Study on Library Books Lending Cooperative Competition Network

Wang Lijun1, Tian Junjie2, Sun Fei1, Zhao Sheng1,*

1. HeBei North University, Hebei, Zhangjiakou, China
2. Zhangjiakou radio & Television University, China

Abstract: By collecting the books lending information in HeBei North University within one year, we build the database of books and their borrowers and the bipartite graph of describing the relationship between the two things in the paper. According to the two kinds of situation, whether the lending relationship exists between books and readers or not, and whether the number of days of borrowing books by the readers exists between books and readers or not, we build unweighted network and weighted network to get the relevant statistical properties respectively. The study found that these statistical properties both obey exponential distribution and the corresponding statistical properties are related in unweighted network and weighted network. It shows that compared with unweighted network, the weighted network not only can describe the cooperative relationship and structure but also can accurately describe the competition result. What’s more, combined with the practice, we discuss the cooperative and competitive relationship within readers and books and between readers and books, and the practical significance owned by various statistical indicators in bipartite graph. Finally combined with the practice, we provide some suggestions for library purchase work.

Keywords: complex network; cooperative competition network; bipartite graph

1. Introduction

Books lending is one of the ways of offering services by the library. Its books lending amount can directly reflect the readers’ demand condition for books. And it is an important indicator of measuring books use efficiency. It is also the important reference factor of books procurement work. Seen from the system perspective, the library is a typical complex system. Books lending network is also a typical complex system. It is also a typical cooperative competition network. By such a kind of process for books lending, we build some kind of connection between books and readers so as to constitute the bipartite graph of library books lending network. Therefore, through the statistic analysis on the bipartite graph constituted by books lending situation, on the one hand, it provides more empirical basises for studying the weighted network, on the other hand it can help the procurement personnel to know the readers’ demands situation better, and grasp the satisfaction degree of readers’ demands from the existing books collected in the library so as to avoid the subjectivity and blindness of books lending. However, previous empirical studies on library books lending network are mostly unweighted network, and lack a certain rationality because the books don’t be classified. Compared with the previous studies on the library network, we have the following improvements in the paper.

(1). In previous studies, we use books bar code to represent books. For the books with the same version, its number of books collected in the library is more than one. If the readers borrow the books with the same version, because each book can use different bar codes to distinguish, the books with the same version may be replaced by different bar codes, then they can be processed into different books. In this paper, we do analysis and process on library bibliotheca database and books repositories to get the call numbers which books bar codes correspond to. Then we get the statistics of borrowing the books with the same version by the readers (the same call numbers), rather than the statistics of borrowing the same books (the same books bar codes). These make up for this shortcoming in the previous studies to a certain extent.

(2). Assume that the time of borrowing a book by the reader is one day or twenty days, then the books with different reading time have obvious different significances on this reader. Therefore, we only do statistics on how many times of books with the same version by the same reader to make it as the weight, so the significance isn’t too big. We do the statistical processing on the time record of books borrowing and returning in library computer system to get the days of borrowing books with the same version by each reader, and take this as the weight of borrowing this book by this reader. If the readers have borrowed the books with the same version for several times, then we add the reading time for each time as its weight.

Therefore, by a reasonable classification on books, we study the unweighted network and the weighted network of bipartite graph about library borrowing relationship so as to get some more accurate results. We put forward and analyze the cooperative competition relationship between books and readers in the paper. We hope that we can provide some new empirical basises for the researches in this aspect and we can provide some references for the development of the library.

2. Statistics amount on the weighted network

The weighted network can be described with the set \( G = (N, W) \) and it includes \( N \) nodes and a set of edges with the weights \( W \). Usually we can use the weighted adjacent matrix \( W \) to represent the weighted network. Among them, the matrix element \( W_{ij} \) represents the edge weight between the two adjacent points. Under normal circumstances, the similarity weight \( w_{ij} \in [0, \infty] \), if \( W_{ij} = 0 \), then it represents that there is no connection between two points. However, the different weight \( w_{ij} \in [0, \infty] \), when \( W_{ij} = \infty \), it represents that there is no connection between two points. When the values of each edge are the same, we can make them be normalized as 1. The weighted network is degraded into the unweighted network. That is to say, the unweighted network is the special case of the weighted network.

