AAKARSHAN BISWAS

INFORMATION TECHNOLOGY IN TOURISM

THE INFO LIBRARY

Information Technology in Tourism

Information Technology in Tourism

Aakarshan Biswas

THE INFO LIBRARY

Published by The InfoLibrary,
4/21B, First Floor, E-Block,
Model Town-II,
New Delhi-110009, India

Information Technology in Tourism
Aakarshan Biswas
ISBN: 978-93-5590-460-7

Table of Contents

Research on Precision Marketing Model of Tourism Industry Based on User's Mobile Behavior Trajectory

Jialin Zhang,[1,2] Tong Wu,[1] and Zhipeng Fan[1,2]

[1]*Harbin University of Commerce, Harbin 150028, China*
[2]*Heilongjiang Provincial Key Laboratory of Electronic Commerce and Information Processing, Computer and Information Engineering College, China*

Correspondence should be addressed to Zhipeng Fan; hsdfzp@126.com

Guest Editor: Jaegeol Yim

With the deep cross-border integration of tourism and big data, the personalized demand of tourist groups is increasingly strong. Precision marketing has become a new marketing mode that the tourism industry needs to pay close attention to and explore. Based on the advantages of big data platform and location-based service, starting from the precise marketing demand of tourism, we design data flow mining technology framework for user's mobile behavior trajectory based on location services in mobile e-commerce environment to get user track data that incorporates location information, consumption information, and social information. Data mining clustering technology is used to analyze the characteristics of users' mobile behavior trajectories, and the precise recommendation system of tourism is constructed to provide support for tourism decision making. It can target the tourist group for precise marketing and make tourists travel smarter.

1. Introduction

1.1. Research Background. Location-based service (LBS) is a kind of value-added service provided by the combination of mobile communication network and satellite positioning system. It obtains the location information of mobile terminal such as latitude and longitude coordinate data through a set of positioning technology and provides it to the communication system, mobile users, and related users to realize various location-related services in military and transportation. As a new mobile computing service in recent years, 80% of the world's information has time and location tags, and location services have developed to the big data stage [1]. Developing location services requires two capabilities: the ability to provide location and the ability to understand location.

The precise marketing information pushed by LBS location service can effectively tap the potential consumer demand and make a scientific and reasonable network marketing strategy based on this, which can further improve the ability of e-commerce enterprises to tap the target customers and potential customers. According to the 2017 China Mobile e-commerce Industry Research Report, the transaction scale of China's e-commerce market reached 20.2 trillion yuan in 2016, an increase of 23.6% compared with the same period of the same period of the year. China's e-commerce market is developing steadily. Among them, the online shopping has a good momentum of development, up from 23.3% in 2015. Huge market potential tempts all walks of life. In 2016, online shopping and B2B e-commerce of small and medium-sized enterprises and enterprises above scale still dominate the Chinese e-commerce market, while online tourism and local life service O2O emerge as bamboo shoots, accounting for 3% and 1.6% of the market, respectively. From 2015 to 2016, the proportion of online tourism market in China's tourism market has greatly increased, the process of product informatization has accelerated, the penetration rate has further improved, the mobile online tourism market has developed rapidly, and consumer understanding and demand and experience of tourism are changing imperceptibly and pursuing higher quality of tourism. With the further development of "Internet+" information technology, the tourism industry has huge room for development, and online travel penetration will also gradually increase.

According to the Statistical Bulletin on National Economic and Social Development of 2016 issued by the State Statistical Bureau, in the whole year of 2016, the number of domestic tourists' trips reached 4.4 billion, an increase of 11.2% over the previous year, and the income of domestic tourism increased by 15.2% to 39390 billion yuan. The number of inbound tourists reached 138.44 million, an increase of 3.5%, and international tourism revenue increased by 5.6% to $120 billion. The number of domestic residents in China has reached 135 million 130 thousand, an increase of 5.7% [2]. With the continuous promotion of the strategic pace of building a well-off society in an all-round way, tourism has become an important part of the people's daily life in China, marking that China's tourism industry has entered the era of mass tourism.

Based on this background, in the mobile e-commerce environment, based on LBS location service, research and analysis of user's mobile behavior trajectory can extract valuable user's mobile behavior features from a large number of mixed dynamic data and integrate the mobile behavior and consumer behavior of tourism users. Based on LBS location, services will integrate the mobile behavior and consumer behavior of tourism users, then excavate the marketing value of consumers, and timely achieve the marketing objectives of enterprises on the appropriate media, so that mobile e-commerce marketing becomes more accurate and effective. Through the research of this subject, the interests of enterprises, consumers, and media can be maximized at the same time, providing personalized products and services for mobile e-commerce, improving consumer loyalty and core competitiveness of mobile e-commerce and bringing higher profits for e-commerce enterprises [3].

1.2. Presentation of Problems. The data of mobile terminal users' historical consumption behavior and location movement process are recorded and stored according to the time series, forming the user's mobile behavior trajectory data, which can be collected by multiple device terminals. The user's mobile behavior trajectory data contain a lot of useful information. Mobile behavior trajectory can express the behavior activities of mobile users in the real world. These activities imply user's interests, hobbies, experiences, and behavior patterns [4]. For example, a user's activities in a week may start from home to work every day, and a user may go to shopping malls, parks, and other places on weekends. Therefore, how to effectively utilize the user's mobile behavior trajectory and extract useful information from the user's mobile behavior trajectory data is very important for the realization of personalized recommendation service.

From the view of consulting a large number of documents, there are more papers on location services than on mobile marketing. However, most of the previous articles on location services focused on the application of natural science, such as surveying and mapping technology, network development, geographic information, and so on. In recent years, the number of cross-research articles on management science, medicine, and agriculture combined with location-based services has begun to increase, most of which are the combination of location-based services and related industries to study the application or technology development of specific industries. For example, the combination of location services and logistics technology can track the journey of packages. Combining location services with electronic maps can provide catering, entertainment, discounts, and other information within a certain range according to the location of users. Combining location service with utility technology can quickly find information such as tap water, gas explosion, and so on. The main keywords of literature research include location service technology, location service system, location service terminal, location service strategy, mobile location service, and so on. Based on location, the services industry is currently considered one of the most dynamic industries. With the rapid development of mobile Internet and Internet of Things technology, more debris time has been transferred to mobile phones, tablets, and smart products [5].

The characteristics of mobile marketing, such as precision, interaction, novelty, and effective delivery, are more and more concerned and recognized by various industries. This paper focuses on the main characteristics of the end-users of the tourism industry, such as frequent location movement, strong sense of sharing, and rich demand for services. First of all, the users of online travel must be mobile users who usually do not stay in a location for a long time, and a high probability of frequent location changes will produce a large number of location data. Moreover, in general, traveling users arrive at an unknown location or move in a series of unfamiliar geographical environments, which makes travel users' demand for location-based services take precedence over personal privacy protection and enable them to obtain real-time user location information. These location data provide favorable conditions for our research. Secondly, the behavior of tourist users is quite different from that of ordinary people. In beautiful scenery and not very familiar environment, users will spontaneously produce self-awareness. Most people share location, photos, and moods through social platforms and micromessaging, and travel companies can access these social data to accurately portray users and provide accurate services for them. Third, travel users need high-quality services to obtain high-quality tourism experience. Scenic spots, accommodation, restaurants, transportation, and finance, including tour guides and their fellow travelers, are also important factors in achieving a high-quality tourism experience. These rich demands for services will generate enormous commercial value [6]. Therefore, this article adopts top-down overall analysis to design ideas from bottom to top. By analyzing the user's characteristics through the trajectory of user's mobile behavior, this paper constructs a travel recommendation system in the mobile point-to-point environment and a precise marketing model in the tourism industry based on the trajectory of user's mobile behavior, so as to provide appropriate services for the appropriate users at the right time and place, in order to provide reference for relevant tourism enterprises to achieve precise marketing.

2. User's Mobile Behavior Trajectory

2.1. User's Mobile Behavior Trajectory Definition. User's mobile behavior trajectory is based on the path that users find frequently in the location mobile path generated by daily life. The location information generated by user's daily behavior is acquired by GPS equipment sampling at a certain time interval, and the spatial position of moving object is represented by Euclidean space coordinates, discrete display in electronic map. Through moving sequence pattern mining, we can find the correlation among these discrete location information points and obtain the user's moving behavior trajectory. This will provide effective support for precision marketing in mobile e-commerce [7]. In this paper, we make the following definitions for user's mobile behavior trajectory.

Definition 1. Location information point: the position information points generated by the user's movement can be obtained by receiving devices such as GPS of mobile terminals. Each position information point indicates a position that the user has arrived at. Suppose an independent location information point is represented as two tuple $P = (Z, T)$, among them Z is the position coordinate, and its structure contains longitude $Z.\ x$ and latitude $Z.\ y$; T is the time information of arrival position Z.

Definition 2. Mobile behavior trajectory: mobile behavior trajectory can be obtained by GPS log. A mobile behavior trajectory consists of a sequence of position information points arranged in order of time attribute T. Suppose L is user's mobile behavior trajectory, then $L = P_1 \longrightarrow P_2 \longrightarrow \cdots \longrightarrow P_n$, where $P_i\,(0 < i \leq n)$ denotes any sampled position information point. Mobile behavior trajectory L satisfies any $0 < i \leq n$, $P_i \cdot T < P_{i+1} \cdot Ta$; n represents the number of location information points and represents it as the length n of mobile behavior trajectory.

Definition 3. Mobile behavior subtrajectory: represents the inclusion or inclusion relationship between two moving behavior trajectories. Suppose that L_1 and L_2 have two trajectories of moving behavior, where $L_1 = a_1 \longrightarrow a_2 \longrightarrow \cdots \longrightarrow a_i, L_2 = b_1 \longrightarrow b_2 \longrightarrow \cdots \longrightarrow \ \ b_n$, If there exists a positive integer $m_1, m_2, \ldots, m_i$, satisfying $1 \leq m_1 < m_2 < \cdots < m_i \leq n$, making $a_1 = b_{m1}, a_2 = b_{m2}, \ldots, a_i = b_{mi}$, then L_1 is said to be the moving behavior subtrajectory of L_2, or L_2 is said to be a moving behavior supertrajectory of L_1. It can be written as $L_1 \subseteq L_2$ or $L_2 \supseteq L_1$. The location information points are adjacent to the mobile behavior subtrajectory and are allowed to be nonadjacent in the original mobile behavior trajectory.

Definition 4. Support degree: the collection of all location information points of moving behavior trajectories constitute a database of mobile behavior trajectories. $\mathrm{DB} = \{L_1, L_2, L_3, \ldots, L_n\}$, where $L_i\,(0 < i \leq n)$ is mobile behavior trajectory and $|\mathrm{DB}|$ is the number of mobile behavior trajectories in the database. The number of mobile behavior trajectory t contained in DB is t of the support in DB:

$$\operatorname{support}(L_i) = \left|\{L \mid L \in D, L_i \subseteq L\}\right|. \tag{1}$$

Definition 5. Frequent Trajectories: when the support degree of the mobile behavior trajectory is greater than or equal to the minimum support threshold, the mobile behavior trajectory is called the frequent trajectory. $\mathrm{FT} = \{l \mid \operatorname{support}(l) \geq \min, l \subseteq L, L \in D\}$, L represents the mobile behavior trajectory sequence and D represents the mobile behavior trajectory sequence set.

The user's mobile trajectory records the user's activity status in the real world, which can reflect the user's behavior preferences and potential intentions to some extent. For example, if a user moves a lot every day, he may be an outdoor sports enthusiast. Through more fine-grained analysis, we can identify users' occupations, taste habits, and so on from their frequent locations and restaurants. Therefore, mining hot spots and planning roads through multiuser mobile trajectory data sharing is an important research content of this paper.

2.2. Classification of User's Mobile Behavior Trajectory. User's mobile behavior trajectory data refer to the sequence of changes in geographic location information caused by user's own motion behavior in a certain time and space environment. These geographic location information points which change with time series can form a user's mobile behavior trajectory data according to the order of occurrence time [8]. According to the different sampling methods, we can classify these user's mobile behavior trajectory data into three categories.

2.2.1. Location Sampling-Based User's Mobile Behavior Trajectory. A trajectory formed by a change in position during the movement of a user can be sampled sequentially according to the change in position. It focuses on the information of location change when the user moves. The data obtained by this method have abundant semantic information and very detailed location change information. We can record the trajectory data of user's mobile behavior based on position sampling by recording discrete variables. The trajectory of user's mobile behavior can be represented by the sequence of sampling points with the change of moving object's position, and it can be formally expressed as

$$L = \{(x_1, y_1, t_1, \ldots), \ldots, (x_i, y_i, t_i, \ldots), \ldots, (x_n, y_n, t_n, \ldots)\}. \tag{2}$$

The location $(x_i, y_i), 1 \leq i \leq n$ denotes the geographical location of the mobile user at the time of t_i, and the location (x_i, y_i) of the mobile user at the time of t_i and the location (x_{i+1}, y_{i+1}) of the time of t_{i+1} are not the same.

Trajectory can be divided into three segments according to the information of stopping point, boarding point, and alighting position, and the trajectory can be preserved according to different semantics and application segments. For example, in the prediction of travel time, it is necessary to delete the stopping point, which may be the vehicle parking or waiting for passengers, in order to measure the trajectory travel time more accurately. For some tasks that analyze the similarity between two users, it is often necessary to use the residence trajectory to reflect the user's region of interest.

2.2.2. Time Sampling-Based User's Mobile Behavior Trajectory. The change of mobile user's behavior is sampled by definite time interval to form the trajectory data of user's mobile behavior, which is called the trajectory of user's mobile behavior sampled according to time. This kind of sampling focused on the change of location information points caused by the change of mobile user's behavior at the same time interval, which has the characteristics of large data volume and wide range. The time-sampled trajectory data of user's mobile behavior is formalized as follows:

$$L = \{(x_1, y_1, t_1, \ldots), \ldots, (x_i, y_i, t_i, \ldots), \ldots, (x_n, y_n, t_n, \ldots)\},$$
$$t_i = t_1 + (i-1)\Delta t, \tag{3}$$

where L is a trajectory data of mobile behavior, Δt is equal interval time, (x_i, y_i), and $1 \leq i \leq n$ denotes the location of the mobile user at any time of t_i. If the time interval between the two sampling points is larger than the threshold value, the trajectory can be divided into two segments through the two sampling points.

2.2.3. User's Mobile Behavior Trajectory Triggered by Events. The trajectory of mobile user's mobile behavior, which is recorded by the system after the sensor event is triggered, is obtained by the event triggering [9]. This sampling method focuses on the event set that triggers the sensor to work. It has the characteristics of short update period and representative sampling objects. Although the behavior of mobile users changes with time, the system does not record the trajectory according to time or position, but only records the trajectory information of mobile users when they produce some specific behavior and trigger sensor events. We can also use discrete variables to record the behavior trajectory of mobile users and formalize it as follows:

$$L = \{(x_1, y_1, t_1, \ldots), \ldots, (x_i, y_i, t_i, \ldots), \ldots, (x_n, y_n, t_n, \ldots)\}. \tag{4}$$

The location (x_i, y_i), $1 \leq i \leq n$, denotes the location of the mobile user at the time of t_i, and the location of the mobile user at the time of t_i, (x_i, y_i) and t_{i+1} can be the same (x_{i+1}, y_{i+1}).

When the trajectory direction changes beyond the threshold value, we can mark the key points according to the direction changes and divide the trajectory into two segments.

2.3. User's Mobile Behavior Pattern Decision. According to the trajectory data of user's movement behavior, the speed of completing the trajectory is calculated by time, and then the user's behavior pattern is determined. Many problems still need to be considered, such as road congestion, construction, and even traffic accidents, which will affect the speed of user behavior. Vehicles travel much faster than people's walking speed on normal roads, but in congested or abnormal roads, the speed difference between vehicles and people's walking speed is not obvious. Therefore, the identification accuracy of trajectory velocity can only be less than 50% through time calculation [10]. In addition, the user may change several different behavior patterns in the same trip, which makes the same user's moving behavior track contain a variety of different speeds. In the overall calculation, if the average speed is obtained, it is obviously not correct to determine the user's behavior patterns. Therefore, it is necessary to divide the user's moving behavior trajectory into several trajectory segments reasonably. By comparing different trajectory segments, we can analyze whether the user has changed the behavior pattern and further improve the recognition accuracy.

How to realize the reasonable division of user movement behavior segments is the problem we want to study. As shown in Figure 1, the walking user and the driving user travel the same way, but the trajectory data of the user's movement behavior are obviously different. We can analyze the following three aspects:

(1) Because the trajectory data of user's moving behavior produced by walking often produce direction change or reciprocating motion, we can divide the trajectory segments according to the change of the trajectory data direction of user's moving behavior. In mobile scenes, people get off a bus, walk to another station to continue to take the bus process, and must pass through a section of walking, although the walking section is short, but still can show obvious direction changes.

(2) The trajectory data of mobile behavior produced by driving users do not change significantly in direction. This kind of characteristic is not affected by traffic conditions. We can train a classification model by the supervised learning method. For example, drivers do not change their direction as freely and frequently as pedestrians do, resulting in a straight line in the trajectory of the user's movement behavior, and the direction of change is not obvious [11].

(3) We can also judge user behavior patterns by the shape of user behavior trajectory data, especially the trajectory of user behavior generated by different user behavior patterns in a journey, which will have obvious morphological changes of trajectory.

3. Analysis of User's Mobile Behavior Trajectory Data

This paper studies the trajectory of user's mobile behavior generated by online travel users during their mobile process. It contains a lot of information to express the personalized behavior of mobile users. We can use data mining methods such as classification, clustering, frequent itemsets, cycle discovery, and anomaly detection to mine and analyze the trajectory of tourism users' mobile behavior.

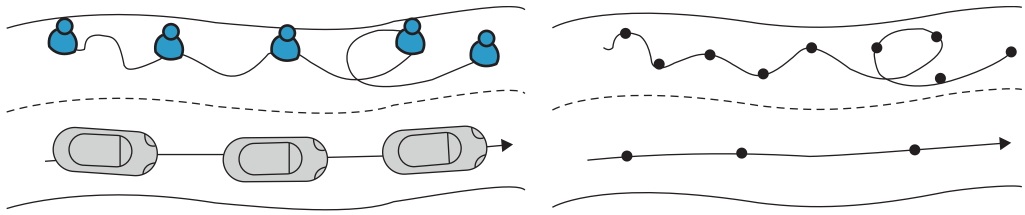


Figure 1: Differences in movement behavior between (a) walking users and (b) driving users.

3.1. Dividing Trajectory Segments. Each user movement behavior trajectory can be regarded as an image data. Structural Similarity Index (SSIM) can effectively measure the similarity of two trajectories, and clustering based on the similarity index is more accurate than traditional clustering based on Euclidean distance index [12]. The accuracy of structure similarity matching is closely related to the reasonableness and validity of the segmentation of user motion trajectory. Therefore, this section mainly studies how to detect the large-angle mutation points in the user's moving behavior trajectory, and how to partition and store the user's moving behavior trajectory records at the mutation points, so as to obtain some trajectory fragments which tend to be stable before clustering.

Each user movement behavior trajectory cannot be a straight line. As the precision of position coordinate recording is higher and higher, the direction of each track will change more and more, especially some subtle direction changes, and the angle of rotation can reflect the degree of change of the track direction. The division of track segments is determined according to the size of the track angle. However, if every corner is stored, it is not conducive to reduce the storage of the corner, and it is not conducive to extract it to divide the trajectory segments. Therefore, by storing the large turning point, we can discover and identify the changes of user behavior or abnormal conditions, which is also conducive to retaining the relatively stable local structure features of user trajectory segments.

We define the turning angle of user's moving behavior trajectory as the turning angle caused by the change of direction of adjacent trajectory segments, which can reflect the movement trend of trajectory and the change of user's behavior [13]. As shown in Figure 2, the angle between the direction changes of the user's moving behavior trajectory can be expressed as α, and the angle of rotation can be divided into outer angle and inner angle, expressed as θ_1 and θ_2, respectively. We set the outer rotation angle θ_1 as a positive value and the inner rotation angle θ_2 as a negative value to facilitate the similarity calculation of the trajectory segments.

As can be seen from Figure 2, the formula for calculating the angle alpha of the direction change is shown in Formula (5), where a, b, and c represent the adjacent and opposite sides of the angle α, respectively.

$$a = \arccos \frac{a^2 + b^2 - c^2}{2ab}. \tag{5}$$

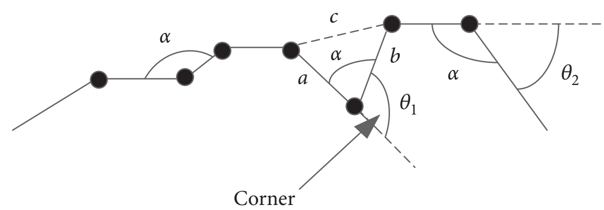


Figure 2: Corner of user's mobile behavior trajectory.

According to the above formula, the formula for calculating the angle theta can be obtained (6):

$$\theta = \begin{cases} 180 - \alpha, & \text{if } (a \times b \geq 0), \\ \alpha - 180, & \text{if } (a \times b < 0). \end{cases} \tag{6}$$

This is the first step to partition the trajectory segments of user's mobile behavior. Using formulas (5) and (6), the trajectory segment partitioning algorithm can be implemented (Algorithm 1).

Some trajectory fragments obtained by calculating rotation angle, setting threshold, and partitioning trajectory fragments can be expressed as a set of several feature attribute vectors. These feature attributes can comprehensively express the local features of a trajectory fragment and the global features of user's moving behavior trajectory. In this section, the trajectory fragment is not simply the expression of coordinate information of the position information points, but extracts the speed, shape, position, rotation angle, acceleration, and other characteristic vectors from it. Using these eigenvectors, we can enhance the accuracy of analyzing the trajectory of user movement. We formally represent the trajectory fragment structure as follows: $\mathrm{TS} = (D, S, A, L)$. In addition to the above four features, we should also calculate the distance, time, and other features, using vector $W = \{W_D, W_S, W_A, W_L\}$ to represent the weight of the four feature vectors.

Since the weights of feature vectors correspond to the eigenvectors of the trajectory segments, their values should be greater than or equal to zero, and the sum of their weights should be 1; we can generally assume that the weights of all feature vectors are equal probability, and we can take the average value of 0.25 as the weights. Similarly, we can adjust the weights of each feature vector according to the sensitivity of the feature vectors of the trajectory fragments in the actual scene. For example, when analyzing the position-sensitive

Step 1: one by one, scanning the location information point sequence in the user movement behavior track;
Step 2: formula (5);
Step 3: formula (6);
Step 4: Set a threshold ω for corner θ, store the corner satisfying $|\theta| > \omega$ as a mutation point, and then divide the track segment according to the position information point of the corner. n is the number of sampling points, and the time complexity of the algorithm is $O(n)$.

Algorithm 1

trajectory fragments, we can focus on the position vectors, and the weights $W_D = W_S = W_A = 0, W_L = 1$ are also feasible.

According to the feature vector and its weight to complete the structural similarity comparison, mainly through the analysis of the differences between the feature vectors of the trajectory segments to complete the comparison [14], according to the definition of the trajectory segment structure, we can define two trajectory segments are $L_i, L_j, 1 \le i \ne j \le n$. The comparison function of two trajectory segments is $D(L_i, L_j)$, velocity vector is $S(L_i, L_j)$, angle vector is $A(L_i, L_j)$, and position vector is $L(L_i, L_j)$. The four comparison functions above constitute the calculation of structural similarity of the trajectory segments, as shown in the following Formulas (7) and (8). The function $N(\cdots)$ denotes the normalization of the distance. Because the range of each eigenvector in the trajectory segment is different, the normalization of the distance is the normalization of the distance of each eigenvector. The SSIM of structural similarity is represented by 1 minus the normalization of the distance:

$$S\left(L_i, L_j\right) = \left(D \times W_D + S \times W_S + A \times W_A + L \times W_L\right), \tag{7}$$

$$\mathrm{SSIM}\left(L_i, L_j\right) = 1 - N\left(S\left(L_i, L_j\right)\right). \tag{8}$$

The structural similarity comparison of trajectory fragments can express the structural differences of each trajectory fragment on the feature vectors. Therefore, the smaller the SSIM value of the trajectory fragments, the greater the SSIM value of the trajectory fragments. Moreover, the distance between the structural similarities of the trajectory fragments is symmetrical, that is, $\mathrm{SSIM}(L_i, L_j) = \mathrm{SSIM}(L_j, L_i)$. Therefore, it can be found that the method based on structural similarity can well reflect the structural differences between trajectory segments.

According to structural similarity, the direction information, speed information, angle information, and position information are compared [15].

(1) The direction vector comparison function $D(L_i, L_j)$ denotes the degree of similarity of two similar trajectory segment L_i, L_j in the direction of motion. As shown in Figure 3(a), φ is the angle between the direction of the trajectory, and the formula for calculating direction vector comparison function is as follows:

$$D\left(L_i, L_j\right) = \begin{cases} \left\|L_i\right\| \times \sin \phi, & \text{if } (0^\circ \le 90^\circ), \\ \left\|L_j\right\|, & \text{if } (90^\circ \le 180^\circ). \end{cases} \tag{9}$$

If two similar trajectory fragments have the same direction and the angle φ is small, the two trajectory fragments tend to be parallel in the same direction, which is called the best state, then the Dir Dist value approaches zero. If two similar trajectory fragments are in opposite directions and the two trajectory fragments with larger angle φ tend to be in reverse parallel, the worst condition is that the Dir Dist value is the length of the trajectory fragments involved in the comparison.

(2) The speed vector comparison function $S(L_i, L_j)$ expresses the trend of user mobility. The velocity vector comparison function is shown in Formula (10), where $S_{\max}(L_i, L_j)$ is $|V_{\max}(L_i) - V_{\max}(L_j)|$, representing the absolute value of the maximum velocity difference between the trajectory segments. Similarly, $S_{\mathrm{avg}}(L_i, L_j)$ and $S_{\min}(L_i, L_j)$ represent the absolute value of the difference between the average velocity and the minimum velocity, respectively. We can judge the difference of velocity vectors from the three aspects of maximum, minimum, and average velocity:

$$S\left(L_i, L_j\right) = \frac{1}{3}\left(S_{\max}\left(L_i, L_j\right) + S_{\mathrm{avg}}\left(L_i, L_j\right) + S_{\min}\left(L_i, L_j\right)\right). \tag{10}$$

(3) The angle vector comparison function $A(L_i, L_j)$ expresses the degree of eigenvalue change caused by the change of direction in the trajectory segment. As shown in Formula (11), where the angle of rotation θ is calculated according to Formula (6), the internal rotation angle is positive and the external rotation angle is negative, the angular distance of the trajectory segment is the cumulative value of many internal corners of the trajectory, and the direction of change within the trajectory segment can determine the value of each angle:

$$A\left(L_i, L_j\right) = \frac{\sum_{1,1}^{P\left(L_i\right), P\left(L_j\right)} \left(\left(\left|\theta_i - \theta_j\right|\right) / \left(\left|\theta_i\right| + \left|\theta_j\right|\right)\right)}{P\left(L_i\right) + P\left(L_j\right)}. \tag{11}$$

Figure 3(b) shows that if each corner of the two trajectory segments rotates to L_i and L_j matches, the value of the angle vector comparison function is 0, which is the best case. If the two trajectory segments turn to L_i and L_j in opposite directions, that is, the two trajectory segments are

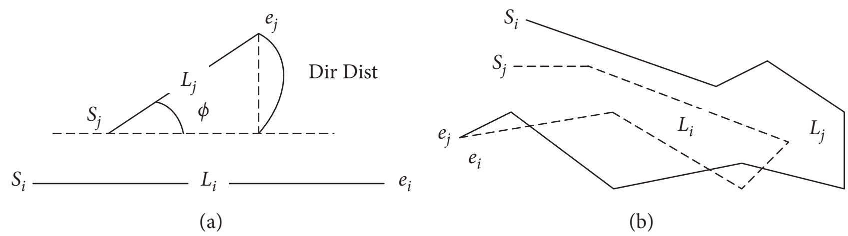


FIGURE 3: Comparison of track direction and rotation angle: (a) direction contrast; (b) corner contrast.

in opposite jagged shape, and the value of the angle vector comparison function is 1, this is the worst case.

(4) For the position vector comparison function $L(L_i, L_j)$, we can use Hausdorff distance to measure the location distance of the trajectory segment, as shown in the following formula:

$$L(L_i, L_j) = \max\left(h(L_i, L_j), h(L_j, L_i)\right), \quad (12)$$

where $h(L_i, L_j) = \max_{a \in L_i}(\min_{b \in L_j}(\mathrm{dist}(a, b)))$ is the direct Hausdorff distance between L_i and L_j, i.e., the maximum distance from a point in L_i to the nearest L_j and $\mathrm{dist}(a, b)$ represents the Euclidean distance function between points.

3.2. Similarity Computation of User's Mobile Behavior Trajectory. At present, we collect and store the trajectories of tourism users' mobile behavior, cluster the typical similar trajectories from these trajectory data, analyze the behavior patterns of user's mobile behavior trajectories, and predict the personalized needs of tourism users based on structural characteristics. Clustering analysis is to divide user behavior trajectory into several groups with high cohesion and low coupling. It requires high similarity of user behavior trajectory in the same group, and low similarity of user behavior trajectory in different groups. The goal of clustering analysis is to find out the trajectory data with the same or similar behavior patterns from the trajectories of some users' mobile behaviors, analyze the personal preferences, consumer demands, and behavior characteristics of the trajectories of tourism users' mobile behaviors, and accurately determine the similarity between trajectories of users' mobile behaviors. At the same time, the trajectories of users' mobile behaviors with high similarity are gathered into one class [16].

Most of the online travel users are in the same scenic spot, similar routes to carry out activities, and most of the resulting mobile behavior trajectory data have local similarity and global dissimilarity. It is difficult to find the personalized characteristics of tourism users by analyzing the complex and large number of users' mobile behavior trajectories and effectively extract users. The analysis of a part of the mobile behavior trajectory is more conducive to finding the information contained in it [17]. Therefore, the trajectory analysis method based on the whole trajectory in traditional research is easy to cause the inaccuracy of trajectory analysis. In this paper, we use structural features to calculate the similarity of user movement behavior. This method needs to calculate every corner of the user's mobile behavior trajectory and find the sampling point with larger rotation angle, which is regarded as the sudden change point of the user's mobile behavior, and then divides the trajectory segment by the sudden change point. In this way, the rotation angle of each trajectory segment obtained will not change significantly, and the trajectory structure tends to be stable. Then, a trajectory model of user's mobile behavior is constructed, which is characterized by trajectory direction, trajectory speed, trajectory angle, and trajectory distance. Taking these features as parameters, threshold values are set to express and adjust the weights of each feature according to the actual application scenarios, and a trajectory similarity algorithm is constructed to calculate the user's movement behavior. The object of this paper is to calculate the structural similarity of some trajectory segments which are divided according to the sudden change points of large turning angles by using the trajectory similarity algorithm constructed with structural features as parameters. It is used to judge the similarity degree of each user's moving behavior trajectory and then completes the feature analysis of user's moving behavior trajectory. The simulation results show that the trajectory similarity calculation algorithm is efficient, the weight adjustment of each structural feature is flexible, and the trajectory analysis results are more in line with the needs of practical application scenarios and have higher application value and practical significance.

On the basis of obtaining the feature vector distance of user's moving behavior trajectory segment, the trajectory segment with high similarity is analyzed, and then the clustering algorithm is used to complete the clustering of user's moving behavior trajectory. By comparing the structural similarity between the trajectory segments and other trajectory segments which are not on the same trajectory, a number of ε-nearest neighbor sets of trajectory segments are formed. The number of ε-nearest neighbor sets is used to determine the midpoint of trajectory segment clustering, and then the trajectory segment clustering is realized. A trajectory segment clustering algorithm based on structural similarity is constructed.

The steps of clustering algorithm based on structural similarity are given in Algorithm 2.

From the analysis of the above algorithms, it can be seen that, in the trajectory segment clustering algorithm based on structural similarity, it is very important to determine the

Step 1: first calculate the corner θ of each track segment sampling point P_i;
Step 2: according to the corner threshold ω, we divide the trajectory of user movement into TS of some track segments.
Step 3: calculate the distance between the trajectory feature vectors based on the weight of the trajectory segment feature vectors.
Step 4: calculate the ε-nearest neighbor set of the track segments with high similarity.
Step 5: the distance clustering segment is centred on the similarity track segment ε-nearest neighbor set.
Step 6: initialize clustering ID and track segment clustering markers.
Step 7: traverse the trajectory fragments, find the core clustering and set the clustering ID, and then add the pointers of these trajectory fragments to a new node in the index tree.
Step 8: determine whether the set center of ε-nearest neighbors meets the specified distance. If it meets the requirement, then add the cluster ID marker to the trajectory fragment, expand the clustering, construct the index tree node, and repeat steps 7 and 8 until all trajectory fragments are traversed.

Algorithm 2

threshold value of ω, ε-nearest neighbor, and the threshold value of σ nearest neighbor number, which can directly affect the time complexity and space complexity of the algorithm. It needs to be verified repeatedly and determined according to the actual application fields. Therefore, we mainly analyze the algorithm qualitatively.

Through repeated verification of the algorithm, in the data analysis of trajectory of travel user's movement behavior, the value of ω cannot be set too small, and if set too small, some characteristic details of trajectory segments will be lost. On the contrary, the value of ω cannot be set too large and cannot effectively identify the abrupt change point or sampling abnormality of the trajectory segment, which directly affects the structure of clustering analysis. Similarly, if the threshold value σ of the number of neighbors is set to be large enough, then no trajectory segment can satisfy the requirement of $|N_{\varepsilon}(L)| \geq \sigma$, and all trajectory segments will be marked as abnormal conditions. On the contrary, if σ is set too small, all the trajectory segments may become the clustering center, so that the trajectory segments will be self-contained and the number of clusters will be too large.

3.3. Discovery of Popular Tourist Attractions. By effectively identifying the location information points in the trajectory data of users' mobile behavior, the feature vectors of the trajectory segments can be extracted, and the semantics of these location information points can be expressed as the route, the scenic spots, and the behavior patterns of an online travel user in the past period of time. By clustering and analyzing the trajectory fragments containing location information points, we can find that the traveling users have a longer time in a certain area, which can be interpreted as the tourist users have a higher degree of interest in a certain scenic spot. Semantic expression is a popular tourist spot with longer stay time for online travel users. In practical scenarios, many traveling users will visit the same or similar scenic spots. From the trajectory of users' mobile behavior and the region of interest, traveling users with similar trajectory and the same region of interest can predict their similar preferences or similar behavior characteristics. These regions of interest frequently stayed by tourist users will appear as overlapping regions in the trajectory of user's mobile behavior. If these overlapping regions are found, the popular scenic spots concerned by tourist users can be found and the users who like these scenic spots can be clustered. And then, dig out the other characteristics of these users to complete the personalized tourist attractions recommendation of similar tourists. We extract the feature parameters of these overlapping areas, such as overlap time and overlap times, which can reflect the similarity between the traveling users. It can identify the tourist attractions that the tourist users are interested in during the mobile process and recommend the most likely popular tourist inventory for other tourist users who have a higher similarity with their user's mobile behavior trajectory, so as to tap the potential preferences of the tourist users [18]. Assuming that travel user A and travel user B share a higher degree of similarity in the trajectory of users' mobile behavior, it can be found that some scenic spots are visited by users A but not by users B. Through mining, it is known that these scenic spots may be of interest to users B. Then, we can recommend these scenic spots to B users through A users, so that these scenic spots become the potential and most likely scenic spots for B users to visit. We can also use the activity sequence to express the popular scenic spots that tourists often visit, and the trajectory of nearby tourists is more instructive [19].

In the process of analyzing mobile user behavior trajectory data with structured eigenvectors, it is not difficult to find that the moving speed of user behavior trajectory is not the same in different time periods, or it is slower in a certain time period, or it is faster in a certain time period. Figure 4 shows the moving speed of the user in different periods of time, in which the trough is formed during the period when the user moves slowly and the peak is formed during the period when the user moves fast, but both trough and peak can indicate the user's continuous generating activity. And, the slow moving trough time contains more user behavior characteristics, so this paper focuses on the behavior characteristics of mobile users in trough situation.

As shown in Figure 4, by comparing the speed, distance, and time of user's moving behavior trajectory, the structured features of mobile users and the behavior differences of tourist users can be clearly analyzed. We focus on the analysis of two dimensions: the speed and the time of the slower wave trough. As shown in Figure 5, the slower the traveling speed of the tourist user, or the less the change of

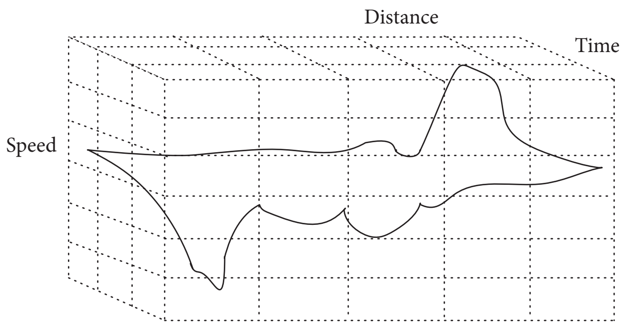


Figure 4: Mobile speed of user behavior trajectory.

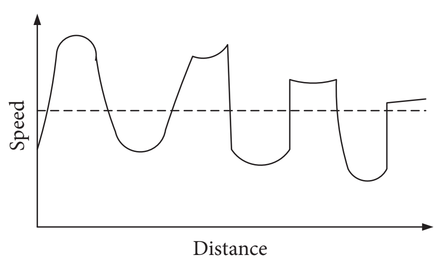


Figure 5: Two-dimensional trajectory analysis of user movement behavior.

the active area in a period of time, the most likely the predictable user behavior characteristic is; that is, the traveling user stays at a certain scenic spot for browsing, resting, or taking photos. The longer the trough, the more attractive the scenic spot is. The more tourists are staying at the same scenic spot, the more scenic spots can be designated as popular tourist attractions.

For example, when a tourist visits a scenic spot, he or she forms a trajectory of the user's movement behavior. Three troughs appear in the trajectory, indicating that the user may have experienced three scenic spots or rest areas, of which the first trough has a shorter experience. It shows that tourists spend less time visiting the first scenic spot, travel faster, continue to move forward at a faster speed after the tour, and spend a little more time watching the tide or taking pictures when they meet the scenic spot of interest. So tourists will slow down, move in a more fixed area, and travel at a slower speed, thus appearing the second trough period, after the tourists continue to move forward; when the formation of the third trajectory speed reached trough state, semantic expression may have two situations. The first is that the tourists reach a certain degree of fatigue or meet a rest area, stop and rest; the second is that the tourists arrive at a well-known scenic spot, gather more tourists, people will stay in a certain position, waiting for sightseeing and photography, moving slowly, and almost stop. The above two semantics can be distinguished by whether the location in the electronic map is a resting area or a scenic spot. However, in the actual tourist attractions, the situation may be more complicated. For example, a tourist is an outdoor sports enthusiast who has good physical strength and likes natural scenery. Because of his fast moving speed, there is little difference between the wave crest and trough of the waveform trajectory formed by the speed and distance. Although his tour speed is fast and his stay time is short, the location he stays in is still the area of interest. In this way, moving objects with similar frequencies in the velocity-distance waveform can be found not only in the known hot spots of the users, but also in the scenic spots that the potential users may be interested in, even in the preferences, occupations, and personality characteristics of the tourist users. It helps to gather tourists with similar preferences and similar personalities to achieve the confluence module [20].

Popular scenic spots refer to scenic spots with long staying time after arrival [21]. In the user's mobile behavior trajectory, the hot spots can be marked as $H = \{H_1, H_2, \ldots, H_n\}, H_j = \{L_i, L_{i+1}, \ldots, L_m\}$, H is used to denote a trajectory fragment. When the traveling user passes through a hot spot area with high interest and stays for a long time, the trajectory fragment moves at a speed close to or far below the normal trajectory speed. We can think that the tourist user has conducted a deep browsing in the scenic spot or some behavior activities have taken place in the scenic spot area. We can analyze the information such as the time of arrival, the time of stay, and so on. The region with dense user access points can be expressed as a popular tourist attraction area with high user access frequency [22].

Because GPS receiving equipment receives satellite signals in vast and open areas with high intensity and good positioning effect, satellite signals in indoor areas will be shielded by the wall, resulting in weak positioning signal and reduced positioning accuracy [23]. Therefore, when analyzing tourists' preference for scenic spots through the status of stay, it is necessary to distinguish between outdoor and indoor scenic spots. The positioning signal of outdoor scenic spots is good and has high precision. It can acquire the location information points at sampling frequency in real time and form the locus of user's movement with dense location information points. The positioning signal of indoor scenic spots is weak, which affects the positioning accuracy. Even when the signal is lost, the location information points cannot be obtained in time according to the sampling frequency requirement, and the space area of the indoor attractions is small, which makes some location information points overlap. This repetitive activity can also find that the tourists are visiting a certain indoor attractions regularly. The popular scenic spots are divided into two types: one is the outdoor scenic spots, such as natural landscape, gardens, playgrounds, and other broad areas, in a longer period of time, can obtain more dense location information points formed by the user's mobile behavior trajectory, recorded as HR_{II}; Another kind is indoor scenic spots, such as restaurants, shopping malls, tourist centers, and other closed areas, in a long period of time, may lose a certain location information point sampling information, but after leaving the area, they can get the location information point again, recorded as HR_I. Firstly, the trajectory data of user's mobile behavior are obtained by sampling the location information points, and then the trajectory data of user's mobile behavior is denoised. Finally, according to the characteristics of the location information points, the HR_I and HR_{II} popular scenic spots domain are divided by the density clustering method. The steps for finding popular scenic spots is given in Algorithm 3.

Input parameters: user movement behavior trajectory Q, minimum speed S, minimum time T, and maximum disturbance threshold MT.
Output parameters: collection of popular scenic spots HR.
Step 1: for $(i = 2,\ i \leq |Q|,\ i{+}{+})$ /∗ $|Q|$ represents the number of location information points∗/
Step 2: $D\ [i-1] \cdot T = \text{cal}\ T\ (p_i - 1,\ p_i)$;
Step 3: $D\ [i-1] \cdot S = \text{cal}\ D\ (p_i - 1,\ p_i)/D\ [i-1] \cdot T$;
Step 4: HR = {}; C = {}; CO = false;
Step 5: for $(j = 2,\ j \leq |T| - 1,\ j{+}{+})$ /∗cycle search indoor attractions area HRI∗/
Step 6: if $(D\ [j-1] \cdot T > \mathrm{T}$ and $D\ [i-1] \cdot S < n * S)$ then
Step 7: $C = \{p_j - 1,\ p_j\}$; /∗record location information point stay area∗/
Step 8: if (not CO) then CO = true;
Step 9: else $C = \text{Update}\ \{p_j - 1,\ p_j\}$; /∗merge the location information points closer to the collection HR∗/
Step 10: else if (CO) then
Step 11: HR = {C}; CO = false; C = {} /∗search outdoor attractions area HRII∗/
Step 12: if $(D\ [j-1] \cdot S \leq S)$ then /∗determine activity intensive areas∗/
Step 13: $C = \{p_j\}$;
Step 14: if (not CO) then CO = true;
Step 15: else if (CO) then
Step 16: last Index = look Ahead (MT, S);
Step 17: if (last Index $\leq j +$ MT) then
Step 18: for $k =$ last Index downto j do
Step 19: $C = \{p_k\}$;
Step 20: $j =$ last Index;
Step 21: else if (time $(C) > T$) then
Step 22: HÈ = $\{C\}$; $C = \{\}$; CO = false;
Step 23: return HR;

Algorithm 3

Because the location information points sampled by GPS are affected by the factors of region, space, and weather, it is easy to have inaccurate positioning or interruption of positioning [24]. When the user enters the indoor scenic spots from the outdoor scenic spots, the short signal interruption will occur, and the positioning data receiving error will easily occur. In order to adapt this error, a maximum disturbance threshold MT is set in the hot spot detection algorithm to enhance the accuracy of location information points.

The outdoor and indoor scenic spots are divided according to the location information of popular scenic spots. The density-based hot spot discovery algorithm is used to retrieve two different types of user residence areas, outdoor and indoor, in the trajectory of user's mobile behavior, and define them as hot spots [25]. The algorithm has four input parameters: trajectory of user's mobile behavior, minimum speed, minimum time, and maximum disturbance threshold. Among them, the threshold of minimum speed is related to the activity speed of tourist users in the scenic spot area. If walking tour, the general speed is 2 to 3 meters per second. If the sudden reaction speed of the location information points slows down significantly, it indicates that the tourist users have arrived at the scenic spot area, and specific user behavior has taken place. On the contrary, if the sudden response speed of the location information points increases over a period of time, it indicates that the travel users have changed their behavior patterns. For example, we can leave the scenic spot and take a sightseeing bus to the next scenic spot, so we can set it according to the sampling time. There is no absolute fixed value in the setting of the minimum time threshold. Generally, if a tourist user stays in a certain area for more than 30 minutes, it can be considered that the tourist user arrives at a scenic spot or rest area, and changes in user behavior have taken place, resulting in a specific activity. The maximum perturbation threshold is only used to express the number of continuous perturbations when the abnormal location information points are sampled. If the number of abnormal location information points is smaller than the perturbation threshold, the abnormal sampling information points can be merged into the normal location information points of the user's mobile behavior trajectory data set. If the number of abnormal location information points sampled in the former HR is larger than the perturbation threshold, it is necessary to preserve the current user's mobile behavior trajectory and then detect the new hot spots after abnormal location information points sampled to form a new user's mobile behavior trajectory. The settings of the minimum time threshold and the maximum disturbance threshold are related to the sampling frequency of the location information points. In the popular scenic spot area detection algorithm, the trajectory of the user's movement behavior is traversed twice. The algorithm complexity is linear order $O(n)$. Among them, n denotes the number of location information points in the trajectory of the user's mobile behavior, and the algorithm can retrieve the hot spots with frequent activities.

4. Travel Recommendation Model

4.1. Application Scene

(1) *Tourist Attractions Recommendation*. In the mobile scenario where the user completes the tourist attraction tour, the tourist user will generate multidimensional information to realize the application scenario recommended by the tourist attraction, which can hide the rarely used dimensions. Focus on the user's city, mobile behavior trajectory in the corner location information points, hot spots and user behavior patterns and other dimensions, accurate analysis of user preferences, the same characteristics of the user recommend the most likely favorite tourist attractions [26].

(2) *Hotel Catering-Related Recommendation*. Tourist users need to visit other applications frequently in the process of touring, such as electronic map applications, O2O applications, restaurants, e-commerce, outdoor equipment, etc., a single visit to each application, users need to frequently exit one application login another application, will cause user inconvenience, inefficiency, etc. According to the application relevance assessment model, we analyze the application of tourism users' preferences. Establish the relationship between these applications, so that users can access a travel APP directly related to other applications they are used to daily life, making travel APP a personalized all-round service platform [27].

(3) *Tourism User Preference Content Recommendation*. Tourist users may often inquire about certain contents, such as outdoor equipment, fitness and health care, restaurants and entertainment, and surrounding scenic spots in the process of using the application. According to the content retrieval evaluation model, users can retrieve keywords, learn user preference content, and discover their hobby characteristics and behavior characteristics. According to these preferences, travel APP can construct a personalized interface for users, giving priority to the information, articles, and news that travel users are interested in [28].

4.2. Travel Recommendation System. With the combination of Bluetooth, WIFI, and other RF communication technologies and mobile terminal devices, mobile point-to-point communication environment has been derived, and many different research topics have also emerged. This research talks about the relationship between mobile attraction recommendation system and social software from the perspective of mobile social software and completes interaction through user comment sharing in mobile point-to-point environment. In this paper, an interactive system of tourism comment information sharing and social networking software is established, which includes three functions: recommendation, reunion, and comment. It is used to explore the interaction between users in mobile point-to-point environment. The preliminary test has been completed in this paper. The experimental results show that the recommendation, convergence, and comment functions of the system can provide precise services for users and provide a basis for further research on the wide application of user behavior trajectory in precise marketing.

This paper focuses on the problem of information sharing and social interaction of tourism mobile recommendation system in mobile point-to-point environment. The system mainly includes three functions: recommendation, convergence, and information sharing. In the recommendation function section, we assume that users will leave comments and other information after visiting a scenic spot. When other users meet with them, they can exchange comments through RF communication technology. These comments are calculated by system algorithm to recommend scenic spots that meet users' interests. In addition, users can also actively send requests to other people to join the information and find similar interests around users to visit a scenic spot. Of course, users can also actively share location, comments, traffic conditions, tourist density, and other related information.

In order to enable the interaction and sharing of information between remote users, the relay mode under mobile point-to-point can be adopted. Every user in the system plays the role of information transmission, that is to say, each user's mobile terminal is a relay node for information transmission and constantly transfers the information they have mastered to the users at a long distance.

4.2.1. System Architecture. The mobile peer-to-peer environment mainly transmits information through the direct transmission between peer-to-peer users and the relay mode assisted by the third party. Using this feature, the system proposed in this paper mainly provides three services: recommendation, convergence, and review. First of all, the main purpose of recommendation service is to recommend scenic spots similar to user's interests to users through user's evaluation information, so that users can have a reference direction for the next destination in the journey, so that users can travel more smoothly. Secondly, the convergence service allows users to initiate a convening activity, gather other interested users around, visit the scenic spots together, or buy specialty goods together, through group buying to get a better price, or to strive for preferential services. Thirdly, evaluation services are divided into general information and specific information. General information is simply the transmission of personal information and specific information, so that the use of convergence services conveys the convening activities of the department of the offensive, through short messages, and the expression of personal information is incompatible; specific information is only for convening activities issued information. To provide the above services, the system architecture is presented in Figure 6.

(1) Interface Module. This module is responsible for the user and the system function docking; through this module, the system function interface is expressed and the user is guided to carry out the operation of various functions.

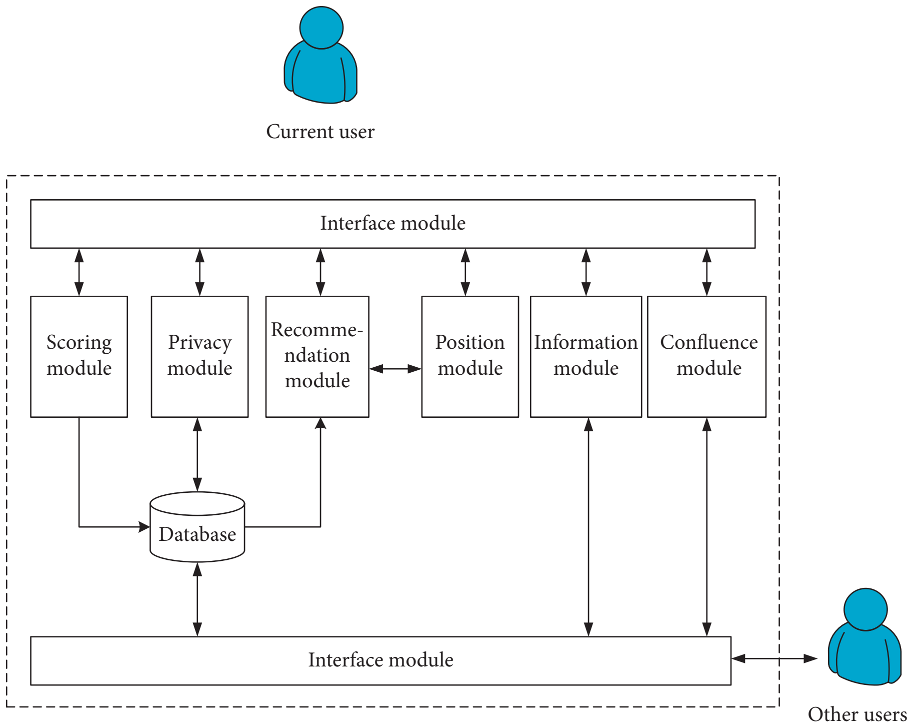


Figure 6: Mobile point-to-point tourism recommendation system architecture.

(2) Scoring Module. At present, the commonly used recommendation system is based on the scoring mechanism, which collects user's scoring data to calculate and provide recommendation services. The scoring module proposed in this study is mainly responsible for recording the user's evaluation of scenic spots. At the initial stage, the recommendation system often faces the problems of incomplete user scoring information, too many items not scored, and the difficulty of calculation caused by the noise of data, resulting in the decline of recommendation accuracy. Therefore, this study classifies scenic spots, requires users in the initial stage, must be based on the type of scenic spots scoring to ensure that individual users in the initial stage, and needs to score for the type of scenic spots to ensure that individual users' scoring information has been scored by the common column.

