection optimization plan was launched based on test result. The first thing to be noted is the allocation of books in general demand. With the most readers, branch I still cannot satisfy demand for books like novels, essays and English learning after optimization. Thus, branch I should be given priority to have these books. Since all the campuses welcomes best-selling books and leisure books, regular book exchange and circulation among branches is necessary. In the case of branch V, demand for reading is motivated by activities like reading appreciation and reading clubs, which in turns activate a demand for books with low circulation rate. This kind of book circulation and reallocation is a win-win strategy that improves book use efficiency and satisfies the demands of other branches at the same time. Second, collection structure of specialized books is discussed. For the schools of liberal arts, branch I has a low degree of demand satisfaction. Education, political theory, European history and other books are inadequate. With sufficient collection space, it is appropriate to increase the purchase of these kinds of books, to optimize the collection and its utilization efficiency. Last, but not least, is the demand for unpopular books. The demands of several schools of branch IV (such as textile engineering and metallurgical Engineering) and School of Physical Education of branch III, are small with a low satisfaction degree. This is because most of their books are unpopular and not collected in other branches. Collection sites for these unpopular books should be adjusted and centralized in school branches, so as to improve utilization efficiency.

This paper puts forward suggestions for collection allocation and adjustment that cater to users' demands. More practical issues concerning ILL circulation and delivery, bookshelf capacity and collection space should also be taken into consideration when launching a branch collection optimization scheme in university libraries. Collection optimization is a process of continuous improvement based on existing the collection system. Through two-way adjustment of centralized procurement and branch allocation, collection optimization aims to achieve a dynamic and intelligent collection system. In addition, collection construction can be combined with recommendation and other library services such as pushing information based on users' demand. It would be very useful to integrate the recommendations for book information and for the allocation of physical books, using real time correlation and data computing, to build a dynamic three-dimensional system for collection optimization and circulation services. The purpose of the system is to enhance data intelligence and service efficiency. It is worth noting that, this empirical research uses Soochow University Library as an example and focuses on the research route of collection optimization. When building the model, issues like compatibility, speed, security and so on, was not taken into account. In addition, this paper simplifies definitions for readers and books, which is likely to influence the accuracy of the results. The collection optimization mentioned in this paper mainly refers to demand predictions by data analysis. An operational plan concerning shelving books or delivery after prediction is not discussed. Moreover, the model doesn't identify the difference between different kinds of books. This is to say that in fact the reading preference may be useful to fiction readers, but not so as for books like subject books. This is one point which the model has oversimplified, one possible solution is to take more details into model so the demands on subject books can be processed well. The model which is used in the paper may only show the relative high demands on subject book, but fail to predict the accurate demands. Therefore, difficulties exist when the plan is put into practice.

In future research, more attention should be put on discovering the relationship between procurement and allocation, as well as a deeper investigation into improving user demand mining and responsiveness studies. Furthermore, whether or not there is a correlation between prediction effect and demand uniqueness is worth discussing further. University libraries mainly provide services for research and teaching. This research found that the distribution of specialized books corresponds with subject distribution in each campus. For cases where students with the same major are on different campuses, this research can also apply. The ILL optimization model, based on bipartite matching, is able to achieve prediction and pre-adjustment of collection resource allocation to some extent. Specifically, this model reaches maximum allocation efficiency by coordinating collection resources and readers' demands. Additionally, the logistics and distribution pressure of ILL is reduced, satisfying both efficiency and cost.

REFERENCES

ALA. (2008). Interlibrary Loan Code for the United States. Retrieved from http://www.ala.org/rusa/guidelines/interlibrary

ALA. (2011). Interlibrary Loan Practices Handbook (C. L. Weible & K. L. Janke Eds. the 3rd edition ed.). Chicago: ALA Editions.

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, *46*(Supplement C), 109–132. doi:10.1016/j.knosys.2013.03.012

British Standards (BSI). (2014). Information and documentation -Library performance indicators (Vol. DIN ISO 11620-2014).

Celikbas, S., & Ekingen Flores Mamondi, F. (2016). Library cooperation in Turkey: The results of a survey of ILL librarians in higher education. *Interlending & Document Supply*, *44*(2), 58–65. doi:10.1108/ILDS-01-2016-0002

Austin Booth, H., & O'Brien, K. (2011). Demand-driven cooperative collection development: Three case studies from the USA. *Interlending & Document Supply*, *39*(3), 148–155. doi:10.1108/02641611111164636

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets* (Vol. 1). Cambridge: Cambridge Univ Press. doi:10.1017/CBO9780511761942

Glover, C., & Langstaff, G. (2006). Floating Collections Within a Consortium: Or, Does It Matter Where the Materials Are Shelved? *PNLA Quarterly*, *71*(1), 12–21.

He, F.-J., Miao, X.-L., Xu, D.-W., Bi, P. (2015). Personalized recommendation of books based on "reader-book" bipartite graph. *Computer Technology And Development*, (5), 1-4.

Jalalimanesh, A., & Yaghoubi, S. M. (2013). Application of social network analysis in interlibrary loan services. *Webology*, *10*(1), 1–6.

Kg, S., & Sadasivam, G.S. (2015). A Survey on Personalized Recommendation Techniques. *International Journal of Emerging Trends & Technology in Computer Science*, *4*(1), 171–173.

Lan, D. (2016). Multi-data set of two diversified book recommendations in the context of big data. *Information Theory and Application*, *49*(02), 69–72.

