2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)

# Collaborative filtering algorithm based on rating prediction and user characteristics

*Na Song*[1*]
*School of Computer Science and Technology*
*Qilu University of Technology*
*Jinan, Shandong*
*1028338294@qq.com*

*Qin Lu*[2*]
*School of Computer Science and Technology*
*Qilu University of Technology*
*Jinan, Shandong*
*54903172@qq.com*

*Zhihao Zhang*[3]
*School of Computer Science and Technology*
*Qilu University of Technology*
*Jinan, Shandong*
*156801551@qq.com*

***Abstract*—Collaborative filtering directly predicts potential favorite items of user based on user's behavior records. It is one of the key technologies in personalized recommendation systems. The traditional similarity measurement method relies on user's rating data in the case of data sparseness, which causes a decrease in the recommendation quality of recommendation systems. To solve this problem, this paper proposes a collaborative filtering algorithm based on item rating prediction and user characteristics. The first step is to select the k nearest neighbor sets of the item using the KNN algorithm, and then calculate the similarity between the items using the improved similarity measurement method, and initially predict the user's rating on the unrated item to improve the sparsity problem. The second step considers the user characteristics when predicting the similarity between users according to the item ratings. Finally, the algorithm combining item-based rating prediction and user characteristics is adopted to make recommendations for the user. The experimental results on MovieLens and Douban datasets show that the proposed collaborative filtering algorithm based on rating prediction and user characteristics can effectively improve the quality of recommendation system compared with the traditional algorithm.**

***Keywords*—collaborative filtering; rating prediction; user characteristics; KNN algorithm; sparsity**

## I. PREFACE

With the fast development of Web services [1] and related technologies, e-commerce platforms continue to grow, people have to spend a lot of energy to obtain the things which they really need from the massive data information, and the "information overload" problem is urgently needed to be solved. The recommendation system [2] came into being. It recommended items that meet the needs of user according to the characteristics of user' interests and purchase behaviors, and realized personalized services, thus efficiently completing search for information.

Collaborative Filtering Algorithm [3,4] is a recommended algorithm for successful application in personalized recommendation system. The basic idea [5] is: by getting the similarity between the basic user and the item rating of target user's interest, forming a neighbor set with the same interests, and finally selecting the nearest neighbor's rating item to make recommendations, that is, the target user. Ratings for unrated items can be approximated by the weighted average of recent item ratings for item.

The traditional collaborative filtering algorithm relies entirely on the user's rating of the item to calculate user's similarity. First, it looks for the $k$ nearest neighbors. The next step is to make recommendations that match the user's situation. However, with the increasing number of users and items in recommendation system, user rating data is very rare. Traditional similarity measurement method only relies on user evaluation data, which makes the accuracy of recommendation system lower. This paper proposes a collaborative filtering algorithm based on the combination of rating prediction and user characteristics. Firstly, the KNN algorithm [3] is used to select the k-nearest neighbor sets of item. The improved similarity measure formula is used to calculate the similarity between items, and user's rating on the unrated items is preliminarily predicted. Then, when calculating the nearest neighbor of the user, consider user characteristics, using item-based and user-based approach to achieve more accurate recommendations for users. Preliminary prediction of item ratings avoids the problem of the same value of all unrated items in the similarity measure (all 0) and the "multiple" and "no mode" issues in the population fill method, and obviously improves the recommendation accuracy of the recommendation system [6]. This paper conducted corresponding experiments on two real datasets: MovieLens [7] public dataset and Douban recommendation dataset. The experimental results show that proposed algorithm effectively predict user's actual rating and further improve the recommendation quality.

## II. RELATED WORK

For the recommendation system rating sparse problem, domestic and foreign scholars have done a great number of researches and proposed proposed various recommendation algorithms. Manotumruksa [8] used a deep collaborative filtering method based on context-aware site recommendation. By using explicit feedback (e.g. user ratings of the venue) to predict the user's rating of the venue, matrix factorization cannot accommodate sparsity issues. Polatidis and Georgiadis [9] proposed a recommended way to improve collaborative filtering, dividing Pearson Correlation Similarity (PCC) into multiple levels. This approach helps users make decisions by providing better quality recommendations, but ignores user profile information. [10] found that the accuracy of the recommendation results is due to the user's assessment of the product, which is a sparse problem in the collaborative filtering algorithm. They proposed a collaborative filtering recommendation algorithm that combines users and projects to

978-1-7281-4328-6/19/$31.00 ©2019 IEEE

DOI 10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00102

alleviate the problem and achieved a certain degree of effect. Kant and Mahara [11] proposed a hybrid approach that combines two algorithms based on user-based CF and item-based CF. The use of dual clustering technology to reduce the number of dimensions, aggregate all users and items into multiple groups, achieved a certain effect.

