Discussion Paper No. 632

AUGMENTING EUCLIDEAN NETWORKS--
THE STEINER CASE

by

Dan Trietsch*

February 1984
(Revised November 1984)**

*Visiting Assistant Professor of Operations Management and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60201.

**The first version of this paper was written while the author was in the Faculty of Management, Tel Aviv University.

The author wishes to thank the referees for their useful remarks.

AUGMENTING EUCLIDEAN NETWORKS--THE STEINER CASE

Abstract

On a Euclidean plane, a network and n points are given. It is required to interconnect points to the network by links of minimal total length. The use of Steiner points is allowed, and connections can be made anywhere along the edges or to the vertices of the network. We prove that the problem can be solved in finite time by methods similar to those used for the Euclidean Steiner tree problem. The problem can be generalized to include flow dependent costs for the various links, or to allow for the connection of several networks. However, even in the form discussed in this paper, it may be useful for problems such as connecting new customers to existing networks (for example, computers, telecommunication, electricity, water, sewage disposal), where the projected flow between the networks does not justify more than the minimal possible investment.

Keywords: Steiner tree, network design.

I. Introduction

The connection of new customers to electrical networks or new exchanges to communications networks is encountered frequently. A network exists and is to be augmented by connecting new nodes (points) to it. This problem can be stated as follows: on a Euclidean plane, a set of points and a network (consisting of straight edges) is given. It is required to connect the points to the network by links of minimal total length.

We assume that links may be connected to the network anywhere along its edges, and also that extra nodes may be added, if by doing so, the total length of the links is reduced. (In effect this also means that our network is embedded in a planar graph, since intersections may be considered as nodes.)

Our problem is actually a generalization of the well known minimal (Euclidean) Steiner tree problem [5,4,1]. A set of n points on a Euclidean plane are to be connected by a network of minimal total length, which may span (up to n-2) additional points, named Steiner Points (if the total length of the network can be reduced by incorporating them). It is easy to show that each Steiner point is of degree (rank) three, and the angles formed between adjacent pairs of arcs anywhere in the tree are at least 120° (120° exactly at the Steiner points). There are many possible configurations (also referred to as "topologies" in the literature) for Steiner trees, which is the inherent reason for the fact that the Steiner tree problem is NP-Hard [2]. (It follows that our generalization is NP-Hard too.) However, for any given configuration an easy construction is available whereby the tree can be drawn using a ruler and compass (see Melzak [5]). The construction may degenerate, which would indicate that no Steiner tree with the configuration involved exists. A useful tool in constructing Steiner trees was suggested by Cockayne [1], and named by him "The Steiner Polygon." The Steiner polygon serves a dual purpose: (i) it defines an area which contains the MST (minimal Steiner tree); (ii) it establishes a cyclic order among the points on its boundary, which makes the construction of the various configurations much easier, by ruling out in advance many seemingly possible topologies. It may also serve to decompose the problem to smaller ones, if it intersects itself.

In this paper we extend the Steiner construction to our generalized version of the problem, thus making it possible to solve it in finite (exponential!) time. We also show that the concept of the Steiner polygon can be partly generalized to our problem.

In order to motivate our discussion, consider the following simple example: the network G is a simple segment and we have to connect two points, 1 and 2, to it. The solution may differ, depending on the location of points 1 and 2 in relation to G. In Figure 1 we illustrate some possibilities.

Insert Figure 1 about here

The solutions to cases a, b and c are obtained by a separate solution for the two points. In case d, one of the points is connected directly to G, and the other is connected through it. In cases e and f, the two points are connected through a (third) Steiner point, such that all three angles formed are exactly 120°. Using results we shall develop, we can show that cases b and c must be optimal (for their arrangements respectively), but in case a, we show by a broken line another tentative solution which must be checked before the optimal solution is identified for sure.

Figure 2 depicts a much more complicated example, where 10 points should be connected to a network which consists of three edges. The optimal solution is given by broken lines, and includes a total of four Steiner points. We shall refer to this example in more detail, towards the end of the paper.

Insert Figure 2 about here

II. The Problem

On a Euclidean plane, let a set N of $n \geqslant 1$ points and a network G(V,A) be given, where V is a set of vertices and A a set of straight edges which span the vertices of V. It is required to connect all the points in N to G in such a manner that the total length of the required links is minimized.

