Journal of Global Optimization **18:** 17–33, 2000.

Approximations for Steiner Trees with Minimum Number of Steiner Points

DONGHUI CHEN[1,*], DING-ZHU DU[1,2,*,‡], XIAO-DONG HU[2,3,‡], GUO-HUI LIN[4,§], LUSHENG WANG[2] and GUOLIANG XUE[4,§]
[1]*Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA (E-mail: dzd@cs.umn.edu)*
[2]*Department of Computer Science, City University of Hong Kong, Kowloon Tong, Hong Kong, China*
[3]*Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing 100080, China*
[4]*Department of Computer Science, The University of Vermont, Burlington, VT 05405, USA*

(Accepted in original form 19 January 2000)

Abstract. Given n terminals in the Euclidean plane and a positive constant, find a Steiner tree interconnecting all terminals with the minimum number of Steiner points such that the Euclidean length of each edge is no more than the given positive constant. This problem is NP-hard with applications in VLSI design, WDM optical networks and wireless communications. In this paper, we show that (a) the Steiner ratio is $1/4$, that is, the minimum spanning tree yields a polynomial-time approximation with performance ratio exactly 4, (b) there exists a polynomial-time approximation with performance ratio 3, and (c) there exists a polynomial-time approximation scheme under certain conditions.

Key words: Steiner trees; Approximation algorithms; VLSI design; WDM optical networks

1. Introduction

Given a set of n terminals $X = \{p_1, p_2, \ldots, p_n\}$ in the Euclidean plane $\mathscr{R}^2$, and a positive constant R, the *Steiner tree problem with minimum number of Steiner points*, denoted by *STP-MSP* for short, asks for a tree T spanning a superset of X such that each edge in the tree has a length no more than R and the number $C(T)$ of points other than those in X, called *Steiner points* [3, 6, 8, 9], is minimized.

In the classical Euclidean Steiner tree problem (which asks for a tree spanning a superset of X such that the total length of the tree, that is the sum of lengths of edges in the tree, is minimized), a Steiner point always has a degree of 3. In the STP-MSP problem, however, degree-2 Steiner points are possible. For example, when $n = 2$ and $|p_1 p_2|$, the Euclidean distance between p_1 and p_2, is larger than R, then the

* Supported in part by the National Science Foundation under grant CCR-9530306.
‡ Supported in part by National Natural Science Foundation of China under grant 19331050 and 973 fundamental research project of China.
§ Supported in part by the Army Research Office grant DAAH04-96-10233 and by the National Science Foundation grants ASC-9409285 and OSR-9350540.

optimal tree is a path containing $\lceil |p_1 p_2|/R \rceil - 1$ Steiner points, each of which has a degree of 2.

The STP-MSP problem has an important application in wavelength-division multiplexing (WDM) optical network design [11, 14]. Suppose we need to connect n sites located at $p_1, p_2, \ldots, p_n$ with WDM optical network. Due to the limit in transmission power, signals can only travel a limited distance (say R) for guaranteed correct transmission. If some of the inter-site distances are greater than R, we need to provide some amplifiers or receivers/transmitters at some locations in order to break it into shorter pieces. The STP-MSP problem also finds applications in VLSI design [2, 7, 15], and the evolutionary/phylogenetic tree constructions in computational biology [9].

Recently, Lin and Xue [12] showed that the STP-MSP problem is NP-hard. They also showed that the approximation obtained from the minimum spanning tree by simply breaking each edge into small pieces within the upper bound has a worst-case performance ratio at most 5. In this paper, we show that this approximation has a performance ratio exactly 4. We also present a new polynomial-time approximation with a performance ratio at most 3 and a polynomial-time approximation scheme under certain conditions.

2. Preliminary

Any shortest optimal solution T for the problem STP-MSP must have the following properties.

(a1) No two edges cross each other.

(a2) Two edges meeting at a vertex form an angle of at least 60°.

(a3) If two edges form an angle of exactly 60°, then they have the same length.

To see (a1), consider two edges ac and bd in T. By contradiction, suppose ac and bd cross at e. Note that quadrangle $abcd$ must have an inner angle of at least 90°. Without loss of generality, assume $\angle abc \geqslant 90°$. Then $\angle bca < 90°$ and $\angle cab < 90°$. Hence $|ab| < |ac|$ and $|bc| < |ac|$, where $|ab|$ denotes the length of edge ab. When edge ac is removed from T, T would be broken into two parts containing vertices a and c, respectively. One of the parts, say the one containing a, contains vertex b. Adding edge bc results in a shorter tree still optimal for STP-MSP. This contradicts the length-minimality of T. Therefore, (a1) holds.

To see (a2), consider two edges ab and bc in T. By contradiction, suppose $\angle abc < 60°$. Then either $\angle cab > 60°$ or $\angle bca > 60°$ and hence either $|bc| > |ac|$ or $|ab| > |ac|$. Using ac to replace either bc or ab would reduce the total length of the tree preserving the vertex set, contradicting the length-minimality of T among optimal solutions for STP-MSP. Therefore, (a2) holds. (a3) can be proved by a similar argument.

The following lemma follows from (a2) and (a3).

LEMMA 1. *There exists a shortest optimal Steiner tree T^* for STP-MSP such that every vertex in T^* has degree at most five.*

Proof. It follows immediately from (a2) that every vertex in a shortest optimal tree T for STP-MSP has degree at most six. Consider a vertex u with degree six in T. By (a2), every angle at u equals 60°. By (a3), all edges incident to u have the equal length.

Next, consider any vertex v with degree d in T. We claim that if v is adjacent to k vertices with degree six, then $d \leqslant 6 - 2k$. In fact, suppose u is adjacent to v with degree six. Then u has two degrees uw and ux such that $\angle wuv = \angle vux = 60°$ and $|uv| = |uw| = |ux|$. Thus, $|vw| = |uw|$ and $|vx| = |ux|$. Replacing uw and ux by vw and vx results in still a shortest optimal tree for STP-MSP. But, v gets two more edges. For all vertices with degree six and adjacent to v, perform the same operation. We will obtain a shortest optimal tree for STP-MSP such that v has degree $d + 2k$. Hence, $d + 2k \leqslant 6$.

