IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 21 April 2023, accepted 9 May 2023, date of publication 15 May 2023, date of current version 30 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3275956

APPLIED RESEARCH

Co-Existing in Differences: Tailoring Travel Itineraries for Tourists With Similar Interests

CHAO YU AND MENGLI YANG
School of Management, Shenyang University of Technology, Shenyang, Liaoning 110870, China
Corresponding author: Mengli Yang (mengli_yang1997@163.com)

This work was supported in part by the Liaoning Provincial Social Science Planning Fund L21CGL021.

ABSTRACT Customized tourism is gaining more and more attention in the era of the prominent personalization of tourists. Private customization has disadvantages such as high cost and resource consumption. To save resources and costs from the three aspects of travel companies, tourists, and society, it is proposed to bring similar tourists together and customize travel itineraries for tourists. Online resource information related to tourism and transportation help companies customize the itinerary more efficiently and accurately, and the itinerary has a detailed schedule. In this method, we first cluster visitors according to their heterogeneous needs. The relevant online information collected is then combined with the needs of each visitor in the tour group to build and solve customized models of the tour itinerary. Finally, the price of the tour itinerary is determined. At the end of this paper, we discussed what level of clustering makes visitors most satisfied. It proves that this approach can be customized for tour groups. Visitors show a certain level of satisfaction with this customized format. However, the applicability of this method needs to be further verified on a larger range of data sets.

INDEX TERMS Customized tourism, heterogeneous demand, tourist clustering, model optimization.

I. INTRODUCTION

The service industry plays an important role in the world economy [1]. Tourism is a kind of service industry, which provides scenic spots, accommodation, catering, and other related services [2]. Tourists' demands for tourism products are increasingly personalized [3], and the personalization of tourism products is getting more and more attention from tourism companies and tourists [4], [5]. With the advent of COVID-19, visitor preferences have changed, and capturing visitor needs has become even more important [6], [7]. Generally speaking, the customization of private tourism must take full account of the individual needs of tourists, but it also increases the cost of tourists and tourism companies. Everyone traveling alone increases the social pressure of traffic and environmental problems. Some scholars have suggested that tourism has strong resilience. With the efforts of the government, the tourism industry, and other parties, it is only a matter of time before the tourism industry recovers [8], [9]. The emergence of COVID-19 has brought about a rethink on tourism, and some scholars have proposed the need to rethink tourism in terms of sustainability, and social and public interest [10], [11]. Scholars have also proposed group tourism [12], [13], in which tourists within a group travel to the same destination by the same means of transport, and tour companies design itineraries for their groups. The group tour usually helps tour companies to secure lower prices for attractions, hotels, and restaurants, which in turn reduces tourist costs; tour companies customize itineraries for the entire group rather than for each individual in the group, which saves company resources; and tourists within the group travel by the same mode of transportation rather than individually, which reduces transportation costs. Compared with private travel, group travel greatly reduces energy consumption and improves social welfare. However, in reality, each tourist has his preference, and the preference of tourists in the same group may produce irreconcilable conflicts, such as one person wanting to go to A scenic spot, while another person refuses to go to A scenic spot. If conflicts of heterogeneous preferences are avoided and tourists with common interests are allowed to travel together, then personalized travel itineraries can be achieved to a greater extent. In this

The associate editor coordinating the review of this manuscript and approving it for publication was Claudia Raibulet.

way, the welfare of tourists, tour companies, and society can also be improved. Studies have shown that a tour group composed of strangers can better improve the quality of group members' interactions, and the interaction opportunities may have a significant positive impact on the satisfaction of tourists' travel experience [14]. In low-familiarity groups, tourists in smaller groups may have higher overall satisfaction [15]. With similar interests of tourists in the group, the tour company can more comprehensively cover each individual's needs when customizing the tour itinerary, increasing the overall group satisfaction. Increasing tourist satisfaction with the tour itinerary can in turn provide a reverse incentive to participate in the tour.

Currently, most travel agencies only offer itineraries, not detailed schedules, including attractions, restaurants, and accommodation. Tourists often know little about the destination and have a high degree of autonomy. It is necessary to accurately customize the detailed schedule taking into account the length of stay, travel time, time window of the place, etc. This gives visitors a clear idea of the itinerary.

It is a very complicated process to arrange a tour route for a tour group, which needs to consider the popularity of the place, the favorable rating, the degree of meeting the comprehensive needs of the tourists in the group and other aspects. Therefore, it is very difficult for the tour company to customize the tour itinerary according to the experience. With the development of tourism and computer technology, the research on smart tourism has attracted great attention [16]. At present, a large number of tourism resources and traffic data have emerged on the tourism websites, such as the price, rating, comments, suggested play time, opening time and the transportation time between places, etc. Tourism companies can make full use of the online resources brought by new technologies to assist them in providing customized services of tourism products. These tourism data can quantify the problem of itinerary customization, and help travel companies more accurately and efficiently customize the itinerary that meets the current trend and the needs of tourist groups.

In summary, this paper proposes to customize tourist itineraries with time schedules for tourists based on their needs after dividing them into different groups using online information. Figure 1 shows the tourist group itinerary customization problem, which is divided into two steps. The first part is to gather tourists with similar interests together by using the clustering method according to the needs of tourists. The second part is to obtain online information related to attractions, restaurants, hotels, etc. for model solving. In the objective function, consider the popularity, favorability and overall satisfaction of each visitor in the team, and consider the scheduling problem of custom modeling. The price of the trip is then determined based on the number of tourists in the group.

Finally, the customization model is illustrated by 20 groups of tourists traveling in Xi Shuang Ban Na, China. This article discusses what level of clustering would make visitors most satisfied; It verifies that team customization can be personalized for different groups according to different needs, which can save the cost of tourists; Tourists express a certain level of satisfaction with the itinerary and price.

II. LITERATURE REVIEW

This article first clusters tourists with similar needs, and then tailors the travel itinerary according to the needs of tourists in the group. This paper divides the problem of tourists' travel itinerary customization into two steps: clustering and itinerary customization. This can avoid non-conciliatory conflicts between group tourists, and can also reduce costs from the aspects of society, enterprises and tourists. At the same time, this article clearly obtains the needs of tourists for itinerary customization, fully considering the preferences of tourists. It combines tourist clustering, group travel itinerary planning and tourism customization, which is a new attempt. At present, there are few academic achievements directly related to this paper, but many scholars have developed the application of the tourism industry from one of these aspects. For example, some scholars have studied the application of clustering in the tourism industry; Some scholars have also studied the problem of tourism itinerary planning with heterogeneous needs in groups; In addition, some scholars have carried out research on the issue of recommending travel itineraries according to the needs of tourists. These studies provide good ideas and support for this paper. Therefore, the paper divides the existing relevant research into three parts: user clustering, travel recommendation, and group travel.

A. USER CLUSTERING

User clustering is the aggregation of similar people together, and current research on user clustering in the tourism industry includes several aspects. On the one hand, the tourist market is subdivided into different categories for different tourist characteristics, and then the corresponding marketing strategies are proposed according to the characteristics of each tourist market. D 'Urso et al. proposed the Bagged Clustering method, which is more stable and flexible, and is applied to the tourism market segmentation to lay a foundation for future tourism marketing [17]. Özel and Kozak clustered the tourists according to their eight travel motives and got five different groups of tourists to segment the Turkish cultural tourism market [18]. Otoo et al. clustered the elderly travelers according to their motives and segmented several heterogeneous markets, which is beneficial for future marketing efforts for each segment [19]. On the other hand research on collaborative filtering by finding similar travelers to make travel recommendations for new tourists. Schiaffino and Amandi matched users with other users with similar interests and preferences and combined user profiles and demographic information to recommend tours and packages to tourists, improving the accuracy of recommendations [20]. Jiang et al. used users' photos and texts to elicit their travel preferences and accurately mine similar users based on similarities in travel preferences as a way to personalize travel recommendations [21]. Esmaeili et al. provide tourists with

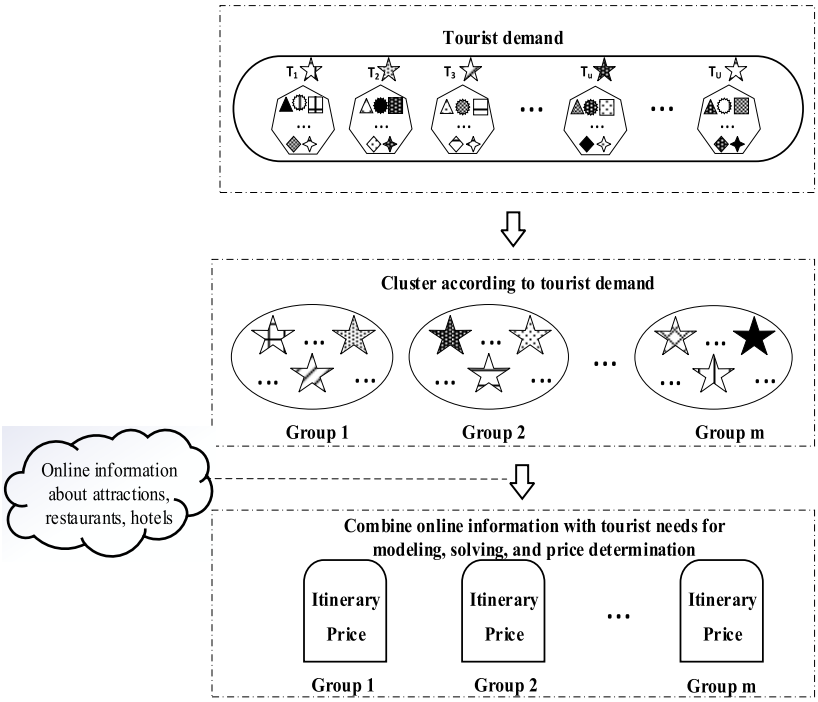

FIGURE 1. Tourist clustering and tour itinerary customization based on heterogeneous needs.

personalized lists of tourist attractions based on similarities in their interests, trust, reputation, relationships, and social communities. This method improves recommendation efficiency and can be used to recommend other products and services in tourism and other social businesses [22]. Some other studies help enterprises manage by clustering tourists. Lee and Han grouped tourists with similar travel itineraries and provided them with services such as tour buses, drivers, and guides, saving tourists' costs [23]. Derek et al. clustered tourists by their activities at the destination, the framework can describe the demographic characteristics of each group and travel patterns, providing specific destination management [24]. Using GPS trajectory data, Liu et al. revealed three micro-temporal and spatial modes of "daytime climbing," "night climbing" and "all-day sightseeing" through cluster analysis, which are of great significance to the tourism management of scenic spots [25].

