Combining User-based and Item-based Collaborative Filtering Techniques to Improve Recommendation Diversity

Jing Wang
School of International Programs
Neusoft Institute of Information Technology, Nanhai
Foshan, China

Jian Yin
Department of Information Science and Technology
Sun-Yat Sen University
Guangzhou, China

Abstract—**Nowadays collaborative filtering technologies are widely used in many websites, while the majority research literatures focused on improving recommendation accuracy. However, it had been recognized that improving recommendation accuracy was not the only requirement for achieving user satisfaction. One important aspect of recommendation quality, recommendation diversity gained focus recently. It was important that recommending a diverse set of items for improving user satisfaction since it provided each user with a richer set of items to choose from and increased the chance of discovering potential interest. In this study, a synthetically collaborative filtering model was proposed, which combined the user-based and item-based collaborative filtering techniques. This model gave each user an option to adjust the diversity of their own recommendation list by using the prevalence rate and novelty rate parameters. Experiments using real-world rating datasets indicated the proposed model had effectively increased the recommendation diversity with little decrease in accuracy and surpassed the traditional collaborative filtering techniques.**

Keywords-recommender system; collaborative filtering; recommendation diversity

I. INTRODUCTION

The immense growth of the World Wide Web has led to the information overload problem. It is difficult for users to quickly obtain what they want from massive information. Recommender systems are designed to overcome this problem by directly recommending information to users according to their history behavior in the systems. Collaborative filtering (CF) is one of the most successful recommendation techniques. In a CF recommender system, there exist some users and items. Preference data of a user in a pure CF-based system is represented as a user-item rating matrix. User-based [1, 2] and item-based [3, 4] CF are classical CF techniques which are widely used because of their simplicity and predicting accuracy. The user-based CF first searches the most similar users of the active user, then predicts the rating of the target item that the active user will give based on his/her average rating and the ratings that his/her neighbors have assigned to the target item. The item-based CF, in contrast, exploits similarities between items based on the ratings they received from users. It makes predictions for the active user based on how (s) he has rated items that are similar to the target item. However, there are two drawbacks in the traditional CF techniques: (1) Most of recommended items are popular. Even if the user does not use the recommender system, such items can also be known by other means. (2)The recommended items may be too similar to each other. For example, suppose that an online store always recommends the books written by the same author, the user may become frustrated for there is little variety. These issues will enable users to feel dissatisfied, even cause users to give up using the system. In recent years the importance has been emphasized that the recommended items should be diverse. Adomavicius and Kwon explore some item ranking techniques which can generate higher aggregate recommendation diversity across all users while maintaining comparable levels of recommendation accuracy [6]. However, these techniques generated recommendation results according to rank the predicted ratings, thus were still unable to resolve the previous problems. Hurley and Zhang regard the tradeoff between similarity and diversity as a binary optimization problem and define a controller to explicitly tune the two metrics for obtaining the optimal tradeoff [7]. Some researchers also adopt the clustering [8] and statistics [9] techniques to improve the recommendation diversity. These methods adjust the diversity of recommendation results only from the perspective of the algorithms, without considering the actual needs of users. In this paper, we propose a synthetically collaborative filtering (SCF) model by combining user-based and item-based collaborative filtering techniques to improve recommendation diversity. Experiments using real-world rating datasets show the proposed model effectively increases the diversity of recommendation result with little decrease in accuracy and surpasses the traditional collaborative filtering technologies.

II. RECOMMENDATION DIVERSITY

A. Definitions

There are two levels to interpret recommendation diversity: individual diversity (ID) and aggregate diversity (AD). Individual diversity refers to the diversity within a recommendation list, which measures the ability to provide diverse items to each individual user. The aggregate diversity refers to the diversity between recommendation lists, which measures the ability to recommend different items to different users. Note that high individual diversity does not necessarily imply high aggregate diversity [5]. To take an extreme example, if the system always recommends the five different best-selling items, the recommendation list for each user is diverse, but only five distinct items are recommended to all users, the recommendation is too centralized. On the other hand, high aggregate diversity does not mean high individual diversity. Suppose a system always recommends the same type of items to a particular user. The recommendation lists between users may be diverse, but the recommendation is poor because the items are too similar for each user. In order to enhance the users' satisfaction, the system must constantly provide diverse items.