If G degenerates to a single point, our problem is reduced to the well known Steiner tree problem. Therefore, we have a generalized version of that problem, and may refer to the optimal solution as the Generalized Minimal Steiner Tree (GMST). Note that the links incorporated in the solution may or may not form a spanning tree by themselves, but together with G, they do span $N \cup V$, and if G is a tree, then the GMST is also a tree. In our discussion, we may refer to G as a single supernode (to which we assign the index 0), so that in a sense the GMST is a tree, even if G contains cycles. However, there is no reason to expect that the GMST will be a proper Steiner tree (let alone an MST) for $N \cup V$.

Simple and Compound GMSTs: If N is ultimately connected to G through one link exactly, we call the GMST a Simple GMST. All other cases are named Compound GMSTs and are actually combinations of partial, simple GMSTs. For example, in Figure 2 the GMST is compound, but it can be broken down to four simple components, namely the connections of $\{1,2,3,4\}$, $\{5,6,7\}$, $\{8\}$ and $\{9,10\}$ to G.

We refer to the case where $|N| = 1$ as "the basic case." In order to solve it, we have to find the nearest arc of A to the node (where each arc includes its endpoints); if the connection is not through an endpoint, it must be by a perpendicular link. Thus, in order to solve the basic case, we have to check up to $|A|$ arcs.

For some less trivial cases, we need the extended Steiner construction, which is applicable to a set $M \leqslant N$ of $m \leqslant n$ nodes connected to G by a simple GMST with m-1 Steiner points exactly. This is actually a full Steiner tree for m + 1 nodes, and we refer to it as a generalized full Steiner tree, or a GFST. Note that for m = 1 a GFST is a single link.

According to Cockayne [1], a full Steiner tree can be represented by a notation which indicates a pairing order of nodes of N (where a pairing implies representing two points by another point on the apex of an equilateral triangle based on the segment associated with the pair). Cockayne also showed that node n (or any other node) can always be left alone and thus become an endpoint of the segment which represents the whole FST at the end of the pairing process. (As an example, take a Cockayne notation such as ((1,2), (3,4)), 5; this would indicate representing 1 and 2 by (1,2), 3 and 4 by (3,4), and finally (1,2) and (3,4) by ((1,2), (3,4)), which, with 5, forms a segment. What we do in our case is simply to leave G, the $(m + 1)^{th}$ node, as such an endpoint, and connect it to the other node--representing G--as per the basic case (thus locating the exact point to which we should make the connection). Note that it is very easy to apply this extended Steiner construction not only to cases where G is a single segment, but to any network, as long as we assume that we are looking for a GFST with a given configuration.

As in the regular Steiner construction, degeneracy may occur, thus indicating that a certain configuration for $M \cup G$ does not exist.

An Example: Let $N = \{1,2,\ldots,10\}$, as depicted in Figure 2, and let $M = \{5,6,7\} \subset N$. It can be shown that for this subset, the GFST depicted in Figure 3 is the GMST, and the

Insert Figure 3 about here

Cockayne notation associated with it is ((5,6),7), 0. Figure 3-a illustrates the extended construction, where nodes 5 and 6 are represented by (5,6), which is represented in turn with 7 by ((5,6),7); ((5,6),7) is connected to G perpendicularly, as per the basic case. (Part b of the figure describes the generalized Steiner polygon for this subset, which we discuss later.)

At this stage, we can present an algorithm which solves our problem in finite time, as follows: Look for all the possible subsets $M \subseteq N$ which can be connected by GFSTs to G; each such M, along with the minimal GFST associated with it implies a super-network (which includes G and the minimal GFST), and a set $N - M$ of nodes which we still have to connect (unless $N - M = \emptyset$; clearly, if we continue in a similar manner (find a subset, etc.), we must ultimately obtain a solution, and the best of these solutions is the required optimum.

It can be shown that the proposed algorithm is exponential in $|N|$, and it is by no means presented here as an efficient algorithm. It is comparable, however, to the regular Steiner tree algorithm [1], since both "check" all the partitions of N.

One feature our algorithm does not make use of--at this stage--is the Steiner polygon. Indeed, it is possible to generalize the Steiner polygon for our case, and we proceed to do so now. However, for this generalization we have to confine ourselves to cases where G is a single segment. This is not too restrictive since simple trees are always connected through single segments or endpoints thereof.

For completeness, we present Cockayne's original Steiner polygon first. We do that in a slightly different manner from Cockayne's own presentation, so that the generalization will be more natural. The definition is an iterative one, as follows:

The Steiner Polygon (Cockayne): For a set N of n points, connect all $n(n-1)/2$ pairs by straight segments, and let P_0 be the convex hull polygon of all the segments (and N). Obviously P_0 is formed by a subset of the segments, and a subset of N is on its boundary. This completes our initial preparations, and we proceed with stage 1.