Now, for each vertex u with degree six, we move only one edge from u to its adjacent vertex. Then every vertex will have degree at most five and the resulting tree is still a shortest optimal tree for STP-MSP. □

Usually, a *spanning tree* is a tree interconnecting the given terminals with edges between given terminals. The shortest spanning tree is called the *minimum spanning tree*. Spanning trees may not be feasible solutions for the problem STP-MSP since some edges may be too long. To make it feasible, we add $\lceil |ab|/R \rceil$ Steiner points to break each edge ab into small pieces of lengths at most R. The resulting tree will be called a *steinerized* spanning tree. The following is an interesting fact.

LEMMA 2. *Every steinerized minimum spanning tree has the minimum number of Steiner points among steinerized spanning trees.*

Proof. Every minimum spanning tree can be obtained from a spanning tree by a sequence of operations that each replaces an edge by another shorter edge. Since the shorter edge needs Steiner points no more than the longer edge needs when we steinerize them. Therefore, the lemma holds. □

It follows easily from the above two lemmas that the steinerized minimum spanning tree is an approximation with performance ratio 5 (see [12]).

3. Steinerized minimum spanning tree

We show the following tight result in this section.

THEOREM 1. *The steinerized minimum spanning tree is a polynomial-time approximation with performance ratio exactly 4.*

The lower bound can be shown by presenting an example as follows. Consider five vertices $v_1, v_2, \ldots, v_5$ of a regular pentagon with each edge of length $1 + \varepsilon$ where ε is a small positive real number such that the distance from the center to

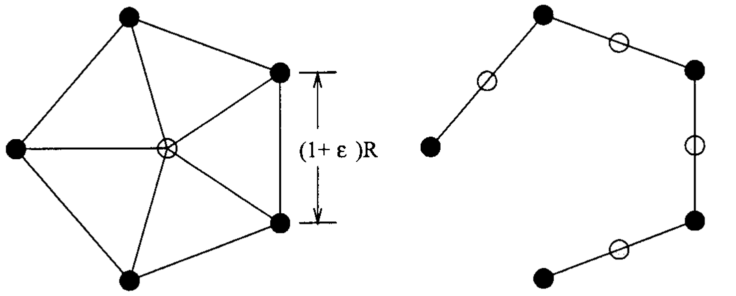

Figure 1.

each vertex is within R (Figure 1). The steinerized minimum spanning tree on $v_1, v_2, \ldots, v_5$ contains four Steiner points. However, every optimal tree for STP-MSP on $v_1, v_2, \ldots, v_5$ contains only one Steiner point. Therefore, the performance ratio of steinerized minimum spanning tree is at least four.

To show the tight upper bound, we need first to study a property of convex path in any shortest optimal tree T for STP-MSP. A path $q_1 q_2 \ldots q_m$ in T is called a *convex* path if for every $i = 1, 2, \ldots, m-3$, $q_i q_{i+2}$ intersects $q_{i+1} q_{i+3}$.

An angle of degree more than 120° will play an important role in the proof of Theorem 1. For simplicity, we call such angles *big* angles.

LEMMA 3. *Let* $q_1 q_1 \cdots q_m$ *be a convex path and* $m \geqslant 2$. *Suppose there are* t *big angles among the* $m-2$ *angles* $\angle q_1 q_2 q_3$, $\angle q_2 q_3 q_4, \ldots, \angle q_{m-2} q_{m-1} q_m$. *Then* $|q_1 q_m| \leqslant (t+2)R$.

Proof. We prove it by induction on m. For $m \leqslant 3$, it is trivial that $|q_1 q_3| \leqslant |q_1 q_2| + |q_2 q_3| \leqslant 2R \leqslant (t+2)R$. Now, suppose $m \geqslant 4$. Consider the convex hull H of points $q_1, q_2, \ldots, q_m$. If at least one of q_1 and q_2 does not lie on the boundary of H, then by the induction hypothesis, any distance between two vertices of the convex hull H is at most $(t+2)R$ and hence any two points lying in H have distance at most $(t+2)R$. Therefore, $|q_1 q_m| \leqslant (t+2)R$.

Next, we may assume that both q_1 and q_m lie on the boundary of H. It follows immediately that whole path $q_1 q_1 \cdots q_m$ lies on the boundary of H (Figure 2(a)). If $\angle q_1 q_m q_{m-1} \geqslant 90°$, then $|q_1 q_m| \leqslant |q_1 q_{m-1}|$ and by the induction hypothesis $|q_1 q_{m-1}| \leqslant (t+2)R$. Hence, $|q_1 q_m|(t+2)R$. Similarly, if $\angle q_2 q_1 q_m \geqslant 90°$, then $|q_1 q_{m-1}| \leqslant (t+2)R$. Therefore, we may assume $\angle q_1 q_m q_{m-1} < 90°$ and $\angle q_2 q_1 q_m < 90°$. It follows that $(m-2) \cdot 180° \leqslant 2 \cdot 90° + (m-t-2) \cdot 120° + t \cdot 180°$. Hence, $m - t - 2 < 3$. This means that the path $q_1 q_1 \cdots q_m$ has at most two angles of degrees not more than 120°.

