Available online at www.sciencedirect.com
ScienceDirect
Theoretical Computer Science 381 (2007) 288–291
Theoretical Computer Science
www.elsevier.com/locate/tcs

Note

Approximating the selected-internal Steiner tree☆

Sun-Yuan Hsieh*, Shih-Cheng Yang

Department of Computer Science and Information Engineering, National Cheng Kung University, No. 1, University Road, Tainan 70101, Taiwan

Received 24 March 2007; received in revised form 22 May 2007; accepted 27 May 2007

Communicated by D.-Z. Du

Abstract

In this paper, we consider a variant of the well-known Steiner tree problem. Given a complete graph $G = (V, E)$ with a cost function $c : E \to \mathbf{R}^+$ and two subsets R and R' satisfying $R' \subset R \subseteq V$, a *selected-internal Steiner tree* is a Steiner tree which contains (or spans) all the vertices in R such that each vertex in R' cannot be a leaf. The *selected-internal Steiner tree problem* is to find a selected-internal Steiner tree with the minimum cost. In this paper, we present a 2ρ-approximation algorithm for the problem, where ρ is the best-known approximation ratio for the Steiner tree problem.

Keywords: Design and analysis of algorithms; Approximation algorithms; Steiner tree; The selected-internal Steiner tree problem; MAX SNP-hard

1. Introduction

In the past years, the Steiner tree problem (STP for short) and its variants have received a lot of focuses because they have many important applications such as VLSI design, network routing, telecommunications, wireless communications, transportation, and so on [2–7,10–12]. There are several well-known variants of the Steiner tree problem, such as the Steiner tree problem on special metric spaces (e.g., the Euclidean metric [8] and the rectilinear metric [9]), the terminal Steiner tree problem [13], and so on. All of the above problems were shown to be NP-complete [8,9,13].

In this paper, we study a variant of the Steiner tree problem. Given a complete graph $G = (V, E)$ with a cost function $c : E \to \mathbf{R}^+$ and two subsets R and R' satisfying $R' \subset R \subseteq V$, the *selected-internal Steiner tree problem* (SISTP for short) is to find a Steiner minimum tree which spans all the vertices in R such that each vertex in R' cannot be a leaf. For convenience, we call such a tree as *optimal selected-internal Steiner tree*, and call the vertices in R' the *demanded terminals*. Since the Steiner tree problem is a special case of SISTP, the NP-completeness and MAX SNP-

☆ An extended abstract of this paper appeared in *Proceedings of 12th Annual International Computing and Combinatorics Conference (COCOON 2006), Lecture Notes in Computer Science 4112*, pp. 449–458, 2006. This work was supported in part by the National Science Council under grants NSC 94-2213-E-006-073 and NSC 95-2221-E-006-076.

* Corresponding author.
E-mail address: hsiehsy@mail.ncku.edu.tw (S.-Y. Hsieh).

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2007.05.035

hardness of the problem[1] follows immediately from the hardness results of the Steiner tree problem. It is conceivable that this problem maybe of practical interest. For example, in a network resource allocation, some specified servers (terminals) must act as transmitters and the others need not have this restriction. Consequently, in a solution tree, some terminals are restricted to be internal vertices and the others can be leaves or internal vertices. Another example is in a sensor network, some nodes might be especially cheap devices that can receive but cannot transmit.

In this paper, we present a 2ρ-approximation algorithm for the problem, where ρ is the best-known approximation ratio for SISTP. The rest of this paper is organized as follows: In the next Section 2, we present an approximation algorithm for the problem. Finally, some concluding remarks are provided in Section 3.

2. An approximation algorithm

In this section, we present an approximation Algorithm A_{SISTP} for SISTP. Some useful definitions and notations are first provided. For a graph G, the *degree* of v in G, denoted by $\deg_G(v)$, is the number of edges incident to v in G. A *path* of *length* k from a vertex v_0 to a vertex v_k in a graph $G = (V, E)$ is a sequence $\langle v_0, v_1, v_2, \ldots, v_k\rangle$ of vertices such that $(v_{i-1}, v_i) \in E$ for $i = 1, 2, \ldots, k$. We use $P_G[u, u']$ to denote a path from u to u' in G. A path is *simple* if all vertices in the path are distinct. The cost function c used throughout this paper is *metric*. For an edge e in a tree T, $c(e)$ is the *cost* of e, and $c(T)$ is the sum of all the edge costs of T.

