

Retraction

Retracted: Research on the Recommendation Algorithm of Rural Tourism Routes Based on the Fusion Model of Multiple Data Sources

Discrete Dynamics in Nature and Society

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

Copyright © 2023 Discrete Dynamics in Nature and Society. 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.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] H. Li, M. Qiao, and S. Peng, "Research on the Recommendation Algorithm of Rural Tourism Routes Based on the Fusion Model of Multiple Data Sources," *Discrete Dynamics in Nature and Society*, vol. 2022, Article ID 2262148, 10 pages, 2022.

Research Article

Research on the Recommendation Algorithm of Rural Tourism Routes Based on the Fusion Model of Multiple Data Sources

Hong Li,¹ Man Qiao,¹ and Shuai Peng ²

¹Henan University of Animal Husbandry & Economy, Zhengzhou 450046, Henan, China

²Guangdong University of Finance and Economics, Guangzhou 510320, Guangdong, China

Correspondence should be addressed to Shuai Peng; 20191021@gdufe.edu.cn

Received 23 November 2021; Accepted 21 March 2022; Published 14 April 2022

Academic Editor: Wei Zhang

Copyright © 2022 Hong Li et al. 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.

Rural tourism has become an important force in implementing the rural revitalisation strategy and accelerating rural economic development. The hectic pace of life has made more and more city dwellers yearn for rural life, and travelling in the countryside has become their weekend choice. However, the current level of rural tourism informationization is low, the publicity is insufficient, the tourists' awareness is low, and the source of customers is seriously insufficient. To this end, this paper proposes a relatively novel multidata source fusion tourism recommendation algorithm, which adopts the idea of tensor orthogonal decomposition and fuses multisource data models to predict the target domain's for rating. The integrated consideration of multiple data sources under the do-it-yourself approach assists the target domain to discover the target user neighbourhood users more quickly and to discover the user's interest degree more accurately. It is worth pointing out that the recommendation algorithm proposed in this paper under the fusion of multiple data sources is not necessarily applicable to data sources with weak correlation, such as travel data sources and music data sources, which are relatively weakly correlated, and the algorithm is slightly weak in making predictions of user preferences.

1. Introduction

With the national economy entering a new normal, tourism has ushered in a golden period of rapid development [1], and rural tourism has become an important part of China's tourism industry. The busy pace of life has made more and more city dwellers yearn for rural life, and travelling in the countryside has become their weekend choice. However, at present, the level of information technology for rural tourism is low, the publicity is not strong enough, tourists are less aware of it, and the source of visitors is seriously insufficient. In this context, rural tourism urgently needs information technology to increase publicity and improve service levels [2].

In practical recommendation systems, the most common method of fusing multiple data sources is matrix decomposition. Decomposed user feature matrix U and item feature matrix V are obtained by training the loss function, and finally the scoring matrix is reduced by matrix inverse

operation [3]. However, in the traditional matrix decomposition process, the data structure information is often lost in the matrix decomposition process due to data sparsity, making the results distorted. Tensor is a way of storing multidimensional data, and the concept of tensor decomposition is based on the idea of matrix completion, which aims to fill in the missing (or unobservable) parts of the target matrix [4]. In simple terms, this means that matrix A is used to approximately evaluate matrix B (there is some inherent correlation between A and B). The invisible parts of the matrix B are filled in by the matrix A .

This study, through the platform of rural tourism products on the malefactor research, found that the content of these rural tourism products is mainly concentrated in the better economy and tourism industry of more developed areas, while the economy in relatively backward areas is not covered [5] or there is also incomplete information, not enough to meet the rising demand of users. For example, the interface of searching for any area on the Nongjiale platform

simply shows a little of the same scenery and cuisine information, there is no user search function on the rural tourism service platform, and the product description is simple, which cannot hook the people's desire to travel [6]. The development of rural tourism products has a very far-reaching significance in promoting the economic development of rural areas, combining field research, literature analysis, and data from existing rural tourism platforms to construct a rural tourism product model and a user model. The aim is to improve the quality and efficiency of users' access to useful information, so that rural tourism products are more accurately submitted to users. Combining users' personalised characteristics to recommend the rural tourism products they really need to have has become a valuable and challenging research topic.

2. Related Work

While China's economy is growing rapidly and people's living standards are steadily rising, the tourism industry, as a sunrise industry, is receiving increasing attention from the government and enterprises [7], and tourism has become an important driver to stimulate consumption and induce rapid upgrading and transformation of the industry. Just as China's tourism industry is moving towards mass tourism and global tourism, the development of information technology and mobile Internet applications has given rise to a new concept of smart tourism for the tourism experience. The combination of tourism and information technology constitutes smart tourism, which is a necessary path for the current development of tourism [8].

Tourism websites, tourism APPs, and tourism WeChat applets are current manifestations of smart tourism, and with the rapid development of smart phones and mobile networks, many smart tourism products have emerged, such as Ctrip, founded in 1999, whose mobile APP was launched in 2010, and its acquisition of the UK-based airfare search platform Skyscanner Limited was in November 2016, which means that Ctrip has started to enter the road of internationalisation [3, 9]. Where to Go was established in 2005, and in 2010 it launched its APP, the same as Ctrip; it is a comprehensive travel APP that provides a collection of business travel management, hotels, airline tickets, holiday booking, and travel information. In October 2015, the two travel giant APPs announced a merger. In 2006, the Ma Hive travel website was launched and became popular with users, and its mobile APP was launched in 2011. Just as smart tourism was being widely promoted and applied, China's rural tourism industry also stepped into a path of rapid development [10]. Previous rural tourism products can no longer meet the tourism needs of the people, and there is an urgent need to introduce intelligent rural tourism into rural tourism. That is why MPPs such as Meiju Countryside, Find the Yard, Go Farm, He Xiang You, Meet the Countryside, and Down to the Country Guest have emerged, and corresponding WeChat platforms such as Countryside Tourism Service Platform, Countryside Tourism Merchant, Nongjia Platform, and Shanghai Nongjia Platform have appeared on WeChat mini-programs.

