Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 6386119, 14 pages
https://doi.org/10.1155/2022/6386119

Research Article

Multi-Objective Optimal Travel Route Recommendation for Tourists by Improved Ant Colony Optimization Algorithm

Haodong Sun,[1] Yanyan Chen ,[1] Jianming Ma,[2] Yang Wang,[1] Xiaoming Liu,[1] and Jiachen Wang[3]

[1]*Beijing Key Lab of Traffic Engineering, College of Metropolitan Transportation, Beijing University of Technology, 100 Pingleyuan, Beijing 100124, Chaoyang District, China*
[2]*Texas Department of Transportation, Austin 78701, TX, USA*
[3]*Faculty of Information Technology, Beijing University of Technology, 100 Pingleyuan, Beijing 100124, Chaoyang District, China*

Correspondence should be addressed to Yanyan Chen; cdyan@bjut.edu.cn

Received 25 May 2022; Revised 30 September 2022; Accepted 14 October 2022; Published 28 October 2022

Academic Editor: Zhixiang Fang

With the swift development of tourism all around the world, it has become vital to improve the recommendation of useful travel information to tourists to assure their convenience and satisfaction. In this paper, we propose a novel multi-objective optimal travel route recommendation framework, which collects tourists' travel trajectories from their mobile phone signaling data. Then, the proposed framework preprocesses the mobile signaling data to transform raw trajectories into tourists' travel sequences. Subsequently, the framework finds the popular attractions and frequent travel routes from the travel pattern sequences by using a frequent pattern mining method. Finally, an improved ant colony optimization (ACO) algorithm with a novel extensible heuristic factor approach is adopted to search the multi-objective optimal travel routes according to the popularity of attractions and travel time of tourists. The experimental results indicate that the proposed framework is efficient in recommending multi-objective optimal travel routes considering tourists' travel time and attractions' popularity while ensuring that the recommended travel route is suitable.

1. Introduction

With the swift development of transportation facilities, especially high-speed rail, tourism travel has gradually become the most important leisure behavior for people during holidays, just like commuting travel is the most important behavior for people on working days. Tourism is a huge industry, which provides various auxiliary services and plays an important part in national economic development and even supports the entire economic development of some countries [1]. Despite the rapid development rate, the tourism industry still has several challenges to deal with [2]. Recently, with the vigorous advance of tourism, the experience of tourists has undergone tremendous changes [1]. Gradually it has become a very important issue on how to make tourists' experience more relaxing and healthier and to improve their satisfaction and happiness.

With the development of location acquisition technology, GPS, the cellular networks, social networks, and location-based services can acquire large amounts of spatiotemporal data in the shape of locus [3–5]. The increasing locus data allow us to explore information that is meaningful for human mobility research and provide a great opportunity to solve many challenging problems such as travel recommendations. Generally, for an unfamiliar tourism destination city, when tourists make travel plans, the extremely concerning issues are where to tour, what to tour, and how to tour [6]. Among them, "how to tour" is the most considerable role of these three crucial issues, that is, how latent tourists plan a satisfactory travel route. Therefore, some leading data mining technologies can help people automatically extract useful information from a large number of trajectory data, thereby aiding tourists make more convenient and economical touring plans [7].

When considering travel recommendations, two main challenges must be noted [8]. First, travel recommendations should be personalized based on users' interests because different users have different preferences for various attractions [9]. Secondly, the recommendation of the travel itinerary is much more complicated than the recommendation of a single attraction since the distance or opening and closing times between the locations of different recommended attractions need to be considered [10]. When there are more attractions in an itinerary, the itinerary will be more elusive and not easy for tourists to understand. Existing travel recommendation methods can be separated into two types according to the services they provide, namely, recommendations based on destinations and recommendations based on routes. The core of the former method is to recommend a single attraction that suits the tourist's interests. Personal tour, for example, was adopted by travel services to aid tourists select the best tour packages according to their hobbies [11]. The latter method places emphasis on helping tourists draw itineraries that include multiple destination attractions. Travel route planning is a troublesome job and always requires a scheme that can offer suitable travel routes based on tourists' preferences [12]. For example, Photo2Trip [13] used geotagged photos to recommend user-defined travel routes based on tourists' preferences and allowed tourists to enter individual interests interactively.

For intricate dynamic environments, multiple factors and multiple objectives should be considered to make an optimal travel route for tourists. However, classical methodologies contain single-objective programming model for making a route of shortest distance [14]. These methodologies are mainly according to heuristic algorithms, such as A^* [15–17], genetic algorithm (GA) [18], simulated annealing (SA) [19], particle swarm optimization (PSO) [20], and ACO [21]. ACO is a famous and popular pathfinding algorithm, which is used to solve the shortest path problem based on graph [22]. Many researchers have applied the ACOs on solving gtraveling salesman problems, scheduling problems [23] and vehicle routing planning problems [24].

In this work, we select the popular attractions by considering the two factors of travel time and attractions' popularity and adopt an improved ACO with a novel extensible heuristic factor approach to find the multi-objective optimal route of shortest distance and highest frequency. In order to improve the search performance of improved ACO and also resolve the problem more efficiently, the devised extensible heuristic factor approach merges three search features, including distance coefficient, frequent pattern coefficient, and damping coefficient. The distance coefficient reflects the distance between two attractions, the frequent pattern coefficient reflects the relationship strength between two attractions, which is obtained through the mobile signaling data mining of tourists, and the damping coefficient reflects that the heuristic factor gradually decreases with the number of iterations. The contribution of this work consists of three parts:

(i) Firstly, we propose a novel multi-objective optimal travel route recommendation framework, which collects tourists' travel trajectories from mobile signaling data.

(ii) Secondly, we employ a frequent pattern mining method to find popular attractions and frequent travel routes from the travel pattern sequences transformed through tourists' raw trajectories.

(iii) Thirdly, we adopt an improved ACO algorithm with a novel extensible heuristic factor to find the multi-objective optimal route of shortest distance and highest frequency according to the popularity of attractions and travel time of tourists.

The remainder of this paper is structured as follows. In Section 2, we present the related research. In Section 3, we sketch out the research framework and propose the methodology. Section 4 shows the experimental results. Section 5 concludes the paper.

2. Related Research

This section provides an overview of some recent travel recommendation research studies and points out the differences between this research study and existing research. Most existing research focused on two aspects: popular attraction recommendation and travel route recommendation.

2.1. Popular Attraction Recommendation. In previous studies of popular attractions and their characteristics, most data source types were according to user-generated text content and geotagged photos on social media, such as online blogs and Flickr data [6]. For example, Zhang et al. [25] proposed a new hybrid travel note discovering model, which used collaboration awareness and trust awareness to generate travel notes that were most suitable for tourists. Hao et al. [26] proposed a probabilistic topic model for finding topics from travel notes, representing locations with suitable topics, recommending destinations for flexible queries, and summarizing the characteristics for a given destination. Ye et al. [27] implemented location association mining from the complementary aspects of association classification and association sorting, comprehensively understood the locations, and adopted the text and geographical characteristics of the locations to develop a joint training model for improvement of classification performance. Xu et al. [28] proposed a new approach to sum up popular information from substantial tourism blog data, investigated a word vector subdivision method, introduced the maximum confidence measure, and collected local features for each popular location to identify the attractions of interest. Yuan et al. [7] implemented the frequent pattern mining method on massive travel blog data to find the city's popular places and introduced a maximum confidence measure to identify popular travel routes helping tourists summarize the tour information and make better travel scheduling for an unfamiliar destination city. Wei et al. [29] developed a pattern-aware trajectory search framework to discover the top K paths through popular attractions, which determines a group of popular attractions and their attractiveness scores by considering not only the popularity of popular scenic spots but also the travel sequential relationships among popular attractions. Hu et al. [30] presented an united framework for abstracting and comprehending unban attractions using the DBSCAN clustering algorithm on Flickr photo data from different cities and discussed the spatiotemporal dynamics and some insights concluded from attractions. Pla-Sacristán et al. [31] used the K + V-DBSCAN clustering algorithm based on Gaussian kernel density to identify most of the major tourist attractions in a certain area properly.

