Approximation ratio 2 for the Minimum number of Steiner Points

Joon-mo Kim

December 10, 2001
Approximation ratio 2 for the Minimum number of Steiner Points

J. Kim *

Dec. 8, 2001

Abstract

This paper provides an approximation algorithm for STP-MSP (Steiner Tree Problem with Minimum number of Steiner Points). Because it seems to be impossible to have a PTAS (Polynomial Time Approximation Schemes), which gives the near optimal solutions, for the problem, the algorithm of this paper is an alternative that has the approximation ratio 2 with $n^{O(1)}$ run time. The importance of this paper is the potential to solve other related unsolved problems. The idea of this paper is to distribute the error allowance over the problem instance so that we may reduce the run time to polynomial bound out of infinitely many cases.

1 Introduction

This work is grounded on Arora’s works for PTAS. While his work is mostly intended to find the near-optimal solutions for optimization problems, this one finds a way to use his big ideas for some other goals of approximations: the goal of finding (i) the approximations for problems that could not have the near-optimal solutions, (ii) the algorithms that are supposed to exist but not yet found, or so. His idea is allowing (adding) infinitesimal amount of error at each step of the running the algorithm, but the idea here is to allot the amount of error allowance in advance and turn it into geometrical margins where the exponential complexity may reduce down

*Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA. E-mail: jkim@cs.umn.edu
to polynomial one: this is embodying an easy structure out of hard one with some flexibility (error allowance).

We apply the methods of PTAS [3, 4] - the Rectangular Partition and the Dynamic Programming (DP) on it - for another Steiner tree problem, STP-MSP (Steiner Tree Problem with Minimum number of Steiner Points).

The problem definition of STP-MSP [9]: Given $n$ terminals on the Euclidean plane and a positive constant, $r$, find a Steiner tree interconnecting all terminals with the minimum number of Steiner points [10, 6, 7, 5] such that the Euclidean length of each edge is $\leq r$.

This problem is NP-hard with applications in VLSI design, WDM optical networks and wireless communications. Reference [9] also mentions about the difference of the problem from the classical Euclidean Steiner tree problem [8], and the applications.

Lin and Xue showed that the STP-MSP problem is NP-hard [11], and that the approximation which uses Minimum Spanning tree has the approximation ratio 5. Later, [9] showed that the same approximation with an improvement has the approximation ratio exactly 4. It also presented a new polynomial time approximation with the approximation ratio 3, and a polynomial-time approximation scheme under certain conditions. Recently, the ratio 2.5 has been reported. However, such ways of geometric analyses for the optimization problems would require new ideas and much efforts whenever only a bit of improvement is to be made. In addition, it has not yet been successful to design a PTAS, which provides $1+\epsilon$ approximation ratio with the use of the rectangular partitions and generalized dynamic programming, for STP-MSP: actually, it looks impossible.

The study of this paper is to acquire the alternative goal of the approximation ratio 2, and to show that the DP could reach it in a polynomial time. We use the methods of PTAS though the resulting approximation ratio is not the desired one, $1+\epsilon$. As a result, our DP over STP-MSP could give the solutions that has at most twice the number of the the Steiner points that the optimal ones are supposed to have, resulting in the ratio 2.

2 Definitions and the Modifications of the Problem instance

This paper has a different viewpoint for STP-MSP from other related works [1, 2, 3], and some new terms are defined in this section to describe it, in addition to the existing terms from [3]. New terms are underlined when it comes up first.
For the approximation, we are to design an algorithm for the loose version of the problem: actually, the condition of finding 'the minimum number of Steiner points' is removed from the problem definition of STP-MSP and it is: Given \( n \) terminals on the Euclidean plane and a positive constant, \( r \), find a Steiner tree interconnecting all terminals with *Steiner points* such that the Euclidean length of each edge is \( \leq r \).

Then, we are to show that we may design an algorithm for the loose version so that the algorithm should produce the feasible solutions such that the number of the *Steiner points* of each solution is \( \leq \) twice of the minimum number of *Steiner points*.