B. TRAVEL RECOMMENDATION

Not much research has been done on travel customization, but some scholars have researched travel itinerary recommendations based on tourists' personalized needs. Travel recommendations can help tourists reduce information search costs and improve decision-making efficiency, making tourists generate more positive experiences and satisfaction. Rodríguez et al. proposed a travel design model with interactive features considering tourists' needs and regional realities [26]. Chiang and Huang proposed a near-automated travel schedule planning service where users can replace unsatisfactory travel units with specific travel units with better recommendation accuracy [27]. Yuan et al. helped tourists select a destination's travel experience from a huge amount of travel blog data so that tourists can arrange their travel plans efficiently [28]. Kotiloglu et al. first use collaborative filtering to identify a set of optional points of interest, and later use an iterative taboo search algorithm to solve the travel route recommendation problem considering constraints such as available days, opening hours, and budget [29]. Liao and Zheng propose designing day trip routes in a time-dependent stochastic environment that can design more realistic and personalized routes for tourists [30]. Ji et al. not only considered tourists' preferences but also the heterogeneity of the spatial structure of attractions to optimize the tourist experience and help tourists travel more enjoyably [31]. Zheng et al. proposed a multi-objective evolutionary algorithm for tourism route recommendation based on the decomposition of two-stage and Pareto hierarchies, which recommends better distribution and diversity of tourist routes, so that the recommended routes can better meet the personalized needs of tourists [32].

C. GROUP TRAVEL

Most tourists prefer to travel with friends, family or others, i.e., in a group, and it is important to tailor a personalized tour to meet the needs of the group tourists. Some scholars have researched group tourism. Garcia et al. propose "aggregation" and "crossover" strategies, where "aggregation" combines the preferences of everyone in the group into one

preference, so that the system can treat a group of tourists as a single tourist to obtain their needs, while "crossover" ensures that the preferences of all tourists in the group are equally considered [33]. Anagnostopoulos et al. tried to find a route that satisfies all tourists in the group to a certain degree, and proposed that the route that one tourist in the group is satisfied with may be very different from the route that all members of the group are satisfied with [34]. Ruiz-Meza and Montoya-Torres proposed a mathematical model of travel routes that take into account individual preferences and group equity, and minimizes the environmental impact of traffic [35]. Liao et al. propose a "join and diverge" strategy for conflicts in tourist groups, allowing members to have some "separate time" to enjoy their respective trips, but also "together time" to create shared memories" [36]. Kargar and Lin cover the locations that all tourists expect to visit in the tour itinerary within the constraints of guaranteed travel distance, time, and cost, and attempt to maximize equity among tourists, which ensures that all members are motivated to participate in group tours [37].

III. PROBLEM DESCRIPTION

Visitors are clustered according to their individual needs, dividing visitors with similar interests into small groups. Tailoring the itinerary to each visitor in the cluster according to their needs can reduce costs for tourists, travel companies and society. Firstly, the personalized needs of tourists are obtained, then the tourists are clustered and grouped according to their personalized needs. Finally, the tourism itinerary customization model is established and solved by combining the collected online information related to tourism resources and the needs of tourists, and the corresponding price is given. In the problem of tourist clustering and tourism itinerary customization based on heterogeneous needs, a tourist mentioned in this paper refers to one or more people with the same needs. For convenience, note $U = \{1, 2, \ldots, m\}$, $L = \{1, 2, \ldots, 7\}$, $U'' = \{1, 2, \ldots, m''\}$, $F = \{1, 2, 3, 4\}$, $K = \{1, 2, 3\}$, $N_0 = \{0\}$, $N_1 = \{1, 2, \ldots, n_1\}$, $N_2 = \{n_1 + 1, n_1 + 2, \ldots, n_2\}$, $N_3 = \{n_2 + 1, n_2 + 2, \ldots, n_3\}$, $N = \{1, 2, \ldots, n_3\}$, $DAY = \{1, 2, \ldots, day\}$.

A description of the meaning of the symbols involved in describing this tourist clustering and trip customization problem is given below.

$T = \{T_1, T_2, \ldots, T_U\}$: The set of tourists with customized tour needs, where T_u denotes the u_{th} visitor, $u \in U$.

$C = \{C_1, C_2, \ldots, C_7\}$: The set of tourist needs, where C_l denotes the l_{th} need, $l \in L$. C_1 denotes hotels, C_2 denotes restaurants, C_3 denotes travel time, C_4 denotes attraction type, C_5 denotes whether the time is adjustable, C_6 denotes must-see attractions, and C_7 denotes no-go attractions.

$D_u = \{d_{u1}, d_{u2}, \ldots, d_{u7}\}$: The value of each demand, each visitor has a different expectation for each demand. Where d_{ul} denotes the value of demand C_l of the visitor T_u, $u \in U$, $l \in L$. The values may be a single value or a collection of discrete values.

$W = (w_1, w_2, \ldots, w_4)$: The weights correspond to different requirements, here only $C_1 \sim C_4$ requirements have the corresponding weights, the weights are calculated from the visitor's score, w_l indicating the weight of the l requirement, $\sum_{l \in L} w_l = 1$.

$G = \{G_1, G_2, \ldots, G_n\}$: The visitors were clustered into n tourist groups.

$T'' = \{T''_1, T''_2, \ldots, T''_{U''}\}$: Tourists in the same group after clustering, where $T''_{u''}$ denotes the u''_{th} visitor in the group, $u'' \in U''$, $T'' \subseteq T$.

$TIME_t$: Tour time for Day t, which is determined by the tour company, $t \in DAY$.

$V = \{v_0, v_1, v_2, \ldots, v_{n_3}\}$: All locations including attractions, restaurants, hotels, where v_j denotes the j_{th} location, $j \in N \cup N_0$.

$comment_j$: Number of comments for the j_{th} location, $j \in N$.

$score_j$: Score for the j_{th} location, $j \in N$.

$stay_j$: Expected length of stay at the j_{th} location, $j \in N$.

ta_j: Arrival time of the j_{th} location, $j \in N \cup N_0$, where ta_0 indicates the arrival time of the origin/destination.

te_j: Departure time of the j_{th} location, $j \in N \cup N_0$, where te_0 indicates the departure time of the origin/destination.

$traffic_{jj'}$: Drive time from the j_{th} location (including origin/destination) to the j'_{th} location (including origin/destination), $j, j' \in N \cup N_0$.

x_j: Decision variable, indicating whether the j_{th} location is selected, with "1" if selected and "0" otherwise, $j \in N$.

$y^t_{jj'}$: Decision variable, whether day t goes from location j (including origin/destination) to location j' (including origin/destination), "1" if yes, "0" otherwise, $j, j' \in N \cup N_0$, $t \in DAY$.

r_j: To express the constraint that there can be no subloops in the trip customization model does not make any sense, $j \in N \cup N_0$.

$[a_j, b_j]$: Time window constraint for the j location, where the time window for the restaurant is set to be within a fixed time (e.g. set = at noon) and the other locations are time window constrained according to the actual opening and closing times, $j \in N$.

$type_j$: The type to which the j site belongs, $j \in N_1$.

cl_j: The level at which the j location is a restaurant, as determined by the per capita price level of local dining, $j \in N_2$.

hl_j: The level when the j location is a hotel, $j \in N_3$.

hp_j: Prices when the j location is a hotel, $j \in N_3$.

$comment_j$, $score_j$, $stay_j$, $type_j$, cl_j, hl_j, hp_j, are online information resources available on the Platform.

The problem to be solved in this paper is how to cluster tourists according to the above variables, and to customize the travel itinerary with time arrangement for each tourism group.

IV. PRINCIPLES AND METHODS

A method of tourist clustering and tour itinerary customization based on heterogeneous demand is given here, and a

specific description of the basic principles and computational steps of the method is given below.

A. TOURIST SIMILARITY CALCULATION METHOD

To subsequently cluster the tourists, the calculation of similarity between tourists is first performed. Different tourists have different needs for their travel itineraries, and different needs are calculated in different ways, and the degree of similarity among tourists is also different. When tourists customize tourism products, they usually put forward requirements related to hotels, restaurants, travel time, attractions, etc. Here the requirements used to calculate similarity are "hotel," "restaurant," "travel time," "Attractions." Before giving the similarity calculation method for different requirements, we define the function $Q(\mathrm{x}, \mathrm{y})$ to express the degree of similarity between x and y below.

$$Q(\mathrm{x}, \mathrm{y}) = 1 - \frac{|x - y|}{|x - y|_{\max}} \tag{1}$$

The following is the similarity calculation method of different demands of tourists.

1) HOTEL

"Hotel" can be divided into different stars, here "3-star," "4-star," and "5-star." In practice, it is found that the price of a hotel of the same star level may vary greatly. Therefore, the way of "hotel level + hotel price" is adopted here to express the similarity of the demand of different tourists for the hotel. When $C_1 =$ "*hotel*," the value of d_{u1} is the set of two discrete values, $d_{u1} = \{d^1_{u1}, d^2_{u1}\}$, where d^1_{u1} represents the demand of the u_{th} tourist for "hotel level" and d^2_{u1} represents the demand of the u_{th} tourist for "hotel price."

The similarity between the expectations of visitors T_u and visitors $T_{u'}$ for "hotel level" is:

$$sim^1_{uu'} = Q(R(d^1_{u1}), R(d^1_{u'1})) \quad u \in U,\ u' \in U \tag{2}$$

where $R(\cdot)$ is a sorting function whose return value is the number of bits of the variable sorted from lowest to highest in the ordered hierarchy.

