

Design of a Hybrid Recommendation Algorithm based on Multi-objective Collaborative Filtering for Massive Cloud Data

Xiaoli Zhou

Sichuan Top IT Vocational Institute, School of Information Engineering, Chengdu, 610000, China

Abstract—The current recommendation technology has some problems, such as lack of timeliness, the contradiction between recommendation diversity and accuracy. In order to solve the problem of lack of timeliness, the time factor is introduced when constructing the self-preference model. The cold start problem in the collaborative filtering algorithm is solved by the hybrid similarity calculation method, and the potential preference model is constructed. The two are fused to obtain a hybrid recommendation algorithm to improve the recommendation performance of the algorithm. For the problem of multi-objective contradiction, the NNIA algorithm is used to further optimize the candidate results of mixed recommendation, and the final recommendation list is obtained. Through verification experiments, the results show that the recall rate and accuracy of the fused preference model are better than those of the non-fused model, and the accuracy is 9.57% and 8.23% higher than that of SPM and PPM, and the recall rate is 9.97% and 7.65% higher, respectively. CBCF-NNIA algorithm has high accuracy and diversity of recommendation, and can provide users with rich and diverse text content to meet their own needs.

Keywords—Self preference; collaborative filtering; potential preferences; mixed recommendation; multi-objective optimization

I. INTRODUCTION

With the advent of the information age, data has become a decisive factor in the development of the industry, and any decision needs to rely on data to speak. In the face of massive data resources, storing them in the cloud to form cloud data that is easy to manage is beneficial for users to access relevant information [1]. The constant development of current Internet technology and communication technology has led to a rapid increase in the amount and speed of information and data dissemination, making it impossible for users to precisely find useful information and making a large amount of information unavailable, which leads to information overload [2]. The personalized recommendation system can analyze the user's interest preferences according to the user's behavior habits, and recommend the content related to the user's interest to the user without the user's initiative to provide information. The excellent performance makes the recommendation system widely used in many fields [3-4]. At present, there are many commonly used recommendation methods. Collaborative Filtering (CF) recommendation is the earliest recommendation method. However, with the growth of data size, data sparsity, system could start and other problems are unavoidable. The hybrid recommendation algorithm can integrate multiple recommendation algorithms, and research the combination of

collaborative filtering and content recommendation method to form a cascade hybrid recommendation algorithm. It is expected that the hybrid recommendation algorithm can effectively alleviate the cold start problem of the system and improve the accuracy of the recommendation results. In view of the contradiction between the diversity and accuracy of the recommendation results, the multi-objective immune optimization algorithm is studied and introduced to further optimize the recommendation list. It is expected that this method can find the best recommendation results, which can achieve the diversity of the results while ensuring the accuracy of the recommendation, and meet the multiple needs of users.

II. RELATED WORK

The research team of Tian proposes a book recommendation system based on a hybrid recommendation algorithm to address the problem of users finding appropriate books quickly. The system combines collaborative filtering with content recommendation algorithms, improves the user item rating matrix, and uses clustering algorithms to solve the data sparsity problem. Through practical application, the results show that the hybrid recommendation method studied can provide users with more accurate book recommendations [5]. Wang scholars have studied a hybrid recommendation algorithm based on interest models in order to improve user satisfaction on e-commerce websites. The algorithm uses collaborative filtering to mine users' potential interests and a content recommendation algorithm to construct a model of users' existing interests, and the two are combined to recommend highly accurate and interesting products for users. Experimental results show that the hybrid recommendation algorithm can provide a better service experience for users [6]. Jiao W et al. faced the problems of data sparsity and cold start in traditional recommendation algorithms and used K-means, weighting for optimization, while introducing adjustment factors to combine collaborative filtering with dichotomous networks. Relevant experimental data show that the hybrid recommendation algorithm studied is operable and has better recommendation accuracy than the comparison algorithm [7]. Liu et al. propose a personalized service recommendation system in order to improve the competitiveness of manufacturing service platforms, and use a hybrid algorithm to solve the composite service problem. In this algorithm, customer preferences are quantified by clustering, and composite services are optimally ranked by ranking genetics. The analysis results of real cases show that the performance of

the studied recommendation algorithm is good and has some practical value [8].

Nafis et al. studied a travel platform based on a hybrid recommendation algorithm in order to help tourists make personalized travel plans. The platform can recommend relevant resources according to tourists' preferred tourist attractions, and through relevant experimental analysis, the platform provides high accuracy of information resources, which can effectively promote the development of read tourism [9]. Pirasteh et al. team members addressed the situation that collaborative filtering has a cold start, which leads to poor recommendation effect, by capturing various similarities between items and finding out hidden preferences in items to alleviate the problem due to the number of items. The team members have been able to mitigate the low quality of recommendations due to insufficient number of items by capturing various similarities between items and finding hidden preferences in items. Simulation experimental data show that the diversity of similarities can provide more reliable results for users and achieve high quality personalized recommendations [10]. Hu The researchers studied an improved particle swarm optimization algorithm based on multiple criteria combined with diverse adaptations, while using bacterial foraging to improve the convergence of the algorithm in the face of problems such as multi-objective contradictions and unsatisfied constraints in solving hybrid recommendation models. The experimental results show that the convergence and diversity of the studied algorithm are better than the comparison algorithm, and the results of the recommended model solving are highly accurate, have wide coverage, and can provide users with diverse results [11]. Ajaegbu scholars propose an algorithm to improve similarity measurement based on traditional measures in order to alleviate the problems of data sparsity and cold start of collaborative filtering. The performance of the algorithm is analyzed on different data, and experimental results show that the algorithm retains the advantages of existing measures while mitigating the disadvantages of traditional methods [12].

