

A research to improve collaborative filtering recommendation algorithm oriented barter electronic trading platform

Ying Liu, Li Zhou, Jie Huang

Department of Computer and Technology, Hangzhou Dianzi University,
310018, Hangzhou, Zhejiang, China
ayinghaha@163.com, july26@hdu.edu.cn, huangjie@hdu.edu.cn

Abstract – This article describes difference between modern barter and traditional e-commerce, makes some appropriate improvement to the user-based collaborative filtering recommendation under the background of barter, introduces a reference range based on users' location and the volume of commodity to improve the recommendatory efficiency and quality, enables it to be better applied in the modern barter system.

Index Terms –Modern barter trading, users' location, volume of commodity, reference range, the collaborative filtering recommendation

1. Introduction

Barter used to be a significant way to trade. It vanishes from business owing to limitations like information exchange and settlement method [1]. But in recent years, these limitations have been overcome due to the high-speed internet. The barter revitalizes and the concept of modern barter springs up.

In the traditional e-commerce transactions, it's hard for customers to find their favoured goods in a short time. So the personalized recommendation system becomes popular. However, it cannot obtain the desired effect if the traditional collaborative filtering recommendation is applied directly to the barter trading platform. Modern barter is different from traditional e-commerce. In order to obtain a satisfying recommendatory result, much needs to be done like combining circumstances of modern barter with the personalized recommendation system.

2. Summary Of The Traditional Collaborative Filtering Recommendation

After nearly 20 years of research and development, recommendation system is equipped with a relatively complete theoretical basis and framework [2]. The researchers provided many kinds of recommendation algorithm. The collaborative filtering is one of the most widely used algorithms with the best effect in practical use.

Collaborative filtering recommendation was born in 1992. It was put forward by Goldberg [3] and was applied to the Tapestry system. As a rudiment of collaborative filtering recommendation, the system showed a new kind of recommended ideas but there are a lot of shortages of technology. Followed there is a collaborative filtering recommendation system based on score, such as GroupLens [4] recommend news and movies. This automation system of

collaborative filtering learn the user's interest through calculating the similarity among users and needn't to search the information content of users. By the time, Users' hidden interests can be found. So it attracts more and more attention from the researchers and more and more widely used in the field of recommendation.

The current collaborative filtering recommendation algorithm mainly divided into two categories: The collaborative filtering recommendation algorithm based on the user and the collaborative filtering recommendation algorithm based on the project [5]. The former is the earliest emergence of collaborative filtering algorithms and also the most used. It computes similarity among users by the rows data (user data) of user - project evaluation matrix and the latter computes similarity among projects by the columns data (project data) of user - project evaluation matrix. Two kinds of collaborative filtering algorithms keep user ratings matrix as a data base and get recommend results by calculating the similarity [6]. At present, most of the collaborative filtering algorithms are on the basis of these two kinds of collaborative filtering algorithms and especially based on the user-based collaborative filtering.

Although these recommendation algorithms have been widely applied in the traditional e-commerce, it needs to make appropriate improvements to get satisfactory recommendatory effect in modern barter trading.

3. Improved Collaborative Filtering Recommendation Algorithm

Different from the traditional e-commerce, two important factors need to be considered in modern barter, that is, users location and commodity value. In the traditional filtering recommendation, these two factors have not been considered seriously. But in modern barter deal, these two factors may influence the success rate of barter transaction. Therefore this research focuses on analysis of circumstances of modern barter and appropriate improvement about recommendatory algorithm based on these factors. The algorithm is mainly divided into two parts: algorithm input and algorithm implementation. The whole process of algorithm is shown in figure 1 below.