(1) Node Strength

In the weighted network, the natural generalization which corresponds to the node degree \( k_i \), is the node strength. It is defined as

\[
S_i = \sum_{j \in N_i} W_{ij}
\]  

(2.6)

Among them, \( N_i \) is neighboring set. Node strength both considers the neighboring number of the nodes, and the weight between nodes and neighboring nodes. It is the integrated embodiment of the node local information. When the edge weight has nothing to do with the topology of the network, function relationship between node strength and node degree is \( s(k) \propto k^\beta \). Among them, \( s(k) \) is the average value of edge weight. When the edge is associated with topological structure, the relationship of node strength and node degree is generally \( s(k) \propto A k^\rho \), or \( s(k) \propto w \), but \( A \propto w \), or \( \beta \neq 1 \).

Node strength distribution $p(k)$ is similar to the role of degree distribution $p(k)$, and it mainly investigates the probability of having node strength $s$ by the node.

2. Weighted Clustering Coefficient

The clustering coefficient reflects the group properties of this vertex between Level 1 neighbors. The closer contact between neighbors becomes, the higher of clustering coefficient of the node is. Based on the clustering coefficient of unweighted network, we develop the definition of clustering coefficient of weighted network in the paper. For example, the clustering coefficient of weighted network defined by Barrat and others is

$$c^w_{ij}(i) = \frac{1}{s_i(k_i - 1)} \sum_{jk} (w_{ij} + w_{ik} - w_{jk}) (2 - a_{ij} - a_{ik} - a_{jk})$$

(2.7)

Onnela and others have considered the geometric average value of the weight on three edges of triangles, and have defined the corresponding clustering coefficient of weighted network.

$$c^w_{ij}(i) = \frac{1}{k_i(k_i - 1)} \sum_{jk} (w_{ij}w_{ik}w_{jk})^{1/3}$$

(2.8)

Among them, $w_{ij}$ is the standardized numerical value of the maximum weight $\max w_{ij}$ through the network, but there are this or that kind of problems in the above definitions.

Petter Holme and others have analyzed the clustering coefficient of weighted network in details, and they have rewritten the form of clustering coefficient defined by Watts-Strogatz.

$$c(i) = \sum_{jk} a_{ij}a_{ik}a_{jk}$$

(2.9)

According to the above expression way, we consider the any edge’s contribution to the clustering coefficient in the triangle, we can write down the clustering coefficient of weighted network.

$$c^w_{ij}(i) = \max_{jk} w_{ij}w_{ik}w_{jk}$$

(2.10)

The clustering coefficient describes the group properties between the neighboring nodes. From the above requirements in article (3), we can find that the weight at this time must be similar weight. The greater the edge weight is, it represents that the closer the connection of the two nodes becomes. From the above definition, we can calculate that we can get the average clustering coefficient $c^w(k)$ of the nodes of taking all the degree values as $k$ and the average clustering coefficient of the network $c^w$. Under normal circumstances, the application of similar weight in the formula (2.10) needs to use the normalization of maximum similar weight. However, when we can make use of network property to make the similar weight be normalized to the $[0, 1]$ interval in advance, we can directly use the weight calculation to omit the second normalization step. This consideration can compare the weighted clustering coefficients of different networks.

In addition to the above basic network statistics amount, people also promote its other network properties on the weighted network. For example, Onnela and others has put forward a systematic way recently. They make the analysis of die body be generalized to the weighted network. Because the edge weight increases the number of dimensions of describing system properties, the corresponding concepts are built. The researches on the special statistical properties on the weighted network is still an important research content in weighted network. Weighted network has introduced the strength of interaction between vertices and has described the diversity of connection, and has increased the abstract describing ability of the network. At the same time, the introduction of edge weight also has greatly enriched the statistical properties of the network. Besides the connection decided by the edge weight, for the weighted network we must also focus on the statistical properties related to the weight, especially the correlation of weight and topology, and this provides a new perspective for us to understand the organizational structure of corresponding system. Many empirical studies show that the weighted network has showed rich statistical properties and power law behaviors. We will introduce some typical practical systems and the corresponding empirical analysis results in the following content.

