Unit-Disk Graphs, 2003; Gao, Zhang

Jie Gao, Stony Brook University. Li Zhang, Hewlett-Packard Labs.

INDEX TERMS: Well-separated pair decomposition, Unit-disk graph, Approximation algorithm.
INDEX TERMS: none.

1 Problem Definition

Well-separated pair decomposition, introduced by Callahan and Kosaraju [3], has found numerous applications in solving proximity problems for points in the Euclidean space. A pair of point sets (A, B) is *c-well-separated* if the distance between A, B is at least c times the diameters of both A and B. A well-separated pair decomposition of a point set consists of a set of well-separated pairs that "cover" all the pairs of distinct points, i.e. any two distinct points belong to the different sets of some pair. In [3], Callahan and Kosaraju showed that for any point set in an Euclidean space and for any constant $c \geq 1$, there always exists a c-well-separated pair decomposition with linearly many pairs. This fact has been very useful in obtaining nearly linear time algorithms for many problems such as computing k-nearest neighbors, N-body potential fields, geometric spanners, approximate minimum spanning trees etc. Well-separated pair decomposition is also shown very useful in obtaining efficient dynamic, parallel, and external memory algorithms.

The definition of well-separated pair decomposition can be naturally extended to any metric space. However, a general metric space may not admit a well-separated pair decomposition with a sub-quadratic size. Indeed, even for the metric induced by a star tree with unit weight on each edge[1], any well-separated pair decomposition requires quadratically many pairs. This makes the well-separated pair decomposition useless for such a metric. However, it has been shown that for unit-disk graph metric, there do exist well-separated pair decompositions with almost linear size, and therefore many proximity problems under unit disk graph metric can be solved efficiently.

Unit-disk graphs [4]. Denote by $\mathrm{d}(\cdot, \cdot)$ the Euclidean metric. For a set of points S in the plane, the unit-disk graph $I(S) = (S, E)$ is defined to be the weighted graph where an edge $e = (p, q)$ is in the graph if $\mathrm{d}(p, q) \leq 1$, and the weight of e is $\mathrm{d}(p, q)$. Likewise, we can define the unit-ball graph for points in higher dimensions.

Unit disk graphs have been used extensively to model the communication or influence between objects [12, 9] and studied in many different contexts [4, 10]. For an example, wireless ad hoc networks can be modeled by unit-disk graphs [6], as two wireless nodes can

[1] A metric induced by a graph (with positive edge weights) is the shortest path distance metric of the graph.

directly communicate with each other only if they are within certain distance. In unsupervised learning, for a dense sampling of points from some unknown manifold, the length of the shortest path on the unit-ball graph is a good approximation of the geodesic distance on the underlying (unknown) manifold if radius is chosen appropriately [14, 5]. By using well-separted pair decomposition, one can encode the all pairs distance approximately by a compact data structure and query efficiently in $O(1)$ time.

Metric space. Suppose that (S, π) is a metric space where S is a set of elements and π the distance function defined on $S \times S$. For any subset $S_1 \subseteq S$, the *diameter* $D_\pi(S_1)$ (or $D(S_1)$ when π is clear from the context) of S is defined to be $\max_{s_1, s_2 \in S_1} \pi(s_1, s_2)$. The *distance* $\pi(S_1, S_2)$ between two sets $S_1, S_2 \subseteq S$ is defined to be $\min_{s_1 \in S_1, s_2 \in S_2} \pi(s_1, s_2)$.

Well-separated pair decomposition. For a metric space (S, π), two non-empty subsets $S_1, S_2 \subseteq S$ are called *c-well-separated* if $\pi(S_1, S_2) \geq c \cdot \max(D_\pi(S_1), D_\pi(S_2))$.

Following the definition in [3], for any two sets A and B, a set of pairs $\mathcal{P} = \{P_1, P_2, \cdots, P_m\}$, where $P_i = (A_i, B_i)$, is called a *pair decomposition* of (A, B) (or of A if $A = B$) if

- for all the i's, $A_i \subseteq A$, and $B_i \subseteq B$;
- $A_i \cap B_i = \emptyset$;
- for any two elements $a \in A$ and $b \in B$, there exists a unique i such that $a \in A_i$, and $b \in B_i$. We call (a, b) is *covered* by the pair (A_i, B_i).

If in addition, every pair in $\mathcal{P}$ is c-well-separated, $\mathcal{P}$ is called a *c-well-separated pair decomposition* (or c-WSPD in short). Clearly, any metric space admits a c-WSPD with quadratic size by using the trivial family that contains all the pairwise elements.

