

## Research Article

# Tourism Destination Preference Prediction Based on Edge Computing

**Bin Deng** , **Jun Xu**, and **Xin Wei**

*Lyceum of the Philippines University, Batangas 4200, Philippines*

Correspondence should be addressed to Bin Deng; [bindeng@lpubatangas.edu.ph](mailto:bindeng@lpubatangas.edu.ph)

Received 26 January 2021; Revised 20 February 2021; Accepted 8 March 2021; Published 17 March 2021

Academic Editor: Hsu-Yang Kung; [kung@mail.npust.edu.tw](mailto:kung@mail.npust.edu.tw)

Copyright © 2021 Bin Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In view of the fact that the important characteristics of tourism destination selection preference are not considered in the current prediction methods of tourism destination selection preference, resulting in low prediction accuracy and comprehensive accuracy and long prediction time, a tourism destination selection preference prediction method based on edge calculation is proposed. This paper uses edge computing to construct the characteristics of tourism destination selection preference and uses a random forest algorithm to select important features and carry out preliminary estimation and ranking. Using the multiple logit selection model, the tourists' preference sequence for tourism destination selection is obtained and sorted and the tourism destination selection preference model is obtained. By calculating the weight value of tourism destination selection preference, the weight set of tourism destination selection preference is determined and the tourism destination selection preference is determined according to the link prediction method to realize the tourism destination selection preference prediction. The experimental results show that the comprehensive accuracy of the proposed method is good, which can effectively improve the prediction accuracy of tourism destination selection preference and shorten the prediction time of tourism destination selection preference.

## 1. Introduction

Tourism is an important part of the modern service industry and a strategic pillar industry of the global economy. It can provide a large number of employment opportunities and good income with strong action and less resource consumption [1]. Promoting the development of tourism is of great practical significance in changing the mode of economic development, adjusting the industrial structure, and expanding domestic demand. The so-called preference is that consumers like one commodity more than another. It is a subjective description method that can reflect a kind of emotion and tendency in people's heart [2]. Preferences have individual characteristics and group characteristics. In other words, a certain group with some common characteristics has similarities in theory. Tourists' preference for tourism destination choice means that some common factors will affect the choice of tourism destination by groups of tourists with specific characteristics. Therefore, there are some similarities in the choice of tourist destinations. That is, they

have similar preferences [3]. Tourism destination preference prediction is an important part of tourism research, tourism planning, and management. Therefore, accurate prediction of tourism destination selection preference can improve the scientific level and effectiveness of tourism research planning and management.

The preference of destination selection directly affects the development trend of outbound tourism. Therefore, it is of great significance to predict the preference of destination selection. At present, a large number of scholars have carried out different degrees of research on the choice of tourism destination and achieved certain theoretical results. In reference [4], a hierarchical Bayesian network with multiple data sources is used to estimate and predict daily source destination tuples. The concept of O-D tuples is proposed. Vehicles in the road network are predicted and tracked by an advanced monitoring system. Aiming at the problem of obtaining the posterior probability of uncertain parameters, a multiprocess hierarchical Bayesian network mechanism in Gaussian space is developed. The model includes the level

and trend components of future traffic volume. This method can meet the demand of forecasting and reduce the uncertainty in the process of estimation and prediction. In reference [5], a destination prediction method based on the initial part of the vehicle trajectory is proposed. The new trajectories are assigned to the most likely clusters, and the similarity grid driven by data is generated by obtaining the trajectory clustering describing user behavior. The final destination of the new trajectory is predicted by using the characteristics of the internal trajectories of the model cluster. The prediction accuracy of this method is high. However, the above methods do not consider the important characteristics of tourism destination selection preference, which leads to the problems of low prediction accuracy and comprehensive accuracy and long prediction time [6].

In view of the above problems, this paper puts forward a prediction method of tourism destination selection preference based on edge computing. It uses the edge calculation to construct the tourism destination selection preference characteristics, uses the random forest algorithm, selects the important features, uses the multiple logit selection model to obtain the tourists' preference sequence for the choice of tourism destinations and sorts them, and obtains the tourism destination selection preference model. By calculating the weight value of tourism destination selection preference, the preference weight set of tourism destination selection is determined, and the tourism destination selection preference prediction is realized. The comprehensive accuracy of the tourism destination preference prediction method is good, which can effectively shorten the prediction time and improve the prediction accuracy [7].

The research contributions of the paper include the following:

- (1) A prediction method based on the preference of tourist destination selection is proposed
- (2) This paper uses edge computing technology to construct tourist destination selection preference features and uses a random forest algorithm to select important features and perform preliminary estimation and ranking
- (3) Using the multiple logit selection model, the preferred sequence of tourists' choice of tourist destinations is obtained and selected, and the model of tourist destination choice preference is obtained
- (4) Determine the weight set of the tourist destination selection preference by calculating the weight value of the tourist destination selection preference and determine the tourist destination selection preference according to the link prediction method to realize the prediction of the tourist destination selection preference

## 2. Edge Computing

*2.1. Concept of Edge Computing.* Edge computing is a new computing mode that performs processing data computation at the edge of the network. In edge computing, the

downstream data of the edge represents the cloud service, and the uplink data represents the Internet service of everything. The "edge" of edge computing refers to any computing and network resources between the data source and the cloud computing center path [8]. Edge computing can reduce the request response time, improve the endurance, reduce network bandwidth, and ensure data security and privacy. The edge computing architecture is as shown in Figure 1.