The similarity between the expectations of visitors T_u and visitors $T_{u'}$ for "hotel prices" is:

$$sim^2_{uu'} = Q(d^2_{u1}, d^2_{u'1}) \quad u \in U,\ u' \in U \tag{3}$$

In calculating the similarity of tourists' demand for "hotel," the importance of "hotel rating" and "hotel price" is measured. The importance of "hotel level" is set as ρ_1 and the importance of "hotel price" is set as ρ_2 and $\rho_1 + \rho_2 = 1$. The similarity between the demand of tourists T_u and tourists $T_{u'}$ for "hotel" is obtained as follows.

$$S^1_{uu'} = \rho_1 sim^1_{uu'} + \rho_2 sim^2_{uu'} \quad u \in U,\ u' \in U \tag{4}$$

2) CATERING

"Dining" is a hierarchical demand, and we consider classifying restaurants into three levels according to their per capita consumption prices: Level 1, Level 2, and Level 3. The level is to express the demand of different tourists for the "restaurant." "The demand value of d_{u2} is a single value. The similarity between the demand for "dining" of visitors T_u and visitors $T_{u'}$ is:

$$S^2_{uu'} = Q(R(d_{u2}), R(d_{u'2})) \quad u \in U,\ u' \in U \tag{5}$$

where $R(\cdot)$ is a sorting function whose return value is the number of bits of the variable sorted from lowest to highest in the ordered hierarchy.

3) TRAVEL TIME

The value of the travel time d_{u3} is a set form with multiple discrete values, where each discrete value represents a travel date for the tourist. The value of whether the travel time is adjustable d_{u5} is a single value and has only two values, "1" for "adjustable" and "0" for non-adjustable.

$$d_{u5} = \begin{cases} 0, & \textit{the travel time is adjustable} \\ 1, & \textit{the travel time is} \text{ non-adjustable} \end{cases} \quad u \in U \tag{6}$$

When both tourists' travel time can be adjusted, the degree of similarity of tourists' travel time is expressed by the overlap of the two tourists' travel dates; when at least one tourist's travel time is not adjustable, the identical travel time of the two is recorded as "1," otherwise it is "0." Therefore, the similarity between the demand for "travel time" of tourist T_u and tourist $T_{u'}$ is:

$$S^3_{uu'} = \begin{cases} 0, & d_{u3} \neq d_{u'3} \textit{ and } d_{u5} + d_{u'5} \leq 1 \\ 1, & d_{u3} = d_{u'3} \textit{ and } d_{u5} + d_{u'5} \leq 1 \\ \dfrac{d_{u3} \cap d_{u'3}}{d_{u3} \cup d_{u'3}}, & d_{u5} = d_{u'5} = 1 \end{cases} \quad u \in U, \quad u' \in U \tag{7}$$

4) ATTRACTIONS

The value of attraction type d_{u4} is a set with multiple discrete values, and there are five attraction types to choose from, namely "natural scenery," "recreation," "folklore," "food and shopping," "history and culture." Visitors will choose the type of attractions they want to visit. The degree of similarity between two tourists is expressed in terms of the degree of overlap between the types of attractions selected by them. However, before this, it is necessary to check whether there is any conflict between the "must-go scenic spots" and "no-go scenic spots" proposed by tourists. If there is any conflict, the similarity of the demand for "scenic spots" is directly recorded as "0"; If there is no conflict, tde corresponding formula is used to calculate the similarity. The value of "must-see" d_{u6} and the value of "don't-see" d_{u7} are both the set form with multiple discrete values. Visitors can specify multiple places they want to see or not. In summary, the similarity between visitors T_u and visitors $T_{u'}$ for

"Attractions" is:

$$S_{uu'}^{4} = \begin{cases} \dfrac{d_{u4} \cap d_{u'4}}{d_{u4} \cup d_{u'4}}, & d_{u6} \cap d_{u'7} = d_{u'6} \cap d_{u7} = \emptyset \\ 0, & else \end{cases} \quad u \in U, \quad u' \in U \tag{8}$$

5) TOTAL SIMILARITY

The degree of influence of different demands on the total similarity value among tourists varies, so the total similarity value is obtained by considering the combination of the similarity of each demand by tourists with its weight. The calculation formula is as follows:

$$S_{uu'} = \sum_{f=1}^{4} w_f S_{uu'}^{f}, \quad u \in U, \ u' \in U \tag{9}$$

where the value of w_f is calculated from the scores given by all visitors, $f \in F$. The sum of all visitors' scores (on a scale of 1-5) for the importance of each of the above four needs was recorded as $grade_1 \sim grade_4$, whereupon we obtained.

$$w_f = \frac{grade_f}{\sum_{f=1}^{4} grade_f}, \quad f \in F \tag{10}$$

B. VISITOR CLUSTERING GROUPING

Based on the above similarity calculation method the total similarity value between two different visitors can be obtained and the visitor similarity matrix is:

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1m} \\ S_{21} & S_{22} & \cdots & S_{2m} \\ \vdots & \vdots & & \vdots \\ S_{m1} & S_{m2} & \cdots & S_{mm} \end{bmatrix}$$

Based on the similarity matrix, the visitors are clustered using the hierarchical clustering method. The distance between classes is calculated using the maximum distance method. The distance between the two furthest points in the two class clusters is used as the distance between the class clusters.

(1) Calculate the similarity matrix based on the information of the visitors, with each visitor as a class in the initial stage, and calculate the similarity.

(2) Combine the two most similar visitors into one class and continue to calculate similarity with the other classes (using the maximum distance method).

(3) The two categories with the greatest similarity continue to be combined, and so on.

(4) Until the similarity between each class is less than a certain threshold δ, stop merging.

At the end of clustering we get several groups of visitors and the visitors belonging to the same group can be guaranteed to have a total similarity above a certain threshold.

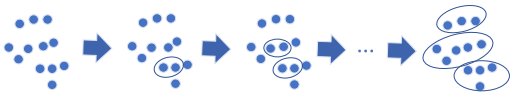

FIGURE 2. Hierarchical clustering process demonstration diagram.

C. CUSTOMIZATION OF TRAVEL ITINERARY

After dividing the tourists into different groups, the tour itinerary needs to be customized for each group of tourists separately. Grouping tourists can reduce the cost and cover the needs of each tourist to a greater extent. In the tourism itinerary customization problem studied in this paper, the objective function of the model needs to be determined first, and then the tourism itinerary modeling is established.

1) DESCRIPTION OF THE OBJECTIVE FUNCTION

During a trip, users want to visit as many places as possible, and the needs of each traveler within the group need to be considered when building a customized itinerary model. In addition to meeting the needs of tourists, the hotness and favorability are also very important, which reflects the popularity of the location and the quality of the tourist experience. So the following three factors are proposed here to be considered in the objective function: hotness, favorability, and tourist satisfaction. The objective function has multiple objectives, which need to be balanced to make the overall objective as optimal as possible.

a: SATISFACTION DEGREE CALCULATION METHOD

When calculating the degree to which a place meets the needs of a tourist group, it is necessary to consider the needs of each tourist in the group. The method of calculating the satisfaction of visitor demand differs for the three categories of locations: attractions, restaurants, and hotels. The degree to which the j_{th} location meets visitor demand is noted as s_j and is calculated as:

$$s_j = \begin{cases} \dfrac{1}{N(T'')} \sum_{u''=1}^{m''} type_j^{u''}, & j \in N_1 \\ \dfrac{1}{N(T'')} \sum_{u''=1}^{m''} Q(R(cl_j), R(d_{u''2})), & j \in N_2 \\ \dfrac{1}{N(T'')} (\rho_1 \sum_{u''=1}^{m''} Q(hp_j, d_{u''1}^{2}) \\ \quad + \rho_2 \sum_{u''=1}^{U''} Q(R(hl_j), R(d_{u''1}^{1}))), & j \in N_3 \end{cases} \tag{11}$$

where $type_j^{u''}$ is calculated as:

$$type_j^{u''} = \begin{cases} 1, & type_j \in d_{u''4} \\ 0, & type_j \notin d_{u''4} \end{cases} \quad j \in N_1, \ u'' \in U'' \tag{12}$$

b: HEAT CALCULATION METHOD

The popularity of a place can be determined by the number of comments on the place. The more comments there are, the

more visitors there are. The number of comments reflects the popularity of the scenic spot. The number of comments differs greatly among the categories of scenic spots, restaurants and hotels. Here, the number of comments is standardized to express the hotness of each location. The hotness of the j_{th} location is noted as h_j, which is calculated as:

$$h_j = \frac{comment^k_{\max} - comment_j}{comment^k_{\max} - comment^k_{\min}} \quad k \in K \tag{13}$$

where $comment^k_{\max}$ indicates the maximum number of comments for the location corresponding to the set N_k and $comment^k_{\min}$ indicates the minimum number of comments for the location corresponding to the set N_k. $comment^k_{\max} = \max\{comment_j | j \in N_k, k \in K\}$, $comment^k_{\min} = \min\{comment_j | j \in N_k, k \in K\}$.

c: FAVORABILITY CALCULATION METHOD

The favorable rating of a location is obtained through the rating of the location, with higher ratings indicating a better experience for the visitor. The ratings between the three categories of attractions, restaurants, and hotels are also represented separately here, with the favorable rating of the j_{th} location noted as p_j, and the specific relationship between the ratings and favorable ratings as:

$$p_j = \frac{score^k_{\max} - score_j}{score^k_{\max} - score^k_{\min}}, \quad k \in K \tag{14}$$

where $score^k_{\max}$ indicates the highest rating for the location corresponding to the set N_k and $score^k_{\min}$ indicates the lowest rating for the location corresponding to the set. $score^k_{\max} = \max\{score_j | j \in N_k, k \in K\}$, $score^k_{\min} = \min\{score_j | j \in N_k, k \in K\}$.