In stage i (for $i=1,\ldots$), P_{i-1} is given. If for any edge on the boundary of P_{i-1}, say $\overline{k,\ell}$, there exists a point $m \in N$ such that $\Delta m,k,\ell$ does not contain any other point of N, and such that $\sphericalangle\, k,m,\ell \geqslant 120^\circ$, then P_i is obtained from P_{i-1} by dropping $\overline{k,\ell}$ and incorporating $\overline{k,m}$ and $\overline{m,\ell}$ in its stead. If no such boundary edge exists, P_{i-1} is the Steiner polygon.

To proceed with our generalized definition, which applies to the case of connecting a set N of n points to a simple network G, consisting of a single edge $\overline{a,b}$, we now define:

The Semi-Generalized Steiner Polygon (SGSP): For a set N of n points, and a network $G(V,A)$, where $V = \{a,b\}$ and $A = \{\overline{a,b}\}$, designated as (super) node 0, connect all $(n+1)n/2$ possible pairs by straight segments, where segments connecting points in N to G, such as $\overline{j,0}$, are obtained as per the basic case, and let P_0 be the convex hull polygon of all the segments. Obviously P_0 is formed by a subset of the $(n+1)n/2$ segments, and possibly a segment $\overline{c,d} - \overline{a,b}$. This completes our initial preparations, and we proceed with stage 1.

In stage i, P_{i-1} is given. For any boundary edge $\overline{k,\ell}$ where $k,\ell \in N$, proceed as in the ungeneralized case, but the internal point m may also be a or b (i.e., $m \in N \cup V$), and $\Delta k,m,\ell$ should not contain any other point of $N \cup G$. If no such boundary edge exists, P_{i-1} is the semi-generalized Steiner polygon.

And finally, in order to define the Generalized Steiner Polygon, we begin with the SGSP, and try to make it even smaller, as follows:

The Generalized Steiner Polygon (GSP): Starting with the SGSP as P_o, we proceed with stage 1. In stage i we have P_{i-1}. If a boundary edge of P_{i-1} such as $\overline{k,\ell}$ where $k \in N$, $\ell \in N \cup G$ and a point $n \in N$ exist in such a manner that no other point of N is in the area defined by the edges $\overline{k,\ell}$, $\overline{k,m}$, $\overline{m,\ell}$ and possibly G (if $\ell = 0$, i.e., $\ell \in G$), and such that $\sphericalangle k,m,\ell \geqslant 120°$, then P_i is obtained from P_{i-1} by dropping $\overline{k,\ell}$ and incorporating $\overline{k,m}$ and $\overline{m,\ell}$ in its stead. If no such edge exists, P_{i-1} is the GSP.

Note: If $\ell = 0$, then $\overline{k,\ell}$ and $\overline{m,\ell}$ (or $\overline{k,0}$ and $\overline{m,0}$) do not necessarily connect to G at exactly the same point. This is the main difference between the GSP and the SGSP or the regular Steiner polygon, and this is the reason that the GSP does not necessarily contain the GMST, even though the SGSP does.

The following three results were introduced by Cockayne [1], where the complete proofs may be found.

Theorem 1 (Cockayne): The MST is completely contained within the Steiner polygon.

It is easy to verify that Theorem 1 can be directly extended to our case if we substitute the MST and the Steiner Polygon with the GMST and the SGSP respectively; hence, we will assume that Theorem 1 includes this variation.

Corollary (Cockayne): If the Steiner polygon intersects itself, it is sufficient to solve the Steiner problem separately for each resulting part of the polygon.

By virtue of the corollary, we may hence assume that the Steiner polygon is not self-intersecting. This implies that there is a well defined cyclic order for its vertices.

Theorem 2 (Cockayne): When applying the Steiner construction to any FST topology, the points on the Steiner polygon need only be considered in their cyclic order, though internal points of the polygon may be inserted between pairs of points.

We find it worthwhile to mention that in his proof of Theorem 1, Cockayne (correctly) states as self evident that P_0 must contain the MST, and proceeds inductively to negate the possibility that P_i may not contain it if P_{i-1} does. Theorems 3 and 4, which follow, are analog to Theorems 1 and 2 respectively, where the GSP replaces the Steiner polygon, and the GMST or the GFST replaces the MST or the FST. The reader can verify that it is not always possible to use the GMST in the generalized case; hence, we state Theorem 3 for simple GMSTs only, and since Theorem 3 is required for the proof of Theorem 4, we have to settle for the minimal GFST there. In this context, we mention that an attempt to generalize the GSP for any G (and not just one segment) fails in the sense that there would be a counter example to Theorem 3.

