European Journal of Operational Research 250 (2016) 745–753

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

Improved algorithms for joint optimization of facility locations and network connections

Xiaofan Lai, Zhou Xu*

Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 22 January 2015
Accepted 5 October 2015
Available online 20 October 2015

Keywords:
Approximation algorithm
Polynomial time algorithm
Facility location
Network connection
Steiner forest

ABSTRACT

This paper studies a k-median Steiner forest problem that jointly optimizes the opening of at most k facility locations and their connections to the client locations, so that each client is connected by a path to an open facility, with the total connection cost minimized. The problem has wide applications in the telecommunication and transportation industries, but is strongly NP-hard. In the literature, only a 2-approximation algorithm is known, it being based on a Lagrangian relaxation of the problem and using a sophisticated primal-dual schema. In this study, we have developed an improved approximation algorithm using a simple transformation from an optimal solution of a minimum spanning tree problem. Compared with the existing 2-approximation algorithm, our new algorithm not only achieves a better approximation ratio that is easier to be proved, but also guarantees to produce solutions of equal or better quality—up to 50 percent improvement in some cases. In addition, for two non-trivial special cases, where either every location contains a client, or all the locations are in a tree-shaped network, we have developed, for the first time in the literature, new algorithms that can solve the problem to optimality in polynomial time.

1. Introduction

Consider a complete undirected graph $G = (V, E)$ where $V = \{1, 2, \ldots, n\}$ denotes the vertex set, E denotes the edge set, and each edge $e \in E$ has a non-negative weight denoted by $\ell(e)$. Let $W \subseteq V$ denote a set of potential facilities, and $J \subseteq V$ a set of clients, where both W and J are not empty. Let k with $1 \leq k \leq |W|$ indicate the maximum number of facilities that are allowed to be opened. Each client in J needs to be served by connecting it to an open facility along a path. The resulting connections can be represented by *a k-median Steiner forest*, which is defined as a collection of at most k trees that covers all the clients, with each tree containing a distinct facility location as the root, where vertices in $V \backslash J$ are Steiner vertices that may or may not be included in the forest. To minimize the total connection cost, we study in this paper *a k-median Steiner forest problem* that aims to find an *optimal k-median Steiner forest* that minimizes the total edge weight. See Fig. 1.

The k-median Steiner forest problem aims to jointly optimize facility locations and network connections. This problem has wide applications, particularly in the telecommunication and transportation industries, where facilities, such as service centers or factories, need to be located and connected to clients by cable or road constructions. Its solutions can also be utilized to construct plans of facility locations combined with vehicle routing or cargo shipping (Carnes & Shmoys, 2011; Prodhon & Prins, 2014; Ravi & Sinha, 2006).

The k-median Steiner forest problem is strongly NP-hard, since it contains the classical Steiner tree problem as a special case with $|W| = k = 1$. As with the classical Steiner tree problem (Vazirani, 2001), it can be assumed without loss of generality that the weight of each edge equals the total edge weight of the shortest path that connects the endpoints of the edge, so that the edge weights satisfy triangle inequality and form a *metric*. This is because each edge of a k-median Steiner forest can always be replaced by the shortest path that connects its endpoints, without increasing the total edge weight of the Steiner forest.

Since the k-median Steiner forest problem is strongly NP-hard, it is of great interest to develop approximation algorithms for it with provable guarantees on running time and solution quality, as well as to identify special cases that are commonly seen in practice and can be solved to optimality by polynomial time algorithms. In this paper, we develop an improved approximation algorithm for the k-median Steiner forest problem, as well as new polynomial time algorithms that can solve two non-trivial special cases of the problem. For a minimization problem, we recall that an algorithm is said to be a ρ-approximation algorithm with an approximation ratio ρ if it runs in polynomial time and can always produce a feasible solution of an objective value no more than ρ times that of an optimal solution, and

* Corresponding author. Tel.: +85234003624; fax: +85223302704.
E-mail address: lgtzx@polyu.edu.hk (Z. Xu).

http://dx.doi.org/10.1016/j.ejor.2015.10.003
0377-2217/

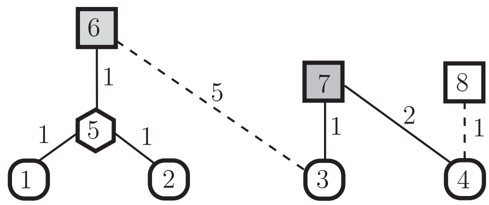

Fig. 1. An instance of the k-median Steiner forest problem with $k = 2$, $V = \{1, 2, \dots, 8\}$, $J = \{1, 2, 3, 4\}$ and $W = \{6, 7, 8\}$: Vertices 5, 6, 7, and 8 are Steiner vertices; the numbers on the edges indicate edge weights; for each edge not shown, its weight equals the total edge weight of the shortest path that connects its endpoints; the optimal k-median Steiner forest (shown in solid lines) opens facilities 6 and 7, and has a total edge weight of 6.

that the approximation ratio ρ is tight if there exists an instance of the problem to which the solution produced by the algorithm is of an objective value exactly equal to ρ times that of an optimal solution.

1.1. Previous work

For the k-median Steiner forest problem, only a 2-approximation algorithm is known in the literature (Carnes & Shmoys, 2011; Ravi & Sinha, 2006). It is based on a Lagrangian relaxation of the problem, and uses a sophisticated primal-dual schema that holds a so-called Lagrangian preserving performance guarantee to construct Steiner forests. (See a detailed review in Section 1 of the online supplementary materials.) As a result, the proof of its approximation ratio is complicated.