Through the study of the above rural tourism wisdom products, we found that these rural tourism wisdom products are mainly concentrated in the more developed tourism industry and the better economic regions, while the relatively more economically backward regions rural tourism wisdom products either do not have complete information or are not covered, and the scope of operation is narrow, the services provided are limited and not well known, and their service content cannot meet the needs of users [11]. For example, the interface of searching for any area on the WeChat platform of Nongjiale is just to show a little scenery and dish information, there is no user search on the WeChat miniprogram of rural tourism service platform, and the product introduction is simple, which cannot hook the people's desire to travel.

Through literature research, the academic world has also seen more research on intelligent rural tourism; for example, [12] published in the China Tourism News about the new mission given to tourism in the new era, [13] conducted an in-depth analysis of the problems in the development of rural tourism information technology and believes that the factors that inhibit the development of rural tourism information technology are publicity, information management level, information infrastructure construction, etc., [14] argues that there is a certain difference between the demand and supply of tourists in rural tourism and that measures should be provided to reduce the difference between the two, [15] designed an ecological service system for rural tourism from three aspects: service design process, interactive experience, and branding, and [16] argues that, with the help of mobile Internet, a seamless connection between tourism enterprises and tourists can be achieved and a bridge for efficient communication between them can be built.

Although smart tourism is a unique concept in China, similar projects with smart tourism have emerged internationally earlier than in China, with USA, Singapore, Korea, England, and other countries being among the more representative ones. The smart wrist system online in the United States in 2005 opened the beginning of smart tourism; the feedback system equipped with radio frequency technology positioning device one by one Mountainwatch was the first to be used by the Colorado Steamboat Ski Resort in the United States, which can provide tourists with real-time consumption and ski routes. Touchwood, a service platform in Seoul, South Korea, is aimed at self-guided, rural travellers, who can use the platform to perceive tourist information [17, 18].

3. Construction of User Ontology Model for Rural Tourism Platform

Personalised recommendation for users is the ultimate goal of an intelligent recommendation system, and the construction of a good user ontology model is a prerequisite for implementing intelligent recommendations. The user ontology model in this study requires the acquisition of data related to the user's interests in the field of rural tourism, and then combining the user's personal information to

eventually build a model that can be recognised by our computer. The construction process of the user personalised interest model [19] is shown in Figure 1.

3.1. Countryside Tourism Platform User Information. The format and quality of the acquired data directly affect the quality of the user model. Currently, there are two common techniques used to obtain user interest information, specifically:

- (1) Display feedback technology: the main way to record user preferences is through the user's evaluation of the corresponding product. This requires the user to actively participate and actively evaluate the product, which takes up more of the user's time and does not allow for good access to the user's personal interest information when the user's participation is not high.
- (2) Implicit feedback technology: it is mainly through the system to view and analyse the user's behavioural data to obtain personalised information about the user's interests. Implicit feedback technology does not require active participation of the user, mainly through the platform's back-end system to obtain the corresponding information of the user, such as the number of visits to a product, the length of stay of each visit (because of the number of visits to products of interest to the user, the length of stay may be longer), the number of searches, etc.

3.2. Personal Information of Rural Tourism Platform Users. Usually, the user's basic personal situation will make the user's interests relatively stable [20]; for example, users with young children have a greater interest in parent-child rural tourism tours, and low-income people will generally choose rural tourism products with lower or free costs.

In this paper, the authors conclude from the corresponding literature, research on existing user information on rural tourism platforms, and field surveys that the personal information affecting users' choice of rural tourism products mainly includes income status, age range, address of residence, user's gender, nature of work, knowledge composition, and education level received.

3.3. User Model Construction. Rural tourism websites currently on the market have functions such as searching and recommending [21], but the following problems still exist:

- (1) The accuracy rate of search and recommendation is not high.
- (2) The correlation between recommended attractions is not great. Through analysis, it is found that the root cause of the problems is the lack of modelling of user information.

In this study, the user model is constructed by using the relevant information generated by the platform users in the process of browsing and purchasing products to

automatically build their interest models. The user model is an abstract representation of the personal information and interest information of the system users. The basic personal information in the user model is relatively stable, while the user's interest information fluctuates greatly with time and product attributes, so how to accurately express the user's interest and facilitate the calculation becomes the key to the implementation of the user model.

Considering the probabilistic topic representation, the core idea is to understand each text message as a mixture of multiple topical features, where each topic is the distribution probability of the corresponding feature [22].

In the rural tourism platform study, the rural tourism product ontology is used to represent the user's interests, which is relatively simple in structure and can be implemented by the keyword vector space representation method. The user model consists of the user's interest set, the user's interest attribute set, the user's attention and weight for each interest in the interest attribute values, and 3 parts, which are the user's interest set, i.e., the product's key attribute set; the user's interest attribute set, i.e., the product attribute value set.

3.4. Dynamic User Interest Model. The user model contains two types of information: static, such as basic information about the user, such as gender, occupation, and knowledge background; and dynamic, which changes over time. In this study, we mainly consider the part of dynamic updates. Users' interests change with the environment and psychological factors, so the common methods used to update dynamic user models are the forgetting function method and the time window method. The forgetting function method is similar to the law of memory forgetting in that, without external stimuli, a user's interest in something will decay over time [23]. In this study, the user interest model is considered to change with the combination of time forgetting and frequency of access. User interest forgetting function is used to dynamically update the user interest model based on the forgetting factor, which is calculated using

$$f(t) = \exp\left(-k \frac{t - T_{\min}}{T_{\max} - T_{\min}}\right), \quad (T_{\min} \leq t \leq T_{\max}), \quad (1)$$

where T_{\min} is minimum forgetting interval and represents the forgetting process, i.e., the buffer period, and the value is the difference between the interest T_s and the interest reference time; T_{\max} is maximum forgetting interval, indicating the decay cycle, i.e., the time required for the interest to decrease to its original value; t is user access interval; T is the time of the last visit on the current date (in days); k is interest decay rate, where the value of k is proportional to the decay rate, here defined as 1 (the specific value can be adjusted according to the actual needs of the user).

