Euclidean Steiner Minimal Trees for 3 Terminals in a Simple Polygon

Pawel Winter *
Dept. of Computer Science, University of Copenhagen
Universitetsparken 1, DK-2100 Copenhagen Ø
DENMARK

June 12, 1995

Abstract

An $O(n)$ time and space algorithm for the Euclidean Steiner tree problem with three terminals in a simple polygon with n vertices is given. Its applicability to the problem of determining good quality solutions for any number of terminals is discussed.

1 Introduction

We consider the following variant of the *Euclidean Steiner tree problem* (ESTP):

- **Given:** A simple polygon P and three terminals t_i, t_j and t_k in P.
- **Find:** *Euclidean Steiner minimal tree* (ESMT) spanning t_i, t_j, t_k in P.

We present an $O(n)$ time and space algorithm for this problem. Our interest in this special case is due to the fact that it is one of the steps toward an efficient heuristic for the ESTP inside a polygon for *any* number of terminals. To justify this claim, we need to consider the obstacle-free case first. The reader is referred to [6] for basic definitions and properties of ESMTs. Obstacle-free ESMTs tend to consist of unions of relatively small full Steiner trees. The greedy concatenation of ESMTs for small, approprietely chosen, subsets of up to 4 terminals proved therefore to yield good quality solutions; ESMTs are ordered by non-increasing savings over the lengths of the Euclidean minimum spanning trees (EMST). One way to identify clusters of terminals is to consider an EMST for all terminals [2]. Subsets of 2, 3 and 4 terminals with connected induced subgraphs of the EMST seem to be reasonable candidates. Another way to identify clusters is to use the Delaunay triangulation of the terminals [10]. Subsets of terminals on a common edge, triangle, and pair of triangles sharing an edge, seems to be reasonable condidates.

ESMTs for terminals inside a simple polygon P most likely consist of unions of small full Steiner trees. However, no exact algorithm for this problem is available. Suppose that clustering into subsets with 2, 3 and 4 terminals is available. We are then left with the problem of finding ESMTs for these small clusters in P. For clusters of size 2, this reduces to the shortest path problem inside P, and can be solved in linear time and space [3, 7]. In this paper we address the problem of determining ESMTs for 3 terminals inside P in linear time. In a companion paper [12], we give a polynomial algorithm for the ESMT with 4 terminals.

The problem of determining reasonable clusters with 2, 3 and 4 terminals inside a simple polygon is far from trivial. The simplest way is perhaps to consider a dual of the geodesic Voronoi diagram for n terminals

*This research was completed during a visit at RUTCOR and DIMACS, Rutgers Univ., New Brunswick, NJ, U.S.A. It was partially supported by AFORS grant F49620-93-1-0041 and NSF grant STC 91-9999.

inside a polygon with k vertices. Aronov [1] gave an $O((k+n)\log(k+n)\log k)$ divide and conquer algorithm for this problem. Another possibility is to determine the visibility graph for the terminals and vertices of P. A low cost subtree of the visibility graph spanning the terminals can be determined by any of several available approximation algorithms for the Steiner tree problem in weighted graphs [6]. A small subset of terminals is considered as a reasonable cluster if the deletion of all other terminals leaves the subtree connected.

The paper is organized as follows. Some basic definitions are given in Section 2. The original problem is reduced in Section 3 to the ESTP for three terminals in a smaller polygon with some particular properties. The three terminals for the reduced problem need not to be identical with the original terminals. In Section 4, the vertices of the smaller polygon are preprocessed in order to determine shortest distances from all vertices to the three terminals. The algorithm for the determination of the ESMT in the reduced polygon is given in Section 5. Conclusions and suggestions for further research are collected in Section 6.

2 Basic Definitions

A polygon is ***simple*** if no point of the plane belongs to more than two edges of P and the only points of the plane that belong to precisely two edges are the vertices of P. A simple polygon P has a well-defined interior $i(P)$ and exterior $e(P)$. A point p is said to be in P if $p \in i(P) \cup P$. A vertex v on P is ***convex*** if its interior angle is less than 180^o. Otherwise it is said to be ***reflex***. The clockwise successor and successor vertices of a vertex v on P will be denoted by v^+ and v^-, respectively. A simple polygon is called a ***c-kite*** if precisely c of its vertices are convex. A polygon P is ***weakly-simple*** if any pair of polygonal chains obtained by partitioning P does not properly intersect (but the chains are permitted to touch each other).

