An Item-based Collaborative Filtering Approach based on Balanced Rating Prediction

Lei Ren
Dept. of Computer Science & Technology
Shanghai Normal University
East China Normal University
Shanghai, China
renlei@shnu.edu.cn

Junzhong Gu, Weiwei Xia
Dept. of Computer Science & Technology
East China Normal University
Shanghai, China
{jzhgu,wwxia}@ica.stc.sh.cn

***Abstract*—As a widespread approach in recommender systems, item-based collaborative filtering can predict an active user's interest for a target item based on his interest and the ratings for those similar items to his visited items. As the effect of human's conformity psychology, an individual user's judgment usually tends to follow the general view. The majority of existing item-based collaborative filtering approaches emphasizes the personalized factor of recommendation separately, but ignores the user's general opinions about items. Aiming at this issue of unbalanced recommendation, this paper proposes a refined item-based collaborative filtering approach which employs a balanced rating prediction method incorporating an individual's personalized need with the general opinions. The experimental result shows an improvement in accuracy in contrast to the classic item-based collaborative filtering.**

Keywords: recommender system; item-based collaborative filtering; balanced rating prediction

I. INTRODUCTION

With the rapid development of web, numerous online information systems have emerged during the last two decades. Facing those numerous information, users have to spend a lot of time and effort on selecting information according to their individual information needs, and this issue is termed as "information overload" [1]. To address the issue, recommender systems have been proposed in academia and industry, and have been becoming a core component in the adaptive online information system [1, 2]. Recommender systems can produce individualized recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options [3].

Item-based collaborative filtering (IBCF) is one of the most promising approaches employed in recommender systems [4]. It follows the assumption that the active user will more likely prefer similar items to those which he has preferred previously. Rating prediction is a key step in IBCF. For a given active user, rating prediction for the target unvisited item is generally a weighted sum of the observed ratings for his visited items similar to the target, in which the importance of each neighbor item is weighted by its similarity respectively [4, 5]. Whereas, as the effect of conformity psychology in social activities, an individual's judgments for things usually tend to follow the general views about them. Specifically in the community of IBCF, most of existing IBCF approaches have merely concentrated on the aspect of personalization in recommendation separately, but the general opinions about items have not been taken into consideration sufficiently.

Aiming at the issue of unbalanced recommendation, a refined item-based collaborative filtering approach based on balanced rating prediction (ICFBP) is proposed in this work. During the phase of rating prediction, it employs a balanced rating prediction method dynamically incorporating the individual user's personalized need with the general opinions about items. The rest of this work is organized as follows: related works are reviewed in section Ⅱ. Next, section Ⅲ describes the proposed approach. Then in section Ⅳ, ICFBP is evaluated experimentally. Finally, conclusions are made in section Ⅴ.

II. RELATED WORKS

Item-based collaborative filtering takes each column of the user-item matrix as the item profile based on which the similarities between items can be induced. The similarity between two given items indicates their distance in rating space with respect to those users who have rated both involved items. For a given item, its k nearest neighbors are selected from the active user's visited items based on their similarities to the target item. And then, for the active user, the rating for each of his unvisited items is estimated through a weighted sum of the observed ratings for its neighbors [4].

The item-based similarity is a key measure used in neighbor selection and rating prediction. It is the reference to select the target item's neighbors, and as well it acts as the weight to measure the importance of the ratings for neighbors in rating prediction. In existing IBCF approaches, two representative similarity measures have been used. The item-based Pearson correlation coefficient (PCC) is the most frequently used one [4, 5], and it indicates the linear correlation of the ratings for two involved items. The item-based Cosine similarity (COSIM) takes the ratings for an item as a vector in the user space, and it can be described as cosine of the angle between two involved rating vectors.

978-1-61284-774-0/11/$26.00 ©2011 IEEE

The other key phase in IBCF is rating prediction. The rating prediction for an active user's unvisited items is carried out based on his historical preference. The unobserved rating for an active user's unvisited item is usually computed as a weighted aggregate of his observed rating. As the effect of human's conformity psychology in daily life, an individual user's judgment for an item tends to follow the general view about it. In other words, generalization is also a useful factor in recommendation except for personalization, and therefore both factors should be balanced. But the classic rating prediction depends only on the ratings for the active user's neighbor items without considering other users' general opinions. Few researchers have noticed the aspect of generalization in recommender systems or made full use of it. In existing research incorporating clustering with recommender systems [6, 7], the average rating of users in a given cluster can represent its local opinion, but helps little in the improvement of accuracy. On the contrary, [8] proposes an E-learning recommender system which predicts the unobserved ratings for unvisited items based on their similarity to good learner's average ratings. But in essence, the proposed approach in [8] is not a personalized recommender system, because the target user's personal preference is not considered in prediction, it can hardly ensure the accuracy.

