ELSEVIER

Discrete Mathematics 260 (2003) 189–196

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Note

The Steiner ratio of several discrete metric spaces

Dietmar Cieslik*

Institut für Mathematik und Informatik, University of Greifswald, Friedrich-Ludwig-Jahn-Strasse 15a, Greifswald 17487, Germany

Received 21 November 2000; received in revised form 17 October 2001; accepted 29 July 2002

Abstract

Steiner's Problem is the "Problem of shortest connectivity", that means, given a finite set of points in a metric space (X,ρ), search for a network interconnecting these points with minimal length. This shortest network must be a tree and is called a Steiner Minimal Tree (SMT). It may contain vertices different from the points which are to be connected. Such points are called Steiner points. If we do not allow Steiner points, that means, we only connect certain pairs of the given points, we get a tree which is called a Minimum Spanning Tree (MST). Steiner's Problem is very hard as well in combinatorial as in computational sense, but, on the other hand, the determination of an MST is simple. Consequently, we are interested in the greatest lower bound for the ratio between the lengths of these both trees:

$$m(X,\rho) := \inf\left\{\frac{L(\text{SMT for } N)}{L(\text{MST for } N)} : N \subseteq X \text{ is a finite set}\right\},$$

which is called the Steiner ratio (of (X,ρ)). We look for estimates and exact values for the Steiner ratio in several discrete metric spaces. Particularly, we determine the Steiner ratio for spaces of words, and we estimate the Steiner ratio for specific graphs.

1. Introduction

The "Problem of shortest connectivity", usually called Steiner's Problem, is to find for a finite set of points in metric space (X,ρ) a network interconnecting these points

* Tel.: +49-3834-864614; fax: +49-3834-864615.
E-mail address: cieslik@mail.uni-greifswald.de (D. Cieslik).

0012-365X/03/$ - see front matter
PII: S0012-365X(02)00762-8

with minimal length. More formally: Let $N \subseteq X$ be a finite set of points. Search a connected graph $G=(V,E)$ with

$$N \subseteq V \tag{1}$$

and

$$L(G)=L(X,\rho)(G)=\sum_{\underline{vv'} inE} \rho(v,v') \tag{2}$$

is as minimal as possible.

Any network solving Steiner's Problem must be a tree, which is called a Steiner Minimal Tree (SMT). It may contain vertices different from the points which are to be connected. Such points are called Steiner points.

The problem of finding an SMT has a long-standing history starting with Gauß in 1836 [4]. Perhaps with the famous book *What is Mathematics* by R. Courant and H. Robbins in 1941 this problem has been popularized under the name of Steiner. A classical survey of Steiner's Problem in the Euclidean plane was given by Gilbert and Pollak [5].

Minimum spanning networks are studied for many years and are solved completely in the case where only the given points must be connected. This is called a Minimum Spaning Tree (MST) problem. Given a finite set N of points in a metric space (X,ρ), an MST for N can be found in time which is polynomially bounded in $|N|$ [10]: For the complete graph $(N, \binom{N}{2})$ with the length-function $f: E \to \mathbb{R}$ defined as $f(\underline{vv'})=\rho(v,v')$, sequentially choose the shortest edge that does not form a circle with edges already chosen until $|N|-1$ edges are chosen.

The novelty of Steiner's Problem is that new points, the Steiner points, may be introduced so that an interconnecting network of all these points will be shorter. Given a set of points, it is a priori unclear how many Steiner points one has to add in order to construct an SMT. The following observations are well-known:

Let (X,ρ) be a metric space and let N be a finite set of points in X. Without loss of generality, the following is true for any SMT $T=(V,E)$ for N

(a) $g_T(v) \geqslant 1$ for each vertex v in V;
(b) $g_T(v) \geqslant 3$ for each Steiner point v in V;
(c) $|V \backslash N| \leqslant |N|-2$,.

where $g_T(v)$ denotes the degree of the vertex v in T.

Therefore, we have now the problem which becomes how many extra points should be added, and where should they be placed to minimize the overall network length. This also shows that it is impossible to solve the problem with combinatorial and geometric methods alone.

Steiner's Problem is one of the most famous combinatorial–geometrical problems. Consequently, in the last three decades the investigations and, naturally, the publications about Steiner's Problem have increased rapidly. Surveys are given by Cieslik [2], Hwang, et al. [8] and Ivanov and Tuzhilin [9]. An introduction of the complete subject has been given by Bern and Graham [1], and by Hildebrandt and Tromba [6].

Whereas Steiner's Problem is very hard as well in combinatorial as in computational sense, the determination of an MST is simple. Consequently, we are interested in the greatest lower bound for the ratio between the lengths of these both trees:

$$m(X,\rho) := \inf\left\{\frac{L(\text{SMT for } N)}{L(\text{MST for } N)}:\ N \subseteq (X,\rho) \text{ is a finite set}\right\},$$

which is called the Steiner ratio (of the space (X,ρ)).