2) MODELING OF TRAVEL ITINERARY CUSTOMIZATION

After clarifying the objective function in the model, the subsequent tour itinerary customization model will be performed. In this model, the starting point and the ending point are the same and fixed. Since the first day of the tour starts at the starting point and each subsequent day starts at the hotel, the departure time of the starting point and the hotel is set to "0." Each day's tour time is constrained accordingly. The dining time of the restaurant needs to be set with a corresponding time window. Each attraction has a time window according to its opening and closing time. The hotel is used as both the end of one day and the start of the next day, so there is no need to set a time window.

The model of this travel itinerary customization problem is represented below. Where α, β, γ indicates the degree of importance of hotness, favorability, and satisfaction.

$$\max z = \sum_{j \in N} (\alpha \cdot h_j + \beta \cdot p_j + \gamma \cdot s_j) \cdot x_j \tag{15}$$

$$\sum_{j \in N} y^1_{oj} = 1 \tag{16}$$

$$\sum_{j \in N} y^{day}_{jo} = 1 \tag{17}$$

$$\sum_{t \in DAY} \sum_{j' \in N} y^t_{jj'} \le 1, \quad j \in N \cup N_0 \tag{18}$$

$$\sum_{j \in N_0 \cup N} y^t_{jj'} = \sum_{j'' \in N_0 \cup N} y^t_{j'j''}, \quad j' \in N_1 \cup N_2, \ t \in DAY \tag{19}$$

$$\sum_{j \in N} y^t_{jj'} = \sum_{j'' \in N} y^{t+1}_{j'j''}, \quad j' \in N_3, \ t \in [1, day-1] \tag{20}$$

$$y^t_{jj'} + y^t_{j'j} \le 1, \quad j, j' \in N, \ t \in DAY \tag{21}$$

$$r_j - r_{j'} + n_3 \cdot y^t_{jj'} \le n_3 - 1, \quad j, j' \in N, \ t \in DAY \tag{22}$$

$$y^t_{jj'} = 0, \quad j, j' \in N_3 \cup N_0, \ j = j', \ t \in DAY \tag{23}$$

$$te_j + traffic_{jj'} - M \cdot (1 - y^t_{jj'}) \le ta_{j'}, \quad j, j' \in N, \ t \in DAY \tag{24}$$

$$ta_{j'} \le te_j + traffic_{jj'} + M \cdot (1 - y^t_{jj'}), \quad j, j' \in N, \ t \in DAY \tag{25}$$

$$ta_j + stay_j - M \cdot (1 - x_j) \le te_j, \quad j \in N \tag{26}$$

$$te_j \le ta_j + stay_j + M \cdot (1 - x_j), \quad j \in N \tag{27}$$

$$te_j = 0, \quad j \in N_3 \cup N_0 \tag{28}$$

$$ta_j \ge a_j, \quad j \in N_1 \cup N_2 \tag{29}$$

$$te_j \le b_j, \quad j \in N_1 \cup N_2 \tag{30}$$

$$\sum_{j \in N} \sum_{j' \in N} y^t_{jj'} \cdot (stay_j + traffic_{jj'}) \le TIME_t, \quad t \in DAY \tag{31}$$

Equation (15) indicates that the objective function combines three aspects of location hotness, favorability, and the degree of satisfying visitor demand; equation (16) and equation (17) specify the start and end points of the route; equation (18) indicates that each location can be selected at most once; equation (19) specifies the connectivity of attractions and restaurants within the same day; equation (20) indicates that the previous day must end with a hotel and the next day begins with a hotel, specifying hotel connectivity; equation (21) indicates that there can be no returning streams; equation (22) indicates that there can be no sub-loops; equation (23) specifies that one must travel from one point to another; equation (24) and equation (25) are used to calculate the arrival time for location j; equation (26) and equation (27) are used to calculate the departure time of location j; equation (28) specifies that the departure time from the starting point and the hotel is "0"; equation (29) and equation (30) are the time window constraints; equation (31) constrains that the daily travel time must not exceed a certain value.

TABLE 1. Tourists and demand information.

ID	the number of people	d_{u1}^{l}	d_{u2}^{l}	d_{u2}	d_{u3}	d_{u5}	d_{u4}	d_{u6}	d_{u7}	Expected price
1	2	4	200	2	2.9-2.12	yes	2,3,4,5			2000
2	3	4	200	2	2.10-2.12	yes	1,2,3,4,5			1900
3	3	4	250	3	2.9-2.12	yes	1,2,3,4,5			1800
4	2	3	150	2	2.9-2.12	yes	1,2,3,4,5			2100
5	4	4	250	1	2.9-2.12	yes	1,2,3,4,5	v_{10}		2000
6	4	5	300	2	2.9-2.12	yes	1,2,3,4,5			2200
7	2	4	200	2	2.8-2.11	yes	1,2,3,4,5		v_{13}	1900
8	2	4	200	2	2.9-2.11	yes	1,2,3,4,5			1600
9	5	4	200	2	2.9-2.12	yes	1,2,3,4			1900
10	3	5	200	3	2.7-2.10	no	2,3,4,5			2000
11	3	3	180	3	2.15-2.18	yes	1,2,3,4,5			2000
12	3	4	180	3	2.15-2.18	yes	1,2,3,4,5			1900
13	4	3	180	2	2.15-2.18	yes	1,2,3,4,5			1850
14	4	3	150	2	2.15-2.18	yes	1,2,3,4,5			2000
15	4	3	150	3	2.16-2.19	yes	1,2,3,4,5			2200
16	4	3	150	3	2.16-2.18	yes	1,2,3,4,5		v_{10}	1600
17	2	3	150	3	2.15-2.18	yes	2,3,4,5			2000
18	2	4	150	3	2.15-2.18	yes	1,2,3,4,5			1800
19	2	3	150	3	2.15-2.18	yes	1,2,3,4,5			1900
20	2	5	300	3	2.13-2.16	no	2,3,4			2300

D. DETERMINATION OF TOUR PRICE

Usually, the more people in the group, the lower the price. The price of the tour is related to the number of tourists willing to participate in the tour at the lowest price with it, if it is within the number of people, the price of the tour product will be set to that lowest price and stop; otherwise, take the next lowest price corresponding to the range of the number of participants; Step two: compare the actual number of tourists willing to participate in the tour at the next lowest price with it. The actual number of tourists willing to join the tour at the second lowest price is compared with it, and if it is within the number range, the price of the tour product is set as that second lowest price and stopped; otherwise, step two is repeated until a feasible price is found.

To obtain the group information of tourists and customize the subsequent itinerary, we first obtained the needs of tourists. In January 2023, we selected 20 tourists who wanted to visit Xishuangbanna from local travel agencies in Shenyang to participate in this study, and collected information on tourists and their needs (TABLE 1). In general, a family or several friends traveling together have the same needs, so they are called ''a tourist'' here. The price determination problem involves the number of participants in the tour, so it is necessary to obtain the number of people included in each tourist. The unit of ''hotel price'' is ''yuan/person/night.'' The number of participants, but whether the tourists join the group tour or not is related to the price. The discount pricing method is used to determine the price of group tourism. Different number of participants correspond to different discounts. The price discount function is shown in Formula (32). According to the tourists' expected tour price to determine the total number of participants under different discount prices. Step one: take the lowest price corresponding to the range of the number of participants, compare the actual

$$Discount(p) = \begin{cases} dis_1, & 0 \le p < p_1 \\ dis_2, & p_1 \le p < p_2 \\ dis_3, & p_2 \le p < p_3 \\ dis_4, & p_3 \le p < p_4 \\ dis_5, & p \ge p_4 \end{cases} \tag{32}$$

V. EXPERIMENT

The proposed model was coded using GAMS (General Algebraic Modeling System) and solved using solver CPLEX on a personal computer with 32GB of running memory,

TABLE 2. Scoring the importance of visitor need.

ID	hotel	catering	travel time	attractions
1	5	5	3	3
2	5	5	3	3
3	5	5	3	3
4	5	5	3	3
5	4	5	3	3
6	4	5	3	4
7	4	5	3	4
8	4	5	3	4
9	4	5	3	4
10	4	4	2	4
11	3	4	2	4
12	3	4	2	5
13	3	4	2	5
14	3	4	2	5
15	3	4	2	5
16	3	3	4	5
17	2	3	4	2
18	2	3	5	1
19	2	3	1	2
20	2	2	1	2

AMDRyzen 7 5800X@3.8GHz, 512G SSD and 64-bit Windows OS and programmed for a maximum execution time of two hours.

A. DATA SETTINGS

Restaurant level is determined according to the per capita consumption price of local restaurants in Xi Shuang Ban Na, with the per capita consumption price of first-class restaurants below RMB 60, second-class restaurants between RMB 60 and 90, and third-class restaurants above RMB 90. The tourism types ''1,'' ''2,'' ''3,'' ''4'' and ''5'' are ''natural scenery,'' ''recreation,'' ''folklore,'' ''food and shopping,'' ''History and Culture.'' Each tourist can propose the attractions that he/she must visit or does not want to visit. Expected travel price refers to the price at which tourists expect to join a group tour.

Tourist similarity is not only related to the similarity degree among the four requirements of ''hotel,'' ''catering,'' ''travel time'' and ''attractions,'' but also related to its importance, which is scored by tourists (full score: 5) (TABLE 2) and calculated according to formula (10).

After obtaining the tourist demand and the weight of each demand, the total similarity of tourists can be calculated and tourists can be grouped. According to the clustering situation, the grouped situation can be obtained and the travel itinerary can be customized for the tourist group. To customize the itinerary to meet the tourists' needs, we need to obtain online information related to tourism resources. The prices, reviews and ratings of attractions and hotels are obtained from ''Go.com,'' the tour time of attractions is obtained from ''Ctrip,'' and the prices, reviews and ratings of restaurants are obtained from ''MeiTuan.'' The latitude and longitude of the locations are called by using excel to the latitude and longitude query interface of Baidu map API, and the full name of the location should be entered when calling to get the online information of the attractions, restaurants and hotels (TABLE 3).

The driving time between the two locations is calculated using Baidu Map API, and the batch route calculation interface of Baidu Map is called through Python language. The longitude and latitude of the two locations are input, and then the driving time between the two locations is obtained (TABLE 4).

