ICIS 2022 Proceedings | AI in Business and Society

Dec 12th, 12:00 AM

An Intelligent Customization Framework for Tourist Trip Design Problems

Le Wang
Xi'an Jiaotong University, wangle19@stu.xjtu.edu.cn

Xi Zhao
Xi'an Jiaotong University, zhaoxi1@mail.xjtu.edu.cn

An Intelligent Customization Framework for Tourist Trip Design Problems

Completed Research Paper

Le Wang
School of Management, Xi'an Jiaotong University
Xi'an, China, 710049
wangle19@stu.xjtu.edu.cn

Xi Zhao*
School of Management, Xi'an Jiaotong University
Xi'an, China, 710049
Zhaoxi1@mail.xjtu.edu.cn

Abstract

In the era of the experience economy, "customized tours" and "self-guided tours" have become mainstream. This paper proposes an end-to-end framework for solving the tourist trip design problems (TTDP) using deep reinforcement learning (DRL) and data analysis. The proposed approach considers heterogeneous tourist preferences, customized requirements, and stochastic traffic times in real applications. With various heuristics methods, our approach is scalable without retraining for every new problem instance, which can automatically adapt the solution when the problem constraint changes slightly. We aim to provide websites or users with software tools that make it easier to solve TTDP, promoting the development of smart tourism and customized tourism.

Keywords: Tourist trip design problems, user-generated content, deep reinforcement learning

Introduction

In the era of the experience economy, tourism has become an indispensable part of leisure life. With the popularization of tourism, travelers increasingly prefer "customized tours" and "self-guided tours" to pre-organized tour routes or standard travel packages (Kotiloglu et al. 2017). But in the face of unfamiliar point-of-interest (POI, e.g., tourist attraction, hotel, and restaurant) and a complex tourism environment (e.g., transportation network, time windows), it is difficult for the traveler to plan a satisfactory tourist trip alone. Therefore, travelers will search for various tourist information before planning a tourist trip. Although travel websites provide a great deal of available tourist information (e.g., POI introduction, online reviews, etc.) to help travelers design their trip routes, this travel information is generally scattered, causing a heavy cognitive load.

To reduce travelers' cognitive overload, various tourist recommendation systems are produced to solve tourist trip design problems (TTDP). However, most recommendation systems present several significant challenges. First, the recommended tourist trip only consists of popular attractions, ignoring traveler heterogeneous and customization requirements. Heterogeneous refers to the different experience preferences among tourists. Such as some travelers prefer natural scenery, while others prefer amusement parks. Customization refers to the constraints specified by travelers according to their requirements. Such as travelers may have different time budgets for their trip, or they may have flexible starting and ending locations or times of the trip. Besides, the route recommended by the existing recommendation system ignores the stochastic distance between two points. For example, in real life, the time required between two points might be different at different times, depending on traffic conditions. Therefore, the research

* Corresponding author.

problem of this paper is how to design a planning method to solve the TTDP problem considering tourist heterogeneous preferences, customized requirements, and stochastic traffic times.

Tourists' actual and heterogeneous preferences are ignored in most existing TTDP studies. Benchmark instances are widely utilized to solve the TTDP problem. While travelers' preference values for attractions are randomly generated in these benchmark instances. Recently, questionnaires have been widely used (Zheng et al. 2020) to collect traveler preferences. However, it may be unreasonable because travelers are usually unfamiliar with upcoming tourist destinations. UGC is significant for the hospitality and tourism industries, where intangible products are hard to evaluate before consumption. Hence travelers tend to rely heavily on e-word of mouth to gain sufficient information and have indirect purchasing experiences to reduce their perceived uncertainty. UGC offers a powerful instrument for understanding tourist preferences (Liu et al. 2017). Existing research has revealed that tourists generally rely on UGCs from the same tourist type, and they rarely bother to look at the UGCs provided by other tourist types (Banerjee and Chua 2016). Therefore, the tourist preference for attractions selection is influenced heavily by the same tourist type or profile (Francesco and Roberta 2019; Wang et al. 2020). In addition, studies found that different types of tourists have different behaviors, such as preference, satisfaction, and evaluation patterns (Xu 2019). However, the heterogeneous preference of different tourist types for attractions is rarely researched.

Customized requirements put forward strict performances on planning methods. The TTDP problem is a large-scale optimization problem considering different constraints. The exact algorithms for VRP are divided into two categories: branch and price (Christofides and Eilon 1969), and dynamic programming (Eilon et al. 1974). Exact methods are often not applicable in practice because they cannot solve problems of moderate size within a reasonable amount of time (El-Sherbeny 2010). Heuristic methods can relatively quickly obtain good feasible solutions without an optimality guarantee (Desaulniers et al. 2014), including classical heuristics (Dantzig and Ramser 1959) and meta-heuristic algorithms (Shaw 1998). However, the solution quality of the heuristic varies in different problem sizes and heavily depends on the initial solution and hyper-parameters of the algorithm. The traditional optimization methods do not scale very well when the number of attractions increases and customization requirements (such as different starting and ending locations, and time budgets, etc.) diverse, which is challenging to satisfy the real-time calculation demand of tourists.

Subsequently, a deep reinforcement learning (DRL) model is proposed to solve the TTDP. As far as I know, Gama and Fernandes (2021) first proposed a reinforcement learning approach to solve the TTDP. Bliek et al. (2022) recently developed DRL algorithms to solve stochastic routing problems. However, both ignored edge information's influence in representation learning. The edge usually represents the traffic time between two locations in the route planning problem, an essential influencing factor. In real situations, traffic congestion is an unavoidable problem, especially during rush hours and holidays. A planning model that does not consider actual traffic conditions is not available. However, existing DRL models mainly extract attraction features as input and ignore the additional information brought by edge features.

To handle the above challenges and research gaps, this paper proposes an intelligent customization framework to solve the TTDP problem considering tourist heterogeneous preferences, customized requirements, and stochastic traffic times. First, a heterogeneous preference module is designed to mine tourists' score of attraction from user-generated content. Tourists' preferences for different attractions can be represented as the user's behavioral intention to experience the attraction. The theory of rational action (TRA) (Fishbein and Ajzen 1977) was developed to predict behavioral intentions. Thus, the attractions score module considering TRA is designed. Subsequently, a new DRL architecture considering edge information is constructed. We not only introduce edge information features into the DRL model but also modify the deep learning architecture to make DRL better represent the tourism environment. Finally, three case studies are presented to verify the effectiveness of our proposed methods for TTDP. These case studies include the actual travel website routes, real travel information obtained based on travelogue, and existing data sets.

