Optimal Planning Method of Rural Tourism Route Based on Multi Constraint and Multi Objective

Qingqing Geng and Yi Liu(✉)

Chongqing College of Architectural and Technology, Chongqing 401331, China
liuyi13808301860@163.com

Abstract. In order to recommend rural tourism routes that best meet users' tourism needs, a multi constraint and multi-objective optimal planning method for rural tourism routes is proposed. Firstly, collect and mine tourist generated data, obtain personalized preferences of tourists, construct a comprehensive evaluation model for tourist attractions, and obtain comprehensive evaluation results for tourist attractions. Secondly, from the perspective of multiple constraints and planning objectives, establish an optimal planning model for rural tourism routes. Finally, by applying the evaluation function method to solve the above construction model, the optimal planning results of rural tourism routes can be obtained. The experimental data shows that after the application of the proposed method, the optimal planning results of rural tourism routes are consistent with the optimal routes, and the minimum number of iterations for solving the optimal planning results of rural tourism routes is 8, fully confirming the better application performance of the proposed method.

Keywords: Multiple Objectives · Rural Tourism · Route Planning · Route Optimization · Tourism Routes · Multi Constraint

1 Introduction

In recent years, with the increasing prosperity of the national economy, people's demand for spiritual aspects has also been increasing. In this trend, tourism has gradually become a popular way of leisure, and as an important driving force for rural development, rural tourism has become an important topic of concern for all sectors of society in recent years. At the same time, a large number of rural tourism areas across the country also showed explosive growth, which to some extent promoted the development of rural tourism industry. Currently, the tourism industry is developing rapidly and occupies an important position in the national economy. The rapid development of the Internet has brought new vitality to the tourism industry, changing the way residents travel [1].

If a person wants to travel at a certain time, they have some basic preferences and constraints, such as planned time constraints, money expenses, types of favorite attractions, peers, accommodation requirements, etc. Therefore, when planning tourist routes for

L. Yun et al. (Eds.): ADHIP 2023, LNICST 547, pp. 75–90, 2024.
https://doi.org/10.1007/978-3-031-50543-0_6

tourists, it is necessary to consider the constraints of tourists from multiple aspects and achieve as many tourist goals as possible. Among the existing research results, the time-based route planning problem model and the tourism route recommendation model based on pattern and preference perception are more widely applied [2]. The former focuses on time and proposes a corresponding route planning problem model. By solving it, a travel route planning result that includes time can be obtained and recommended to users; The latter constructed a system architecture based on the proposed travel route recommendation, and then modeled the mobility patterns of each user. Finally, a tourism route recommendation plan was proposed to develop tourism routes for target users and recommend personalized services. The above two methods consider fewer influencing factors and cannot obtain the optimal rural tourism route, which reduces tourists' interest and restricts the subsequent development of rural tourism. Therefore, a multi constraint and multi-objective optimal planning method for rural tourism routes is proposed.

2 Research on the Optimal Planning Method of Rural Tourism Route

2.1 Collection and Mining of Visitor Generated Data

Each tourist has different preferences, and their corresponding rural tourism route planning models also have some differences. In order to meet the personalized needs of tourists, the data generated by tourists are collected and mined to obtain tourists' preferences, which lays a solid foundation for the establishment of the optimal planning model of the follow-up rural tourism routes.

The data generated by tourists are mainly image information and text information. Among them, image features are used to express the collected image information, as shown below:

First, global image features

Color feature: color feature is a feature based on pixels, that is, all pixels in the image contribute to the statistical results of the image feature. Therefore, color feature cannot describe the local changes of the image well.

Texture feature: The advantage of texture feature is that it is calculated statistically in the area containing multiple pixels, rather than based on pixel features. In pattern matching, it will not fail to match due to local deviation.

Shape features: The extraction of shape features is very important for object recognition and retrieval [3]. In order to retrieve or recognize objects or objects in images, it is very effective and necessary to extract the shape information of objects or objects. However, the extraction and application of shape features still face great challenges. The most common method, Hough transform, has a large amount of computation, and cannot effectively detect the line segments in the image, as well as the lines including the line segments, and cannot complete the detection of line segments with fewer pixels.

Second, local image features

SIFT feature: SIFT feature is a local image feature based on key points, first proposed by David Lowe, and is considered to be one of the most robust local image features. SIFT features first find the key points in the image and express them in vector form one by one. These feature vectors have the characteristics of rotation invariance, scale invariance, and illumination invariance.

