

Research Article

Personalized Travel Recommendation Based on the Fusion of TGI and POI Algorithms

Lu Fan¹ and Wenliang Zhang² 

¹School of Henan College of Transportation, Zhengzhou, Henan 450000, China

²School of Zhengzhou University, Zhengzhou, Henan 450001, China

Correspondence should be addressed to Wenliang Zhang; zwl@zzu.edu.cn

Received 26 November 2021; Revised 20 December 2021; Accepted 27 December 2021; Published 24 January 2022

Academic Editor: Mohammad R Khosravi

Copyright © 2022 Lu Fan and Wenliang Zhang. 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.

On the way to travel, the public expect to get a tourism experience with low cost, convenient travel, and high comfort. At the same time, they also have different tourism needs such as history and culture, natural landscape, and food shopping. To address the problem that traditional travel route recommendation algorithms have limited accuracy and only analyze text or pictures alone, we propose a personalized travel route recommendation algorithm that integrates text and photo information from travelogues and obtain the historical tourism footprint of tourists by analyzing travel notes. According to the frequency and cooccurrence of scenic spots in the travel notes and the number of photos taken by each scenic spot, the popularity of scenic spots and the interest preferences of various types of tourists are analyzed. Under the given starting and ending points or passing points, the optimal tourism route generation method is designed. Experiments on the real data set of Ctrip Travel website show that the recommendation accuracy of this algorithm is significantly improved compared with the traditional algorithm which only uses travel notes text or photos. Compared with the algorithm that only considers the popularity of interest points or tourists' interest preferences, the accuracy of the route recommended by the algorithm is improved. Compared with the algorithm that only considers the cooccurrence of scenic spots or only considers the influence of photos, this algorithm can obtain a better popularity score of scenic spots. This method integrates the two kinds of information including picture and text, fully considering the interest of users with high practicability.

1. Introduction

When traveling to an unfamiliar city, users usually select points of interest (POI) first and then make a travel itinerary based on their interests and the time [1]. For example, when a person comes to a new place, he must be interested in visiting the most POIs in the shortest time. Therefore, the recommendation of tourist attraction is conducive to promoting the rapid development of tourism.

With the development of information technology, the Internet is becoming an important source for people to plan their travels [2]. In the field of tourism, various forms of tourism temporal and spatial trajectory data have been formed, such as GPS trajectory, BeiDou navigation information, and check-in records. These data and a large number of travel experience,

travel photos, and other data shared by users jointly form tourism big data [3].

Scientific travel route planning can not only help travelers formulate their travel routes according to their time and budget but also improve their travel experience [4]. Based on the problems encountered by current users in travel planning, tourism route planning came into being. To get a high-quality solution in travel planning, we need to consider many factors and establish corresponding evaluation models according to different standards [5, 6]. For example, Rahimi and Xin further extended the existing work by studying the periodicity of space and time in user check-in data and proposed two new recommendation algorithms [7]. Zhang et al. studied some representative topic model extraction methods based on the spatial and temporal

features of short text [8]. Bin et al. proposed a neural multi-context modeling framework (NMMF) combined with rich heterogeneous tourism information [9].

When using a single type of user generated content for tourism route planning, there is often great uncertainty, so it is difficult to ensure the accuracy of the user trajectory [10]. Therefore, the comprehensive use of various content to more accurately mine the user's historical trajectory has become the focus of current research [11]. Feng and Qian studied a new method to help users digest a large number of available opinions in an easy way [12]. Marrese-Taylor et al. used a multisource social media integration method to integrate fragmented tourism information from many aspects to recommend routes to users [13]. Hu et al. proposed a new framework called scenic planner for travel route recommendation, including scenic road network modeling and scenic spot route planning [14].

More and more travelers are posting their travelogues, including experiences, suggestions, and experiences, to online social media platforms [15]. Travelogues contain a large number of unique experiences and reliable travel advice from different travelers, providing an excellent reference for travel planners to select their routes [16]. As a result, some studies today make more accurate and personalized travel route recommendations by uncovering graphic information on travelogues (TGI) and finding out travelers' preferences, habits, and the popularity of attractions [17].

Lim et al. used the positioning information and shooting time in the picture to calculate the preference of tourists and the popularity of scenic spots [18]. Peng et al. recommended scenic spot areas according to the clustered areas by using social media pictures [19]. However, they only analyzed the picture information and ignored the text. Instead, the massive travel notes released by social media are full of pictures and texts. Coincidentally, Murphy and Banerjee paid more attention to the text information in travel notes [20]. Tai et al. and Lu et al. did not combine user behavior habits, interest preferences, and route popularity, and the personalization of route planning results was not high [21, 22]. In addition, based on privacy protection and other considerations, photos in travelogues often deliberately hide some attribute data [23]. This paper uses the number of pictures taken by tourists to calculate the tourist preference and scenic spot popularity, supplemented by the text description of the travel notes to make up for the lack of some attribute data. Therefore, the comprehensive use of picture and text information can often obtain more accurate recommendation results [24]. For example, Arain et al. extracted semantic information of tourist attractions and user preferences by using photos with geographical labels and user context information [25]. Huang developed a heuristic algorithm calculating context similarity, which can be used in photo data and GPS track [26].