Related Work

In this section, we will review the related literature, specifically on (1) tourist trip design problem, (2) DRL algorithm for routing optimization problem, and (3) user-generated content. By comparing with the existing literature, we will summarize our main contributions.

TTDP-Related Researches

TTDP is an extension of tourism's orienteering problem (OP). The problem aims to design tour routes for travelers according to their preferences and requirements to maximize total score while considering numerous constraints (Vansteenwegen and Van Oudheusden 2007; Zheng et al. 2020). TTDP is a convenient and cost-effective exploration of a tourist destination for many travelers (Kotiloglu et al. 2017). Recently, the study of trip planning and personalized tour guide has attracted significant attention (Kotiloglu et al. 2017). Rodriguez et al. (2012) proposed a multiple-criteria decision-making method to design a customized trip with many conflicting objectives. To improve the operating efficiency of the algorithm with large numbers of attractions, Kotiloglu et al. (2017) proposed a “filer-first, tour-second” framework to generate personalized tour recommendations for tourists based on online data sources. They first identified a subset of attractions by collaborative filtering method, and then Iterated Tabu Search algorithms were introduced to generate tours. In addition, considering that tourism activities are usually group-oriented, and the group's preferences may be completely different, Zheng and Liao (2019) proposed a colony optimization and evolution algorithm to handle personalized tourism route design for heterogeneous tourism groups. However, existing TTDP solution methods almost utilized benchmark instances while ignoring the importance of travelers’ preference for attractions. In addition, traditional algorithms have poor scalability and need to be retrained on a per-problem.

DRL Algorithms for Routing Optimization Problems

DRL is a combination of deep learning (DL) (Law et al. 2019) and reinforcement learning (RL) (Kaelbling et al. 1996). DL has a powerful representational ability (Law et al. 2019) to represent the real, complex, and dynamic tourism environment. While RL, inspired by behavioral psychology, can be used to simulate the human’s ability of goal-oriented self-learning (Shin et al. 2012).

The first attempt to use a neural network to tackle routing optimization problems is the Hopfield-Network (Hopfield and Tank 1985). However, it is trained to solve one instance and has little advantage over traditional methods. Based on the sequence-to-sequence model widely used in natural language processing, Vinyals et al. (2015) proposed the pointer network (PN) trained with supervised learning for routing optimization problems. However, supervised learning depends on optimal solutions such as labels, which are hard to obtain for routing optimization problems. To overcome these limitations, Bello et al. (2016) first proposed to use DRL to train the PN in an unsupervised manner, which learns from the reward without the optimal solution in advance. This opens the way to solving broader routing optimization problems. Nazari et al. (Nazari et al. 2018) replaced the LSTM encoder of the PN with a simple graph embedding layer and first introduced dynamic elements to the attention mechanism. Recently, inspired by the transformer architecture (Vaswani et al. 2017) utilized in natural language processing (NLP), Deudon et al. (2018), Kool et al. (2018), Gama and Fernandes (2021) applied the multi-head attention mechanism of the transformer to tackle the routing optimization challenge by modifying the encoder of the PN. The original transformer was designed for NLP, but different from NLP, edge information is one of the essential elements in routing optimization problems. Therefore, the architecture for routing optimization problems that do not leverage the edge information may perform poorly when the edge information is essential.

User-Generated Content

UGCs have a significant influence on tourist tour route planning. Previous research has revealed that tourists’ preferences and behaviors in attraction selection are heavily influenced by tourist types or purposes (Francesco and Roberta 2019; Wang et al. 2020). Researchers (Banerjee and Chua ; Wang et al. 2020) found different types of tourists have different evaluation and satisfaction patterns for the same hotel. For example, a minor flaw found by one tourist type may be unacceptable for another type (Banerjee and Chua 2016). Xu (2019) proved that the overall satisfaction of hotels varies with the type of tourists. In addition, when tourists select attractions, they are more likely to be guided generally by the content from the same tourist types (Wang et al. 2020). Banerjee and Chua (2016) found that most tourists only pay attention to the ratings of their type of tourists and proved that tourists rarely make an effort to read the UGC posted by different types of tourists.

Although most of the research focuses on the differences of preference of hotel selection, the studies of tourist heterogeneity in route design are few. Most recent studies about route design are based on questionnaires (Liao and Zheng 2018; Rodriguez et al. 2012; Zheng et al. 2020; Zheng and Liao 2019). Such as, Zheng et al. (2020) collected tourist information through questionnaires and oral interviews in situations and airports. Intuitively, we aversion the people chatting us up at the station and airport. And tourist are usually unfamiliar with the upcoming tourist destinations, so tourists' preferences for attractions given by questionnaire are biased and unreal. UGCs have a large amount of data and are public and generated voluntarily by actual consumers (Raghupathi et al. 2015). This study shed light on the heterogeneity preference and satisfaction of tourist types in TTDP based on UGCs.

An Intelligent Customization Framework for TTDP

In this section, we first define some preliminaries of TTDP and describe the overall design framework. Then, we describe the framework in detail.

Problem Description

The TTDP problem can be represented as a given graph $G = (V, E)$, where V represents a set of attractions and E represents the travel time t_{ij} between each pair of attractions. V is composed of two subsets, which are starting and ending locations $[v_0, v_1]$, and N attractions $[v_2, v_3, \dots, v_{N+1}]$. Each attraction has attribute information, which contains the longitude and latitude coordinates of geographical location $[long_i, lat_i]$, a visiting time window $[open_i, close_i]$ with opening time (the earliest visiting time) and closing time (the latest time to end the visit), and visit duration d_i. In addition to attraction's attribute information, it also contains traveler customization and personalization information. Customization information refers traveler's initial location v_1 and final location v_2, start time t_{start} and end time t_{end}, and days of travel D. Personalization information refers to the score or reward r_i that travelers assign to each attraction based on their historic preferences. Besides, the time between the two locations is uncertain and stochastic, affected by traffic conditions.

Based on the given information and constraints, the objective of TTDP is to find a multi-day tourist route sequence with the maximum sum of scores without repeat visits. To solve this problem, we proposed an intelligent customization framework depicted in Figure 1. First, the attractions score module is designed based on TRA to obtain tourist heterogeneous preferences. Then a route design module is constructed based on the DRL model considering edge information to solve the optimal tour route to account for stochastic traffic times.

