

Geometric spanners with few edges and degree five

Michiel Smid

School of Computer Science
Carleton University
Ottawa, Ontario, Canada K1S 5B6
E-mail: michiel@scs.carleton.ca

Abstract

An $O(n \log n)$ -time algorithm is presented that, when given a set S of n points in \mathbb{R}^d and an integer k with $0 \leq k \leq n$, computes a graph with vertex set S , that contains at most $n - 1 + k$ edges, has stretch factor $O(n/(k+1))$, and whose degree is at most five. This generalizes a recent result of Aronov *et al.*, who obtained this result for two-dimensional point sets.

Keywords: Computational geometry, spanners, minimum spanning trees.

1 Introduction

Given a set S of n points in \mathbb{R}^d and a real number $t \geq 1$, a graph G with vertex set S is called a t -spanner for S , if for any two points p and q in S , there exists a path in G between p and q whose length is at most t times the Euclidean distance $|pq|$ between p and q . Here, the length of a path is defined to be the sum of the Euclidean lengths of all edges on the path. A path in G between p and q whose length is at most $t|pq|$ is called a t -spanner path. The *stretch factor* (or *dilation*) of G is defined to be the smallest value of t for which G is a t -spanner.

The problem of constructing a t -spanner for any given point set has been studied intensively. Salowe (1991) and Vaidya (1991) were the first to show that, for any constant $t > 1$ and for any constant dimension $d \geq 2$, a t -spanner with $O(n)$ edges can be computed in $O(n \log n)$ time, where the constant factors in the Big-Oh bounds depend on the stretch factor t and the dimension d . Since then, many more algorithms have been discovered that compute spanners with $O(n)$ edges and that have other properties; see the survey papers by Eppstein (2000), Gudmundsson and Knauer (2006), and Smid (2000).

Das and Heffernan (1996) considered a dual version of the spanner problem: Given a bound on the number of edges, what is the smallest stretch factor that can be obtained? They present an $O(n \log n)$ -time algorithm that constructs, when given any set S of n points in \mathbb{R}^d and any real constant $\epsilon > 0$, a graph that contains at most $(1 + \epsilon)n$ edges, whose degree is

at most three, and whose stretch factor is bounded by a constant that only depends on ϵ and d . Das and Heffernan left open the problem of determining the smallest possible stretch factor when $n + o(n)$ edges are allowed.

Since any graph with a finite stretch factor is connected, it must have at least $n - 1$ edges. Let S be a set of n points in the plane that are regularly spaced around a circle. Eppstein (2000) shows that every connected graph with vertex set S and consisting of $n - 1$ edges (i.e., every spanning tree of S) has stretch factor $\Omega(n)$. He also shows that, for any set S of n points in \mathbb{R}^d , the minimum spanning tree has stretch factor $O(n)$; in fact, the stretch factor can easily be shown to be at most $n - 1$; see Lemma 2 below.

Aronov *et al.* (2005) generalize these results for the case when the points are in \mathbb{R}^2 . They show that, for any set S of n points in the plane and for any integer k with $0 \leq k \leq n$, in $O(n \log n)$ time, a graph with vertex set S and consisting of $n - 1 + k$ edges can be computed, whose stretch factor is $O(n/(k+1))$. They also show that there exists a set S of n points, such that every connected graph with vertex set S and consisting of $n - 1 + k$ edges has stretch factor $\Omega(n/(k+1))$.

The algorithm of Aronov *et al.* is based on properties of the minimum spanning tree and the Delaunay triangulation. In particular, it uses the facts that, for two-dimensional point sets, (i) these structures can be computed in $O(n \log n)$ time, (ii) the stretch factor of the Delaunay triangulation is bounded by a constant, and (iii) the Delaunay triangulation is a planar graph. As a result, their analysis is only valid for two-dimensional point sets: First, in dimension $d \geq 3$, it is unlikely that the minimum spanning tree can be computed in $O(n \log n)$ time; see Erickson (1995). Second, for $d \geq 3$, no non-trivial upper bound on the stretch factor of the Delaunay triangulation is known. Finally, again for $d \geq 3$, the Delaunay triangulation is not a planar graph; in particular, it may have $\Theta(n^2)$ edges.