The shortest path between two points u and v in a polygon P will be denoted by $P(u,v)$. $P(u,v)$ is a unique polygonal chain and its interior vertices are reflex vertices of P.

3 Polygon Reductions

Consider a triangulation of the simple polygon P. Remove from it (one at a time) outermost triangles (sharing only one side) provided that the remaining triangles contain all terminals. Let P_1 denote the polygon determined by the boundary of the exterior face of the reduced triangulation.

Lemma 1 *ESMT for t_i, t_j, t_k in P is in P_1.*

Proof. Assume that there is no ESMT completely in P_1. Let T denote the ESMT for the terminals in P. Let e be an edge of P_1 crossed by at least one edge of T. Let v_1^r denote the rightmost intersection of e with T (when looking from the interior of P_1). Follow the edge of T away from P_1. When reaching a vertex, take the rightmost edge. Continue in this manner until the edge e is reached again. Let v_1^l denote this cross-point. If there are additional intersections of e with T to the left of v_1^l, let v_2^r be the rightmost one among them. Determine v_2^l as above. Keep repeating this until all cross-points are covered by line segments $v_1^r v_1^l, v_2^r v_2^l, \ldots$. Replace the portions of T beyond e by line-segments $v_1^r v_1^l, v_2^r v_2^l, \ldots$. The resulting network is not longer than T, it spans all terminals, and it does not cross e. This process can be repeated for all edges of P_1 crossed by T, a contradiction. □

Consider a weakly-simple polygon P_2 consisting of the concatenation of the shortest paths $P_1(t_i,t_j)$, $P_1(t_j,t_k)$, $P_1(t_k,t_i)$ (Fig. 1).

Lemma 2 *ESMT for t_i, t_j, t_k in P is in P_2.*

Proof. There always exists an obstacle-avoiding ESMT within the minimum length perimeter enclosing all terminals (see [9]). This is valid for any number of terminals and any number of polygonal obstacles. In the case of 3 terminals inside a polygon P, $e(P) \cup P$ can be considered as a single obstacle. Obviously, shortest paths $P_1(t_i,t_j), P_1(t_j,t_k)$ and $P_1(t_k,t_i)$ define a weakly-simple polygon P_2 inside P_1 with the minimum length perimeter enclosing all terminals. □

If $i(P_2) = \emptyset$, then the unique ESMT for t_i, t_j, t_k in P overlaps with the edges of P_2. We therefore assume in the following that $i(P_2) \neq \emptyset$. Let P_3 denote the smallest polygon containing $i(P_2)$. Let q_u, $u = i, j, k$, denote a vertex on P_3 where the path to t_u (overlapping with P_2) begins. This path can be of zero-length, i.e., $q_u = t_u$. The vertex q_u will be referred to as a *semi-terminal*. Note that P_3 is a concatenation of the shortest paths $P_2(q_i, q_j)$, $P_2(q_j, q_k)$ and $P_2(q_k, q_i)$. Furthermore, P_2 contains P_3.

Rather than determining ESMT for t_i, t_j, t_k in P, one can look for an ESMT for q_i, q_j, q_k in P_3. It will together with $P_2(t_i, q_i)$, $P_2(t_j, q_j)$ and $P_2(t_k, q_k)$ yield ESMT for t_i, t_j, t_k in $P_2 \subseteq P_1 \subseteq P$.

