A Large-Scale Collection Reselection Method Oriented by Library Space Rebuilding Based on the Statistics of Zero-Borrowing Books

Lei Huang(✉)

Library, Shanghai University of Electric Power, Shanghai 200090, China
huanglei@shiep.edu.cn

Abstract. Taking the relocation of Pudong branch Library of Shanghai University of Electric Power and comparing the difference between the goals of collection reselection oriented by space rebuilding and traditional collection weeding, this paper explores a large-scale collection reselection method based on the statistics of zero-borrowing books in recent *X* years, which takes "volume" as the unit, pointing out that the *X* value depends on the Library's reserved space and disciplines nature first, and discussing the using precautions of this method. The practice shows that the core collection selected by this method can meet the readers borrowing needs, ensure the reasonable discipline collection structure and the professional literature charactered by energy and electricity, which has some value for university libraries' large-scale collection adjustment oriented by space rebuilding in the digital transformation of libraries.

Keywords: Space rebuilding · Collection reselection · Zero-borrowing · Circulation statistics

1 Introduction

With the rapid development of the Internet technologies, significant changes have occurred in readers' reading behavior and learning habits: from paper reading to electronic reading, from individual learning to collaborative learning [1]. In the face of this change, university libraries in China are actively carrying out space rebuilding to meet the changing needs of readers [2, 3], such as Tongji University's continuous space rebuilding of its Sino-German Library [4], science and technology reading room [5], Hubei branch library [6]; and phased transformation of Library of Shanghai University of Finance and Economics [7]. The library of Shanghai Electric Power University has also actively responded to this trend. In the construction of its new library in Pudong Campus (after this referred to as the "new library"), it has fully considered the readers' demand for space and made a large-scale adjustment to the collection distribution: taking its Pudong branch as an example, it has a collection of more than 300000 volumes Chinese books (October 2017 data). We planned to put 150000 volumes Chinese books into the compact stack room on the first floor of the Pudong branch to release more space

H.-R. Ke et al. (Eds.): ICADL 2021, LNCS 13133, pp. 115–123, 2021.
https://doi.org/10.1007/978-3-030-91669-5_10

for collaborative learning and group discussion. Therefore, the Pudong branch library should select 150 thousand volumes as core collections from more than 30000 volumes Chinese books and allocate them to the "core stack of Pudong" when moving to the new library. The remaining 150 thousand Chinese books need to be drawn off and assigned to the "compact stack of Pudong".

For the collection construction of Pudong New Library of Shanghai Electric Power University, the selected collection is expected to achieve the following goals: 1) Meeting the readers borrowing needs basically; 2) Ensuring the reasonable distribution of disciplines in the collection structure of the new library; 3) Ensuring energy and power literatures as much as possible which are the key professional characters of the university; 4) The reselection method should be feasible and conducive to large-scale operations.

In this paper, we propose a large-scale collection reselection method and reach the above goals at the same time.

The rest of this paper is organized as follows: Sect. 2 surveys related work on collection reselection methods. Then, we introduce our method including the establishment of the conditions for selecting books and book extraction data generation in Sect. 3. In Sect. 4, we analyze whether the results can meet the expected goals followed by discussion in Sect. 5. A conclusion is drawn in Sect. 6.

2 Related Work

Through literature researching, it is found that most studies on the reselection of library collections are aimed at book weeding and collection optimizing [8], but not on the selection methods aiming at the rebuilding of digital library space. The emphases of the two are different because of their different goals. Book weeding is a process of weeding from the whole collection, and more focus on security [9]; The space rebuilding is the relayout of the collection, emphasizing efficiency. The reselection method to optimize the collection by weeding has high costs of time and labor, and the processes are complex [10], which can not meet the requirements of large-scale selection guided by space rebuilding in library digitalization.

Based on the practice of moving the Pudong branch of Shanghai Electric Power University Library to the new library in 2018, this paper attempts to achieve a method suitable for large-scale operation through the statistics of zero-borrowing books(volume) in recent X years and can meet the above objectives at the same time.