Equation (1) considers the case of a forgotten decay cycle, i.e., ($T_{\min} \leq t \leq T_{\max}$), when $t > T_{\max}$ or $t < T_{\min}$ is not involved, as defined in this study.

When $t < T_{\min}$, $f(t) = 1$, it means that the interest is not decreasing.

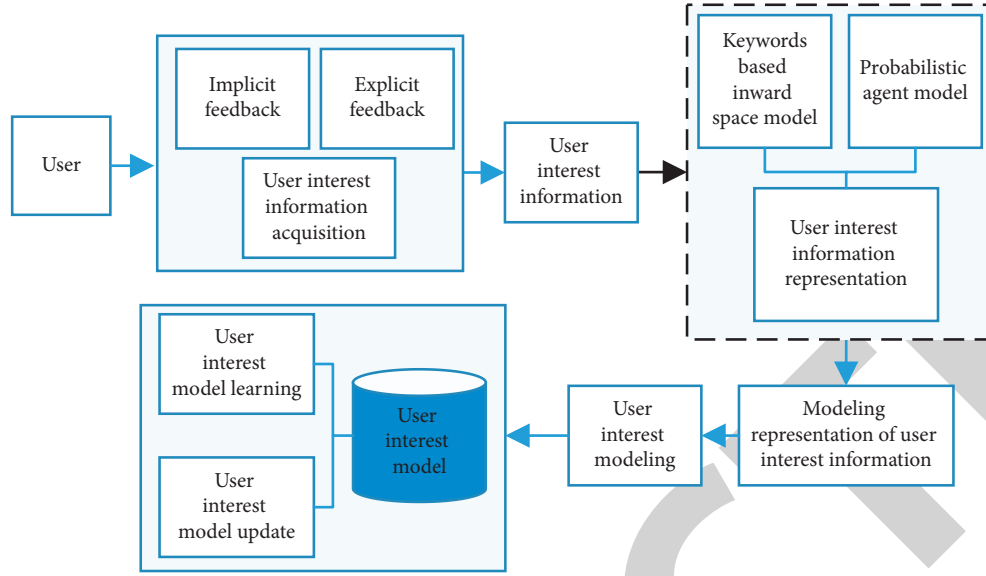


FIGURE 1: Selected photographic records of participants.

When $t > T_{\max}$, $= 0$, it means that the user has lost the interest and is removed from the user model.

The exponential forgetting function allows users' interest weights to decay according to the length of the time interval. This forgetting function takes into account the laws of human memory and treats interest as a special kind of memory.

The dynamic user interest weighting formula that combines the frequency of visits to the dynamic user interest model can consider not only the one factor of time, but also the frequency of visits to the interest. In this study, the dynamic user interest model [24] is constructed by combining the time forgetting factor and the frequency of access to the interest, as shown in (1)

$$v_i(t, N_i) = \begin{cases} 1, \\ \exp\left(-\frac{1}{N_i}\right) * fi(t), \\ 0, \end{cases} \quad (2)$$

where N_i indicates the frequency of the user's visit to the i th interest point; $v_i(t, N_i)$ indicates the weight of the user's interest in the i th interest point at time t . The initial value of the user's interest in all interests is 0, the weight of the user's interest in "parent-child" recreation at the current moment is V_j , and according to the time cycle of rural tourism, the minimum forgetting interval is 30 days and the maximum forgetting interval is one year, i.e., 365 days.

4. Tensor Decomposition-Based Fusion Model for Multiple Data Sources

4.1. Tensor Models. The reason for applying the idea of tensor decomposition to the collaborative filtering recommendation algorithm is that collaborative filtering

essentially evaluates the preference information of most users and then makes a recommendation for a particular user. Once the set of users' rating vectors is constructed, the target user's rating vectors from other rows are used to predict the target user's rating of the product (filling in the missing values of the target user's rating in one row). A tensor is a multilinear vector space, where first-order tensors can be considered as vectors and second-order tensors can be represented by matrices. Tensors larger than second order are uniformly called higher order tensors. The rank of a tensor is defined as if a tensor can be expressed as an outer product of N vectors, and then the rank of the tensor is said to be N . The rank of a tensor means a tensor of rank one, which can be expressed by the outer product of vectors [25]. Figure 2 shows a third-order tensor, whose outer product is of the form $x_1 \cdot x_2 \cdot x_3$.

The cp decomposition of a tensor is based on the basic concept of a rank one tensor, which is a decomposition of a multiorder tensor into the form of an outer product of multiple rank one tensors. For example,

$$X = \sum_{r=1}^K x_1^r \cdot x_2^r \cdot \dots \cdot x_N^r, \quad (3)$$

where K denotes the number of rank one quantities.

In the description of this paper, we need to consider predicting the scoring data in the target data source from the feature vectors of the secondary data source, and here we enhance the correlation between multiple data sources with the help of tensor decomposition techniques between the target and secondary data sources.