In this paper, we show that, by using a minimum spanning tree of a bounded degree spanner for S (as opposed to a minimum spanning tree of the point set itself), the result of Aronov *et al.* is in fact valid for any constant dimension $d \geq 2$. Moreover, we show that this result can be obtained by a graph having degree five.

2 Properties of the minimum spanning tree of a spanner

Let S be a set of n points in \mathbb{R}^d , let $t \geq 1$ be a real number, and let G be an arbitrary t -spanner for S . Let T be a minimum spanning tree of G . In this section, we prove some properties of T that will lead to our generalization of the result by Aronov *et al.*

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Copyright 2006, Australian Computer Society, Inc. This paper appeared at Computing: The Australasian Theory Symposium (CATS2006), Hobart, Australia. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 51. Barry Jay and Joachim Gudmundsson, Eds. Reproduction for academic, not-for profit purposes permitted provided this text is included.

These properties basically state that T has “approximately” the same properties as an exact minimum spanning tree of the point set S .

Lemma 1 *Let p and q be two distinct points of S . Then every edge on the path in T between p and q has length at most $t|pq|$.*

Proof. Let P be the path in T between p and q , and let (x, y) be an arbitrary edge on P . We will prove by contradiction that $|xy| \leq t|pq|$. Hence, we assume that $|xy| > t|pq|$.

Let Q be a t -spanner path in G between p and q . Since the length of Q is at most $t|pq|$, every edge of Q has length at most $t|pq|$. In particular, (x, y) is not an edge of Q . We may assume without loss of generality that x is between p and y on the path P . Starting at x , follow the path P towards p , and let x' be the first vertex that is on Q . Similarly, starting at y , follow the path P towards q , and let y' be the first vertex that is on Q . Let P' be the subpath of P between the vertices x' and y' , and let Q' be the subpath of Q between the vertices x' and y' . Then, P' and Q' do not have any edge in common, and these two subpaths form a simple cycle in G that contains the edge (x, y) .

Let G' be the graph obtained from T , by adding all edges of Q' (that are not in T yet), and deleting the edge (x, y) . Then G' is a connected subgraph of G on the point set S , and, since the weight of Q' is less than the weight of (x, y) , the weight of G' is less than the weight of T . This is a contradiction and, thus, we have shown that $|xy| \leq t|pq|$. ■

Lemma 2 *The minimum spanning tree T of the t -spanner G is a $(t(n-1))$ -spanner for S .*

Proof. Let p and q be two distinct points of S , and let P be the path in T between p and q . By Lemma 1, each edge of P has length at most $t|pq|$. Since P contains at most $n-1$ edges, it follows that the length of P is at most $t(n-1)|pq|$. ■

Lemma 3 *Let m be an integer with $1 \leq m \leq n-1$, and let T' and T'' be two vertex-disjoint subtrees of T , each consisting of at most m vertices. Let p be a vertex of T' , let q be a vertex of T'' , and let P be the path in T between p and q . If x is a vertex of T' that is on the subpath of P within T' , and y is a vertex of T'' that is on the subpath of P within T'' , then*

$$|xy| \leq (2t(m-1) + 1)|pq|.$$

Proof. Let P' be the subpath of P between p and x . By Lemma 1, each edge of P' has length at most $t|pq|$. Since P' contains at most $m-1$ edges, it follows that this path has length at most $t(m-1)|pq|$. On the other hand, since P' is a path between p and x , its length is at least $|px|$. Thus, we have $|px| \leq t(m-1)|pq|$. A symmetric argument can be used to show that $|qy| \leq t(m-1)|pq|$. Therefore, we have

