

## Research Article

# Cross-Border E-Commerce Personalized Recommendation Based on Fuzzy Association Specifications Combined with Complex Preference Model

Dan Xiang <sup>1</sup> and Zhijie Zhang <sup>2</sup>

<sup>1</sup>Business School, Zhijiang College of Zhejiang University of Technology, Hangzhou, Zhejiang, China

<sup>2</sup>School of Electrical Engineering, University of Jinan, Jinan 250022, China

Correspondence should be addressed to Zhijie Zhang; [cse\\_zzj@ujn.edu.cn](mailto:cse_zzj@ujn.edu.cn)

Received 1 September 2020; Revised 27 September 2020; Accepted 7 October 2020; Published 21 October 2020

Academic Editor: Yi-Zhang Jiang

Copyright © 2020 Dan Xiang and Zhijie Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since cross-border e-commerce involves the export and import of commodities, it is affected by many policies and regulations, resulting in some special requirements for the recommendation system, which makes the traditional collaborative filtering recommendation algorithm less effective for the cross-border e-commerce recommendation system. To address this issue, a simple yet effective cross-border e-commerce personalized recommendation is proposed in this paper, which integrates fuzzy association rule and complex preference into a recommendation model. Under the constraint of fuzzy association rules, a hybrid recommendation model based on user complex preference features is constructed to mine user preference features, and personalized commodities recommendation is realized according to user behavior preference. Compared with the traditional recommendation algorithm, the improved algorithm reduces the impact of data sparsity. The experiment also verifies that the improved fuzzy association rule algorithm has a better recommendation effect than the existing state-of-the-art recommendation models. The recommendation system proposed in this paper has better generalization and has the performance to be applied to real-life scenarios.

## 1. Introduction

The recommendation system refers to the use of e-commerce websites to provide customers with product information and suggestions. The system can help users decide what products to buy and simulate sales staff to help customers complete the purchase process. The personalized recommendation is to recommend information and commodities that users are interested in based on the user's interest characteristics and purchase behavior. As a new industry in e-commerce, cross-border e-commerce has dramatically expanded its data scale as e-commerce, and users are facing an increasingly serious problem of information overload. As an effective means to solve this problem, the e-commerce recommendation system has achieved certain results in both academia and business [1–3]. However, since cross-border e-commerce

involves the export and import of goods, it is affected by many policies and regulations, resulting in some special requirements for the recommendation system, which makes the traditional collaborative filtering recommendation algorithm less effective for the cross-border e-commerce recommendation system [4].

The recommendation system collects user behavior information, feature information, and context information and then pushes the appropriate commodity information to the user through a certain recommendation algorithm at the appropriate time [5]. It is an intelligent system that can help users filter and look for products. In addition, the recommendation system can provide users with a list of products that may have preferences when users have not entered search keywords or users have not defined their own needs, which improves the efficiency of users in filtering products

from massive commodity information. On the other hand, it can intelligently analyze the personalized needs of users and provide a customized recommendation list for each user, so it can promote the long-tail effect of information [6].

The recommendation systems generally use the following methods: content-based recommendation, knowledge-based recommendation, rule-based recommendation, and collaborative filtering recommendation. In this paper, we mainly introduce the fuzzy association recommendation algorithm. Currently, all major e-commerce platforms have applied recommendation systems [7]. Amazon, the largest shopping platform, is an e-commerce company that applied recommendation systems earlier. Its online system has applied a variety of collaborative filtering (CF) recommendation algorithms, which can intelligently recommend products for users based on user behavior information such as searching, browsing, and purchasing on the website [8]. At least 35% of sales benefits in Amazon is from the recommendation system. Therefore, more and more e-commerce platforms are using CF for the recommendation system. In addition to the CF recommendation algorithm, other common recommendation algorithms also include knowledge-based (KB) recommendation algorithms and context-based (CB) recommendation algorithms. Compared with these two recommendation algorithms, the CF recommendation algorithm has the advantages of low implementation difficulty, low data dependence, and high recommendation accuracy [9].

As we all know, the core of the recommendation system is the recommendation algorithm [9]. The CF recommendation algorithm is the most widely used in the world. Traditional user-based collaborative filtering recommendation (User-CF) uses the nearest neighbor search model in user-commodity rating data to complete personalized recommendations. User-item rating data generation, nearest neighbor search (NNS), and product recommendations are the three most important steps of CF recommendation algorithm. The accuracy of the nearest neighbor search is an important factor affecting the quality of CF recommendations. In practical scenarios, CF recommendation often faces the cold-start problem, which is related to the sparsity of item information used in the recommendation algorithm. Recently, a lot of research has been devoted to solving the cold-start problem. Zafar et al. [8] combined collaboration data with context data as a solution to the cold-start problem, where this type of model contains three data sources: user, item, and content. By adopting a probability model, the influence of collaboration data and content naturally appears in a given data source; Wendi et al. [9] proposed six models for understanding new users available in the User-CF recommendation system, which will select a series of items for each new user; Shudong et al. [10] proposed a novel association rules to expand user information through so as to avoid the cold-start problem; Tian et al. [11] proposed a classification algorithm combined with similar technology and prediction mechanism, which provides the necessary means for retrieval and recommendation; Wu et al. [12] proposed a recommendation framework based on tightly coupled CF model and deep learning neural network

to solve the cold-start problem and proved the feasibility and effectiveness of the model through experiments on Netflix movie recommendation data; Li et al. [13] proposed a recommendation model DTMF that integrates scoring values and review texts, which can reduce the sparsity of scoring data by applying text analysis and opinion mining methods, and the scoring matrix decomposition model also can improve recommendation quality. In addition, there are many studies to solve the problem of data sparsity from the perspective of classification and clustering. Aleksandra et al. [14] adopted the clustering analysis for users in different contexts to reduce the sparsity of user-item rating data, thereby improving the effect of the CF algorithm; Jiang et al. [15] proposed the user clustering based on domain knowledge and designed a two-stage recommendation model. From the above analysis, it can be clearly seen that the recommendations based on collaborative filtering are reasonable, but it is not the best choice in a cross-border e-commerce recommendation system, so the recommendation accuracy will also be lost. The real data in reality is complex and ambiguous, and the number of sets in the dataset does not all meet the amount demand for a recommendation system. Therefore, the association rule-based recommendation is more suitable for real-life scenarios.

