A Survey for Self-Driving Travel Itinerary Problem

Lei Li \textsuperscript{1*}, Jiaoman Du\textsuperscript{2}

\textsuperscript{1}Department of Science and Engineering, Hosei University, Koganei, Tokyo, Japan; \textsuperscript{2}Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan

ABSTRACT

Travel Itinerary Problem is a kind of travel plan for travel events, such as sight spot, flight and hotel timetable arrange, in a chronological order of date and time. A self-driving travel itinerary consists of a tour with one or more stops that a traveler takes. These destinations are visited follow one another sequentially in time in a way of self-driving. This survey discusses models and algorithms for recent research of the self-driving travel itinerary problem.

Keywords: Travel itinerary; Algorithms; Functionalities; Self-drive

INTRODUCTION

Travel Itinerary Problem (TIP) is a kind of travel plan for travel events, such as sight spot, flight and hotel timetable arrange, in a chronological order of date and time. In general, planning itinerary problem for TIP includes the following categories of decisions: (i) the determination of the tour (the sequence of stops visited by traveler), (ii) the specification of the transport modes and numbers for each section of the itinerary. The latter decision is closely dependent on the former, since for each given sequence of stops, the time and the number of visited sites each day are different, and the total cost is different.

In particular, this survey will focus on Self-driving Travel Itinerary Problem (STIP), and discusses some models and algorithms for recent research of the self-driving travel itinerary problem. A self-driving travel itinerary consists of a tour with one or more stops that a traveler takes. These destinations are visited follow one another sequentially in time in a way of self-driving. As characteristics of STIP, in the first, determination of travel tour is necessary for STIP. Second, the resting locations (hotels) are selected under two constraints include resting beak and best viewing time window, along with time arrangement while the time arrangement is more flexible for STIP, further, we can adjust dynamically the tour and time arrangement of travel events. Due to TIP is a travel way from origin to interesting places by self-drive, there is no the choice about transport modes.

In the area of smart travel optimization, a key problem arising from multi-scenic journey is to optimally determine an integrated travel scheme that satisfies tourists’ requirements (e.g., hotel selection around different attractions, routing, etc). The problem attracts much attention from both practitioners and researchers along with the availability of large-scale online data sources including travel sites data (location, opening/closing time window), hotel data (reservation price, service evaluation), transport data (network distribution, trip modes), etc. Several prevailing online trip webs such as Trip Advisor, Priceline, Ctrip and Airbnb, provide travel planning suggestions (i.e., hotel reservation, online ticket book). However, none of them provides online integrated scheme for tourists, which may generate some problems or conflicts, i.e., invalid reservation for hotels due to the mismatched itineraries, leading to partially optimized strategy and even impracticable travel recommendations. Numerous existing studies also take into account the personalized travel guidelines for tourists, typically involving three main functionalities: (i) recommending functionality to generate a list of travel sites for each different tourist profile \cite{1,2}; (ii) routing optimization functionality to generate tourist routes for interesting travel sites \cite{3}; (iii) customization functionality to allow tourists to modify the generated personalized route \cite{4}. Nonetheless, most of them separately consider the different functionalities of travel planning. In addition, the existing travel planning problems such as classical traveling salesman problem were mainly formulated with the optimization objective of cost minimization Veenstra et al., \cite{5} and utility (trip satisfaction) maximization.
In this section we overview the travel salesman problem with hotel selection problem simultaneously considers the functionalities of routing, hotel selection, and time scheduling as well as the optimal departure time at the origin under two types of personalized considerations: best site-viewing time windows and rest requirements by using multiple real-world data sources. To the best of our knowledge, this problem has not been addressed in the literature. To solve this two problems, two heuristic algorithms, namely, dynamic programming and clustering based heuristic algorithm and branch-and-bound and clustering-based heuristic algorithm, are designed to verify the effectiveness and efficiency of proposed algorithms. There are two problems: travel salesman problem with hotel selection and orienteering problem with hotel selection which are closely approximating the proposed self-driving scheme planning problem. Details about the two problems will be stated in the following.

TRAVEL SALESMAN PROBLEM WITH HOTEL SELECTION

In this section we overview the travel salesman problem with hotel selection. This problem is an extension of the TSP which have been developed by Castro et al., [11-14]. The TSPH can be stated in the following. A visitor departs at an origin, visits a set of geographically dispersed customers and hotels, and finally returns to the origin to minimize the total travel time. During this process, the following conditions should be satisfied: (1) the time budget for each trip; (2) every trip should start and end in one of the available hotels; (3) every customer is visited exactly once. Here, the trip refers to an ordered set of customers together with a starting and an ending hotel. A model formulation is presented in [15], where the objective function minimizes the total travel time. The number of trips is not a decision variable in this formulation, but a given parameter of the problem. An alternative formulation could be to minimize the number of trips first and then minimize the travel time. Constraint guarantees: (1) the tour starts from the initial hotel, (2) every trip starts and ends in one of the available hotels, (3) every trip starts in the same hotel, (4) every trip should start and end in one of the available hotels. In the model formulation presented by Divsalar et al., [17], the objective function maximizes the total collected score. Constraint guarantees: (1) the tour starts from the initial hotel, (2) the tour ends at the final ending hotel, (3) each trip starts and ends in one of the available hotels, (4) verify that if a trip ends in a given hotel, the next trip starts in the same one, (5) determine the connectivity inside each trip, (6) ensure that every vertex is visited at most once, (7) limit each trip with a time budget. If visiting a vertex requires a certain service time, this time can easily be modeled as a part of all travel times to and/or from the vertex, (8) sub-tour elimination constraints.

