A Collaborative Filtering Methods Based on Expected Utility

Tao Zhang¹, Kun Zhao¹,*, Xianhua Zhang¹ and Yan Zou¹

¹School of Information, Yunnan University of Finance and Economics, Kunming, China

*Corresponding author e-mail: 472365373@qq.com

Abstract. In recent years, there has been a variety of improved collaborative filtering methods, which promote accuracy of recommendation by analysing too much user or item information like individual characteristics, application scenarios and so on. This increases the time complexity of the algorithm and leads to low efficiency of recommendation. Therefore, how to improve the quality of recommendation through analysing a small amount of effective information has attracted more attention in the research of collaborative filtering methods. This paper proposes a collaborative filtering method based on expected utility from the perspective of the economic significance of user rating. Firstly, the method establishes the corresponding similarity criterion and calculates the similarity between users or items by scoring expectation. Then, the nearest neighbour set is calculated and the final prediction is got. Finally, our method is validated by the experimental results using movielens-1m data set. The results show that the collaborative filtering method proposed in this paper is simpler and more feasible compared with the traditional IBCF, KNN and other mainstream methods, and can improve the efficiency and quality effectively of recommendation only using less information. This work has significance both in theoretical research and practical application value.

1. Introduction

Collaborative filtering [1, 2] is the earliest and most widely used recommendation technology and a hotspot of network technology research. It is a kind of recommendation technology based on ratings, which uses the known users' ratings of items in the system (generally refers to various recommended objects, such as various commodities in the electronic commerce system, news, reports, articles and other network information in the news website, etc.) to generate recommendations by searching and predicting users' ratings of relevant items by the nearest neighbor. Collaborative filtering technology can solve the problem of "information overload" in network applications by providing personalized recommendations [3, 4]. Currently, many websites have successfully used collaborative filtering recommendation systems such as Amazon, Google and Yahoo.

In the research of collaborative filtering recommendation algorithm, improving the accuracy of prediction is the core goal of the algorithm research. From the collaborative filtering recommendation algorithm, we can see that the similarity and the information of relevant users or items used in the calculation of fitting score have a direct impact on the calculation of the predicted value, which is related to its accuracy. Therefore, how to understand the similarity between items or users, and on the basis of this, how to filter the data used in the calculation equation and how to improve the calculation method, constitute the basic starting point of collaborative filtering technology. This paper reviews the
literature at home and abroad, and finds that the improvement of collaborative filtering recommendation algorithm is mainly reflected in the improvement of similarity calculation model, and the basic ideas of improvement can be divided into four categories. The first kind is k-Nearest-Neighbor (KNN) [5, 6], which selects the first k items or users with the highest similarity as neighbors to calculate the prediction value according to the similarity calculated from the scoring data, and takes into account the influence of the selection of neighbors on the accuracy of the prediction results. The second type is model-based recommending technique. By using the corresponding data analysis methods, such as matrix decomposition [7, 8], principal components analysis (PCA), data mining [10, 11] and machine learning [12], the user's data analysis model is established and selected. The core of this method is to establish and select the data analysis model of users. The commonly used models are based on regression model [13, 14], association rule model [10, 11], Bayesian classification model [15, 16], support vector machine model [17], clustering algorithm [18, 19] and latent semantic model [20, 21]. The third type is based on the characteristics of user ratings recommendation methods, according to the behavior characteristics of user ratings to build a similarity model, such as Bobadilla et al. proposed a similarity measure based on singularity [15], which judges the similarity of users by analyzing some "different" ratings in the user history rating data. The fourth type is Hybrid Collaborative Filtering [9, 10]. This method improves the neighbor selection and similarity calculation method by using other information (including the individual characteristics of users and items) and other recommendation technologies, which can not only improve the accuracy of recommendation effectively, but also provide a broader research space for collaborative filtering technology, so it has been widely concerned. For example, the context-aware recommendation method is a typical hybrid filtering technique. According to Dey's definition of context, the context contains all the information that describes the entity state involved in the interaction between the user and the application, including, of course, the user in the recommendation system, the recommendation items, even the location and the application itself [11, 13].