4.2. Basic Ideas of the Algorithm. It is assumed that the user has had relevant Internet operations on multiple data sources and has generated a rating matrix in the corresponding data source. The set of data sources is $\mathfrak{R}_i = \{r_1, r_2 \dots r_i \dots\}$, and the user's rating matrix is

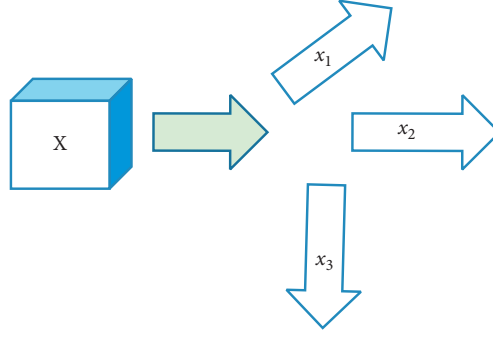


FIGURE 2: Third-order tensor.

represented by the set q , where $X_1 = \{x_{1p_1}', \dots, x_{D_1}'\}$ represents the user's rating of item $Y_i = \{y_{D_1}^1, \dots, y_{D_1}'\}$.

Here, X , Y denote the corresponding user ratings of items in different data sources, which may be in the same data source or in different data sources where the user ratings of items overlap, e.g., the user ratings of items in the data source and in the data source at the same time in the corresponding data source, and there is also the possibility that the user ratings of items in each data source are unrelated. The data sources are also independent of each other with no overlap [26]. The data sources are considered to be the target data sources, and other relevant data sources are considered to be the secondary data sources.

Assuming that there is partial overlap and interpopulation of users between data sources, the corresponding ratings made by users in different domains are abstracted into n -order tensors according to the idea of tensor, and the tensors in this example all belong to rank one tensor, and the vector model diagram is shown in Figure 3.

4.3. Model Building. The multisource tourism information fusion model proposed in this paper is based on tensor decomposition, and the user's rating matrix in the target data source is regarded as a reorganization of the approximation matrix of the secondary data source through the idea of matrix complementation, and then the elements of the approximation matrix are used as the unobservable part of the target matrix for rating estimation [27].

The main objective of the multidata source fusion travel recommendation algorithm is to construct a global rating matrix model, where the matrix fuses the global rating matrices of users in different data sources. The matrix fusion process is shown in Figure 4.

From the above model, we abstract the three data sources into a third-order tensor model, assuming a global data domain of \mathfrak{R} , the third-order tensor is described as $\mathfrak{R} = [[M, T, B]] = \sum_{r=1}^R M_{D_r} \circ T_{D_r} \circ B_{D_r}$, where M, T, B correspond to the user ratings in the three data sources of movies, travel, and data, respectively. According to the tensor decomposition model, the fusion of multiple data sources can be normalised into a minimisation solution problem. The solution formula is as follows:

$$\min(M, T, B) = \frac{1}{2} \|\mathfrak{R} - [[M, T, B]]\|^2 + \frac{\lambda_M}{2} \|M\|_F^2 + \frac{\lambda_T}{2} \|T\|_F^2 + \frac{\lambda_B}{2} \|B\|_F^2, \quad (4)$$

where $[[M, T, B]] = \sum_{r=1}^R M_{D_r} \circ T_{D_r} \circ B_{D_r}$, a regularization term is added to prevent overfitting of the training results, λ represents the regularization parameter, and the model is trained using stochastic gradient descent.

To ensure that the local optimum result can be achieved in finite time, it needs to be shown that partial derivatives exist for the C data sources in their respective directions. The proof procedure is as follows:

$$\frac{\partial f}{\partial M} = (X_{(1)} - [[M, T, B]]_{(1)}) (B \Theta T) + \lambda_M M. \quad (5)$$

Similarly it can be shown that the partial derivatives of T and B in the respective directions are

$$\frac{\partial f}{\partial T} = (X_{(2)} - [[M, T, B]]_{(2)}) (M \Theta B) + \lambda_T T, \quad (6)$$

$$\frac{\partial f}{\partial B} = (X_{(3)} - [[M, T, B]]_{(3)}) (M \Theta T) + \lambda_B B.$$

4.4. Travel Recommendation Algorithm Design. Since the user's scoring matrix in a multidata source environment may not be a regular tensor, it is not possible to use the tensor decomposition model directly, so here it is necessary to introduce an invertible transformation to ensure that the ones in different domains can be transformed into the same dimensional information matrix; i.e., the matrix product of the $[M, T, B]$ score components is expressed in the form of

$$Y_K = M \sum_k B^T + E_k, \quad (7)$$

where A, B denote the tensor matrix vectors in the auxiliary data sources. $\sum_k = \text{diag}(C_k)$ denotes the diagonal matrix of $R \times R$ of the target data sources. E_k denotes the residual terms of the model training. The global scoring matrix for multiple data sources can be obtained by minimizing the objective function:

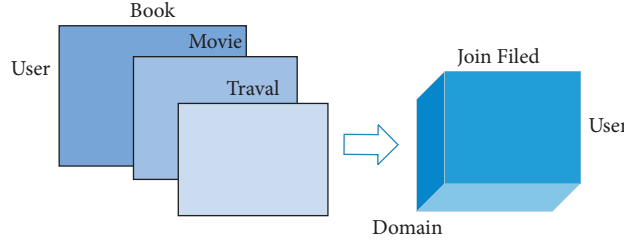


FIGURE 3: Data fusion model.