Color SIFT feature: Color SIFT feature was proposed by Alaa E. Abdel Hakim and Aly A. Farag. This feature considers both color and geometric shape information when describing objects or objects in the image. The color invariance is guaranteed by the model of Geusebroek et al., while the geometric invariance is guaranteed by the SIFT feature, so that the feature makes good use of the discriminability of color information, and guarantees the color invariance (the stable feature that the color of the object surface in the scene is independent of the object shape is called color invariance).

The collected text information is described by text feature units. The representation unit of text information can be words or phrases from the perspective of statistical machine learning. The perspective of natural language processing is to define various syntactic rules. This research mainly introduces the expression of text information from the perspective of statistical machine learning. In addition, Chinese text expression involves word segmentation. Compared with phrases, word segmentation is much easier than phrase segmentation. Therefore, when processing Chinese data, words are often used as text features to express the content of the text. Text features are mainly as follows:

First, word frequency

Word frequency is the number of times a word appears in a document. Text expression based on word frequency is proposed based on the assumption that words with high frequency have a relatively important impact on the expression of document content. But information retrieval theory believes that words with low frequency may also contain more information. Therefore, simply using word frequency to describe the content of a document is sometimes inaccurate.

Second, document frequency

Document frequency refers to the number of documents containing a word in the entire dataset. Calculate the document frequency for each feature in the document set. If the value is large, it indicates that many documents contain this feature, and then it is considered that this feature has no discrimination; If the value is very small, it indicates that very few documents contain this word, and this feature is not representative for this document set. The features in these two extreme cases cannot well express the content of the document [4].

Third, TF-IDF

Compared with text expression based on word frequency and document frequency, TF-IDF is a very effective text expression method, proposed by Salton. TF is the word frequency, which is used to calculate the ability of each word in the document to describe the document content; IDF is called anti document frequency, which is used to calculate the ability of each word in a document to distinguish content differences between documents, and is inversely proportional to document frequency DF. TF-IDF is based on the assumption that words that appear many times in a document will also appear many times in documents expressing similar meanings, and vice versa. Therefore, the larger the TF and the smaller the IDF, the better the distinguishability of this feature, and the better the content of such documents can be expressed. The TF-IDF calculation formula is:

$$\Omega = \#\left(t_k, d_j\right) * \log_2 \frac{|T_r|}{\#T_r(t_k)} \tag{1}$$

In Eq. (1), Ω represents the TF-IDF value; $\#\left(t_k, d_j\right)$ represents the number of occurrences of the word t_k in article d_j; $|T_r|$ represents the number of documents; $\#T_r(t_k)$ represents the total number of documents where the word t_k appears.

Based on the collected tourist generated data mentioned above, apply the Location Topic model to conduct in-depth mining on it. The Location Topic model is a statistical model used for text analysis and topic modeling. It combines the geographic location information and thematic information of documents, aiming to reveal the correlation between space and themes in text data. This model assumes that the document generation process includes two main steps: topic selection and word generation. In the topic selection step, the document selects topics from a topic distribution that is related to the geographic location of the document. In the word generation step, the document generates specific words based on the word distribution of the selected topic. In this way, the Location Topic model can combine themes with geographic locations to reveal spatial thematic patterns present in text data. This is very useful for many application scenarios, such as social media analysis, regional feature analysis of news reporting, etc.

Due to the inevitable connection between tourist generated data and location, location information was taken into account when establishing the model. It is not difficult to imagine that the following information usually appears in tourist generated data: travel expenses, local characteristics, travel time, transportation, accommodation, playing mood, and so on. Among them, travel expenses, transportation, accommodation, and entertainment mood are all irrelevant information to the location, while local characteristics are closely related to the location. Considering the above domain knowledge, the Location Topicl model first divides the possible semantics in tourist generated data into two categories when modeling, one is location related and the other is location independent. Based on this, semantic mining of tourist generated data is carried out. The process of generating a document in the LT model is as follows: firstly, the generation of the document is simplified to the generation of paragraphs. When several paragraphs are all generated, a set of tourist generated data is also completed. The advantage of this is that each segment can be constructed according to different locations. For each paragraph, first determine the type of theme in the paragraph, which semantic themes are related to the location, and which semantic themes are not related to the location. For each topic, generate words based on their distribution on the topic, thus forming a discourse.

In addition, with the spatio-temporal data becoming more and more accessible, the analysis and mining of spatio-temporal data has gradually become a new research field and hotspot. The partition group framework extracts common parts from scattered route sets by dividing long routes into small segments for clustering. The concept of route pattern uses route mining to analyze the user's behavior in time and space, such as the geographical region the user passes, the time the user walks, etc.

The above process completes the collection and mining of visitor generated data, determines user preferences, and provides support for subsequent research.