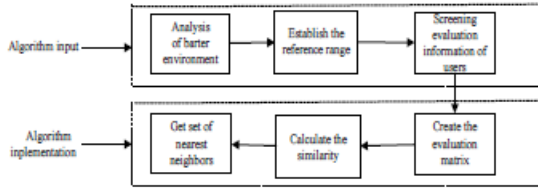


Figure 1 The process of algorithm implementation

A. The analysis based on barter environment

Users' location refers to users' geographic location. In the modern barter, user location will directly influence barter. At first, finding out a closer object will reduce the cost of goods transport and shorten the trading hour which can largely affect the exchange intention of users. Second, goods of barter may be second-hand product whose quality is prior. If users are nearby, they can inspect the goods quality as they want before trading. Finally, the geographical position information is easy to get according to the user's IP position.

Commodity value is goods value matching degree to exchange. Whether the commodity value can match may also influence exchange intention of users. In the process of trading, when bartering, the user expects to get a relatively equal swap in his subconscious. Therefore, commodity with similar values is easy to raise the interest of both parties. If the value gap between the two goods is too wide, it will cost more for users, which brings inconvenience for the transaction. The value of goods is filled in by the user and properly modified by the mediation. It will be fixed with the users' agreements and related data getting is also very easy.

B. The reference of scope based on environmental factors

Therefore, a reference range can be provided to the existing recommend algorithm combining these two elements. Specific ideas are shown in figure 2 below. In the rough sketch, green square represent for the commodity for the user who wants to be recommended, there are many other items around it based on distance, other squares represent for goods that value matches, triangles are not matching. There are three circles which represent for three different scopes and to choose the best range for reference range. In the reference range, most matching products can be got and too big or too small are not appropriate.

The establishment of reference range can be calculated and determined by the weight ratios of the two elements. Specific steps are as follows:

- 1) Distributing a lot of concentric circles according to the distance of the user's location. For instance there is a new scope every 10 km (specific numerical values as the case may be). The range of weight ratio decreased with distance.
- 2) Creating a group for each range and all users' data within each range are stored in it. There are user's geographic location and the value of goods to trade.
- 3) Goods of each group are matched in the quantity with the target user. The weight of goods decreases as the differences between the matching values.

4) Calculating weights of goods in each group and multiplied by the weight ratio of the users. Accumulating and then multiplied by the weight of the range. The last result is the weight of each group.

5) Comparing weights of all groups and the maximum value represents the reference range.

Specific calculation method can be represented by the following formula:

$$I_i = w_i \sum_{j=1}^m a_j b_j$$

, i represents different scope, j

represents quantity of commodity in range i , I_i represents weight of different scope, w_i represents the weight ratio of different range, a_j represents user's weight ratio of each goods, b_j represents weights of goods.

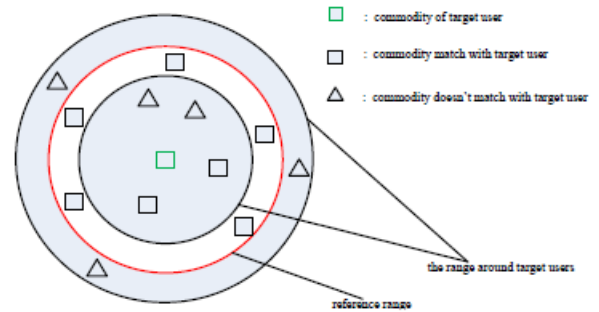


Figure 2 The establishment of reference range

C. Screening evaluation information of users

After the work ahead, the reference range can be used in collaborative filtering recommendation method and make some improvements on recommendation systems to make it can be applied in the modern barter. The introduction of reference range is to improve the input of algorithm. By the former description of the collaborative filtering recommendation, it can be seen that the user rating matrix form is the key. In the past, this matrix is by looking up the data in all associated with the user in the system or project to conclude in either user-based collaborative filtering recommendation or collaborative filtering recommendation algorithm based on the project. As the site scale expands unceasingly, visitors in the recommendation system and the dimensions of the attributes are becoming more and more complex [7], these make the user ratings, evaluation of the behaviour about data collection become more difficult which lead to a recommendation system efficiency and recommend quality seriously decline. The introduction of the reference range just could change this situation.