Personalized tourism recommendation is more difficult to analyze and mine useful information from numerous tourism data. At present, many methods have many deficiencies in recommendation quality and speed. The POI+TGI model mainly considers the generated content of users and has significant advantages in preference extraction and fast route generation. This paper proposes a personal trip recommendation

based on interest and popularity (PTRIP) algorithm. The algorithm comprehensively considers the text description and photo data in online travel notes, uses the cooccurrence information of scenic spots and the number of photos taken by tourists for a certain type of POI to calculate the popularity of scenic spots and tourist preference for various scenic spots, respectively, and, combined with the time cost of tourists conversion between POIs, generates a tourism route transfer map for tourism route recommendation. Compared with the original method, this algorithm uses the text and picture information in the travel notes at the same time, comprehensively considers the two factors of tourist preference and scenic spot popularity, and effectively improves the accuracy of recommendation.

2. Basic Definition

This paper uses the orientation problem to return an optimal travel route for the user, considering the interest preference of users and the popularity of POI. The personalized recommendation method proposed in this paper incorporates two social factors: preferences of user in travelogue graphic information and interpersonal interest similarity [27]. Therefore, we first introduce the user interest factors. Then, we derive the objective function of the proposed personalized recommendation model. Finally, the training method of the model is given. In the following, we present our definitions and methods in detail. We take Wuhan, China, as the empirical research object of this article. This city has many types of scenic spots, such as natural scenery, urban prosperity, and historical heritage, which is very attractive.

Let a directed weighted POI transfer graph be $G = \langle V, E \rangle$, where V is the set of nodes and E is the set of edges. A node $p \in V$ represents a POI, and each p has the category attribute Cat_p (e.g., beach and castle), longitude, and latitude. The value on node p represents the score of POI_p , while C denotes the set of all POIs. (c_i, pop_i) is the attribute of node p_i , where c_i denotes the category and pop_i denotes the popularity. Each directed edge (p_i, p_j) represents a feasible route between two POIs, and the weights on the edges represent the travel time (in h) spent to visit the two POIs consecutively. An example of a POI transfer map is shown in Figure 1. Each p represents a scenic spot in Wuhan. The specific digital source and path analysis will be described in the next chapter.

Definition 1. Given a visitor t , its POI preferences of category c_i can be expressed as Equation (1).

$$\text{Int}(t) = \langle \text{Int}(t, c_1), \text{Int}(t, c_2), \dots, \text{Int}(t, c_{|C|}) \rangle. \quad (1)$$

Definition 2. The time from p_i to p_j is defined as Equation (2).

$$T^{\text{travel}}(p_i, p_j) = \frac{\text{Dist}(p_i, p_j)}{\text{speed}}. \quad (2)$$

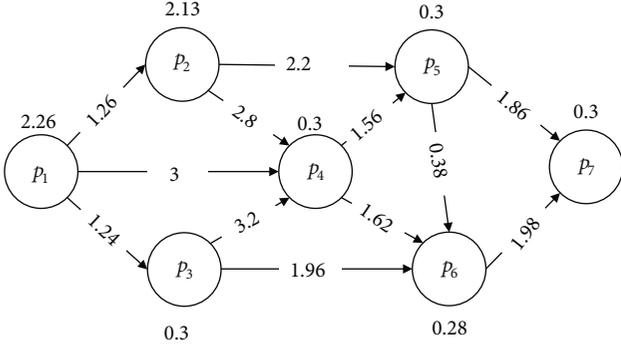


FIGURE 1: An example of POI transfer graph.

$\text{Dist}(p_i, p_j)$ denotes the distance between p_i to p_j . Its value is the actual distance recorded by Gaode Map. Generally, the bus and self-driving tours are greatly affected by traffic conditions, while the walking time of tourists is within the predictable range. Therefore, this paper uses the walking time between POIs as the travel time and takes speed = 5 km/h.

Definition 3. Given a user u and the set of POIs he/she has visited, S_u defines his/her historical travel footprint in chronological order, as in Equation (3).

$$S_u = \left((p_1, t_{p_1}^a, t_{p_1}^d), (p_2, t_{p_2}^a, t_{p_2}^d), \dots, (p_n, t_{p_n}^a, t_{p_n}^d) \right). \quad (3)$$

Each triplet $(p_x, t_{p_x}^a, t_{p_x}^d)$ consists of three elements: the arrival time $t_{p_x}^a$ at p_x and the departure time $t_{p_x}^d$ from p_x . The first photo taken by the user at each POI is the arrival time, and the last photo is the departure time. Thus, the visit time of u at p_x (i.e., the stay time) can be defined by the difference between $t_{p_x}^a$ and $t_{p_x}^d$. Similarly, for the travel sequence S_u , $t_{p_x}^a$ and $t_{p_x}^d$ represent the start and end time, respectively. Simplicity, we put S_u as $S_u = (p_1, \dots, p_n)$.