The above researches had the following shortcomings: 1) In the case that the dataset is very sparse [12], the user only needs to rely on user-item rating to accurately predict the rating of unrated item; 2) In calculation of the nearest neighbor of target user, the general score of the group of users with different characteristics is not considered [13], because the characteristics of user, such as age, gender and occupation, will have different views on the same thing. The user characteristic attribute [14] is stable. User characteristics are introduced when computing user similarity, and the interference of users who are not related to the target user in similar users is reduced. 3) Only consider on user-based or item-based collaborative filtering algorithms, the results will have some bias [15]. We selected the KNN algorithm [16] that suitable for classifying sparse events is used to find the nearest neighbor, and improved similarity measure formula is to fill in the empty data of the item-rating matrix. Then, user characteristics are introduced when computing user similarity. Finally, the item-based and user-based methods are linearly merged to make recommendations for users [17]. The algorithm in this paper effectively reduces the sparseness of the rating data and makes the recommendation results for users more accurate.

## III. KNN ALGORITHM AND IMPROVED SIMILARITY MEASUREMENT METHOD

Collaborative filtering recommendation's definition is as follows [18]: the user rating data can be represented by a $u \times i$ order matrix $A(u,i)$, the $u$ line is the user $u$, the $i$ column is the item $i$, and $R_{(m,n)}$ which is the intersection of the row $m$ and column $n$ represents the rating of the user m on item $n$. We can see the matrix in Table I.

First, the nearest neighbor set KNN(X) is obtained by the KNN algorithm [19], and the similarity between user $i$ and user $j$ is calculated using the improved similarity measurement, which is represented as $sim(i,j)$.

TABLE I. USER-ITEM RATING MATRIX

| | $Item_1$ | ... | $Item_p$ | ... | $Item_i$ |
|---|---|---|---|---|---|
| $User_1$ | $R_{1,1}$ | … | $R_{1,p}$ | … | / |
| ... | … | … | … | … | … |
| $User_n$ | $R_{n,1}$ | … | / | … | $R_{n,i}$ |
| ... | … | … | … | … | … |
| $User_u$ | / | … | $R_{u,p}$ | … | $R_{u,i}$ |

### A. Determining the nearest neighbor using KNN algorithm

KNN (K-nearest neighbor) algorithm [19], namely K-nearest neighbor algorithm. The algorithm proposed by Cover and Hart is a relatively mature algorithm in theory. The advantage is that the stability is good and the use is convenient. The basic idea about KNN algorithm: Calculate distance between the target user $x$ to be classified and each training user according to the distance function, select the smallest $K$ users from the target user to be classified as the K-nearest neighbors of $x$, and finally judge the $x$ nearest neighbors according to $x$ Category. This article uses the classic Euclidean distance [20] formula to classify users or items.

The paper denotes an $n$-dimensional user vector $I = \{i_1, i_2, \dots i_n\}$, and $J$ denotes an $n$-dimensional user vector $J = \{j_1, j_2, \dots j_n\}$. The distance between two user sets is represented by $d(I,J)$, as shown in equation (1).

$$d(I,J)=\sqrt{/\textstyle\sum_{k=1}^{n} (R_{ik}-R_{jk})/} \tag{1}$$

$U$ denotes an n-dimensional item vector, and $V$ denotes an $n$-dimensional item vector. The distance between two item sets is represented by as shown in equation (2).

$$d(U,V)=\sqrt{/\textstyle\sum_{k=1}^{n} (R_{uk}-R_{vk})/} \tag{2}$$

The pseudo code of KNN algorithm can be described as follows:

