Rank of the Longest Edge in the Greedy Triangulation

Matthew T. Dickerson*
Department of Mathematics and Computer Science
Middlebury College

April 8, 1992

Abstract

For a uniform distribution of n points in a disc, we show that the expected value of the rank of the longest edge in the greedy triangulation is $\Theta(n^{4/3})$. Based on this result, on the triangulation compatibility test method of Gilbert [3], and on the enumeration algorithm of Dickerson and Drysdale [2], we then give a new greedy triangulation algorithm which for uniform distributions in a disc requires $O(n^{4/3} \log n)$ time and $O(n^{4/3})$ space in the expected case.

1 Introduction

Given a set S of n points, the *rank* of x (written $r(x)$) is the number of pairs of points in S separated by a distance of x or less. Formally:

$$r(x) = \#\{(u,v) : d(u,v) \leq x; u,v \in S; u \neq v\}$$

In this paper, we prove the following theorems:

Theorem 1 *Let S be a set of n points chosen in a uniform random distribution on a disc. Let x be the length of the longest edge in the greedy triangulation of S. The expected value of $r(x)$ is $\Theta(n^{4/3})$.*

Theorem 2 *Let S be a set of n points chosen in a uniform random distribution from the interior of a convex r-gon. Let x be the longest edge in the GT of S. The expected value of $r(x)$ is $\Theta(n^2)$.*

Based on Theorem 1, we desribe a new algorithm for computing the greedy triangulation of a set of n points in the plane. For a uniform random distribution in a disc, the algorithm requires $O(n^{4/3} \log n)$ time and $O(n^{4/3})$ space in the expected case, which is asymptotically faster than the best previously known greedy triangulation algorithms of [4] and [10] with only a slight ($O(n^{1/3})$ factor) additional cost in storage. Our algorithm makes use of the enumeration method of Dickerson and Drysdale [2] and the triangulation compatibility test of Gilbert [3].

1.1 Background

Triangulating a set of points is a problem of extreme importance in computatinal geometry, and the applications of triangulations to other fields are far too numerous to list. (See sections 5.1 and 6.2 of [11]). For many applications, a *minimum weight triangulation* (MWT) is most desirable. Unfortunately, although it has been shown how to compute the MWT for convex polygons in $O(n^2)$ time [3,7], no efficient algorithm is known for computing the MWT in the general case. Efficiently computable *approximations* to the MWT are therefore sought. One such possibility is the *Greedy Triangulation* (GT). The Greedy Triangulation method adds at each stage the shortest possible edge that is compatible with previously added edges. A number of the properties of the GT have been discussed [6,7,9,10]. For example, although neither the GT nor the Delaunay triangulation (DT) yields the MWT [5,8], the GT appears to be the better of the two at approximating it. In fact, for convex polygons the GT approximates the MWT to with a constant factor [6].

*This study was supported in part by the funds of National Science Foundation, RII-8610679.

A straightforward approach to computing the greedy triangulation is to compute all $\binom{n}{2}$ distances, sort them, and then build the GT an edge at a time by examining each pair in order of length and adding or discarding it based on its compatibility with the edges already added. It is easy to see that this method requires $O(n^2)$ space and $O(n^2 \log n + n^2\psi(n))$ time where $\psi(n)$ is the time required to test for compatibility [11]. A naive test would compare each new potential edge to each of the existing edges (of which there are at most $O(n)$) for an $O(n^3)$ time algorithm. Gilbert [3] presented a new structure and algorithm for a $O(\log n)$ time compatibility test, thus improving the algorithm's overall time complexity to $O(n^2 \log n)$, without affecting space complexity. Manacher and Zobrist [10] have since given an $O(n^2)$ expected time and $O(n)$ space greedy triangulation algorithm which makes use of a probabilistic method for *pretesting* compatibility of new edges; Goldman [4] has provided an $O(n^2 \log n)$ worst case time, $O(n)$ space algorithm which makes use of generalized Delaunay triangulations; and for the special case of n vertex convex polygons, Lingas [7] has presented an $O(n^2)$ time and $O(n)$ space algorithm.