Definition 4. The POI scores in this paper are obtained by combining attraction popularity and visitor preferences by weighting. The score of POI_{p_i} for c_i is defined in Equation (4).

$$\text{score}(p_i) = \alpha \cdot \text{Int}(t, c_i) + (1 - \alpha) \text{pop}_i, \quad (4)$$

where α is a weight adjustment parameter to adjust the weight of preference and the popularity of POI in the tour route?

Definition 5. Given a tourist t , the total number of route attractions n and the set of POI scores, the visitor gain is defined as Equation (5).

$$\text{profit}(t) = \frac{\sum_{i=1}^n \text{score}(p_i)}{\sum_{i=1}^{n-1} \sum_{j=2}^n T^{\text{travel}}(p_i, p_j)}. \quad (5)$$

3. Recommendation Method

3.1. Method Framework. The travel route recommendation algorithm proposed in this paper is divided into data preprocessing, POI mode, association graph construction, and tourist interest preference learning and route recommendation. The POI transfer graph is constructed offline and learns interest preferences of tourists. The POI and the interest preferences of tourists are obtained by analyzing the cooccurrence information of attractions and photo data in the travelogue. The route recommendation is conducted online. Based on the personal information entered by tourists, the number of expected attractions and the designated tour points, the PTRIP algorithm is used to recommend the routes with the highest benefits to tourists, considering the POI popularity and tourists' preferences. The detailed framework is shown in Figure 2. The basis for tourist itinerary recommendations mainly comes from visitor information, design tour sites, budget number of attractions, and profit of route. The first three bases mainly refer to the subjective will of the referee; the last basis is the problem to be solved by the algorithm proposed in this paper. Route profit is calculated by popularity of POI and visitor preference, and the two indicators have their weights. Moreover, the POI mode is the most important part in the travel itinerary recommendation model.

3.2. Construct the POI Transfer Graph. The POI transfer graph is constructed offline. Treating all POIs as nodes on the way, the travel routes can be generated by visiting the directed edges in the graph consecutively.

(1) Map the photo

The web travelogues shared by tourists contain textual description information such as travel routes, travel feelings, and their photos taken at each attraction. The travelogue number and tourist number can be extracted from them. The structure of the photo data shared by the user conclude Photo ID, User ID, Time, Longitude, Latitude, and Category. Based on the longitude and latitude of each photo, the distance of each POI can be calculated by using Formula Haversine. If the result is less than 200 meters, it is assumed that the photo is taken at this POI. And the list of POI is $S_t = (p_1, p_2, \dots, p_n)$. Meanwhile, the time cost of inter-POI transition can be calculated in walking mode.

(2) The popularity of POI

The popularity of POI is calculated by combining the number of photos in the historical travelogues shared by visitors and the cooccurrence information of attractions by weighting, as in Equation (6).

$$\text{pop}(p) = \beta \cdot \frac{N(p)}{N_{\max}} + (1 - \beta) \cdot \frac{F(p)}{F_{\max}}. \quad (6)$$

In Equation (6), $N(p)$ is the number of photos taken by visitors to POI_p ; N_{\max} is the maximum number of photos taken by visitors to POI_p ; $F(p)$ is the number of times

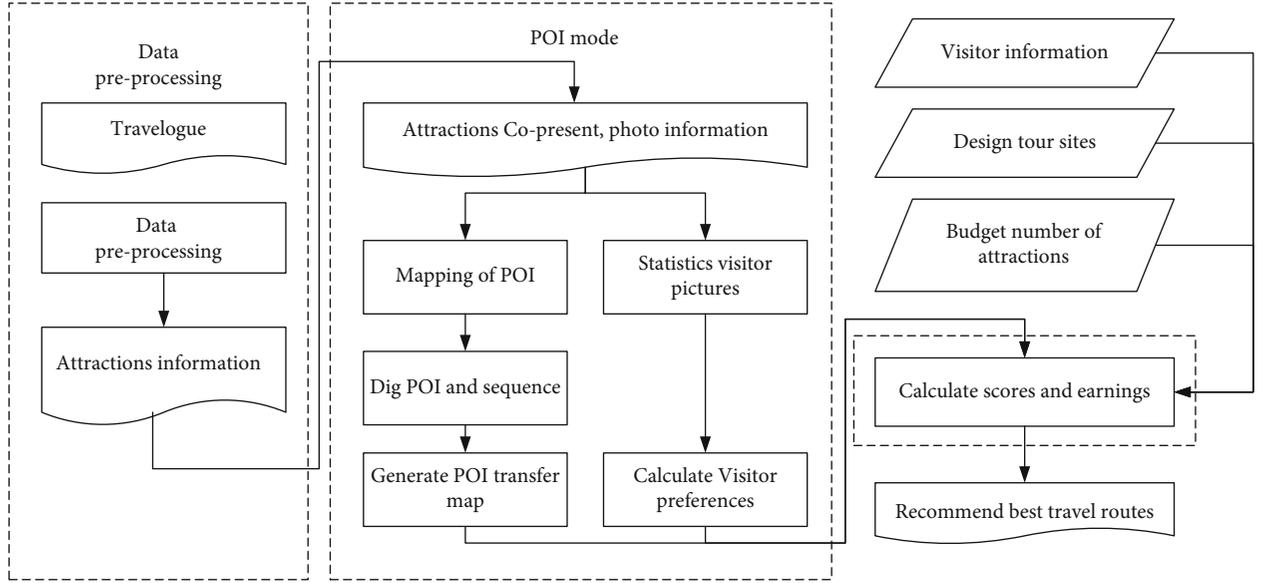


FIGURE 2: Framework for travel itinerary recommendations.

