Expert Systems With Applications 112 (2018) 321–330

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Optimization approaches to support the planning and analysis of travel itineraries

Admilson Alcântara da Silva[a], Reinaldo Morabito[b,*], Vitória Pureza[b]

[a] Department of Statistics, State University of Pará, Tv. Djalma Dutra, s/n, Telégrafo, Belém, PA, Brazil
[b] Department of Production Engineering, Federal University of São Carlos, Rodovia Washington Luís, km 235, São Carlos, SP, Brazil

ARTICLE INFO

Article history:
Received 28 December 2017
Revised 29 May 2018
Accepted 19 June 2018
Available online 22 June 2018

Keywords:
Travel itinerary optimization
Profitable tour problem
Traveling salesman problem with priority prizes
Mixed integer programming
Tabu search
Multivariate statistical analysis
Tourism activities

ABSTRACT

This paper addresses the problem of elaborating travel itineraries considering the visitors´profiles, travel distances and costs. Other parameters such as how much the tourists value the attractions offered are particularly taken into account, as well as their preferences regarding the order of visiting these attractions. This problem can be seen as a traveling salesman problem with profits that also considers priority prizes. It is described by an optimization model based on mixed integer programming, which aims to generate itineraries that maximize the total value of the attractions visited and minimize the total travel cost involved. We present results for sets of instances based on randomly generated and real data by applying mathematical programming techniques and a tailored tabu search algorithm. To obtain the input parameters of the problem for the real cases, some statistical techniques are used to analyze the data collected, such as multivariate correspondence analysis. The resulting solutions illustrate the potential application of the proposed approach to support planning and formatting decisions of travel itineraries.

1. Introduction

The World Tourism Organization (OMT, 2014) estimates that tourism activities are directly responsible for 6%–8% of jobs generated worldwide. In Brazil, these activities generated a turnover of BR$11.9 billion (approximately US$4.4 billion) in 2014, representing a market of six million tourists traveling around the country (MTUR, 2015). These data motivate and justify initiatives that lead to improvements in the sector, especially to planning travel itineraries.

A travel itinerary can be defined as a planned itinerary of a tourist activity, which includes a detailed description of the attractions and activities of the destinations visited, as well as the duration and specification of times and services included. The itinerary includes the trip and the stay with a sequence of activities carried out in a destination or in various locations. In the case of travel agencies, the first step of arranging consists of finding out about the visitors´preferences and satisfaction. According to Silva (2017), each target audience has different levels of interest in the available attractions and often shows preferences regarding the order in which these visits should be undertaken during the tour. For example, some audiences prioritize visits to museums in the first orders while others want these visits to occur in the last orders of the itinerary. In addition, quite a few tourists would agree to pay more, to a certain extent, for changes in the chosen attractions of the itinerary and/or the order of their visits in the itinerary. Therefore, it is important to value peopleś preferences of attractions and their order, and consider them with the costs involved so as to be able to plan itineraries that simultaneously maximize customer satisfaction and minimize costs.

Various studies have been developed to help organize personalized tours. These studies have detected some new aspects that should be considered in order to offer each tourist the most appropriate itinerary when planning a route. Some of these studies help tourists to search for information, based on their preferences (Colineau & Wan, 2001; Camacho, Borrajo & Molina, 2001). Others also consider the existing travel offer at that time (Jakkilinki, Georgievski, & Sharda, 2007). Some studies can also be highlighted that offer recommendations about the destination to be visited or activities to be carried out, such as the travel recommender system (Ricci, 2002), user-oriented adaptive systems for planning tourist visits (Castillo et al., 2008), adaptive recommender systems for travel planning (Mahmood, Ricci, Venturini, & Höpken, 2008), and expert systems for tourists (P. Vansteenwegen, Souffriau, VandenBerghe, & Van Oudheusden, 2011). Some studies use multi-criteria techniques to consider the different objectives considered in planning a tourist trip (Godart, 1999), based on combinatorial optimization techniques for travel planning.

* Corresponding author.
E-mail address: morabito@ufscar.br (R. Morabito).

https://doi.org/10.1016/j.eswa.2018.06.045

Studies were also found in the literature based on the location of the tourist, using global positioning systems, to keep them constantly informed about activities nearby and guide them in a specific itinerary. Other studies, besides location, consider the user profile (Yu, Spaccapietra, Cullot, & Aufaure, 2003) or the context of the tourist (Schmidt-Belz, Laamanen, Poslad, & Zipf, 2003; Van Setten, Pokraev, & Koolwaaij, 2004; Zipf, 2002). Others consider all these aspects at the same time (Kramer et al., 2007; Ten Hagen, Modsching, & Krarner, 2005). Souffriau and Vansteenwegen (2010) carried out a thorough literature review and indicated the functionalities of different studies of this nature; one of the most complete of which is Vansteenwegen et al. (2011).