After determining the reference range, all data within the scope of reference can be selected according to the stored information of each group. Then only the similarity of each user who is near the target user within the reference range should be calculated to find the closest neighbour set with target users instead of calculating with all the users in the whole system. Thus greatly reduces the workload and improves the efficiency of recommendation.

D. The process of collaborative filtering algorithm based on the improvement of reference

Collaborative filtering recommendation algorithm based on the user is currently one of the most common in the application of personalized recommendation technology, Its main idea is to first find users who have the same hobby with target users and then recommend projects they are interested in to target user. If one user is adjacent to target user, the hobby of target user can be predict according to its interest [8].

The process can be divided into the following three steps after adding reference range into the algorithm:

1) Build user-project evaluation matrix

Assumed that the number of users in reference range is M and the number of goods is N. There is a matrix of M*N to represent the evaluation situation, element r_{ij} in the matrix represents the user i' interest in the j project [9]. Matrices are shown in figure 2 below:

$$\begin{bmatrix} r_{11} & \dots & r_{1i} & \dots & r_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ r_{i1} & \dots & r_{ii} & \dots & r_{in} \\ \dots & \dots & \dots & \dots & \dots \\ r_{m1} & \dots & r_{mi} & \dots & r_{mn} \end{bmatrix}$$

Figure 3 The evaluation matrix

2) Calculating the nearest neighbors

In order to find the nearest neighbour project of target project, the similarity among projects must be calculated. This is a key part of the collaborative filtering recommendation algorithm based on the user, there are three methods [10]:

① cosine-based similarity

This is the most basic method of similarity calculation, N project according to the user of the evaluation can be seen as the N dimensional space vector. Similarities between two users can be obtained by computing the vector Angle cosine:

$$sim = \cos(\vec{i}, \vec{j}) = \frac{\vec{j} \cdot \vec{i}}{\|\vec{i}\| \times \|\vec{j}\|}$$

② related similarity

This method calculates the similarity on the basis of a common score, setting I_{ij} is a common item set of user I and user j, the similarity showed as follow:

$$sim = \frac{\sum_{c \in I_{ij}} (r_{ic} - \bar{r}_i)(r_{jc} - \bar{r}_j)}{\sqrt{\sum_{c \in I_{ij}} (r_{ic} - \bar{r}_i)^2} \sqrt{\sum_{c \in I_{ij}} (r_{jc} - \bar{r}_j)^2}}$$

r_{ic} represents the score of user I to project c, \bar{r}_i represents average score of user I to project and \bar{r}_j represents average score of user j to project.

③ Modified cosine similarity

The two methods of similarity calculation ahead don't consider to different users' evaluation scale problem, it is improved in the revised method and the set of users I and j are taken into accounted , expression is as follow:

$$sim = \frac{\sum_{c \in I_{ij}} (r_{ic} - \bar{r}_i)(r_{jc} - \bar{r}_j)}{\sqrt{\sum_{c \in I_i} (r_{ic} - \bar{r}_i)^2} \sqrt{\sum_{c \in I_j} (r_{jc} - \bar{r}_j)^2}}$$

r_{ic} represents the score of user I to project c, \bar{r}_i represents average score of user I to project and \bar{r}_j represents average score of user j to project.

3) Produce recommendation

After finding out the nearest neighbours of target user through the above three methods, the project evaluation set can be arranged to get. Then put the items in descending order, finally the located in the front row of data are the outcomes of the recommendation system that users can intuitively feel. Combining collaborative filtering algorithm and reference range to apply in the modern barter has important significance to improve the efficiency and quality of recommendation.