In this section, we review the existing literature on travel scheme planning with regard to travel routing optimization, hotel selection in travel planning, and solution methods.

TRAVEL ROUTING OPTIMIZATION

Regarding to the travel scheme problem, there are some popular problems such as travel salesmen problem, orienteering problem which are similar to the focused problem. The Travel Salesman Problem (TSP) is the most widely known problem in the field of routing and scheduling problem and combination optimization problem. Nowadays, this problem has been studied in extensive areas such as printing press scheduling problem [19], home health care problem [20], and hot rolling scheduling problem [21], and among others. The TSP can be stated as follows: a traveler departs at an origin, visits each customer exactly once, and finally returns to the origin, aiming to minimize the total travel cost. Today, various varieties associated with TSP have been emerged such as time dependent TSP [22], TSP with time windows [23], multiple TSP [24], pickup-and-delivery TSP [25] and TSP with drone [26]. The solution methods mainly consist of exact algorithms and approximate approaches. Regarding to the exact algorithm, Dantzig et al., [27] in 1954 is the first to propose the formulation in order to solve the TSP where they adopted a strategy consisting of initially relaxing constraints and the integrality requirements. The branch-and-bound algorithm is also the common used algorithm for TSP which are studied by Carpaneto [28-30]. Many algorithms based on the assignment problem relaxation have been devised. Some of the best known are those of Little et al., [31-33]. The 2-matching lower bound algorithm for TSP is proposed by Miliotis [34,35]. Afterwards, this work is followed by Crowder [36-40]. Generally, the Branch and Bound method and Held-Karp lower bound are the main approaches for solving the TSP and ensuring the optimality. There exist some heuristics to solve the approximate solutions for TSP. The approximately approaches main consist of tour construction approaches i.e., closest neighbor heuristic [41], greedy heuristic [41], insertion heuristic [42], christofide
heuristic [42], tour improvement i.e., 2-opt, 3-opt [41], k-opt [43,44]), tabu search [41], simulated annealing [41], and genetic algorithm [41]), ant colony optimization [45], and among others. Focusing on routing optimization in travel scheme planning, the orienteering problem Chao et al., [46] that is quite relevant with the targeted problem in this study, aims to determine a subset of nodes to visit as well as the optimal visiting sequence among these visiting nodes, such that the total collected score (i.e., utility) is maximized under given time and/or cost budgets. The orienteering problem can also be seen as a combination with knapsack problem and travel salesman problem, since it involves both vertex selection and shortest Hamiltonian path determination among the selected set of sites. In the literature, there are formulations with the similar concept of the orienteering problem, including selective traveling salesman problem Laporte et al., [47], maximum collection problem Butt et al., [48], bank robber problem Awerbuch et al., [49], and among others. Extensive variants of the orienteering problem have been investigated such as the team orienteering problem [50], team orienteering problem with time windows Vansteenwegen et al., [51], and time-dependent orienteering problem Verbeeck et al., [52]. Surveys that situate the orienteering problem among other types of route problems can be found in studies by Vansteenwegen et al., [53-55], in which details about problem descriptions, benchmark instances, solutions approaches, and applications of the orienteering problem can be found. The above studies take into account wide varieties of orienteering problem, but that none of the studies have addressed the integrated travel scheme planning problem aiming at tourists’ requirements in reality travel. Not much studies take into account the travel scheme planning problem for self-driving. Li et al., [56] formulate a travel itinerary problem for public transportation to minimize the total cost which aims to provide the transport alternatives’ decisions of visiting sequence of cities, flight/train number choice, under some time constraints. This study uses an implicit enumeration algorithm to solve the optimal combination of tour and numbers. Hashima et al., [57] study the self-drive tourism route problem using a goal programming model. This problem determined the route sequence among the travel sites in seven cities and utilized LINGO12.0 software to minimize the travel distance, minimize accommodation costs and maximize the total number of tourist attractions under the travel time constraint. Chang et al., [58,59] designed an automatic travel itinerary planning system for the domestic area to provide the travel route under some time constraints where the system considers some factors (i.e., departure time, days to travel, departure/destination points) affecting the travel preferences of users. Masoodian et al., [58] designed a graphical travel itinerary visualization system instead of conventional travel itineraries list using personal digital assistant type devices. These systems rely on large computer displays for visualization of itinerary information during the pre-travel planning and preparation phase.

