Explore for a day? Generating personalized itineraries that fit spatial heterogeneity of tourist attractions

Haipeng Ji[a]   Weimin Zheng[a]   Xinyi Zhuang[a]   Zhibin Lin[b]

[a] School of Management, Xiamen University, 422 South Siming Road, 361005, Xiamen, China

[b] Durham University Business School, Mill Hill Lane, Durham DH1 3LB, United Kingdom

Explore for a day? Generating personalized itineraries that fit spatial heterogeneity of tourist attractions

Abstract: Recommender systems are widely adopted by firms as an innovative personalization tool across various industries. Most of the existing tour recommender systems treat the spatial structure of tourist attractions as a single type, which neglects the spatial heterogeneity among these attractions. This study attempts to address this problem by modeling the spatial heterogeneity in the design of personalized trips. We propose a two-phase heuristic approach, which involves an improved artificial bee colony algorithm and a differential evolution algorithm. The results of a field experiment confirm that our new model outperforms the benchmark models in maximizing customer utilities.

Keywords: Recommender system; Tourist trip design problem; Heuristic approach; Personalization; Spatial structure; Tourism attraction

1. Introduction

Planning a one-day exploration in various places of interest at an unfamiliar destination is time-consuming. A tour recommender system can help users plan a dream holiday exploration by automatically generating a personalized travel plan that suits their needs and preferences. A recommender system generally uses content-based and/or collaborative filtering algorithms by considering the attributes of a product that a consumer liked or purchased in the past (content-based) or the similarity between a consumer and others according to their historical like or purchase data (collaborative filtering) (Lee & Hosanagar, 2019). The design of a tour recommendation is more complex than that of a product recommendation, as it involves recommending the attractions or points of interest (POIs) and a travel route that connects the POIs (Kotiloglu, et al., 2017).

Developers of a tour recommendation have to consider numerous real-life constraints, among which are the temporal and spatial constraints (Lim, et al., 2019). In terms of the spatial structure, there are two basic types of attractions, namely the node- and line-shaped attractions, which often exist simultaneously in a tourism destination. For line-shape attractions, such as greenway, coastline, river, and street, the profits (or utility) for tourists are associated with arcs, rather than vertices and their spatial position changes when they finish visiting the attractions. Despite the advances

in tour recommender systems, most studies treat all attractions as the same in terms of spatial structure and ignore the multiple entrances/exits of attractions, thereby restricting the modeling of vertices or arcs. As a result, these systems may fail to capture the practical properties of attractions with large areas and multiple entrances/exits. For example, Yellowstone National Park has five entrance stations, each of which is far from the others. Therefore, choosing appropriate stations to enter and leave the park while planning the trip is essential. Moreover, previous studies put relatively little emphasis on optimizing the time for exploration or enjoyment at each attraction, which is a pivotal part of tourist experiences and must not be ignored in the recommender system design (Zheng & Liao, 2019).

This study aims to tackle the above problems by considering tourist attractions' spatial structure and the time spent at each attraction for generating a personalized travel itinerary. Specifically, we consider the heterogeneity of attractions' spatial structure, that is, the inhomogeneity and complexity of spatial structure, and categorized it into three types: (a) POI (node-shape attraction with a unique and identical entrance and exit), (b) line of interest (LOI, line-shape attraction with only one entrance/exit at both ends), and (c) area of interest (AOI, large area with multiple entrances/exits). In addition, the time spent at attractions is considered an optimization variable. The complexity of such a problem results from the correlation among attraction selection, sequencing determination, choice of entrances/exits and time allocation, and multiple constraints. We tackle this mixed tourist trip design problem by developing a two-phase heuristic approach (HA), which involves an improved artificial bee colony (ABC) algorithm and a differential evolution algorithm (DEA). Our approach differs from existing trip design methods in several ways. First, our proposed approach applies a variant-form nectar with four pheromones to code the solution. Second, it designs a cell array embedded in different dimension matrices for improved storing and managing of the relationship among vertices. Third, it improves the search ability and optimization performance of the algorithm by adjusting the evolution structure of ABC and adding a new group of bees. Finally, it also improves the evolution structure by introducing an adaptive evolutionary parameter to reach an equilibrium of the solution quality and algorithm efficiency.

This study makes three major contributions to information systems literature. First, this study advances recommender system design by offering an improved approach that integrates spatial heterogeneity with other features and influencing factors, whereas

most previous recommender system studies mainly focus on recommending POIs, without considering the unique spatial features of each POI. Second, our system design incorporates the actual available entrances and exits of each attraction; thus, it avoids unnecessary detours and allows more time for the users to explore and enjoy their visit. Third, our design further considers the access order, access time, and visitors' personal preferences. As a result, the recommendations generated from our approach can closely match users' preferences while meeting their time budgets, which outperform those generated from the baseline methods.

The remaining sections of this paper are organized as follows. Section 2 reviews the design of recommender systems and, specifically, tourist trip design. Section 3 formalizes a high-efficient mathematical model to address the mixed tourist trip design problem. Section 4 presents our proposed HA framework. Section 5 examines the effectiveness of our method through a field experiment. Finally, Section 6 discusses the results and offers directions for further studies.

2. Literature review

2.1. *Tourism recommender system*

The design of a recommender system is usually based on one or both of the two basic modeling approaches: content-based and collaborative filtering algorithms (Lee & Hosanagar, 2019). Content-based systems focus on the attributes of a product that a consumer liked or purchased in the past, whereas collaborative filtering recommenders are based on the similarity between a consumer and others according to their historical like or purchase data. Many recommender systems use a combination of both approaches. Research on the POI recommender system has explored four major techniques, including the collaborative filtering-based, matrix factorization-based, probabilistic, and link-based models (Si, et al., 2019). First, both user-based and friend-based collaborative filtering have been adapted for recommending POIs. Second, various matrix factorization models have been leveraged, together with a combination of multiple latent factors. Third, probabilistic models have been used to capture the different influencing factors, including social, sequential, geographical, and temporal influences on user preferences to generate recommendations. Fourth, link-based models have been used to present graphs of the links between users and POIs for recommendations. Integrating the various factors with algorithmic techniques improves the tour recommendations (Si, et al., 2019).

The design of a tour recommendation should consider numerous real-life constraints, in addition to that of a product or POI recommendation. The temporal and spatial constraints are the key influencing factors that should be integrated into the design (Lim, et al., 2019). In a typical day-trip excursion at a destination, a day reflects the constraint of time, and the destination offers the space for tourist exploration. The temporal constraints include the limit of time allocated for the tour including time at attractions and travel between attractions. The design thus needs to incorporate various factors, such as crowdedness and time for queuing at the attractions, transport modes used, traffic congestion, and uncertainty in time for travel (Gavalas, et al., 2015; Zhang, et al., 2016). Moreover, the tour recommender system has to consider factors such as user demographics (Cheng, et al., 2011), traveling group size (Chen, et al., 2013), real-time location, and user's personal interests (Taylor, et al., 2018). Contextual factors such as the day, time, season, and weather should also be taken into account (Majid, et al., 2015).

2.2. Spatial constraints for a tourist trip design

The spatial constraints for a tourist trip design include the need for the start and end at certain locations, including the location of the hotel the tourist stays, stations, attractions, and specifically the entrance and exit of an attraction. Most studies abstract tourist attractions as vertices and assume that tourists enter and leave attractions at the same location (see a summary in Table 1). These works are sufficient for designing routes at attractions with single entrance/exit, such as museums, galleries, small squares, or parks. However, the practical properties of other types of attractions, such as greenway, coastline, river, and street, are not mined. When tourists visit such attractions, the profits are associated with arcs, rather than vertices (Lu, et al., 2017; Lu & Shahabi, 2015), and their spatial position changes when they finish visiting the attractions (Souffriau, et al., 2011). Given these differences, abstracting attractions as vertices may not be feasible or ideal in practice (Eby & Molnar, 2002).

Several recent studies treat attractions as arcs and regard tourist trip design problem as a variant of arc orienteering problem. For example, Souffriau et al. (2011) formulated the cycle route planning and proposed a heuristic solution approach and presented a mathematical optimization model. Verbeeck et al. (2014) extend the arc orienteering problem, consider the different profits of various direction arcs, and introduce a branch cutting method to solve the bicycle journey design problem with the same end and starting points. Lu and Shahabi (2015) introduce a variant of arc

orienteering problem and used a set of meta-heuristic algorithms to tackle the problem in search of the optimal travel path in large-scale road networks. Lu et al. (2017) proposed a two-time-dependent arc orienteering problem, under which the travel time and benefit value depend on time.