B. Metrics

1) Aggregate diversity

Aggregate diversity can be measured by the Hamming distance [11] or coverage [6]. Hamming distance measures the difference between top-N places of users' recommendation lists. Coverage measures the percentage of different items that the recommender system is able to generate for all users. Since we intend to measure the performance of the recommender systems based on the top-N recommended items lists, in this paper we use the coverage as the metric of aggregate diversity, as follows:

$$\mathrm{AD} = \frac{1}{|I|}\left|\bigcup N_u\right| \tag{1}$$

where I is the item set and $|I|$ is the number of items in the set. N_u is the recommendation result which is recommended to the user u. $|\cup N_u|$ is the total number of items recommended across all users. The greater the value is, and the more diverse (more personalized) items are recommended to users.

2) Individual diversity

Denoting the recommended items for user u as $\{i_1, i_2 \ldots i_N\}$, the individual diversity to user u is given as the average dissimilarity of all pairs of recommended items in the recommendation list N. That is,

$$ID_u = \frac{1}{N(N-1)/2}\sum_{i\in N}\sum_{j\in N, j\neq i} d(i,j)$$
$$d(i,j) = 1 - sim(i,j) \tag{2}$$

where $sim(i,j)$ is the similarity between items i and j, which can be obtained either directly from the input ratings or from item meta-data. $d(i,j)$ refers to the dissimilarity, which is defined as 1 minus the $sim(i,j)$. Here we assume that this similarity is symmetric ($sim(i,j) = sim(j,i)$). Mean individual diversity can be further averaged over all users. The lower this quantity is, the more diverse items are recommended to the users.

III. SYNTHETICALLY COLLABORATIVE FILTERING MODEL

A. Item Popularity

In the CF recommender system, item popularity (IP) measures the users' rating frequency for each item in the system. Given item *i,* the IP can be calculated through the ratio of the number of users who have rated the item to the total number of users in the system, as follows,

$$IP_i = \frac{|R_i|}{|U|} \tag{3}$$

where $|R_i|$ is the number of users who have rated the item i. U is the user set and $|U|$ is the number of users in the set. Items with high popularity are called popular items, otherwise known as the long tail items.

To demonstrate the distribution of the popular items and long tail items in the recommendation list, we compare the user-based collaborative filtering (UBCF) and item-based collaborative filtering (IBCF) algorithm in the MoviesLens 100k dataset which contains 100,000 ratings (1-5 scales) from 943 users on 1682 movies (items) [18]. Table 1 records the number of items belonging to the 500 popular items, 500 long tail items, and the distinct items. *N* is the size of recommendation list.

TABLE I. POPULAR AND LONG TAIL ITEMS

N	UBCF			IBCF		
	popular items	long tail items	distinct items	popular items	long tail items	distinct items
10	395	46	788	265	49	536
20	+55	+21	+154	+108	+82	+353
30	+76	+29	+225	+164	+165	+620
40	+86	+33	+257	+191	+219	+754
50	+95	+35	+283	+202	+255	+849

As illustrated by Table 1:

1) When N is small, UBCF and IBCF tend to recommend popular items.

2) As N increases, UBCF tends to recommend popular items, while IBCF tends to recommend long tail items.

3) As N increases, compared to UBCF, IBCF can recommend more distinct items.

B. SCF Recommendation Model

If a recommender system only recommends popular items to a user, even if the user is strongly interested in these items, it is poor recommendation because it is easy for the user to find them by other means. On the other hand, if a recommender system only recommends long tail items, which are novel for users, but recommendation accuracy may decrease because few users have preferred them. A good recommender system not only provides diverse recommendation but also ensures recommendation accuracy. As observed from Table 1, UBCF and IBCF compensate for recommending popular and long tail items. On such consideration, we combine UBCF and IBCF approaches to a synthetically collaborative filtering model, as follows:

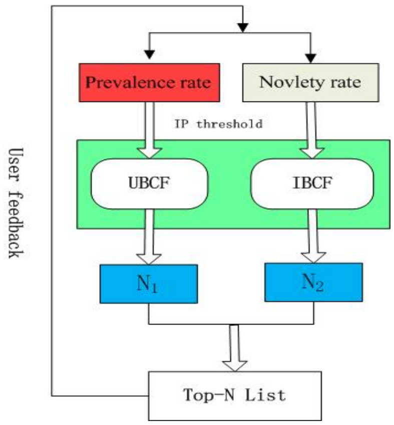

Figure 1. SCF Model

In SCF recommendation model the set of recommendation list N is composed of two subsets which are generated from UBCF and IBCF respectively. One subset contains N_1 items that best match the user's stated preferences. Another subset includes N_2 items that take into account the user's divers need. Assume that i is one of the items in the system, the recommendation list N is composed as follows:

$$N = \{i \mid i \in N_1 \bigcup N_2, N_1 \bigcap N_2 = \Phi\} \qquad (4)$$

In SCF recommendation model there are three parameters: prevalence rate, novelty rate and item popularity threshold, which are our diversity adjusting strategies. The prevalence rate controls the number of the popular items, and novelty rate controls the number of the long tail items. They can be inputted by users according to their desire. The item popularity threshold can be dynamically retrieved according to the number of users in the system. The SCF recommendation model can accept users' feedback and provide more diverse recommendation result, thus it may improve the users' experience and satisfactory.

C. SCF Recommendation Algorithm

Algorithm Steps Input：user-item rating matrix， item popularity vector， the size of recommendation list N， prevalence rate α, novelty rate β and item popularity threshold δ Output：a recommendation list with N items
1 Begin 2 Generating $N*\alpha$ popular items(item popularity is greater than δ) by UBCF, denoted as N_1 3 Generating $N*(1-\alpha)$ items by IBCF, denoted as N_2, where $N*\beta$ ($\beta \leqslant (1-\alpha)$) belonging to the long tail items (item popularity is less than δ) , and the remainder items are produced by IBCF directly 4 Return N_1+ N_2 5 End

Figure 2. SCF Recommendation Algorithm

IV. EMPIRICAL RESULTS

A. Dataset

We evaluate our model on the MovieLens 100K datasets, which contains 100,000 ratings (1-5 scales) from 943 users on 1682 movies and each user has rated at least 20 items [18]. There are 18 genres of the movies, such as Action, Adventure, Animation, Children's, Comedy, Crime, Documentary, Drama, Fantasy, Film-Noir, Horror, Musical, Mystery, Romance, Sci-Fi, Thriller, War, and Western.

B. Recommendation Diversity

In order to evaluate the recommendation diversity we compare our SCF model with UBCF [2] and IBCF [3] techniques. In UBCF and IBCF, the similarities between users or items are calculated using Pearson Correlation Coefficient (PCC). A threshold-based Neighborhood [12, 13] is used for selecting similar users in UBCF (The threshold is set as 0.6). In our SCF model, the prevalence rate(α) is set as 0.3, novelty rate (β) is set as 0.3 and the item popularity threshold(δ) is set as 0.005(according the number of users, about 1000, and items with less than five ratings belonging to the long tail items).

1) Aggregate diversity

In order to evaluate the aggregate diversity we vary the size of recommendation list. As shown in Figure. 3, with the increase of N, the aggregate diversity of three approaches increase accordingly. With any size of recommendation list, the SCF model obtains highest aggregate diversity. For example when the size of recommendation list is 5, the aggregate diversity of SCF model is more than 0.5, which indicates that the system can recommend more than half of the distinct items in the system. When the size of recommendation list is 30， the aggregate diversity of SCF model is above 0.8， far better than UBCF (0.6023) and IBCF (0.6837). With the increase of N, the aggregate diversity of IBCF increases quickly. When the size of recommendation list is 30, the aggregate diversity of IBCF has exceeded UBCF. As mentioned earlier, with the increase of the size of recommendation list, IBCF tends to recommend long tail items, so it can recommend more distinct items.

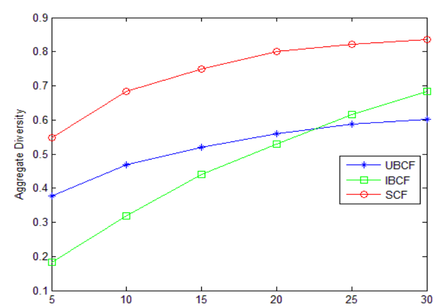

Figure 3. Aggregate diversity

2) Individual diversity

Evaluating individual diversity is in order to avoid recommending too many similar items for each user. Here the similarity means that the contents of items are similar. The dataset provides the genres of movies, so we use the similarity formula of item genres [14] to calculate the similarities of items. When comparing the individual diversity, the recommendation list N is set as 5. We generated five subsets from the original dataset by 5 fold cross-validation, denoted as data1 to data5. As observed from Figure.4, in general, all approaches were able to provide significant individual diversity, which means that it is reasonable to adopt item genres to calculate the similarities of items. The individual diversity of our SCF model is above 0.88, and UBCF is slightly lower, between 0.87 and 0.88. The individual diversity of IBCF is between 0.83 and 0.85, which weaves in the different datasets.

