An E-commerce Personalized Recommendation Algorithm based on Fuzzy Clustering

Yan Hou[1], Shuling Yang[2]
[1]Management School of Jilin Normal University, Siping, China
[2]Foreign Language School of Jilin Normal University, Siping China

***Abstract*—The key to the performance of recommendation system lies in which recommendation technology to choose. Personalized recommendation technology has attracted the extensive attention of more and more merchants and researchers with its humanization and potential commercial value, and earliest and most widely used personalized recommendation technology. However, with the increasing number of users and commodities, the recommendation system is facing some problems, such as the difficulty of ensuring the real-time performance and the decline of recommendation quality. Many scholars put forward many solutions to the shortcomings of collaborative filtering. So as to obtain better recommendation effect. Based on different angles, this paper adopts fuzzy clustering recommendation algorithm. Using fuzzy clustering algorithm to increase population diversity can avoid the lack of diversity of adjacent users; The feature of flexible partition of fuzzy c-means clustering algorithm is used to reduce the dimension of data, so as to effectively solve the problem of recommendation quality degradation caused by data sparsity.**

***Keywords*—Fuzzy clustering, e-commerce, personalized recommendation algorithm, collaborative filtering, data set**

I. INTRODUCTION

With the development of e-commerce, online shopping has entered people's daily life. In the face of massive data, it is difficult for online shoppers to quickly and accurately select goods that really meet their needs [1]. How to help online shoppers quickly find goods or information, win customers' favor, maintain stable customer relations and promote consumption has become a new problem faced by e-commerce, and the recommendation system came into being [2].

With the information technology and the continuous deepening of Internet technology, e-commerce based on virtual economy is occupying a more and more important position in the development of social economy, and has gradually developed into the backbone of China's emerging industries [3]. However, with the gradual expansion of the scale of e-commerce market, more and more manufacturers have joined the ranks of e-commerce, followed by increasingly fierce competition among merchants [4]. How to occupy a place in the increasingly fierce e-commerce war and get a share of the huge profits is an important issue that all e-commerce businesses have to face. The expansion of e-commerce [5]. After clustering the objects with clustering technology, Users with high similarity gather in the same group. When using collaborative filtering algorithm, it is not necessary to find the nearest neighbor of the target user in the whole user space, and reduce the search space to several clusters close to the target user.However, in the face of such a variety of commodity information, how users can quickly and accurately select the commodities they need has also become the most concerned topic of users. In this context, e-commerce recommendation system came into being.[6]. For users, the e-commerce recommendation system can recommend personalized goods according to their preferences and interests, so that users can quickly and accurately select the goods they need in the massive commodity information, reduce the time wasted in the commodity selection process, and provide users with more humanized services. Therefore, whether for businesses or users, e-commerce recommendation system has great development prospects and application value.

Based on the fuzzy c-means clustering algorithm, this paper optimizes the algorithm for e-commerce personalized recommendation.The specific recommendation process is as follows: firstly, the improved fuzzy c-means clustering algorithm is used to cluster the user item scoring matrix to reduce the matrix dimension, then the to reduce the sparsity of the matrix, and then the user-based collaborative filtering algorithm is used to complete the project recommendation for the target user[7]. The recommendation algorithm proposed in this paper can quickly and accurately recommend interested products to users according to a user's historical score [8]

II. RELATED WORK

The concept of personalized recommendation technology was put forward in the 1990s. Since then, scholars from various countries have invested a lot of manpower and resources in this field and made considerable progress. After the concept of recommendation system was put forward, it has been widely used in e-commerce websites and created great economic value for them. In its publicity, Amazon claims that its core competitiveness is not to provide users with the lowest price of goods, but to provide users with the goods they need most. The application of recommended technology has increased Amazon's sales by nearly one-third [9].

The technical core of search engine service providers represented by literature [10] is based on search technology and exists as separate products. For recommendation technology, it exists as a functional part of e-commerce website system to improve user experience and improve the turnover of e-commerce websites. With the popularity of e-commerce, Nowadays, some companies provide recommendation services for e-commerce companies.

Literature [11] uses item similarity to recommend goods for users. After a user browses or purchases a product, Amazon keeps the user's browsing and purchase records, displays what other products the user browsed the product has purchased for the user at the next login, and displays these products to the user first.

Literature [12] provides users with two reading methods, namely, recommending the latest news messages and the content of interest to users according to the system analysis. The video recommendation on Google's video websites also adopts the same collaborative filtering recommendation based on items.

These algorithms have their own characteristics, and their applicable data sets and system environment are different. Many algorithms can improve the accuracy of recommendation to a certain extent. In addition, indicators to evaluate the effectiveness of recommendation system, such as diversity and novelty, have attracted people's attention.