2.2 Construction of Comprehensive Evaluation Model for Scenic Spots

The quality of scenic spots is one of the main influencing factors of rural tourism route planning, so before the establishment of the optimal planning model of rural tourism routes, it is necessary to build a comprehensive evaluation model of scenic spots to obtain the comprehensive evaluation results of scenic spots.

The construction of a comprehensive evaluation model for scenic spots is mainly divided into three steps. Firstly, it is necessary to analyze and summarize the influencing factors of the comprehensive rating of scenic spots. A comprehensive analysis and induction of the influencing factors of the comprehensive evaluation of scenic spots is an important step in the construction of a comprehensive evaluation model for scenic spots. Otherwise, it will greatly affect the objectivity and reliability of the final comprehensive rating of the tourist attraction, and thus greatly affect the tourist experience of the final planned route. Secondly, use random forest calculation to get the importance ranking of all the influencing factors summarized in the first step, select some influencing factors according to the importance ranking, and build a more objective comprehensive evaluation model of scenic spots than the simple tourist rating [5]. Finally, the weight coefficients of several influencing factors selected in the second step are calculated by using the information entropy assignment method, and the comprehensive evaluation model of scenic spots is obtained.

Drawing on existing research and further analysis of practical problems, this paper analyzes the influencing factors related to scenic spot rating from the perspectives of scenic spots, transportation and tourists. In fact, when analyzing the rating of scenic spots, one of the factors can be selected to measure the rating of scenic spots, or multiple factors can be selected to comprehensively measure the rating of scenic spots through previous experience. However, such measurement method is relatively not objective and comprehensive. Therefore, this section selects random forests to obtain the importance ranking of these factors, and constructs a comprehensive evaluation model of scenic spots according to the importance ranking. Thus, relatively objective comprehensive scores of scenic spots can be obtained.

The influencing factors of comprehensive evaluation of scenic spots are as follows:

(1) Scenic spot factors

For travel, scenic spots are closely related to tourists and routes. From the perspective of scenic spots, the relevant influencing factors are: scenic spot resources, tourist seasons, scenic spot tickets, specific categories of scenic spots, levels of scenic spots and recommended travel duration of scenic spots. These factors will affect the planning of tourism routes to some extent. This paper mainly collected the relevant data of scenic spot resource indicators, travel duration and scenic spot level. The specific data of scenic spot resources were obtained through the tourism census conducted by the National Bureau of Statistics, and the scenic spot resources were evaluated from various aspects to obtain the final data.

(2) Traffic factors

Traffic accessibility is the main consideration of traffic factors. According to the existing research, tourists mainly gather in a certain area centered around the city. If this range is exceeded, the greater the cost of transportation time is, the fewer tourists will be. A large number of relevant studies have proved that the more densely distributed the scenic spots are, the more likely they will be selected and planned by tourists. That is to say, if a scenic spot has better traffic accessibility than other scenic spots, then the scenic spot is more likely to be selected, and the degree of accessibility is measured by time, expressed as

$$T_i = \frac{\sum_{j=1}^{n} L_{ij}}{V} \tag{2}$$

In Eq. (2), T_i represents the total transportation time from the i-th scenic spot to the other scenic spots; n represents the total number of rural tourist attractions; L_{ij} represents the distance between attraction i and attraction j; V represents the speed of the mode of transportation.

(3) Tourist factors

The purpose of planning tourist routes is to serve tourists. Therefore, the evaluation results of tourists on tourist attractions have great reference significance in planning tourist routes. Here we mainly collect the data of comprehensive environmental score. The comprehensive environmental score mainly refers to the tourists' evaluation of the infrastructure, transportation, health and environment of the scenic spot through the indicators of very good, good, average, poor and very poor. The corresponding scores are 5, 4, 3, 2 and 1 respectively. The average score is the comprehensive environmental score, and the calculation formula is:

$$H_i = \frac{\sum_{j=1}^{4} \Upsilon_{ij}}{4} \tag{3}$$

In Eq. (3), H_i represents the comprehensive environmental rating of the i-th scenic spot; Υ_{ij} represents the corresponding scores for infrastructure, transportation, health, and scenic environment.

After analyzing the influencing factors, usually several important factors can be selected based on previous experience to construct a model and calculate the comprehensive rating of the scenic spot. However, such an approach may not be objective and comprehensive enough. Therefore, it is necessary to obtain the importance ranking of multiple factors through specific calculations before constructing a comprehensive evaluation model for scenic spots. The random forest method has two main advantages: the first is that it can measure and analyze the utility of multiple variables in a large amount, mainly because it can improve the accuracy of analysis and measurement, but will not increase the amount of calculation significantly; The second point is that the processing results of missing and unbalanced data are relatively stable, mainly because of its low sensitivity to multivariate collinearity. The calculation of variable importance is to first change the value of independent variable of out of pocket data in random forest, and then measure the importance of the factor by the mean reduction of classification accuracy of the observed variable before and after the change [6–8].