**HOTEL SELECTION IN TRAVEL PLANNING**

One aspect of travel scheme planning that is of much importance in practice is hotel selection. Orienteering problem with hotel selection (OPHS), a variant of the orienteering problem, was initiated in Vansteenwegen et al., (2012) [15,60]. The goal is to determine a fixed number of trips that connect given travel sites to maximize the sum of collected utility scores with preceived utility-like scores for each site and a set of hotels for choice [61]. OPHS restricts each trip’s length and requires that the trip should start and end in one of the hotels [12]. The other similar work, the travelling salesperson problem with hotel selection (TSPHS), has al been developed by Vansteenwegen et al., [11,12,15,61]. These studies provided extensive algorithms to deal with the TSPHS.

**SOLUTION METHODS**

There are streams of solving methods proposed to generate travel planning solutions in the literature. A skewed variable neighborhood search algorithm Divsalar et al., [60] that consists of a constructive initialization procedure and an improvement module to solve OPHS. In the improvement module, a solution with a slightly worse objective value is also accepted compared to a regular variable neighborhood search approach. In fact, the skewed variable neighborhood search algorithm has been successfully applied for solving orienteering problem Seykli et al., [62], team orienteering problem [63], bi-objective orienteering problem [64], multi-period orienteering problem with multiple time windows [65], and team orienteering problem with time windows [66]. A memetic algorithm to solve OPHS is proposed in Divsalar et al., [18]. The memetic algorithm consists of two levels: one level is to find a sequence of intermediate hotels by a genetic component, on the other level, six local search moves are embedded in a variable neighborhood structure to deal with the selection and visiting sequence determination of the sites between the hotels. Except for the above solving methods, dynamic programming is also commonly used for solving routing planning. Especially, in travel salesmen problem and its variants, the exact and heuristic dynamic programming algorithms are frequently applied to solve the optimal or near optimal solutions [67-70]. But, very few studies utilize dynamic programming algorithm in travel scheme planning. Only recently, Lu et al., [12] designed a hybrid dynamic programming and memetic algorithm to solve the traveling salesman problem with hotel selection where the dynamic programming is applied to find an optimal hotel sequence for a given tour, and the memetic algorithm is used to search solutions by using three crossover operators.

**DYNAMIC PROGRAMMING AND CLUSTERING-BASED HEURISTIC ALGORITHM**

Dynamic programming has been widely used to solve routing problem and is able to guarantee the global optimality in a small-scale network. However, for general applications in a large-scale network, DP algorithm shows exponential time complexity and consumes much computation resources. For solving the large-scale problem, numerous studies considered available algorithms combining the advantages of several methodological modules [71-73]). In this work, we propose a DP and clustering
based heuristic algorithm combining multiple modules with further modifications. The proposed solution method decomposes the problem into three parts: (i) sites clustering; (ii) routing optimization; (iii) hotel selection and time scheduling. The proposed heuristic algorithm consists of three modules: (i) a multi-categorical attribute K-means clustering algorithm to cluster the travel sites; (ii) routing optimization where DP algorithm is iteratively executed for each cluster of sites without considering hotel selection and the tours connection representation is adopted to deal with the problem of multiple tours; (iii) hotel selection and time scheduling based on the obtained routes by performing the constraint inspection procedure.

INTEGRATED SELF-DRIVING TRAVEL SCHEME PLANNING

In Du et al., [9,10], authors proposed an integrated self-driving travel scheme planning problem as a novel travel planning problem. The ISTSP can make a personalized and integrated travel scheme for self-driving travel among multiple travel sites and hotels, which incorporates routing, hotel selection and time scheduling functionalities for total cost minimization. In addition, the authors presented a mathematical formulation that considers travel utility maximization as a simultaneous optimization objective besides traditionally focused cost minimization objective. For solving large-scale travel network, the authors proposed two heuristic algorithms: DP and clustering-based heuristic algorithm and branch-and-bound and clustering-based heuristic algorithm to solve the ISTSP. Numerical illustrative experiments verify the validity of the model and the solution methods for practical use in the travel industry. For future study explorations, an interesting extension of this study is to consider time dependent attributes in terms of cost and time, since the travel time and cost are stationary in the proposed model. Further, it is an interesting problem to explore a Pareto-optimal set of travel scheme by considering a multi-objective travel planning problem with the comprehensive aim of waiting time minimization, travel cost minimization, and travel utility maximization. It is worthy of investigation about the effectiveness of the proposed heuristics when embedded into some well-recognized solving methods for multi-objective programming such as NSGAI [74], and among others.

CONCLUSION

In addition, authors also plan to investigate multiple data sources (e.g., search engine data, social media information, trip comments, travel guidelines, etc) enabled site selection models to better capture tourist’s trip preferences in the travel scheme planning. Since there are both structured and unstructured data for site selection, some state-of-the-art data analytics (e.g., statistics learning, machine learning, text mining, and also deep learning) that show promising effectiveness will be used.

REFERENCES