For a given problem instance of STP-MSP, suppose we know a set of the Steiner points that is the solution of the loose version, then we may draw circles such that each of the Steiner points is the center of a circle, and name the resulting set of circles on the plane as a *Steiner-Cover*. A *Steiner-Cover\textsubscript{OPT}* is a *Steiner-Cover* when the corresponding set of the Steiner points is the optimal solution: the solution for the original STP-MSP problem. Every circle in this paper has the constant radius \( r \). The size of a Steiner-Cover is the the number of its circles.

In this paper, we are to check if the terminals are inside or outside of a circle, not to consider the distance from the terminals to the portals as in other related works [1, 2, 3]. We do not shift the terminals into grids because even a slight movement of a terminal may cause a far different result. Minimizing the unit distance of the grid does not help the situation. But we use grids for the analysis purpose, i.e., we put the line-separators and portals along the grids: their definitions are following.

The *bounding box* of a set of nodes is the smallest *rectangle* enclosing them. A *rectangle* is an axis-aligned rectangle which is a partition of the bounding box. The size of the rectangle is the length of its longer edge. A *line-separator* of a *rectangle* is a straight line segment which is parallel to the shorter edge of the rectangle: the line-separator partitions the rectangle into two of at least \( 1/3 \)rd the area each. For example, if the rectangle’s width \( W \) is greater than its height, then a line-separator is any vertical line in the middle \( W/3 \) of the rectangle. Now we define a recursive partition of a rectangle over which the dynamic program runs.

**Definition 1** (1\( /3 \): 2\( /3 \)-tiling) A 1\( /3 \): 2\( /3 \)-tiling of a rectangle \( R \) is a binary tree (a hierarchy) of sub-rectangles of \( R \). The rectangle \( R \) is at the root. If the size of \( R \) is \( \leq 1 \), then the hierarchy contains nothing else. Otherwise the root contains a line-separator for \( R \), and has two subtrees that are 1\( /3 \): 2\( /3 \)-tilings of the two rectangles into which the line-separator divides \( R \).

Note that *rectangles* at depth \( d \) in the tiling form a partition of the root *rectangle*. The set of all *rectangles* at depth \( d + 1 \) is a refinement of this partition obtained by putting a *line-separator* through each depth \( d \) *rectangle* of size \( \geq 1 \). The area of any
depth $d$ rectangle is at most $(2/3)^d$ times the total area. The following proposition is therefore immediate.

**Proposition 1** If a rectangle has width $W$ and height $H$, then its every $1/3 : 2/3$-tiling has depth at most $\log_{1.5} W + \log_{1.5} H + 2$

**Definition 2** (portals) A portal in a $1/3 : 2/3$-tiling is any point that lies on the perimeters of rectangles in the tiling. If $m$ is any positive integer then a set of portals $P$ is called $m$-regular for the tiling if there are exactly $m$ equidistant portals on the line-separator of each rectangle of the tiling. (We assume that the end-points of the line-separator are also portals. In other words the line-separator is partitioned into exactly $m - 1$ equal parts by the portals on it)

The points are the union of the Steiner points and the given terminals. Either two points or a side of a rectangle and a point are said to be **connectable** if they are within the distance $r$. If each point in a rectangle has a connectable pair, the set of points is also called **connectable**, and **connected** means that all the points of the problem instance have connectable neighboring points.

In view of the PTAS, in order to run the DP, we need to let the circles cross with the portals whenever they pass line-separators, but only a few circles could do that. To give it a feasibility, we are to change the condition so that, when a circle and a portal is chosen, only one point of the circle should cross the portal: though many other points of the same circle may cross other portals, we may disregard such crossings as redundancies. In addition, to keep the bounded number of the combinations, we put a device along the perimeter of each circle, i.e., we mark $p$ equidistant points along the perimeter and give them index from 0 to $p-1$. (See Figure 1) Then let the circle meet the portal only at one of the indexed-points. The feasibility of this weak condition will be supported later. Note that there may exist many redundant choices that overlap each other completely, but we just randomly select one of them: ignoring the redundancies does not cause any difference.

**Definition 3** ($m$-light Steiner-Cover) Let $m \in Z^+$ and $S$ be a $1/3 : 2/3$-tiling of the bounding box and $P$ be an $m$-regular set of portals on this tiling. Then a Steiner-Cover in which each circle crosses with at least one portals in $P$ at its indexed-points is $m$-light with respect to $S$.