III. THE PROPOSED APPROACH

As discussed in section Ⅱ, the issue of unbalanced rating prediction is a crucial factor affecting the quality of recommendation. Focusing on improving the accuracy of rating prediction, the proposed ICFBP reforms the classic IBCF by employing a balanced rating prediction method. ICFBP can be taken as a variant of the classic IBCF, and its working flow can be divided into four phases:

- Based on the historical ratings for items, similarities between items are measured by PCC or COSIM;
- For the target item, those *k*-percentage-nearest items of the active user's visited items are selected as its neighbors according to their similarities;
- To predict the rating for each of the active user's unvisited items, a balanced rating prediction method is employed, in which the active user's personalized rating is dynamically incorporated with the general rating for the target item;
- For a given user, the ranked items with corresponding top-*n* estimated ratings will be recommended to him.

In ICFBP, the set $U=\{u_1,u_2,\ldots,u_i,\ldots,u_m\}$ denotes the user set with m users, and the set $T=\{t_1,t_2,\ldots,t_j,\ldots,t_n\}$ denotes the item set with n items. All users' historical observed ratings are represented and stored in the user-item rating matrix $M(m\times n)$, in which the element $r_{ij}\in M$ indicates the rating for item t_j by user u_i. The value of rating r_{ij} is usually defined as an integer varying in a given range such as 1 to 5, or a continuous value is optional. The higher the rating is, the more the user prefers the item.

The similarities between different items can be measured in item-based PCC or COSIM. The item-based PCC [4, 5] is defined as

$$w_{ij}=\frac{\sum_{u\in U_{ij}}(r_{u,i}-\overline{r}_i)(r_{u,j}-\overline{r}_j)}{\sqrt{\sum_{u\in U_{ij}}(r_{u,i}-\overline{r}_i)^2}\sqrt{\sum_{u\in U_{ij}}(r_{u,j}-\overline{r}_j)^2}}, \tag{1}$$

where U_{ij} is the set including those users having rated both item i and j, and r_{ui} denotes the rating for item i by user u, and $\overline{r}_i$ is the average rating of item i. The item-based COSIM can be formulized as

$$w_{ij}=\frac{\sum_{u\in U_{ij}}r_{u,i}r_{u,j}}{\sqrt{\sum_{u\in U_{ij}}r_{u,i}^2}\sqrt{\sum_{u\in U_{ij}}r_{u,j}^2}}. \tag{2}$$

In most of existing IBCF approaches, the neighborhood size for the target item is set as an absolute threshold. Whereas the amount of different users' visited items does not distribute identically, and an identical absolute threshold can hardly fit all users' rating distribution, which means the neighborhood size is a user-dependent parameter. So in ICFBP, the size of neighborhood k is set as a relative percentage of the amount of each user's visited items.

Based on the discussion in section Ⅱ, the aspects of both personalization and generalization of recommendation should be considered in rating prediction. ICFBP employs a balanced rating prediction method incorporating the active user's personalized rating with the general rating for the target item. The proposed rating prediction method can be taken as a linear combination of personalized rating and general rating for the target item, and it can be formulized as

$$r_{ui}=\alpha r_{ui}'+(1-\alpha)g_i=\alpha\cdot\frac{\sum_{j\in T_{ui}}r_{uj}w_{ij}}{\sum_{j\in T_{ui}}\left|w_{ij}\right|}+(1-\alpha)g_i, \tag{3}$$

where r_{ui}' denotes the personalized rating for item i by user u, which is estimated by the classic rating prediction method and represents the aspect of personalization in recommendation. On the other hand, the general rating for item i is defined as g_i. For its insensitivity to outlier, ICFBP takes the median of the ratings for the target item as the general rating.

To retain the balance between personalized rating and general rating, a parameter α is introduced as a weight to measure their corresponding significance in rating prediction. Moreover for the diversity of users, the opinion about a given item usually varies with the user's background. Especially for controversial items, different users can hardly reach a consensus about them. The variability of ratings can indicate the extent of users' disagreement about it, and more variability generally means less reliability for the general rating. So as a practical measure for variability, the statistical dispersion is employed in rating prediction as the weight for general ratings, which is denoted as the weight α in (3).

The weight α can be measured as one of the following three relative dispersion measures. The first one is the coefficient of variation (CV), which defines the ratio of standard deviation to the mean of a data set. For ICFBP, the CV of ratings for item i can be defined as

$$c_i = \frac{\sigma_i}{\overline{r_i}} = \frac{\sqrt{\frac{1}{|U_i|}\sum_{u \in U_i}(r_{ui} - \overline{r_i})^2}}{\overline{r_i}}, \tag{4}$$

where σ_i denotes the standard deviation of the ratings for item i, and $\overline{r_i}$ is its average rating. The set U_i includes those users having rated item i.