If $\angle q_{m-2} q_{m-1} q_m$ is a big angle, then by the induction hypothesis, $|q_1 q_{m-1}| \leqslant ((t-1)+2)R$. Therefore, $|q_1 q_m| \leqslant |q_1 q_{m-1}| + |q_{m-1} q_m| \leqslant (t+2)R$. Similarly, if $\angle q_1 q_2 q_3$ is a big angle, then $|q_1 q_m| \leqslant (t+2)R$. Therefore, we may assume $\angle q_{m-2} q_{m-1} q_m \leqslant 120°$ and $\angle q_1 q_2 q_3 \leqslant 120°$. They are the only two angles not big

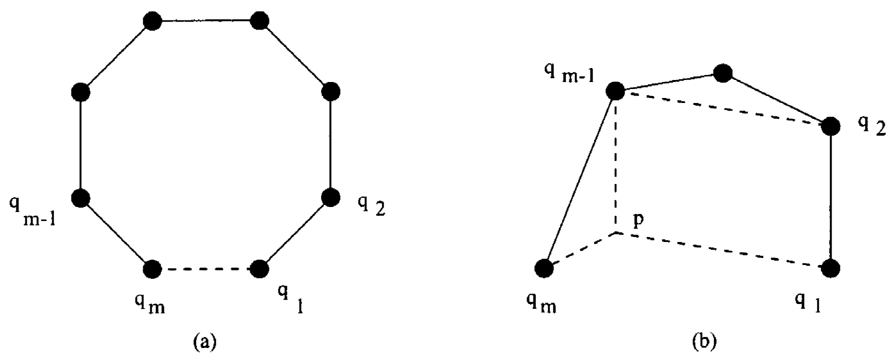

Figure 2.

on the path $q_1 q_1 \cdots q_m$. Now, draw a parallelogram $q_1 q_2 q_{m-1} p$ as shown in Figure 2(b). Since $\angle q_1 q_2 q_{m-1} \leqslant q_1 q_2 q_3 \leqslant 120°$, we have $\angle q_2 q_{m-1} p \geqslant 60°$. Moreover, $\angle q_2 q_{m-1} q_m \leqslant \angle q_{m-2} q_{m-1} q_m \leqslant 120°$. Thus, $\angle pq_{m-1} q_m \leqslant 60°$. It follows that

$$|pq_m| \leqslant \max(|pq_{m-1}|, |q_{m-1} q_m|) = \max(|q_1 q_2|, |q_{m-1} q_m|) \leqslant R .$$

Therefore,

$$|q_1 q_m| \leqslant |q_1 p| + |pq_m| = |q_2 q_{m-1}| + |pq_m| \leqslant (t+1)R + R = (t+2)R . \qquad \square$$

LEMMA 4. *In a shortest optimal tree T for STP-MSP, there are at most two big angles at a vertex with degree three, there is at most one big angle at a vertex with degree four, and there is no big angle with degree five*

Proof. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_d$ are all angles at a vertex with degree d and k (>0) of them are big angles. Since each angle is of at least 60°, we have $360° = \alpha_1 + \alpha_2 + \cdots + \alpha_d > (d-k) \cdot 60° + k \cdot 120°$. Thus, $6 > (d-k) + 2k = d + k$, i.e., $5 \geqslant d + k$. The lemma follows immediately from this inequality. $\square$

Note that every leaf in a Steiner tree is a terminal. A Steiner tree is *full* if every terminal is a leaf. If a Steiner tree is not full, then we can always find a terminal with degree more than one which enable us to break the tree at this terminal. In this way, every Steiner tree can be broken into several small full Steiner trees. Those small full Steiner trees are called *full components* of a Steiner tree.

LEMMA 5. *Consider a shortest optimal tree T for STP-MSP. Suppose T is a full Steiner tree. Let s_i denote the number of Steiner points with degree i in T. Then*

$$3s_5 + 2s_4 + s_3 = n - 2$$

where n is the number of terminals.

Proof. Since T has totally $s_5 + s_4 + s_3 + s_2 + n - 1$ edges, we have $5s_5 + 4s_4 + 3s_3 + 2s_2 + n = 2(s_5 + s_4 + s_3 + s_2 + n - 1)$. Hence, $3s_5 + 2s_4 + s_3 = n - 2$. $\square$

Consider a shortest optimal tree T for STP-MSP. Suppose T is a full Steiner tree on n terminals. The following fact is easily seen.

(b1) T has exactly n convex paths; each connects two terminals,

(b2) each terminal appears in exactly two convex paths in T, and

(b3) each angle at a Steiner point appears in those n convex paths exactly once.

Now, we are ready to show Theorem 1.

Proof of Theorem 1. By Lemma 1, there exists a shortest optimal tree T^* for STP-MSP in which every vertex has degree at most five.

First, we assume that T^* is a full Steiner tree. Let s_i denote the number of Steiner points with degree i in T^*. By Lemma 5, $n = 3s_5 + 2s_4 + s_3 + 2$. Consider a spanning tree T_S consisting of $n-1$ edges each connecting two terminals at endpoints of a convex path in T^* (Figure 3). By Lemma 3, each edge ab in T_S has length upper-bounded by $(t+2)R$ where t is the number of big angles on the convex path connecting a and b. Hence, we need at most $(t+1)$ Steiner points to steinerize edge ab. By Lemma 4, the spanning tree T_S can be steinerized by at most $s_4 + 2s_3 + 2s_2 + n - 1$ Steiner points. By Lemma 2, any steinerized minimum spanning tree contains at most $s_4 + 2s_3 + 2s_2 + n - 1$ Steiner points. Clearly,

$$\begin{aligned} s_4 + 2s_3 + 2s_2 + n - 1 &= 3s_5 + 3s_4 + 3s_3 + 2s_2 + 1 \\ &\leqslant 3(s_5 + s_4 + s_3 + s_2) + 1\,. \end{aligned}$$

If $s_5 + s_4 + s_3 + s_2 > 0$, then $s_4 + 2s_3 + 2s_2 + n - 1 \leqslant 4(s_5 + s_4 + s_3 + s_2)$. If $s_5 + s_4 + s_3 + s_2 = 0$, then $T_S = T^*$. Therefore, in either case, every steinerized minimum spanning tree contains at most $4(s_5 + s_4 + s_3 + s_2)(= 4 \cdot C(T^*))$ Steiner points.