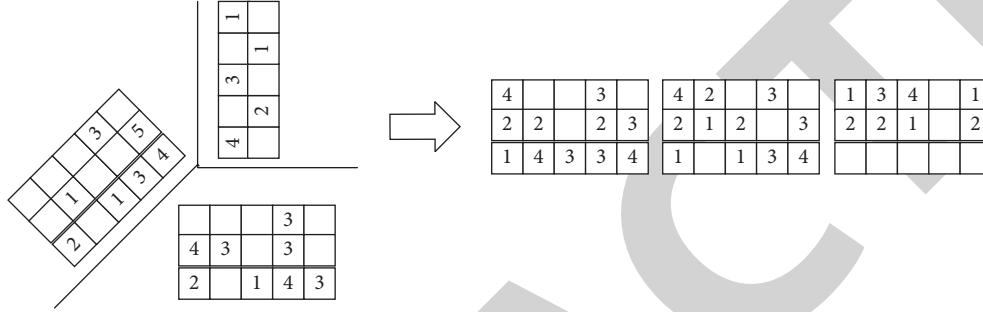


FIGURE 4: Matrix fusion process.

$$U \approx \min(M, T, B) = \frac{1}{2} \left\| Y_K - M \sum_K B^T \right\|^2 + \frac{\lambda_M}{2} \|M\|_F^2 + \frac{\lambda_T}{2} \|T\|_F^2 + \frac{\lambda_B}{2} \|B\|_F^2. \quad (8)$$

Here it is assumed that $X_D = \{X_1, X_2, \dots, X_K\}$ is a scoring matrix for K different domains, where X_K has $N \times M_k$ dimensions, N denotes the number of users under the current dimension, and M_k denotes the number of items in the K th data source.

The vectors under different data sources can be iteratively trained and fused under the same data source. With the help of a unified data source model, the similarity of the target data source domain is calculated and the nearest neighbours are selected based on this to obtain global recommendation results [9].

When the global user rating matrix is obtained the target user can be selected to calculate the nearest neighbours for predicted ratings and obtain the final set of recommendation results. The rating prediction formula is as follows:

$$P_{u,i} = \bar{R} + \frac{\sum_{u \in U(u)} \text{sim}(u, u) * (R_u - \bar{R}_u)}{\sum_{u \in U(u)} |\text{sim}(u, u')|}, \quad (9)$$

where $U(u)$ represents the global data rating matrix, $\text{sim}(u, u)$ represents the similarity of users on the global data domain, and \bar{R}_u represents the average user rating.

5. Analysis of Experimental Results

5.1. Experimental Data Sets. In order to obtain sufficient experimental data to verify the feasibility of the algorithm, the experimental dataset was processed as follows: firstly, 1039 attractions in six Chinese cities (Beijing, Shanghai,

Guangzhou, Guilin, Hangzhou, and Haikou) were crawled from a domestic travel website using the dynamic agent technique of the PythonScrapy crawler framework, which contains 8720 users and 23,305 rating records of the attractions by users; each user can make a range of ratings for the attractions, along with 1,500 travel tips, corresponding types of attractions, the number of visitors to the attractions under different seasons, etc. The information is mined to analyse the relevant characteristics, including gender, age, occupation, travel time, as well as the route, type, and rating of the attraction, and stored in the database for subsequent analysis.

The types of attributes and the corresponding number of ratings for the attractions are shown in Table 1.

The user ratings for the scenic spots are shown in Table 2.

5.2. Experimental Results and Discussion. The experimental dataset was divided into a training set and a test set in the ratio of 8 : 2, and the results were validated by evaluating the data from the training set and using the data from the test set. The test dataset contains over 17 attraction types (historical, scenic, natural, human, etc.) [11].

5.2.1. Effect of Weighting Parameters. The rating similarity and attribute similarity of attractions are combined through (9) to construct a global similarity of attractions, and the proportion in which these parameters are adapted is a focused point of investigation for this section of the

TABLE 1: Types of attractions.

POI_ID	POI_NAME	Rating	Type
1	The Great Wall	456	History, scenery
2	Palace Museum	432	History, humanity
3	XiHu	112	Scenery, park
4	Nanshan Temple	321	Buddhism
5	Tiananmen Square	222	Politics, park
6	Fenjiezhou Island	532	Landscape, park

TABLE 2: User-attraction rating scale.

User_ID	POI_ID	Rating	Times stamp
189	212	3	865454
231	331	2	823213
123	213	4	872443
412	23	3	865231
11	321	5	865454
6	221	3	843233

experiment. From (9), $\mu_a + \mu_p = 1$, to ensure a single controllable experimental variable; here we propose a hypothesis of $\alpha = \mu_a$, and then $\mu_p = 1 - \alpha$, so that a single variable can be controlled to observe the effect of the weighting parameters on the experimental results. In this experiment, the uniform parameter α was taken in the range $[0, 1]$, and the performance of the algorithm was observed by adjusting the different neighbourhood users selected.

As can be seen from Figure 5, the horizontal coordinate represents the range of the parameter values, and the vertical coordinate represents the MAE values, which change differently between different numbers of neighbours k as the parameter goes from 0 to 1. The MAE of the algorithm is optimal when the weight parameter $\alpha = 0.6$ and the number of neighbours $k = 60$. The MAE value decreases at the beginning as the weight parameter α increases and starts to increase when it exceeds $\alpha = 0.6$. This is mainly because the algorithm gradually ignores the evaluation of attraction attributes when the weight parameter exceeds 0.6. In order to find the optimal set of neighbours for the target item, we set the range of neighbours for the target item to 60 and use three different weight control parameters $\alpha = 0.061$ to observe the influence of different number of neighbours on the recommendation result, which was observed by using three different weight control parameters; $\alpha = 0.061$.