2.2. Travel Route Recommendation

2.2.1. Data Source Types of Travel Route Mining. In the study of travel route finding and recommendation, the data collected are mostly image data with GPS location and travel note data on social media websites shared by tourists [6]. Kurashima et al. [32] developed a travel route recommendation approach, employing the photographers' histories on social photo sharing websites. Recommendations are implemented by a probabilistic photographer behavior framework associating topic models and Markov models, which consider both the user's current place and their preference. Considering the past and/or future trips, as well as the free time that can be spent in future trips, the recommendation approach outputs a group of personalized travel plans that conform to the user's preference, current location, spare time, and transportation modes. An et al. [33] devised to further consider specific user information or attributes automatically detected in large-scale photos to make personalized travel recommendations. Finally, a probabilistic Bayesian model that further entailed mobile recommendation in the field was investigated. Guo et al. [34] proposed a multi-faceted cross media correlation approach to connect piecemeal tourism information, such as discovering hot attractions and popular travel routes from crowd contributed data and travelogues. Basiri et al. [2] used crowdsourcing trajectory data based on pattern recognition method to detect interest points and guide tourists in attractions and ambient services, especially advisory services. Yong et al. [35] developed a cost-aware probabilistic matrix factorization model with Gaussian prior which aims to explore the cost preferences and tourists' interests synchronously from the mass travel logs. Majid et al. [36] proposed a new context-aware personalized approach for recommending tourist places that were related to tourists in the given context based on users' shared geotagged photos on social media sites and predicted tourists' preferences in an unfamiliar city precisely from users' travel history in one location. Lou et al. [37] used the sentiment-based POI (Points-of-Interest) Ming and Recommendation algorithm to mine the POIs with obvious sentimental attributes, and then recommend the POIs to other users. Gao et al. [38] presented a travel guide system W2Go (Where to Go), which could automatically identify and sort the attractions for tourists. They proposed a new automatic attraction ranking approach by using the photo tag and geographic tag information on Flickr.

Previous studies on tourism knowledge discovery and recommendation are mainly based on social media data and travel blogs shared by tourists [6]. However, there are two major shortcomings of these data types: (1) the rarefaction of the sample results in a fragmentary and improper description of the travel recommendation; (2) the completeness and precision of the travel blogs may also mislead the consequences of the analysis because the travel notes are usually written by memory from the tourists after their trip [39]. On the contrary, mobile phone signaling data can be collected for a large number of samples at a lower cost; meanwhile, it contains the accurate time and location information of the users [3]. Therefore, in this research, we proposed a novel framework that preprocesses the mobile signaling data to transform raw trajectories into tourists' travel sequence and finds the popular attractions and the frequent travel sequences.

2.2.2. Optimization Algorithms to Travel Route Planning. In this work, the travel route planning is closely related to traveling salesman problem (TSP). The TSP has become one of the most classical problems in combinatorial optimization research. Its purpose is to find the shortest route that can cross all cities under given city coordinates and solve other practical problems by expanding TSP. Therefore, the research on TSP has very crucial theoretical and practical significance [40]. A significant issue on the research of TSP is how to enhance accuracy and practicability of the algorithm. Several methods have been successfully implemented to TSP, including GA [41], PSO [42], SA [19], and ACO [43]. Deng et al. [44] combined cellular GA with SA to improve the optimization performance. Dong et al. [45] proposed a hybrid GA with variable proximity search, and its performance has been proven to be a variant of TSP, in which there are multiple salesmen and tasks. Zhou et al. [46] proposed a comparative research of an improved GA and PSO for the multi-TSP. Harmanani and Ghosn [47] proposed an efficient SA model that employs two different neighborhood structures: exchange and shift and rotation. Yang et al. [40] combined entropy weight learning strategy, nucleolus game strategy, and mean filtering to improve population diversity and prevent getting into local optimum. Tuani et al. [48] adopted an adaptive method for multi-phase ant colony population and employed the control parameters of evolutionary ACO algorithm to locate the approximate optimal solutions. Skinderowicz [22] proposed a new focused ACO, the core element of which is to control the difference between the newly constructed solution and the previously selected solution. This mechanism led to more focused search process and greatly improved the search efficiency. Stodola et al. [49] presented an adaptive ACO model based on the node clustering method applied to the TSP. The proposed model implemented three new techniques (contains node clustering, adaptive pheromone volatilization, and termination conditions) to improve the overall performance and reduce the execution time and negative impacts.

The advantages of GA mainly lie in parallelism and global search ability, including remarkable astringency and the ability to find the optimal solution. However, there are also disadvantages of weak local optimization abilities and easy to get into the local optimization. In PSO, there are few parameters that need to be restructured, so it is easy to carry out. Compared with other algorithms for solving TSP, its solving capacity is insufficient. SA has a good capacity to step out of the local optimization, but the selection of cooling time will immediately influence the capability of the method. The advantages of ACO primarily involve robustness and solving capability. However, due to the lack of premier pheromone in the inchoate stage, the astringency rate will be affected. In this paper, we introduce the relationship strength between attractions mined from mobile phone signaling data as the frequency coefficient into the initial heuristic pheromone to improve the convergence rate and finally get the optimal route with higher frequency and shorter distance.

3. Methodology

3.1. Questions and Framework

3.1.1. Research Questions. In order to make better tour planning, tourists usually collected various information online, for example, attractions, routes, times, and hotels, and the main purpose is to select the target attraction because it has a critical impact on other travel tasks, such as travel route scheduling [7]. Therefore, potential tourists who have not been to the destination city may mainly want to obtain the following information from other experiences: (1) where to tour; (2) how to tour. The first question is about the selection of tourist attractions, which attempts to select the most popular and most desirable attractions from the attractions of the entire city based on the experiences of other tourists. The second problem is based on the first problem: after selecting the appropriate attractions, schedule an optimal and effective recommended travel route.

In this paper, we answer the first question by implementing the frequent pattern mining algorithm to discover the popular attraction in the target city based on travel pattern sequence data. Then, the improved ACO algorithm is adopted to design an optimal recommended travel route from the selection of popular attractions to answer the second question. The results provided by our framework may facilitate tourists to better decide where to tour and how to tour in an unfamiliar city.

3.1.2. Research Framework. In this point, we will present the framework for recommending the multi-objective optimal travel route based on improved ACO algorithm. We are interested in not only discovering popular attractions and frequent travel routes but also scheduling an optimal travel route for tourists. Figure 1 shows our framework diagram. The left column exhibits data sources, including mobile phone signaling data and attraction location data. The middle column explains the pivotal modules of our model. After using the K-means clustering algorithm to identify tourists [50], we extract tourist travel sequences from tourist mobile phone trajectory data by spatial matching attraction location data. Next, popular attractions and frequent travel routes are mined based on a frequent pattern mining algorithm. Finally, a travel route recommendation is scheduled by the improved ACO algorithm for tourists. The right column presents the output of the research framework.

3.2. Data Collection

3.2.1. Mobile Phone Signaling Data. In this research, mobile phone signaling data are provided by China telecommunication operator. To deal with the big data, we investigate Hadoop and Spark to store and process the big data. The mobile phone signaling data contain the details of a phone call or other telecommunication transaction, mainly including the location information of base station, call start time, call duration, and call type (such as voice and SMS).

3.2.2. Attraction Location Data. In this study, we gathered the tourism attraction location data from tourism websites (e.g., http://www.tripadvisor.com) by querying keywords (e.g., city names) and then collected the attraction location information (i.e., latitude and longitude) into a dataset.

3.2.3. Spatial Matching. In this paper, we employ the spatial matching algorithm to transform the user's spatiotemporal trajectory into the user's travel spatiotemporal sequence between attractions. To deal with this problem, we refer to the Thiessen polygon, where each polygon theoretically represents the approximate coverage of each base station's GSM network signal. Based on the ArcGIS tool, a table of data of spatial correlations is constructed between base station data and attraction location data.