Literature [13] proposed the personalized recommendation of user preference on the energy propagation model based on bipartite graph, and found that the results of the recommendation system are related to user degree and user preference

After paying attention to the diversity of recommendations, literature [14] proposed that the user commodity purchase relationship should be integrated into the network structure of energy transmission and heat conduction of bipartite graph to provide users with diversity recommendations. The user item label three part graph method based on network structure mentioned in the paper integrates content information into recommendations. The impact of user interest time migration is taken into account.

By analyzing that user emotion classification can affect the personalized recommendation results of e-commerce, literature [15] proposes to classify user emotion by using computer intelligence technology and machine learning technology, and then make personalized recommendation to users according to the emotion classification results. The system adopts Na ï ve Bayes, SVM, decision tree and other methods in IMDB and twitter, Comparative experiments are carried out on four data sets of hotel reviews and Amazon reviews to verify that the decision tree classification algorithm has better emotional classification effect, and the method is compared with personalized recommendation technology. The results effectively improve the effect of personalized recommendation.

In the design this paper first designs the system framework for the requirements of the system, and analyzes and designs the system level according to the requirements, so that the system can meet the functional requirements [16]. Then, each process analysis of the system is designed and described in detail, which is designed from the foreground and background of the system. Using fuzzy c-means clustering algorithm as one of the core parts of the system, the main data tables and related fields of the database are described and explained in detail. This paper describes and tests the implementation of the system from several of the most common and important functions of the system. The e-commerce recommendation system based on Personalized Recommendation designed and implemented in this paper can realize the main functions of the e-commerce system, analyze the information provided by users and the data obtained by the system, recommend personalized goods to users, and achieve the goal designed in this paper [17].

III. RESEARCH ON FUZZY CLUSTERING ALGORITHM

A. Introduction to fuzzy clustering

Clustering is a way of data processing. Its purpose is to divide the disordered data into several data groups according to the similarity, so that the data feature similarity in the same group is the highest. Clustering is an unsupervised classification. There is no rule limit before classification, but takes similarity as the only criterion for classification [18]. Clustering algorithm has been proposed for a long time. Clustering technology is generally divided into two categories: hard clustering technology and soft clustering technology. Soft clustering technology is the fuzzy clustering mentioned in this paper. Hard clustering is to completely divide the items into a certain class, which is "either 0 or 1" according to mathematics. However, for some samples with unclear membership, it may belong to both this class and another class. Using hard clustering to divide such data may cause problems. Fuzzy clustering can solve this problem well. It allocates the membership interval of 0 to 1 for each sample. Instead of completely assigning a sample to a class, it allocates the membership of different classes, which can cluster the sample data between classes more effectively [19].

This paper uses the fuzzy c-means clustering algorithm (FCM), deeply studies and analyzes the FCM clustering principle, introduces it into the recommendation technology, and completes the recommendation function together with the collaborative filtering technology.

B. Fuzzy c-means clustering algorithm

Fuzzy c-means clustering (FCM) algorithm is derived from the model of hard c-means clustering algorithm. The biggest difference is to assign a weight value m to the membership u ij. The mathematical reasoning process and clustering process of FCM are as follows:

FCM clustering initial function (1):

$$\min J_m(U,C)=\sum_{i=1}^{n}\sum_{j=1}^{c}u_{ij}^{m}d_{ij}^{2} \tag{1}$$

The relationship is as follows (2):

$$\sum_{i=1}^{n} = 1, 1 \le j \le n$$
$$u_{ij} \in [0,1], 1 \le j \le n, 1 \le i \le c$$
$$0 < \sum_{j=1}^{c} u_{ij} < n, 1 \le i \le c \tag{2}$$

In the initial function, the general weight factor M > 1. In order to find the optimal solution of the objective function, the reasoning process is as follows

In order to obtain the optimal solution, the constraint conditions of extreme value can be used , in $\sum_{i=1}^{c} u_{ij} = 1$ Under the condition of $\sum_{i=1}^{n}\sum_{j=1}^{c} u_{ij} d_{ij}^{2}$ Lagrange function can be constructed to solve the minimum value of. Let the Lagrange function be (3):

$$F = \sum_{i=1}^{n}\sum_{j=1}^{c} u_{ij}^{m} d_{ij}^{2} + \sum_{j}^{n} \lambda(\sum_{i}^{c} u_{ij} - 1) \tag{3}$$

because $\sum_{i=1}^{c} u_{ij} = 1$ therefore, the newly added part is 0, which does not affect the initial function.

In Lagrange function λ And Uij, the membership degree can be calculated as (4):

$$u_{ij} = \frac{1}{\sum_{r=1}^{c}\left[\frac{d_{ij}(k)}{d_{rj}(k}\right]^{\frac{2}{m-1}}}, \forall j, r, d_{rj}(k) > 0 \tag{4}$$

Where k represents the k-th iteration

The mathematical formula for calculating the k-th clustering center is as follows (5):

$$C_i(k+1) = \frac{\sum_{j=1}^{n} u_{ij}^{m}(k) x_j}{\sum_{i=1}^{n} u_{ij}^{m}(k)} \tag{5}$$

Solving the optimal solution of the objective function is a repeated process. Each time the membership matrix and cluster center matrix are obtained, the distance between the two cluster centers needs to be compared, as shown in formula (6)

$$\left\| C^{(k+1)} - C^{(k)} \right\| \le \varepsilon \tag{6}$$

Only when the distance between the two clustering centers is less than the value of the termination standard set in advance will the iteration be ended, otherwise the iterative solution will continue.