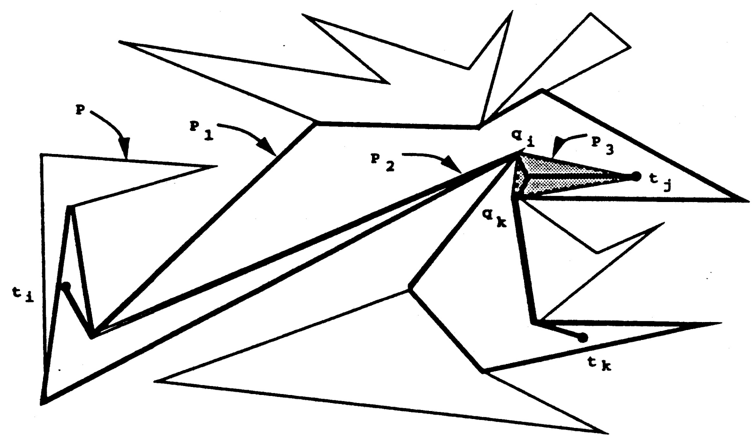

Figure 1: Nested polygons with the ESMT

Lemma 3 *P_3 is a 3-kite.*

Proof. All interior vertices on $P_3(q_i, q_j)$ must be reflex. Assume to the contrary that v is the first interior convex vertex on $P_3(q_i, q_j)$. There is a point v^* in the interior of the edge vv^+ such that the line segment $v^- v^*$ does not intersect P_3. Furthermore, $|v^- v^*| + |v^* v^+| < |v^- v| + |vv^+|$, a contradiction. By a similar argument, it follows that all interior vertices of $P_3(q_j, q_k)$ and $P_3(q_k, q_i)$ must be reflex. Since every simple polygon must have at least 3 convex vertices, it follows that q_i, q_j, q_k are all convex vertices. □

A line L is said to be an ***interior tangent*** of a c-kite P at a ***touch-point*** $v \in P$ iff one of the following cases occurs.

- v is reflex and the two edges of P incident with v are on the same side of L.
- v is convex and the two edges of P incident with v are on the opposite sides of L.
- v is an interior point of an edge of P and this edge overlaps with L.

Lemma 4 *Every c-kite P, $c \geq 3$, has exactly $c - 2$ interior tangents for any fixed slope $r \in [-\infty, \infty]$. In particular, a 3-kite has exactly 1 interior tangent for any fixed slope.*

Proof. The sum of interior angles of any simple polygon is $(n-2)\pi$. This follows from the fact that every triangulation of a simple polygon has $n - 2$ triangles. Each triangle contributes to the total sum of angles by π.

Let α_i, $1 \le i \le c$, denote interior angles of convex vertices. Let $\beta_j + \pi$, $1 \le j \le n - c$ denote interior angles of reflex vertices. Then

$$\sum_{i=1}^{c} \alpha_i + \sum_{j=1}^{n-c} \beta_j = (n-2)\pi - (n-c)\pi = (c-2)\pi$$

Since α_i and β_j denote maximal rotation of interior tangents at respectively convex and reflex vertices, and the slope interval at a particular vertex does not overlap but has a common point with slope interval of next vertex on the polygon. □

4 Preprocessing of 3-Kites

We need to associate shortest distances $d_i(v), d_j(v), d_k(v)$ between every vertex v on a 3-kite P_3 and semi-terminals q_i, q_j, q_k (Fig. 2). The distance $d_i(v)$ for every vertex v on $P_3(q_i, q_j)$ is determined by the following procedure.

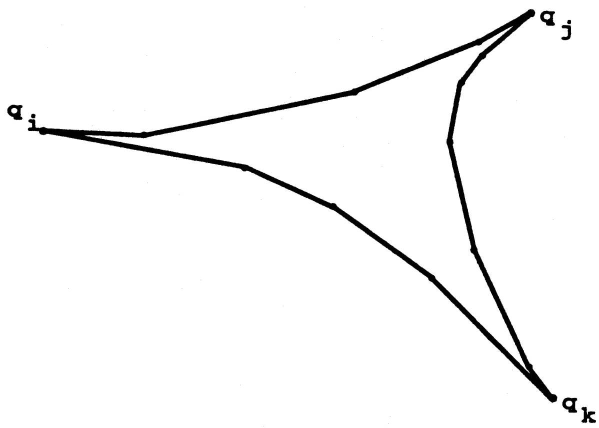

Figure 2: 3-kite P_3

$v := q_i$; $d_i(v) := 0$;
repeat
  $v := v^+$;
  $d_i(v) := d_i(v^-) + |v^- v|$;
until $v = q_j$;

Naturally, $d_j(v)$ for every vertex v on $P_3(q_i, q_j)$ is determined by a procedure traversing $P_3(q_i, q_j)$ in the opposite direction. Similar procedures determine $d_j(v)$ and $d_k(v)$ for every vertex v on $P_3(q_j, q_k)$ as well as $d_k(v)$ and $d_i(v)$ for every vertex v on $P_3(q_k, q_i)$.

