Better approximation bounds for the network and Euclidean Steiner tree problems

Alexander Zelikovsky *

Abstract

The network and Euclidean Steiner tree problems require a shortest tree spanning a given vertex subset within a network $G = (V, E, d)$ and Euclidean plane, respectively. For these problems, we present a series of heuristics finding approximate Steiner tree with performance guarantee coming arbitrary close to $1 + \ln 2 \approx 1.693$ and $1 + \ln \frac{2}{\sqrt{3}} \approx 1.1438$, respectively. The best previously known corresponding values are close to 1.746 and 1.1546.

Keywords: Combinatorial problems, approximation algorithms, Steiner trees.

1 Introduction

Let $G = (V, E, d)$ be a graph with a vertex set V, an edge set E and distance function $d : E \to R^+$. A tree T is called a *Steiner tree* of S, $S \subset V$, if S is contained in the vertex set of T.

Network Steiner Problem (NSP). Given G and S, find the shortest Steiner tree (also called the *Steiner minimal tree*) of S.

This problem is NP-complete [9], so many approximation algorithms for Steiner minimal trees appeared in the last two decades. The quality of an approximation algorithm is measured by its performance ratio: an upper bound on the ratio between the achieved length and the optimal length.

NSP belongs to $MAXSNP$-class [3], so the constant factor approximation algorithm exists [13] and for some $\epsilon > 1$, ϵ-approximation is NP-complete [1]. To find such ϵ is an important open problem.

Without loss of generality we may assume that G is complete, i.e. for any $u - v$-path in G, there is an edge $(u, v) \in E$. Moreover, the length of any edge in G coincides with the length of the minimal path between its ends. G_S denotes a subgraph of G induced by the set S.

Euclidean Steiner Problem (ESP). Given a point set S in Euclidean plane, find the shortest Steiner tree spanning S.

*Dept of CS, Thornton Hall, UVA, Charlottesville, VA 22903-2442 email: alexz@cs.virginia.edu. Research partially supported by Volkswagen Stiftung.

For ESP, the graph G_S is the same as for NSP.

A well-known minimum spanning tree heuristic for the Steiner tree problem approximates a Steiner minimal tree with a minimum length spanning tree (MST) of a complete graph G_S. Du and Hwang [7], Takahashi and Matsuyama [13] proved that the exact performance ratios of this heuristic are equal to $\frac{2}{\sqrt{3}} \approx 1.1547$ and 2 for ESP and NSP, respectively.

Two better heuristics appeared recently while consideration of so called $k-$restricted Steiner trees. The approximation guarantee of the generalized greedy heuristic is bounded by 11/6 [15] for NSP and 1.1546 for ESP [8]. The series of evaluation heuristics achieves better guarantees while increasing of their runtime. Their performance ratio converges to a value close to 1.746 [4, 5] for NSP.

The main result of this paper is the following

Theorem 1 *There is a polynomial-time* $(1 + \ln r_2)$*-approximation scheme* $\{A_k, \quad k = 2, 3, ...\}$ *for the Steiner tree problems, where* r_2 *is a performance ratio of the minimum spanning tree heuristic.*

Corollary 1 *There is a polynomial-time* $(1 + \ln 2)$*- and* $(1 + \ln \frac{2}{\sqrt{3}})$*-approximation scheme for NSP and ESP, respectively.*

In the next section we describe a general framework for different heuristics solving Steiner tree problems and approximation algorithms A_k. In section 3 we prove the performance guarantee for these algorithms.

2 The greedy contraction framework

First we introduce some denotations: $Smt(S)$ and $smt(S)$ are a Steiner minimal tree of S and its length, respectively. For a complete graph G_S, $Mst(S)$ denotes MST of G_S, and $M(S)$ denotes its length. $Smt(S)$ may in general contain vertices of $V \backslash S$. So any Steiner tree contains the set S of *terminal vertices* and some additional vertices. A Steiner tree is called a *full* Steiner tree, if it does not contain internal terminal vertices. If a Steiner tree is not full, then we can split it into the union of edge-disjoint full Steiner subtrees which are called *full components*.