POI_p was mentioned in the travelogue; F_{\max} is the maximum number of times POI_p was mentioned in the travelogue.

3.3. Interest Preferences of Tourists. We propose a time-based user interest preference from the historical travel footprint of users. When one visits a POI, he stays there for a certain amount of time. From the historical travel footprints of all users, the visit time (i.e., stay time) of each user at each POI is calculated according to Definition 4, so that the average time required for any user to visit POI can be calculated. In this paper, $\bar{V}(p)$ is the average visit time at POI_p for any user, as Equation (7).

$$\bar{V}(p) = \frac{1}{n} \sum_{u \in U} \sum_{p_x \in S_x} (t_{p_x}^d - t_{p_x}^a) \sigma(p_x = p); \forall p \in P. \quad (7)$$

In Equation (7), U is all users, n is the number of users accessing p , and $\sigma(p_x = p) = \begin{cases} 1, & p_x = p \\ 0, & \text{else} \end{cases}$.

The average access time of a user at each POI does not truly reflect his interest preference for the POI. Therefore, we propose a time-based interest preference. The preference of user for the category attribute of POI is given by Equation (8):

$$\text{Int}(t) = \sum_{u \in U} \frac{t_{p_x}^d - t_{p_x}^a}{\bar{V}(p_x)} \sigma(\text{Cat}_{p_x} = c); \forall c \in C. \quad (8)$$

In Equation (6), Cat_p is the category attribute, and σ

$$\text{Cat}_{p_x} = c = \begin{cases} 1, & \text{cat}_{p_x} = c \\ 0, & \text{esle} \end{cases}.$$

Equation (8) determines the interest of user in category attribute of a particular POI. Relative to the average access

time of all users at the same POI, it is calculated based on the time cost by the user at each POI with category attribute. In other words, a user may spend more time visiting the POI that he is interested in, which in turn determines the interest of users in such POIs.

3.4. PTRIP Algorithm. Orienteering problem (OP) has already widely used in travel route recommendation. In a directed band power diagram $G(V, E)$, V is the set of all points on the graph, and E is the set of all edges on the graph. Each point has a corresponding score which can be expressed as a gain. And each edge has a corresponding weight which represents the travel time between two points. The start and end points are specified, and some points are selected from diagram G , and a path is planned through these points and the specified start and end points, while maximizing the score under the condition that the total weight of the path does not exceed a certain time budget.

In this paper, we propose the PTIR route recommendation algorithm based on the TGI and POI. PTIR can provide a route with the highest score and satisfied time budget, i.e., $R = \{p_1, p_2, \dots, p_N\}$. Time budget is calculated by function $\text{Cost}(p_x, p_y) = T(p_x, p_y) \bar{V}(p_y)$. From this, it follows that the travel route recommendation model in this paper can be expressed integer programming problem satisfying multiple constraints:

$$\text{Max} \sum_{i=2}^{N-1} \sum_{j=2}^N x_{i,j} \text{profit}(p_i). \quad (9)$$

In Equation (9), $x_{i,j} = 1$ indicates the route from i to j , i.e., (p_i, p_j) , or $x_{i,j} = 0$.

$$\sum_{j=1}^N x_{1,j} = \sum_{i=1}^{N-1} x_{i,N} = 1, \quad (10)$$

$$\sum_{j=2}^N x_{k,j} = \sum_{i=1}^{N-1} x_{i,k} \leq 1; \forall k = 2, 3, \dots, N-1, \quad (11)$$

$$\sum_{i=1}^{N-1} \sum_{j=2}^N \text{Cost}(i, j) x_{i,j} \leq B, \quad (12)$$

$$2 \leq u_i \leq N, \quad (13)$$

$$u_i - u_j + 1 \leq (N-1)(1 - x_{i,j}); \forall i, j = 2, 3, \dots, N. \quad (14)$$

Equation (9) is the objective function that maximizes the POI popularity and user interest preferences in the recommended route. Equations (10)–(14) are constraints. Equation (10) ensures that the starts at p_1 and ends at p_n ; Equation (11) ensures that the itinerary is coherent and that each POI in the itinerary is visited only once; Equation (12) ensures that the time spent on the trip is within budget; Equation (13) and Equation (14) ensure that there are no subcircuit routes in this integer programming problem. The lpsolve (BERKELAAR M, 2004) linear programming package is used to solve the proposed integer programming problem.