As shown in Figure 6, the overall performance of the algorithm's MAE value is low after fusing the optimal weight parameters, and the MAE of the algorithm gradually decreases as the number of project neighbours increases and gradually stabilizes when the number of neighbours exceeds 50. The experimental results show that when the number of neighbours is chosen around 40, the algorithm can achieve the optimal recommendation result. Comparing different algorithms, this section compares the algorithms in this paper (RACF with the traditional user-based collaborative filtering algorithm (IBCF) and the improved item-based collaborative filtering algorithm (IITEM-CF)). 800 users were selected in the experimental dataset, which contained a large

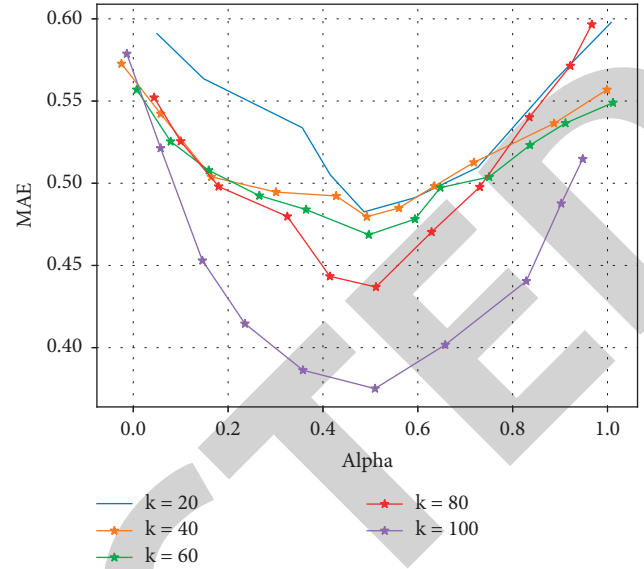


FIGURE 5: Effect of weighting parameters.

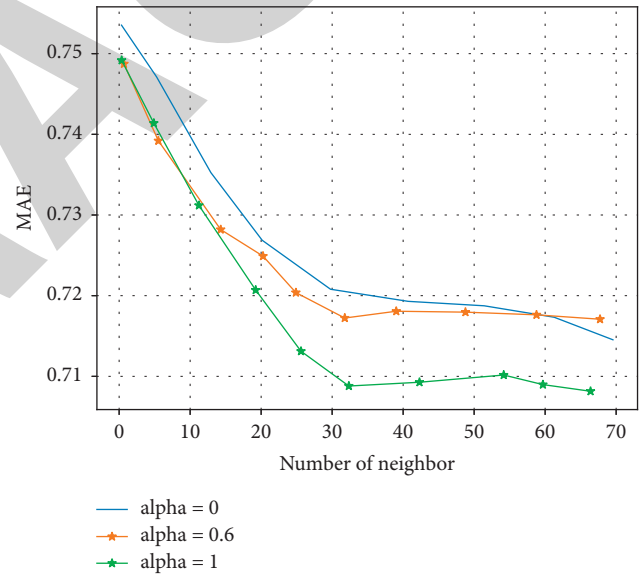


FIGURE 6: Effect of number of residences.

number of unrated scenic items. The experimental results are shown in Figure 7.

From Figure 7, it can be seen that the MAE values of the algorithms in this paper are relatively low when crossing the other two algorithms. Traditional algorithms IBCF and ITEM-CF have difficulty in obtaining the nearest neighbours of the target items due to the lack of basis for calculating the similarity matrix due to the absence of a large amount of rating data. In this paper, the algorithm uses a combination of attraction scores and project attributes to calculate the global similarity in the absence of project scores, combined with the inherent project attributes to assist in the calculation of global similarity, which alleviates the problem of data sparsity to a certain extent.

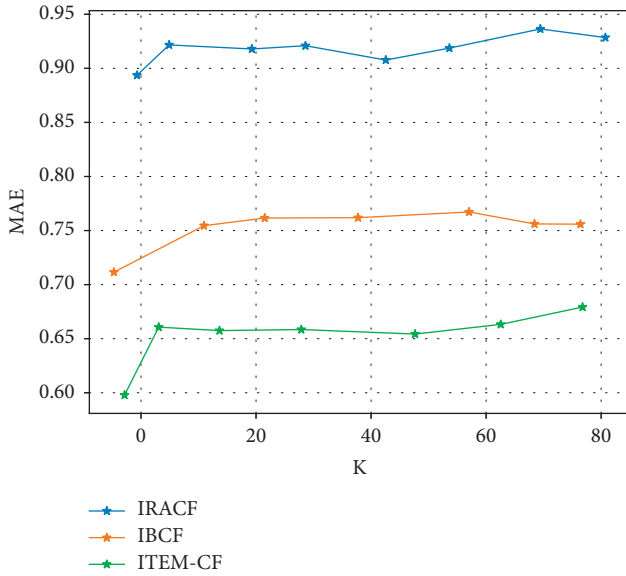


FIGURE 7: Comparison results of the same algorithm.

5.3. Real Life Examples. Based on the tourism data of domestic Ctrip, we used the dynamic agent technology of Python scratch crawler framework to capture 1039 scenic spots in 6 cities in China (Beijing, Shanghai, Guangzhou, Guilin, Hangzhou, and Haikou) from a domestic tourism website. It contains 8720 users and 23305 scoring records of scenic spots. Each user can store the scenic spots within the range of [1, 5]. At the same time, there are 1500 tourism strategies, corresponding scenic spot types, number of scenic spots visited in different seasons, etc.

In total, the experiments in this section involve three data sources: the movie data sources, book data sources, and travel data sources, using movies and books as secondary data sources to make predictions on the target data source, the travel domain. The experiments in this paper are divided into the following tasks.

- (1) The movie data source and the book data source, respectively, are used as auxiliary data sources to calculate the similarity between users, to predict users who did not make a rating, to calculate the nearest neighbours, to calculate the correlation coefficient using the modified cosine similarity, and thus to predict the items that populate the user who did not make a rating.
- (2) Based on the similarity between users calculated from the auxiliary data sources, predictions were made to the target according to the source of the items that were not rated.
- (3) Fusion of the target data source and the auxiliary data source was carried out. The algorithm in this paper is used to score the target data source.
- (4) Without differentiating data sources, the global data are considered, and the movie data source, book data source, and travel data source are considered as the overall target data domain.