At the same time, the author notes that in the use of expectations and other related concepts to improve collaborative filtering algorithm, although some scholars have introduced such concepts as information entropy, information gain, expectations to improve the algorithm, but its ideas basically belong to the category of mixed recommendation. Its information entropy, information gain, expectation and other related concepts are mainly used to analyze the individual characteristic information of users or items, in order to further improve the neighbor selection and similarity calculation method according to this information. For example, in Ref. [12], based on the context-aware model, the information gain theory is introduced to reduce the attributes of many context factors, so as to calculate the weights of different context attributes by extracting the important context information which has great influence on the recommendation results, and then to improve the similarity calculation method to improve the recommendation quality. In Ref. [16], based on the similarity of item attributes and item ratings, the information entropy is introduced to construct the balance factor, so as to deal with the similarity calculation of new items and non-new items differently, and to solve the cold start problem caused by new items. In Ref. [17], we use mathematical expectation to improve the traditional collaborative filtering method for the attribute information of the item, such as "style" and "age" of the movie.

Therefore, the above four improved technologies can effectively improve the accuracy of recommendation, but in the process of recommendation, we need to use a lot of information about individual characteristics of users and items, complex mathematical models and application scenarios, which greatly improve the algorithm time complexity and reduce the efficiency of recommendation. Therefore, how to improve the recommendation efficiency and quality while using less information has become an important issue in the research of recommendation algorithms, especially in the current context of large data applications, the study of this issue is of great practical significance. Therefore, this paper puts forward the research premise: Under the premise of only using the existing scoring data, we seek effective methods to improve the recommendation efficiency and quality, so that the algorithm can obtain high-quality recommendation results in a simple and efficient manner. Therefore, this paper introduces the concept of expectation and improves the collaborative filtering algorithm in the following two aspects: One is to replace the actual score and other information except the score
with the score expectation as the direct basis for calculating the similarity between items or users; The other is to establish a criterion to judge whether the two items are similar or not by the expected utility, which can be used as a direct basis for neighbor selection in the calculation of fitting score. Based on this, this paper proposes a collaborative filtering recommendation method based on expected utility.

2. Related Work

2.1. Collaborative filtering recommendation algorithm

The general form of collaborative filtering recommendation problem is: Let $U$ be the user set with $m$ users, and $I$ be the item set that may be recommended to users (such as books, movies, CDs, and other commodities) with $n$ items. All users of a recommendation system rate relevant items as $r_{ij}$, forming a $U \times I$ user-item matrix, as shown in Table 1. Collaborative filtering is based on these known ratings, predicting the ratings $r_{u,p}$ of unknown ratings $p$ by any user $u$.

| Table 1. User-item Matrix |
|---------------------------|
| User U | I | ...... | j | ...... | n |
| ...... | ...... | ...... | ...... | ...... | ...... |
| i | $r_{i,j}$ | ...... | $r_{ij}$ | ...... | $r_{i,n}$ |
| ...... | ...... | ...... | ...... | ...... | ...... |
| m | $r_{m,i}$ | ...... | $r_{mj}$ | ...... | $r_{m,n}$ |

Collaborative filtering can generate recommendations through both item-based and user-based approaches. The technology to generate recommendations by analyzing the similarity between items is called Item-Based Collaborative Filtering (IBCF), and the technology to make recommendations by analyzing the similarity between users is User-Based Collaborative Filtering (UBCF). Both IBCF and UBCF produce recommendations through two basic processes: One is to calculate the similarity between items or users; Secondly, the predicted value is obtained by calculating the fitting score according to the known score of the predicted user on the basis of the obtained similarity. Similarity and fitting score calculation constitute the two basic technical means by which recommendation system generates recommendations.