3. Construction of cooperative competition network with library books lending relationship

In this paper, data are from the situation of borrowing books by the readers in the library of HeBei North University within one year. During this period, the total number of books which are borrowed by the readers is 83959 (namely, different book bar codes), and the number of kind for the books with different contents (namely, different call numbers) is 51084, and the total number of readers is 12610. These data have laid the substantial evidence foundation for the library books lending network research. First we have introduced the statistics situation of library raw data. Each reader corresponds to a reader bar code and each book corresponds to a book bar code. If a reader borrows a book, it corresponds for that a reader bar code connects a book bar code. In data format of books repositories, a call number corresponds to a number of different books bar codes. Because the library will provide more than one books with the same version (Specifically it means that the book name, author, press, publishing time are all the same.) for the readers to borrow, but we will use the same call number to identify these books with the same version (it is convenient for the readers to find.) We adopt different bar codes to distinguish each book (It is convenient for the librarians to conduct the books statistics.). In actual situation of borrowing books by the readers, assume that the reader has borrowed two books, books bar codes are different, but the content of the two books are exactly the same (the same call number). Therefore, we should consider them as the same book. Because of this, in order to show the situation of borrowing the same book by the readers more reasonably, we use the call number to replace books bar code so as to establish the library books lending network through SQL program in the paper. It contains two types of nodes in the bipartite graph of the library. For one kind, it represents the call numbers of books, and we call them as the projects. For another kind, it represents the readers who borrow the books, and we call them as the nodes. If the books lending relationship exists between projects and nodes, we connect the two with a line to constitute an edge. In the researches on the weighted network, we take the days of borrowing books by the readers as the edge weight value of this edge. In Figure 1, we use Netdraw to draw the weighted bipartite graph which contains four projects. In the entire bipartite graph, a kind of cooperative competition relationship is shown between the projects. That is to say, the lending amount of all books shows the service level of providing the books lending by the library, and the books quality level and the popularity degree form the competitive relationship between books. On the other hand, the cooperative competition relationship between the books borrowed by the same reader has formed. That is to say, these books jointly constitute the knowledge system of the reader. However, because of the reader’s limited energy, it makes the competition of scrambling for the services for readers among these books be formed. In each project, a relationship with both cooperation and competition...
is also shown between the nodes. If several readers borrow the same book, it will constitutes the borrowing value of this book. At the same time, because there are more readers and relatively few readers, the process that the readers borrow the same book jointly also forms the competition of books borrowing among the readers.

![Figure 1 Bipartite Graph](image)

**Figure 1** Bipartite Graph (In the figure, square represents the project, namely book, and dot represents node, namely reader, and the weight value on ligature represents the reading time of borrowing the book by the reader. represents.)

4. Bipartite graph property of library books lending network

4.1 Project size and node project degree

Project size refers to how many nodes are contained by a project. That is to say, the same book is borrowed and read by how many readers within one year. It roughly shows the competition strength size of the book. In Figure 2 it shows that the cumulative probability distribution of project size obtained through the statistics under the single logarithmic coordinate. From this we know that the project size obeys the exponential decay distribution, and its fitting curve equation is $P = 0.9519e^{-0.25s}$. 

![Figure 2 Cumulative Probability Distribution Graph of Project Size](image)

**Figure 2** Cumulative Probability Distribution Graph of Project Size

Combined with Figure 2 we can see that under the form of single logarithmic coordinate except for the tail end whose impact factor is very small, the cumulative probability distributions of project size are all approximate to a straight line. Table 1 is the distribution for the number of times of books borrowed by the readers. Among them, at most two books have been borrowed more than 70 times, 54.62% of books have been borrowed with the number of times range of two and nine, and 36.70% of books have been borrowed once.

