Approximation Hardness of the (1, 2)-Steiner Tree Problem

Mathias Hauptmann*

Abstract

We give a survey on the approximation hardness of the Steiner Tree Problem. While for the general metric case, Chlebik and Chlebikova [CC02] prove a lower bound of $\approx 1.01$, this result does not seem to apply to bounded metrics. We show that combining an $L$-reduction from the Bounded Degree Vertex Cover Problem to the (1,2)-Steiner Tree Problem due to Bern and Plassmann [BP99] with approximation hardness results of Berman and Karpinski [BK03], one obtains the currently best known lower bound of $\approx 1.0026$ for the approximability of the (1,2)-Steiner Tree Problem.

1 Introduction

Given a graph $G = (V, E)$, a cost function $c : E \rightarrow \mathbb{R}_+$ and a subset $S \subseteq V$ of the vertices of $G$, a Steiner Tree $T$ for $S$ in $G$ is a subtree of $G$ that all vertices from $S$. The elements of $S$ are called terminals. The Steiner Tree Problem (STP) is: Given $G$, $c$ and $S$ as above, find a Steiner Tree $T$ for $S$ in $G$ of minimum cost $c(T) = \sum_{e \in E(T)} c(e)$.

The Steiner Tree Problem is equivalent to the Metric Steiner Tree Problem: Given a finite metric space $(V, d)$ and a set of terminals $S \subseteq V$, find a tree $T = (V_T, E_T)$ with $S \subseteq V_T \subseteq V$ such as to minimize $d(T) := \sum_{(u,v) \in E(T)} d(u, v)$.

The Steiner Tree Problem is one of the fundamental and most important network design problems with applications ranging from transportation networks, energy supply and broadcast problems to VLSI design and Internet Routing.

In his seminal paper [Kar72], Richard Karp has shown NP-Hardness of the Steiner Tree Problem. Hence there is little hope for polynomial time algorithms solving the STP to optimality. One is therefore interested in efficient approximation algorithms.

*Dept. of Computer Science, University of Bonn. Email: hauptma@cs.uni-bonn.de
A simple Minimum Spanning Tree heuristic already gives a 2-approximation [TM80]. Within the last two decades, substantial improvements upon this were made, yielding better and better approximation ratios for the Steiner Tree Problem. We only mention some of the most important results. The first polynomial-time algorithm with approximation ratio strictly less than 2 was proposed by Zelikovsky [Zel93], achieving an $11/6$-approximation based on a greedy approach. Berman and Ramaiyer [BR94] achieve an approximation of $\frac{11}{6} \approx 1.78$. Karpinski and Zelikovsky [KZ97a] combine these methods with a sophisticated preprocessing approach, obtaining an approximation ratio of $\approx 1.644$. By iterating the algorithm from [KZ97a], Hougardy and Prömel [HP99] obtain an approximation ratio of 1.598. The currently best known Steiner Tree approximation algorithm is due to Robins and Zelikovsky [RZ00] and achieves a ratio of $1 + \frac{\ln(2)}{2} \approx 1.55$.

Several special cases of the Steiner Tree Problem have been considered in the literature. We only mention three of them. Arora [Aro98] gives a polynomial time approximation scheme (PTAS) for geometric instances in fixed dimension. Karpinski and Zelikovsky [KZ97b] give a PTAS for the $\epsilon$-Dense Steiner Tree Problem. Their method was subsequently used by the author [Hau04, Hau07] to obtain polynomial time approximation schemes for dense versions of the Price Collecting Steiner Tree Problem, $k$-Steiner Tree and Group Steiner Tree Problem.

The $(1,2)$-Steiner Tree Problem is defined as the Steiner Tree Problem restricted to the case of $(1,2)$-metrics, i.e. metrics of the form $d: V \rightarrow \{0, 1, 2\}$. This problem was first considered by Bern and Plassmann [BP89] who proved MAX SNP-hardness and constructed a $4/3$-approximation algorithm. The currently best-known polynomial-time approximation algorithm for the $(1,2)$-Steiner Tree Problem is due to Robins and Zelikovsky [RZ00] and achieves an approximation ratio of $\approx 1.28$.