The specific analysis process of the importance of influencing factors is as follows:

(1) First, analyze and measure each decision tree through out of pocket data, and then record the deviation of each out of pocket data analysis and measurement. Use $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_b$ to represent the deviation value of each decision tree;ourist factors
(2) The second step is to generate new out of pocket data. The specific operation method is to perform unrestricted noise interference on the sample feature P on the data outside the original bag. Then use the old data to test the new data to get the new error, which is expressed as $\varepsilon_{11}, \varepsilon_{12}, \cdots, \varepsilon_{1b}$;
(3) Finally, calculate the importance of variables. The specific operation method is to subtract the deviation measured by analyzing the data outside the new bag from the deviation measured by analyzing the data outside the old bag, and then calculate the average value of the sum. Set the number of decision trees to, and the formula for calculating the importance of variables is:

$$\omega_i = \frac{(\varepsilon_{11} - \varepsilon_1) + (\varepsilon_{12} - \varepsilon_2) + \cdots + (\varepsilon_{1b} - \varepsilon_b)}{M} \tag{4}$$

In Eq. (4), ω_i represents the importance value of the influencing factors for the comprehensive evaluation of the i-th scenic spot.

Formula (4) is used to calculate the importance ranking of multiple influencing factors, and select several relatively important influencing factors to build a comprehensive evaluation model of scenic spots. If A/B/C is used to represent the selected factors, the expression of the comprehensive evaluation model of scenic spots is:

$$Y_i = (\alpha * a_i + \beta * b_i + \chi * c_i) * 100\% \tag{5}$$

In Eq. (5), Y_i represents the comprehensive evaluation result of the i-th scenic spot; a_i, b_i and c_i represent the relative scores of the i-th scenic spot regarding the influencing factors A/B/C; α, β and χ represent the weights of the influencing factors A/B/C for the i-th scenic spot.

The above process completes the construction of the comprehensive evaluation model of scenic spots and the acquisition of the comprehensive evaluation results of scenic spots, which facilitates the establishment of the optimal planning model of the follow-up rural tourism routes.

2.3 Construction of Optimal Planning Model for Rural Tourism Routes Based on Multi-objective and Multi Constraint

In this section, we mainly solve the problem of helping tourists to determine the tour sequence of scenic spots and generate a tourism route suitable for tourists after they determine their hotel and the scenic spots to travel, on the premise of meeting the tourists' multi constraint and multi tourism goals. This section establishes a multi constraint and multi-objective tourism route planning model.To facilitate our calculation, we first set up reasonable model assumptions according to the actual problems, and then define the relevant constraints to establish the model [9]. When tourists need to plan a tour, they need to input their own personalized preferences, such as the start time and end time of the tour, the type of scenic spots they want to visit, the estimated travel costs, peer groups and other information. After receiving the request from tourists, we recommend hotels and scenic spots with higher scores to users according to the hotel scenic spot scoring model shown in the previous section, and then use the tourism route planning model to plan the travel for tourists. Under the condition of meeting the constraints of tourists, we recommend appropriate tourism routes for tourists.

For the convenience of calculation, this article needs to set reasonable assumptions based on the actual travel environment, so that unimportant factors do not affect the planning results of the route. On the basis of collecting a large amount of tourism industry growth data, tourist data volume, and other information, combined with recent tourism trends, under the guidance of experts, the following assumptions are made:

(1) Tourists always start from their hotel. Because the luggage and articles taken by tourists are stored in the hotel, tourists will always return to the hotel after the tour, regardless of the possibility of changing the hotel.
(2) By default, tourists use private cars for self-help travel, regardless of fuel fees, tolls, parking fees and other possible accidental charges during driving. That is, the toll is not considered in the model.
(3) The starting price of the hotel is taken as the price for staying in the hotel, and the hotel is not full, that is, tourists can always book the hotel they want to stay in.
(4) There is no change in ticket prices for tourist attractions in off season and peak season, and there is no preferential price for students, teachers and other different groups.
(5) The food cost of tourists during the travel is not considered.
(6) The travel time of tourists only includes travel time and travel time, without considering other time that may be spent, tourists will not visit the same scenic spot repeatedly, and the time of each tourist destination is always fixed.
(7) Tourists always enter the scenic spot during its opening hours, regardless of the closure of the scenic spot.
(8) The influence of weather, traffic jam and other conditions on the day of travel is not considered.