Let A_{STP} denote the best-known approximation algorithm for STP with ratio $\rho = 1 + \frac{\ln 3}{2} \approx 1.55$ [14], and also let $S_A = (V_A, E_A)$ be the Steiner tree returned by A_{STP}. For a Steiner tree T of the instance $I = (G, R, R', c)$ of the problem SISTP, a vertex $v \in V(T)$ is said to be a *demand-leaf* if v is a leaf of T and $v \in R'$. Hereafter, we assume that $|R \setminus R'| \geq 2$ if $R' \neq \emptyset$ (to make sure that the solution of SISTP exists).

Lemma 1. *Let T be a Steiner tree of the instance $I = (G, R, R', c)$ of the problem SISTP such that $|R \setminus R'| \geq 2$. If v is a demand-leaf of T, then there is an internal vertex $m_v \in V(T)$ satisfying one of the following two conditions:* (1) $\deg_T(m_v) = 2$ *and* $m_v \notin R'$, *and* (2) $\deg_T(m_v) \geq 3$.

Proof. If neither Condition (1) nor Condition (2) holds, then the resulting tree T is a path $\langle v, v_1, v_2, \ldots, v_{n-1}\rangle$ such that $v, v_1, v_2, \ldots, v_{n-2}$ are all in R'. Then, $|R \setminus R'| \leq 1 < 2$, which contradicts to the assumption that $|R \setminus R'| \geq 2$. □

The skeleton of our algorithm is first to apply A_{STP} to obtain a Steiner tree $S_A = (V_A, E_A)$ spanning R, and then transform it to a selected-internal Steiner tree based on Lemma 1 to make each demand-leaf of S_A to be an internal vertex. We next present our approximation algorithm, namely Algorithm A_{SISTP}. We call the two vertices m_v and t_v selected by Algorithm A_{SISTP} for each demand-leaf v the *medium vertex* and the *target vertex* of v, respectively.

It is not difficult to show the following result, which is useful for analyzing the approximation ratio of our algorithm in Theorem 5.

Lemma 2. *Let $v_1, v_2, \ldots, v_l$ be an order of the demand-leaves of S_A handled by Algorithm A_{SISTP}. Then, the paths $P_{S_A}[v_1, m_{v_1}], P_{S_A}[v_2, m_{v_2}], \ldots, P_{S_A}[v_l, m_{v_l}]$ are pairwise edge-disjoint.*

Lemma 3. *Suppose that v is a demand-leaf of the current tree T, which is being handled by Algorithm A_{SISTP}. Then, the target vertex t_v does not belong to $P_{S_A}[v, m_v]$.*

Proof. According to Line 8 of the algorithm, it is clear that t_v does not belong to $P_T[v, m_v]$, i.e., t_v is not a vertex in $P_T[v, m_v]$. Assume that $P_T[v, m_v] = \langle v_0(=v), v_1, v_2, \ldots, v_k(=m_v)\rangle$. It can be shown by induction that the vertices of $P_{S_A}[v, m_v]$ are contained in $P_T[v, m_v]$ and the relative order of the vertices in $P_{S_A}[v, m_v]$ are retained in $P_T[v, m_v]$, i.e., $P_{S_A}[v, m_v] = \langle v_0, v_{i_1}, v_{i_2}, \ldots, v_{i_j}, v_k\rangle$, where $0 < i_1 < i_2 < \cdots < i_j < k$. Therefore, t_v is not a vertex in $P_{S_A}[v, m_v]$. □

Lemma 4. *Let $P = \langle v_1, v_2, \ldots, v_{k-1}, v_k\rangle$ be a path of a graph $G = (V, E)$ with a metric cost function $c : E \to \mathbf{R}^+$, and let $P' = \langle v_1, v_2, \ldots, v_{k-2}, v_{k-1}\rangle$. Then, $c(v_1, v_k) - c(v_{k-1}, v_k) \leq c(P')$, where $c(P') = \sum_{j=1}^{k-2} c(v_j, v_{j+1})$.*

[1] It was shown that if any MAX SNP-hard problem has a polynomial-time approximation scheme (PTAS), then P = NP [1]. In other words, it is very unlikely that for a MAX SNP-hard problem to have a PTAS.

Algorithm $A_{SISTP}(G, R, R', c)$

Input: A complete graph $G = (V, E)$ with a metric cost function $c : E \to \mathbf{R}^+$, and two subsets R and R' satisfying $R' \subset R \subseteq V$ and $|R \setminus R'| \geq 2$ if $R' \neq \emptyset$.
Output: A selected-internal Steiner tree T_s.