Approximation Hardness. The NP-hardness result for the Steiner Tree Problem by Karp [Kar72] was based on a polynomial-time reduction from the Satisfiability Problem (SAT) (cf. section 2). This reduction also works for the $(1,2)$-Steiner Tree Problem (Corollary 2.1). It turns out that Karp's construction also gives an L-reduction from Max-B-Occ-Max-3SAT to the Steiner Tree Problem and hence implies APX-hardness of the STP. Here Max-B-Occ-Max-3SAT denotes the restricted version of Max-3SAT where each variable occurs at most $B$ times in the formula.

Bern and Plassmann [BP89] proved that for each $B$, there is an L-reduction from Vertex Cover Problem in graphs of maximum degree $B$ to the $(1,2)$-Steiner Tree Problem.

First explicit lower bounds for the approximability of the STP could be
obtained by combining these results with hardness results of Berman and
Karpinski [BK98a, BK98b, BK03] (cf. sections 2 and 3).

In 2001 Thimm [Thi01] announced a hardness result for the Steiner Tree
Problem which was based on starting directly from Hastad’s result [Has97],
yielding a lower bound of 1.00617. However, this result was subsequently
corrected and improved by Chlebik and Chlebikova [CC02]. They proved
that the Steiner Tree Problem is NP-hard to approximate within 1.01063.
The techniques of Chlebik and Chlebikova are based on gadget constructions
involving edge-weighted graphs. Currently this approach does not seem to
be applicable to the (1, 2)-Steiner Tree Problem.

In this paper we concentrate on the (1, 2)-Steiner Tree Problem which
is defined as the Steiner Tree Problem restricted to metric spaces with the
only distances being 0, 1 or 2, and which will be denoted as (1, 2)-STP. To
our knowledge, the currently best known lower bound for approximability
of the (1, 2)-STP is 1.0026, by combining the L-reduction from the Bounded
Degree Vertex Cover Problem due to Bern and Plassman [BP89] with the
hardness results from Berman and Karpinski [BK98b].

The rest of the paper is organized as follows: In section 2 we describe
the NP-hardness proof (in the exact setting) given by Karp [Kar72] and the
extension of this result to an L-reduction from the Max-B-Occ-Max-3SAT
to the (1, 2)-Steiner Tree Problem. Combined with the hardness results of
Berman and Karpinski [BK98b, BK03] this gives a lower bound of \( \approx 1.0014 \)
for the approximability of the (1, 2)-Steiner Tree Problem (Corollary 2.2).

In section 3 we describe the Bern-Plassmann result [BP89] which gives
an L-reduction from the Bounded Degree Vertex Cover Problem to the (1, 2)-
STP. Combined with hardness results for the Bounded Degree vertex Cover
Problem in [BK03], this gives a lower bound of \( \approx 1.0026 \).

2 First Hardness Result:
Reduction from Max-SAT

The decision version of the Steiner Tree Problem was already proved being
NP-complete by Karp [Kar72], using the following reduction from the Satis-
fiability Problem (SAT): Given a boolean formula in conjunctive normal
form \( \varphi = C_1 \land \ldots \land C_m \) with clauses \( C_1, \ldots, C_m \) and variables \( x_1, \ldots, x_m \), one
constructs a graph \( G = G_\varphi = (V, E) \) and a terminal set \( S \subseteq V \) as follows:
For each clause \( C_j \), \( S \) contains a terminal \( t_j \), for each variable \( x_i \) there exists
a three-vertex path \( P_i = v_{i,0} - v_i - v_{i,1} \) with \( v_i \) being a terminal as well,
furthermore there is an additional vertex \( u \). We connect all vertices \( v_{i,0}, v_{i,1} \),

3
to \( u \) by one edge, furthermore we take all edges \( \{v_i, \alpha, t_j\} \) such that literal \( x_i^{\alpha} (\alpha \in \{0, 1\}) \) occurs in clause \( C_j \). The construction is shown in figure 1, where terminals are drawn black and other vertices white.