According to the calculation, we get that the average project size $< T > = 9.44$. That is to say, the same book has been borrowed on average 9.44 times within one year. By statistics the categories of top hundred books with the maximum number of times of books borrowing are Chinese literature, English language literature and computer. What’s more, in the category of computer, there are respectively seven and four books about Matlab and C++. And they account for most of books about the computer category books borrowed by the readers. Similarly, among the four books about mathematics category, there are three books about arobility theory and mathematical statistics. Therefore, the library should increase the corresponding books procurement, and try to meet the needs of readers.

![Figure 3 Cumulative Probability Distribution Graph of Project Degree](image)

**Figure 3** Cumulative Probability Distribution Graph of Project Degree

<table>
<thead>
<tr>
<th>number of times of books borrowed by the readers</th>
<th>79,71</th>
<th>50-59</th>
<th>40-49</th>
<th>30-39</th>
<th>20-29</th>
<th>10-19</th>
<th>2-9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of books</td>
<td>2</td>
<td>5</td>
<td>13</td>
<td>42</td>
<td>268</td>
<td>3115</td>
<td>27903</td>
<td>18750</td>
</tr>
<tr>
<td>percentage</td>
<td>0.00</td>
<td>0.01</td>
<td>0.03</td>
<td>0.08</td>
<td>0.53</td>
<td>6.10</td>
<td>54.62</td>
<td>36.70</td>
</tr>
</tbody>
</table>

Table 1 Distribution Table for the Number of Times of Books Borrowed by the Readers

<table>
<thead>
<tr>
<th>books categories</th>
<th>novels of Chinese literature</th>
<th>English language</th>
<th>computer</th>
<th>British and American literature</th>
<th>mathematis</th>
<th>physics</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of books</td>
<td>29</td>
<td>20</td>
<td>16</td>
<td>11</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2 Distribution Table for the Categories of Top Hundred Books

Node project degree refers to how many books are borrowed from the library by one read within one year in total. It shows the reader’s competition strength size of borrowing books. Distribution of reader project degree describes the distribution situation of readers about library books borrowing. Figure 3 is the cumulative probability distribution of node project degree. Except for the tail end whose impact factor is very small, the distribution curve of project degree cumulative probability is approximate to $P = 0.9555e^{-0.50s}$. Combined with


243.3

© 2016 DEStech Publications, Inc.
doi:10.12783/jissn.1544-8053/13/6/243
Table 3 we can know that 43.24% of readers have the number range of books borrowing with two and nine within one year. From statistical data, we can get that the average project degree of each reader is 14.16. That is to say, each reader has borrowed 14.16 books on average within one year. By doing statistics on the top hundred readers with the maximum number of times of books borrowing, we can know that among them there are four teachers, 71 postgraduates, 24 other students. From a practical standpoint we can explain such phenomenon that the teachers have rich knowledge, and more of them are engaged in the specific areas, therefore, they borrow books relatively specializedly, and their economic strengths allow them to buy more books. However, postgraduates’ knowledge is relatively thin and they have relatively strong thirst for knowledge. About the studies and researches they all need to look up a lot of related literatures. And at the same time, postgraduates have large limits of authority (ten books) than other students (ten books), therefore, their limits of authority are relatively higher than other students.

<table>
<thead>
<tr>
<th>number of times of books borrowing by the readers</th>
<th>1</th>
<th>2-9</th>
<th>10-19</th>
<th>20-29</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>79-165</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of readers</td>
<td>1040</td>
<td>5452</td>
<td>3010</td>
<td>1491</td>
<td>750</td>
<td>434</td>
<td>202</td>
<td>111</td>
<td>57</td>
<td>62</td>
</tr>
<tr>
<td>percentage</td>
<td>8.25</td>
<td>43.24</td>
<td>23.87</td>
<td>11.82</td>
<td>5.95</td>
<td>3.44</td>
<td>1.60</td>
<td>0.88</td>
<td>0.45</td>
<td>0.49</td>
</tr>
</tbody>
</table>

### 4.2 Point strength distribution

By the observation we know that the library has the detailed records of borrowing and returning about each book borrowed by each reader. We can use SQL and ACCESS to do the statistics processing on data to get the reading time. If the reader borrows a book, we establish an edge between the two and at the same time we take the reader’s reading time of borrowing the book through the statistics as the edge weight of this edge. In this way we establish the weighted bipartite graph of library borrowing relationship. In the weighted network, the natural promotion which corresponds to node degree is the node strength. Node strength distribution \( p(k) \) is similar to the role of the degree distribution \( p(k) \). It mainly investigates the probability of having node strength \( S \) by the node. It is defined as

\[
S_i = \sum_{j \in N_i} w_{ij}
\]

Among them, \( N_i \) is the neighboring set of the node \( i \). Node strength both considers the neighboring number of the nodes, and the weight between nodes and neighboring nodes. It is the integrated embodiment of the node local information.