In the era of big data, the data has the trend of explosion. There are many overlapped intersections between datasets, and the boundary between them becomes more and more fuzzy. Therefore, the problem is that Boolean association rules only deal with Boolean data, but they cannot solve the problem of big data perfectly. Therefore, the fuzzy association rules become important at this time. If we can get fuzzy association rules, we can solve the problem of recommendation perfectly. In recent years, more and more scholars at home and abroad study fuzzy association rules for the recommendation system [16–19]. Chun et al. [20] proposed a recommendation algorithm based on fuzzy association rules and proved its effectiveness. Zhang et al. [21] introduced a fuzzy c-means algorithm to obtain association rules in 2016 and proposed improved fuzzy association rules based on the Apriori algorithm combined with fuzzy c-means. Yera et al. [22] proposed a novel multigroup parallel multimutation particle swarm optimization algorithm to search multiple recommendation products in parallel. Zhang et al. [23] proposed a fuzzy data recommendation algorithm based on association rules, where the regression model was added to the recommendation process of fuzzy association rules, and the data can be statistically analyzed. Feng et al. [24] proposed a recommendation algorithm to obtain appropriate membership functions and useful fuzzy association rules from a database with a wide range of uncertain data. Meanwhile, the fuzzy Apriori algorithm was improved to learn membership function from the genetic process so that effective association rules could be obtained from uncertain data, and its effectiveness had been proved. Xiao and Shen [25] proposed a fuzzy reasoning method based on fuzzy association rules in 2015 and applied it to evaluate fuzzy association rules. Ahmadian et al. [26] proposed a fuzzy recommendation algorithm based on association rules. A discretization improved fuzzy Apriori was

used to classify the data of each dimension, and then the quantitative results were expressed in the form of a nominal variable matrix. The association rules and fuzzy model were combined to establish the quantitative association rules with fuzzy rules, and a fuzzy recommendation algorithm was realized. It is applied to predict the recommendation product.

Therefore, from the above analysis, it can be seen that the fuzzy association rule is an implementation to solve the problem of the fuzzy boundary in the dataset. Since there are many data without a clear boundary in cross-border e-commerce, fuzzy association rules are widely used. At present, the improvement and application of fuzzy association rules has been a hot research direction. With the continuous development of big data, the association rules will play a greater role in the field of cross-border e-commerce. The use of association rules can get more rules, which is a better aid for the recommendation system.

This paper first summarizes and combs the basic theory of the existing recommendation system. On this basis, it combines the user preferences to commodity attributes and contextualized user preferences to meet the system requirements of the cross-border e-commerce recommendation system. In order to verify the effectiveness of the improved recommendation algorithm, the cross-border e-commerce enterprise orders a dataset to analyze the improved algorithm. Compared with the traditional collaborative filtering recommendation algorithm, the improved algorithm reduces the impact of data sparsity. The experiment also verifies that the improved fuzzy association rule algorithm has a better recommendation effect than the traditional collaborative filtering.

## 2. Association Rule

As an important research content in the cross-border e-commerce enterprise, the association rule has been promoted in various industries, and it has quickly become a very popular research field, where the more typical algorithms are Apriori and FP-growth [26]. In the era of big data, the boundaries between data attribute values cannot be strictly divided. Therefore, the transformation of quantitative data in real data into discrete data has always been the key to mining association rules of quantitative attribute data.

Association rules based on Apriori algorithm is the most classic algorithm in recommendation system and data-mining technology, which can discover some hidden relationships between two products through a certain rule [27]; the original pseudocode is offered in Table 1. Association rules can be represented by  $X \longrightarrow Y$ , where  $X$  represents the antecedent of the association rules and  $Y$  represents the consequent of the association rules. In addition, there are definitions of support degree and confidence degree in the association rule algorithm, and the calculation equations are shown as follows:

$$\text{Support}(X \longrightarrow Y) = \frac{P(X \cup Y)}{P(I)}. \quad (1)$$

The support degree determines the frequency of a given dataset.  $I$  is denoted as the total transaction set.

$$\text{Confidence}(X \longrightarrow Y) = \frac{P(X \cup Y)}{P(X)}. \quad (2)$$

The confidence degree represents the probability that  $Y$  is deduced from the association rule  $X \longrightarrow Y$  when the antecedent  $X$  occurs.

Fuzzy association rules need to solve the following problems: (1) Find the required frequent itemsets in cross-border e-commerce datasets by an iterative method. (2) The association rules are found from the fuzzy frequent itemsets. This is the same as the process of association rules, but the found datasets are ordinary datasets and fuzzy datasets. The mining of association rules is to find the lowest support and the lowest confidence interval in a cross-border e-commerce dataset, where the frequent itemset is found by calculating candidate sets, and then strong association rules are obtained through frequent candidate sets. As we all know, mining the core of frequent itemset association rules is also a step in the recommendation system that affects efficiency. Therefore, the association rule algorithm mainly has the following process:

(1) Calculate all frequent itemsets: according to equation (1), if the support of itemsets is not less than the minimum support, then it is frequent itemsets. The recommendation system in cross-border e-commerce needs to obtain all the frequent sets from 1 to  $n$ , which is all the frequent itemsets we require. (2) Generate strong association rules from frequent itemsets: as shown in equation (2), calculate association rules from all frequent itemsets, then calculate the confidence intervals of all rules, and choose which confidence intervals are not less than the lowest confidence interval. Ultimately, the obtained rules are the required strong association rules. Apriori for association rule is shown in Figure 1.