The Steiner ratio is a parameter of the considered space and describes the approximation ratio for Steiner's Problem. The quantity $m(X,\rho)L(\text{MST for } N)$ would be a convenient lower bound for the length of an SMT for N in (X,ρ); that means, roughly speaking, $m(X,\rho)$ says how much the total length of an MST can be decreased by allowing Steiner points.

The ultimative goal is to determine or at least to estimate the Steiner ratio for many spaces.

2. The range of the Steiner ratio

What are the values which the Steiner ratio of a metric space can achieve? The following two facts are known, compare [3]:

Theorem 2.1 (Moore in [5]). *For the Steiner ratio of every metric space*

$$1 \geqslant m(X,\rho) \geqslant \tfrac{1}{2}$$

holds.

In other terms, for any finite set N of points in a metric space the length of an MST for N is less than two times of the length of an SMT for N. Moreover,

Corollary 2.2. *Let N be a finite set of n points in a metric space (X,ρ). Then*

$$L(MST\ \ for\ N) \leqslant 2\left(1-\frac{1}{n}\right)L(SMT\ \ for\ N).$$

Next we show that the lower bound 0.5 is the best one over the class of all metric spaces: Let $G=(V,E)$ be a star with n leaves. All edges have length one. The leaves form the set N of given points. Then an MST for N has the length $2(n-1)$ and an SMT with the internal vertex of the star as Steiner point has the length n. Hence, the ratio between the two lengths is $n/(2n-2)$ and if n tends to infinity we have the assertion.

We stated, that graphs are the finite metric spaces. Consequently,

Corollary 2.3. *The lower bound $\frac{1}{2}$ is the best possible one for the Steiner ratio of all metric spaces, even for spaces of finite cardinality.*

3. Spaces which achieve the extreme values of the Steiner ratio

In the next examples, we describe classes of metric spaces in which the Steiner ratio achieve the value 1 or 0.5.

I. In the following spaces Steiner's Problem is as easy as finding a minimum spanning tree.

Let (X,ρ) be a metric space.

A metric ρ is called an ultrametric if

$$\rho(v,w) \leqslant \max\{\rho(v,u), \rho(w,u)\} \tag{3}$$

for any points u, v, w in X. It is easy to see that

Lemma 3.1. *The following is true for all ultrametric spaces* (X,ρ):

$$\textit{If } \rho(v,u) \neq \rho(w,u), \textit{ then } \rho(v,w) = \max\{\rho(v,u), \rho(w,u)\}.$$

That means that all triangles in (X,ρ) *are isosceles triangles where the base is the shorter side.*

Now, we prove

Theorem 3.2. *The Steiner ratio of an ultrametric space equals one.*

Proof. Let $T=(V,E)$ be an SMT for N. Let Q denote the set of all Steiner points in T, i.e., $Q=V\backslash N$. Suppose that Q is nonempty.

There is a Steiner point q in Q such that q is adjacent to two vertices v and v' in N. Otherwise, each vertex in Q is adjacent to at most one vertex in N. The set Q induces in T a subgraph $G'=(Q,E')$, for which it follows

$$\begin{aligned} |E'| &= \tfrac{1}{2}\sum_{v\in Q} g_{G'}(v) \\ &\geqslant \tfrac{1}{2}\sum_{v\in Q} (g_T(v)-1) \\ &\geqslant \tfrac{1}{2}\sum_{v\in Q} 2 \\ &= |Q|. \end{aligned}$$

This contradicts the fact that the forest G' has at most $|Q|-1$ edges.

Using 3.1, we may assume that $\rho(v,v')=\rho(v,q)$. The tree $T'=(V,E\backslash\{\underline{vq}\}\cup\{\underline{vv'}\})$ has the same length as T, and it is an SMT for N, too. If $g_{T'}(q)\geqslant 3$ we repeat this procedure. If $g_{T'}(q)=2$ we find an SMT with a smaller number of Steiner points than T, since no Steiner point has degree smaller than 2. Hence, we proved, that Steiner's Problem in an ultrametric space is the same as finding an MST. Consequently, we have the assertion. □

II. Note, that we define the Steiner ratio as the greatest lower bound of the ratios of SMT— by MST-lengths, since it is not sure that there is a finite set of points in a space which ratio achieves the Steiner ratio.

For a finite metric space (X, ρ) there is a finite set N of points such that

$$m(X, \rho) = \frac{L(\text{SMT for } N)}{L(\text{MST for } N)}. \tag{4}$$

Consequently, the Steiner ratio may be defined as a minimum problem.