**Algorithm 1**：KNN algorithm implementation

**Input:** Training samples

**Output：** Predict the k neighbor sets closest to the target

1： Enter the test samples set $X = \{x_i \ / i = 1,2,3 \cdots, n\}$ and determine the value of the parameter $k$

2： **For** $i = 1,2,3 \cdots, n$

3： Calculate the *dist()*

4： **If** $i \leqslant k$

5： Put $x_i$ in the nearest neighbor set

6： **Else if**

7： **Delete** $x_i$

8： Select the $K$ points with the smallest distance

9： **End if**

10： **End for**

### B. Improved similarity measurement method

In field of collaborative filtering, Pearson correlation coefficient [21] can measure similarities between users or between items. A similar user or collection of items largely affects the accuracy of the final predicted rating. To describe the similarity of users or item collections more accurately, we use the improved Pearson correlation coefficient as the similarity measure to calculate users' similarity. Its value is proportional to the similarity between different users [22].

The measurement method [21]: the item set jointly rating by user $i$ and user $j$ is denoted as $I_{ij}$, and the similarity $sim(i,j)$ between user $i$ and user $j$ is measured by the improved and the similarity $sim(i,j)$ between user $i$ and user $j$ is measured by improved Pearson correlation coefficient as equation (3) :

$$sim(i,j)=\frac{\Sigma_{c\in I_{ij}}(R_{i,c}-\bar{R_i})(R_{j,c}-\bar{R_j})*d(I,J)}{\sqrt{\Sigma_{c\in I_{ij}}(R_{i,c}-\bar{R_i})^2}\sqrt{\Sigma_{c\in I_{ij}}(R_{j,c}-\bar{R_j})^2}*d(I,J)} \tag{3}$$

Where $R_{i,c}$ is the rating of user $c$ for item $c$, $R_{j,c}$ represents the rating of user $j$ for item $c$, and $\bar{R_i}$ and $\bar{R_j}$ denote average rating of user $i$ and user $j$ on item.

Similarly, we can use the size of the improved Pearson correlation coefficient to represent the degree of similarity between items. The larger the value, the greater the similarity between items. The measurement method: Let the user set that jointly evaluates item $u$ and item $v$ be represented as $U_{uv}$, the similarity *sim(u,v)* between item u and item v is measured by improved Pearson correlation coefficient as equation (4) :

$$sim(u,v)=\frac{\Sigma_{d\in U_{uv}}(R_{d,u}-\bar{R_u})(R_{d,v}-\bar{R_v})*d(U,V)}{\sqrt{\Sigma_{d\in U_{uv}}(R_{d,u}-\bar{R_u})^2}\sqrt{\Sigma_{d\in U_{uv}}(R_{d,v}-\bar{R_v})^2}*d(U,V)} \tag{4}$$

Where $R_{d,u}$ is the rating of user $d$ for item $u$, $R_{d,v}$ represents the rating of the user $d$ for the item $v$, $\overline{R_u}$ and $\overline{R_v}$ denote the average ratings of the item $u$ and the item $v$.

## IV. COLLABORATIVE FILTERING ALGORITHM BASED ON RATING PREDICTION AND USER CHARACTERISTICS

In this section, we will introduce the collaborative filtering algorithm based on rating prediction and user characteristics. We follow the following three steps to complete [23]: (1) Collecting data of user preference, rating information and user characteristics are obtained through the dataset; (2) Determine the nearest neighbor set, and using the improved Pearson correlation coefficient to get similarity after classifying with KNN algorithm; (3) Completing the recommendation, and predicting final rating based on the nearest neighbor collection.

### A. User characteristics

In MovieLens dataset and Douban recommendation dataset, information such as user profile information, user ratings, and movie types are included. This basic information helps predict user interest and better personalizes the user [24]. To improve the recommendation quality of the recommendation algorithm, we considers important user characteristics [25] when calculating user similarity, and selects three basic characteristics of gender, age and occupation. We considers making a movie recommendation for the user as an example, to improve the collaborative filtering algorithm, and improve recommendation accuracy.

In our real life, the choice of same things between different genders is significant. In terms of films, for men, most of the types of films they like are science fiction war films, while women are more inclined to romantic romance. In this paper, the gender male and female of users are represented by numbers 1 and 0 respectively, and the set: gender $=\{1,0\}$.