B. IMPACT OF GROUP SIZE ON TOURISTS' SATISFACTION

According to Table 2 and Equation (10), the weights of ''hotel,'' ''catering,'' ''travel time'' and ''attractions'' are 0.25, 0.30, 0.19 and 0.26. Combining the weights with tourists' demand, the similarity matrix of tourists is calculated from equation (2) to equation (9) (TABLE 5).

TABLE 3. Basic information on locations.

No.	longitudes	longitude	price	comment	score	type	level	playtime	Time window
v_0	100.773531	21.97709497	/	/	/	0	/	/	/
v_1	101.262951	21.93160414	75	22293	4.60	1	/	2h	[8:00-18:00]
v_2	100.9018419	22.02432247	28	16179	4.80	1	/	2h	[8:30-17:00]
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
v_{27}	100.827892	22.01342459	42.00	109.00	4.90	5	/	2.00	[15:00-21:00]
v_{28}	100.825066	22.01403803	60.00	112.00	5.00	5	/	2.00	[15:00-21:00]
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
v_{32}	100.828792	22.01370999	75	4293	4.40	catering	2	/	[11:30-14:30]
v_{33}	100.8017	22.03104696	63	347	4.60	catering	2	/	[11:30-14:30]
v_{34}	100.7966121	21.9949973	92	2281	4.40	catering	1	/	[11:30-14:30]
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
v_{86}	100.8129437	21.98456416	525	628	4.80	hotel	3	/	/
v_{87}	100.8243393	22.01536903	445.5	1473	4.50	hotel	4	/	/
v_{88}	100.83051	22.01443995	525	531	4.70	hotel	4	/	/

TABLE 4. Driving time between locations (unit: hour).

No.	0	1	2	3	4	⋯	78	⋯	88
0	0.00	0.73	0.18	0.07	1.50	⋯	0.10	⋯	0.10
1	0.73	0.00	0.60	0.68	0.83	⋯	0.65	⋯	0.65
2	0.18	0.60	0.00	0.15	1.40	⋯	0.10	⋯	0.10
3	0.07	0.70	0.15	0.00	1.50	⋯	0.07	⋯	0.07
4	1.50	0.83	1.40	1.50	0.00	⋯	1.40	⋯	1.40
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
78	0.10	0.65	0.10	0.07	1.40	⋯	0.00	⋯	0.03
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
88	0.10	0.65	0.10	0.07	1.40	⋯	0.03	⋯	0.00

After obtaining the visitor similarity matrix, the clustering results were obtained by clustering the visitors using the hierarchical clustering method (Figure 3).

This section discusses the variation in tourist satisfaction under different groupings. With the same number of people, the more groupings and the smaller the number of people in each group, the more personalized and expensive the tour route is likely to be. Too much or too little grouping may lead to tourist dissatisfaction, so this section explores to what extent the grouping approach makes tourists most satisfied. When everyone is grouped into one group, the minimum similarity of tourists within the group is 0.3; when every-

TABLE 5. Visitor similarity matrix.

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	0.9	0.76	0.84	0.76	0.8	0.87	0.9	0.9	0.6	0.53	0.59	0.68	0.65	0.5	0.5	0.56	0.57	0.5	0.45
2	0.9	1	0.76	0.85	0.76	0.81	0.89	0.91	0.9	0.55	0.58	0.64	0.73	0.71	0.56	0.56	0.5	0.62	0.56	0.41
3	0.76	0.76	1	0.7	0.7	0.75	0.73	0.76	0.76	0.65	0.69	0.75	0.54	0.51	0.66	0.66	0.61	0.73	0.66	0.6
4	0.84	0.85	0.7	1	0.7	0.75	0.82	0.85	0.84	0.44	0.64	0.57	0.79	0.81	0.66	0.66	0.61	0.6	0.66	0.31
5	0.76	0.76	0.7	0.7	1	0.75	0.73	0.76	0.76	0.35	0.39	0.45	0.54	0.51	0.36	0.36	0.31	0.43	0.36	0.3
6	0.8	0.81	0.75	0.75	0.75	1	0.78	0.81	0.8	0.52	0.44	0.5	0.59	0.56	0.41	0.41	0.36	0.47	0.41	0.56
7	0.87	0.89	0.73	0.82	0.73	0.78	1	0.95	0.87	0.55	0.58	0.64	0.73	0.71	0.56	0.56	0.5	0.62	0.56	0.41
8	0.9	0.91	0.76	0.85	0.76	0.81	0.95	1	0.9	0.55	0.58	0.64	0.73	0.71	0.56	0.56	0.5	0.62	0.56	0.41
9	0.9	0.9	0.76	0.84	0.76	0.8	0.87	0.9	1	0.49	0.53	0.59	0.68	0.65	0.5	0.5	0.45	0.57	0.5	0.45
10	0.6	0.55	0.65	0.44	0.35	0.52	0.55	0.55	0.49	1	0.62	0.68	0.47	0.44	0.59	0.59	0.64	0.65	0.59	0.66
11	0.53	0.58	0.69	0.64	0.39	0.44	0.58	0.58	0.53	0.62	1	0.94	0.85	0.83	0.9	0.93	0.92	0.91	0.98	0.48
12	0.59	0.64	0.75	0.57	0.45	0.5	0.64	0.64	0.59	0.68	0.94	1	0.79	0.76	0.84	0.87	0.86	0.98	0.91	0.54
13	0.68	0.73	0.54	0.79	0.54	0.59	0.73	0.73	0.68	0.47	0.85	0.79	1	0.98	0.75	0.78	0.77	0.76	0.83	0.33
14	0.65	0.71	0.51	0.81	0.51	0.56	0.71	0.71	0.65	0.44	0.83	0.76	0.98	1	0.77	0.8	0.8	0.79	0.85	0.31
15	0.5	0.56	0.66	0.66	0.36	0.41	0.56	0.56	0.5	0.59	0.9	0.84	0.75	0.77	1	0.95	0.87	0.86	0.92	0.46
16	0.5	0.56	0.66	0.66	0.36	0.41	0.56	0.56	0.5	0.59	0.93	0.87	0.78	0.8	0.95	1	0.9	0.89	0.95	0.46
17	0.56	0.5	0.61	0.61	0.31	0.36	0.5	0.5	0.45	0.64	0.92	0.86	0.77	0.8	0.91	0.9	1	0.89	0.95	0.5
18	0.57	0.62	0.73	0.6	0.43	0.47	0.62	0.62	0.57	0.65	0.91	0.98	0.76	0.79	0.86	0.89	0.89	1	0.94	0.52
19	0.5	0.56	0.66	0.66	0.36	0.41	0.56	0.56	0.5	0.59	0.98	0.91	0.83	0.85	0.92	0.95	0.95	0.94	1	0.46
20	0.45	0.41	0.6	0.31	0.3	0.56	0.41	0.41	0.45	0.66	0.48	0.54	0.33	0.31	0.46	0.61	0.5	0.52	0.46	1

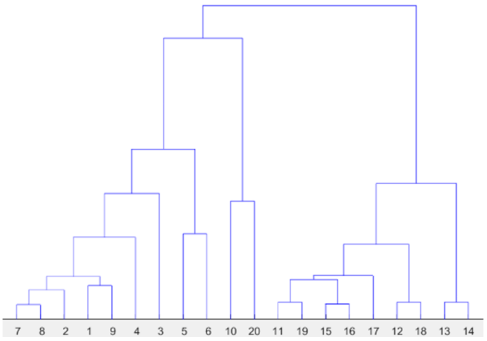

FIGURE 3. Visitor clustering map.

one is grouped into two groups, the minimum similarity of tourists within the group is 0.7, 0.3; when everyone is grouped into three groups, the minimum similarity of tourists within the group is 0.7, 0.75, 0.66; and when everyone is grouped into four groups, the minimum similarity of tourists within the group is 0.7, 0.7, 0.75 and 0.66. It can be seen that the more groups there are, the higher the similarity of tourists within the group will be. This section investigates the satisfaction values of 20 tourists for the route, the price, the number of people, and the total satisfaction value for different grouping cases. For the same number of people, as the number of groups increases, the route becomes more and more personalized, and the satisfaction of tourists to the route gradually increases; the satisfaction of tourists to the price gradually decreases; the satisfaction of tourists to the number of tourists gradually increases. The total satisfaction was affected by route, price and number of tourists. The total satisfaction showed a trend of increasing first and then decreasing. When the tourists were divided into three groups, the total satisfaction reached the highest value. In practice, tourists should be divided into several groups according to the similarity, price and number of group tourists.

C. ANALYSIS OF ITINERARY CUSTOMIZATION RESULTS FOR DIFFERENT TOURIST GROUPS

Based on the discussion about grouping, the grouping with the highest satisfaction in this case was chosen to discuss the results of travel itinerary customization with different needs, that is, participants were divided into 3 groups. According to the actual attention of tourists to popularity, praise and satisfaction, experts used the analytic hierarchy method to obtain the weights: $\alpha = 0.4$, $\beta = 0.3$, and $\gamma = 0.3$. Combined with the information about tourists' needs, it was decided to customize the itinerary for 4 days, and the maximum tour time per day was set to 4 hours, 10 hours, 10 hours, and 4 hours. Formula (14)∼Formula (30) is used to model the itinerary customization problem. The itinerary of the group of two tourists (TABLE 6) and the route view (Figure 5, Figure 6 and Figure 7) are obtained. In the figure, red, green, blue and

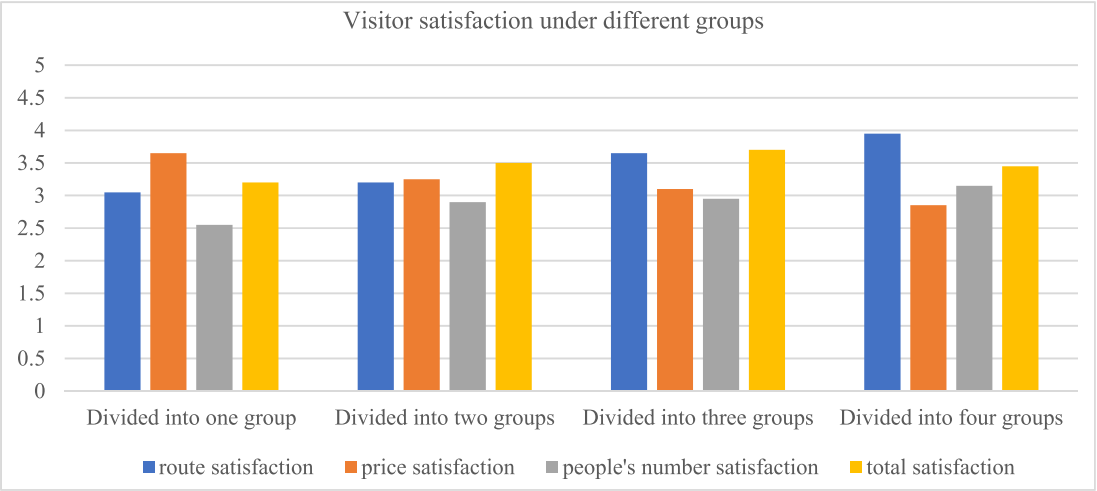

FIGURE 4. Visitor satisfaction under different groups.