The traditional Pearson Correlation Coefficient, Cosine Similarity and Jaccard Coefficient are usually used to calculate the similarity. Taking the IBCF recommendation as an example, the Pearson correlation coefficient and cosine similarity equations for calculating the similarity of the two items $i$ and $j$ are given by Equation (1) and Equation (2) [18].

$$sim(i, j) = \frac{\sum_{u \in U'} (r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{u \in U'} (r_{u,i} - \bar{r}_i)^2 \times \sum_{u \in U'} (r_{u,j} - \bar{r}_j)^2}}$$ (1)

$$sim(i, j) = \frac{\sum_{u \in U'} (r_{u,i} \times r_{u,j})}{\sqrt{\sum_{u \in U'} r_{u,i}^2 \times \sum_{u \in U'} r_{u,j}^2}}$$ (2)

Where $U'$ is a user set that has scored both items $i$ and $j$. $r_{u,i}$ is the score of the user $u$ obtained for item $i$, $\bar{r}_i$ is the average of all the scores obtained for item $i$; $r_{u,j}$ is the score of the user $u$ obtained for item $j$ and $\bar{r}_j$ is the average of all the scores obtained for item $j$.

Predicted values are usually calculated by similarity fitting scoring equation. Equation (3) gives a simple fitting scoring equation commonly used in IBCF [13], which is used to calculate user $u$'s predictive scoring $p_{u,p}$ for item $p$.

$$p_{u,p} = \frac{\sum_{i \in U'} r_{u,i} \cdot sim(i, p)}{\sum_{i \in U'} sim(i, p)}$$ (3)
Where $I'$ is the item set evaluated by the target user $u$, $r_{u,i}$ is the known score of the target user $u$ on the item $i$, $i \in I'$.

2.2. Expected utility and scoring expectations

According to the economic explanation, the market is the place where commodities are exchanged. In the new era when network technology is widely used in various fields such as electronic commerce and information services, it can be considered that a recommendation system records the information related to the activities of goods and information services carried out by users in a specific market. Therefore, for the convenience of discussion, the items recommended to users in the recommendation system are collectively referred to as products, such as products may refer to websites, web pages, hotels, tourist attractions, and a variety of goods and services. These products and their potential audience and related activities constitute a specific "market".

From this perspective, a user's rating of a product can be understood as a measure of the utility of a product in a particular market to its user [5]. The higher a user's rating of a product, the stronger the product's ability to meet its needs, and hence the greater its usefulness. Based on this, this paper has the following definitions.

**Definition 1 Expected Utility**

From the point of view of product, the expected value of a product's possible score in the market is called the expected utility of the product; From the user's point of view, the expected value of a user's possible rating of an item on the market is called the user's expected utility.

In the case of using integral value scoring (assuming a 5-point system with $r=1, 2, 3, 4, 5$), the expected utility of the product and the expected utility of the user can be calculated according to the usual expected calculation method, as shown in Equation (4) and Equation (5), respectively.

$$E(i) = \sum_{r=1}^{5} r \times p_{i,r} \quad , \quad i \in I$$

$$E(j) = \sum_{r=1}^{5} r \times p_{j,r} \quad , \quad j \in U$$

Where $p_{i,r}$ is the probability that item $i$ obtains a score of $r$, and $p_{j,r}$ is the probability that user $j$ obtains a score of $r$, both of which are referred to as scoring probabilities. The scoring probability can be calculated according to the standardised likelihood ratio, that is, the frequency at which an item is obtained or a user gives a score $r$, that is, the proportion of the number of samples scored to the total number of samples, as shown in Equation (6) and Equation (7), respectively.

$$p_{i,r} = \frac{\text{Num}(r)}{\text{Num}(U_i)}, \quad r=1,2,3,4,5$$

$$p_{j,r} = \frac{\text{Num}(r)}{\text{Num}(I_j)}, \quad r=1,2,3,4,5$$

Where $U_i$ is a user set that has scored item $i$ and $I_j$ is a user set that has been evaluated by user $j$. $\text{Num}(U_i)$ is the total number of ratings for item $i$, i.e. The number of users who scored item $i$. $\text{Num}(I_j)$ is the total number of ratings given by user $j$, i.e., the number of items evaluated by user $j$. $\text{Num}(r)$ is the number of users whose score is $r$ in Equation (6) and the number of items whose score is $r$ in Equation (7).

