Discrete Applied Mathematics 161 (2013) 618–632

Contents lists available at SciVerse ScienceDirect

Discrete Applied Mathematics

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

Tight compact models and comparative analysis for the prize collecting Steiner tree problem

Mohamed Haouari [a,b,*], Safa Bhar Layeb [c], Hanif D. Sherali [d]

[a] Department of Industrial Engineering, Faculty of Engineering, Ozyegin University, Istanbul, Turkey
[b] Princess Fatimah Alnijris Research Chair for AMT, College of Engineering, King Saud University, Saudi Arabia
[c] Department of Industrial Engineering, Ecole Nationale d'ingénieurs de Tunis, University of El Manar, Tunisia
[d] Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, USA

ARTICLE INFO

Article history:
Received 1 December 2010
Received in revised form 17 July 2011
Accepted 24 September 2011
Available online 22 October 2011

Keywords:
Steiner tree
MTZ subtour elimination constraints
Reformulation–Linearization Technique (RLT)
Mixed-integer programming

ABSTRACT

We investigate a generalized version of the prize collecting Steiner tree problem (PCSTP), where each node of a given weighted graph is associated with a prize as well as a penalty cost. The problem is to find a tree spanning a subset of nodes that collects a total prize not less than a given quota Q, such that the sum of the weights of the edges in the tree plus the sum of the penalties of those nodes that are not covered by the tree is minimized. We formulate several compact mixed-integer programming models for the PCSTP and enhance them by appending valid inequalities, lifting constraints, or reformulating the model using the Reformulation–Linearization Technique (RLT). We also conduct a theoretical comparison of the relative strengths of the associated LP relaxations. Extensive results are presented using a large set of benchmark instances to compare the different formulations. In particular, a proposed hybrid compact formulation approach is shown to provide optimal or very near-optimal solutions for instances having up to 2500 nodes and 3125 edges.

1. Introduction

The impressively rapid development witnessed by the telecommunication sector during the last two decades has prompted the investigation of many new challenging network design problems. In particular, there has been an ever-increasing interest in the investigation of models and solution strategies for the *Prize Collecting Steiner Tree Problem* (PCSTP). This network optimization problem has numerous relevant applications arising in the design of fiber-optic networks, local access networks, and cable television, to quote just a few. In this paper, we address a unifying PCSTP model that is formally defined as follows. We are given a connected, undirected graph $G = (V, E)$, where $V = \{0, \ldots, n\}$ is the node set, with node 0 being a specified *root node*, and E is the edge set, along with a nonnegative edge weight c_e associated with each edge $e \in E$, a prize p_j and a penalty γ_j associated with each node $j \in V^* = V \setminus \{0\}$, and a preset prize quota Q. A feasible PCSTP solution is a tree $T(\mathcal{S}) = (\mathcal{S}, E(\mathcal{S}))$ that spans a to-be-determined node subset $\mathcal{S} \subseteq V$ such that: (i) $0 \in \mathcal{S}$, and (ii) $\sum_{j\in\mathcal{S}} p_j \geq Q$. The PCSTP requires finding a feasible tree that minimizes the sum of the weights of the edges in the tree plus the sum of the penalties of those nodes that are not covered by the tree. Clearly, the PCSTP is $\mathcal{NP}$-hard since the well-known *Steiner Tree Problem* (STP) can be cast as a PCSTP by setting $Q = |T|$ (where $T \subset V$ refers to the set of *terminals*), $p_j = 1$ for all $j \in T$, $p_j = 0$ for all $j \in V \setminus T$, and $\gamma_j = 0$ for all $j \in V^*$.

To the best of our knowledge, the PCSTP in the foregoing form has only been addressed in a recent paper by Haouari et al. [16], where the authors described preprocessing procedures for reducing the graph size, an optimization-based heuristic, as well as an exact solution approach. A simplified variant of the PCSTP that considers only penalties but no prizes

* Corresponding author. Fax: +90 216 755 585.
E-mail addresses: mh6368@yahoo.com, Mohamed.Haouari@ozyegin.edu.tr (M. Haouari).