3.3. Frequent Pattern Mining Method. In this paper, a frequent pattern mining method is employed to find the correlation between popular attractions in cities, and then all the correlations are presented in a travel network. The algorithm needs to predefine the minimum frequency of items, and the output result is the item-set whose frequency appears in the dataset and is not less than the threshold value. Here, the frequency of a pattern (*fp*) related to the dataset (DS) is defined as the proportion of *fp* vectors in the *DS* containing the item-set, which is called the minimum support in data mining. The target of frequent pattern mining in this framework is to output the corresponding frequent pattern given a dataset and predefined minimum support of the item-set. Let $FP^{(n)}$ (DS) represent the mined frequent n-item-set in DS. Therefore, $FP^{(1)}$ (DS) denotes the group of popular attractions, and $FP^{(2)}$ (DS) shows the group of frequent relationships between these popular attractions. In other words, the frequent pattern mining method mainly

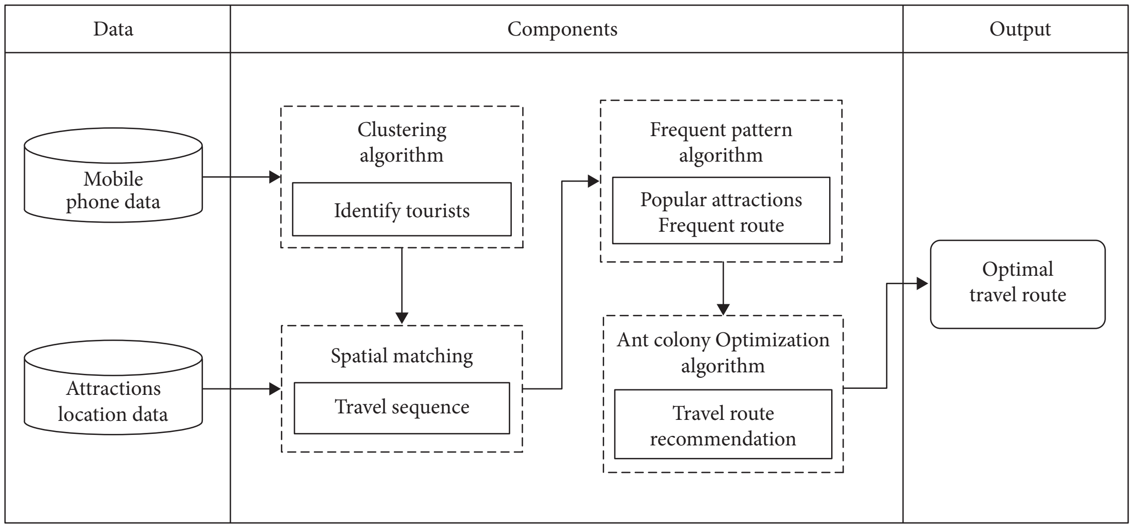


FIGURE 1: The research framework.

focuses on finding the frequent 1-itemsets and 2-itemsets in DS.

The frequent pattern mining algorithm can only be used to discover frequent item-sets, not to find association rules. In general, association rules are used to mine the association relationships hidden behind the transactions, often accompanied by support and confidence. When making tourist attraction recommendations, we want to explore the correlation probability of the association between attraction A and attraction B and whether it is necessary to recommend both for tourists. Therefore, here we introduce the maximum confidence algorithm. Given a frequent 2-itemset, the metric of maximum confidence measure is defined as

$$\begin{aligned} \theta_{\{A,B\}} &= \frac{\operatorname{supp}(\{A,B\})}{\min\{\operatorname{supp}(\{A\}),\ \operatorname{supp}(\{B\})\}}, \\ &= \max\{\Pr(B|A),\ Pr(A|B)\}, \\ \theta_{\{A,B\}} &\geq \theta_0\ \theta_0 \in [0,1], \end{aligned} \tag{1}$$

where A and B represent different attractions and θ_0 denotes the maximum confidence threshold. Without loss of generality, if $\theta_{\{A,B\}} \geq \theta_0$ and $\operatorname{supp}(\{A\}) \geq \operatorname{supp}(\{B\})$ are satisfied, then attraction A is named for the master node of attraction B. If a tourist wants to visit attraction A, then this tourist would also be recommended to visit attraction B together.

3.4. Improved Ant Colony Optimization Algorithm. The ACO algorithm is a metaheuristic global optimization intelligent search approach that seeks optimal travel path through distributed collaboration [51–54]. This algorithm has the characteristics of diversity and positive feedback, in which diversity enables the ant colony to innovate, and the positive feedback enables the ant colony to learn and strengthen, so as to retain the excellent information and finally obtain the optimal solution. It has advantages in solving the combinational optimization problems such as large-scale travel agent problems and resource allocation problems such as scheduling. In the following, we will introduce a formal definition of the basic ACO.

In the implementation of the approach, travel path node selection and pheromone updating are the two most important steps. Supposing ant k is located at node i at time t, the probability of travel path node selection is calculated by the following formula, and the roulette model is used to select the next node. The probability of its traveling to the next node j on node i is given by

$$P_{ij}^{k} = \begin{cases} \dfrac{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)}{\sum_{s\in C}\tau_{is}^{\alpha}(t)\eta_{is}^{\beta}(t)}, & j \in C, \\ 0, & j \notin C, \end{cases} \tag{2}$$

$$\eta_{ij}(t) = \frac{1}{d_{ij}}, \tag{3}$$

where C is the group of nodes that ants can choose at the next step, $\tau_{ij}(t)$ is the pheromone value of the path from node i to node j, $\eta_{ij}(t)$ is the heuristic value, which presents the expectation degree of ants traveling from node i to node j, and α and β represent the relative weight parameters that control pheromone value and heuristic value, respectively.

Each ant also secretes pheromones in the process of employing the pheromone and the pheromone on the travel path will accumulate. To prevent too many pheromones from drowning the heuristic value, the pheromone on the travel path will update after each ant has completed a cycle, and the pheromone value on the travel path at time $t+1$ is updated according to

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \rho\triangle\tau_{ij}(t), \tag{4}$$

$$\triangle\tau_{ij}(t) = \begin{cases} \dfrac{Q}{L_k}, & \text{ant}k\ \text{passed.through.path}\langle i,\ j\rangle, \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where ρ is the pheromone volatility coefficient, $1-\rho$ is the residual factor, $\triangle\tau_{ij}(t)$ is the enhancement of pheromone on the path $<i, j>$ in this cycle, Q is the strength of the pheromone, and L_k is the full length of the travel path traveled by the ant in this cycle.

The heuristic value of the basic ACO algorithm is reciprocal ratio to the distance from the current node to the next node, which prompts the ant to select a shorter length path [55]. However, considering the frequent relationship strength between nodes, the distance between two nodes is short, but the transfer probability is not necessarily the maximum. Meanwhile, in the later stage of the search, to accelerate the astringency speed, it is necessary to weaken the impact of heuristic value on the travel path selection. Therefore, in this paper, we employed a novel approach to improve the heuristic value function by referring to the frequent pattern coefficient λ and the damping coefficient ϕ. The novel heuristic value function is presented as follows:

$$\eta_{ij}(t) = \frac{1}{d_{ij}} * \gamma * \varphi, \tag{6}$$

$$\gamma = \frac{\text{supp}(fp\langle i, j\rangle)}{\sum_{s\in C}\text{supp}(fp\langle i, s\rangle)}\ j \in C, \tag{7}$$

$$\varphi = \begin{cases} \dfrac{N_{\max} - N}{N_{\max}}, & N \neq N_{\max,} \\ \dfrac{1}{N_{\max}}, & N = N_{\max,} \end{cases} \tag{8}$$

where $\text{supp}(fp\langle i, j\rangle)$ denotes the frequency of path $<i, j>$, $N_{\max}$ represents the maximum number of iterations, and N is the current iteration number.

4. Empirical Study

4.1. Data Description. The targeted tourism city for the experimental application of this framework is Xiamen, Fujian Province, China, and the mobile phone signaling data were provided by China Mobile Xiamen Operator. The mobile phone signaling data include more than 6 billion records from 10 million users in Xiamen city for a month in 2015. There are more than 5,500 telecommunication base stations in the city of Xiamen. Besides, we crawled the attraction location data from http://www.tripadvisor.com to structure the attraction location dataset. For the city of Xiamen, there are a total of 256 geographical locations (tourist attractions) extracted, among which 5 duplicate attractions were removed, and a total of 251 attraction location data were finally obtained. Furthermore, we presented the base stations and attractions of Xiamen on a map together using ArcGIS (Figure 2(a)). Meanwhile, we employed the spatial matching algorithm in the ArcGIS tool to establish the spatial correlation data table between the base station and the attraction through the theoretical signal coverage of the base station (Thiessen polygon) and the location of the attraction. Besides, we showed the Thiessen polygon and attractions of Xiamen on a map together using ArcGIS (Figure 2(b)).