To explain the algorithm, we take Figure 1 for example and give set $P = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7\}$, as in Table 1. Given a tourist t_1 and the set of POIs, he has visited $S_{t_1} = \{p_1, p_2\}$. The POIs visited by tourist t_1 and the number of photos at each POI are shown in Table 2. The average number of photos at each POI calculated from Equation (5) based on historical tourist data is shown in Table 3, and the popularity of each POI calculated from Equation (4) is shown in Table 4. The time cost required for a visitor t_1 to transfer between POIs and the rating value of each POI are represented by the values of the directed edges and the values of the nodes in Figure 1.

We assume that p_4 is a mandatory site for tourists to visit and plan to go to 4 attractions. The amount of interest preference can be calculated by Equation (8). The result is $\text{Int}(t_1) = \langle 2.8, 2.67, 0, 0, 0 \rangle$, and from this, we can get 13 routes based on PTRIP algorithm, as shown in Table 5.

And the benefits of seven routes are calculated by Equation (5). The results are the following: $\text{pro}(R_1) = 0.89$, $\text{pro}(R_2) = 0.86$, $\text{pro}(R_3) = 0.53$, $\text{pro}(R_4) = 0.52$, $\text{pro}(R_5) = 0.64$, $\text{pro}(R_6) = 0.49$, $\text{pro}(R_7) = 0.48$, $\text{pro}(R_8) = 0.64$, $\text{pro}(R_9) = 0.49$, $\text{pro}(R_{10}) = 0.70$, $\text{pro}(R_{11}) = 0.19$, $\text{pro}(R_{12}) = 0.18$, $\text{pro}(R_{13}) = 0.17$. Finally, the R_1 route that has the highest profit may be recommended to visitor t_1 , i.e., Yellow Crane Tower, Riverbank Park, Wuhan University, and Tumultuan Lin. Combined with the reality of tourism websites, this route is adopted more frequently, which preliminarily proves the effectiveness of the algorithm.

4. Experimental Results and Analysis

4.1. Experimental Data. In this paper, we use 2,638 travelogues obtained from the Ctrip Travel website with “Wuhan” as the keyword as the experimental dataset. After data preprocessing, the dataset contains two main aspects: 116,396 photos of Wuhan and its surrounding areas, including the

TABLE 1: The properties of POI.

POI	Category properties
Yellow Crane Tower (p_1)	Ancient ruins and buildings (c_1)
Riverbank Park (p_2)	Natural scenery (c_2)
Guiyuan Buddhist Temple (p_3)	Folk religion (c_3)
Wuhan University (p_4)	Humanities, academia (c_4)
Tumultuan Lin (p_5)	Bussiness, street(c_5)
Hubu Lane (p_6)	Ancient ruins and buildings (c_1)
East Lake (p_7)	Natural scenery (c_2)

TABLE 2: The visited POI of t_1 .

POI	Number of photos
P1	35
P2	20

TABLE 3: Average photos of POI.

POI	Average photos
p_1	12.5
p_2	7.5
p_3	4.5
p_4	13.8
p_5	6.3
p_6	11.7
p_7	8.2

TABLE 4: The popularity of POI.

POI	pop(p)
p_1	1.00
p_2	0.87
p_3	1.00
p_4	1.00
p_5	1.00
p_6	0.93
p_7	1.00

number of the travelogue to which the photos belong and the location where they were taken, and the cooccurrence statistics of 168 POIs in the travelogue, as well as 5,238 historical single-day travel routes and the actual distance information between the connected POIs in the routes. This experiment uses the leave-one-out cross-validation method commonly used in recommendation system validation to experimentally validate the algorithm, which loops the records in the specified dataset as the test set or training set, respectively, and calculates the predicted conclusions of each loop in a comprehensive manner to derive the measurement index.

TABLE 5: Initial recommended route of Wuhan.

Route	POI passed
R_1	$\{P_1, P_2, P_4, P_5\}$
R_2	$\{P_1, P_2, P_4, P_6\}$
R_3	$\{P_1, P_3, P_4, P_5\}$
R_4	$\{P_1, P_3, P_4, P_6\}$
R_5	$\{P_1, P_4, P_5, P_6\}$
R_6	$\{P_1, P_4, P_5, P_7\}$
R_7	$\{P_1, P_4, P_6, P_7\}$
R_8	$\{P_2, P_4, P_5, P_6\}$
R_9	$\{P_2, P_4, P_5, P_7\}$
R_{10}	$\{P_2, P_4, P_6, P_7\}$
R_{11}	$\{P_3, P_4, P_5, P_6\}$
R_{12}	$\{P_3, P_4, P_5, P_7\}$
R_{13}	$\{P_3, P_4, P_6, P_7\}$

4.2. *Algorithm Accuracy Analysis.* The accuracy of recommendation is the most basic metric for evaluating algorithms. In this paper, the precision and recall are used as the criteria to measure the algorithm performance. The calculation formulas are as Equations (15) and (16).