It can be seen that, in the recommendation process, the first problem is the key step to improve the recommendation accuracy and the performance of the recommendation is also determined by the first step. It may be affected by various aspects: firstly, too large candidate itemsets are generated. For example, if we have  $10^4$  frequent 1-item sets during the calculation process, when calculating the candidate 2-item sets, the number of 2-item candidate sets that need to be generated is  $10^7$ . In addition, in order to calculate the minimum support of each frequent candidate set, we need to repeat the traversal data many times when looking for frequent itemsets from the candidate set.

## 3. Fuzzy Association Specification Combined with Complex Preference Model

**3.1. Problem Description.** It is given that there are  $p$  users and  $q$  commodities in the cross-border e-commerce shopping guide platform, and the user set is  $u = \{u_1, u_2, \dots, u_p\}$ , the itemset is  $i = \{i_1, i_2, \dots, i_q\}$ , the user-item rating matrix is  $V$ , and  $v_{ij}$  represents the user  $u_i$ 's rating result of item  $i_j$ ,

TABLE 1: Pseudocode for Apriori.

**Algorithm 1.**

The original Apriori algorithm.

- (1)  $L_1 = \{\text{large 1-itemsets}\};$
- (2) **for** ( $k = 2; L_{k-1} \neq \emptyset; k++$ ) **do begin**
- (3)  $C_k = \text{apriori}(L_{k-1});$  //New candidates
- (4) **For all** transactions  $t \in D$  **do begin**
- (5)  $C_t = \text{Subitemset}(C_k, t)$   
//Candidates itemset in  $t$
- (6) **For all** candidates  $c \in C_t$  **do**
- (7)  $c.\text{count}++;$
- (8) **end**
- (9)  $L_k = \{c \in C_k | c \geq \text{min support}\};$
- (10) **end**
- (11)  $\text{Result} = \bigcup_k L_k;$

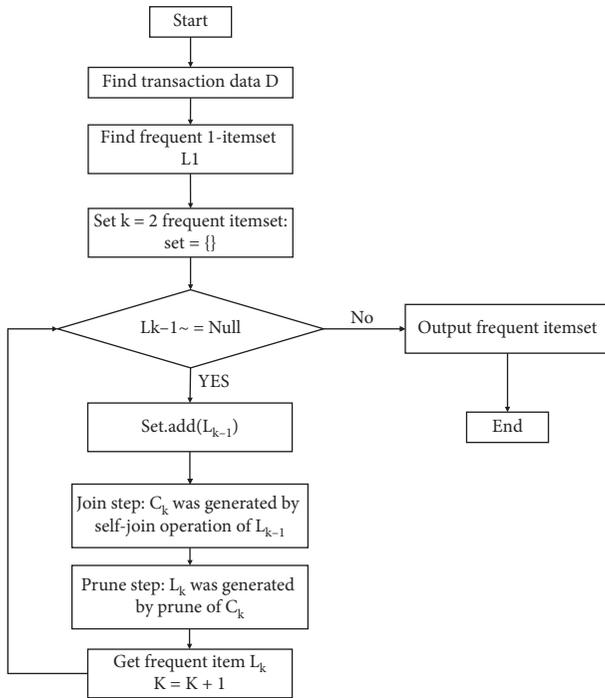


FIGURE 1: Apriori for association rule.

where  $1 \leq i \leq p$  and  $1 \leq j \leq q$ .  $v_{ij}$  can be Boolean or a quantitative number. If user  $u$  does not grade for item  $i$ , then  $v_{ij} = 0$ . In a cross-border e-commerce network, each user may purchase multiple products. The user's preference relationship matrix  $T = C(p_{ij})$ , where  $p_{ij}$  is the user  $u_i$ 's preference degree of item  $i_j$ . If users A and B have similar attributes, the element  $p_{ij}$  in the preference relationship matrix represents the strength of the user's preference; otherwise,  $c = 0$  indicates that the user and the item have no preference relationship.

**3.2. Improved Fuzzy Association Specification.** Section 2 has pointed out that all candidate itemsets from 1 to  $k$  are calculated by iteration for the discovery of all fuzzy frequent

items, and all frequent candidate itemsets are obtained by calculating the support degree of candidate sets. However, most of the rules are calculated from all the fuzzy frequent candidate sets in the actual operation. The confidence intervals of these rules are finally compared with the lowest confidence interval, where all the strong association rules are obtained, but there are large errors. If the interval partition is too small, there will be too many partition intervals. Thus, the support degree is too low in the same interval, resulting in too few rules generated in this interval. When the value of the fuzzy number is divided into multiple intervals, some data will always be lost. However, we cannot predict whether the missing data will affect the results, so we cannot predict whether it will affect the accuracy of the minimum support. Therefore, define a fuzzy support rate  $F_s$  and fuzzy confidence  $F_c$ . For any fuzzy set  $X = \{x_1, x_2, \dots\}$ , the fuzzy support rate of  $X$  can be expressed as  $F_s$ , and its equation is written as follows:

$$F_s(x) = \frac{\sum_{j=1}^n \prod_{m=1}^p t_j(x_m)}{n} \quad (3)$$

If  $F_s(x)$  is not less than the minimum support rate  $\text{min}_F$ , given by the user, then  $x$  is a fuzzy frequent attribute set. The fuzzy support rate of the fuzzy association rule  $X \rightarrow Y$  is  $F_s(x, y)$ ; it can be denoted as

$$F_s(x, y) = \frac{\sum_{j=1}^n \prod_{m=1}^{p+q} t_j(x_m, y)}{n}, \quad (4)$$

where  $y$  is any item in cross-border e-commerce. So the fuzzy confidence degree of the fuzzy association rule  $X \rightarrow Y$  is  $F_c$ , which can be denoted as  $F_s(x)/F_s(y) \geq \text{min\_conf}$ .