Age also plays a big role in people's choice of things. For example, people of different ages have different hobbies in movies. Children like cartoons, young people like romance films and science fiction films, and old people like war films. Therefore, when calculating the nearest neighbor of user, the characteristic attribute of user age must be considered. In this paper, by analyzing the differences in preferences of users of different ages in the dataset, and combined with social network age group classification based on deep learning [26], the user age is roughly divided into six stages. The user's age collection is $age=\{age_1, age_2, age_3, age_4, age_5, age_6\}$ . Where $age_i$ indicates the stage $i$ ($i\in[1,6]$), $age_1$ indicates under 16 years old, $age_2$ means 17~23 years old, $age_3$ means 24~35 years old, $age_4$ means 36~45 years old, $age_5$ means 45~60 years old, $age_6$ means 60 years old or older.

Occupations [20] can be divided into many types, and users of different occupations have different interests. There are 21 kinds of users in the dataset, such as administrators, artists, doctors, etc. The occupational collection is $occ=\{occ_1, occ_2, \cdots occ_{21}\}$ $occ_m$ indicates the occupation type $m$ ($m\in[1,21]$).

User's characteristic set is $\{d_1, d_2, d_3\}$, and d represents user's gender, age, and occupation. The value corresponding to each characteristic attribute of user $i$ is $\{a_1, a_2, a_3\}$, and value corresponding to each characteristic attribute of user $j$ is $\{b_1, b_2, b_3\}$, and similarity between user characteristic attributes can be obtained as a formula (5):

$$sim_t(i,j)=\frac{1}{1+\sqrt{\Sigma_{m=1}^{n}(i_m-j_m)^2}} \tag{5}$$

Where $i_m$ and $j_m$ represent the characteristic attribute $m$ of user $i$ and user $j$, respectively. This paper considers three characteristic attributes of gender, age, and occupation, and $m$ denotes characteristics' numbers, $m$=3 here.

### B. Finding the nearest neighbors

The recommendation quality is reduced due to the extremely sparse user rating data, this paper proposes a linear combination [27] of item-based rating prediction and user characteristic similarity method. The biggest characteristic of this paper is introduction of user characteristics in calculation of user similarity and linear integration of item-based and user-based methods.

First, the similarity between two items is calculated. The similarity $sim(u,v)$ of the item $u$ and the item $v$ is calculated by the formula (4) given in Section III, and target user's rating on unrated item $P_{item}(R_{i,u})$ can be preliminarily predicted, the predict formula is as follows:

$$\mathrm{P_{item}}(\mathrm{R_{i,u}})=\overline{\mathrm{R_u}}+\frac{\Sigma_{\mathrm{i\in KNN(u)}}\mathrm{sim(u,v)}*(\mathrm{R_{iv}}-\overline{\mathrm{R_v}})}{\Sigma_{\mathrm{i\in KNN(u)}}|\mathrm{sim(u,v)}|} \tag{6}$$

Where $R_{iv}$ denotes the rating of target user $i$ to nearest neighbor user $j$ of target item $u$, and KNN($u$) is the set of $k$ nearest neighbor items of target item $u$.

Then, the similarity formula (3) between the user characteristic attributes is added to the formula (1), and the improved user characteristic-based similarity $Sim'(i,j)$ is further calculated.

$$Sim'(i,j)=\frac{\sum_{c\in I_{ij}}(R_{i,c}-\bar{R}_i)(R_{j,c}-\bar{R}_j)*d(I,J)*sim_t(i,j)}{\sqrt{\sum_{c\in I_{ij}}(R_{i,c}-\bar{R}_i)^2}\sqrt{\sum_{c\in I_{ij}}(R_{j,c}-\bar{R}_j)^2}*d(I,J)} \tag{7}$$

The target user's rating $P_{user}(R_{i,u})$ for the unrated item can be initially predicted.

$$P_{user}(R_{i,u})=\overline{R}_i+\frac{\sum_{j\in KNN(i)} Sim'(i,j)*(R_{ju}-\overline{R}_j)}{\sum_{j\in KNN(i)}|Sim'(i,j)|} \tag{8}$$

Where $R_{ju}$ is the rating value of the nearest neighbor $j$ of target user to item $v$, KNN($i$) is the set of $k$ nearest neighbor users of target user $i$.