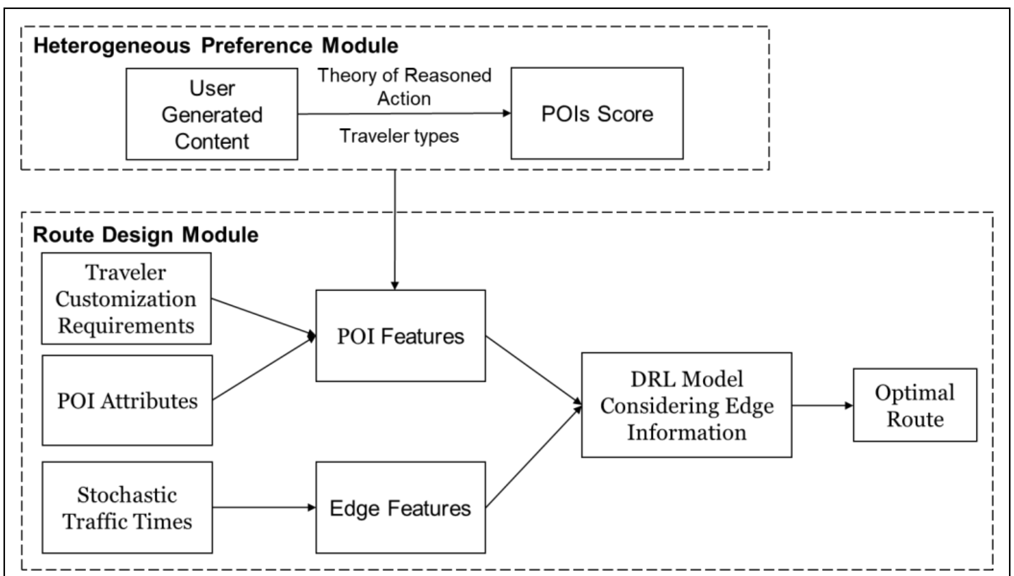

Figure 1. Intelligent Customization Framework for TTDP

Heterogeneous Preference Module

The purpose of heterogeneous preference module is to obtain tourists' scores of attraction. Xiang et al. (2017) demonstrated that UGC from a single website could not be an abundant source of quality data because different websites may possess unique characteristics. Therefore, to improve the generalization of research results, this study analyses data of UGC from two popular tourism websites, TripAdvisor.com and Ctrip.com. Developing machine learning algorithms, such as text and word frequency analysis, provides powerful tool support for preference analysis of different tourist types. Hence, we analyze the heterogeneous preference of different tourist types hidden in UGC by machine learning algorithms.

From a theoretical perspective, TRA is utilized to obtain the heterogeneous preference of attraction. Initially, TRA was developed to predict behavioral intentions (Fishbein and Ajzen 1977). It suggests that an individual's behavioral intention depends on their attitude about the behavior and subjective norms. Attitude indicates a person's emotional polarity, positive or negative, toward the behavior (Miller 2005). At the same time, subjective norms mean the influence of other people's opinions on the behavior (Miller 2005). Simply, the TRA can be expressed as $\mathrm{BI} = \mathrm{w}_1 \times A + w_2 \times SN$ (Huang et al. 2017; Rehman et al. 2007). Where BI is behavioral intention; A and SN present the attitude and the subjective norms related to performing the behavior, respectively; and w_1 and w_2 are the weight of AB and SN, which depend on the individual.

In our paper, tourist preference can be represent as the user's behavioral intention to experience the attraction. The emotional score of tourist can be used to measure the attitude, and the frequency of visits by all tourist can be used to measure subjective norms (Huang et al. 2017). In the following, we will analyze how to obtain the attraction score for different tourist types based on TRA in detail. The traveler types are denoted as k=[1,2,3,4]=[couples, families, friends, solo].

Attitude: Travel online reviews contain numerical ratings and text comments, both used to express tourists' emotional score with attraction. The number rating represents the overall evaluation of the attraction by tourists. While, the review text is the supplement of the number rating, which provides a more granular evaluation. The combination of the two will reflect the tourists' emotion towards the attraction more accurately. Therefore, the mean of the numerical rating and text sentiment to represent the user emotional score to measure attitude.

Numerical rating is based on the 5-point rating scale, where 5 represents the highest degree of satisfaction, and 1 represents the lowest degree of satisfaction. We first calculate the normalized the number rating. The normalized rating for traveler type k on the attraction p_i can be calculated by Eq. (1).

$$rating_i^{*k} = \frac{rating_i^k - \min\limits_{i=h,\dots,h+n-1} rating_i^k}{\max\limits_{i=h,\dots,h+n} rating_i^k - \min\limits_{i=h,\dots,h+n-1} rating_i^k} \tag{1}$$

where $rating_i^k = \sum_{m=1}^{M} q_{i,m}^k / M$, $q_{i,m}^k$ is the rating value for attraction p_i given by the mth traveler of the k traveler type, and M is the total number of travelers who post the rating.

Text sentiment analysis is used to mine process text comments. We store the text comments of travelers type k on the attraction p_i in a text document (txt_i^k), and then after text pre-processing (e.g., tokenization and elimination of stop words), the module package of SnowNLP in Python is used to calculate sentiment value ($sentiment_i^k$) of traveler type k on the attraction p_i. Similar to the Eq. (1), the normalized sentiment value ($sentiment_i^{*k}$) also can be calculated.

Then, the normalized attitude (A_i^{*k}) of traveler type k on the attraction p_i is calculated by Eq. (2).

$$\mathrm{A}_i^{*k} = w_1 \times rating_i^{*k} + w_2 \times sentiment_i^{*k} \tag{2}$$

where w_1, w_2 ($w_1 + w_2 = 1$) are the weights of $rating_i^{*k}$ and $sentiment_i^{*k}$. Travelers can flexibly assign the weights based on their preference and their profile. In general, $w_1 = w_2 = 0.5$.

Subjective norms: The subjective norms (SN_i^k) of traveler type k on the attraction p_i is obtained to extract the mentioned frequency of attractions from the travelogues using the TF-IDF algorithm. We store the travelogues of traveler type k in a text document ($note_i^k$), then after text pre-processing, the module package of sklearn in Python is used to calculate the SN_i^k assigned to the attraction p_i by traveler type k. Similar to Eq. (1), the normalized subjective norms (SN_i^{*k}) can be calculated.

Finally, the attraction score $score_i^k$ of traveler type k on the attraction p_i will be calculated based on Eq. (3).

$$score_i^k = \theta_1 \times A_i^{*k} + \theta_2 \times SN_i^{*k} \tag{3}$$

where θ_1, θ_2 ($\theta_1 + \theta_2 = 1$) are the weights of A_i^{*k} and SN_i^{*k}. In general, $\theta_1 = \theta_2 = 0.5$.

Route Design Module

This section describes the route design module of TTDP from a DRL perspective. First, the TTDP problem is constructed as an MDP formulation. And then, the network structure and training process of DRL will be described.