In this study, we randomly selected 100,000 users from 10 million anonymous mobile phone users as the study subject. Before processing the data, we preprocessed the mobile signaling data, such as duplicated records, missing records, invalid records, and data drift. After preprocessing the data, we extracted 31 features from their spatiotemporal trajectories which mainly included period, active days, weekdays, monthly distance, daily distance, daily centroid distance, overlap, total number of base stations interacted, daily base stations interacted, interaction time, daily records generated, total number of trips, number of daily trips, average trip distance, average trip speed, average travel time, stop duration, and so on. Then, we employed the K-means clustering algorithm to extract tourist clusters using tourist characteristics and identified 11,366 tourists who visited many attractions during their short stay in Xiamen. Next, the travel trajectory sequence of tourists between attractions was filtered out from the spatiotemporal trajectory of these tourists, and we removed the records where the tourists stayed in the attraction for less than 5 minutes. Finally, the day is taken as the segmentation point to cut the tourist travel sequence and get the daily travel sequence vector for each tourist.

4.2. Popular Tourism Attractions. In the beginning, the frequent pattern mining method is employed to mine the popular attractions in Xiamen. The result shows that the frequency of attractions to their corresponding ranks follows the general power-law distribution. In this study, by fitting the frequency of the attractions and the corresponding ranking to the logarithm, it is found that the linear characteristics of the first few points are not very strong, while most of the points afterward are almost a straight line with a negative slope. Overall, the determining coefficient of the fitting function (i.e., R^2) is 0.9727, and the residual sum of squares (RSS) is 0.0043, showing a very good fitting effect (Figure 3). This feature shows that the distribution of attractions in Xiamen tourism is unbalanced. A few scenic spots are very popular (often visited by most tourists), while the others are not. In addition, the top 50 frequent attractions visited by tourists in Xiamen city are shown in Figure 4(a). It can be found that the Xiamen Ferry Terminal is the most popular attraction with more than 8,000 visits. Besides, we displayed the top 50 popular attractions on the city map which showed that the popular attractions are mainly distributed in the Gulangyu Scenic Area and along the beach (Figure 4(b)).

4.3. Tourism Area Identification. After using the frequent pattern mining method, we adopted the max-confidence threshold, $\theta_0 = 0.25$, in selecting popular travel sequences

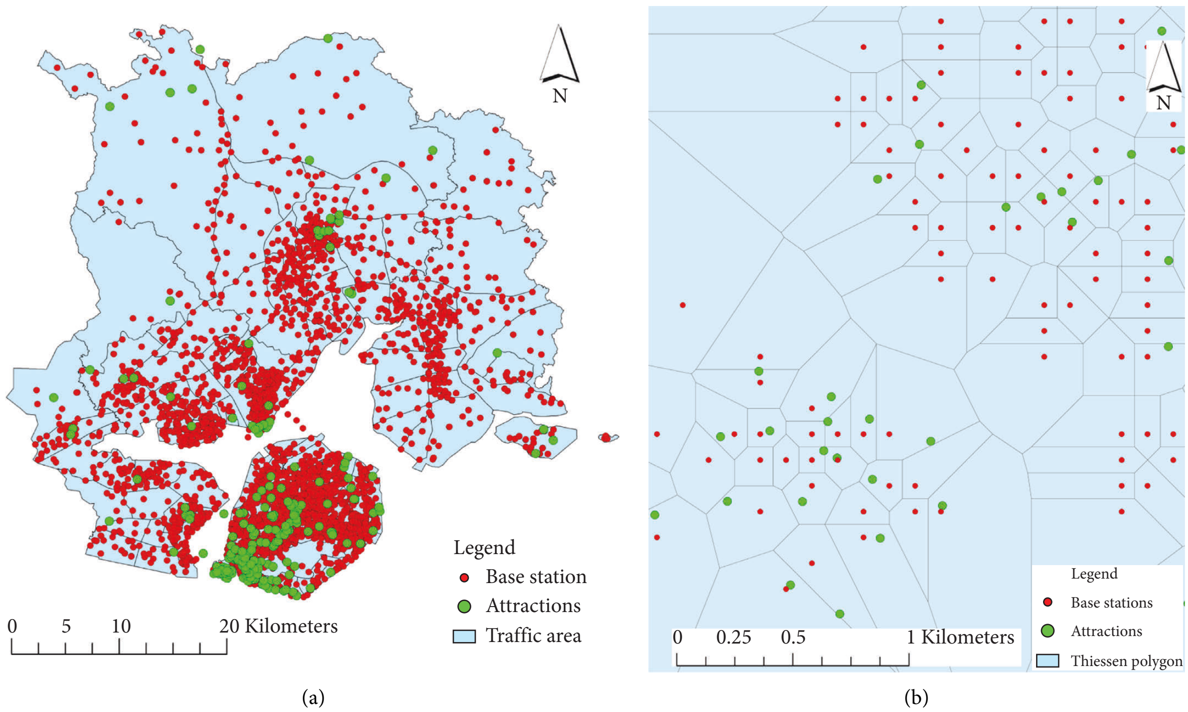


FIGURE 2: Distribution of base station and attractions in Xiamen. (a) Base station and attractions. (b) Thiessen polygon and attractions.

with high correlations and sketching out a popular travel network which was comprised of popular attractions (i.e., network nodes) and the relationships (e.g., network edges) in Xiamen (Figure 5). In Figure 5, each node denotes a popular attraction, and the node size indicates its popularity. The edges describe that there are relationships between attractions, and the edge thickness means the strength of the relationship (the thicker the edge is, the stronger the relationship is). Furthermore, we possibly divide the travel network into three tourism areas (Gulangyu tourism area, Zhongshan Road tourism area, and Beach tourism area) and assign the popular attractions to each tourism area in Xiamen to verify the result of the identification (Figure 6). It is shown that the scenic spots are clustered in a data-driven way according to the travel sequence of tourists from mobile phone signaling data instead of using spatial geographical proximity. For example, some green attractions and purple attractions are very close to one another spatially, but they are divided into different tourism areas because there are few tourists visiting those attractions together. It is helpful to package these popular attractions in the same tourism area into one travel plan for potential tourists. Meanwhile, by setting a relatively high value of max-confidence threshold $\theta_0 = 0.35$, tourists can get popular travel routes for identifying tourism areas (Figures 7(a)–7(c)). As we have seen, only the scenic spots with high frequency and relationships have been reserved in the same tourism area to arrange a travel path.