0166-218X/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2011.09.012

(i.e., with zero quota) was originally introduced by Bienstock et al. [2] and has received a great deal of attention in recent years. Approximation algorithms were proposed by Bienstock et al. [2], Goemans and Williamson [8], and Johnson et al. [17]. In addition, Lucena and Resende [20] provided a 0–1 programming formulation and used it to derive strong lower bounds. Later, Ljubic et al. [19] proposed a branch-and-cut approach. Finally, da Cunha et al. [6] implemented a Lagrangian non-delayed relax-and-cut algorithm to generate primal and dual bounds. A closely related zero-quota variant is the so-called *Prize Collecting Generalized Minimum Spanning Tree Problem* (PC-GMSTP). In this variant, the node set is partitioned into clusters, and the problem consists of designing a tree that is incident to exactly one node in each cluster while minimizing the total cost minus the sum of the prizes corresponding to the nodes selected for the tree. The PC-GMSTP has been addressed by Pop [26], who proposed several formulations as well as an exact solution procedure. Also, Golden et al. [9] designed local search-based heuristics, a genetic algorithm, as well as an exact branch-and-cut algorithm for this problem. A second special case of the PCSTP with no penalties has been investigated by Johnson et al. [17], who proposed approximation algorithms, and by Haouari and Chaouachi [14], who described a Lagrangian-based genetic algorithm. Furthermore, several other variants of PCSTP have been investigated. In particular, a generalized multi-quota version has been addressed by Haouari et al. [15], who presented an integer programming formulation along with a variety of Lagrangian relaxation-based lower bounds. Also, a PCSTP variant with revenues, a budget constraint, and hop constraints has been investigated by Costa et al. [4,5]. In these two papers, the authors designed heuristics and an exact branch-and-cut algorithm, respectively. Finally, a variant with node degree dependent costs has been recently studied by Gouveia et al. [10]. Other variants of the PCSTP studied in the literature are described in [3].

For the sake of convenience, we first introduce a generic *directed* formulation of the PCSTP. Toward this end, we consider the bidirected graph $B = (V, A)$ that is obtained from G by replacing each edge $e = \{i, j\} \in E$ with two directed arcs (i, j) and (j, i) (with corresponding weights $c_{ij} = c_{ji} = c_e$). The root node 0 has an in-degree equal to zero. Thus, we can pose the PCSTP as requiring to find an optimal arborescence in B that is rooted at node 0. Denote by x_{ij}, $(i, j) \in A$, the binary variable that takes the value 1 if arc (i, j) belongs to the arborescence, and 0 otherwise. We also define a binary variable $y_j, j \in V$, which takes on a value of 1 if node j belongs to the arborescence, and 0 otherwise. Using these definitions, a basic formulation of the PCSTP can be stated as follows:

$$\textbf{PCSTP}: \quad \text{Minimize} \sum_{(i,j)\in A} c_{ij}x_{ij} + \sum_{j\in V^*} \gamma_j(1 - y_j) \tag{1}$$

subject to:

$$y_0 = 1, \tag{2}$$

$$\sum_{j\in V^*} p_j y_j \geq Q, \tag{3}$$

$$\sum_{i:(i,j)\in A} x_{ij} = y_j, \quad \forall j \in V^*, \tag{4}$$

$$x \equiv (x_{ij})_{(i,j)\in A} \text{ defines an arborescence in } G \text{ rooted at } 0, \tag{5}$$

$$x_{ij} \in \{0, 1\}, \quad \forall (i, j) \in A, \tag{6}$$

$$y_j \in \{0, 1\}, \quad \forall j \in V^*. \tag{7}$$

The objective (1) is to minimize the sum of the installation costs and the penalties for non-covered nodes. Constraint (2) requires that the root node is covered by the arborescence. Constraint (3) enforces the total collected prizes to be at least equal to the preset quota. Constraint (4) asserts that each covered non-root node (that is, $y_j = 1$ for $j \in V^*$) has exactly one incoming arc incident to it. Otherwise, if $y_j = 0$, then node j has no incident arc. Constraint (5) (together with (6)) enforces that x is the incidence vector of an arborescence in G rooted at 0. There are several valid alternatives for expressing (5) in a formal way. For instance, we can use the so-called *generalized subtour-elimination constraints* (GSEC) [20]:

$$\sum_{(i,j)\in A(S)} x_{ij} \leq \sum_{j\in S} y_j - y_k, \quad \forall k \in S \subseteq V^*, \; |S| \geq 2, \tag{8}$$

where $A(S) \subseteq A$ denotes the set of arcs having both ends in S.

Also, we can alternatively use the so-called *cut* (or, *connectivity*) *constraints* [19]:

$$\sum_{(i,j)\in \delta^+(S)} x_{ij} \geq y_k, \quad \forall S \subset V, \; 0 \in S, \; k \in V \setminus S, \tag{9}$$

where $\delta^+(S) = \{(i, j) \in A : i \in S \text{and} j \notin S\}$. Likewise, for later use, we define the set $\delta^-(S) = \{(i, j) \in A : i \notin S \text{ and } j \in S\}$. In the sequel, we shall refer to the models that are derived through substituting (8) or (9) in place of (5) as the *Directed Subtour Elimination Formulation* (**DSEF**), and the *Direct Cut Formulation* (**DCF**), respectively.

Clearly, both **DSEF** and **DCF** include an exponential number of constraints, which necessitate a row-generation solution process. Indeed, several effective cutting-plane-based approaches have been successfully implemented for solving large-scale PCSTP problems and its variants (see for example [19,16]). However, the development and implementation of these