MDP Formulation

In solving TTDP, a preceding attraction determines the area to locate its succeeding attraction, and the subsequent attraction locating is influenced by the previous attractions. Therefore, the route generation process of TTDP can be regarded as a sequential decision on which attraction to select, which can be naturally formulated as the form of MDP.

Let $\mathcal{S}$ be the state space and $\mathcal{A}$ be the action space. Each state $s_t \in \mathcal{S}$ represents the partial route Y_t constructed at time step t, where Y_t contains all visited attractions until step t, and Y_0 refers to an empty set. Each action $a_t \in \mathcal{A}$ represents selecting attraction p_j at step t, and a_0 refers to the starting location. Executing the action a_t (adds attraction to the end of the partial route Y_t), the state transits to the next state s_{t+1} (partial route Y_{t+1}), and receives the reward (score) r_{a_t} at the same time. This process continues until the ending location is selected to obtain the completed tour route $Y_T = [a_0, a_1, \dots, a_T]$ and the cumulative reward $R(Y|X) = \sum_{t=1}^{T} r_{a_t}$ of tour route Y under the given a TTDP instance X.

From a DRL perspective, we assume that the tour route sequence $Y_T = [a_0, a_1, \dots, a_T]$ is generated step by step through the interaction between agent and environment. The tourism environment contains TTDP instance information X and state s_t information. Instance information X mainly consists of all attractions information (i.e., location, time window, visit duration, and travel time), traveler preference information (i.e., the score of attractions), and requirement information (i.e., the starting and ending location and time) given by travelers. The state information s_t is defined as the environment information under the partial route Y_{t+1}. The agent, relying on a policy $\pi_\theta(Y|X)$, seeks to generate a tour route sequence Y given a TTDP instance information X. According to the probability chain rule (Sutskever et al. 2014), the policy $\pi_\theta(Y|X)$ can be factorized as Eq. (4).

$$\pi_\theta(Y|X) = \prod_{t=1}^{T} \pi_\theta(a_t|X, s_t) \tag{4}$$

where $\pi_\theta(a_t|X, s_t)$ is a single-step assignment policy parameterized by parameter θ, which is a probability distribution over candidate attractions a_t given a TTDP instance X and the state s_t information.

Network Architecture

Policy $\pi_\theta(Y|X)$ is the main element of the travel agent with parameters θ . We design the policy network as an encoder-decoder architecture depicted in Figure 2. The encoder network first takes the instance information and environment state as inputs to obtain a representation for each attraction. The decoder produces the route sequence Y_T of input attractions step by step. Next, we introduce the proposed encoder and decoder details, respectively.

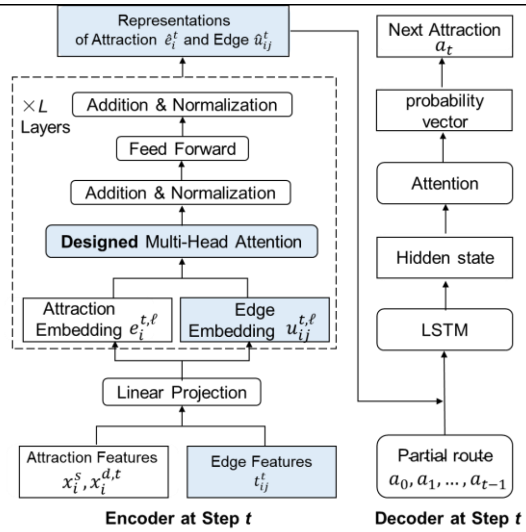

Figure 2. An Encoder-Decoder Architecture considering edge features. The right-angle box represents the data or vector, the rounded box represents various deep learning layers, and the blue represents the innovation of the model in this paper.

Encoder. The purpose of the encoder is to produce a representation for each attractions given the raw features (e.g., location, score, time budget, etc.) of a problem instance. The first layer of the encoder is (1) a linear projection layer that projects the raw features into a higher-dimension space. Then (2) an attention mechanism is proposed to capture the interaction between arbitrary two attractions for better feature extraction.

(1) In the linear projection layer, we output attraction and edge features' embedding vectors. Like Gama and Fernandes (2021), we divide attraction features into static and dynamic groups since some features change dynamically during the route generation, such as time budgets. We process the static and dynamic features separately before concatenating. Specifically, for step t, we map static features x_i^s, dynamic features $x_i^{d,t}$, and edge features t_{ij}^t into d-dimensional hidden features e_i^s, $e_i^{d,t}$, and g_{ij}^t based on Eq. (5).

$$e_i^s = \tanh(W_s x_i^s + b_s),\ e_i^{d,t} = \tanh\left(W_d x_i^d + b_d\right),\ g_{ij}^t = \tanh\left(W_t t_{ij}^t + b_t\right) \tag{5}$$

where W_s, W_s, W_u, b_s, b_d and b_u are the parameters of the linear projection layers. Then the feature embedding $e_i^t = \left[e_i^s, e_i^{d,t}\right]$ of step, t is obtained by concatenating static and dynamic feature embedding. Finally, the feature embedding vector $[e_0^t, e_1^t, \ldots, e_{H+N+1}^t]$ of the attraction set $[v_0, \ldots, v_{H+N+1}]$ and edge embedding vector g_{ij}^t is obtained.

(2) The attention mechanism layer consists of L stacks of identical layers that include a multi-head attention (MHA) mechanism and a feed-forward network (FFN) (Gama and Fernandes 2021), and both are followed by a residual-connection and layer normalization (LN), as shown in Eqs. (6)-(7).

$$\tilde{e}_i^{t,\ell} = LN^\ell(\hat{e}_i^{t,\ell} + MHA_i^\ell\left(e_0^{t,\ell-1}, \ldots, e_{H+N+1}^{t,\ell-1}\right)) \tag{6}$$

$$e_i^{t,\ell} = LN^\ell(\tilde{e}_i^{t,\ell} + FFN^\ell(\tilde{e}_i^{t,\ell})) \tag{7}$$

where ℓ indicates the index of the identical layer. For the MHA sub-layer, our architecture is extended to edge feature representation. To make it easier for readers to understand how to extend edge features in MHA, we first introduce the original MHA (the black solid lines in Figure 3) and then further introduce the MHA with edge feature (the black dotted lines in Figure 3). To simplify, we use one head of MHA as an example to describe in the following.

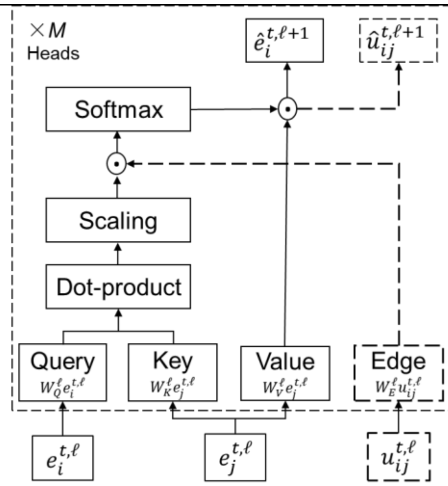

Figure 3. MHA with Edge Feature. The black dotted lines are the difference from the original MHA.