IV. Design of e-commerce platform recommendation model framework based on Fuzzy Clustering

Fuzzy clustering algorithm is very different from the traditional recommendation algorithm. It is to establish a user group database to find the specified users, analyze the contents that the specified users are interested in, and then make recommendations. Evaluate a certain content based on the comprehensive comparison information of similar customers, so as to form a new system, analyze the understanding and interest of the specified user in a certain information, and judge whether it should be pushed. Because there are many recommended contents in modern e-commerce platforms, the design of a model framework needs to combine a variety of recommendation technologies at the same time [20].

The e-commerce platform recommendation model based on fuzzy clustering is shown in Figure 1.

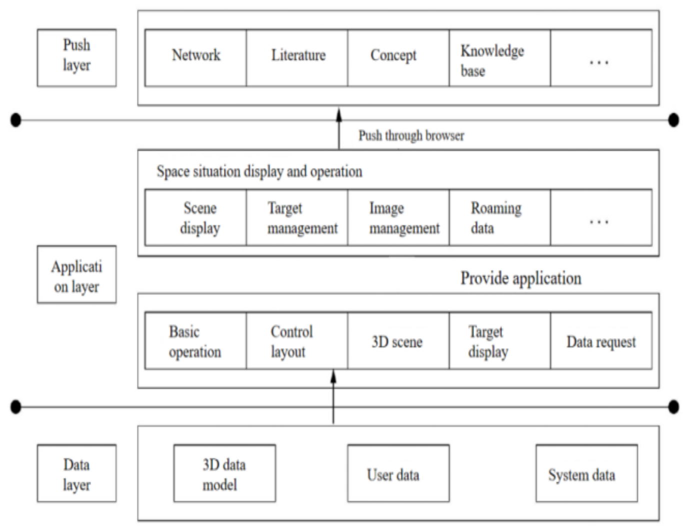

Fig.1 framework of e-commerce platform recommendation model based on Fuzzy Clustering

It can be seen from Figure 1 that electronic information is pushed by establishing the relationship between data layer, application layer and push layer. The data in the database mainly includes three types: 3D data model, user data and system data. The data analyzed from the data layer is transferred to the application layer, and the central system makes an overall analysis of its content, After a series of basic operations, decide whether the content should be pushed to the user. The pushed content includes network, literature, concept and knowledge base.

V. Recommendation design based on Improved Fuzzy Clustering Algorithm

A. Improved algorithm flow chart design

The flow chart of personalized recommendation algorithm based on improved fuzzy c-means clustering is shown in Figure

2,

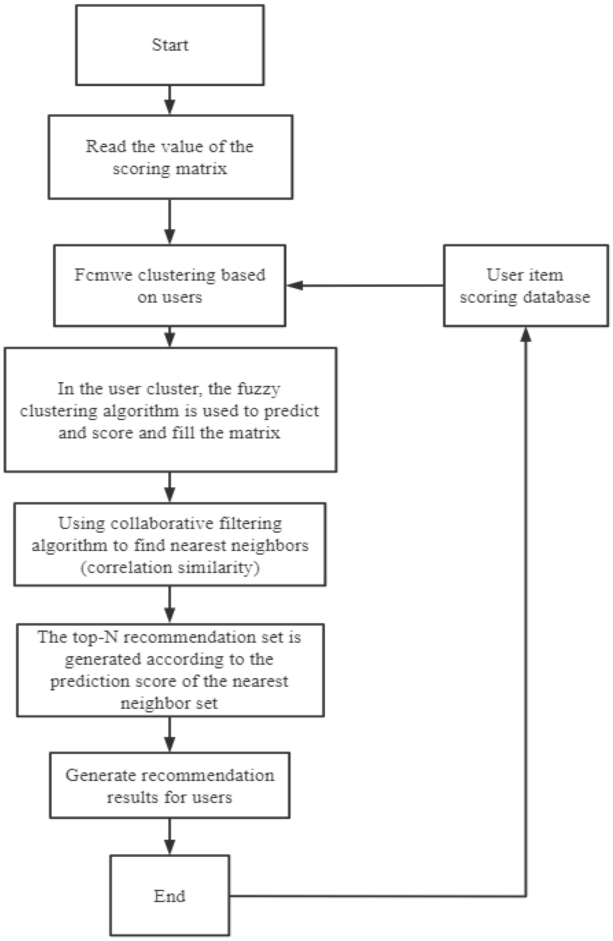

Fig. 2 flow chart of improved algorithm

Firstly, the data set is transformed into matrix form, then the matrix is classified by fcmwe algorithm, then the data is filled with fuzzy clustering algorithm, and then the nearest neighbor of the target user is found and the nearest neighbor set is generated by collaborative filtering algorithm.