The second measure is the coefficient of mean deviation (CMD) which defines the ratio of mean deviation to mean, and its definition can be denoted as

$$c_i = \frac{\delta_i}{\overline{r_i}} = \frac{\frac{1}{|U_i|}\sum_{u \in U_i}|(r_{ui} - \overline{r_i})|}{\overline{r_i}}, \tag{5}$$

where δ_i is the mean deviation of ratings for item i, and the definition of other symbols is same to those in (4).

The third one is the quartile coefficient of dispersion (QCD) which characterize the dispersion in terms of quartile, and it can be formulized as

$$c_i = \frac{Q_{i3} - Q_{i1}}{Q_{i3} + Q_{i1}}, \tag{6}$$

where Q_{i3} and Q_{i1} denote the first and third quartile of ratings for item i. More contents about the above three measures can be found in [9]

Finally, all the unvisited items can be regarded as the candidates for recommendation. ICFBP can produce the top-n recommendation [10]. According to the estimated ratings for those candidates, ICFBP produces a ranked item list with n top-ranked items.

IV. Experimental Evaluation

To empirically optimize the parameter and prove the effectiveness of ICFBP, three experiments are conducted in this section. Focusing on the improvement of recommendation quality, the proposed approach is compared with the classic item-based collaborative filtering.

In following experiments, the MovieLens dataset is adopted as training and test data. The MovieLens dataset was collected by the GroupLens research project at the University of Minnesota, and has been widely adopted by the research in the community of collaborative filtering. The dataset contains 100,000 anonymous discrete ratings scaling from 1 to 5 for approximately 1682 movies by 943 users. Each user in the data set has at least 20 ratings, and each item has at least one rating. The sparsity of MovieLens reaches 93.7% which is appropriate for testing the proposed approach's sensitivity for sparsity. To ensure that each rating in the dataset can be tested once and only once, 10-fold cross-validation is used as the test and partition method.

For accuracy metrics, mean absolute error (MAE) is employed as predictive accuracy metric. To test the significance of results, paired t-test is employed at the significance level of 0.05 (5%).

In following experiments, the classic item-based collaborative filtering approaches based on PCC (IBCF-PCC) and COSIM (IBCF-COS) are set as the benchmark, and the proposed approaches with three kind of weight schemes employing PCC (ICFBP-CV-PCC, ICFBP-CMD-PCC, ICFBP-QCD-PCC) and COSIM (ICFBP-CV-COS, ICFBP-CMD-COS, ICFBP-QCD-COS) are compared with the benchmark.

Firstly to optimize the neighborhood size, the neighborhood experiment tests the benchmark and the proposed approach in different level of neighborhood size. The result is demonstrated in Fig. 1.

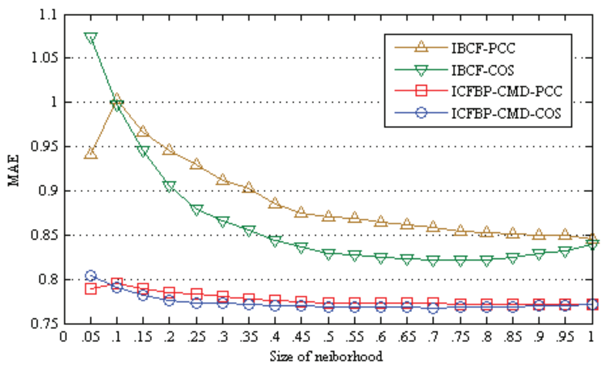

Figure 1. Impact of neighborhood size

As Fig. 1 indicates, with the size of neighborhood increasing, MAE of the classic IBCF and ICFBP employing COSIM and PCC tends to decrease. The approaches employing COSIM reach the lowest MAE around 70% (0.7) of neighbors, and then more neighbors help nothing in improving accuracy. But the approach employing PCC need more neighbors to achieve prediction. Moreover COSIM-based approaches present more accurate recommendation than PCC-based ones. So in following experiments COSIM is employed and the neighborhood size is specified as 70% (0.7).

Secondly to test the effectiveness of ICFBP, the accuracy of ICFBP is compared with the one of classic IBCF. The result is illustrated as Fig. 2 and Tab. Ⅰ.