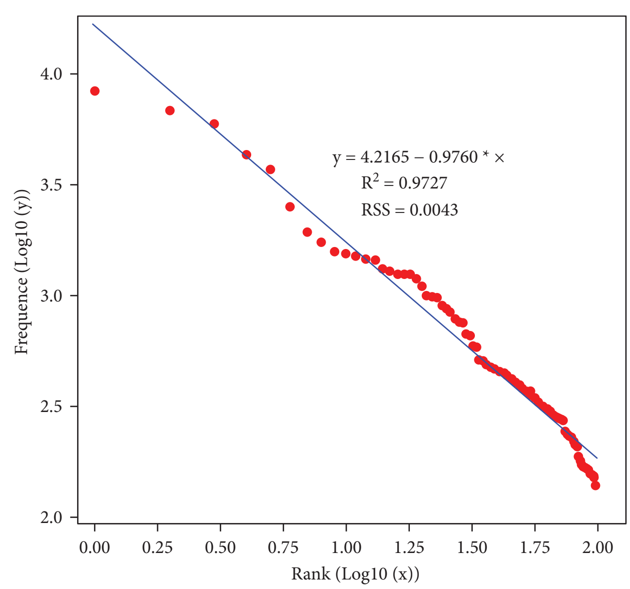
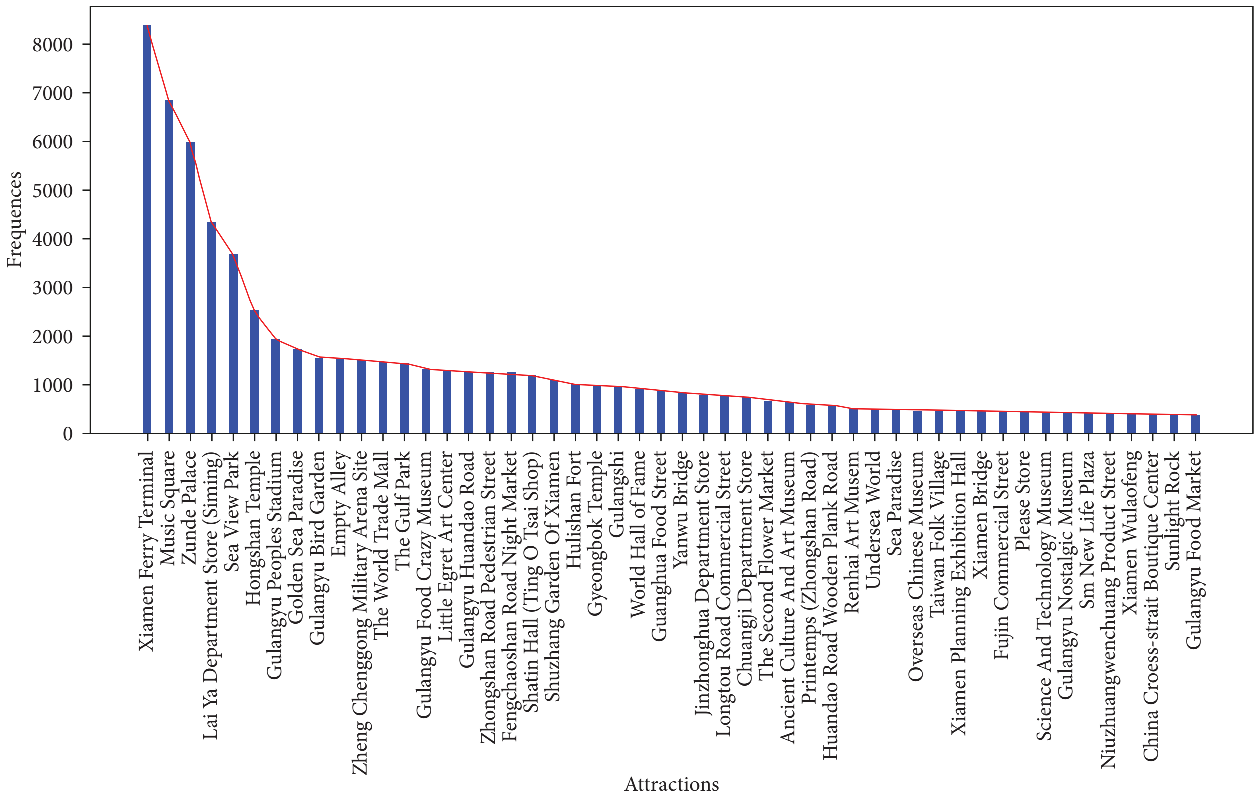
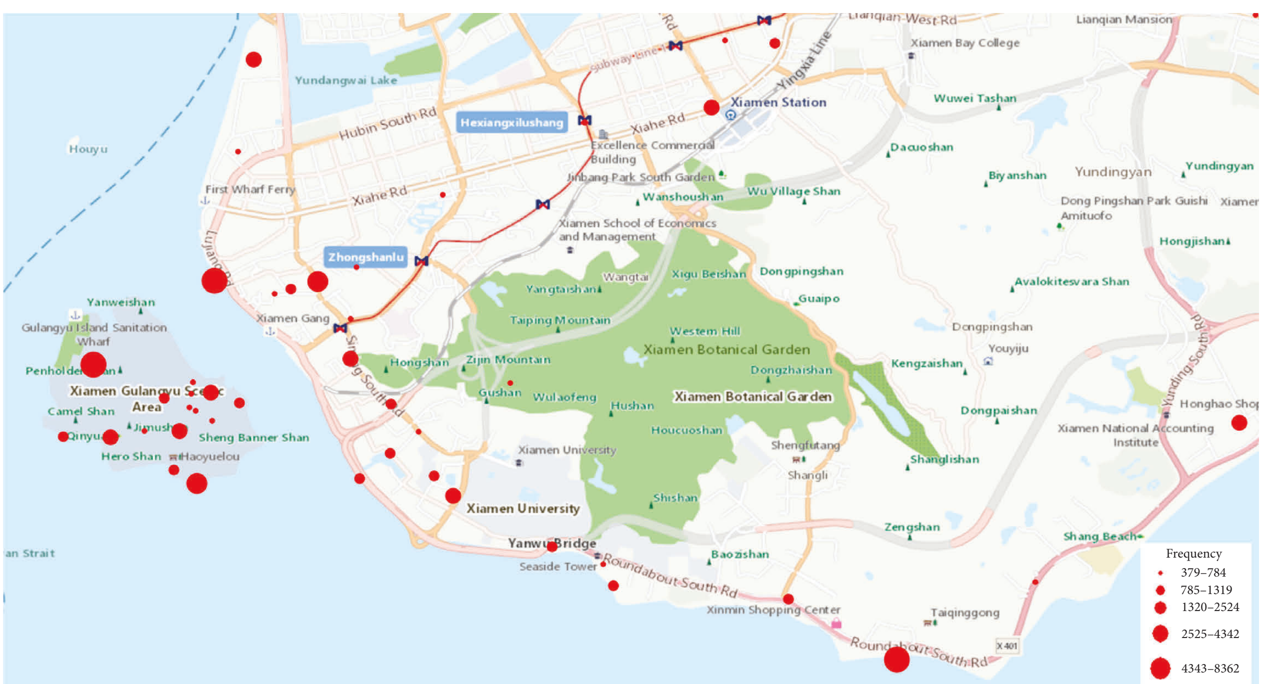


FIGURE 3: Frequencies of attractions and their rank relationship chart.

4.4. Optimal Route Recommendation. After mining the popular attractions and frequent travel sequences and identifying the tourism area, we adopt the improved ACO algorithm to schedule an optimal travel route in each tourism area for potential tourists. In this paper, we have taken one day as a segmentation point to cut the tourist travel sequence and obtain the daily travel sequence. We hope to plan an optimal travel route to visit as more popular attractions as possible in the tourism area in one day (tour time <8 hours); therefore, we take the shorter total distance and higher frequency as the ultimate multi-objective travel route.



(a)



(b)

FIGURE 4: Popular tourism attractions in Xiamen. (a) Top 50 popular attractions. (b) Top 50 popular attractions on the city map.

By mining the tour time of each tourist in each attraction through mobile phone signaling data, we can calculate the average tour time recommended by each attraction and the number of tourists in each attraction. Through the brute force search approach, we can get the combination of visiting the popular attractions as much as possible within the specified tour time. We selected 13 popular attractions and 12 popular attractions, respectively, from Gulangyu tourism

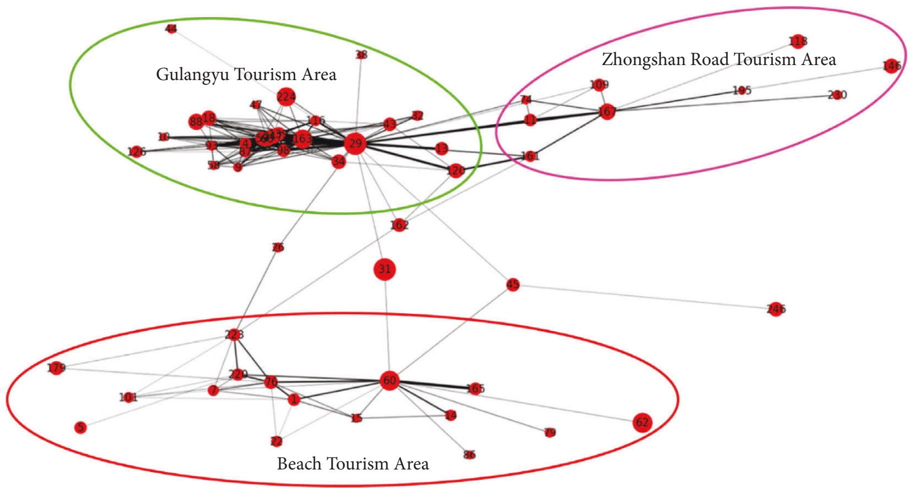


FIGURE 5: Tourism areas identified from the travel network ($\theta_0 = 0.25$).