As well as travel planning problems, other studies by De Choudhury et al. (2010), Basu Roy et al. (2011) and Yoon, Zheng, Xie, and Woo (2012) calculate optimal itineraries according to some reward metric. Most of these papers focus on the data mining aspect of travel planning and planning itineraries. In line with this interest, this paper presents a mathematical programming model to adequately represent a problem of planning travel itineraries, which can be useful for travel agencies and other tourism organizations, such as State Departments of Tourism. The model is based on formulations of two variants of the well-known Traveling Salesman Problem (TSP) in operations research: the particular case of the Travelling Salesman Problem with Profits known as the Profitable Tour Problem (PTP) (Archetti, Speranza, & Vigo, 2014; Balas, 1987; Feillet, P. Dejax, & Gendreau, 2005), and the Travelling Salesman Problem with Priority Prizes (TSPPP) (Morabito et al., 2014). Basically, the problem of planning travel itineraries, called here the Profitable Tour Problem with Priority Prizes (PTPPP), consists of choosing places to visit and defining a travel itinerary considering these places in order to maximize tourist preferences and satisfaction and minimize the cost of the itinerary. It is considered that customer satisfaction depends on how much customers value the places chosen and also how much they value the order of visitation of these points, as defined by the itinerary.

To solve the PTPPP, an exact branch-and-cut algorithm is used in an optimization software and a heuristic tabu search algorithm (Glover & Laguna, 1997) is also developed. Heuristic methods are necessary to treat larger instances of the problem, since, as most vehicle routing problems of interest, the travelling salesman problem and both variants belong to the NP-hard class. Therefore, it is considered difficult to solve in various real situations because of the sizes of these problem instances. To obtain the input parameters of the problem for the real cases, some statistical techniques are used to analyze the data collected, such as multivariate correspondence analysis. These techniques are useful to classify the tourists and determine parameters for each tourists' class. To the best of our knowledge, there are no other studies concerning travel itineraries using optimization approaches in the same line of research explored in this study.

This article is organized as follows. In Section 2, the problem of planning travel itineraries is posed based on the PTP and the TSPPP. The mathematical modeling of the PTPPP is presented in Section 3. In Section 4, the heuristic algorithm of the proposed tabu search is outlined. Section 5 initially describes the data collection process and statistical data analysis to obtain the parameters of the problem, and then presents the computational results obtained from the application of the proposed approach using real examples from the city of Belém, State of Pará in Brazil and also randomly generated examples. Finally, the conclusions are drawn in Section 6 and some future research perspectives are suggested.

2. Problem definition

As mentioned, the problem of planning travel itineraries can be seen as a variant of the TSP. The TSP roughly consists of finding a minimum distance route that visits exactly once n spatially distributed points, and then returns to the starting point. TSP is one of the most studied problems in network optimization and has a wide range of practical applications (Applegate, Bixby, Chvatal, & Cook, 2006; Christofides, 1979; Laporte, 2010) and often transcends the logistical scope. From the seminal works by Dantzig, Fulkerson, and Johnson (1954) and Dantzig and Ramer (1959), a variety of models and solution methods have been proposed to effectively represent and solve instances of different sizes. Many TSP variants have been reported and studied in the literature and mathematical programming models and solution methods are often applied to analyze these variants.

A variant of the TSP is particularly relevant to this research, specifically the TSPPP studied in Morabito et al. (2014). It can be seen as a TSP extension, where all clients (nodes) have to be visited by the salesman, but the order of customer visits is directly taken into account in the objective function of the problem. Specifically, a priority prize value p_{ki} is received if point i is visited in order k of the route, while a travel cost c_{ij} is incurred due to the displacement of the stretch between i and j. It should be emphasized that in this variant, a visit prize value p_i can also be included that the traveling salesman receives when visiting point i, regardless of the order of visitation k, as well as the priority prize p_{ki} that he/she collects when he/she visits point i in the k^{th} position of the route. Another variant of the TSP considered here, the PTP (Archetti et al., 2014; Feillet et al., 2005), associates a prize p_i at each point i and a travel cost c_{ij} between each i and j, and it consists of obtaining a route that visits the subset of points that maximizes the total collected prize, while travel costs are minimized. In other words, it is not required to visit all the points as in the TSP and TSPPP.

The problem of planning travel itineraries can be formulated by combining aspects of these two TSP variants, PTP and TSPPP. Given a set of available tourist attractions, p_i prize values are collected when attraction i is visited, p_{ki} prize values are collected when attraction i is visited in order k of the itinerary and travel costs c_{ij} are spent on the displacement between points i and j. In other words, the problem consists of obtaining a route that visits a subset of attractions that maximizes the total profit, as a result of the difference in prizes collected and travel costs. This problem, referred to here as the PTPPP, adopts a more general objective function than the TSPPP, with opposite criteria, looking for solutions that somehow consider quality of service to customers and delivery priorities, maximizing received services and minimizing delivery costs of the services.

3. Model formulation

The PTPPP can be represented by a directed graph $G(N, A)$ with node set $N = \{1, 2, \ldots, n\}$, such that the first node ($i = 1$) represents the touristś starting point (e.g., the hotel where he/she is staying), while the others $n - 1$ nodes represent the available attractions. Set A contains the arcs (i, j) connecting nodes i and j, $i \neq j$, of set N. The problem can be formulated as a mixed integer programming model, as follows:

Parameters:

p_{ki}: valuation (priority prize) of the tourist visiting attraction $i \in N$ in order k ($1 \leq k \leq n$).

p_i: valuation (visit prize) of the tourist visiting attraction $i \in N$, regardless of the order of the visit.

c_{ij}: travel cost of arc $(i, j) \in A$.

Decision variables:

$$X_{ij} = \begin{cases} 1, & if\ arc\ (i,j) \in A\ is\ traveled \\ 0, & otherwise \end{cases}$$