International Journal of u- and e- Service, Science and Technology
Vol.8, No. 10 (2015), pp.203-210
http://dx.doi.org/10.14257/ijunesst.2015.8.10.20

The University Library's Recommending System with the Personalized Recommending Functions

Jihong Wei

College of Civil Engineering and Architecture, Harbin University of Science and Technology, Harbin, 150080, China
1070402243@qq.com

Abstract

Recently, with the rapid development of science and cultural areas, the written and electronic amount of books in the university libraries increases sharply. When readers use the traditional searching system which is based on the collaborative filtering method, it is usually difficult for them to find out their interested books due to a large number of results from the system. Aiming at this problem, the paper points out a sort of personalized recommending system. This system optimizes the collaborative filtering method based on the information of the users, uses the new collaborative filtering method based on the classification of users and books, and analyzes and orders a variety of recommending information after filter. From the experiment, it concludes that compared with the traditional recommending system the personalized system targets towards different types of readers. The numbers of books which feed back to readers have decreased a lot. What's more, after a comprehensive analysis and order of the various recommending information, the average recommending accuracy will make further improvement.

Keywords: *the library recommending system, personalized recommendation, collaborative filter*

1. Introduction

In 1990s, the personalized recommendation technology has been pointed out formally as an independent concept. The studies of collaborative filter, recommending system and personalized search at home and aboard have been conducted for a long time. At aboard, Tapestry is the earliest case that is based on the collaborative system, which depends on mutual recognized users from a small community to share its opinions. Group Lens provides a method to recommend movies and news based on the anonymous collaborative filter, whose basic principle is to render the using community to cooperate together. Designed by MIT, Ringo is a system to recommend music and movies based on emails and World Wide Web. In 2006, Golde and Huberman have analyzed users' labels, users' behaviors and the collaborative system which is based on the labels to calculate the targets' relationship. With the wide spread of the Internet, the competition of different industries become more and more intensely. The personalized recommending system has become a powerful tool to improve its competing competence, whose key function is to increase the users' adhesion. The personalized recommending system has already drawn the attention of some enterprises and has brought them immense profit. According to Venture Beat's statistics, Amazon's recommending system provides 35% of the sale value. Sixty percent of the movies in Netflix are from the personalized recommendation.

ISSN: 2005-4246 IJUNESST

Existing personalized recommendation system undoubtedly already in the major online shopping platform has achieved great success [1]. However, in the huge resources of the library, the users' personalized studies are very few. At home, as early as 2005 Zeng Qinghui has launched the e-book recommendation system based on collaborative filter. They realized collaborative filter through evaluation of the users. It does not have the time-varying function, and cannot be achieved personalized recommendation [2]. In 2007, Sun Yizhou proposed a user clustering personalized book recommendation system, which also uses the collaborative filtering techniques to achieve a personalized recommendation through the user clustering. However, when a new account or a new book joins into the system, the recommendation system will not work [3]. Ma Yan proposed an adaptive collaborative filtering system to solve the problem of time-varying, remaining personalized recommendations unfilled. This thesis tries to extract and analyze the users' information in the university libraries, build up the informational module, and optimize collaborative filtering method based on users' information [4]. A new application based on the classification of users and books has been used. A variety of recommending information has been analyzed and ordered to suit personalized book search functions in university libraries, so as to improve the existing library searching system.

2. Recommendation based on the Collaborative Filter

Recommendation based on the collaborative filter can be divided into two categories: the user-based collaborative filter and target-based collaborative filter. Filtering based on a user's system mainly focuses on users' preference information on the project (generally in scores), to find out a neighbor set which is similar to the current user. The common method is to use the "K-neighbor", which is based on historical preference information of the K neighbors and makes recommendations for the current users. The latter is firstly to use all of the user preference information on the target or the information, find the similarity among the targets. Based on the users' history of preference information, the similar targets are recommended to the users [5].

Similarity computing is a key element in collaborative filtering. The traditional computing similarities include: Pearson correlation and cosine similarity, modified cosine similarity and the like. Since this thesis does not involve users' rating information, we use the cosine similarity metrics for the calculation of the similarity of users' information, as follows:

$$sim(i, j) = \cos(I, J) = \frac{\sum_{c=1}^{n} R_{i,c} R_{j,c}}{\sqrt{\sum_{c=1}^{n} R^2_{i,c}} \sqrt{\sum_{c=1}^{n} R^2_{j,c}}}$$

$R_{i,c}$ represents the number of books borrowed from user *i* which belongs to *c* classification. *n* means that all the numbers of books' classification. The books' classification adopt in the method of CLC, a total of 22 categories.

3. Personalized Recommending System

The main purpose of this study is how to apply a personalized recommendation system to the routine system of the university library, and adapt the working processes. This personalized book recommendation system can recommend their interested books to the readers at a fast speed. Through logging in the system reads the basic information, establish the users' information model, identify their nearest

neighbors through the filter system, and recommend the books borrowed by the nearest neighbors.