FIGURE 6: Identified tourism areas on the city map.

area (27 scenic spots) and Beach tourism area (17 scenic spots). The recommended visiting time and popularity for each attraction are shown in Table 1.

After selecting the popular attractions, we employ the improved ACO to explore the multi-objective optimal travel route. Setting the parameters for the improved ACO algorithm is crucial to the performance of the algorithm, and numerous researchers have conducted a large number of studies on the optimization of the algorithm parameters. In order to reduce the complexity of the algorithm, this paper adopts an empirical and experimental approach to select the optimal combination of parameters by setting different parameters for simulation experiments and analyzing the benefits and drawbacks of the experimental results [55]. The test method takes a group of values for each parameter. In each experiment, only one parameter is changed, and other parameters remain unchanged. The number of iterations, run time, frequency, and the route length are tested. In this paper, set the parameters as follows: $m = 50$, $N_{\max} = 500$, $Q = 100$, $\alpha = 1$, $\beta = 3$, and $\rho = 0.2$. In addition, to fully verify the performance of the improved ACO, it is compared with a series of effective intelligent optimization approaches, i.e., GA, PSO, SA, and dynamic programming (DP). Comparative data are shown in Table 2.

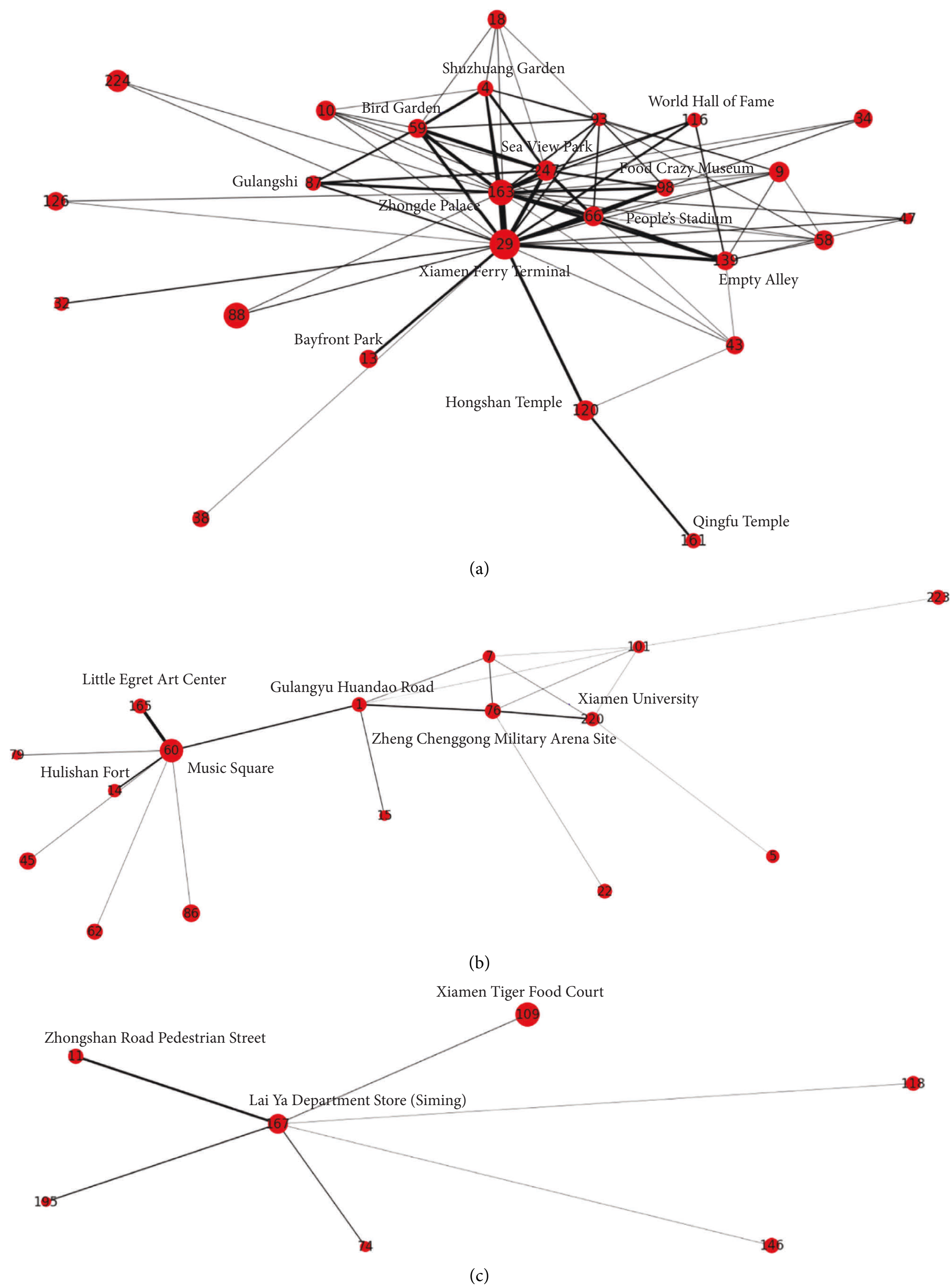


Figure 7: Tourism areas identified in Xiamen. (a) Gulangyu tourism area travel network (θ_0 = 0.35). (b) Beach tourism area travel network (θ_0 = 0.35). (c) Zhongshan Road tourism area travel network (θ_0 = 0.35).

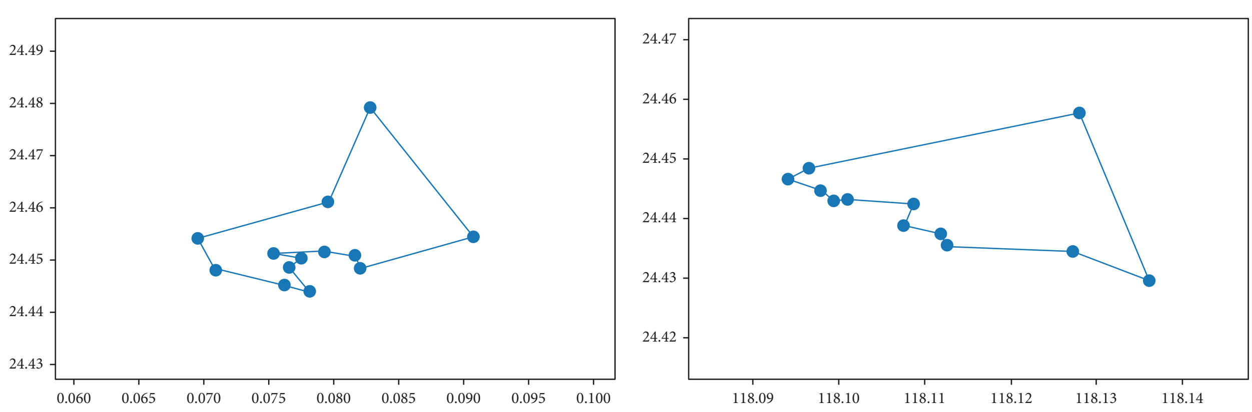
According to Table 2, the improved ACO outperforms other intelligent optimization methods on travel route frequent. In terms of run time, the advantages of improved ACO are magnified significantly than GA and similar to PSO, SA, and DP. Meanwhile, compared with PSO, the improved ACO is superior to it on iteration. To sum up, the improved ACO with a new exploratory heuristic information approach can facilitate the multi-objective optimization approach to find better and more suitable travel route. Also, we plot the optimal route on a map (Figure 8).

Table 1: The recommended visiting time and popularity for each attraction (visiting time: min).

Tourism area	Attraction name	Popularity	Visiting time
	Xiamen Ferry Terminal	8362	35
	Zunde Palace	5984	45
	Sea View Park	3692	48
	Hongshan Temple	2524	28
	Gulangyu People's Stadium	1937	49
	Gulangyu Bird Garden	1572	47
Gulangyu tourism area	Empty Alley	1541	46
	The Gulf Park	1441	19
	Gulangyu Food Crazy Museum	1319	41
	Shuzhuang Garden of Xiamen	1095	38
	World Hall of Fame	902	39
	Longtou Road Commercial Street	763	28
	Haoyueyuan	288	16
	Sum	31420	479
	Music Square	6853	86
	Zheng Chenggong Military Arena Site	1501	42
	Little Egret Art Center	1286	39
	Gulangyu Huandao Road	1249	32
	Fengchaoshan Road Night Market	1245	59
	Shatin Hall (Ting O Tsai Shop)	1185	44
Beach tourism area	Hulishan Fort	1000	28
	Huandao Road Wooden Plank Road	585	29
	Overseas Chinese Museum	476	48
	Xiamen University	354	30
	Wenzeng Road Strange Slope	271	13
	Xiamen University History Exhibition Hall	169	24
	Sum	16174	474

Table 2: Performance analysis of each optimization algorithm.