2 Key Results

In [7], it is shown that for the metric induced by the unit-disk graph on n points and for any constant $c \geq 1$, there does exist a c-well-separated pair decomposition with $O(n \log n)$ pairs, and such a decomposition can be computed in $O(n \log n)$ time. It is also shown that the bounds can be extended to higher dimensions. The following theorems state the key results for two and higher dimensions.

Theorem 1 *For any set S of n points in the plane and any $c \geq 1$, there exists a c-WSPD $\mathcal{P}$ of S under the unit-disk graph metric where $\mathcal{P}$ contains $O(c^4 n \log n)$ pairs and can be computed in $O(c^4 n \log n)$ time.*

Theorem 2 *For any set S of n points in $\mathbb{R}^k$, for $k \geq 3$, and for any constant $c \geq 1$, there exist a c-WSPD $\mathcal{P}$ of S under the unit-ball graph metric where $\mathcal{P}$ contains $O(n^{2-2/k})$ pairs and can be constructed in $O(n^{4/3} \operatorname{polylog} n)$ time for $k = 3$ and in $O(n^{2-2/k})$ time for $k \geq 4$.*

The difficulty in obtaining a well-separated pair decomposition for unit-disk graph metric is that two points that are close in the space are not necessarily close under the graph metric. The above bounds are first shown for the point set with constant-bounded density, i.e. a point set where any unit disk covers only a constant number of points in the set. The upper bound on the number of pairs is obtained by using a packing argument similar to the one used in [8].

For a point set with unbounded density, one applies a clustering technique similar to the one used in [6] to the point set and obtains a set of "clusterheads" with a bounded density. Then the result for bounded density is applied to those clusterheads. Finally, the well-separated pair decomposition is obtained by combining the well-separated pair decomposition for the bounded density point sets and for the Euclidean metric. The number of pairs is dominated by the number of pairs contructed for a constant density set, which is in turn dominated by the bound given by the packing argument. It has been shown that the bounds on the number of pairs is tight for $k \geq 3$.

3 Applications

For a pair of well-separated sets, the distance between two points from different sets can be approximated by the "distance" between the two sets or the distance between any pair of points in different sets. In other words, a well-separated pair decomposition can be thought as a compressed representation to approximate the $\Theta(n^2)$ pairwise distances. Many problems that require to check the pairwise distances can therefore be approximately solved by examining those distances between the well-separated pairs of sets. When the size of the well-separated pair decomposition is sub-quadratic, it often gives us more efficient algorithms than examining all the pairwise distances. Indeed, this is the intuition behind many applications of the geometric well-separated pair decomposition. By using the same intuition, one can apply the well-separated pair decomposition in several proximity problems under the unit-disk graph metric.

Suppose that S, d is a metric space. Let $S_1 \subseteq S$. Consider the following natural proximity problems.

- **Furthest neighbor, diameter, center.** The furthest neighbor of $p \in S_1$ is the point in S_1 that maximizes the distance to p. Related problems include computing the *diameter*, the maximum pairwise shortest distance for points in S_1, and the *center*, the point that minimizes the maximum distance to all the other points.

- **Nearest neighbor, closest pair.** The nearest neighbor of $p \in S_1$ is the point in S_1 with the minimum distance to p. Related problems include computing the *closest pair*, the pair with minimum shortest distance, and the *bichromatic closest pair*, the pair that minimizes distance between points from two different sets.

- **Median.** The median of S is the point in S that minimizes the average (or total) distance to all the other points.

- **Stretch factor.** For a graph G defined on S, its stretch factor with respect to the unit-disk graph metric is defined to be the maximum ratio $\pi_G(p,q)/\pi(p,q)$ where π_G, π are the distances induced by G and by the unit-disk graph, respectively.

All the above problems can be solved or approximated efficiently for points in the Euclidean space. However, for the metric induced by a graph, even for planar graphs, very little is known other than solving the expensive all-pairs shortest path problem. For computing diameter, there is a simple linear time method that achieves a 2-approximation[2] and a $4/3$-approximate algorithm with running time $O(m\sqrt{n \log n} + n^2 \log n)$, for a graph with n vertices and m edges, by Aingworth *et al.* [1].

By using the well-separated pair decomposition, it is shown in [7] that one can obtain better approximation algorithms for the above proximity problems for the unit-disk graph metric. Specifically, one can obtain almost linear time algorithms for computing 2.42-approximation and $O(n\sqrt{n \log n}/\varepsilon^3)$ time algorithms for computing $(1+\varepsilon)$-approximation for any $\varepsilon > 0$. In addition, the well-separated pair decomposition can be used to obtain an $O(n \log n/\varepsilon^4)$ space distance oracle so that any $(1+\varepsilon)$ distance query in the unit-disk graph can be answered in $O(1)$ time.