(3) Transport Module. Because this research system is built in the mobile point-to-point environment, user scoring, information, and other need to be obtained and transmitted through the transmission function; this study uses Bluetooth transmission technology to achieve the transmission of related functions. This module enables the system to automatically exchange scoring data through Bluetooth without interfering with the user when they meet, so as to achieve the purpose of collecting data. In terms of scoring exchange mechanism, this study currently uses unlimited scoring exchange method, when users meet, the exchange of all the scoring data held by both sides. However, information that has not yet been scored is not helpful to the recommendation system. Therefore, in the scoring exchange module, it is assumed that only the user has scored more than five scenic spots before the exchange, while the other scoring data obtained by others is more than five items before passing on to other users. On the other hand, the transmission module has the search function and can discover other users around the user; when the user wants to pass its ideas to the surrounding users, it can be completed through this module.

(4) Recommendation Module. This study analyzes the recommended operation by exchanging accumulated score data. This recommendation module uses collaborative recommendation and Pearson correlation coefficients to perform recommendation operation. The formula is shown in (13). Suppose that $U(i,a)$ is used to predict the possible degree of preference of i to a scenic spots. $F_j(a)$ is the score of user j for a attractions and $\overline{F_i}$ is used to score the average score of holder i. $\overline{F_j}$ is the average score of user j, and $\mathrm{sim}(i,j)$ is the similarity between user i and user j calculated by Pearson correlation coefficient:

$$U(i,a) = \overline{F_i} + \frac{\sum_j \mathrm{sim}(i,j) \times \left| F_j(a) - \overline{F_j} \right|}{\sum_j \mathrm{sim}(i,j)}. \tag{13}$$

In the process of recommendation, the recommendation module first calculates Pearson correlation coefficient, calculates the first 20 items of scoring data which are most similar to users, and then runs the subsequent recommendation algorithm. Finally, the user's predicted value of a certain scenic spot is obtained by weighted average of these scoring data and similarity, and five scenic spots with the highest predicted value are recommended to users for reference.

(5) Position Module. This module can use Bluetooth GPS receiver to receive satellite signals and select the local latitude

and longitude values to determine the location of the user. Finally, combined with the processing results of the recommendation module, the electronic map shows the location of each recommendation site and the location of the user.

(6) Information Module. The concept of user's active comment can be added in the system; through the transmission of information, users can express their personal ideas to other users around. Adding the function of transmitting information in this part, the user can not only transmit the new information but also transmit the received information to other users in the transmission range.

(7) Convergence Module. Since the system is designed in a mobile peer-to-peer environment, a mobile convergence function is derived from the concept of mobile social networks. Through this function, travelers can dynamically search for other users with the same goals and preferences. Through the transmission of information, travelers can share the requested information to the surrounding users and thus find travelers willing to act together.

(8) Privacy Module. One of the focuses of mobile social software is to explore the interaction between users, but not everyone is willing to interact with others, so this study adds privacy considerations. This module can provide users whether to allow all other users or only allow some friends to search their own location through the system; through the privacy settings, users can not be disturbed by other users to carry out system operations but also to observe whether there is a willingness to interact between users.

J2ME can be chosen as the development platform of the system, and Bluetooth technology is the basic wireless transmission technology commonly available in mobile terminals. Therefore, it is feasible to use Bluetooth technology as a transmission tool. In order to expand the scope of information transmission, WIFI wireless network can also be considered as a transmission medium, which can effectively solve the problem of short transmission distance and unstable signal of Bluetooth.

5. Conclusion

Advanced GPS devices enable people to record their location histories with GPS trajectories. The trajectory of users' mobile behavior means to a certain extent that a person's behavior and interests are related to their outdoor activities, so we can understand the users and their locations and their correlation according to these trajectories. This information enables accurate travel recommendations and helps people to understand a strange city efficiently and with high quality. By measuring the similarity of different user location histories, the similarity between users can be estimated and personalized friend recommendation can be realized. The user stereoscopic user portrait can be portrayed through the integration of user movement behavior trajectory and social information. This paper takes the trajectory data of tourism users' mobile behavior as the research object and constructs the tourism precise marketing model. In the process of obtaining the trajectory of user movement, the characteristics of mobile user behavior track data are taken into account. The sensitivity of various features in the trajectory analysis process is adjusted by weight. The structured feature vectors and popular scenic spots discovery methods of user's mobile behavior trajectory are fully studied by clustering and collaborative filtering techniques, which lay a foundation for constructing the application model of tourism precision marketing.

Acknowledgments

This research work was supported by the Ministry of Education Humanities and Social Sciences Planning Fund Project (No. 18YJAZH128) and Research Project of Harbin University of Commerce (No. 18XN022).

References

[1] Y. Yuan and M. Raubal, "Measuring similarity of mobile phone user trajectories–a Spatio-temporal Edit Distance method," *International Journal of Geographical Information Science*, vol. 28, no. 3, pp. 496–520, 2014.
[2] Z. Sun and X. (Jeff) Ban, "Vehicle classification using GPS data," *Transportation Research Part C: Emerging Technologies*, vol. 37, pp. 102–117, 2013.
[3] D. Wang, "Approaches for transportation mode detection on mobile devices," in *Proceedings of Seminar on Topics in Signal Processing*, pp. 77–82, 2014.
[4] S. Hong and A. Vonderohe, "Uncertainty and sensitivity assessments of GPS and GIS integrated applications for transportation," *Sensors*, vol. 14, no. 2, pp. 2683–2702, 2014.
[5] M. Lin and W.-J. Hsu, "Mining GPS data for mobility patterns: a survey," *Pervasive and Mobile Computing*, vol. 12, pp. 1–6, 2014.
[6] S. Khajehzadeh, H. Oppewal, and D. Tojib, "Mobile coupons: what to offer, to whom, and where?," *European Journal of Marketing*, vol. 49, no. 5-6, pp. 851–873, 2015.
[7] H. Li, "Review on state-of-the-art technologies and algorithms on recommendation system," in *Proceedings of the International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2016)*, p. 7, Wuhan Zhicheng Times Cultural Development Co., Xi'an, China, 2016.
[8] Htet Htet Hlaing, "Location-based recommender system for mobile devices on University campus," in *Proceedings of 2015 International Conference on Future Computational Technologies (ICFCT'2015); International Conference on Advances in Chemical, Biological & Environmental Engineering (ACBEE) and International Conference on Urban Planning, Transport and Construction Engineering (ICUPTCE'15)*, p. 7, Universal Researchers in Science and Technology; Universal Researchers in Science and Technology, Singapore, March 2015.

[9] W. Wörndl and B. Lamche, "User interaction with context-aware recommender systems on Smartphones," *icom*, vol. 14, no. 1, 2015.

[10] L. O. Colombo-Mendoza, R. Valencia-García, G. Alor-Hernández, and P. Bellavista, "Special issue on context-aware mobile recommender systems," *Pervasive and Mobile Computing*, vol. 38, pp. 444-445, 2017.

[11] L. O. Colombo-Mendoza, R. Valencia-García, A. Rodríguez-González, G. Alor-Hernández, and J. J. Samper-Zapater, "RecomMetz: a context-aware knowledge-based mobile recommender system for movie showtimes," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1202–1222, 2015.

[12] W.-S. Yang and S.-Y. H. iTravel, "A recommender system in mobile peer-to-peer environment," *Journal of Systems & Software*, vol. 86, no. 1, pp. 12–20, 2013.

[13] T. Pessemier, D. Simon, K. Vanhecke, B. Matté, E. Meyns, and L. Martens, *Context and Activity Recognition for Personalized Mobile Recommendations*, Springer, Berlin, Germany, 2014.

[14] J. Zeng, F. Li, Y. Li, J. Wen, and Y. Wu, "Exploring the influence of contexts for mobile recommendation," *International Journal of Web Services Research*, vol. 14, no. 4, pp. 33–49, 2017.

[15] A. Majid, L. Chen, G. Chen, H. T. Mirza, I. Hussain, and J. Woodward, "A context-aware personalized travel recommendation system based on geotagged social media data mining," *International Journal of Geographical Information Science*, vol. 27, no. 4, pp. 662–684, 2013.

[16] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, "Mobile recommender systems in tourism," *Journal of Network and Computer Applications*, vol. 39, pp. 319–333, 2014.

[17] S. K. Hui, J. J. Inman, Y. Huang, and J. Suher, "The effect of in-store travel distance on unplanned spending: applications to mobile promotion strategies," *Journal of Marketing*, vol. 77, no. 2, pp. 1–16, 2013.

[18] L. Liu, J. Xu, S. S. Liao, and H. Chen, "A real-time personalized route recommendation system for self-drive tourists based on vehicle to vehicle communication," *Expert Systems with Applications*, vol. 41, no. 7, pp. 3409–3417, 2014.

[19] J. P. Lucas, N. Luz, M. N. Moreno, R. Anacleto, A. Almeida Figueiredo, and C. Martins, "A hybrid recommendation approach for a tourism system," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3532–3550, 2013.

[20] Z. Bahramian, R. Ali Abbaspour, and C. Claramunt, "A CONTEXT-AWARE tourism recommender system based ON a spreading activation method," *ISPRS–International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4/W4, pp. 333–339, 2017.

[21] M. Nilashi, K. Bagherifard, M. Rahmani, and V. Rafe, "A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques," *Computers & Industrial Engineering*, vol. 109, pp. 357–368, 2017.

[22] I. Cenamor, T. de la Rosa, S. Núñez, and D. Borrajo, "Planning for tourism routes using social networks," *Expert Systems with Applications*, vol. 69, pp. 1–9, 2017.

[23] K. Meehan, T. Lunney, K. Curran, and A. McCaughey, "Aggregating social media data with temporal and environmental context for recommendation in a mobile tour guide system," *Journal of Hospitality and Tourism Technology*, vol. 7, no. 3, pp. 281–299, 2016.

[24] Z. Shi and A. B. Whinston, "Network structure and observational learning: evidence from a location-based social network," *Journal of Management Information Systems*, vol. 30, no. 2, pp. 185–212, 2014.

[25] Q. Lu, "Mobile e-commerce precision marketing model and strategy based on LBS," *E-Business Journal*, no. 4, pp. 20-21, 2014.

[26] P. J. Danaher, M. S. Smith, K. Ranasinghe, and T. S. Danaher, "Where, when, and how long: factors that influence the redemption of mobile phone coupons," *Journal of Marketing Research*, vol. 52, no. 5, pp. 710–725, 2015.

[27] N. M. Fong, Z. Fang, and X. Luo, "Geo-conquesting: competitive locational targeting of mobile promotions," *Journal of Marketing Research*, vol. 52, no. 5, pp. 726–735, 2015.

[28] S. A. Shad and E. Chen, "Precise location acquisition of mobility data using cell-id," *International Journal of Computer Science Issues*, vol. 9, no. 3, 2012.

Tourism Information Data Processing Method Based on Multi-Source Data Fusion

YaoGuang Li and HeChi Gan

College of Culture and Tourism, Jeonju University, Jeonju 55069, Jeollabub-do, Republic of Korea

Correspondence should be addressed to HeChi Gan; hechigan@163.com

Academic Editor: Mu Zhou

Urban social civilization and the quality of life of residents are gradually improved, and the development scale and trend of the leisure tourism industry have been growing. This paper constructs a multi-source data fusion model based on an ensemble learning algorithm, uses Ctrip 2020 open data set to train the model, and then obtains the tourism information data processing and prediction results. This paper takes the data of Ctrip as the training set and compares the trained model with the data of tunic and Feizhu. In this paper, sensor detection technology is used to analyze many famous scenic spots in China, including tourist type, gender, and location. The results show that tourism feature extraction results are consistent with data from trending flying bamboo, tunics, and other websites, according to the results of a multi-source fusion of tourism information. Among them, in the data of the first half of 2020, the prediction accuracy of the model after data processing is about 62%. Affected by the epidemic situation, the accuracy of the model is low. In the second half of the year, the prediction accuracy is 78%, which can be used to fuse tourism information in a short time. Therefore, the data show that the model has high learning ability and high trend prediction ability in tourism data processing, which can provide necessary information support for tourists.

1. Introduction

Tourism information processing technology usually uses a POI model related to specific human social activities to represent a group of points in arc tourism information. This paper analyzes the distribution and agglomeration characteristics of leisure tourism space and discusses how to coordinate the relationship with urban construction and development, which plays a vital role in the construction of livable cities, the development of tourism, and the sustainable development of cities.

At present, many studies refer to the urban landscape and tourism data. For example, Chen y divides the urban leisure space into four categories by analyzing the world's major parks [1]. Shang CF discusses and analyzes the research status of the concept of leisure tourism at home and abroad. As a new data source and research idea, it can further analyze the urban leisure tourism space [2]. Oberoi analyzes the distribution characteristics of tourism information services in Guangzhou by using the nearest neighbor distance method and other spatial analysis methods [3]. In terms of coupling analysis of tourism information data and scenic spot data, Bernardi PD uses landscape pattern index, gravity model, gravity model, and coupling analysis to analyze the urban spatial distribution structure of Anhui Province Based on DMSP/OLS image, statistical data, and tourism information data [4]. PON W C selects tourism information data and NPP remote sensing data for kernel density analysis and further analyzes the characteristics of the urban spatial structure of Wuhan City [5].

In data fusion, a new artificial intelligence technology proposed by poetry y a proposes that multi-source data fusion is that a central server coordinates multiple clients to complete a learning task without disclosing data [6]. Ramadhan GR proposes a user-level differential privacy training algorithm, which effectively reduces the possibility of recovering personal information from the transmission model by adding privacy protection to the aggregation algorithm [7]. On the other hand, Huyan w proposed a differential privacy hybrid model, which partitions users by their trust preferences, to reduce the size of the required user base [8]. Oezturan m combines gradient selection and secret sharing algorithm, which greatly improves the communication efficiency while ensuring user privacy and data security [9]. In terms of resource optimization in multi-source data fusion, Liu P considered multi-source data fusion through the wireless network and proposed the problem of optimizing energy consumption and global multi-source data fusion time [10–11]. The above research in the data source processing is not fully integrated, so the tourism data prediction cannot achieve accurate prediction.

This paper constructs a multi-source data fusion model based on an ensemble learning algorithm, uses Ctrip 2020 open data set to train the model, and then obtains the tourism information data processing and prediction results [11–12]. This white paper uses Ctrip data as a training set and compares the trained model with tunic and flying bamboo data. This treatise mainly analyzes some of China's famous scenic spots, such as tourist type, gender, and location. According to the results of the multi-source fusion of tourism information, the results of tourism feature extraction are consistent with the data of Feizhu, tunic, and other websites in the trend.

2. Tourism Data Processing and Multi-Source Data Fusion

2.1. Distributed Tourism Data Processing Method. In the information age, massive tourism data will bring many problems to the centralized data processing mode with the cloud computing model as the core [13]. First, the processing mode of uploading all data to the cloud will not only cause low efficiency but also cause additional bandwidth overhead, at the same time, the network delay will also increase [14]. Second, due to the improvement of users' privacy awareness, the data of edge devices are likely to leak in the upload link, and the security of personal privacy cannot be guaranteed [15]. The distributed data processing mode can effectively solve the delay and efficiency problems of traditional cloud computing [16]. At the same time, aiming at the problem of "data island", Google proposed a new concept of "multi-source data fusion" for the first time [17]. The model is individually trained on multiple edge devices using training samples, and sharing of multi-source information is achieved by aggregating model parameters without disclosing user privacy. In addition, due to the diversity of edge devices, the data collected by the devices exhibits diversity in annotations, semantics, and existing formats [18]. For example, for the description of the same object, there are text class, picture class, video class, and audio class data, and multi-modal data widely exists [19]. Different modal data can describe the target object from many aspects [20]. By eliminating redundant data and fusing various data sources for correlation and supplementary analysis, more valuable new information can emerge from the data, to achieve the effect of $1 + 1 > 2$. Multimedia data collected from the Internet and mobile devices are typical unstructured data, which is significantly different from the traditional structured data format [21]. Therefore, the processing of multi-source heterogeneous data collected by different edge devices becomes an urgent problem in big data research [22]. In the traditional multi-source heterogeneous data fusion algorithm, data centralized processing has the risk of data privacy leakage in practical application [23]. Therefore, there are still many problems in multi-source heterogeneous data processing without disclosing user privacy: first, due to enterprise competition and user privacy protection awareness, data exchange is blocked for a long time, and information sharing cannot be realized, so the value of heterogeneous data cannot be fully exploited. Secondly, the neural network is used to process the data, and the model designed according to the data cannot be changed once it is determined [24]. However, in edge computing, there are differences in data structure and the number of types collected by edge devices. If we design a neural network for the data of each network edge device, the workload is huge [25–26]. At the same time, the model can only be applied to a single node or the edge device with the same data characteristics as the node, the universality is not high, and it can not give full play to the value of other heterogeneous data in the Internet of things [26]. To solve the problem of multi-source heterogeneous data fusion without disclosing user privacy in edge computing, this paper proposes a multi-source heterogeneous data fusion algorithm based on multi-source data fusion [27–30]. Starting from the data structure characteristics collected by edge devices, combined with tensor tucker decomposition theory, this paper studies the adaptive processing of multi-source heterogeneous data model on different edge devices, and solves the single adaptability problem caused by the disunity of the heterogeneous data model in multi-source data fusion.

2.2. Multi-Source Data Fusion System Model. In this paper, we aim at the problem of heterogeneous data fusion without data exchange, consider the introduction of multi-source data fusion to edge computing, and learn the potential characteristics of multi-user without invading user privacy. The basic architecture of its intelligent sensor system is shown in Figure 1.

In this framework, the system is composed of an edge node, Internet of things, and cloud server. The edge node connects with the cloud server through the Internet of things (such as gateway and router). Multi-source data fusion is a distributed learning framework, in which the original data is collected and stored on multiple edge nodes, and the model training is performed at the nodes, and then the model is gradually optimized through the interaction between node n and cloud server h. The formula is as follows

$$\mathrm{H} = \frac{\sum_{j=1}^{k}\sum_{h=1}^{k}\sum_{t=1}^{n_j}\sum_{r=1}^{n_h}\left|y_{ij} - y_{hr}\right|}{2n^2u} \quad (1)$$

Based on the above framework, multi-source data fusion can use local data from multiple independent edge nodes sharing model, and use model transmission instead of data transmission to avoid the risk of user privacy leakage in the process of transmitting the collected original data from multiple edge nodes to the cloud server.

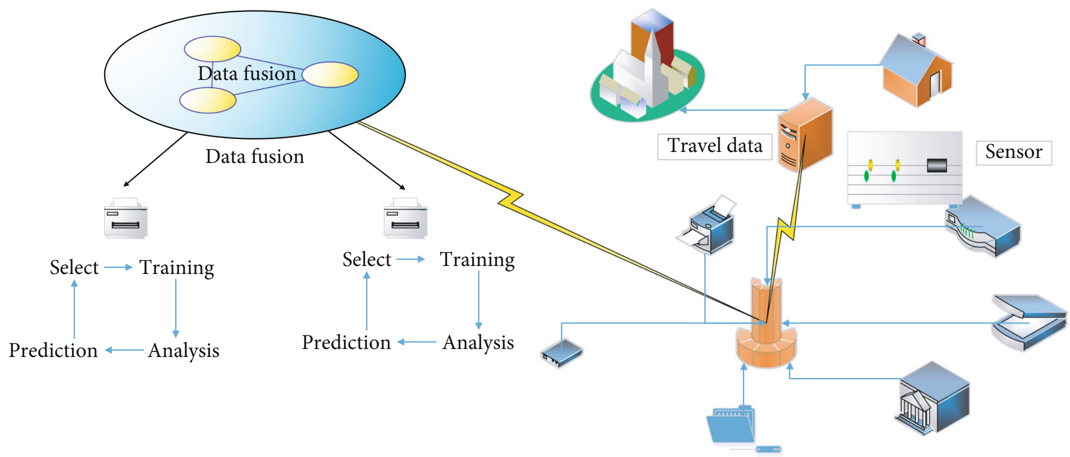


Figure 1: Multivariate data fusion framework diagram.

The multi-source heterogeneous data fusion algorithm proposed in this paper is mainly divided into feature extraction modules, feature fusion modules, and feature decision modules. The feature extraction module is composed of feature extraction sub-networks corresponding to various heterogeneous data. In the initialization stage, the central control node randomly initializes the network parameters of the feature extraction module, feature fusion module, and feature decision module, and sends them to the edge node T, where the fusion parameter is C.

$$\sum_T = \mathrm{C}(\max(\sigma_i, \upsilon, 0))$$
$$\mathrm{E}_j = \frac{1/2u_j \sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\left|y_{ji}-y_{jr}\right|}{n_j^2} \tag{2}$$

In the model training stage, after receiving the model from the central control node, the edge node selects the corresponding feature extraction module according to the data set structure of the local node and uses the local data set to train the feature extraction module, feature fusion module, and feature decision module. The termination condition of a new round of edge node training is that the number of local node training rounds exceeds the given number of training rounds. After training, the training model is returned to the central control node for model aggregation. The first practice test is to use the average aggregation algorithm for the function fusion module and the function determination module. For the feature extraction module, the average sub-module is extracted according to the corresponding feature extraction submodule to ensure that the same mode data extract features are similar. Finally, the updated model is redistributed to the edge nodes for a new round of training. This paper assumes that the heterogeneous data to be processed are audio, visual, and textual data. In the feature extraction module, according to the features of different modes, different feature extraction sub-networks are used to extract the features of audio, visual, and text information. Audio and visual feature sub network: for audio information and visual information, covered acoustic analysis framework and face facial expression analysis framework is used to sample and extract features from MoSi data sets (sampling frequency is 100 Hz and 30 Hz, respectively). In this section, it is assumed that the heterogeneous data features to be processed are audio data features

$$\mathrm{V}_{jh} = \frac{\sum_{z=1}^{hj}\sum_{r=1}^{n_h}\left|y_{ji}-y_{hr}\right|}{n_j n_h\left(u_j+u_h\right)} \tag{3}$$

Visual data features:

$$\mathrm{V}_{nb} = \sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\left(p_j s_h + p_h s_j\right) D_{jh} \tag{4}$$

Text data features:

$$\mathrm{V}_0 \int_0^{\infty} dF_j(y) \int_0^{y} (y-x)\, dF_h(x) \tag{5}$$

After the feature fusion module, the feature output is v0. Next, taking the above assumption as the basic condition, the basic principle of the proposed heterogeneous data fusion algorithm based on tucker decomposition is described. The first mock exam module introduces a high tensor W with heterogeneous data feature space, and each mode of the tensor corresponds to spatial mapping of heterogeneous data characteristics. Therefore, when fusing the features of each heterogeneous data, the high-order tensor w can not only introduce the features of other heterogeneous data modes for correction but also memorize the ongoing heterogeneous data modal features. When the features of the heterogeneous data to be processed are P and S, the memory unit W is a third-order tensor, and the three dimensions of the tensor corresponding to the feature spaces of the three heterogeneous data features. In the heterogeneous data feature fusion proposed in this section, the memory unit with the heterogeneous data feature can be obtained by modular multiplication of the heterogeneous data feature and the feature space corresponding to the memory unit, and further feature fusion operation can be carried out. The fusion operation is mainly divided into three stages. The memory unit W is modulo multiplied with the heterogeneous data feature AZ along the first order to obtain a new memory unit with AZ feature.

$$W_{jh} \frac{\mathrm{AZ}_{jh} - P_{jh}}{\mathrm{AZ}_{jh} + P_{jh}} \tag{6}$$

Secondly, memory unit W is modulo multiplied with heterogeneous data features along the second-order to obtain memory unit w with AZ features. Finally, the memory unit W is modularly multiplied with the heterogeneous data features along with the third order, and finally, the fusion tensor with the three features is obtained Z。 The specific process can be expressed as follows:

$$E_{\mathrm{AZ}} = \sum_{j=1}^{k} G_{jj} p_j s_j \tag{7}$$

Among them:

$$P_j = \int_0^{\infty} dF_h(y) \int_0^{y} (y - x) dF_j(y) \tag{8}$$

For the fused data, this chapter uses the traditional full connection layer to make decisions based on the global characteristics, including the prediction of regression model and the probability prediction of a classification model. In this module, the L1 norm loss function is used to measure the error between the target value and the predicted value. Its et expression is

$$\mathrm{E}_T \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(p_j s_h + p_h s_j \right) D_{jh} \left(1 - L1_{jh} \right) \tag{9}$$

The expression of NL is:

$$\mathrm{NL}_t = \tanh\left(w_c, x_t + u_c(r_t \Theta h_{t-1}) + b_c\right) \tag{10}$$

Then, it is assumed that n edge nodes participate in the training of the shared model, and all edge nodes collect m kinds of heterogeneous data. In the initialization stage, according to the collected m kinds of heterogeneous data, the cloud designs the corresponding feature extraction module F, feature fusion module I and feature decision module C. Then the shared model G can be expressed as:

$$\mathrm{G} = \mathrm{F} * \mathrm{f/C} = \frac{\sqrt{1/M \sum_{i=1}^{n} (I_{it} - I_{it})^2}}{I_{it}} \tag{11}$$

Where $*$ represents the model splicing operation. Specifically, in the feature extraction module F, the corresponding feature extraction sub-network MF is designed from kinds of heterogeneous data, which can be expressed as:

$$1\mathrm{n}\left(\frac{MF_{it}}{MF_{it} - 1}\right) = \alpha + \beta\, 1\mathrm{n} MF_{it} - 1 + \varphi X_{it} - \tau_t \tag{12}$$

Where x is the feature extraction sub-network of the I heterogeneous data. In feature fusion module I, a high-order tensor with spatial dimension characteristics of heterogeneous data is constructed. After training, the parameters of the tensor expanded along the I module can reflect the spatial dimension characteristics of the I heterogeneous data. In the feature decision-making module C, through the training of the fused heterogeneous data features, the potential relationship between heterogeneous data is mined in a deeper level, and the feature expression of the model in multi-source heterogeneous data is improved. Due to the diversity of data collected by edge devices, the data processed by each edge node is different. Therefore, in the premise of not disclosing user privacy, multi-source heterogeneous data fusion has the problem of insufficient adaptability. The size of the tensor after fusion is consistent with that of memory unit W. Therefore, when the factor matrix satisfies the square matrix constraint, there is an identity relationship between the core tensor and the original tensor in the spatial dimension. Using this feature, in the initialization phase, the global is further set, and the feature extraction sub-network fi feature graph is defined

$$1\mathrm{n}\left(\frac{FI_{it}}{FI_{it} - 1}\right) = \alpha + \beta\, 1\mathrm{n} FI_{it} - 1 + v_i + \mathfrak{J}_t \tag{13}$$

Thus, the problem of heterogeneous data fusion caused by the uncertainty of heterogeneous data in edge computing is solved. In the model training stage, n edge nodes participate in the training according to the heterogeneous data types they have

$$N_t z_t \Theta h_{t-1} + (1 - z_t) \Theta h_t \tag{14}$$

The corresponding feature extraction sub-network FZ is adaptively selected for training. In the feature fusion module, because the feature graph of the feature extraction sub-network r is set to FR in the global initialization definition stage, the tensor size after heterogeneous data fusion is constrained to a fixed value

$$\begin{aligned} \mathrm{R} &= \frac{\mathrm{FR}}{1 - \mathrm{FZ}} \\ \mathrm{FR} &= -\frac{1}{T} 1\mathrm{n}(1 + \beta) \end{aligned} \tag{15}$$

Suppose that all the heterogeneous data types to be processed are N, X, and H, respectively, and the three dimensions of the corresponding memory unit w correspond to the feature spaces of the three heterogeneous data. The heterogeneous data types collected by node 1 and node n are different. In the model training stage, the features are obtained, respectively

$$f(\mathrm{N})=\frac{1}{Nh}\sum_{i=1}^{N}k\left(\frac{X_i-x}{h}\right)$$
$$\mathrm{W}=k(\mathrm{H})=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \tag{16}$$

According to the number of heterogeneous data types on the node, the feature fusion stage is divided into two parts: first, memory unit W is modulo multiplied with AF feature along the first order to get a new memory unit w with AF feature. Secondly, the memory unit w multiplies the VF features along the second-order to obtain the fusion tensor with the above two heterogeneous data features Z。 The process can be expressed as follows:

$$\mathrm{W1}_j\sum_i c_{ij}u_{(\mathrm{AF}|\mathrm{AV})} \tag{17}$$

W1 is the fusion of VF features based on AF features. In this process, the model first uses the memory unit to memorize the AF features, and obtains the model with AF features, and takes this as a priori condition for the fusion of VF features. Thus, in the process of model training, the memory unit can not only learn the spatial dimension features of each heterogeneous data but also capture the potential relationship between different heterogeneous data. The training mechanism on node n is similar to that on node 1. The above process can be expressed as follows:

$$k_{t1}[\mathrm{N}]=\sum_j\cos\left(w_i^1,w_j^2\right)/\mathrm{K} \tag{18}$$

Where n is the node K in the j-round global iteration, using the heterogeneous data collected locally, the learning rate is η The gradient descent algorithm is used to get the local model. In the model aggregation stage, because each edge node uses the adaptive selection mechanism of feature extractor to train the feature extraction module, it is necessary to merge the feature extraction sub-networks selected and trained by each edge node, and then use the average aggregation algorithm to get the shared model with global heterogeneous data features

$$u_{(j|i)}=w/\mathrm{N}*_{ij}A_i \tag{19}$$

Where a is a shared model with global characteristics obtained by aggregating the local models on n edge nodes through the joint average algorithm. J is the number of all heterogeneous samples and the total number of all heterogeneous samples on the edge node.

3. Multivariate Fusion Data Model

3.1. Methods. This paper constructs a multi-source data fusion model based on an ensemble learning algorithm, uses Ctrip 2020 open data set to train the model, and then obtains the tourism information data processing and prediction results. This paper takes the data of Ctrip as the training set and compares the trained model with the data of tunic and Feizhu. In this paper, sensor detection technology is used to analyze many famous scenic spots in China, including tourist type (youth, adult), gender, and location.

3.2. Data Acquisition and Processing. Through a comprehensive comparison of other online tourism service platforms such as tunics and Hitake, this paper has the highest number of online comments on http://Trip.com/'s Saikei Wetland, and Ctrip's online comments are rich in content and have interference factors. Few, positive and negative comments, and most of them are the true feelings of tourists. Therefore, it is common and effective to choose Ctrip online game reviews as your data source. Through Python technology, this paper grabs the online comments of tourists in Xixi Wetland scenic spot on Ctrip from May 1, 2020, to May 1, 2021, removes the invalid online comments that deviate from the research theme, and collates 3052 effective online comments, and makes specific research and analysis of the collected online comments. In addition, in cooperation with Xixi Wetland Tourism Development, the research group carried out an offline questionnaire survey on tourists' comprehensive satisfaction of Xixi National Wetland Park during the long holidays of May Day and national day in 2020, and obtained 102 valid questionnaires, which fully investigated the customer source market and tourists' satisfaction of Xixi National Wetland Park, as a supplementary means of network big data survey, this paper puts forward more relevant and targeted suggestions for the future development of Xixi National Wetland Park.

According to the text data of tourists' online reviews and the comprehensive satisfaction offline questionnaire, the research results of Xixi Wetland Park's tourism image perception are relatively consistent. When analyzing the influencing factors of tourism perception experience, there are many overlapping problems, that is, the important concerns when improving the tourism brand image. For example, given the insufficient setting of public toilets, rest chairs, and free drinking water points, drinking water, and lunch supply points should be added accordingly; For the lack of a single tour line of battery car and battery boat in the park and long waiting time for tourists to enter and leave the park, we should provide parking guidance service; Continuously improve service awareness and management level, and focus on ticket and catering prices, safety management of cruise terminal, queuing guidance, increase the number of sanitation personnel, increase recreational equipment, strengthen the integration of tourism formats and products, and improve the effective supply of products.

Figure 2: Word frequency statistics for comment texts of famous scenic spots.

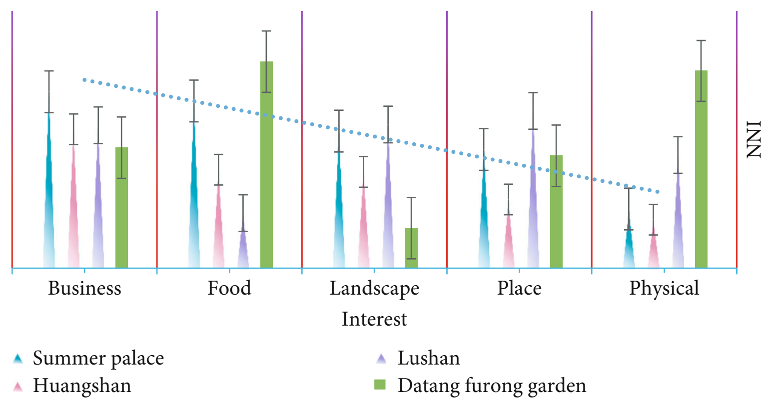


Figure 3: Overall NNI of leisure and tourism space.

4. Results and Discussion

4.1. Multi-Source Information Fusion Tourism Information Processing Method. Use Python to perform word frequency statistics on the comment texts of scenic spots and famous scenic spots in China, and extract the top 50 high-frequency feature words, as shown in Figure 2. Key terms such as "travel" and "attractions" are ranked high, indicating that projects that attract tourists to Tiger Beach are primarily special venues and ropeways across the ocean. The most attractive to tourists is the third-ranked "performance", such as "dolphin show". In addition, the characteristic words such as "getting tickets", "buying tickets" and "tickets" indicate that tourists pay great attention to the purchase of tickets, and the innovation and improvement of tourist ticket purchase methods in scenic spots should be highly valued. Compared with China's famous scenic spots, "shows" ranked first, indicating that watching performances are the main activity of tourists in Shengya.

As shown in Figure 3, the overall NNI (neural network intelligence) of leisure tourism space is less than 1, and the NNI of different types of space is also less than 1, among which catering services and commercial services are lower, indicating that these two types are more inclined to cluster development, and the cluster scale is larger. Catering and business services have a strong economic effect on agglomeration and are affected by factors such as economic level, population density and convenience of transportation. The NNI of sports leisure and cultural entertainment services were 0.35 and 0.52, respectively, which also showed a general trend of agglomeration. The NNI of scenic spot service is the largest, which is 0.46. Because of the influence of natural ecological conditions and urban open space, the agglomeration degree of all scenic spots is relatively low.

As shown in Figure 4, it has passed the significance test at the level of 5%, and the influence coefficient is positive, indicating that the innovation and technological progress of the tourism industry will play a promoting role in the growth of local tourism consumption without considering other factors. As the main platform of scientific and technological innovation in the industry, tourism colleges and research institutes, represented by coastal areas, have formed a certain scale. A large number of tourism colleges and universities have formed mature curriculum and perfect talent training systems, with rich research results, which are applied to the development and design of tourism creative products and tourism service supply management.

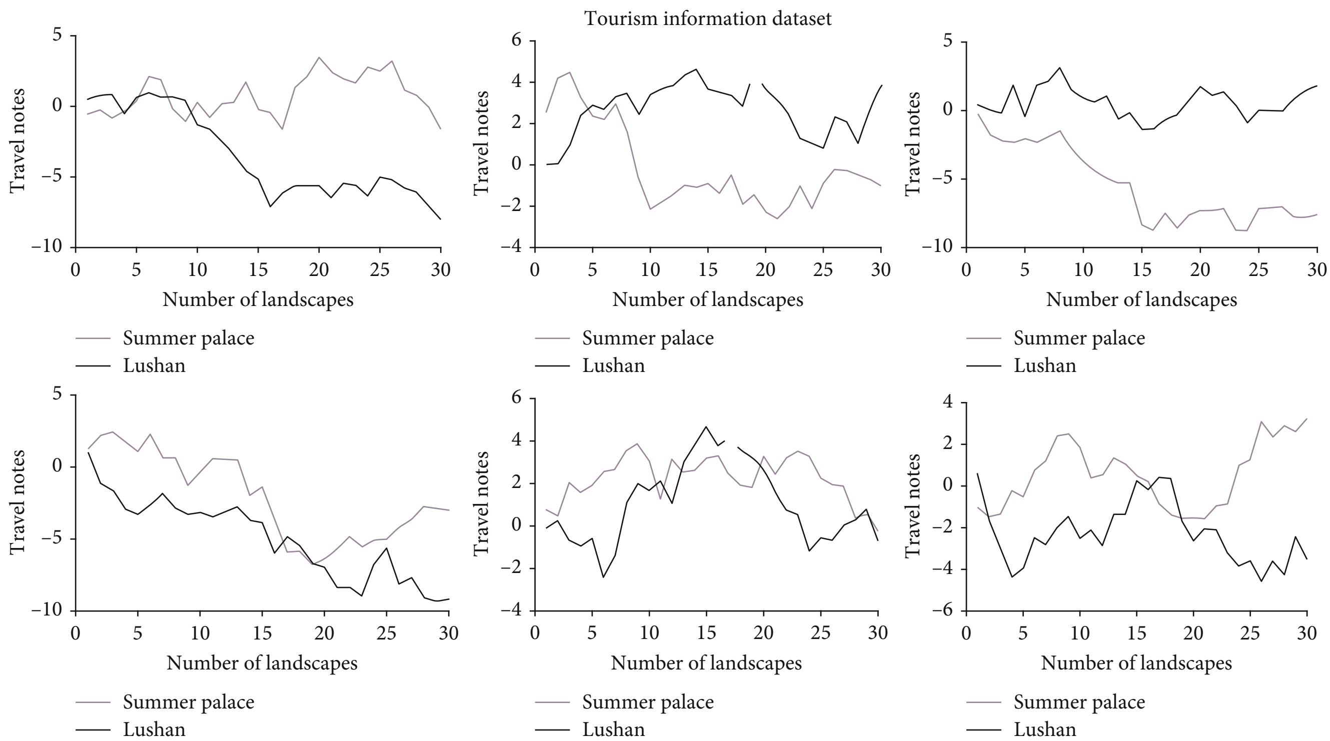


Figure 4: Tourism Information Dataset.

Under the special evaluation content, the target layer, criterion layer, and index layer of urban physical examination are established, and the index system reflecting the urban natural background and the operational signs of tourism information is constructed, as shown in Table 1. At the same time, according to the national standards of relevant indicators, the vertical establishment of planning expectation value and the horizontal comparison of the level of cities at the same level, the standard value and reference interval is determined, and the evaluation and calculation are carried out according to the positive and negative direction of the indicators. According to the data sources and evaluation contents of each index, the calculation methods can be divided into two categories: tourism information spatial analysis and statistical data analysis. For the smaller-scale spatial data of urban districts and counties and below, the methods of tourism information computing geometry, buffer analysis and overlay analysis can be used, such as calculating the coverage of public service facilities through buffer analysis and overlay analysis. Using the statistical data, we can summarize the static data of the basic elements of the city, calculate the average value, per capita value, proportion, number statistics, and coverage analysis.

As shown in Figure 5, the distribution and total amount of tourism information are always increasing. To analyze the change direction and trend of the center of gravity of urban space and leisure tourism space more intuitively, the migration trajectory of the center of gravity of urban space is drawn by using tourism information. It shows that the overall development of urban space and leisure tourism space is first to the northwest, then to the southwest, and the spatial growth is mainly in the southwest cities.

Table 1: The Index System of Tourism Information Operation Signs.

Item	Summer palace	Huangshan	Lushan	Datang Furong garden
Business	4.8	3.78	3.89	3.28
Food	4.55	2.68	1.5	5.62
Landscape	3.73	2.62	3.91	1.09
Place of interest	3.23	1.87	4.28	3.06
Physical education	1.61	1.32	3.08	5.39

Among the famous scenic spots in China, there are "dinosaur museum", "Ocean Museum" and "undersea tunnel". The reason why the "dinosaur museum" ranks high is that it is far away from the other four museums, which is also one of the important reasons why "walking" ranks high. As shown in Figure 6, the second-highest frequency word in the two scenic spots is "children", which indicates that the two scenic spots have a strong attraction for children and are the choice of many families for parent-child travel. In addition, "a lot of people", which ranks third among the famous scenic spots in China, ranks 50th among the characteristic words of tiger beach. The main reason is that Tiger Beach covers a large area and tourists are more evacuated. China's famous scenic spots are more attractive to tourists.

As shown in Table 2, tourists have a higher perception of tourism attractions and experience of tiger beach and Shengya, but a lower perception of tourism environment and facilities and services, and there are obvious differences

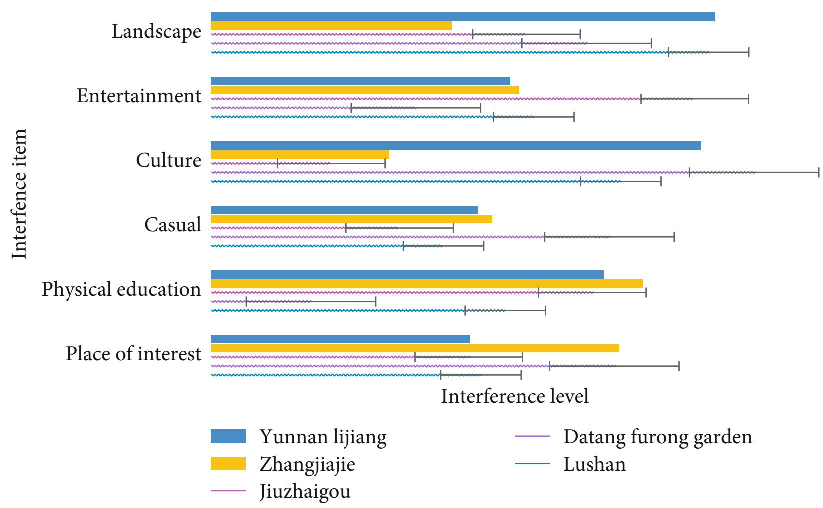


Figure 5: Distribution and Total Change Trend of Tourism Information.

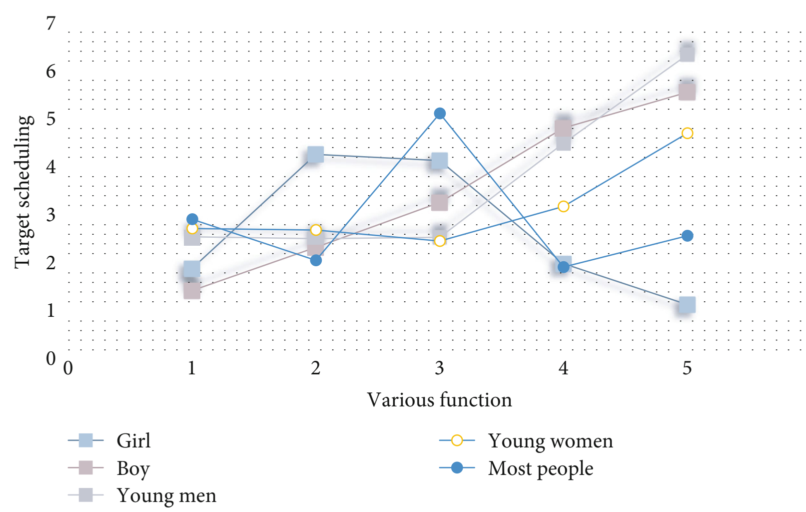


Figure 6: Projects that attract tourists among China's famous attractions.

Table 2: Perception of tourist attractions and tourist experience.

Item	Landscape	Place of interest	Physical education	Casual
Summer palace	0.8	0.12	1.62	0.86
Huangshan	2.06	3.46	2.95	3.87
Lushan	5.23	3.86	3.47	3.19
Datang Furong	4.1	5.08	4.55	2.36
Jiuzhaigou	1.64	1.76	3.54	4.43
Zhangjiajie	4.66	2.31	2.96	5.06

Table 3: Tourists' cognitive evaluation and emotions.

Item	Datang Furong garden	Jiuzhaigou	Zhangjiajie	Yunnan Lijiang
Place of interest	4.96	3.17	5.02	3.18
Physical education	1.23	4.69	5.31	4.83
Casual	4.9	2.32	3.46	3.28
Culture	6.68	1.48	2.19	6.02
Entertainment	2.52	5.95	3.79	3.68

between the two scenic spots. In the overall perception of scenic spots, tiger beach and Shengya have little difference. The two scenic spots belong to marine theme scenic spots, in which leisure and entertainment are all around the ocean, which is a romantic ocean trip for tourists; Children are attracted by all kinds of interesting marine creatures and performances in the scenic area, which is also a good choice for parent-child travel. Among the tourist attractions, tiger beach and Shengya are quite different, which indicates that tourists pay more attention to the tourist attractions, and

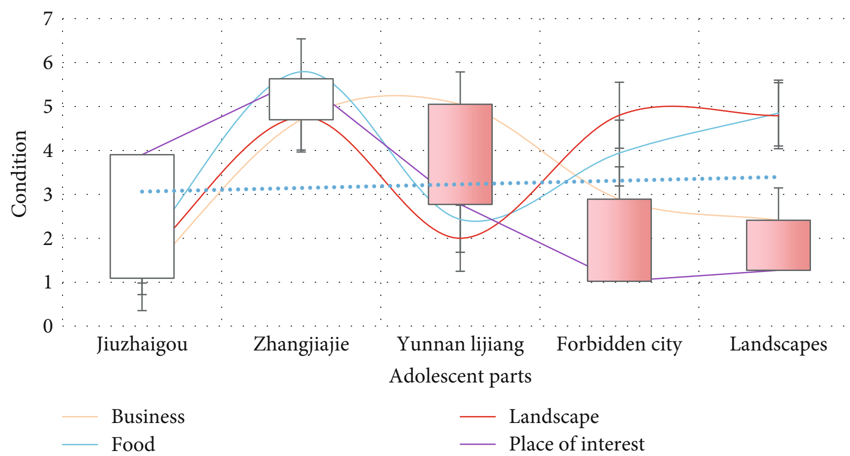


Figure 7: The overall trend of tourism information.

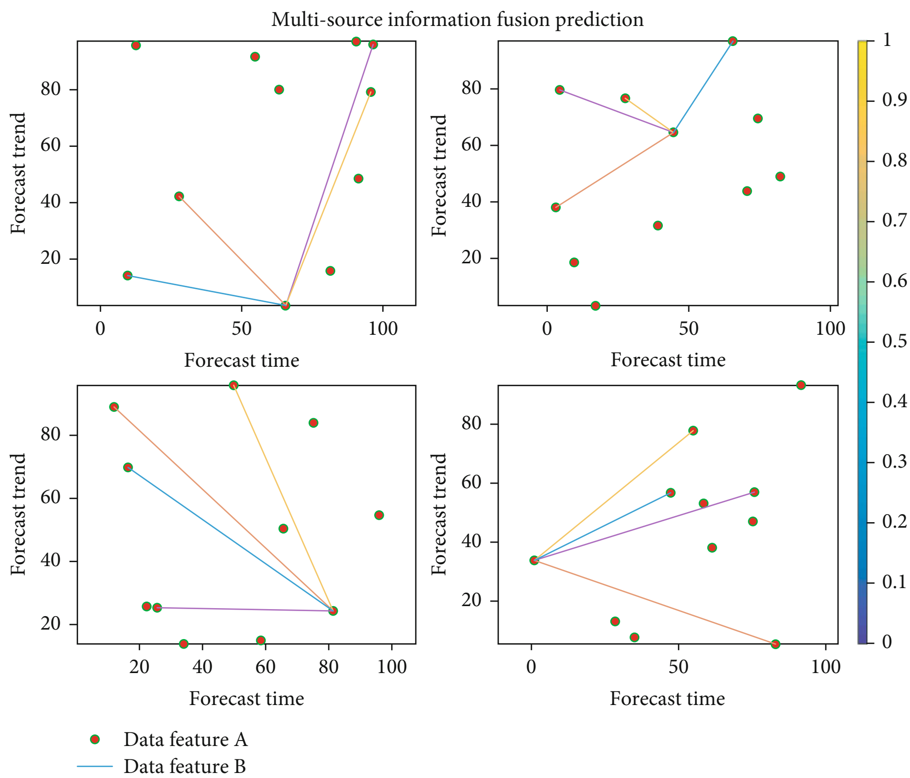


Figure 8: Multi-source information fusion prediction.

Shengya is more attractive. Tourists have a higher perception of Shengya theme venues and special performances.

As shown in Table 3, among the famous scenic spots in China, tourists' cognitive evaluation is more positive and there are not too many negative emotions. This shows that tourists are more satisfied with the service of famous scenic spots in China. Tourists' perception of the tourism environment and facilities of tiger beach and Shengya is weak. In the level of tourism environment perception factors, there are obvious differences in social environment perception. Among them, China's famous scenic spots are located in urban areas and are very close to Xinghai Square, which makes "Xinghai Square" stand out. Dalian Xinghai Square itself is one of the scenic spots that Dalian must visit. Tourists will take two scenic spots as a one-day tour plan, while Tiger Beach Ocean Park is relatively remote, this is also one of the important reasons why tourists sometimes choose Shengya when they choose two marine theme scenic spots.

As shown in Figure 7, through the coupling analysis of tourism information data and scenic spot data, it can be found that the overall tourism information is gradually decreasing from the central urban area to the surrounding areas, and there is a high coupling between the two data, and the areas showing coupling differences can also reflect the buildings with the same nature, such as stations, airports, industrial development zones, and other large areas. There is

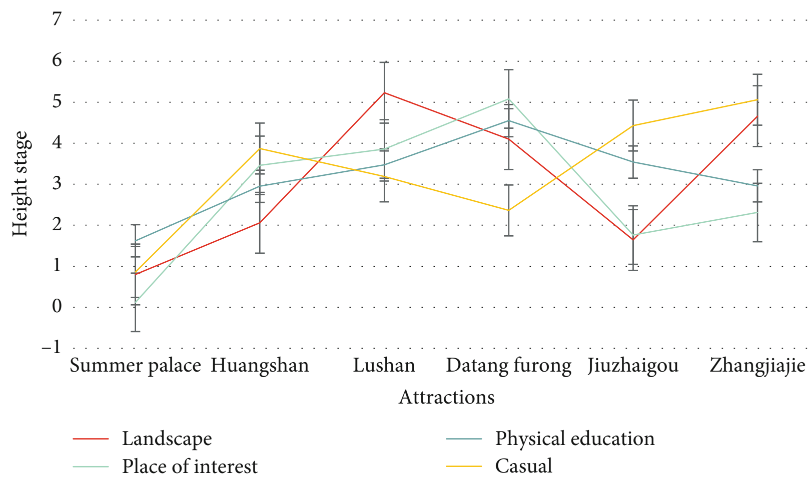


FIGURE 9: The center of gravity of urban space and leisure tourism space.

a wide range of low low coupling around the central city because the economic construction of suburban towns is relatively backward, so the value of tourism information and various leisure tourism service facilities are low, which is in line with the reality.

The center of gravity of urban space and leisure tourism space is shown in Figure 8. This paper finds that with the increase of human activity intensity and urban construction, the area of tourism information increases year by year, and the direction of gravity shift is from southwest to northwest. The results show that the multi-method comprehensive analysis based on big data can better describe the spatial distribution attributes and spatiotemporal evolution characteristics of the region. For example, this paper objectively and comprehensively reveals the agglomeration characteristics of leisure tourism space and provides indispensable reference information for land space planning. But at the same time, there are some limitations. The analysis of the characteristics of regional spatial agglomeration in this paper is more about the results, and the analysis of its driving factors or influencing factors is still lacking.

As shown in Figure 9, according to the results of the multi-source fusion of tourism information in this model, the results of tourism feature extraction are consistent with the data of Feizhu, tunic, and other websites in the trend. Among them, in the data of the first half of 2020, the prediction accuracy of the model after data processing is about 62%. Affected by the epidemic situation, the accuracy of the model is low. In the second half of the year, the prediction accuracy is 78%, which can be used to fuse tourism information in a short time. Therefore, the data show that the model has high learning ability and high trend prediction ability in tourism data processing, which can provide necessary information support for tourists.

4.2. Discussion. Given the target redundancy in the reconstruction results of specific scenic spots, this paper constructs a multi-source data fusion model based on an ensemble learning algorithm. In the field of 3D reconstruction, the irrelevant objects in a specific scene are eliminated by data fusion to achieve 3D scene reconstruction! Firstly, the lightweight algorithm is used to extract and match the features of different types of feature points, and the point clouds at different times are fused to complete the reproduction of the point cloud map. Then, for the irrelevant targets that may exist in the constructed point cloud map, with the help of multi-source sensor data and deep learning application technology in the field of computer vision, the target detection and elimination are carried out in three-dimensional space. For the two different processes of point cloud map modeling and target detection, the point cloud registration method is used to fuse them, and finally, the scene reappearance in the scenic spot environment is completed! The experimental results show that the method based on multi-source data fusion can effectively combine the two processes of 3D modeling and target detection, and complete the construction of a point cloud map without redundant targets in scenic spots.