As can be seen from Figure 1, the edge algorithm structure diagram is mainly composed of 4 parts, each part has many branches, and the algorithm is more complicated. User terminals are all kinds of Internet of things devices, such as mobile phones, driverless cars, cameras, etc. These terminals collect a large amount of data and upload the data to the nearest edge device for storage and calculation. Edge devices are a large number of nodes on the edge of the network, such as routers, switches, base stations, or edge servers. These devices store and precalculate the data from the terminal and upload the processed results to the cloud computing center. Cloud Computing Center has powerful computing power, which can perform complex calculations on the acquired data and return the final settlement results to the user terminal [9].

*2.2. Relationship between Edge Computing and Cloud Computing.* Cloud computing is suitable for nonreal-time, long-period data, and business decision-making scenarios, while edge computing has an irreplaceable role in real-time, short-period data, and local decision-making scenarios [10]. Therefore, the edge computing model and the cloud computing model can complement each other according to different needs.

Edge computing is combined with the existing cloud computing centralized processing model cloud edge collaboration, which migrates part or all of the computing tasks performed by the original cloud computing model to the network edge devices, forming a computing mode of "business application at the edge, management in the cloud," which completes the transformation of computing, network and storage capacity from cloud to network edge. For example, video surveillance is a typical application scenario. Due to the need for social security management and public security, many cameras are deployed in the streets to monitor the crowd in real time. However, due to a large number of cameras and real-time data generation, if all the data are uploaded to the cloud computing center, the network bandwidth and cloud computing center will be under great pressure. Therefore, the collected data can be preprocessed at the nearest storage device of each camera, and the useful data can be filtered out and transmitted to the cloud computing center, which then calculates according to the received data [11]. Therefore, edge computing and cloud computing complement each other to help realize the interconnection of all things.

*2.3. Edge Computing Advantages.* Edge computing can effectively reduce the delay of computing system, reduce the data transmission bandwidth, relieve the pressure of cloud

computing center, improve availability, and protect data security and privacy [12]. Compared with cloud computing, edge computing has the following advantages:

- (1) Edge computing is the first access to data, which enables edge computing to process a large number of real-time data at the edge of the network, without having to upload all the data to the cloud computing center. This reduces the energy consumption pressure of cloud computing centers and also reduces the load of network bandwidth.
- (2) Edge computing processes data near the edge of the network to reduce network latency and speed up service response.
- (3) Users' private data can be stored and processed on edge devices without network transmission, which reduces the risk of data leakage and improves data security and privacy.

To sum up, the edge computing model is more suitable for processing the data generated by users in the mobile edge terminal. The low latency and high efficiency meet the needs of people in tourism today. For the prediction of tourism destination choice preference, edge computing is undoubtedly the most matching calculation model.

### 3. Tourism Destination Preference Theory

**3.1. Tourist Destination.** Tourist destination refers to non-residential area, which is the place where tourists stay and visit for a short time, also known as a tourist destination or tourist resort. The formation of a tourism destination needs six elements: food, housing, transportation, tourism, shopping, and entertainment. Tourism destination is a specific geographical area managed by a unified destination management organization, which can be a specific scenic spot, or a town, or an area of a country, or the whole country, or even a larger place [13]. The common characteristics of tourist destinations are attraction, comfort, and accessibility, a specific area with certain tourism resources, facilities, and transportation conditions, which can attract a certain number of tourists to carry out tourism activities.

**3.2. Tourism Preference Theory.** Tourism preference refers to the influence of tourists' personality characteristics on tourism behavior. Personality characteristics include interest, hobby, ability, temperament, and personality. The study of tourism preference can be carried out from the aspects of tourists' age, occupation, educational background, gender, and living environment. According to psychology, attitude influences preference [14]. Tourism preference depends on the intensity and complexity of tourists' attitude. The stronger the attitude, the more influence the preference: the more complex the attitude, the easier to form preference. The formation process of tourism preference is as shown in Figure 2.

It can be seen from Figure 2 that the process of forming tourism preferences is more logical, and the strength of tourists' attitudes depends on the needs of

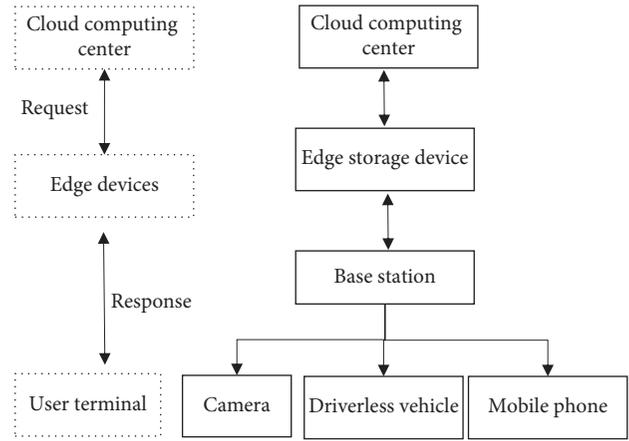


FIGURE 1: Edge computing architecture.

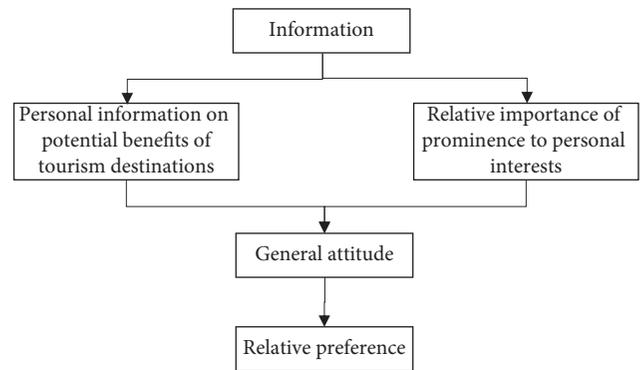


FIGURE 2: Process diagram of tourism preference formation.

tourists and the outstanding attributes of the destination. Once a person has a certain need, he will naturally have a desire, which will affect the strength of his attitude. For example, in a holiday resort, climate, comfort, and environment may be very important for one tourist, but for another tourist, price, transportation, and entertainment facilities are the most important attributes.