As we all know, whether it is to use hard-threshold methods or fuzzy soft-threshold methods to divide quantitative data into closed intervals from 0 to 1, the original database will be divided into a new database, where this may involve data preprocessing processing, such as deleting or filling incomplete data or deleting data that deviates far from the actual true value. Therefore, according to the definition and theorem in Section 2, the improved fuzzy model (see Figure 2) firstly calculates the fuzzy support rate of all 1-item candidate sets and then selects those with fuzzy support rate not less than the minimum support degree, which are all 1-item fuzzy frequent candidate sets. By combining all 1 fuzzy frequent candidate sets, all 2 frequent candidate sets are obtained; by calculating the support fuzzy rate of all 2 frequent candidates, select those candidate sets whose support rate is not less than the lowest support rate, namely, joining 2 frequent items. By combining all 2 frequent candidate sets, all 3 candidate sets are obtained; in other words, we use 3-item fuzzy candidate sets from 2-item fuzzy frequent candidate sets. According to the properties of user item, all its subsets must be frequent if the 3-item fuzzy candidate set is frequent; then, calculate the fuzzy support of the remaining 3-item fuzzy candidate set and select the one with the support not less than  $\text{min}_F$ , which is the 3-item

fuzzy frequent candidate set. As the iteration proceeds, the  $k$  frequent candidate sets are calculated, and the calculation is stopped until the  $k$  items are frequently empty; all the frequent candidate sets obtained in this way are the sets of frequent candidate sets we require. The association rules are calculated through all fuzzy frequent candidate sets, and then by calculating their confidence interval, select the association rules not less than the given minimum support min\_conf, which is the required strong association rule.

**3.3. Personalized Recommendation.** In the cross-border e-commerce shopping guide platform, the user's access behavior can be recorded through the basic content layer and the content-type layer. According to the user's preference, the commodity consistent with the user's preference is obtained through the association rule and recommended to users.

Assuming that  $A$  and  $B$  are content itemsets and  $A_p$  and  $B_p$  are, respectively, the categories of  $A$  and  $B$ , the double-layer association rule set is  $F = \{A \rightarrow B \& A_p \rightarrow B_p\}$ , where  $A \rightarrow B$  represents the association rules of the basic content layer and  $A_p \rightarrow B_p$  represents the association rules of the content-type layer.

The personalized recommendation process in a cross-border e-commerce shopping guide platform can be divided into multiple steps. All the high support itemsets whose support exceeds the given threshold in the content-type data required are obtained, and the calculation formula of the content-type association rules' support degree can be denoted as

$$T_{A_p \rightarrow B_p} = \frac{M(A_p \cap B_p)}{M}, \quad (5)$$

where  $A_p \cap B_p$  is the set of content-type items;  $M$  is the total amount of items. If the support degree of the itemset  $Y$  containing item  $X$  exceeds the established threshold, then  $Y(X)$  is a high support itemset; namely,

$$\frac{|Y(X)|}{M} \geq T_{\min}, \quad (6)$$

where  $T_{\min}$  represents the minimum support of content-type association rules. In addition, for each set  $S$ , the confidence degree of all nonempty subsets  $S'$  of  $S$  is obtained as follows:

$$\zeta_c = \frac{T_c}{T_s}, \quad (7)$$

where  $T_c$  is the length of the confidence interval of each set;  $T_s$  is the average length of confidence interval; if  $\zeta_c$  exceeds the given minimum confidence threshold, the content-type association rule  $C: S' \rightarrow (S \rightarrow S')$  is generated. All the high support itemsets in the basic content table with support exceeding the threshold are obtained, and the content association support is denoted as

$$T_{A \rightarrow B} = \frac{|M(A \cap B)|}{|M|}, \quad (8)$$

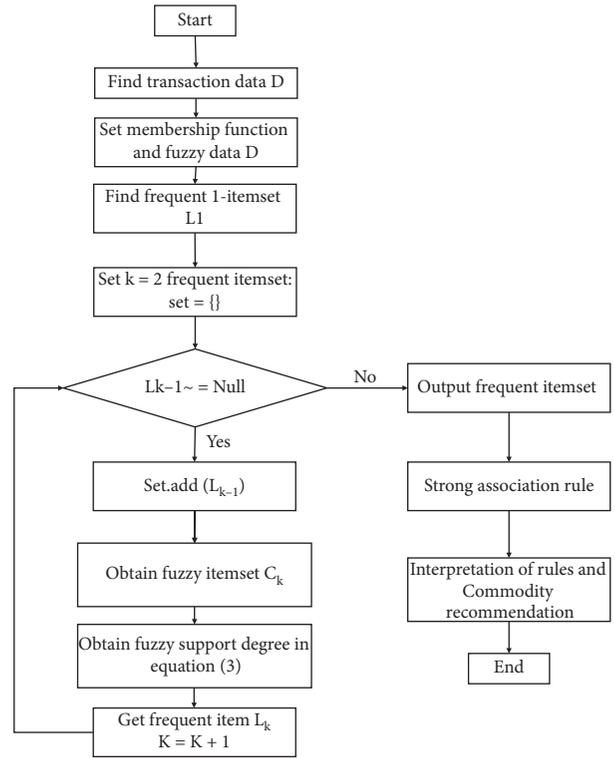


FIGURE 2: Improved fuzzy association rule.

where  $M(A \cap B)$  represents the type itemset of content. Similarly, the itemset is called the high support itemset  $Y(E)$  when it meets  $|Y(E)|/|M| \geq T_{\min}$ .