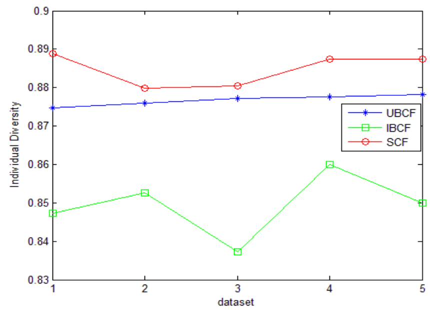

Figure 4. Individual diversity

C. Recommendation Accuracy

Recommendations from the long tail of the popularity distribution of items are valuable. On the other hand, recommendation accuracy tends to decrease towards the long tail items which often have fewer ratings and are more difficult to predict. In this paper, we use precision [15] to measure recommendation accuracy, as follows:

$$\text{precision} = \frac{|T \cap L|}{L} \tag{4}$$

Where L is the recommendation list and T is the test set. The precision measures the ratio of common items in T and L to the recommendation list.

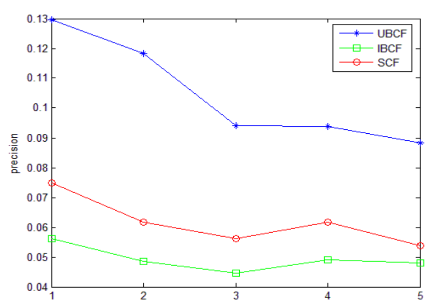

Figure 5. Accuracy of of different approaches

As shown in Figure.5, the accuracy of UBCF is highest, the accuracy of IBCF is lowest, and that of SCF is between them. As known well, there is a tradeoff between accuracy and diversity because high accuracy may often be obtained by safely recommending to users the most popular items, which can clearly lead to the reduction in diversity. And conversely, higher diversity can be achieved by trying to recommend highly idiosyncratic or personalized items for each user. These items often have less ratings and are inherently more difficult to predict, thus such recommendations may lead to a decrease in recommendation accuracy [5].

D. Parameters

In our SCF model presented in this paper we introduce two parameters: prevalence rate α and novelty rate β. The prevalence rate α is used to adjust the number of popular items. As shown in Figure.6 (a), when α is 0.2(data2 and data4) or 0.3(data5), the aggregate diversity is highest. As shown in Figure.6 (b), with the increase of α, the accuracy increases. In order to improve the recommendation diversity simultaneously with little decrease in recommendation accuracy, α is set 0.3 in our SCF model.

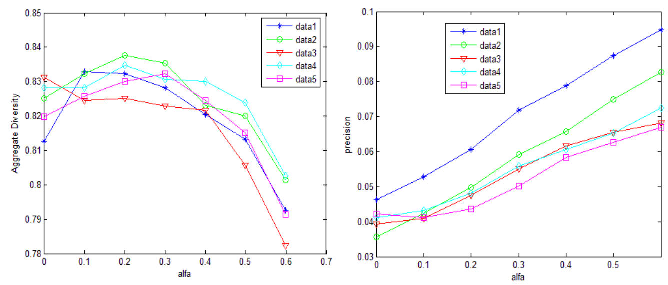

Figure 6. The effect of α to diversity and accuracy（β=0.3，N=30）

The novelty rate β is used to adjust the number of long tail items. As shown in Figure.7 (a), when β is 0.3(data2 to data5) or 0.2(data1), the aggregate diversity is highest. As shown in Figure.7 (b), with the increase of β, the recommendation accuracy tends to decrease but the variation is small. In order to get highest aggregate diversity, the β is set 0.3 in our SCF model.