4. The Experimental Results

The experimental data in this article is obtained in a barter trading platform, the Eeba Net. The experimental data is divided into four groups and they are recommended hit ratio obtained by testing 4 times under the premise of traditional recommendation algorithm(abbreviated TRA), reasonable reference range provided in algorithm(abbreviated RRA), too small reference range provided in algorithm(abbreviated SRA), too large reference range provided in algorithm(abbreviated LRA). Hit rate refers to the goods provided by the recommendation system can cause the risk of user interest and it directly reflects the users' satisfaction to recommendatory effect. The Specific experimental data is as follows in figure 4 and polygonal line rendering is shown in figure 5.

hit rate	TRA	RRA	SRA	LRA
first time	0.57	0.68	0.15	0.24
Second time	0.42	0.75	0.24	0.15
third time	0.67	0.82	0.17	0.18
fourth time	0.48	0.71	0.31	0.26

Figure 4 Specific experimental data

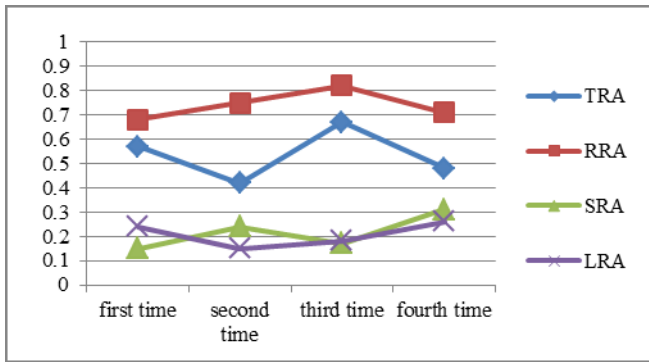


Figure 5 Polygonal line renderings

It can be seen clearly according to the figures above: the recommendatory hit rate is between 40% and 70% without reference range; when the reference range is too big or too small the recommendatory effect has not been improved and got even worse, the hit rate under 30% by the time; but after the reasonable reference range is selected, recommendatory effect is truly improved and the hit rate is more than 70%.

5. Conclusion

Recommendation system in traditional e-commerce has become increasingly mature, combining other areas with recommend technology will be a research hot spot as constant expansion of application domain and it has very broad prospects in application. This article describes difference between modern barter and traditional e-commerce, introduces a reference range under the background of barter to make some improvement to the user-based collaborative filtering recommendation and enable it to be better applied in the modern barter system. Therefore providing the reference

range under the background of barter has improved the input part of collaborative filtering recommendation algorithm. Combined with modern barter has important significance to improve the recommendatory efficiency and quality.

Acknowledgment

This work is supported by the Important Science & Technology Specific Projects of Zhejiang Province of China (2010 C13022).

References

- [1] Jing Wang, "The application of modern barter industry based on the improved Apriori algorithm," *Financial Perspective*, pp. 122-124, March 2012.
- [2] Lei Ren, "The key technology research of recommendation system," East China Normal University, pp. 1-11, 2012
- [3] Goldberg D, Niehols D, Oki B M, et al, "Using Collaborative Filtering to Weavean Information Tapestry," *Communications of the ACM*, pp. 61-70, 1992.
- [4] Resnick P, Iakovou N, Sushak M, et al, *GroupLens: An open architecture for collaborative filtering of netnews*, Press, pp. 175 - 186, 1994
- [5] Sobecki J, "Implementations of Web-based Recommender Systems Using Hybird Methods," *International Journal of Computer Science & Applications*, vol.3, pp. 52-64, 2006
- [6] Xiaoyuan Su, Taghi M. Khoshgoftaar, "A Survey of Collaborative Filtering Techniques," *Advances in Artificial Intelligence*, Vol.7, 2009
- [7] Goldberg K, Roeder T, Gupta D, et al, "Eigentaste: A Constant Time Collaborative Filtering Algorithm," *Information Retrieval*, vol.4, no.2, pp. 133-151, 2001
- [8] Wei Hu, "The application of collaborative filtering algorithm in the recommendation system," *Computer Age*, vol.11, pp.15-20, 2009
- [9] Peiyong Xia, "The collaborative filtering algorithms in personalized recommendation technology," *Ocean University of China*, pp. 30-32, 2011
- [10] Hongwei Ma, Guangwei Zhang, Peng Li, "Collaborative filtering recommendation algorithm," *Small Miniature Computer*, vol.7, no.7, pp.1281-1287, 2009