Table 1 Studies on personalized tour design

Authors	Contributions	Models	Factors
Lee, et al. (2009)	An ontological recommendation multi-agent.	Ant colony optimization	Context information; Tourists' requirements; Tainan City travel ontology.
Rodriguez et al. (2012)	A tool that formalizes a mathematical model and interactive multi-criteria technique.	Tabu search	Multiple tourists' objectives; Interactive process with the tourist.
Hsu et al. (2012)	An intelligent recommender system for tour decision-making.	None (Based on Google API)	Tourist preference prediction.
Tsai and Chung (2012)	A route recommender system based on tourist behavior and real-time information.	Route generation algorithm	Tourist behaviors similarity; Current facility queuing situation.
Liu, et al. (2014)	A recommender system that focuses on real-time personalized tour design.	Route generated algorithm	Real-time traffic information.
Cenamor, et al. (2017)	A system based on information gathered from social networks.	Automated planning approach	User expectations for POIs; POIs popularity.
Kotiloglu, et al. (2017)	A framework named "Filter-first, Tour-second".	Iterated tabu search	Mandatory points; Optional points.
Sun and Lee (2017)	A four-phase framework based on contents gathered from photo-sharing social networks.	Tour recommendations by sharing photos approach	Landmark topics; User characterization.
Zheng, et al. (2017)	A combination of difference evolution algorithm and a genetic algorithm	Genetic algorithm and differential evolution	Aesthetic fatigue; Variable sightseeing value.
Liao and Zheng (2018)	A stochastic environment that is time-dependent in the tourist trip design problem.	Heuristic algorithm based on random simulation	Time-dependent stochastic environment like travel times and wait times.
Zheng and Liao (2019)	A heuristic approach using Pareto optimality to meet group member preferences.	Nondominated sorting heuristic algorithm	Heterogeneous preferences of group members.
Zheng, et al. (2020a)	A two-level heuristic approach with consideration of hotel selection.	Memetic algorithm	Hotel selection.
Zheng, et al. (2020b)	A model that considers transport mode choice in the day itinerary design	Nondominated sorting heuristic algorithm	Transport mode.

The two types of attractions (node- and line-shape) may exist simultaneously in tourism destinations, which make the tour itinerary design a typical combination of orienteering problem and arc orienteering problem (Vansteenwegen, et al., 2011; Zheng, et al., 2020a). Gavalas et al. (2017) abstract the multiple-day itinerary design problem with node- and line-shape attractions and proposed the first metaheuristic approaches to tackle it. Mrazovic et al. (2017) also model this issue and introduce a variable

neighborhood search to deal with this problem.

Despite the progress made in the recommender system research, existing studies continue to ignore the discussion and solution of spatial heterogeneity. First, prior literature focuses on the information at the tourist and attraction levels to recommend tours (Cenamor, et al., 2017; Hsu, et al., 2012; Sun & Lee, 2017). Most studies assume that all the attractions have the same spatial structure and that tourists enter and leave attractions at predefined locations (Cenamor, et al., 2017; Hsu, et al., 2012; Liao & Zheng, 2018; Sun & Lee, 2017; Zheng, et al., 2020a; Zheng, et al., 2017). This situation is impractical in many cases, especially for attractions with large areas and multiple entrances/exits (e.g., Yellowstone National Park). Considering that the paths (travel distance) from different exits of an attraction to different entrances of another attraction are completely different (Huang, et al., 2020), which further affects tourist behavior, ignoring the choice of attractions' entrances/exits may lead to potentially infeasible or suboptimal solutions (travel time increased and travel experience decreased). For example, although Liu et al. (2014) and Liao and Zheng (2018) take traffic/travel time into account when modeling, ignoring the spatial heterogeneity of attractions results in the underperformance of their approach in solving our problem. Second, existing studies on the mixed orienteering problem and the mixed tourist trip design problem assign a fixed value to the time spent on the vertex (Gavalas, et al., 2017; Mrazovic, et al., 2017). However, in reality, the lengths of time that each tourist wishes to take at a vertex may vary. Hence, vertex time allocation should also be optimized in line with tourists' characteristics (Zheng, et al., 2020a; Zheng & Liao, 2019).

Moreover, space variables, such as attraction, sequencing, and the entrances/exits of an attraction) are discrete variables, whereas time is a continuous one as the simultaneous optimization of spatial and temporal structure factors present additional modeling challenges. Consequently, we consider the following improvements. First, we fully consider the heterogeneity of tourist attractions' spatial structure and the duration spent at each vertex by proposing a two-phase HA to design additional reasonable trips. Second, with the recognition of the complexity of the mixed tourist trip design problem, we use multiple methods to reach an equilibrium of the solution quality and algorithm efficiency.

3. Mathematical model

A mathematical model can be developed to introduce the research problem. Table 2 lists the description of variables used. Let V be the set of vertices, including attractions

$\varLambda_j \in V$ such that $\varLambda_1 \in V_I$, $\varLambda_M \in V_F$, $\{\varLambda_2,\cdots,\varLambda_{M-1}\} \subset V_A$, the arrival time at $\varLambda_1$ is set to τ. For the attractions visited at each stage $\varLambda_j$, its actual visit time t_j^s should not take place outside its time window $[to_i, tc_i]$ where to_i is the opening time and tc_i is the closing time. However, the earlier or the later arrival time t_j^a will cause unnecessary waiting time or miss the favorite attraction. The time budget for the trip equals to T_{max}, which contains the duration of $\varLambda_j$ and the travel time between $\varLambda_{j-1}$ and $\varLambda_j$. For clarity, an example of a five-stage route ($M = 5$) is shown in Fig.1, where the solid red lines and dots represent a visit to a vertex (a stage), the dotted lines represent the road between each stage. Compared with the models built, when we consider the impact of attractions' heterogeneity on tourist trip design, our model introduces the number of entrances and exits of vertices and the influence of the selected entrances and exits on the travel time between two adjacent stages.
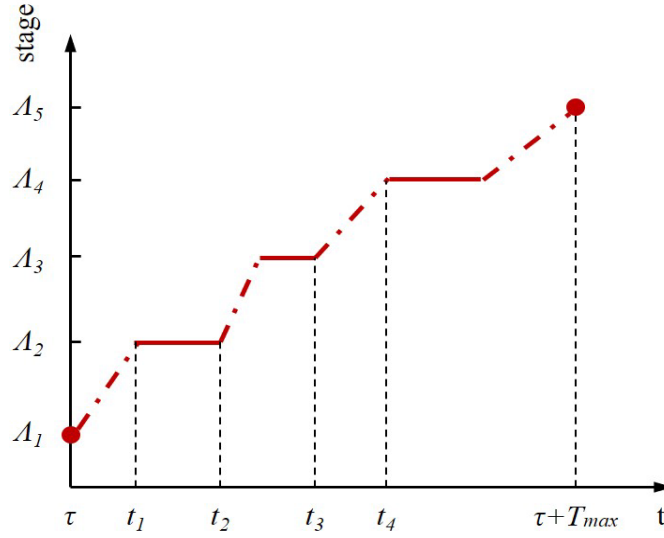

Fig.1 Illustration of a tourist route

Table 2 Mathematical variables

Variable	Description
V_I	Set of the initial starting locations of the destination, $i = 1, 2, ..., N_1$
V_A	Set of the attractions of the destination, $i = 1, 2, ..., N_2$
V_F	Set of the final ending locations of the destination, $i = 1, 2, ..., N_3$
V	Set of vertices, $V = V_I \cup V_A \cup V_F$
N	Number of vertices, $N = N_1 + N_2 + N_3$
EN_i^k	kth entrance of v_i
EX_i^k	kth exit of v_i
K_i^{EN}	Number of entrances for v_i

K_i^{EX}	Number of exits for v_i
T_{max}	Budgeted time available for the tourist
τ	Arrival time at the destination
n_i	Number of discrete visits to vertex v_i
M	Number of total stages in the trip (i.e., the sum of n_i, $M = \sum n_i$, $i = 1, 2, ..., N$)
Λ_j	Vertex visited at the jth stage, $j = 1, 2, ..., M$
$[to_i, tc_i]$	Time windows of v_i
$t(\Lambda_j, \Lambda_{j+1})$	Travel time between Λ_j and Λ_{j+1}
t_j^a	Arrival time at vertex Λ_j
t_j^s	Actual start time visiting vertex Λ_j
t_j^e	Departure time from vertex Λ_j
p_i	Tourist's preference value for v_i, $p_i \in [0,1]$
t_i	Average time spent at v_i by previous tourists
x_{ij}	If the tourist visits v_i at the jth stage, set $x_{ij} = 1$; otherwise, 0
y_{ij}	If a visit to v_i is followed by a visit to v_j, set $y_{ij} = 1$; otherwise, 0

3.1. Model objective

The model's objective is to maximize the utility of tourists under numerous constraints. Considering that tourists may repeatedly visit a landmark attraction during their tour (Tsai & Chung, 2012), let n_i represent the number of discrete visits to vertex v_i. In Eq. (1), integer M denotes the number of the stages, and N represents the number of vertices at the attraction:

$$M = \sum_{i=1}^{N} n_i \qquad (1)$$

The utility at each stage obtained by an individual tourist is mainly determined by the vertex visited at that stage (Λ_j). It is closely related to the length of time spent at the corresponding vertex and p_i for Λ_j. In tourist destinations, a unit time stay in an attraction is regarded as a commodity (Rugg, 1973). According to the law of diminishing marginal utility, marginal satisfaction decreases as consumers purchase more of the same product (Brewer & Venaik, 2010). Owing to the aesthetic fatigue of tourists in a similar way, as the duration a tourist stays at the same vertex increases, the marginal utility decreases (Liao & Zheng, 2018). Under this consideration, the utility gained by the tourist at jth stage can be calculated by Eq. (2) according to the duration that needs to be optimized and the preference value provided by tourists. In this equation, $MS_i(t)$ denotes the marginal sensation acquired by the tourist from v_i at moment t, a non-negative decreasing function of time. We set x_{ij} as a 0–1 discrete variable: if the tourist visits v_i at the jth stage, $x_{ij} = 1$; otherwise, 0. t_j^s indicates the start time at vertex Λ_j, whereas t_j^a is arrival time at Λ_j. These two values are usually unequal because the time windows of the vertices may result in waiting time. Thus, t_j^s