For each frequent itemset  $D$ , the confidence degree of all nonempty subsets  $D'$  is obtained as  $\zeta_c = T_{D'}/T_D$ , where  $T_D$  is the length of the confidence interval of each frequent itemset. If  $\zeta_c$  is higher than the minimum confidence threshold, the content-type-based association rule  $F: D' \rightarrow (D \rightarrow D')$  is generated. In the content-type association rules, the rule which is the content-type association rule with preference type is extracted, and the corresponding commodities based on content rules with higher support and confidence are obtained and recommended to users.

## 4. Experiment and Result Analysis

**4.1. Design and Implementation of Our Proposed Recommendation System.** The proposed recommendation system is implemented so as to validate and evaluate its performance in the distributed computing platform, which is mainly composed of a distributed system and distributed batch processing framework. The distributed system can provide large-scale data storage system, and the distributed batch processing framework Map-Reduce can provide a distributed application development interface. So the distributed computing platform can effectively improve the efficiency of our proposed algorithm, make full use of the

computing power of all servers, and enhance the scalability of our proposed algorithm [26, 28–30].

*Map-Reduce* is divided into two parts: data-mapping *Map* and data-reduction *Reduce*. *Map* has the function of decomposing tasks into multiple tasks, and *Reduce* has the function of gathering the processing results of multiple tasks after decomposition to obtain the final processing results. During this period, the data are saved in the form of key value,  $\langle \text{key}, \text{value} \rangle$ . The detailed process of Map-Reduce is as follows. Firstly, *Map-Reduce* takes the block as a unit and divides the input data into several blocks by using its splitter tool. The size of those divided blocks can be set by oneself, and then all data blocks are copied. In task assignment, any copy program can be selected as the main-control program, which can complete work scheduling and monitoring. The main-control program assigns tasks to other workers and assigns *Map* or *Reduce* task to idle workers. In the mapping of the key-value pairs, the data of the worker assigned by *Map* task is read in the input subblock. The record provided by *Map-Reduce* is adopted to extract the key-value pairs and use the  $\text{map}(\cdot)$  function to map and obtain the key-value pair and then save it in memory. And then, the block function is used to divide the key-value pairs into several blocks, which are written to the local memory at intervals. At the same time, the location information of each block on the memory is sent to the main-control program and then is sent to the workers for *Reduce* task through the main-control program. The worker for *Reduce* task reads the data according to the location information transmitted by the main-control program. After reading the data, it sorts the intermediate data and arranges the values with the same key together. After sorting, the worker transfers the intermediate values of each corresponding same key to the completed *Reduce* function, gathers the values together, generates new key-value pairs, and then writes them to the output file.

**4.2. Datasets and Parameter Setup.** The data source in our experiment is part of the product data and user behavior data from Jing Dong Mall, which is the largest independent business-to-consumer e-commerce business in China by transaction volumes [24–42]. The product data contains information about the product number, product attributes, and the brand category of the product. The user behavior data includes information about the user number, the product number of the user interaction, the interaction behavior, and the purchase indicators. The experiment uses the dataset of the cross-border e-commerce shopping guide platform with more purchase records. Firstly, the data is preprocessed, and the users are grouped according to the purchase records to get the records of each user's own purchase records. Delete the users whose purchase records are less than 100 times, and delete the commodities that have been purchased less than 60 times. After data preprocessing, a user-item matrix is established. If a commodity/item is purchased by a user, the corresponding matrix element is 1.

In order to simulate the actual commodity recommendation in the experiment, the dataset is sorted according

to the user's purchase time and is divided into a training set and testing set. That is to say, for all users, the purchase records are sorted according to time, 70% of the purchase records are taken as the training set, and the remaining 30% are used as the testing set. To sum up, the research dataset is shown in Table 2.

**4.3. Comparison Algorithms and Quantitative Evaluation Index.** When testing the performance of the proposed personalized recommendation algorithm in cross-border e-commerce shopping guide platform, the proposed algorithm is compared with several existing state-of-the-art algorithms, including Bayesian algorithm (Bay\_RM), random walk algorithm (Ran\_RM), and graph-node algorithm (Gra\_RM). The Bayesian algorithm establishes the optimization function under the feedback of users. It assumes that the weight of commodities purchased by users exceeds the weight of commodities not purchased. Random walk algorithm obtains parts of the random paths by randomly walking in the network and then describes it by collecting low dimensional vectors of the network nodes. The graph-node algorithm firstly transforms the users and commodities into the graph model, obtains the low dimensional vector identification of users and commodities, and then completes the recommendation.

The Bayesian algorithm is a relatively good algorithm in the existing recommendation algorithms. Random walk algorithm and graph-node algorithm are widely used recently. So the experiment takes these three algorithms as comparative algorithms.

The recommendation algorithms select the evaluation index according to the accuracy rate, recall rate, and the mean absolute error (MAE). The accuracy rate is the ratio of the accurate recommended commodities to the total recommended commodities, and the recall rate is the ratio of the accurate recommended commodities to the total commodities.

The calculation equation of average accuracy is as follows:

$$\text{Accuracy} = \frac{\sum_{i=1}^N \left[ \frac{\sum_{j=1}^k q_j(i)/m_j}{N} \right]}{N}, \quad (9)$$

where  $m_j$  is the number of commodities purchased by user  $j$  in the cross-border e-commerce shopping guide platform;  $N$  is the total number of users; and  $q_j$  is the commodities purchased by users.