The above clustering based real-time strategy is clustering and the process of searching for the nearest neighbor is also carried out on the matrix. Through the above accuracy experiment, although this method can reduce the search space and solve the real-time problem of recommendation, the recommendation accuracy is reduced. It is not advisable to improve the recommendation efficiency at the expense of reducing the accuracy. In addition to the user scoring matrix, you can also use the item attributes. Using this information can effectively improve the accuracy of the system.

B. Hierarchical fuzzy clustering

The hierarchical clustering algorithm simulates the tree structure. Through multiple iterations, all sample points form the tree structure. It is very convenient to select the nodes of the tree as the clustering result. At present, there are two kinds of hierarchical clustering algorithms, namely, clustering algorithm through splitting and clustering algorithm through aggregation. The basic idea of the split based clustering algorithm is that in the initial stage of the algorithm, all samples are regarded as a cluster, and some data are segmented in one iteration. When the stop criterion is reached, the algorithm stops. The basic idea of clustering algorithm based on aggregation is completely opposite to that of splitting algorithm. In the initial stage of the algorithm, each sample is an independent cluster. In one iteration, the two samples are aggregated into a new cluster, and then enter the next round of aggregation process. When only one cluster is left or other shutdown criteria are reached, the algorithm ends. On the whole, the computation of hierarchical clustering algorithm is much less than that of partition based algorithm, but the biggest problem is that hierarchical clustering algorithm is not traceable, which means that when a sample point is merged or split, it can not return to the state before the operation is completed. This strategy greatly reduces the calculation scale, but the non traceability makes it necessary to ensure the correctness of each iteration, otherwise the errors will accumulate and amplify all the time, and finally affect the results [22].

Hierarchical clustering algorithm is widely used in practice. The focus of this method is to determine the distance between clusters. Table 1 below lists the commonly used formulas that can be used to calculate the clustering distance

Table 1 Calculation Method of hierarchical clustering distance

Hierarchical clustering method	Distance formula
Shortest distance method	Euclidean distance
Longest distance method	Euclidean distance
Average Linkage Clustering	Square Euclidean distance
Center of gravity method	Flat Euclidean distance
Sum of squares of deviations	Square Euclidean distance

Because various distance calculation methods are different, have unique characteristics, and the calculation complexity is different due to the actual situation, it is necessary to measure and analyze various distance calculation methods in the application process in order to achieve satisfactory performance. At this stage, there is no clear research on the relationship between the calculation method that can measure the selected distance and the final result, so we can only adopt the method with excellent research and comparison results for many times, and generally follow two basic principles:

1) By analyzing a large number of data, we can get experience, and then improve the hit rate of selecting the correct model and calculation method through experience, so as to reduce the number of experiments.

2) Test all existing algorithms, and analyze the detailed calculation method adopted under a specific index by determining the index authority.

C. Improved recommendation algorithm based on Fuzzy Clustering

Initialize the number of clustering categories C, where 2 ≤ C ≤ n, and the value of C is generally determined according to the empirical value. Set the initial fuzzy classification matrix R (0) and iterate successively according to the following steps, where I = 0,1,2... Is the number of iterations, and N is the number of data divided in the data set.

For R (1), calculate the cluster center matrix, where $V^{(I)} = (V_1^{(I)}, V_2^{(I)}, ..., V_c^{(I)})^T$,according to formula (7)

$$V_i^{(I)} = \frac{\sum_{K=1}^{N} (r_{ik}^{(I)})^q u_k}{\sum_{K=1}^{N} (r_{ik}^{(I)})^q} \tag{7}$$

Adjust the fuzzy classification matrix R (1) to obtain (8)

$$r_{ik}^{(I+1)} = \frac{1}{\left\{ \sum_{j=1}^{c} \left[\left(\frac{\left| u_k - V^{(I)}{}_i \right|}{\left| u_k - V^{(I)}{}_j \right|} \right)^{\frac{2}{q-1}} \right] \right\}} (k = 1,2,...,n;\ j = 1,2,...c) \tag{8}$$

Clustering results: clustering center matrix V (I) and fuzzy classification matrix R * (9)

$$R^* = \begin{bmatrix} r_{11}^* & r_{12}^* & ... & r_{1n}^* \\ r_{21}^* & r_{22}^* & ... & r_{2n}^* \\ ... & ... & ... & ... \\ r_{n1}^* & r_{n2}^* & ... & r_{nn}^* \end{bmatrix} V^* = (V_1^*, V_2^*, ...V_C^*)^T \tag{9}$$

For ∀ UK ∈ u, if in column K of R *, if according to (10)

$$r_{ik}^* = \max(r_{jk}^*) \tag{10}$$

Then the object UK is classified as class I, indicating that the object UK belongs to class I with the greatest membership, and the UK is divided into class I.