can be obtained as follows (see Eq. (3)):

$$u_j = \int_{t_j^s}^{t_j^e} \left\{ \sum_{i=1}^{N} \left[MS_i(t) \cdot p_i \cdot x_{ij} \right] \right\} dt \qquad (2)$$

$$t_j^s = \max\left[t_j^a, to_i \right] \qquad (3)$$

We suppose that tourists obtain zero utility while waiting and during traffic. Thus, the utility can be obtained as follows:

$$u = \sum_{j=1}^{M} u_j \qquad (4)$$

3.2. Model constraints

A tour route design has personalized and permanent technical constraints. Implementing permanent technical constraints ensures the validity and real meaning of the designed routes, as shown in Eqs. (5)–(10), whereas implementing personalized constraints to ensure the trip is within the time budget limit is the premise to meet an individual's needs and preferences (Rodriguez, et al., 2012), as illustrated in Eq. (11). Specifically, Eqs. (5) and (6) set a limit that a tourist starts her/his trip at one of the initial starting locations and ends the trip at one of the final arrival locations. Eq. (7) guarantees that only one attraction can be visited per stage, from the second one to the $M-1$th stages of the entire trip:

$$\sum_{v_i \in V_I} x_{i1} = \sum_{v_j \in V_F} x_{jM} = 1 \qquad (5)$$

$$\sum_{j=1}^{M} \sum_{v_i \in V_I} x_{ij} = \sum_{j=1}^{M} \sum_{v_i \in V_F} x_{ij} = 1 \qquad (6)$$

$$\sum_{v_i \in V_A} x_{ij} = 1, j = 2, 3, \cdots, M-1 \qquad (7)$$

Equations (8)–(9) ensure the connectivity of time and path, where y_{ij} is a 0–1 discrete variable. If a visit to v_i is followed by a visit to v_j, then we set y_{ij} to 1; 0, otherwise. If a tourist visits v_i at the jth stage and enters v_i from the kth entrance, then we set 0–1 discrete variable $g_{ijk}=1$; otherwise, $g_{ijk}=0$. Similarly, if a tourist visits v_i at the jth stage and leaves v_i from the kth exit, then we set 0–1 discrete variable $h_{ijk}=1$; otherwise, $h_{ijk}=0$. Eq. (10) restricts that a tourist enters through one of the entrances to visit the vertex and then leaves through one of the exits. Eq. (11) limits the total visit time of the trip to a maximum time budget T_{max}, where t_M^a is the arrival time at Λ_M and

τ is the time that the tourist starts the trip:

$$t_j^e + t\left(\Lambda_j, \Lambda_{j+1}\right) = t_{j+1}^a, \left(\forall j = 1, 2, \cdots, M-1\right) \tag{8}$$

$$\sum_{v_i \in V_I \cup V_A} y_{ij} - \sum_{v_k \in V_A \cup V_F} y_{jk} = 0, \forall v_j \in V_A; v_i \neq v_j, v_j \neq v_k \tag{9}$$

$$\sum_{k=1}^{K_i^{EN}} g_{ijk} = \sum_{k=1}^{K_i^{EX}} h_{ijk} = x_{ij} \tag{10}$$

$$t_M^a \leq \tau + T_{\max} \tag{11}$$

4. Solution approach

The problem proposed in this study is a generalization of the orienteering problem, which has been proven to be NP-hard (class of problems that are at least as hard as the hardest problems in NP) (Golden, et al., 1987). This implies that deriving the optimal solution of the model in a limited time is difficult. Moreover, our model considers the influence of spatial heterogeneity on tour recommendation, which makes it more complex and increases the risk of falling into local optimum. Therefore, we propose a two-phase HA that combines improved ABC and DEA, including the preprocessing and evolution phases. ABC algorithm, first introduced by Karaboga (2005). DEA is specifically applicable to solving continuous optimization problems. The comprehensive framework of HA is illustrated in Fig. 2. The details of each phase are presented in Sections 4.1 and 4.2.
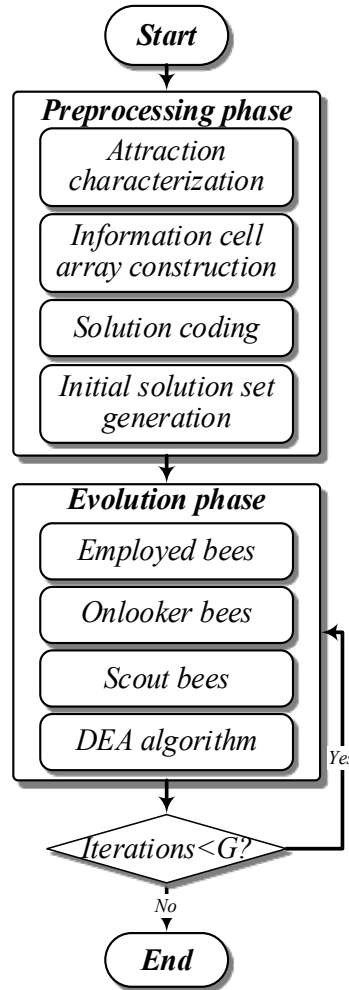
Fig. 2 Procedure of the proposed method

4.1. Preprocessing phase

In the preprocessing phase, four tasks are conducted: attraction characterization, information cell array construction, solution coding, and initial solution set (**ISS**) generation.

As previously mentioned, a tour route may contain attractions with different spatial structures: POI, LOI, and AOI. For example, the route as shown in Fig. 3 (a) means that the tourist starts her/his trip from the initial start location (v_1) and then successively visits v_3, v_7, v_5, v_6, and v_2. Finally, the trip ends at the final end location (v_4). For the visited attractions, v_3 and v_7 are POIs, v_5 is an AOI, and v_6 and v_2 are LOIs. To facilitate the subsequent processing of the approach, each entrance and exit of vertices should also be characterized. Therefore, the information of attractions in Fig. 3 (a) is transformed to Fig. 3 (b).
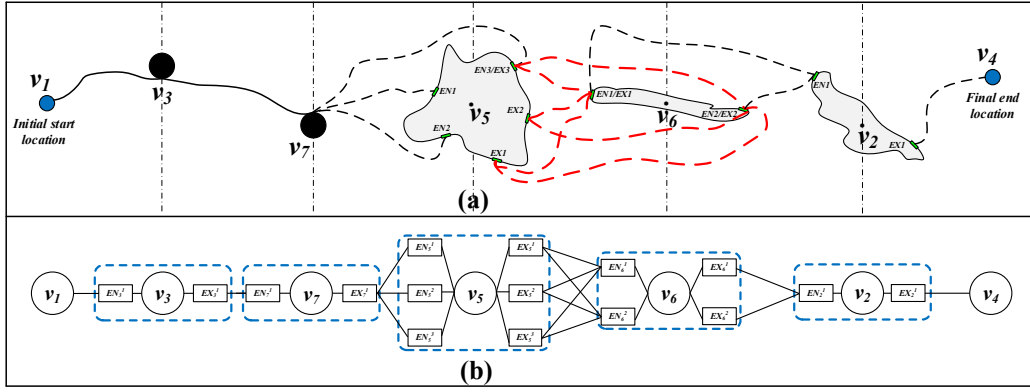
Fig. 3 Example of attraction characterization

Certain vertices contain multiple entrances and exits. This leads to multiple paths between a pair of vertices. For example, six paths exist between the pair (v_5, v_6) (as shown in red lines in Fig. 3 (a)). Choosing different entrances/exits indicates that different path selections are available, which may correspond to different travel times. This information can significantly affect trip optimization. Most previous studies use a regular matrix to store information without considering the multiple entrances/exits of attractions. Given that the existing methods are incapable of solving the problem concerned, we design a cell array embedded in different dimension matrices, which denote a distance matrix between two vertices. For clarity, we provide a specific example to illustrate such a structure, as shown in Fig. 4 (a). The structure contains the distance information among vertices $(v_1, v_2, …, v_7)$. As displayed in Fig. 3, six paths exist between the pair (v_5, v_6) and two paths between the pair (v_6, v_2). Thus, two cell arrays ($v_5 \rightarrow v_6$, shown in the red grid; $v_6 \rightarrow v_2$, presented in the blue grid) denote two matrices with different dimensions, as illustrated in Fig. 4 (b).
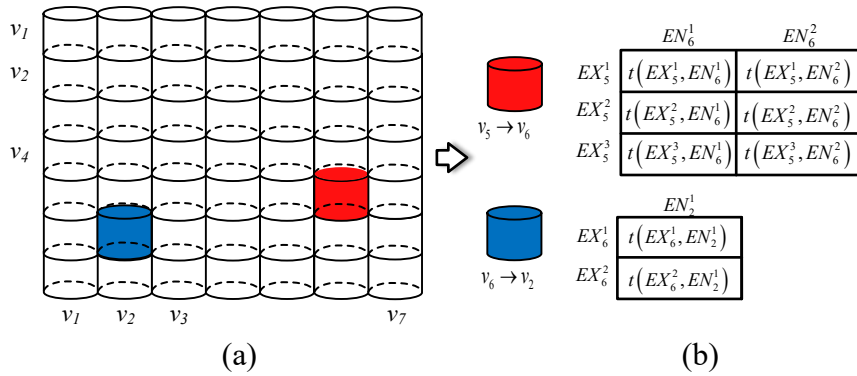

Fig. 4 Cell array embedded in different dimension matrices

After characterizing the attractions and constructing the information cell array, we must code the solutions. Most evolution algorithms require an advanced determination of solution dimensions (Geiger & Eskandari, 2008), which cannot be applied to the present problem owing to possible variations in the vertices that a tourist visits. To code

the solutions, Zheng et al. (2017) design a double-layer, variable-length chromosome that involves vertex selection, sequencing, and time allocation. In addition to these three elements, our study must determine the entrances/exits of the chosen vertices. Thus, we introduce a variable-form nectar with four pheromones to code the solutions. The two upper pheromones are the vertex selection of the route and the time spent at each chosen vertex, whereas the two lower pheromones indicate the entrances and exits of the chosen vertices. This solution coding is illustrated with an example in Fig. 5, which depicts that the tourist starts the trip at v_1, then successively visits v_3, v_7, v_5, v_6, v_2, and v_4, where the trip ends. At v_3, v_7, v_5, v_6, and v_2, the time spent are 69, 18, 6, 43, and 26 minutes, respectively. The entrance and exit for each chosen vertex are $\left(EN_1^1, EX_1^1\right)$, $\left(EN_3^1, EX_3^1\right)$, $\left(EN_7^1, EX_7^1\right)$, $\left(EN_5^3, EX_5^2\right)$..., $\left(EN_4^1, EX_4^1\right)$, as shown in the red lines in Fig. 5 (a).

