European Journal of Operational Research 000 (2017) 1–9

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

The Set Orienteering Problem

Claudia Archetti[a,*], Francesco Carrabs[b], Raffaele Cerulli[b]

[a] Department of Economics and Management, University of Brescia, C.da S. Chiara 50, Brescia, Italy
[b] Department of Mathematics, University of Salerno, Via Giovanni Paolo II 132, Fisciano, Italy

ARTICLE INFO

Article history:
Received 1 June 2017
Accepted 8 November 2017
Available online xxx

Keywords:
Routing
Orienteering problem
Matheuristic

ABSTRACT

In this paper, we study the Set Orienteering Problem which is a generalization of the Orienteering Problem where customers are grouped in clusters and a profit is associated with each cluster. The profit of a cluster is collected only if at least one customer from the cluster is visited. A single vehicle is available to collect the profit and the objective is to find the vehicle route that maximizes the profit collected and such that the route duration does not exceed a given threshold. We propose a mathematical formulation of the problem and a matheuristic algorithm. Computational tests are made on instances derived from benchmark instances for the Generalized Traveling Salesman Problem with up 1084 vertices. Results show that the matheuristic produces robust and high-quality solutions in a short computing time.

1. Introduction

Routing problems with profits received significant attention in recent years, as witnessed by the large literature surveyed in three recent papers, i.e., Archetti, Speranza, and Vigo (2014), Gunawan, Lau, and Vansteenwegen (2016) for routing problems with profits where the profit is associated with nodes of a graph, and Archetti and Speranza (2014) for problems where the profit is associated with arcs or edges of a graph. As witnessed by these three surveys, the literature on 'node' routing problems, i.e., the first of the two classes mentioned above, is much wider than the one on arc routing problems. In particular, among all node routing problems with profits studied in the literature, the most widely known is undoubtedly the Orienteering Problem (OP) where a profit is associated with each customer and the objective is to find a single vehicle tour maximizing the profit collected from visited customers and such that the duration of the tour does not exceed a maximum time limit. The profit of each customer can be collected at most once. The problem was first introduced in Tsiligirides (1984) and many variants of the problem have been studied, as described in Archetti et al. (2014), Gunawan et al. (2016). One of the most recent variants is the one presented in Angelelli, Archetti, and Vindigni (2014) where customers are grouped in clusters, a profit is associated with each cluster and it is collected only if all customers belonging to the cluster are visited. The authors called this problem the Clustered Orienteering Problem (COP). They present a mathematical formulation and two solution algorithms. They also describe practical applications related to the distribution of mass products where customers are retailers belonging to different supply chains.

In this paper, we are interested in a variant of the OP which shows some analogies with the one studied in Angelelli et al. (2014). In particular, we study the problem where customers are grouped in clusters. A profit is associated with each cluster and is collected only if at least one customer from the cluster is visited. The objective is to find the vehicle route that maximizes the collected profit and such that the corresponding duration does not exceed a given threshold. We call this problem the Set Orienteering Problem (SOP). The SOP finds application in mass distribution products, as for the COP, where a different distribution plan is sought. In particular, consider the case where customers belong to different supply chains and the carrier stipulates contracts with chains. Then, instead of having to serve all retailers belonging to the chain with which the contract has been stipulated, as happens in the COP, in the SOP the carrier may choose to serve only one customer from the chain (and, implicitly, serving the entire quantity demanded by the chain). This way, the carrier may be able to offer a better price for the service. The inner distribution among all retailers in the chain will be then organized internally. Thus, the SOP presents an alternative to the distribution strategy applied in the COP which may be advantageous both for the carrier and for the chains. Another application arises when customers are clustered in areas and the service to each area is made by delivering the entire quantity required by all customers in the area to a single customer, the one that is visited. This happens also when private customers group together to reach large quantity orders, and thus

* Corresponding author.
E-mail addresses: claudia.archetti@unibs.it, archetti@eco.unibs.it (C. Archetti), fcarrabs@unisa.it (F. Carrabs), raffaele@unisa.it (R. Cerulli).

https://doi.org/10.1016/j.ejor.2017.11.009
0377-2217/

hopefully a lower price. Typically, in this case, the delivery is made to a single location.

The contribution of this paper can be summarized as follows. We introduce the SOP, present a formal description and a mathematical formulation. We then propose a matheuristic algorithm for its solution which is tested on instances derived from benchmark instances for the Generalized Traveling Salesman Problem (GTSP) with up to 1084 vertices. In particular, we first show the performance of the algorithm on small instances by comparing the results obtained from the matheuristic with optimal solutions. In addition, we test the performance of the matheuristic on large instances for which the optimal solution is known. We then present an exhaustive study of the contribution of the MILP embedded in the matheuristic. The results show that the contribution of the MILP is more evident on large instances. Moreover, even if the MILP makes the overall algorithm slower, computing times remain reasonable even on the largest instances.

The paper is organized as follows. A formal description of the problem together with a mathematical formulation are presented in Section 2. The matheuristic algorithm is described in Section 3 while computational results are presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Problem description and formulation

As the SOP is a generalization of the OP, we first provide a formal description of the OP.