Algorithm	Travel route	Frequency	Distance (km)	Run time (s)	Iteration
ACO	[3, 7, 0, 1, 5, 9, 2, 4, 11, 8, 6, 10, 12]	1445	10.78	3.94	117
GA	[5, 9, 2, 4, 8, 11, 6, 10, 12, 3, 7, 0, 1]	1435	10.71	140.62	39
PSO	[7, 0, 1, 5, 9, 2, 4, 8, 11, 6, 10, 12, 3]	1435	10.71	2.59	243
SA	[11, 6, 10, 12, 3, 7, 0, 1, 5, 9, 2, 4, 8]	1435	10.71	3.16	8831
DP	[5, 8, 11, 4, 9, 2, 6, 10, 12, 3, 7, 0, 1]	1248	11.06	1.43	—



(a) (b)

Figure 8: Optimal travel routes in Xiamen. (a) Optimal travel route in the Gulangyu tourism area. (b) Optimal travel route in the Beach tourism area.

5. Conclusion

In this work, we design a novel research framework by investigating the historical mobile signaling data of previous tourists to make an optimal travel route recommendation for potential tourists. Firstly, we propose a novel multi-objective optimal travel route recommendation framework, which collects tourists' travel trajectories from mobile signaling data. Secondly, we employ a frequent pattern mining method to find popular attractions and frequent travel routes from the travel pattern sequences transformed through tourists' raw trajectories. Thirdly, we adopt an improved ACO algorithm with a novel extensible heuristic factor to find the multi-objective optimal route of shortest distance and highest frequency according to the popularity of attractions and travel time of tourists. Finally, the experimental results indicate that the proposed framework is efficient in recommending multi-objective optimal travel routes considering tourists' travel time and attractions' popularity while ensuring that the recommended route is suitable. However, in this research, we have not considered the actual transportation accessibility between the attractions, the travel time on the trip, the crowd level and carrying capacity of the attractions, and the user's personalized preferences. Meanwhile, the proposed improved ACO has some limitations: the frequency coefficient is investigated to increase the astringency rate of the model, but it also brings the complexity of parameter adjustment. Meanwhile, with the increase of graph network complexity, the running speed of the model will increase greatly. The future research will build the optimal recommendation of multi-objective dynamic tourism travel route model combined with the actual environment, using reinforcement learning algorithms and multi-agent simulation to implement real-time dynamic travel route recommendations [56–59].

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Haodong Sun, Yanyan Chen, and Xiaoming Liu were responsible for research conception and design. Haodong Sun and Yanyan Chen were responsible for data collection. Haodong Sun and Jiachen Wang were responsible for analysis and discussion of results. Haodong Sun, Jianming Ma, and Yang Wang were responsible for draft manuscript writing. All authors approved the final version of the manuscript.

Acknowledgments

This research was funded by the National Key Research and Development Program of China (2018YFB1601302).

References

[1] Malik and Kim, "Optimal travel route recommendation mechanism based on neural networks and particle swarm optimization for efficient tourism using tourist vehicular data," *Sustainability*, vol. 11, pp. 3357–12, 2019.

[2] A. Basiri, P. Amirian, A. Winstanley, and T. Moore, "Making tourist guidance systems more intelligent, adaptive and personalised using crowd sourced movement data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 2, pp. 413–427, 2017.

[3] X. Zhu, T. Sun, H. Yuan, Z. Hu, and J. Miao, "Exploring group movement pattern through cellular data: a case study of tourists in hainan," *ISPRS International Journal of Geo-Information*, vol. 8, pp. 74–2, 2019.

[4] D. Cao, K. Zeng, J. Wang et al., "BERT-based deep spatial-temporal network for taxi demand prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9442–9454, 2022.

[5] J. WangWangWangWang, H. Han, H. Li, S. M. He, P. Kumar Sharma, and L. Chen, "Multiple strategies differential privacy on sparse tensor factorization for network traffic analysis in 5G," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1939–1948, Mar 2022.

[6] S. Du, H. Zhang, H. Xu, J. Yang, and O. Tu, "To make the travel healthier: a new tourism personalized route recommendation algorithm," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 9, pp. 3551–3562, 2018.

[7] H. Yuan, H. Xu, Y. Qian, and Y. Li, "Make your travel smarter: summarizing urban tourism information from massive blog data," *International Journal of Information Management*, vol. 36, no. 6, pp. 1306–1319, 2016.

[8] L. Hang, S.-H. Kang, W. Jin, and D.-H. Kim, "Design and implementation of an optimal travel route recommender system on big data for tourists in jeju," *Processes*, vol. 6, pp. 133–8, 2018.

[9] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: a survey," *GeoInformatica*, vol. 19, no. 3, pp. 525–565, 2015.

[10] S. K. Zhao and M. R. Lyu, "A survey of point-of-interest recommendation in location-based social networks," in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pp. 25–30, 2015, https://arxiv.org/abs/1607.00647.

[11] F. Lorenzi, S. Loh, and M. Abel, "PersonalTour: a recommender system for travel packages," in *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, Lyon, France, August 2011.

[12] P. Resnick and H. R. Varian, "Recommender systems," *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.

[13] C. W. Xin Lu, J.-M. Yang, Yanwei pang, and lei Zhang, "Photo2Trip: generating travel routes from geo-tagged photos for trip planning," in *Proceedings of the ACM Multimedia 2010 International Conference*, pp. 622–625, Firenze, Italy, October 2010.

[14] C. Ntakolia and D. V. Lyridis, "A comparative study on Ant Colony Optimization algorithm approaches for solving multi-objective path planning problems in case of unmanned surface vehicles," *Ocean Engineering*, vol. 255, Article ID 111418, 2022.

[15] Y. Singh, S. Sharma, R. Sutton, D. Hatton, and A. Khan, "A constrained A^* approach towards optimal path planning for an unmanned surface vehicle in a maritime environment

containing dynamic obstacles and ocean currents," *Ocean Engineering*, vol. 169, pp. 187–201, 2018.

[16] R. Song, Y. Liu, and R. Bucknall, "A multi-layered fast marching method for unmanned surface vehicle path planning in a time-variant maritime environment," *Ocean Engineering*, vol. 129, pp. 301–317, 2017.

[17] R. Song, Y. Liu, and R. Bucknall, "Smoothed A^* algorithm for practical unmanned surface vehicle path planning," *Applied Ocean Research*, vol. 83, pp. 9–20, 2019.

[18] H. M. Harmanani, F. Drouby, and S. B. Ghosn, "A parallel genetic algorithm for the open-shop scheduling problem using deterministic and random moves," in *Proceedings of the 2009 Spring Simulation Multiconference*, SpringSim 2009, San Diego, California, USA, March 2009.

[19] M. F. Demiral and A. H. Işik, "Simulated annealing algorithm for a medium-sized TSP data," in *Proceedings of the Artificial Intelligence and Applied Mathematics in Engineering Problems*, pp. 457–465, Springer International Publishing, Antalya, Turkey, April 2020.

[20] P. Pongchairerks and V. Kachitvichyanukul, "A two-level particle swarm optimisation algorithm for open-shop scheduling problem," *International Journal of Computing Science and Mathematics*, vol. 7, no. 6, p. 575, 2016.

[21] S. Mirjalili, J. Song Dong, and A. Lewis, "Ant colony optimizer: theory, literature review, and application in AUV path planning," in *Nature-Inspired Optimizers Studies in Computational Intelligence*pp. 7–21, Berlin, Germany, 2020.

[22] R. Skinderowicz, "Improving Ant Colony Optimization efficiency for solving large TSP instances," *Applied Soft Computing*, vol. 120, p. 108653, 2022.