![Figure 4](image)

**Figure 4 Cumulative Probability Distribution Graph of Project Point Strength**

Book point strength is represented as the total number of days of books borrowed by the readers. Accurately it shows that the competition strength size of books. In Figure 4 it shows the accumulative distribution of books node strength, and we can use the exponential function to fit very well. Through the statistical analysis, the book average node strength in the library books lending network is 106.74. That is to say, each book is borrowed 106.74 days on average within one year. In Table 4 it lists the specific situation of number of days of books borrowed by the readers. The reading time of 92.29% of books is within 300 days which is less than one year’s time. This shows that the books in the library can meet the needs of readers basically. Because the book with the larger point strength represents that the book with longer time of being borrowed, its competition strength is relatively larger in the lending process, and it shows the readers’ borrowing interests and trends, we will emphatically analyze it. From Table 5 we can know that English, computer and other learning books are still ranked top among the books. Literature books have significantly decreased accordingly. It shows the readers’ concept of taking the learning as the majority and taking the entertainment as the supplementary. At the same time, we take English and computer as the key of learning. And it shows that they have become the indispensable tools in learning. What’s more, the number of power engineering, mechanics, physics, wireless and telecom technology is relatively uniform. It shows that the readers with various professional categories will all borrow the relevant professional books. It further illustrates the accuracy of the empirical work.

<table>
<thead>
<tr>
<th>number of books</th>
<th>14431</th>
<th>9740</th>
<th>8635</th>
<th>10204</th>
<th>4139</th>
<th>1955</th>
<th>960</th>
<th>425</th>
<th>415</th>
</tr>
</thead>
<tbody>
<tr>
<td>percentage</td>
<td>28.25</td>
<td>19.07</td>
<td>16.90</td>
<td>19.97</td>
<td>8.10</td>
<td>3.82</td>
<td>1.88</td>
<td>0.83</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>point strength</th>
<th>0-30</th>
<th>31-60</th>
<th>61-100</th>
<th>101-200</th>
<th>201-300</th>
<th>301-400</th>
<th>401-500</th>
<th>501-600</th>
<th>601-2176</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of books</td>
<td>14431</td>
<td>9740</td>
<td>8635</td>
<td>10204</td>
<td>4139</td>
<td>1955</td>
<td>960</td>
<td>425</td>
<td>415</td>
</tr>
<tr>
<td>percentage</td>
<td>28.25</td>
<td>19.07</td>
<td>16.90</td>
<td>19.97</td>
<td>8.10</td>
<td>3.82</td>
<td>1.88</td>
<td>0.83</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Table 5 Top Hundred Books with Maximum Point Strength (The books have the longest time of being reading.)

<table>
<thead>
<tr>
<th>category</th>
<th>English language</th>
<th>computer</th>
<th>mathematics</th>
<th>power engineering</th>
<th>mechanics</th>
<th>physics</th>
<th>British and American literature</th>
<th>telecom technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>29</td>
<td>25</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
In Figure 5 it shows the cumulative probability distribution graph of reader point strength. We can see that the main part of the distribution obeys the exponential distribution. Reader point strength is represented as the total number of days of borrowing books by the readers. More accurately it shows that the competition strength size of borrowing books by the readers. Through the statistics we get that the average node strength of the readers is 423.44 days. The reader with the maximum node strength is a teacher from College of Power Engineering. By analyzing the top hundred readers with maximum point strength, we find that among these readers there are 29 teachers, 70 postgraduates, and one student. And compared with the unweighted network, the difference exists partly in that the limits of authority for the time of teachers’ borrowing books (90 days) is longer than the postgraduates (The time of doctors’ borrowing books is 90 days and the time of masters’ borrowing books is 30 days) . Therefore, the total time of borrowing books can be increased accordingly.