$$\begin{aligned} |xy| &\leq |xp| + |pq| + |qy| \\ &\leq t(m-1)|pq| + |pq| + t(m-1)|pq|, \end{aligned}$$

completing the proof of the lemma. ■

3 A graph with $n + O(k)$ edges and stretch factor $O(n/k)$

Let S be a set of n points in \mathbb{R}^d , and let k be an integer with $1 \leq k \leq n$. Fix a constant $t > 1$, and let G be a

t -spanner for S whose degree is bounded by a constant that only depends on the dimension d . Clearly, the minimum spanning tree T of G has bounded degree as well. Thus, T contains a *centroid edge*, i.e., an edge whose removal from T yields two subtrees, each consisting of at most αn vertices, for some constant $\alpha < 1$ that depends on the degree of T . In fact, a centroid edge can be computed in $O(n)$ time. By repeatedly choosing a centroid edge in the currently largest subtree, we can remove $\ell = O(k)$ edges from T , and obtain vertex-disjoint subtrees T_0, T_1, \dots, T_ℓ , each containing $O(n/k)$ vertices. Observe that the vertex sets of these subtrees form a partition of S . Let X be the set of endpoints of the ℓ edges that are removed from T . Then, the size of X is at most 2ℓ , which is $O(k)$.

We define G' to be the graph with vertex set S that is the union of

1. the trees T_0, T_1, \dots, T_ℓ , and
2. a t -spanner G'' for the set X , consisting of $O(k)$ edges.

We first observe that the number of edges of G' is bounded from above by $n - 1 + O(k)$.

Lemma 4 *The graph G' has stretch factor $O(n/k)$.*

Proof. Let p and q be two distinct points of S . Let i and j be the indices such that p is a vertex of the subtree T_i and q is a vertex of the subtree T_j .

First assume that $i = j$. Let P be the path in T_i between p and q . Then, P is a path in G' . By Lemma 1, each edge on P has length at most $t|pq|$. Since T_i contains $O(n/k)$ vertices, the number of edges on P is $O(n/k)$. Therefore, since t is a constant, the length of P is $O(n/k) \cdot |pq|$.

Now assume that $i \neq j$. Let P be the path in T between p and q . Let (x, x') be the edge of P for which x is a vertex of T_i , but x' is not a vertex of T_i . Similarly, let (y, y') be the edge of P for which y is a vertex of T_j , but y' is not a vertex of T_j . Then, both (x, x') and (y, y') are edges of T that have been removed when the subtrees were constructed. Hence, x and y are both contained in X and, therefore, are vertices of G'' . Let P_i be the path in T_i between p and x , let P_{xy} be a t -spanner path in G'' between x and y , and let P_j be the path in T_j between y and q . The concatenation Q of P_i , P_{xy} , and P_j is a path in G' between p and q .

Since both P_i and P_j are subpaths of P , it follows from Lemma 1 that each edge on P_i and P_j has length at most $t|pq|$. Since T_i and T_j contain $O(n/k)$ vertices, it follows that the sum of the lengths of P_i and P_j is $O(n/k) \cdot |pq|$. The length of P_{xy} is at most $t|xy|$ which, by Lemma 3, is also $O(n/k) \cdot |pq|$. Thus, the length of Q is $O(n/k) \cdot |pq|$. ■

We take for G the t -spanner of Das and Heffernan (1996). This spanner can be computed in $O(n \log n)$ time, and each vertex has degree at most three. Given G , its minimum spanning tree T can be computed in $O(n \log n)$ time. Since a centroid edge can be computed in $O(n)$ time, the subtrees T_0, T_1, \dots, T_ℓ can be computed in $O(n \log n)$ time. Finally, we take for G'' the t -spanner of Das and Heffernan. This spanner G'' can be computed in $O(k \log k) = O(n \log n)$ time, and each vertex has degree at most three.