**Figure 1:**

If \( \varphi \) is satisfiable, then there exists a Steiner Tree for \( S \) in \( G \) of length \( 2m \), otherwise every Steiner Tree has length at least \( 2m + 2 \). This establishes NP-hardness of the STP. Furthermore, essentially the same construction works for the \( (1, 2) \)-STP: In the above construction, let all vertices being connected by an edge have length 1 and all other distances be 2. Hence we obtain:

**Corollary 2.1.** The Steiner Tree Problem is NP-hard even for \((1, 2)\)-metric instances.

This construction can also be used to obtain an L-reduction from the problem Max-3-OCC-MAX-3SAT to the Steiner Tree problem, which then gives an explicit approximation lower bound.

**Theorem 2.1.** The construction of Karp as described above yields an L-reduction

\[
\text{Max-3-OCC-MAX-3SAT} \leq L \text{ Steiner Tree Problem}
\]

with parameters \( \alpha = 15 \) and \( \beta = 1 \).

**Proof:** Consider an instance \( \varphi = C_1 \land \ldots \land C_m \) of Max-3-OCC-MAX-3SAT with clauses \( C_1, \ldots, C_m \) and variables \( x_1, \ldots, x_n \). Given an assignment \( \beta: \{x_1, \ldots, x_n\} \rightarrow \{0, 1\} \), let the tree \( T_\beta \) in \( G_\varphi \) consist of all edges \( \{u, v_i, \beta(x_i)\}, \{v_i, \beta(x_i), v_i\} \), for each satisfied clause \( C_j \) an edge connecting it to the vertex corresponding to a satisfying true literal and for each unsatisfied clause a path of length 2 connecting it to \( u \). Then \( T_\beta \) is a Steiner Tree.
for the terminal set $S_{\varphi} = \{t_1, \ldots, t_m, v_1, \ldots, v_n\}$ in the graph $G_{\varphi}$ of cost
\[
c(T_{ij}) = 2n + 2m - |\{ j \in \{1, \ldots, m\} : \beta(C_j) = 1\}|
\]
On the other hand, let $T$ be an arbitrary Steiner Tree for the terminal set $S_{\varphi}$ in $G$. We call such tree a normal form tree if it satisfies the following conditions:

1. Each clause vertex $t_j$ has degree 1.
2. Each variable vertex $v_i$ is connected to exactly one of the vertices $v_{i,0}, v_{i,1}$.
3. If $\{v_{i,j,0}, v_{i,j,1}\}$ is an edge of $T$, then $\{v_{i,j}, u\}$ is an edge of $T$ as well.

In polynomial time each tree $T$ can be transformed into a normal form tree $T'$ such that $c(T') \leq c(T)$. Hence we may assume $T$ to be in normal form. We will now construct an assignment $\beta_T$ as follows:

\[
\beta_T(v_i) = \begin{cases} 
0 & \text{if only } \{v_i, v_{i,0}\} \text{ is an edge of } T \\
1 & \text{if only } \{v_i, v_{i,1}\} \text{ is an edge of } T \\
\alpha & \text{if both } \{v_n, v_{i,0}\}, \{v_i, v_{i,1}\} \text{ are edges of } T
\end{cases}
\]