Mean absolute error (MAE) is also used to measure the accuracy of the prediction score, and the convergence of the iteration is verified. MAE is a method to evaluate the quality of the recommendation system. The smaller the value of MAE, the higher the prediction accuracy of the algorithm; otherwise, the worse the prediction accuracy. The equation of MAE is as follows:

$$\text{MAE} = \frac{\sum_{i=1}^C |P_i - Q_i|}{C}, \quad (10)$$

where  $P_i$  is the predicted score of item  $i$ ;  $Q_i$  is the true score of item  $i$ ; and  $C$  is the number of user's scores on item  $i$  in the testing set.

**4.4. Ablative Analysis.** In order to verify that the cross-border e-commerce personalized recommendation model of fuzzy association specification (FAS) combined with complex preference (CP) proposed in this paper can promote the recommendation performance, the performance of the complete algorithm with its simplified version is analyzed. Fuzzy association specification combined with complex preference is abbreviated as FASC\_RM, fuzzy association rule is abbreviated as FA\_RM, complex preference is abbreviated as CP\_RM, and the original recommendation algorithm is abbreviated as RM, as shown in Tables 3 and 4 that show the ablation analysis results of different modules.

It can be seen that FASC\_RM achieves higher recommendation performance than the simplified version. Specifically, as for FAS\_RM and CP\_RM in the same dataset, the accuracy of FASC\_RM is improved by 3.2%, and the recall rate and MAE are decreased by 2.9% and 10.3%, respectively. According to the experimental results, the proposed model can significantly improve the recommendation performance.

According to Table 3, the recommendation accuracy, recall rate, and average accuracy of the Bayesian algorithm are all lower than those of the other three algorithms, while the evaluation indexes of the proposed algorithm are higher than those of other comparison algorithms, which shows that it has high recommendation accuracy. It can be seen from the above experiments that the prediction recommendation error of the recommendation algorithm proposed in this paper is smaller, the recommendation accuracy rate is higher, and the recommendation quality of the recommendation system can be significantly improved.

In order to verify the effect of the recommendation algorithm proposed in this paper on different datasets, the experiment divides the dataset used in Table 2 into 6 parts and compares them with different comparison algorithms. The experimental results are shown in Table 5. It can be seen from Table 5 that as the dataset increases, the data becomes more and more sparse, and the error of the original recommendation algorithm becomes larger and larger, but the proposed algorithm in this paper still has a good accuracy rate. This is because although the dataset becomes larger, the number of users increases, the number of products becomes larger, the rating data becomes more, and the user rating data for the same commodity becomes less, the number of commodity types has not changed and the user's preference still exists. Combined with the existing state-of-the-art preference model in this paper, the prediction accuracy on datasets with different sizes is higher than the comparison

TABLE 2: Experimental dataset.

|                | User   | Items  | Purchase record |
|----------------|--------|--------|-----------------|
| Training set   | 82564  | 696561 | 1598452         |
| Testing set    | 23511  | 233322 | 825632          |
| Validation set | 1258   | 25811  | 35187           |
| Total          | 107333 | 955694 | 2459271         |

TABLE 3: Characterization of different recommendation model.

| Model   | FAS | CP |
|---------|-----|----|
| FASC_RM | √   | √  |
| CP_RM   | ×   | √  |
| FAS_RM  | √   | ×  |
| RM      | ×   | ×  |

TABLE 4: Ablative analysis.

| Model   | MAE   | Recall | Accuracy |
|---------|-------|--------|----------|
| FASC_RM | 0.087 | 0.1215 | 0.8325   |
| CP_RM   | 0.141 | 0.1328 | 0.8215   |
| FAS_RM  | 0.102 | 0.1285 | 0.7922   |
| RM      | 0.05  | 0.1508 | 0.7798   |

recommendation algorithm, so our proposed recommendation algorithm is more practical.

**4.5. Qualitative and Quantitative Analysis.** The user's purpose in practical applications is to get product recommendations, not just to know the rating status of the products. In order to verify the recommendation accuracy rate of our proposed algorithm in this paper, this paper calculates the proportion of the products that users have grading and the top 10 products with high ratings, calculates the accuracy rate of recommendation under different user-number sets, and make a comparison with the three recommendation algorithms. The results are shown in Tables 6 and 7; it can be seen that the accuracy of our proposed algorithm in this paper is 8.22% higher than that of the original Bay\_RM algorithm. As can be seen from these experimental data, our proposed algorithm is the best on the whole. With the increase of products, its precision and recall rate did not decrease much.

On the basis of the above analysis, the recommendation accuracy, recall rate, and average accuracy of 1258 users are compared, respectively, using the proposed algorithm in this paper, Bayesian algorithm, random walk algorithm, and graph-node algorithm. The results are shown in Table 5. Compared with the traditional Bayesian algorithm, the random walk algorithm recommendation algorithm, the improved algorithm reduces the impact of data sparsity and has a better recommendation effect.