In the next section we will describe metric spaces with Steiner ratio equal 0.5. As a consequence of 2.2 we have

Theorem 3.3. *Let (X, ρ) be a metric space with Steiner ratio $\frac{1}{2}$. Then there does not exist a finite set of points in X which achieves the Steiner ratio.*

4. Sequence spaces

Let A be a finite alphabet. Its elements will be called letters. A word over A is a finite sequence of letters. Let d be a positive integer. Then A^d denotes the set of all sequences of length d. We define the Hamming metric ρ_H over A^d in

$$\rho_H((a_1, \ldots, a_d), (b_1, \ldots, b_d)) = |\{i\colon\ a_i \neq b_i \text{ for } i = 1, \ldots, d\}|.$$

We investigate specific sets of words to find an upper bound for the Steiner ratio of a sequence space (A^d, ρ_H): Let us assume that $|A| > 1$ and $d > 1$. Let a and b be two different letters. Then consider the words w_i which only consists of the letter a, except the ith position where the letter b is located, $i = 1, \ldots, d$.

For $i \neq j$ it holds $\rho_H(w_i, w_j) = 2$. Hence, $L(\text{MST for } \{w_1, \ldots, w_d\}) = 2(d-1)$. The word $w = a \ldots, a$ has distance 1 to any w_i. Consequently, the star with the center w and the leaves w_i, $i = 1, \ldots, d$ is an SMT which has length d.[1] Now, we have

Theorem 4.1. *For the Steiner ratio of the sequence space (A^d, ρ_H), $|A| > 1$, $d > 1$, it holds*

$$\frac{1}{2} \leqslant m(A^d, \rho_H) \leqslant \frac{d}{2(d-1)}.$$

The value $d/2(d-1)$ tends to $\frac{1}{2}$ if the dimension d of the space runs to infinity. Since in all applications of sequence spaces,[2] the dimension is a great number, we may assume that

$$m(A^d, \rho_H) \approx \tfrac{1}{2} \tag{5}$$

if $d \gg 1$.

[1] That this tree is really an SMT, we see by the fact that 1 is the smallest positive distance in the space.
[2] Namely in the consideration of molecular biology datas.

The set of all words over the alphabet A is denoted by $A^\star$. That means

$$A^\star = \bigcup_{d=0,1,\ldots} A^d.$$

The Levenshtein, or edit distance, between two words of not necessarily equal length is the minimal number of "edit operations" required to change one word into the other, where an edit operation is a deletion, insertion, or substitution of a single letter in either word. [3] To determine the Steiner ratio of the phylogenetic space, consider the words w_i which consists of the letters a, except the ith position where another letter b is located, $i=1,\ldots,d$. Then define the set

$$N(d) = \{w_i : |w_i| = d, i = 1,\ldots,d\}$$

of d points.

For $i \neq j$ it holds $\rho_L(w_i,w_j) = 2$. Hence, $L(\text{MST for } N(d)) = 2(d-1)$.

The word $w = a\ldots a$ has distance 1 to any w_i. Consequently, the star with the center w and the leaves w_i, $i=1,\ldots,d$ is an SMT for $N(d)$ which has length d. [4] Hence,

$$m(A^\star,\rho_L) \leqslant \frac{d}{2(d-1)} \tag{6}$$

for all positive integers $d \geqslant 2$. Now, we have found a metric space which achieves the lower bound $\frac{1}{2}$ for the Steiner ratio. [5]

Theorem 4.2. *For the Steiner ratio of the phylogenetic space* [6] $(A^\star,\rho)$, $|A|>1$, *it holds*

$$m(A^\star,\rho_L) = \tfrac{1}{2}.$$

5. The Steiner ratio of graphs

Let $G=(V,E)$ be a connected graph, then we may consider G as a metric space, where the distance $\rho(v,v')$ is defined in the way that it is the number of edges in a shortest path between the vertices v and v' in G. The Steiner ratio gets the form

$$m = m(G) = \min\left\{\frac{L(\text{SMT for } N)}{L(\text{MST for } N)} : N \subseteq V\right\}. \tag{7}$$

[3] In general, the sequence space (A^d,ρ_H) with Hamming distance is not a subspace of the phylogenetic space $(A^\star,\rho_L)$: Consider the two words $v = abab\ldots ab$ and $w = baba\ldots ba$ of length d, whereby d is an even integer, then $\rho_L(v,w)=2$ but $\rho_H(v,w)=d$.

[4] Similar facts we use above to estimate the Steiner ratio of the sequence space.

[5] But we do not have a finite set N of points such that

$$\frac{L(\text{SMT for } N)}{L(\text{MST for } N)} = \frac{1}{2}.$$

Furthermore, such set cannot exist, compare 3.3.

[6] And all spaces with a distance or a similarity measure for $A^\star$, compare [11].