3 Condition Establishment and Data Generation

3.1 The Establishment of the Conditions for Selecting Books

Considering the predicted learning space and the future collection growth rate, the new library should have 150000 books estimatedly, not more to ensure students learning space and not less to meet students reading needs, and these 150000 books also should meet the four selected goals in the introduction at the same time. In practice, the library of Shanghai University of Electric Power has adopted the "reverse selection" operation to extend the opening hours of the library as long as possible, that is, the selected core

collections will continue to be on the shelves for readers to borrow, while the non-core collections will be removed from the shelves. Therefore, the extraction data generated in this paper is the "non-core collection" data, and the corresponding operation objects based on the extraction data are also the "non-core collection" books. Each library can flexibly select the operation objects according to the specific situation of the library.

Through a variety of experiments, it is found that the data based on the statistical method of zero-borrowing books (volume) in recent X years can better meet the four selection goals. In contrast, other methods, such as trying to take the total number of borrowing times or borrowing rate of a kind of book lower than a "certain value" as the condition of book extraction, have the following disadvantages: it doesn't consider the time factor - for example, a certain book used to be hot, but now no one is interested in it—although this kind of book borrowing times are high, it should be removed from the shelves; second, the "certain value" is difficult to be defined objectively. However, other more complex methods of book extraction by subject division are not feasible and difficult to large-scale operate. More importantly, after many experiments, the number of books extracted by other methods is far less than the target value of 150000.

To sum up, on the basis of comparing various methods, the conditions for book selection of the "non-core collection" are finally determined as follows: "zero-borrowing books(volume) in recent X years, excluding the books collected after the year of Y", and the schematic diagram is shown in Fig. 1.

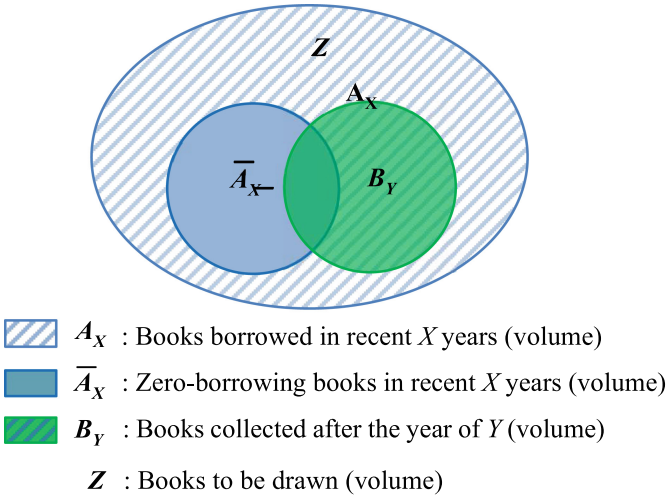

Fig. 1. Non-core collections (books to be drawn) diagram.

In Fig. 1, Z is the "non-core collection", that is, books to be drawn out, A_X is "zero-borrowing books in recent$X$ years", and B_Y is "books collected after the year of Y". As can be seen from Fig. 1, the formula of book selection can be written as follows:

$$Z = A_X - B_Y \tag{1}$$

The Formula1 shows that the books to be withdrawn should consider the condition of zero borrowing quantity as well as consider the factor of collection time. The "books collected after the year of Y" should be eliminated so as to ensure that the new books collected after the year of Y will not be withdrawn from the shelves, and the values of X and Y mainly depend on the predicted collection space of stack rooms and the distribution of the disciplines.

3.2 Book Extraction Data Generation

According to Formula1, the extraction data is generated as follows:

1) First, setting the book extraction conditions for the overall collections as "zero-borrowing books in recent 5 years, excluding the books collected after January 1st, 2013", namely, $Z_0 = A_5 - B_{2013}$. At this time, X is 5, Y is 2013, and 150000 books drawn list Z_0 are generated;
2) Second, optimizing some Classes with large quantity by CLC (Chinese Library Classification, CLC) according to the different disciplines. For example, in this university characterized by energy and power, the books of Class T are industrial and technical books accounting for a largest collections, its conditions(the values of X, Y) for drawing book should be different, set as "zero-borrowing books in recent 5 years, excluding the books collected after January 1st, 2016", that is, $Z_T = A_5 - B_{2016}$. At this time, X is 5, Y is 2016, and the book extraction list Z_T is generated. Besides, Class TP3 computer books in Class T are further optimized according to their subject half-life [11, 12] and the expected growth of this kind of book collection. The Class TP3 book extraction condition is set as "zero-borrowing books in recent 3 years, excluding the books collected after January 1st, 2016", that is, $Z_{TP3} = A_3 - B_{2016}$. At this time, X is 3, Y is 2016, and the book extraction list Z_{TP3} is generated. After Class T, Class O and Class H are further optimized too, which is not repeated here.