Now, suppose T^* is not a full Steiner tree. Then T^* can be decomposed into several full components $T_1, T_2, \ldots, T_k$. For each full component T_j, by the above argument, we know that the steinerized minimum spanning tree on terminals in T_j contains at most $4 \cdot C(T_j)$ Steiner points. Note that the union of steinerized minimum spanning trees each for terminals in a full component is a steinerized spanning tree

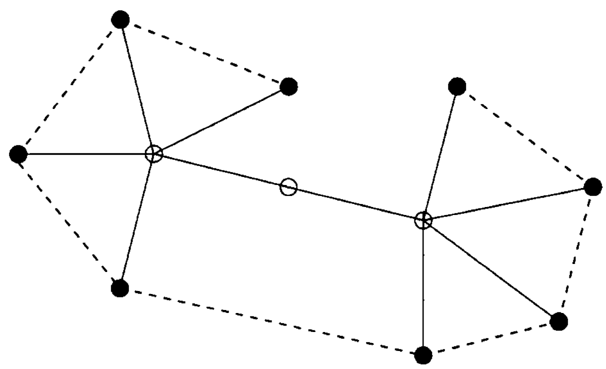

Figure 3.

for all terminals. By Lemma 2, the number of Steiner points in T^* is at most $4\sum_{j=1}^{k} C(T_j) = 4 \cdot C(T^*)$. □

4. 3-Approximation

Let T^* be a shortest optimal tree for STP-MSP with Steiner points of degrees at most five. Suppose $T_1, T_2, \ldots, T_k$ are all full components of T^*. In the proof of Theorem 1, we showed that the steinerized minimum spanning tree on terminals in T_j contains at most $3 \cdot C(T_j) + 1$ Steiner points. Now, we study when this upper bound can be improved.

LEMMA 6. *Let T^* be a shortest optimal tree for STP-MSP with property that every Steiner point has degree at most five. Let T_j be a full component of T^*. Then the following hold:*

(c1) The steinerized minimum spanning tree on terminals in T_j contains at most $3 \cdot C(T_j) + 1$ Steiner points.

(c2) If T_j contains a Steiner point with degree at most four, then the steinerized minimum spanning tree on terminals in T_j contains at most $3 \cdot C(T_j)$ Steiner points.

(c3) If the steinerized minimum spanning tree on terminals in T_j contains an edge between two terminals, then it contains at most $3 \cdot C(T_j)$ Steiner points.

Proof. (c1) and (c3) follow immediately from the proof of Theorem 1. Next, we show (c2). Let s_i be the number of Steiner points with degree i in T_j. Let n_j be the number of terminals in T_j. Note that there are exactly n_j convex paths in T_j. Choose any $n_j - 1$ of them and connect two endpoints of each path. We will obtain a spanning tree. Its steinerization is denoted by T_S. Now, assume u is the Steiner point with degree at most four. If there is a big angle at u, then we choose $n_j - 1$ convex paths not containing the big angle. If there is no big angle at u, then we can choose any $n_j - 1$ convex path. With this choice, we would have $C(T_S) \leqslant s_4 + 2s_3 + 2s_2 - 1 + (n_j - 1) \leqslant 3(s_4 + s_3 + s_2) = 3C(T_j)$. □

To design a new approximation, we need to study how to find whether three or four terminals can be connected to a common Steiner point.

Note that an angle of less than 90° is *acute* and an angle of more than 90° is *obtuse*. A triangle is *acute* if its three angles are all acute. A triangle is *obtuse* if it has one obtuse angle. A triangle is *right* if it has one right angle.

LEMMA 7. *If a triangle abc is acute or right, then the minimum disk (i.e., the disk of minimum radius) for covering the triangle abc is the one bounded by the circle circumscribing abc. If a triangle abc is obtuse or right, then the minimum disk for covering the triangle abc is the one whose diameter is the longest edge of triangle abc.*

Proof. Suppose *ab* is the longest edge of the triangle *abc*. When a disk covers *abc*, we can always arrange the boundary of the disk passing through a and b. If ab is not a diameter of the disk and c is not on its boundary, then we can shrink the disk still covering *abc*. □

LEMMA 8. *Four terminals a, b, c, d can be covered by a disk of radius R if and only if each of four triangles abc, bcd, cda, dab can be covered by a disk of radius R.*

Proof. The 'only if' part is trivial. It suffices to show the 'if' part. Without loss of generality, we may assume that *abcd* form a convex quadrilateral. In fact, if *abcd* does not form a convex quadrilateral, then one of them must lie in the triangle of other three, which can be covered by a disk of radius R.

Consider the longest edge of complete quadrilateral *abcd*. (Note: A complete quadrilateral has six edges.) If this longest edge is not a diagonal, say *ab*, then compare $\angle acb$ with $\angle adb$. Without loss of generality assume $\angle acb \leqslant \angle adb$. Then, the minimum disk covering triangle *abc* also covers point d (Figure 4(a)). Next, we may assume that the longest edge of complete quadrilateral *abcd* is a diagonal, say *ac*, and consider following cases.

Case 1. Triangles *abc* and *acd* are obtuse or right. In this case, $\angle abc$ and $\angle cda$ are obtuse or right. Therefore, the disk with a diameter *ac* covers a, b, c, d (Figure 4(b)).

Case 2. Either triangle *abc* or *acd* is acute, say triangle *abc*. If triangle *abd* is also acute, then compare $\angle acb$ with $\angle adb$. Without loss of generality assume $\angle acb \leqslant \angle adb$. Then, the minimum disk covering triangle *abc* also covers point d (Figure 4(a)). Similar argument may apply to the subcases that triangle *bcd* is acute and that $\angle bdc \geqslant 90°$ or $\angle adb \geqslant 90°$. Moreover, $\angle cbd \leqslant \angle cba < 90°$ and $\angle dba \leqslant \angle cba < 90°$. Therefore, the remainder is that $\angle bad \geqslant 90°$ and $\angle dcb \geqslant 90°$. In this subcase, the disk with a diameter *bd* covers a, b, c, d. □

The proof of Lemma 8 is constructive. We can actually use the proof to find the Steiner point to connect the four terminals when it exists.

Now, we present the following approximation algorithm.

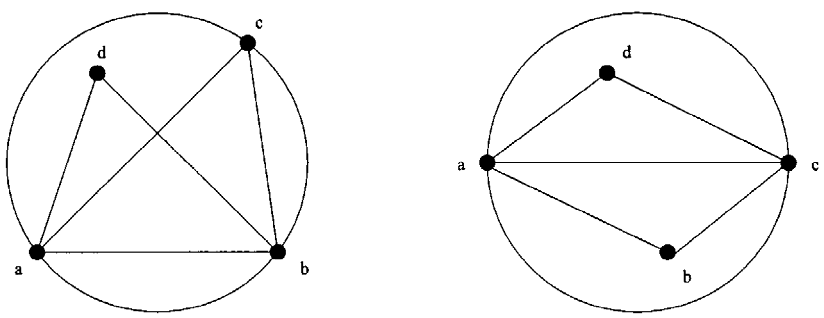

Figure 4.