In formula (11)

$$\left\| u_k - V^* \right\| = \min(\left\| u_k - V_j^* \right\|) \tag{11}$$

If ∀ UK ∈ u, the object UK is divided into class I, indicating that the object UK is closest to the cluster center vector Vi.

VI. EXPERIMENTAL VERIFICATION AND ANALYSIS OF RECOMMENDED RESULTS

A. Experimental scheme

There are many conditions that affect the performance of the recommendation algorithm. In the fuzzy clustering algorithm, the following can affect the accuracy of recommendation: first, the accuracy of recommendation will vary with the number of determined clustering centers [23]; Second, it is related to the number of nearest neighbors selected in each recommendation; Third, the accuracy of recommendation results will be different under different sparsity degrees. Based on this, in order to ensure the credibility of the experiment, this paper carries out experiments from the above three aspects to verify the performance of the improved fcmwe algorithm in this paper by comparison:

(1) When determining the number of cluster centers, the accuracy of the recommendation algorithm in different cases is compared by changing the number of nearest neighbors, so as to verify the impact of selecting different number of nearest neighbors on the recommendation accuracy.

(2) When the number of nearest neighbors is constant, different cluster centers are selected for experiments to verify the impact of the number of different cluster centers on the recommendation performance.

(3) To verify whether the data sparsity has an impact on the recommendation results, we only need to observe the recommendation accuracy in the data sets with different sparsity in the above two groups of experiments.

In order to ensure the accuracy of the simulation experiment in this paper, the selected movielens data set is divided into five parts on average, marked as a, B, C, D and e respectively. Each experiment selects four of them as the training set and the remaining one as the test set for five experiments, and then takes the average value of these five experiments as the final result, so as to reduce the error. The specific data set division method is shown in Table 2.

Table 2 Data set division

Number of experiments	Training set	Test set
1	ABCD	E
2	ABCE	D
3	ABDE	C
4	ACDE	B
5	BCDE	A

B. Experimental results analysis

The number of cluster centers is determined to be 20, and then the number of nearest neighbors selected each time is gradually changed to 5, 10, 15, 20, 25, 30, 35 and 40. Mae values are calculated using three recommendation algorithms: traditional FCM recommendation algorithm, improved fcmwe recommendation algorithm (not pre filled with slopeone) and improved fcmwe recommendation algorithm (pre filled with slopeone), The advantages and disadvantages of the algorithm in this paper are verified by data comparison. The experimental results are shown in tables 3 to 5.

Table 3 average absolute error of traditional FCM recommended algorithm

Number of neighbors	Test set A	Test set B	Test set C	Test set D	Test set E	Mae mean

5	0.9037	0.9265	0.9246	0.9076	0.8985	0.9121
10	0.8976	0.8904	0.8864	0.8915	0.8912	0.8914
15	0.8808	0.8867	0.8751	0.8680	0.8769	0.8775
20	0.8756	0.8799	0.8751	0.8800	0.8746	0.8742
25	0.8636	0.8782	0.8666	0.8788	0.8635	0.8701
30	0.8757	0.8704	0.8660	0.8774	0.8716	0.8722
35	0.8774	0.8834	0.8841	0.8753	0.8808	0.8802
40	0.8859	0.8920	0.8894	0.8874	0.8948	0.8899

Table 4 average absolute error of improved fcmwe recommended algorithm (slope one is not used)

Number of neighbors	Test set A	Test set B	Test set C	Test set D	Test set E	Mae mean
5	0.9002	0.9145	0.9106	0.9112	0.9012	0.9088
10	0.8951	0.8796	0.8796	0.8796	0.8796	0.8796
15	0.8710	0.8796	0.8749	0.8749	0.8636	0.8703
20	0.8631	0.8629	0.8622	0.8657	0.8612	0.8631
25	0.8626	0.8585	0.8629	0.8710	0.8667	0.8644
30	0.8767	0.8724	0.8745	0.8732	0.8716	0.8756
35	0.8931	0.8830	0.8839	0.8906	0.8831	0.8868
40	0.8919	0.8920	0.8885	0.8887	0.8989	0.8919

Table 5 average absolute error of improved fcmwe recommended algorithm (using slope one)

Number of neighbors	Test set A	Test set B	Test set C	Test set D	Test set E	Mae mean
5	0.9006	0.9016	0.9005	0.9015	0.9012	0.9011
10	0.8808	0.8813	0.8818	0.8805	0.8806	0.8810
15	0.8893	0.8879	0.8876	0.8878	0.8887	0.8883
20	0.8419	0.8486	0.8482	0.8491	0.8487	0.8489
25	0.8526	0.8512	0.8513	0.8511	0.8523	0.8644
30	0.8709	0.8696	0.8692	0.8707	0.8756	0.8686
35	0.8860	0.8776	0.8774	0.8781	0.8759	0.8770
40	0.8620	0.8608	0.8605	0.8613	0.8609	0.8611

According to the number of selected nearest neighbors and the change of corresponding Mae value, the MAE broken line comparison diagram of the three recommended algorithms is generated, as shown in Figure 3.