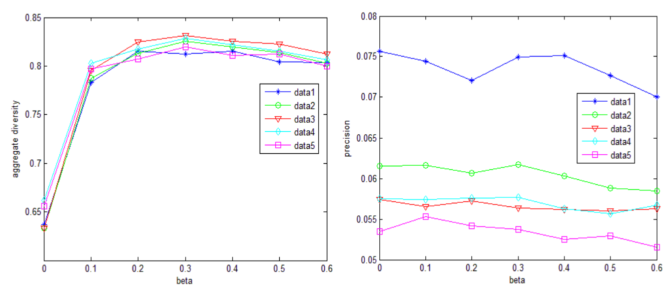

Figure 7. The effect of β to diversity and accuracy（α=0.3，N=30）

V. CONCLUSION

In this paper we analyze the deficiency of the traditional CF recommender system in the aspect of recommendation effect-that is the lack of recommendation diversity. In order to solve this problem, we first give clear definition of recommendation diversity. Then we present the item popularity metric, through which the items are divided into popular items, long tail items and items in between. Through analysis and comparison of experiments we find that user-based collaborative filtering technique is more inclined to recommend popular items and item-based collaborative filtering technique is more inclined to recommend long tail items. Inspired by the experimental results, we propose a synthetically collaborative filtering model by combining user-based and item-based collaborative filtering techniques. In this model, the control parameters can be inputted by users and the item popularity threshold can be dynamically retrieved based on the number of users in the system. The recommendation model we have proposed can accept users' feedback, so it meets the users' needs and provide more diverse recommendation result. Experiments using real-world rating datasets indicate that the proposed model effectively increases the recommendation diversity with little decrease in recommendation accuracy and surpasses the traditional collaborative filtering techniques.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (61033010), Natural Science Foundation of Guangdong Province (S2011020001182), Research Foundation of Science and Technology Plan Project in Guangdong Province (2009B090300450, 2010A040303004, and 2011B040200007), Research Foundation of Neusoft Institute of Information Technology, Nanhai (NN100507)

REFERENCES

[1] D. Goldberg, D. Nichols, B.M. Oki, D. Terry. "Using collaborative filtering to weave an information Tapestry". Communications of the ACM Special issue on information filtering, vol.35,No.12, pp.61-70, December 1992

[2] P.Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl. "GroupLens: An open architecture for collaborative filtering of net news". In: Proceedings of ACM Conference on Computer Supported Cooperative Work, Chapel Hill: ACM, New York, USA, pp.175-186,1994

[3] B. Sarwar, G. Karypis, J. Konstan, J. Riedl. "Item-based collaborative filtering recommendation algorithms". In: Proceedings of the 10th international conference on World Wide Web, pp. 285-295,2001

[4] G. Linden, B. Smith, and J. York. "Amazon.com recommendations: item-to-item collaborative filtering". IEEE Internet computing, vol. 7, no.1, pp.76-80, 2003

[5] G.Adomavicius, Y.O. Kwon . "Improving aggregate recommendation diversity using ranking-based technique". IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 5, pp. 896-911, 2012

[6] J. L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl. "Evaluating collaborative filtering recommender systems". ACM Transactions on Information Systems, vol. 22, no.1, pp.5-53,2004

[7] N.Hurley, M. Zhang. "Novelty and diversity in top-N recommendation—analysis and evaluation". ACM Transactions on Internet Technology, vol.10, no.4, pp.14:1-14:30, 2011

[8] T. Aytekin, M. O. Karakakaya. "Clustering-based diversity improvement in top-N recommendation". Journal of Intelligent Information Systems, pp.1-18, 2013

[9] M. Zhang, N. Hurley. "Statistical modeling of diversity in top-N recommender systems". In: IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, pp. 490-49, 2009

[10] C. N. Ziegler, S. M. Mcnee, J. A. Konstan, G. Lausen. "Improving recommendation lists through topic diversification". In: Proceedings of the 14th international conference on World Wide Web, pp 22-32, 2005

[11] G. Adomavicius. "Maximizing aggregate recommendation diversity: a graph-theoretic approach". In: Proceeding of the first Workshop on Novelty and Diversity in Recommender System, pp 3-10, 2011

[12] J. L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl. "Empirical analysis of design choices in neighborhood-based collaborative filtering algorithms", Information Retrieval, vol.5, no.4, pp.287-310, 2002

[13] H. Ma, I. King, R. M. Lyu. "Effective missing dataprediction for collaborative filtering". In: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, pp 39-46, 2007

[14] J. Wang, J. Yin. "An optimized item-based collaborative filtering recommendation algorithm". Journal of Chinese Computer Systems vol.31, no.12, pp. 2337-2342, 2010

[15] B. Sarwar, G. Karypis, J. Konstan, J. Riedl. "Application of dimensionality reduction in recommender system—a case study". . In: Proceedings of ACM Web KDD Workshop, 2000

[16] J. S. Breese, D. Heckerman, C. Kadie. "Empirical analysis of predictive algorithms for collaborative filtering". In: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, pp.43-52, 1998

[17] L. Lu, M. Medo. C. H. Yeung, et al. "Recommender aystems" . Physics Reports, 519, pp.1-49, 2012

[18] http://movielens.umn.edu/