This section recommends hotels and scenic spots to tourists according to the hotel scenic spot rating model to help tourists plan the tourist routes that meet their preferences. The goal of route planning is to maximize the total score of routes and the number of scenic spot types, and time, cost, etc. are constraints of the model [10–12]. To facilitate the subsequent description of the article, based on the above assumptions, the constraints of the construction model are defined as follows:

(1) Tourist information: Each tourist has their own preference constraints, including time constraints UT, cost constraints UC, and other information.
(2) Attraction Information: Each recommended attraction set $S = \{s_1, s_2, \cdots, s_n\}$ should include information such as recommended travel time ST, attraction rating SS, attraction ticket price SP, attraction type SG, suitable season SD, suitable audience SC, longitude SL, latitude SA. For different attractions, tourists will prioritize the attraction with the highest total constraint score.
(3) Play route: If the edge set of the path is represented by R, then $R = \{r_1, r_2, \cdots, r_K\}$. The weight of any edge is the travel time T between two scenic spots. According

to the assumption in the previous text, tourists arrive between two different scenic spots by car. The shortest driving time obtained from Baidu Maps is used as the distance time between the two places, and the driving time is represented as the weight of each edge.

(4) Route map: The route map refers to the network information of tourist attractions during their travels. If Q is used, then $Q = (S, R)$, S represents the set of tourist attractions, and R represents the edge set between attractions. For example, the route composed of tourists departing from Hotel A, visiting attractions B, C, D, E, and F, and returning to Hotel A is the tourist's itinerary.

(5) Total travel path: If TR represents the total travel path for tourists, then $TR = \{h_0, s_k, s_{k+1}, \cdots, s_{k+n}\}$, which includes the hotels where tourists stay and the tourist attractions. s_k represents the scenic spots that tourists have passed by, and the time of arrival at the scenic spot s_k is represented by $TRT(s_k)$.

(6) Total cost of play TRF: Based on the above assumptions, in this article, except for the hotel fees and ticket prices of tourists, other expenses of tourists are not considered, and only individual expenses are considered. The total cost of a tourist's total travel path $TR = \{h_0, s_k, s_{k+1}, \cdots, s_{k+n}\}$ is the ticket price of all attractions plus the price of the tourist's hotel stay.

(7) Time spent in a single scenic spot: the time spent in a scenic spot is the average time recommended on the website.

(8) Total travel time: the total travel time of tourists includes the driving time of the user from the hotel to the target scenic spot plus the travel time of the target scenic spot plus the time of the last return to the hotel.

(9) Attraction Total Constraint Score: For each recommended attraction, there is a total constraint score for the attraction, which is calculated by the hotel attraction rating model mentioned above as the total constraint score SRS for that attraction. If the total score of a tourist attraction is higher, it indicates that it can better meet the personalized preferences of tourists.

(10) Total score of route TRS: After a tourist chooses a tourist route, $TR = \{h_0, s_k, s_{k+1}, \cdots, s_{k+n}\}$ will have a total score of the route. The calculation method for the total score of the route is the sum of the scores of the various scenic spots that the tourist has passed through.

(11) Effective route: The effective route for tourists is the route that meets the total travel time constraints of tourists, and the cost does not exceed the maximum cost of tourists' budget. If one item does not meet the needs of tourists, the route cannot be said to be effective.

According to the above constraints, the optimal planning model of rural tourism routes is established, and the expression is

$$\begin{cases} \max TRS(TR) \\ \max SG(TR) \\ s.t.\ TRT(TR) \le UT \\ TRF(TR) \le UC \end{cases} \tag{6}$$

The above process completed the construction of the optimal planning model of rural tourism routes, laying the foundation for the realization of the final research goal.

The construction process of the optimal planning model for rural tourism routes is shown in Fig. 1.

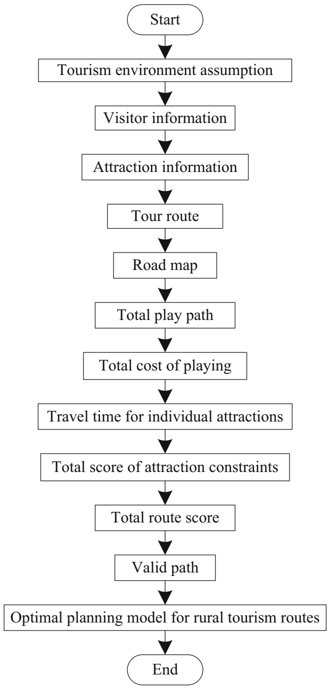

Fig. 1. Construction process of the optimal planning model for rural tourism routes

2.4 Solution to the Optimal Planning Model of Rural Tourism Routes

There are many methods to solve the multi constrained multi-objective programming problem, and the evaluation function method is one of them. Its purpose is to transform the multi-objective optimization into the single objective optimization and then solve it [13]. By using the evaluation function $u[f(x)]$, a multi constraint multi-objective planning problem - the optimal planning model for rural tourism routes - is transformed into the following problem:

$$(Pu)\begin{cases} \min\ u[f(x)] \\ s.t.\quad x \in X \end{cases} \tag{7}$$

In Eq. (7), (Pu) represents the auxiliary calculation function; $f(x)$ represents the optimal planning objective for rural tourism routes; X represents the optimal planning target set for rural tourism routes.