Use A_{STP} to find a Steiner tree $S_A = (V_A, E_A)$ spanning R in G.
$S'_A \leftarrow S_A$.
if there is a demand-leaf in S'_A **then**
  for each demand-leaf v in the current tree S'_A **do**
    Select the nearest vertex $m_v \in V(S'_A)$ satisfying one of the following condition:
    (1) $\deg_{S'_A}(m_v) = 2$ and $m_v \notin R'$
    (2) $\deg_{S'_A}(m_v) \geq 3$. ▷ The existence of m_v is ensured by Lemma 1
    Choose a vertex $t_v \in V(S'_A)$ which is adjacent to m_v but does not belong to the path $P_{S'_A}[v, m_v]$.
    $E(S'_A) \leftarrow E(S'_A) \cup \{(v, t_v)\}$.
    $E(S'_A) \leftarrow E(S'_A) \setminus \{(m_v, t_v)\}$.
Return $T_s \leftarrow S'_A$.

Proof. Since c is metric, $c(v_1, v_k) - c(v_{k-1}, v_k) \leq c(v_1, v_{k-1})$, $c(v_1, v_{k-1}) - c(v_{k-2}, v_{k-1}) \leq c(v_1, v_{k-2})$, $c(v_1, v_{k-2}) - c(v_{k-3}, v_{k-2}) \leq c(v_1, v_{k-3}), \ldots, c(v_1, v_3) - c(v_2, v_3) \leq c(v_1, v_2)$. By combining these inequalities, we have that $c(v_1, v_k) - c(v_{k-1}, v_k) \leq c(v_1, v_2) + c(v_2, v_3) + \cdots + c(v_{k-3}, v_{k-2}) + c(v_{k-2}, v_{k-1}) = c(P')$. □

Let $L_R = \{v \mid v \in R \text{ is a leaf of } S_A\}$ and $L_{R'} = \{v \mid v \text{ is a demand-leaf of } S_A\}$. Note that $L_{R'} \subseteq L_R$. Define $\phi = \min\{|L_R \setminus L_{R'}|, |R \setminus R'| - 2\}$. We now show our main result.

Theorem 5. *Let $e_1, e_2, \ldots, e_i$ denote the first i smallest-cost edges of $S_A = (V_A, E_A)$. Algorithm A_{SISTP} is a $(2 - \frac{\sigma}{c(S_A)})\rho$-approximation algorithm for SISTP, where ρ is the best-known approximation ratio of the Steiner tree problem and σ is defined as follows:*

$$\sigma = \begin{cases} c(S_A) & \text{if } R' = \emptyset, \\ \sum_{i=1}^{\phi} c(e_i) & \text{otherwise}. \end{cases}$$

Moreover, $0 \leq \frac{\sigma}{c(S_A)} \leq 1$.

Proof. It is clear that Algorithm A_{SISTP} correctly constructs a selected-internal Steiner tree. We now analyze the approximation ratio. Let T_s, T^*, and S^* be the output of A_{SISTP}, the optimal solution of SISTP and the optimal solution of STP, respectively. Since S_A is the output of Algorithm A_{STP}, we have that $c(S_A) \leq \rho c(S^*)$. Since T^* is a feasible solution of STP, $c(S^*) \leq c(T^*)$. Therefore, $c(S_A) \leq \rho c(T^*)$.

Next, we consider the following two cases according to the demanded-terminal set R'.

CASE 1: $R' = \emptyset$. Then, $c(T_s) = c(S_A) \leq \rho c(T^*) = (2 - \frac{c(S_A)}{c(S_A)})\rho c(T^*)$. The theorem holds.

CASE 2: $R' \neq \emptyset$ and $|R \setminus R'| \geq 2$. According to Algorithm A_{SISTP}, $c(T_s) = c(S_A) + \sum_{v \in L_{R'}}(c(v, t_v) - c(m_v, t_v))$. According to Lemmas 3 and 4, we know that $\sum_{v \in L_{R'}}(c(v, t_v) - c(m_v, t_v)) \leq \sum_{v \in L_{R'}} c(P_{S_A}[v, m_v])$. Define Q to be the set obtained by selecting arbitrary ϕ (=min$\{|L_R \setminus L_{R'}|, |R \setminus R'| - 2\}$) elements from $L_R \setminus L_{R'}$. Note that

$$Q = \begin{cases} \emptyset & \text{if } \phi = 0, \\ L_R \setminus L_{R'} & \text{if } \phi = |L_R \setminus L_{R'}|, \\ \text{a proper subset of } L_R \setminus L_{R'} & \text{otherwise}. \end{cases}$$

If we transform $L_{R'} \cup Q$ into internal vertices using Algorithm $A_{SISTP}(G, R, Q \cup R', c)$, then the resulting tree remains a selected-internal Steiner tree. (Note that the algorithm actually transform only $L_{R'}$ into internal vertices.) By above observation together with Lemma 2, we have $\sum_{v \in L_{R'}} c(P_{S_A}[v, m_v]) + \sum_{v \in Q} c(P_{S_A}[v, m_v]) \leq c(S_A)$.