The complexity of attitude usually depends on the amount of information about the tourist destination. The more information, the more complex the attitude, the easier it is to form tourism preference, while a simple attitude is difficult to change [15]. For example, if a person has a negative attitude towards outbound tourism, it is difficult to persuade him to form an outbound travel preference.

In the process of forming attitude, two levels should be considered, namely, standard cognition level and interest evaluation level. Through the standard cognition, we can form the feeling and trust to the tourism products; through the benefit evaluation, when a tourist considers the alternative tourist destination, he should first estimate the visible or potential benefits that the destination can give him. After evaluation, plus the salient points of each interest, tourists can decide which tourism destination can fully meet their needs and purposes, and then form corresponding preferences and finally make decisions [16].

**3.3. Tourism Destination Mapping Theory.** Tourism destination mapping theory, namely tourism destination terrain image theory, refers to the comprehensive image of tourism destination obtained by people directly through tourism activities or indirectly through television, magazines, advertising, Internet, and other media. This image has a certain impact on tourists' choice of destination. Through the construction of the model, the author explores the map of tourist destination from the aspects of tourists' psychological characteristics, experience, perception, demand, education level, media, and marketing, etc., and summarizes and modifies the original image before the start of tourism activities, and finally forms the evaluation image after the tourism behavior combining with the experience image in the process of tourism activities. The formation of the initial impression of a tourist destination depends on the role of advertising, among which the introduction advertising, persuasive advertising, and reminder advertising have a great external impetus to the formation of the image.

**3.4. Tourism Decision Theory.** Tourism decision-making refers to the process in which tourists select relevant tourism information, formulate tourism plans or tourism plans, and finally put them into practice. The whole process is composed of tourism needs, tourism information, tourism routes, tourism budget, tourism mode, tourism decision, and tourism travel, etc. Among them, the tourism route is the key factor affecting tourism decision-making, which not only determines the division of tourism regions but also determines the tourism budget and tourism experience. Tourists will perceive the tourism information after hoof selection and then make tourism decisions through the influence of other factors such as preference.

The formation of tourism decision-making is the result of the interaction of individual needs, motives, and preferences of tourists and social pressure. The tourism destination selection opportunity group model is as shown in Figure 3.

The optional part is called the overall opportunity group, and the part sensed by tourists is called the induction opportunity group. Due to the limitation of certain conditions, only a part of the induction opportunity that tourists will consider is called the consideration opportunity group [17]. When considering the overlap between opportunity group and opportunity group, which can provide tourists in the market at that time, it is the part selected by tourists, which is called the possibility opportunity group. Therefore, only when the destination (perceived opportunity) is perceived by tourists and the destination (accessibility opportunity) is within the scope of tourists' economic ability can they enter into tourism decision-making and become the real tourism destination (real opportunity).

## 4. Tourism Destination Preference Prediction Based on Edge Computing

**4.1. Edge Computing Prediction Process.** The process of edge computing and prediction is mainly divided into two parts:

the cloud processing part and edge prediction part. The operation process of the two parts will be described, respectively.

- (1) Cloud processing: the collected tourism destination selection preference data is preprocessed in the cloud, and the tourism destination selection preference features are constructed. The important features of tourism destination selection preference features are selected by using a random forest algorithm, and preliminary estimation and ranking are carried out. Based on the multinomial logit selection model, the preference sequence of tourist destination choice is obtained, and the preference model of tourist destination selection is sorted. The model is saved and loaded to the edge [18].
- (2) Edge prediction: on the basis of the tourism destination selection preference model, the weight value of tourism destination selection preference is calculated, and the set of tourism destination selection preference weight is determined. Through the link prediction method to determine the tourism destination selection preference, using the tourism destination selection preference prediction formula, the tourism destination selection preference is predicted, thus realizing the tourism destination selection preference prediction. When the preference of tourism destination choice changes, the edge end will call the prediction model corresponding to the new model. When there is a big deviation between the predicted value and the real value, the edge end will make appropriate adjustments.

### 4.2. Tourism Destination Preference Model

**4.2.1. Construct Preference Characteristics of Tourism Destination Selection.** Feature construction refers to the generation and selection of features from the original information, also known as feature extraction. In the cognition and prediction of things, the selection feature is particularly prominent. The characteristics of tourism destination selection preference constructed in this paper are as shown in Figure 4.

The characteristics of tourism destination selection preference can be divided into tourist characteristics, tourism destination characteristics, and the interaction between them. The basic characteristics of passenger characteristics are mainly described from the granularity of passenger operation number, passenger expense number, passenger days, and passenger time. Among them, the granularity of passenger operation number mainly represents the number of tourist destination clicks in a specific period of time. The granularity of the number of tourist expenses refers to the number of tourist attractions tickets purchased in a specific period of time. Passenger days granularity refers to the number of days that tourists stay at the destination. Passenger time granularity refers to the time interval between passengers arriving at the destination recently [19].

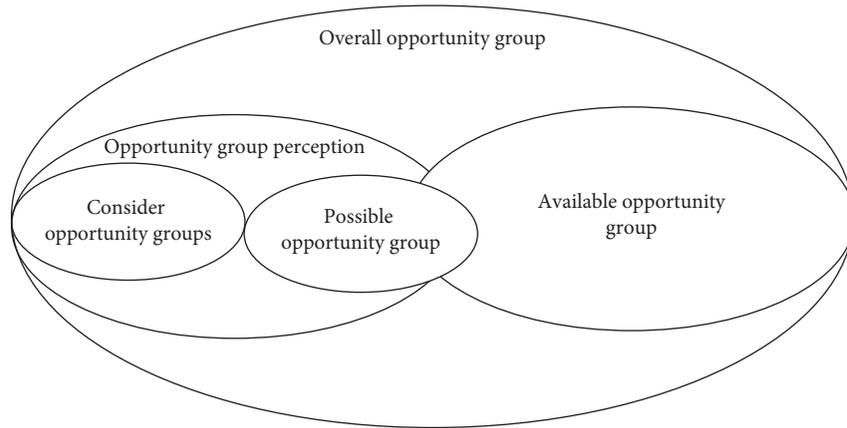


FIGURE 3: Tourism destination selection opportunity group model.