By comparing the online comment text data of tourists and the offline comprehensive satisfaction survey questionnaire, it is found that the survey results of the tourism image perception of West Wetland Park are relatively consistent. Based on multi-source data integration of leisure travel space based on tourist information data and tourist attraction data, it is found that the leisure travel space in the central city area is distributed in a clear central multipoint layout, and the overall leisure travel is there. In the urban area, there is a trend of high agglomeration of scenic spots, which shows that scenic spots are the most concentrated areas of leisure tourism, while the agglomeration effect of peripheral counties is weak. In addition, the overall distribution pattern of leisure tourism space shows obvious characteristics of denseness and sparseness and a trend from northwest to southeast.

5. Conclusions

The multi-source fusion of tourism information of this model shows that the results of this model are consistent with the data of Feizhu, tunic, and other websites in the trend. Among them, in the data of the first half of 2020, the prediction accuracy of the model after data processing is about 62%. Affected by the epidemic situation, the accuracy of the model is low. In the second half of the year, the prediction accuracy is 78%, which can be used to fuse tourism information in a short time. Therefore, the data show that the model has high learning ability and high trend prediction ability in tourism data processing, which can provide necessary information support for tourists. Based on the online comments on Ctrip, combined with the survey results of tourists' comprehensive satisfaction in the same period, the online tourism big data and offline questionnaire verify each other, comprehensively analyze the tourism image perception of Xixi National Wetland Park, and innovate the means and methods of tourism image perception research, However, there are still some deficiencies in the number and period of tourists' online comments. In the follow-up research, it is necessary to further increase the number of online tourism big data and offline questionnaire survey samples, and enrich the correlation analysis of tourists' attributes, temporal and spatial behavior characteristics and tourism image perception, try to reduce the error caused by the difference of online and offline survey samples.

References

[1] Y. Chen, "A survey on industrial information integration 2016-2019," *Journal of Industrial Integration and Management*, vol. 5, no. 4, pp. 56–59, 2020.

[2] C. F. Shang, Y. F. Wang, and J. L. Du, "Information integration for motor generation," *Current Opinion in Physiology*, vol. 8, no. 4, pp. 116–120, 2019.

[3] A. Oberoi, M. Arora, and V. K. Garg, "A novel approach for dynamic information integration," *International Journal of Reasoning-based Intelligent Systems*, vol. 13, no. 2, pp. 76-77, 2021.

[4] P. de Bernardi, A. Bertello, and R. Shams, "Logics hindering digital transformation in cultural heritage strategic management: an exploratory case study," *Tourism Analysis*, vol. 24, no. 3, pp. 315–327, 2019.

[5] W. C. Poon and K. Y. Koay, "Hong Kong protests and tourism: Modelling tourist trust on revisit intention," *Journal of Vacation Marketing*, vol. 4, no. 4, pp. 1356–1359, 2021.

[6] Y. A. Poetra, "Upaya Pemerintah Dalam Mengkomunikasikan Tradisi Malamang Menjadi Objek Pariwisata Budaya Di Kabupaten Padang Pariaman," *Jurnal Pustaka Budaya*, vol. 5, no. 2, pp. 52–61, 2018.

[7] G. R. Ramadhan and I. Buchori, "Strategi Integrasi Sistem Transportasi Umum Dalam Menunjang Pariwisata Kota Yogyakarta," *Jurnal Pengembangan Kota*, vol. 6, no. 1, pp. 84-85, 2018.

[8] W. Huyan and J. Li, "Research on rural tourism service intellectualization based on neural network algorithm and optimal classification decision function," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 4, pp. 1–21, 2021.

[9] M. Özturan, M. Mutlutürk, B. Çeken, and B. Sarı, "Evaluating the information systems integration maturity level of travel agencies," *Information Technology & Tourism*, vol. 21, no. 2, pp. 237–257, 2019.

[10] P. Liu, H. Zhang, J. Zhang, Y. Sun, and M. Qiu, "Spatial-temporal response patterns of tourist flow under impulse pre-trip information search: from online to arrival," *Tourism Management*, vol. 73, no. 8, pp. 105–114, 2019.

[11] M. Zhou, Y. Li, M. J. Tahir, X. Geng, Y. Wang, and W. He, "Integrated statistical test of signal distributions and access point contributions for Wi-fi indoor localization," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 5057–5070, 2021.

[12] G. Xiao, Q. Cheng, and C. Zhang, "Detecting travel modes using rule-based classification system and Gaussian process classifier," *IEEE Access*, vol. 7, pp. 116741–116752, 2019.

[13] X. Guangnian and Z. Wang, "Empirical Study on Bikesharing Brand Selection in China in the Post-Sharing Era," *Sustainability*, vol. 12, p. 3125, 2020.

[14] T. Arenas, M. N. Martínez, H. Xu, O. Morales, and M. Chávez, "Integrating VSM and network analysis for tourism strategies-case: Mexico and the Chinese outbound market," *Systemic Practice and Action Research*, vol. 32, no. 3, pp. 315–333, 2019.

[15] R. Strulak-Wójcikiewicz, N. Wagner, A. Łapko, and E. Hącia, "Applying the business model canvas to design the E-platform for sailing tourism," *Procedia Computer Science*, vol. 176, no. 5, pp. 1643–1651, 2020.

[16] M. Shiraishi, H. Ashiya, A. Konno et al., "Development of real-time collection, integration, and sharing Technology for Infrastructure Damage Information," *Journal of Disaster Research*, vol. 14, no. 2, pp. 333–347, 2019.

[17] K.-Y. Kwon, "A study on the path-dependent in Korea and China's early cultural industry policy," *The Journal of Chinese Cultural Research*, vol. 39, no. 6, pp. 88-89, 2018.

[18] D. E. Setyowati, Y. Antariksa, E. Haryati, H. Haryono, and I. P. P. Salmon, "Local tourism? Why not! Integrating tourism geographic spatial in Ngawi Regency, Indonesia," *Modern Applied Science*, vol. 14, no. 9, pp. 1–8, 2020.

[19] V. V. Yavorska, I. V. Hevko, V. A. Sych, O. I. Potapchuk, and K. V. Kolomiyets, "Features of application of information technologies in modern tourism," *Journal of Geology Geography and Geoecology*, vol. 28, no. 3, pp. 591–599, 2019.

[20] W. Zhu, Y. Hou, E. Wang, and Y. Wang, "Design of Geographic Information Visualization System for marine tourism based on data mining," *Journal of Coastal Research*, vol. 103, no. 1, pp. 30–34, 2020.

[21] K. Y. Chen and S. Y. Yang, "A cloud information monitoring and recommendation multi-agent system with friendly interfaces for tourism," *Applied Sciences*, vol. 9, no. 20, pp. 43–48, 2019.

[22] I. Ramos-Soler, "AM Martínez-Sala, Campillo-Alhama C. ICT and the sustainability of world heritage sites. Analysis of senior Citizens' use of tourism apps," *Sustainability*, vol. 11, no. 11, pp. 32-33, 2019.

[23] S. L. Ratnasari, E. N. Susanti, W. Ismanto, R. Tanjung, D. C. Darma, and G. Sutjahjo, "An experience of tourism development: how is the strategy?," *Journal of Environmental Management and Tourism*, vol. 11, no. 7, pp. 1877–1886, 2020.

[24] T. Melikh, D. Voit, and D. Archybisova, “Aquacultural Integration in recreational tourism: features of development and management of coastal territories,” *Baltic Journal of Economic Studies*, vol. 5, no. 5, pp. 84–95, 2020.

[25] M. N. Rifa'I, “Integrasi Pariwisata Halal di Kota Malang,” *FALAH Jurnal Ekonomi Syariah*, vol. 4, no. 2, pp. 194–196, 2019.

[26] G. Manyara, “Regional tourism in inter-governmental authority on development: a comparative policy and institutional best practice approach,” *International Journal of Tourism Policy*, vol. 9, no. 1, pp. 50–52, 2019.

[27] M. Zhou, Y. Long, W. Zhang et al., “Adaptive genetic algorithm-aided neural network with channel state information tensor decomposition for indoor localization,” *IEEE Transactions on Evolutionary Computation*, 2021.

[28] M. Abd, “The integration time-driven activity-based costing (TDABC) and events approach their role in decision-making and their effect on tourism,” *African Journal of Hospitality Tourism and Leisure*, vol. 8, no. 2019, pp. 51–53, 2019.

[29] R. Peng, Y. Lou, M. Kadoch, and M. Cheriet, “A human-guided machine learning approach for 5G smart tourism IoT,” *Electronics*, vol. 9, no. 6, pp. 947–949, 2020.

[30] Z. Lv, D. Chen, R. Lou, and Q. Wang, “Intelligent edge computing based on machine learning for smart city,” *Future Generation Computer Systems*, vol. 115, pp. 90–99, 2021.

Design and Management of Control System for Rural Tourism Network Information Based on MVC Model

Hanlin Dong [1,2]

[1] *Suan Sunandha Rajabhat University, Bangkok, Thailand*
[2] *Shandong Vocational Institute of Fashion Technology, Taian 271000, China*

Correspondence should be addressed to Hanlin Dong; 3216003469@mail2.gdut.edu.cn

Academic Editor: Fazlullah Khan

Tourism has developed into an industry with a powerful momentum of development in the world today. With the development of information technology, information has become a powerful driving force to promote the prosperity and development of the tourism industry and the entire society. The introduction of the tourism information system can significantly improve the service level, operation level, and management level of the tourism industry, thereby accelerating tourism development. The research of this article is to help villages establish a set of MVC-based rural tourism information service systems that can promote the development of rural tourism. First of all, this article conducts demand research on tourists, rural scenic spots, and ancient villages to discover the problems in rural tourism development. Secondly, this paper combines the problems existing in constructing the various subsystems of the rural tourism information system, combined with the fuzzy comprehensive evaluation method. In this paper, we propose the rural tourism system architecture based on MVC. The system architecture consists of the user layer, service layer, business layer, and data layer. It describes the system's implementation process from two aspects: the system's interface design and its deployment model. Finally, the network topology structure of the rural tourism information service system based on MVC is drawn. Finally, the system's deployment is implemented according to the network topology structure.

1. Introduction

With the accelerated pace of urbanization and the intensification of competition, urban residents increasingly prefer to travel to the countryside. These are coupled with the continuous construction of rural roads and transportation facilities, providing unprecedented convenience for rural tourism. Rural tourism can be divided into traditional rural tourism and modern rural tourism [1, 2]. Traditional rural tourism appeared after the industrial revolution. Urban residents from rural areas mainly embodied it by going home to visit relatives [3]. Traditional rural tourism has a particular economic impact on the local area. However, it has not effectively promoted the economic development of the local rural villages. It cannot provide local employment opportunities and improve the local rural financial environment [4].

In today's society, in the western developed countries, high and new technology is widely promoted and applied in tourism development, tourism management, etc. It has improved the work efficiency of tourism areas and the tourism experience of tourists [5, 6]. Concepts such as electronic maps, satellite guides, ticketless travel, and virtual travel were quickly adopted by the tourism industry [7]. Grezel proposed the critical design elements of the smart tourism system, which mainly include related background information such as tourism resources, culture, and language. Therefore, to better understand the needs of tourists and design the smart tourism system according to the needs of tourists [8], Ricci proposed a web-based travel recommendation system, using a use-case-oriented approach to provide quality services for tourists to make plans [9]. IBM's "Smart Hotel" project pursues humanized brilliant experience service. It proposes four robust solutions, including systematic monitoring and management, electronic keys, network unsubscription, and desktop cloud [10].

At present, smart tourism, mobile tourism, and rural tourism have flourished in a certain way. However, tourists can only passively perceive tourism information during the travel process [11]. Moreover, the information transmitted between the government, various tourism-related enterprises, and tourists is not smooth. It weakens the tourist's sense and interactivity [12]. In order to tap the needs and pain points of tourists' travel information services, it is necessary to meet the needs of users. Tourism has developed into an industry with a robust development momentum in the world today [13]. The research of this article is to help villages establish a set of MVC-based rural tourism information service systems that can promote the development of rural tourism. In order to tap the needs and pain points of tourists' travel information services, it is necessary to be guided by user needs. It can be achieved by combining with the network information technology developed by the times to provide tourists with a full range of services before, during, and after travel.

The rest of the paper is in accordance with the following pattern. In Section 2, network informationization is studied. In Section 3, network informationization of rural tourism based on the MVC model is given. In Section 4, the rural network tourism system testing is elaborated. Finally, the paper is concluded in Section 5.

2. Network Informationization of Rural Tourism Based on MVC Model

2.1. Informatization of Rural Tourism Network. Rural tourism began in the 1830s, and after the 1980s, rural tourism began to develop on a large scale. Now, it has a considerable scale in some western developed countries. In some highly urbanized regions and countries, rural tourism can account for 10%–25% of all tourism activities [14–16]. The development of rural tourism has effectively changed the phenomenon of rural economic downturn. The contribution of rural tourism to the local economy and the significance of local development have been well proven. In many countries, it is agreed that rural tourism is the driving force of economic development and economic diversification in remote rural areas [17, 18].

In addition to the support of national policies, rural tourism development also has a profound background. Nowadays, the economic level of urban residents has improved, and their leisure time has increased. Also, urban residents' physical and mental needs to return to nature are more urgent [19, 20]. The open space, fresh air, beautiful environment, and rich local culture in the rural areas can meet the desire of urban tourists to return to nature and return to the basics. Rural tourism is rapidly formed and developed under such conditions [21, 22]. Today, rural tourism is in the ascendant. After a long development period, rural tourism has developed from the initial spontaneous stage to the present conscious stage. Rural tourism can greatly promote rural economic development. Rural tourism development actively utilizes the agricultural natural environment, agricultural production and management activities, and human resources [23, 24]. After planning and design, the formation of leisure tourism and holiday park with pastoral pleasures can effectively perform agricultural production functions and increase agricultural income.

2.2. MVC Model. The MVC pattern is a software architecture pattern. It is a software architecture pattern that separates the three modules of view, controller, and model. The advantage of this design is that system developers and system designers can perform their maintenance [25, 26]. Therefore, it improves the reuse rate of system code and also improves the scalability of system applications. The most significant advantage of the system is that it brings great convenience to system development. MVC is the ideal way to use three different parts to construct a software or component. It provides a powerful object separation mechanism, makes the program more object-oriented, and handles the design of the software architecture and the development of the program. The core idea of the model is to combine effectively "model," "view," and "controller" [27, 28]. The model is used to store data objects. The view provides the data display object for the model. The controller is responsible for specific business logic operations and is responsible for matching various operations performed by the "view layer" to the corresponding data of the "data layer" and displaying the results. The structure of the MVC model is shown in Figure 1.

The user interacts with the view page, and some requests input by the user are first received by the controller. Then, it is responsible for selecting the corresponding model for processing. The model processes the user's request through business logic and returns the processed data [29, 30]. Finally, the controller selects the appropriate view to format the returned data. The separation between the three components allows a model to be displayed in multiple different views. When the user changes the data in the model through the controller of a particular view, all other views that depend on the data should reflect these changes. In short, no matter what data changes occur at any time, the controller will notify these changes to all associated views and update the related displays [31, 32]. It is a change propagation mechanism of the model.

3. Network Informationization of Rural Tourism Based on MVC Model

The rural tourism information service system is designed from functional design, class design, and database design. For the functional design of the system, first, analyze the system structure of rural tourism. Then, draw the overall functional structure diagram of the system. Finally, decompose the subsystems of the rural tourism information service system [33, 34]. To design the class of the system, first determine the class of the system and then determine the relationship between the system classes to draw a class diagram. The conceptual design and the logical design carry on the database design to the system using the database ER

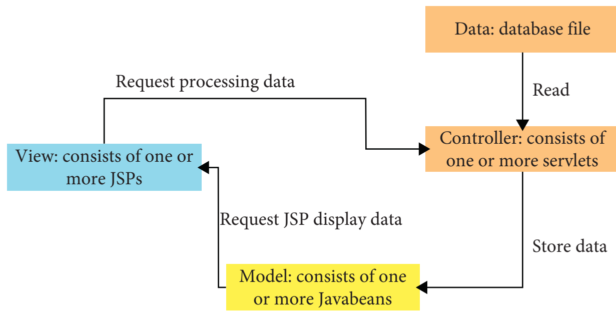


Figure 1: MVC model structure diagram.

diagram and the database table to design the system database [35, 36]. A theory provides a very effective tool for intelligent information processing. It is a new mathematical method for dealing with inaccurate, uncertain, and incomplete data.

In the fuzzy comprehensive evaluation method, determining the weight is mainly based on the AHP analytic method. The weight calculation results in more accurate values [37, 38]. The AHP analytic hierarchy process is a decision-making method that combines quantitative and qualitative analysis. This method is mainly proposed for problems with multiple structural layers and is affected by multiple factors. First of all, we must determine the evaluation indicators, and then, calculate the weights of indicators at all levels. Then, measure the degree of influence of the changes of each factor on the overall system to comprehensively detect the effectiveness of the evaluation results [39, 40]:

$$\begin{aligned} \mathrm{d}'(t,x) &= \int_{-\infty}^{+\infty} \mu(\tau - px, x)\mathrm{d}p, \\ \mu(\tau, p) &= \sum_{i}^{Nx} \mathrm{d}(t = \tau + px_i, x_i)\Delta x_i. \end{aligned} \tag{1}$$

When the fuzzy comprehensive evaluation method is used in this paper, firstly, AHP is used to establish an analytic hierarchy model for the effectiveness of internal control operation, which is divided into 5 first-level indicators and 17 second-level indicators:

$$\begin{aligned} \left|f^{(j)} - f^{(j-1)}\right| &\le \tau, \\ \left|f^{(j)} - f^{(b)}\right| &\le \tau. \end{aligned} \tag{2}$$

A detailed description of the secondary indicators is used as the specific question of the questionnaire. The degree of membership of the indicators at all levels is calculated based on the questionnaire results [41, 42]. At the same time, the evaluation set is established. Then, the index comparison matrix is established according to the analysis model. Invite experts to score the importance of the comparison matrix, use the AHP analytic method to calculate the weights of indicators at all levels, and check the consistency of the weight values [43–45]. Under the premise of consistent inspection results, the evaluation results of the indicators at all levels are calculated through the fuzzy comprehensive evaluation mathematical formula. They combine each indicator's weight and the evaluation results to comprehensively calculate the index scores of the effectiveness of internal control operations at all levels. Take the column corresponding to each indicator in the scoring table as the comparison sequence of factors:

$$D(v) = \sum C\left(v_i, v_j\right), \quad \text{where } v_i, v_j \in V. \tag{3}$$

The correlation degree represents the relationship between the comparison sequence and the reference sequence. The formula for solving the correlation analysis is as follows [46]:

$$D(v_0) \le D(v_1) \le D(v_2) \le \mathrm{L} \le D(v_n). \tag{4}$$

Because the dimensions of many selected evaluation indicators are not the same, it is more difficult to compare. It requires dimensionless processing for all initial indicator values:

$$D(v_{n+1}) = \min\{D(v_n) + C(v_i, w)\}, \quad \text{where } v_i \in U, w \in V - U. \tag{5}$$

After the cascade structure is constructed, the evaluation indicators of each level are compared with each other to determine the importance of the factors contained in each level. Then, the judgment matrix is determined according to the Saaty 1–9 scale method:

$$\begin{aligned} &x'_j \le x_j \le x''_j, \quad (j = 1, 2, \ldots, n), \\ &g_j \le g''_j, \quad (j = 1, 2, \ldots, n), \\ &h'_j \le h_j(x), \quad (j = 1, 2, \ldots, n), \\ &w'_j \le w_j(x) \le w''_j, \quad (j = 1, 2, \ldots, n). \end{aligned} \tag{6}$$

According to the above steps, scoring each indicator factor of the TV drama evaluation is weighted and summed for each level:

$$Q(x, q) = \frac{f}{f_0} + \sum_j p_x\left(x_j\right) + q\left[\sum_j p_g\left(g_j\right) + \sum_j p_h\left(h_j\right) + \sum_j p_w\left(w_j\right)\right]. \tag{7}$$

It is possible to finally calculate the total score of each index factor that affects the running status of the TV series, which also reflects the degree of influence of the index factor on the comprehensive evaluation of the running status of the TV series:

$$x^{(j+1)} = x^{(j)} + s_j d^{(j)}. \tag{8}$$

The construction of the evaluation index system is the basis for evaluating the image of a tourist destination. Different subjects have significant differences in the selection of evaluation indexes for the same evaluation object. Therefore, when constructing the evaluation index system, the following principles should be followed. According to the process mentioned above, after obtaining the fuzzy evaluation matrix of various aspects and the weight of each index, a multilevel fuzzy comprehensive evaluation should be carried out on the evaluated object.

4. Rural Tourism Network System Test

4.1. Test Environment and Program. The testing method used in this article is the unit testing method, which is a method used in many software testing processes now, i.e., when the functional module of a certain unit is designed. It is tested, and if a problem is found, it is immediately modified. After the design of the entire system is completed, it will be tested as a whole. It is often the content of functional testing, performance testing, etc., that the user needs to perform. Software testing is a process with strict procedures and should follow specific rules. In a nutshell, software testing should meet the following basic principles. The design work of the spatial database of the rural tourism network system is shown in Figure 2.

A very important job before system development is spatial database design, which is mainly used to record the spatial information of data and the primary attributes of primitives. The spatial system database mainly includes two aspects: one is the primary geographic database, and the other is the tourism thematic spatial database.

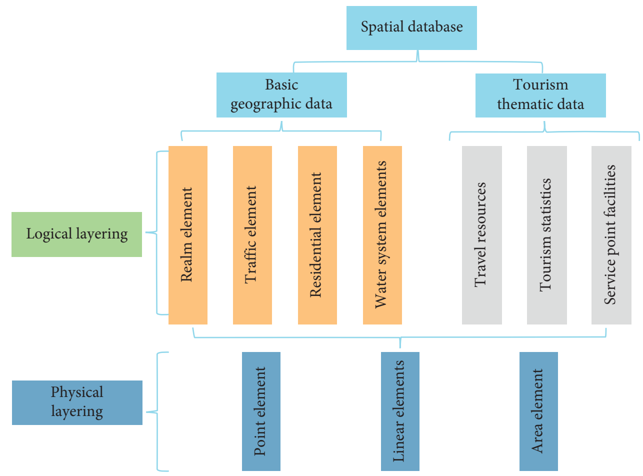


Figure 2: Spatial database of the rural tourism network system.

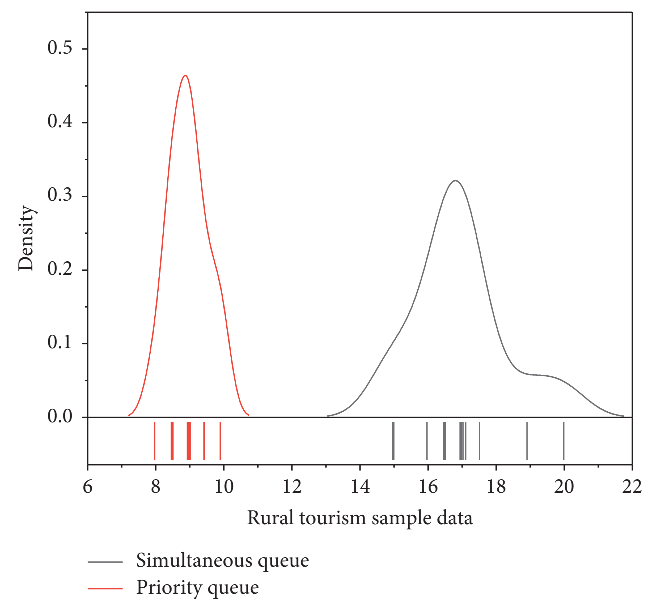


Figure 3: Comparison between the priority request scheduling strategy and the case of sending requests at the same time.

4.2. Test Results. A system function test is a test for the analysis of user functional requirements. It is the most basic design requirement of the entire system. In the previous design process, we have introduced the functions of each unit, and here is the overall function of the system: testing and analysis. In the function test of rural electric tourism based on MVC mode, we mainly test the main function modules of the system.

According to the principle of performance response time, the user's response speed to the page is divided into four evaluation levels. Here, less than 2 seconds indicates that the page response is fast, while 2–5 seconds indicates that the response speed is okay. The 5 seconds indicates the response speed plodding but acceptable. After more than 8 seconds, the page does not respond, and the user may send the request again or leave the site. From user behavior, the increase in page response speed is conducive to improving user experience. The test environment sends 5 requests with priority 1 and 5 requests with priority 2 on the page. Figure 3 shows that the priority request scheduling strategy is compared with the case of sending requests simultaneously, and the request with priority 1 is responded to the time required.

Figure 4 is based on the front-end framework implemented in this article, sending 50 requests to the server and comparing the time required for the page to get the server's response and load to complete.

A geographic information system (GIS) is a computer system that collects, displays, stores, analyzes, manages, and applies geographic information. It is a general technology for analyzing and processing massive geographic data. A variety of geospatial entity data and their relationships are the objects of GIS processing and management. It includes spatial positioning data, attribute data, graphic data, and remote-sensing image data. It is mainly used to analyze and process various phenomena and processes to solve complex management, decision-making, and planning issues. Figure 5 shows the test results of the regional spatial distribution of the rural tourism network.

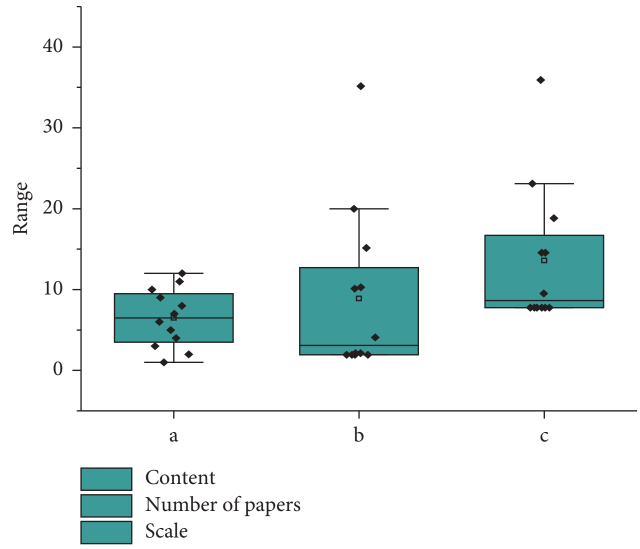


Figure 4: The page gets the server's event response result.

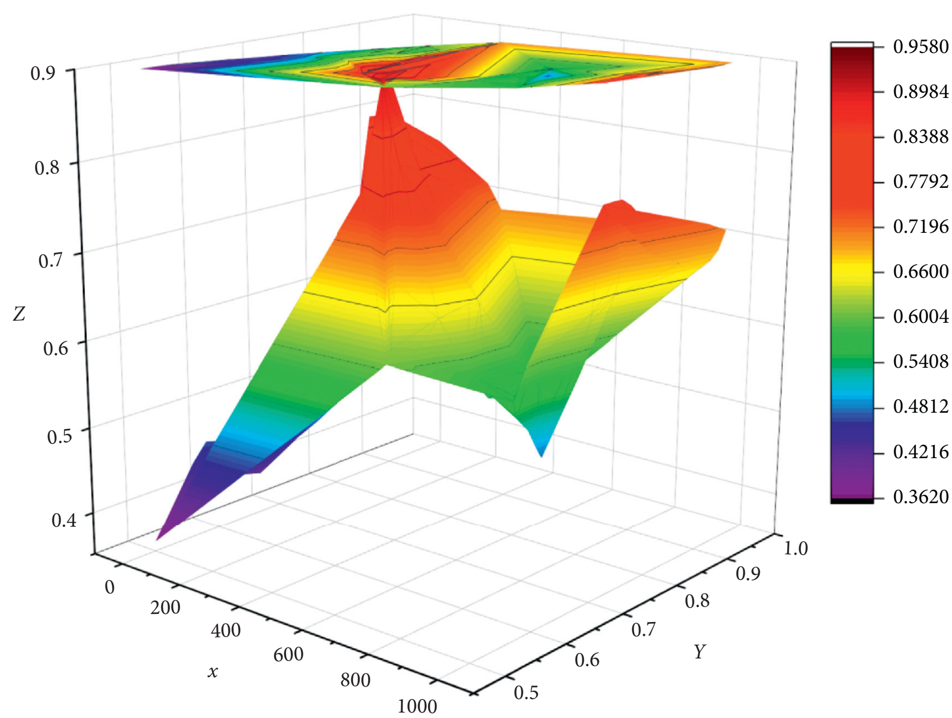


Figure 5: Test of the regional spatial distribution of rural tourism network.

Data maintenance users can import, create, draw, and electronic export maps in various formats for the database in the map control management module. On the client side, visitors can zoom out, zoom in, pan, and display all types of travel-related maps they need. The system has an eagle-eye navigation function.

5. Conclusions

Tourism is an industry with intensive information demand. The unprecedented development and advancement of mobile information technology have promoted the development of mobile tourism information services. The ways of tourism information services have become more diverse and intelligent. Due to the continuous increase of the urban population, social pressure and life pressure is gradually increasing. People are more willing to choose the form of rural tourism to pursue the experience of returning to nature and relaxing the body and mind. With the improvement of living standards, people expect the scenery of scenic spots during travel and receive high-quality services and get a good experience in obtaining travel information. This article studies the development of rural tourism in the context of smart tourism, starting from the three stages of tourism, including food, housing, transportation, travel, shopping, and entertainment. Entertainment is starting from the three stages of tourism. Experience is the starting point, exploring the needs of tourists for rural tourism information services. Combining the overall framework of the information service system provides a reference for the tourism management and tourism information service design of rural scenic spots. Tourism has developed into an industry with a powerful momentum of development in the world today. The introduction of the tourism information system can significantly improve the service level, operation level, and management level of the tourism industry, thereby accelerating tourism development. The research of this article is to help villages establish a set of MVC-based rural tourism information service systems that can promote the development of rural tourism.

References

[1] T. Dax, D. Zhang, and Y. Chen, “Agritourism initiatives in the context of continuous out-migration: comparative perspectives for the alps and Chinese mountain regions,” *Sustainability*, vol. 11, no. 16, p. 4418, 2019.
[2] F. Mancini, G. M. Coghill, and D. Lusseau, “Quantifying wildlife watchers’ preferences to investigate the overlap between recreational and conservation value of natural areas,” *Journal of Applied Ecology*, vol. 56, 2019.
[3] G. M. Robinson and B. Song, “Rural transformation: cherry growing on the guanzhong plain, China and the adelaide hills, south Australia,” *Journal of Geographical Sciences*, vol. 29, no. 5, pp. 675–701, 2019.
[4] S. Yang, “Eco-agricultural economic development strategy based on improving the eco-cultural tourism environment in rural areas along the coast,” *Journal of Coastal Research*, vol. 104, no. sp1, 2020.
[5] L. Zhao and R. Hou, “Human causes of soil loss in rural karst environments: a case study of Guizhou, China,” *Scientific Reports*, vol. 5, 2019.
[6] J. Chlachula, “Between sand dunes and hamadas: environmental sustainability of the thar desert, west India,” *Sustainability*, vol. 13, no. 7, p. 3602, 2021.
[7] A. Ib, B. Ys, B. Dva et al., “A 3d additive manufacturing approach for the validation of a numerical wall-scale model of catalytic particulate filters,” *Chemical Engineering Journal*, vol. 405, 2020.
[8] A. Fg, C. Mfb, B. Mvc, A. Cb, A. Mf, and C. Frb, “On the difference between real-time and research simulations with ctipe,” *Advances in Space Research*, vol. 64, no. 10, pp. 2077–2087, 2019.
[9] J. L. Sánchez, M. Padrino, E. Martinez, and M. Florez, “The use of peleg’s equation to model water absorption in triticale (x triticosecale wittmack) seeds magnetically treated before soaking,” *Romanian Journal of Physics*, vol. 64, no. 3-4, 2019.
[10] G. Vesentini, A. M. P. Barbosa, D. C. Damasceno et al., “Alterations in the structural characteristics of rectus abdominis muscles caused by diabetes and pregnancy: a comparative study of the rat model and women,” *PLoS One*, vol. 15, no. 4, Article ID e0231096, 2020.
[11] M. L. Elsayed, O. Mesalhy, R. H. Mohammed, and L. C. Chow, “Transient and thermo-economic analysis of med-mvc desalination system,” *Energy*, vol. 167, no. JAN.15, pp. 283–296, 2019.
[12] M. F. Anjos and M. Vieira, “Mathematical optimization approach for facility layout on several rows,” *Optimization Letters*, vol. 6, no. 12020.
[13] M. F. Anjos and M. Vieira, “Mathematical optimization approaches for facility layout problems: the state-of-the-art and future research directions,” *Operations Research*, vol. 59, no. 1-2, pp. 89–91, 2019.
[14] S. K. Verma, D. Torres-Sanchez, K. R. Hernández-MArtínez, V. P. Malviya, and B. A. Rivera-Escoto, “Geochemistry of eocene felsic volcanic rocks from the mesa virgen-calerilla, zacatecas, Mexico: implications for the magma source and tectonic setting,” *Geological Journal*, vol. 56, 2021.
[15] W. M. Gilliland, H. A. Prince, A. Poliseno, A. Kashuba, and E. P. Rosen, “Ir-maldesi mass spectrometry imaging of human hair to characterize longitudinal profiles of the antiretroviral maraviroc for adherence monitoring,” *Analytical Chemistry*, vol. 91, no. 16, 2019.
[16] D. P. Salgado Terêncio, L. F. Sanches Fernandes, R. M. Vitor Cortes, J. P. Moura, and F. A. Leal Pacheco, “Can land cover changes mitigate large floods? a reflection based on partial least squares-path modeling,” *Water*, vol. 11, no. 4, p. 684, 2019.
[17] M. C. dos Santos, J. L. R. Scaini, M. V. C. Lopes et al., “Mefloquine synergism with anti-tuberculosis drugs and correlation to membrane effects: biologic, spectroscopic and molecular dynamics simulations studies,” *Bioorganic Chemistry*, vol. 110, p. 104786, 2021.
[18] A. Yr, A. Sh, Z. A. Peng, B. Mh, and A. Zx, “Self-paced and auto-weighted multi-view clustering-sciencedirect,” *Neurocomputing*, vol. 383, pp. 248–256, 2020.
[19] S. Chang, J. Hu, T. Li, H. Wang, and B. Peng, “Multi-view clustering via deep concept factorization,” *Knowledge-Based Systems*, vol. 217, no. 2, p. 106807, 2021.
[20] R. Jayendiran, S. Campisi, M. Viallon, P. Croisille, and S. Avril, “Hemodynamics alteration in patient-specific dilated ascending thoracic aortas with tricuspid and bicuspid aortic valves,” *Journal of Biomechanics*, vol. 110, 2020.
[21] I. Belot, D. Vidal, R. Greiner, M. Votsmeier, R. E. Hayes, and F. Bertrand, “Impact of washcoat distribution on the catalytic performance of gasoline particulate filters as predicted by lattice Boltzmann simulations,” *Chemical Engineering Journal*, vol. 406, 2020.
[22] V. Brancato, M. Ventre, R. L. Reis, and P. A. Netti, “Decellularized matrices for tumor cell modeling,” *Methods in Cell Biology*, vol. 157, pp. 169–183, 2020.
[23] D. Spieler, C. Namendorf, T. Namendorf, M. von Cube, and M. Uhr, “Donepezil, a cholinesterase inhibitor used in alzheimer’s disease therapy, is actively exported out of the brain by abcb1ab p-glycoproteins in mice,” *Journal of Psychiatric Research*, vol. 124, pp. 29–33, 2020.
[24] D. P. S. Terêncio, F. A. L. Pacheco, L. F. Sanches Fernandes, and R. M. V. Cortes, “Is it safe to remove a dam at the risk of a sprawl by exotic fish species?” *Science of the Total Environment*, vol. 771, 2021.

[25] G. Baldin, A. Ciccullo, S. Rusconi et al., "Long-term data on the efficacy and tolerability of lamivudine plus dolutegravir as a switch strategy in a multi-centre cohort of hiv-1-infected, virologically suppressed patients," *International Journal of Antimicrobial Agents*, vol. 54, no. 6, pp. 728–734, 2019.

[26] M. Morais, A. I. Oliva-Avilés, M. Matos et al., "On the effect of electric field application during the curing process on the electrical conductivity of single-walled carbon nanotubes–epoxy composites," *Carbon*, vol. 150, 2019.

[27] D. Pavkovic, P. Prljan, M. Cipek, and M. Krznar, "Crossaxis control system design for borehole drilling based on damping optimum criterion and utilization of proportionalintegral controllers," *Optimization and Engineering*, vol. 22, 2021.

[28] W. Jiang, G. C. Ye, D. H. Zou, and Y. Yan, "Mechanism configuration and innovation control system design for power cable line mobile maintenance robot," *Robotica*, vol. 39, pp. 1–13, 2020.

[29] S. Cheng, L. Li, Y.-G. Liu, W.-B. Li, and H.-Q. Guo, "Virtual fluid-flow-model-based lane-keeping integrated with collision avoidance control system design for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–10, 2020.

[30] X. Xu, X. Liu, and Y. Zhang, "Design of a digital control system of gyroscope," *Modern Physics Letters B*, vol. 3, 2021.

[31] Q. Shen, C. Yue, H. C. Goh, and D. Wang, "Active fault-tolerant control system design for spacecraft attitude maneuvers with actuator saturation and faults," *IEEE Transactions on Industrial Electronics*, vol. 66, 2019.

[32] M. Aghanouri, A. Ghaffari, and N. D. Serej, "Image based high-level control system design for steering and controlling of an active capsule endoscope," *Journal of Intelligent & Robotic Systems*, vol. 94, no. 1, 2019.

[33] P. Zítek, J. Fier, and T. Vyhlídal, "Dynamic similarity approach to control system design: delayed pid control loop," *International Journal of Control*, vol. 92, 2019.

[34] S. S. Yang, Y. H. Jang, M. H. Park, S. C. Park, and H. J. Kim, "Design and implementation of active access control system by using nfc-based eap-aka protocol," *Wireless Personal Communications*, vol. 118, no. 6, 2021.

[35] Q. Hong, C. Booth, M. Karimi, M. Sun, and B. Marshall, "Design and validation of a wide area monitoring and control system for fast frequency response," *IEEE Transactions on Smart Grid*, vol. 99, p. 1, 2020.

[36] M. Aull and K. Cohen, "A nonlinear inverse model for airborne wind energy system analysis, control, and design optimization," *Wind Energy*, vol. 24, 2021.

[37] R. Markam and A. K. Bajpai, "Functionalization of ginger derived nanoparticles with chitosan to design drug delivery system for controlled release of 5-amino salicylic acid (5-asa) in treatment of inflammatory bowel diseases: an in vitro study," *Reactive and Functional Polymers*, vol. 149, no. Apr, pp. 104520-104521, 2020.

[38] T. L. Grigorie, S. Khan, R. M. Botez, M. Mamou, and Y. Mébarki, "Design and experimental testing of a control system for a morphing wing model actuated with miniature bldc motors," *Chinese Journal of Aeronautics*, vol. 33, no. 4, 2020.

[39] K. Yamakawa, A. Augustinus, G. Batigne et al., "Design and implementation of detector control system for muon forward tracker at alice," *Journal of Instrumentation*, vol. 15, no. 10, Article ID T10002, 2020.

[40] F. Gaetani, P. Primiceri, G. Antonio Zappatore, and P. Visconti, "Hardware design and software development of a motion control and driving system for transradial prosthesis based on a wireless myoelectric armband," *IET Science, Measurement & Technology*, vol. 13, no. 3, pp. 354–362, 2019.

[41] W. Peng, J. Liu, and Y. Wang, "Design of an achromatic projection system on a curved surface for enlarging view using polarization control metasurface," *Optical Engineering*, vol. 58, no. 4, p. 1, 2019.

[42] X. Zhai, X. U. Yu, and Y. U. Zhijin, "Design and performance simulation of a novel liquid Co_2 cycle refrigeration system for heat hazard control in coal mines," *Journal of Thermal Science*, vol. 28, no. 3, pp. 195–205, 2019.

[43] Y. Zuobin, S. I. Yuanping, M. A. Jianfeng, J. Wenjie, X. U. Shengmin, and L. Ximeng, "P2hbt: partially policy hidden e-healthcare system with black-box traceability," *Chinese Journal of Electronics*, vol. 30, no. 2, 2021.

[44] G. Zhong, Y. Cui, M. Qian, X. Yan, and H. Wang, "A closed-loop brain stimulation control system design based on brain-machine interface for epilepsy," *Complexity*, vol. 2020, Article ID 3136715, 15 pages, 2020.

[45] S. Peng, X. Jiang, Y. Tang et al., "Recoverable autonomous sonde for subglacial lake exploration: electronic control system design," *Annals of Glaciology*, vol. 3, pp. 1–17, 2021.

[46] Y. V. Mitrishkin, P. S. Korenev, N. M. Kartsev, E. A. Kuznetsov, A. A. Prokhorov, and M. I. Patrov, "Plasma magnetic cascade multiloop control system design methodology in a tokamak," *Control Engineering Practice*, vol. 87, pp. 97–110, 2019.

Optimization of Tourism Information Analysis System Based on Big Data Algorithm

Jing Yang,[1] Bing Zheng,[1] and Zhenghua Chen[2]

[1]*School of Information Engineering, Hainan Vocational University of Science and Technology, Haikou 571126, China*
[2]*School of of Finance and Economics, Hainan Vocational University of Science and Technology, Haikou 571126, China*

Correspondence should be addressed to Bing Zheng; 21227082@zju.edu.cn

Academic Editor: Wei Wang

On the basis of ecological footprint theory and tourism ecological footprint theory, the sustainable development indexes such as ecological footprint, ecological carrying capacity, ecological deficit, and ecological surplus of the research area were calculated and the long-term change pattern of each index was analyzed. This paper shows that the ecological footprint of the research area increases year by year, but the ecological footprint is always smaller than the ecological carrying capacity, indicating that the area is still in the state of sustainable development. However, the per capita ecological surplus shows a decreasing trend year by year, indicating that the sustainable development of the region is getting worse. This paper proposes a reordering method of tourist attractions based on heterogeneous information fusion, and realizes the retrieval and reordering of tourist attractions based on user query and fusion of heterogeneous information, so as to help users make travel decisions. In view of the shortage of tourism commercial websites to passively provide scenic spot information, this paper puts forward a scenic spot retrieval method based on query words to enable users to obtain scenic spot information according to their needs, and constructs a tourist consumer data analysis system. The preprocessing methods and methods adopted by the data preprocessing module are analyzed in detail, and the algorithms used in the travel route analysis and consumer spending ability analysis are described in detail. The data of tourism consumers are analyzed by this system, and the results are evaluated.

1. Introduction

Tourism information is the basic data guarantee of tourism informatization, so it is particularly important to obtain high-quality tourism data resources. Tourism information not only exists in a large number of government tourism departments and tourism companies, but also with the rapid development of social media in recent years, and there are also a large number of tourism information resources available on the major social media websites. However, it is difficult for people to accurately obtain data from the vast amount of travel information, especially for the diversity of travel destinations or dazzling many travel routes. The current tourism information service simply displays the information of tourist attractions on the Internet, and its disadvantages are mainly reflected in the following aspects: (1) the information is not comprehensive enough, and the rich tourism information shared on social media is not fully utilized; (2) information is only passively presented to users and cannot be searched and filtered according to users' needs; (3) ignore user needs, fail to consider user context information, and fail to achieve customized and personalized services for users. Without intelligent consideration of the comprehensiveness and preferences of the information, it is often impossible to recommend satisfactory results for users. These disadvantages have seriously restricted the acquisition of high-quality tourism information, the recommendation of destinations in line with the needs and personalized planning, and customization of a series of tourism behaviors. Therefore, in addition to mastering huge tourism information resources, tourism informatization also needs to deal with mine and analyze these massive data professionally.

No matter at home or abroad, tourism route planning is an important part of tourism planning and design [1, 2] and the core of travel agency management. With the continuous development and innovation of tourism, many new tourism modes have emerged, such as self-help travel, self-driving travel, and experiential tourism. Tourists can choose scenic spots and routes according to their preferences, and their tour preferences are more personalized. However, tourist routes planned by travel agencies are mostly for the purpose of making profits, and the design of relevant tourist routes is mainly based on factors such as time, traffic, and scenic spots, and other factors in specific profit space in specific areas [3, 4]. At present, tourism route planning mostly focuses on regional tourism planning [5, 6], which is aimed at a large tourist scenic spot. Within the scenic spot, there are a set of scenic spots located in different spatial positions. Different sightseeing modes can be obtained through tour sequence design of these scenic spots. However, the route planning method for scenic spots is not applicable to the planning of a large area such as scenic spots in the city [7]. In the field of tourism, users usually share their experiences and comments after a trip, forming a large number of user-generated content including user comments, photos, travel notes, and other information. These data provide great convenience for trip planning [8, 9]. While there may be noise or bias in a single comment or travel note, incorporating a large amount of user-generated content as a whole can effectively capture the essence of a site. Therefore, in an increasing number of studies, spatial analysis and data mining technologies are used to analyze these contents [10, 11], to obtain users' relevant preferences and historical track information, and to find the similarity between tourists, so as to realize the recommendation of tourist routes [12, 13]. With the advent of Web 2.0 era, online multimedia sharing websites have become popular. The information uploaded by users contains a large number of travel-related contents, which can be widely used in the tourism system. Therefore, in recent years, many intelligent tourism systems have been established to achieve accurate retrieval or personalized recommendation functions through the analysis and mining of tourism multimedia information, thus making travel more convenient and faster [14, 15]. Wikitravel is an early tourism information system before the advent of Web 2.0 era, providing users with open, complete, real-time, and trusted tourism information [16, 17]. It provides online travel services, devotes itself to mining high-quality travel photos from image sharing websites, and designs a user interface with search function and map positioning function, which can provide help for users to plan their path and travel [18, 19]. By analyzing more than 110,000 images with geographical marks on Flickr, the visual perspective of scenic spots is generated by using the images, and the diversity of scenic spot search results is satisfied [20, 21]. Combined with text, geotagged images, and video, the scenic spot summary is generated, and then personalized scenic spot summary is recommended to users by means of query [22, 23]. Another type of system focuses on the retrieval and recommendation of tourism multimedia information. Through the analysis of photos and text information, as well as the relevant knowledge in Yahoo travel, the popular scenic spots are recommended for users, and the summary information of scenic spots is returned to users [24]. Through the acquired images and travel notes, the routes within and between scenic spots are mined, so as to provide users with travel route planning [25, 26]. A mobile travel search framework is proposed, which can display multiple perspectives of scenic spots based on image information to users through compression transmission technology [27, 28]. Low resolution query images are processed by remote server, and then scenic spots are identified and searched, and the corresponding scenic spots are reconstructed from the 3d perspective based on the photo set [29]. In order to improve the performance of the system and overcome the limitations of traditional recommendation algorithms, a hybrid recommendation algorithm has gradually become popular. It uses two or more recommendation algorithms by mixing, weighting, switching, cascading, and feature combination to make full use of the advantages of each recommendation algorithm to obtain higher performance. The most common examples of hybrid recommendation algorithms are the hybrid collaborative filtering algorithm and other recommendation algorithms to alleviate cold start and data sparsity problems.

A reordering method of tourist attractions based on heterogeneous information fusion is proposed. For analysis of the current attraction retrieval method and the deficiency of the need to solve the problem, this chapter then introduces the algorithm implementation of the block diagram; then, it is introduced based on the content and the heterogeneous information mining method based on the score, as well as to the scenic spots, based on the query of the initial sorting method, based on the content of the resort sorting method, and based on the score of the reorder adjustment method. Finally, the objective and subjective experiments verify that the proposed method of scenic spot reordering based on heterogeneous information fusion can efficiently obtain scenic spot information based on user query. Data stored in the database may reflect noise, anomalies, or incomplete data objects. These objects may have some adverse effects on the analysis process, resulting in the over-adaptation of data to the constructed knowledge model or the failure of mining analysis. As a result, the patterns found can be very inaccurate. Data cleaning methods and data analysis methods to deal with data noise and outlier mining methods to find and analyze abnormal situations are required.

2. Optimization Method and Framework of Tourism Information Big Data Analysis

The ecological carrying capacity and sustainability of an area are affected by many aspects, such as meteorology, hydrology, geology, environment, social economy, and other industries and fields. The change of ecological carrying capacity and sustainability is not only related to the state of each single factor, but also to the result of the interaction of all factors. This chapter will introduce the monitoring methods and principles of all aspects of the scenic and historic interest area and its surrounding areas in detail from three aspects of ecological carrying capacity and sustainability assessment. Among them, the three aspects mainly include the mining and analysis of the laws of the long time series of tourism ecological environment elements, the evaluation of ecological carrying capacity and sustainable is based on the ecological footprint theory, and the trend prediction and early warning based on the big data technology. The specific frame system is shown in Figure 1. First of all, the long time series multisource heterogeneous original data are collected and sorted. On this basis, the typical tourism ecological environment elements are extracted based on the extraction principles and methods of thematic information. Secondly, for extraction of tourism ecological environment factors of single factor, long time series, the trend of the discussion, and analysis, the main tourism ecological environment factors, including land use cover, vegetation coverage, biodiversity, landscape vulnerability, climate comfort level, the level of economic development, the industrial support ability, tourist reception capacity, swim in proportion, and attractive tourism resources, lay the foundation for subsequent driving force of the ecological footprint analysis. Thirdly, the ecological footprint theory, which is widely adopted and highly recognized in the world, is used to discuss the long-term changes of sustainable development indicators such as ecological footprint, ecological carrying capacity, and ecological deficit or ecological surplus in the research area. Finally, the big data analysis method is adopted and two kinds of time series prediction models (the ARIMA model and LSTM model) are used to predict and warn the future sustainable development status of the research area.

2.1. Analysis of Tourism Ecological Environment Elements. Vegetation coverage is the ratio of the vertical projection of stems, leaves, branches, and other vegetation onto the ground to the total area of the statistical area, which is usually expressed by percentage. Like the normalized difference vegetation index, it is one of the important indicators to measure the growth status of surface vegetation, and it is of great significance to the regional ecological environment assessment. The calculation methods of vegetation coverage mainly include pixel dichotomy model, regression model, and vegetation index method. Among them, the vegetation coverage calculated based on binary pixel model has been widely used by researchers. The assumptions in the binary model divide each pixel into two parts, one with vegetation coverage and the other with no vegetation coverage. The electromagnetic spectrum information observed by the remote sensor is calculated by the linear weighted sum of the two parts, and the weight of each part is related to the area proportion of the part in the whole pixel, then the vegetation coverage of the pixel is equal to the percentage of the surface covered by vegetation in the whole pixel area.

Because the change of land use cover type directly affects the change of biodiversity, the land use cover map obtained by remote sensing can reflect the difference of regional biodiversity to a certain extent. The weights of various surface feature types are shown in Table 1, including cultivated land, construction land, woodland, water area wetland, grassland, and unused land.

The ratio of tourists to residents refers to the ratio between the total number of tourists to the scenic spot and the total number of local residents, which is an indicator reflecting the psychological carrying capacity of local residents. The development of tourism influences the social culture of local residents to a certain extent, which has positive and negative effects. On the positive side, more tourists will promote the local economic development, increase the employment rate of local residents, and greatly improve people's life. The negative aspect is that the increase of tourists will put pressure on the local ecological environment resources and make the living environment of local residents worse than before. In general, the greater is the density of tourists, the stronger the impact. From the perspective of tourism development, the most essential value of tourism resources lies in its attraction to the whole tourism market, that is, its ability to attract tourists. The core competitiveness of tourism resources development is the attraction of tourism resources, which is affected by many factors, including the quality and richness of tourist attractions, tourist traffic conditions, reception facilities, services, and accommodation. The quantitative expression of the only objective standard to measure this value is the number of visitors the tourist resource can attract.

In order to make up for the limitation of search engine using text to retrieve images, visual features are used to reorder images to make up for the semantic gap between text and image. In addition, due to the different sensitivity of images to different visual features, multivisual features are combined to generate mixed features for reordering. At the same time, in order to ensure the correlation between query words and reordered images, a reordering framework based on the graph model is adopted to complete the reordering of image search results, so as to help users get the most relevant images from a large number of search engine returned images. The proposed image search reordering method based on the mixed feature graph model is to mix the visual features and then use the reordering framework based on the graph model to complete the reordering. Figure 2 shows the block diagram of the image search reordering method based on the hybrid feature graph model, which is mainly divided into two parts: learning hybrid feature and graph-based reordering. In the offline learning stage, all images, after visual features are extracted, use potential semantic analysis to learn the mixed features based on visual feature fusion. In the online sorting stage, a query word is given, and the initial sorting result of the returned image is obtained after matching with the text information of the image. Then, the similarity between the images based on the mixed features is calculated under the reordering framework of the graph model to complete the construction of the graph model and finally the reordering result is given.