Therefore, the expected utility of a item can be interpreted as the overall recognition of a product in the market. The expected utility of the product is large, which indicates that the overall acceptance of the product by the market users is high, so the overall score of the product is relatively high. Conversely, the expected utility is small, indicating that the overall acceptance of the product is low. Similarly, users' expected utility reflects a user's overall recognition of the products on the market. The expected utility is large, which indicates that the user has a high degree of acceptance of the products in the market. On the contrary, the expected utility is small, which indicates that the overall acceptance
of the products on the market is low.

**Define 2 Score Expectations**

Further, according to the adopted scoring system, the expected $1 \times p_{i,1}$, $2 \times p_{i,2}$, $\ldots \ldots$ of each score $r (r=1, 2, \ldots)$ that may be obtained for an item $i$ may be expressed as a vector $(e_{i,1}, e_{i,2}, \ldots)$, wherein $e_{i,r}=r \times p_{i,r}$ is referred to as the scoring expectation of the item $i$, which reflects the constituent characteristics of the expected utility of the item $i$. Likewise, similarly, the vector $(e_{j,1}, e_{j,2}, \ldots)$ formed by the expectation of each score given by the user $j$ is referred to as the scoring expectation of the user $j$. Similarity between rating expectations also constitutes a similarity relationship between items or users.

3. **Collaborative Filtering Recommendation Method Based on Expected Utility**

3.1. **Collaborative filtering recommendation algorithm**

Collaborative filtering recommendation, in fact, assumes that the user has a certain behavior tendency as the basic premise of recommendation generation. For example [13], UBCF generally assumes that users with similar interests to the target user will also share their interests and hold similar views and perceptions of the objects of their respective interest; IBCF assumes that the user's future interest in information will remain consistent with past interest in information. According to this idea, this paper further assumes that the user has the behavior tendency to keep consistent with the expected utility in the evaluation and selection of the item, that is, the user will give consistent or similar ratings to the items with the same or similar expected utility; For the items with large differences in expected utility, the scores are correspondingly different.

Based on this assumption and the above-mentioned understanding of the expected utility, the analysis of the expected utility of the target item in the market and the evaluation of the item which is close to its expected value constitute a way of user behavior in item evaluation and selection. From the above discussion, the following basic judging criteria are proposed:

1) The expected similarity of ratings constitutes a measure of similarity between items or users.

2) The similarity of two items or two users is significant if and only if their expected utility falls into a neighborhood at the same time. I.e.

$$
\text{sim}(i,j) = \begin{cases} 
\text{sim}'(i,j), & |E(i)-E(j)| \leq \varepsilon \\
0, & |E(i)-E(j)| > \varepsilon 
\end{cases}
$$

(8)

Where $\text{sim}'(i,j)$ is the similarity between two items or two users as expected by the score.

It can be seen from the above equations that if the neighborhood value $\varepsilon$ is too large, the above criterion is meaningless for neighbor screening. In the case of a 5-point system, the range of expected utility is $[1, 5]$. Therefore, if $\varepsilon = 4$, the neighbor items are selected and filtered out, all possible items are covered; If $\varepsilon = 3$, the $(4+5+5+5+4)/25=96\%$ of the possible items are covered when the expected utility of the item is evenly distributed over the period $[1, 5]$; If $\varepsilon = 2$, it covers $(3+4+5+4+3)/25=76\%$ of the possible items; If $\varepsilon = 1$, then $(2+3+3+3+2)/25=52\%$ of the possible items are covered, and nearly half of them are screened out.

For the convenience of discussion, the similarity based on rating expectation is denoted as REBS (rating expectation based similarity), and the neighbor based on expected utility is denoted as EUBN (expected utility based neighbor).

3.2. **Basic algorithms**

Based on the above discussion and taking the item recommendation as an example, the basic steps for predicting the user $u$'s rating of the item $p$ are as follows.