ALGORITHM A. For input set X of n terminals, sort all $n(n-1)/2$ possible edges between the n terminals in length increasing order $e_1, e_2, \ldots, e_{n(n-1)/2}$. Initially, set $T_A = (X, \emptyset)$ and $i = 1$. Then do the following:

Step 1 **while** $|e_i| \leqslant R$ **do begin**
  if e_i connects two different connected components of T_A
  then put e_i into T_A;
  $i := i + 1$;
end-while
Step 2 **for** each subset of four terminals b, b, c, d respectively in four connected components of T_A **do**
  if there exists a point s within distance R from a, b, c and d
  then put the 4-star, consisting of four edges sa, sb, sc, sd, into T_A;
Step 3 **while** $i \leqslant n(n-1)/2$ **do begin**
  if e_i connects two different connected components of T_A
  then put e_i into T_A;
  $i := i + 1$;
end-while
return T_A

Clearly, this algorithm runs in $O(n^4)$ time.

THEOREM 2. *Let T^* be optimal tree for STP-MSP and T_A an approximation produced by Algorithm A. Then $C(T_A) \leqslant 3C(T^*)$.*

Proof. Denote by $T^{(i)}$ the T_A at the beginning of Step i in the algorithm A. Suppose $T^{(3)} - T^{(2)}$ contains k 4-stars. Then

$$C(T_A) \leqslant C(T_S) - 2k$$

where T_S is a steinerized minimum spanning tree on all given terminals. Let T^* be a shortest optimal tree for STP-MSP with Steiner points of degrees at most five. Suppose T^* has g full components $T_1, T_2, \ldots, T_g$. We construct a steinerized spanning tree T as follows: Initially, put $T^{(2)}$ into T. For each full component T_j ($1 \leqslant j \leqslant g$), add to T the steinerized minimum spanning tree H_j for terminals in T_j. If T has a cycle, then destroy the cycle by deleting some edges and Steiner points of H_j. An important observation is that if H_j does not contain an edge between two terminals, then a Steiner point must be deleted for destroying a cycle in $H_j \cup T^{(2)}$. From this observation and by Lemma 6, we have

$$C(T_S) \leqslant 3C(T^*) + h$$

where h is the number of full components T_j's with properties that every Steiner point in T_j has degree five and $T_j \cup T^{(2)}$ has no cycle. Hence,

$$C(T_A) \leqslant 3C(T^*) + h - 2k\,.$$

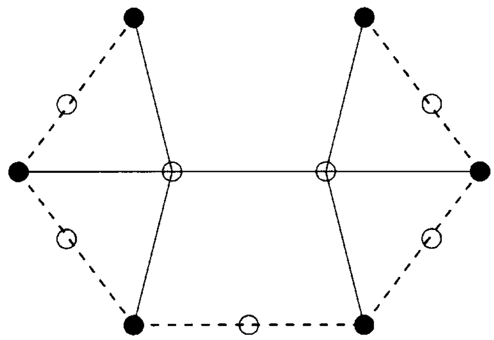

Figure 5.

It suffices to show $h \leqslant 2k$.

Suppose $T^{(2)}$ has p connected components. Then, $T^{(3)}$ has $p - 3k$ connected components $C_1, C_2, \ldots, C_{p-3k}$. Now, we construct a graph H with vertex set X and the following edges: First, we put all edges of $T^{(2)}$ into H. Then consider every full component T_j $(1 \leqslant j \leqslant h)$ with properties that every Steiner point in T_j has degree five and $T_j \cup T^{(2)}$ has no cycle. If T_j has only one Steiner point, then this Steiner point connects to five terminals which must lie in at most three C_i's. Hence, among them there are two pairs of terminals; each pair lie in the same C_i. Connect the two pairs with two edges and put the two edges into H. If T_j has at least two Steiner points, then there must exist at least two Steiner points each connecting to four terminals. We can also find two pairs of terminals among them such that each pair lies in the same C_i. Connect the two pairs with two edges and put the two edges into H. Clearly, H has at most $p - 2h$ connected components. Since every connected component of H is contained by a C_i, we have $p - 3k \leqslant p - 2h$. Therefore, $h \leqslant 3k/2$. □

What is the exact value of the performance ratio of Algorithm A? It is still open. What we know is that this value is between 2.5 and 3. The lower bound 2.5 can be shown by the instance in Figure 5.

5. Polynomial-time approximation scheme

In this section, we consider a variation of STP-MSP. The input and the constraint are the same. Instead of minimizing the number of Steiner points in the tree, we minimize the number of total points (both Steiner points and given terminals) in the tree. Obviously, the decision versions of the two problems are identical. The new version is called the *Steiner tree problem with minimum number of total points* (STP-MTP). We construct a polynomial time approximation scheme when the given set of terminals satisfies certain conditions.

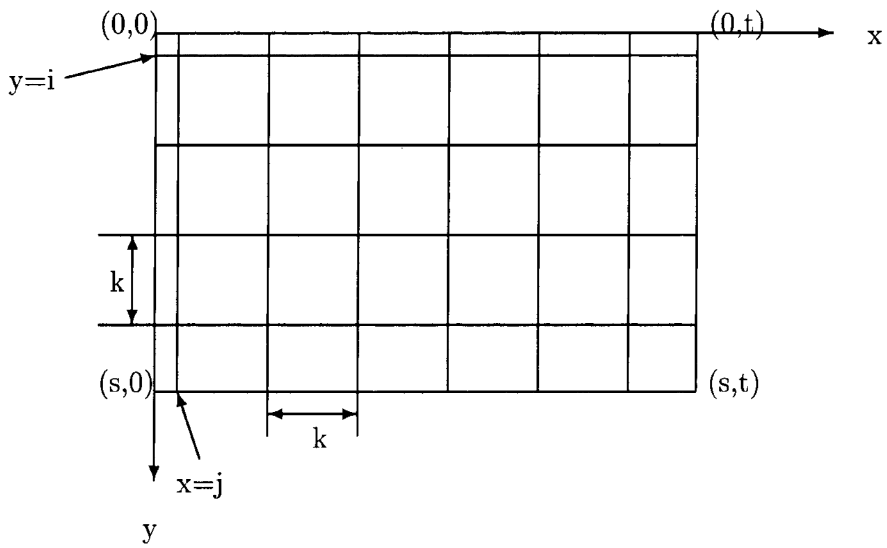

Figure 6. The rectangle with partition $P_{i,j}$ of size k.