$$\text{precision} = \frac{|P_r \cap P_v|}{|P_r|}, \quad (15)$$

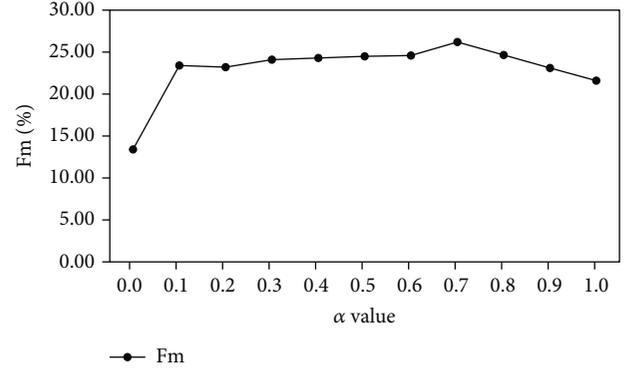
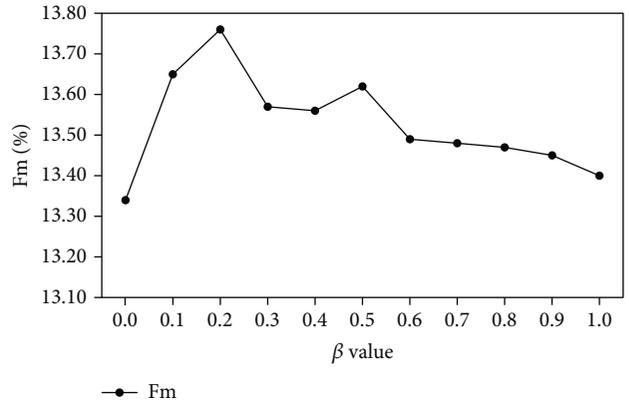
$$\text{recall} = \frac{|P_r \cap P_v|}{|P_v|}. \quad (16)$$

Precision represents the probability that a user is interested in the recommended route, and the recall represents the probability that one preferred POI is recommended, and the higher the precision and recall, the better the recommendation. P_r represents the set of POIs in the recommended route, and P_v represents the set of POIs in the real travel sequence that the user has visited. To better verify the recommendation quality of the algorithm in this paper, the metrics F_m are introduced, as Equation (17).

$$F_m = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (17)$$

The value of α is used to determine the weight assignment of visitor interest preferences and POI popularity when calculating route profit. For a given number of tour points and a specified number of tour points, the effect of α on the recommendation accuracy is shown in Figure 3.

When $\alpha = 0$, the weight of $\text{pop}(p)$ is 1, and the recommendation of visitor route is just based on the popularity of POIs. When $\alpha = 1$, the weight of $\text{Int}(t, c_i)$ is 1, and the recommendation of visitor route is just based on the interest preferences of tourists. As shown in Figure 4, the accuracy of route recommendation tends to increase and then decrease with the increase of α . It shows that the recommendation

FIGURE 3: The effect of α on the recommendation accuracy.FIGURE 4: The effect of β on the recommendation accuracy.

effect of considering popularity of POI and visitor preference is better than that of considering only one of them, and the best recommendation result is achieved when $\alpha = 0.7$.

The value of β is used to determine the weight distribution between the number of photos and the cooccurrence information of attractions when calculating POI popularity. In the case of route recommendation using POI popularity only, the effect of β value on the accuracy of route recommendation is shown in Figure 4, given the number of attractions visited and the specified tour points.

When $\beta = 0$, the weight of $\text{Int}(t)$ is 1, calculation of POI popularity based only on the cooccurrence data of attractions in travelogues. When $\alpha = 1$, the weight of $\text{Int}(t, c_i)$ is 1, calculation of POI popularity based only on the number of photos in the historical travelogues shared. As shown in Figure 4, the accuracy of route recommendation fluctuates with the variation of β . The lowest values at both ends of the curve, which means the popularity of POI calculated by considering the cooccurrence of attractions and the number of photos in the travelogue is more accurate. And the best recommendation result is achieved when $\beta = 0.2$.

Comparing the analysis of Figures 3 and 4, it is found that β has a small effect on the accuracy, fluctuating in the range of 1%, while the change of α makes the accuracy fluctuate in the range of 10%. From this, we can find that the calculation of POI popularity is related to the number of photos and photo cooccurrence information, but the weight

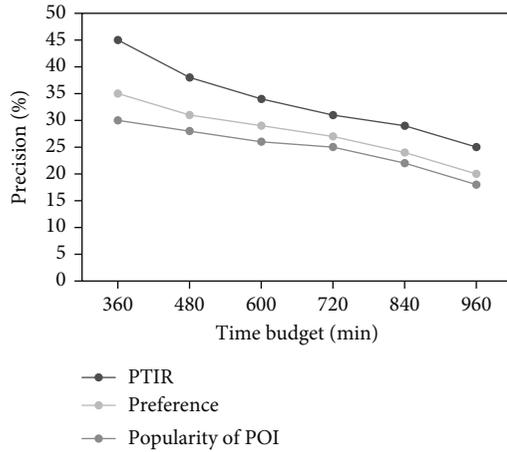


FIGURE 5: The experimental result of precision.