On the same plane, we put two Steiner-Covers: one is the copy of the other and each circle of the copy is slightly shifted. Then, we may pick up a pair of circles: one is the original and the other is the shifted copy of it. For each pair, we call the circle-distance as the circle areas which are not overlapped. The set-distance is the union of all the circle-distances of all the pairs.
Figure 1: Indexed-points

3 Ratio 2 and the Structure Theorem

The idea of this section is to measure the approximation ratio: we compare the minimum-sized \emph{m-light Steiner-Cover} with the \emph{Steiner-Cover}$_{\text{OPT}}$ in terms of the number of their circles.

First, we show why our goal is ratio 2. Because the idea of our DP is randomly hitting the desired feasible solutions out of many possible cases, the number of the trials for the desirable ratio of $1 + \epsilon$ turns out to be unbounded. However, fortunately, the geometric properties of the problem let us hold the less-desirable ratio 2, i.e., we found a way to bound the number of the possible cases by a polynomial for the ratio.

The work of this paper is grounded on the following properties [9] of an optimal solution tree for a STP-MSP problem: (i) No two edges cross each other. (ii) Two edges meeting at a vertex form an angle of at least $60^\circ$. (iii) If two edges form an angle of exactly $60^\circ$, then they have the same length.

The \emph{Steiner-Cover$\text{APX}$} is one of the \emph{m-light Steiner-Cover} and has the \emph{set-distance} of $\epsilon$ from \emph{Steiner-Cover}$_{\text{OPT}}$. However, we show next that the DP does not ensure to generate \emph{Steiner-Cover$\text{APX}$} during its bounded run time. Thus, we aim at another feasible solution, \emph{Steiner-Cover$_{2\times\text{OPT}}$}, which has twice as many circles as \emph{Steiner-Cover$_{\text{OPT}}$} does. We show later that there exist many \emph{Steiner-Cover$_{2\times\text{OPT}}$}, for which the DP could reach.

If we temporarily suspend the requirement that the circles should cover the terminals, then the \emph{Steiner-Cover$\text{APX}$} is obtainable by the DP in terms of the \emph{set-distance}. However, once we get back to the condition over the same \emph{Steiner-Cover$\text{APX}$} then there would exist the uncovered terminals in any small area of the \emph{set-distance} and the DP makes more circles to cover those. So the \emph{Steiner-Cover$\text{APX}$} is not assured to be reached by the DP. Such case clearly shows when there exist terminals right on the perimeters of the circles of the \emph{Steiner-Cover$_{\text{OPT}}$}. (See Figure 2).
Approximation by the Distance

To cover terminals

One more covering circle

--- Given Circle

...... Covering Circle

• Terminal Point

Figure 2: Difficulty for approximation by $\epsilon$ movements

Figure 3: $c$-band

Now, we are to consider the case that two covering-circles of which the two centers are within the distance of $r$ cover a given-circle except for one small part, and make some definitions for this case as follows. Because the centers of the covering-circles become the neighboring Steiner points, they should lie within the distance $r$; the shaded small circle represents the area that the centers of the two covering-circles may lie. (See Figure 4)

For a circle, picture that the lines of the perimeters got to have a constant width, $w$, then it forms a doughnut; we call it a circular-band. For the two covering-circles, if both circles are replaced by circular-bands: we call it a $c$-band; see Figure 3. The constant width, $w$, of the $c$-band is determined so that: referring Figure 4, (1) the inner-perimeter of $c$-band nearest to the point ‘$a$’ should lie above ‘$a$’, (2) the same for ‘$b$‘ symmetrically, and (3) the cross point of the inner-perimeters nearest to the point ‘$c$‘ lies below ‘$c$‘.

Later, we mean ‘a pair of circles inside a $c$-band’ as the case that each of the two circular-band of the $c$-band includes one circle inside. The inner-area is the union of the two areas inside the two inner-perimeters.

Lemma 1 ($m$-cover; covering with a margin) For a given-circle whose center is a Steiner point of a shortest optimal solution tree and the terminals on it, there exist a $c$-band such that its inner-area covers all the given terminals.

Proof:
The part of given circle
---
Covering circle
------
Perimeter of band

Inner perimeter of band

Outer perimeter of band

Figure 4: Covering with enough margins

The diameter of the small circle and the radius of the big circles are length $r$.