The key of the geometric weighting method is to create a new evaluation function to simplify the multi-objective optimization problem and then solve it [14, 15]. The specific method of linear weighting method is to define different weight coefficients υ_i for each optimization objective $f_i(x)$ based on the different importance of each optimization objective in the minds of decision-makers, and then convert multi-objective optimization into single objective optimization, with the expression:

$$u[f(x)] = \sum_{i=1}^{m} \upsilon_i f_i(x) \tag{8}$$

Then the solution steps of the optimal planning model of rural tourism routes are as follows:

(1) Unify dimensions and standardize the objective functions $f_i(x)$. Then calculate the two maximum values of each optimization objective function: $f_{i-\max}(x)$ and $f_{i-\min}(x)$. Specific application functions:

$$\psi_i(x) = \frac{f_i(x) - f_{i-\min}(x)}{f_{i-\max}(x) - f_{i-\min}(x)} \tag{9}$$

In Eq. (9), $\psi_i(x)$ represents the normalization result of the objective function.

(2) Based on the different importance levels of different optimization objectives in the minds of decision-makers, the weight coefficient υ_i that matches the optimization objective function $\psi_i(x)$ is given, and the evaluation function is represented as

$$u[f(x)] = \prod_{i=1}^{m} \psi_i(x)^{\upsilon_i} \tag{10}$$

Solve single objective optimization problems. By solving the optimal solution x^* of the single objective optimization problem, the non inferior solution of the multi-objective optimization problem before transformation is obtained. The two objectives of the multi constraint multi-objective programming problem are: the maximum number of tourist attractions and the highest overall score of tourist attractions. The weight coefficients representing their importance will also change with the individual needs and preferences of tourists. Therefore, choosing the geometric weighting method to solve the multi-objective optimization problem of scenic spot screening has to some extent improved the reliability of the final route in the practical application of the model.

Through the above process, the multi constraint and multi-objective planning problem is solved, and the optimal planning results of rural tourism routes are obtained, providing better rural tourism experience for tourists.

3 Experiment and Result Analysis

3.1 Selection of Experimental Objects

This paper selects a town and village as the research area of short-term tourism route planning. Its good geographical location and unique cultural history have created a wealth of tourism resources. Natural ecology, rural ancient towns, characteristic blocks, exhibition halls and other tourism resources are concentrated, with good geographical combination conditions, which enables tourists to experience more types and combinations of scenic spots in a short time.

The distribution of scenic spots of experimental subjects is shown in Fig. 2.

Fig. 2. Distribution of scenic spots of experimental subjects

As shown in Fig. 2, the number of scenic spots in the rural area, the experimental object, is large, but the distribution is not obvious, which meets the application performance test requirements of the proposed method.

3.2 Experimental Data Acquisition and Pre-processing

For the collection of tourist attraction data, tourist route data and travel time data, use Python's urllib library to obtain relevant information from the corresponding platform. Urllib is Python's built-in HTTP request library, which provides a way to obtain web page data from a specified URL address, and then parse it to obtain the desired data.It includes the request module, the exception handling module error, the URL parsing module parse, and the robots.txt parsing module robotparser. The methods of parsing data include direct processing, JSON parsing, regular expression processing, Beautiful Soup parsing, PyQuery parsing, and XPath parsing. In this study, we mainly select direct processing, JSON parsing and regular expression processing for the acquired data, and finally select relational data Oracle to store the collected data in a unified way. In addition to the tourist route data, the construction of the scenic spot association map also needs the taxi track data. In order to make the data meet the requirements of the subsequent algorithm, the original taxi GPS data needs to be processed, including data cleaning, starting and ending point extraction, and grid processing.

Due to the limitation of research space, only the traffic time data processing process is described in detail, as shown below:

(1) Request Web Service API key.

Create a new application in the Gaode Map Console, click the "Add New Key" button on the created application, enter the application name in the pop-up dialog box, select the binding service as "Web Service API", and click the Submit button to obtain the key.