where $\alpha \in \{0, 1\}$ is such that $x_i^\alpha$ occurs more often than $x_i^{1-\alpha}$ in $\varphi$. If we start from MAX-SAT, we have $\frac{n}{3} \leq \text{opt} \leq m$. For Max-3-OCC-MAX-3SAT we obtain $n/3 \leq m \leq n$ and therefore $\text{smt} = 2n + 2m - \text{opt}(f) \leq 8m - \text{opt}(f) \leq 16\text{opt} - \text{opt} = 15\text{opt}$, which establishes $\alpha = 15$. For a clause $C$, let $\beta_T(C)$ denote the truth value assigned to $C$ by $\beta_T$. We observe that

\[
\text{opt}(f) - |\{C \text{ clause} : \beta_T(C) = 1\}| \leq m - (2(n + m) - c(T)) = c(T) - (2n + m) \leq c(T) - \text{opt}(G_f, S_f)
\]

which directly implies $\beta = 1$, thus completing the proof.

This result will now be combined with existing lower bound results for the approximability of Max-3OCC-MAX-3SAT in order to derive explicit lower bounds for the Steiner Tree Problem.

We consider first the following hardness result for the problem Max-E3-Lin-2 obtained by Hastad [Has97]. Max-E3-Lin-2 is the following problem: given a system of linear equations modulo 2 with exactly 3 variables per equation, find an assignment such as to maximize the number of satisfied equations.

5
Theorem 2.2. (Hastad 1997 [Has97]) For every $\epsilon \in (0, 1/4)$ and sufficiently large integer $k \geq k(\epsilon)$ the following problem is NP-hard: Given an instance of Max-E3-Lin-2 consisting of $n$ equations with exactly $2k$ occurrences of each variable, decide if at least $(1 - \epsilon)n$ or at most $(1/2 + \epsilon)n$ equations are satisfied by the optimum assignment. Equivalently: For E3-Lin-2 with $2n$ equations and $n$ variables it is NP-hard to decide if $\text{MaxLin}(f) \leq (1 + \epsilon)n$ or $\text{MaxLin}(f) \geq (2 - \epsilon)n$.

Let BOCC-Ek-Lin-2 denote the special case of Max-E3-Lin-2 where each equation contains precisely $k$ variables and each variable occurs at most $B$ times. Berman and Karpinski obtained explicit hardness results for the problem $3OCC-E2-Lin-2$:

Theorem 2.3. (Berman, Karpinski 1998 [BK98b]) It is NP-hard for instances of $3OCC-E2-Lin-2$ with $336n$ equations to decide whether $\text{opt} \leq (331 + \epsilon)n$ or $\text{opt} \geq (332 + \epsilon)n$.

Berman and Karpinski further improved these bounds, obtaining the following result.

Theorem 2.4. (Berman, Karpinski 2003 [BK03]) It is NP-hard to approximate E3-OCC-E2-Lin-2 within $\frac{112}{111} - \epsilon$.

In order to combine these results such as to achieve lower bounds for the STP, we use a simple reduction from $3OCC-E2-Lin-2$ to $2OCC-E2-SAT$ by replacing each linear equation $x + y = 0/1$ by a set of 4 clauses. Starting from the first result of Berman and Karpinski [BK98b], it is NP-hard for $3OCC-E2-Lin-2$ instances with $336n$ equations and $n$ variables to decide whether $\text{opt} \leq (331 + \epsilon)n$ or $\text{opt} \geq (332 + \epsilon)n$. This yields $672n$ clauses and $n$ variables in the associated $2OCC-E2-SAT$ instance, and using the above reduction we obtain an instance of the Steiner Tree Problem with $4n + 672n - 2 = 1348n$ edges, $(672 + 3)\cdot n + 1 = 675\cdot n + 1$ nodes and $673\cdot n + 1$ terminals, where the cost of an optimum Steiner Tree is $\text{smt} = 2n + (2^1 + 1)\cdot m = \text{MaxLin}(f) = 2n + 3 \cdot 336 \cdot n = \text{MaxLin}(f) = 1010 \cdot n - \text{MaxLin}(f)$. Here $f$ denotes the $3OCC-E2-Lin-2$ instance we start from.