1) Calculating the expected utility and scoring expectation of each item in the system according to Equation (4), (6) and Definition 2, respectively.
2) Calculate the similarity between any two items in the system according to the similarity calculation equation (e.g., Equation (1) or (2)), and calculate the similarity between any two items in the system according to the score expectation.

3) For the item \( p \) to be predicted, according to the preset neighborhood value \( \epsilon \), selecting the item whose expected utility is in the neighborhood of the expected utility of the item \( p \) and the user \( u \) has scored as a neighbor, and recording the item as \( N_{u,p} \).

\[
N_{u,p} = \{ i | i \in I_u \cap |E(p) - E(i)| \leq \epsilon \}
\]  
(9)

Where \( I_u \) is the itemset rated by user \( u \), \( E(p) \) and \( E(i) \) are the predicted item \( p \) and the expected utility of the itemset rated by user \( u \), respectively.

4) According to the score of user \( u \) in \( N_{u,p} \) and its similarity with the predicted item \( p \), the predicted value of user \( u \) to \( p \) is calculated according to the fitted scoring equation (such as Equation (3)).

Wherein the steps (1) and (2) can be performed in a pretreatment and used repeatedly after one treatment.

4. Experiment

This paper is based on item similarity (IBCF). This is because, according to relevant studies, although the principles of IBCF and UBCF are the same, IBCF is generally superior to UBCF in terms of recommendation quality.

Both IBCF and UBCF, traditional collaborative filtering techniques generate recommendations through two basic processes: One is to calculate the similarity between items or users; Secondly, the predicted value is obtained by calculating the fitting score according to the known score of the predicted user on the basis of the obtained similarity. Similarity and fitting score calculation constitute the two basic technical means by which recommendation system generates recommendations.

4.1. Quality standard

MAE (Mean Absolute Error) and MSE (Mean Squared Error) are used as the quality standards to test the accuracy of the algorithm.

\[
MAE = \frac{\sum_{i,j} |p_{i,j} - r_{i,j}|}{n}
\]  
(10)

\[
MSE = \frac{\sum_{i,j} (p_{i,j} - r_{i,j})^2}{n}
\]  
(11)

Where \( n \) is the total number of items measured, \( p_{i,j} \) is the predicted value of item \( j \) by user \( i \), and \( r_{i,j} \) is the actual score thereof. The smaller the values of MAE and MSE, the higher the accuracy of the algorithm. For the test method, the quality of MAE can be further distinguished by MSE when the MAE values are close to each other.

4.2. Datasets

The experiment was carried out using all the data from the Movie Lens-1M dataset. The Movie Lens-1M dataset has 6,040 users rating 1,000,209 films on a five-point scale. The minimum number of films evaluated per user was 20 (out of a total of 86 users) and the maximum was 2,314 (out of a single user).

4.3. Experimental design

The basic idea of the experiment is to use the traditional IBCF method, KNN method and EUBN-based method to predict the ratings of each movie in the test set based on the traditional item scoring similarity and the REBS proposed in this paper, and then calculate the MAE and MSE of the predicted results for each method to compare their recommendation quality. The main experiments are shown in Table 2. In the experiment, the expected utility is normalized to \([0, 1]\), and the similarity neighborhood of the expected utility is also selected in this interval.
In the experiment, cosine similarity Equation (2) was used to calculate the similarity between items in two ways: the actual score and the expected score.

### Table 2. Experiment Description

| Method Name | Basis of Similarity Calculation | Neighbor Item Selection Method | Purpose of the Experiment |
|-------------|---------------------------------|--------------------------------|---------------------------|
| IBCF        | Item score                       | Generate recommendations for neighbors for all scoring items | The method is compare with the method EUBN in this paper |
| IBCF-KNN    |                                 | Choose the k items with the highest similarity to generate recommendations for neighbors, and take k=5, 10, 15,..., 50 for experiments | The k value of each of the two calculation modes of the similarity can be obtain under the condition of better recommendation quality |
| REBS-KNN    | Score expectation                | According to the set neighborhood value $\varepsilon$, the neighbor items are dynamically selected to generate recommendations according to the judgment criteria of this paper, and the recommendations are taken as $\varepsilon=0.05, 0.075, 0.1,..., 0.25$ for experiments | The value capable $\varepsilon$ of obtain a better recommend quality is obtained |
| EUBN        |                                 |                                |                           |

### 4.4. Analysis of experimental results

The experimental results for IBCF-KNN, REBS-KNN, and EUBN are shown in Figure 1, Figure 2, and Figure 3, respectively.