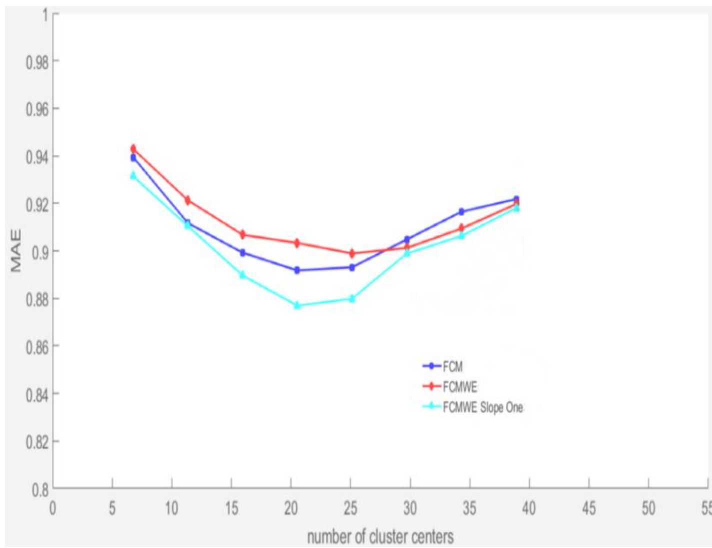

Fig. 3 comparison results of the first group of experiments

It can be clearly seen from the comparison figure that the number of cluster centers is fixed. With the increase of the number of nearest neighbors selected, the MAE value tends to decrease gradually, but when the number of nearest neighbors reaches a certain degree, the MAE value tends to increase. It can be seen that the selection of the number of nearest neighbors has an impact on the recommendation accuracy. Through the comparison of the three recommendation algorithms, it can be seen that fcmwe algorithm can effectively improve the recommendation accuracy, but with the increase of the number of nearest neighbors selected, the recommendation effects of the three algorithms tend to be the same.

In each experiment, the fixed number of nearest neighbors is 20, and then the number of cluster centers is gradually changed to 5, 10, 15, 20, 25 and 30. Mae values are calculated by using three recommendation algorithms: traditional FCM recommendation algorithm, improved fcmwe recommendation algorithm (not pre filled with slope one) and improved fcmwe recommendation algorithm (pre filled with slope one), Compare and observe the changes of MAE value, and then analyze the impact of the change of cluster center on the recommendation accuracy. The experimental results are shown in tables 6 to 8

Table 6 average absolute error of traditional FCM recommended algorithm

Number of neighbors	Test set A	Test set B	Test set C	Test set D	Test set E	Mae mean
5	0.8479	0.8483	0.8471	0.8488	0.8488	0.8481
10	0.8522	0.8509	0.8505	0.8517	0.8517	0.8512
15	0.8602	0.8590	0.8584	0.8595	0.8595	0.8592
20	0.8636	0.8622	0.8619	0.8621	0.8621	0.8625
25	0.8812	0.8796	0.8795	0.8799	0.8799	0.8801
30	0.8908	0.8914	0.8921	0.8905	0.8905	0.8911
35	0.9122	0.9110	0.9104	0.9115	0.9115	0.9112
40	0.9250	0.9239	0.9234	0.9238	0.9238	0.9241
45	0.9363	0.9347	0.9346	0.9350	0.9350	0.9352
50	0.9483	0.9470	0.9466	0.9477	0.9477	0.9473

Table 7 average absolute error of improved fcmwe recommended algorithm (slope one is not used)

Number of neighbors	Test set A	Test set B	Test set C	Test set D	Test set E	Mae mean
5	0.8351	0.8339	0.8333	0.8345	0.8337	0.8341
10	0.8402	0.8389	0.8385	0.8394	0.8390	0.8392
15	0.8521	0.8507	0.8505	0.8514	0.8508	0.8511
20	0.8524	0.8535	0.8540	0.8536	0.8530	0.8533
25	0.8551	0.8564	0.8568	0.8559	0.8564	0.8561
30	0.8763	0.8767	0.8761	0.8759	0.8775	0.8765
35	0.8794	0.8805	0.8811	0.8799	0.8793	0.8801
40	0.9111	0.9121	0.9110	0.9109	0.9106	0.9112
45	0.9226	0.9236	0.9225	0.9240	0.9228	0.9231
50	0.9310	0.9318	0.9319	0.9312	0.9311	0.9314

Table 8 average absolute error of improved fcmwe recommended algorithm (using slope one)