Therefore it is NP-hard for instances of the Steiner Tree Problem with $1348 \cdot n$ edges, $675 \cdot n + 1$ nodes and $673 \cdot n + 1$ terminals to decide whether $\text{smt} \leq 1010 \cdot n - (332 - \epsilon) \cdot n = (678 + \epsilon) \cdot n$ or $\text{smt} \geq 1010n - (331 + \epsilon)n = (679 - \epsilon) \cdot n$. This gives the following hardness result for the Steiner Tree Problem.

Corollary 2.2. It is NP-hard to approximate the Steiner Tree Problem within $\text{A.R.} \frac{679}{678} - \epsilon \approx 1.0014 - \epsilon$. The same hardness result holds for the $(1,2)$-Steiner Tree Problem.

6
3 Second Hardness Result: Reduction from Bounded-Degree Vertex Cover

In this section we consider hardness results for the Steiner Tree Problem that are based on reductions from the Bounded Degree Vertex Cover Problem. Given a graph $G = (V, E)$, a vertex cover of $G$ is a subset $C \subseteq V$ of vertices such that each edge $e$ has at least one end-vertex in $C$ (in which case $C$ is said to cover $e$). For a given nonnegative integer $B$, the $B$-VC is the Vertex Cover Problem restricted to graphs of maximum degree at most $B$.

In 1989 Bern and Plassmann [BP89] constructed an $L$-reduction from the Bounded Degree Vertex Cover Problem to the $(1, 2)$-Steiner Tree Problem. We will combine this reduction with explicit lower bounds for the approximability of the $B$-VC due to Berman and Karpinski [BK98a, BK98b].

Let us first state the result of Bern and Plassmann.

**Theorem 3.1.** [BP89] For each nonnegative integer $B$, there is an $L$-Reduction from the $B$-VC to the $(1, 2)$-Steiner Tree Problem with parameters $\alpha = B/2, \beta = 1$.

**Proof:** Given a graph $G = (V, E)$ with vertex degree bounded by $B$, an instance of the Steiner Tree Problem is constructed consisting of a graph $H = (V_H, E_H)$ and terminal set $S \subseteq V_H$ as follows: For each edge $e$ of $G$ we introduce a vertex $v_e$, furthermore for each vertex $u$ of $G$ a vertex $v_u$, hence $V_H = \{v_e | e \in E\} \cup \{v_u | u \in V\}$. For each $u \in V$ we let $\{v_u, v_e\}$ be an edge of $H$, furthermore we add all the edges $\{v_u, v_w\}$ for $u, w \in V$. The terminal set is defined as $S := \{v_e | e \in E\}$.

Let us analyze the construction: First, let $U \subseteq V$ be a vertex cover for the graph $G$. We can construct a Steiner Tree $T_U$ by taking $\{v_u | u \in U\}$ as the set of Steiner Points, connecting each edge vertex $v_e$ to a node $v_u$ such that $u \in U$ covers edge $e$ and adding the edges of a spanning tree of length $|U| - 1$ for the set $\{v_u | u \in U\}$. The cost of the tree $T_U$ equals the number of its edges, which is $|E| + |U| - 1$. Hence we obtain the following lower bound for the cost of an optimum Steiner tree: $\text{smt}(G', E) = |E| + |VC| - 1$. Furthermore, if $T$ is an arbitrary Steiner Tree with $|E(T)| - 1$ edges, we may assume that each edge of $G$ has degree 1 in $T$ (if the degree is 2, the neighbors are $u, v$ and $e = \{u, v\}$, so we can replace one of the edges $\{e, u\}, \{e, v\}$ by $\{u, v\}$). Then the Steiner nodes of $T$ define a cover $U_T$ in $G$ of size $|U_T| = \text{smt}(G', E) - |E| - 1$. If $G$ has maximum degree $\Delta_G = B$ then $|E| \leq B/2 |V|$, and the size of an optimum vertex cover $VC_{opt}$ can be bounded as follows: $|VC_{opt}| \leq |E| \leq (B/2) \cdot |V|$. This implies $\text{opt}(G', E) \leq$
\[(B/2)\cdot|V|+|VC_{opt}|-1 \leq (B/2)\cdot|V|+(B/2)\cdot|V|-1 = B\cdot|V|-1 \leq B \cdot \text{opt}(G).\]