Original MHA. For a layer ℓ, given the feature embedding vector $[e_0^{t,\ell}, e_1^{t,\ell},..., e_{H+N+1}^{t,\ell}]$, the attention score $a_{ij}^{t,\ell}$ between $e_i^{t,\ell}$ and $e_j^{t,\ell}$ is obtained based on Eq.(8) through the softmax normalization of the correlation scores c_ij^(t,l), which is calculated by scaled dot-product attention. Then the attraction feature representation vector $\hat{e}_i^{t,\ell+1}$ is calculated based on Eq. (9) through the weighted sum of the vector $W_V^\ell e_j^{t,\ell}$ and the attention score $u_{ij}^{t,\ell}$.

$$u_{ij}^{t,\ell} = softmax\left(c_{ij}^{t,\ell}\right), \text{with } c_{ij}^{t,\ell} = \frac{W_Q^\ell e_i^{t,\ell} \cdot W_K^\ell e_j^{t,\ell}}{\sqrt{d}}, \forall i,j \in [0, \dots\dots, H+N+1] \tag{8}$$

$$\hat{e}_i^{t,\ell+1} = \sum\nolimits_j u_{ij}^\ell W_V^\ell e_j^{t,\ell} \tag{9}$$

where $\sqrt{d}$ is a scaling constant, W_Q^ℓ, W_K^ℓ and W_V^ℓ are parameters of the ℓ*-th* layer, which are usually called a query, key, and value.

Designed MHA with edge feature. The MHA with edge feature is extended to edge feature representation better to capture the interaction or correlation of each attractions pair. The basic idea of MHA with edge feature is to fuse edge feature embedding g_{ij} with the correlation score c_{ij} by addition, thereby improved attention scores u_{ij}. Then, the attraction feature representation $\hat{e}_i^{t,\ell+1}$ and edge representation $\hat{g}_{ij}^{t,\ell+1}$ are formed, see Eqs. (10)-(11).

$$u_{ij}^{t,\ell} = softmax(c_{ij}^{t,\ell}), \text{with } c_{ij}^{t,\ell} = \left(\frac{W_Q^\ell e_i^{t,\ell} \cdot W_K^\ell e_j^{t,\ell}}{\sqrt{d}}\right) + W_E^\ell g_{ij}^{t,\ell} \tag{10}$$

$$\hat{e}_i^{t,\ell+1} = W_e \sum\nolimits_j u_{ij}^{t,\ell} W_V^\ell e_j^{t,\ell} + b_e, \hat{g}_{ij}^{t,\ell+1} = W_g u_{ij}^{t,\ell} + b_g \tag{11}$$

where $W_E^\ell e_{ij}^{t,\ell}$ denotes the edge value, W_e and W_g are the transformation matrix and b_e and b_g are the bias vector. Adding edge information into the calculation of attention score can add more spatial information for attractions. After the last layer of MHA, the representation vector $[\hat{e}_0^t, \hat{e}_1^t,..., \hat{e}_{H+N+1}^t]$ of all attractions and the edge representation $\hat{u}_{ij}^t$ are output, which are utilized to represent each attraction and as input to a decoder.

Decoder. Decoding happens sequentially. At step t, the decoder outputs the attraction a_t based on the attractions and edge representation vectors outputted by the encoder and the summary information of preciously selected attractions. The decoder consists of an (1) LSTM layer and (2) an attention mechanism layer. At step t, the LSTM layer is utilized to obtain the summary information h_t^d of preciously selected attractions. Then based on an attention mechanism, the probability distribution of the next candidate attractions is generated.

Training Method of Policy Gradient

We expect the best policy $\pi_\theta^*(Y|X)$ can generate the tour route with the maximum score with high probability. To find an optimal policy $\pi_\theta^*(Y|X)$, the goal of DRL is to maximize the expected cumulative reward $\mathbb{E}_{Y\sim\pi_\theta}R(Y|X)$ given a problem instan ce X, defined as Eq.(12).

$$J(\theta|X) = \mathbb{E}_{Y\sim\pi_\theta}R(Y|X) \tag{12}$$

We use the policy gradient method and gradient descent to maximize $J(\theta)$. The gradient of Eq. (12) is formulated using a well-known REINFORCE algorithm (Williams 1998) with Monte Carlo sampling, as shown in Eq. (13).

$$\nabla_\theta J(\theta|X) \approx \frac{1}{B}\sum_{i=1}^{B}\big(R(Y_i|X_i) - b(X_i)\big)\,\nabla_\theta log\pi_\theta(Y_i|X_i) \tag{13}$$

where B is the batch size. $b(X)$ denotes a baseline function, which plays the role of critic and is used to reduce training variance. If the actions perform better than the baseline, the policy will increase the probability of these actions. In this paper, $b(X)$ demonstrates the batch average cumulative reward.

Case studies

As a demonstration of the proposed tour route designation framework, we focused on Beijing in China. Beijing is considered one of China's top ten tourism cities released by the 2018 Global Tourism Destination Marketing Summit and World Culture Tourism Forum (Hou et al. 2019). As China's capital city, government center and business hub, Beijing is the most international and busiest cities in China and has the most significant transportation hub and network.

Heterogeneous Preference Analysis

Online reviews on TripAdvisor.com and travelogues on Ctrip.com are collected as UGC, used to calculate attraction scores. For online review, we first crawled the UGC of attractions from 2010/5/1 to 2020/5/31 of the selected 64 attractions. Through data cleansing, i.e., deleting data with missing values, we collected 63228 reviews from 30928 travelers. For travelogues, we total crawled 9000 travelogues.