Therefore, we have that $c(T_s) = c(S_A) + \sum_{v \in L_{R'}} (c(v, t_v) - c(m_v, t_v))$ (by the algorithm) $\leq c(S_A) + \sum_{v \in L_{R'}} c(P_{S_A}[v, m_v])$ (by Lemmas 3 and 4) $\leq c(S_A) + (c(S_A) - \sum_{v \in Q} c(P_{S_A}[v, m_v])) \leq 2c(S_A) - \sum_{i=1}^{|Q|} c(e_i)$ (by the fact that $e_1, e_2, \ldots, e_{|Q|}$ is the first $|Q|$ smallest-cost edges in S_A) $= 2c(S_A) - \sum_{i=1}^{\phi} c(e_i) = (2 - \frac{\sum_{i=1}^{\phi} c(e_i)}{c(S_A)})c(S_A) \leq (2 - \frac{\sum_{i=1}^{\phi} c(e_i)}{c(S_A)})\rho c(T^*)$ (by $c(S_A) \leq \rho c(T^*)$). Therefore, $\frac{c(T_s)}{c(T^*)} \leq (2 - \frac{\sum_{i=1}^{\phi} c(e_i)}{c(S_A)})\rho$. It is clear that $0 \leq \frac{\sigma}{c(S_A)} \leq 1$. □

We next analyze the time complexity of Algorithm A_{SISTP}. Let $f(n, m)$ be the time complexity of an approximation algorithm A_{STP} for the Steiner tree problem, where n and m are the numbers of vertices and edges of the input graph G, respectively. Hence, constructing S_A in Step 1 takes $f(n, m)$ time. Based on the linear-time depth-first-search to find the medium vertex m_v for each demand-leaf v, the overall for-loop (lines 4–10) can be implemented to run in $O(|R'|(|V_A| + |E_A|)) = O(|R'||V_A|) = O(n^2)$ time. Therefore, we have the following result.

Theorem 6. *Algorithm A_{SISTP} can be implemented to run in time $O(n^2) + f(n, m)$.*

3. Concluding remarks

In this paper, we develop an approximation algorithm for the selected-internal Steiner tree problem. A future work is to extend our result to obtain a better approximation algorithm or study the case where the cost function is not metric.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy, Proof verification and the hardness of approximation problems, Journal of the Association for Computing Machinery 45 (1998) 501–555.
[2] M. Bern, Faster exact algorithms for Steiner tree in planar networks, Networks 20 (1990) 109–120.
[3] A. Borchers, D.Z. Du, The k-Steiner ratio in graphs, SIAM Journal on Computing 26 (3) (1997) 857–869.
[4] X. Cheng, D.Z. Du, Steiner Trees in Industry, Kluwer Academic Publishers, Dordrecht, Netherlands, 2001.
[5] D.Z. Du, On component-size bounded Steiner trees, Discrete Applied Mathematics 60 (1995) 131–140.
[6] D.Z. Du, J.M. Smith, J.H. Rubinstein, Advance in Steiner Tree, Kluwer Academic Publishers, Dordrecht, Netherlands, 2000.
[7] L. Foulds, R. Graham, The Steiner problem in phylogeny is NP-complete, Advances in Applied Mathematics 3 (1982) 43–49.
[8] M. Garey, R. Graham, D. Johnson, The complexity of computing Steiner minimal trees, SIAM Journal on Applied Mathematics 32 (1977) 835–859.
[9] M. Garey, D. Johnson, The rectilinear Steiner problem is NP-complete, SIAM Journal on Applied Mathematics 32 (1977) 826–834.
[10] D. Graur, W.H. Li, Fundamentals of Molecular Evolution, Second ed., Sinauer Publishers, Sunderland, Massachusetts, 2000.
[11] S. Hougardy, H.J. Prömel, A 1.598 approximation algorithm for the Steiner tree problem in graphs, in: Proceedings of the 10th Annual ACM–SIAM Symposium on Discrete Algorithms, SODA, 1999, pp. 448–453.
[12] A.B. Kahng, G. Robins, On Optimal Interconnections for VLSI, Kluwer Publishers, 1995.
[13] C.L. Lu, C.Y. Tang, R.C.T. Lee, The full Steiner tree problem, Theoretical Computer Science 306 (2003) 55–67.
[14] G. Robins, A. Zelikovsky, Improved Steiner tree approximation in graphs, in: Proceedings of the 11th Annual ACM–SIAM Symposium on Discrete Algorithms, SODA, 2000, pp. 770–779.