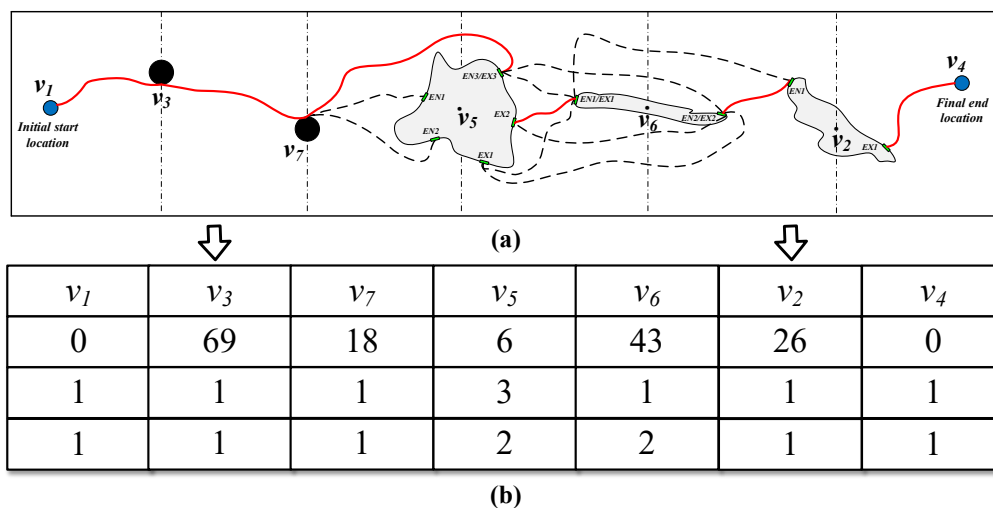| v_1	v_3	v_7	v_5	v_6	v_2	v_4
0	69	18	6	43	26	0
1	1	1	3	1	1	1
1	1	1	2	2	1	1

(b)

Fig. 5 Examples of solution coding

Each employed bee (*EB*) corresponds to a food source (solution). The quality of **ISS** strongly influences the performance of our approach. To ensure the diversity of solutions, *SN* (population size) solutions are generated based on the constraints as presented in Section 3.2, and they are randomly assigned to the *EBs*.

4.2. Evolution process

The goal of the evolution process is to obtain solutions with greater utility, matching an individual tourist's needs. As described in Section 4.1, a solution is coded as a variable-form nectar with four pheromones, including the selection and sequencing of vertices, the length of time at each vertex, and the choice of vertices' entrances and exits. We use an improved ABC algorithm and a DEA to evolve these four variables. The improved ABC algorithm optimization includes three discrete variables: selection of vertices, sequencing of vertices, and the choice of vertices' entrances and exits. The

DEA is applied to optimize the visit time for the corresponding vertices.

The ABC algorithm is developed by Karaboga (2005) based on honeybees' behavior. In the ABC system, *EBs* and onlooker bees (*OBs*) select food sources based on their experiences and nestmates and then adjust their positions. Scout bees (*SBs*) fly and randomly select food sources without using experiences. The ABC system combines local and global search methods. Local search methods are used by *EBs* and *OBs*, whereas global search methods are used by *SBs*. Given the simplicity, flexibility, and robustness of the ABC algorithm, it has been extensively used in optimization problems with multiple variables (Cura, 2014; Karaboga, et al., 2014; Kıran, et al., 2013; Martín-Moreno & Vega-Rodríguez, 2018; Omkar, et al., 2011; Pan, et al., 2011; Szeto, et al., 2011). Realizing that our study involves interacting elements, we improve ABC in two ways. (1) Four neighborhood structures are embedded in ABC to enhance the local search ability. (2) A new group of bees named "employed scout bee (*e-SB*)" is designed to further optimize the choice of entrances/exits of attractions.

4.2.1. EB process

A solution is denoted as a variable-form nectar with four pheromones. In this study, the *EBs* focus on the optimization of the first pheromone, that is, the selection and sequencing of vertices. Each *EB* is linked with a particular food source (solution). At each iteration, *EBs* search for new food sources and assess their fitness (the utility of the solution). When the food source position cannot be improved further, ABC algorithm abandons the food source after a predetermined number of iterations *limit* (Karaboga, 2005). However, to solve the mixed tourist trip design problem, we suggest the extract-insert search strategy to allow the algorithm a greater opportunity to escape from a local optimum. All four applied neighborhood structures are introduced: insertion (select a vertex and then insert it into the optimal location), inversion (reverse the sequence between two vertices), swap (select two vertices and then exchange their positions), and extract-insert (exclude a random number of vertices and insert non-included vertices). For the detailed illustration of these neighborhood structures, we refer to Cura (2014).

We suppose that the solution loaded on an *EB* is the current solution (*CS*), and the fitness value of CS (*f(CS)*) represents the utility of the corresponding solution, which can be obtained through Eq. (4). First, the *EB* randomly selects one of the above neighborhood structures to search for a new solution (*NS*). The fitness value of *NS* (*f(NS)*) is also calculated based on Eq. (4). If (*f(NS) > f(CS)*), then a good solution is found. Subsequently, the *CS* of the *EB* is replaced with the *NS*, and the number of trials

($t(EB)$) is reset to 0; otherwise, the *EB* keeps the current solution, and $t(EB)$ is added by 1 (i.e., $t(EB) = t(EB)+1$).

4.2.2.   OB process

After the *EB* searching behavior is completed, *EBs* return to the hive with information about the food sources (solutions) and head to the dance area to share the information. *OBs* waiting in the nest selects a food source according to such information. Specifically, an *OB* chooses a solution based on the probability values, which can be obtained by using the fitness values that *EBs* provided. In view of this purpose, the roulette wheel selection method can be adopted. This fitness-based selection technique is proposed by Goldberg (1989). Hence, many onlookers are attracted to rich sources, resulting in positive feedback behavior. The probability value p_m with which *EB* is selected by an *OB* can be calculated using Eq. (12), where CS_m means the solution loaded on EB_m, and *SN* denotes the population size:

$$p_m = \frac{f(CS_m)}{\sum\limits_{m=1}^{SN} f(CS_m)}. \tag{12}$$

After an *OB* selects an *EB* and its corresponding solution, a neighborhood solution is determined by randomly using one of the above neighborhood structures again. Its fitness value is computed using Eq. (4). Following the *EB* process, a greedy selection is used between the originally selected solution (CS_m) and the neighborhood solution. If the neighborhood solution dominates CS_m, then CS_m is replaced by the neighborhood solution.

4.2.3.   SB process

In the general ABC algorithm, unemployed bees who randomly select their food sources are named scout bees. In addition to general *SBs* (*g-SBs*), a new group named *e-SB* is designed to further optimize the choice of entrances/exits of attractions. The details of both groups are presented in the following:

(1) *g-SBs*. As mentioned earlier, if the solution to an *EB* is unable to be improved through a predetermined number of trials (named abandonment criteria, *limit*), that is $t(EB) = limit$, then the *EB* becomes *g-SB*, and its solution is abandoned. The converted *g-SB* begins to randomly seek a new solution, which can be assigned to the corresponding *EB*, whose solution has been abandoned. Hence, poor solutions are discarded, and negative and positive feedbacks are balanced.