The k-median Steiner forest problem is a joint optimization problem that takes into account decisions on both facility locations and network design (Contreras & Fernández, 2012). Among many facility location problems that have been extensively studied in the literature (Aardal, van den Berg, Gijswijt, & Li, 2015; Prodhon & Prins, 2014), the k-median problem is the most relevant one, which aims to open at most k facilities and *directly* connect each client to an open facility by an edge with the total edge weight minimized. For the k-median problem, Charikar, Guha, Tardos, and Shmoys (1999) achieved the first constant approximation ratio of 6.67, and recently, based on a breakthrough made by Li and Svensson (2013), Byrka, Pensyl, Rybicki, Srinivasan, and Trinh (2015) achieved the current best approximation ratio of $2.61 + \epsilon$ for any $\epsilon > 0$. The existing 2-approximation algorithm for the k-median Steiner forest problem (Carnes & Shmoys, 2011) followed the approach of a 6-approximation algorithm developed by Jain and Vazirani (2001) for the k-median problem. Jain and Vazirani (2001) reduced the k-median problem to an *uncapacitated facility location problem* by relaxing the constraint of opening at most k facilities, and penalizing the opening of each facility by a Lagrangian multiplier. They then applied a primal-dual schema to obtain a 3-approximation of the uncapacitated facility location problem, and transformed it to a 6-approximation of the k-median problem. This approach has been improved by Jain, Mahdian, and Saberi (2002) and Arya et al. (2004) to achieve approximation ratios of 4 and $3 + \epsilon$ for the k-median problem.

Among various network design problems that have been extensively studied in the literature (Contreras & Fernández, 2012; Li & Balakrishnan, 2015), the classical Steiner tree problem is the most relevant one, which, as mentioned earlier, is a special case of the k-median Steiner forest problem with $|W| = k = 1$, aiming to minimize the total cost of connecting all the clients to a given facility. It is well-known that a minimum spanning tree of the subgraph induced by the clients can lead to a 2-approximation of the Steiner tree problem (Vazirani, 2001). Moreover, approximation ratios smaller than 2 have been achieved by Zelikovsky (1993),Karpinski and Zelikovsky (1997),Prömel and Steger (2000),Robins and Zelikovsky (2000), and with the current best approximation ratio being $\ln(4) + \epsilon < 1.39$, recently achieved by Byrka, Grandoni, Rothvoss, and Sanità (2013). Moreover, Ravi (1994) proposed a primal-dual schema that achieves an approximation ratio of 2 for a Steiner forest problem, which aims to connect clients by a given number of trees with the total edge weight minimized. This problem is equivalent to a special case of the k-median Steiner forest problem, where each vertex contains a facility, i.e., $W = V$.

Solutions to the k-median Steiner forest problem are often used to construct approximations of other problems that jointly optimize facility locations and network design (Drexl & Schneider, 2015). Ravi and Sinha (2006) studied a location-shipping problem that aims to open at most k facilities and install cables of sufficient capacity for shipping cargo from clients to facilities. By combining the 2-approximation of the k-median Steiner forest problem and a ρ-approximation of the k-median problem, they constructed a $(\rho + 2, 2)$ bicriteria approximation with a total cost at most $\rho + 2$ times that of the optimal solution, and with a total of at most $2k$ facilities opened. Carnes and Shmoys (2011) studied a k-location-routing problem that aims to assign depots (facilities) to k vehicles and to route the vehicles to serve clients, with the total routing cost minimized. By duplicating each edge of the 2-approximation of the k-median Steiner forest problem, they obtained a collection of k tours with each tour starting and ending at an open facility, and proved that such a tour collection is a 2-approximation of the k-location-routing problem. Moreover, Xu, Xu, and Xu (2013) studied a special case of the k-location-routing problem, where vertices are located in a tree-shaped network and edge weights represent lengths of shortest paths on the tree. Such a tree shaped network appears in several manufacturing and logistics applications (Asano, Katoh, & Kawashima, 2001; Chen, Campbell, & Thomas, 2008; Chhajed & Lowe, 1992; Karuno, Nagamochi, & Ibaraki, 1996; Wang, Lim, & Xu, 2006), including those in rural or water transportation systems (Tsitsiklis, 1992; Xu, Lai, Lim, & Wang, 2014). For this special case, solutions to the k-median Steiner forest problem and solutions to the k-location-routing problem are one-to-one correspondence, and therefore these two problems are equivalent. Xu et al. (2013) developed an algorithm that can solve this special case of the problem to optimality in $O(n2^{6k})$ time. However, when k is part of the input, whether or not this special case has a polynomial time algorithm still remains open.

1.2. Our results

For the k-median Steiner forest problem, we have developed a new 2-approximation algorithm. It is simpler than the existing 2-approximation algorithm of Carnes and Shmoys (2011), consisting of only an $O(n^2)$-time transformation from a minimum spanning tree of the clients and a new vertex that replaces all the facilities. This extends the well-known result that a minimum spanning tree of the clients can lead to a 2-approximation of the Steiner tree problem. Compared with the existing 2-approximation algorithm, our new algorithm has an improved approximation ratio of $2 - 1/|J|$, which is not only tight but easier to be proved, and it can always produce solutions of equal or better quality, the quality improvement being up to 50 percent in some cases.

Moreover, we have developed new polynomial time algorithms that can solve two non-trivial special cases of the k-median Steiner forest problem to optimality. For a special case where each vertex contains a client, i.e, $J = V$, we show that it is equivalent to a problem of finding a minimum weighted basis for a matroid, and that its optimal solution can be obtained in polynomial time by a transformation from a minimum spanning tree of the clients. This result is interesting because the same special case for the k-median problem, where $J = V$, is still strongly NP-hard. For the other special case, where vertices are located in a tree-shaped network, we develop a dynamic programming algorithm that, for the first time in the literature, can solve the problem to optimality in polynomial time, a