In order to determine $d_k(v)$ for every vertex v on $P_3(q_i, q_j)$, initialize $d_k(v)$ to $|vq_k|$ for all $v \in P_3(q_i, q_j)$. Values of $d_k(v)$ for vertices of $P_3(q_i, q_j)$ not visible from q_k must then be updated. This is done in two scans of $P_3(q_i, q_j)$. First, the vertices of $P_3(q_i, q_j)$ not visible from q_k are scanned from q_i toward q_j. Next, the vertices of $P_3(q_i, q_j)$ not visible from q_k are scanned from q_j toward q_i. Only the first scan is described as they are; analogous. Let `LeftTurn`(v, q, q_k) be a predicate that returns true iff points v, q, q_k make a left turn at q. In particular, if these three points are colinear (or two of them overlap), then `LeftTurn`(v, q, q_k) = `False`. It is well-known that `LeftTurn` can be implemented so that it requires $O(1)$ time [8].

```
v := q_i^+; q := q_i;
repeat
   while LeftTurn(v, q, q_k) do q := q^-;
   d_k(v) := |vq| + d_k(q); v := v^+;
until q = q_k;
```

Similar two-way scans are applied to determine $d_i(v)$ and $d_j(v)$ for all vertices v on $P_3(q_j, q_k)$ and $P_3(q_k, q_i)$ respectively.

5 ESMT for Semi-Terminals

The determination of the ESMT for the semi-terminals q_i, q_j, q_k in the 3-kite P_3 can be divided into 2 separate cases by assuming that the unique Steiner point s is located on P_3 (in particular, it can overlap with one of the semi-terminals) or in $i(P_3)$. The best solutions for each of these two cases are compared, and the better one is the overall ESMT.

The first case is simple. The Steiner point s cannot be in the interior of an edge of P_3; two of the three edges incident with s meet at less than 120^o. Suppose that s overlaps with a vertex v. Shortest paths from v to the semi-terminals have total length $d_i(v) + d_j(v) + d_k(v)$.

The second case is more interesting. Three chains of the ESMT incident with the Steiner point s in the interior of P_3 meet at 120^o. Each of these chains touches P_3, and then follows P_3 until reaching semi-terminals q_i, q_j, q_k. Let sv_i, sv_j, sv_k denote the edges incident with s.

Lemma 5 *The points v_i, v_j, v_k are vertices of P_3. Furthermore, the edges sv_i, sv_j, sv_k overlap with interior tangents at v_i, v_j, v_k, respectively.*

Proof. Suppose that v_i is not a vertex of P_3. The chain from s through v_i in the ESMT must continue along the edge containing v_i. At least a portion of this path in the neighborhood of v_i must be visible from a point s^* on sv_i. Let v_i^* denote a point in this neighborhood. Then $|s^*v_i| + |v_iv_i^*| > |s^*v_i^*|$, a contradiction.

Suppose that the edge sv_i is not overlapping with an interior tangent at v_i. If v_i is a convex vertex, then sv_i is not completely in P_3, a contradiction. Assume that v_i is reflex. By an argument similar to that in the first part of the proof, a shorter network exists, a contradiction. □

Lemma 5 suggests a trivial $O(n^3)$ algorithm for a shortest tree spanning q_i, q_j, q_k with a Steiner point s in $i(P_3)$. Construct a non-degenerate Steiner tree for each triple v_i, v_j, v_k of vertices of P_3 (or decide that it does not exist). Check for intersections with P_3. If no intersections, connect each q_i, q_j, q_k to the closest point among v_i, v_j, v_k. Take the shortest of such trees (if any).

Instead of the exhaustive enumeration of all triples of vertices, one only needs to consider triples of vertices that admit interior tangents making 120^o with each other.

- **Initialization:** Let v_i denote the vertex of P_3 touched by the vertical interior tangent L_i. If L_i overlaps with an edge of P_3, let v_i be its first end-point (when P_3 is traversed in clockwise direction). Traverse the vertices of P_3 in clockwise direction, beginning at v_i, until reaching a vertex v_j which admits an interior tangent L_j making 120^o with L_i. Continue the clockwise traversal of P_3 until reaching a vertex v_k which admits an interior tangent L_k making 120^o with L_j.
- **Iteration:** Determine a non-degenerate Steiner tree for v_i, v_j, v_k (or decide that it does not exist). If the edges overlap with interior tangents at v_i, v_j, v_k, connect each semi-terminal q_i, q_j, q_k to the closest of vertices v_i, v_j, v_k. Save the tree found if its length is less than the length of the best solution found so far (if any).
- **Sweep:** Interior tangents L_i, L_j, L_k are rotated (counterclockwise) around their touch-points v_i, v_j, v_k until one overlaps with an edge of P_3. Such rotations amount to finding the smallest angle between

interior tangents L_i, L_j, L_k and edges $v_iv_i^+, v_jv_j^+, v_kv_k^+$. Suppose that L_i is the first interior tangent overlapping with an edge. Replace v_i by v_i^+.