By summarizing the achievements of domestic and foreign researchers, researchers have proposed different improvement measures for the shortcomings of existing recommendation algorithms. Among them, the hybrid recommendation algorithm can effectively integrate the advantages and disadvantages of different recommendation algorithms to achieve accurate and efficient target recommendation, but the contradiction between accuracy and diversity needs further research. Therefore, the research will combine content-based recommendation and collaborative filtering to form a cascaded hybrid recommendation algorithm in order to achieve highly accurate recommendation results. At the same time, in the face of the contradiction between diversity and accuracy, the immune optimization algorithm is introduced into the recommendation results. It is expected that the multi-objective hybrid recommendation algorithm studied can achieve accurate recommendation and multiple needs of users.

III. HYBRID RECOMMENDATION ALGORITHM FOR CLOUD DATA BASED ON MULTI-OBJECTIVE COLLABORATIVE FILTERING

A Recommendation Algorithms based on Own Preference Models

The rapid development of information technology has intensified the problem of data overload, and the massive amount of data has caused many problems for users. In the constant search for solutions, personalized recommendation has emerged, and recommendation systems can effectively filter distracting information and improve data usage [13-14]. Collaborative filtering is a classical algorithm in recommendation systems. The core idea is to find the nearest neighbors by calculating the similarity, and to find the content of interest to the user in the nearest neighbors and recommend it to the target user. Commonly used calculation formulas are such as modified cosine similarity, Pearson, Jaccard similarity, etc. Through similar user groups, some potentially preferred content will also be recommended to the target customer, reflecting the diversity of recommendation results of collaborative filtering [15-16]. As the data expands, the likelihood of different users being interested in the same content among themselves gradually decreases, and the problem of data sparsity cannot be avoided. In addition, when a new user or item appears, relevant attribute information does not exist and the system is unable to make a recommendation resulting in a cold start problem. In the field of textual data recommendation, Content Based Recommendations (CBR) is based on the analysis of users' historical data, and by constructing interest models, the similarity between the interest model and the content of the item reflects the user's preference for the content. However, the recommendation results lack some diversity and cannot effectively explore the potential preferences of users [17]. Therefore, in the field of text data recommendation, research will combine CBR and CF to form a hybrid recommendation algorithm, effectively integrate the advantages of the two, and achieve personalized and diversified content recommendation. At the same time, there is a contradiction between the diversity and accuracy of the recommendation results. In order to alleviate the contradiction between the two, the text data recommendation problem is transformed into a multi-objective optimization problem through modeling, and the immune optimization algorithm is used to optimize the candidate text set to generate the final recommendation list.

The Self Preference Model (SPM) is built on the basis of CBR, which can intuitively reflect the user's own interests, and the whole construction process is shown in Fig. 1. The text features are obtained by pre-processing the user's browsing history, using a vector space model to represent the text features, introducing a time factor to adjust the text model, and finally obtaining the user's own preference model.

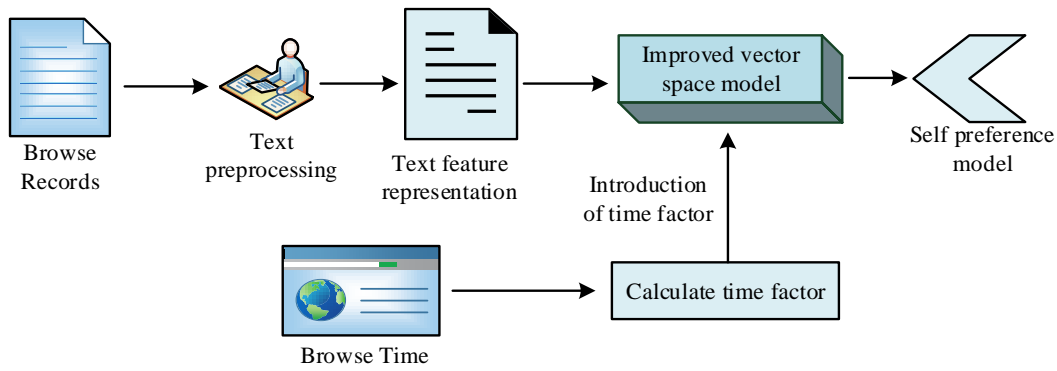


Fig. 1. Construction process of self-preference model.

In the text pre-processing stage, special characters such as emoticons and face characters are removed, and then the text is divided into sequences of word order using the word separation technique jieba. To avoid excessive word sequences that increase the computational effort, meaningless word sequences such as conjunctions, auxiliaries and stop words are usually screened out of the feature word sequences, while verbs, nouns and adjectives are retained in the word sequences. Representation of text features is a key technique in CBR, where key content features are used to represent text information and build a text model. Suppose there is a text dataset $H = \{h_1, \dots, h_a, \dots, h_m\}$, m is the total amount of text, after the textual sub-word processing, the main feature word order of the text set is g dimensional composition vector $F = \{f_1, \dots, f_j, \dots, f_g\}$, the feature vector of any text h_a is $h_a = \{w_{a1}, \dots, w_{aj}, \dots, w_{ag}\}$, w_{aj} represents the weight of the feature word in the text h_a , the higher the weight value, the more important the feature word f_j in the text h_a . The TF-IDF (Term Frequency-Inverse Document Frequency) representation is often used to calculate the weights of feature word sequences in different texts, and the TF is used to count the frequency of feature words in a text set.

$$TF(a, j) = \frac{f(a, j)}{\max Others(a, j)} \quad (1)$$

In eq. (1), $TF(a, j)$ is the frequency of the feature word in the text h_a , $f(a, j)$ is the absolute frequency of the feature word in the text h_a , and $\max Others(a, j)$ is the most frequent feature word in the text h_a . The IDF is used to represent the distribution of the feature word in the document and to count the frequency of the feature word, which is calculated in eq. (2).