Contraction of a full Steiner tree T means reducing to 0 the lengths of edges of G_S connecting terminal vertices of T. We denote by S/T the result of contraction. Note that contraction reduces the value $M(S)$, i.e. $M(S/T) \leq M(S)$.

For all the Steiner tree problems, the following greedy contraction framework is successfully used in approximations.

Greedy contraction framework (GCF)

(1) repeat until $M_0(S) = 0$
    (a) find a full Steiner tree T^* in *a class* K which minimizes
        a criterion function $f(T)$: $T^* \leftarrow arg\min_{T \in K} f(T)$.
    (b) insert T^* in $LIST$.
    (c) contract T^*, $S \leftarrow S/T^*$.
(2) reconstruct an output Steiner tree from trees of $LIST$.

To determine a performance guarantee of an algorithm A embedded in GCF we may bound the following two ratios:

$$a_1 = \frac{smt_{\tilde{K}}}{smt},$$

where $smt_{\tilde{K}}$ is the the minimum tree length in the family $\tilde{K}$ containing all Steiner trees with full components belonging to K, and

$$a_2 = \frac{st_A}{smt_{\tilde{K}}},$$

where st_A is the length of the output Steiner tree of A. Thus, a performance ratio $pr(A)$ equals to their product $pr(A) = a_1 \cdot a_2$.

Many famous heuristics can be embedded in this framework considering different definitions of a class K and a criterion function f.

The minimum spanning tree heuristic. The class K consists of all paths in G and $f(T) = d(T)$. The exact bounds for a_1 were proved in [7, 13]. An equality $a_2 = 1$ follows from the famous fact that the greedy algorithm finds exact MST of a graph.

The Rayward-Smith's heuristic (RSH) [12]. The class K contains stars and $f(T) = \frac{d(T)}{r-1}$, where r is the number of leaves of T. For NSP, exact upper bounds $a_1 \leq 5/3$ and $a_1 \cdot a_2 \leq 2$ were proved in [15, 16] and [14], respectively.

The generalized greedy heuristic (GGH) [16]. The class K consists of trees with k terminal vertices and $f(T) = d(T) - (M(S) - M(S/T))$. For NSP, the same (as for RSH) exact upper bound $a_1 \leq 5/3$ [15, 16] holds. For $k = 3$, the upper bound $a_1 \cdot a_2 \leq 11/6$ [16] may be not exact. Berman [6] found an instance of NSP such that $a_1 \cdot a_2 > 5/3$ and conjectured that $a_1 \cdot a_2 \leq 43/24$ ($k = 3$). For ESP, $a_1 \cdot a_2 \leq 1.1546$ ($k = 128$) [8].

In this paper we present

The relative greedy heuristic (RGH). The class $K = K_k$ contains all trees with at most k terminal vertices. The criterion function f is defined slightly different than for GGH.

$$f(T) = \frac{d(T)}{M(S) - M(S/T)} \tag{1}$$

Now we describe the class $K = K_k$ and review known facts about it. A Steiner tree is called $k-$*restricted* if all its full components have at most k terminal vertices. Let the shortest $k-$restricted Steiner tree for the set S, denoted by $Smt_k(S)$, have the length $smt_k(S)$. Note, that $Smt_2(S) = M(S)$.

Below we give constants for ESP in brackets, they should be less than for NSP. Let $r_k = sup\{smt_k(S)/smt(S)\}$. The bound for the MST-heuristic implies $r_2 = 2$ [13] ($r_2 = \frac{2}{\sqrt{3}}$ [7]). As mentioned above, $r_3 = 5/3$ [15, 16] (conjectured value ≈ 1.073 [8]), $r_4 \leq \frac{3}{2}$ and $r_8 \leq \frac{4}{3}$ [4]. Moreover, $r_{2^k} \leq 1 + \frac{1}{k}$ [8]. Unfortunately, even when $k = 4$, computing optimal $k-$restricted Steiner tree is NP-hard [9]. For $k = 3$, this problem is still open although V. Arora et al. [2] applied so called backtrack greedy approach and conjectured that it gave polynomial-time approximation scheme.