In order to verify the proposed method and ensure the higher correlation performance of the image in front, the image number is selected here. In different cases, find its

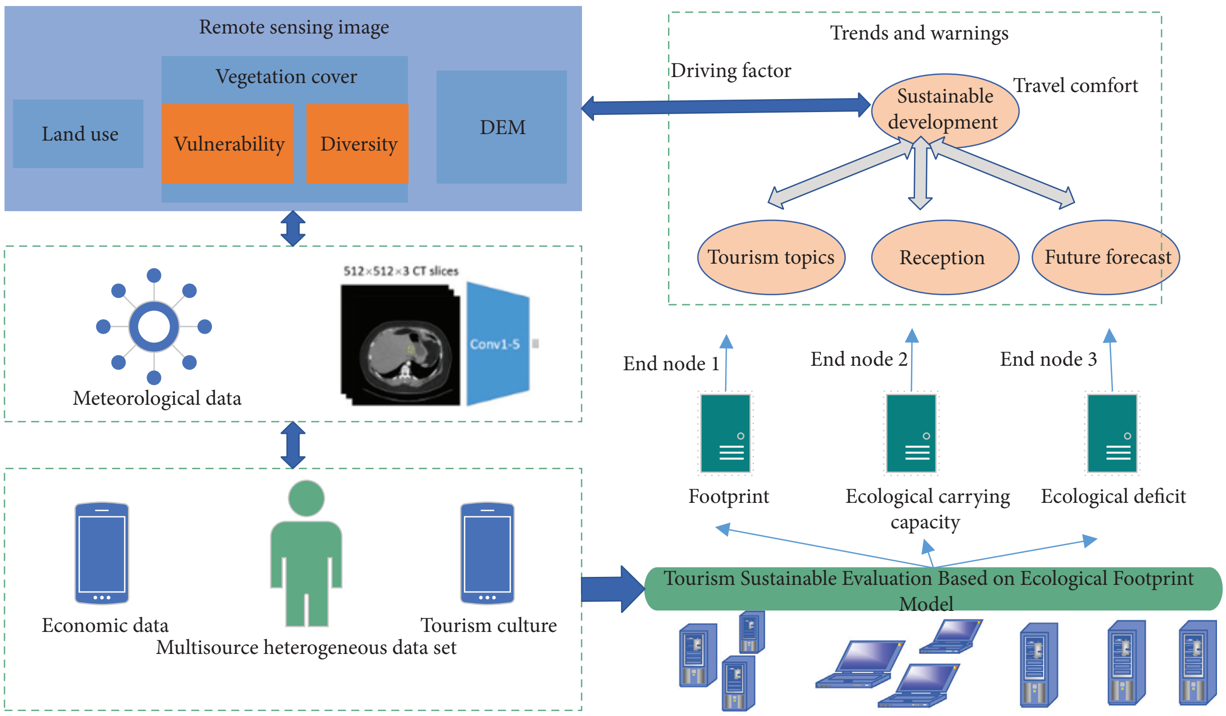


Figure 1: Frame diagram of big data analysis of tourism information system.

Table 1: Biodiversity weight table.

Types of objects	Woodland	Grass	Waters	Arable land	Construction land	Unused
Weight	0.32	0.24	0.26	0.13	0.03	0.02

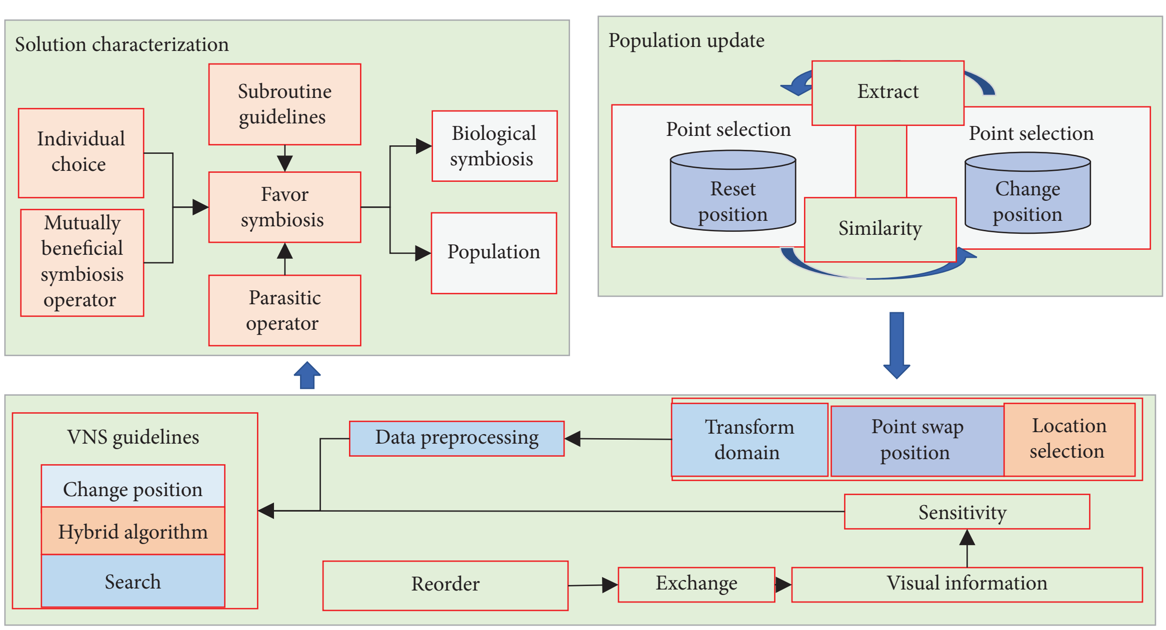


Figure 2: Block diagram of image search reordering based on the hybrid feature graph model.

NDCG value. The change in the NDCG value is shown in Figure 3.

The NDCG value is the average of the NDCG values for the 20 query words used in the experiment. It can be seen from the figure that the effectiveness of the proposed method is still higher than that of the comparison method, which again indicates that the proposed method can ensure a strong correlation between the reordering images and the

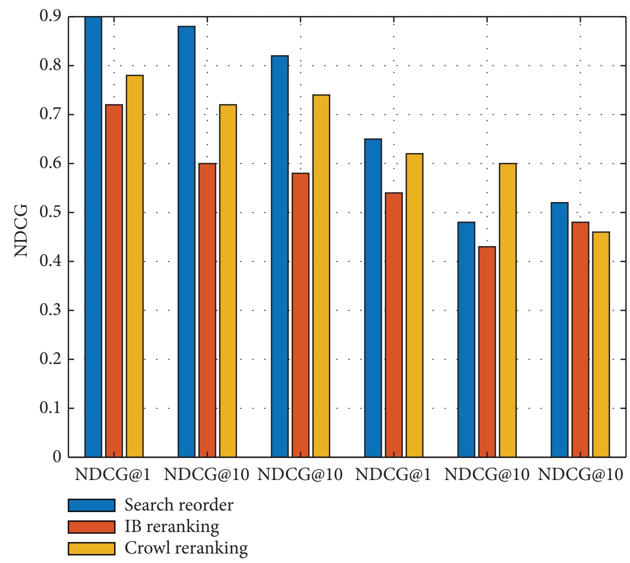


Figure 3: Comparison of reordering results.

previous ones. At the same time, it can be seen that as the value increases, the NDCG value decreases. This is because after reordering, images with low correlation are ordered in the back, thus affecting the NDCG value, which also meets the requirements of reordering.

2.2. Research on Big Data Analysis Algorithm of Tourism Route Planning. In order to carry out short-time travel route planning, this paper uses the basic idea map solution such as short-time travel route planning algorithm. The detailed solution steps are as follows:

(1) First, the user enters six input values according to his/her own time arrangement and travel preference: departure date, duration of visit, departure location, must-see scenic spots, category of scenic spots of interest, and transportation mode
(2) Secondly, use the input values in the previous step and combine with the machine learning regression algorithm to carry out the training of short-time travel route scoring model
(3) Then, according to the input value of tourists, the routes in the short-time tour route library are conditionally screened and the routes in the screening results are scored using the short-time tour route scoring model
(4) Finally, sort the scores in the previous step and output the top 10 tourist routes with the highest scores

In view of the shortcomings of ID3 algorithm, our solution idea is to record the number of records satisfying the condition of the path from the node to the root node while generating the node of the decision tree, so as to solve the predictive ability problem of the decision tree in the case of missing attribute values.

The data processing capacity of the ID3 algorithm is reduced by introducing minimum support and minimum confidence. In fact, the training sample data set have many rules, but not all of the rules has practical application value, therefore, through the association rules in the analysis of the concept of minimum support and minimum confidence, introducing part will not often appear data filtering for the rest of the operating data of ID3 algorithm.

Rules extracted by the ID3 algorithm can solve the type of data object attribute values are given identification problem, but if not all, of the given attribute values, then the rule is obtained by decision tree will not be able to give the judgment result, and we are through at the node of decision tree, at the same time record the content from the node to the path of the root node of the condition number of records, can effectively solve the missing attribute value of the decision tree's ability to predict:

(1) The support degree of decision attribute to category identification attribute is used to reduce the scale of training set processed by the ID3 algorithm

    In the process of data processing, the results often have a lot of redundancy. The most important manifestation is that in the generated decision tree, the decision tree has too many branches, which makes the tree too large and too cumbersome. In this way, the decision information obtained will be too complex and complicated, and many unnecessary rules will be generated. The minimum support degree is used to effectively control the number of possible branches of decision nodes layer by layer.

    When the support degree is less than the specified minimum support decision makers in the tuple filter, its basic idea is the decision tree branch. If the current layer corresponds to a subset of a tuple of the category attribute support small, which is the next layer of support smaller, the corresponding tuples in this group do not need to generate a new branch. It can be seen that in the process of decision tree generation, the pruning of branches that may be generated but have less practical value is carried out according to the value of each decision attribute. Since the size of the child value to some extent represents the amount of effective information contained in the attribute, our processing work is to a large extent based on the amount of information contained in the decision attribute. Since the information contained in the discarded data is relatively small, the above processing ensures the information content of the data to a certain extent.

(2) The redundant branches of the decision tree generated by the minimum confidence level scavenging D3 algorithm

By using the minimum support degree, we reduce the amount of data processing in the process of decision tree generation. However, in the construction of decision tree, the knowledge contained in some branches of the decision tree is too unreliable to have application value. We use the minimum confidence to cut off the branches with low confidence in the decision tree generated by the ID3 algorithm, so as to reduce the scale of the decision tree and make the generated decision tree more practical.

We combine the above two improvements on the ID3 algorithm in one processing process to obtain the following ID3 improved algorithm, as shown in Algorithm 1.

3. Big Data Analysis and Forecast of Tourism Information

Since the per capita ecological footprint reflects the amount of resources consumed by a single individual in an area, the per capita ecological carrying capacity represents the sum of the amount of resources that the area can provide to each individual and the amount of resources needed to deal with the waste it produces. Therefore, from the perspective of regional development, per capita ecological deficit and per capita ecological surplus appear to be very critical and important. In this paper, the time series prediction of the calculated per capita is carried out by using the ARIMA model to explore the regional sustainable development capacity allocated to each individual in the context of rapid regional development and increasingly strong tourism trend. In order to enhance the accuracy, science and rationality of the ARIMA prediction model, input data sets were divided into the training set (80%) and test set (20%).

The ADF test results are shown in Table 2, where the ADF test result is −0.656688, which is greater than the threshold value of significance level from 1% to 10% given. Therefore, the null hypothesis of the existence of unit root is accepted, which further verifies the instability of 1nXt sequence. Therefore, a first-order difference is made for 1nXt to check its trend characteristics, as shown in Figure 4. It can be seen from Figure 4 that the first-order difference series has no trend characteristics and the time trend is basically eliminated, so it can be considered as a stationary series. The unit root test was carried out on D (1nXt), and the results showed that the null hypothesis of unit root was rejected at the significance level of 0.01, indicating that the first-order difference sequence was stationary series. The test results are shown in Table 3. Therefore, $d = 1$.

Figure 5 shows a block diagram of a tourist attraction reordering method based on social media heterogeneous information fusion. First, heterogeneous travel information was taken from social media, building a database by scraping user reviews and ratings from Tripadvisor, images from Flickr, and site descriptions from Wikitravel. Then, data preprocessing is carried out, denoising heterogeneous information and structural analysis. Secondly, when positioning city is given, users input query words according to their own needs, and the initial ranking results will be given according to text information matching. Then, by analyzing the heterogeneous information of tourism in social media, the reordering framework based on the graph model is used to reorder the features of text and image fusion of scenic spots, and finally, the final ranking of scenic spots is carried out based on the numerical information.

4. Results Analysis

First, experiments are designed to discuss the influence of content-based operations and vision-based operations on the construction of video traceability relationship. Among them, content-based operation detector can be used to find the video of content change, while the visual-based operation detector can be used to find the video of visual perception change from the perspective of human visual perception. For the time-scale transformation detector, the duration statistics are introduced to pseudoclassify similar videos, which can improve the detection speed and accuracy. Figure 6 shows the use of duration in time-scale transformation detection. Known videos (a) and (b) are divided into a group according to the duration of the video, and the detection results show that the video (b) has additional shots. Since videos (a) and (b) have the same grouping in a group of similar videos, they are considered to be the same version, so there is no need to compare them with other groups of videos to save detection time. And this will not detect a slightly longer video (b) as an added shot. Therefore, without the reference of original video, it is very necessary to detect its preclassification.

In Figure 7, the recall rate and precision of the content-based detector are plotted. Here, five thresholds are given for each detector to observe the change of recall rate and precision. Similarly, it can be seen that recall rate and precision rate tend to change with the change of threshold, and in the actual detection of video traceability relationship, the threshold value can be set according to the demand for recall rate or precision rate. There is still room for improvement in the recall and precision of spatial information coverage detection and time scale transformation detection. In the time-scale transformation detection, the matching algorithm of similar video frames can be improved and the

```
Input: R:a set of noncategorical attributes
  D: the categorical
  T: training set
Output: a decision tree
Begin
  If T is null
    Return empty flags or single data point flags
If the records in T all have the same classification mark
    Return the classification value with a single node flag
Assign threshold and confidence
For all attribute X in D
    Calculate the obtained value of (x, T)
    Let g be the maximum value obtained by (x, T)
      Let x = Xi
  Let W be the attribute with the greatest gain
      Calculate the percentage of each class in the data set and the subset of different values of the decision attribute W;
  If the percent value < x then
    Not counted in the queue
  End if
End
```

Algorithm 1: D3 NEW AND ID3 IMPROVED ALGORITHM.

Table 2: Results of the 1nXt unit root test.

ADF test statistic	−0.656677	1.5% critical value	−3.808548
		5.5% critical value	−3.020683
		10.5% critical value	−2.650414
	One-sided p values		

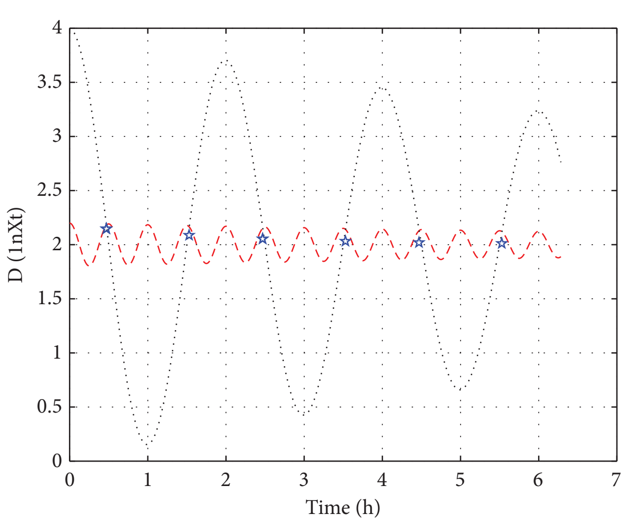


Figure 4: First-order difference sequence diagram.

matching accuracy can be improved. At the same time, in the spatial information coverage detection algorithm, when the coverage information block is very small, it is filtered as noise, resulting in the failure to detect the information coverage.

Table 4 shows the accuracy of some video detectors. For the spatial information superposition detector, it is necessary to find the appropriate threshold. Here, according to empirical values, five different thresholds are selected to draw the recall and precision of operational behavior detection. The audio conversion detector also gives a good performance value, because most audio conversions in video are full audio replacement, which can be easily detected using existing techniques. For visual based detection, including detection of spatial color change, spatial scale transformation, and visual quality detection, the detector can easily detect the change of spatial scale and spatial color, which has a good accuracy. Moreover, the detector can easily detect the change of spatial scale and spatial color, so it has a good accuracy.

The test uses cross validation to avoid the overfitting of the model. The test data set is divided into 10 parts, one part of which is taken out as the test set and the other nine parts are the training set. The average of the ten test scores is finally output as the final score of the model.

As can be seen from Figure 8, due to the different styles of scenic spots in different cities, the precision and recall rates of the same query words are different. The results of all reordering methods are better than that of the reference method TF/IDF. The results of reordering based on multiple hidden features and reordering based on heterogeneous information fusion are better than that of single modal information, which indicates that the reordering result of multimodal fusion is better than that of single modal information in the reordering problem setting in this chapter. The results of reordering based on the characteristics of multiple hidden topics are good because recall and precision are mainly used to examine content-based relevance. At the same time, the reordering result of multihidden theme features is superior to the reordering result of single mode text or image and their mean value, which means that multihidden theme features not only integrate the important information of text and visual features, but also excavate the potential information between the two kinds of information. Heterogeneous information fusion based on social media sites reordering mainly combines the hidden theme features

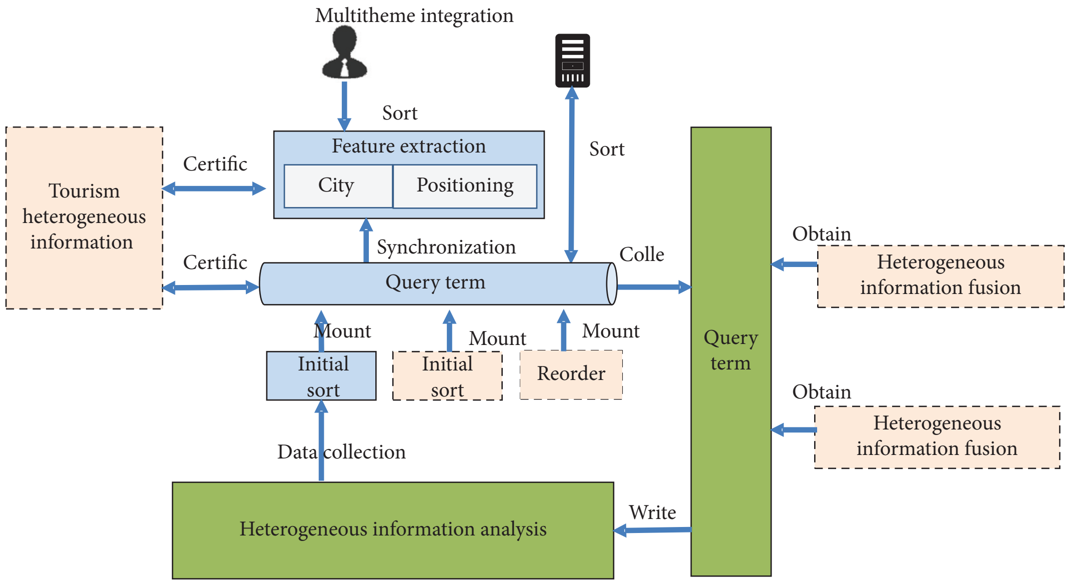
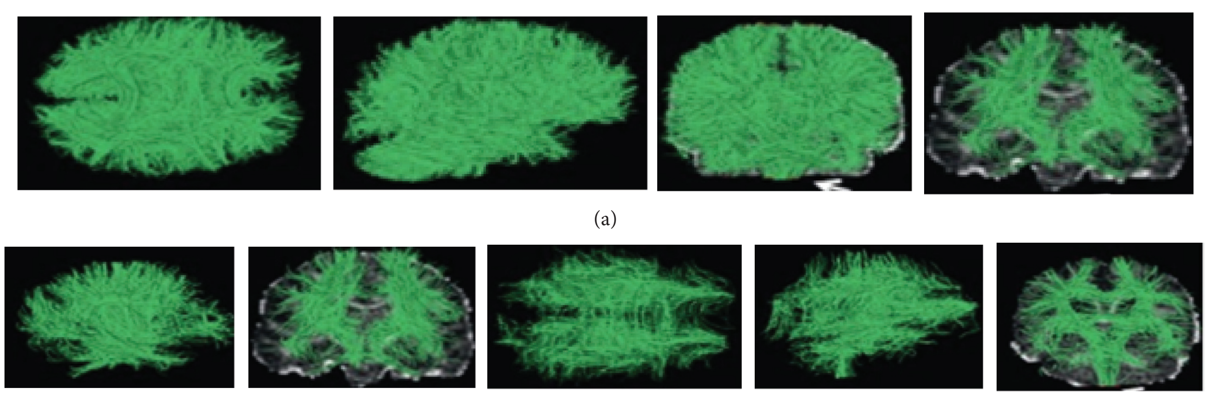


Figure 5: Reordering block diagram of tourist attractions based on heterogeneous information fusion.



(a)

(b)

Figure 6: Time-scale transformation detection: (a) duration of 65 seconds; (b) duration of 155 seconds.

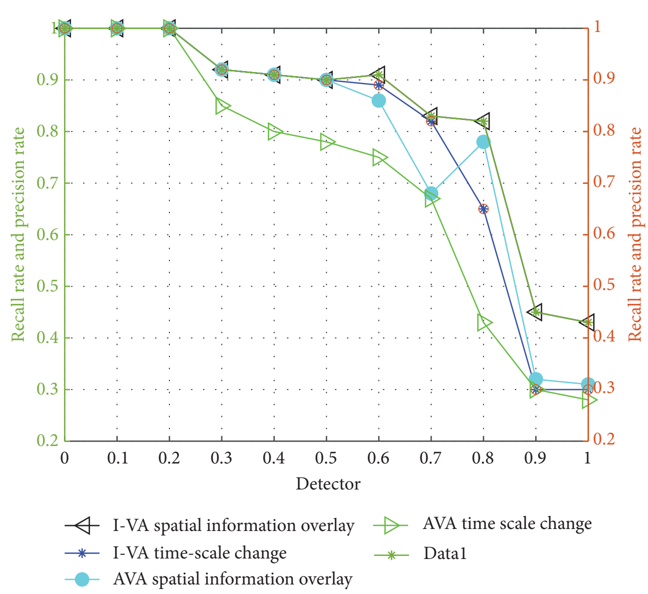


Figure 7: Graph of recall and precision of content-based detectors.

Table 3: Unit root test results.

ADF test statistic	−0.686677	1.5% critical value	−3.82832
		5.5% critical value	−3.029683
		10.5% critical value	−2.655418
	One-sided p values		

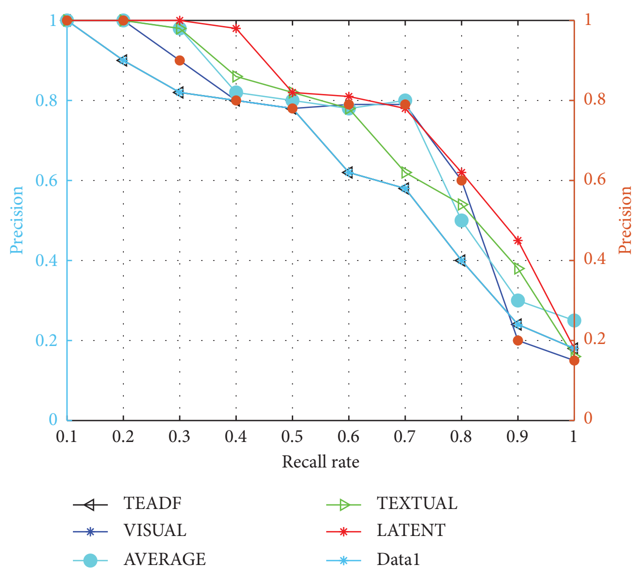


Figure 8: Precision and recall curves of all methods.

Table 4: The recall rate and precision rate of the detection subset of user operation behavior.

	Audio conversion (%)	Space scale (%)	Visual quality (%)
Precision	99.2	78.4	99.2
Recall	97.3	99.5	95.6

Table 5: Statistics of travel data of real and reliable users in different cities.

City	Number of users	
	Total	Number of tourist attractions over 5
Beijing	4255	326
London	16953	1258
New York	16172	1303
In Paris	14466	1392
Singapore	11578	503
Xi'an	3457	162

and user ratings reorder, which can be seen from the diagram. Sometimes, more implicit characteristics of reorder results are better than the result of the heterogeneous information fusion. Because the recall and precision of this experiment are related to the content, the result of reordering based on multiple hidden features is better than the final reordering based on heterogeneous information fusion.

Table 5 gives a summary of the real and reliable user travel data obtained in different cities. In this experiment, 80 users were selected from each city to test the proposed model. For each user, that is, the user has been to the scenic spots and has not been to the scenic spots. In the scenic spots that the user has been to, the scenic spots are classified into two categories, one as marked scenic spots, and the other as unmarked scenic spots mixed into the scenic spots that have not been to. Similarly, in order to verify accuracy, the comparison methods are PR and U-CF to compare with our PAS model. Here, U-CF helps users find people who have been to the same scenic spots as them, and similar users who have been to other scenic spots are considered as the scenic spots that this user may go to, and are recommended to him. The percentage of the recommended data that contains the sites the user has visited is used as the accuracy of the real data authentication. So, precision N is the proportion of positive samples are evaluated in the first N data. The evaluation criterion is the proportion of the first N recommended results containing samples visited by users, and the value of PrecisionL0 is calculated, where the number of marked positive samples varies from 1 to 4.

5. Conclusion

For the first time, the collective wisdom that exists in travel information has been harnessed to personalize recommendations of attractions, as social media travel information contains a wealth of experiences for travelers. In order to make full use of the collective wisdom in the Internet, structured knowledge can be extracted from the collective wisdom. In order to solve the problem of sparsity and diversity of tourism data, the preferred scenic spots in the current city are obtained by explicit feedback from users. Combining with the attractions of collective wisdom and user feedback, similar classification problem to solve the problem of recommended is established, which can adaptively adjust the weights of multimodal information, get similar sites, and interact with the user. Then, combined with the user positioning information in one of the user contexts, the candidate recommended scenic spots are screened again, and the personalized scenic spots are recommended to the user finally, so as to complete the personalized recommendation of tourist scenic spots based on collective wisdom. The long-time series early warning and prediction method of big data are used to evaluate the sustainable development ability of the research area in the future. The ARIMA model was used to predict the per capita ecological surplus. The results show that the per capita ecological surplus of the research area will decrease year by year in the next 10 years; that is, the local natural ecological resources are more and more difficult to meet the needs of human society and economic development. Aiming at the task of classification and discovery of big data, the ID3 algorithm, the basic method of the travel recommendation decision tree algorithm, is studied, and it has been improved from two aspects to improve the effectiveness of the algorithm, reduce the amount of data processing, and enhance the prediction ability. The practical application proves the effectiveness of the algorithm. The preprocessing methods and methods adopted by the data preprocessing module are analyzed in detail, and the algorithms used in the travel route analysis and consumer spending ability analysis are described in detail. The data of tourism consumers are analyzed by this system, and the results are evaluated.

Acknowledgments

This work was supported by the Philosophy and Social Science Planning Research General Subject of Hainan Province, the research of the influence mechanism of Hainan tourists' three-degree cognition of online short-term rent platform, grant no. HNSK(YB)18-96, and the Higher Education Scientific Research General Project of Hainan Province, the research of the adoption behavior of Hainan tourists on short-term rent platform in the perspective of the whole region tourism, grant no. Hnky2018-103.

References

[1] G. N. S. Navneet, "A novel algorithm for big data classification based on lion optimization," *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 7, pp. 1525–1532, 2017.

[2] W. Daopeng, F. Jifei, and F. Hanliang, "Research on optimization of big data construction engineering quality management based on RNN-LSTM," *Complexity*, vol. 2018, Article ID 9691868, 16 pages, 2018.

[3] S. Zhai, J. Duan, and X. Ai, "Research on the third party logistics system and economic performance optimization based on big data analysis," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 11, pp. 301–308, 2017.

[4] S. Chakraborti and S. Dey, "Analysis of competitor intelligence in the era of big data: an integrated system using text summarization based on global optimization," *Business & Information Systems Engineering*, vol. 61, no. 3, pp. 345–355, 2019.

[5] D. Li, Z. L. Deng, and Z. Cai, "Statistical analysis of tourist flow in tourist spots based on big data platform and DA-HKRVM algorithms," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 87–101, 2020.

[6] G. Stefanos, L. Moritz, G. Tais, S. Vanhuysse, B. Johnson, and E. Wolff, "Normalization in unsupervised segmentation parameter optimization: a solution based on local regression trend analysis," *Remote Sensing*, vol. 10, no. 2, pp. 222–234, 2018.

[7] J.-Y. Park, G. Kim, and H.-J. Oh, "A study on tourism resource strategy of film location using social bigdata based on SNS trend analysis of jeonju area," *The Journal of the Korea Contents Association*, vol. 16, no. 11, pp. 477–487, 2016.

[8] X. Gao and J. Cai, "Optimization analysis of urban function regional planning based on big data and GIS technology," *Boletin Tecnico/technical Bulletin*, vol. 55, no. 11, pp. 344–351, 2017.

[9] Y. Wu, T. Z. Xia, R. YuYi, and J. Wang, "Characteristics and optimization of core local network: big data analysis of football matches," *Chaos, Solitons & Fractals*, vol. 138, pp. 110136–110145, 2020.

[10] H. Zuo, "Development and robust testing of big data analysis system based on optimized Java technique and parallel oracle database," *Revista de la Facultad de Ingenieria*, vol. 32, no. 2, pp. 180–189, 2017.

[11] Y. Chen, X. J. Luo, and X. Yao, "A medical big data analysis algorithm based on access control system," *International Journal of Reasoning-Based Intelligent Systems*, vol. 10, no. 1, pp. 51–63, 2018.

[12] C. Lu and F. Wen, "GIS information feature estimation algorithm based on big data," *International Journal of Information and Communication Technology*, vol. 15, no. 2, pp. 198–213, 2019.

[13] X. Xia, "Fast search of art culture resources based on big data and cuckoo algorithm," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 127–138, 2020.

[14] H. Shin, "Analysis of subway passenger flow for a smarter city: knowledge extraction from seoul metro's "untraceable" big data," *IEEE Access*, vol. 8, pp. 69296–69310, 2020.

[15] J. Al-Sawwa and S. A. Ludwig, "Parallel particle swarm optimization classification algorithm variant implemented with Apache Spark," *Concurrency Practice & Experience*, vol. 32, no. 2, 2020.

[16] M. Tang and Z. Wu, "Research on the mechanisms of big data on consumer behavior using the models of C2C e-commerce and countermeasures," *African Journal of Business Management*, vol. 9, no. 1, pp. 18–34, 2015.

[17] W. Wang, "Deployment and optimization of wireless network node deployment and optimization in smart cities," *Computer Communications*, vol. 155, no. 4, pp. 117–124, 2020.

[18] R. Xu, L. Z. Wen, B. LiLu, and X. Wang, "Ensemble with estimation: seeking for optimization in class noisy data," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 2, pp. 231–248, 2020.

[19] X. Gao and Y. Wang, "Optimized integration of traditional folk culture based on DSOM-FCM," *Personal and Ubiquitous Computing*, vol. 24, no. 2, pp. 273–286, 2020.

[20] H. Guo, L. J. Liu, and L. Zhao, "Big data acquisition under failures in FiWi enhanced smart grid," *IEEE Transactions on Emerging Topics in Computing*, vol. 7, no. 3, pp. 420–432, 2019.

[21] B. D. Rouhani, A. Mirhoseini, E. M. Songhori et al., "Automated real-time analysis of streaming big and dense data on reconfigurable platforms," *ACM Transactions on Reconfigurable Technology and Systems*, vol. 10, no. 1, pp. 8.1–8.22, 2016.

[22] Y. Uematsu, K. Y. Fan, and W. LinLv, "SOFAR: large-scale Association network learning," *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 4924–4939, 2019.

[23] J. Fan, K. H. Song, F. YanLiu, and W. Lian, "Real-time manifold regularized context-aware correlation tracking," *Frontiers of Computer Science*, vol. 14, no. 2, pp. 334–348, 2020.

[24] S. Sharma, J. LuBmann, and J. So, "Controller independent software-in-the-loop approach to evaluate rule-based traffic signal retiming strategy by utilizing floating car data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 9, pp. 3585–3594, 2019.

[25] Y. Li, Y. J. Bao, Z. GongWu, and Y. Zheng, "Mining the most influential -location set from massive trajectories," *IEEE Transactions on Big Data*, vol. 4, no. 4, pp. 556–570, 2018.

[26] X. Xu, F. R. Gu, and S. WanQi, “Multi-objective computation offloading for Internet of Vehicles in cloud-edge computing,” *Wireless Networks*, vol. 26, no. 3, pp. 1611–1629, 2020.
[27] T. Kim, W. Li, A. Behm et al., “Similarity query support in big data management systems,” *Information Systems*, vol. 88, no. Feb., pp. 101455.1–101455.23, 2020.
[28] S. Knoch, S. Herbig, F. Kosmalla et al., “Sensor-based human-process interaction in discrete manufacturing,” *Journal on Data Semantics*, vol. 9, no. 1, pp. 21–37, 2020.
[29] X. Zhao, W. Meng, and J. Su, “Parameter optimization of single sample virtually expanded method,” *The International Arab Journal of Information Technology*, vol. 16, no. 6, pp. 988–994, 2019.

5

Research on the Evaluation Model of Urban Tourism Management Efficiency with Uncertain Linguistic Information

Song-Mao Wang,[1] Liang-Yan Fang,[2] and Feng Deng[1]

[1] *Xinjiang University, Urumqi, Xinjiang 830047, China*
[2] *Urumqi Vocational University, Urumqi, Xinjiang 830047, China*

Correspondence should be addressed to Song-Mao Wang; wsm0316@sohu.com

Academic Editor: Xiaofei Zhao

We investigate the multiple attribute decision making problems for evaluating the urban tourism management efficiency with uncertain linguistic information. We utilize the uncertain linguistic weighted averaging (ULWA) operator to aggregate the uncertain linguistic information corresponding to each alternative and get the overall value of the alternatives and, then rank the alternatives and select the most desirable one(s). Finally, a numerical example for evaluating the urban tourism management efficiency with uncertain linguistic information is used to illustrate the proposed model.

1. Introduction

China's tourism industry has experienced the stages of formation, rapid growth, and steady development since 1978 and now has developed into one of the fastest growing industries and a new growth point of China's national economy. As the core and foundation of the development of China's tourism industry, urban tourism has made an important contribution to the rapid prosperity and growth of China's tourism industry in the past 35 years [1, 2]. The industrial status of tourism industry in the urban economy has risen to become a key industry and one of the pillar industries of urban tertiary industry and has also become the core of urban vigorous development of tertiary industry so as to optimize the urban industrial structure [3, 4]. So in-depth study of the development of urban tourism industry is not only the objective requirement of promoting the development of china's tourism industry but also the reality need of optimizing the allocation of the tourism system so as to optimize the urban industrial structure and promote the urban economic growth. Although the overall urban tourism industry has shown a scene of prosperity and development, we are concerned about the differences in the development of urban tourism industry, especially the regional differences in tourism industry of the eastern, central, and western cities showing the different characteristics of the development [5, 6]. The essence of these different characteristics is the macroperformance under the influence of microscopic mechanism of the development of the tourism industry. Therefore, we can analyse the microinfluencing factors under the macroscopic characteristics of the development of urban tourism industry in order to study the breakthrough point to grasp, refine, and identify the mode of development of urban tourism of the growth center of eastern, central, and western regions. To make out what factors are the key factors which contribute to the urban development of the tourism industry and what factors are the restrictive factors which hinder the development of the tourism industry in different development modes [7, 8]. Distinguishing the mode of development of urban tourism with a breakthrough point of its factors and refining and summarizing the characteristics of the development paths and effects of urban tourism industry in the growth centers of eastern, central, and western cities will play a demonstration effect for the other cities with similar industrial development conditions in the eastern, central, and western regions and provide reference of development mode and development path selection for other cities to improve the tourism industry system and enhance and optimize the structure of tourism industry [9, 10].

In this paper, we investigate the multiple attribute decision making problems [11, 12] for evaluating the urban tourism management efficiency with uncertain linguistic information. We utilize the uncertain linguistic weighted averaging (ULWA) operator to aggregate the uncertain linguistic information corresponding to each alternative and get the overall value of the alternatives and then rank the alternatives and select the most desirable one(s). Finally, a numerical example for evaluating the urban tourism management efficiency with uncertain linguistic information is used to illustrate the proposed model.

2. Preliminaries

Let $S = \{s_i \mid i = 1, 2, \ldots, t\}$ be a linguistic term set with odd cardinality. Any label, s_i, represents a possible value for a linguistic variable, and it should satisfy the following characteristics [13, 14]: (1) the set is ordered $s_i > s_j$, if $i > j$; (2) there is the negation operator $\operatorname{neg}(s_i) = s_j$ such that $j = t + 1 - i$. For example, S can be defined as

$$S = \{s_1 = \text{extremely poor}, s_2 = \text{very poor}, s_3 = \text{poor}, s_4 = \text{medium}, s_5 = \text{good}, s_6 = \text{very good}, s_7 = \text{extremely good}\}. \quad (1)$$

Definition 1 (see [15]). Let ULWA : $\widetilde{S}^n \rightarrow \widetilde{S}$; if

$$\mathrm{ULWA}_\omega(\widetilde{s}_1, \widetilde{s}_2, \ldots, \widetilde{s}_n) = \omega_1 \times \widetilde{s}_1 \oplus \omega_2 \times \widetilde{s}_2 \oplus \cdots \omega_n \times \widetilde{s}_n, \quad (2)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ is the weighting vector of uncertain linguistic variables $\widetilde{s}_i$ ($\widetilde{s}_i \in \widetilde{S}$, $i = 1, 2, \ldots, n$) with $\omega_i \in [0, 1]$, $\sum_{i=1}^{n} \omega_i = 1$, then the function ULWA is called the uncertain linguistic weighted averaging (ULWA) operator of dimension n.

Definition 2 (see [15]). Let $\widetilde{s}_1 = [s_{\alpha_1}, s_{\beta_1}]$ and $\widetilde{s}_2 = [s_{\alpha_2}, s_{\beta_2}]$ be two uncertain linguistic variables, and let $\operatorname{len}(\widetilde{s}_1) = \beta_1 - \alpha_1$ and $\operatorname{len}(\widetilde{s}_2) = \beta_2 - \alpha_2$, and then the degree of possibility of $\widetilde{s}_1 \geq \widetilde{s}_2$ is defined as [14]

$$p(\widetilde{s}_1 \geq \widetilde{s}_2) = \frac{\max(0, \operatorname{len}(\widetilde{s}_1) + \operatorname{len}(\widetilde{s}_2) - \max(\beta_2 - \alpha_1, 0))}{\operatorname{len}(\widetilde{s}_1) + \operatorname{len}(\widetilde{s}_2)}. \quad (3)$$

3. Research on the Evaluation Model of Urban Tourism Management Efficiency with Uncertain Linguistic Information

In this section, we investigate the multiple attribute decision making problems for evaluating the urban tourism management efficiency with uncertain linguistic information. Let $A = \{A_1, A_2, \ldots, A_m\}$ be a discrete set of alternatives and $G = \{G_1, G_2, \ldots, G_n\}$ the set of attributes; $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ is the exponential weighting vector of the attributes G_j ($j = 1, 2, \ldots, n$), where $\omega_j \in [0, 1]$, $\sum_{j=1}^{n} \omega_j = 1$. Suppose that $\widetilde{R} = (\widetilde{r}_{ij})$ is the decision matrix, where $\widetilde{r}_{ij} \in \widetilde{S}$ are the uncertain linguistic variables.

In the following, we apply the ULWA operator to MADM forevaluating the urban tourism management efficiency with uncertain linguistic information.

Step 1. Utilize the decision information given in matrix $\widetilde{R}$ and the ULWA operator

$$\widetilde{r}_i = \mathrm{ULWA}_\omega(\widetilde{r}_{i1}, \widetilde{r}_{i2}, \ldots, \widetilde{r}_{in}), \quad i = 1, 2, \ldots, m, \quad (4)$$

to derive the collective overall preference values $\widetilde{r}_i$ ($i = 1, 2, \ldots, m$) of the alternative A_i, where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ is the weighting vector of the attributes.

Step 2. To rank these collective overall preference values $\widetilde{r}_i$ ($i = 1, 2, \ldots, m$), we first compare each $\widetilde{r}_i$ with all the $\widetilde{r}_j$ ($j = 1, 2, \ldots, m$) by using (2). For simplicity, we let $p_{ij} = p(\widetilde{r}_i \geq \widetilde{r}_j)$, and then we develop a complementary matrix as $P = (p_{ij})_{m \times m}$, where

$$p_{ij} \geq 0, \quad p_{ij} + p_{ji} = 1, \quad hp_{ii} = 0.5, \quad i, j = 1, 2, \ldots, m. \quad (5)$$

Summing all the elements in each line of matrix P, we have

$$p_i = \sum_{j=1}^{m} p_{ij}, \quad i = 1, 2, \ldots, m. \quad (6)$$

Step 3. Rank all the alternatives A_i ($i = 1, 2, \ldots, m$) and select the best one(s) in accordance with p_i ($i = 1, 2, \ldots, m$).

Step 4. End.

4. Illustrative Example

With the rapid development of world tourism, the phenomena of cities “tourlized” and urbanization of tour destination are becoming more and more prominent. All these promote the function and facilities of cities to cater for a urban tourist in the course of urban plan and construction, which meet the requirement of tourism and meanwhile form the convergence and development of tourist economic elements. As the carrier of tourist development, tourist urban management is the key problem to be solved in the urban development of our country, and it would bring significant influence to the development of economy and society of our country. In this section, we present an empirical case study of evaluating the urban tourism management efficiency with uncertain linguistic information. There is a panel with five possible urban tourism cities A_i ($i = 1, 2, 3, 4, 5$) to select. The team of experts must take a decision according to the following four attributes: (1) G_1 is the informatization level; (2) G_2 is the financial level; (3) G_3 is the service level; (4) G_4 is the business level. The five possible urban tourism cities A_i ($i = 1, 2, 3, 4, 5$) are to be evaluated using the linguistic term set S by the decision makers under the above four attributes (whose weighting vector $\omega = (0.20, 0.30, 0.40, 0.10)^T$) and construct, respectively, the decision matrix as follows: $\widetilde{R} = (\widetilde{r}_{ij})_{5\times 4}$ and

$$\widetilde{R} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array} \begin{array}{c} \begin{array}{cccc} G_1 & G_2 & G_3 & G_4 \end{array} \\ \begin{pmatrix} [s_4, s_6] & [s_5, s_6] & [s_3, s_4] & [s_2, s_4] \\ [s_2, s_4] & [s_2, s_3] & [s_4, s_5] & [s_4, s_5] \\ [s_4, s_5] & [s_3, s_5] & [s_1, s_2] & [s_5, s_6] \\ [s_5, s_6] & [s_4, s_5] & [s_4, s_6] & [s_2, s_3] \\ [s_3, s_4] & [s_5, s_6] & [s_3, s_5] & [s_5, s_6] \end{pmatrix} \end{array} \quad (7)$$

In the following, we apply the ULWA operator to MADM forevaluating the urban tourism management efficiency with uncertain linguistic information. To get the most desirable urban tourism cities, the following steps are involved.

Step 1. Utilizing the ULWA operator, we obtain the collective overall preference values $\widetilde{r}_i$ of the urban tourism cities A_i $(i = 1, 2, 3, 4, 5)$:

$$\widetilde{r}_1 = [s_{3.21}, s_{5.18}], \quad \widetilde{r}_2 = [s_{3.74}, s_{4.82}], \quad \widetilde{r}_3 = [s_{4.72}, s_{5.46}],$$
$$\widetilde{r}_4 = [s_{3.36}, s_{5.57}], \quad \widetilde{r}_5 = [s_{4.62}, s_{6.03}]. \quad (8)$$

Step 2. Rank these overall preference values $\widetilde{r}_i$ $(i = 1, 2, 3, 4, 5)$; we first compare each $\widetilde{r}_i$ with all the $\widetilde{r}_j$ $(j = 1, 2, 3, 4, 5)$ by using (3) and then develop a complementary matrix:

$$P = \begin{bmatrix} 0.500 & 0.720 & 0.180 & 0.459 & 0.132 \\ 0.280 & 0.500 & 0.010 & 0.350 & 0.390 \\ 0.820 & 0.990 & 0.500 & 0.758 & 0.450 \\ 0.541 & 0.650 & 0.242 & 0.500 & 0.222 \\ 0.868 & 0.610 & 0.550 & 0.778 & 0.500 \end{bmatrix}. \quad (9)$$

Summing all the elements in each line of matrix P, we have

$$p_1 = 2.037, \quad p_2 = 1.108, \quad p_3 = 3.564,$$
$$p_4 = 2.246, \quad p_5 = 3.663. \quad (10)$$

Step 3. Rank all the urban tourism cities A_i $(i = 1, 2, 3, 4, 5)$ in accordance with the values p_i $(i = 1, 2, 3, 4, 5)$: $A_5 \succ A_3 \succ A_4 \succ A_1 \succ A_2$. Thus, the most desirable urban tourism city is A_5.

5. Conclusions

With the rapid development of tourism and city, the development of tourism in promoting the city has also brought a series of problems: the integration of tourism resources among the cities is more difficult, lacking of urban tourism regional planning, the chaos of urban tourism functions and other features, tourism functional area is difficult to develop and so on, and the internal region has also suffered more severe competitive pressures. How to cultivate and improve the tourist city of travel services functional areas, in order to improve tourism services, and better development of urban tourism are becoming the most urgent problems to be solved in each tourist city currently. In this paper, we investigate the multiple attribute decision making problems for evaluating the urban tourism management efficiency with uncertain linguistic information. We utilize the uncertain linguistic weighted averaging (ULWA) operator to aggregate the uncertain linguistic information corresponding to each alternative and get the overall value of the alternatives and then rank the alternatives and select the most desirable one(s). Finally, a numerical example for evaluating the urban tourism management efficiency with uncertain linguistic information is used to illustrate the proposed model.

Acknowledgment

The work was supported by the Xinjiang Social Science Fund (13BGL035).

References

[1] X. Wang, "Model for tourism management with 2-tuple linguistic information," *Advances in Information Sciences and Service Sciences*, vol. 3, no. 4, pp. 34–39, 2011.

[2] C. Yan, "An improved city tourism efficiency evaluation model based on data envelopment analysis," *Journal of Digital Content Technology and Its Applications*, vol. 7, no. 2, pp. 343–349, 2013.

[3] H. U. Xiaotao, "The application of GIS technology in tourism resources management," *JDCTA*, vol. 7, no. 3, pp. 308–314, 2013.

[4] G. Xiong and M. Jiang, "The analysis and evaluation on tourism competitiveness of Poyang lake ecological economic zone of china based on TOPSIS method," *International Journal of Advanced Computer Technology*, vol. 5, no. 4, pp. 720–726, 2013.

[5] W. Min, "Supply chains for tourism enterprises management based on information management system," *International Journal of Advancements in Computing Technology*, vol. 5, no. 5, pp. 1000–1010, 2013.

[6] G. D. Zhuang and H. Zhang, "Innovation strategies and technology for tourism industry based on case-of-tourism theory," *AISS*, vol. 5, no. 5, pp. 97–105, 2013.

[7] Y. C. Dong, W. C. Hong, Y. F. Xu, and S. Yu, "Selecting the individual numerical scale and prioritization method in the analytic hierarchy process: a 2-Tuple fuzzy linguistic approach," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 1, pp. 13–25, 2011.

[8] Z.-P. Fan, B. Feng, and W.-L. Suo, "A fuzzy linguistic method for evaluating collaboration satisfaction of NPD team using mutual-evaluation information," *International Journal of Production Economics*, vol. 122, no. 2, pp. 547–557, 2009.

[9] Z. P. Fan and Y. Liu, "A method for group decision-making based on multi-granularity uncertain linguistic information," *Expert Systems with Applications*, vol. 37, no. 5, pp. 4000–4008, 2010.

[10] Y. Dong, Y. Xu, and H. Li, "On consistency measures of linguistic preference relations," *European Journal of Operational Research*, vol. 189, no. 2, pp. 430–444, 2008.

[11] Y. C. Dong, Y. F. Xu, H. Y. Li, and B. Feng, "The OWA-based consensus operator under linguistic representation models using position indexes," *European Journal of Operational Research*, vol. 203, no. 2, pp. 455–463, 2010.

[12] Z. P. Fan, B. Feng, Y. H. Sun, and W. Ou, "Evaluating knowledge management capability of organizations: a fuzzy linguistic method," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3346–3354, 2009.

[13] F. Herrera and L. Martínez, "A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 31, no. 2, pp. 227–234, 2001.

[14] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 746–752, 2000.

[15] Z. Xu, "Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment," *Information Sciences*, vol. 168, no. 1–4, pp. 171–184, 2004.

6

Research on Construction of a Cloud Platform for Tourism Information Intelligent Service Based on Blockchain Technology

Cao Wei, Qinan Wang, and Chengying Liu

School of Geography Information System and Tourism, Chuzhou University, Chuzhou 239000, China

Correspondence should be addressed to Cao Wei; lcy@chzu.edu.cn

Academic Editor: Hongju Cheng

The development of global tourism has put forward new requirements for the construction of smart tourism. More and more travel-related data have reached the level of TB or even PB, which has brought great difficulties to tourism management. This article explores the use of big data technologies such as genetic algorithms to explore massive travel data and establish a comprehensive tourism information service platform for governments, enterprises, tourists, and scientific research institutions. The overall design of an industrial information service platform based on big tourist data is proposed. The overall function, data source, data standard, and application scope of the platform are all focused on. The traceability and nontampering of blockchain technology can also help passengers retain and verify their identity information. From the perspective of the design goals of the system, in general, the time required for repeated authentication will greatly reduce air ticket bookings, accommodation reservations, and bill verification, and improving efficiency is the only way to establish a "trust ecology." Architecture design, distributed architecture, and intelligent service design, as well as the key implementation technology of service platform construction, route recommendation algorithm and tourism information big data mining, research and analysis on the construction of tourism information intelligent service.

1. Introduction

At present, China has become the largest tourist market in the world [1, 2]. Among them, outbound tourism consumption ranks first in the world, and inbound tourism ranks third in the world [3]. Driven by the new Internet information technology such as mobile communication, large data service, and cloud computing, especially the proposal of "intelligent earth," the concepts of "intelligent city." "Intelligent tourism" and "intelligent service" have come into being one after another. At the same time, the rapid integration of these new Internet applications and tourism industry will also promote the development of tourism information services to intelligent services [4]. Based on the technology of large data, cloud computing, and artificial intelligence, this research is based on the users' behavior analysis and realizes the information analysis and processing of mass tourist users [5]. Through the mobile Internet and other media, real-time online interaction is carried out with users, so as to provide users with more intelligent and convenient route decision travel information clothes, Business.

In recent years, the National Tourism Administration has vigorously promoted the construction of intelligent tourism. In 2011, intelligent tourism was listed as an important work of tourism in the 12th Five-Year Plan; in 2012, 18 pilot cities of intelligent tourism and 22 pilot units of intelligent scenic spots were determined; in 2013, 14 intelligent tourist cities were determined [6]. 2014 was identified as the year of wisdom tourism. According to studies, China's tourism industry began to enter the era of big data in 2013. In the rush hour, the hotel has a daily booking capacity of several hundred thousand and, at the same time, produces massive tourism data such as information collection, consumer reviews, and product recommendations [7]. Big data technology can dig and analyze these huge amounts of data, which will make wisdom travel like a duck to water. The impact of big data on the tourism industry is all-around, and the whole industry management decision-making mode has changed accordingly [8]. In 2015, the State Council issued the "program for the promotion of big data development" and systematically deployed large data development work [9, 10]. In the same year, Gui Zhou started the construction of the first large data comprehensive test area in China. Many scholars began to study the integration of large data and intelligent tourism. In this context, the author combed the relevant literature in order to better clarify the relationship between the two and provide reference for the further development of intelligent tourism. In this paper, we show how an iterative development process that incorporates client-developer joint exploration of partial designs facilitates the development of client understanding of their needs [11, 12].