TABLE 6. Results of tour itinerary customization for different tourist groups.

Visitor groups		Specific itinerary
Visitor groups1	Day1	[14:00]v1—[14:15]v31—[15:34]v23—[16:38]v18—[17:40]v70
	Day2	[8:00]v70—[8:18]v10—[11:38]v45—[13:43]v3—[15:49]v28—[17:49]v62
	Day3	[8:00]v62—[8:39]v2—[11:08]v16—[12:25]v32—[14:28]v9—[16:29]v30—[17:31]v65
	Day4	[8:00]v65—[9:12]v26—[10:16]v7—[11:58]v1
Visitor groups2	Day1	[14:00]v1—[14:22]v22—[16:22]v16—[17:40]v62
	Day2	[8:00]v62—[8:02]v18—[9:21]v31—[10:40]v23—[11:45]v45—[13:50]v3—[15:56]v28—[17:57]v62
	Day3	[8:00]v62—[8:19]v10—[11:38]v57—[13:40]v11—[15:45]v29—[17:48]v65
	Day4	[8:00]v65—[8:03]v9—[10:04]v30—[11:08]v1
Visitor groups3	Day1	[14:00]v1—[14:04]v23—[15:09]v18—[16:29]v31—[17:48]v63
	Day2	[8:00]v63—[8:02]v16—[9:37]v19—[11:33]v57—[13:40]v26—[15:45]v7—[17:34]v62
	Day3	[8:00]v62—[8:06]v3—[10:14]v9—[12:14]v44—[14:20]v30—[15:39]v22—[17:57]v65
	Day4	[8:00]v65—[8:04]v11—[10:07]v1

purple represent the tour routes of the first, second, third and fourth days respectively.

The customization of the above two tourist groups is a four-day tour to Xishuangbanna, and the tourists have slightly different demands. Under the same parameter Settings and calculation time, the two kinds of tour customization results are different. Both tour itineraries are for 17 locations, 13 of which are the same and 4 are different. The degree of difference between the two itineraries reached 24%, which shows that this itinerary customization method can customize a personalized tour itinerary according to the different needs of tourists.

D. VISITOR ACCEPTANCE ANALYSIS

The price discount function here is:

$$Discount(p) = \begin{cases} 100\%, & 0 \le p < 5 \\ 95\%, & 5 \le p < 10 \\ 90\%, & 10 \le p < 15 \\ 85\%, & 15 \le p < 20 \\ 80\%, & p \ge 20 \end{cases}$$

The corresponding prices for the three groups of group tours at different numbers of people are obtained from the price discount function.

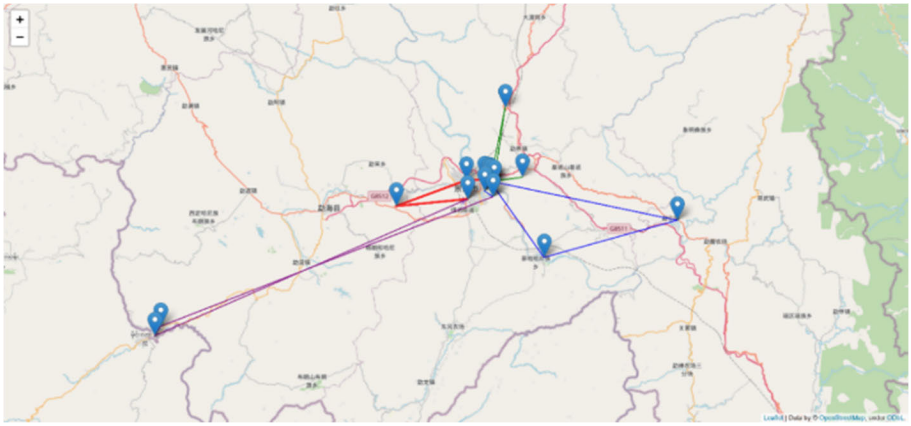

FIGURE 5. Visitor group 1 tour route map.

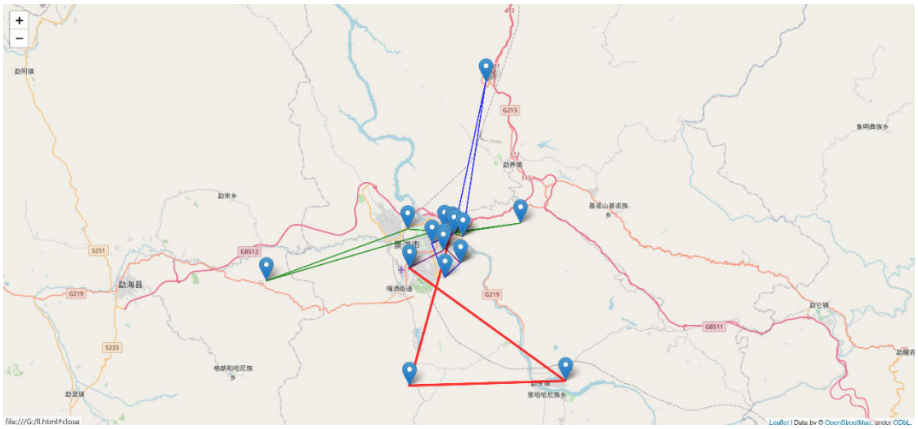

FIGURE 6. Visitor group 2 tour route map.

TABLE 7. Prices for different groups.

Group 1			Group 2			Group 3		
rebates	price	number of people	rebates	price	number of people	rebates	price	number of people
100%	2340	0	100%	2286	0	100%	2445	0
95%	2223	0	95%	2172	4	95%	2323	0
90%	2106	4	90%	2057	6	90%	2200	2
85%	1989	12	85%	1943	13	85%	2078	2
80%	1872	22	80%	1829	22	80%	1956	5

In group 1, according to the expected price of each tourist, 22 people accept the price of 1872, which has already reached the number required for an 80% discount, so the group can form a group at the price of 1872; Similarly, in group 2, it is possible to form a group for 1829; In group 3, only the number of people in the original price of 2445 reached the number of people required for the discount, so the group was formed at the price of 2445.

Generally speaking, whether tourists are willing to participate in group tours is related to the route and price after the group. However, in practice, some tourists do not have a clear and clear understanding of the route and price, so here we provide tourists with individual travel routes and prices (TABLE 8), as a reference, to help tourists determine whether they are willing to participate in group travel.

TABLE 8. Individual itinerary and price.