Number of neighbors	Test set A	Test set B	Test set C	Test set D	Test set E	Mae mean
5	0.8311	0.8299	0.8293	0.8298	0.8304	0.8301
10	0.8320	0.8310	0.8306	0.8315	0.8309	0.8312
15	0.8490	0.8480	0.8479	0.8481	0.8485	0.8483
20	0.8497	0.8500	0.8512	0.8501	0.8505	0.8503
25	0.8516	0.8512	0.8507	0.8511	0.8524	0.8514
30	0.8649	0.8661	0.8645	0.8656	0.8648	0.8652
35	0.8711	0.8697	0.8695	0.8704	0.8698	0.8701
40	0.8864	0.8877	0.8860	0.8870	0.8865	0.8867
45	0.8907	0.8917	0.8906	0.8921	0.8909	0.8912
50	0.9019	0.9028	0.9026	0.9015	0.9016	0.9021

The MAE value comparison curves of the three algorithms are generated from the above three data tables, as shown in Figure 4.

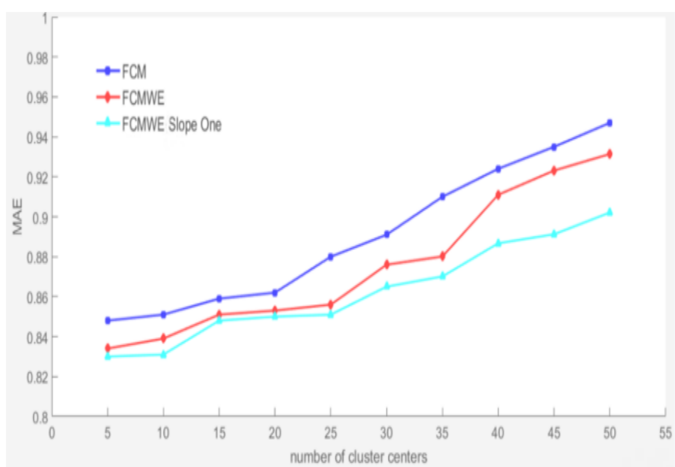

Fig. 4 comparison results of the second group of experiments

It can be seen from the graph in Figure 4 that the different number of selected cluster centers will have an impact on the recommendation results. With the increase of the number of selected cluster centers, the MAE value gradually increases, and the recommendation results tend to become worse. Therefore, the selection of the number of cluster centers will affect the recommendation results. It can also be seen from the above two comparison figures that the three recommendation algorithms have certain differences in recommendation accuracy. Compared with the other two, fcmwe algorithm has higher accuracy and achieves the purpose of optimizing the algorithm. From the above two sets of experimental data, it can be seen that the MAE values obtained by experimental simulation are different for different test sets. Therefore, it can be seen that different data sparsity will affect the accuracy of recommended results.

VII. Conclusions

The traditional collaborative filtering recommendation has some disadvantages, such as poor real-time, accuracy to be improved, and sensitivity to sparse data. The process of clustering objects is the process of clustering objects with high similarity[24]. The core idea adopted is similar to the recommendation technology, that is, to find similar objects. After clustering the objects with clustering technology, Users with high similarity gather in the same group. When using collaborative filtering algorithm, it is not necessary to find the nearest neighbor of the target user in the whole user space, and reduce the search space to several clusters close to the target user. Therefore, introducing clustering technology into the recommendation system can effectively improve the real-time requirements of the recommendation system [25].

In the whole recommendation system, the information that can be used to seek similar relationships is limited. The traditional recommendation algorithm mainly recommends based on the user item scoring relationship, and does not apply all the relationships in the recommendation system. The hybrid recommendation algorithm considers the user attribute and commodity attribute in recommendation, The similarity calculated by user attributes and commodity attributes and the similarity calculated by user scoring information are fused [26], which can not only improve the recommendation accuracy, but also solve some cold start problems. However, this hybrid method is derived from the computational similarity of commodity attributes, and the commodity attribute matrix itself is a matrix with high sparsity, The contribution of fusion only by calculating similarity to the improvement of accuracy is limited. Therefore, this paper proposes hybrid recommendation based on fuzzy clustering, which effectively improves the accuracy of hybrid recommendation [27].

Acknowledgments

This work is supported by 2019 Doctoral Research Startup Project of Jilin Normal University(Project No. 2019037).