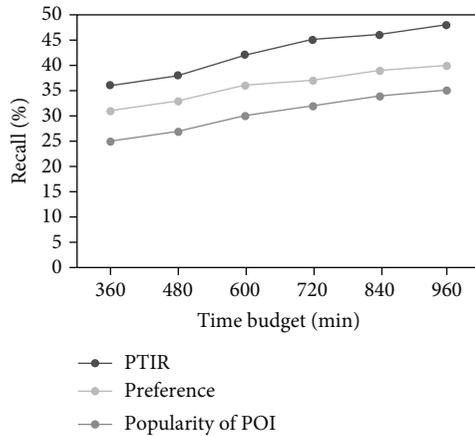


FIGURE 6: The experimental result of precision.

between them is not very important. The focus of the recommended route is on the subjective will of the tourists, so personalized customization is the future development direction of customized travel routes.

To verify the effectiveness of the PTIR algorithm, this paper uses the traditional travel route recommendation algorithm as a control, in which the recommendation algorithm considering only POI popularity and the recommendation algorithm considering only user interest preference are used as the metric, respectively. Under different time budgets B , the traditional algorithm is compared with PTIR, a travel route recommendation algorithm based on POI popularity and user interest preferences, and the experimental results are shown in Figures 5 and 6.

Figure 5 shows the difference in precision between the PTIR algorithm and the traditional algorithm. The precision of PTIR algorithm is much higher than algorithms that only consider user interest or only consider POI popularity. Figure 6 shows the difference in recall between the PTIR algorithm and the traditional algorithm. The recall of PTIR algorithm has the same situation with recall rate indicator. Among them, the accuracy and recall accuracy of the algorithm considering user interest are higher than the POI pop-

ularity only. One of the influencing factors is that both the algorithm PTIR and the algorithm that considers only the user's interest consider the user's interest because users prefer to visit places that interest them. The high accuracy and high recall of algorithm PTIR indicate that the algorithm proposed in this paper can recommend routes that reflect real travel sequences of users more accurately. It shows that when recommending travel itineraries to tourists, they should be guided by interests of users.

Overall, the precision increases with the increase of time budget, while the recall rate is the opposite. There are large uncertainties in the process of personalized tourism recommendation. In addition to the popularity of POI and preference of visitor, it is also necessary to consider time budget of them. By controlling the time budget and comprehensively considering the accuracy and recall of the algorithm, the time budget point corresponding to the best experimental result can be found. Then, while planning the tourist route for tourists, it is suggested to travel time.

5. Conclusion

To improve the accuracy of travel route recommendation and make comprehensive use of the graphic information in travel notes, this paper proposes a personalized route recommendation algorithm PTRIP.

Firstly, the algorithm uses the scenic spot cooccurrence information and photo data shared in the online travel notes to calculate the POI popularity and tourists' interest preferences and then comprehensively uses the above information to construct a personalized travel route recommendation framework to recommend the optimal travel route to tourists.

Finally, the experimental verification is carried out by using the real data set shared on the Ctrip Travel website. It is proved that the recommendation accuracy of the PTRIP algorithm proposed in this paper is significantly higher than that of the traditional recommendation algorithm which only uses the cooccurrence information of text scenic spots and also higher than that of the original algorithm which only uses the photo information of tourists.

The accuracy of PTRIP algorithm is much higher than the traditional algorithm considering only the popularity of POI. It is also better than the traditional algorithm that only considers tourists' preferences. Moreover, the comprehensive use of graphic information in travel notes can maximize the use of the information recorded in travel notes on the one hand and make up for the incomplete basic attributes of tourist photos caused by privacy and other reasons [28]. The POI popularity score quality calculated by PTRIP algorithm is also higher than the traditional algorithm considering text or picture alone. Experiments show that PTRIP algorithm can effectively make comprehensive use of the graphic information of Travel Notes published in social media to make more accurate personalized travel route recommendation.

The proposal of global tourism, smart tourism, and other strategies and the proliferation of user shared content have not only brought opportunities but also greater challenges

to tourism route planning. The planning method based on user generated content is not perfect. In real life, tourists may have multiple tourism needs to be optimized at the same time. How to efficiently solve the multiobjective optimization problem of personalized tourism route recommendation will be the next research direction.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interests.