The overall framework of the personalized recommending system discussed in this paper mainly includes three modules: users' information model, personalized recommendation model and integrated sorting algorithm. As it is shown below:

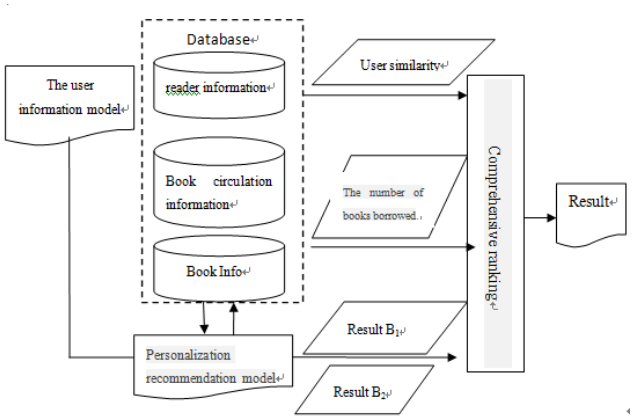

Figure 1. The Overall Framework Diagram of Personalized Book Recommendation System

3.1 The Users' Information Model

This thesis chooses the university library as the study object. The main information which determines whether the user is interested in the search results includes: the user's searching terms, the user type (student or teacher), the professional structure, the personal knowledge, the history of loan information. In practical application process, we find that users of the history of borrowing information may well reflect their interest in reading. Contrary to several other user information, historical records can transform into machine-readable models through a user-vector. Therefore, we can define the users as a vector: $\mathbf{U}(\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n)$. $\mathbf{b}_i$ indicates that the property of the users. Based on borrowing record filtering, it represents the number of the books of borrowing in history. While in the classification of filtering based on the books, it represents the number of books which users borrow the respective categories.

3.2 The Personalized Recommending Model

3.2.1 The Collaborative Filter based on the Borrowing Records of the Users: The collaborative filter based on the borrowing record means to calculate the similarity through the history records of every users, and then recommend the books to the current user borrowed by the similar users.

Unlike university users of e-commerce users, the information is relatively comprehensive and reliable, and the knowledge structure is clear. Therefore, based on prior knowledge of the structure of the filter, the further exploration of university library personalized book recommendations work has a very important practical significance. Collaborative filtering application of data analysis techniques, by generating a predicted user a preferring Top-N list of objectives to help users find the possible favorite books.

Firstly, we create a list that contains *m* users $U=\{u_1,u_2,\ldots,u_m\}$; and a list of books that contains *n* elements $I=\{i_1,i_2,\ldots,i_n\}$; corresponding to each user there is a lending

borrowed books numbered lists i_{ui}, which can be a empty set. To match book numbers in the book lists and book inventories, we can find the borrowing lists based on the name of the books. At last, we can decide the similarity through comparing the rate of the same title of the books in the borrowing lists in the total numbers of the list of two users. The calculation formula can be expressed as follows:

$$sim(u_1,u_2)=\frac{B_{u1}\bigcap B_{u2}}{\max(|B_{u1}|,|B_{u2}|)}$$

u_1 u_2 represents two different users, B_{u1} B_{u2} represent any two users on the title of a collection of books to borrow. The final range is between 0 and 1, when two users borrowed books all the same, the value is 1, indicating that they are highly similar, when two users borrowed books are not the same, the results of the formula 0, indicating a lower degree of similarity of these two users.

3.2.2 The collaborativeFilter Based on the Books' Classification: In accordance with the methods described above, there are more records to borrow from the same readers which are relatively similar, but not identical to borrow record readers similarity is very low. This is true to some extent, but clearly not representative of all the circumstances, such as when two users although not borrow the same books, but the borrowed books are similar, such as all about "Business English", "java language tutorial "a category of books or other, then we believe that these two users should also be similar. In order to achieve this similarity comparison, this paper points out the collaborative filtering method based on classification.

Similarly, first we create a list that contains users, as well as a list of books containing classified elements. Each user has a loan which has been borrowed books corresponding classification list, which can be both zero. Their angle is determined by comparison of the similarity of the two user vectors. The calculation formula can be expressed as follows:

$$sim(i,j)=\cos(I,J)=\frac{\sum_{c=1}^{n}U_{i,c}U_{j,c}}{\sqrt{\sum_{c=1}^{n}U^2{}_{i,c}}\sqrt{\sum_{c=1}^{n}U^2{}_{j,c}}}$$

$U_{i,c}$ and $U_{j,c}$ represent the number of books that the user *i* and user *j* has borrowed. *n* indicates the number of book classification, where we use the the first category of "Chinese Library Classification" and determine the classification numbers for 22 categories. The similarity of the results of such calculations in the range of 0-1. The greater the value, the higher the similarity of two user instructions. Only when two users to borrow books at the ratio of the number of books each classification are the same, the only similarity is 1.