The OP is defined on a complete directed graph $G = (V, A)$ where $V = \{0\} \cup C$. Vertex 0 represents the depot from which the vehicle starts and ends its tour. C is the set of customers. A profit p_i is associated with each customer $i \in C$ and is collected if and only if customer i is visited by the vehicle. Moreover, a cost c_{ij} is associated with each arc $(i, j) \in A$. The objective is to find the tour that maximizes the collected profit and such that the associated cost (or duration) does not exceed a maximum value T_{max}.

In this paper we study a variant of the OP which we call the *Set Orienteering Problem* (SOP). In the SOP, customers in C are grouped in clusters C_g with $g = 1, \dots, l$ such that $\bigcup_{g=1}^{l} C_g = C$ and $C_g \cap C_h = \emptyset$, $\forall C_g, C_h \in \mathcal{P}$ where $\mathcal{P} = \{C_1, \dots, C_l\}$ is the set of clusters. A profit p_g is associated with each cluster and is collected if and only if at least a customer $i \in C_g$ is visited in the tour. The profit of each cluster can be collected at most once. As in the OP, the objective is to find the tour that maximizes the collected profit and such that the associated cost does not exceed T_{max}. In the following, we assume that costs c_{ij} satisfy the triangle inequality. In this case, as shown in Laporte and Norbert (1983), an optimal solution always exists where one vertex per cluster at most is visited. This property is used in the solution method presented in Section 3.

In order to present a mathematical formulation for the SOP, we need the following notation. For any subset of vertices $S \subset V$, we define $\delta^+(S) = \{(i, j) \in A : i \in S, j \notin S\}$ and $\delta^-(S) = \{(i, j) \in A : i \notin S, j \in S\}$. For the ease of presentation, in the following we will use the notation $\delta^+(i)$ and $\delta^-(i)$ when $S = \{i\}$. The decision variables are the following:

- $y_i =$ binary variable equal to 1 if vertex $i \in V$ is visited by the vehicle, and 0 otherwise,
- $x_{ij} =$ binary variable equal to 1 if arc $(i, j) \in A$ is traversed by the vehicle, and 0 otherwise,
- $z_g =$ binary variable equal to 1 if the profit of cluster C_g is collected and 0 otherwise.

The mathematical programming formulation of the SOP is the following:

$$\max \sum_{j \in \mathcal{P}} p_g z_g \tag{1}$$

$$\text{s.t.} \sum_{(i,j) \in \delta^+(i)} x_{ij} = y_i \qquad i \in V, \tag{2}$$

$$\sum_{(j,i) \in \delta^-(i)} x_{ji} = y_i \qquad i \in V, \tag{3}$$

$$\sum_{(i,j) \in \delta^+(S)} x_{ij} \geq y_h \qquad S \subseteq V \setminus \{0\},\ h \in S, \tag{4}$$

$$\sum_{(i,j) \in A} c_{ij} x_{ij} \leq T_{\max}, \tag{5}$$

$$z_g \leq \sum_{i \in C_g} y_i \qquad C_g \in \mathcal{P}, \tag{6}$$

$$y_i \in \{0, 1\} \qquad i \in V, \tag{7}$$

$$z_g \in \{0, 1\} \qquad C_g \in \mathcal{P}, \tag{8}$$

$$x_{ij} \in \{0, 1\} \qquad (i, j) \in A. \tag{9}$$

The objective function (1) maximizes the collected profit. Constraints (2) and (3) ensure that one arc enters and one arc leaves each visited vertex. Subtours are eliminated through (4). Constraint (5) is the maximum duration constraint on the route while (6) imposes that the profit of cluster C_g is collected only if at least one customer $i \in C_g$ is visited in the tour. Finally, (7)–(9) are variable definitions.

Note that formulation (1)–(9) has an exponential number of subtour elimination constraints (4). A formulation with a polynomial number of constraints is obtained by introducing arc flow variables u_{ij}, representing the amount of flow crossing the edge (i, j), and substituting (4) with the following constraints:

$$\sum_{j \in V} u_{ji} - \sum_{j \in V} u_{ij} = y_i \quad i \in V \setminus \{0\}, \tag{10}$$

$$u_{ij} \leq (n-1) x_{ij} \quad (i, j) \in A, \tag{11}$$

$$y_0 = 1 \tag{12}$$

$$u_{ij} \geq 0 \quad (i, j) \in A, \tag{13}$$

where $n = |V|$. We note that this formulation of subtour elimination constraints has been proposed in Gavish and Graves (1978) for the Traveling Salesman Problem (TSP) and its performance has been recently assessed in Öncan, Altinel, and Laporte (2009).

3. A matheuristic for the SOP

In this section, we describe the heuristic algorithm we have designed for the solution of the SOP. It is a matheuristic algorithm which is composed by the following two phases:

1. *Phase 1: construction of an initial solution.*
2. *Phase 2: Tabu search.*

It is a matheuristic algorithm as the tabu search makes use of a MILP formulation when it struggles in finding a new non-tabu feasible solution. In the following we refer to the algorithm as MASOP – a MAtheuristic for the SOP.

We now describe in details the two phases. We define the following notation. Given a tour T, we denote as $C(T) \subseteq \mathcal{P}$ the set of