$$IDF(j) = \log\left(\frac{M}{M(j)}\right) \quad (2)$$

In eq. (2), M represents the number of all texts and $M(j)$ represents the number of texts where the feature term j is present. $IDF(j)$ The larger the number, the lower the number of occurrences of the feature word in the text set. The formula for calculating the weight of a feature word w_{aj} is shown in eq. (3).

$$w_{aj} = TF(a, j) * IDF(j) \quad (3)$$

After calculating the weights of all feature words, the weight matrix HM of the text set H is obtained as shown in eq. (4).

$$HM = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1g} \\ w_{21} & w_{22} & \dots & w_{2g} \\ \vdots & \vdots & \dots & \vdots \\ w_{m1} & w_{m1} & \dots & w_{mg} \end{bmatrix} \quad (4)$$

Using the text that the user has read as a feature vector, the weight matrix HM_u for the set of texts that the user u has read is shown in eq. (5).

$$HM_u = \begin{bmatrix} w_{u11} & w_{u12} & \dots & w_{u1g} \\ w_{u21} & w_{u22} & \dots & w_{u2g} \\ \vdots & \vdots & \dots & \vdots \\ w_{um1} & w_{um1} & \dots & w_{umg} \end{bmatrix} \quad (5)$$

As users' interests change, text recommendations need to be time-sensitive, and analyzing users' recent behavioral data can improve prediction accuracy. Therefore, the study introduces a time-factor adjustment model matrix to obtain users' real-time interest preferences. Define the set of texts read by the user u H_u , t_0 is the current time, t_i is the user's browsing time, and the time factor is calculated by eq. (6).

$$\varepsilon_i = \frac{e^{-(t_0-t_i)}}{\sum_{h_{ij} \in H_u} e^{-(t_0-t_j)}} \quad (6)$$

In eq. (6) e is a constant. The introduction of a time factor to adjust the weights of the weight matrix HM_u gives the user's own preference model as $SPM_u = (W1_{u1}, W1_{u2}, \dots, W1_{ug})$, where $W1_{ua} = \sum_{h_{aa} \in H_u, w_{uaj} \in HM_u} \epsilon_i * w_{uaj}$.

B Hybrid Recommendation Algorithm with Fused Preference Models

In practice, if only one's own preferences are considered, the recommendation results lack a certain degree of richness, and it is necessary to explore the potential preferences of users to improve user satisfaction [18]. The Potential Preference Model (PPM) is built by extracting the user's browsing history to form a matrix of user behavior and finding nearest neighbors through a mixture of behavioral and content similarity calculations to solve the problem of not being able to categories similar users due to text diversity [19]. When using CF for recommendation, it is transformed to recommend the feature words of interest to the nearest neighbors, which can effectively avoid the cold start problem, and the construction process is shown in Fig. 2.

In the user behavior matrix, suppose the user set is $U = \{u_1, u_2, \dots, u_n\}$, n is the total number of users and the text set is $H = \{h_1, h_2, \dots, h_m\}$, according to the user's browsing history, when the user u reads the text a , then $x_{ua} = 1$, and vice versa, the user's behaviour matrix $X_{n \times m}$ can be expressed as eq. (7).

$$X_{n \times m} = \begin{bmatrix} x_{11} & \dots & x_{1a} & \dots & x_{1m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{u1} & \dots & x_{ua} & \dots & x_{um} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & \dots & x_{na} & \dots & x_{nm} \end{bmatrix} \quad (7)$$

The user's level of interest will change over time. To calculate the similarity of user behavior if only the same behavior exists between users is considered, ignoring the time

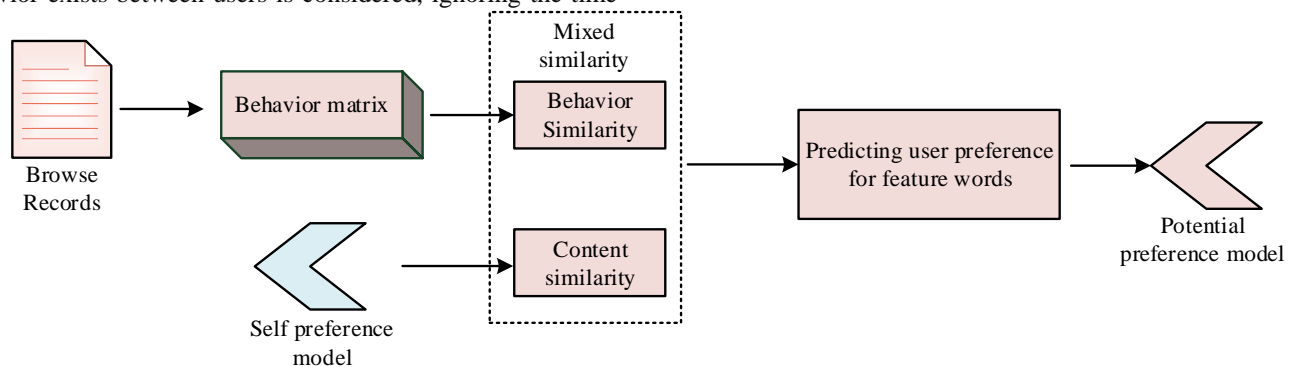


Fig. 2. Potential preference model construction process.

difference in user browsing will affect the level of interest in text recommendations, the time difference in user browsing behavior is defined as time decay, and its formula is eq. (8).