4.3 Comparison of Unweighted Network and Weighted Network

In the bipartite graph, the project size in unweighted network corresponds to the number of times of books borrowed by the readers, and the book point strength in weighted network corresponds to the number of days of books borrowed by the readers. They all show the competition strength of books. Similarly, the project degree in unweighted network corresponds to the number of times of borrowing books by the readers, and the book point strength in weighted network corresponds to the total number of days of borrowing books by the readers. They also show the competition strength of borrowing books by the readers. Then whether there is a certain correlation between the two pairs of distributions in unweighted and weighted form? First of all, known from the above text we can know that the distributions which they obey are similar. This conclusion is the same with previous studies. The reason of producing this kind of mechanism is that it has further explanation in network evolution model. Second, as shown in Figure 6 and 7, they are both positive correlation, and they both obey the power-law distribution. It shows that if there are more number of times of books borrowed by the readers, the number of days of books borrowed by the readers also increases accordingly. Similarly, if there are more number of times of borrowing books by the readers, the number of days of borrowing books by the readers also increases accordingly. By studying the statistical relationship between degree and point strength we can know that when the node is involved in multiple projects, then this node has even larger competition strength. Combined with the practical example, because the library has limited the number and time of borrowing books by the readers, in order to borrow the books which are needed more by the readers, they usually return the books which are not much useful for themselves. Then even though a reader borrows a book, but it doesn’t represent that it has the reading value for the readers. What’s more, the length of number of days for the reading also reflects that books have different reading values for the readers, therefore, the weighted network can reflect the real situation more carefully than the unweighted network. These all show that in bipartite graph the unweighted network and the weighted network both can describe the relationship of this kind of cooperative competition relationship and its structure in the network, but only the weighted network can describe this kind of competition result more accurately.

5. Conclusion

By doing the statistics on the specific information of library books lending in HeBei North University within one year, we get the related characteristics of library books lending bipartite graph, including the project size distribution, the node project degree distribution and point strength distribution. They all obey the exponential distribution. Known from the nature of exponential distribution, its connection number of most nodes is roughly the same. The nodes whose connection number is much higher than the average number or much lower than the average number are rare. In this paper it shows that the library books lending in the library is relatively uniform, and the books lending is more perfect, and on the whole it can better meet the needs of readers’ books borrowing. Someone once questioned in the paper why the library distribution doesn’t conform to the power-law distribution. If it conforms to the power-law distribution it will show that the library has more disadvantages. For example, many people borrow the same book or a lot of books are borrowed by few people. At the same time, compared with the original number of readers or the number of library books, the increase of the number of readers or the increase of the number of books borrowed the readers are relatively few, therefore, it doesn’t meet the characteristics of the power-law distribution. That is to say, at the beginning of the network, there are only a few nodes and a large number of nodes will be added in it after that.

In addition, in bipartite graph the project size distribution and node project degree all have the positive correlation with corresponding book point strength distribution and reader point strength distribution, and on the double logarithmic coordinate axis it is approximate to a straight line. This illustrates that it conforms to the power-law distribution very well. What’s more, known from the above text, the weighted network is the same with the unweighted network, and it can describe the cooperative relationship and structure, but it can more accurately describe the network competition result.

Know from the analysis of relevant statistics amount, in the process of borrowing books, the readers insist on the concept of relying mainly on learning, and apart from learning they still read Chinese and foreign literature books to enrich the life. Among them, English and computer books are most favored by the readers ( The competition strength is the biggest.), especially English learning, matlab, C++ and other computer programming category and mathematical statistics category books, therefore, the library should increase the relevant type books. In readers group, we find that the strength of borrowing books by the teachers is far greater than the students. It also fully implies the teachers always stand at the forefront of discipline, and they actively improve their own academic levels. At the same time, the postgraduates have a strong thirst for knowledge, and become the dominant force of the readers group.

References