TABLE 5: Quantitative indicators of different algorithms.

| Datasets | Bay_RM |          |       | Ran_RM |          |       | Gra_RM |          |       | FASC_RM |          |       |
|----------|--------|----------|-------|--------|----------|-------|--------|----------|-------|---------|----------|-------|
|          | Recall | Accuracy | MAE   | Recall | Accuracy | MAE   | Recall | Accuracy | MAE   | Recall  | Accuracy | MAE   |
| A        | 0.582  | 0.75     | 0.121 | 0.576  | 0.75     | 0.132 | 0.689  | 0.78     | 0.099 | 0.621   | 0.83     | 0.087 |
| B        | 0.625  | 0.80     | 0.174 | 0.602  | 0.79     | 0.166 | 0.725  | 0.81     | 0.089 | 0.585   | 0.75     | 0.099 |
| C        | 0.548  | 0.76     | 0.227 | 0.548  | 0.75     | 0.152 | 0.705  | 0.75     | 0.095 | 0.581   | 0.81     | 0.082 |
| D        | 0.615  | 0.79     | 0.101 | 0.625  | 0.75     | 0.099 | 0.709  | 0.82     | 0.093 | 0.588   | 0.85     | 0.094 |
| E        | 0.637  | 0.83     | 0.092 | 0.608  | 0.81     | 0.088 | 0.668  | 0.87     | 0.698 | 0.605   | 0.88     | 0.098 |
| F        | 0.485  | 0.69     | 0.185 | 0.459  | 0.666    | 0.182 | 0.491  | 0.71     | 0.211 | 0.445   | 0.64     | 0.017 |

TABLE 6: Comparison of the recall rate of different methods.

| Model   | Commodities' number |                  |                  |                  |
|---------|---------------------|------------------|------------------|------------------|
|         | $5 \times 10^3$     | $10 \times 10^3$ | $15 \times 10^3$ | $20 \times 10^3$ |
| FASC_RM | 0.5812              | 0.6670           | 0.7123           | 0.7658           |
| Gra_RM  | 0.5631              | 0.5848           | 0.6012           | 0.5231           |
| Ran_RM  | 0.3978              | 0.4065           | 0.4675           | 0.4323           |
| Bay_RM  | 0.3129              | 0.3767           | 0.3898           | 0.3711           |

TABLE 7: Comparison of the accuracy of different methods.

| Model   | Commodities' number |                  |                  |                  |
|---------|---------------------|------------------|------------------|------------------|
|         | $5 \times 10^3$     | $10 \times 10^3$ | $15 \times 10^3$ | $20 \times 10^3$ |
| FASC_RM | 0.9324              | 0.9168           | 0.8953           | 0.8786           |
| Gra_RM  | 0.8541              | 0.8622           | 0.8724           | 0.8798           |
| Ran_RM  | 0.8434              | 0.8110           | 0.8485           | 0.8643           |
| Bay_RM  | 0.7473              | 0.7556           | 0.7241           | 0.7007           |

## 5. Conclusion

Since cross-border e-commerce involves the export and import of commodities, it is affected by many policies and regulations, resulting in some special requirements for the recommendation system, which makes the traditional collaborative filtering recommendation algorithm less effective for the cross-border e-commerce recommendation system. To address this issue, a simple yet effective cross-border e-commerce personalized recommendation is proposed in this paper, which integrates fuzzy association rule and complex preference into a recommendation model. Compared with the traditional collaborative filtering recommendation algorithm, the improved algorithm reduces the impact of data sparsity. The experiment also verifies that the improved fuzzy association rule algorithm has a better recommendation effect than the traditional recommendation models.

## Data Availability

The labeled datasets used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

- [1] G. Sansonetti, "Point of interest recommendation based on social and linked open data," *Personal and Ubiquitous Computing*, vol. 18, no. 4, pp. 1–16, 2019.
- [2] X. Li, H. Hong, and L. Liao, "A time-aware personalized point-of-interest recommendation via high-order tensor factorization," *ACM Transactions on Information Systems*, vol. 35, no. 4, pp. 1–23, 2017.
- [3] L. Jiang, S. J. Pan, and Q. Yang, "A unified framework of active transfer learning for cross-system recommendation," *Artificial Intelligence*, vol. 245, pp. 38–55, 2017.
- [4] X. Zhang, L. Zhang, F. Cheng, Y. Jin, and K. Tang, "Pattern recommendation in task-oriented applications: a multi-objective perspective [application notes]," *IEEE Computational Intelligence Magazine*, vol. 12, no. 3, pp. 43–53, 2017.
- [5] E. Duan, "Application of computational intelligence to improve education in smart cities," *Sensors*, vol. 18, pp. 328–334, 2018.
- [6] Z. Yu, R. Wong, and C. H. Chi, "Efficient role mining for context-aware service recommendation using a high-performance cluster," *IEEE Transactions on Services Computing*, vol. 10, no. 6, pp. 914–926, 2015.
- [7] W. Jun, "Crowd sensing-enabling security service recommendation for social fog computing systems," *Sensors*, vol. 17, no. 8, pp. 1744–1756, 2017.
- [8] A. Zafar, "A robust missing data-recovering technique for mobility data mining," *Applied Artificial Intelligence*, vol. 31, no. 3, pp. 1–14, 2017.
- [9] J. I. Wendi, X. Wang, and F. Zhu, "Time-aware conversion prediction," *Frontiers of Computer Science*, vol. 11, no. 4, pp. 1–15, 2017.
- [10] L. Shudong, "User modeling for point-of-interest recommendations in location-based social networks: the state of the art," *Mobile Information Systems*, vol. 20, no. 18, pp. 1–13, 2018.
- [11] H. Tian, Y. Cai, Y. Huang et al., "Distributed steganalysis of compressed speech," *Soft Computing*, vol. 21, no. 3, pp. 795–804, 2017.
- [12] O. Wu, A. L. Sadek, and D. Landa-Silva, "An evolutionary strategy with machine learning for learning to rank in information retrieval," *Soft Computing*, vol. 24, no. 11, pp. 58–69, 2018.
- [13] D. Bra, "Challenges in user modeling and personalization," *IEEE Intelligent Systems*, vol. 32, no. 5, pp. 76–80, 2017.
- [14] A. Klačnja-Milićević, "Enhancing e-learning systems with personalized recommendation based on collaborative tagging techniques," *Applied Intelligence*, vol. 1, pp. 1–17, 2017.
- [15] L. Jiang and C. C. Yang, "User recommendation in healthcare social media by assessing user similarity in heterogeneous network," *Artificial Intelligence in Medicine*, vol. 81, pp. 63–77, 2017.