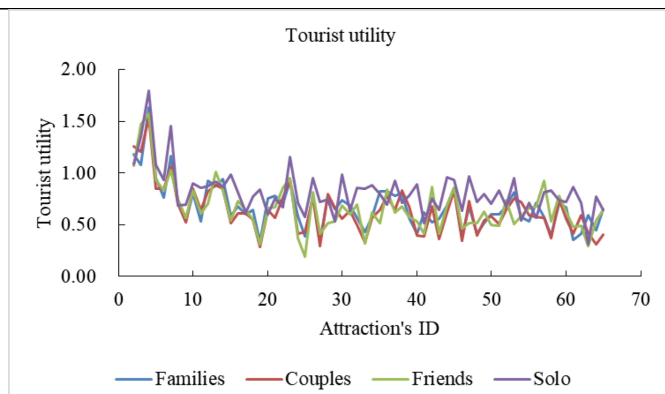

Figure 4. Attraction Score for Attractions with respect to the Four Traveler Types

By processing the UGC on TripAdvisor.com and Ctrip.com, we obtain the normalized score for each attraction with respect to the four traveler types, as shown in Figure 4. As shown in Figure 4, for the same attraction, the normalized scores of the solo are mostly higher than the other three traveler types. Therefore, solo travelers prefer a high-values pattern (Radojevic et al. 2015). While the normalized scores for the same attraction of friends and couples are lower due to their low-values pattern. The difference of these scores between couples and friends are the smallest. Based on the scores, ID4 (the Palace Museum) is the attraction with highest score for four tourist types, which are 1.63 (families), 1.53 (couples), 1.57 (friends), and 1.80 (solo). While ID19 (Ming Tombs) is the attraction with lowest score for families and couples, which are 0.34 (families) and 0.28 (couples); ID25 (Happy Valley) is the attraction with lowest score (0.19) for friends; and ID63 (Monument to the People's Heroes) is the attraction with lowest score (0.31) for solo. The above phenomenon indicates that the scores are different among different types of travelers and also proves that it is necessary to consider heterogeneous tourist types.

DRL Experimental Design

The effectiveness of our proposed DRL methods is verified in three instance sets. The first and second instance sets are generated based on website routes and travelogue information. While the third instance set is existing datasets with optimal solutions. In the following, we will detail how to generate the first and the second instance sets.

Instances Generation

For instance, it should have attraction attributes (attraction's ID, locations, sightseeing duration, and time window), attraction score, traffic time matrix, travel days, and start and end locations and times. For the instance sets of website routes and travelogue, the attraction attributes, traffic time matrix, starting and ending locations, and starting time are generated similarly. In contrast, the attraction score, travel day, and ending time are generated differently.

Attraction attributes. Attraction attributes mainly include attraction's ID, sightseeing duration, and time window. For attractions, we select 460 attractions. Subsequently, each attraction's sightseeing duration and time window are crawled from Ctrip.com, Meituan.com, and Baike.baidu.com. attraction's locations (longitude and latitude) are crawled through the API (Application Programming Interface) of amap.com.

Traffic time matrix. Considering the real-time traffic changes, the traffic time matrix is climbed every hour from 8. AM to 8. PM. In this study, we consider taxis the only transportation between each pair of attractions. In practice, we can easily modify the environment code to add more kinds of transportation options.

Starting and ending locations. Although railway stations, high-speed rail stations, and airports are often used as the starting and ending locations of a tour route, we randomly generated the starting and ending points within Beijing in consideration of the diverse requirements of travelers. The Starting and ending locations IDs are assigned 0 and 1, respectively.

Starting times. The starting is a random generation from [8. AM, 12.AM].

Travel day and ending time. For the instance set of website routes, the travel day and ending time are generated based on the website routes and travelogue information, respectively.

Attraction score. The website routes are the same for any tourist, so their attraction scores are generated based on the average score of the four travelers type. The travelogues are divided according to traveler type. Thus their attraction scores are developed according to traveler type.

Training, Validation, and Test Datasets

For the training dataset, the attraction attributes, start and end locations/times are generated in the same way described in Section of Instance Generation. The traffic time matrix is randomly generated between the maximum and minimum values of the traffic time between each pair of attractions. The travel day is randomly generated between the maximum and minimum values of all travel days. Similarly, attraction scores are randomly generated between the maximum and minimum values of four traveler types. The validation and test datasets are generated based on Section of Instance Generation. To ensures that the three data sets are different, the datasets of training, validation, and test are randomly generated based on different random seeds, which are [0, 1000], [1000, 2000], [2000, 3000], respectively.

Metrics

To evaluate the performance of our proposed model, the total score of the solution is exploited as the evaluation metric, which is widely utilized for TTDP (Divsalar et al. 2013; Divsalar et al. 2014; Gama and Fernandes 2021; Zheng et al. 2020). For each generated instance set, the reported score is the mean value of the test sets within that group. Besides, the waste time, average gap, and average time were also selected as metrics for different instance sets according to relevant research. A detailed description of these metrics is given in the Section of DRL performance evaluation.

Implementation Detail

We implement our model in PyTorch and use an Adam optimizer. The setting of hyper-parameters is the same of Gama and Fernandes (2021).

DRL Performance Evaluation

Compared with the Website Routes instance set

In this section, we compare our proposed method with the actual routes to verify the practicability of our proposed method. Through the Ctrip.com website, two types of actual routes are obtained: popular routes (abbreviated to Popular) and free travel products (shortened to Product)1. In total, we climbed 3,074 routes, of which the popular routes account for 74, while the free travel products account for 3,000 routes. We randomly generated the instance based on the Section of Instance Generation. In addition, since users will specify the attractions they must visit, we will randomly select 0%, 25%, and 50% attractions of the natural route to form a must-visit list. These attractions in the must-visit list will be planned in the final solution.

Dataset	Must-visit	Score		Wasted time(hour)	
		Our	Website	Our	Website
Popular	0%	16.06	8.48	1.94	2.64
	25%	15.05	8.48	2.51	2.60
	50%	13.63	8.48	2.28	2.60
Product	0%	41.56	20.52	4.68	6.84
	25%	38.22	20.52	5.55	7.10
	50%	33.18	20.52	7.05	7.36

Table 1. Comparison Results with website routes

As shown in the Table 1, we compared the total score and wasted time of our proposed method with the website routes. The wasted time on website routes refers to waiting for the attractions to open. While the wasted time of our proposed method also includes the time that the planning route ends earlier than the end time given by the tourists. We can find that our proposed methods' score and wasted time are superior to the website routes under the must-visit list of different lengths. As the list length of must-visit increases from 0 to 50%, our proposed method performs progressively worse. That's because user-specified attractions can deviate from the user's starting and ending locations, resulting in suboptimal routes. However, even the worst results of our proposed model are better than the website routes.

Compared with the Travelogue Instance Set

First, we compare and analyze the travelogue instance set with existing reinforcement learning methods. In addition to the reinforcement learning algorithm proposed in this paper, we compared Gama and Fernandes (2021) and Bello et al. (2016) on total score metrics.

As shown in Table 2, these four travelers' highest and lowest scores are solo (21.30) and couples (16.30), respectively, which is consistent with the traveler's comment pattern, such as solo travelers prefer high-couple prefers low-values pattern. In addition, the results, to some extent, prove the effectiveness and rationality of the proposed method. The results show that the proposed method can achieve better solutions for various traveler types than other baseline methods.