Theorem 3: If the GSMT is simple, it is completely contained by the GSP.

For the proof, which is similar to that of Theorem 1, see [6] (the first version of this paper).

Theorem 4: When applying the Extended Steiner construction to a GFST (assuming one exists), the points on the GSP need only be considered in their cyclic order, though internal points may be inserted between pairs.

Proof: Since the GFST is an FST, it follows that it is simple and G is placed correctly in the cyclic order. Hence, Cockayne's original proof for Theorem 2 holds here as well.

We now return to our general problem for any G. Our premise is that G consists of a finite number of connected segments, and for any subset $M \subseteq N$ and full topology, we can easily locate the best segment to connect through. However, in the choice of M, we can sometimes save much time by intelligent inspection.

Let P_0 denote the convex hull polygon of N (as in its Steiner polygon definition). Let P be the Steiner polygon for N. Finally, let Q denote the Steiner polygon as defined for $N \cup V$, then:

Theorem 5: When solving for any set M, only those edges of G which are accessible from P_0 by straight uninterrupted lines need be considered.

Proof: Trivial. []

Theorem 6: If Q is partitioned by various chains of edges in G to some disjoint faces, the GMST can be obtained by solving separately for the edges of each such face and the subset of N which is contained in it.

Proof: If we suppose that there is a direct connection between two such faces, two possibilities exist: (i) This connection is by an edge which intersects G; (ii) The connection does not intersect G, but passes (partly) outside Q. In case (i), it is clear that the GMST could still be found by separate solutions for the two faces. In case (ii), an inductive procedure, such as the one used in the proof of Theorem 3, will rule this possibility out. []

Theorem 7: If P is intersected by G to some disjoint parts, it suffices to solve the problem separately for the points of N in each of these parts.

Proof: Clearly the condition is sufficient (but not necessary) for Theorem 6 to hold. □

Note: Theorem 5 will then serve to identify the edges of A which need be considered for each subset of N.

Theorems 5 and 6 may be used in a straightforward manner to facilitate the solution procedure. Theorem 7 is not strictly necessary since its applicability implies the applicability of Theorem 6; however, when solving manually, it is sometimes clear at a glance that Theorem 7 applies (e.g., case b in Figure 1), and it is easier to apply than Theorem 6.

III. An Example:

In the example depicted in Figure 2, we have $N = \{1,2,\ldots,10\}$, $V = \{a,b,c\}$, $A = \{\overline{a,b},\ \overline{b,c},\ \overline{b,d}\}$, $P = \overline{1,2,3,4,5,6,7,8,9,10,1}$ and $Q = \overline{a,1,2,3,d,5,6,7,c,8,9,10,a}$. By Theorem 7, $\{5,6,7\}$ can be solved separately, with possible connections to $\overline{b,c}$ or $\overline{b,d}$. $N - \{5,6,7\}$ should be solved with possible connections to all three edges in A. However, Theorem 6, which in this case is indeed better than Theorem 7, allows us to partition $N - \{5,6,7\}$ further. Thus, we can solve for $\{1,2,3,4\}$ separately, considering only connections to $\overline{a,b}$ or $\overline{b,d}$; and $\{8,9,10,11\}$ can be solved while considering only connections to $\overline{a,b}$ or $\overline{b,d}$. We have already discussed the case of $M = \{5,6,7\}$; the GSP for this case is depicted in Figure 3-6, and it imposes the cyclic order which we have used.

Conclusion

We have demonstrated a finite procedure to solve the network augmenting problem, by extending to it known techniques used in the Steiner tree problem, including the Steiner construction which can be executed by a ruler and compass. Much work remains to be done. For instance, specific heuristics should be developed to parallel those known for the Steiner tree problem. The problem may be extended to augmenting a network with rectilinear distances, or a network embedded within a graph--cases for which versions of the Steiner tree problem exist. The problem may also be generalized further in two directions at least. One such generalization is to assign costs to the edges according to their flows (see Gilbert [3], or Trietsch and Handler [7]). Another generalization is to allow more than one network to be given. We call this version the Network Connection Problem, and in a subsequent paper, we intend to show that it is also solvable in finite time. Finally, we note that the two generalizations we mentioned do not rule each other out.

References

[1] E. J. Cockayne, On the Efficiency of the Algorithm for Steiner Minimal Trees, SIAM J. Appl. Math. 18 (1970), pp. 150-159.

[2] M. R. Garey, R. L. Graham and D. S. Johnson, The Complexity of Computing Steiner Minimal Trees, SIAM J. Appl. Math. 32 (1977), pp. 835-859.