Figure 5: Covering : Analytic view

By the properties [9] mentioned earlier in this section and the fact that the center of the given-circle is such a Steiner point of the optimal case, the angle between any two neighboring edges must be $> 60^\circ$: an edge is a line between the Steiner point and a terminal.

See Figure 5. Put three points on the perimeter of a given-circle so that they make three fans of angles $150^\circ, 150^\circ$ and $60^\circ$, respectively. Now, consider a big fan with the angle $150^\circ$ and draw a line between the two vertex on the perimeter. Name the center of the line as $c$. Setting $c$ as the center, we may draw a new circle that cover the big fan because the distances from $c$ to the center of the given-circle, to each of the vertex and to the arc of the fan are all $< r$. Symmetrically it applies to the other big fan. By calculations, the distance between $c$ and another $c$ in the other big fan is $2r \cdot \cos 75^\circ \cos 15^\circ = 1/2 \cdot r$. Therefore, the $c$-band in the Lemma 1 could be formed based on the two new circles.
Theorem 1 (Structure Theorem) A set of terminals on $\mathbb{R}^2$ has infinitely many Steiner-Covers$_{2\times OPT}$ which m-cover the Steiner-Cover$_{OPT}$ of the set and could be reached by the DP. The set also has an associated $1/3 : 2/3$-tiling of the bounding box such that the Steiner-Cover$_{2\times OPT}$ is m-light for this tiling, where $m = O(\log n)$ and $L$ is the size of the bounding box. (See Figure 6)

Proof:

By Lemma1 (m-cover), for a Steiner-Cover$_{OPT}$, there exist corresponding Steiner-Cover$_{2\times OPT}$ such that each circle, $C_{opt}$, from the Steiner-Cover$_{OPT}$ is m-covered by a pair of distinct circles from the Steiner-Cover$_{2\times OPT}$ if we put them on the same plane; name the pair as pair$_{2\times OPT}$. A pair$_{2\times OPT}$ has a c-band that includes a set of infinitely many other pairs of distinct circles that we name as pair$_{kin}$; See Figure 7. Each of pair$_{kin}$ covers all the terminals on $C_{opt}$ because it includes the inner-area, and so it functions the same as the pair$_{2\times OPT}$: now pair$_{2\times OPT}$ is just one of pair$_{kin}$. This helps us to lead to the fact that there exist m-light Steiner-Cover of which all the circles are ones of pair$_{kin}$. Now we define Steiner-Cover$_{kin}$ as the generalization of Steiner-Cover$_{2\times OPT}$ such that each of the m-covering pair is a pair$_{kin}$.

As the DP runs and the tiling is built up, each c-band area of Steiner-Cover$_{2\times OPT}$ is to be crossed by line-separators. Because the width of c-band is a constant, $w$, we
Infinitely many kin pairs

c-band perimeters

defined to be

c-h band perimeters

may choose \( d (\leq w) \) as the inter-portal distance so that there exist portals along the line-separators lying inside a c-band.

Additionally, we may set the number of the indexed points, \( p \), large enough so that there exist \( m \)-light Steiner-Covers which are \( \text{Steiner-Covers}_{kin} \) at the same time; \( \text{Steiner-Covers}_{kin} \) could be imagined based on the supposed \( \text{Steiner-Cover}_{OPT} \) on the same plane. That is, there may exist \( m \)-light \( \text{Steiner-Covers}_{kin} \), which are \( m \)-light \( \text{Steiner-Covers} \) and the kin of \( \text{Steiner-Cover}_{2\times OPT} \) at the same time. A kin of \( \text{Steiner-Cover}_{2\times OPT} \) covers \( \text{Steiner-Cover}_{OPT} \) as \( \text{Steiner-Cover}_{2\times OPT} \) does, and next section shows that a \( m \)-light \( \text{Steiner-Covers}_{kin} \) could be reached by the DP.

The circles in a \( \text{Steiner-Cover}_{OPT} \) does not overlap more than five times as explained below and so the number of crossings between a line-separator and circle perimeters is bounded by a constant. Thus, in this problem, we do not have to consider the cases that the number of the crossings is more than \( m \).