$$t_a = 1 / (1 + \delta |t_{ua} - t_{ra}|) \quad (8)$$

In eq. (8), δ represents the time decay factor, t_{ua} and t_{ra} represent the time that users u and r spent viewing the same text a , and $|t_{ua} - t_{ra}|$ represents the time difference between users u and r viewing the text a . The smaller the difference, the greater the similarity between users. At the same time, the influence of the external environment can also lead to user interest, for example, the coverage of popular events will make users interested in the field for a short period of time, and after the heat has passed, users' interest will return to its original state. Therefore, we need to take the external heat factor into account, and define the heat difference between users reading the same text at different times as k_a , which is calculated by eq. (9).

$$k_a = 1 / \log(1 + |N(a)|) \quad (9)$$

In eq. (9), $N(a)$ represents the set of users who have viewed the text a . The more people read the text, the hotter the text is and the more likely users are to view the text due to its hotness, and the larger $N(a)$ is, the greater the difference in hotness k_a is. Therefore, the formula for calculating the similarity of user behavior, taking into account the temporal differences in the generation of browsing behavior and the heat of the text, is eq. (10).

$$sim_{act}(u, r) = \frac{\sum_{a \in H(u) \cap H(r)} \eta * t_a + (1 - \eta) * k_a}{\sqrt{|H(u)| |H(r)|}} \quad (10)$$

In eq. (10), $H(u)$ and $H(r)$ represent the set of texts viewed by the user u and r , and η represents the moderator, with larger values indicating a greater influence of temporal differences and smaller values indicating a greater influence of text hotness, $\eta \in [0,1]$. Content similarity is calculated by the own preference model, and the user's content similarity is calculated by eq. (11).

$$sim_{con}(u,r) = \frac{PSM_u * PSM_r}{|PSM_u| \times |PSM_r|} \quad (11)$$

By combining the behavioural similarity with the content similarity through a weighting factor, the formula for the mixed similarity is eq. (12).

$$sim(u,r) = \chi sim_{act}(u,r) + (1-\chi) sim_{con}(u,r) \quad (12)$$

In eq. (12), χ represents the weighting factor and $\chi \in [0,1]$. When $\chi = 0$ only content features are considered, the larger χ is, the greater the degree of influence of user behavioural features. The set of nearest neighbours of the user u is $N(u)$ and the potential preference model is $PPM_u = (W2_{u1}, W2_{u2}, \dots, W2_{ug})$, where $W2_{ua} = \sum_{r \in N(u)} \frac{sim(u,r)}{\sum_{r \in N(u)} sim(u,r_x)} W1_{ra}$ is used to find the nearest neighbors of the user's mixed similarity.

SPM can reflect the user's own preferences and has obvious personalization characteristics, while PPM can expand the user's interests and has diversity characteristics. The fusion of the two can ensure that the recommended content can meet the user's own preferences and enhance the diversity of content. When fusing SPM and PPM, the measurement between personalization and diversity should not be simply mixed. $\max()$ function can take the feature vector with greater weight in the two models as the final preference vector of the user, which is selected as the fusion function, and the Fusion Preference Model (FPM) constructed is $FPM_u = (W3_{u1}, W3_{u2}, \dots, W3_{ug})$, where $W3_{ua} = \max(W1_{ua}, W2_{ua})$.

A fusion preference model is used to calculate the similarity between users, select users with greater similarity as nearest neighbors, predict the interest of users in candidate text sets from the browsing records of nearest neighbors, and recommend results according to the degree of preference, and the preference prediction formula is eq. (13).

$$P(u,a) = \sum_{r \in N(u)} sim(u,r) x_{ra} \quad (13)$$

In eq. (13), x_{ra} indicates whether the user r has generated a view record for the text a , and if so, $x_{ra} = 1$, and if not, $x_{ra} = 0$.

C NNIA-based Multi-objective Optimization Recommendation Algorithm

In recommendation systems, there is a conflicting status quo between the diversity and accuracy of recommendation results. To alleviate the conflict between the two, the text data recommendation problem is transformed into a multi-objective optimization problem through modelling. Multi-objective optimization problems are common in everyday life. In solving multi-objective problems, the objectives are so interlocked that finding the best solution that satisfies all the objectives is impossible, so a trade-off between the objectives is needed to find the relatively better solution [20]. In the field of textual data recommendation, the first consideration is the accuracy of the recommendation results. High accuracy can improve user satisfaction and also measure the level of interest in the recommendation results. The objective function for accuracy is chosen as the similarity matching function and is calculated by eq. (14).

$$\begin{cases} y_v = \frac{1}{L} \sum_{c \in R} 1 - sim(c, P_u) \\ sim(c, P_u) = \frac{1}{|P_u|} \sum_{a \in P_u} sim(c, a) \end{cases} \quad (14)$$

In eq. (14), L represents the length of the list of all recommendations, P_u represents the set of texts that the user u has viewed, and $sim(c, a)$ represents the similarity between the text c and the text a . y_v The smaller the value, the more similar the recommendation is to the text the user has viewed and the better the accuracy of the recommendation. To help users find more interesting text data, diversity is taken into account and the objective function of diversity is eq. (15).

$$y_D = \frac{1}{L(1-L)} \sum_{c, a \in R, c \neq a} sim(c, a) \quad (15)$$

From eq. (15), the smaller y_D is, the smaller the similarity between the texts in the recommendation list, which means the richness and diversity of the recommendation content is better. Considering diversity and accuracy, it is found that the smaller the y_v and y_D , the better. Therefore, the final objective function is minimization, and the specific expression is eq. (16).

$$Min : Y = \{y_v, y_D\} \quad (16)$$

The Non-dominated Neighborhood Immune Algorithm (NNIA) has excellent evolutionary performance and is used to solve multi-objective problems due to its fast convergence and robustness. A certain proportion of each superior antibody is cloned, and the population is then crossed and mutated to enhance the diversity of the population and avoid local optima. During the crossover operation, the same gene loci in the two parent individuals are inherited, and the remaining positions are randomly crossed with a probability of [0,1]. The mutation operation is carried out for the parent individuals. The crossover and mutation operations are shown in Fig. 3.