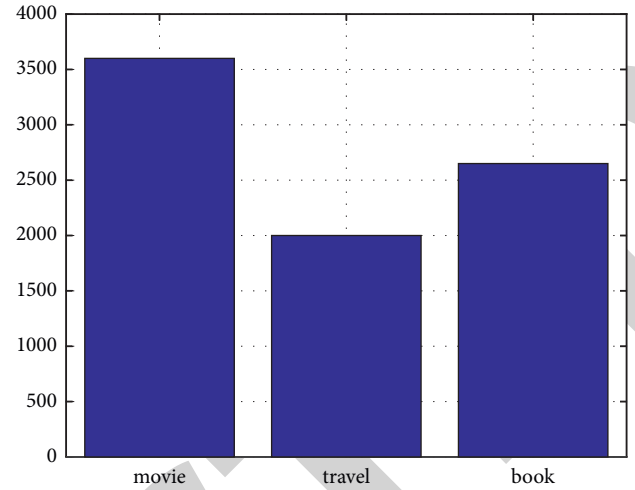


FIGURE 8: Distribution of test data statistics.

- (5) Without distinguishing the data sources, the global data is considered, and the movie data source, book data source, and tourism data source are regarded as the overall target data domain, and the similarity is calculated for the target user in the target data domain and the nearest neighbours are selected, and the missing rating items of the target user are filled to make the user's rating of the attractions in the tourism data source.

The data set division in the experiments is uniformly constructed in the form of proportional division; i.e., the ratio of the training set to the test set is 8:2, and the algorithm model is trained based on the data in the training set to predict the users' scores for the unrated items in the test set. Firstly, as shown in the figure, we conducted a statistical analysis of the scores in the secondary data sources to verify the correlation between the data sources, and the experimental results are shown in Figure 8.

Figure 8 shows the number of users who have rated books in the different data sources. Since the number of users selected is fixed (5000), we can observe that there is a certain coverage of users' behaviour in the different data sources. The main reason for this phenomenon is that if a user has marked books on the topic of history and humanities several times in the book data source, he will also pay more attention to movies related to history and humanities in the movie field, and by sending similar users in the secondary data source, it helps the target data source field to discover similar users in the target field faster. This is also a prerequisite for fusion modelling of multiple data sources.

Figure 9 shows the performance of the algorithm in the case of different data sources.

In Figure 9, the horizontal coordinate indicates the number of neighbours of the selected target user, and the vertical coordinate indicates the value of the system average absolute error MAE. From the above figure, it can be seen that the algorithm's prediction in the single data source environment is difficult to find suitable neighbours in effective time due to the sparsity of the data, resulting in low

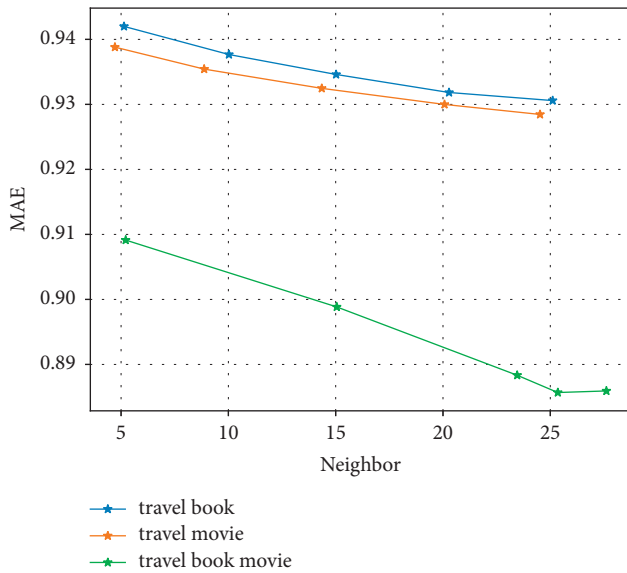


FIGURE 9: Performance of the algorithm in this paper with the same data source.

recommendation accuracy. As new data sources are added, the algorithm's MAE value starts to decrease compared to the single data source environment, and the algorithm's MAE value is lowest when fusing global data sources and reaches the optimal result when the number of neighbours is 25.

6. Conclusions

This paper proposes a relatively novel travel recommendation algorithm with fusion of multiple data sources. The algorithm adopts the idea of tensor orthogonal decomposition and fuses multiple data models to predict the rating of the target domain. The integrated consideration of multiple data sources under the use of the target domain is to assist the target domain to discover the target user neighbourhood users more quickly and more accurately to discover the user's interest degree. It is worth pointing out that the recommendation algorithm proposed in this paper under the fusion of multiple data sources is not necessarily applicable to data sources with weak correlation. The algorithm has the lowest MAE value when fusing global data sources and reaches the optimal result when the number of neighbours is 25.

Data Availability

The dataset used in this paper is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by the National Social Science Fund Item, Research on Ecological Aesthetics in line with

the traditional culture of unique ethnic minorities in Yunnan (no. 20BMZ164).