The number of crossings is upper bounded by these three factors: (i) For a grid of unit length \( r \) on a plane, we may draw circles for every grid point as its center point, then the circles completely cover all the area. Therefore no more than five overlappings are needed to cover any unit area, and so \( \text{Steiner-Cover}_{OPT} \) does not need more than five overlappings to cover any unit area: See Figure 8. As a result, the number of crossings over a length, \( 2r \), of a line-separator is bounded by a constant, which is 8: the number of crossings at the center looks like 10 in Figure 8, but 2 crossings should be carried over to other parts. (ii) Constant \( L \), the size of the bounding box. (iii) Constant \( r \), the radius of the circles.

Therefore, the maximum number of crossings could be written as \( \frac{8L}{r} \) and so, differently from related works [1, 2, 3], we may ignore the case that the number of
crossings is larger than $m$ because we may choose $m(>\frac{8L}{T})$ and $m$ may be set as $m = O(\log n)$.

4 The Dynamic Programming (DP)

The Structure Theorem guarantees the existence of the $m$-light Steiner-Cover$_{\text{kin}}$ and its associated tilings $S$, where $m = O(\log n)$. Therefore, finding the minimum sized $m$-light Steiner-Covers is the approximation with ratio 2. Remind that the ratio 2 is for the worst case. By the Proposition 1, the depth of any such tiling is at most $O(\log n)$. This section describes the DP that finds both $S$ and m-light Steiner-Cover$_{\text{kin}}$ in a polynomial time, $n^{O(1)}$.

Generally speaking, the DP’s works are to enumerate all the polynomially many combinations of choices and simultaneously build up the minimum cost structures (connected graphs) for each choice. For a rectangle, its circles are determined by the choices of portals and indexed-points, and the DP considers if the centers of the circles would work as the Steiner points that will satisfy the problem later at the top rectangle. For the connectivity concern, if all the points inside the rectangle is connectable, the corresponding case could be accepted for further works at the upper level rectangles. For the complexity concern, if a circle’s center (the Steiner point) does not make a
necessary connection but a redundant one, the circle can be removed from the further works.

When two small rectangles are combined to form a bigger one, the locations of the portals on the small rectangles would not meet exactly with the portals on the bigger one. Then we may move the portal of the smaller rectangle to the nearest one of the bigger rectangle together with its circle. (See Figure 9) This movement is acceptable because the width of $c$-band gives some margin of freedom for moving. We may control the distance of the two neighboring portals on the line-separator of the top most rectangle, and as long as the distance is less enough than the width of the $c$-band, the short movement is acceptable. This part is analogous to the use of the bridges in the related works [1, 2, 3].

Each combinatorial case of a rectangle has an entry in the DP's lookup table and it may point to the entries of the smaller rectangles to form the connectable point set with the minimum number of circles (Steiner points): the minimum number of circles used for each case is recorded in the table entry accordingly. Actually, we check all the polynomially many cases of random circle allocations over the plane. To keep the number of cases to be polynomial, we utilized the error allowance by moving the circles in the smaller rectangles when a bigger rectangle is formed. Actually, we could keep the numbers of the combinatorial cases (choices from $m$ portals) in every rectangle to be the same while the rectangles get bigger at every stage.
At the end of the DP, for each case of the circle allocations in the top most rectangle - the *bounding box* - we check if it is a Steiner-Cover and compare all the table entries of the rectangle to find the one with the minimum number of circles, which is the final result. Checking if a given circle allocation is a solution for the problem is in polynomial time because the problem is NP-Complete.

Now we view the work of the DP in the analytic way. The work of the DP is bottom-up approach, but it is easy to view the procedure from the final stage to the start, i.e., top-down way. The final result of the DP is the rectangle, which is the *bounding box* for the given point set, and which contains the *m-light structure* that satisfies the problem with ratio 2.

We partition the *bounding box* that includes the problem instance with the repeated applications of $1/3 : 2/3$-tiling until all the bottom most rectangles include small number of terminals so that we can find the *connectable* set of points with the brute-force algorithm inside each of the rectangles in polynomial time. That is, for each of the bottom most rectangles, we consider all the possible $2^m$ choices of portals along the surrounding *line-separators* and the $p$ choices of indexed-points along the perimeter of a circle. For each case of the combinations, the DP draws the corresponding circles and checks if the resulting points are *connectable*.