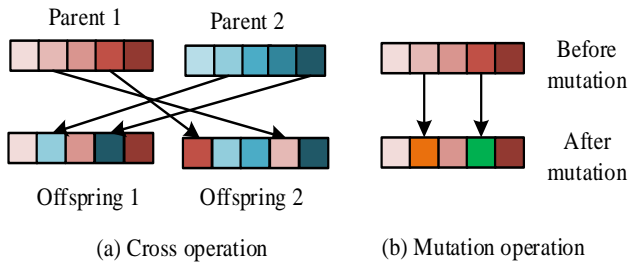


Fig. 3. Crossover and mutation operation.

Using the NNIA algorithm, it is necessary to encode the texts in the candidate set with real numbers. Let the encoding of an antibody be $z\{z_1, K, z_b, K, z_l\}$, z denotes any user, z_b denotes the recommended text number, and each text has a unique number. The candidate text set obtained from the fusion preference model is optimized by crossover and variation operations to produce the final recommendation list. The

specific process of text data recommendation using NNIA algorithm is shown in Fig. 4. The user text candidate set data program generated by the fusion preference model is used to encode the candidate text set in real numbers. Set the maximum number of iterations, population size, crossover and mutation probability and other related parameters to randomly generate the initial antibody population. Identify the dominant antibody and clone the dominant antibody. Place the cloned individuals in the temporary population. Select the individuals with the lowest density according to the population density to enter the new population and update the dominant population. The new dominant population is cloned, crossed, and mutated to continue to update the dominant population until the maximum number of iterations is reached, and finally the best recommendation list of text recommendation is output.

The NNIA algorithm can be used to generate a series of recommendation lists with different recommendation targets, personalized to the specific needs of the user, increasing user satisfaction and loyalty.

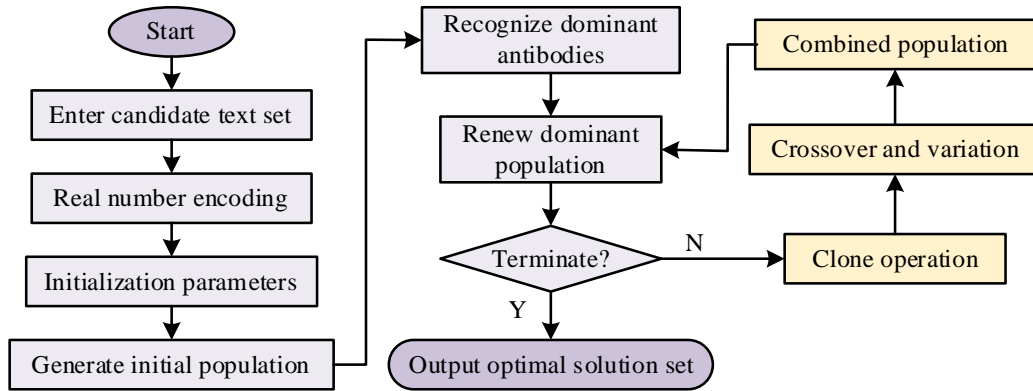


Fig. 4. NNIA algorithm text recommendation process.

IV. PERFORMANCE ANALYSIS OF MULTI-OBJECTIVE COLLABORATIVE FILTERING-BASED DATA-MIXING RECOMMENDATION ALGORITHMS

A Performance Validation of Hybrid Recommendation Algorithms Incorporating Preference Models

In order to determine the values of the parameters present in the latent preference model (PPM), namely the temporal

decay factor δ , the text heat modifier η and the mixed similarity weighting factor χ , Fig. 5 shows the variation in the accuracy of the recommendation results for different temporal decay factors δ and modifiers. η

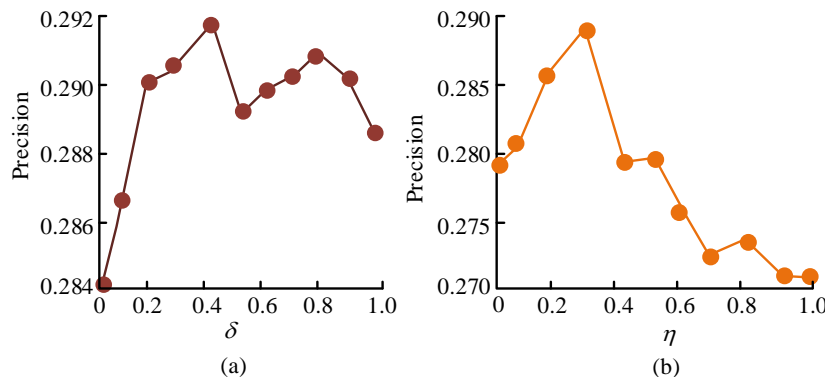


Fig. 5. Recommended accuracy when δ and η have different values.

As can be seen in Fig. 5(a), the accuracy of the recommendations fluctuates as the value of δ increases, and the experimental results show that there is a maximum accuracy of the recommendations at $\delta=0.4$. In Fig. 5(b), as the value of η increases, the recommended accuracy curve increases to a peak and then gradually decreases, showing the maximum accuracy at $\eta=0.3$. In the case of $\delta=0.4$ and $\eta=0.3$, the best weighting factor χ was analysed, and the accuracy curves for different values of χ are shown in Fig. 6.