After the above operation, the book extraction list of 150000 "non-core collection" books is generated.

4 Results and Analysis

4.1 Analysis of Readers Borrowing Needs Guarantee

As mentioned above, the book extraction lists generated can meet the space requirements of the new library firstly, but whether it can meet the 4 selection goals in the introduction still needs further verification. Taking the book list Z_0 as an example, the condition of book selection is "zero-borrowing books in recent 5 years, excluding the books collected after January 1st, 2013". In other words, as long as the books borrowed within the past five years will be retained, and the books collected after January 1st, 2013 (inclusive) will also be retained on the shelves, that is, the book list itself is generated from the perspective of meeting the readers borrowing needs. From the standpoint of probability, books borrowed within recent 5 years are more likely to be borrowed again due to the

similar professional distribution and similar needs of students in the same school. In addition, all the newer books are reserved. Therefore, the selection goal of "Meeting the borrowing needs of readers basically" can be achieved by the book extraction lists.

4.2 Analysis of Disciplines Distribution Reasonability

Before verifying whether the book list can "guarantee the reasonable distribution of disciplines in the collection structure of the new library", many staff members expressed concern: according to experience, the most books borrowed by students are art novels, while the number of professional books seems to be small. If we use the book list as the condition of the new book collection, there will be a large number of professional books be taken off the shelves. However, a large number of literary novels are kept on shelves, which affects the disciplines distribution of the new library and leads to the unreasonable collection structure for our school characterized by energy and power. Therefore, before the implementation of book selection, we comparatively analyzed the distribution of different disciplines before and after book selection. The results are shown in Fig. 2.

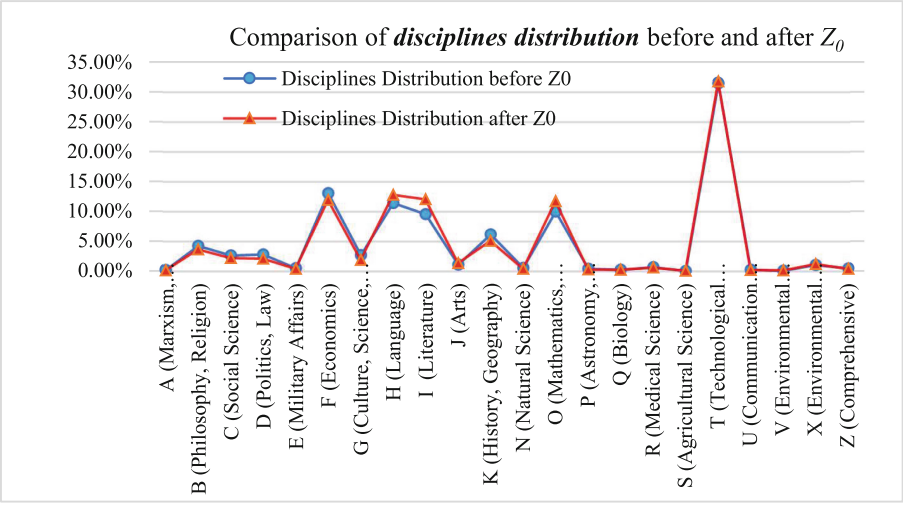

Fig. 2. Comparison of disciplines distribution before and after Z_0

Figure 2 shows that the distribution of all disciplines in the collection after the book extraction is basically the same as before. The major categories of books with a significant increase after the book extraction are Class H Language, I Literature, O Mathematical and Physical Sciences and Chemistry. This shows that the students have a high borrowing amount in these major categories of books because according to the book extraction condition Z_0, all the books borrowed in the past five years will be kept on the shelves. Therefore, which kind of books account for a high proportion after drawing books indicates that students have a high borrowing amount and a great demand for such books. However, the proportion of Class I literature and art novels increased the

most, which is in line with people's experience expectations. Still, the ratio of Class T industrial technology with school characteristics has not been weakened at all, which is quite different from the previous staff's worry or feeling based on experience alone, which also shows that the school students' demand for professional books is beyond expectations.