ID	Itinerary	price
1	Day1: [14:00] v1 → [14:05] v23 → [15:11] v29 → [17:17] v70 Day2: [8:00] v70 → [8:21] v31 → [9:39] v9 → [11:42] v45 → [13:59] v16 → [15:00] v22 → [17:18] v65 Day3: [8:00] v65 → [8:04] v11 → [10:07] v25 → [12:07] v32 → [14:09] v18 → [15:10] v28 → [17:10] v62 Day4: [8:00] v62 → [8:03] v30 → [9:20] v21 → [11:40] v1	2510
2	Day1: [14:00]v1 → [14:05]v23 → [15:23]v31 → [16:41]v30 → [17:43]v65 Day2: [8:00] v65 → [8:03] v9 → [10:11] v3 → [12:17] v45 → [14:19] v18 → [15:20] v28 → [17:20] v62 Day3: [8:00] v62 → [8:17] v22 → [10:18] v16 → [11:40] v1	1734
3	Day1: [14:00] v1 → [14:05] v22 → [16:23] v16 → [17:41] v65 Day2: [8:00] v65 → [8:03] v9 → [10:04] v11 → [12:07] v44 → [14:13] v3 → [16:19] v62 Day3: [8:00] v62 → [8:02] v18 → [9:06] v29 → [11:09] v30 → [14:11] v57 → [14:13] v70 Day4: [8:00] v70 → [8:21] v31 → [9:40] v23 → [10:44] v1	2258
4	Day1: [14:00] v1 → [14:11] v3 → [16:19] v30 → [17:49] v79 Day2: [8:00] v79 → [8:45] v26 → [9:46] v12 → [12:29] v56 → [15:14] v7 → [16:41] v31 → [18:00] v65 Day3: [8:00] v65 → [8:03] v29 → [10:06] v9 → [12:09] v45 → [14:11] v18 → [15:13] v27 → [17:14] v62 Day4: [8:00] v62 → [8:17] v22 → [10:38] v23 → [11:42] v1	2274
5	Day1: [14:00] v1 → [15:12] v7 → [15:45] v26 → [17:57] v76 Day2: [8:00] v76 → [8:02] v9 → [10:20] v31 → [11:38] v54 → [13:46] v3 → [15:55] v4 → [17:59] v66 Day3: [8:00] v66 → [8:23] v10 → [11:42] v45 → [13:42] v25 → [15:43] v28 → [17:43] v62 Day4: [8:00]v62 → [8:01]v18 → [9:05]v23 → [10:26]v16 → [11:48]v1	2321
6	Day1: [14:00] v1 → [14:22] v16 → [15:43] v29 → [17:46] v65 Day2: [8:00]v65 →[8:04]v19 →[10:06]v4 →[12:09]v45 →[14:11]v18 →[15:12]v28 →[17:12]v62 Day3: [8:00] v62 → [8:02] v9 → [10:10] v3 → [12:16] v32 → [14:33] v22 → [16:51] v63 Day4: [8:00]v63 → [8:02]v30 → [9:20]v31 → [10:39]v23 → [11:43]v1	2353
7	Day1: [14:00]v1 → [14:05]v23 → [15:09]v30 → [16:27]v31 → [17:47]v62 Day2: [8:00] v62 → [8:02] v9 → [10:06] v29 → [12:12] v32 → [14:14] v18 → [15:16] v27 → [17:17] v70 Day3: [8:00] v70 → [8:17] v22 → [10:18] v16 → [11:35] v45 → [13:40] v3 → [15:46] v28 → [17:49] v65 Day4: [8:00] v65 → [9:12] v26 → [10:16] v7 → [11:58] v1	2401
8	Day1: [14:00]v1 → [14:05]v23 → [15:23]v31 → [16:41]v30 → [17:43]v65 Day2: [8:00] v65 → [8:03] v9 → [10:11] v3 → [12:17] v45 → [14:19] v18 → [15:20] v28 → [17:20] v62 Day3: [8:00] v62 → [8:17] v22 → [10:18] v16 → [11:40] v1	1734
9	Day1: [14:00] v1 → [14:11] v3 → [15:21] v23 → [17:25] v62 Day2: [8:00] v62 → [8:03] v11 → [10:04] v9 → [12:06] v34 → [14:24] v21 → [16:42] v65 Day3: [8:00]v65 →[9:12]v26 →[10:16]v7 →[12:04]v45 →[14:06]v18 →[15:09]v4 →[17:13]v70 Day4: [8:00] v70 → [8:17] v16 → [9:18] v22 → [11:40] v1	2393
10	Day1: [14:00] v1 → [14:05] v23 → [15:23] v31 → [16:42] v66 Day2: [8:00]v66 →[8:04]v11 →[10:05]v9 →[12:08]v44 →[14:11]v30 →[15:14]v29 →[17:17]v65 Day3: [8:00] v65 → [8:04] v19 → [10:07] v25 → [12:08] v57 → [14:09] v18 → [15:10] v28 → [17:10] v62 Day4: [8:00] v62 → [8:17] v22 → [10:18] v16 → [11:40] v1	2686
11	Day1: [14:00] v1 → [15:12] v7 → [15:45] v26 → [17:57] v65 Day2: [8:00] v65 → [8:03] v9 → [10:06] v18 → [11:10] v23 → [12:15] v44 → [14:32] v16 → [15:33] v22 → [17:50] v70 Day3: [8:00] v70 → [8:03] v19 → [10:10] v3 → [12:16] v57 → [14:22] v29 → [16:40] v31 → [18:00] v62	2545

TABLE 8. *(Continued.)* **Individual itinerary and price.**

	Day4: [8:00]v62 → [8:03]v4 → [10:06]v30 → [11:10]v1	
12	Day1: [14:00] v1 → [14:06] v18 → [15:09] v30 → [16:13] v23 → [17:17] v62 Day2: [8:00] v62 → [8:02] v9 → [10:21] v16 → [11:38] v32 → [13:44] v3 → [15:50] v28 → [17:51] v70 Day3: [8:00] v70 → [8:03] v14 → [11:35] v57 → [13:55] v31 → [15:53] v7 → [17:41] v65 Day 4: [8:00] v65 → [8:43] v2 → [11:27] v1	2461
13	Day1: [14:00]v1 → [14:05]v23 → [15:09]v30 → [16:12]v18 → [17:13]v62 Day2: [8:00] v62 → [8:02] v9 → [10:20] v31 → [11:40] v45 → [13:23] v3 → [15:51] v28 → [17:52] v70 Day3: [8:00] v70 → [8:17] v22 → [10:18] v16 → [11:36] v38 → [13:39] v4 → [15:41] v11 → [17:44] v65 Day 4: [8:00] v65 → [8:23] v10 → [11:47] v1	2214
14	Day1: [14:00] v1 → [14:11] v3 → [16:19] v30 → [17:22] v76 Day2: [8:00] v76 → [9:18] v7 → [12:05] v32 → [14:07] v18 → [15:23] v21 → [17:41] v65 Day3: [8:00]v65 →[8:03]v29 →[10:06]v9 →[12:09]v45 →[14:26]v16 →[15:43]v28 →[17:43]v62 Day4: [8:00] v62 → [8:17] v22 → [10:38] v23 → [11:42] v1	2497
15	Day1: [14:00] v1 → [15:12] v7 → [16:45] v26 → [17:57] v65 Day2: [8:00] v65 → [8:18] v22 → [10:19] v16 → [11:36] v44 → [13:42] v3 → [15:50] v9 → [17:51] v72 Day3: [8:00]v72 →[8:32]v15 →[10:46]v31 →[12:06]v57 →[14:07]v18 →[15:08]v28 →[17:08]v62 Day4: [8:00] v62 → [8:15] v21 → [10:34] v23 → [11:38] v1	2305
16	Day1: [14:00] v1 → [14:22] v22 → [16:42] v30 → [17:45] v62 Day2: [8:00] v62 → [8:17] v16 → [9:38] v23 → [10:42] v18 → [11:43] v57 → [13:49] v3 → [15:57] v9 → [18:00] v65 Day3: [8:00] v65 → [9:18] v7 → [10:45] v31 → [12:00] v1	1852
17	Day1: [14:00] v1 → [14:22] v22 → [16:40] v18 → [17:42] v72 Day2: [8:00] v72 → [8:01] v9 → [10:02] v11 → [12:05] v44 → [14:25] v31 → [15:45] v27 → [17:46] v62 Day3: [8:00] v62 → [8:04] v23 → [9:08] v30 → [10:11] v25 → [12:12] v57 → [14:14] v19 → [16:18] v65 Day4: [8:00] v65 → [8:18] v16 → [9:39] v29 → [11:43] v1	2424
18	Day1: [14:00] v1 → [14:04] v30 → [15:08] v23 → [16:29] v16 → [17:46] v62 Day2: [8:00]v62 →[9:12]v12 →[11:15]v7 →[12:30]v56 →[15:04]v28 →[17:04]v70 Day3: [8:00]v70 →[8:06]v3 →[10:14]v11 →[12:17]v57 →[14:37]v31 →[15:57]v27 →[18:00]v65 Day4: [8:00]v65 → [8:04]v18 → [9:07]v9 → [11:11]v1	2501
19	Day1: [14:00] v1 → [15:12] v7 → [15:45] v26 → [17:57] v65 Day2: [8:00] v65 → [8:18] v22 → [10:19] v16 → [11:36] v44 → [13:42] v3 → [15:50] v9 → [17:51] v72 Day3: [8:00]v72 →[8:32]v15 →[10:46]v31 →[12:06]v57 →[14:07]v18 →[15:08]v28 →[17:08]v62 Day4: [8:00] v62 → [8:15] v21 → [10:34] v23 → [11:38] v1	2305
20	Day1: [14:00] v1 → [14:22] v16 → [15:42] v11 → [17:45] v65 Day2: [8:00] v65 → [9:18] v7 → [9:51] v26 → [12:03] v57 → [14:06] v9 → [16:10] v23 → [17:14] v62 Day3: [8:00]v62 →[8:02]v19 →[10:21]v31 →[11:41]v44 →[13:41]v25 →[15:41]v27 →[17:44]v63 Day4: [8:00] v63 → [8:18] v22 → [10:36] v18 → [11:41] v1	2584

To determine whether our method meets the needs of tourists in the group for travel routes, and whether the group travel model wins tourists in terms of price, we surveyed 20 tourists who were satisfied with the route and price given the tourist's individual travel route and price, and whether they would like to participate in the group tour rather than the individual tour. According to the survey results, the satisfaction rate of tourists with group tour routes reached 3.8 (out of 5), 4.2 (out of 5) on group tour prices, and 17 out of 20 tourists chose to participate in group tours, bringing the total number of participants to 50. It can be seen that compared with individual tourism, tourists are more accepting of the mode of

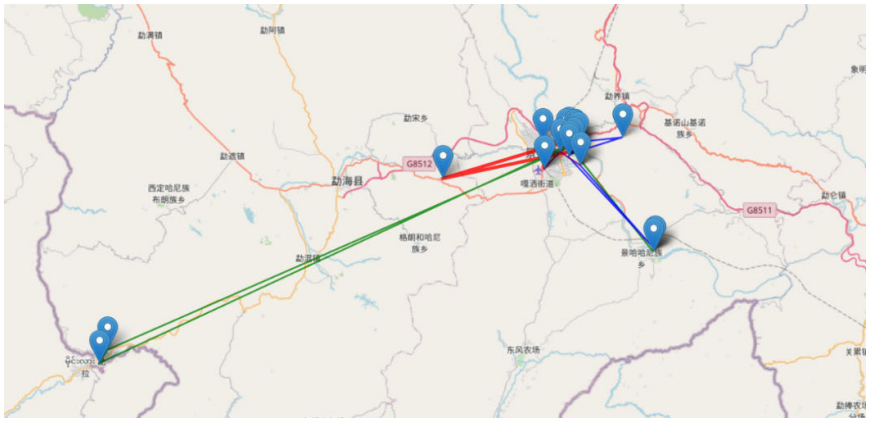

FIGURE 7. Visitor group 3 tour route map.

group tourism, which confirms the practicality of the method in this paper.