The rest of this paper is organized as follows. Section 2 discusses the concept of large data and intelligent tourism. The task joint release model based on a genetic algorithm is discussed in Section 3. Construction of an intelligent tourism application model under large data background is discussed in Section 4. Section 5 concludes the paper with a summary and future research directions.

2. The Concept of Large Data and Intelligent Tourism

The concept of big data comes with the growth of unstructured data. In 2011, the McKinsey Global Institute defined large data as a data group of data groups [13] that exceeded the grasp, storage, management, and analysis capabilities of the traditional database software tools. Processing large data requires special technology, namely, large data technology, including large-scale parallel processing database, data mining grid, distributed file system, distributed database, cloud computing platform, Internet, and extensible storage system. Internet companies represented by Google, Facebook, LinkedIn, and Microsoft have gradually introduced various types of big data processing systems. Smart tourism is a new concept that evolved from the digital earth. Mole, an assistant professor at the College of the Holy Cross, defines intelligent tourism as the use of mobile digital connectivity to create a more intelligent, meaningful, and sustainable association between tourists and the city. Domestic scholars have also formed several views on the concept of intelligent tourism [14].

Relying on the accurate and advanced information data platform, and integrating the tourism industry information, according to the travel records and consumption preference of tourists, it provides tourists with personalized tourism services and realizes the private customization of tourism. The smart tourism application model based on the big data perspective is shown in Figure 1.

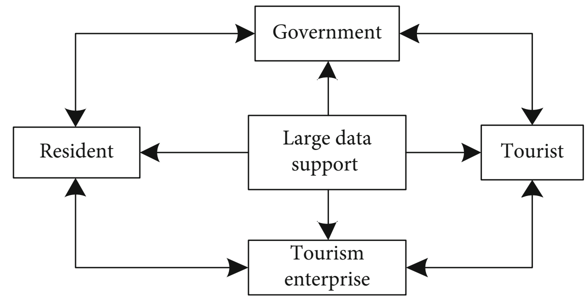


Figure 1: Smart tourism application model based on the big data perspective.

The first view is that intelligent tourism is a variety of changes brought by the application of some new technologies to tourism; the second view is that intelligent tourism is a modernization of promoting tourism service, improving tourism experience, innovating tourism management, and optimizing the use of tourism resources. Engineering; the third view is that intelligent tourism is a high-level form and stage of tourism; the fourth view is that intelligent tourism is a new mode of operation to improve the experience and participation of tourists; and there are other different views. With regard to the relationship between large data and intelligent tourism, most scholars regard big data as a background of intelligent tourism and identify the role of large data on the development of intelligent tourism. However, some scholars regard big data technology as one of the components of intelligent tourism [15, 16]. Liang Changyong believes that large data mining technology is intelligent tourism. The core. Li Yunpeng and so on pointed out that intelligent tourism is a large amount of real-time data formed by tourists on the Internet. Using big data, it can further guide the practice of tourism industry and better serve tourists [17, 18].

3. Task Joint Release Model Based on a Genetic Algorithm

3.1. Model Preparation. A genetic algorithm is a search algorithm with self-adaptability and self-organization ability developed based on the natural selection and evolution mechanism of the biological world. It is widely used to solve complex optimization problems. The genetic algorithm first randomly generates a set of possible solutions to the optimization problem and encodes each possible solution [19]. The set of possible solutions is called a population, and each possible solution in the population is called an individual. Each individual has a corresponding fitness value, which is used to measure the "good or bad" of the solution to the problem represented by the individual. Imitating the principle of "survival of the fittest" in the biological world, according to the size of fitness, select a number of better individuals from the initial population to participate in crossover and mutation operations [20, 21].

The crossover operation generates a new individual by swapping and recombining a part of the codes of the two individuals, similar to the offspring inheriting the parent's genes. The mutation operation generates new individuals by randomly changing the coding bits of an individual, thereby increasing the diversity of the population. The selection, crossover, and mutation operations are performed iteratively several times or until a specific termination rule is met, and the individual with the highest fitness in the population is the approximate optimal solution of the optimization problem. In recent years, genetic algorithms have been successfully applied in pattern recognition, machine learning, image processing, and intelligent control [22].

3.2. The Basic Idea of the K-Means Algorithm. The K-means algorithm is one of the most widely used clustering algorithms. The algorithm uses parameters to divide objects into clusters so that the clusters have a high degree of similarity, but the similarity between clusters is low. The algorithm first randomly selects objects. Each object initially represents the average value or center of a cluster [23, 24]. For each remaining object, it is assigned to the nearest cluster according to its distance from the center of each cluster, and then each object is recalculated. For the average value of the cluster, the process is repeated until the criterion function converges. The criterion function is as follows:

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} |x - \bar{x}_i|^2, \tag{1}$$

where E is the average value of the cluster. The description of the K-means algorithm is as follows:

(1) Randomly select a record as the initial cluster center

$$W(u_i, v_i) = W(i) = \begin{bmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & w_{\text{in}} \end{bmatrix}. \tag{2}$$

(2) Calculate the distance between each record and the cluster centers, and use the closest cluster as the class to which the point belongs [25]

$$\beta = \begin{pmatrix} \beta_0(u_1, v_1) & \beta_1(u_1, v_1) & \cdots & \beta_k(u_1, v_1) \\ \beta_0(u_2, v_2) & \beta_1(u_2, v_2) & \cdots & \beta_k(u_2, v_2) \\ \cdots & \cdots & \cdots & \cdots \\ \beta_0(u_n, v_n) & \beta_1(u_n, v_n) & \cdots & \beta_k(u_n, v_n) \end{pmatrix}, \tag{3}$$

where β is the estimated value, k is the number of independent variables, n is the spatial sample, and u is the weight given to the data point n when describing the model for position i.

3.3. Weight Function and Bandwidth Determination. The weight function is often determined by the Gaussian function in practice, and the expression method is as follows:

$$W_{ij} = e^{-(1/2)\left(d_{ij}/b\right)^2}, \tag{4}$$

where b refers to the bandwidth. If the data of point i is observed, the weight of other points will decrease with the increase of distance according to the Gaussian curve. Bandwidth b, then the weight of points far enough away from point i will tend to zero. The choice of bandwidth b can be determined by the cross-confirmation method and calculated for a given b value:

$$cv(b) = \sum_{i=1}^{n} \left[y_i - \hat{y}_{\neq i}(b) \right]^2, \tag{5}$$

where y is the dependent variable after the observation value and the width b is removed, and then the weighted least square method is used to obtain the fitted value of the dependent variable y. The selection result is shown in the following formula:

$$cv(b_0) = \min_{b>0} cv(b). \tag{6}$$

Calculate the centroid of each cluster (the mean value of the cluster point) and the distance between each object and these central objects, and redivide the corresponding objects according to the minimum distance. Repeat this step until equation (3) no longer changes significantly [26].

3.4. Model Establishment. One advantage of the genetic algorithm is that it does not require special knowledge of the problem domain to be solved, so the process of solving the problem with the genetic algorithm is basically the same. The following figure shows the processing flow of the genetic algorithm used in this paper for cluster analysis [27, 28].

The evolution process in genetic algorithms is based on the coding mechanism, and coding has a great influence on the performance of the algorithm, such as search capabilities [29, 30]. Commonly used encoding methods include floating-point number encoding and binary encoding. In contrast, binary encoding has stronger search capabilities than floating-point number encoding. In addition, binary encoding has the advantages of simple crossover and mutation operations. Therefore, binary encoding is used here. The solution of the clustering problem is the center of each cluster. For the clustering problem of points in space, each individual in the genetic algorithm includes a binary component corresponding to the real number component, and the number of binary digits contained in each component is calculated by the method [31, 32].

In this way, each individual includes binary bits. The first binary code string of length in the individual is decoded by the decoding function [33, 34].

Randomly select a point from the set of points to be classified as a solution of the problem and encode it. Repeat the selection of Psize times, and Psize is the set population size [35].

In the process of the genetic algorithm, the fitness function is used as the basis and the fitness value of each individual in the population is used to search [36, 37]. Therefore, the selection of fitness function is very important and directly affects the convergence speed of the genetic algorithm. For each individual, use the same method as the K-means algorithm to divide the cluster and recalculate the center of each

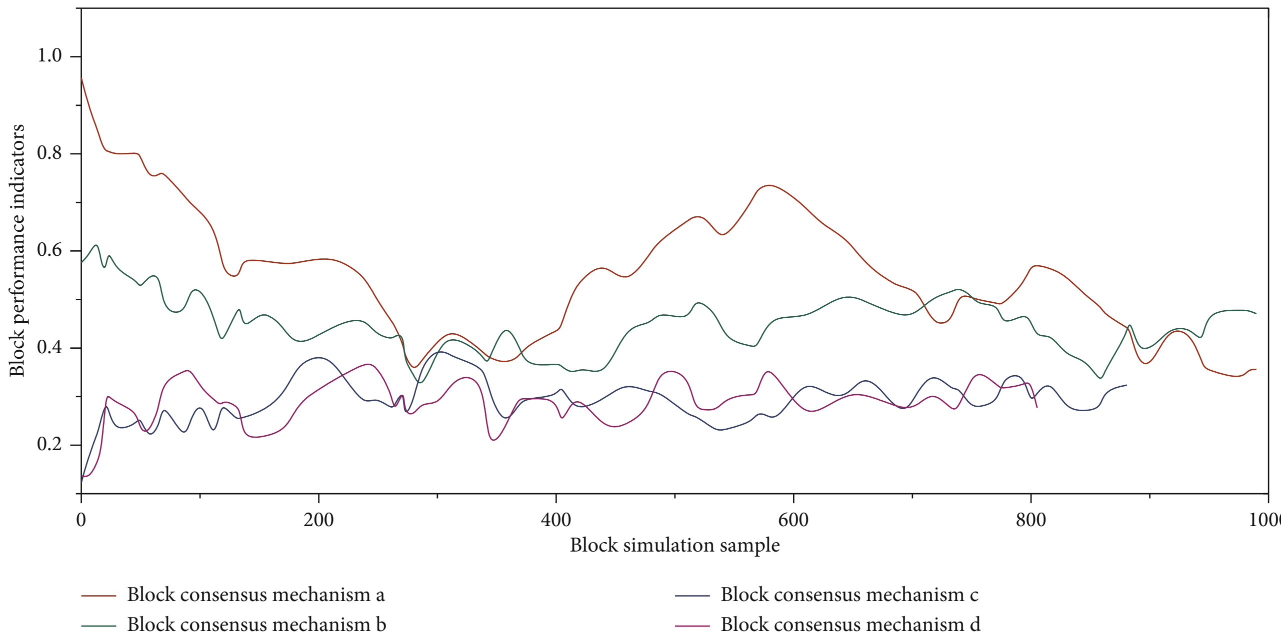


Figure 2: Changes in blockchain performance indicators under different consensus mechanisms.

cluster, and then use the sum of the distances between the points in each cluster and the corresponding cluster center as the judgment cluster division. The smaller the quality criterion function, the better the quality of clustering. The mathematical expression is [38]

$$\Delta(G_1, G_2, \cdots, G_K) = \sum_{i=1}^{K} \sum_{x_j \in G_i} \left\| x_j - c_i \right\|. \tag{7}$$

The purpose of the genetic algorithm is to search for the cluster center that minimizes the value, so the fitness function [39]. In order to improve the search speed, after each cluster division is obtained, the corrected cluster center is used to replace the original cluster center in the individual [40, 41].

4. Construction of the Intelligent Tourism Application Model under Large Data Background

The foundation of smart tourism in the context of big data is the smart tourism application model based on the big data platform. The smart tourism application model based on the large data platform can provide the latest tour route quotations, the most favorable discount ticket information, the most reasonable travel advice, and the most detailed tourist information.

In order to identify the relationship between tourism stakeholders and their relations as a breakthrough point, the intelligent tourism application model of the large data perspective is built around the application objects and the demand for mutual relations for intelligent tourism. The application objects of smart tourism mainly include government tourism department, tourism enterprise, community residents, and tourists. Compared with the traditional application model, it only faces government, tourism enterprises, and tourists. Smart tourism also includes community residents in the application of smart tourism. Intelligent tourism not only provides services for tourists, tourism enterprises, and government but also promotes the coordination of tourism management and service and the economic development of destinations, so as to realize the friendly relationship between tourists and community residents. Using the Internet of Things technology, massive database, cloud computing technology, and scientific analysis, the intelligent tourism application model collects a large amount of data produced in the tourism activities in time, effectively integrates the tourism information of the regional tourism destination in recent years, establishes the related models, and predicts the future of the region. Changes in blockchain performance indicators under different consensus mechanisms are shown in Figure 2.

The flow of tourism, transmission to the server platform, and setup of a monitoring and forecasting system for tourism activity data can effectively predict the flow direction of tourism destinations and the trend of future flow and rationally guide the tourist attractions, tourism enterprises, and government tourism departments to put the corresponding manpower into service and prevent the group. The occurrence of a sexual event. The framework of the intelligent tourism application model based on big data is shown in Figure 3.

The virtual service platform of intelligent tourism is set up in the provincial capital cities, and the information channels of tourist accommodation, travel agencies, tourism enterprises, tourism colleges, and tourist attractions are used to realize the information circulation of various stakeholders, and the hidden value function of the large database will be greatly played, and the information sharing platform will be constructed. The model platform is based on the massive data of various provincial tourism events at the provincial level in recent years. The basic data are classified into basic data, basic data of tourist attractions, basic data of

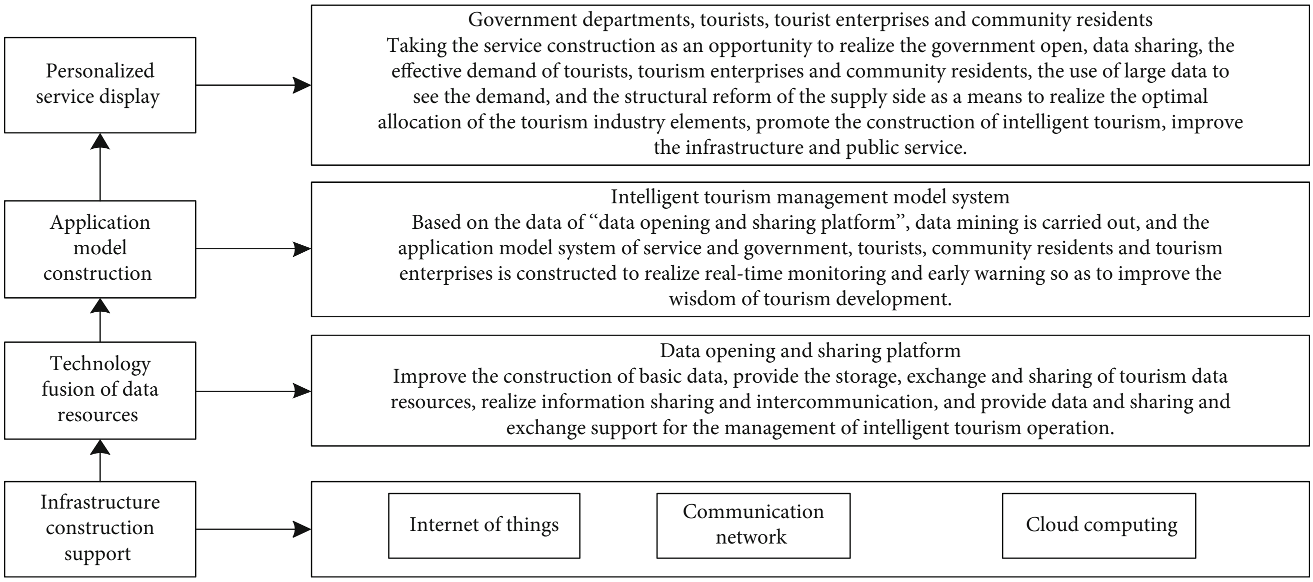


Figure 3: Framework of the smart tourism application model based on big data.

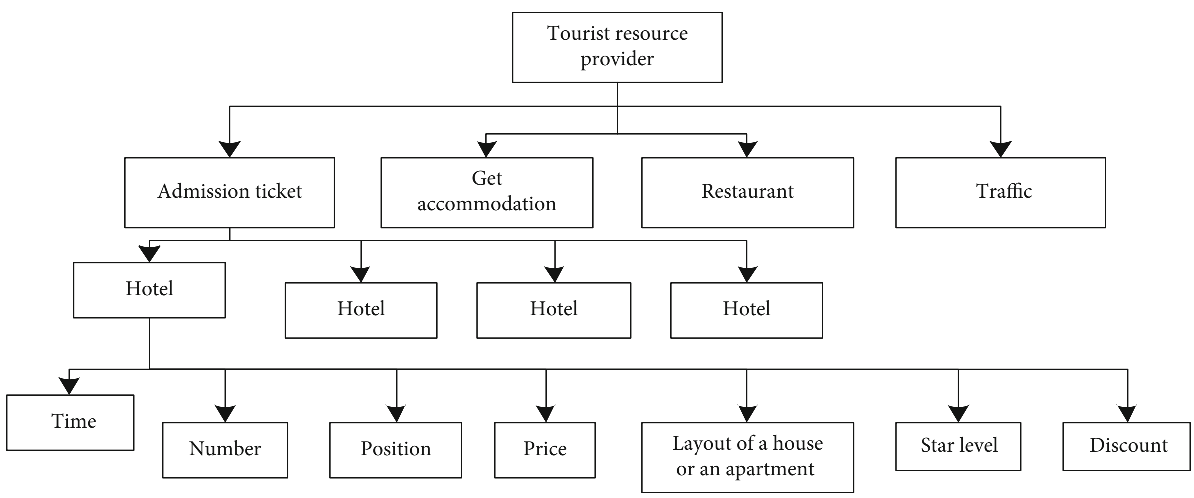


Figure 4: Basic information on hotel resources.

government tourism departments, and basic data of community residents, and then the standardized data are introduced into the application of intelligent tourism.

The data platform system of the model. At the same time, through the Internet of Things technology, all kinds of basic data are accurately tagged on the map of each province, and the provincial city is divided into several regional tourist destinations. The spatial coordinates of all kinds of basic data are tagged in various regional tourist destinations, and space grid numbers are carried out, and a variety of related prediction models can be used. It can realize real-time monitoring and early warning for a certain area tourist destination and a certain time period and predict the possibility of various events in the region. The basis of the monitoring and early warning system is that the basic data of tourism enterprises, tourist attractions, government tourism departments, and community residents include the type, time, and location of events. The combination definition is a structure with recursive nature; that is, the inventory can be combined repeatedly to form a more complex and richer product system. The inventory combination model is very suitable for the statistical price and discount treatment. The basic information and necessary attributes of the accommodation are shown in Figure 4.

The data processing platform of the application model can be used to discover the time and time of various types of events through the mathematical model of cloud computing. Location rules can automatically predict the flow and flow of tourists in different periods of the region by the icon of a number of axes, including the real-time tourist flow of regional tourist destinations and the monitoring and early warning of a specific day, week, month, year, or custom time period. In order to achieve the above planning objectives, we must first define a good inventory management basis, and the inventory is the basis for the calculation of the whole tourism level, and the inventory management hierarchy diagram of the tourism system is shown in Figure 5.

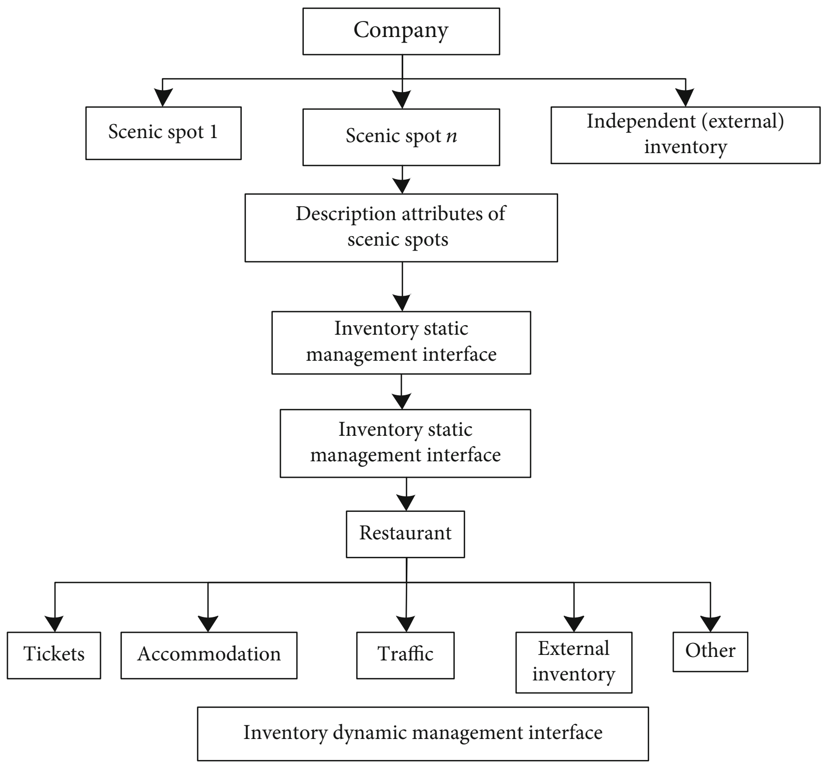


Figure 5: An inventory management hierarchy map of the tourism system.

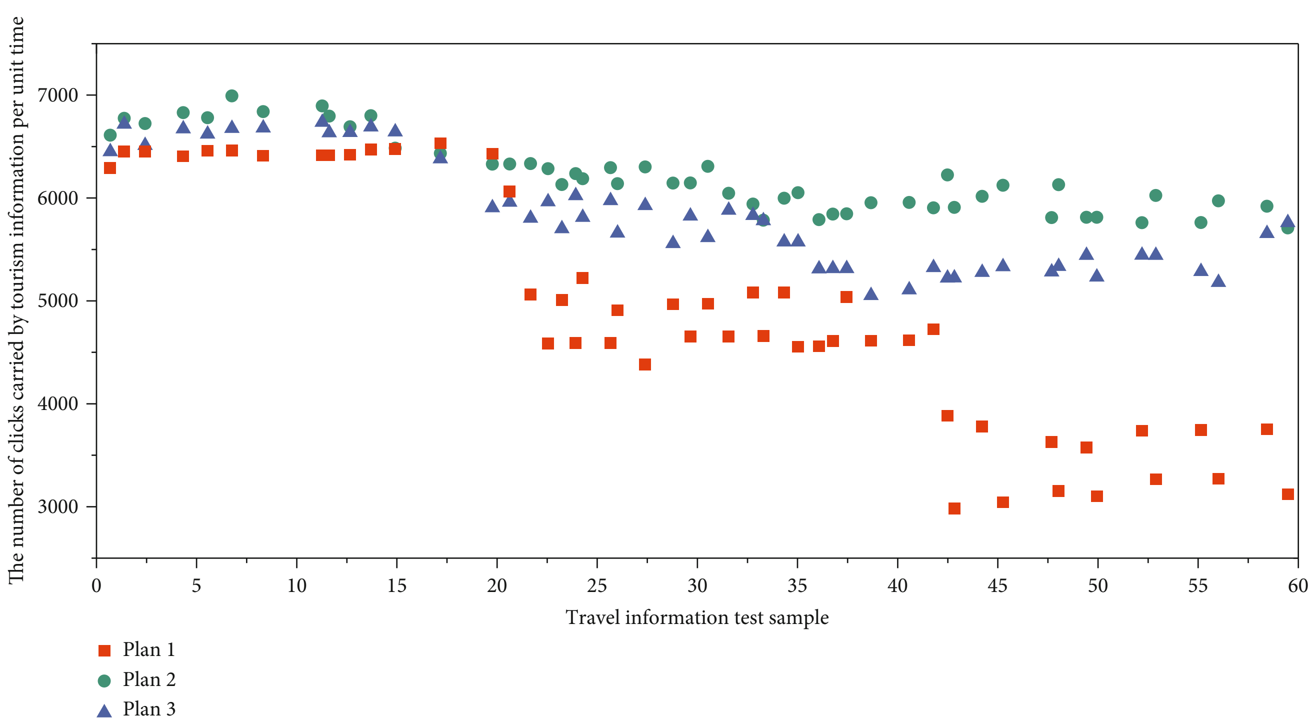


Figure 6: The number of clicks carried by tourism information per unit time.

At present, the resource information provided by the provider of resources is divided into four categories: admission, accommodation, catering, and transportation, of which each element is subdivided into specific resource providers and resource attributes. The number of clicks carried by tourism information per unit time is shown in Figure 6.

The whole inventory system mainly uses the tree structure to manage the data, and the inventory management

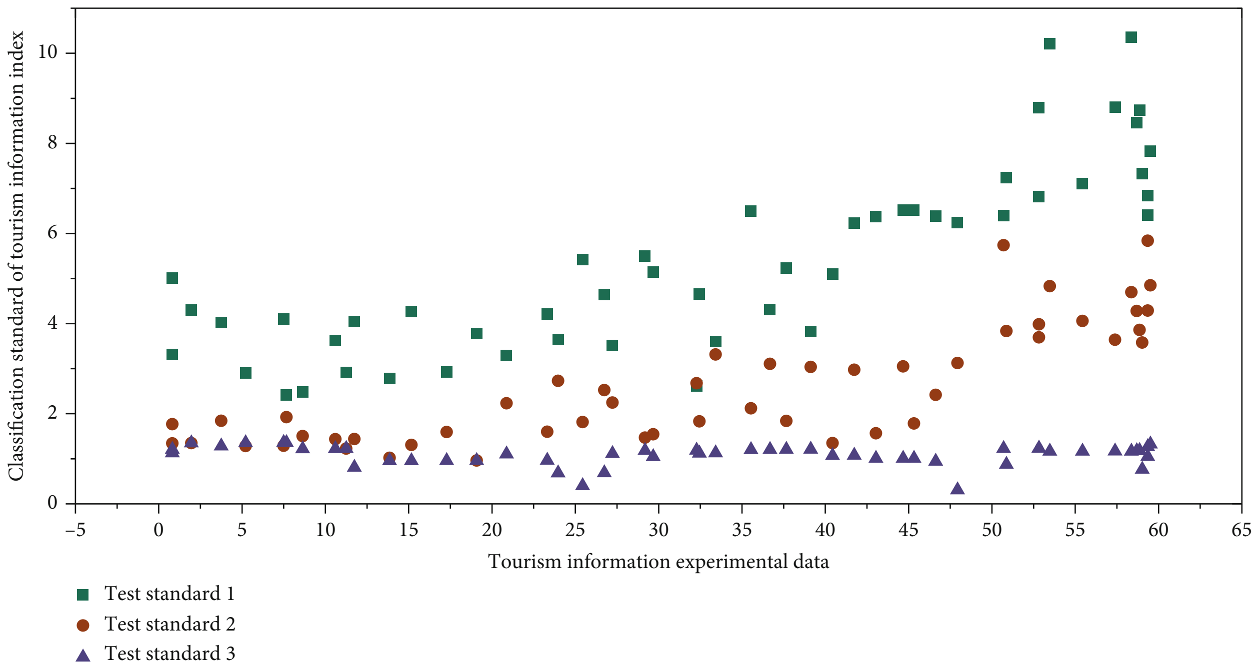


FIGURE 7: Classification standard of the tourism information index.

interface determines the basic elements of all the tourism elements, including the name of the commodity provider, the price attribute, the quantity type, the time attribute, the position attribute, and the quality attribute of the product. The definition of the inventory portfolio determines that those goods can be combined and sold together to form a set of services. The classification standard of the tourism information index is shown in Figure 7.

The information of the specific inventory is classified according to specific categories. This is the storage and processing of specific data in inventory management. The dynamic structure of the inventory management is to deal with the inventory information of the dynamic goods. One of the characteristics of the tourism inventory is that the product produces new and effective products automatically every day so that the product's reservation and ability can be calculated.

5. Conclusions

A smart tourism system is a highly complex cloud system platform that can automatically help users set up travel planning. This choice is based on a series of priority selection rules. The selection and setting of these rules have been studied by the market and the users, which have a fairly accurate selection effect. After using the system, users only need to enter their own travel plans. They can get out of the tedious travel strategy and get a complete, efficient, feasible, and optimized tour line immediately. It saves users valuable time and helps users save costs and costs. The online implementation of the system will bring a new user experience mode to the tourism industry.

Acknowledgments

This work was supported by the General Teaching research project by the Education Department of Anhui Province in 2019 (No. 2019jyxm0447) and the Major Teaching research project by Chuzhou University in 2019 (No. 2019jyz026).

References

[1] H. Wu, X. Wu, Q. Ma, and G. Tian, "Cloud robot: semantic map building for intelligent service task," *Applied Intelligence*, vol. 49, no. 2, pp. 319–334, 2019.

[2] Z. Rong-Hao, "Research on the temperature rise test system construction of for high altitude and low voltage switchgear," *Environmental Technology*, vol. 19, no. 2, pp. 31–44, 2019.

[3] D. Hong, "Research on the construction of new energy micro grid in countryside," *Electrical Engineering*, vol. 10, no. 12, pp. 30–41, 2015.

[4] C. Zhang and S. Chang, "Research on the strategy for hospital medical records information sharing service," *Basic & Clinical Pharmacology & Toxicology*, vol. 118, 2016.

[5] X. Wang, C. Wang, C. Jiang, L. Yang, Z. Li, and X. Zhou, "Rule optimization for real-time query service in software-defined internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–11, 2015.

[6] N. Khademi, “Intelligent transportation system user service selection and prioritization: hybrid model of disjunctive satisfying method and analytic network process,” *Transportation Research Record*, vol. 2189, no. 1, pp. 45–55, 2010.

[7] C. Tomasetti and B. Vogelstein, “Variation in cancer risk among tissues can be explained by the number of stem cell divisions,” *Science*, vol. 347, no. 6217, pp. 78–81, 2015.

[8] R. Draghia, F. Letourneur, C. Drugan et al., “Metachromatic leukodystrophy: identification of the first deletion in exon 1 and nine novel point mutations in the arylsulfatase A gene,” *Human Mutation*, vol. 9, no. 3, pp. 234–242, 1997.

[9] C. Bollmeyer, J. D. Keller, C. Ohlwein et al., “Towards a high-resolution regional reanalysis for the European CORDEX domain,” *Quarterly Journal of the Royal Meteorological Society*, vol. 141, no. 686, pp. 1–15, 2015.

[10] C. Gómez, S. Vega-Quiroga, F. Bermejo-Pareja, M. J. Medrano, E. D. Louis, and J. Benito-León, “Polypharmacy in the elderly: a marker of increased risk of mortality in a population-based prospective study (NEDICES),” *Gerontology*, vol. 61, no. 4, pp. 301–309, 2015.

[11] J. Kim, Y. Jeon, and H. Kim, “The intelligent IoT common service platform architecture and service implementation,” *The Journal of Supercomputing*, vol. 74, no. 9, pp. 4242–4260, 2018.

[12] S. Hu, Y. Wang, W. Wang et al., “Ag nanoparticles on reducible CeO2 (111) thin films: effect of thickness and stoichiometry of ceria,” *Journal of Physical Chemistry C*, vol. 119, no. 7, pp. 3579–3588, 2015.

[13] G. R. Griffith and H. M. Burrow, “The value of research: using the Impact Tool to evaluate realised and anticipated benefits of the Cooperative Research Centre for Beef Genetic Technologies,” *Animal Production Science*, vol. 55, no. 2, p. 133, 2015.

[14] C. S. Pimentel, J. McKenney, P. N. Firmino, T. Calvão, and M. P. Ayres, “Sublethal infection of different pine species by the pinewood nematode,” *Plant Pathology*, vol. 69, no. 8, pp. 1565–1573, 2020.

[15] K. Han, “On a stochastic construction of the kinematics in discrete space–time,” *Canadian Journal of Physics*, vol. 93, no. 5, pp. 496–502, 2015.

[16] W. Ting, “Construction of intelligent communication service based on artificial intelligence,” *Electrical Engineering*, vol. 5, no. 12, pp. 19–34, 2019.

[17] L. Shun-jie, “Research on development countermeasures for farmer specialized cooperative on the background of new rural construction,” *Journal of Investigative Medicine*, vol. 63, no. 8, article S66, 2015.

[18] C. Hongyu, “Research on the application of honeycomb structure dynamic algorithm in intelligent information management system,” *Electrical Engineering*, vol. 9, no. 2, pp. 20–34, 2015.

[19] S. S. Rautaray and A. Agrawal, “Vision based hand gesture recognition for human computer interaction: a survey,” *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1–54, 2015.

[20] J. M. Roman-Belmonte, H. de la Corte-Rodriguez, and E. C. Rodriguez-Merchan, “How blockchain technology can change medicine,” *Postgraduate Medicine*, vol. 130, no. 4, pp. 420–427, 2018.

[21] H. Subramanian, “Decentralized blockchain-based electronic marketplaces,” *Communications of the ACM*, vol. 61, no. 1, pp. 78–84, 2017.

[22] H. I. Ozercan, A. M. Ileri, E. Ayday, and C. Alkan, “Realizing the potential of blockchain technologies in genomics,” *Genome Research*, vol. 28, no. 9, pp. 1255–1263, 2018.

[23] S. Hall, B. Poller, C. Bailey et al., “Use of UV fluorescence-based simulation in evaluation of personal protective equipment worn for first assessment and care of a patient with suspected high consequence infectious disease,” *Journal of Hospital Infection*, vol. 23, no. 6, pp. 13–24, 2018.

[24] Y. Zhang, R. H. Deng, X. Liu, and D. Zheng, “Blockchain based efficient and robust fair payment for outsourcing services in cloud computing,” *Information Sciences*, vol. 462, pp. 262–277, 2018.

[25] K. Gammon, “Experimenting with blockchain: can one technology boost both data integrity and patients' pocketbooks?,” *Nature Medicine*, vol. 24, no. 4, pp. 378–381, 2018.

[26] V. Sharma, I. You, F. Palmieri, D. N. K. Jayakody, and J. Li, “Secure and energy-efficient handover in fog networks using blockchain-based DMM,” *IEEE Communications Magazine*, vol. 56, no. 5, pp. 22–31, 2018.

[27] L. Li, J. Liu, L. Cheng et al., “CreditCoin: a privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2204–2220, 2018.

[28] K. N. Khaqqi, J. J. Sikorski, K. Hadinoto, and M. Kraft, “Incorporating seller/buyer reputation-based system in blockchain-enabled emission trading application,” *Applied Energy*, vol. 209, pp. 8–19, 2018.

[29] B. Hamdaoui, M. Alkalbani, T. Znati, and A. Rayes, “Unleashing the power of participatory IoT with blockchains for increased safety and situation awareness of smart cities,” *IEEE Network*, vol. 34, no. 2, pp. 202–209, 2020.

[30] H. E. Pence, “Blockchain: will better data security change chemical education?,” *Journal of Chemical Education*, vol. 97, no. 7, pp. 1815–1818, 2020.

[31] D. Jiang, G. Li, Y. Sun, J. Kong, and B. Tao, “Gesture recognition based on skeletonization algorithm and CNN with ASL database,” *Multimedia Tools and Applications*, vol. 78, no. 21, pp. 29953–29970, 2019.

[32] Y. He, G. Li, Y. Liao et al., “Gesture recognition based on an improved local sparse representation classification algorithm,” *Cluster Computing*, vol. 22, Supplement 5, pp. 10935–10946, 2019.

[33] W. Cheng, Y. Sun, G. Li, G. Jiang, and H. Liu, “Jointly network: a network based on CNN and RBM for gesture recognition,” *Neural Computing and Applications*, vol. 31, no. S1, pp. 309–323, 2019.

[34] Y. Sang, H. Shen, Y. Tan, and N. Xiong, “Efficient protocols for privacy preserving matching against distributed datasets,” in *Information and Communications Security. ICICS 2006. Lecture Notes in Computer Science, vol 4307*, P. Ning, S. Qing, and N. Li, Eds., pp. 210–227, Springer, Berlin, Heidelberg, 2006.

[35] F. Long, N. Xiong, A. V. Vasilakos, L. T. Yang, and F. Sun, “A sustainable heuristic QoS routing algorithm for pervasive multi-layered satellite wireless networks,” *Wireless Networks*, vol. 16, no. 6, pp. 1657–1673, 2010.

[36] J. Li, N. Xiong, J. H. Park, C. Liu, S. MA, and S. E. Cho, “Intelligent model design of cluster supply chain with horizontal cooperation,” *Journal of Intelligent Manufacturing*, vol. 23, no. 4, pp. 917–931, 2012.

[37] Z. Chen, D. Chen, Y. Zhang, X. Cheng, M. Zhang, and C. Wu, “Deep learning for autonomous ship-oriented small ship detection,” *Safety Science*, vol. 130, article 104812, 2020.

[38] Z. Liu, B. Hu, B. Huang, L. Lang, H. Guo, and Y. Zhao, "Decision optimization of low-carbon dual-channel supply chain of auto parts based on smart city architecture," *Complexity*, vol. 2020, Article ID 2145951, 14 pages, 2020.

[39] L. Dong, W. Wu, Q. Guo, M. N. Satpute, T. Znati, and D. Z. du, "Reliability-aware offloading and allocation in multilevel edge computing system," *IEEE Transactions on Reliability*, vol. 9, no. 2, pp. 1–12, 2019.

[40] J. Hu, Y. Sun, G. Li, G. Jiang, and B. Tao, "Probability analysis for grasp planning facing the field of medical robotics," *Measurement*, vol. 141, pp. 227–234, 2019.

[41] W. Wei, H. Song, W. Li, P. Shen, and A. Vasilakos, "Gradient-driven parking navigation using a continuous information potential field based on wireless sensor network," *Information Sciences*, vol. 408, no. 2, pp. 100–114, 2017.

A Classification Method of Tourism English Talents Based on Feature Mining and Information Fusion Technology

Xin Wei [1,2]

[1]Graduate School, Nueva Ecija University of Science and Technology, Cabanatuan, Nueva Ecija 3100, Philippines
[2]English Language Teaching Unit, Lu'an Vocational and Technical College, Lu'an 237100, China

Correspondence should be addressed to Xin Wei; xin615290@lvtc.edu.cn

Academic Editor: Muhammad Usman

With the rapid development of the Internet, text data has become one of the major formats of big data tourism and improves the quality and promotes the optimization and upgradation of tourism English talents. This paper proposes a model of tourism English talent resources based on data mining techniques using a big data framework. The characteristic distribution structure model is built to identify and blend the characteristics of tourism English talent resources. Connection feature mining and information fusion are combined to share data and schedule resources during the talent training process. Initially, the proposed research work uses a cloud storage system for developing intercultural communicative competence of tourism English talents. Next, the optimal scheduling design of tourism English talent training resource's big data is carried out. Finally, the fuzzy clustering method deals with the adaptive clustering of tourism English talent resource distribution big data. The simulation findings show that the proposed method has high precision and big data computation efficiency. Moreover, it can successfully mentor the development of a new framework of tourism English talent training.

1. Introduction

The training goal of tourism English talents is located in the technical application of the tourism industry and the compound management talents to cultivate the middle level of the tour guide industry, the hotel industry and the middle level, and the management talents above the middle level. Therefore, the training target of the Higher Vocational Tourism English talents is the technical nature of the talent type, the professionalism of knowledge ability, and the graduate's direction [1]. At present, a lot of travel agencies and hotel enterprises urgently need strong English listening, speaking, reading, writing, translation comprehensive ability, familiarity with the tourism industry-related policies, policies, good tourism, hotel management, and other professional knowledge and skills, with cross-cultural interdisciplinary talents. Tourism English majors must be closely integrated into the market because the professional orientation of the students is suitable for the needs of the current society; it depends highly on whether the orientation of the majors and the arrangement of the courses are reasonable and whether they have the current travel agency and the knowledge structure required by the hotel. Direct trade relations are established with many countries and regions. The number of the top five hundred multinational companies in the world has been increasing. Also, the number of multinational corporations and the amount of outbound tourism in the world ranks are increasing [2]. Therefore, there is a strong ability to express English, the managers who are familiar with the tourism knowledge and tour guide business, the tour guides, and the market demand. To improve tourism English, the quality of talents needs to optimize the training mode of tourism English talents and promote the improvement of the quality of tourism English talents [3].

Given the strong practical characteristics of a tourism English major, the most effective way to train its talents can only be the combination of production and learning. Therefore, in addition to strengthening the construction of training rooms, internships, outside schools, and training bases, big data scheduling and integration of tourism English talent resources should also be carried out according to the requirements of the quality, knowledge, and ability of social posts. The data mining methods for tourism English talents mainly include blind convolution algorithm, feature extraction method, time-frequency analysis method, and statistical analysis method. These methods cannot eliminate the convolution effect of the mining channel to singular features. There are several data mining tools that companies can use to translate raw data into actionable insights. These include statistics, artificial intelligence, hidden Markov models, metalearning, genetic algorithm, machine learning, and decision tree. It affects the accuracy of big data features mining in the distribution of big tourism English talent resources. In [4], a new information mining and scheduling algorithm for tourism English talent resources based on differential cumulate function feature mining is proposed. To improve the ability of information collection and feature analysis of tourism English talent resources, the distributed resources information of tourism English talent resources under a big data environment is fused and processed. However, the algorithm has high computational complexity and a big implementation cost [5–7].

In this research work, an effective and intelligent computational model is proposed. In view of the above problems, a new model of tourism English talents training is proposed based on big data, and it constructs a big data mining model of tourism English talents. Data mining is widely used in the field of artificial intelligence (AI), marketing, government intelligence (GI), services, and advertising. There are some other industries like crime agencies, retail healthcare, e-commerce, telecom, biological data analysis, and information retrieval like communication systems. Data mining is used to examine or explore the data using queries. These queries can be fired on the data warehouse. In the cloud storage system, the optimal scheduling design of tourism English talent training resource big data is carried out, and the fuzzy clustering method is used to deal with the adaptive clustering of tourism English talent resource distribution big data. This research work adopts the attribute distribution structure model to classify and fuse the features of tourism English talent resources, and it combines the methods of association feature mining and information fusion to share data and schedule resources in the process of talent training. Finally, the performance test is carried out through the simulation experiment, which shows the superior performance of this method in improving the big data analysis ability of tourism English talents training.

The rest of the paper is organized as follows: tourism English talent resource distribution model is explained in Section 2; big data feature extraction is discussed in Section 3, and, in Section 4, experimental results and discussion are discussed. Finally, we demonstrate a conclusion and future work in Section 5.

2. Tourism English Talent Resource Distribution Model

In this section, we elaborate the new cross-cultural communicative competence model that is proposed in the book titled "Intercultural Communication in Context" and the elements of cross-cultural communicative competence of tourism English talents.

2.1. The Formation of Cross-Cultural Communication Ability Elements. The formation of cross-cultural communication ability elements includes four elements, that is, knowledge factors, emotional factors, mental activity characteristics, and situational characteristics. Cross-cultural communication elements of tourism English talents are shown in Figure 1. The knowledge in the cross-cultural communication skills mentioned in the figure refers to the degree of understanding of the target's culture in the party of communication, and the cross-cultural communication ability is directly proportional to its understanding. Among them, knowledge factors include cultural values, beliefs and behaviors, verbal and nonverbal scripts, simplified and rigid cognition, and ethnocentrism. These factors can have a positive or negative impact on the communicators. Emotional factors refer to the attitudes adopted by the communicators in dealing with communicative objects from different cultural groups, that is, approaching or alienating. In the intercultural communication activities, inevitably there will be fear or anxiety. Therefore, if they are willing to communicate, the motivation of their communication is a more important emotional factor in cross-cultural communication. The ability to deal with stress and tolerance is positively related to intercultural communicative competence, and knowledge and affective factors in cross-cultural communicative competences interact with and support each other. The more knowledge, the less anxiety, and the more cross-cultural, the more arguments. The stronger the communicative motivation is, the stronger the communicative motivation is, and the more chances of obtaining relevant experience are, the more intercultural communication knowledge becomes more and more abundant.

Similarly, mental activity factors are a comprehensive manifestation of knowledge and emotional factors, including verbal and nonverbal communication and role play. Having more opportunities to practice using the target language is an important factor to enhance their abilities, and attention should also be paid to a series of auxiliary nonlanguage symbols such as body language, spatial language, and time language of the communication partner in communication. Role play is related to context. It is the use of verbal and nonverbal symbols in the destination culture according to their role identities. Moreover, for situational characteristics, the environment has a great influence on both sides of the communication. Situational characteristics

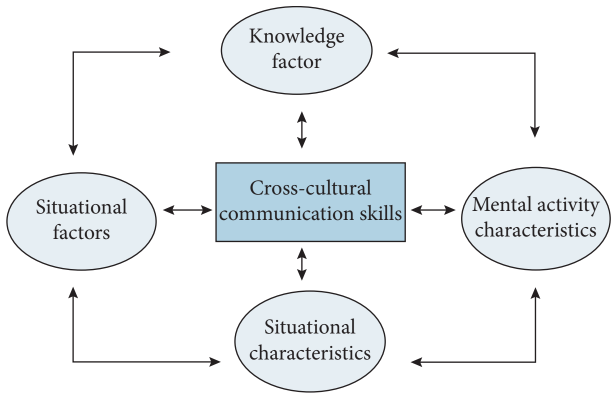


Figure 1: Cross-cultural communication elements of tourism English talents.

mean the real context of cross-cultural communication. The situational characteristics affecting their communicative competence are environmental context, prior contact, status differences, and third-party interference. Therefore, individuals have increased knowledge and experience in cross-cultural communication, and, as a result, their motivation has increased and they have actively participated in communicative activities, and vice versa, forming a learning cycle of good success.

2.2. The Elements of Cross-Cultural Communicative Competence of Tourism English Talents. The cross-cultural communicative competency elements constitute the first element of the domestic study's cross-cultural communicative competencies and are summarized into four communicative competence systems: basic communicative competence systems, emotional and relational competence systems, plot competence systems, and strategic capacity system. The elements of cross-cultural communicative competence of tourism English talents are shown in Figure 2.

From Figure 2, basic communicative competence system refers to the communicative ability that an individual needs to have in order to achieve communicative purposes, including language competence, and wants to connect with social cultural norms. These include verbal skills, nonverbal abilities, cultural abilities, interpersonal skills, and cognitive abilities, where the tourism English talent cultivation model believes that emotional competence refers to the communicator's recognition and understanding of the other party. Empathy needs to recognize the differences between individuals and cultures fully understand the self, suspend the self, eliminate the separation of oneself from the environment, conceiving for others, and finally preparing for empathy go through these six stages. Relational ability refers to the communication strategies used by the communicating parties in the process of communication. For example, the need for mutual communication between the two parties should be the basis for establishing relationships. The generation of consensus is the prerequisite for communication, and it is also related to cultural orientation and values.

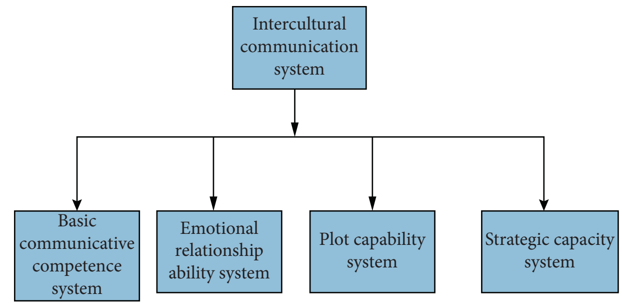


Figure 2: The elements of cross-cultural communicative competence of tourism English talents.

Similarly, the plot ability system comes to the plot competence system; it generally refers to the fixed sequence of interactions in a particular cultural environment. There are unique verbal and nonverbal rules in each plot. It proposes four aspects of ability, that is, common sense in communication, and also refers to the ability of a script communicator who directs behaviors to strive to achieve communicative purposes. The ability to follow the rules of communication in a particular context is appropriate to handle social scene responses. Moreover, strategy capability system proposed by the tourism English talent training model is different from the general communication strategy, but it is the remedial measures adopted by the communicators in the face of language barriers or verbal skills and clearly shows that strategic competence is an important part of the communicative competence.

3. Big Data Feature Extraction

In this section, we elaborate the big database model of tourism English talent resources distribution and the principles of big data mining for big-scale tourism English talent resources distribution. The feature fusion of tourism English talent resources optimized by big data fusion has good information mining capabilities.

3.1. Big Database Model of Tourism English Talent Resources Distribution. The big data characteristic information of tourism English talent resources distribution is usually expressed as a set of nonstationary broadband time series. The big-scale tourism English talent database big data resource information mining and scheduling is realized by using the method of time series detection [8]. Firstly, a big database model of tourism English talent resource distribution is constructed. In the construction of big data flow model of tourism English talent resource distribution, the task code is distributed to a number of data nodes by using distributed code execution mode, and the result is returned to the client after execution. It can effectively avoid the movement and transmission of cached data in the big database of tourism English talent resources distribution and use single computer or network computer to access the big amount of data and realize the evaluation and prediction of tourism resources information amount. The big database of tourism English talent resources distribution uses I/O, USB, and disk and other devices to achieve efficient read-write access [9]. Thus, the access and management model of tourism English talent resources distribution database is shown in Figure 3.

According to the analysis of Figure 2, suppose that the big database data set of tourism English talent resources distribution is the number of big data sets $X = \{x_1, x_2, \ldots, x_n\}$ of tourism English talent resources distribution, and the joint distribution feature vectors in n are all N-dimensional vector X containing c big data classes of resource distribution. The clustering center of the big data class of $v_i = \{v_{i1}, v_{i2}, \ldots, v_{ip}\}$ resource distribution is big data. If the joint distribution function of nonstationary time sampling $\{x(t_1), \ldots, x(t_n)\}$ and the joint distribution function of $\{x(t_1 + \tau), \ldots, x(t_n + \tau)\}$ are autocorrelation in the big database of tourism English talent resources distribution, the big data features of tourism English talent resource distribution are expressed as follows:

$$\theta_1(k+1) = \theta_1(k) - \mu \mathrm{Re}[y(k)\varphi^*(k)]. \tag{1}$$

The spectral features of the big database of tourism English talent resources distribution are analyzed by using aliasing spectrum ambiguity function, and the function relationship between the response variable and the forecast variable $\mathrm{X} = (\mathrm{x}_1, \ldots, \mathrm{x}_n, \ldots, \mathrm{x}_{n+m})$ is determined. $\langle R_1, +, \times \rangle$ and $\langle R_2, \oplus, \otimes \rangle$ are closed loop of the two eigenvector sets. Because the big data features of tourism English talent resource distribution database are combined with probability distribution function mapping $f\colon R_1 \mapsto R_2$, for any $m_i \in R_1$, the big data information feature extraction equation of tourism English talent resource distribution can be obtained as follows:

$$\begin{aligned}
\dot{x} &= V \cos\theta \cos\psi_V,\\
\dot{y} &= V \sin\theta,\\
\dot{z} &= -V \cos\theta \sin\psi_V,\\
\dot{\vartheta} &= \omega_y \sin\gamma + \omega_z \cos\gamma,\\
\dot{\psi} &= \frac{\left(\omega_y \cos\gamma - \omega_z \sin\gamma\right)}{\cos\vartheta},\\
\dot{\gamma} &= \omega_x - \tan\vartheta\left(\omega_y \cos\gamma - \omega_z \sin\gamma\right),
\end{aligned} \tag{2}$$

where x, y, and z indicate the initial frequency of big data in the big database of tourism English talent resources distribution, ψ_V indicates the big data training set of tourism English talent resources distribution, θ denotes the solution vector Q in the space of the talent resource distribution variable, ϑ is the instantaneous amplitude of the complex signal $z(t)$ of the big tourism English talent resource distribution database, and γ represents the statistical characteristics in the big tourism English talent database [10].

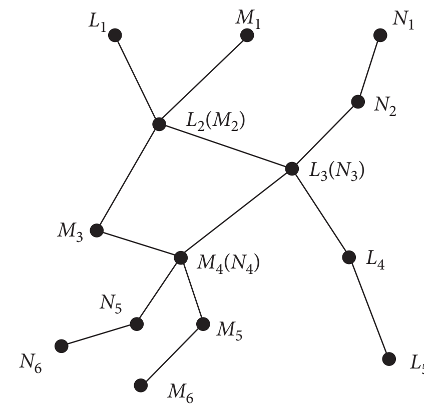


Figure 3: Distribution structure model of tourism English talent resources.

3.2. Principles of Big Data Mining. Based on constructing the distributed data structure model of tourism English talent training resources, the big data mining process is carried out and the big data distributed structure model of tourism time series algorithm provides regression algorithms that are used for optimizing and forecasting of continuous values like sales, overtime, and temperature. This algorithm can predict the trends that are based only on the original data set that is used to create a model. English talent resources distribution is designed. The computer data set under the condition of source distribution big data is fused and analyzed [11]. The schematic block diagram of big data mining for big-scale tourism English talent resources distribution is shown in Figure 4.