Another interesting class of problems is that of determining the expected *rank* of certain interpoint distances. Salowe [13] has a number of nice results in this area, including: 1) relationships between the rank of a particular distance x and the rank of a distance $2x$, and 2) relationships between the ranks $r_\infty(x)$ and $r_p(x)$ for the L_∞ and L_p metrics.

In this paper, we are interested in the following problem dealing with ranks of edges in the GT:

Problem 1 *Let S be a set of n points in the plane, and x the length of the longest edge in the greedy triangulation. What is the expected value of $r(x)$?*

This problem is particularly interesting in light of recent results from Dickerson and Drysdale [2] and Salowe [13] on the problem of shallow interdistance enumeration. Dickerson and Drysdale have presented an $O(n \log n + k \log n)$ time algorithm for enumerating in sorted order the k smallest of $\binom{n}{2}$ interdistances for n points on the plane. And Salowe has presented an $O(n \log n + k)$ time algorithm for enumerating (not necessarily in sorted order) the k smallest distances for n points in d dimensions. Let k be the rank of the longest edge in the Greedy Triangulation. With an efficient enumeration algorithm and the $O(\log(n))$ time edge test method of Gilbert [3], it is possible to contruct the GT in $O(n \log n + k \log n)$ time and $O(k)$ space. (We will discuss this method in slightly more detail in a later section.) If k is $o(n^2/\log n)$, then we have a new algorithm for computing the GT which is asymptotically faster than the previously best known GT algorithm of Manacher and Zobrist. As was stated in Theorem 1, this proves to be the case for uniform distributions on a disc.

2 Rank for a Uniform Distribution on a Disc

In this section we prove Theorem 1. The proof comes in two parts. In Theorem 3 we give the expected length x of the longest edge in a greedy triangulation for a uniform distribution on a unit disc. Then in Theorem 4 we give the expected rank of this distance x. We begin with Theorem 3.

Theorem 3 *Let S be a set of n points chosen in a uniform random distribution on a unit disc. The expected length of the longest edge in the greedy triangulation of S is $\Theta(1/n^{1/3})$.*

Proof of Theorem 3 The proof of this theorem comes in three parts. We first give the expected length x of the longest edge on the convex hull. We then show that with probability 0 as $n \to \infty$ all greedy edges sufficiently far from the hull will be shorter than x, implying that the longest edge is on or near the hull. Finally, we deal with edges near but not on the hull.

We begin with the expected *average* length of a convex hull edge. Assume a unit disc. As $n \to \infty$, the convex hull CH of S approximates the disc, and thus the circumference of CH approaches 2π. For a uniform distribution in a disc, the expected number of points on the convex hull is $\Theta(n^{1/3})$ [12]. The expected average length of a hull edge is therefore $c/n^{1/3}$ for some constant c. The following lemma states that the expected length of the longest edge in the convex hull is of the same order.

Lemma 1 *Let S be a set of n points chosen in a uniform random distribution on a unit disc, and let CH be the convex hull of S. For any positive constant ϵ, the probability of there being an edge in CH of length $\Omega(1/n^{1/3-\epsilon})$ approaches 0 as $n \to \infty$.*

proof Consider a polar angle from the center of the disc to each convex hull point. The points in S form a uniform continuous distribution on the disc, and so the polar angles to the points of CH also form a uniform continuous distribution from 0 to 2π. Let u and v be two points of CH, but with (u,v) not necessarily an edge of CH. Then (u,v) is an edge of CH if and only if no other points of CH between u and v. On a unit disc, for points u and v near the circumference of the disc, the polar angle between u and v will be greater than $d(u,v)$. Let w be a random point from CH, other than u and v. It follows that the probability of w falling between u and v is greater than $d(u,v)/2\pi$, and conversely the probability q' of w not falling between u and v is less than $1-d(u,v)/2\pi$. Therefore, the probability that *no* convex hull point falls between u and v, and thus (u,v) is an edge in CH, is less than $(1-\frac{d(u,v)}{2\pi})^h$ where h is the number of points on CH.