(2) *e-SBs*. *e-SBs* are special bees that are designed to optimize the entrances and exits of attractions. Such entrances and exits are represented as the two lower

Each *e-SB* only corresponds to a single *EB* in the whole evolution process, and its optimization is based on the solution of the *EB*. Specifically, *e-SBs* optimize the route whose vertex selection and sequencing have been optimized by *EBs* and *OBs*. *e-SBs* seek good solutions according to the following steps. First, the paths set between each pair of adjacent vertices (Λ_1, Λ_2), (Λ_2, Λ_3), …, $(\Lambda_{M-1}, \Lambda_M)$ are determined, and the path sets are denoted as $S_1, S_2, …, S_{M-1}$. Second, the shortest path (p_i^*) of the *i*th path set (S_i) is found. Finally, the starting and ending points of p_i^* are marked as the exit and entrance vertices Λ_i and Λ_{i+1}, respectively. For clarity, the pair (v_5, v_6) shown in Fig. 3 is taken to illustrate this process. Six paths are observed between pair (v_5, v_6), and the travel time of the path between EX_5^3 and EX_6^1 is the shortest among the six paths. Thus, EX_5^3 and EX_6^1 are selected as the exit of v_5 and the entrance of v_6, respectively (Fig. 6).

v_1	v_3	v_7	v_5	v_6	v_2	v_4
0	69	18	6	43	26	0
1	1	1	3	1	1	1
1	1	1	2	2	1	1

v_1	v_3	v_7	v_5	v_6	v_2	v_4
0	69	18	6	43	26	0
1	1	1	3	1	1	1
1	1	1	3	2	1	1

Fig. 6 Example of an e-SB

4.2.4. DEA process

Typically, each iteration involves all four processes: *EBs*, *OBs*, *SBs*, and DEA. The optimization of vertex entrance/exit and time allocation is based on that of vertex selection and sequencing. In addition, in the early stage of evolution, the optimization intensity of vertex selection and sequencing is strong. Thus, the optimization of vertex entrance/exit and time allocation in the early stage of evolution cannot effectively improve the quality of final solutions but can reduce the efficiency of the algorithm. To reach an equilibrium of solution quality and algorithm efficiency, we improve the evolution structure by introducing the adaptive evolutionary parameter (p_e), which can be calculated using Eq. (13). In this equation, G is the iteration times of the algorithm, and *Iter* refers to the current iterations. p_e increases as the number of iterations increases, indicating that the optimization intensity of vertex entrance/exit and time allocation increases gradually:

$$p_e = \frac{1}{G - Iter + 1} \qquad (13)$$

5. Field experiment

5.1. Field

The field selected for the experiment is Kulangsu (or Gulangyu), a tiny island of merely 1.88 km^2 located southwest of Xiamen City in China (Fig. 7). The island is a UNESCO World Cultural Heritage Site, surrounded by various attractions such as heritage buildings, beaches, gardens, rugged terrain, and historical sites. The only access to the island is two dedicated ferries for tourists, operating between Kulangsu and Xiamen, with the third one open only to local residents (shown as red dots in Fig. 7). The island is car-free, walking is the only way for tourists to tour around the island. Approximately 90% of the tourists visiting the island chose the one-day tour option, according to official statistics (Zheng & Liao, 2019). A dilemma exists between the number of attractions in Kulangsu and the limited time tourists have for their tour, which makes it an ideal case for our study.
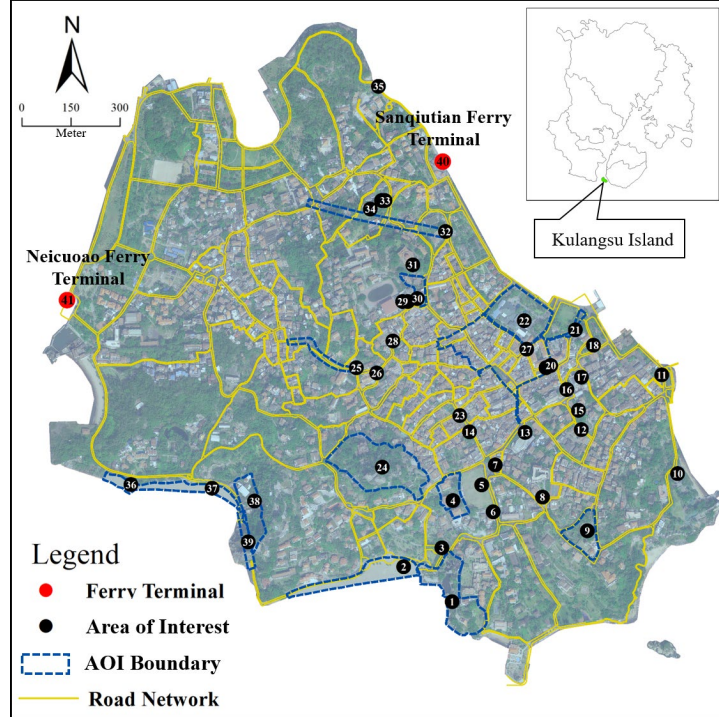

Fig. 7 Map of Kulangsu Island

5.1.1. Basic information about the attractions

Kulangsu is dotted with plenty of attractions. This study selects 39 of the most popular ones for the study. Locations of the 39 attractions are indicated in Fig. 7. The number of entrances and exits for each attraction can significantly affect route planning. The information is listed in the fifth column of Table 3. Note that certain attractions are open areas (e.g., Gangzaihou Seaside Resort). Thus, tourists can enter and leave from anywhere. For these attractions, the number of entrances and exits can be considered infinite. In addition, the average time spent by earlier tourists (t_i) at each attraction can

influence the development of the initial solution in the preprocessing phase. The responses from a survey with tourists and Kulangsu tourism service staff members serve as data on t_i. We randomly conducted several interviews with tourists leaving Kulangsu. They wrote down two kinds of information: (1) the attractions they had visited, and (2) the time they had used at each attraction. We deleted the values that are too large or too small to calculate the average time spent at each attraction. The sixth column in Table 3 presents the results.

Table 3 Attractions in Kulangsu

No	Name	Spatial structure	Time window	Number of entrances/exits	t_j (min)
v_1	Shuzhuang Garden	AOI	[05:00–21:30]	2, 2	60
v_2	Gangzaihou Seaside Resort	AOI	[00:00–24:00]	Infinite, Infinite	10
v_3	International Calligraphy and Carving Gallery	POI	[08:15–18:15]	1, 1	15
….	…..	….	…..	…..	…..
v_{36}	Gulang Rock	POI	[00:00–24:00]	1, 1	10
v_{37}	Merihua Beach	AOI	[00:00–24:00]	Infinite, Infinite	10
v_{38}	Shell Museum	AOI	[08:00–18:00]	1, 1	60
v_{39}	Gusheng Tunnel	LOI	[00:00–24:00]	2, 2	5
v_{40}	Sanqiutian Ferry Terminal	Ferry Terminal	[00:00–24:00]	1, 1	--
v_{41}	Neicuoao Ferry Terminal	Ferry Terminal	[07:20–18:40]	1, 1	--

5.1.2. *Basic information of the participants*

We recruited 100 tourists to participate in our survey at Sanqiutian and Neicuoao Ferry Terminals on August 11, 19, and 27, 2019. We conducted a simple oral interview with the participants. We described to the participants the 39 attractions using pictures and collected the willingness respondents rated to visit each attraction based on a scale from "0" (no interest to visit the attraction) to "1" (the highest interest to visit the attraction). Subsequently, the respondents recorded their time budget. Respondents' demographic information was requested, too. Among all our participants, 37 were male, and 63 were female; 45 were recruited at Sanqiutian Ferry Terminal, whereas 55 were gathered at Neicuoao Ferry Terminal; 39 were recruited for the first survey, 34 during the second, and the remaining 27 were invited for the final survey. Table 4 lists the aforementioned tourist information.

Table 4 Sample tourist preferences and time budget

Tourist	Gender	Preferred attractions	Time budget
1	M	[.50, .36, .29, ..., .20, .23, .29]	5 hours, [8:00–13:00]
2	F	[1.0, .40, .28, …, .94, 1.0, .90]	5 hours,

			[8:00–13:00]
…	…	…	…
100	M	[.63, .88, .64, …, 1.0, .74, .81]	12 hours, [8:00–20:00]

5.2. Algorithm parameters

The performance of our approach can be considerably affected by algorithm parameters, including SN, G, the numbers of EBs (N_{eb}), OBs (N_{ob}), g-SBs ($N_{g\text{-}sb}$), and e-SBs ($N_{e\text{-}sb}$) in the colony, abandonment criteria (*limit*), and differential evolution rate (F_s). Inappropriate SN and G values may reduce the performance of the approach: too small SN and G may increase the risk of reaching a local optimum, whereas too large values may cause computational inefficiency. In general, N_{eb} and N_{ob} are the same and equal to SN, that is, $N_{eb} = N_{ob} = SN$. Moreover, N_{sb} is usually set to 1 (Guo & Zhang, 2017; Kıran, et al., 2013). In this study, a new group of bees called *e-SBs* is designed to optimize the choice of entrances/exits of attractions. $N_{e\text{-}sb}$ equals to N_{eb}, as *e-SB* is mapped one-to-one with *EB*. The parameter *limit* is set to $SN \times D$, where D is the dimension of the solution and equals to the number of vertices (N). The parameters of our algorithm are presented in Table 5.