- **Termination** The circular sweep terminates when the interior tangents have been rotated at least 120^o; otherwise perform next **Iteration.**

Theorem 1 *The ESMT for three terminals in a simple polygon can be determined in $O(n)$ time and space.*

Proof. To prove the correctness of the algorithm for the ESMT for three semi-terminals q_i, q_j, q_k in P_3, we only need to observe that all triples of vertices on P_3 which admit interior tangents meeting at 120^o will be generated during the circular scan. The correctness of the ovarall algorithm then follows immediately.

Linear time complexity is quite straightforward to establish. Triangulation of P can be done in linear time [4]. Determination of P_1 from P involves the deletion of vertices of degree 2 provided that the remaining triangulation contains all terminals. This can be done in $O(n)$ time. Shortest paths between the three terminals can be determined in linear time [3, 7]. Hence, P_2 can be obtained from the triangulated P_1 in linear time. Determination of P_3 from P_2 is trivial. Distances from all vertices of the kite to each semi-terminal q_i, q_j, q_k can be found in $O(n)$ time. When the Steiner point s overlaps with one of the vertices, $O(n)$ time is needed to find the solution. In the case when s is in $i(P_3)$, the circular sweep generates as many triples as there are vertices on the 3-kite. For each triple, the corresponding Steiner tree can be determined in constant time. Also constant time is needed to check if the edges of the Steiner tree overlap with interior tangents of P_3. Finally, it is obvious that the space complexity of the algorithm is $O(n)$. □

6 Conclusions

We presented an $O(n)$ time and space algorithm for the ESTP for three terminals inside a simple polygon with n vertices. There is a number of interesting issues that remain open or will be addressed in companion papers. One is to obtain an efficient time and space algorithm for the same problem but with four terminals[12]. Another problem is how to preprocess the polygon and any number of terminals in order to identify reasonable clusters of up to four terminals. This would lead to a general heuristic for the ESTP for any number of terminals inside a polygon. Such a heuristic would be based on the concatenation of small ESMTs. Determination of ESMTs for small subsets of terminals in presence of several (convex) obstacles inside a simple polygon is also of interest. In this context, Steiner visibility graphs introduced in [11] could prove useful. Finally, we mention the problem of preprocessing a simple polygon so that three and/or four terminals querries for ESMTs can be answered efficiently.

References

[1] B. Aronov, On the geodesic Voronoi diagram of point sites in a simple polygon, *Algorithmica* 4 (1990) 109-140.

[2] J.E. Beasley, A heuristic for the Euclidean and rectilinear Steiner tree problem, Working paper, Management School, Imperial College, London (1989).

[3] B.M. Chazelle, A theorem on polygon cutting with applications, *Proc. of the 23-th IEEE Symp. on Foundations of Computer Science* (1982) 339-349.

[4] B.M. Chazelle, Triangulating a simple polygon in linear time, *Disc. Comp. Geom.* 6 (1991) 485-524.

[5] L.J. Guibas and J. Stolfi, Primitives for the manipulation of general subdivisions and the computation of Voronoi diagrams, *ACM Trans. on Graphics* 4 (1985) 74-123.

[6] F.K. Hwang, D.S. Richards and P. Winter. *The Steiner Tree Problem*, North-Holland (1992).

[7] D.T. Lee and F.P. Preparata, Euclidean shortest paths in the presence of rectilinear barriers. *Networks* **14** (1984) 393-410.

[8] F.P. Preparata and M.I. Shamos, *Computational Geometry: An Introduction*, Springer, Berlin (1985).

[9] J.S. Provan, An approximation scheme for finding Steiner trees with obstacles. *SIAM J. Comput.* **17** (1988) 920-934.

[10] J. MacGregor Smith, D.T. Lee and J.S. Liebman, An $O(n \log n)$ heuristic for the Steiner minimal tree problem on the Euclidean metric, *Networks* **11** (1981) 23-29.

[11] P. Winter, Euclidean Steiner minimal trees with obstacles and Steiner visibility graphs, *Discrete Applied Mathematics* **47** (1993) 187-206.

[12] P. Winter, Euclidean Steiner minimal trees for 4 terminals in a simple polygon, submitted for publication.

[13] P. Winter and J. MacGregor Smith, Steiner minimal tree for three points with one convex polygonal obstacle, *Ann. Oper. Res.* **33** (1991) 577-599.