Substituting $cn^{1/3}$ for h and letting $d(u,v) = a'/n^{1/3-\epsilon}$ for some constant a', we solve in the limit for q':

$$\begin{aligned} q' &\leq \lim_{n\to\infty}(1-\frac{a}{n^{1/3-\epsilon}})^{cn^{1/3}} \\ &= e^{-acn^{\epsilon}} \end{aligned} \tag{1}$$

where $a = a'/2\pi$. For some particular pair of points (u,v) on CH, with $d(u,v) = a'/n^{1/3-\epsilon}$, we now have the probability q' that (u,v) is a convex hull edge, and the probability $1-q'$ that (u,v) is not an edge of CH. Now we must find the probability q that there is *no* edge of length $\Omega(1/n^{1/3-\epsilon})$ on CH, or equivalently that *every* pair of points separated by a distance of $\Omega(a'/n^{1/3-\epsilon})$ is *not* an edge of CH. The expected number of *pairs* of hull point is given by $\binom{h}{2}$ or $kn^{2/3}$ for some new constant k. We can now give an equation for q and solve in the limit using L'Hopital's rule.

$$\begin{aligned} q &= (1-q')^{kn^{2/3}} \\ &= e^{h(n)} \end{aligned} \tag{2}$$

where

$$\begin{aligned} h(n) &= \lim_{n\to\infty} \frac{\ln(1-e^{-acn^{\epsilon}}}{kn^{-2/3}}) \\ &= \lim_{n\to\infty}(\frac{-3ac\epsilon}{2})(\frac{1}{1-e^{-acn^{\epsilon}}})(\frac{kn^{2/3+\epsilon}}{e^{acn^{\epsilon}}}) \\ &= 0 \end{aligned} \tag{3}$$

for any positive constant ϵ. This gives us $q = e^0 = 1$ is the probability that no pair of point separated by a distance of $\Omega(a'/n^{1/3-\epsilon})$ is an edge of CH. Conversely, the probability that there *is* a long convex hull edge is 0, and the proof of Lemma 1 is complete. □

We now examine edges in the interior of GT (edges sufficiently far from the Convex Hull) and show that the probability of there being an interior GT edge of length $\Omega(1/n^{1/3})$ goes to 0. The implication of this claim is that the longest edge in a Greedy Triangulation will be on (or very near) the hull. This is an intuitively satisfying result. Every convex hull edge, regardless of its length, must be an edge in every triangulation. Furthermore, as pointed out in [10], hull edges are larger than typical interior edges, and triangles near the hull are in general thinner. Nonetheless, it is not a trivial point and a proof is necessary.

Lemma 2 *Let S be a set of n points chosen in a uniform random distribution on a unit disc, and let GT be a greedy triangulation of S. As $n\to\infty$, the probability of there existing an edge in GT of length at least $c/n^{1/3}$ which is a distance of at least $c/n^{1/3}$ from the convex hull approaches* 0 .

proof Let $u,v \in S$ be a pair of points such that $d(u,v) \geq c/n^{1/3}$ and the distance from (u,v) to the convex hull of S is also at least $c/n^{1/3}$. Consider the disc C_{uv} with segment $\overline{uv}$ as its diameter. C_{uv} is cut into two half-discs by $\overline{uv}$. If both of these half-discs contain a point from S, then (u,v) is not a GT edge because there is a shorter edge intersecting $\overline{uv}$ that will be chosen first. Since the entire disc C_{uv} falls within our unit disc, the probability of a particular random point from S falling in a particular one these half-discs is proportional to the area of that half-disc, which is at least $c'/n^{2/3}$ for some new constant c'. The probability p'_{uv} that *all* n of the points fall *outside* that half-disc is therefore:

$$p'_{uv} \leq (1-(c'/n^{2/3}))^n \tag{4}$$

Solving in the limit we get:

$$\lim_{n\to\infty}(1-c'/n^{2/3})^n = \lim_{n\to\infty}((1-c'/n^{2/3})^{n^{2/3}})^{n^{1/3}} = e^{-c'n^{1/3}}. \quad (5)$$