As can be seen in Fig. 6, the best accuracy of the recommended results is obtained at $\chi=0.7$. The experimental analysis finally resulted in the accurate values of each parameter when constructing the PPM, i.e. $\delta=0.4$, $\eta=0.3$, $\chi=0.7$. In order to investigate whether the recommendation performance of the fusion preference model (CBCF) is better than before fusion, the study conducted comparison experiments with SPM and PPM, and Fig. 7 shows the comparison results of the accuracy and recall of the three models' recommendations.

As can be seen in Fig. 7(a), the trend of all three models decreases as the amount of recommended text increases, but the accuracy of CBCF is always higher than the comparison algorithms, with a maximum increase of 9.57% and 8.23% over the accuracy of SPM and PPM. As can be seen in Fig. 7(b), the pattern of change in recall is opposite to the pattern of change in accuracy, with the higher the number of recommended texts, the greater the recall. the recall of CBCF is higher than the other two models, with a 9.97% and 7.65% improvement over the SPM and PPM recall. The fusion of

models can effectively improve the accuracy of recommendations. Fig. 8 shows the F value and recommended coverage results of the three models.

As can be seen from Fig. 8(a), the change trend of the F value curve of the three models is similar. With the increase of the recommended text, the F value rises to the highest value, and then decreases gradually. The maximum F value of CBCF is 0.493, and the maximum values of SPM and PPM are 0.438 and 0.461 respectively. Compared with the comparison model, the F value of CBCF is 0.055 and 0.032 higher. Fig. 8(b) shows the results of recommendation coverage. It can be seen that with the increase of the number of recommended texts, the coverage of the three models is on the rise, and the coverage of CBCF is always higher than that of the comparison model. Compared with SPM and PPM, the coverage increased by 10.04% and 16.45%. Relevant data shows that the fusion of models can effectively improve recommendation performance.

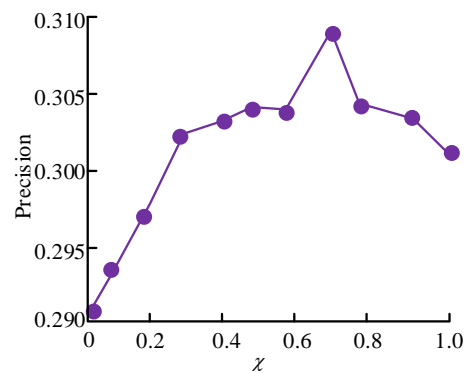


Fig. 6. χ Recommended accuracy rate for different values.

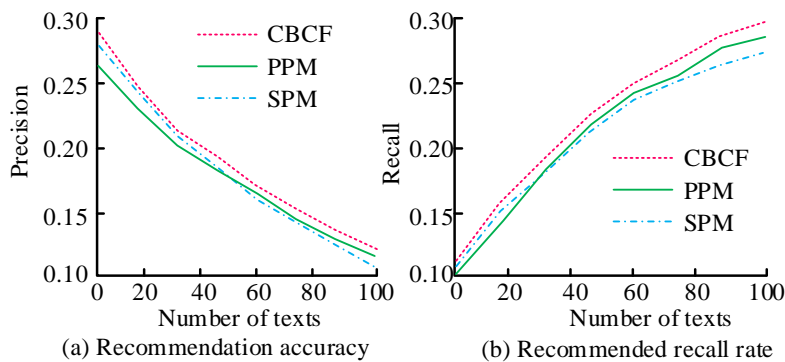


Fig. 7. Comparison results of accuracy and recall rate recommended by three models.

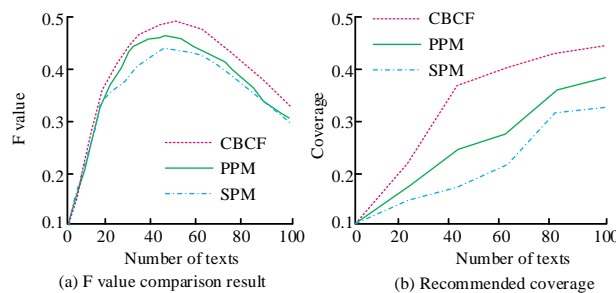


Fig. 8. F value and recommended coverage results of the three models.

B Analysis of the Effectiveness of the Hybrid Recommendation Algorithm based on NNIA Multi-objective Optimization

To verify the performance of the NNIA multi-objective optimization recommendation algorithm, the maximum number of iterations of the algorithm was set to 300, the number of nearest neighbors to 25 and the size of the dominant population to 50. The top 50 non-dominated solutions were selected as the recommendation result solution in NNIA according to the degree of density, and one user was randomly selected for presentation, and the results are shown in Fig. 9.

The distribution of the user 27 recommended solution set in the target space is illustrated in Fig. 9, where it can be seen that the NNIA optimized recommendation algorithm can effectively obtain 50 different solutions. The leftmost solution is the one with the highest accuracy and the worst diversity among all solutions, while the rightmost solution is the one with the best diversity and the lowest accuracy among all solutions. From the distribution of the solution set, as the value of the diversity function increases, the value of the accuracy function decreases, indicating that there is a contradiction between the two indicators in the recommendation process, which is in line with the law of the recommendation process and verifies the effectiveness of the NNIA multi-objective algorithm.

To analyses the performance of the studied NNIA multi-objective optimized hybrid recommendation algorithm (CBCF-NNIA), it was compared with Recommendations Based on User Browsing (BUB), Personalized recommendation based on user interest (PBUI) and Hybrid recommendation with fused preference model (CBCF) in terms of accuracy, diversity of recommendation results. The accuracy and diversity of recommendation results are compared. The accuracy and diversity of recommendation results are compared.