Table 5 Algorithm parameters

	SN	G	N_{eb}	N_{ob}	$N_{g\text{-}sb}$	$N_{e\text{-}sb}$	*Limt*	F_s
Value	15	10000	15	15	1	15	615	0.1

5.3. Model performance evaluation

Considering the personal characteristics, preferences, and constraints of the 100 participants (Table 4), the essential information of Kulangsu (Table 3), and algorithm parameters (Table 5), we designed tour routes using our method along with four baseline methods, namely, iteration local search (ILS), standard genetic algorithm (sGA), particle swarm optimization (PSO), and ant colony optimization (ACO). To prevent random errors, each algorithm creates the route for each tourist 30 times and averages the total utility of results 30 times (Karaboga, 2005). Figure 8 shows the average utility obtained by each tourist through these five methods.
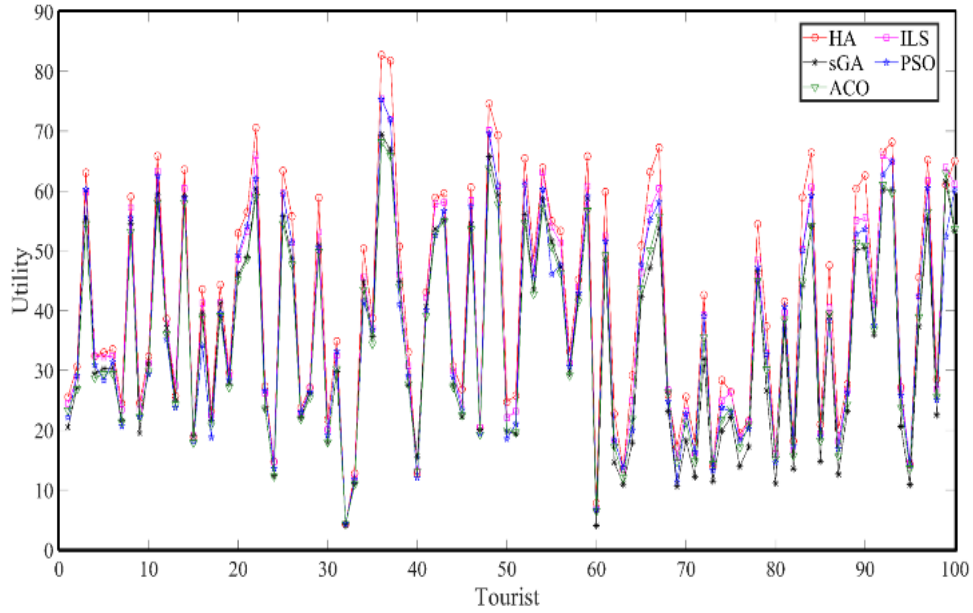
Fig. 8 Average utility by methods (HA, ILS, sGA, PSO, and ACO)

Several paired sample *t*-tests were run to examine whether any statistical differences exist among the utility achieved by the five methods of HA, ILS, sGA, PSO, and ACO. Table 6 lists the means and standard deviations of the utility generated from the five methods. Table 7 shows the results of the paired t-tests. For the first pair (HA–ILS), the gap mean was 2.442, and HA reached a significantly higher utility ($M = 40.811$, $SD = 19.352$) than ILS ($M = 38.369$, $SD = 17.887$) ($t(100) = 10.222$, $p < 0.05$). Analogously, the second (HA–sGA), third (HA–PSO), and fourth pairs (HA–ACO) showed an obvious advantage of our proposed method over sGA, PSO, and ACO in improved utility.

Table 6 Basic statistics of paired samples

		Mean	N	Standard deviation	Standard error mean
Pair 1	HA	40.811	100	19.352	1.935
	ILS	38.369	100	17.887	1.789
Pair 2	HA	40.811	100	19.352	1.935
	sGA	35.006	100	17.301	1.730
Pair 3	HA	40.811	100	19.352	1.935
	PSO	36.847	100	17.675	1.768
Pair 4	HA	40.811	100	19.352	1.935
	ACO	35.416	100	16.449	1.645

Table 7 Results of paired sample t-tests

	Paired differences				t	df	Sig. (2-tailed)
	Mean	Standard	Standard	95% Confidence			

		deviation	error mean	interval of the difference				
				Lower	Upper			
Pair 1	HA–ILS	2.442	2.389	0.239	1.968	2.916	10.222	99	0.000***
Pair 2	HA–sGA	5.805	3.655	0.365	5.080	6.530	15.883	99	0.000***
Pair 2	HA–PSO	3.964	2.751	0.275	3.418	4.510	14.407	99	0.000***
Pair 3	HA–ACO	5.395	3.690	0.369	4.663	6.127	14.621	99	0.000***

$*p < 0.05, **p < 0.01, ***p < 0.001.$

5.4. Model efficiency evaluation

We now compared our proposed approach with the above four algorithms (ILS, GA, PSO, and ACO) to analyze the efficiency of HA. The two tourists from each group with a time budget of 4, 8, and 12 hours were selected as the test samples, and we ran each algorithm 30 times. To depict the relationship between the result and iterations, the optimization history of HA is provided in Fig. 9. As our approach pays attention to the optimization of discrete and continuous variables and provides more opportunities to escape from local optimality, the efficiency of the algorithm may be limited. Therefore, we flexibly adjusted the number of iterations to 2,000 (denoted as LHA) and kept other parameters of the algorithm unchanged, to achieve a better trade-off between efficiency and performance.
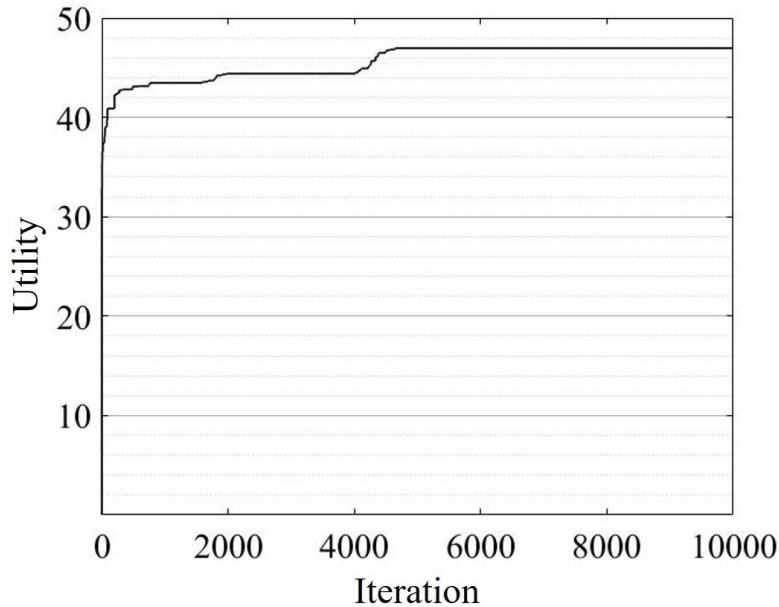

Fig. 9 Optimization history of HA

Table 8 shows the results of five methods for various time budgets T_{max} and the corresponding number of tourists where the average utility U and the average computational time T are reported. The performance of HA and LHA is optimal for all

the runs, especially in optimizing the route for a longer time budget, whereas ILS shows great advantages in computational time. The structure of our approach, which focuses on finding the best choice of entrance and exit and the duration at attractions, becomes very time consuming to calculate.

Table 8 Comparison of efficiency between different methods

		HA		LHA		ILS		GA		PSO		ACO	
T_{max}	No.	U	T	U	T	U	T	U	T	U	T	U	T
4H	28	23.53	10.93	23.21	2.54	20.83	0.09	18.09	3.20	18.02	2.75	19.35	3.37
4H	77	21.87	11.42	21.56	2.54	21.44	0.06	17.69	3.16	20.87	2.67	20.34	3.37
8H	16	44.00	13.67	43.11	3.39	38.39	0.05	38.86	3.55	39.30	3.41	35.00	3.70
8H	58	47.71	15.60	46.64	3.17	42.73	0.12	40.29	3.53	41.63	3.34	43.14	3.38
12H	22	72.86	13.46	71.53	3.47	61.38	0.09	58.28	3.97	58.18	4.25	62.55	3.55
12H	35	82.77	13.79	82.17	3.60	72.13	0.10	65.53	4.03	65.52	4.29	72.04	4.39

5.5. Case demonstration

The results indicate the advantage of our approach in accumulating more utility for the tourist over the other four methods, indicating that it can help improve the effectiveness of the tour recommender system by considering the heterogeneity of tourist attractions' spatial structure and the time spent at each attraction. By contrast, most previous studies abstract attractions as pure vertices or pure arcs, ignoring the multiple entrances/exits of attractions, leading to unnecessary detours or additional time en route between attractions. In general, longer time at attraction and less time en route provide greater utility (Zheng, et al., 2020a). Our approach reflects the actual situation of attractions' spatial structure and incorporates time optimization, thus helping tourists avoid unnecessary detours or extra time en route between attractions.

For validation, we performed a comparative test to make a distinction between our approach and ILS presented by Gavalas, et al. (2016). For example, the first tourist in Table 4 plans to visit Kulangsu for 5 hours (from 8:00 to 13:00). We employed HA and ILS to create routes for this tourist, as shown in Fig. 10 (the left shows our proposed HA, whereas the right indicates ILS). As displayed in the figures, compared with ILS, HA designs routes that reduce the amount of time spent on the road, thereby increasing the number and time of visits to attractions for achieving additional utility, as presented in Table 9.