For these choices of G and G'' , the graph G' has stretch factor $O(n/k)$, it contains $n - 1 + O(k)$ edges, and it can be computed in $O(n \log n)$ time. We analyze the degree of G' : Consider any vertex p of G' . If $p \notin X$, then the degree of p in G' is equal to the degree of p in T , which is at most three. Assume that

$p \in X$. The graph G'' contains at most three edges that are incident to p . Similarly, the tree T contains at most three edges that are incident to p , but, since $p \in X$, at least one of these three edges is not an edge of G' . Therefore, the degree of p in G' is at most five. Thus, each vertex of G' has degree at most five.

Let c be a constant such that the graph G' contains at most $n - 1 + ck$ edges.

4 The main result

We are now ready to prove the main result of this paper. Let S be a set of n points in \mathbb{R}^d , and let k be an integer with $0 \leq k \leq n$. Consider the constant c that was introduced above.

First assume that $k < c$. Let G be a t -spanner for S , for some constant t , in which each vertex has degree at most three, and let G' be a minimum spanning tree of G . Then, G' has $n - 1 \leq n - 1 + k$ edges, degree at most three and, by Lemma 2, the stretch factor of G' is at most $t(n - 1)$, which is $O(n/(k + 1))$.

If $c \leq k \leq n$, then we apply the results of Section 3 with k replaced by k/c . This gives a graph G' with at most $n - 1 + k$ edges, degree at most five, and stretch factor $O(n/(k + 1))$. Thus, we have proved the following result:

Theorem 1 *Let S be a set of n points in \mathbb{R}^d , and let k be an integer with $0 \leq k \leq n$. In $O(n \log n)$ time, a graph with vertex set S can be computed that has the following properties:*

1. *The graph contains at most $n - 1 + k$ edges.*
2. *The graph has stretch factor $O(n/(k + 1))$.*
3. *Each vertex of the graph has degree at most five.*

Aronov *et al.* (2005) gave an example of a set S of n points in the plane, such that every connected graph with vertex set S and consisting of $n - 1 + k$ edges has stretch factor $\Omega(n/(k + 1))$. Therefore, the result in Theorem 1 is optimal. An interesting open problem is whether the degree can be reduced.

Acknowledgements

The author would like to thank the members of the Algorithms Seminar Group at Carleton University for pointing out a serious error in a previous version of this paper.

References

- Aronov, B., de Berg, M., Cheong, O., Gudmundsson, J., Haverkort, H. & Vigneron, A. (2005), Sparse geometric graphs with small dilation, in 'Proceedings of the 16th International Symposium on Algorithms and Computation', Lecture Notes in Computer Science, Springer-Verlag, Berlin.
- Das, G. & Heffernan, P. J. (1996), 'Constructing degree-3 spanners with other sparseness properties', *International Journal of Foundations of Computer Science* **7**, 121–135.
- Eppstein, D. (2000), Spanning trees and spanners, in J.-R. Sack & J. Urrutia, eds, 'Handbook of Computational Geometry', Elsevier Science, Amsterdam, pp. 425–461.
- Erickson, J. (1995), On the relative complexities of some geometric problems, in 'Proceedings of the 7th Canadian Conference on Computational Geometry', pp. 85–90.

Gudmundsson, J. & Knauer, C. (2006), Dilation and detours in geometric networks, in T. F. Gonzalez, ed., 'Handbook on Approximation Algorithms and Metaheuristics', Chapman & Hall/CRC, Boca Raton.

Salowe, J. S. (1991), 'Constructing multidimensional spanner graphs', *International Journal of Computational Geometry & Applications* **1**, 99–107.

Smid, M. (2000), Closest-point problems in computational geometry, in J.-R. Sack & J. Urrutia, eds, 'Handbook of Computational Geometry', Elsevier Science, Amsterdam, pp. 877–935.

Vaidya, P. M. (1991), 'A sparse graph almost as good as the complete graph on points in K dimensions', *Discrete & Computational Geometry* **6**, 369–381.