## C. Final recommendations

Now, we linearly combined the item-based rating prediction and the user characteristics, the final rating [28] is calculated, and recommendations are generated based on the size of rating. The formula (6) is combined with the formula (8) by dynamic balance parameter $\alpha$.

$$P(R_{i,u})=\alpha*P_{item}(R_{i,u})+(1-\alpha)*P_{user}(R_{i,u}) \tag{9}$$

In the formula, α is a dynamic balance weight with a range of [0,1]. The magnitude of $\alpha$ determines the degree of dependency on item-based ratings and user-based ratings.

# V. EXPERIMENTS

## A. Datasets

Through experiments, we verify the performance of the collaborative filtering algorithm based on user characteristics and item rating prediction proposed in this paper. The dataset comes from the MovieLens site are used. The dataset provide by Douban are used. The MovieLens dataset is divided into 3 parts by size. This paper selects the MovieLens-1M dataset. The number of movies rated by each user in the data set exceeds 20, and the ratings were divided into five values (1, 2, 3, 4, 5). The value is proportional to the user's preference [29]. The Douban dataset includes two datasets of books and movies. We use the movie dataset. The specific data used in the experiment is shown in Table II.

TABLE II. Experimental data analysis

| | MovieLens | Douban |
|---|---|---|
| Total number of users | 6040 | 1000 |
| Total number of movies | 3900 | 37252 |
| Total rating | 100209 | 372200 |
| Sparse level | 0.937 | 0.991 |

## B. Metrics

Predictive accuracy [30] is an important indicator to recommend the recommended ability of an algorithm. In this paper, the mean absolute error (MAE) and root mean square error (RMSE) are used to measure the accuracy of recommendation results. The MAE reflects difference between predicted value and true value, while RMSE represents a measure of goodness of fit. The smaller the value of these two experimental evaluation indicators, the higher the accuracy of the recommendation.

Assuming that the original rating set is $R=\{r_1, r_2, \cdots, r_n\}$, and predicted rating set is $P=\{p_1, p_2, \cdots, p_n\}$, then the calculation formula of MAE is as follow in equation (10):

$$MAE=\frac{\Sigma_{i=1}^{N}|r_i-p_i|}{N} \tag{10}$$

The formula for calculating RMSE is shown in equation (11):

$$RMSE=\frac{\sqrt{\Sigma_{i=1}^{N}(r_i-p_i)^2}}{N} \tag{11}$$

## C. Experimental results

### 1) The value of dynamic balance parameter α

The dynamic balance factor is the balance factor of our proposed collaborative filtering algorithm, which affects the quality of the recommendation results. The prediction was performed according to the improved algorithm to see the effect of the balance parameter $\alpha$ on MAE and RMSE. Figures 1 and 2 show the changes of MAE and RMSE with $\alpha$ on different datasets.

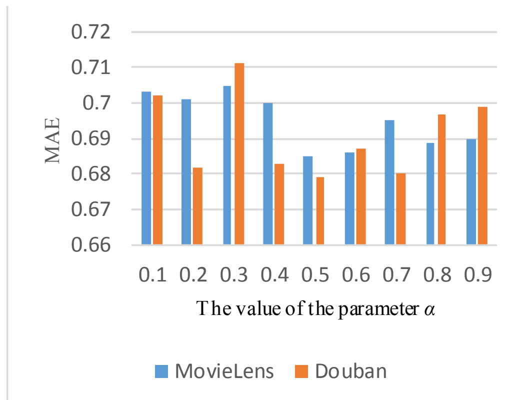


Fig. 1. The effect of parameter $\alpha$ on MAE on two datasets

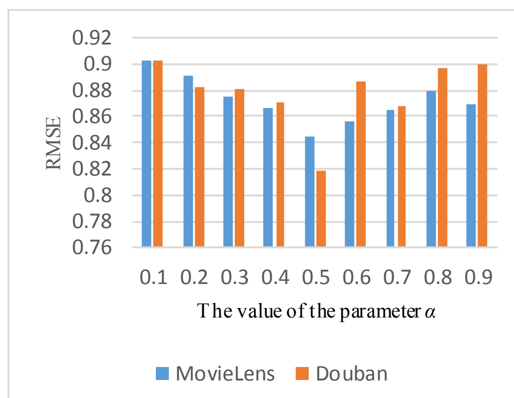


Fig. 2. The effect of parameter $\alpha$ on RMSE on two datasets

From the figures 1 and 2, we can see the effect of balance parameter $\alpha$ on MAE and RMSE. Whether it is on MovieLens dataset or Douban dataset, when $\alpha$=0.5, the recommended effect is the best.