GGH and RGH approximate $k-$restricted Steiner minimal trees. Berman and Ramayer suggested a more complicated way to approximate such trees [5]. They construct a family of *evaluation* heuristics B_k and bound a performance ratio of B_k with the following value

$$pr(B_k) \le r_2 - \sum_{i=3}^{k} \frac{r_{i-1} - r_i}{i-1}.$$

Knowing bounds for r_{2^k} we may count an upper bound for $\lim_{k\to\infty} pr(B_k)$ which is close to 1.746. Since r_k is not evaluated for k unequal to powers of 2, this value might be less. The next step [10] decreases the bound for Berman/Ramayer algorithm by $\frac{1}{48} = 0.02$.

For RGH, $\lim_{k\to\infty} a_1 = \lim_{k\to\infty} r_k = 1$. In the next section we prove that $a_2 \le 1 + \ln 2$. In other words, it induces a polynomial $(1+\ln 2)$-approximation scheme for STP. Again we are not sure that $1 + \ln 2$ is an exact upper bound for this scheme.

3 The performance guarantee

In this section we determine a performance ratio of a relative greedy heuristic A_k while it approximates a minimal k-restricted Steiner tree. In other words, we bound the value $a_2 = a_2(A_k)$. For brevity, below we omit S in the following denotations: $smt(S)$, $smt_k(S)$, $M(S)$.

Lemma 1 $a_2(A_k) \le 1 + \ln \frac{M}{smt_k}$.

Proof. We need the following denotations. Let T be the output tree of RGH A_k and $T = T_1 \cup \cdots \cup T_l$ be a partition of T into its full components, i.e. $LIST = \{T_1, ..., T_l\}$. We denote the length of T_i by $d_i = d(T_i)$, $M_0 = M(S)$. Recursively, $M_i = M(S/T_i)$ and

$$m_i = M_{i-1} - M_i \tag{2}$$

So the criterial function (1) equals $f(T_i) = \frac{d_i}{m_i}$.

We also partition the optimal k-restricted Steiner tree T^* into its full components: $T^* = \cup T_j^*$. It is easy to see that

$$f(T^*) \ge \min d(T_j^*) \ge f(T_1),$$

since $T_1 = arg \min f(T)$. So we obtain

$$\frac{d_1}{m_1} \le \frac{smt_k}{M_0}$$

We may apply the same argument to the k-restricted Steiner tree after contraction of T_1.

$$\frac{d_2}{m_2} \le \frac{smt_k(S/T_1)}{M_1} \le \frac{smt_k}{M_1}$$

Inductively we obtain

$$\frac{d_i}{m_i} \le \frac{smt_k}{M_{i-1}}, \quad i = 1, ..., l \tag{3}$$

Now we apply an analysis technique due to Leighton and Rao [11] to prove Inequality (5). Substituting Equality (2) into (3), we get

$$M_i \le M_{i-1}(1 - \frac{d_i}{smt_k}) \tag{4}$$

Unraveling (4), we obtain

$$M_r \leq M_0 \prod_{i=1}^{r} (1 - \frac{d_i}{smt_k}).$$

Taking natural logarithm on both sides and simplifying using the approximation $\ln(1+x) \leq x$, we obtain

$$\ln \frac{M_0}{M_r} \geq \frac{\sum_{i=1}^{r} d_i}{smt_k} \tag{5}$$

Note that $M_l = M_{l-1} - m_l = 0$, since Step (1) of GCF interrupts when $M(S) = 0$. So we can choose r such that $M_r > smt_k \geq M_{r+1}$. We partition m_{r+1} into two parts $m_{r+1} = m^* + m'$ such that $m^* = M_r - smt_k$ and $m' = smt_k - M_{r+1}$. We also partition d_{r+1} in the same proportion, i.e. $d_{r+1} = d^* + d'$ and $\frac{d_{r+1}}{m_{r+1}} = \frac{d^*}{m^*} = \frac{d'}{m'}$ ($d' = 0$ if $m' = 0$). Denote $M^*_{r+1} = M_r - m^*$.