A set X of terminals is *c-local* if in the minimum spanning tree of X the length of the longest edge is at most c times of the length of the shortest edge. Without loss of generality, we assume that the distance between any pair of terminals in X is at least 1 and $c \geqslant 1$. We are interested in the case where $R < c$.

5.1. THE BASIC IDEA

The basic idea of our algorithm is to combine the shifting technique in [4] with a local optimization method. We design a set of partitions, each of them partitions the whole area enclosing all terminals into many rectangular *cells* (mostly squares) of some *constant* size. (See Figure 6.) Each cell is further divided into *interior* and *boundary* areas as in Figure 7. Then, with respect to each partition, we organize the terminals contained in the interior area of each cell into several groups such that the

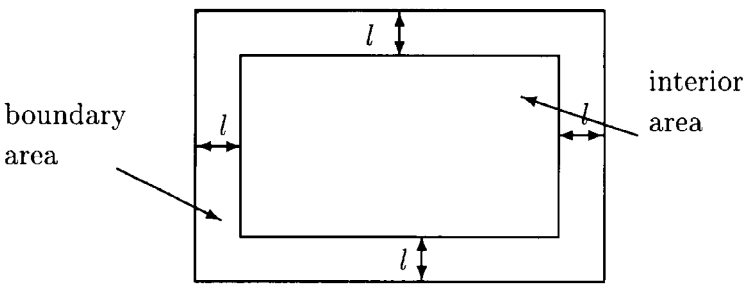

Figure 7. The interior and boundary areas. The width of the boundary areas is $l = (2 + 3 \log k)c$.

distance between any two groups is greater than c, and construct an optimal solution (a local Steiner tree) for each group. The collection of all the *local Steiner trees* in a cell form a *local Steiner forest* for the cell. After that, we connect all the local Steiner forests and the terminals contained in the boundary areas using the spanning tree approach. Finally, we select a partition which yields an optimal global solution among all the partitions.

5.2. PARTITION STRATEGY

First, we focus on the partitions. Without loss of generality, assume that the set of terminals X is contained in a rectangle *Rec* with corners $(0, 0)$ $(s, 0)$, $(0, t)$, and (s, t), as shown in Figure 6. For any positive integer k, a *partition* of size k is a grid in which adjacent horizontal/vertical lines are separated by a distance k. Clearly, there are k^2 different partitions of size k, depending on the positions of the top horizontal line and the leftmost vertical line. We use $P_{i,j}$, where $0 \leqslant i, j < k$, to denote the partition in which the top horizontal line and the leftmost vertical line are $y = i$ and $x = j$, respectively. The grid partitions the rectangle *Rec* into many cells, most of which are squares of size $k \times k$. Thus, each cell contains at most k^2 terminals in X. Each cell is divided into an *interior area* and a *boundary area*, with a boundary of *width* $l = (2 + 3 \log k)c$. (See Figure 7.)

5.3. THE APPROXIMATION SCHEME

Let X be the set of terminals in the plane, P be a partition, and $X_P \subseteq X$ be the set of terminals in the interior areas. An edge is a *crossing* edge if it is not completely contained in any interior area of a cell. A *stem* in a Steiner tree T is a path in T such that every vertex in the path is degree-2 Steiner points except that the two vertices at the ends are terminals. A stem is a *crossing* stem if at least one of the terminals is in the boundary area. Let T^P be an optimal solution of STP-MTP for X_P. T^{min} denotes an optimal solution of STP-MTP for X and $\overline{C}(T)$ denotes the total number of points in the tree T, i.e. $\overline{C}(T) = C(T) + n$ when T contains n terminals. Since X_P is a subset of X, we have

$$\overline{C}(T^P) \leqslant \overline{C}(T^{min}) . \tag{1}$$

In our algorithm, we deal with one cell at a time. Recall that the terminals in the interior area of a cell are divided into several groups and an optimal solution is constructed for each group. In order to show how to correctly group the terminals in an interior area, let us consider an optimal solution of STP-MTP T^P for X_P. We need to modify T^P into a forest F^P such that each tree in F^P is completely included in the interior areas of some cell. Note that each interior area of a cell may contain more than one tree in F^P. Define the *distance* between two trees to be the shortest distance between any pair of terminals in the two trees. We further require that the distance between any pair of trees in F^P is greater than c.

LEMMA 9. *Let P be a partition and T^P be an optimal solution of STP-MTP for X_P. T^P can be modified into a forest F^P such that each tree in F^P is completely in an interior area of a cell for P and the distance between any pair of trees in F^P is at least c. Moreover, the total cost $\overline{C}(F^P)$, which is the sum of the costs of all the trees in F^P, is at most $\overline{C}(T^P)$. Thus,*

$$\overline{C}(F^P) \leqslant \overline{C}(T^P) \leqslant \overline{C}(T^{min}) .$$

Proof. First, we eliminate the stems with length greater than c from T^P. The distance between any pair of resulting trees is greater than c since T^P is optimal. For each tree T_i in the forest obtained above, we reconstruct an optimal tree connecting the terminals in T_i. Without loss of generality, we can assume that each stem in the reconstructed trees has length at most c. (Otherwise, we can repeat the procedure and further decompose the forest.)

Now we prove that each tree in the forest obtained above is completely in an interior area of a cell. It suffices to show that there is no Steiner point in the boundary area.