The distribution of disciplines in the new library is almost consistent with that in the original library: the difference between the distribution of disciplines in the new library and the original library is in 22 categories, 14 categories fluctuate between 0–0.5%, only 3 categories fluctuate between 0.5–1%, and 5 categories fluctuate between 1–2.5%. Assuming that the distribution of disciplines in the original collection is reasonable, it can show that the book extraction list Z_0 meets the target of "Ensuring the reasonable distribution of disciplines in the collection structure of the new library".

4.3 Analysis of Key Disciplines Guarantee

Shanghai University of Electric Power is a University with the characteristics of energy and electric power. This is not only reflected in the collection structure, that is, the absolute advantage of the collection proportion of the Class T industrial technology books, but also reflected in the distribution of the key disciplines that can best represent the characteristics of energy and electric power in Class T. To verify whether the book list can guarantee the energy and power characteristic key disciplines of the school, we take the original collection as the reference and introduce the retention rate η,

$$\eta = q_i(\text{volumes on shelves})/Q_i(\text{total volumes}) \tag{2}$$

Further subdividing the book extraction list Z_T, the results are shown in Table 1.

Table 1 presents that the retention rate of TK energy and power engineering and TL atomic energy technology, the two most representative disciplines with energy and power characteristics, is in the front row, and TK is in the first place. The retention rate of the books that are highly related to the school's energy and power characteristics, such as the mechanical industry, TM electrical technology, TQ chemical industry, and other books, is more than 50%. In comparison, the retention rate of TK energy and power engineering books is 70%, followed by TL atomic energy technology books, which is also the micro reflection of the school's increasing research on nuclear power engineering in terms of borrowing quantity. To sum up, the book list can meet the selection goal of "Ensuring the key characteristic professional literature of energy and power".

5 Discussion

5.1 The Book Extraction Data by the Unit of "volume" Can Automatically Generate the Number of Copies Reserved for Each Kind of Book

Considering the given selection goals, this paper has tried many other methods and finds that, besides the convenience of data processing, this method processes data by the unite of "volume" instead of "kind" has an important but not easily perceived advantage: the

Table 1. Detailed data in retention rate of Z_T.

Subclasses in class T	Total number of volumes	Volumes to be drawn	Volumes retained on shelves	Retention rate η
T	28	17	11	**39.29%**
TB	4556	2181	2375	**52.13%**
TD	70	32	38	**54.29%**
TE	155	82	73	**47.10%**
TF	118	70	48	**40.68%**
TG	3025	1521	1504	**49.72%**
TH	5312	2323	2989	**56.27%**
TJ	63	24	39	**61.90%**
TK	1459	439	1020	**69.91%**
TL	211	66	145	**68.72%**
TM	12826	5646	7180	**55.98%**
TN	19677	10514	9163	**46.57%**
TP	39715	19477	20238	**50.96%**
TQ	1649	666	983	**59.61%**
TS	1520	781	739	**48.62%**
TU	5380	2918	2462	**45.76%**
TV	215	152	63	**29.30%**
SUM	95979	46909	49070	**51.13%**

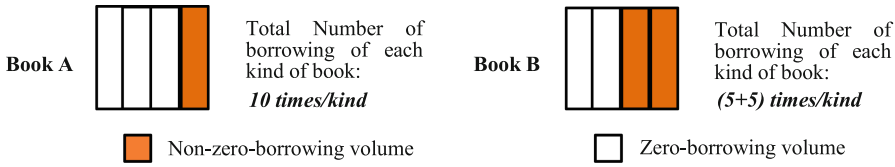

Fig. 3. Schematic diagram of retention of copies of different books.

number of copies of each kind of book will be automatically generated. As shown in Fig. 3.

As shown in Fig. 3, assuming that there are four copies of book A, should one or two be retained? If only one volume has been borrowed 10 times in the past five years, and the other three volumes have not been borrowed, then the three volumes will be withdrawn, and the only one volume will be retained. Another case: there are four copies of book B, two of which have been borrowed five times, while the other two volumes have not been borrowed, two volumes will be retained. Therefore, although the total borrowing times of the two kinds of books are the same, the probability of book A being borrowed only

by one reader is higher, while the probability of book B being borrowed by different readers is higher, for which it is speculated that the reason may be related to whether the different kinds of books are of professional concentration or universal dispersion by their reader groups. Therefore, even if the total borrowing times of the two kinds of books are equal, the number of copies retained could be different, which skillfully avoids the unreasonable "one size fits all" decision in the number of copies of different kinds of books.