Theorem 5.1. *Let G be a (connected) graph. Then for the Steiner ratio of G*

$$\tfrac{1}{2} \leqslant m(G) \leqslant 1$$

holds.

These bounds are the best possible ones.

Proof. The inequalities are obvious. The last statement will be proved by the following two theorems. □

Now, we give a little collection of known Steiner ratios for (connected) graphs:

Theorem 5.2. *The value for the Steiner ratio of complete graphs, paths and cycles equals* 1.

Theorem 5.3. *Let G be a star with k leaves. Then*

$$m(G) = \frac{k}{2(k-1)}.$$

Proof. Considering, the leaves as the set of given points we find an MST of length $2(k-1)$ and an SMT of length k. Hence,

$$m(G) \leqslant \frac{k}{2(k-1)} = \frac{1}{2 - 2/k}. \tag{8}$$

It is easy to see that all other sets of given points do not form a smaller value for the Steiner ratio. □

Lemma 5.4. *Let G be a graph in which no vertex has a degree greater than three. Then*

$$m(G) \geqslant \tfrac{3}{4}.$$

Proof. Let an SMT for a finite set of vertices be given. If there is a Steiner point used then we have a subset $N = \{v_1, v_2, v_3\}$ which creates a star consisting of three edges from v_1, v_2 and v_3 to the common Steiner point v.

Say that $\rho(v_2, v_3)$ is greater than both $\rho(v_1, v_2)$ and $\rho(v_1, v_3)$. Then

$$L_M := L(\text{MST for } N) = \rho(v_1, v_2) + \rho(v_1, v_3).$$

The SMT for N has a length L_S less than L_M. Then

$$\begin{aligned} 4L_S &= 4(\rho(v_1, v) + \rho(v_2, v) + \rho(v_3, v)) \\ &= 2(\rho(v_1, v) + \rho(v, v_2)) + 2(\rho(v_2, v) + \rho(v, v_3)) \\ &\quad + 2(\rho(v_3, v) + \rho(v, v_1)) \end{aligned}$$

$$\geqslant 2(\rho(v_1, v_2) + \rho(v_2, v_3) + \rho(v_3, v_1))$$

$$\geqslant 2L_M + 2\rho(v_2, v_3)$$

$$\geqslant 2L_M + \rho(v_1, v_2) + \rho(v_1, v_3)$$

$$\geqslant 3L_M. \qquad \square$$

Theorem 5.5. *Let G be an $m \times n$-grid, whereby m, n are positive integers. Then*

$$m(G) = \begin{cases} 1 & \text{if } n = 1 \text{ or } m = 1, \\ 0.75 & \text{if } n = 2 \text{ or } m = 2, \\ 0.666\ldots & \text{if otherwise.} \end{cases}$$

Proof. The first assertion follows from 5.2 and the second from 5.4. In the last case we embed the grid in the affine plane with rectilinear norm: That means that $\|(x, y)\| = |x| + |y|$. This embedding prefers the distances for the vertices of the grid. Consequently, the Steiner ratio of the grid is at least the Steiner ratio of the plane. And for the plane with rectilinear distance the Steiner ratio $\frac{2}{3}$ is known [7]. $\square$

References

[1] M.W. Bern, R.L. Graham, The shortest network problem, Sci. Amer. 260 (1989) 84–89.
[2] D. Cieslik, Steiner Minimal Trees, Kluwer Academic Publishers, Dordrecht, 1998.
[3] D. Cieslik, The Steiner Ratio, Kluwer Academic Publishers, Dordrecht, 2001.
[4] C.F. Gauß, Briefwechsel Gauß-Schuhmacher, in: Werke Bd. X,1, Göttingen, 1917, pp. 459–468.
[5] E.N. Gilbert, H.O. Pollak, Steiner minimal trees, SIAM J. Appl. Math. 16 (1968) 11–29.
[6] S. Hildebrandt, A. Tromba, The Parsimonious Universe, Springer, Berlin, 1996.
[7] F.K. Hwang, On Steiner minimal trees with rectilinear distance, SIAM J. Appl. Math. 30 (1976) 104–114.
[8] F.K. Hwang, D.S. Richards, P. Winter, The Steiner Tree Problem, North-Holland, Amsterdam, 1992.
[9] A.O. Ivanov, A.A. Tuzhilin, Minimal Networks—The Steiner Problem and Its Generalizations, CRC Press, Boca Raton, FL, 1994.
[10] J.B. Kruskal, On the shortest spanning subtree of a graph and the travelling salesman problem, Proc. Amer. Math. Soc. 7 (1956) 48–50.
[11] J. Setubal, J. Meidanis, Introduction to Computational Molecular Biology, PWS Publishing Company, Massachusetts, 1997.