References

- [1] M. Wang, "Research on convergence of rural tourism and related industries based on the perspective of holistic tourism: practice analysis from China," *Ekoloji*, vol. 28, no. 107, pp. 4295–4302, 2019.
- [2] P. Yochum, L. Chang, T. Gu, and M. Zhu, "Linked open data in location-based recommendation system on tourism domain: a survey," *IEEE Access*, vol. 8, pp. 16409–16439, 2020.
- [3] Y. Yi, C. Zhao, and J. Fu, "Research on the integration and development of modern agriculture and rural tourism based on AHP—take Yangjia town in Mianyang city as an example," *Open Journal of Social Sciences*, vol. 7, no. 7, pp. 76–87, 2019.
- [4] D. Wu, C. Zhang, L. Ji, R. Ran, H. Wu, and Y. Xu, "Forest fire recognition based on feature extraction from multi-view images," *Traitement du Signal*, vol. 38, no. 3, pp. 775–783, 2021.
- [5] H. Zhang, Y. Duan, and Z. Han, "Research on spatial patterns and sustainable development of rural tourism destinations in the yellow river basin of China," *Land*, vol. 10, no. 8, p. 849, 2021.
- [6] J. Zhang and L. Dong, "Image monitoring and management of hot tourism destination based on data mining technology in big data environment," *Microprocessors and Microsystems*, vol. 80, Article ID 103515, 2021.
- [7] G. Alexandridis, A. Chrysanthi, G. E. Tsekouras, and G. Caridakis, "Personalized and content adaptive cultural heritage path recommendation: an application to the Gournia and Çatalhöyük archaeological sites," *User Modeling and User-Adapted Interaction*, vol. 29, no. 1, pp. 201–238, 2019.
- [8] Q. Li, S. Li, S. Zhang, J. Hu, and J. Hu, "A review of text corpus-based tourism big data mining," *Applied Sciences*, vol. 9, no. 16, p. 3300, 2019.
- [9] F. Randelli and F. Martellozzo, "Is rural tourism-induced built-up growth a threat for the sustainability of rural areas? The case study of Tuscany," *Land Use Policy*, vol. 86, pp. 387–398, 2019.
- [10] H. Li, D. Zeng, L. Chen, Q. Chen, M. Wang, and C. Zhang, "Immune multipath reliable transmission with fault tolerance in wireless sensor networks," in *Proceedings of the International Conference on Bio-Inspired Computing: Theories and Applications*, pp. 513–517, Springer, Singapore, 2016 October.
- [11] I. Ibrahim, K. Zakariya, and N. A. Wahab, "Satellite image analysis along the kuala selangor to sabak bernam rural tourism routes," in *Proceedings of the IOP Conference Series: Earth and Environmental Science*, vol. 117, no. 1, Article ID 012013, United Kingdom, 2018.
- [12] G. Bambi, S. Iacobelli, G. Rossi, P. Pellegrini, and M. Barbari, "Rural tourism to promote territories along the ancient roads of communication: case study of the rediscovery of the st. Francis's ways between fv," *European Countryside*, vol. 11, no. 3, pp. 462–474, 2019.
- [13] R. Beridze, I. Kordzaia, M. Diasamidze, and N. Beridze, "Sustainable rural tourism development recovery from covid-19 (in adjara region)," *Globalization and Business*, vol. 10, pp. 287–294, 2020.
- [14] S. Görmüş, E. Atmış, H. B. Günşen, N. K. Özkazanç, and M. Artar, "The importance of mapping natural and cultural routes in rural tourism: bartın case," *International Journal of Agricultural and Natural Sciences*, vol. 10, no. 1, pp. 32–38, 2017.

- [15] I. Trišić, S. Štetić, D. Privitera, and A. Nedelcu, "Wine routes in Vojvodina Province, Northern Serbia: a tool for sustainable tourism development," *Sustainability*, vol. 12, no. 1, p. 82, 2020.
- [16] T. Wijijayanti, Y. Agustina, A. Winarno, L. Istanti, and B. Dharma, "Rural tourism: a local economic development," *Australasian Accounting, Business and Finance Journal*, vol. 14, no. 1, pp. 5–13, 2020.
- [17] C. J. A. Mitchell and M. Shannon, "Establishing the routes to rural in-migrant proprietorship in a Canadian tourism region: a mobilities perspective," *Population, Space and Place*, vol. 24, no. 3, p. e2095, 2018.
- [18] L. Wang, C. Zhang, Q. Chen et al., "A communication strategy of proactive nodes based on loop theorem in wireless sensor networks," in *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 160–167, IEEE, Wanzhou, China, 2018, November.
- [19] M. A. Uwaisy, Z. K. A. Baizal, and M. Y. Reditya, "Recommendation of scheduling tourism routes using tabu search method (case study bandung)," *Procedia Computer Science*, vol. 157, pp. 150–159, 2019.
- [20] S. Li, S. Takahashi, K. Yamada, M. Takagi, and J. Sasaki, "Analysis of SNS photo data taken by foreign tourists to Japan and a proposed adaptive tourism recommendation system," in *Proceedings of the 2017 International Conference on Progress in Informatics and Computing (PIC)*, pp. 323–327, IEEE, Nanjing, China, 2017, December.
- [21] W. Zheng, Z. Liao, and Z. Lin, "Navigating through the complex transport system: a heuristic approach for city tourism recommendation," *Tourism Management*, vol. 81, Article ID 104162, 2020.
- [22] L. Zhang, "Intelligent tourism route recommendation method based on big data," *International Journal of Autonomous and Adaptive Communications Systems*, vol. 13, no. 4, pp. 329–341, 2020.
- [23] J. Xu and J. Xu, "Research on the design of online travel service recommendation system based on data analysis," in *Proceedings of the International Conference on Cognitive Based Information Processing and Applications (CIPA 2021)*, pp. 1042–1046, Springer, Huainan, China, 2022 August.
- [24] X. Zheng, Y. Luo, L. Sun, Q. Yu, J. Zhang, and S. Chen, "A novel multi-objective and multi-constraint route recommendation method based on crowd sensing," *Applied Sciences*, vol. 11, no. 21, p. 10497, 2021.
- [25] L. Feng, "Design of tourism intelligent recommendation model of mount tai scenic area based on knowledge graph," in *Proceedings of the 2020 International Conference on E-Commerce and Internet Technology (ECIT)*, pp. 241–244, IEEE, Zhangjiajie, China, 2020, April.
- [26] C. Cao, Y. Tang, D. Huang, W. Gan, and C. Zhang, "IIBE: an improved identity-based encryption algorithm for WSN Security," *Security and Communication Networks*, vol. 2021, Article ID 8527068, 8 pages, 2021.
- [27] C. Yuan and M. Uehara, "An optimal route recommendation method for a multi-purpose travel route recommendation system," in *Proceedings of the International Conference on Broadband and Wireless Computing, Communication and Applications*, pp. 398–408, Springer, Fukuoka, Japan, October 2019.