[3] E. N. Gilbert, Minimum Cost Communication Networks, Bell System Technical Journal 46 (1967), pp. 2209-2227.

[4] E. N. Gilbert and H. O. Pollak, Steiner Minimal Trees, SIAM J. Appl. Math. 16 (1968), pp. 1-29.

[5] Z. A. Melzak, On the Problem of Steiner, Canad. Math. Bull. 4 (1961), pp. 143-148.

[6] D. Trietsch, Augmenting Euclidean Networks--The Generalized Steiner Case, working paper No. 799/84, IIBR, Tel Aviv University, Faculty of Management.

[7] D. Trietsch and G. Y. Handler, The Gilbert and Pollak Conjecture--A Generalization, working paper No. 763/R/84, IIBR, Tel Aviv University, Faculty of Management.

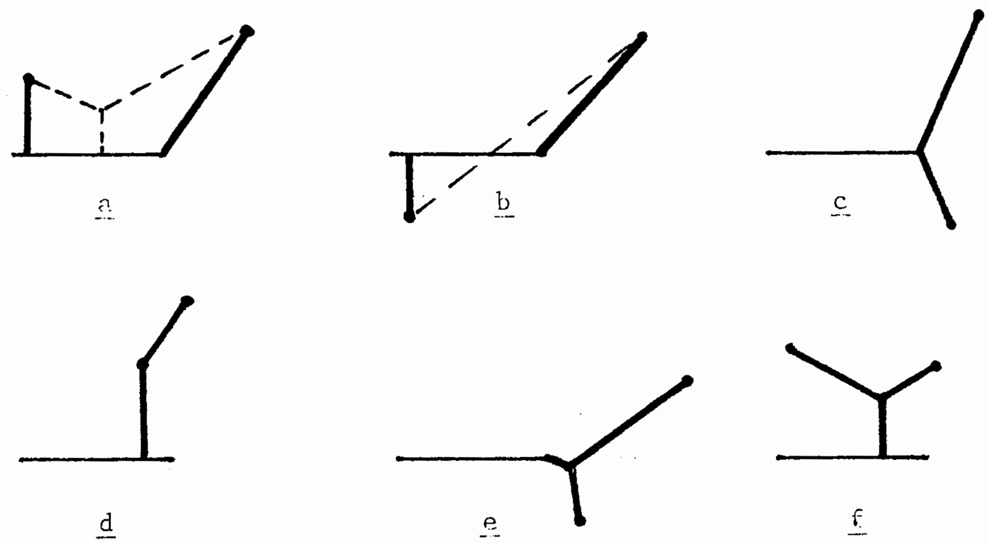

Figure 1

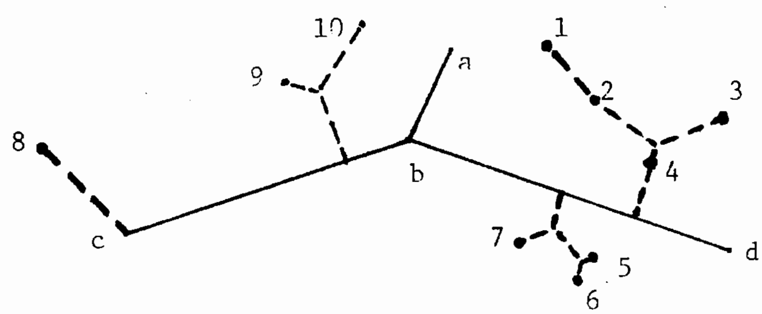

Figure 2

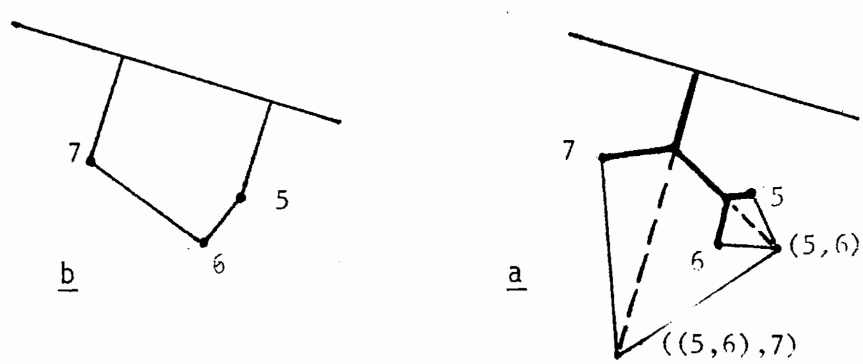

Figure 3