Fast data and resource information mining and scheduling are carried out in the big database of tourism English talent resources distribution to realize the integration of resource distribution big data [12]. Let R be the trust relation with quaternion (E_i, E_j, d, t) in the big data feature data of big-scale tourism English talent resources distribution. The information state equation of data classification attribute $A = \{A_1, A_2, \ldots, A_m\}$ and the information state equation of big data feature data integration of big-scale tourism English talent resources distribution are obtained as follows:

$$\begin{aligned}
P_i(t) &= \sum_{n=1}^{N} \frac{A}{r} e^{-jkr} R_{in} \frac{1}{r} e^{-ikr},\\
P_i(t) &= \frac{A}{r^2} \sum_{n=1}^{N} e^{-j2kr} a_{in} e^{j\psi_{in}}.
\end{aligned} \tag{3}$$

The time-frequency characteristics of the big database signal of tourism English talent resource distribution are calculated, and the big data structural model of the talent training information base resource distribution is constructed [13]. The big data model of resource distribution is obtained as follows:

$$x(t) = \sum_{i=0}^{p} a(\theta_i) s_i(t) + n(t). \tag{4}$$

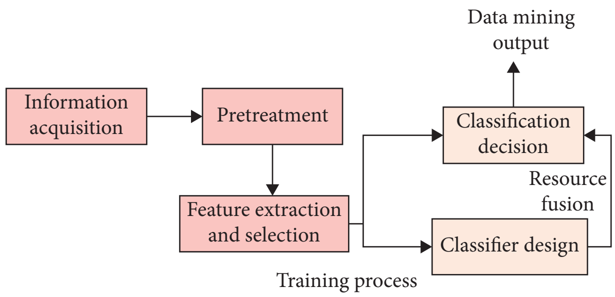


FIGURE 4: Principles of big data mining for big-scale tourism English talent resources distribution.

The big data characteristic data of tourism English talent resource distribution are sampled and updated in sequence, and the spectrum of nonstationary broadband signal $z(t)$ in the big database of tourism English talent resource distribution is obtained. Furthermore, time-frequency analysis is used to carry out transient analysis [14]. According to the estimation of time frequency, the probability function density expression of the accuracy of big data characteristic data for the distribution of big tourism English talent resources is as follows:

$$p(y|\alpha,\theta) = \sum_{k=1}^{K} \alpha_k p_k\left(y|\mu_k, \sum k\right). \quad (5)$$

Big tourism English talent resource distribution big data is a sample set composed of n samples. There are m indexes in each big data sample of tourism English talent resource distribution; then the index characteristic vector of j sample is $x_j = \left\{x_{1j}, x_{2j}, \ldots, x_{mj}\right\}^T$, tourism English. The a posteriori probability of big data is estimated as $p(x_0)$:

$$P_{ij}(k) = \frac{\left(l_j(k) - l_i(k)\right)\eta_{ij}(k)}{\sum_{j\in N_i(k)}\left(l_j(k) - l_i(k)\right)\eta_{ij}(k)}. \quad (6)$$

According to the results of big data mining, data sharing and resource scheduling in the process of talent training are carried out by combining association feature mining and information fusion methods.

3.3. Optimization of Cross-Cultural Communicative Competence. Constructing English for tourism in the context of big data is not exactly equivalent to the Basic English content. It has its uniqueness, except listening, speaking, reading, writing, and translation of Basic English training. In addition to their abilities, they must also develop their professional English skills, relevant cross-cultural thinking skills, and communication skills. Therefore, in the tourism English curriculum, student's cross-cultural communication skills should be cultivated to meet the requirements of the current society, disciplines, and students themselves. From the study of intercultural communicative competence by various scholars at home and abroad, cross-cultural communicative competence is a comprehensive, multidimensional system such as knowledge system, relationship system, and behavior system. Based on these researches on the elements of intercultural communication, and combining the characteristics of the curriculum of tourism English under the big data environment, a cross-cultural communicative competence framework for tourism English teaching in the context of big data is constructed. The framework of cross-cultural communicative competence of English tourism talents is shown in Figure 5.

From Figure 5, we can observe that, for a student majoring in tourism management, English communication skills have become a core skill. Compared with other professional students, English communication ability is a requirement within the category of professional competence. Therefore, colleges and universities pay great attention to tourism English teaching in the context of big data. However, in practice, students find that they have a high level of enthusiasm for learning professional English, but the results are not obvious. The bigger reason is that travel professional English is different. Basic English requires global awareness and industry awareness, which is the basis for developing cross-cultural communication skills. The global awareness mentioned here includes cross-cultural awareness and cross-cultural thinking. Cross-cultural awareness can also be called cultural awareness [15]. Similarly, In the process of cross-cultural communication, cultural conflicts occur everywhere. In order to shorten the conflict period, it is necessary to adapt to the rapid recovery period. This is a process of entering a steady state through a series of adjustments from a relatively unstable state [16]. The theory of adaptation to cross-cultural adaptation process can be summarized as a model curve [17]. The adaptability plays an important role in the exchange process. This is a turning point from uncertainty decision rather than acceptance. Adaptability includes two elements: psychological adaptability and cultural adaptability [18, 19].

Moreover, in the reform of foreign language classroom teaching, cultivating student's communicative application ability is the focus of foreign language circles in recent years. The University English Classroom Requirements with a target meaning also highlights language knowledge and application capabilities. For the purpose of tourism English, language knowledge and cultural knowledge are the basic teaching contents of the teacher's class. In addition, the ability of the first three media is in communication practice. Communicative practice ability is the highest level of intercultural communication ability. It includes communicative competence, nonverbal communication, and communication skills. The communicative practice ability emphasizes the use of linguistic, nonverbal, and related strategies in practical situations to solve practical cross-cultural communication problems. The ability of a large communication strategy is the core competitiveness of tourism services.

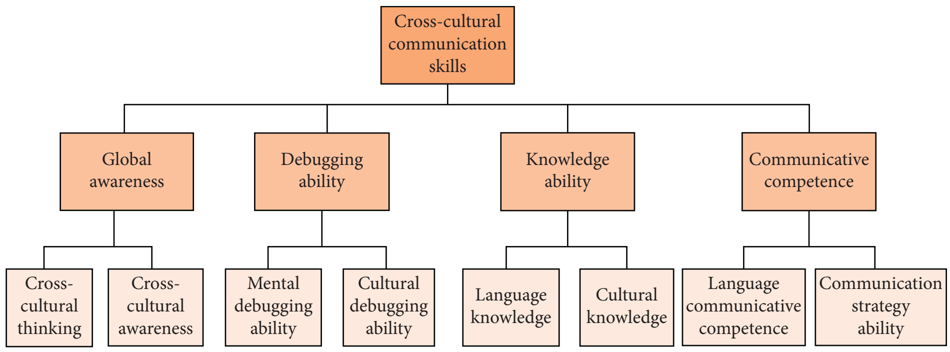


Figure 5: Framework of cross-cultural communicative competence of English tourism talents.

3.4. Feature Classification and Information Fusion of Tourism English Talent Resources. The fuzzy clustering method is used to deal with the adaptive clustering of tourism English talent resources distribution, the attribute distribution structure model is used to classify the features and information fusion of tourism English talent resources, and the time-frequency analysis method is adopted. By extracting the big data feature of resource distribution, the information state function of big data feature data of tourism English talent resource distribution is presented as follows:

$$\begin{aligned}\dot{x}_i &= f_i(x_i, u_i) D\left(x_i, A_j(L)\right)\\ &= \min\left\{D\left(x_i, A_j(L)\right)\right\},\end{aligned} \tag{7}$$

where $x_i \in R^n$, and the state vector of big data feature data representing the distribution of tourism English talent resources is presented. Based on the above analysis, the feature extraction of resource distribution big data in the big database of tourism English talent resources distribution is realized, and the big scale is extracted. The big data training sample of tourism English talent resource distribution in the database, the time-frequency feature of the extracted tourism English talent resource distribution big data is rearranged by sliding time window, and the tourism English talent database distribution big data is carried out. The expression of the frequency domain model is as follows:

$$x_n = [x(0), x(1), \ldots, x(N-1)]^T. \tag{8}$$

The big data fusion algorithm of tourism English talents training is improved by phase spectrum compensation, and the attribute weights of random linear access channel for resource big data fuzzy decision fusion of tourism English talent training are obtained as follows:

$$F = \begin{bmatrix} \omega^0 & \omega^0 & \omega^0 & \cdots & \omega^0 \\ \omega^0 & \omega & \omega^2 & \cdots & \omega^{k-1} \\ \omega^0 & \omega^2 & \omega^4 & \cdots & \omega^{2(k-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \omega^0 & \omega^{k-1} & \omega^{2(k-1)} & \cdots & \omega^{(k-1)^2} \end{bmatrix} = \prod_{0 \le j < i \le k-1} \left(\omega^i - \omega^j\right). \tag{9}$$

The vector model structure of big data fuzzy decision fusion for tourism English talents training is expressed as follows:

$$\begin{cases} \dot{m}_i(t) = -a_i m_i(t) + b_i\left(p_1(t-\sigma), p_2(t-\sigma), \ldots, p_n(t-\sigma)\right), \\ \dot{p}_i(t) = -c_i p_i(t) + d_i m_i(t-\tau). \end{cases} \tag{10}$$

Combined with the association feature mining and information fusion method, data sharing and resource scheduling are carried out in the training process, so that the data fusion capacity is improved [15, 16].

3.5. Big Data Fusion and Optimization Mining of English Talent Resource Distribution. In the tourism English talent database, the prediction of packet information can be as follows: $\{\lambda_i\colon 1 \le i \le S\}$, criterion $\left\{R_j\colon 1 \le j \le L\right\}$, and density prior information sliding time window rearrangement of big data for tourism English talent resources distribution calculation is expressed as follows:

$$
\begin{aligned}
N_{i,j} &= \left(f(A)_{i,1} \otimes f(B)_{1,j}\right) \oplus \left(f(A)_{i,2} \otimes f(B)_{2,j}\right) \oplus \cdots \oplus \left(f(A)_{i,h} \otimes f(B)_{h,j}\right) \\
&= \left(\left(K_A \cdot \left(A_{i,1} + rp\right) \cdot k_A^r \bmod n\right) \cdot \left(K_B \cdot \left(B_{1,j} + rp\right) \cdot k_B^r \bmod n\right)\right) \oplus \cdots \oplus \\
&\quad \left(\left(K_A \cdot \left(A_{i,h} + rp\right) \cdot k_A^r \bmod n\right) \cdot \left(K_B \cdot \left(B_{h,j} + rp\right) \cdot k_B^r \bmod n\right)\right) \\
&= \sum_{u=1}^{h} \left(\left(K_A \cdot \left(A_{i,u} + rp\right) \cdot k_A^r \bmod n\right) \cdot \left(K_B \cdot \left(B_{u,j} + rp\right) \cdot k_B^r \bmod n\right)\right) \bmod n \\
&= \sum_{u=1}^{h} \left(\left(\frac{s \cdot l}{k_A^r} \cdot \left(A_{i,u} + rp\right) \cdot k_A^r \bmod n\right) \cdot \left(\frac{s \cdot l}{k_B^r} \cdot \left(B_{u,j} + rp\right) \cdot k_B^r \bmod n\right)\right) \bmod n \\
&= (s \cdot l)^2 \sum_{u=1}^{h} \left(\left(A_{i,u} + rp\right) \cdot \left(B_{u,j} + rp\right)\right) \bmod n \\
&= (s \cdot l)^2 \sum_{u=1}^{h} \left(A_{i,u} \cdot B_{u,j} + \left(A_{i,u} + B_{u,j}\right) \cdot rp + r^2 p^2\right) \bmod n.
\end{aligned}
\tag{11}
$$

It can be seen that the nearest-neighbor matching method using the time-frequency eigenvector can find out the potential matching of the big data features of the tourism English talent resources:

$$
\begin{aligned}
&\left(V(a_1, \ldots, a_m)^{(\alpha_1, \ldots, \alpha_m)}\right)^{-1} V(b_1, \ldots, b_m)^{(\beta_1, \ldots, \beta_m)} \\
&= \left(\left(V(a_1, \ldots, a_m)^{-1} V(b_1, \ldots, b_m)\right)^{\left(\alpha_1^{-1}, \ldots, \alpha_m^{-1}\right)^T}\right)^{(\beta_1, \ldots, \beta_m)}.
\end{aligned}
\tag{12}
$$

The fuzzy feature balance metric of the differences between j and c categories in the big data sample of tourism English talent resources distribution can be expressed as follows:

$$
f_j\left(u_j, s, w\right) = \sum_{h=1}^{c} \left\{ u_{hj}^2 \left[\sum_{i=1}^{m} \left[w_i \left(r_{ij} - s_{ih} \right) \right]^2 \right]^{\alpha/2} \right\}. \tag{13}
$$

The big data signal analysis of tourism English talent resources distribution is extended to the disturbed environment. For the time series of tourism English talent resources $\{x(t_0 + i\Delta t)\}$, $i = 0, 1, \ldots, N-1$, the fast resource information of the big database of tourism English talent resources distribution is established. The mining and scheduling model objective function is,

$$
\min\{f(u, s, w) = f_1(u_1, s, w), f_2(u_2, s, w), \ldots, f_n(u_n, s, w)\}. \tag{14}
$$

4. Result Analysis and Discussion

In this section, we compare the performance and utility of the proposed model. The efficiency of the proposed model was initially tested using the optimization of tourism English talent training mode is realized. In order to improve the ability of fast resource information mining and scheduling, a sliding time window rearrangement of big data prior information of tourism English talent resource distribution in big tourism English talent database is carried out.

4.1. Experimental Setup. In this paper, each algorithm was evaluated through experiments, and broad data computational approaches were used, that is, the open-source Apache Hadoop and the Apache Spark platform. First of all, we created a Spark cluster that uses physical nodes, that is, 80 nodes. To determine the basic software and hardware specifications for the experiment, all computing nodes have been simplified, and their configuration is shown in Table 1. For example, operating systems Ubuntu 18, Spark 3.01, and Hadoop 3.2 are used in experimental configurations. In all 80 nodes, one node was used as a master node or as a processing node, and the remaining 79 nodes were used as slaves in the cluster. Apache Hadoop's software library is a tool used in basic programming models to spread vast data collection through computer clusters. For each local server calculation and storage, it is scheduled to reach multiple computers. Apache Spark is a hierarchical programming engine for large-scale data processing that provides high-level APIs for Java, Python, R, and Scala and is used as an advanced map driver. Apache Spark enables a wide range of high-level tools, including SQL Spark for structured data analysis, MLlib for apprenticeships, and GraphX for graphics processing.

4.2. Simulation Experiment and Result Analysis. In order to test the application performance of this algorithm in realizing the big data mining of tourism English talent resource distribution and the optimization of training mode, the simulation experiment is carried out. In the experiment, 100 pieces of big data sample data of tourism English talent resource distribution are selected, the number of talent attribute samples is 50, and the dimension of time-frequency

Table 1: Configuration detail of cluster.

Hardware	Number of cluster nodes	80
	Hard disk	500 GB
	No. of CPU cores	8
	Processor	3.0 GHz × 4
	Connectivity	100 Mbps ethernet LAN
	RAM size	30 GB
	Data block size	256 Mb
Software	Operating system	Ubuntu 18
	Java-JDK	V-1.8
	Hadoop	V-3.0
	Spark	V-3.0.1
	Open-PDC	V-1.5
	Python	V-3.8

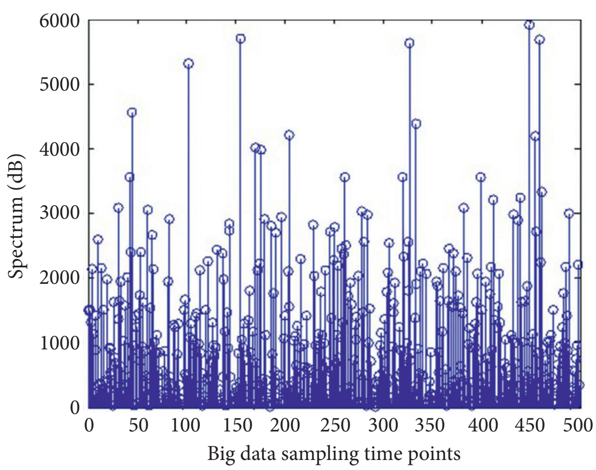


Figure 6: Big data mining results of tourism English talent training.

feature state space of the big database of tourism English talent resource distribution is 10, which is sampled according to the characteristic data. According to the above simulation experimental environment design and parameter design, the big data feature extraction and resource information mining of tourism English talent resource distribution are carried out, and the data mining is obtained. The effect of some parameters, that is, learning rate and activation functions, on accuracy is summarized in Table 1, that is, performance of different phenotype. We have optimized the parameters settings for all the learning algorithms for accuracy-based analysis. The result is shown in Figure 6.

The results of the analysis of Figure 6 show that the data mining of tourism English talent training resource distribution using this method is of high accuracy. The fast resource information mining of big data for tourism English talent resources distribution can be realized. SQL Server Data Mining offers Office 2007 data mining additives for the detection of data patterns and relationships. This is also helpful in improving the analysis. The add-in called Data Mining Client for Excel is used first to plan, compile, analyze, handle, and forecast data. In order to compare the performance of scheduling, different methods are used to mine the big data of tourism English talent

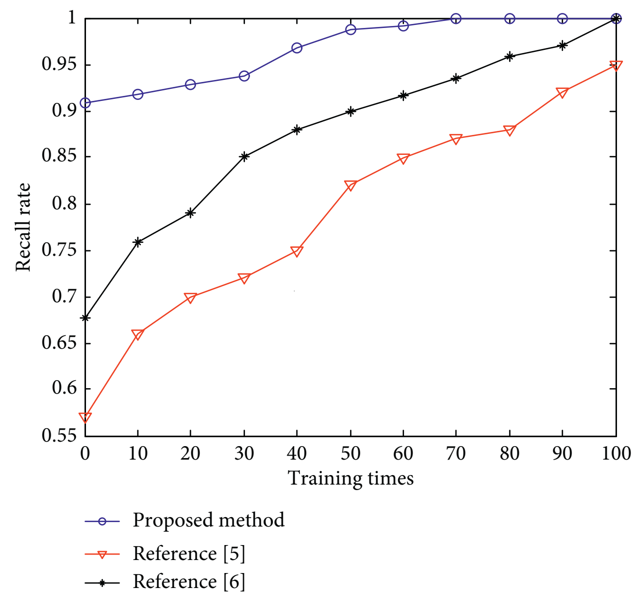


Figure 7: Big data analysis performance test.

resources distribution, and the recall test results are shown in Figure 7.

The results of the analysis in Figure 7 show that this method has good accuracy and big data clustering ability for tourism English talent resource data mining, which can effectively guide the construction of a new model for tourism English talent training.

5. Conclusions

In this paper, a new model of tourism English talent training based on big data is proposed, and the big data mining model of tourism English talents is constructed. Traditional methods use basic algorithms to predict the future. However, it does not provide reliable findings as compared to data mining. In the cloud storage system, the optimal scheduling design of tourism English talent training resource big data is carried out, and the fuzzy clustering method is used to deal with the adaptive clustering of tourism English talent resource distribution big data.

Moreover, the attribute distribution structure model is constructed to classify and fuse the features of tourism English talent resources, and the methods of association feature mining and information fusion are combined to share data and schedule resources in the process of talent training. Time-frequency features of tourism English talent resource distribution big data are extracted to realize the fast mining and positioning of the data. The simulation results show that this method has good accuracy and big data clustering ability, which can effectively guide the construction of a new model of tourism English talent training. This model has a good application value in promoting the promotion of tourism English talent training model. In the future, big data research needs to deal with the relationship between data and theory. The combination of data driving and theoretical driving is conducive to the benign development of the mutual promotion of theory and data.

Acknowledgments

The paper was supported by Anhui Province Overseas Study Visiting Project of Universities Excellent Talents Cultivation in 2017 (Grant no. GXFX2017237); Internet-Based Teaching Management Model Innovation Research, Anhui Vocational and Adult Education Association Education Scientific Research Planning Project (Grant no. AZJXH17156); Provincial Quality Engineering Practical English Excellent Open Course of Anhui Provincial Education Department (Grant no. 2017KFK136); Research on the Professional Development of Public English Teachers from the Perspective of Innovative Development of Higher Vocational Education, Key Project of Humanities and Social Sciences in Universities of Anhui Province (Grant no. SK2018A0832); Research on the Construction and Development of "2 + 2 + 2" Practical English Mixed Teaching Curriculum System in Higher Vocational Colleges (Grant no. SK2018A1170); Provincial Quality Engineering Teaching Research Project of Anhui Provincial Department of Education: Higher Vocational English Teaching Research Based on EOP + PE Ability Cultivation from NA Perspective (Grant no. 2019JYXM0613); and Provincial Quality Engineering Teaching Research Project of Anhui Provincial Education Department in 2020: Study on Practical English Course Ideological and Political Construction from the Perspective of "Taoism" (Grant no. 2020KCSZYJXM228).

References

[1] T.-z. Wen, A.-q. Xu, and G. Chneg, "Multi-fault diagnosis method based on improved ENN2 clustering algorithm," *Control and Decision*, vol. 30, no. 6, pp. 1021–1026, 2015.

[2] A. Kumar, R. Pooja, and G. K. Singh, "Design and performance of closed form method for cosine modulated filter bank using different windows functions," *International Journal of Speech Technology*, vol. 17, no. 4, pp. 427–441, 2014.

[3] N. Rajapaksha, A. Madanayake, and L. T. Bruton, "2D space-time wave-digital multi-fan filter banks for signals consisting of multiple plane waves," *Multidimensional Systems and Signal Processing*, vol. 25, no. 1, pp. 17–39, 2014.

[4] Y. Jiang, F. L. Chung, S. Wang, Z. Deng, J. Wang, and P. Qian, "Collaborative fuzzy clustering from multiple weighted views," *IEEE Transactions on Cybernetics*, vol. 45, no. 4, pp. 688–701, 2015.

[5] A. Bi and S. Wang, "Transfer affinity propagation clustering algorithm based on kullback-leiber distance," *JEIT*, vol. 38, no. 8, pp. 2076–2084, 2016.

[6] B. N. Alahmad and S. Gopalakrishnan, "Energy efficient task partitioning and real-time scheduling on heterogeneous multiprocessor platforms with QoS requirements," *Sustainable Computing: Informatics and Systems*, vol. 1, no. 4, pp. 314–328, 2011.

[7] H. Wang, H. Jin, J. Wang, and W. Jiang, "Optimization approach for multi-scale segmentation of remotely sensed imagery under k-means clustering guidance," *Acta Geodaetica et Cartographica Sinica*, vol. 44, no. 5, pp. 526–532, 2015.

[8] O. Shi-feng, Y. Gao, and Z. Xiao-hui, "Adaptive combination algorithm and its modified scheme for blind source separation," *Journal of Electronics & Information Technology*, vol. 33, no. 5, pp. 1243–1247, 2011.

[9] G. Tian, H. E. Ke-qing, J. Wang et al., "Domain-oriented and tag-aided web service clustering method," *Chinese Journal of Electronics*, vol. 43, no. 7, pp. 1266–1274, 2015.

[10] W. Tao, C. Lifei, and G. Gongde, "High-dimensional data clustering algorithm with subspace optimization," *Journal of Computer Applications*, vol. 34, no. 8, pp. 2279–2284, 2014.

[11] Y. Lei, X. Yu, S. Yue et al., "Research on PSO-based intuitionistic fuzzy kernel clustering algorithm," *Journal of Communication*, vol. 4, no. 5, Article ID 2015099, 2015.

[12] B. Zhang, H. Jie, G. Ma et al., "Mixture of probabilistic canonical correlation analysis," *Journal of Computer Research and Development*, vol. 52, no. 7, pp. 1463–1476, 2015.

[13] C. Sun, C. Yang, S. Fan et al., "Energy efficient distributed clustering consensus filtering algorithm for wireless sensor networks," *Information and Control*, vol. 44, no. 3, pp. 379–384, 2015.

[14] X.-y Guo, "Simulation and analysis on uncertain attenuation property of underwater acoustic signal for oil field pipe," *Computer Simulation*, vol. 31, no. 3, pp. 118–121, 2014.

[15] W. Zhang and Q. Chen, "Network intrusion detection algorithm based on HHT with shift hierarchical control," *Computer Science*, vol. 41, no. 12, pp. 107–111, 2014.

[16] H. Luo, Y. Qu, and Y. Yu, "Oscillation criteria of second order neutral delay emden-fowler equations with positive and negative coefficients," *Acta Mathematicae Applicatae Sinica*, vol. 40, no. 5, pp. 667–675, 2017.

[17] Z. Duo, G. Xu, X. Chen, and K. Yuan, "Rational non-hierarchical quantum state sharing protocol," *Computers, Materials & Continua*, vol. 58, no. 2, pp. 335–347, 2019.

[18] Q. He, S. Yu, H. Xu et al., "A weighted threshold secret sharing scheme for remote sensing images based on chinese remainder theorem," *Computers, Materials & Continua*, vol. 58, no. 2, pp. 349–361, 2019.

[19] F. Guo, P. Liu, W. Ren et al., "Research on the relationship between garlic and young garlic shoot based on big data," *Computers, Materials & Continua*, vol. 58, no. 2, pp. 363–378, 2019.

Optimization Algorithm of Tourism Security Early Warning Information System Based on Long Short-Term Memory (LSTM)

Lei Feng[1] and Yukai Hao[1,2]

[1]School of Management, Wuhan University of Technology, Wuhan 400070, China
[2]School of Economics and Management, Tibet University, Lhasa 850000, China

Correspondence should be addressed to Yukai Hao; haoyukai@whut.edu.cn

Academic Editor: Syed Hassan Ahmed

Tourism safety is the focus of the tourism industry. It is not only related to the safety of tourists' lives and property, but also related to social stability and sustainable development of the tourism industry. However, the security early warning of many scenic spots focuses on the response measures and remedial plans after the occurrence of security incidents, and the staff of many scenic spots have limited security awareness and information analysis ability, which is prone to lag in information release, and do not pay attention to the information of potential security problems. Therefore, this paper studies the optimization algorithm of the tourism security early warning information system based on the LSTM model and uses the recurrent neural network and LSTM to improve the processing and prediction ability of time-series data. The experimental results show that the number of three hidden layers in the tourism security early warning information system based on the LSTM model can reduce the training time of the model and improve the performance. Compared with the tourism safety early warning information system based on the BP neural network, it has better accuracy and stability, has better processing and prediction ability for time series data, and can monitor and analyze data scientifically in real-time and dynamically analyze data.

1. Introduction

Tourism has gradually become one of the important industries. As people no longer meet the basic needs of life, more and more people begin to pursue high-quality life. Tourism has gradually become one of the important industries. More and more people begin to pursue tourism life [1]. In addition, the tourism resources are constantly developed and utilized, and the tourism environment and content are constantly changing. In recent years, in addition to characteristic cultural city tourism, there are also natural and cultural tourism, marathon competition in natural tourism areas, etc. These emerging tourism projects not only attract more tourists but also improve the economic growth of the tourism industry [2, 3]. However, there are many unfortunate events that tourists encounter in the process of tourism, such as abrupt weather changes in marathon competitions in natural scenic spots, stampede on the Bund of Shanghai caused by too many tourists, tsunami, kidnapping of tourists, and constant theft in tourist areas, which have caused serious adverse effects on the tourism industry and restricted the sustainable development of the tourism industry [4, 5]. Therefore, tourism security has become a highly valued and concerned issue in various countries and regions, and tourism security early warning has become an inevitable trend of tourism development. With the development and application of intelligent information technology, many scenic spots will collect and release the safety information and early warning information of scenic spots through intelligent wearable products and corresponding apps based on big data, such as the information of dangerous areas of scenic spots and the number of visitors to scenic spots. Although the danger that some tourists will encounter in the scenic spots can be avoided to a certain extent, there are some problems in many scenic spots, such as untimely release of safety early warning information, low safety awareness of staff, and error in judgment of corresponding information. At the same time, the focus of many tourism safety measures, implementation plans, and methods in scenic spots is that after the occurrence of tourism safety events, the corresponding information release channel is narrow, and there is a lack of relevant knowledge reserve and mature and stable response plan in advance warning [6]. This shows that the development of tourism security early warning in the tourism industry can no longer meet the needs of the development of the tourism industry. Therefore, the tourism security early warning information system that the tourism industry needs to build can conduct accurate and scientific information analysis on the collected security information of relevant scenic spots in real-time and effectively, and output the information analysis results in time and improve the efficiency of safety early warning information in scenic spots.

This paper studies the optimization algorithm of the tourism security early warning information system based on the LSTM model. Compared with the traditional tourism security early warning methods, the artificial neural network has better fault tolerance and stronger robustness. It can quickly process data and find the corresponding optimal solution, and its nonlinear thinking can well deal with the relationship between many factors. Compared with the BP neural network, the LSTM model can better process temporal information and realize the purpose of real-time processing tourism safety early warning information. This paper is mainly divided into the following three parts. The first part introduces the development and related concepts of tourism security early warning information system and the development and application of the LSMT recurrent neural network. The second part constructs a tourism early warning information system based on the BP neural network and introduces the recursive neural network and LSTM to optimize the algorithm of the tourism early warning information system. In addition, the corresponding tourism security early warning information indicators are constructed by integrating various factors of tourism security. In the third part, the optimization algorithm of the tourism security early warning information system based on the LSTM model is trained and tested, and the simulation results are analyzed.

2. Related Work

The tourism security early warning information system contains a complex system of many influencing factors, which can evaluate and analyze various security indicators of tourism destination and determine the change trend of the system composed of the overall environment of tourism destination, so as to early warn and eliminate the security incidents that may occur in the security system [7, 8]. The tourism safety early warning information system can promote the sustainable development of tourism destination and improve the satisfaction of tourists' experience and personal safety and has important guiding significance in the long-term development of tourism industry and social economy, natural environment, and social stability [9]. Therefore, the research of the tourism security early warning information system has always been the focus of attention. Tourism safety factors are diversified, and their external manifestations can be roughly divided into natural disasters, diseases, crimes, traffic safety, and others. Many of them are uncontrollable, but scenic spots can still analyze some potential risk factors according to the analysis of relevant information. Some scholars have proposed an Intelligent Tourism early warning system for the stampede in scenic spots, that is, to analyze and guide the monitored data through intelligent services and processing functions [10]. This method is more suitable for use in urban scenic spots, and its early warning focuses on the tourism safety problems caused by human factors. According to the characteristics of natural scenic spots, some scholars proposed to establish the risk identification and evaluation model of natural scenic spots through the combination of the GIS and Bayesian network model [11]. This method has strong pertinence and can clarify the scope of risk and improve the accuracy of tourism safety early warning, but it needs long-term effective data as the basis of decision-making, which greatly increases the time cost. Some scholars proposed to build a safety early warning system based on the BP neural network. Its modeling is relatively simple and can obtain information analysis results in a relatively short time [12]. However, the BP neural network is weak in the analysis of time series information, and its output early warning results tend to static analysis. And with the increase of the types of risk factors, its accuracy is also affected to a certain extent. In addition, some research on tourism security early warning mostly focuses on the application mechanism of the artificial neural network in tourism security emergencies, which provides theoretical support and lays a solid theoretical foundation for tourism security research [13]. In addition, according to the current situation of tourism environment, researchers put forward to explore the ecological deterioration and sudden environmental security problems caused by tourism activities from an ecological perspective, predict the ecological environment status of tourism destinations, and make targeted preventive measures and rescue plans [14]. However, from the aspect of tourism security crisis early warning and management, the tourism security early warning system based on the BP neural network still has many deficiencies in processing time series data and needs to be further optimized.

The main objects of tourism security early warning system are tourists or local residents [15, 16]. Therefore, the information it provides is more detailed, which has a good effect in the security of outbound tourism. However, some travellers ignore early warning information or do not pay attention to relevant early warning information in time, and they do not pay attention to early warning information and suggestions [17]. It should be noted that the tourism safety early warning information system does not specifically establish a long-term safety early warning information system for tourism, but is issued by the relevant meteorological bureau, and the Safety Supervision Bureau and other departments carry out classified early warning for natural disasters and social security events, and the information subject is not limited to tourists [18, 19]. A similar tourism security early warning information system has been established, and the warning language is relatively mild. This paper briefly introduces the tourism destination countries, but it does not clearly classify the contents of early warning [20]. At the same time, citizenship does not connect tourism services, so few people pay attention to the released tourism security early warning information [21, 22]. Researchers have been constantly trying and studying, hoping to build a more scientific tourism security early warning information system [23].

2.1. Construction and Optimization of Tourism Security Early Warning Information System Based on LSTM. The tourism safety early warning information system is used to predict and warn the changes of scenic spots in the future from multiple dimensions according to the reasonable index system and scientific methods. Therefore, the influencing factors of tourism safety early warning information are diversified and nonlinear. In this paper, the LSTM model is used to construct the tourism safety early warning information system, which improves the processing ability of the system to temporal information, so as to realize the purpose of real-time dynamic information supervision and analysis. Figure 1 shows the flowchart of the tourism security early warning information system based on the LSTM model.

2.2. Construction of Tourism Safety Early Warning Neural Network Information System. The artificial neural network which simulates the connection between human neurons can process the relevant signals, obtain the data signal prediction model, and solve the nonlinear data prediction and other related problems. Therefore, this paper selects the BP neural network as the foundation of the tourism security early warning information system and extracts the implicit relationship of the static data that need to be analyzed and predicted [24]. The neurons of the BP neural network can connect multiple inputs but only have one output node, as shown in Figure 2.

The input layer of the multilayer perceptron is represented as L_{in} and $x = (x_1, x_2, x_3)^T$, the hidden layer as L_{hidden} and $h = (h_1, h_2, h_3, h_4)^T$, and the output layer as L_{out} and $y = (y_1, y_2, y_3)$. The process of data transmission from input layer to output layer is forward propagation, as shown in

$$h = f\left(W^{1T} x + b^1\right), \tag{1}$$

$$y = f\left(W^{2T} x + b^2\right). \tag{2}$$

The weight matrix of data transfer is expressed as $W^1 \in R^{3\times4}$; the weight of data transfer connection between hidden layer and output layer is expressed as $W^2 \in R^{4\times3}$, and the bias of output layer is expressed as $b^2 \in R^{3\times1}$. The activation function f is shown in

$$f = \frac{1}{1 + \exp(-x)}. \tag{3}$$

In the learning process of the BP neural network, the weights and thresholds are modified by the gradient method. The iterative formulas after the rest are shown in

$$\begin{cases} \Delta W(n+1) = -\eta \dfrac{\partial E}{\partial W(n)} + a \cdot \Delta W(n), \\ \\ \Delta \theta(n+1) = -\eta \dfrac{\partial E}{\partial \theta(n)} + a \cdot \Delta \theta(n), \end{cases} \tag{4}$$

$$\begin{cases} \Delta W(n+1) = W(n+1) - W(n), \\ \Delta \theta(n+1) = \theta(n+1) - \theta(n). \end{cases} \tag{5}$$

Although the BP neural network can solve the nonlinear problem, the data processed by the BP neural network have no correlation on the time line; that is, it cannot process the time series characteristic data related to the last moment or earlier data. An LSTM is designed based on the structure of the BP neural network. The structure of function of network memory, that is, it can effectively remember the data characteristics of time dimension in training data, and its structure is shown in Figure 3.

The difference between the LSTM and BP neural network is that the nodes between the hidden layers are interconnected, so that the input of the hidden layer at that time contains two parts, that is, the output transmitted by the input layer at this time and the output of the hidden layer at the last moment, so that the hidden layer contains the information memory at this time and at the last moment or earlier [25]. The transfer process is shown in

$$h_t = f\left(W^{1T} x_t + W^{hT} h_{t-1} + b^1\right), \tag{6}$$

$$y_t = f\left(W^{2T} h_t + b^2\right). \tag{7}$$

And b^2 represents the output layer offset. When the time is t, the input is x_t, the hidden layer output is h_t, and the output is y_t; when the time is $t-1$, the hidden layer output is h_{t-1}. The activation function formula is shown in

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{8}$$

The sum of the losses is the total loss function. Let the loss function be labeled as the negative log likelihood function, as shown in

$$L^t = -y^t \log o^t. \tag{9}$$

Then, the total loss function of the sequence is shown in

$$L\left(\left\{x^1, \ldots, x^t\right\}, \left\{y^1, \ldots, y^t\right\}\right) = \sum L^t. \tag{10}$$

The development of LSTM can be transformed into the corresponding BP neural network and trained by the BPTT algorithm [26, 27]. If it is necessary to train t time data, the LSTM is expanded into a BP neural network with t hidden layer. It can be seen from Figure 2 that the parameters in the same position after the expansion of the

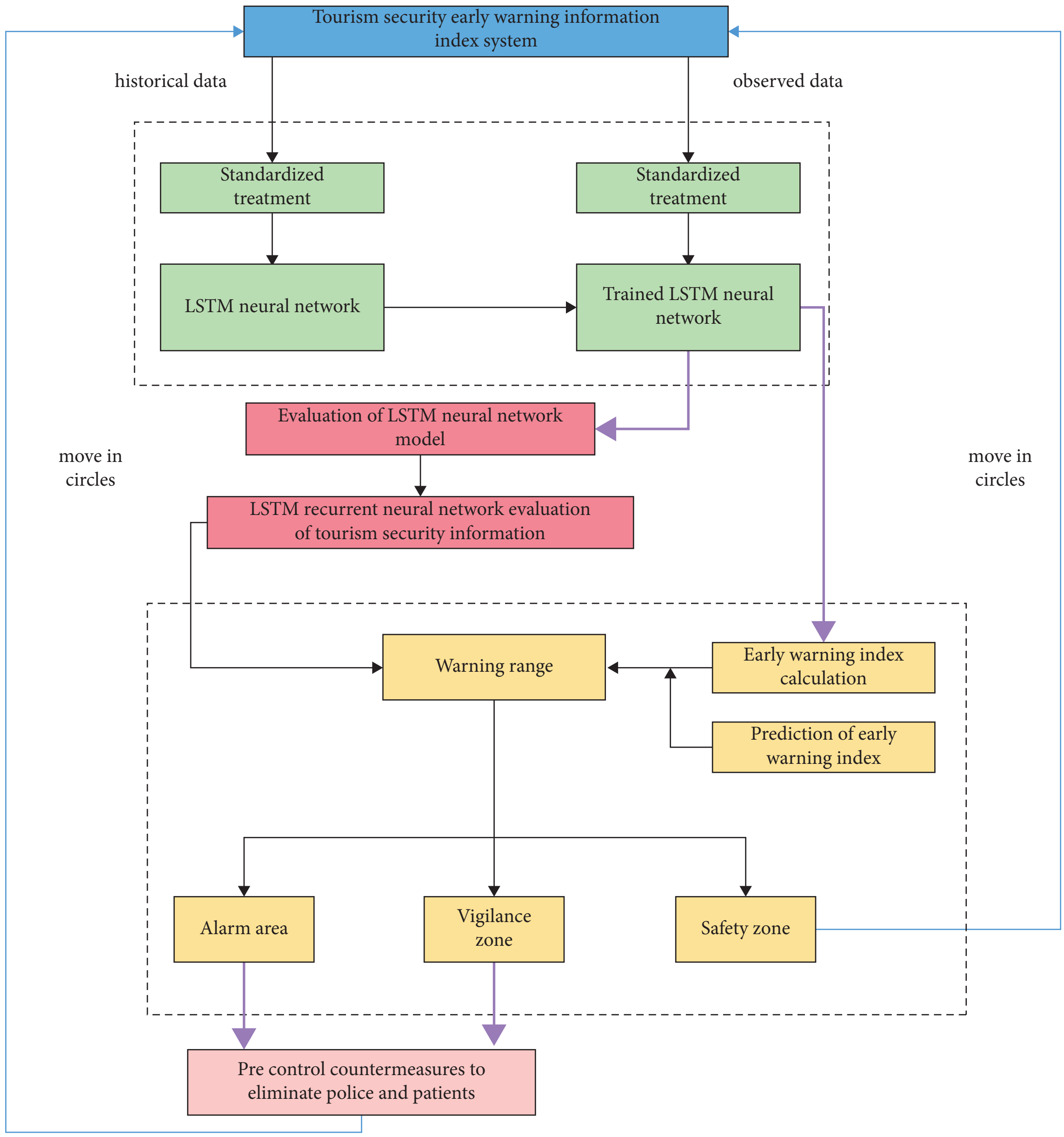


FIGURE 1: Flowchart of the tourism security early warning information system based on the LSTM model.

LSTM are shared with each other, while the BP neural network is not shared, so the LSTM greatly reduces the learning and training parameters. According to the relevant theory, it can be considered that the length of the sequence data that can be processed by LSTM is infinite, and the hidden layer of processing information is also infinite. Therefore, in the actual information transmission process, the information in the hidden layer neurons will gradually weaken and lose due to the extension of time, that is, the gradient vanishing phenomenon. This leads to the decrease of prediction performance. In addition, the length of the delay window needs to be determined in advance when the LSTM is trained, which improves the difficulty of automatically obtaining the optimal parameters in practical application.

2.3. Optimization of Tourism Security Early Warning Information System Based on LSTM. In order to solve the above problems of LSTM, so as to make its memory long-term, Figure 4 shows the LSTM structure development diagram.

It can be seen from the figure that the LSTM neural network introduces controllable, so as to avoid the gradient disappearance problem of LSTM. There are four layers in LSTM cell, including forgetting gate, input layer, output layer, and update layer, which interact with each other in a special way. And we provide corresponding continuous writing, reading, and reset functions. That is to say, the LSTM neural network adds a C state for long-term information memory on the basis of the recurrent neural network, which is the unit structure of LSTM. If the time at this time is t, when the forgetting gate layer controls the amount

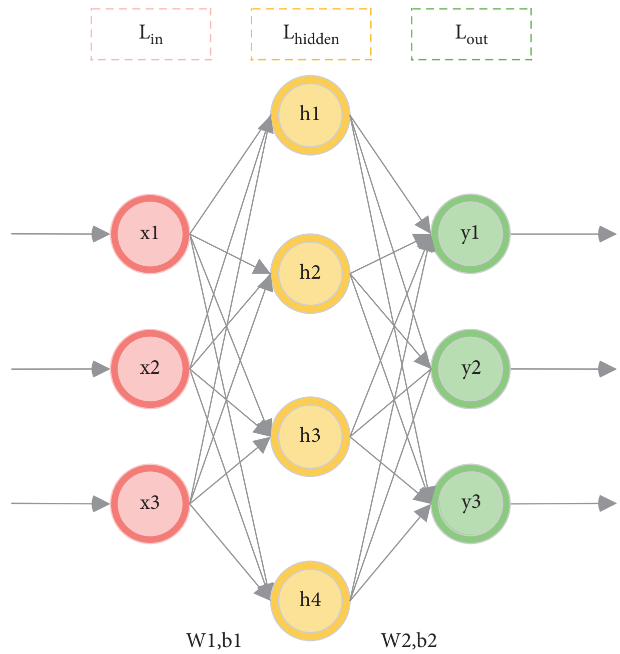


FIGURE 2: BP neural network simple model.

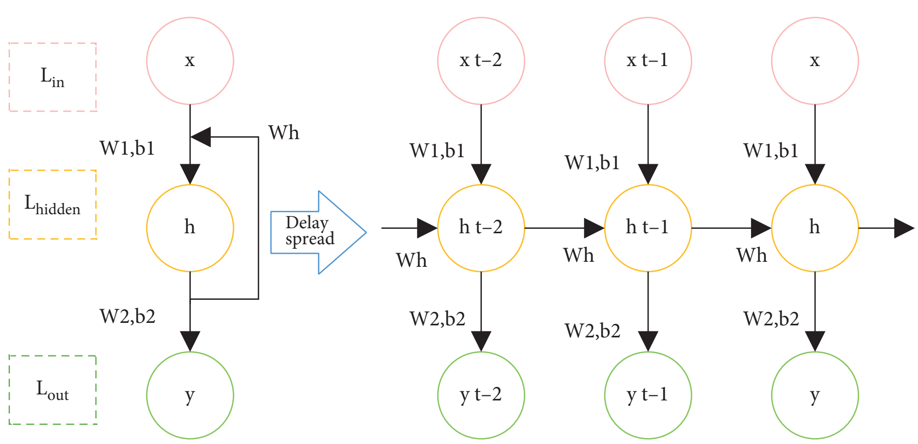


FIGURE 3: The structure of network memory function.

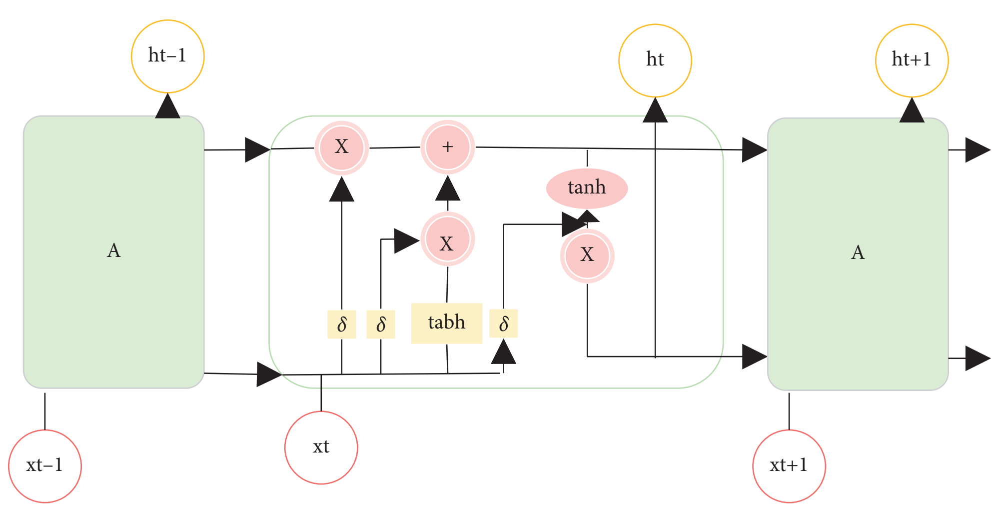


FIGURE 4: LSTM structure expansion.

of information transferred from the previous unit state c_{t-1} to the current c_t state,

$$f_t = \sigma\left(w_f * \left[h_{t-1}, x_{t-1}\right] + b_f\right). \tag{11}$$

The main purpose of the input gate is to filter information to avoid useless information entering the current c_t state. The sigmoid layer and tanh layer of the input gate can update the state. The formulas are shown in

$$i_t = \sigma\left(w_i + \left[h_{t-1}, x_t\right] + b_i\right), \tag{12}$$

$$g_t = \tanh\left(w_c * \left[h_{t-1}, x_t\right] + b_c\right), \tag{13}$$

where σ is sigmoid function and the numerical range is [0, 1]. After that, the last moment state c_{t-1} and f_t are multiplied to update the state. The useless information is filtered out, and a new value $i_t * g_t$ is added. The corresponding adjustment is made according to the individual update state. The formula is shown in

$$c_t = f_t * c_{t-1} + i_t * g_t. \tag{14}$$

The output gate is used to control the current output affected by long-term storage information. It mainly determines the cell to be output through sigmoid layer, then sets the cell in [−1, 1] range by tanh, and multiplies the corresponding output gate. Finally, the determined output part is output, as shown in

$$o_t = \sigma\left(w_0 * \left[h_{t-1}, x_t\right] + b_0\right), \tag{15}$$

$$h_t = o_t * \tanh\left(c_t\right). \tag{16}$$

It can be seen from the above formula that sigmoid function is the activation function of input, output, and forgetting gates with values in [0, 1]. The other activation function, tanh, as shown in formula (6), is also commonly used in the input and output gates of LSTM, and its monotonicity is more consistent with the characteristics of neurons in neural networks. Figure 5 shows two kinds of activation function diagrams.

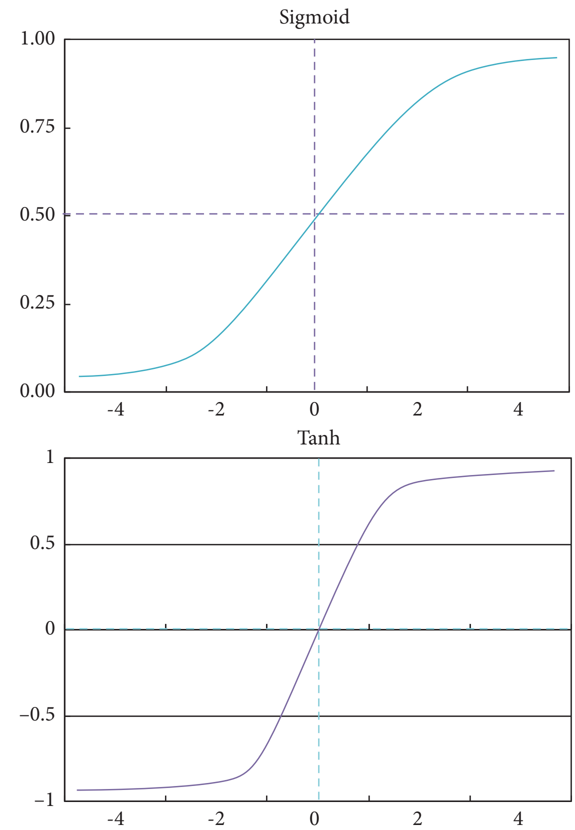


Figure 5: Two activation function diagrams.

2.4. Index Construction of Tourism Security Early Warning Information System. Tourism safety includes many influencing factors. According to the relevant analysis and induction, the index of tourism early warning information system in this paper is three levels, that is, the first level is tourism safety early warning, and the second level has four indicators, that is, the stability of tourism natural disasters, the safety of Tourism travel facilities, the safety of tourism destinations, and the safety of tourism environmental pollution. In addition, each first level indicator also contains three levels of impact factors.

Tourism safety early warning mode is divided into excellent level, good level, qualified level, and critical level. Among them, the excellent level indicates that the overall environment of the tourism destination has high security, there is no hidden danger, and there is no need to worry about the occurrence of emergencies. Good level means that the overall environment of the tourism destination has high security. Although there may be potential safety hazards and the possibility of small-scale security emergencies, the probability of occurrence is very small, and there are sound and mature treatment plans and remedial measures. From the perspective of realistic probability, potential tourism safety accidents may occur. However, the impact of such accidents can be effectively controlled within the corresponding range, and there is a good response plan, which requires tourists to have a certain degree of cognition and knowledge of potential safety hazards. Tourists who do not have this condition are not encouraged to enter the tourist destination. The critical level means that the tourism destination has a high probability of serious tourism safety accidents, and because there is no corresponding treatment plan and measures, once a safety accident occurs, it will have serious or even catastrophic consequences for tourists and the tourism destination. Tourists should be prevented from entering the tourism destination within this level. In order to show the four tourism safety early warning modes more intuitively, the four level alarms are matched with corresponding early warning signals, that is, the safety level early warning signal is green, the good level early warning signal is blue, the qualified level early warning signal is orange, and the critical level signal light is red. Figure 6 shows the warning value and discrimination mode of tourism security early warning.

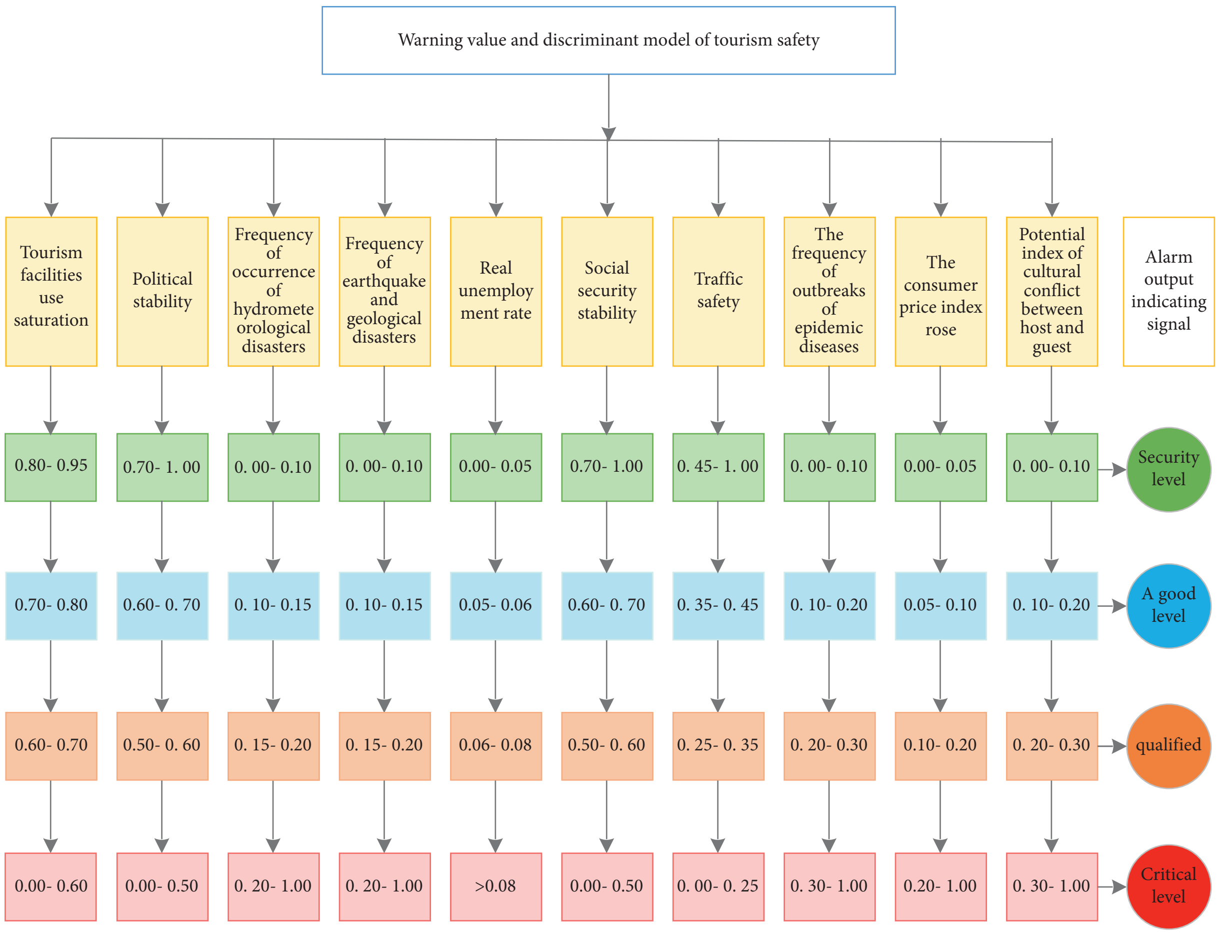


Figure 6: Warning value and discriminant model of tourism safety.