- [16] Y. Zhang, S. Wang, K. Xia, Y. Jiang, and P. Qian, "Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion," *Information Fusion*, vol. 66, pp. 170–183, 2021.
- [17] K. Xia, Y. Zhang, Y. Jiang et al., "TSK fuzzy system for multi-view data discovery underlying label relaxation and cross-rule & cross-view sparsity regularizations," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, 2020.
- [18] Y. Zhang and J. Cai, "Fuzzy clustering based on automated feature pattern-driven similarity matrix reduction," *IEEE Transactions on Computational Social Systems*, vol. 99, pp. 1–10, 2020.
- [19] Y. Zhang, H. Ishibuchi, and S. Wang, "Deep takagi-sugeno-kang fuzzy classifier with shared linguistic fuzzy rules," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 3, pp. 1535–1549, 2018.
- [20] A. U. Joanna, "HyRA: a hybrid recommendation algorithm focused on smart POI. Ceuti as a study scenario," *Sensors*, vol. 18, pp. 890–902, 2018.
- [21] C. Zhang, L. Tiantian, R. Zhibin, H. Zheng, and J. Yang, "Taxonomy-aware collaborative denoising autoencoder for personalized recommendation," *Applied Intelligence*, vol. 24, no. 11, pp. 1256–1269, 2019.
- [22] R. Yera and L. Martínez, "A recommendation approach for programming online judges supported by data preprocessing techniques," *Applied Intelligence*, vol. 24, no. 19, pp. 369–377, 2017.
- [23] P. Zhang, Z. Zhang, T. Tian, and Y. Wang, "Collaborative filtering recommendation algorithm integrating time windows and rating predictions," *Applied Intelligence*, vol. 24, no. 1, pp. 24–37, 2019.
- [24] S. Feng, "Merging user social network into the random walk model for better group recommendation," *Applied Intelligence*, vol. 24, no. 16, pp. 2009–2026, 2018.
- [25] T. Xiao and H. Shen, "Neural variational matrix factorization for collaborative filtering in recommendation systems," *Applied Intelligence*, vol. 24, no. 11, pp. 1298–1310, 2019.
- [26] M. Rahman and J. C. Oh, "Graph bandit for diverse user coverage in online recommendation," *Applied Intelligence*, vol. 2-3, pp. 1–17, 2017.
- [27] S. Xing, "Points-of-interest recommendation based on convolution matrix factorization," *Applied Intelligence*, vol. 24, no. 11, pp. 63–79, 2017.
- [28] C. Song, "Toward privacy-preserving personalized recommendation services," *Engineering*, vol. 18, 2018.
- [29] L. Xing, Q. Ma, and L. Jiang, "Microblog user recommendation based on particle swarm optimization," *China Communications*, vol. 14, no. 5, pp. 134–144, 2017.
- [30] R. Logesh and V. Subramaniaswamy, "A reliable point of interest recommendation based on trust relevancy between users," *Wireless Personal Communications*, vol. 20, 2017.
- [31] A. Sajad, M. Majid, and A. Mohsen, "Incorporating reliable virtual ratings into social recommendation systems," *Applied Intelligence*, vol. 24, no. 2, pp. 13–23, 2018.
- [32] Q. Pan and X. Wang, "Independent travel recommendation algorithm based on analytical hierarchy process and simulated annealing for professional tourist," *Applied Intelligence*, vol. 64, no. 12, pp. 2100–2209, 2017.
- [33] P. M. Chuan, "Link prediction in co-authorship networks based on hybrid content similarity metric," *Applied Intelligence*, vol. 12, no. 3, pp. 119–129, 2017.
- [34] A. M. Dakhel, H. T. Malazi, and M. Mahdavi, "A social recommender system using item asymmetric correlation," *Applied Intelligence*, vol. 48, no. 3, pp. 527–540, 2018.
- [35] X. Ren and J. Song, "Context-aware probabilistic matrix factorization modeling for point-of-interest recommendation," *Neurocomputing*, vol. 241, pp. 38–55, 2017.
- [36] J. G. DuBois and A. Srinivasan, "Predicting Trust and Distrust in Social Networks," in *Proceedings of the Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, IEEE, Boston, MA, USA, pp. 418–424, October 2011.
- [37] K. Oh, J.-W. Kim, S.-W. Kim, and K. Lee, "A probability-based trust prediction model using trust-message passing," in *Proceedings of the 22nd International Conference on World Wide Web Companion*, pp. 161–162, Rio de Janeiro, Brazil, November 2013.
- [38] A. Jiliang Tang, "Recommendations in Signed Social Networks," in *Proceedings of the 25th International World Wide Web Conference (WWW)*, pp. 31–40, Montreal, Canada, April 2016.
- [39] B. Hui Fang, "Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pp. 30–36, Québec, Canada, July 2014.
- [40] J. T. Suhang Wang, "Exploring implicit hierarchical structures for recommender systems," in *Proceeding of the international joint conference on artificial intelligence(IJCAI)*, Buenos Aires, Argentina, July 2015.
- [41] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel recommendation model regularized with user trust and item ratings," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 7, pp. 1607–1620, 2016.
- [42] D. Y. Jiliang Tang and S. Wang, "Recommendation with social dimensions," in *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, pp. 251–257, Phoenix, AZ, USA, February 2016.