The next step is to use the collaborative filtering algorithm based on user characteristics and item rating prediction to calculate impact on MAE and RMSE under different $K$ values of the KNN algorithm.

### 2) Comparison of algorithm MAE on different datasets

Our improved collaborative filtering algorithm based on user characteristics and project score prediction compares the two algorithms proposed in [9] and [11]. In face of extremely sparse user rating data, these two algorithms improve the prediction accuracy. In this paper, the KNN algorithm suitable for sparse event classification is used to determine the nearest neighbor of the target, and better results are obtained.

Firstly,we compares the improved algorithm with the MAE of the algorithm in the literature [9] and the literature [11] on the MovieLens dataset. We find that the improved algorithm proposed in this paper works best regardless of the number of nearest neighbors. The experimental results are shown in Figure 3:

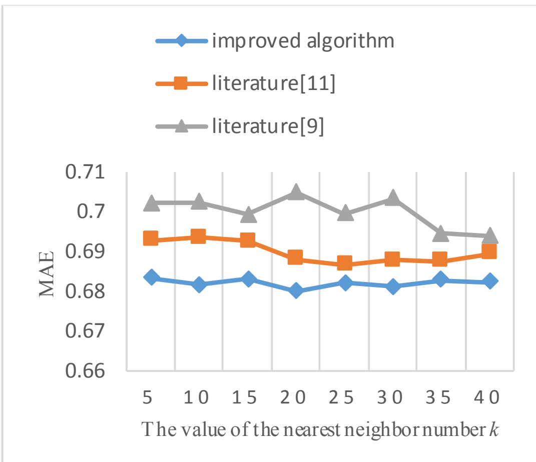


Fig. 3. Comparison of MAE of three algorithms on the MovieLens dataset

Similarly, this paper also carried out a comparative experiment of three algorithms on the Douban network dataset. We still choose the effect of three algorithms on MAE under different nearest neighbors. The experimental results are as follows:

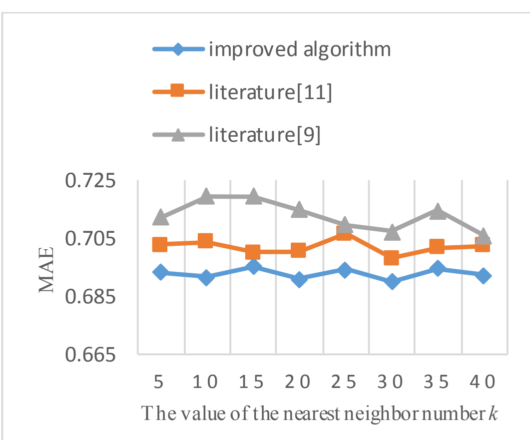


Fig. 4. Comparison of MAE of three algorithms on the Douban dataset

*3) Comparison of algorithm RMSE on different datasets*

On the MovieLens dataset, the improved algorithm is compared with RMSE of algorithm in the literature [9] and the literature [11]. Among them, the improved collaborative filtering method has larger square root error and the recommended effect is better. The experimental results are shown in Figure 5:

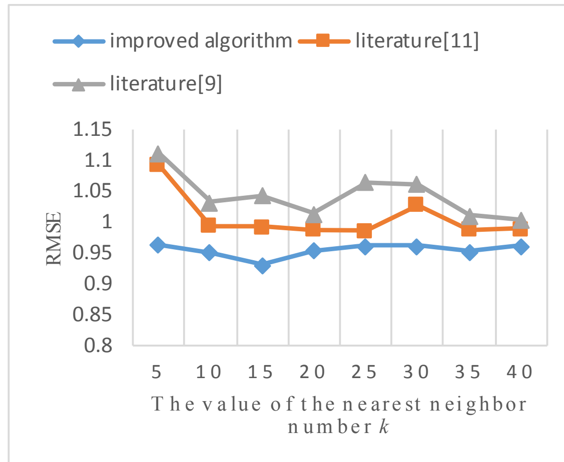


Fig. 5. Comparison of RMSE of three algorithms on the MovieLens dataset

On Douban dataset, the improved algorithm is compared with the RMSE of the algorithm in the literature [9] and the literature [11]. As with the MovieLens dataset, overall, the improved algorithm is more dominant. Our results are as follows:

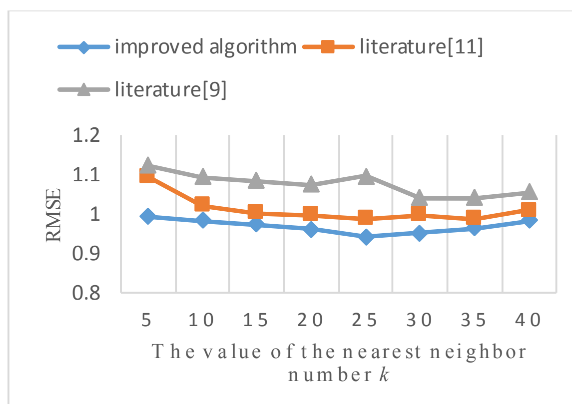


Fig. 6. Comparison of RMSE of three algorithms on the Douban dataset

Experiments has been carried out on the MovieLens public dataset and Douban recommendation dataset. We compare the two algorithms in the literature [9] and the literature [11]. It is obvious from the above figure that the improved collaborative filtering algorithm is excellent in the case of MAE and RMSE under different neighbor numbers. The two algorithms in the literature [9] and the literature [11] verify the effectiveness of the improved algorithm .

## VI. CONCLUSIONS AND FUTURE WORK

This paper first analyzes in depth the user's scoring data for the project is very rare, the current solution and its advantages and disadvantages, We propose a collaborative filtering algorithm based on rating prediction and user characteristics. Firstly, KNN algorithm is finding the set of neighboring users, and then similarity measurement method is used to calculate similarity between target item and similar item, and the prediction rating of item is initially completed to solve the sparsity problem. Then, according to different characteristics of the user, the degree of preference for item is different, and the similarity of target user's neighbor set is improved; Finally, the dynamic weighting parameter $\alpha$ is used to linearly combine the item-based and user-based methods to predict the final rating of the target user and generate recommendations. Through experiments, we get that the algorithm improves the recommendation quality to some extent. In the next work, We will study how to solve the problem that the user rating data sparsity is less than 0.6 and apply the improved algorithm to books, music and more.


## ACKNOWLEDGMENT

This work was supported in part by Shandong Provincial Natural Science Foundation, China (No.ZR2017LF021), Key Research and Development Plan Project of Shandong Province (No.2017CXGC0614).


## REFERENCES


[1] Chen F, Lu C, Wu H, et al. A Semantic Similarity Measure Integrating Multiple Conceptual Relationships for Web Service Discovery[J]. Expert Systems with Applications An International Journal, 2017, 67(C):19-31.

[2] Jannach D , Zanker M , Felfernig A , et al. Recommendation Systems: An Introduction[M]. Cambridge University Press, 2010.

[3] Hegde A , Shetty S K . Collaborative Filtering Recommendation System[J]. Esrsa Publications, 2015, V4(3).

[4] Herlocker J L , Konstan J A , Terveen L G , et al. Evaluating collaborative filtering recommendation systems[J]. ACM Transactions on Information Systems, 2004, 22(1):5-53.

[5] Li Weilin, Wang Chengliang, Wen Junhao. Collaborative filtering algorithm based on comments and scoring[J]. Journal of Computer Applications, 2017, 34(2): 361-364.

[6] Bouadjenek M R, Pacitti E, Servajean M, et al. A Distributed Collaborative Filtering Algorithm Using Multiple Data Sources[J]. 2018.

[7] Harper F M , Konstan J A . The MovieLens Datasets: History and Context[M]. ACM, 2015.

[8] Manotumruksa J . Deep Collaborative Filtering Approaches for Context-Aware Venue Recommendation[C]// the 40th International ACM SIGIR Conference. ACM, 2017.