Lancaster, F. W. (1995). The evaluation of library services: A concise review of the existing literature. *Investigación Bibliotecológica*, *9*(018), 25–37. doi:10.22201/iibi.0187358xp.1995.18.3829

Li, Y. (2010). Research on the Distribution Mode of Public Library [Master's dissertation]. Zhejiang University, Hangzhou.

Li, H. (2016). Service Innovation of Public Library with internet. *Economic Research Guide*, (25), 154-161.

Li, N. N., & Zhang, N. (2009, February). The bipartite network study of the library book lending system. *Proceedings of the International Conference on Complex Sciences* (pp. 773-782). Springer.

Lindy, C. D., Oktaviani, E., Willy, , & Prasetio, Y. L. (2015). Library System Development Implementing Integrated Book Circulation for Interlibrary Loan. *Procedia Computer Science*, *59*(Supplement C), 326–330. doi:10.1016/j.procs.2015.07.574

Liu, X. (2016). Bridge the gap between network-based inference method and global ranking method in personal recommendation. *Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA)* (pp. 1802-1806). IEEE Press. doi:10.1109/ICIEA.2016.7603879

Leon, L., & Kress, N. (2012). Looking at resource sharing costs. *Interlending & Document Supply*, *40*(2), 81–87. doi:10.1108/02641611211239542

Mangiaracina, S., Cocever, C., Chiandoni, M., & Arabito, S. (2014). Assessing the effectiveness of a national resource sharing system. Interlending & Document Supply, 42(2/3), 98-104. doi:10.1108/ILDS-12-2013-0039

McGrath, M. (2016). Interlending and document supply: A review of the recent literature; 93. *Interlending & Document Supply*, *44*(4), 178–181. doi:10.1108/ILDS-06-2016-0019

Ochola, J. N. (2002). Use of circulation statistics and interlibrary loan data in collection management. *Collection Management*, *27*(1), 1–14. doi:10.1300/J105v27n01_01

Pedersen, W. A. (2010). The Paper Divide. *The Serials Librarian*, *59*(3-4), 281–301. doi:10.1080/03615261003619078

Pernot, E., Roodhooft, F., & Van den Abbeele, A. (2007). Time-Driven Activity-Based Costing for Inter-Library Services: A Case Study in a University. *Journal of Academic Librarianship*, *33*(5), 551–560. doi:10.1016/j.acalib.2007.06.001

Posner, B. (2007). Library resource sharing in the early age of Google. *Library Philosophy and Practice*, (June), 1–10.

Seal, R. A. (2016). Resource sharing begins at home: Opportunities for library partnerships on a university campus. *Interlending & Document Supply*, *44*(3), 127–134. doi:10.1108/ILDS-01-2016-0003

Soochow-University-Library. (2016). Lending and returning books. Retrieved from http://library.suda.edu.cn/menu_detail.jsp?id=7&iType_id=2–5

Stein, J. (2001). Measuring the Performance of ILL and Document Supply: 1986 to 1998. *Performance Measurement and Metrics*, *2*(1), 11–72. doi:10.1108/EUM0000000007215

Stevens, R. E. (1974). A study of interlibrary loan. College &amp. *Research Libraries*, *35*(5), 336–343. doi:10.5860/crl_35_05_336

Sidorko, P., Wong, R., Tai, A., & Wong, E. (2006). User-initiated resource sharing in Hong Kong universities: Planning to reality with HKALL. *Interlending & Document Supply*, *34*(2), 60–69. doi:10.1108/02641610610669750

Van Den Hoogen, S., & Parrott, D. (2012). Communication, collaboration and cooperation: An evaluation of Nova Scotia's borrow anywhere, return anywhere (BARA) multi-type library initiative. *Journal of Academic Librarianship*, *38*(6), 321–325. doi:10.1016/j.acalib.2012.08.004

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge: Cambridge university press. doi:10.1017/CBO9780511815478

Benn, J. (2009). Who uses the interlibrary loan and document delivery service and what do they request? A case study at the University of Western Australia. *Interlending & Document Supply*, *37*(1), 41–45. doi:10.1108/02641610910938613

Wikipedia. (2017). Interlibrary loan. Retrieved from https://en.wikipedia.org/wiki/Interlibrary_loan

Zhang, X., & Jiang, S. (2012). Personalized Recommendation Algorithm Based on Weighted Bipartite Graph. Computer Application, (3), 654-657+678.

Zhou, T., Ren, J., Medo, M., & Zhang, Y.-C. (2007). Bipartite network projection and personal recommendation. *Phys. Rev. E*, *76*(4), 46–115. doi:10.1103/PhysRevE.76.046115 PMID:17995068

Yue Ma is a Master graduate student at the School of Information Management, Central China Normal University since September 2017. He received a bachelor's degree in the field of information resource management from Soochow University in June 2017. He has published articles about library and information service in other Chinese academic journals.

Jingxian Han has been an Associate research librarian since 2017. Her major is library science, and she has worked in academic libraries for more than ten years. She works in the circulation department in the library of Soochow University and is in charge of ILL. She has published more than ten articles in library and information science academic journals in China.

Zhuozhuo Li (corresponding author) is an associate professor at the Society School in Soochow University since 2012. She received her Ph.D. degree in the field of library and information science in Wuhan University of China in 2009. She is a visiting scholar at University of Illinois at Champaign Urbana (UIUC). She has published more than thirty articles and books in China. Her research interests include information resource sharing and data services.