1 https://vacations.ctrip.com/list/freetravel/sc1.html?startcity=1

Traveler type	Our	Gama and Fernandes (2021)	Bello et al. (2016)
Families	17.19	16.6	14.36
Couples	16.30	16.1	14.1
Friends	17.29	17.2	15.79
Solo	21.30	19.9	20.54

Table 2. Comparison Results of TTDP. The values in this table represent scores of the solution.

Compared with Existing TTDP-HS Datasets

TTDP-related research scarcely focuses on hotel selection, an essential component of a multi-day trip. McKercher et al. (2012) believed that the hotel location significantly impacts the choice of tourist attractions and time allocation. Lau and Mckercher (2006) stated that hotel location affects the movement patterns of tourists. In turn, tourists' tour route planning impacts their hotel selection. Zheng et al. (2020) found whether the hotel's location is near the attractions is also an essential factor affecting the hotel selection. It is infeasible to design multi-day tour itineraries without considering the hotel selection. However, the relationship between hotel selection and trip design renders TTDP-HS more complex than general TTDP. In this section, we analyze the proposed method for the TTDP-HS. Based on the existing benchmark instances dataset of TTDP-HS, the results of Skewed Variable Neighborhood Search (SVNS) (Divsalar et al. 2013) and Memetic Algorithm (MA) (Divsalar et al. 2014) are given. As shown in Table 3, the first column indicates the number of hotels and trips in each data set. The second column indicates the number of instances in each set. Then the third column displayed the maximum number of the total number of feasible sequences of hotels (TNFS) in each set. Afterward, the average gaps (Average Gaps) between the results obtained by different methods and optimal solutions are presented, that is

$$\frac{optimal\ result - model\ result}{optimal\ result} \times 100 \tag{14}$$

Due to the randomness structure in the MA method, it is applied to each instance three times. The gap describes how far the results are far from the optimality, which is the percentage difference between the method result and the optimal solution. And the last column shows the average computational time (Average Time) spent in solving each instance set by applying different methods.

Data sets	Instances	Max TNFS	Average Gaps (%)				Average Time (s)		
			MA		SVNS	DRL	MA	SVNS	DRL
			Avg	Best					
17 hotel-4 trip	22	4.91×10^3	1.92	1.32	2.66	1.04	6.63	4.95	0.6
17 hotel-5 trip	22	8.35×10^4	2.22	1.4	4.42	1.65	5.41	4.2	0.6
17 hotel-6 trip	22	1.42×10^6	2.55	1.39	6.86	1.46	4.78	4.29	0.6
17 hotel-8 trip	13	4.10×10^8	3.66	2.95	15.54	2.78	5.16	53.62	0.6
17 hotel-10 trip	9	1.19×10^{11}	5.03	3.78	–	3.72	5.04	–	0.6
Average	17.6	2.38×10^{10}	3.08	2.17	7.37	2.13	5.40	16.77	0.6

Table 3. Comparison Results of TTDP-HS

As shown in Table 3, our DRL model outperforms the SVNS and MA in average gaps. Although our DRL model does not outperform the MA' best gap on 17 hotel-5 trips and 17 hotel-6 trips, the average results achieved by the DRL model are much better than SVNS and MA results. In computation time, the computation time of the DRL is always faster than MA. The results show that for instances with a large number of TNFS, the performance of DRL proposed in our study is advantageous. Therefore, our proposed method can be applied directly by tourism websites to generate tour routes in low latency according to the constraints of traveler input.

Conclusions

The TTDP is an essential research topic in tourism management. A customized tour route design algorithm helps tourism websites provide high-quality and customized services to improve tourist satisfaction and competitiveness. The availability of data on websites offers an unprecedented opportunity to study the heterogeneity of tourists in TTDP and explore how to design a more customized, reasonable, and practical tour route for tourists. Our research objective was to design a tour route with the highest score for tourists while accounting for their heterogeneity and personalized requirement. We proposed a customized route design framework –the DRL model– combining big data analysis and artificial intelligence. We performed an empirical study to interpret the proposed framework's practicality and compare it with state-of-the-art methods. The comparison results show that the DRL model outperforms all the baseline models in designing customized, reasonable tour routes. Significantly, the proposed model can be applied directly by tourism websites.

However, our study still has some limitations. First, although this paper considers heterogeneous preferences, it does not realize personalized preferences. Furthermore, the heterogeneous preferences of users are not verified by user usage experiments. In addition, DRL can solve more complex objective function design problems, but only single objective function problems are analyzed in this paper. Finally, the crowding of attractions has become a severe problem restricting the further development of tourism, which is not taken into account in this paper. In the future, we will conduct further research to address the above limitations.

Acknowledgments

The corresponding author is a Tang Scholar. This work is supported by the National Natural Science Foundation of China (Grant No. 72231007), The Key Research and Development Projects of Shaanxi Province (Grant No. 2020ZDLGY09-08).

References

Banerjee, S., and Chua, A. Y. K. 2016. "In Search of Patterns among Travellers' Hotel Ratings in Tripadvisor," Tourism Management (53), pp. 125-131.

Bello, I., Pham, H., Le, Q. V., Norouzi, M., and Bengio, S. 2016. "Neural Combinatorial Optimization with Reinforcement Learning," arXiv preprint arXiv:1611.09940).

Christofides, N., and Eilon, S. 1969. "An Algorithm for the Vehicle-Dispatching Problem," Journal of the Operational Research Society (20:3), pp. 309-318.

Dantzig, G. B., and Ramser, J. H. 1959. "The Truck Dispatching Problem," Management science (6:1), pp. 80-91.

Desaulniers, G., Madsen, O. B., and Ropke, S. 2014. "Chapter 5: The Vehicle Routing Problem with Time Windows," in Vehicle Routing: Problems, Methods, and Applications, Second Edition. SIAM, pp. 119-159.

Deudon, M., Cournut, P., Lacoste, A., Adulyasak, Y., and Rousseau, L.-M. 2018. "Learning Heuristics for the Tsp by Policy Gradient," International conference on the integration of constraint programming, artificial intelligence, and operations research: Springer, pp. 170-181.

Divsalar, A., Vansteenwegen, P., and Cattrysse, D. 2013. "A Variable Neighborhood Search Method for the Orienteering Problem with Hotel Selection," International Journal of Production Economics (145:1), pp. 150-160.

Divsalar, A., Vansteenwegen, P., Sorensen, K., and Cattrysse, D. 2014. "A Memetic Algorithm for the Orienteering Problem with Hotel Selection," European Journal of Operational Research (237:1), pp. 29-49.