2.5. *Test Results of Tourism Security Early Warning Information System Based on LSTM*

2.5.1. Optimization Test Results of Tourism Security Early Warning Information System Based on LSTM. In the LSTM algorithm, the time step represents the length of the index sequence that can be used, which has a certain impact on the model. Therefore, under the condition that the algorithm remains unchanged, the performance of the algorithm with the step size of 4, 45, and 90 is tested, as shown in Figure 7.

It can be seen from the results in the figure that the LSTM algorithm will continuously improve the corresponding prediction performance with the increase of time step. When the time step increases to a certain length, the accuracy of LSTM algorithm decreases. In addition, the time step can reflect the correlation length of the data in the time series. If the time step is too short, the correlation information between the data will be insufficient, which will reduce the prediction effect of the algorithm. When the step length is too long, it will reduce the correlation between the data because of too much redundant data, thus reducing the prediction accuracy of the algorithm, so the selection of the step length algorithm is very important.

According to the LSTM recurrent algorithm, the increase of the number of layers will improve the learning

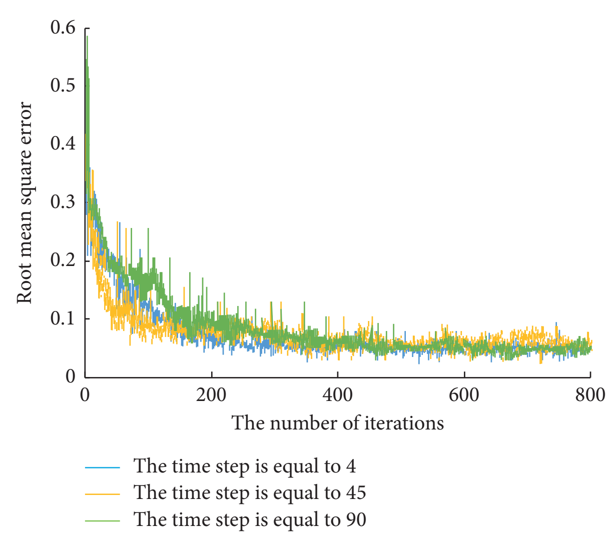


Figure 7: LSTM recursive neural network algorithm does not synchronize the long performance test.

performance, but layers will also lead to the improvement of the complexity of the algorithm system, affect its convergence speed, consume more time in the sample training, and increase the difficulty of training. Therefore, this paper tests

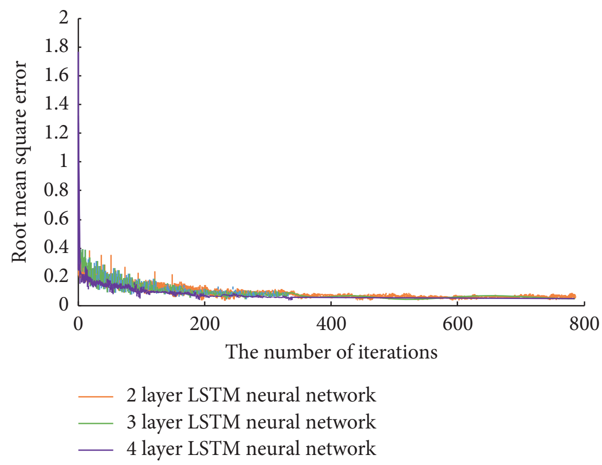


Figure 8: LSTM recursive neural network algorithm was used to test the level dependent performance.

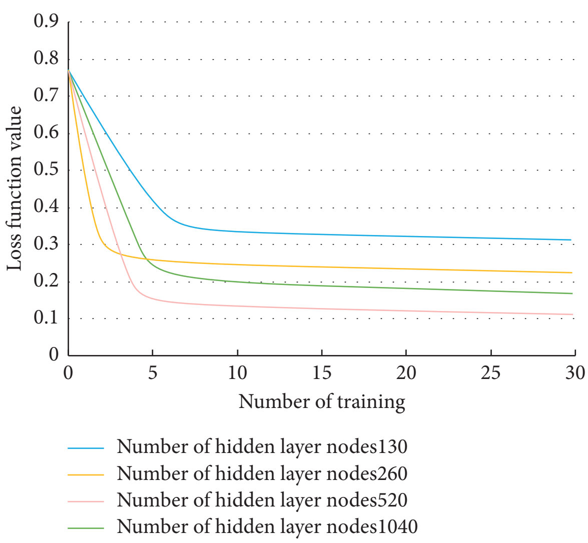


Figure 9: The hidden layer of the LSTM recursive neural network algorithm contains different numbers of node loss function values.

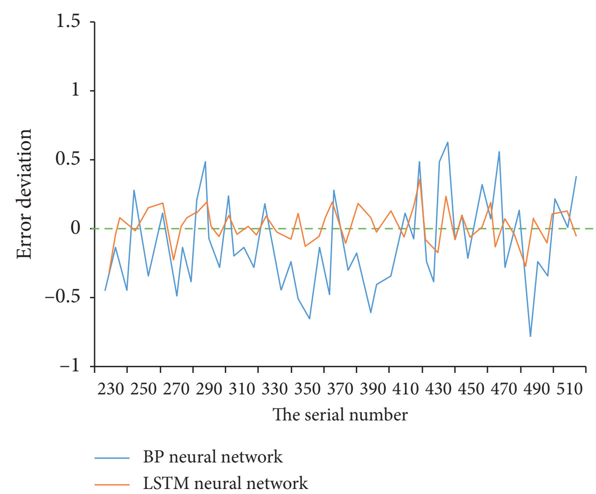


Figure 10: The error comparison graph of the BP neural network algorithm and LSTM recursive neural network algorithm for tourism safety index prediction results.

the performance of the LSTM algorithm, as shown in Figure 8.

In the figure that the convergence effect of LSTM improves with the increase of layers, but the corresponding training and testing time is also longer and longer. And when the number of layers of LSTM increases to four, the improvement of its performance is not obvious, but it takes a long time. Therefore, considering the needs of all aspects, the three-layer LSTM algorithm is the most appropriate. As shown in Figure 9, the hidden layer of LSTM algorithm contains different numbers of node loss function values.

It can be seen from the figure that when the number of hidden layer nodes reaches 520, the loss function value of the LSTM algorithm reaches the minimum. Compared with the loss function of the hidden layer with 130 and 260 nodes, it can be seen that with the increase of the number of nodes, the corresponding loss function value decreases significantly. This shows that when the number of hidden layer nodes is large enough, the fitting performance of the LSTM algorithm can be brought into full play.

2.6. Simulation Test Results of Tourism Security Early Warning Information System Based on LSTM. As shown in Figure 10, it is the error comparison chart of the BP neural network algorithm and LSTM algorithm for tourism safety index prediction results. In the results of the figure, the prediction result of the LSTM algorithm is closer to the real value. The algorithm has a large error in the prediction results of individual values, mainly because the BP neural network is prone to the problem of local optimal solution. Therefore, in terms of accuracy and stability, the prediction accuracy of LSTM for time series data is higher and the stability is better.

As shown in Figure 11, it is an early warning analysis model information system based on LSTM. Tourism destination security is a complex dynamic change, so the input value of the tourism security early warning information system can be not only discrete variables but also continuous

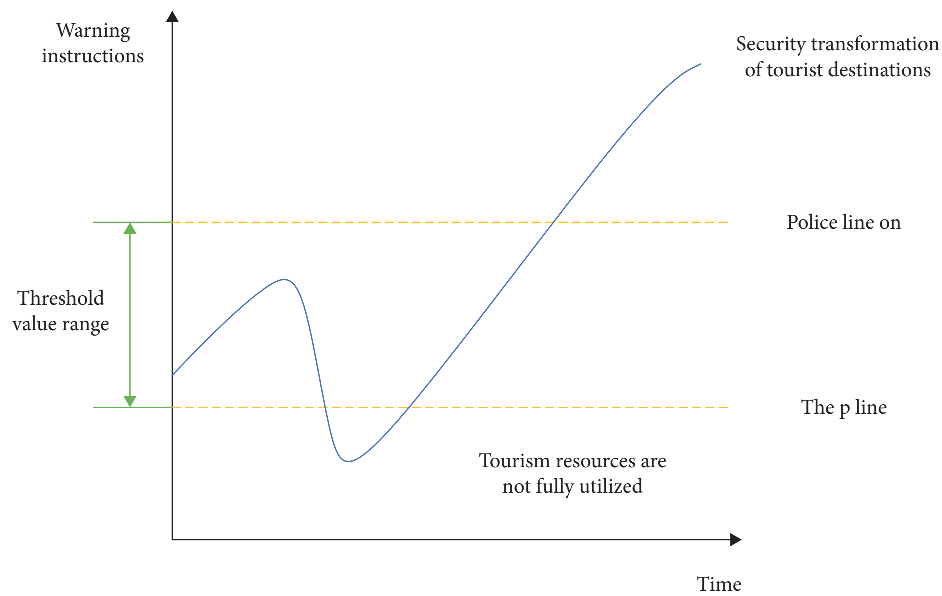


Figure 11: An early warning analysis model of the tourism security early warning information system based on the LSTM recursive neural network.

Table 1: Based on LSTM recursive neural network tourism security early warning information system simulation test results.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Alarm indication
1	0.95	0.74	0.089	0.06	0.02	0.72	0.44	0.04	0.03	0.12	(1000) green
2	0.87	0.61	0.06	0.11	0.05	0.67	0.36	0.15	0.06	0.11	(1000) green
3	0.74	0.66	0.14	0.16	0.08	0.67	0.44	0.20	0.09	0.16	(0100) blue
4	0.67	0.51	0.16	0.17	0.07	0.54	0.34	0.28	0.12	0.32	(0010) orange
5	0.82	0.67	0.08	0.11	0.05	0.68	0.39	0.15	0.06	0.14	(0100) blue
6	0.58	0.58	0.16	0.20	0.09	0.50	0.27	0.31	0.15	0.22	(0010) orange
7	0.62	0.48	0.27	0.20	0.12	0.50	0.25	0.36	0.22	0.32	(0001) red
8	0.57	0.41	0.21	0.22	0.15	0.42	0.28	0.44	0.27	0.51	(0001) red

variables, and the output value belongs to Boolean discrete vector.

The security status of tourism destination is divided into different levels, and the output value information system based on LSTM is set as a vector between 0 and 1. When the m-th index element represents 1 and the other index elements represent 0, the security of tourism destination is in a certain level. The simulation test results the tourism security early warning information system as shown in Table 1.

3. Conclusion

With the continuous development of economy, people begin to see the difference of the world through tourism on the basis of meeting the basic life. However, what is not matched with the booming tourism industry is the tourism security early warning information system. Tourism security is a comprehensive problem composed of many factors, which is not only related to the life and property safety of tourists, but also related to social stability, and the development and protection of tourism resources. At present, the tourism is relatively backward, focusing on the remedial measures and treatment after the occurrence of security incidents, which cannot play the role of early warning to reduce disaster losses. Therefore, this paper studies the optimization algorithm of the tourism security early warning information system based on LSTM. On the basis of the tourism security based on the BP neural network, it uses recurrent neural network and LSTM to optimize the system algorithm, so as to improve the ability of the early warning information system to process and predict the time series data. The experimental results show that the learning ability and convergence effect of LSTM model will improve with the increase of the number of hidden layers, but when it increases to a certain number, the increase of learning ability and convergence effect is not obvious. Therefore, it is necessary to set an appropriate number of hidden layers for the LSTM model to improve its performance. The tourism security early warning information system based on the LSTM model has better accuracy and stability than the tourism security early warning information system based on the BP neural network algorithm, has better processing and prediction ability for time series data, and is more in line with the needs of the tourism security early warning information system. In addition, compared with other methods, the tourism security early warning information system based on the LSTM model can be applied to a wider range, whether it is a tourist city scenic spot or a tourist natural scenic spot, or it can be combined with intelligent wearable devices for data collection and analysis. However, the experimental data in this paper are mainly for the analysis of the indicators of the scenic spot, so the index system needs to be further improved. In the future development, the tourism safety early warning information system between scenic spots should be connected with each other to strengthen the information circulation. At the same time, set up a tourism safety early warning information subsystem for economically underdeveloped scenic spots to reduce the cost of tourism safety early warning information system on the basis of ensuring the safety of scenic spots and tourists.

Acknowledgments

This work was supported by the Ministry of Science and Technology of the People's Republic of China, project number: 2018YFB0804300.

References

[1] H. Benbrahim, H. Hachimi, and A. Amine, "Deep transfer learning with Apache spark to detect covid-19 in chest x-ray images," *Romanian Journal of Information Science and Technology*, vol. 23, no. S, pp. S117–S129, 2020.

[2] P. Zhao, Y. Liu, H. Liu, and S. Yao, "A sketch recognition method based on deep convolutional-recurrent neural network," *Journal of Computer-Aided Design & Computer Graphics*, vol. 30, no. 2, pp. 217–224, 2018.

[3] R. McKinley, R. Wepfer, F. Aschwanden et al., "Simultaneous lesion and brain segmentation in multiple sclerosis using deep neural networks," *Scientific Reports*, vol. 11, no. 1, pp. 1087–1111, 2021.

[4] R. A. Bhuiyan, S. Tarek, and H. Tian, "Enhanced bag-of-words representation for human activity recognition using mobile sensor data," *Signal, Image and Video Processing*, vol. 2021, Article ID 1907-4, 8 pages, 2021.

[5] H. N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing internet of things: opportunities, challenges and enabling technologies," *Enterprise Information Systems*, vol. 14, no. 9-10, pp. 1279–1303, 2020.

[6] G. Gui, Z. Zhou, J. Wang, F. Liu, and J. Sun, "Machine learning aided air traffic flow analysis based on aviation big data," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4817–4826, 2020.

[7] B. M. H. Abidine, L. Fergani, B. Fergani, and M. Oussalah, "The joint use of sequence features combination and modified weighted SVM for improving daily activity recognition," *Pattern Analysis & Applications*, vol. 21, no. 1, pp. 119–138, 2018.

[8] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," *Mobile Networks and Applications*, vol. 25, no. 2, pp. 743–755, 2020.

[9] A. Amaya, P. P. Biemer, and D. Kinyon, "Total error in a big data world: adapting the TSE framework to big data," *Journal of Survey Statistics and Methodology*, vol. 8, no. 1, pp. 89–119, 2020.

[10] F. Y. Zhou, L. P. Jin, and J. Dong, "A review of convolutional neural networks," *Chinese Journal of Computers*, vol. 40, no. 6, pp. 1229–1251, 2017.

[11] Z. Xu, C. Cheng, and V. Sugumaran, "Big data analytics of crime prevention and control based on image processing upon cloud computing," *Journal of Surveillance, Security and Safety*, vol. 1, no. 1, pp. 16–33, 2020.

[12] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2298–2304, 2017.

[13] X. Li, Z. Zhao, and F. Liu, "Big data assimilation to improve the predictability of COVID-19," *Geography and Sustainability*, vol. 1, no. 4, pp. 317–320, 2020.

[14] M. Geist, P. Petersen, M. Raslan, R. Schneider, and G. Kutyniok, "Numerical solution of the parametric diffusion equation by deep neural networks," *Journal of Scientific Computing*, vol. 88, no. 1, pp. 1–37, 2021.

[15] M. Mathew, D. Karatzas, and C. V. Jawahar, "Docvqa: A dataset for vqa on document images," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2200–2209, Waikola, HI, USA, August 2021.

[16] A. E. R. ElSaid, J. Karns, Z. Lyu, D. Krutz, A. Ororbia, and T. Desell, "Improving neuroevolutionary transfer learning of deep recurrent neural networks through network-aware adaptation," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, pp. 315–323, Prague, Czech Republic, March 2020.

[17] A. Vernotte, M. Välja, M. Korman, G. Björkman, M. Ekstedt, and R. Lagerström, "Load balancing of renewable energy: a cyber security analysis," *Energy Informatics*, vol. 1, no. 1, 5 pages, 2018.

[18] M. Kend and L. A. Nguyen, "Big data analytics and other emerging technologies: the impact on the Australian audit and assurance profession," *Australian Accounting Review*, vol. 30, no. 4, pp. 269–282, 2020.

[19] J. M. Johnson and T. M. Khoshgoftaar, "The effects of data sampling with deep learning and highly imbalanced big data," *Information Systems Frontiers*, vol. 22, no. 5, pp. 1113–1131, 2020.

[20] A. T. Fadi and B. D. Deebak, "Seamless authentication: for IoT-big data technologies in smart industrial application systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2919–2927, 2020.

[21] R. Feng, D. Grana, and N. Balling, "Uncertainty quantification in fault detection using convolutional neural networks," *Geophysics*, vol. 86, no. 3, pp. M41–M48, 2021.

[22] C. Baskin, N. Liss, E. Schwartz et al., "Uniq: uniform noise injection for non-uniform quantization of neural networks," *ACM Transactions on Computer Systems*, vol. 37, no. 1–4, pp. 1–15, 2021.

[23] A. Skolik, J. R. McClean, M. Mohseni, P. Van Der Smagt, and M. Leib, "Layerwise learning for quantum neural networks," *Quantum Machine Intelligence*, vol. 3, no. 1, pp. 1–11, 2021.

[24] M. Aboelmaged and S. Mouakket, "Influencing models and determinants in big data analytics research: a bibliometric

analysis," *Information Processing & Management*, vol. 57, no. 4, Article ID 102234, 2020.

[25] J. L. Leevy and T. M. Khoshgoftaar, "A survey and analysis of intrusion detection models based on cse-cic-ids2018 big data," *Journal of Big Data*, vol. 7, no. 1, pp. 1–19, 2020.

[26] A. Glowacz, "Fault diagnosis of electric impact drills using thermal imaging," *Measurement*, vol. 171, Article ID 108815, 2021.

[27] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the internet of things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

The Construction Path and Mode of Public Tourism Information Service System Based on the Perspective of Smart City

Hongyan Ma [1,2]

[1]School of Humanities and Social Sciences, Xi'an Jiaotong University, Xi'an, Shaanxi Province 710049, China
[2]School of Business, Xi'an Fanyi University, Xi'an, Shaanxi Province 710105, China

Correspondence should be addressed to Hongyan Ma; maria78@stu.xjtu.edu.cn

Academic Editor: Wei Wang

Under the framework of smart city, starting from the demand for urban public tourism information services, drawing on the new public service theory, customer perceived value theory, and basic information service theory, combined with previous research results, using literature analysis, questionnaire survey, and other methods, and starting from the carrier level of the public tourism information service system, this study analyzed the public tourism information services in detail. This study combed the current status quo and problems of the public tourism information service system for qualitative analysis and summed up some common shortcomings and deficiencies. Along with the results of the questionnaire, this paper analyzed the existing problems in the current public tourism information service system and identified the demands of tourists for public tourism information services. Based on this, the research puts forward the idea of constructing the public tourism information service system under the background of smart city, including the system framework, characteristics, service media and mode, operation mode, management mechanism, and other aspects. The study concluded that there is a need to construct theoretical and research models of the impact of public tourism services on the quality of destination from the perspective of smart city, to use the structural equation model for quantitative analysis, to verify the model hypothesis, and to consider that public tourism services have a significant positive impact on the behavioral and psychological responses of tourists. On the basis of combing the literature and starting from the overall situation, this study puts forward a systematic integration idea, which is of theoretical significance.

1. Introduction

The rapid development of information technology has laid an important material foundation for economic and social development, made processing and dissemination of information more convenient, made the media more diversified, changed the traditional service model, extended the industrial chain, and stimulated some new industries [1]. Thus, information resources have increasingly become an important factor of production, integrating with other industries. Tourism industry is one of the information-intensive industries [2]. The comprehensiveness and nonstorability of tourism products make informatization important, especially for tourism enterprises. Since the 1980s, information technology has been applied to business operations of China's tourism enterprises. After the rise of a large number of travel e-commerce sites, the independence of the Tourism Information Center and the National Tourism Administration "Golden Travel" is represented by "Ctrip Travel Network." The implementation of the "project" has laid to a solid foundation for China's tourism informatization construction [3].

The development of informatization has opened up a new development path for tourism. First, the promotion of tourism through online marketing is different from the traditional passive model that solely relies on travel agencies and word-of-mouth. It actively promotes tourism resources by providing information to every potential tourist who needs relevant information through the Internet and expands the popularity of tourism resources; second, the construction of tourism informatization enhances the reception capacity of tourist destinations and increases the satisfaction of tourists; then, a large number of tourism e-commerce sites have emerged in the construction of tourism informatization, seize the gaps in market demand, develop business management models, and enrich the tourism market. In addition, tourism informatization also plays an important role in protecting tourism resources, facilitating tourism services, protecting the rights and interests of tourists, and promoting the development of tourism products. It has become an important way for the transformation and upgrading of the tourism industry [4].

In the context of the rapid advancement of information technology and the boom of smart city construction, this study combs the status quo and problems of public tourism information service system, analyzes and evaluates the demand of tourists for public tourism information services, and combines the concepts and practices of smart city construction [5]. This study also provides ideas for constructing a public tourism information service system for smart cities, provides new directions for the transformation and upgrading of tourist cities, and lists some domestic cities as typical cases. It also analyzes their construction practices in the public tourism information service system in the process of smart city construction and identifies deficiencies and provides suggestions for countermeasures [6]. It provides important theoretical exploration and practical guidance value. The innovation of this research lies in (1) putting the tourists' demand perspective on the dominant position of the public tourism information service system. The public tourism information service system has public attributes, and general research focuses on the supply subject and supply mechanism. Less attention has been paid to this demand, which has caused repeated construction of the public tourism information service system and waste of resources. On the other hand, it has also lowered the satisfaction of tourists, especially individual tourists, which has affected the development of the local tourism industry [7]. This research starts from the reality and needs to be reflected in the questionnaire and is based on the perspective of tourists. (2) The theoretical model and the research model of the impact of tourism public services on destination brand relationship quality are constructed. A total of 36 factors in three dimensions forming a rectangular factor relationship structure model in the actual work of the tourism industry, around its development scale, and exploring new paths and methods for future research are constructed.

2. The Rise of Smart Cities and the Transformation of Tourist Cities

As early as 1992, Singapore first proposed the construction of a smart island, focusing on enhancing the city's information technology foundation [8]. In 2008, Industrial BM Company proposed the concept of "smart earth," thinking that "the world is becoming smaller, flatter, and smarter." Thus, the concept of "smart city" has also evolved. "Smart city" is built on the basis of "digital city," along with the wisdom concept of contracting BM, using advanced information technology for convenience and to communicate and operate more efficiently, advocate economic health and reasonable sustainability, harmonious and safe, comfortable life, and intelligent information management technology [9].

As soon as the smart city was proposed, it quickly gained wide attention and recognition. The governments of the United States, Europe, Singapore, South Korea, Japan, Hong Kong, and Taiwan launched smart city construction [10]. For example, as early as 2006, "Garden City" Singapore launched a ten-year "Smart Country 2015" (iN2015) plan; in July 2009, the Japanese government proposed after "e-Japan" and "u-Japan" [11] the updated version of the national informatization strategy, the medium and long-term information technology development strategy "i-Japan (Smart Japan) Strategy 2015," and designated Yokohama City in Kanagawa Prefecture, Toyota City in Aichi Prefecture, Toyota City in Kyoto Prefecture, and Toyota City in Fukuoka Prefecture. Four districts in Kitakyushu City plan to pilot smart cities; in October 2009, the European Union announced a new energy research investment plan, which will invest 11 billion euros for the "smart city" project and develop low-carbon housing in 25–30 cities and transportation [12].

2.1. The Construction of Smart City. The construction of smart cities in China presents four distinctive characteristics: first, the development of smart cities is highly valued and the enthusiasm for participation is increasing; hence, the number is increasing; second, the overall development among the cities is uneven, and the cities in the eastern coastal areas have a clear lead; third, the foundation of smart governance is solid, and the smart industry and infrastructure are steadily advancing. Finally, the government's service capabilities are more prominent, but the construction of typical applications and organizational systems lags significantly [13]. The rise of smart city construction around the world has brought a qualitative leap to the improvement of the city's overall informatization level. From the "smart city construction planning" and related documents issued by various cities, the basic construction of smart cities can be seen. The above all follow a certain model, combining the current status and long-term development planning of the city itself and constructing the three major systems of "bottom base layer-intermediate exchange layer-upper application layer" from the bottom up, which integrates citizens, tourists, enterprises, and governments in the city. Changing the traditional and backward operation and service modes of various roles, the intelligentization of life, production, and management is realized. The basic architecture of a smart city is shown in Figure 1.

With the transformation of China's economic development mode and the tourism industry, especially the increasing personalization and diversified needs of tourists, China's outstanding tourism cities have shown serious shortages in the reception capacity of individual tourists; at the same time, in the process of urbanization, problems such as environmental pollution, waste of resources, and population expansion have also spurred the transformation of China's outstanding tourism cities in search of new development methods.

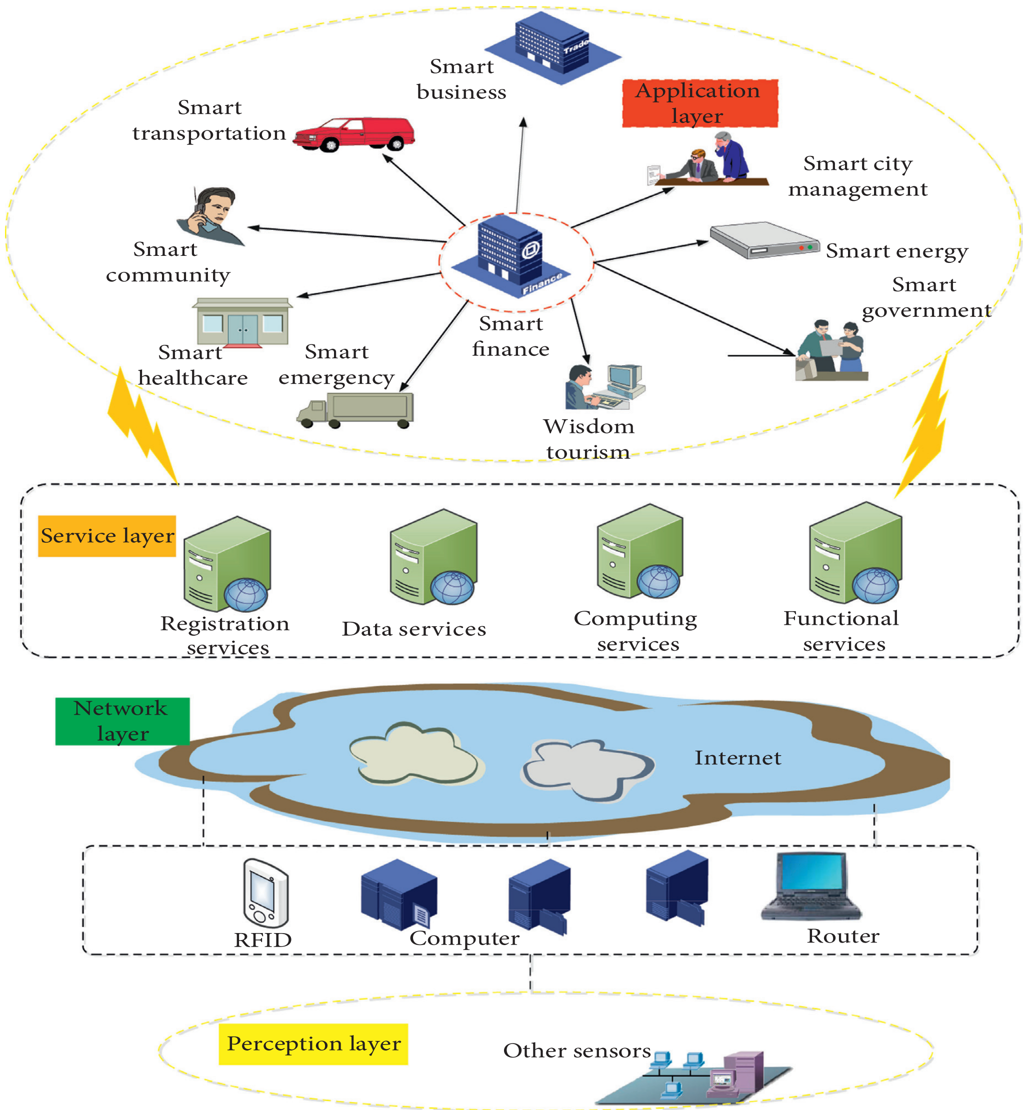


FIGURE 1: Basic architecture of a smart city.

The purpose of the transformation of China's outstanding tourism cities is to further optimize the urban tourism environment and improve urban tourism functions and the level of urban tourism management and service quality. But, its essence is to promote urban construction by expanding tourism consumption demand and to improve the city's opening degree, civilization quality, and international level [14]. In order to pursue transformation and achieve sustainable development, tourism cities have also been actively explored. Development paths such as ecologicalization, cultural creativity, low-carbon type, and informatization have become alternative directions and breakthroughs in the transformation of many tourist cities. The proposal of smart city provides a new path for the transformation and upgrading of tourist cities.

Smart transformation is in line with the general trend of tourism informatization development and is a concentrated expression of the integrated development of tourism and information technology industries. With the help of intelligent concepts and means, the conditions and support that the tourism industry of an excellent tourist city depends on can be comprehensively and systematically improved, thereby creating a good environment for the development of tourism.

2.2. Construction of Smart City Information Service Platform. In the construction of smart cities, in order to complete the storage and processing of massive data, it is necessary to build a processing platform that carries all applications. This platform has become the "smart city core platform." The smart city core information service platform promotes services such as triple play, office integration, intelligent interconnected buildings, digital media, city management, smart cards, collaboration, and centralized operation centers in areas such as residence, business, medical, education, and government affairs. Smart work spaces, smart transportation, smart buildings, smart energy, and smart society are connected together to form an overall smart system, as shown in Figure 2.

A cloud computing integrated platform with powerful data analysis capabilities is required to realize calculation,

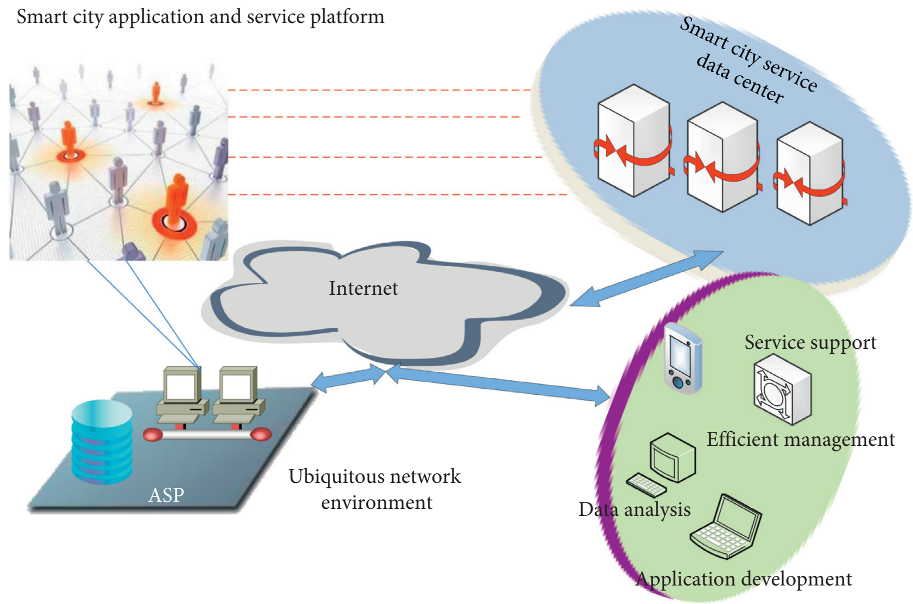


Figure 2: Smart city information service platform construction.

storage, analysis, processing, and decision-making of massive data. The "Smart City Core Platform" is different from ordinary industrial DCs and cloud computing centers and can achieve the following functions: providing running server hosting for the operation of all application systems and data space for the storage of all massive data. To meet the needs of city management, the core platform must be connected to the main administrative departments of the government; to meet the needs of citizens services, the core platform must be logically isolated from the Internet in order to meet the needs of information services. Various channels such as the Internet have established interaction mechanisms with citizens. Based on the Internet, using various information technologies, under the ubiquitous network environment, through the efficient management and application development of the city, the integrated application of the smart city service platform and smart city service data center is realized.

2.3. Construction Model Based on Internet of Things. As an application of the Internet, the Internet of Things is regarded as an important symbol of a "smart city." The country has proposed the Internet of Things strategy. After more than two years of promotion and development, the Internet of Things is no longer an abstract concept. The government and enterprises have gradually become more rational and clear in understanding the Internet of Things and have accumulated some construction experience. The core of the smart city is the construction of the Internet of Things system, so the smart city information service platform is divided into three levels: "perception layer," "network layer," and "application layer." Among them, the "perception layer" is to assemble sensors to real objects such as power grids, automobiles, buildings, household appliances, and water supply systems to a certain degree so as to achieve the connection of objects; the "network layer" is connected to wireless networks through interfaces and operates [15]. Through the integration of the city's unified network with telecommunications networks, radio grids, and the Internet, the city's unified network's basic environment and effective information transmission service highway are provided to the smart city public platform and various business application systems. The "network layer" is divided into "nearby," "transmission," and "remote transmission," which has two sublevels; the application layer is an application service system built on the smart city public platform, connecting the three major application sources of government, enterprises, and the public and providing them with a livable urban environment, security prevention and control, application services in areas such as life security, public services, and industrial optimization. The smart city construction model based on the Internet of Things is shown in Figure 3.

The industrialization of the Internet of Things and the development of R&D applications are accelerating, speeding up the realization of real-time control and precise management of the world of the Internet of Things, and its ability to support economic operations and carry social public services is further improved and expanded. Through intelligent information equipment, efficient integration, interconnected ubiquitous networks, education, medical care, social management services, entertainment, socializing, etc., the necessity of life from "things" to "people" is as the center, from one-way diffusion shifts to the direction of participation and interaction. A new lifestyle with a wider range, a smarter, and a better experience gives people fingertips.

3. Analysis of the Demand Survey Results of Public Tourism Information Service System

The public tourism information services must rely on certain carriers to realize its service function, and its service efficiency is often reflected by the service level of these carriers in reality [15]. The carriers of public tourism information

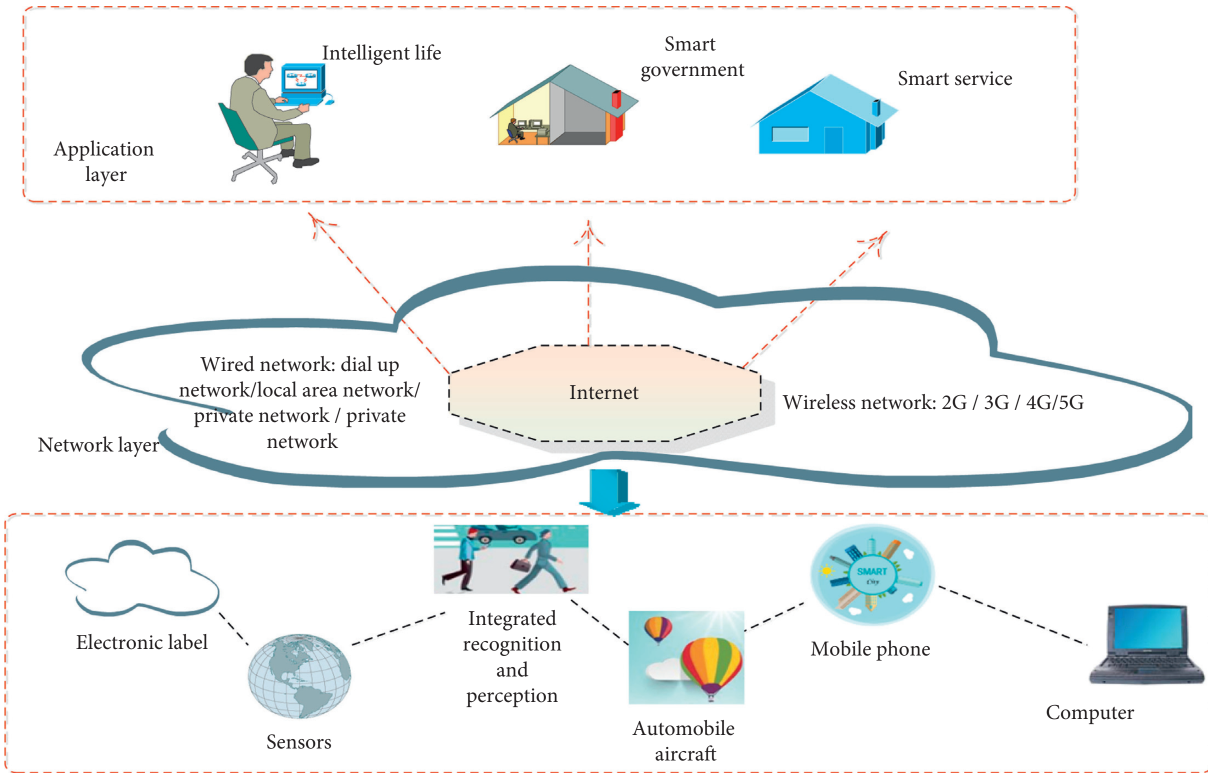


Figure 3: IoT-based construction model.

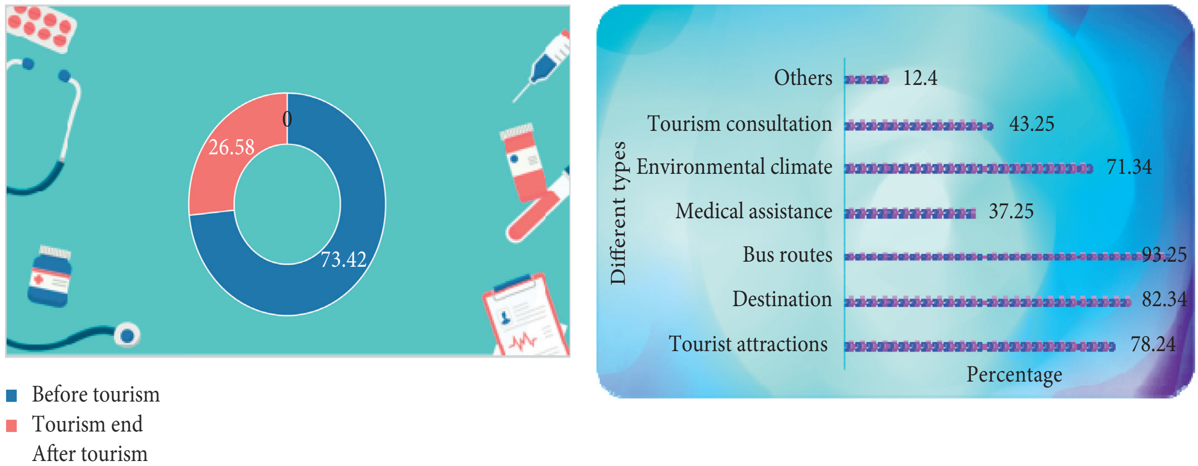
services generally include website (official website group of government part of tourism destination and public service information of commercial tourism website), tourism distribution center (naturally formed tourism distribution center and constructed tourism distribution center), the tourism consulting service site (the free consulting site set by tourism enterprises and the consulting site set by the Tourism Management Department), and the tourism call center. Reflecting the basic situation of the number, function, and intelligence of public tourism information service carriers, we can see the basic level and existing problems of public tourism information services [16].

3.1. Analysis of Tourism Information Demand. The demand for public tourism information services is high, accounting for 73.42%, but there is also information demand for tourism, accounting for 26.58%, nearly a quarter (Figure 4(a)). Therefore, information services in the process of tourism activities are also necessary.

Among the respondents, food, accommodation, shopping, bus routes, tourist attractions, language, and folk customs are still the main contents corresponding to the six elements of traditional tourism, such as food, accommodation, travel, shopping, and entertainment. However, 71.34% accounted for environmental and meteorological information (Figure 4(b)). It also becomes an important concern. However, the popularity of information content is very high, accounting for 55.25%. Only 13.24% of the respondents preferred to customize the content of public tourism information services, and the choice between the two accounts for 32.51% (Figure 5(a)). On the one hand, it is consistent with the public attribute of the public tourism information services, and it also shows that tourists still have a certain tendency towards customized services and do not want to fully popularize them.

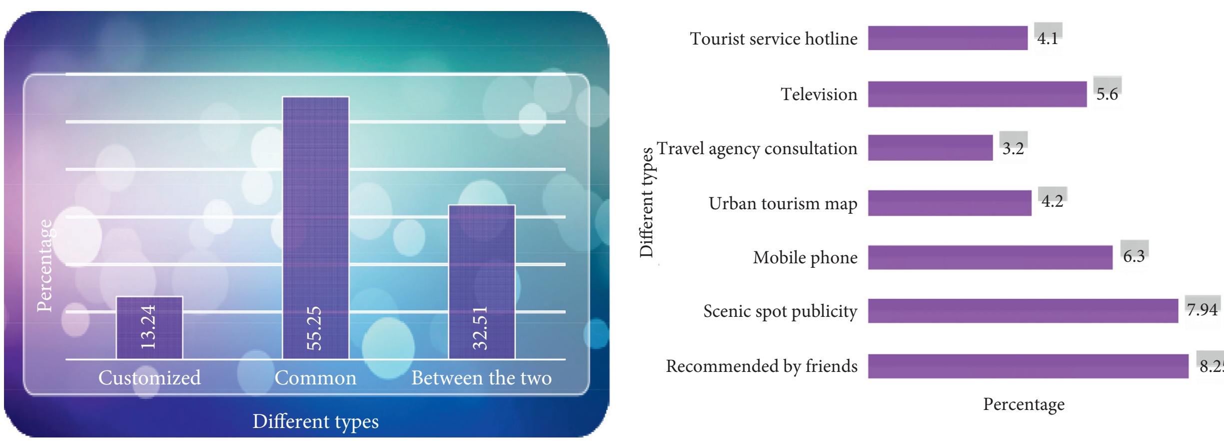
3.2. Analysis of the Ways and Forms of Tourists' Demand for Tourism Information Services. Among the existing public tourism information service channels, tourists have the strongest perception of the importance of the Internet, reaching 10.5%. Mobile phones rank fourth, while the tourism service hotline ranks last (Figure 5(b)). It can be seen that tourists have a strong demand for tourism information on the Internet and mobile phones, but although the tourism service hotline has been opened nationwide, it does not get the attention of tourists. In contrast, tourists like the form of network interaction the most, followed by on-site consultation and telephone consultation (Figure 6). In the survey, 88.5% of the respondents have such demands, so the demands in this regard cannot be ignored.

3.3. Reliability Test and Data Purification. After ensuring the validity, trend, and distribution of sample data meet the requirements, this study also needs to test the reliability and validity of the scale used in empirical analysis. In the reliability test of the scale of the impact of public tourism services on the quality of destination brand relationship, this study uses the widely used corrected-item total correlation (CITC) index and internal consistency Cronbach's alpha coefficient to analyze.



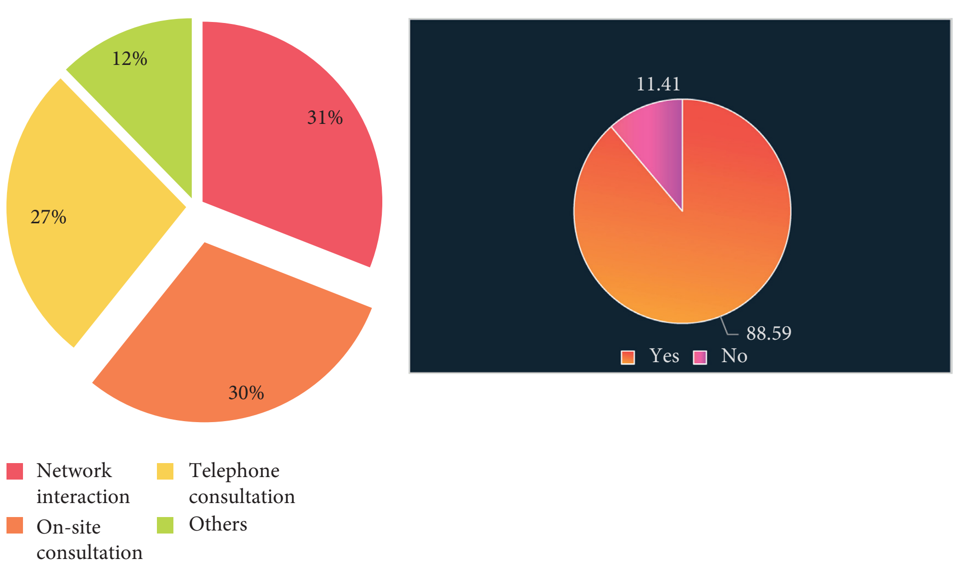
(a) (b)

Figure 4: (a) Distribution stage of tourism information demand. (b) Content analysis of tourism information demand.



(a) (b)

Figure 5: (a) Universality analysis of tourists' demand for tourism information services and (b) ranking of importance of tourism information service channels.



(a) (b)

Figure 6: (a) Ranking of tourists' preference for consulting services and (b) investigation of tourists' mood sharing needs.

In this study, after reliability analysis, the main standards used for data purification are as follows: corrected-item total correlation index (CITC) to purify the data and each item. When the strict CITC standard *a* is less than 0.5, the measurement item should be deleted, but it is also a relatively loose standard. For example, Lu Wendai considered that if the CITC index of the item is greater than 0.3, and the line with the requirements need further analysis. In this study, the scale of CITC is selected as 0.5, the test standard of Cronbach's α coefficient [17]. Bashynska considers that the coefficient above 0.6 is acceptable, while others consider that the coefficient above 0.7 is the minimum acceptable value. In this study, the strict coefficient is more than 0.7 as the standard [18].

Using the sample data collected in this survey, SPSS 20.0 software was used in this study. Firstly, the CITC index of each item should be greater than 0.5; otherwise, the item should be deleted. In Table 1, except the CITC index of S01 (0.518), all of them are greater than 0.5, and the deletion of S01 here will lead to unsuitable analysis in the structural equation model, and the CITC index of S01 is also too close to 0.5, so we reserve it here [19].

Secondly, Cronbach's *a* index of the whole scale is calculated. The result shows that Cronbach's α index is 0.938, and the *a* index after removing a certain question item will not change significantly. In addition, Cronbach's α index of 12 factor variables in the scale was calculated.

It can be seen that Cronbach's *a* index of the overall scale of the impact of each factor of public tourism services on each factor of destination brand relationship quality and the scale of each factor is above 0.700.

4. Construction of Public Tourism Information Service System Based on Smart Cities

4.1. Factor Analysis. In SPSS 20.0, factor analysis was conducted again for reliability analysis and purified data. When the factor loading of the item is greater than 0.5 and the cumulative proportion of explanatory variance is greater than 50%, it is up to the standard. According to the analysis of the previous chapter and the design of the questionnaire structure, all items are divided into three dimensions, four parts, and thirty-six items. Factor analysis is carried out on the public services of tourism, the behavioral response of tourists, the psychological response of tourists, and the quality of destination brand relationship. The factor loads of each item, Kaiser-Meyer-Olkin value, Bartlett's sphericity test (Butler's sphericity test), and Cronbach's *a* index were observed.

4.1.1. Factor Analysis of Public Tourism Services. Firstly, we extract the main components. Secondly, the KMO test is carried out. Finally, Bartlett's sphericity test (Butler's sphericity test) was carried out. The testing results are shown in Table 2.

Among them, factor F1 represents the traffic service in public tourism services, and the factor load of the measurement items is A02 (0.956), A01 (0.924), and A03 (0.943), the KMO value is 0.765, Bartlett's sphericity test (Butler's ball test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.938. Factor F2 represents hotel service in public tourism services, and the factor load of measurement items is B02 (0.918), B01 (0.922), and B03 (0.873), the KMO value is 0.735, Bartlett's sphericity test (Butler's ball type test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.889. Factor F3 represents tourism service in public tourism services, and the factor load of measurement items is C02 (0.914), C01 (0.913), and C03 (0.872), the KMO value is 0.730, Bartlett's sphericity test (Butler's test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.824.

4.1.2. Response Factor Analysis of Tourists' Behavior. Firstly, we extract the main components. Secondly, the KMO test is carried out. Finally, Bartlett's sphericity test (Butler's sphericity test) was carried out. The results are shown in Table 3.

Among them, factor F4 represents the reliability factor in the tourists' behavior response, and the factor load of the measurement items is L02 (0.892), L01 (0.842), and L03 (0.823), the KMO value is 0.694, Bartlett's sphericity test (Butler's ball test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.825. Factor F5 represents the competency factor in the tourists' behavior response. The factor load of measurement items is M03 (0.917), M02 (0.846), and M01 (0.832), the KMO value is 0.661, Bartlett's sphericity test (Butler's ball test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.827. Factor F6 represents the communication factor in the tourists' behavior response. The factor load of measurement items is N02 (0.831), N01 (0.782), and N03 (0.742), the KMO value is 0.634, Bartlett's sphericity test (Butler's ball test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.750.

4.1.3. Analysis of Psychological Response Factors of Tourists. Firstly, we extract the main components. Secondly, the KMO test is carried out. Finally, Bartlett's sphericity test (Butler's sphericity test) was carried out. Table 4 shows the testing results.

Among them, factor F7 represents the safety factor in tourists' psychological response, and the factor load of measurement items is R02 (0.896), R01 (0.862), and R03 (0.840), the KMO value is 0.744, Bartlett's globosity test (Butler's globosity test) *P* value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.814. Factor F8 represents the trust factor in tourists' psychological response, and the factor load of measurement items is S02 (0.890), S03 (0.783), and S01 (0.725), the KMO value is 0.592, Bartlett's sphericity test (Butler's sphericity test) *P* value of is significant at the level of 0.001, and the overall Cronbach's α index (see Table 1) is 0.734. Factor F9 represents the politeness factor in tourists' psychological response, and the factor load of the measurement items is T02

Table 1: Reliability analysis.

Variables	Item	Total correlation corrected by item	Cronbach's α value after deleting the item	Cronbach's α
Smart transportation	A01, A02, A03	0.882	0.930	0.938
Smart hotel	B01, B02, B03	0.893	0.881	0.889
Smart tourism	C01, C02, C03	0.793	0.759	0.824
Reliability	L01, L02, L03	0.649	0.768	0.825
Communication	M01, M02, M03	0.626	0.719	0.825
Winning willfulness	N01, N02, N03	0.503	0.771	0.750
Security	R01, R02, R03	0.742	0.801	0.814
Trustworthiness	S01, S02, S03	0.518	0.829	0.734
Politeness	T01, T02, T03	0.612	0.817	0.792

Table 2: Factor analysis of public tourism services.

Number	Item	F1	F2	F3
A01	The traffic at the destination is convenient and fast	0.924		
A02	The process of purchasing this travel and transportation product is smooth	0.956	—	—
A03	In this travel traffic, I feel the advantages of Internet plus	0.943		
B01	Destination hotel is clean and comfortable		0.918	
B02	The process of purchasing the hotel accommodation product is smooth	—	0.922	—
B03	I feel the advantages of Internet plus in this trip accommodation		0.873	
C01	Destination quality is high and intelligent			0.911
C02	The process of purchasing products for this tour is smooth	—	—	0.914
C03	In this tour, I feel the advantages of Internet plus			0.872

Table 3: Response factor analysis of tourists' behavior.