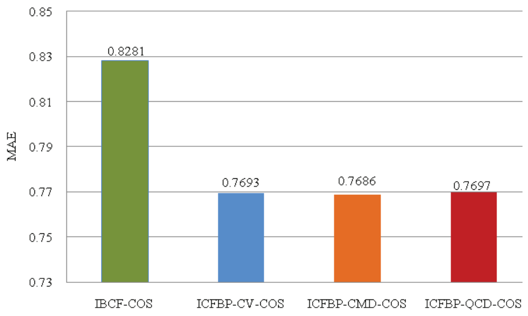

Figure 2. Comparison of the classic prediction and the balanced prediction method

TABLE I. STATISTICAL SIGNIFICANCE OF THE RESULT

t-test(α=0.05 df=18 one-tail)	*p-value*	*t-value*
ICFBP-CV-COS vs IBCF-COS	1.2572e-6< 0.05	-6.7506
ICFBP-CMD-COS vs IBCF-COS	1.1540e-6 < 0.05	-6.7947
ICFBP-QCD-COS vs IBCF-COS	2.2950e-6 < 0.05	-6.4449
ICFBP-CV-COS vs ICFBP-CMD-COS	0.5348 > 0.05	0.0886
ICFBP-CV-COS vs ICFBP-QCD-COS	0.4821 > 0.05	-0.0454
ICFBP-CMD-COS vs ICFBP-QCD-COS	0.4490 > 0.05	-0.1301

The result shows that ICFBP is more accurate than the classic IBCF. We also notice that the difference of ICFBP employing different weight schemes is not statistical significant which indicates the equivalence of the proposed weight schemes.

To test the impact of sparsity, a sparsity experiment is conducted, in which 10 rating matrices are randomly exacted from the original MovieLens data set. The number of users in each matrix varies from 10% to 100% of the original number of users by step of 10%, and similar process is applied to the number of items in each rating matrix. The result is presented in Fig. 3 and Tab. Ⅱ.

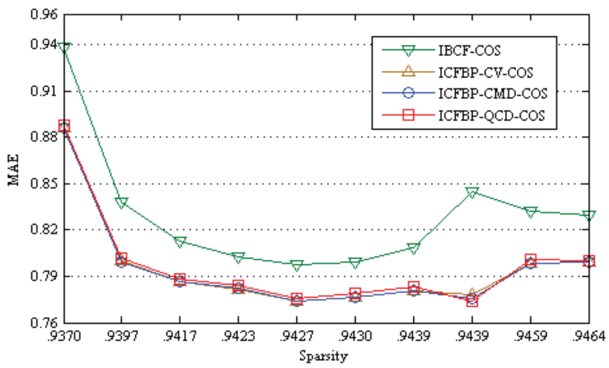

Figure 3. Impact of the sparsity of training data set

TABLE II. VARIANCE FOR THE RESULT OF SPARSITY EXPERIMENT

	IBCF-COS	**ICFBP-CV-COS**	**ICFBP-CMD-COS**	**ICFBP-QCD-COS**
Variance of MAE	0.0016	0.0010	0.0010	0.0010

The result indicates that, with the sparsity of user-item matrix varying, ICFBP presents lower variance for MAE than the classic IBCF, so ICFBP is more stable for sparsity.

V. CONCLUSIONS

Aiming at the issue of unbalanced recommendation, an item-based collaborative filtering approach based on balanced rating prediction is proposed in this work. Based on the distribution of ratings for the target item, ICFBP retains a dynamic balance between personalization and generalization of recommendation. The experimental result indicates that the proposed approach can provide more accurate and stable recommendation than the classic IBCF.

ACKNOWLEDGMENT

This work is partly supported by the grant from STCSM project (No. 09510703000 and No. 08511500303) and SHNU project (No. A-3101-10-033). The authors also wish to thank GroupLens for providing the MovieLens dataset.

REFERENCES

[1] P. Resnick, et al., "GroupLens: an open architecture for collaborative filtering of netnews," in Proceedings of the 1994 ACM conference on Computer supported cooperative work, Chapel Hill, North Carolina, United States, 1994, pp. 175-186.

[2] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. on Knowl. and Data Eng., vol. 17, pp. 734-749, 2005.

[3] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," User Modeling and User-Adapted Interaction, vol. 12, pp. 331-370, 2002.

[4] B. Sarwar, et al., "Item-based collaborative filtering recommendation algorithms," in Proceedings of the 10th international conference on World Wide Web, Hong Kong, Hong Kong, 2001, pp. 285-295.

[5] G. Linden, et al., "Amazon.com recommendations: item-to-item collaborative filtering," Internet Computing, IEEE, vol. 7, pp. 76-80, 2003.

[6] K. Taek-Hun, et al., "Improving Prediction Quality in Collaborative Filtering Based on Clustering," in Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on, 2008, pp. 704-710.

[7] B. Sarwar, et al., "Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering," in Proceedings of the Fifth International Conference on Computer and Information Technology, 2002.

[8] K. I. Bin Ghauth and N. A. Abdullah, "Building an E-learning Recommender System Using Vector Space Model and Good Learners Average Rating," in Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on, 2009, pp. 194-196.

[9] D. Freedman, et al., Statistics. New York; London: W. W. Norton, 2009.

[10] J. L. Herlocker, et al., "Evaluating collaborative filtering recommender systems," ACM Trans. Inf. Syst., vol. 22, pp. 5-53, 2004.