As can be seen from Figure 2, for IBCF-KNN, both MAE and MSE of the predicted result at k=15 are optimal; As can be seen from Figure 3, for REBS-KNN, when k=25, MAE and MSE of the prediction result are optimal; As can be seen from Figure 4, when EUBN is taken to be $\varepsilon=0.075$, the MAE of the prediction result is the smallest, and the MSE of the prediction result also reaches a lower level. Figure 4-a and Figure 4-b compare MAE and MSE of four methods, IBCF, IBCF-KNN, REBS-KNN, and EUBN, respectively (the horizontal axis is % of the number of test set samples). K=15 and 25 for IBCF-KNN and REBS-KNN, respectively, and $\varepsilon=0.075$ for EUBN.

Figure 4-a and 4-b show that that application of the KNN method to the conventional IBCF method and the recommendation based on similar score expectation proposed herein both significantly improve the quality of the recommendation; The EUBN method can further improve the recommendation quality by distinguishing whether the items are similar or not through the expected utility neighborhood. For both REBS-KNN and IBCF-KNN, although the MAE values of both
methods are approximately [0.72, 0.73], REBS-KNN is slightly lower than IBCF-KNN, and the quality improvement of REBS-KNN is significantly better than that of IBCF-KNN in terms of MSE (see Figure 4-b). In the experiment, when EUBN is $\varepsilon = 0.075$, it is the best of the four algorithms, and MAE is slightly higher than 0.71, which is about 0.01 better than that of REBS-KNN when $k = 25$.

5. Conclusion

According to the traditional similarity calculation and recommendation method, similarity has relativity, that is, any two items or users have a certain degree of similarity, the difference lies in the size of similarity. KNN method is actually a typical application of this understanding. In this paper, the expected utility is introduced and the criterion of judging similarity according to its neighborhood is given. The core of the criterion is that there is a qualitative difference between items or users in a certain sense, which can be defined quantitatively according to the expected utility. For example, the item with a lower expected utility is clearly not similar to the item with a higher expected utility.

This understanding is not only significant in theory, but also can effectively improve the quality of recommendations in a more convenient and easy way. For example, if $k$ neighbors ($N > k$) are selected from $N$ items, the KNN method needs to be processed $\Sigma_{i=1}^{k}(N - i) = (2N - k - 1) \times k / 2 \geq N$ times (when $k \geq 2$, $N > k$), while this method only needs to be processed $N$ times regardless of the neighborhood value. And its recommendation quality is not only better than KNN method, but also close to or even better than many mainstream methods. For example, the MAE of Slope One is about 0.72 [14], and the MAE based on the singular similarity model is only about 0.74 level [15]. Therefore, this method has high practical application value.

In addition, based on the expected utility, the ways to further improve the recommendation quality can also be explored. For example, the combination of users' expected utility and products' expected utility and the generation of recommendations based on the distribution information of ratings' expected utility constitute the future research direction.

6. Acknowledgments

This work was financially supported by the National Natural Science Foundation of China, under Grant No.71462036, and the Projections of Science Research Foundation of Yunnan University of Finance and Economics, under Grant No. 80059900226.

References

[1] Goldberg D, Nichols D, Oki BM, et al. Using collaborative filtering to weave an information Tapestry[J]. Communication of the ACM, 1992, 35(12): 61-70.
[2] Resnick P, Varian HR. Recommender systems[J]. Communications of the ACM, 1997, 40(3): 56-58.
[3] Borchers A, Herlocker J, Konstan J, et al. Ganging up on information overload[J]. Computer. 1998, 31(4):106-108.