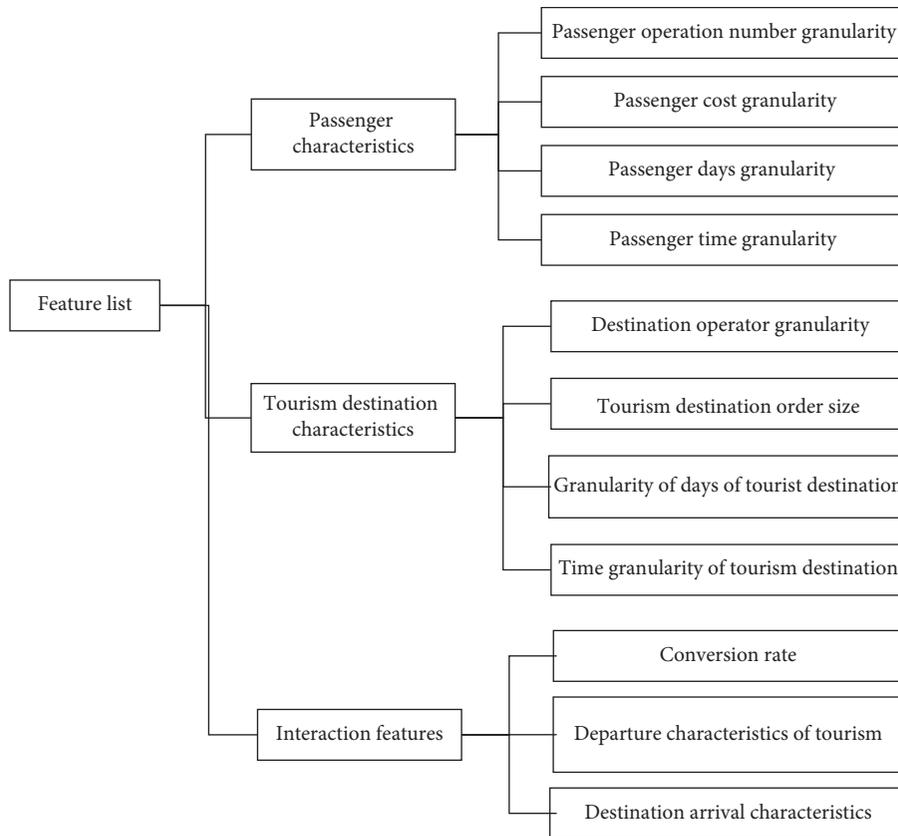


FIGURE 4: Preference characteristics of tourism destination selection.

Tourism destination characteristics are similar to the characteristics of tourists, but the calculation object is changed from tourists to tourism destinations. For example, the tourism destination operands granularity represents the number of times a tourist destination is queried by clicking at a specific time. The granularity of tourism destination order number refers to the purchased quantity of tourism destination in a specific time period, the tourism destination days granularity refers to the arrival days of tourism destination, and the tourism destination time granularity refers

to the time interval between one tourism destination and another.

In addition to the above-mentioned basic characteristics, the interactive characteristics of tourists and tourism destinations also include the conversion rate, the characteristics of departure, and the arrival of tourists' choice of tourism destinations [20]. Among them, the conversion rate refers to the number of times a passenger chooses a tourist destination to set out divided by the number of times a tourist chooses a tourist destination. The characteristics of a

tourist's choice of a tourist destination are mainly expressed as the date and time of departure or arrival.

Because there are many interactive features between tourists and tourist destinations, the most important features must be selected from the large number of features constructed, so as to reduce the number of features, improve the prediction accuracy of the model, and reduce the running time. The goal of the tourist preference sequence is to predict the preference selection model of tourist destination choice. Considering the operation efficiency and prediction accuracy of the algorithm, this paper proposes a pointwise sorting algorithm combining random forest and gradient lifting decision tree. In this paper, we propose a method to fuse a single prediction model and introduce a variety of prediction factors as far as possible, and finally, we can get a more accurate preference prediction model. Therefore, this paper uses a random forest algorithm [21] to select important features, and the steps are as follows:

- (i) Step 1: preliminary estimation and ranking.
  - (1) The characteristic variables of random forest were sorted in descending order according to the importance of variables
  - (2) According to the determined deletion ratio, the unimportant indexes in the current feature variables are eliminated to obtain new features
  - (3) A new random forest is established by using the new features generated, and the importance of each feature is calculated and sorted
  - (4) Repeat the above three steps until  $m$  features are left
- (ii) Step 2: according to the feature set obtained in the first step and the established random forest, the out-of-bag error rate is calculated. Finally, the selected feature set is the feature set with the lowest out-of-bag error rate.

According to the input feature training data of interaction between tourists and tourism destination, the first  $m$  features of all the features constructed in this paper are selected and input into multiple logit selection model by using the ranking relationship between cross validation method and the importance of the obtained features.