As can be seen in Table I, the mean accuracy of the recommendation lists of 9 out of 10 users outperformed the other three comparison algorithms, and the maximum accuracy of the recommendation lists of the remaining 1 user outperformed the comparison algorithms. The experimental

data shows that the recommendation list given by the CBCF-NNIA algorithm studied is not necessarily the most accurate, but can give more accurate recommendations than the other three comparison algorithms.

Diversity can indicate the variability between texts in a list. The smaller the diversity value, the greater the similarity between texts and the richer the content. Table II shows the recommended diversity results for several algorithms.

As can be seen in Table II, the mean value of the diversity of the recommendation lists of 8 out of 10 users is lower than the other three comparison algorithms, and the minimum value of the diversity of the recommendation lists of the remaining 2 users is lower than the comparison algorithms. The experimental data suggests that the CBCF-NNIA algorithm does not give the best diversity of recommendation lists, but can give richer recommendations than the other three comparative algorithms. The study selected 50 sets of experimental data to calculate the average accuracy and average diversity of several recommendation algorithms, and the overall comparison results are shown in Table III.

As can be seen in Table III, the accuracy and diversity of the recommended results of the studied CBCF-NNIA algorithm outperform the comparison algorithm, which provides users with textual content that meets their needs as well as being rich and diverse.

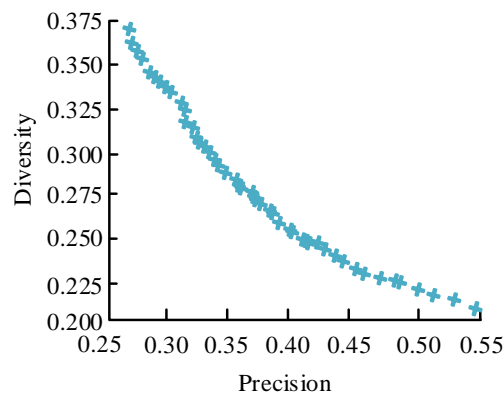


Fig. 9. Non dominated solution of user 27.

TABLE I. RECOMMENDATION ACCURACY OF SEVERAL RECOMMENDATION ALGORITHMS

User ID	CBCF-NNIA			CBCF	BUB	PUBI
	min	max	mean			
300	0.2	0.4	0.32	0.3	0.2	0.3
301	0.2	0.5	0.39	0.3	0.2	0.2
302	0.1	0.4	0.26	0.2	0.1	0.2
303	0.3	0.7	0.45	0.6	0.3	0.5
304	0.2	0.4	0.27	0.2	0.2	0.2
305	0.2	0.6	0.36	0.3	0.2	0.3
306	0.3	0.7	0.51	0.4	0.3	0.4
307	0	0.5	0.23	0.2	0.1	0.2
308	0.5	0.9	0.73	0.7	0.6	0.7
309	0	0.2	0.14	0.1	0	0.1

TABLE II. RECOMMENDATION DIVERSITY RESULTS OF SEVERAL ALGORITHMS

User ID	CBCF-NNIA			CBCF	BUB	PUBI
	min	max	mean			
300	0.36	0.45	0.39	0.40	0.42	0.43
301	0.32	0.42	0.36	0.36	0.37	0.40
302	0.24	0.29	0.26	0.28	0.30	0.39
303	0.30	0.39	0.36	0.38	0.38	0.41
304	0.25	0.33	0.28	0.31	0.29	0.38
305	0.29	0.33	0.31	0.31	0.32	0.35
306	0.26	0.31	0.29	0.32	0.34	0.37
307	0.28	0.38	0.33	0.40	0.40	0.45
308	0.27	0.32	0.29	0.30	0.33	0.31
309	0.31	0.37	0.35	0.36	0.39	0.47

TABLE III. OVERALL COMPARISON RESULTS OF SEVERAL RECOMMENDED ALGORITHMS

/	CBCF-NNIA	CBCF	BUB	PUBI
Average accuracy	0.35	0.32	0.29	0.30
Average diversity	0.29	0.31	0.36	0.34

V. DISCUSSION AND CONCLUSION

The research determines the parameter values in the PPM model through the accuracy changes of the recommendation results under different parameter values. The results show that the recommended accuracy of PPM is the best when $\delta = 0.4$, $\eta = 0.3$, $\chi = 0.7$. The PPM considers the impact of user browsing time difference, external environment and other user behavior characteristics on user interest, and the recommendation results will be more accurate. By analyzing the accuracy, recall, F value and coverage of the recommendation results, the recommendation performance of the three models is compared. The results show that the various index values of CBCF are higher than those of the comparison model. This is because CBCF is a model integrating SPM and PPM models, which can ensure that the recommended content not only meets the user's own preferences, but also improves the content diversity. When the model is fused, the personalization and diversification are measured, and the feature vector with larger weight in the model is used as the user's final preference vector, so as to construct the CBCF model. Therefore, CBCF model combines the advantages of the two models and can effectively improve its recommendation performance. The performance of CBCF-NNIA algorithm is analyzed from the diversity and accuracy of recommendation results. The results show that the algorithm can better take into account the diversity and accuracy of recommendation results, and meet the user's personalized and rich and diverse needs of text. In order to solve the contradiction between recommendation diversity and accuracy, the problem is transformed into a multi-objective optimization problem. The NNIA algorithm has fast convergence speed and good robustness. It uses crossover and mutation operations to enhance the diversity of the population and avoid the occurrence of local optimization. Introducing

NNIA into CBCF can effectively obtain the content list with high accuracy and good diversity.