At the upper level rectangles, the *line-separators* are determined by the $1/3 : 2/3$-tiling. We have to find the *connectable* point set for an upper level rectangle out of the combinations of three factors: (i) the *line-separators* inside the rectangle, (ii) the portals along it, and (iii) the choices of the indexed-points of a circle.

For each case of the combinations at a rectangle, we know that we have already computed the *connectable* set of points for both of the left and right sub-rectangles. So, we concatenate the two sets of points from the two sub-rectangles to get the *connectable* set of points of the current rectangle. In this way, up to the root of the tiling, we can search all the possible cases of the *connectable* sets of points: at the root, the DP selects the final result which is the *connected* set of points that has the minimum number of Steiner points(circles).

Now, we need to show the number of entries of the lookup table for this DP is polynomial and the run time for each of the entries is poly time. An entry is indexed by the triple: (a) A *rectangle*, (b) A set of $k_1 (\leq 4m)$ portals along the perimeter of the *rectangle*, and (c) The choices of $k_1$ circle positions, i.e., the permutation of size $k_1$ out of the $p$ indexed-points, $\{0,1,2,...,p-1\}$.

For (a), the number of distinct rectangles is at most $\binom{n}{4}$. For (b), Each *rectangle* has 4 sides which are the parts of the *line-separators* of some upper level *rectangles*. The $m$ portals on the *line-separator* are evenly spaced, so they are completely deter-
mined once we know the line-separators. But the number of choices of a line-separator is at most the number of pairs of points, which is $\binom{n}{2}$. This accounts for the factor $O((n^2)^4) = O(n^8)$. Furthermore, once we have identified the set of $\leq 4m$ portals on the four sides, the number of ways choosing a set of $k_1$ portals is $\binom{4m}{k_1}$. So the choices in (b) is $n^8 \times \sum_{k=1}^{c} \binom{4m}{k}p_k = O(n^{12} \times (2p)^{4m})$

where $m = O(\log n)$, $p$ is a constant, and $c = 4 \cdot \frac{8L}{r}$.

Then, we consider the running time over the lookup table. At the bottom level rectangles, drawing circles and checking if the choice is connectable is a constant time because we may let the rectangles have a bounded number of terminals. For the upper level rectangles, the DP constructs the connectable set of points by the comparisons of the two sub-rectangles, bounded number of movements along the line-separators, concatenation of the two, and checking if the points are connectable. Because the dominant term of the run time is for the connectability checking and others are all constant terms, the run time for the upper level rectangles is at most $n^2$. Therefore the run time of the DP is upper bounded by $n^2 \times \text{Size of the table}$, which is still $n^{O(1)}$.

As a result, the DP checks all the $m$-light Steiner-Covers so that the $m$-light Steiner-Coverkin, of which the existence in any chosen rectangle is ensured by the Structure Theorem, shall be found at the least.

5 Conclusion

As a result, the picture is to approximate the problem for another goal which is not optimal but feasible. This alternative goal has twice of the optimal value and so the ratio became 2 instead of $1 + \epsilon$, but now it is obtainable by the DP in a polynomial time.

If the DP is applied to the real problems, the terminals would not be mathematical points, but has small areas each. So it would be possible to get near optimal solutions with the ratio far less than 2. For such practical purposes, the change of the definition for $c$-band is the only work to do.

This problem has an important near kin, the ‘Minimum Disc Cover’, and possibly more. Such related problems would be solved by the method of this paper with
minor modifications, resulting in the best ratio ever. It should be stressed that this method would provide consistent similar approaches for the related problems, which usually require far different studies each other if geometric analysis is used for the approximations.

In another view, we do not have any clue for the approximation for STP-MSP with the ratio of less than 2. Further, the fact that the objective value is an integer would give us too good results to accept if we succeeded in having a PTAS: for any instance of the problem, we can always choose a small enough $\epsilon$, with which the PTAS gives the optimal solutions. It would be an indication that having a PTAS is impossible. Actually, this challenge does not have any clue at the moment.

The work of this paper is for (i) from the Introduction, and some of the next works will be for (ii).

References