[23] W. Deng, J. Xu, and H. Zhao, "An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem," *IEEE Access*, vol. 7, pp. 20281–20292, 2019.

[24] J. E. Bell and P. R. McMullen, "Ant colony optimization techniques for the vehicle routing problem," *Advanced Engineering Informatics*, vol. 18, no. 1, pp. 41–48, 2004.

[25] S. Zhang, R. Zhang, X. Liu, and H. Sun, "A Personalized Trust-Aware Model for Travelogue Discovering," in *Proceedings of the IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, Macau, China, December 2012.

[26] R. C. Q. Hao and C. Wang, Y. Pang and L. Zhang, Equip tourists with knowledge mined from travelogues," in *Proceedings of the 19th International Conference on World Wide Web*, Raleigh North Carolina USA, April 2010.

[27] M. Ye, R. Xiao, W. C. Lee, and X. Xie, "On theme location discovery for travelogue services," in *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR 2011, Beijing, China, July 2011.

[28] H. Xu, H. Yuan, B. Ma, and Y. Qian, "Where to go and what to play: towards summarizing popular information from massive tourism blogs," *Journal of Information Science*, vol. 41, no. 6, pp. 830–854, 2015.

[29] L.-Y. Wei, W.-C. Peng, and W.-C. Lee, "Exploring pattern-aware travel routes for trajectory search," *ACM Transactions on Intelligent Systems and Technology*, vol. 4, no. 3, pp. 1–25, 2013.

[30] Y. Hu, S. Gao, K. Janowicz, B. Yu, W. Li, and S. Prasad, "Extracting and understanding urban areas of interest using geotagged photos," *Computers, Environment and Urban Systems*, vol. 54, pp. 240–254, 2015.

[31] E. Pla-Sacristán, I. González-Díaz, T. Martínez-Cortés, and F. Díaz-de-María, "Finding landmarks within settled areas using hierarchical density-based clustering and meta-data from publicly available images," *Expert Systems with Applications*, vol. 123, pp. 315–327, 2019.

[32] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, "Travel route recommendation using geotagged photos," *Knowledge and Information Systems*, vol. 37, no. 1, pp. 37–60, 2013.

[33] Y. Y. C. An, J. Cheng, W. H. Yen Ta Huang, Hsu, and Hong Yuan Mark Liao, "Personalized travel recommendation by mining people attributes from community-contributed photos," *Proceedings of the 19th International Conference on Multimedea*, 2011.

[34] T. Guo, B. Guo, Y. Ouyang, Z. Yu, J. C. K. Lam, and V. O. K. Li, "CrowdTravel: scenic spot profiling by using heterogeneous crowdsourced data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 6, pp. 2051–2060, 2017.

[35] G. Yong, L. Qi, X. Hui, A. Tuzhilin, and C. Jian, "Cost-aware travel tour recommendation," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, USA, August 2011.

[36] 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.

[37] P. Lou, G. Zhao, X. Qian, H. Wang, and X. Hou, "Schedule a Rich Sentimental Travel via Sentimental POI Mining and Recommendation," in *Proceedings of the IEEE Second International Conference on Multimedia Big Data*, BigMM), Taipei, Taiwan, April 2016.

[38] Y. Gao, J. Tang, R. Hong, Q. Dai, T.-S. Chua, and R. Jain, "W2Go," in *Proceedings of the international conference on Multimedia - MM'10*, Firenze, Italy, 2010.

[39] N. Mou, R. Yuan, T. Yang, H. Zhang, J. Tang, and T. Makkonen, "Exploring spatio-temporal changes of city inbound tourism flow: the case of Shanghai, China," *Tourism Management*, vol. 76, Article ID 103955, 2020.

[40] K. Yang, X. You, S. Liu, and H. Pan, "A novel ant colony optimization based on game for traveling salesman problem," *Applied Intelligence*, vol. 50, no. 12, pp. 4529–4542, 2020.

[41] F. Yu, X. Fu, H. Li, and G. Dong, "Improved roulette wheel selection-based genetic algorithm for TSP," in *Proceedings of the 2016 International Conference on Network and Information Systems for Computers*, ICNISC), Wuhan, China, April 2016.

[42] I. Dubey and M. Gupta, "Uniform mutation and SPV rule based optimized PSO algorithm for TSP problem," in *Proceedings of the International Conference on Electronics & Communication Systems*, pp. 168–172, Coimbatore, India, February 2017.

[43] M. Dorigo, "Member, ZEEE, V. Maniezzo, and albert, "ant system: optimization by a colony of cooperating agents,".

[44] Y. Deng, J. Xiong, Q. Wang, and A. A. R. Hosseinabadi, "A hybrid cellular genetic algorithm for the traveling salesman problem," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–16, Article ID 6697598, 2021.

[45] X. Dong, H. Zhang, M. Xu, and F. Shen, "Hybrid genetic algorithm with variable neighborhood search for multi-scale multiple bottleneck traveling salesmen problem," *Future Generation Computer Systems*, vol. 114, pp. 229–242, 2021.

[46] H. Zhou, M. Song, and W. Pedrycz, "A comparative study of improved GA and PSO in solving multiple traveling salesmen problem," *Applied Soft Computing*, vol. 64, pp. 564–580, 2018.

[47] H. M. Harmanani and S. B. Ghosn, "An efficient method for the open-shop scheduling problem using simulated annealing," in *Information Technolog: New Generations Advances in Intelligent Systems and Computing*, pp. 1183–1193, Springer, Berin, Germany, 2016.

[48] A. F. Tuani, E. Keedwell, and M. Collett, "Heterogenous adaptive ant colony optimization with 3-opt local search for the travelling salesman problem," *Applied Soft Computing*, vol. 97, Article ID 106720, 2020.

[49] P. Stodola, P. Otřísal, and K. Hasilová, "Adaptive ant colony optimization with node clustering applied to the travelling salesman problem," *Swarm and Evolutionary Computation*, vol. 70, Article ID 101056, 2022.

[50] H. Sun, Y. Chen, J. Lai, Y. Wang, and X. Liu, "Identifying tourists and locals by K-means clustering method from mobile phone signaling data," *Journal of Transportation Engineering, Part A: Systems*, vol. 147, no. 10, 2021.

[51] D. L. Luo and S. X. Wu, "Ant colony optimization with potential field heuristic for robot path planning," *XI Tong Gong Cheng Yu Dian Zi Ji Shu/systems Engineering & Electronics*, vol. 32, no. 6, pp. 1277–1280, 2010.

[52] G. Yi, Z. Feng, T. Mei, P. Li, W. Jin, and S. Chen, "Multi-AGVs path planning based on improved ant colony algorithm," *The Journal of Supercomputing*, vol. 75, no. 9, pp. 5898–5913, 2019.

[53] Y. L. Zhang and X. M. Niu, "Simulation Research on Mobile Robot Path Planning Based on Ant Colony Optimization," *Computer Simulation*, vol. 28, no. 6, pp. 231–234, 2011.

[54] M. X. Zhou, "Improved ant colony algorithm with planning of dynamic path," *Computer Science*, vol. 40, no. 1, pp. 314–316, 2013.

[55] Q. Luo, H. Wang, Y. Zheng, and J. He, "Research on path planning of mobile robot based on improved ant colony algorithm," *Neural Computing & Applications*, vol. 32, no. 6, pp. 1555–1566, 2019.

[56] G. Leguizamón and Z. Michalewicz, "A new version of ant system for subset problems," in *Proceedings of the Congress on Evolutionary Computation*, Washington, DC, USA, July 1999.

[57] H. Nezamabadi-pour, S. Saryazdi, and E. Rashedi, "Edge detection using ant algorithms," *Soft Computing*, vol. 10, no. 7, pp. 623–628, 2005.

[58] A. Shmygelska and H. H. Hoos, "An ant colony optimisation algorithm for the 2D and 3D hydrophobic polar protein folding problem," *BMC Bioinformatics*, vol. 6, no. 1, p. 30, Feb 14 2005.

[59] T. S. Feng Ling, X. Zhu, Q. Chen, X. Tang, and Xin Ke, "Mining travel behaviors of tourists with mobile phone data: a case study in Hainan," in *Proceedings of the 2016 2nd IEEE International Conference on Computer and Communications(ICCC)*, IEEE, Chengdu, China, October 2016.