The rapid development of Internet technology has made the problem of information overload increasingly serious, and recommendation systems are one of the effective techniques to improve information utilization. Faced with the shortcomings in existing recommendation algorithms and the conflict between recommendation accuracy and diversity, the study combines recommendation algorithms with multi-objective optimization algorithms and proposes an NNIA multi-objective optimization hybrid recommendation algorithm. Experimental tests determine the relevant parameters of the hybrid recommendation model. The results of the comparison experiments showed that the accuracy and recall of the hybrid recommendation algorithm were higher than those of the model before the hybridization, with the accuracy improving by 9.57% and 8.23% over SPM and PPM respectively, and the recall improving by 9.97% and 7.65% respectively. The results are consistent with the personalized recommendation law. By exploring the performance of CBCF-NNIA recommendation accuracy and diversity, the results show that the algorithm can effectively obtain content lists with high accuracy and good diversity, which can meet the multiple needs of different users and improve user satisfaction. Although the research has achieved certain results, there are still many shortcomings. The recommendation process mainly considers textual content data, and subsequently, factors such as text labels and categories are also taken into account, while the problem of data sparsity has not been effectively solved, which will be investigated in the future.

REFERENCES

- [1] H. Cao, "Personalized web service recommendation method based on hybrid social network and multi-objective immune optimization." *Journal of Information Processing Systems*, vol. 17(2), pp. 426-439, 2021.
- [2] H. Pan, Z. Zhang, "Research on context-awareness mobile tourism e-commerce personalized recommendation model". *Journal of Signal Processing Systems*, vol. 93(2), pp. 147-154, 2021.
- [3] S. Li, R. Chen, C. Sun, H. Yao, X. Cheng, Z. Li, X. Kang, "Region-aware neural graph collaborative filtering for personalized recommendation." *International Journal of Digital Earth*, vol. 15(1), pp. 1446-1462, 2022.
- [4] Y. Liu, X. You, S. Liu, "Multi-ant colony optimization algorithm based

- on hybrid recommendation mechanism.” *Applied Intelligence*, vol. 52(8), pp. 8386-8411, 2022.
- [5] Y. Tian, B. Zheng, Y. Wang, Y. Zhang, Q. Wu. “College Library Personalized Recommendation System Based on Hybrid Recommendation Algorithm.” *Procedia Cirp*, vol. 83(4), pp. 490-494, 2019.
- [6] Y. Wang, Y. Zhu, Z. Zhang, H. Liu, P. Guo. “Design of Hybrid Recommendation Algorithm in Online Shopping System.” *Journal of New Media*, vol. 3(4), pp. 119-128, 2021.
- [7] W. Jiao, B. Ma, L. Zhu “Hybrid Recommendation Algorithm Based on Improved Collaborative Filtering and Bipartite Network”, *ICBDT 2020: 2020 3rd International Conference on Big Data Technologies*. vol. 11(4), pp. 80-84, 2020.
- [8] Z. Liu, L. Wang, X. Li, S. Pang. “A multi-attribute personalized recommendation method for manufacturing service composition with combining collaborative filtering and genetic algorithm.” *Journal of Manufacturing Systems*, vol. 58(A), pp. 348-364, 2021.
- [9] F. Nafis, K. Al Fararni, B. Aghoutane, A. Yahyaouy, J. Riffi, A. Sabri. “Hybrid recommender system for tourism based on big data and AI: A conceptual framework.” *Big Data Mining and Analytics*, vol. 4(1), pp. 47-55. 2021.
- [10] P. Pirasteh, M. R. Bouguelia, K. C. Santosh, “Personalized recommendation: an enhanced hybrid collaborative filtering.” *Advances in Computational Intelligence*, vol. 1(4), pp. 1-8, 2021.
- [11] Z. Hu, Y. Lan, Z. Zhang, X. Cai, “A many-objective particle swarm optimization algorithm based on multiple criteria for hybrid recommendation system.” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 15(2), pp. 442-460, 2021.
- [12] C. Ajaegbu, “An optimized item-based collaborative filtering algorithm.” *Journal of ambient intelligence and humanized computing*, vol. 12(12), pp. 10629-10636, 2021.
- [13] Na L, Ming-xia L, Hai-yang Q, Hao-long S. A hybrid user-based collaborative filtering algorithm with topic model. *Applied Intelligence*, 2021, 51(11): 7946-7959.
- [14] Aljunid M F, Huchaiah M D. An efficient hybrid recommendation model based on collaborative filtering recommender systems. *CAAI Transactions on Intelligence Technology*, 2021, 6(4): 480-492.
- [15] Y. Xin, B. Henan, J. Niu, W. Yu, H. Zhou, X. Ji, P. Ye, “Coating matching recommendation based on improved fuzzy comprehensive evaluation and collaborative filtering algorithm.” *Scientific Reports*, vol. 11(1), pp. 1-14, 2021.
- [16] L. Xie, Z. Hu, X. Cai, W. Zhang, J. Chen, “Explainable recommendation based on knowledge graph and multi-objective optimization.” *Complex & Intelligent Systems*, vol. 7(3), pp. 1241-1252, 2021.
- [17] F. Rezaimehr, C. Dadkhah “A survey of attack detection approaches in collaborative filtering recommender systems.” *Artificial Intelligence Review*, vol. 54(3), pp. 2011-2066, 2021.
- [18] L. Duan, W. Wang, B. Han, “A hybrid recommendation system based on fuzzy c-means clustering and supervised learning.” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 15(7), pp. 2399-2413, 2021.
- [19] Thakker U, Patel R, Shah M. “A comprehensive analysis on movie recommendation system employing collaborative filtering.” *Multimedia Tools and Applications*, vol. 80(19), pp. 28647-28672, 2021.
- [20] J. Chen, B. Wang, Z. Ouyang, Z. Wang, “Dynamic clustering collaborative filtering recommendation algorithm based on double-layer network.” *International Journal of Machine Learning and Cybernetics*, 2021, vol. 12(4), pp. 1097-1113, 2021.