We now move from the probability p'_{uv} that one particular half-disc of C_{uv} is empty to the probability p_{uv} that *at least one* of these half-discs is empty, which is a necessary but not sufficient condition for (u,v) to be a greedy edge. The probability that a particular half-disc contains a point is $1-p'_{uv}$, and so the probability that both contain a point is $(1-p'_{uv})^2$ and the probability that at least one does not contain a point is :

$$p_{uv} \leq 1-(1-p'_{uv})^2 \leq 2e^{-c'n^{1/3}}, \quad (6)$$

We now have the probability p_{uv} that (u,v) meets a necessary condition for being a GT edge. We apply this probability globally to all such possibly pairs of points. The number of interior pairs separated by a distance of $c/n^{1/3}$ is certainly less than n^2. Therefore, the probability that none of the pairs meets the necessary condition (equivalently that all of the pairs fail to meet the condition) is given by:

$$p \geq (1-p_{uv})^{n^2}. \quad (7)$$

We substitute for p_{uv} and solve in the limit as $n \to \infty$.

$$p \geq \lim_{n\to\infty}(1-2e^{-c'n^{1/3}})^{n^2} = e^{h(n)} \quad (8)$$

where

$$h(n) = \lim_{n\to\infty}\frac{\ln(1-2e^{-c'n^{1/3}})}{n^{-2}} = \lim_{n\to\infty}\frac{(\frac{1}{1-2e^{-c'n^{1/3}}})(-2e^{-c'n^{1/3}})(-c'/3)n^{-2/3}}{-2n^{-3}} = 0. \quad (9)$$

Substituting into Equation 8 we have $p \geq e^{h(n)} = e^0 = 1$, the probability that none of the interior pairs meet the necessary condition for being a GT edge. This completes the proof of Lemma 2. □

We now need only deal with pairs of points which are close to the hull but not on it. Let $u, v \in S$ with (u,v) "close" to the convex hull of S. We note that one of the half-discs of C_{uv} still falls entirely within the hull and so our reasoning in Lemma 2 holds for this half-disc. Assume that a percentage ϵ of the other half-disc also lies within the hull. Then in Equation 4 we replace constant c' with $\epsilon c'$. For any constant percentage ϵ, we can then simply replace c' with a new constant c'' and our Lemma still holds.

However as $\epsilon \to 0$ (that is, as our pair (u,v) gets arbitrarily close to the hull) we have a new situation and the above analysis no longer applies. But here we note that at least one hull edge passes through C_{uv} very close to the diameter. We divide this into two possibilities. Either a single hull edge spans the entire outer half-disc of C_{uv}, or a hull point falls within the outer half-disc of C_{uv}. In the first case, a single hull edge spans C_{uv} very close to the diameter and we thus have a hull edge whose length is at least $d(u,v)-\epsilon'$ for some small ϵ' depending on ϵ. Thus our interior edge, even if it is a greedy edge, is no more than ϵ' longer than some hull edge, whose maximum expected length we showed in Lemma 1 to be $\Theta(1/n^{1/3})$. In the second case, we have a convex hull point in C_{uv}. But this reverts back to the previous case given in Lemma 2 except that now we *know* that one of the half-discs, namely the one closer to the hull, is not empty.

This completes the proof of Theorem 3. □

We have now shown that the expected length of the longest edge of the GT of a set of n points in a uniform random distribution on a unit disc is $\Theta(1/n^{1/3})$. It remains to be shown that the expected rank of that longest edge is $\Theta(n^{4/3})$.

Theorem 4 *Let S be a set of n points chosen in a uniform random distribution on a unit disc. The expected value of $r(1/n^{1/3})$ is $\Theta(n^{4/3})$.*

Proof Of the $\binom{n}{2}$ possible distances, we need to count the number of distances of size $O(1/n^{1/3})$. We begin with a lemma of Manacher and Zobrist [10].