Eilon, S., Watson-Gandy, C. D. T., Christofides, N., and de Neufville, R. 1974. "Distribution Management-Mathematical Modelling and Practical Analysis," IEEE Transactions on Systems, Man, and Cybernetics:6), pp. 589-589.

El-Sherbeny, N. A. 2010. "Vehicle Routing with Time Windows: An Overview of Exact, Heuristic and Metaheuristic Methods," Journal of King Saud University-Science (22:3), pp. 123-131.

Fishbein, M., and Ajzen, I. 1977. "Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research," Philosophy and Rhetoric (10:2).

Francesco, G., and Roberta, G. 2019. "Cross-Country Analysis of Perception and Emphasis of Hotel Attributes," Tourism Management (74), pp. 24-42.

Gama, R., and Fernandes, H. L. 2021. "A Reinforcement Learning Approach to the Orienteering Problem with Time Windows," Computers & Operations Research (133), p. 105357.

Hopfield, J. J., and Tank, D. W. 1985. "“Neural” Computation of Decisions in Optimization Problems," Biological cybernetics (52:3), pp. 141-152.

Hou, Z., Cui, F., Meng, Y., Lian, T., and Yu, C. 2019. "Opinion Mining from Online Travel Reviews: A Comparative Analysis of Chinese Major Otas Using Semantic Association Analysis," Tourism Management (74), pp. 276-289.

Huang, X., Luo, L., and Peng, H. 2017. "A Recommendation Approach Based on the Theory of Reasoned Action," 2017 12th International Conference on Computer Science and Education (ICCSE): IEEE, pp. 419-422.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. 1996. "Reinforcement Learning: A Survey," Journal of artificial intelligence research (4), pp. 237-285.

Kool, W., Van Hoof, H., and Welling, M. 2018. "Attention, Learn to Solve Routing Problems!," arXiv preprint arXiv:1803.08475).

Kotiloglu, S., Lappas, T., Pelechrinis, K., and Repoussis, P. P. 2017. "Personalized Multi-Period Tour Recommendations," Tourism Management (62), pp. 76-88.

Lau, G., and Mckercher, B. 2006. "Understanding Tourist Movement Patterns in a Destination: A Gis Approach," Tourism & Hospitality Research (7:1), pp. 39-49.

Law, R., Li, G., Fong, D. K. C., and Han, X. 2019. "Tourism Demand Forecasting: A Deep Learning Approach," Annals of Tourism Research (75), pp. 410-423.

Liao, Z., and Zheng, W. 2018. "Using a Heuristic Algorithm to Design a Personalized Day Tour Route in a Time-Dependent Stochastic Environment," Tourism Management (68), pp. 284-300.

Liu, Y., Teichert, T., Rossi, M., Li, H., and Hu, F. 2017. "Big Data for Big Insights: Investigating Language-Specific Drivers of Hotel Satisfaction with 412,784 User-Generated Reviews," Tourism Management (59), pp. 554-563.

McKercher, B., Shoval, N., Ng, E., and Birenboim, A. 2012. "First and Repeat Visitor Behaviour: Gps Tracking and Gis Analysis in Hong Kong," Tourism Geographies (14:1), pp. 147-161.

Miller, K. 2005. "Communication Theories," USA: Macgraw-Hill).

Nazari, M., Oroojlooy, A., Snyder, L., and Takác, M. 2018. "Reinforcement Learning for Solving the Vehicle Routing Problem," Advances in Neural Information Processing Systems, pp. 9839-9849.

Radojevic, T., Stanisic, N., and Stanic, N. 2015. "Solo Travellers Assign Higher Ratings Than Families: Examining Customer Satisfaction by Demographic Group," Tourism Management Perspectives (16), pp. 247-258.

Raghupathi, D., Yannou, B., Farel, R., and Poirson, E. 2015. "Customer Sentiment Appraisal from User-Generated Product Reviews: A Domain Independent Heuristic Algorithm," International Journal of Interactive Design and Manufacturing (9:3), pp. 201-211.

Rehman, T., McKemey, K., Yates, C., Cooke, R., Garforth, C., Tranter, R., Park, J., and Dorward, P. 2007. "Identifying and Understanding Factors Influencing the Uptake of New Technologies on Dairy Farms in Sw England Using the Theory of Reasoned Action," Agricultural systems (94:2), pp. 281-293.

Rodriguez, B., Molina, J., Perez, F., and Caballero, R. 2012. "Interactive Design of Personalised Tourism Routes," Tourism Management (33:4), pp. 926-940.

Shaw, P. 1998. "Using Constraint Programming and Local Search Methods to Solve Vehicle Routing Problems," International conference on principles and practice of constraint programming: Springer, pp. 417-431.

Shin, M., Ryu, K., and Jung, M. 2012. "Reinforcement Learning Approach to Goal-Regulation in a Self-Evolutionary Manufacturing System," Expert Systems with Applications (39:10), pp. p.8736-8743.

Sutskever, I., Vinyals, O., and Le, Q. V. 2014. "Sequence to Sequence Learning with Neural Networks," Advances in neural information processing systems, pp. 3104-3112.

Vansteenwegen, P., and Van Oudheusden, D. 2007. "The Mobile Tourist Guide: An or Opportunity," Or Insight (20:3), pp. 21-27.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. 2017. "Attention Is All You Need," Advances in neural information processing systems, pp. 5998-6008.

Vinyals, O., Fortunato, M., and Jaitly, N. 2015. "Pointer Networks," Advances in Neural Information Processing Systems (28), pp. 2692–2700.

Wang, L., Wang, X.-k., Peng, J.-j., and Wang, J.-q. 2020. "The Differences in Hotel Selection among Various Types of Travellers: A Comparative Analysis with a Useful Bounded Rationality Behavioural Decision Support Model," Tourism Management (76).

Williams, R. J. 1998. "Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning," Machine Learning (8:3-4), pp. 229-256.

Xiang, Z., Du, Q., Ma, Y., and Fan, W. 2017. "A Comparative Analysis of Major Online Review Platforms: Implications for Social Media Analytics in Hospitality and Tourism," Tourism Management (58), pp. 51-65.

Xu, X. 2019. "Does Traveler Satisfaction Differ in Various Travel Group Compositions? Evidence from Online Reviews," International Journal of Contemporary Hospitality Management (30:3), pp. 1663-1685.

Zheng, W., Ji, H., Lin, C., Wang, W., and Yu, B. 2020. "Using a Heuristic Approach to Design Personalized Urban Tourism Itineraries with Hotel Selection," Tourism Management (76), p. 103956.

Zheng, W., and Liao, Z. 2019. "Using a Heuristic Approach to Design Personalized Tour Routes for Heterogeneous Tourist Groups," Tourism Management (72), pp. 313-325.