3.2.3 Recommendation based on hybrid filter: Loan records and books based on classification of these two filtering methods, there are some limitations, for example, based on collaborative filtering borrow record is relatively accurate, but in practice, due to borrow books less the same situation, looking for Neighbors process often leads to the lack of relevant data the similarity of most users is 0, and the similarity of the neighbors are often in a very small range of values, the correlation is not easy to distinguish neighbor size. The recommended books based on the classification, although there is no problem in the above calculation of the similarity of, but because of the books classification is very general, often a category of books million copies itself to dollars, the same category of books which relate to the content is vastly different. To make up for these shortcomings, this paper attempts to mix these two methods, the specific method is to use two methods were nearest neighbors, each neighbor recommended books in a certain order to store a large list

of recommendations to the two of them, then it was taken from two lists a fixed result, re-deposited into the final list of results.

3.3 The Comprehensive Order of the Recommending Results

Both filtering methods described above will produce two different recommendation results:

1. Based on the current borrowing record, we generated the number of K neighboring users' collection of books B_1.
2. Based on the books' classification, we recommend the number of K neighboring users' collection of books B_2.

Obviously, if the result of all the above recommended to the user, the result set will be very large, recommended quality will be greatly affected. So the search results filtering and sorting process is very important and often is the key factor in the success of the entire recommending system. To address the above problems, this paper presents a complex method in the system after the completion of the recommended books recommended to base nearest neighbor sorting through the number of library books to borrow comprehensive sort.

3.3.1 The Recommending Process of the Results: Firstly, we get the following set of three books: the collection of books B3 by search terms, and the above B_1 and B_2. As it is shown in the following:

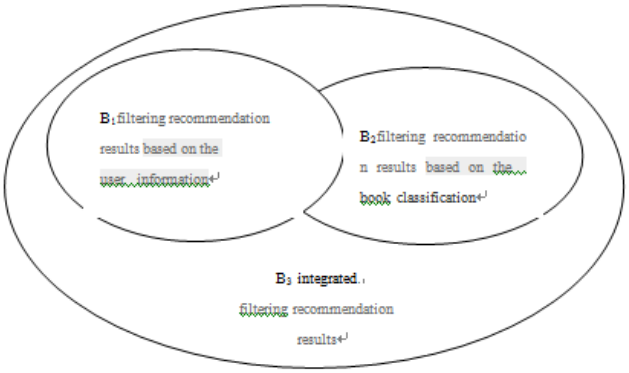

Figure 2. The Graph of Recommending Results

For the graph b in the three collections, its final recommendation results include two parts:

1. Based on lending recorded recommend results B1: Firstly to calculate K neighbors, and make order of the recommendation results for each neighborhood by the number of books borrowed from small to large order. Sorting methods can be shared, the main idea is to first generate one and the same length of the list to be sorted array and the number corresponding to borrow books assigned to the array, then the array by comparing the results to adjust the order of recommendation.
2. Recommended classification based on the results of the book B2: First calculate the K neighbors, also in accordance with the number of books to borrow sort the results by each neighbor after the results stored in B2 an orderly, sorting method is the same as the above. Due to the above two types of scores of different types, it is unsuitable for direct sum. It can be determined by the relationship between the two results a constant N, take a turn repeat the result in the final recommendation result. Its pseudo-code implementation process is as follows:

```
for (int i=1;i<Math.min(bl.size(), b2.size())/N; i++)
{
```

```
b3.addAll(bl.subList((i-1)*N, i*N));
b3.addAll(b2.subList((i-1)*N, i*N));
}
return b3;
```

4. The Experiment and Analysis of the Results

Under normal circumstances, a recommendation system by measuring the prediction is accurate and complete; we can evaluate the merits of the system. Recommended applications in various evaluation methods used are not the same. In order to verify the effectiveness of the system to recommend this article to a university in Harbin book circulation data for the experiment. Since personalized books recommended this recommendation is based on collaborative filtering-based, but at the same time did not involve user evaluation information, and therefore not suitable MAE using traditional methods. By dividing the training of users and test a user, the system was repeated Hitrate test. The experiment was as follows: (1) Assessing the effective of the recommendation and verifying the overall effectiveness of the entire personalized search system; (2) Checking the case of changes in the number of Top-N, the affection of recommending effectiveness.