Lemma 3 (Manacher and Zobrist) *For any point in a set of u.d. points let $e_k (1 \le k < n)$ be the mean distance to the k^{th} nearest neighbor. Then*

$$(k/n)^{1/2} \le e_k \le 2(k/n)^{1/2}. \tag{10}$$

Substituting $k/4$ for k in Equation 11 we get:

$$(1/2)(k/n)^{1/2} \le e_{k/4} \le (k/n)^{1/2}. \tag{11}$$

Thus at a distance of $(k/n)^{1/2}$ from any given point, we expect between $k/4$ and k neighbors, or simply $\Theta(k)$ neighbors. Setting this distance to $1/n^{1/3}$, which is the expected length of the longest convex hull edge, we solve for k. Solving $(k/n)^{1/2} = \Theta(1/n^{1/3})$ we get k is $\Theta(n^{1/3})$. So the expected number of neighbors within a distance of $1/n^{1/3}$ from a point in S is $\Theta(n^{1/3})$. We now simply multiply by n, the number of points in S, to get the expected rank of $1/n^{1/3}$. We now have $r(1/n^{1/3})$ is $\Theta(n(n^{1/3})) = \Theta(n^{4/3})$. □

The proof of Theorem 1 is now complete. The expected rank of our longest convex hull edge is $\Theta(n^{4/3})$.

3 A New Greedy Triangulation Algorithm

Based on the Theorems 1 and 2, we now present a new algorithm for computing the greedy triangulation for a set of points on the plane. The algorithm is based on the enumeration algorithm of Dickerson and Drysdale [2] and the edge testing method of Gilbert [3]. Dickerson and Drysdale presented an algorithm for enumerating in order the k shortest distances in a set of n points in time $O(n \log n + k \log n)$ and space $O(n+k)$. We use this algorithm to enumerate the possible pairs in order, and for each pair we use the $O(\log n)$ method of Gilbert to test to see if that pair becomes in edge in the GT. The algorithm is terminated when the triangulation is complete. It is a trivial matter to determine when to terminate because the number of edges in any triangulation is known to be $3n - 3 - C(S)$ where $C(S)$ is the number of edges on the convex hull of S. Though $C(S)$ could be easily computed, the number of edges is already known as a biproduct of the algorithm of [2] which first computes the Delaunay triangulation.

The analysis of this algorithm is simple. We halt when the last GT edge has been added. The total number of edges tested is precisely the rank k of this last edge, the amortized time to enumerate each edge is $O(\log n)$ [2], and the time required to test each edge is also $O(\log n)$ using the method of Gilbert. This yields an $O(k \log n)$ time, $O(k)$ space algorithm for the greedy triangulation. From Theorem 1, the expected value of k for a uniform distribution in a disc is $\Theta(n^{4/3})$ and so we have an expected case $O(n^{4/3} \log n)$ time $O(n^{4/3})$ space GT algorithm. This is asymptotically faster than the best known algorithms of Manacher and Zobrist [10] which requires $O(n^2)$ time in the expected case, or of Goldman [4] which requires $O(n^2 \log n)$ time in the worst case. Both the algorithms of [10] and [4] require only $O(n)$ space, compared with this algorithm which requires $O(n^{4/3})$ space in the expected case.

4 Rank in an R-gon

We now prove Theorem 2, showing that for points drawn from a uniform random distribution in an r-gon, the expected rank of the longest edge in the GT is $\Theta(n^2)$.

Proof of Theorem 2 Consider our set S of points in a uniform random distribution in a r-gon. Without loss of generality, we orient the r-gon so that a longest edge is on the left side. Let $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ be the two leftmost points in S. Since the x and y coordinates of these points are independent, the expected value of $|y_1 - y_2|$ is a constant c depending entirely on the r-gon. Now since p_1 and p_2 are the leftmost points in the set, either (p_1, p_2) is an edge in the greedy triangulation of S or there exists a longer edge which includes the leftmost of these points and "spans" the other point. But this implies that the expected value of the longest edge is at least as long as our constant c, which independent of n. It follows that a constant fraction of the $\binom{n}{2}$ interpoint distances will be greater than c and a constant fraction less than c, and thus the rank of c is $\Theta(n^2)$.

