Discrete Optimization 21 (2016) 85–117

Contents lists available at ScienceDirect

Discrete Optimization

www.elsevier.com/locate/disopt

Binary Steiner trees: Structural results and an exact solution approach

Frauke Liers, Alexander Martin, Susanne Pape*

FAU Erlangen-Nürnberg, Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany

ARTICLE INFO

Article history:
Received 20 November 2014
Received in revised form 25 May 2016
Accepted 26 May 2016
Available online 24 June 2016

Keywords:
Steiner tree
Degree constraint
Steiner ratio
Integer programming
Polytope
Branch-&-cut

ABSTRACT

In this paper we study the Steiner tree problem with degree constraints. Motivated by an application in computational biology we focus on binary Steiner trees in which all node degrees are required to be at most three. It is shown that finding a binary Steiner tree is $\mathcal{NP}$-complete for arbitrary graphs. We relate the problem to Steiner trees without degree constraints as well as degree-constrained spanning trees by proving approximation ratios. Further, we give integer programming formulations for this problem on undirected and directed graphs and study the associated polytopes for both cases. Some classes of facets are introduced. Based on this study a branch-&-cut approach is developed and evaluated on biological instances coming from the reconstruction of phylogenetic trees. We are able to solve nearly all instances up to 200 nodes to optimality within a limited amount of time. This shows the effectiveness of our approach.

1. Introduction

Scientific or engineering applications often require the solution of optimization problems. Over the years the Steiner tree problem in graphs and its variants have taken an increasingly important role. Many real-life applications in network design in general and VLSI design (very large scale integration) in particular use Steiner trees to model and solve their problems, see [1] for example.

Given an undirected graph $G = (V, E)$ and a terminal set $T \subseteq V$, a **Steiner tree** for T is a subset $X \subseteq E$ that spans all nodes in T. A Steiner tree may contain Steiner nodes of the set $S = V \setminus T$. The **degree-constrained Steiner tree problem** for G is stated as follows: Given a cost function $c : E \to \mathbb{R}_+$ and degree requirements $b(v) \in \mathbb{N}$ for all $v \in V$, find a minimum cost Steiner tree such that the degree at each node $v \in V$ is less than or equal to $b(v)$.

* Corresponding author.
E-mail addresses: frauke.liers@math.uni-erlangen.de (F. Liers), alexander.martin@math.uni-erlangen.de (A. Martin), susanne.pape@math.uni-erlangen.de (S. Pape).

http://dx.doi.org/10.1016/j.disopt.2016.05.006

The Steiner tree is called **binary** if all nodes in X have a degree less than or equal to three. Given a cost function $c : E \to \mathbb{R}_+$, the **binary Steiner tree problem** is to find a minimum cost binary Steiner tree.

The Steiner tree problem without any degree constraints has been extensively studied in the literature, see for example [2] for an exact dynamic programming approach, Takahashi & Matsuyama [3] and Rayward-Smith [4] for two heuristic algorithms, Maculan [5] and Goemans & Myung [6] for different integer programming formulations, Chopra & Rao [7], Chopra & Rao [8] and Goemans [9] for a polyhedral study of the Steiner tree polytope or Lucena & Beasley [10], Chopra et al. [11] and Koch & Martin [12] for branch-&-cut approaches. For a survey compare for example Winter [13] or the books of Hwang et al. [14] and of Prömel & Steger [15]. While the introduction of additional degree constraints has been received growing attention for spanning trees (see for example [16–18]), studies about degree-constrained Steiner trees are rare. As a subcase of degree-constrained network design problems [19–24] introduced bicriteria approximation algorithms, approximating both the objective value and the degree simultaneously. Besides, Furer & Raghavachari [25]; Khandekar et al. [19] introduced a related problem of finding a Steiner tree of minimal degree. However, up to our knowledge there do not exist theoretical studies, exact algorithms or heuristics for degree-constrained Steiner tree problems. In this paper, we consider the subcase of binary Steiner trees.

Binary Steiner trees are very important in biological and evolutionary questions. According to current theories of evolution, all species share a common history and are linked by common ancestors. These ancestral relationships can be represented by evolutionary trees such as tree alignments or phylogenetic trees, compare for example Zhang et al. [26] and Jarvis et al. [27] for a recent study. From the viewpoint of biologists, the terminals of a binary Steiner tree represent the given taxa, for example extant species or biomolecular sequences. Steiner nodes are the extinct taxa, i.e. the common ancestors, and the edge length represents the evolutionary time or number of mutations between the taxa. The principle of ‘Maximum Parsimony’ involves the identification of a phylogenetic tree that requires the smallest number of evolutionary changes, compare Camin & Sokal [28], Eck & Dayhoff [29] or Fitch [30]. Typically, evolutionary trees are required to be binary, since it is assumed that a phylogenetic tree is a bifurcation tree due to Darwin’s theory of the origin of species. Today there exist many different methods for constructing phylogenetic trees, including Maximum Parsimony, distance methods [31,32], maximum likelihood [33], and Bayesian inference [34,35]. Overviews of tree making algorithms are given by Felsenstein [36], Clote & Backofen [37], Hall [38], Penny & Hendy [39] and Swofford et al. [40]. Unfortunately, inferring such trees is a difficult problem, see [36] for a general introduction in the problem of inferring phylogenetic trees.

The connection between phylogenetic trees and Steiner trees has already been mentioned in the literature, compare Hwang et al. [14] and Cavalli-Sforza & Edwards [41]. Lu et al. [42] and Foulds & Graham [43] proved that the so called **Steiner problem in phylogeny** is $\mathcal{NP}$-hard. Cieslik [44] modeled phylogenetic trees as Steiner trees in a special metric space called sequence space and proved some theoretical results. However, up to our knowledge the degree constraints in evolutionary trees have not been considered in the context of Steiner trees so far. Further, integer programming formulations for phylogenetic trees have only been established for the minimum-evolution criterion, a special distance method [45], but not for parsimony methods and phylogenetic Steiner trees. This paper presents some theoretical and computational results on binary Steiner trees that may help to construct evolutionary trees.

The paper is organized as follows: In Section 2 we show that already the existence problem for binary Steiner trees is $\mathcal{NP}$-complete in general. Some approximation ratios based on Steiner trees without degree-constraints and binary spanning trees are presented in Section 3. We introduce two integer programming formulations in Section 4. The second one is a bidirected version of the first and used for the implementation of the branch-&-cut approach. The binary Steiner tree polytope is defined and some classes of valid and facet-inducing inequalities are considered in Section 5. Separation routines for these inequality classes are presented. In Section 6 we introduce a primal heuristic. Based on the polyhedral study we develop a branch-&-cut algorithm in Section 7 and discuss computational results.