Number	Item	F1	F2	F3
L01	The commitment of the tourism product or service provider of the destination has fulfilled	0.842		
L02	The tourism products or services I buy can complete consumption or experience	0.894	–	–
L03	The provider of tourism products or services at the destination is reliable	0.823		
M01	When I have problems, I know whom to talk to and can communicate with		0.832	
MO2	Internet and the social platform play a great role in my tourism communication	–	0.846	–
M03	No problem of smooth or improper communication in my whole tourism activity		0.917	
N01	The scenic spot of the destination can provide corresponding products or services			0.782
NO2	The service personnel at the destination is competent for his/her work	–	–	0.831
N03	Internet plus or intelligent building of N03 destination can meet my needs			0.742

Table 4: Psychological effects of tourists.

Number	Item	F1	F2	F3
R01	The social security and safety facilities at the destination are good	0.862		
R02	I did not feel any threat or danger in this tour	0.896	–	–
R03	The travel products and services I buy are safe	0.840		
S01	The provider of tourism products or services at the destination is reliable		0.725	
S02	I am treated equally and honestly in tourism activities	–	0.890	–
S03	I did not encounter any dishonesty in this tour		0.783	
T01	The tourism product or service provider of the destination is civilized and polite			0.802
TO2	I did not encounter any impolite or uncivilized events in this tour	–	–	0.837
T03	I feel that the civilization and politeness of the destination are very high			0.792

(0.837), T01 (0.802), and T03 (0.792) from large to small, the KMO value is 0.659, Butler's ball test P value is significant at 0.001 level, and the overall Cronbach's α index (see Table 1) is 0.792.

The problems and shortcomings of the public tourism information service system are the factors that hinder the development of tourism industry into a "modern service industry more satisfactory to the people." The main reason lies in the inappropriate framework and mode of the transformation and upgrading of tourism services, the failure to fully consider the domestic environment and conditions, the failure to find a good foothold and entry point, and the lack of many ideas that are beneficial to practice [20]. For example, although the international experiential public tourism information service mode enhanced, international exchange activities were active. Some

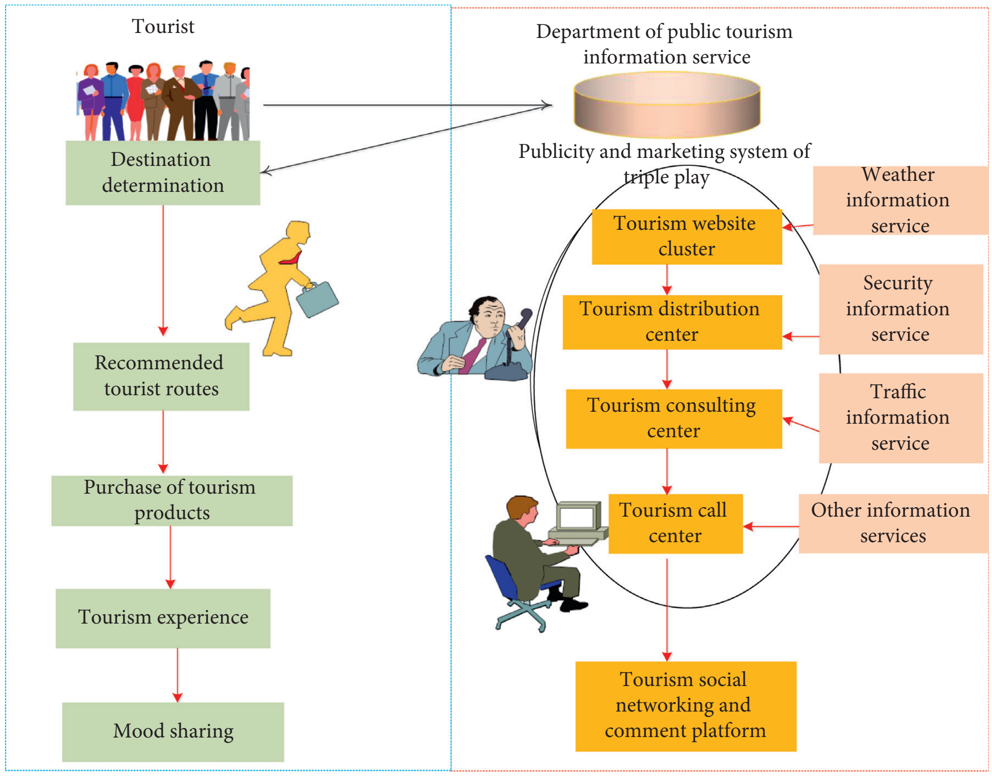


Figure 7: Public tourism information services for smart city.

tourist destinations learn a full range of foreign public tourism information service modes and means, which are applied to domestic construction, regardless of the environmental conditions of the region, resulting in the mode not applicable and not working. Therefore, the most important thing in the construction of the public tourism information service system is applicability, which is based on the existing basis and conditions.

4.2. Construction of the Public Tourism Information Service System for Smart City. On the basis of the system, platform, and operation management platform built for the construction of smart city, the improvement of the public tourism information service system is realized, the problems existing in the traditional public tourism information service system are changed, and the shortcomings are summed up. The public tourism information service system of face-to-face smart city is more humanized in terms of overall construction and basic characteristics and service media and methods intellectualization.

The traditional public tourism information service system is based on the perspective of tourism supply, including four parts: tourism website cluster, tourism distribution center, tourism consulting service center (station), and tourism call center. Each part has several subsystems or divided into several forms. It is built, operated, and managed by each department independently, lacking information interoperability and cooperation. With the rapid development of smart city construction, smart ideas provide new directions and methods for urban management and services and gradually get wide recognition and support. Smart cities also provide system, platform, and operation management support for the transformation and upgrading of the public tourism information service system. Therefore, the construction of the public tourism information service system for smart city is of great practical significance and very feasible.

From the analysis of the statistical results, it shows that tourists' dissatisfaction with the public tourism information service is mainly concentrated in the aspects of incomplete content, poor timeliness of updating, poor interoperability, and single form; the demand for the quantity and quality of public tourism information can be summarized as the characteristics of desirability, timeliness, comprehensiveness, system, relevance, and accuracy. These characteristics need to change the tradition. The public tourism information service system of China has become a key target to improve the satisfaction of tourists and a key index to build a new service system.

4.3. Mode of the Public Tourism Information Service System for Smart City. With the rapid development of smart city construction, smart ideas provide new directions and methods for urban management and services and gradually get wide recognition and support. Smart cities also provide system, platform, and operation management support for the transformation and upgrading of the public tourism information service system. Therefore, the construction of the public tourism information service system for smart city is of great practical significance and very feasible.

To build a public tourism information service system for smart city, with the help of the concept and construction foundation of smart city, starting with the improvement of the carrier of public tourism information services, improve the channels of public tourism information services, optimize the means of services, endow it with intelligent core, and cover the needs of the public tourism information service system reflected by the questionnaire. The public tourism information service system for smart city is shown in Figure 7.

5. Conclusion

This paper systematically analyzes and summarizes the problems and shortcomings of the existing public tourism information service system and, through the form of questionnaire, realizes the information demand characteristics of tourists, especially the individual tourists. Solving these problems in combination with the needs is the key to improving the public tourism information service system. To solve these problems and deficiencies, the proposal and construction of smart city is the key direction. With the consideration of the current situation, this paper analyzes the foundation platform of smart city for the construction of the public tourism information service system, mainly including the understanding foundation, physical foundation, and management foundation. On the basis of smart city, it is scientific and feasible to realize the construction of the public tourism information service system. This paper presents a systematic model, starting from the needs of tourists and aiming at the characteristics of tourists' demand for information before, during, and after tourism; the tourism supplier is extended to the whole society and tourism-related industrial departments. The government departments, as the main body of the service providers, realize the interconnection of various suppliers and information under the security mode and realize the intelligent communication under the stimulation of the city, and the innovation of service media, means, and management mode is constantly realized.

Acknowledgments

This work was supported by the major projects of the National Social Science Fund in 2018: The history, current situation, and future of information philosophy (project no. 18zda027).

References

[1] Y. Li, C. Hu, C. Huang, and L. Duan, "The concept of smart tourism in the context of tourism information services," *Tourism Management*, vol. 58, pp. 293–300, 2017.

[2] J. Wang, C. Jiang, K. Zhang, T. Q. Quek, Y. Ren, and L. Hanzo, "Vehicular sensing networks in a smart city: principles, technologies and applications," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 122–132, 2017.

[3] R. Díaz-Díaz, L. Muñoz, and D. Pérez-González, "Business model analysis of public services operating in the smart city ecosystem: the case of smart Santander," *Future Generation Computer Systems*, vol. 76, no. 6, pp. 198–214, 2017.

[4] V. A. Memos, K. E. Psannis, Y. Ishibashi, B.-G. Kim, and B. B. Gupta, "An efficient algorithm for media-based surveillance system (EAMSuS) in IoT smart city framework," *Future Generation Computer Systems*, vol. 83, no. 3, pp. 619–628, 2018.

[5] M. M. Rathore, A. Paul, W.-H. Hong, H. Seo, I. Awan, and S. Saeed, "Exploiting IoT and big data analytics: defining smart digital city using real-time urban data," *Sustainable Cities and Society*, vol. 40, no. 4, pp. 600–610, 2018.

[6] Y. Wu, W. Zhang, J. Shen, Z. Mo, and Y. Peng, "Smart city with Chinese characteristics against the background of big data: idea, action and risk," *Journal of Cleaner Production*, vol. 173, no. 3, pp. 60–66, 2018.

[7] H. Kim, L. Mokdad, and J. Ben-Othman, "Designing UAV surveillance frameworks for smart city and extensive ocean with differential perspectives," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 98–104, 2018.

[8] M. Gohar, M. Muzammal, and A. Ur Rahman, "SMART TSS: defining transportation system behavior using big data analytics in smart cities," *Sustainable Cities and Society*, vol. 41, no. 1, pp. 114–119, 2018.

[9] P. G. V. Naranjo, Z. Pooranian, M. Shojafar, M. Conti, and R. Buyya, "FOCAN: a fog-supported smart city network architecture for management of applications in the internet of everything environments," *Journal of Parallel and Distributed Computing*, vol. 132, no. 2, pp. 274–283, 2019.

[10] N. Komninos, C. Kakderi, A. Panori, and P. Tsarchopoulos, "Smart city planning from an evolutionary perspective," *Journal of Urban Technology*, vol. 26, no. 2, pp. 3–20, 2019.

[11] K. Zhang, J. Ni, K. Yang, X. Liang, J. Ren, and X. S. Shen, "Security and privacy in smart city applications: challenges and solutions," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 122–129, 2017.

[12] X. Li, P. S. W. Fong, S. Dai, and Y. Li, "Towards sustainable smart cities: an empirical comparative assessment and development pattern optimization in China," *Journal of Cleaner Production*, vol. 215, no. 5, pp. 730–743, 2019.

[13] Z. Khan, Z. Pervez, and A. G. Abbasi, "Towards a secure service provisioning framework in a smart city environment," *Future Generation Computer Systems*, vol. 77, no. 7, pp. 112–135, 2017.

[14] V. Della Corte, C. D'Andrea, I. Savastano, and P. Zamparelli, "Smart cities and destination management: impacts and opportunities for tourism competitiveness," *European Journal of Tourism Research*, vol. 1, no. 7, pp. 7–27, 2017.

[15] I. Lopez-Carreiro and A. Monzon, "Evaluating sustainability and innovation of mobility patterns in Spanish cities. analysis by size and urban typology," *Sustainable Cities and Society*, vol. 38, no. 8, pp. 684–696, 2018.

[16] D. Grimaldi and V. Fernandez, "The alignment of university curricula with the building of a smart city: a case study from

Barcelona," *Technological Forecasting and Social Change*, vol. 123, no. 3, pp. 298–306, 2017.

[17] G. Viale Pereira, M. A. Cunha, T. J. Lampoltshammer, P. Parycek, and M. G. Testa, "Increasing collaboration and participation in smart city governance: a cross-case analysis of smart city initiatives," *Information Technology for Development*, vol. 23, no. 3, pp. 526–553, 2017.

[18] I. Bashynska and A. Dyskina, "The overview-analytical document of the international experience of building smart city," *Business: Theory and Practice*, vol. 1, no. 9, pp. 228–241, 2018.

[19] W. Ejaz, M. Naeem, A. Shahid, A. Anpalagan, and M. Jo, "Efficient energy management for the internet of things in smart cities," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 84–91, 2017.

[20] V. Moustaka, A. Vakali, and L. G. Anthopoulos, "A systematic review for smart city data analytics," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–41, 2018.

Research on Intelligent Tourism Information System Based on Data Mining Algorithm

Jia Du

Lanzhou University of Arts and Science, Lanzhou 730000, China

Correspondence should be addressed to Jia Du; 1000247@luas.edu.cn

Academic Editor: Fazlullah Khan

Smart tourism purposes symbolize a new idea of IT application to increased competition and satisfaction of all stakeholders, including visitors as co-creators of tourism products and co-promoters of a destination. To improve the effect of smart tourism, this paper improves the common big data technology through algorithm enhancement to improve the intuitive effect of big data. We construct big data visualization technology and realize real-time online visualization of tourism data. In the spark-distributed environment, we use the conventional *K* clustering technique to improve the final output utilizing clustering means. The research results show that the smart tourism information system based on big data constructed in this paper can meet actual tourism information needs and user experience needs. The outcomes of the experimental results show that the proposed predictor significantly outperforms based on the improved algorithm.

1. Introduction

Smart travel is developed on traditional sightseeing services which is new way of gathering information [1]. It encompasses the most modern computer technology in terms of software and hardware. In the fields of e-commerce, online banking, tourism, medical care, etc., we momentarily witness situations in which data mining is successful [2]. Many data have been gathered in the tourism sector with the popularization of tourist information. Data mining and visualization technology is, therefore, required to process tourist information to assess tourism effects thoroughly and in detail and give efficient support for tourism industry. In the field of scientific literature, [3] the concept of intelligent tourism destinations has become a hot topic, and constructive arguments are not just in tourism. This notion is largely linked to the development of modern IT, which resulted in the changing of tourism purchasing behavior, leading to the implementation of new information systems for destination management organizations (DMOs), to govern the destination efficiently. To exchange knowledge and skills, the task is to provide visitors in real time a personal service and link all parties. In smart travel, it is not a passive way for users to be provided with what information they browse. Rather, it is to provide users with correct scenic spot information through a personalized recommendation method, which is very important [4].

In literature, a number of articles [4, 5] have proposed different smart travel system and guidance. For instance, in [6], the author provides examples of this aspect to help the user find the information if the user has insufficient knowledge of a certain aspect of tourism resources. Similarly in [7], the authors describe that how the knowledge-based recommendation system replicates depending on the user's operational behavior and addresses trip knowledge. The author also suggests an improvement strategy; however, it is very difficult to apply it because this knowledge base needs often to be updated. Moreover, the literature [8–11] conducted an in-depth study on the navigation of the intelligent itinerary based on research and design constraints. These systems use technologies such as WebGIS and Apriori algorithm. This is an interactive platform where users can change the travel schedule according to their wishes until they are satisfied.

In this research, we proposed an intelligent recommendation system for travel services and introduced the analysis results to users using big data technology. The study focuses on the function of large data algorithms in the intelligent tourism system. We also improve the visualization of big data technologies on the intelligent tourist information system. The functional structure of the intelligent tourist information system is developed according to current requirements. The proposed system performance is assessed in combination with experiments to enhance the marketing effect of intelligent tourism and the user experiences. The major contributions of the paper are as follows to improve the big data-based algorithms with the *k*-means algorithm.

(i) The proposed model examines the smart tourism system using digital technology when traveling
(ii) The proposed model improves the visualization of big data technologies
(iii) The proposed model analyzes a huge amount of data with experiments to enhance the marketing effect of intelligent tourism
(iv) The performance of the suggested model is thoroughly assessed using the powerful computing abilities of the spark cluster to handle the intelligent tourism system efficiency and quality

The rest of the paper is organized as follows. In Section 2, a proposed system model is designed for smart tourism. The big data application analytics analysis is conducted in Section 3. The experimental results and discussion are further summarized in Section 4. Finally, Section 5 concludes the paper with a summary and future research directions.

2. Tourism Distribution Model

In this section, we elaborate the framework of the smart tourism platform, the improved mining algorithm for visualized data of tourism, and the paradigm of communication competence in the communicative capability of tourist skills.

2.1. Framework of Smart Tourism. The smart travel ecosystem is built around the smart travel platform. The overall framework of smart tourism is shown in Figure 1. From Figure 1, the framework is composed of five parts i.e., smart tourism platform, tourism management department, tourism enterprise, traveler, and security system. The tourist industry and the department for tourism manage the smart tourism platform and manipulate tourism businesses using the appropriate information received from tourism companies on the platform. Tourism businesses, as the market's major body, receive different useful information from smart tourism platforms to give visitors tailored tourism services and product launches. Travelers, as customers, may easily receive information and services supplied by travel firms via smart travel platforms. The security system enables the creation and subsequent safe and stable functioning of the smart tourist system in terms of policies, money, skills, and procedures.

Technical framework of the smart tourism platform is shown in Figure 2. From Figure 2, the technical framework of the smart travel platform is divided into four levels i.e., data collection layer, network communication layer, data center layer, and business application layer. The data collection layer is the direct source of data information, which is composed of cameras, sensors, smart phones, and other information input terminals. The network communication layer is the transmission path of data information and makes the transmission of data information between layers possible. The data center layer is responsible for the preservation of data information, classifying the data, and further mining and analysis. The business application layer uses the valuable information provided by the data center layer to provide rich business functions to different users.

2.2. Improved Mining Algorithm of Visualized Data of Smart Tourism. In the traditional coordinate system, all the axes cross each other. Figure 3 shows the parallel coordinates in the 6-dimensional space. Before displaying, the algorithm first converts the source data. Generally, the data types can be divided into two types: sequence data and actual data. The sequence data is represented by real numbers. It can also be represented by an integer, but in this case, the meaning of continuity is different from the definition of mathematical perspective. The view allows users to roughly identify the data distribution of each attribute, especially different types of data are displayed in different colors, and can more clearly indicate the differences between different types of data.

Furthermore, parallel coordinates according to their features are suited for continuous data and cannot be used for discrete data. In the event of real and discrete data, the processing of the data in the actual number type has to be continuous. The so-called parallel coordination concept consists of the mapping of N-dimensional reservoirs in parallel N axes to line segments. The N-points that cross the line and the N-axis represent the N-dimensional data of the data point. A linear independent equation [12] can represent the polyline produced by this line segment:

$$\frac{x_1 - a_1}{u_1} = \frac{x_2 - a_2}{u_2} = \cdots = \frac{x_n - a_n}{u_n}. \tag{1}$$

Among them, the slope is

$$m_i = \frac{u_{i+1}}{u_i}. \tag{2}$$

The following formula represents the intercept on the x_{i+1} axis in the $x_{i+1}x_i$ plane [13]:

$$b_i = \left(a_{i+1} - m_i a_i\right). \tag{3}$$

Here, if the distance between parallel axes is 1, then the coordinates of the intersection point are

$$\left(\frac{i}{(1 - m_i)}, \frac{b_i}{(1 - m_i)}\right), i = 1, 2, \ldots, n - 1. \tag{4}$$

The point in the rectangular coordinate of the plane is mapped to the parallel coordinate as a line segment. In

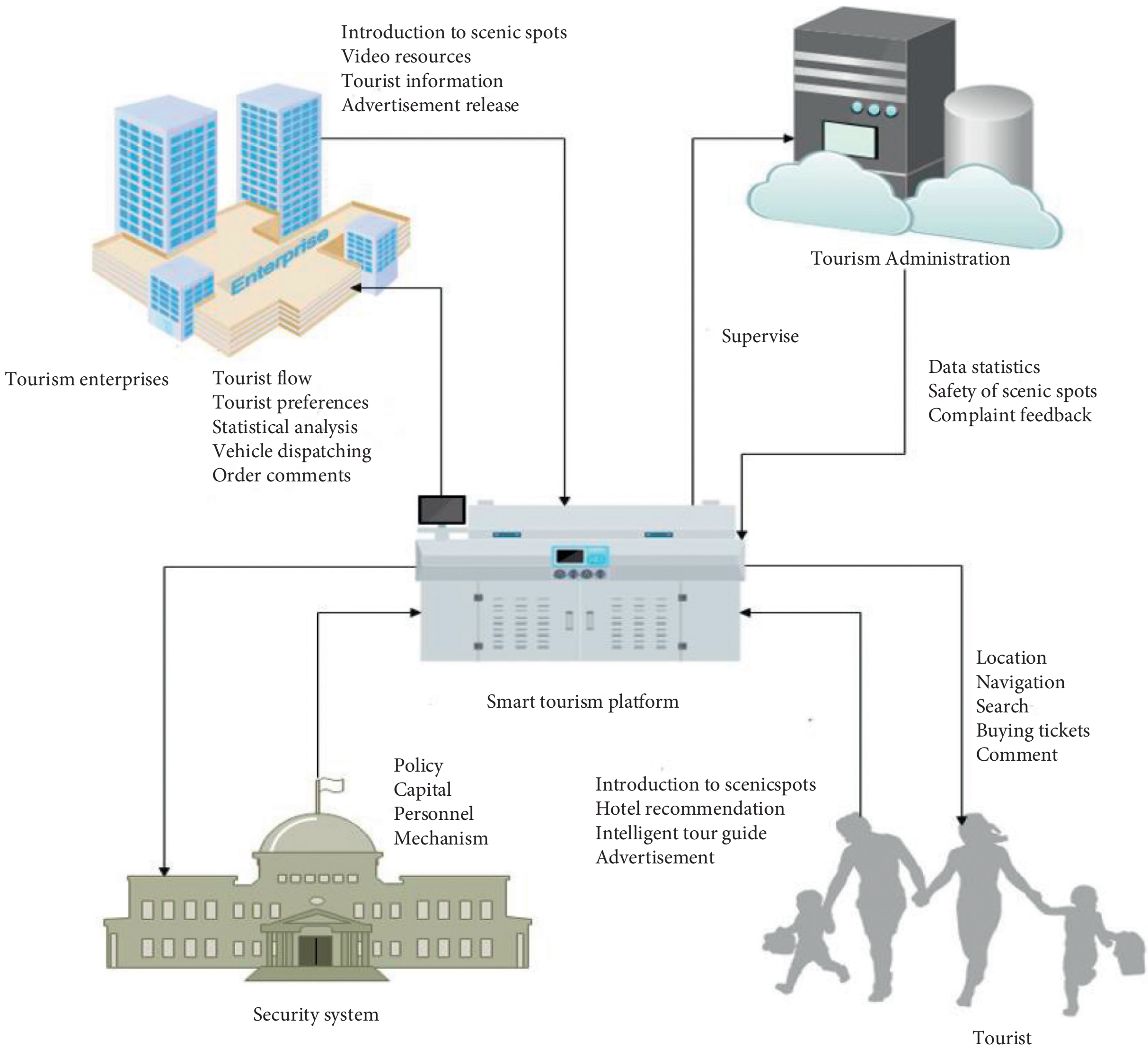


FIGURE 1: The overall framework of smart tourism.

planar rectangular coordinates, multiple points on a straight line are mapped to parallel coordinates by multiple line segments that intersect one point so that a point in parallel coordinates can correspond to a straight line in planar rectangular coordinates. The mathematical relationship between rectangular coordinates and parallel coordinates, mutual transformation, distance comparison, etc., has been proved very early. In short, the above two-dimensional situation can be expanded in a multidimensional space, and these mapping relationships provide a basis for understanding the multidimensional dataset in the parallel coordinate system [14]. Correspondence between parallel coordinates and rectangular coordinates is shown in Figure 4.

2.3. Intelligent Information System for Smart Tourism. The advantages of the parallel coordinate method are mainly obvious. Users can quickly convert from the traditional rectangular coordinate system to the parallel coordinate system and can better use the two-dimensional plane to display multidimensional data, and compared with other data visualization methods, such as scatter plots, it is simple and intuitive for users to receive. Data preprocessing is an important step of knowledge discovery in the database, which helps to improve the quality of data and ensure the accuracy and performance of the results. Data preprocessing is a broad field that includes different strategies and technologies related in many complex ways. In view of the ongoing research and the unique characteristics of the Iris data, the focus is on solving the data conversion problem. Data conversion is the conversion of all values used for variables. In other words, the variable value of the object is used for the conversion of each object. Normalization and standardization are general methods of variable transformation. Moreover, the general normalization method is to linearly transform the data, map the corresponding value to a new interval, and perform the following minimum and maximum normalization, as shown in the following formula [15]:

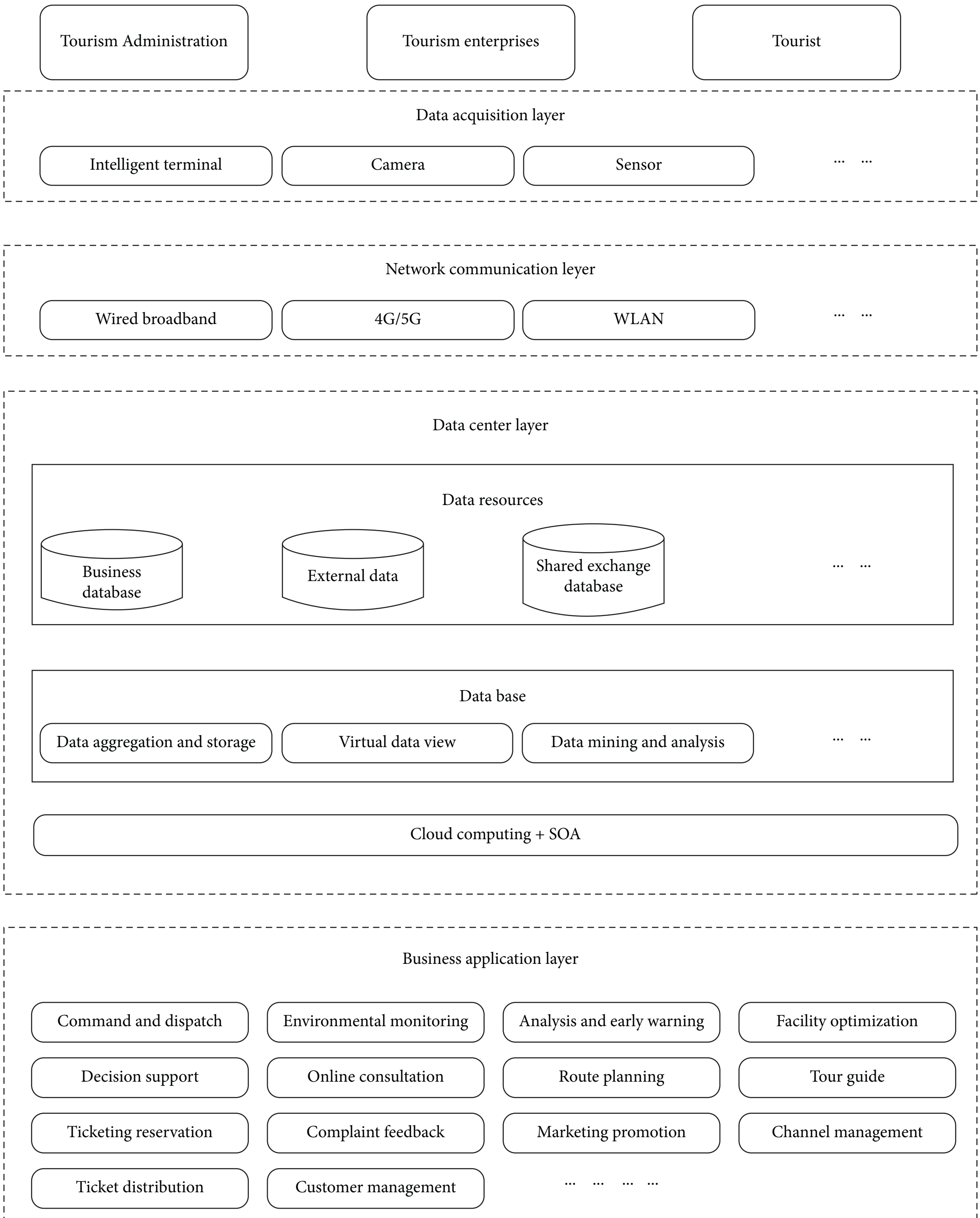


Figure 2: Technical framework of the smart tourism platform.

$$x'_{ij} = \frac{x_{ij} - \mathrm{Mix}_j}{\mathrm{Max}_j - \mathrm{Min}_j}(\mathrm{newMax} - \mathrm{newMin}) + \mathrm{newMin}. \quad (5)$$

In the formula, x_{ij} is the original data, Max_j is the maximum value of the *J*th column in the array, Min_j is the minimum value in column *J*, *newMin* and *newMax* are the new maximum and minimum values, respectively, which are two constants, and $x_{ij}^{'}$ is the value in the new interval after conversion. In the data mining samples, each variable represents various properties of the sample, and the measurement unit is different. In this way, variables with large absolute values are affected by it, while variables with small

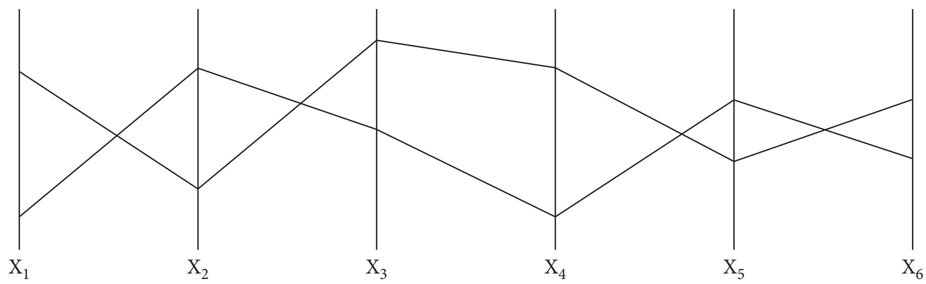


Figure 3: Parallel coordinates in the six-dimensional space.

absolute values may disappear. In order to ensure that the variables are in the same position in the analysis, the data can be centralized and standardized conversion [16]. Taking P as the total number of variables and n as the total number of samples (records), the average value of the Jth variable is as follows:

$$\overline{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \quad j = 1, 2, \ldots, p. \tag{6}$$

The center transformation of the nth data of the jth variable:

$$x'_{ij} = x_{ij} - \overline{x}_j, \quad i = 1, 2, \ldots, n. \tag{7}$$

Through this conversion, the average value of each variable is 0. In other words, the value of each variable is the same base point. Standardization is based on centralization and transformation to make the range of changes of various variables equal. This paper mainly adopts standard deviation standardization. The standard deviation of the ith variable is as follows [17]:

$$S_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ij} - \overline{X}_j\right)^2}. \tag{8}$$

Deviation standardization:

$$x'_{ij} = \frac{x_{ij} - \overline{x}_j}{S_j}, \quad i = 1, 2, \ldots, n. \tag{9}$$

After the conversion, the average value of each variable is 0, and the standard deviation is 1.

The cluster separation point also substantially impacts the average value and standard deviation, such that the above conversion has to be adjusted. First, the average is substituted with the average value. The standard difference is then substituted with the absolute default. In the following, you will find the median value of the jth variable:

$$\begin{aligned} \mu_j &= x_{ij}, \quad i = \frac{n}{2} + 1;\ N \text{ is an odd number},\\ \mu_j &= \frac{1}{2}\left(x_{ij} + x_{i+1j}\right), \quad i = \frac{n}{2};\ N \text{ is even}. \end{aligned} \tag{10}$$

In a nutshell, if there are odd values, the central value is the middle value. If there is an even value, the middle value is the average of the middle two number s [18]. The average absolute baseline deviation of the jth variable is shown below:

$$\sigma_j = \frac{1}{n}\sum_{i=1}^{n}\left|x_{ij} - k\mu_j\right|. \tag{11}$$

In the formula, μ_j is the mean or median. The multiplication factor k is used to stretch the distribution range near the central value of the data, and since the number of even numbers is even in the normalization of the data using Iris, using $k = 2$ has a better effect. For the data n of the jth variable, the median absolute standard deviation is standardized [19]:

$$x'_{ij} = \frac{x_{ij} - \overline{x}_j}{\sigma_j}, \quad i = 1, 2, \ldots, n. \tag{12}$$

2.4. Application Systems of Smart Tourism. In addition, typical techniques of converting standardization include very bad standardization and very poor standardization. Some of the following typical methods for data conversion are commonly used for visual mining. Moreover, similar replacement conversions are needed for Iris data. Otherwise, you will not get the intended outcome. It should be mentioned that when other sets of data are different, the standard conversion of data preparation should be changed by the circumstances of a given collection of data. In this experiment, the dataset is utilized to conduct a unified digital process in a category, while assessing the clustering algorithm, and to differentiate the indistinguishable categories. Generally, such processing does not impact clustering method implementation and the visualization effect might be extremely obvious. Similarly, this column's value attribute is prepared to prevent undue impact on other values of the attribute. At this time, the processing procedure cannot employ standardization. The smallest normalization process will have a great impact on other dimensional attributes, so it is necessary to use standard deviation conversion or absolute standard deviation standard conversion. At the same time, it must be ensured that the conversion method has the least impact on other attributes. Of course, in the classification algorithm, the category attribute is only used to identify the category, so it does not require preprocessing [20]. Since the quality of a cluster analysis process depends on the choice of measurement standards, the measurement standards must

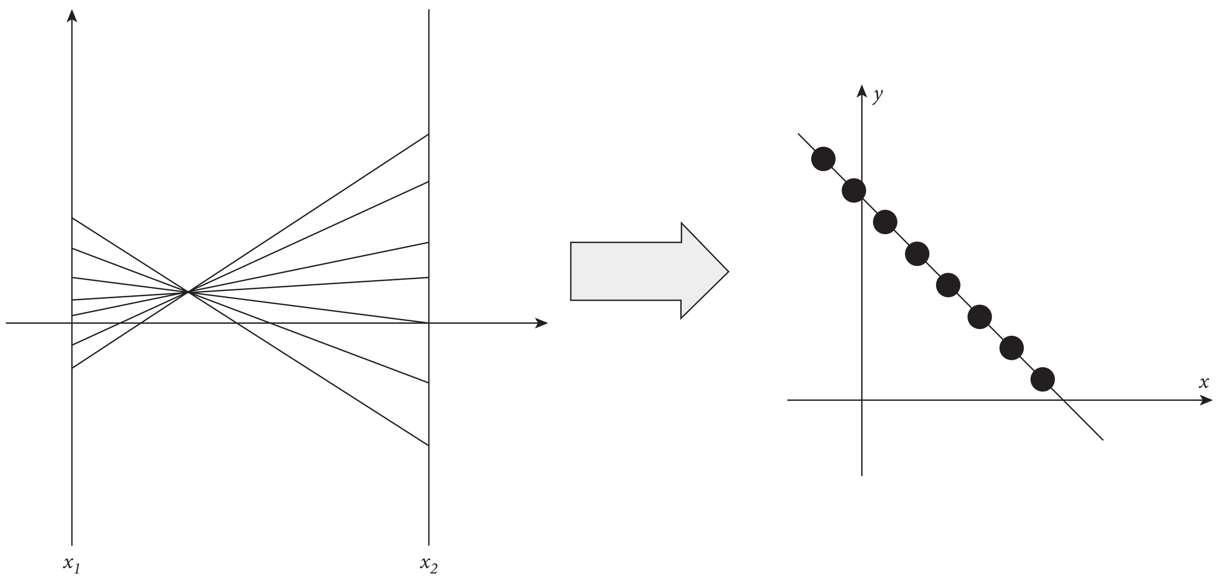


Figure 4: Correspondence between parallel coordinates and rectangular coordinates.

be chosen carefully. The distance between classes is defined to illustrate the relationship between the two types.

(1) The shortest distance method is as follows [21]:

$$D_s(C_a, C_b) = \min\{d(x, y) | x \in C_a, y \in C_b\}. \quad (13)$$

(2) The group average method is as follows:

For the group average version of hierarchical clusters, the closeness of two clusters is defined as the average closeness of all points of different clusters:

$$D_s(C_a, C_b) = \frac{\sum d(x, y)}{m_i \cdot m_j} | \{x \in C_a, y \in C_b\}. \quad (14)$$

Among them, m_i and m_j are the size and number of classes C_a and C_b, respectively. In addition to the above two methods, there are methods for explaining the distance between classes, including the sum of squares method, the longest distance method, the center method, and the class average method. Of course, when the data is complex, the differential sum of the squares method can be considered [22]. There are two most common methods for measuring this proximity i.e., Euclidean distance and cosine similarity:

$$d(x_i, y_j) = \sqrt{\sum_{k=1}^{n} (x_{ik} - y_{jk})}. \quad (15)$$

Among them, n is the dimension. It means that the smaller the distance value, the greater the similarity in the vicinity:

$$Sim(x_i, y_j) = \frac{\sum_{k=1}^{n} x_{ik} y_{jk}}{\sqrt{\left(\sum_{k=1}^{n} x_{ik}^2\right)\left(\sum_{k=1}^{n} y_{jk}^2\right)}}. \quad (16)$$

The nearest neighbor method is also called passive learning, and its advantages are as follows:

(1) The distribution of attribute values does not need to be known beforehand. Most alternative typing approaches nevertheless require previous knowledge of the distribution of the property value.

(2) Suitable for gradual education.

(3) The classification precision of the KNN is typically greater than other classification methods since there is no requirement to create explicit rules.

3. Big Data Application Analytics

The beginning point and destination of smart travel construction are various smart apps based on big data. Big data is useful only via its discovery and analysis, and companies can benefit only by applying big data to applications. As a result, following the construction of the data collecting layer, network communication layer, and data center layer, various smart apps are built to provide relevant services for travel management departments, travel firms, and travelers to realize genuine smart travel. It is the smart travel application system, as indicated in Figure 5. The overall structure of the user recommendation system is shown in Figures 6 and 7, which show the process of interacting with the user, and the B/S mode is adopted.

4. Experimental Results and Discussion

During the construction of the aforementioned system, this article validates the performance of the big data smart tourism system built in this paper and performs system tests based on the actual scenario of smart tourism. Furthermore, as the system input, this paper collects huge tourist data from the Internet to assess the data mining effect of the smart tourism system developed in this research. Furthermore, this article divides the data into 59 sets and assesses the influence of system data mining in the form of expert grading. The

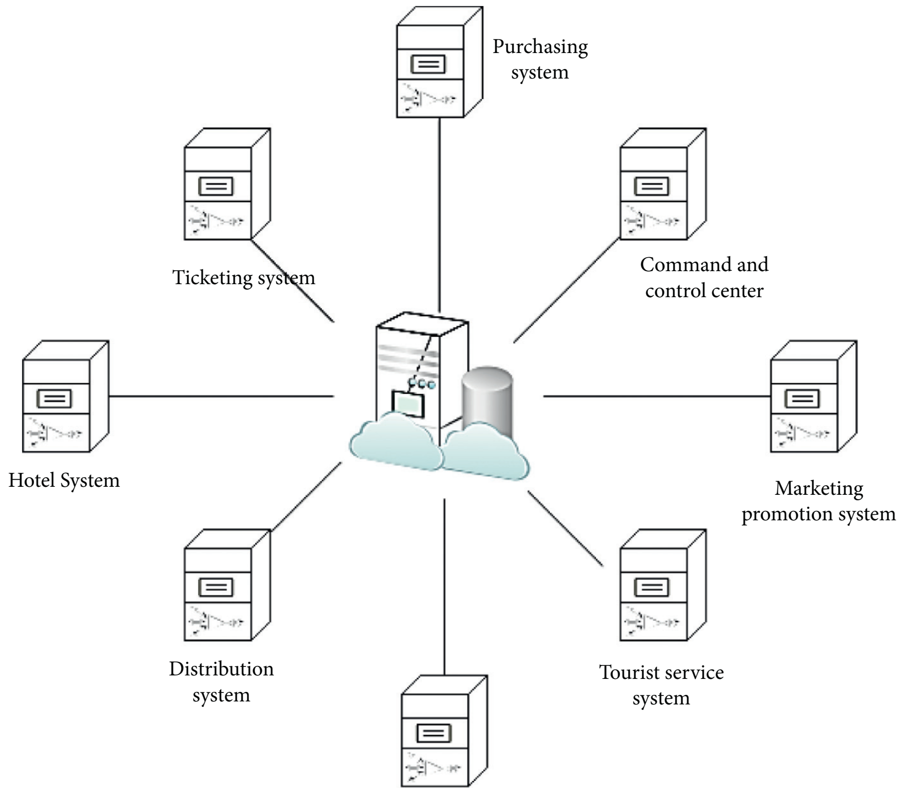


Figure 5: Application systems of smart tourism.

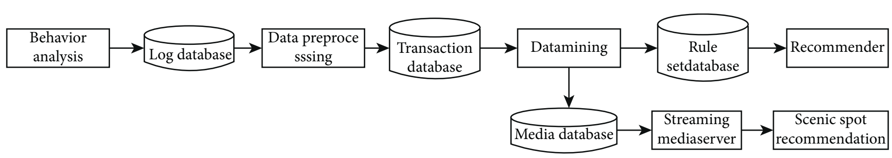


Figure 6: User recommendation system structure.

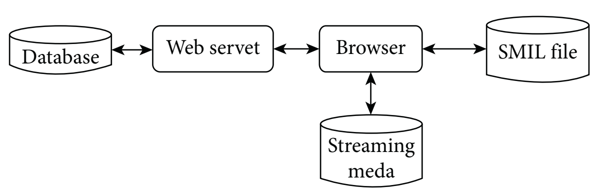


Figure 7: User interaction system architecture.

acquired findings are presented in Table 1 and Figure 8 below.

According to the study findings, the smart tourism system developed in this article has a significant impact on the smart mining of tourist data. On this premise, this paper combines the actual scenario with simulation research to validate the user experience satisfaction of the system created in this article. The acquired findings are presented in Table 2 and Figure 9.

According to the aforementioned results, the big data smart tourism system built in this article has good data mining effects, and the smart tourism system can effectively improve user experience, with high user satisfaction. This is also the unavoidable future path of tourism growth.

Table 1: Statistical table of evaluation of data mining effect of the smart tourism system.

No.	Data mining	No.	Data mining	No.	Data mining	No.	Data mining
1	93.99	16	96.71	31	90.12	46	87.91
2	94.57	17	92.14	32	87.03	47	96.85
3	88.13	18	92.09	33	95.76	48	95.05
4	87.32	19	90.20	34	90.55	49	96.33
5	89.54	20	92.26	35	93.99	50	92.89
6	89.69	21	90.73	36	94.45	51	87.69
7	96.84	22	91.40	37	88.49	52	95.45
8	91.44	23	93.27	38	86.98	53	92.58
9	87.46	24	96.03	39	90.21	54	89.33
10	91.64	25	95.60	40	90.91	55	87.33
11	93.09	26	90.50	41	88.53	56	90.04
12	87.31	27	86.01	42	87.09	57	93.26
13	88.93	28	92.14	43	87.48	58	94.58
14	88.67	29	87.45	44	93.08	59	94.66

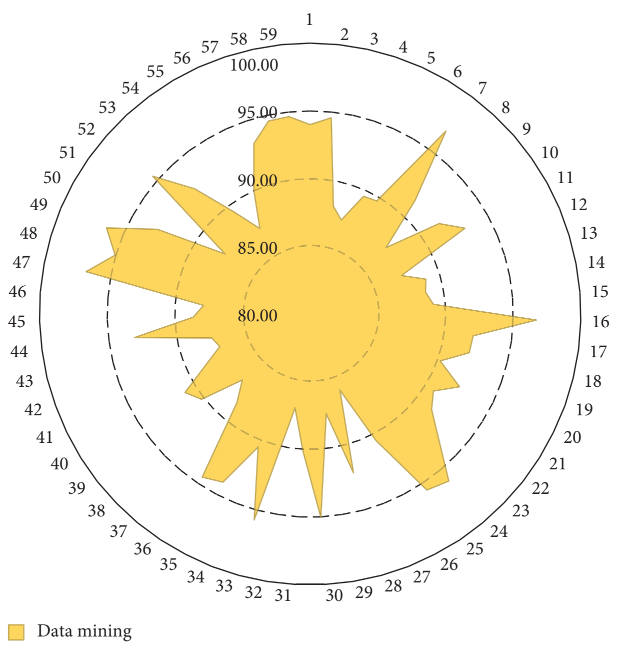


Figure 8: Statistical diagram of evaluation of data mining effect of the smart tourism system.

Table 2: Statistical table of user experience satisfaction of the smart tourism system.

No.	Satisfaction	No.	Satisfaction	No.	Satisfaction	No.	Satisfaction
1	87.50	16	88.72	31	91.63	46	88.09
2	86.60	17	92.64	32	92.12	47	87.42
3	93.91	18	85.89	33	89.89	48	91.70
4	91.16	19	85.49	34	89.96	49	85.86
5	92.00	20	89.50	35	84.11	50	85.28
6	91.77	21	87.06	36	88.75	51	93.96
7	87.67	22	84.36	37	85.98	52	91.04
8	85.65	23	93.72	38	91.44	53	92.12
9	85.54	24	86.07	39	89.80	54	92.12
10	93.88	25	93.13	40	91.45	55	90.94
11	85.81	26	91.94	41	89.82	56	89.64
12	85.79	27	92.71	42	88.10	57	84.88
13	88.05	28	89.69	43	90.86	58	89.89
14	88.60	29	86.25	44	87.90	59	86.08
15	91.54	30	91.76	45	91.12		

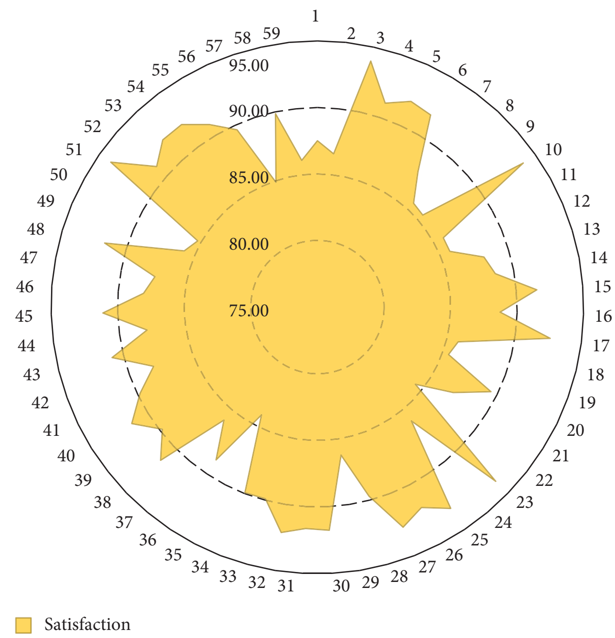


FIGURE 9: Statistical diagram of user experience satisfaction of the smart tourism system.

5. Conclusion

In this paper, we propose a new tourism system model using big data technology and create the large-scale tourist ability mining model. We provide the new paradigm examination of the management for tourism development. In terms of data, traveling firms must rely on big data technologies if they are to become better. Tourism is an information-intensive sector, with information resources depending on the operation and administration of companies. Therefore, an unanticipated increase for companies by using big data technologies to investigate traveling data. Finally, the experimental investigation also shows that the system built in this work has some effect.

Acknowledgments

This study was supported by Lanzhou University of Arts and Sciences Distinguished Young Scientific Research Talent Cultivation Project (2018JCQN009) and Talent Innovation and Entrepreneurship Project of Lanzhou City (2018-RC-68).

References

[1] A. K. Tripathy, P. K. Tripathy, N. K. Ray, and S. P. Mohanty, "iTour: the future of smart tourism: an IoT framework for the independent mobility of tourists in smart cities," *IEEE consumer electronics magazine*, vol. 7, no. 3, pp. 32–37, 2018.

[2] E. Sigalat-Signes, R. Calvo-Palomares, B. Roig-Merino, and I. García-Adán, "Transition towards a tourist innovation model: the smart tourism destination," *Journal of Innovation & Knowledge*, vol. 5, no. 2, pp. 96–104, 2020.

[3] H. Lee, J. Lee, N. Chung, and C. Koo, "Tourists' happiness: are there smart tourism technology effects?" *Asia Pacific Journal of Tourism Research*, vol. 23, no. 5, pp. 486–501, 2018.

[4] C. Koo, L. Mendes-Filho, and D. Buhalis, "Guest editorial," *Tourism Review*, vol. 74, no. 1, pp. 1–4, 2019.

[5] T. Zhang, C. Cheung, and R. Law, "Functionality evaluation for destination marketing websites in smart tourism cities," *Journal of China Tourism Research*, vol. 14, no. 3, pp. 263–278, 2018.

[6] M. A. C. Ruíz, S. T. Bohorquez, and J. I. R. Molano, "Colombian tourism: proposal app to foster smart tourism in the country," *Advanced Science Letters*, vol. 23, no. 11, pp. 10533–10537, 2017.

[7] M. A. C. Ruíz, S. T. Bohorquez, and J. I. R. Molano, "Colombian tourism: proposal app to foster smart tourism in the country," *Advanced Science Letters*, vol. 23, no. 11, pp. 10533–10537, 2017.

[8] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the Internet of things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[9] I. Guerra, F. Borges, J. Padrão, J. Tavares, and M. H. Padrão, "Smart cities, smart tourism? The case of the city of porto," *Revista Galega de Economia*, vol. 26, no. 2, pp. 129–142, 2017.

[10] Y. Topsakal, M. Bahar, and N. Yüzbaşioğlu, "Review of smart tourism literature by bibliometric and visualization analysis," *Journal of Tourism Intelligence and Smartness*, vol. 3, no. 1, pp. 1–15, 2020.

[11] S. Joshi, "Social network analysis in smart tourism driven service distribution channels: evidence from tourism supply chain of Uttarakhand, India," *International Journal of Digital Culture and Electronic Tourism*, vol. 2, no. 4, pp. 255–272, 2018.

[12] F. Femenia-Serra, B. Neuhofer, and J. A. Ivars-Baidal, "Towards a conceptualisation of smart tourists and their role

within the smart destination scenario," *Service Industries Journal*, vol. 39, no. 2, pp. 109–133, 2019.

[13] C. Koo, F. Ricci, C. Cobanoglu, and F. Okumus, "Special issue on smart, connected hospitality and tourism," *Information Systems Frontiers*, vol. 19, no. 4, pp. 699–703, 2017.

[14] H. Abdel Rady and A. Khalf, "Towards smart tourism destination: an empirical study on sharm el sheikh city, Egypt," *International Journal of Heritage, Tourism and Hospitality*, vol. 13, no. 1, pp. 78–95, 2019.

[15] F. Femenia-Serra, B. Neuhofer, and J. A. Ivars-Baidal, "Towards a conceptualisation of smart tourists and their role within the smart destination scenario," *Service Industries Journal*, vol. 39, no. 2, pp. 109–133, 2019.

[16] T. Pencarelli, "The digital revolution in the travel and tourism industry," *Information Technology & Tourism*, vol. 22, no. 3, pp. 455–476, 2020.

[17] C. J. P. Abad and J. F. Álvarez, "Landscape as digital content and a smart tourism resource in the mining area of cartagena-La unión (Spain)," *Land*, vol. 9, no. 4, pp. 1–22, 2020.

[18] P. M. da Costa Liberato, E. Alén-González, and D. F. V. de Azevedo Liberato, "Digital technology in a smart tourist destination: the case of porto," *Journal of Urban Technology*, vol. 25, no. 1, pp. 75–97, 2018.

[19] J.-J. Hew, G. W.-H. Tan, B. Lin, and K.-B. Ooi, "Generating travel-related contents through mobile social tourism: does privacy paradox persist?" *Telematics and Informatics*, vol. 34, no. 7, pp. 914–935, 2017.

[20] Z. Ghaderi, P. Hatamifar, and J. C. Henderson, "Destination selection by smart tourists: the case of Isfahan, Iran," *Asia Pacific Journal of Tourism Research*, vol. 23, no. 4, pp. 385–394, 2018.

[21] T. T. Nguyen, D. Camacho, and J. E. Jung, "Identifying and ranking cultural heritage resources on geotagged social media for smart cultural tourism services," *Personal and Ubiquitous Computing*, vol. 21, no. 2, pp. 267–279, 2017.

[22] P. Del Vecchio and G. Passiante, "Is tourism a driver for smart specialization? Evidence from Apulia, an Italian region with a tourism vocation," *Journal of Destination Marketing & Management*, vol. 6, no. 3, pp. 163–165, 2017.