Suppose there are Steiner points in the boundary area. Call a Steiner point with degree greater than 2 a *real Steiner point*. Note that the distance between two cells is $(2 + 3 \log k)c$. It is easy to see that no terminals in distinct cells are connected in the above resulting forest. Otherwise, there must be a Steiner point s which is at least $1.5 \cdot c \cdot \log k$ away from any boundary line. To reach any boundary line, s has to create at least $2^{1.5 \log k} = k^{1.5}$ real Steiner points and $k^{1.5}(\lceil c/R \rceil - 1)$ degree-2 Steiner points. Now, we remove all those $k^{1.5} \cdot \lceil c/R \rceil$ Steiner points in the boundary areas and use them to connect the disconnected subtrees with distance less than c in the corresponding 4 neighbor cells. (See Figure 8. s is in the shadowed area. At most $k \cdot \lceil 1/r \rceil$ Steiner points are required to be added in each of the eight boundary segments of the four cells.)

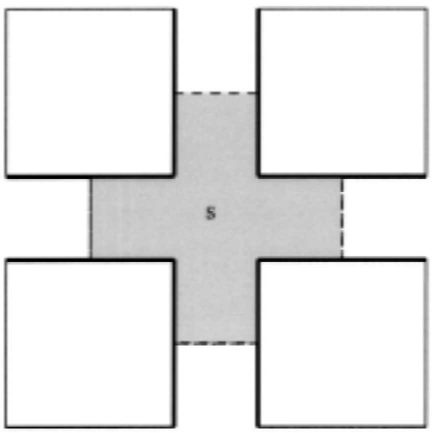

Figure 8. The eight boundary segments.

Since no two cells are connected, we can move the Steiner points in the boundary areas back to the interior areas. In this way, all Steiner points in the boundary area can be eliminated. □

It is difficult to compute the forest F^P, since T^P is unknown. Nevertheless, we can construct a forest which is similar to F^P. Consider the terminals in the interior area of some fixed cell. By Lemma 9, if the distance between two terminals is at most c, then they must belong to the same tree of F^P. Thus, we can group the terminals by forming a minimum-cost spanning tree of these terminals and then deleting the edges longer than c. Therefore, we get a set of (spanning) trees $\{S_1, \ldots, S_m\}$, consisting of degrees of length at most c. We call these trees the *c-spanning trees*. Let Y_i, $i = 1, \ldots, m$, be the set of terminals contained in the c-spanning tree S_i. Clearly, the terminals in the same group Y_i belong to the same tree of the forest F^P. The converse is not necessarily true. Namely, terminals in different groups Y_i's may also belong to the same tree of F^P. In other words, to find the best way of grouping the terminals, we have to consider all possible ways merging the groups $Y_1, \ldots, Y_m$. After each such possible merge, we obtain a local Steiner forest by constructing an optimal solution for every *new* group. We are interested in a local Steiner forest with the minimum cost among all possible merges for each cell.

Let forest $\hat{F}^P$ denote the collection of the minimum-cost local Steiner forests, one for each cell. $\hat{F}^P$ has the following properties.

LEMMA 10. *(i) Each tree in $\hat{F}^P$ is completely contained in the interior area of a cell*; *(ii) The distance between any pair of trees T_i and T_j in $\hat{F}^P$ is greater than c*; *and (iii) The total cost of the forest $\hat{F}^P$ is at most $\overline{C}(F^P)$. Thus,*

$$\overline{C}(\hat{F}^P) \leqslant \overline{C}(F^P) \leqslant \overline{C}(T^{min}) .$$

Suppose that there are m groups in a cell. Using the method in [16], we can compute a minimum-cost local Steiner forest in $O(2^m M(|Y|))$ time, where $M(Y)$ is the time to construct an optimal solution for the set of terminals Y, which is exponential in the size of Y.

5.4. AN EXACT ALGORITHM FOR STP-MTP

Let Y be a set of $|Y|$ terminals. Without loss of generality, we can assume that the terminals in Y are leaves in the tree. The number of possible topologies (the degree can be unbounded) for Y is at most $|Y|!$. Consider a fixed topology T for Y. If the number of candidate points for each internal vertex in T is at most m, then a modification of a standard dynamic programming algorithm finds an optimal solution for the fixed topology T in $O(|Y|m)$ time [5].

LEMMA 11. *The number of candidate points for each internal vertex is at most $(|Y|\sqrt{2}k/R)^{3^{|Y|-1}}$ if terminals in Y are in a square of size k by k.*

Proof. Let T^{min} be an optimal solution for the fixed topology T. Consider an internal vertex v at the bottom whose children are leaves in T. Without increasing the number of Steiner points, we can move the point assigned to v such that the distance between v and v_i ($i = 1$ and 2) is Rh_i, where h_i's are integers, and v_1 and v_2 are some children of v. Thus, the number of candidate points for v is at most $|Y|^2 \times (\sqrt{2}k/R)^2$. The height of T is at most $|Y| - 1$. For a vertex of height i, the number of candidate points is denoted as $f(i)$. Then $f(i) \leqslant f(i-1)^2 \times |Y|^2 \times (\sqrt{2}k/R)^2 \leqslant f(i-1)^3$. Therefore, for any internal node, the number of candidate points is at most $(|Y|\sqrt{2}k/R)^{3^{|Y|-1}}$. □

From the above discussion, it is easy to see that $M(Y) = O(|Y|!(|Y|\sqrt{2}k/R)^{2^{|Y|}})$.

5.5. CONNECTING THE LOCAL FORESTS AND BOUNDARY POINTS

We can construct a Steiner tree for X from the forest $\hat{F}^P$ as follows. Fix a minimum-cost spanning tree T_S for X and add degree-2 Steiner points to ensure that the length of each edge is at most R. Note that each stem in T_S has length at most c since X is c-local. Let E_P denote the set of crossing edges in T_S. Construct a graph G_P by adding all the crossing edges in E_P to $\hat{F}^P$ and adding degree-2 Steiner points to ensure that the length of each edge is at most R. It is easy to see that

LEMMA 12. *G_P is connected.*

Now, we are ready to introduce our algorithm, which in fact computes $G_{P_{i,j}}$ for every possible partition $P_{i,j}$, selects a $G_{P_{i,j}}$ with the smallest cost, and prunes the selected $G_{P_{i,j}}$ into a tree. See Figure 9.