Table 9 Information about tour routes designed by HA and ILS

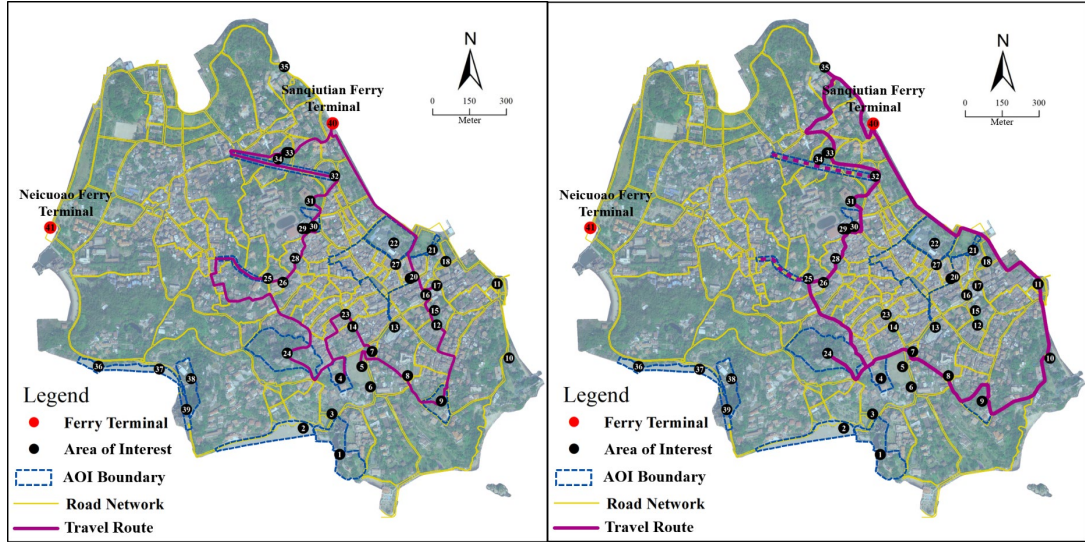| Method	Number of attractions visited	Time spent in the attractions	Time spent on the road	Utility
HA	21	243	57	25.49
ILS	13	225	75	24.26

Fig. 10 Tour routes designed by HA and ILS

Given the full consideration of the attractions' spatial structure, the route designed by our approach can avoid "backtracks" as much as possible. These so-called "backtracks" refer to repeating the route taken to reduce the marginal utility of tourists or cause additional traffic time consumption. For example, both the routes designed by HA and ILS (Fig. 10) include the sub-route "The Bagua Building (Organ Museum) (v_{31}) → Longshan Tunnel (v_{32}) → Epigraphy on Restoration of the Sanhe Taoist Temple (v_{33})." Our approach considers that Longshan Tunnel (v_{32}) is a line-shape attraction and avoids "backtrack" by optimizing the choice of v_{32}'s entrance and exit. By contrast, the route designed by ILS has a "backtrack," which may reduce the marginal utility of the tourist. Figure 11 illustrates the comparison of these two approaches (the left shows our proposed HA, whereas the right indicates ILS). The dotted line in the right figure represents that the tourist walked twice in Longshan Tunnel.
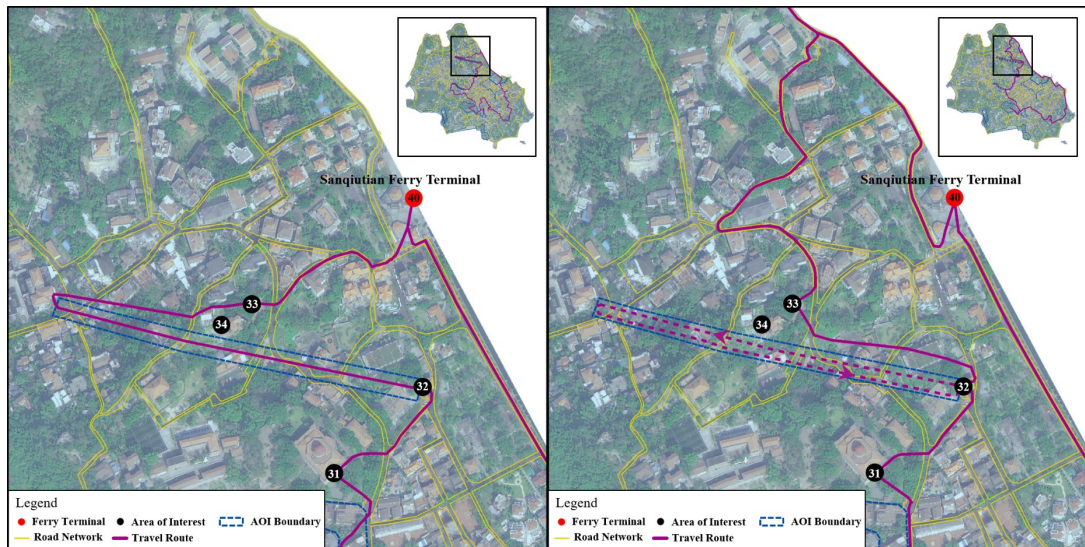

Fig. 11 Tour routes designed by HA and ILS (backtracks)

In addition, certain tourist attractions contain multiple entrances/exits. Therefore, choosing appropriate stations to enter and leave attractions is important in planning trips. Our approach considers the multiple entrances and exits of attractions to avoid unnecessary detours. For example, both the routes designed by HA and ILS (Fig. 10) contain the sub-route "Yu Garden (v_9) → Xiamen Music School (v_8)." Yu Garden (v_9) has three entrances (EN_9^1, EN_9^2, and EN_9^3) and three exits (EX_9^1, EX_9^2, and EX_9^3) (Fig. 12). Our approach optimizes the choice of entrances/exits of attractions. The tourist enters v_9 from EN_9^3 and leaves the attraction from EN_9^1, which is the nearest exit from v_8 (left of Fig. 12). In the route designed by ILS, the tourist enters v_9 from EN_9^2 and leaves the attraction from the same station, adding an unnecessary detour (the yellow curves in the right of Fig. 12). Detours can undoubtedly increase the travel time between attractions. For example, the traffic time between v_9 and v_8 is 1 minute in the route designed by HA, but 4 minutes in the route presented by ILS.
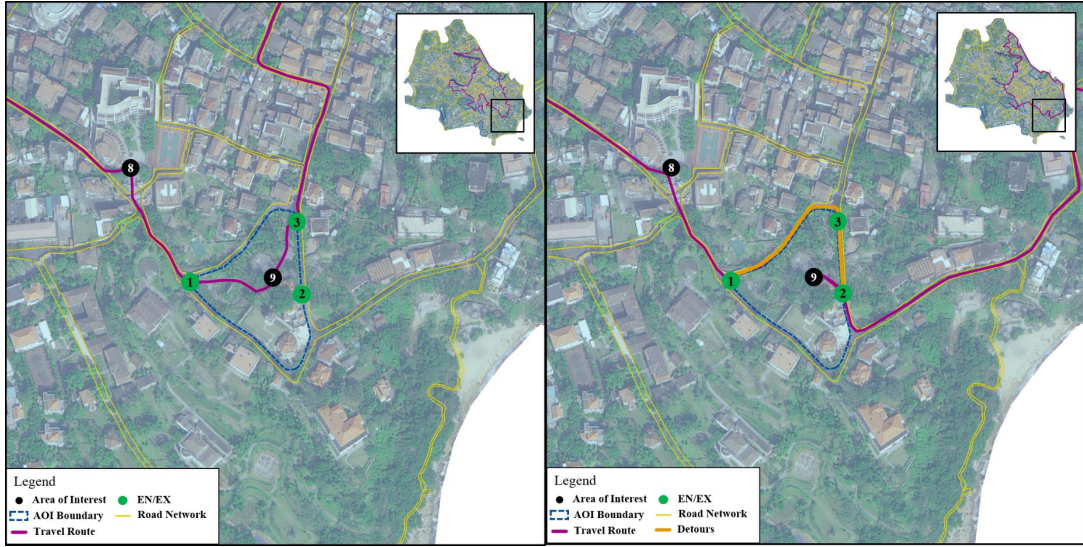

Fig. 12 Tour routes designed by HA and ILS (detours)

6. Discussion and conclusions

Providing information services such as fit-for-purpose recommendations is essential for firms to successfully compete in today's market environment (Feng, et al., 2015; Gorgoglione, et al., 2019). One of the neglected issues in the tour recommender research is the heterogeneity of attractions' spatial structure. Failing to consider it may lead to potentially infeasible or suboptimal recommendations for the users. This study proposes a novel model with solutions coded using a variant-form nectar with four pheromones, optimized with variables combining improved ABC and DEA. Our proposed model adopts various measures to reach an equilibrium of solution quality and algorithm efficiency (e.g., improving the evolution structure, embedding

neighborhood search structure, and adding a new group of bees in the ABC algorithm). We test our new model along with four benchmark models using a field experiment that verifies the superior performance of our model.

This study offers an improved model of tourist trip design that integrates spatial heterogeneity. The design provides significant improvement for tour recommender systems that optimize tourism experience; thus, our model can help tourism organizations provide more enjoyable trips. This is important because system design should be user-centered, and user well-being is of paramount importance. Service providers can integrate our proposed model to improve their existing recommender systems. In addition, our model can be further integrated with other intelligent systems, such as tourist personal conversational agents, which can be applied across multiple channels such as websites, smartphones, kiosks, and service robots to enhance tourist experience, satisfaction, and loyalty.

There are some limitations in this study, which are worth considering in future research. First, although we provided the evidence that the time spent in an attraction can be regarded as a commodity, the utility of stay with time requires further exploration, based on different tourist characteristics and destination characteristics. Second, for island destinations that have multiple ferry terminals, the choices of terminals affect the route structure, thereby increasing the complexity of the design problem. Future studies should take into consideration the choices of ferry terminals in the trip design for island destinations. Finally, as tourists may change their minds owing to the changes in the weather condition, traffic condition, personal issues, tiredness, or mood, future researchers may consider the development of a dynamic tour recommender system that can adjust in real time to the changes of the contexts and tourist preferences.