(2) Splice HTTP request URL

Fill in the required request parameters according to the URLs of different APIs. The key applied in the first step should be sent together as a required parameter.

(3) Receive the data returned from the HTTP request (in JSON format), and parse the JSON data using Python

The Walking Path Planning API can obtain data for walking commuting options up to 100 km. The URL is https://restapi.amap.com/v3/direction/walking?parameters. Parameters indicates the request parameter to be entered.

Through the above process, the collection and processing of experimental data are completed, providing support for the smooth progress of subsequent experiments.

3.3 Analysis of Experimental Results

Select the time-based route planning problem model as the comparison method 1, and the tourism route recommendation model based on pattern and preference perception as the comparison method 2, and design the comparative experiment of rural tourism route optimal planning. The application effect of the proposed method is shown by the results of the optimal planning of rural tourism routes and the number of solving iterations.

In order to clearly show the optimal planning results of rural tourism routes, 10 rural tourist attractions are selected as tourist destinations, and the optimal planning results of rural tourism routes are shown in Fig. 3.

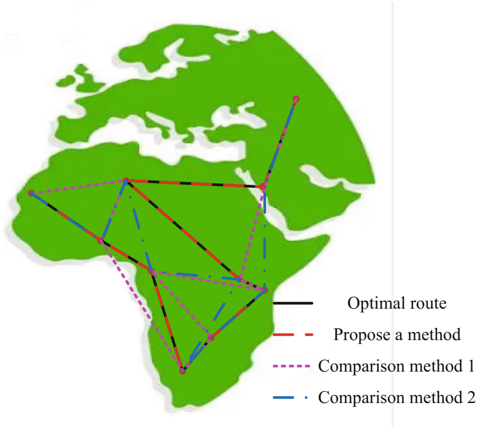

Fig. 3. Schematic diagram of optimal planning results of rural tourism routes

As shown in Fig. 3, the proposed Location Topic model exhibits higher accuracy in the optimal planning of rural tourism routes. Compared with comparison methods 1 and 2, the optimal planning results of rural tourism routes obtained by this method are consistent with the actual optimal routes. There is a significant difference between the optimal planning results of rural tourism routes obtained by comparing methods 1 and 2 and the actual optimal routes.

This indicates that the method proposed in this paper can more accurately capture the correlation between geographical location and themes, thereby providing more reliable and accurate results in rural tourism route planning. By combining geographic location information and thematic information, our model can better understand the characteristics and thematic needs of different regions, and consider these factors into the planning process of the optimal route.

This high-precision rural tourism route planning result is of great significance to tourists. It can help them better arrange their itinerary, choose the best travel route, and obtain a richer and more satisfying travel experience. At the same time, it also has a positive impact on the development and promotion of rural tourism. Accurate route planning can attract more tourists, enhance the attractiveness and competitiveness of rural tourism, and promote the sustainable development of rural tourism.

The number of iterations for solving the optimal planning results of rural tourism routes obtained through experiments is shown in Table 1.

Table 1. Iterations of Solving the Optimal Planning Results of Rural Tourism Routes

Working condition	Propose method	Comparison method 1	Comparison method 2
1	12	20	26
2	10	25	35
3	9	27	34
4	8	30	36
5	15	19	28
6	14	20	27
7	13	24	25
8	9	21	21
9	10	23	20
10	11	18	24

As shown in the data in Table 1, under different experimental conditions, the number of iterations for solving the optimal planning results of rural tourism routes obtained after the application of the proposed method is less than that of the comparison methods 1 and 2.

Specifically, the comparison methods 1 and 2 generally have higher iteration times for solving under each experimental condition, while the proposed method has relatively fewer iteration times for solving. This indicates that the proposed method has higher efficiency in solving the optimal planning of rural tourism routes.

Reducing the number of solving iterations means that our method can find the optimal solution more quickly, thereby saving computational resources and time costs. This is very important for practical applications, especially in large-scale rural tourism route planning.

Therefore, based on the data analysis results in Table 1, we can conclude that the proposed method has significant advantages in the efficiency of optimal planning of rural tourism routes, and can solve the optimal solution more quickly, providing users with more efficient and convenient rural tourism planning services.

4 Conclusion

The vigorous development of the economy has brought progress to the tourism industry. The rapid popularization of internet technology has led to the emergence of massive amounts of tourism information on the internet, and the methods for obtaining tourism information have also become diverse. Currently, the tourism information on the internet has reached a point where people cannot browse it one by one. In this situation, when tourists want to travel to a strange Urban tourism, it is difficult to determine their own travel routes according to the travel information on the Internet. Therefore, helping tourists plan a travel route that meets their basic preferences and improving their decision-making efficiency is of great significance.