These denotations allows us to put down (5) for M^*_{r+1}

$$\ln \frac{M_0}{M^*_{r+1}} \geq \frac{\sum_{i=1}^{r} d_i + d^*}{smt_k}$$

Note that $\frac{d_i}{m_i} \leq 1$. Indeed, $T_i = arg \min f(T)$, but $f(e) = 1$ for any nonzero edge $e \in Mst(S)$. Therefore,

$$d(T) \leq \sum_{i=1}^{r+1} d_i + M_{r+1} = \sum_{i=1}^{r} d_i + d^* + d' + M_{r+1}$$

Thus, we are ready to finish the proof of Lemma:

$$a_2(A_k) = \frac{d(T)}{smt_k} \leq$$

$$\frac{\sum_{i=1}^{r} d_i + d^*}{smt_k} + \frac{d' + M_{r+1}}{smt_k} \leq$$

$$\ln \frac{M_0}{M^*_{r+1}} + \frac{m' + M_{r+1}}{smt_k} = \ln \frac{M_0}{smt_k} + 1 \quad \diamondsuit$$

To prove Theorem 1 we only need to note that

$$\frac{M}{smt_k} \leq \frac{M}{smt} \quad \diamondsuit$$

4 Conclusion

In this paper we introduced a greedy conraction framework for a wide class of heuristics for the Steiner tree problem. A series of new *relative greedy* heuristics embedded in this framework has a better approximation ratio which tends to $1+\ln r_2$, where r_2 is the approximation ratio for the MST-heuristic.

We applied relative greedy heuristics to the Steiner tree problems in networks and in Euclidean plane. The performance guarantee of these heuristics comes close to $1 + \ln 2 \approx 1.693$ and $1+\ln \frac{2}{\sqrt{3}} \approx 1.1438$, respectively. These values improved the best previously known bounds of 1.746 and 1.1546.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. In *Proc. of 33d Annual IEEE Symp. on Foundations of Computer Science*, 14–23, 1992.

[2] V. Arora, V. Santosh, H. Saran and V. Vazirani. A limited-backtrack greedy schema for approximation algorithms. Manuscripts, 1994.

[3] M. Bern and P. Plassmann. The Steiner problems with edge lengths 1 and 2. *Inform. Process. Lett.* 32: 171–176, 1989.

[4] P. Berman and V. Ramaiyer. An approximation algorithm for the Steiner tree problem. *Tech. Rep.#CS-91-05, The Pennsylvania State University*, 1991.

[5] P. Berman and V. Ramaiyer. Improved approximations for the Steiner tree problem. In *Proc. of 3d Annual ACM-SIAM Symp. on Discrete Algorithms*, 325–334, 1992.

[6] P.Berman. Personal communication. 1993.

[7] D. Du and F.K. Hwang. An approach for proving lower bounds: solution of Gilbert-Pollak's conjecture on Steiner ratio. In *Proc. of 31st Annual IEEE Symp. on Foundations of Computer Science*, 76–85, 1990.

[8] D. Du, Y. Zhang and Q. Feng. On better heuristic for Euclidean Steiner minimum trees. In *Proc. of 32nd Annual IEEE Symp. on Foundations of Computer Science*, 431–439, 1991.

[9] R. M. Karp. Reducibility among combinatorial problems. In Miller and Thatcher (eds.), *Complexity of Computer Computations*, Plenum Press, New York, 85–103, 1972.

[10] M. Karpinski and A. Zelikovsky. Further improvements of Steiner tree approximations. Manuscripts, 1994.

[11] F. T. Leighton and S. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with application to approximation algorithms. In *Proc. of 29th Annual IEEE Symp. on Foundations of Computer Science*, 422–431, 1988.

[12] V. J. Rayward-Smith, The computation of nearly minimal Steiner trees in graphs, *International J. Math. Ed. Sci. Tech.* 14: 15–23, 1983.

[13] H. Takahashi and A. Matsuyama. An approximate solution for the Steiner problem in graphs. *Math. Japonica*, 24: 573–577, 1980.

[14] B. M. Waxman and M. Imase. Worst case-performance of Rayward-Smith's Steiner tree heuristic. *Inform. Process. Lett.* 29: 283–287, 1988.

[15] A. Z. Zelikovsky. The Steiner problem in graphs without optimal vertices. *Matemat. Issled.* 121: 44–49, 1991 (in Russian).

[16] A. Z. Zelikovsky. An 11/6-approximation algorithm for the network Steiner problem. *Algorithmica* 9: 463–470, 1993.