[9] Polatidis N , Georgiadis C K . A multi-level collaborative filtering method that improves recommendations[J]. Expert Systems with Applications, 2016, 48:100-110.

[10] Liu F M , Li H X , Dong P . A Collaborative Filtering Recommendation Algorithm Combined with User and Item[J]. Applied Mechanics and Materials, 2014, 513-517:1878-1881.

[11] Kant S , Mahara T . Merging user and item based collaborative filtering to alleviate data sparsity[J]. International Journal of System Assurance Engineering and Management, 2016.

[12] Koren Y . Factorization meets the neighborhood: A multifaceted collaborative filtering model[C]// Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008. ACM, 2008.

[13] Fa-Sheng L , Ying H . A collaborative filtering recommendation algorithm based on user characteristic attribute and cloud model[J]. Computer Engineering & Science, 2014.

[14] Luo X , Zhou M , Xia Y , et al. An Efficient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems[J]. IEEE Transactions on Industrial Informatics, 2014, 10(2):1273-1284.

[15] Wang Juan, Xiong Wei, School of statistics. Collaborative Filtering Recommender System Based on Matrix Factorization and Its Application[J]. Statistics & Decision, 2019.

[16] Chen X , Liu X , Huang Z , et al. RegionKNN: A Scalable Hybrid Collaborative Filtering Algorithm for Personalized Web Service Recommendation[C]// 2010 IEEE International Conference on Web Services. IEEE Computer Society, 2010.

[17] Jian W, He J, Kai C, et al. Collaborative Filtering and Deep Learning Based Recommendation system For Cold Start Items[J]. Expert Systems with Applications, 2017, 69:29-39.

[18] Luo Xin, Ouyang Yuanxin, Xiong Wei, et al. Optimization of K-nearest neighbor collaborative filtering algorithm based on similarity support[J]. Chinese Journal of Computers, 2010, 33(8): 1437-1445.

[19] Shi B, Han L, Yan H. Adaptive clustering algorithm based on kNN and density[J]. Pattern Recognition Letters, 2018, 104:37-44.

[20] Zhang W, Tian F, An S, et al. Research on electronic nose system based on continuous wide spectral gas sensing[J]. Microchemical Journal, 2018, 140:1-7.

[21] Zeng Q, Dong F, Shi F, et al. Apparent diffusion coefficient maps obtained from high b val-ue diffusion-weighted imaging in the preoperative evaluation of gliomas at 3T: comparison with standard b value diffusion-weighted imaging[J]. European Radiology, 2017, 27(12):5309-5315.

[22] Tao Wang, Gunasekaran Manogaran, Minghui Wang. Framework for social tag recommendation using Lion Optimization Algorithm and collaborative filtering techniques[J]. Cluster Computing, 2019(2).

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

[24] Samundeeswary K, Krishnamurthy V. Comparative study of recommendation systems built using various methods of collaborative filtering algorithm[C]// International Conference on Computational Intelligence in Data Science. 2018.

[25] Jing Z, Peng Q, Sun S, et al. Collaborative filtering recommendation algorithm based on user preference derived from item domain characteristics[J]. Physica A Statistical Mechanics & Its Applications, 2014, 396(2):66-76.

[26] Guimaraes R G , Lopes Rosa R , De Gaetano D , et al. Age Groups Classification in Social Network Using Deep Learning[J]. IEEE Access, 2017:1-1.

[27] Bayraktar T. Expected Number of Real Roots for Random Linear Combinations of Orthogonal Polynomials Associated with Radial Weights[J]. Potential Analysis, 2018:1-13.

[28] Zhang Z, Kudo Y, Murai T. Neighbor selection for user-based collaborative filtering using covering-based rough sets[J]. Annals of Operations Research, 2017, 256(2):359-374.

[29] Sun Guangfu, Wu Le, Liu Qi, et al. Collaborative Filtering Recommendation Algorithm Based on Time Series Behavior[J]. Journal of Software, 2013(11):2721-2733.

[30] Zhang F , Gong T , Lee V E , et al. Fast Algorithms to Evaluate Collaborative Filtering Recommendation Systems[J]. Knowledge-Based Systems, 2016:S0950705115005079.