THEOREM 3. *The performance ratio of the algorithm in Figure 9 is* $1 + [16(4 + 3\log k)c/k]$.

Proof. Consider the stems in the minimum spanning tree for X. Since the boundary area of each cell consists of at most $4(2 + 3\log k)ck$ terminals, each terminal of a crossing stem can be inside a boundary area at most $4(2 + 3\log k)ck$ times under the k^2 partitions. Since the length of a stem is at most c, a stem can be a crossing stem at most $4(4 + 3\log k)ck$ times. Therefore, the total cost of the $k^2 G_{P_{i,j}}$'s is bounded as follows:

$$\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \overline{C}(G_{P_{i,j}}) \leqslant k^2\overline{C}(T^{min}) + \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \overline{C}(E_{P_{i,j}})$$
$$\leqslant k^2\overline{C}(T^{min}) + 4(4 + 3\log k)ck \times \overline{C}(T_s)\,.$$

From Theorem 1, we know that at least one partition yields a solution with cost less than or equal to $1 + [16(4 + 3\log k)c/k]$ times of the optimum. □

Construct a minimum-cost spanning tree T_S for X.
for each possible partition $P_{i,j}$
  begin
    Find the set of crossing edges $E_{P_{i,j}}$.
    for each cell
      Compute a minimum-cost local Steiner forest.
    Let $\hat{F}^{P_{i,j}}$ be the set of all local Steiner forests.
    Construct the graph $G_{P_{i,j}} = E_{P_{i,j}} \cup \hat{F}^{P_{i,j}}$.
    Add degree-2 Steiner points if necessary.
  end.
Select a $G_{P_{i,j}}$ with the smallest cost among all partitions.
Prune $G_{P_{i,j}}$ into a tree.

Figure 9. Algorithm 1.

COROLLARY 1. *There exists a polynomial time approximation scheme for STP-MTP when the set of terminals is c-local.*

COROLLARY 2. *Suppose T_B is produced by the algorithm in Figure 9 and T^* is an optimal tree for STP-MSP. Then*

$$\frac{C(T_B)}{C(T^*)} \leqslant 1 + \frac{16(4 + 3\log k)c}{k} + \frac{16(4 + 3\log k)c}{k} \cdot \frac{4n}{C(T_S)}$$

where T_S is a steinerized minimum spanning tree for the same set of n terminals. That is, there exists a polynomial time approximation scheme for STP-MSP when the given set of terminals is c-local and the minimum spanning tree on n terminals has length at least $(1 + \alpha)nR$ for some positive constant α.

6. Discussion

One of the reasons that we are so interested in the problem STP-MSP is that no geometric optimization problem has been found to be MAX-SNP-hard. STP-MSP may be the one. In fact, Arora's approach [1] does not work for STP-MSP. It is an open problem whether STP-MSP has a polynomial-time approximation scheme.

Acknowledgement

The authors wish to thank Professor Rich Pollack for his comments on a previous version of this paper.

References

1. Arora, S. (1996), Polynomial time approximation schemes for Euclidean TSP and other geometric problems, *Proceedings of 37th FOCS*.
2. Chiang, C., Sarrafzadeh, M. and Wong, C.K. (1990), A powerful global router: based on Steiner min-max tree, in *Proc. ICCAD-89*; also *IEEE Transactions on Computer-Aided Design* 19: 1318–1325.
3. Du, D.-Z. and Hwang, F.K. (1992), A proof of Gilbert-Pollak's conjecture on the Steiner ratio, *Algorithmica* 7: 121–135.
4. Hochbaum, D.S. and Maass, W. (1985), Approximation schemes for covering and packing problems in image processing and VLSI, *J. ACM* 32: 130–136.
5. Hwang, F.K. and Richards, D.S. (1992), Steiner tree problems, *Networks* 22: 55–89.
6. Garey, M.R. Graham, R.L. and Johnson, D.S. (1977), The complexity of computing Steiner minimal trees, *SIAM Journal on Applied Mathematics* 32: 835–859.
7. Gilbert, E.N. (1967), Minimum cost communication networks, *Bell System Technical Journal* 9: 2209–2227.
8. Gilbert, E.N. and Pollak, H.O. (1968), Steiner minimal trees, *SIAM Journal on Applied Mathematics* 16: 1–29.
9. Hwang, F.K., Richard, D. and Winter, P. (1992), The Steiner minimum tree problems, *Annals of Discrete Mathematics*, Vol. 53, North-Holland.
10. Karpinski, M. and Zelikovsky, A. (1997), New approximation algorithms for the Steiner tree problems, *Journal of Combinatorial Optimization* 1(1): 47–65.
11. Li, C.-S., Tong, F.F., Georgiou, C.J. and Chen, M., Gain equalization in metropolitan and wide area optical networks using optical amplifiers, *Proc. IEEE INFOCOM'94*, pp. 130–137, June 1994.
12. Lin, G.-H. and Xue, G.L. (1999), Steiner tree problem with minimum number of Steiner points and bounded edge-length, *Information Processing Letters* 69: 53–57.
13. Rubinstein, J.L. and Weng, J.F. (1997), Compression theorems and Steiner ratios on spheres, *Journal of Combinatorial Optimization* 1(1): 67–78.
14. Ramamurthy, B., Iness, J. and Mukherjee, B., Minimizing the number of optical amplifiers needed to support a multi-wavelength optical LAN/MAN, *Proc. IEEE INFO-COM'97*, pp. 261–268, April 1997.
15. Soukup, J. (1975), On minimum cost network with nonlinear costs, *SIAM Journal on Computing* 29: 571–581.
16. Wang, L. and Jiang, T. (1996), An approximation scheme for some Steiner tree problems in the plane, *Networks* 28: 187–193.
17. West, D.B. (1996), *Introduction to Graph Theory*, Prentice Hall.