4.1. The Preparation of the Experiment Statistics

This thesis uses a Harbin University Library from 2007 to 2010 the real flow of data for testing. A total of 12 000 including user data, book data 380 000, 300 000 circulation data for three years. Since the amount of data is too large, requires computing time is too long, this article did experiment to take 80% of every library user data is recorded to train the model, 20% tested. Training data set includes 200 users, a total of 13,368 loan records, 8,236 kinds of books; and the test data set includes 200 users, 3516 borrowing records, 1,366 kinds of books. The contents of which the user table by user number, serial number, etc to borrow books; borrow record sheet for the user ID, title, etc to borrow books; books contents of the table by the book title, loan number, classification and other components.

4.2. The Evaluating Standard of the Experiment

In this paper, the recommended system Hitrate common evaluation methods: the chances of data by comparing the test results appear in the recommendation or personalized recommendation system search system to evaluate the performance of the system, which is defined as follows:

$$hr(ui) = \frac{\left|T_{ui} \bigcap X_{ui}\right|}{\left|T_{ui}\right|}$$

T_{ui} represents the results, and X_{ui} represents the statistics.

4.3. The Experiment Results and Analysis

Because direct access to user satisfaction survey by the form, the results obtained are subjective and inconvenient for K values for debugging. Because every time K value adjustment, would need to be re-tested scoring recommendation result, it will take enormous time and labor costs. Therefore, this experiment Hitrate commonly used test.

In order to verify the reliability of the recommendation system, and determine the best value of K, we are taking a different K values (take 1,3,5,10), carried out

repeated testing, and finally found a 5 K take effect is more well, it Hitrate @ n test results are as follows:

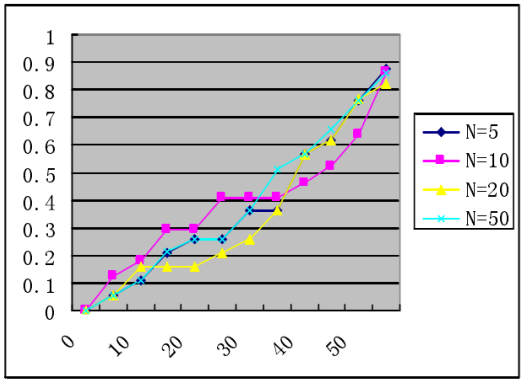

Figure 3. The Hitrate@n Graph based on the Sorted Mixed Recommendation

n represents the number of returned results. N represents a comprehensive range of continuous values when sorting.

From the above results, we can see that the test results are significantly better than the use of sorted and unsorted Hitrate value is higher, and the curve showing certain regularity, recommend effect with returns substantially linearly related to the number of results. Also based on the filtering effect is slightly better than the book catalog-based lending record filtering effect, it recommended mixed with high accuracy, and all test results, for different values, the effect of mixing the recommended less. In this paper, the research proved personalized book recommendation system; the accuracy and reliability of the overall recommendation system have reached a relatively satisfactory result, in order to achieve the purpose of the personalized book recommendations.

5. Conclusion

This paper focuses on the borrowing records based on user-based collaborative filtering techniques and classification of books personalized books collaborative filtering recommendation system applications. Nearest neighbors on the recommended method places the result of the sort generated based collaborative filtering, according to the number of nearest neighbors to borrow much to order. And also to study the methods recommended by the results of a comprehensive sorting process: by borrowing records and books based on user classification of collaborative filtering to produce neighbor users, all recommended books list of results, in accordance with the sort nearest neighbors on the basis of these books again lending to the number of books being sorted. The highest number of neighbors users borrow books lined the front position. This order from a certain extent, reflects the user's overall interest in books, in order to optimize the accuracy of the recommendation system is recommended.

Due to the current practice in the University library, almost no readily available book recommendation system, so the pilot phase of this article was not able to compare with another preferred method, the paper failed to demonstrate overall superiority recommendation system. In addition, the paper on the results of the recommended books sorting algorithm is relatively simple, how deep level mining user's preferences and interest in reading, this article is the direction of future research.

References

[1] D. Goldberg, D. Nichols, B. Oki, *et al.,* "Using Collaborative Filtering to Weave an Information Tapestry", Communications of the ACM, vol. 35, no. 12, **(1992)**, pp. 61-70.
[2] J. Konstan, B. Miller, D. Maltz, *et a1*., "GroupLens: Applying Collaborative Filtering to Usenet News", Communications of the ACM, vol. 40, no. 3, **(1997)**, pp. 77-87.
[3] U. Shardanand and P. Maes, "Social information filtering: Algorithms for automating 'WordofMouth'", Proc Conf Human Factors in Computing Systems Denver, **(1995)**, pp. 210-217.
[4] S. Badrul, K. George, K. Joseph, *et al.,* "Item-Based Collaborative Filtering Recommendation Algorithms", The Pennsylvania State University, **(2001)**, pp. 285-295.
[5] G. Adomavieius and A. Tuzhilin, "Toward the next generation of reeommender systems: A survey of the state of the art and Possible extensions", IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, **(2005)**, pp. 734-749.