References

Brewer, P., & Venaik, S. (2010). GLOBE practices and values: A case of diminishing marginal utility? *Journal of International Business Studies, 41*, 1316-1324.

Cenamor, I., de la Rosa, T., Núñez, S., & Borrajo, D. (2017). Planning for tourism routes using social networks. *Expert Systems with Applications, 69*, 1-9.

Chen, Y.-Y., Cheng, A.-J., & Hsu, W. H. (2013). Travel recommendation by mining people attributes and travel group types from community-contributed photos. *IEEE Transactions on Multimedia, 15*, 1283-1295.

Cheng, A.-J., Chen, Y.-Y., Huang, Y.-T., Hsu, W. H., & Liao, H.-Y. M. (2011). Personalized travel recommendation by mining people attributes from community-contributed photos. In *Proceedings of the 19th ACM international conference on Multimedia* (pp. 83-92): ACM.

Cura, T. (2014). An artificial bee colony algorithm approach for the team orienteering problem with time windows. *Computers & Industrial Engineering, 74*, 270-290.

Eby, D. W., & Molnar, L. J. (2002). Importance of scenic byways in route choice: A survey of driving tourists in the United States. *Transportation Research Part A: Policy and Practice, 36*, 95-106.

Feng, H., Tian, J., Wang, H. J., & Li, M. (2015). Personalized recommendations based on time-weighted overlapping community detection. *Information & Management, 52*, 789-800.

Gavalas, D., Kasapakis, V., Konstantopoulos, C., Pantziou, G., & Vathis, N. (2017). Scenic route planning for tourists. *Personal and Ubiquitous Computing, 21*, 137-155.

Gavalas, D., Kasapakis, V., Konstantopoulos, C., Pantziou, G., Vathis, N., & Zaroliagis, C. (2015). The eCOMPASS multimodal tourist tour planner. *Expert Systems with Applications, 42*, 7303-7316.

Gavalas, D., Konstantopoulos, C., Mastakas, K., Pantziou, G., & Vathis, N. (2016). Efficient metaheuristics for the mixed team orienteering problem with time windows. *Algorithms, 9*, 1-21.

Geiger, C. D., & Eskandari, H. (2008). A fast Pareto genetic algorithm approach for solving expensive multiobjective optimization problems. *Journal of Heuristics, 14*, 203-241.

Goldberg, D. (1989). *Genetic algorithms in search, optimization, and machine learning*. Inc., Boston, MA, USA: Addison-Wesley Longman Publishing Co.

Golden, B. L., Levy, L., & Vohra, R. (1987). The orienteering problem. *Naval Research Logistics, 34*, 307-318.

Gorgoglione, M., Panniello, U., & Tuzhilin, A. (2019). Recommendation strategies in personalization applications. *Information & Management, 56*, 103143.

Guo, K., & Zhang, Q. (2017). A discrete artificial bee colony algorithm for the reverse logistics location and routing problem. *International Journal of Information Technology & Decision Making, 16*, 1339-1357.

Hsu, F.-M., Lin, Y.-T., & Ho, T.-K. (2012). Design and implementation of an intelligent recommendation system for tourist attractions: The integration of EBM model, Bayesian network and Google Maps. *Expert Systems with Applications, 39*, 3257-3264.

Huang, X., Li, M., Zhang, J., Zhang, L., Zhang, H., & Yan, S. (2020). Tourists' spatial-temporal behavior patterns in theme parks: A case study of Ocean Park Hong Kong. *Journal of Destination Marketing Management Science, 15*, 100411.

Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization. In: Technical Report TR06, Erciyes University, Engineering Faculty, Computer Engineering Department.

Karaboga, D., Gorkemli, B., Ozturk, C., & Karaboga, N. (2014). A comprehensive survey: Artificial bee colony (ABC) algorithm and applications. *Artificial Intelligence Review, 42*, 21-57.

Kıran, M. S., İşcan, H., & Gündüz, M. (2013). The analysis of discrete artificial bee colony algorithm with neighborhood operator on traveling salesman problem. *Neural Computing and Applications, 23*, 9-21.

Kotiloglu, S., Lappas, T., Pelechrinis, K., & Repoussis, P. P. (2017). Personalized multi-period tour recommendations. *Tourism Management, 62*, 76-88.

Lee, C.-S., Chang, Y.-C., & Wang, M.-H. (2009). Ontological recommendation multi-agent for Tainan City travel. *Expert Systems with Applications, 36*, 6740-6753.

Lee, D., & Hosanagar, K. (2019). How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment. *Information Systems Research, 30*, 239-259.

Liao, Z., & Zheng, W. (2018). Using a heuristic algorithm to design a personalized day tour route in a time-dependent stochastic environment. *Tourism Management, 68*, 284-300.

Lim, K. H., Chan, J., Karunasekera, S., & Leckie, C. (2019). Tour recommendation and trip planning using location-based social media: A survey. *Knowledge and Information Systems, 60*, 1247-1275.

Liu, L., Xu, J., Liao, S. S., & Chen, H. (2014). A real-time personalized route recommendation system for self-drive tourists based on vehicle to vehicle communication. *Expert Systems with Applications, 41*, 3409-3417.

Lu, Y., Jossé, G., Emrich, T., Demiryurek, U., Renz, M., Shahabi, C., & Schubert, M. (2017). Scenic routes now: Efficiently solving the time-dependent arc orienteering problem. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 487-496): ACM.

Lu, Y., & Shahabi, C. (2015). An arc orienteering algorithm to find the most scenic path on a large-scale road network. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 46): ACM.

Majid, A., Chen, L., Mirza, H. T., Hussain, I., & Chen, G. (2015). A system for mining interesting tourist locations and travel sequences from public geo-tagged photos.

Data & Knowledge Engineering, 95, 66-86.

Martín-Moreno, R., & Vega-Rodríguez, M. A. (2018). Multi-objective artificial bee colony algorithm applied to the bi-objective orienteering problem. *Knowledge-Based Systems, 154*, 93-101.

Mrazovic, P., Larriba-Pey, J. L., & Matskin, M. (2017). Improving mobility in smart cities with intelligent tourist trip planning. In *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)* (Vol. 1, pp. 897-907): IEEE.

Omkar, S., Senthilnath, J., Khandelwal, R., Naik, G. N., & Gopalakrishnan, S. (2011). Artificial bee colony (ABC) for multi-objective design optimization of composite structures. *Applied Soft Computing, 11*, 489-499.

Pan, Q.-K., Tasgetiren, M. F., Suganthan, P. N., & Chua, T. J. (2011). A discrete artificial bee colony algorithm for the lot-streaming flow shop scheduling problem. *Information Sciences, 181*, 2455-2468.

Rodriguez, B., Molina, J., Perez, F., & Caballero, R. (2012). Interactive design of personalised tourism routes. *Tourism Management, 33*, 926-940.

Rugg, D. (1973). The choice of journey destination: A theoretical and empirical analysis. *The Review of Economics Statistics*, 64-72.

Si, Y., Zhang, F., & Liu, W. (2019). An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features. *Knowledge-Based Systems, 163*, 267-282.

Souffriau, W., Vansteenwegen, P., Berghe, G. V., & Oudheusden, D. V. (2011). The planning of cycle trips in the province of East Flanders. *OMEGA, 39*, 209-213.

Sun, C. Y., & Lee, A. J. T. (2017). Tour recommendations by mining photo sharing social media. *Decision Support Systems, 101*, 28-39.

Szeto, W. Y., Wu, Y., & Ho, S. C. (2011). An artificial bee colony algorithm for the capacitated vehicle routing problem. *European Journal of Operational Research, 215*, 126-135.

Taylor, K., Lim, K. H., & Chan, J. (2018). Travel itinerary recommendations with must-see points-of-interest. In *Companion Proceedings of the The Web Conference 2018* (pp. 1198-1205): International World Wide Web Conferences Steering Committee.

Tsai, C.-Y., & Chung, S.-H. (2012). A personalized route recommendation service for theme parks using RFID information and tourist behavior. *Decision Support Systems, 52*, 514-527.

Vansteenwegen, P., Souffriau, W., & Van Oudheusden, D. (2011). The orienteering problem: A survey. *European Journal of Operational Research, 209*, 1-10.

Verbeeck, C., Vansteenwegen, P., & Aghezzaf, E. H. (2014). An extension of the arc orienteering problem and its application to cycle trip planning. *Transportation Research Part E, 68*, 64-78.

Zhang, C. Y., Liang, H. W., & Wang, K. (2016). Trip recommendation meets real-world constraints: POI availability, diversity, and traveling time uncertainty. *ACM Transactions on Information Systems, 35*, 1-28.

Zheng, W., Ji, H., Lin, C., Wang, W., & Yu, B. (2020a). Using a heuristic approach to

design personalized urban tourism itineraries with hotel selection. *Tourism Management, 76*, 103956.

Zheng, W., & Liao, Z. (2019). Using a heuristic approach to design personalized tour routes for heterogeneous tourist groups. *Tourism Management, 72*, 313-325.

Zheng, W., Liao, Z., & Lin, Z. (2020b). Navigating through the complex transport system: A heuristic approach for city tourism recommendation. *Tourism Management, 81*, 104162.

Zheng, W., Liao, Z., & Qin, J. (2017). Using a four-step heuristic algorithm to design personalized day tour route within a tourist attraction. *Tourism Management, 62*, 335-349.