Furthermore, if \(T\) is a Steiner tree then \(|U_T| - \text{opt}(G)| = |c(T)| - |E| - 1 - (\text{smt} - |E| - 1) = c(T) - \text{smt}.\) Hence we obtain an L-reduction from the B-VC to the Steiner Tree Problem:

\[B-VC \leq L \text{ STP \ with parameters } \alpha = B/2, \beta = 1.\]

\[\square\]

Berman and Karpinski [BK98a, BK98b] obtained the following hardness results for bounded-degree versions of the Vertex Cover Problem.

**Theorem 3.2. (Berman, Karpinski 1998 [BK98a, BK98b])**

The Vertex Cover Problem is NP-hard to approximate within \(\frac{26}{25} - \epsilon\) in graphs \(G\) with maximum degree \(\Delta_G = 4\) and within \(\frac{145}{144} - \epsilon\) in graphs \(G\) with \(\Delta_G = 3\).

More precisely, they proved the following: For \(\epsilon \in (0, 1/2)\) it is NP-hard to decide whether an instance of the problem 3MIS (Maximum Independent Set in graphs with maximum degree 3) with 284 \cdot n nodes has maximum independent set of size below \((139 + \epsilon) \cdot n\) or above \((140 - \epsilon) \cdot n\). Since independent sets are complements of vertex covers, one obtains the following equivalent formulation for the Vertex Cover Problem: It is NP-hard to decide whether an instance of the problem 3VC with 284n nodes has a Minimum Vertex Cover of size below \((144 + \epsilon)n\) or above \((145 - \epsilon)n\).

For graphs with maximum degree \(\Delta_G = 4\) Berman and Karpinski obtain the following result: For \(\epsilon \in (0, 1/2)\) it is hard to decide whether an instance of 4MIS with 152n nodes has a maximum independent set of size below \((73 + \epsilon)n\) or above \((74 - \epsilon)n\). Equivalently, it is hard to decide whether the minimum VC is of size below \((78 + \epsilon)n\) or above \((79 - \epsilon)n\).

Combining these results with the Bn-Plussmann reduction, we obtain the following: The Berman-Karpinski graph for 3MIS has at most 3 \cdot 284 \cdot n / 2 = 432 \cdot n edges, therefore it is NP-hard to decide for the (1,2)-Steiner Tree Problem whether \(\text{smt} \leq 432n + (144 + \epsilon)n - 1 = (576 + \epsilon)n - 1\) or \(\text{smt} \geq 432n + (145 - \epsilon)n - 1 = (577 - \epsilon)n - 1\). Therefore the (1,2)-Steiner Tree Problem is NP-hard to approximate with approximation ratio \(577/576 - \epsilon \approx 1.0013 - \epsilon\).

The graph for 4MIS has at most 4 \cdot 152n / 2 = 304n edges, therefore it is NP-hard to decide whether \(\text{smt} \leq 304 \cdot n + (78 + \epsilon) \cdot n - 1 = (382 + \epsilon) \cdot n - 1\) or \(\text{smt} \geq 304 \cdot n + (79 - \epsilon) \cdot n - 1 = (383 - \epsilon) \cdot n - 1\). Therefore the (1,2)-Steiner Tree Problem is NP-hard to approximate with approximation ratio \(383/382 - \epsilon \approx 1.0026 - \epsilon\).
**Corollary 3.1.** It is NP-hard to approximate the $(1, 2)$-Steiner Tree Problem within $A.R. 1.0026 - \epsilon$. 

**References**