[4] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6):734-749. [doi: 10.1109/TKDE.2005.99]

[5] Deshpande M, Karypis G. Item-based top-n recommendation algorithms. ACM Transactions on Information Systems, 2004, 22(1):143-177.

[6] Ouyang Y X, Xiong Z, Yuan M. The Effect of Similarity Support in K-Nearest-Neighborhood Based Collaborative Filtering[J]. CHINESE JOURNAL OF COMPUTERS, 2010, 33(8):1437-1445.

[7] Hernando A, Bobadilla J, Ortega F. A non-negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model[J]. Knowledge-Based Systems. 2016:188-202.

[8] Guan X, Li C T, Guan Y. Matrix Factorization With Rating Completion: An Enhanced SVD Model for Collaborative Filtering Recommender Systems[J]. IEEE Access. 2017:27668-27678.

[9] Deng Z J, Zhang F, Wang S P S. Shilling attack detection in collaborative filtering recommender system by PCA detection and perturbation [J]. Proceedings of 2016 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR). 2016:213-218.

[10] Zhang D, Wang X R, Wang D, et al. Research on Personalized Recommendation of Smart Phones Applications Based on Association Rules and Collaborative Filtering Algorithms[J]. Proceedings of the Twelfth International Conference on Management Science and Engineering Management. 2019: 379-391.

[11] Maryam Khanian Najafabadi, Mohd Naz'ri Maharin, et al. Improving the accuracy of collaborative filtering recommendations using clustering and association rules mining on implicit data[J]. Computers in Human Behavior. 2017: 113-128.

[12] Fu M S, Qu H, et al. A Novel Deep Learning-Based Collaborative Filtering Model for Recommendation System[J]. IEEE Transactions on Cybernetics. 2018:1-13.

[13] Nilashi, M, Jannach, D, et al. Clustering- and regression-based multi-criteria collaborative filtering with incremental updates. Information Sciences, 2015, 293, 235–250.

[14] S.Subbulakshmi, K.Ramar, et al. Web Service Recommendation Based on Semantic Analysis of Web Service Specification and Enhanced Collaborative Filtering[J]. Intelligent Systems Technologies and Applications. 2018:54-65.

[15] Jiang W J, Wei Q, et al. A Dynamic Bayesian Network Based Collaborative Filtering Model for Multi-stage Recommendation[J]. Advances in Fuzzy Logic and Technology 2017.2017:290-301.

[16] Luo C, Zhang B, Xiang Y, Qi M. Gaussian-Gamma collaborative filtering: A hierarchical Bayesian model for recommender systems[J]. Journal of Computer and System Sciences. 2017

[17] Zhou W, Wen J H, Gao M, et al. A Shilling Attack Detection Method Based on SVM and Target Item Analysis in Collaborative Filtering Recommender Systems[J]. Knowledge Science, Engineering and Management (KSEM 2015). 2015:751-763.

[18] Liu X J. An improved clustering-based collaborative filtering recommendation algorithm[J]. Cluster Computing-The Journal of Networks Software Tools and Applications. 2017, Vol.20(No.2):1281-1288.

[19] Urszula Kuzielewska. Collaborative Filtering Recommender Systems Based on k-means Multi-clustering[J]. Contemporary Complex Systems and Their Dependability. 2019:316-325.

[20] Qi X, Wu W, Huang Y, Huang, et al. Rating Correlated Topic Model: An Improved Latent Semantic Model for Collaborative Filtering. Journal of Computational Information Systems, 2014:7259–7267.

[21] Defu Lian, Yong Ge, Fuzheng Zhang, Nicholas Jing Yuan, Xing Xie, Tao Zhou, Yong Rui. Scalable Content-Aware Collaborative Filtering for Location Recommendation[J]. IEEE Transactions on Knowledge and Data Engineering. 2018, Vol.30(No.6):1122-1135.