References

- [1] R. A. Abbaspour and F. Samadzadegan, "Time-dependent personal tour planning and scheduling in metropolises," *Expert Systems with Applications*, vol. 38, no. 10, pp. 12439–12452, 2011.
- [2] T. Majeed, A. Stämpfli, A. Liebrich, and R. Meier, "What is of interest for tourists in an alpine destination: personalized recommendations for daily activities based on view data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 11, pp. 4545–4556, 2020.
- [3] Z. Ning, X. Hu, Z. Chen et al., "A cooperative quality-aware service access system for social Internet of Vehicles," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2506–2517, 2018.
- [4] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: a survey," *Geo Informatica*, vol. 19, no. 3, pp. 525–565, 2015.
- [5] H. Sun, Z. Zheng, J. Chen, and M. R. Lyu, "Personalized web service recommendation via normal recovery collaborative filtering," *IEEE Transactions on Services Computing*, vol. 6, no. 4, pp. 573–579, 2013.
- [6] D. Gavalas, V. Kasapakis, C. Konstantopoulos, G. Pantziou, N. Vathis, and C. Zaroliagis, "A personalized multimodal tourist tour planner," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia*, pp. 73–80, Melbourne, Australia, 2014.
- [7] S. M. Rahimi and W. Xin, *Location Recommendation Based on Periodicity of Human Activities and Location Categories*, Springer, Berlin Heidelberg, 2013.
- [8] J. Zhang, X. Hu, Z. Ning et al., "Energy-latency tradeoff for energy-aware offloading in mobile edge computing networks," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2633–2645, 2018.
- [9] C. Bin, T. Gu, Z. Jia, G. Zhu, and C. Xiao, "A neural multi-context modeling framework for personalized attraction recommendation," *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 14951–14979, 2020.
- [10] S. Banerjee and A. Y. K. Chua, "In search of patterns among travellers' hotel ratings in TripAdvisor," *Tourism Management*, vol. 53, pp. 125–131, 2016.
- [11] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user Interest and social circle," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 7, pp. 1763–1777, 2014.
- [12] H. Feng and X. Qian, "Mining user-contributed photos for personalized product recommendation," *Neurocomputing*, vol. 129, pp. 409–420, 2014.
- [13] E. Marrese-Taylor J. D. Velásquez et al., "A novel deterministic approach for aspect-based opinion mining in tourism products reviews," *Expert Systems with Applications*, vol. 41, no. 17, pp. 7764–7775, 2014.
- [14] B. Hu, G.-P. Gao, L. L. He, X.-D. Cong, and J.-N. Zhao, "Bending and on-arm effects on a wearable antenna for 2.45 GHz body area network," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 378–381, 2016.
- [15] C. Chen, X. Chen, Z. Wang, Y. Wang, and D. Zhang, "Scenic-Planner: planning scenic travel routes leveraging heterogeneous user-generated digital footprints," *Frontiers of Computer Science*, vol. 11, no. 1, pp. 61–74, 2017.
- [16] A. Majid, L. Chen, G. Chen, H. T. Mirza, I. Hussain, and J. Woodward, "A context-aware personalized travel recommendation system based on geotagged social media data mining," *International Journal of Geographical Information Science*, vol. 27, no. 4, pp. 662–684, 2013.
- [17] A. S. Tewari and A. G. Barman, "Sequencing of items in personalized recommendations using multiple recommendation techniques," *Expert Systems with Application*, vol. 97, pp. 70–82, 2018.
- [18] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized tour recommendation based on user interests and points of interest visit durations," in *Proceedings of the 24th International Conference on Artificial Intelligence*, pp. 1778–1784, California, 2015.
- [19] H. Peng, B. Hu, Q. Shi et al., "Removal of ocular artifacts in EEG—an improved approach combining DWT and ANC for portable applications," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 600–607, 2013.
- [20] H. C. Murphy, M. M. Chen, and M. Cossutta, "An investigation of multiple devices and information sources used in the hotel booking process," *Tourism Management*, vol. 52, pp. 44–51, 2016.
- [21] C. H. Tai, D. N. Yang, L. T. Lin, and M.-S. Chen, "Recommending Personalized Scenic Itinerary with Geo-Tagged Photos," in *Proceedings of 2008 IEEE International Conference on Multimedia and Expo*, pp. 1209–1212, Hannover, Germany, 2008.
- [22] X. Lu, C. Wang, J. M. Yang, Y. Pang, and L. Zhang, "Photo 2Trip: generating travel routes from geo-tagged photos for trip planning," in *Proceedings of the 18th International Conference on Multimedia*, pp. 25–29, Firenze, Italy, 2010.
- [23] G. Cui, J. Luo, and X. Wang, "Personalized travel route recommendation using collaborative filtering based on GPS trajectories," *International Journal of Digital Earth*, vol. 11, no. 3, pp. 284–307, 2018.
- [24] X. Hu, J. Cheng, M. Zhou et al., "Emotion-aware cognitive system in multi-channel cognitive radio ad hoc networks," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 180–187, 2018.
- [25] Q. A. Arain, H. Memon, I. Memon et al., "Intelligent travel information platform based on location base services to predict user travel behavior from user-generated GPS traces," *International Journal of Computers and Applications*, vol. 39, no. 3, pp. 155–168, 2017.
- [26] H. Huang, "Context-aware location recommendation using geotagged photos in social media," *International Journal of Geo-Information*, vol. 5, no. 11, p. 195, 2016.

- [27] B. Zheng, H. Su, K. Zheng, and K. Zhou, "Landmark-based route recommendation with crowd intelligence," *Data Science and Engineering*, vol. 1, no. 2, pp. 86–100, 2016.
- [28] D. Gavalas, C. Konstantopoulos, K. Mastakas et al., "A survey on algorithmic approaches for solving tourist trip design problems," *Journal of Heuristics*, vol. 20, no. 3, pp. 291–328, 2014.