**4.2.2. Based on the Multiple Logit Selection Model.** The multinomial logit model is a random utility model, which is usually used as the main method to describe the choice of tourist destinations [22]. In the multilogit model, the utility of passenger  $t$  choosing a tourist destination  $j$  is defined as follows:

$$U_t^j = \mu_t^j + \xi_t^j, \quad j = 0, 1, \dots, n, \quad (1)$$

In formula (1),  $\mu_t^j$  represents the expected utility constant, and  $\xi_t^j, j = 0, 1, \dots, n$  represents independent noise. Generally, the linear model is established to express  $\mu_t^j$  by the characteristics of tourist destination selected by passenger  $t$ . For example,

$$\mu_t^j = \beta^T a_t^j. \quad (2)$$

In formula (2),  $\beta$  represents the parameter vector, and  $a_t^j$  represents the characteristics of the tourist destination  $j$  selected by the passenger  $t$ . Use the tourist destination selection preference feature to construct  $a_t^j$ . The noise  $\xi_t^j$  is defined as a double exponential distribution. The specific formula is as follows:

$$F_\xi(x) = P(\xi_t^j \leq x) = e^{-(\delta x + \gamma)}. \quad (3)$$

In formula (3),  $\gamma$  represents the Euler constant defined as the Euler constant, and  $\delta$  represents a positive scale parameter, and is often set to 1. Define  $C_t \subset N$  as the preference set of traveler  $t$  to choose a tourist destination, and the probability of choosing a tourist destination  $j \in C_t$  is as follows:

$$P_t(j) = \frac{e^{\mu_t^0}}{\sum_{i \in C_t} e^{\mu_t^i} + e^{\mu_t^0}}. \quad (4)$$

In formula (4),  $\mu_t^0$  represents the nonselection utility, and  $e^{\mu_t^i}$  represents the preference weight of the traveler  $t$  choice of the travel destination  $j$ . Then the probability of the traveler  $t$  choosing the travel destination  $j$  can be compared with the probability of choosing other preferences—destination selection preference sequence.

**4.2.3. Preference Model of Tourism Destination Selection.** Aiming at the problem of passengers' preference selection and ranking of tourist destinations, historical data is obtained again to construct the training set  $D$  to train the ranking model. The specific solution process is as follows:

- (i) Step 1: in order to make the tourism destination selection preference model get a better evaluation ranking effect, the target ranking value is defined as follows:

$$y_{ij} = 2R_{ij} - 1. \quad (5)$$

In formula (5),  $R_{ij}$  represents the predicted probability of tourists' preference for travel destination selection. A training set is obtained as follows:

$$D = [x_s, y_s]. \quad (6)$$

In formula (6),  $x_s$  represents the  $f$  dimensional characteristics of travelers' choice of destinations.

- (ii) Step 2: establish regression function, which is expressed as follows:

$$\sum_{(x_s, y_s)} [y_s - g(x_s)]^2. \quad (7)$$

The independent variables and dependent variables are continuous variables, and the data obtained from the questionnaire is the score value, which is a linear regression with the mean value of the data (Algorithms 1 and 2)

- (iii) Step 3: use random forest algorithm to train formula (7), and the algorithm process is as follows:
- (iv) Step 4: use the initial training result obtained in the previous step as the initial value input of the tourist destination selection preference model. The algorithm process is as follows:

According to the above method, the tourists' preference model for tourism destination selection is finally obtained.

**4.3. Realize the Prediction of Tourism Destination Choice Preference.** In order to predict the preference of tourism destination choice, when looking for the sequence of tourists' preference for a tourism destination, this paper combines the trust degree of tourists in tourism destination and the similarity of tourism destination selection preference as the weight value of tourism destination selection preference. The tourism destination selection preference is determined by the link prediction method, and the tourism destination selection preference is predicted by using the prediction formula of tourism destination selection preference. The specific methods are as follows.

**4.3.1. Calculate the Preference Weight Value of Travel Destination Selection.** In the process of tourism destination selection preference prediction, this paper not only considers the impact of similarity of tourism destination selection preference on the prediction results but also considers the influence of tourists' trust in tourism destination on tourism destination selection preference. The harmonic weight is used to select the weight value of tourism destination selection preference. The harmonic weight is determined by the selection similarity and passenger trust [23]. When the trust degree of tourists to tourism destination is relatively large, the weight value of trust degree is relatively large; when the trust degree of tourists to tourism destination is relatively small, the weight value of similarity is relatively large. Therefore, the calculation formula of harmonic weight is expressed as follows:

$$W(u_i, u_j) = \begin{cases} \alpha_1 \text{sim}(u_i, u_j) + \beta_1 T(u_i, u_j), & T(u_i, u_j) \geq \sigma, \\ \alpha_2 \text{sim}(u_i, u_j) + \beta_2 T(u_i, u_j), & T(u_i, u_j) < \sigma. \end{cases} \quad (8)$$

In formula (8),  $W(u_i, u_j)$  represents the reconciliation weight of mobile passenger  $u_i$  and mobile passenger  $u_j$ ,  $\text{sim}(u_i, u_j)$  represents the score similarity of mobile passenger  $u_i$  and mobile passenger  $u_j$ ,  $T(u_i, u_j)$  represents the trust degree of mobile passenger  $u_i$  and mobile passenger  $u_j$ , and  $\alpha_1, \alpha_2, \beta_1, \beta_2$  represents the reconciliation factor, and  $\alpha_1 + \beta_1 = 1$ ,  $\alpha_2 + \beta_2 = 1$ ,  $\sigma$  represent the trustworthiness threshold.

In determining the preference weight set of tourism destination selection, the  $o\%$  harmonic weight is selected as the preference weight set of tourism destination selection. When  $o = 0$ , the preference weight set of tourism destination selection is empty set; when  $o = 100$ , the preference weight

set of tourism destination selection includes all passengers in the data set. Therefore, if the value of  $o$  is too small or too large, the ideal result can not be obtained. In this paper, the reasonable  $o$  value is selected according to the results of many experiments.