Based on this situation, this article studies the tourism route planning problem based on multiple constraints and objectives, establishes a comprehensive rating model for scenic spots and a tourism route planning model, and plans tourism routes for tourists. The experimental data shows that the proposed method greatly improves the accuracy and efficiency of optimal planning for rural tourism routes, providing better route planning services for tourists.

In terms of future prospects, with the vigorous development of the economy and the progress of the tourism industry, the following trends can be foreseen:

Firstly, with the continuous development of intelligent technology, future tourism planning will become more intelligent and personalized. By combining artificial intelligence and Big data technology, the intelligent system can provide each user with customized travel route planning according to their preferences, interests and needs to meet their personalized travel needs.

Secondly, combining real-time data and user feedback will become an important direction for future tourism planning. By obtaining real-time information on transportation, weather, and scenic spot congestion, combined with user feedback and evaluations, tourism routes can be adjusted and optimized in a timely manner, providing more accurate and practical planning services.

In addition, cross-border cooperation and resource integration are also one of the trends in future tourism planning. Collaborating with transportation departments, hotel and scenic area management departments, and sharing data and resources can provide more comprehensive and efficient tourism planning services.

Finally, sustainable development and ecological protection will become important considerations for future tourism planning. Pay attention to protecting the natural environment and cultural heritage, promote the development of Sustainable tourism, and ensure the sustainable impact of tourism on society, economy and environment.

In summary, future tourism planning will be more intelligent and personalized, combining real-time data and user feedback, cross-border cooperation and resource integration, focusing on sustainable development and ecological protection. The method proposed in this article provides useful exploration and foundation for the development of future tourism planning.

References

1. Su, J., Li, J.: Decentralized intelligent search of tourist routes based on check-in data. Int. J. Mobile Comput. Multimedia Commun. **12**(3), 16–21 (2021)
2. Castillo, S.A., Hornillos, N.B., Val, P.A., et al.: Machine learning to predict recommendation by tourists in a Spanish province. Int. J. Inf. Technol. Decis. Mak. **21**(4), 1297–1320 (2022)
3. Zyrianov, A.I., Zyrianova, I.S.: Planning of the interregional tourist route in the Urals. Quaestiones Geographicae **40**(2), 109–118 (2021)
4. Chen, C., Zhang, S., Yu, Q., et al.: Personalized travel route recommendation algorithm based on improved genetic algorithm. J. Intell. Fuzzy Syst. Appl. Eng. Technol. **40**(3), 4407–4423 (2021)
5. Han, S., Liu, C., Chen, K., et al.: A tourist attraction recommendation model fusing spatial, temporal, and visual embeddings for flickr-geotagged photos. Int. J. Geo-Inf. **10**(1), 20–26 (2021)
6. Ishizaki, Y., Koyama, Y., Takayama, T., et al.: A route recommendation method considering individual user's preferences by Monte-Carlo tree search and its evaluations. J. Inf. Process. **29**(11), 81–92 (2021)
7. Zhang, Q., Liu, Y., Liu, L., et al.: Location identification and personalized recommendation of tourist attractions based on image processing. Traitement du Signal: Signal Image Parole **38**(1), 197–205 (2021)
8. Liu, X., Jan, N.: Personalized recommendation algorithm of tourist attractions based on transfer learning. Math. Probl. Eng. **25**(15), 1–7 (2022)
9. Guowei, L.: Optimization method of tourist route in peak period based on ant colony algorithm. Comput. Simul. **37**(12), 353–357 (2020)
10. Zhang, Y., Tang, Z.: Cross-modal travel route recommendation algorithm based on internet of things awareness. J. Sens. **56**(10), 981–989 (2021)
11. Li, X.F.: Personalized planning method of rural tourism route based on landscape gene. J. Hebei North Univ.: Nat. Sci. Ed. **37**(07), 47–51 (2021)
12. Li, C.J.: Design of ecotourism personalized route planning system based on weather forecast data. Mod. Electron. Tech. **44**(04), 103–106 (2021)
13. Luo, Y.: SWOT analysis of new trend of rural tourism development in Xinjiang. Asian Agric. Res. **14**(6), 51–54 (2022)
14. Couto, G., Pimentel, P.M., Batista, M.D.G., et al.: The potential of rural tourism development in the Azores islands from the perspective of public administration and decision-makers. WSEAS Trans. Environ. Dev. **17**(5), 713–721 (2021)
15. Uri, N., Svitlica, A.M., Brankov, J., et al.: The role of rural tourism in strengthening the sustainability of rural areas: the case of Zlakusa village. Sustainability **13**(12), 1–23 (2021)