References

[1] Zhao Qing. Research on Personalized Recommendation Algorithm of chemical product e-commerce based on genetic fuzzy clustering [J]. Bonding, 2020, 44 (11): 4

[2] Zhu Zhihui, Zhu Meifang. Research on e-commerce personalized recommendation algorithm based on genetic fuzzy clustering [J]. Journal of Jiujiang University: Natural Science Edition, 2019, 034 (001): 61-65

[3] Li Qingxia, Wei Wenhong, Cai Zhaoquan. E-commerce personalized recommendation algorithm for hybrid user and project collaborative filtering [J]. Journal of Sun Yat sen University: Natural Science Edition, 2016, 55 (5): 6

[4] Zhang Kaisheng, song Wenwei, Li Huizhen. Parallel recommendation algorithm based on fuzzy clustering [J]. Journal of Shaanxi University of Technology: Natural Science Edition, 2019 (4): 57-61

[5] Zhao Hua, Lin Zheng, Fang AI, et al. A recommendation algorithm based on knowledge tree and its application in mobile e-commerce [J]. 2021 (2011-6): 54-58

[6] Ji Xiaoyan, Li Yulong. Personalized recommendation algorithm based on improved clustering in Hadoop environment [J]. Journal of Lanzhou Jiaotong University, 2017, 036 (001): 70-76

[7] Du Xi Xi, Liu Huafeng, Jing Liping. A superposition joint clustering recommendation model integrating social networks [J]. Journal of Shandong University: Engineering Edition, 2018, 48 (3): 7

[8] Zhang Yinghui, Li Xue. Tourism recommendation algorithm based on fuzzy clustering [J]. Computer technology and development, 2016, 026 (012): 99-102

[9] Zhang Yinghui, Li Xue. Tourism recommendation algorithm based on fuzzy clustering [J]. Computer technology and development, 2016, 26 (12): 99-102, 4 pages in total

[10] Gao Tianying, Zhang Zhigang, Li Guoyan, et al. A collaborative filtering recommendation algorithm based on user attribute and website type clustering [J]. Journal of Tianjin urban construction university, 2018, 24 (1): 6

[11] Wang Xi, Wang Yanming, Wang, et al. An improved collaborative filtering recommendation algorithm [J]. Modern computer (Professional Edition), 2017, 14 (14): 10-15

[12] Dong Hui, Fang Xiao, Ma Jian, et al. Mobile e-commerce user item clustering collaborative filtering recommendation algorithm based on situational awareness [J]. Journal of Guangxi University for Nationalities (NATURAL SCIENCE EDITION), 2018, 24 (02): 67-74

[13] Shi Yingying, Ge Wancheng, Wang Liangyou, et al. Research on improvement of K-means clustering personalized recommendation algorithm [J]. Information and communication, 2016, No. 157 (01): 19-21

[14] Sun Kele, Deng Xianrong. An improved o2o e-commerce recommendation model based on gradient lifting regression algorithm [J]. Journal of Anhui University of architecture and Architecture: Natural Science Edition, 2016

[15] Zhang Yinghui, Li Xue. Tourism recommendation algorithm based on fuzzy clustering [J]. Computer technology and development, 2016 (12): 99-102, 4 pages in total

[16] Wang Min, Ji Shaochun. Research on personalized recommendation of Digital Library Based on fuzzy clustering and fuzzy pattern recognition [J]. Modern information, 2016, 36 (04): 52-56

[17] Wang Xiaojun. Distributed hybrid collaborative filtering method in recommendation system [J]. Journal of Beijing University of Posts and telecommunications, 2016, 39 (002): 25-29

[18] Zhao Yan, Wang Yamin, Liu huailiang. Research on personalized Resource Recommendation Model Based on tag network clustering [J]. 2021 (2014-4): 179-183

[19] Wang Zhaokai, Li Yaxing, Feng Xupeng, et al. Personalized information recommendation based on deep belief network [J]. Computer Engineering, 2016, 42 (010): 201-206

[20] Hu Chaoju, sun Keni. Research on Personalized Recommendation Based on user fuzzy clustering [J]. Software guide, 2018, 017 (002): 31-34

[21] Li Haoyang, Fu Yunqing. Collaborative filtering recommendation algorithm based on tag clustering and project topic [J]. Computer science, 2018, v.45 (04): 247-251

[22] Wang Xiaojun, Fu Chao. Using fuzzy blocking to improve the scalability and accuracy of collaborative filtering [J]. Journal of Beijing University of Posts and telecommunications, 2017 (01): 78-82

[23] Zhang Yanju, Lu Chang. IFCM slope one collaborative filtering recommendation algorithm under missing data [J]. 2021 (2020-9): 185-188

[24] Tan Libin, Tang dunbing, Chen Weifang, et al. Open design decision-making method with large-scale user participation [J]. Computer integrated manufacturing system, 2020, 26 (4): 9

[25] Mu Jun. community network data crawler algorithm based on association rule mining [J]. Microelectronics and computer, 2018, 035 (008): 105-108

[26] Liu Jingping, Li Ping. A fuzzy cognitive collaborative filtering algorithm [J]. Computer engineering and science, 2018, 040 (005): 898-905

[27] Zhou Chaojin, Wang Yuzhen. Research on personalized recommendation of agricultural products based on improved collaborative filtering algorithm [J]. Journal of Shaoyang University (NATURAL SCIENCE EDITION), 2017, 014 (006): 23-31