**4.3.2. Determine the Preference Weight Set of Tourism Destination Selection.** The higher the similarity between tourists and tourism destination selection preferences, the greater the possibility of tourists' preference in tourism destination selection. Therefore, when determining the weight set of tourism destination selection preference to be predicted, this paper selects the tourism destination selection preference with correlation degree of  $\varepsilon\%$  as the weight set of tourism destination selection preference.

**4.3.3. Predict Tourism Destination Preference.** Combining the reconciliation weight with the tourism destination selection preference prediction formula, the final weighted average prediction formula is expressed as follows:

$$r_{ui,s} = \bar{r}_{ui} + k \sum_{\substack{u_j \in S_{u_j} \\ s \in N_{ui}}} [W(u_i, s, u_j) \times (r_{ui,s} - \bar{r}_{ui})]. \quad (9)$$

In formula (9),  $N_{ui}$  represents the travel destination selection preference weight set of mobile traveler  $u_i$ ,  $S_{u_j}$  represents the travel destination selection preference weight set of mobile traveler  $u_j$ , and  $k$  represents a standardized factor.

Through the above steps, using edge computing to construct tourism destination selection preference features, using random forest algorithm to select important features of tourism destination selection preference features, and preliminary estimation and ranking, using multiple logit, in this paper, we get the sequence of tourists' preference for tourism destination selection and rank the preference model of tourism destination selection, and then get the preference model of tourism destination choice. This paper calculates the weight value of tourism destination selection preference, determines the weight set of tourism destination selection preference, determines the tourism destination selection preference through link prediction method, and predicts the tourism destination selection preference by using the tourism destination selection preference prediction formula, so as to realize the tourism destination selection preference prediction.

## 5. Experimental Analysis

**5.1. Experimental Environment and Data.** In order to verify the effectiveness of the edge computing-based tourism destination selection preference prediction method, the experiment uses a computer configured with Inter E1400 2.0 GHz processor, 8.00 G memory, 800 G hard disk, and 64-bit Windows 10 operating system. In this paper, we obtain the data of about 3000 tourists' preference for tourism destination through a questionnaire and extract the

```

(i) Input: data set  $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$ , parameter:  $K: 0 < K \leq f, M_{RF}: M_{RF} > 0$ 
(ii) for  $t = 1$  to  $M_{RF}$  do
(iii)  $D_t \subseteq D$  #Random sampling sample  $D_t$ , making  $|D_t| = |D|$ .
(iv) #Establish a full tree ( $d = \infty$ )Cart and randomly select  $K$  features when the tree is decomposed downward
(v) end for
(vi)  $T(\cdot) = (1/M_{RF}) \sum_{t=1}^{M_{RF}} h_t(\cdot)$ 
(vii) return  $T(\cdot)$ 

```

ALGORITHM 1: Random forest.

```

(i) Input: data set  $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$ , parameter:  $\alpha, M_B, d, K_{RF}, M_{RF}$ 
(ii)  $F \leftarrow$  Random Forests ( $D_t, K_{RF}, M_{RF}$ )
(iii) Initialization:  $r_i = y_i - F(x_i)$  for  $i = 1$  to  $n$ 
(iv) for  $t = 1$  to  $M_B$  do
(v)  $T_t \leftarrow$  Cart ( $\{(x_1, r_1), \dots, (x_n, r_n), f, d\}$ ) #Build a tree Cart with a depth of  $d$ , and enter all the features  $f$  and  $\{r_i\}$ .
(vi) for  $i = 1$  to  $n$  do
(vii)  $r_i \leftarrow r_i - \alpha T_t(x_i)$  #Update the residual of each sample  $x_i$ .
(viii) end for
(ix) end for
(x)  $T(\cdot) = F(\cdot) + \alpha \sum_{t=1}^{M_B} T_t(\cdot)$ . #Merge the regression tree  $T_1, \dots, T_M$  and the  $F$  obtained by the random forest
(xi) return  $T(\cdot)$ 

```

ALGORITHM 2: Initialize gradient boosted regression trees (squared loss).

historical data of destination preference of these tourists. Taking the above data as a data record of tourists' choice of tourism destination products at different times, the bootstrap method is used to repeatedly sample and generate 1000000 simulated data. Each data include passenger's age, surname, occupation, education background, location, destination selection, travel date, reservation date, etc. The experimental environment is carried out under the joint support of multiple terminal machines.

**5.2. Evaluation Index.** This article uses mean absolute error (MAE) and  $F$  index ( $F$  measure) as evaluation indicators. MAE measures the accuracy of prediction by calculating the deviation between the predicted travel destination preference and the actual travel destination preference. The smaller the MAE value, the higher the accuracy of the prediction. Assuming that the predicted tourism destination selection preference set is denoted as  $p_i = \{p_1, p_2, \dots, p_n\}$ , and the corresponding actual tourism destination selection preference set is denoted as  $q_i = \{q_1, q_2, \dots, q_n\}$ , then MAE is defined as follows:

$$MAE = \frac{\sum_{i=1}^n |p_i - q_i|}{n} \quad (10)$$

$F$  index is a comprehensive evaluation form of precision rate and recall rate. The larger the  $F$  index, the higher the comprehensive accuracy. This article redefines the precision rate and recall rate according to the evaluation standard of experimental results as follows:

The precision rate quantifies the accuracy of the prediction of the preference of tourists to the travel destination. It can be expressed as the ratio of the effective, relevant

preference to all predicted preferences. The precision rate can be expressed as follows:

$$P_z = \frac{N_{rs}}{N_s} \quad (11)$$

The recall rate reflects the comprehensiveness of the tourism destination selection preference prediction. The recall rate is the ratio of the passenger's related tourism destination selection preference to the selection preference of all tourism destinations. Its form can be expressed as follows:

$$R_z = \frac{N_{rs}}{N_r} \quad (12)$$

The  $F$  index is expressed as follows:

$$F = \frac{2 \times P_z \times R_z}{P_z + R_z} \quad (13)$$

In order to verify the prediction accuracy of the tourism destination selection preference prediction method based on edge computing, the reference [4] method, the reference [5] method and the proposed method were used to compare the average absolute deviation results of different methods as shown in Figure 5.