VI. CONCLUSION

In this paper, we propose a customized travel itinerary for groups with similar interests. First, visitors are clustered according to their heterogeneous needs to determine visitor groupings. Then, aiming at the popularity of the location, the praise rate, and the satisfaction of the needs of tourists in the group, the online information resources related to tourism and transportation are used to construct and solve the customized model of the travel itinerary. Finally determine the price of the itinerary, which is a detailed route with a specific schedule. The method uses data from attractions, restaurants, hotels, transportation, etc. to digitize and quantify the itinerary customization problem and help tourism companies improve customization efficiency. In addition to considering the needs of visitors, the customization process also takes into account the popularity and praise of each location, which is in line with the actual habits of tourists when choosing a location. Because when they travel, they usually consider whether the place is popular and whether the place has a better rating. Traveling as a group enables easier negotiation with suppliers, resulting in lower prices, while companies save costs and resources by sending only one customizer to customize the itinerary for the group. Multi-person tours typically cost less than individual tours and increase the opportunity to interact with others. Tourists within a group use the same mode of transport, eliminating the need for one person and one vehicle, reducing carbon emissions and reducing traffic pressure.

This approach can clearly capture the changes in tourist demand after the COVID-19, while meeting the requirements of sustainable tourism development. This method has certain practicality and can provide tourists with more satisfactory travel routes and prices at the same time. This article briefly discusses the use of this method, but in practical application, the degree to which tourists can be grouped can ensure the satisfaction of both routes and prices, and further experiments on a larger scale and more times are needed. In practice, travel companies can also determine the grouping method based on experience, which requires cooperation between the method and the tourism company.

REFERENCES

[1] M. Delgado and K. G. Mills, "The supply chain economy: A new industry categorization for understanding innovation in services," *Res. Policy*, vol. 49, no. 8, Oct. 2020, Art. no. 104039.

[2] H. K. Lee and Y. Fernando, "The antecedents and outcomes of the medical tourism supply chain," *Tourism Manag.*, vol. 46, pp. 148–157, Feb. 2015.

[3] V. Cho, "A comparison of three different approaches to tourist arrival forecasting," *Tourism Manag.*, vol. 24, no. 3, pp. 323–330, Jun. 2003.

[4] K. F. Hyde and R. Lawson, "The nature of independent travel," *J. Travel Res.*, vol. 42, no. 1, pp. 13–23, 2003.

[5] G. I. Alptekin and G. Buyuközkan, "An integrated case-based reasoning and MCDM system for web based tourism destination planning," *Exp. Syst. Appl.*, vol. 38, no. 3, pp. 2125–2132, Mar. 2011.

[6] M. Aronica, P. Pizzuto, and C. Sciortino, "COVID-19 and tourism: What can we learn from the past?" *World Economy*, vol. 45, no. 2, pp. 430–444, Feb. 2022.

[7] X. Fan, J. Lu, M. Qiu, and X. Xiao, "Changes in travel behaviors and intentions during the COVID-19 pandemic and recovery period: A case study of China," *J. Outdoor Recreation Tourism*, vol. 41, Mar. 2023, Art. no. 100522.

[8] L. Yu, P. Zhao, J. Tang, and L. Pang, "Changes in tourist mobility after COVID-19 outbreaks," *Ann. Tourism Res.*, vol. 98, Jan. 2023, Art. no. 103522.

[9] F. Fusté-Forné and N. Michael, "Limited tourism: Travel bubbles for a sustainable future," *J. Sustain. Tourism*, vol. 31, no. 1, pp. 73–90, Jan. 2023.

[10] S. Gössling, D. Scott, and C. M. Hall, "Pandemics, tourism and global change: A rapid assessment of COVID-19," *J. Sustain. Tourism*, vol. 29, no. 1, pp. 1–20, Jan. 2021.

[11] F. Higgins-Desbiolles, "Socialising tourism for social and ecological justice after COVID-19," *Tourism Geographies*, vol. 22, no. 3, pp. 610–623, May 2020.

[12] D. Herzog, C. Laß, and W. Wörndl, "TourRec: A tourist trip recommender system for individuals and groups," in *Proc. 12th ACM Conf. Recommender Syst.*, Sep. 2018, pp. 496–497.

[13] K. H. Lim, "Recommending and planning trip itineraries for individual travelers and groups of tourists," in *Proc. 26th Int. Conf. Automated Planning Scheduling*, 2016, pp. 115–120.

[14] H. Lin, M. Zhang, D. Gursoy, and X. Fu, "Impact of tourist-to-tourist interaction on tourism experience: The mediating role of cohesion and intimacy," *Ann. Tourism Res.*, vol. 76, pp. 153–167, May 2019.

[15] L. Su, J. Cheng, and Y. Huang, "How do group size and group familiarity influence tourist satisfaction? The mediating role of perceived value," *J. Travel Res.*, vol. 60, no. 8, pp. 1821–1840, Nov. 2021.

[16] U. Gretzel, M. Sigala, Z. Xiang, and C. Koo, "Smart tourism: Foundations and developments," *Electron. Markets*, vol. 25, no. 3, pp. 179–188, Sep. 2015.
[17] P. D'Urso, L. De Giovanni, M. Disegna, and R. Massari, "Bagged clustering and its application to tourism market segmentation," *Exp. Syst. Appl.*, vol. 40, no. 12, pp. 4944–4956, Sep. 2013.
[18] Ç. H. Özel and N. Kozak, "Motive based segmentation of the cultural tourism market: A study of Turkish domestic tourists," *J. Quality Assurance Hospitality Tourism*, vol. 13, no. 3, pp. 165–186, Jul. 2012.
[19] F. E. Otoo, S. Kim, and Y. Choi, "Understanding senior tourists preferences and characteristics based on their overseas travel motivation clusters," *J. Travel Tourism Marketing*, vol. 37, no. 2, pp. 246–257, Feb. 2020.
[20] S. Schiaffino and A. Amandi, "Building an expert travel agent as a software agent," *Exp. Syst. Appl.*, vol. 36, no. 2, pp. 1291–1299, Mar. 2009.
[21] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model-based collaborative filtering for personalized POI recommendations," *IEEE Trans. Multimedia*, vol. 17, no. 6, pp. 907–918, Jun. 2015.
[22] L. Esmaeili, S. Mardani, S. A. H. Golpayegani, and Z. Z. Madar, "A novel tourism recommender system in the context of social commerce," *Exp. Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113301.
[23] G. H. Lee and H. S. Han, "Clustering of tourist routes for individual tourists using sequential pattern mining," *J. Supercomput.*, vol. 76, no. 7, pp. 5364–5381, 2020.
[24] M. Derek, E. Wozniak, and S. Kulczyk, "Clustering nature-based tourists by activity. Social, economic and spatial dimensions," *Tourism Manag.*, vol. 75, pp. 509–521, Dec. 2019.
[25] W. Liu, B. Wang, Y. Yang, N. Mou, Y. Zheng, L. Zhang, and T. Yang, "Cluster analysis of microscopic spatio-temporal patterns of tourists movement behaviors in mountainous scenic areas using open GPS-trajectory data," *Tourism Manag.*, vol. 93, Dec. 2022, Art. no. 104614.
[26] B. Rodríguez, J. Molina, F. Pérez, and R. Caballero, "Interactive design of personalised tourism routes," *Tourism Manag.*, vol. 33, no. 4, pp. 926–940, Aug. 2012.
[27] H.-S. Chiang and T.-C. Huang, "User-adapted travel planning system for personalized schedule recommendation," *Inf. Fusion*, vol. 21, pp. 3–17, Jan. 2015.
[28] H. Yuan, H. Xu, Y. Qian, and Y. Li, "Make your travel smarter: Summarizing urban tourism information from massive blog data," *Int. J. Inf. Manag.*, vol. 36, no. 6, pp. 1306–1319, Dec. 2016.
[29] S. Kotiloglu, T. Lappas, K. Pelechrinis, and P. P. Repoussis, "Personalized multi-period tour recommendations," *Tourism Manag.*, vol. 62, pp. 76–88, Oct. 2017.
[30] Z. Liao and W. Zheng, "Using a heuristic algorithm to design a personalized day tour route in a time-dependent stochastic environment," *Tourism Manag.*, vol. 68, pp. 284–300, Oct. 2018.
[31] H. Ji, W. Zheng, X. Zhuang, and Z. Lin, "Explore for a day? Generating personalized itineraries that fit spatial heterogeneity of tourist attractions," *Inf. Manag.*, vol. 58, no. 8, Dec. 2021, Art. no. 103557.
[32] X. Zheng, B. Han, and Z. Ni, "Tourism route recommendation based on a multi-objective evolutionary algorithm using two-stage decomposition and Pareto layering," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 2, pp. 486–500, Feb. 2023.
[33] I. Garcia, L. Sebastia, and E. Onaindia, "On the design of individual and group recommender systems for tourism," *Exp. Syst. Appl.*, vol. 38, no. 6, pp. 7683–7692, Jun. 2011.
[34] A. Anagnostopoulos, R. Atassi, L. Becchetti, A. Fazzone, and F. Silvestri, "Tour recommendation for groups," *Data Mining Knowl. Discovery*, vol. 31, no. 5, pp. 1157–1188, Sep. 2017.
[35] J. Ruiz-Meza and J. R. Montoya-Torres, "Tourist trip design with heterogeneous preferences, transport mode selection and environmental considerations," *Ann. Oper. Res.*, vol. 305, nos. 1–2, pp. 227–249, Oct. 2021.
[36] Z. Liao, W. Zheng, and Z. Lin, "Time apart while together: A smart trip design for group travelers," *Ann. Tourism Res.*, vol. 93, Mar. 2022, Art. no. 103374.
[37] M. Kargar and Z. Lin, "A socially motivating and environmentally friendly tour recommendation framework for tourist groups," *Exp. Syst. Appl.*, vol. 180, Oct. 2021, Art. no. 115083.

CHAO YU was born in Shenyang, Liaoning. She received the bachelor's, master's, and Ph.D. degrees in management from Northeastern University, China, in 2010, 2012, and 2016, respectively. She is currently an Associate Professor with the Department of Engineering Management, School of Management, Shenyang University of Technology, focusing on operations management and decision analysis.

MENGLI YANG was born in Langfang, Hebei. She is currently pursuing the master's degree in management science and engineering with the Shenyang University of Technology, China. Her research interest includes tourism service management.