5.2 The Book Extraction Data Should Be Verified by the Rationality of Collection Disciplines Structure

The actual verification of the book extraction data generated by this method shows that, except for Class A books of Marx, Lenin, Mao Zedong, and Deng Xiaoping monographs, we decided to keep almost all of them because of their essential political core status and made necessary manual intervention. For other categories, the core collection selected by this method can meet readers borrowing needs and ensure the reasonable structure of the subject collection and the characteristics of the school energy and power professional literature. Therefore, it can be inferred that the seemingly simple condition of zero-borrowing book (in volume) actually contains multiple laws, which reveals the objective relationship between the borrowing rate and the disciplines distribution and the guarantee of key disciplines from a simple perspective. However, with the rapid development of the Internet, the impact of electronic literature, the migration of students' reading habits, whether the borrowing rate can truly reflect the students' reading needs should be further paid close attention to. In other words, the collection data generated by the statistical method of zero-borrowing books (in volume) can only be put into practice after the rationality verification of the collection disciplines structure and manual spot check.

6 Conclusion

Library space rebuilding is an active measure taken by China university libraries to meet the changing needs of readers and the changes of readers' reading behavior and learning habits in the Internet era, which makes it inevitable to adjust the existing collection on a large scale in the digital transformation of libraries.

Taking the relocation of the Pudong branch of Shanghai Electric Power University Library, this paper proposed a large-scale collection selection method based on the statistics of zero-borrowing books (in volume) in recent$X$ years. It points out that the X value firstly depends on the predicted collection space of stack rooms, then the nature of the disciplines such as half-life period etc., and also discusses what should be paid attention to in using this method. The practice shows that the core collection selected by this method can meet the readers borrowing needs and ensure the reasonable disciplines distribution and the characteristic professional literature of energy and power.

It should be pointed out that, before the large-scale implementation of this method, it is necessary to verify the rationality of disciplines distribution of the core collection and manual sampling check. On the long run, multiple methods should be taken to avoid the potential risk of unbalancing core collection, such as searching books on APP, e-display

of non-core books information, assigning any returned books to the core collection including non-core books borrowed before, etc.

References

1. Su, L., Fang, F.: Hot spot and trend analysis of collaborative learning in China: a study of scientific knowledge map based on core journals. Adult Educ. **12**, 7–12 (2019)
2. Zhengqiang, L., Qiao, L.: A review of the research on library space reconstruction. Library (10), 55–59 (2018)
3. Yongjie, Z.: Review of the research on the space reconstruction and innovation of domestic library. J. Univ. Libr. Inf. **37**(3), 114–120 (2019)
4. Case introduction of German Library of Tongji University Library Homepage. https://www.lib.tongji.edu.cn/index.php?classid=11977&newsid=19664&t=show. Accessed 15 Jan 2021
5. Space transformation of science and technology reading room of Tongji University Library Homepage. https://www.lib.tongji.edu.cn/index.php?classid=11977&newsid=19668&t=show. Accessed 15 Jan 2021
6. Library space transformation (under construction) of Shanghai North Campus of Tongji University Library Homepage. https://www.lib.tongji.edu.cn/index.php?classid=11977&newsid=19669&t=show. Accessed 15 Jan 2021
7. Linlin, S., Li, Z.: Opening ceremony of academic database, academic exhibition, sharing space opening ceremony and Millennium Village Survey exhibition of Shanghai University of Finance and economics, and world top 500 enterprise exhibition held in the library. Res. Libr. Inf. Work Shanghai Univ. **29**(4), 100–101 (2019)
8. Na, X., Hong, Y., Limei, X.: Introduction and characteristics analysis of the old document removal project of Wesley University Library in USA. Libr. Inf. Work **59**(8), 38–41 (2015)
9. Yingchun, C.: The construction and empirical study of the collection reselection system. Libr. Constr. (3), 36–40 (2010)
10. Hailing, Y.: A new approach to the elimination of old books in library collection. Libr. Constr. (1), 66–72 (2015)
11. Liya, W.: Half life analysis of computer science based on CNKI. Libr. Inf. (1), 100–105 (2015)
12. Lijun, Z.: Statistics and analysis of zero lending books -- Taking the Chinese books in the branch library of South China Agricultural University as an example. Libr. Circles (3), 66–70 (2014)