It can be seen from Figure 5 that under different tourist destination selection preference sets, the average absolute deviation of reference [4] method is 0.68, the average absolute deviation of reference [5] method is 0.73, and the average absolute deviation of the proposed method only 0.59. Therefore, compared with the reference [4] method and reference [5] method, the average absolute deviation of the proposed method is smaller, and the prediction accuracy of

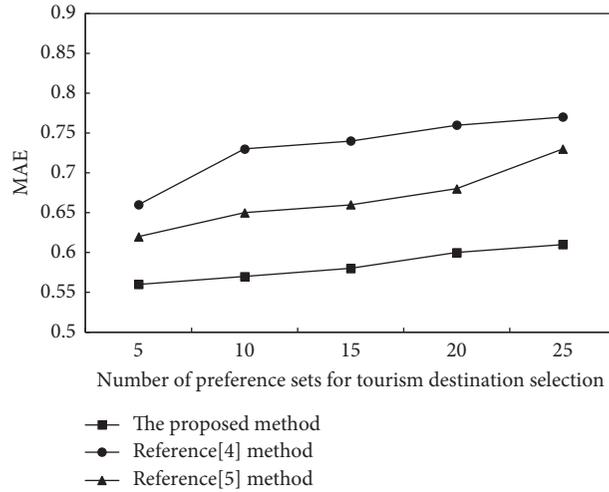


FIGURE 5: Comparison of mean absolute deviation of different methods.

tourism destination preference is higher because the proposed method uses edge computing to construct the characteristics of tourism destination selection preference and determines the weight set of tourism destination selection preference by calculating the value of tourism destination selection preference weight, so as to predict the tourism destination selection preference, which can effectively improve the accuracy of tourism destination selection preference.

**5.3. Comparison of Comprehensive Accuracy of Tourism Destination Choice Preference Prediction.** To further verify the comprehensive accuracy of the tourism destination selection preference prediction method based on edge computing, reference [4] method, reference [5] method, and the proposed method are used to compare the  $F$  index results of different methods as given in Figure 6. Reference [4] is the daily origin-target tuple estimation and prediction using multilayer Bayesian networks. Reference [5] is the destination prediction based on the trajectory distribution model.

It can be seen from Figure 6 that under different tourist destination selection preference sets, the  $F$  index of reference [4] method is 0.35, the  $F$  index of reference [5] method is 0.68, and the  $F$  index of the proposed method is 0.83. Therefore, compared with reference [4] method and reference [5] method, the proposed method has larger  $F$  index and higher comprehensive accuracy of tourism destination selection preference prediction because the proposed method calculates the weight value of tourism destination selection preference, determines the weight set of tourism destination selection preference, uses the link prediction method to determine the tourism destination selection preference, and uses the tourism destination selection preference prediction formula to realize the better prediction of tourism destination selection, so as to effectively improve the comprehensive accuracy of tourism destination selection preference.

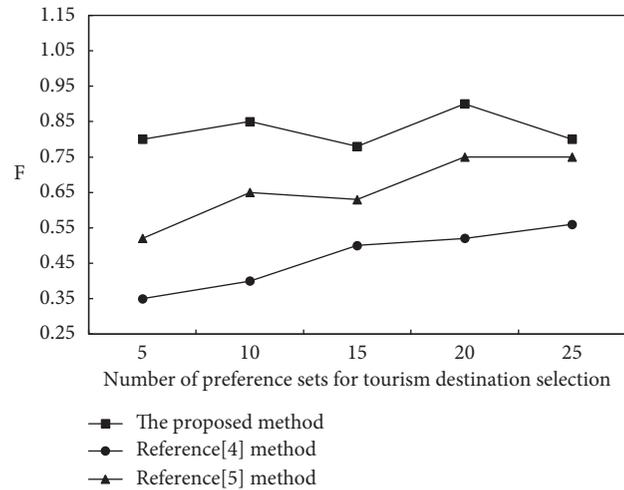


FIGURE 6: Comparison of  $F$  indexes of different methods.

**5.4. Comparison of Travel Destination Preference Prediction Time.** On this basis, the prediction time of tourism destination selection preference prediction method based on edge computing is verified. Reference [4] method, reference [5] method, and the proposed method are used to compare the prediction time of tourism destination selection preference of different methods as shown in Table 1.

According to the data in Table 1, when the number of experiments reaches 5, the average tourism destination selection preference prediction time of reference [4] method is 6.33 s, and the average tourism destination selection preference prediction time of reference [5] method is 9.15 s, and the average tourism destination preference prediction time of the proposed method is only 3.16 s. It can be seen that compared with reference [4] method and reference [5] method, the tourism destination preference prediction time of the proposed method is shorter. This is due to the proposed method; from a large number of tourism destination selection preference features, a random forest algorithm is used to select the important features of tourism destination

TABLE 1: Comparison of prediction time of tourism destination selection preference with different methods.

Number of experiments	The proposed method	The reference [4] method	The reference [5] method
1	3.34	5.98	8.79
2	3.18	5.87	8.93
3	2.97	5.76	9.65
4	3.23	7.05	9.96
5	3.09	6.99	8.97
Average value	3.16	6.33	9.15

selection preference features; based on multiple logit, the choice model can get the sequence of tourists' preference for tourism destination selection and sort the tourism destination selection preference model to obtain the tourism destination selection preference model, thus effectively shortening the prediction time of tourism destination selection preference.