5 Summary and Open Problems

We have shown that the expected value of the rank of the longest edge in uniform distributions in a disc as $\Theta(n^{4/3})$. We used this result in an application of the enumeration algorithm of Dickerson and Drysdale [2] to present a new algorithm for the greedy triangulation which in the expected case is asymptotically faster than the best known previous result. Unfortunately, for uniform distribution in an r-gon we showed that the expected rank of the longest edge is $\Theta(n^2)$ and thus the GT algorithm given in this paper is not as efficient as those of Manacher and Zobrist [10], and Goldman [4]. We also note here that the enumeration algorithm of Salowe [13] without modification would not suffice as a basis for this greedy triangulation algorithm. The algorithm of [13] requires that k be known in advance, whereas the algorithm of [2] does not have this requirement.

We end with a few comments and questions. First, it would be interesting to explore this problem for other distributions. The normal distribution is one which has received some attention. Intuitively, we would expect that the normal distribution would allow for longer edges in the greedy triangulation. In an r-gon, this will once again lead to an expected rank of $\Theta(n^2)$ though with a higher constant. But on a disc we might expect something better. The number of points on the convex hull for a normal distribution is $O(log^{1/2}n)$ [12] giving us convex hull edges whose average length is $\Omega(1/log^{1/2}n)$. Will this lead to an efficient algorithm?

We might also ask if results for the greedy triangulation algorithm can be improved yet further. It seems unlikely that there exists an $O(n \log n)$ algorithm, as there does for the Delaunay triangulation, but can we improve on the method given here in either the worst or the expected case?

References

[1] M. Dickerson and R.L. Drysdale, "Fixed radius search problems for points and segments." *Information Processing Letters* 35 (1991) 269-273.

[2] M. Dickerson and R.L. Drysdale, "Enumerating k distances for n points in the plane." *Proceedings of the Seventh Annual ACM Symposium on Computational Geometry* (1991) 234-238.

[3] P. Gilbert, "New results in planar triangulations." MS Thesis, University of Illinois, Urbana, IL, 1979.

[4] S. Goldman, "A Space Efficient Greedy Triangulation Algorithm." *Information Processing Letters* 31 (1989) 191-196.

[5] D. Kirkpatrick, "A note on Delaunay and optimal triangulations." *Information Processing Letters* 10 (1980) 127-128.

[6] C. Levcopoulos and A. Lingas, "On Approximating Behavior of the Greedy Triangulation for Convex Polygons." *Algorithmica* 2 (1987) 175-193.

[7] A. Lingas, "On Approximating Behavior and Implementation of the Greedy Triangulation for Convex Planar Point Sets." *Proceedings of the Second Annual ACM Symposium on Computational Geometry* (1991) 72-79.

[8] E. Lloyd, "On triangulations of a set of points in the plane." *Proceedings of the 18th FOCS* (1977) 228-240.

[9] G. Manacher and A. Zobrist, "Neither the greedy nor the Delaunay triangulation of the planar set approximates the optimal triangulation." *IPL* 9 (1979) 31-34.

[10] G. Manacher and A. Zobrist, "Probabilistic methods with heaps for fast-average-case greedy algorithms." *Advances in Computing Research* vol. 1 (1983) 261-278.

[11] F. Preparata and M. Shamos, "Computational Geometry: an introduction." Springer-Verlag, 1985.

[12] H. Raynaud, "Sur L'Enveloppe Convexe des Nuages des Points Aléatiores dans R^n, I." *J. Applied Probability* 7(1970) 35-48.

[13] J. S. Salowe, "Shallow interdistance selection and interdistance enumeration." *Proceedings WADS '91, Lecture Notes in Computer Science*, Springer-Verlag, 1991, 117-128.