## 6. Conclusion

This paper proposes a tourism destination selection preference prediction method based on edge computing, which gives full play to the advantages of edge computing. It can effectively shorten the prediction time of tourism destination selection preference and has high comprehensive accuracy performance. However, in the process of tourism destination selection preference prediction, this method lacks the collection of tourism destination selection preference data and cannot verify the tourism destination selection preference model. Therefore, in the next step of research, we need to collect the data of tourism destination selection preference in real time so as to verify the effect of the tourism destination selection preference model and further improve the model to make the prediction results more accurate.

## Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that they have no conflicts of Interest.

## References

- [1] R. Steiger, E. Posch, G. Tappeiner, and J. Walde, "The impact of climate change on demand of ski tourism - a simulation study based on stated preferences," *Ecological Economics*, vol. 170, Article ID 106589, 2020.
- [2] V. C. Sander, G. C. Caspar, and W. Bert van, "Simulation study on impacts of high aviation carbon taxes on tourism: application of portfolio vacation choice model," *Transportation Research Record*, vol. 2449, no. 1, pp. 64–71, 2018.
- [3] K. Lindner-Cendrowska and K. Błażejczyk, "Impact of selected personal factors on seasonal variability of recreationist weather perceptions and preferences in Warsaw (Poland)," *International Journal of Biometeorology*, vol. 62, no. 1, pp. 113–125, 2018.
- [4] Y. Ma, R. Kuik, and H. Zuylen, "Day-to-Day origin-destination tuple estimation and prediction with hierarchical bayesian networks using multiple data sources," *Transportation Research Record*, vol. 2343, no. 1, pp. 51–61, 2018.
- [5] P. C. Besse, B. Guillouet, J.-M. Loubes, and F. Royer, "Destination prediction by trajectory distribution-based model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2470–2481, 2018.
- [6] D. Li, G. Wu, J. Zhao, W. Niu, and Q. Liu, "Wireless channel identification algorithm based on feature extraction and BP neural network," *Journal of Information Processing Systems*, vol. 13, no. 1, pp. 141–151, 2017.
- [7] X. Tongle, W. Yingbo, and C. Kang, "Tailings saturation line prediction based on genetic algorithm and BP neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 30, no. 4, pp. 1947–1955, 2016.
- [8] R. Zhao, M. Jin, P. Ren, and Q. Zhang, "Stable two-sided satisfied matching for ridesharing system based on preference orders," *Journal of Supercomputing*, vol. 76, no. 10, pp. 1–19, 2016.
- [9] T. Ren, S. Liu, G. Yan, and H. Mu, "Temperature prediction of the molten salt collector tube using BP neural network," *IET Renewable Power Generation*, vol. 10, no. 2, pp. 212–220, 2020.
- [10] M. Lozano-Oyola, F. J. Blancas, M. González, and R. Caballero, "Sustainable tourism tags to reward destination management," *Journal of Environmental Management*, vol. 250, no. Nov.15, pp. 1–11, 2019.
- [11] L. Wu, J. Zhang, and A. Fujiwara, "Dynamic analysis of Japanese tourists' three stage choices: tourism participation, destination choice, and travel mode choice," *Transportation Research Record*, vol. 2322, pp. 91–101, 2018.
- [12] S. Wan, R. Gu, T. Umer, K. Salah, and X. Xu, "Toward offloading Internet of vehicles applications in 5G networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, pp. 1–9, 2020.
- [13] A. Rossi, G. Barlacchi, M. Bianchini, and B. Lepri, "Modelling taxi drivers' behaviour for the next destination prediction," *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1–10, 2019.
- [14] Z. Gao, K. X. Xue, and S.H. Wan, "Multiple discrimination and pairwise CNN for view-based 3D object retrieval," *Neural Networks*, 2020, <https://arxiv.org/abs/2002.11977>.
- [15] J. Lv, Q. Sun, Q. Li, and L. Moreira-Matias, "Multi-scale and multi-scope convolutional neural networks for destination prediction of trajectories," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3184–3195, 2020.
- [16] X. Xu, B. Shen, X. Yin et al., "Edge server quantification and placement for offloading social media services in industrial cognitive IoT," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2910–2918, 2020.
- [17] X. Xian, H. Ye, X. Wang, and K. Liu, "Spatiotemporal modeling and real-time prediction of origin-destination traffic demand," *Technometrics*, vol. 63, pp. 1–20, 2019.

- [18] X. Duan, D. Guo, and C. Qin, "Image information hiding method based on image compression and deep neural network," *Computer Modeling in Engineering & Sciences*, vol. 124, no. 2, pp. 721–745, 2020.
- [19] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.
- [20] A. Faghih-Imani and N. Eluru, "A finite mixture modeling approach to examine New York City bicycle sharing system (CitiBike) users' destination preferences," *Transportation*, vol. 47, no. 2, pp. 529–553, 2020.
- [21] H. Alibabai and H. S. Mahmassani, "Dynamic origin-destination demand estimation using turning movement counts," *Transportation Research Record*, vol. 2085, no. 1, pp. 39–48, 2018.
- [22] G. Feng, Y. Han, Y. Wang, and J. Cheng, "Modeling and simulation analysis of QoE perception of network quality based on user preferences," *Computer Simulation*, vol. 37, no. 4, pp. 356–360, 2020.
- [23] K. F. Chu, A. Y. S. Lam, and V. O. K. Li, "Deep multi-scale convolutional LSTM network for travel demand and origin-destination predictions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3219–3232, 2019.