The bottleneck of the above algorithms turns out to be the approximation of the shortest distance between $O(n \log n)$ pairs. The algorithm in [7] only contructs well-separated pair decompositions without computing an accurate approximation of the distances. The approximation ratio and the running time are dominated by that of the approximation algorithms used to estimate the distance between each pair in the WSPD. Once the distance estimation has been made, the rest of computation only takes almost linear time.

For a general graph, it is unknown whether $O(n \log n)$ pairs shortest path distances can be computed significantly faster than all pairs shortest path distances. For the planar graph, one can compute $O(n \log n)$ pairs shortest path distance in $O(n\sqrt{n \log n})$ time by using separators with $O(\sqrt{n})$ size [2]. This method extends to the unit-disk graph with constant bounded density since such graphs enjoy similar separator property as the planar graphs [13]. As for approximation, Thorup [15] recently discovered an algorithm for planar graphs that can answer any $(1+\varepsilon)$-shortest distance query in $O(1/\varepsilon)$ time after almost linear time preprocessing. Unfortunately, Thorup's algorithm uses balanced shortest-path separators in planar graphs which do not obviously extend to the unit-disk graphs. On the other hand, it is known that there does exist planar 2.42-spanner for a unit-disk graph [11]. By applying Thorup's algorithm to that planar spanner, one can compute 2.42-approximate shortest path distance for $O(n \log n)$ pairs in almost linear time.

4 Open Problems

The most notable open problem is the gap between $\Omega(n)$ and $O(n \log n)$ on the number of pairs needed in the plane. Also, the time bound for $(1+\varepsilon)$-approximation is still about $\widetilde{O}(n\sqrt{n})$ because of the lack of efficient methods for computing $(1+\varepsilon)$-approximate shortest distance between $O(n)$ pairs of points. Any improvement to the algorithm for that

[2] Select an arbitrary node v and compute the shortest path tree rooted at v. Suppose that the furthest node from v is of distance D away. Then the diameter of the graph is no longer than $2D$, by triangular inequality.

problem will immediately lead to improvement to all the $(1+\varepsilon)$-approximate algorithms presented in this paper.

5 Experimental Results

None is reported.

6 Data Sets

None is reported.

7 URL to code

None is reported.

8 Cross References

Applications of Geometric Spanners, Separators, Sparse Graph Spanners

9 Recommended Reading

References

[1] D. Aingworth, C. Chekuri, and R. Motwani. Fast estimation of diameter and shortest paths (without matrix multiplication). In *Proc. 7th ACM-SIAM Symposium on Discrete Algorithms*, pages 547–553, 1996.

[2] S. R. Arikati, D. Z. Chen, L. P. Chew, G. Das, M. H. M. Smid, and C. D. Zaroliagis. Planar spanners and approximate shortest path queries among obstacles in the plane. In J. Díaz and M. Serna, editors, *Proc. of 4th Annual European Symposium on Algorithms*, pages 514–528, 1996.

[3] P. B. Callahan and S. R. Kosaraju. A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. *J. ACM*, 42:67–90, 1995.

[4] B. N. Clark, C. J. Colbourn, and D. S. Johnson. Unit disk graphs. *Discrete Mathematics*, 86:165–177, 1990.

[5] B. Fischl, M. Sereno, and A. Dale. Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9:195–207, 1999.

[6] J. Gao, L. J. Guibas, J. Hershberger, L. Zhang, and A. Zhu. Geometric spanners for routing in mobile networks. *IEEE Journal on Selected Areas in Communications Wireless Ad Hoc Networks (J-SAC)*, 23(1):174–185, January 2005.

[7] J. Gao and L. Zhang. Well-separated pair decomposition for the unit-disk graph metric and its applications. In *Proc. of 35th ACM Symposium on Theory of Computing (STOC'03)*, pages 483–492, 2003.

[8] L. Guibas, A. Ngyuen, D. Russel, and L. Zhang. Collision detection for deforming necklaces. In *Proc. 18th ACM Symposium on Computational Geometry*, pages 33–42, 2002.

[9] W. K. Hale. Frequency assignment: Theory and applications. *Proceedings of the IEEE*, 68(12):1497–1513, 1980.

[10] H. B. H. III, M. V. Marathe, V. Radhakrishnan, S. S. Ravi, D. J. Rosenkrantz, and R. E. Stearns. NC-approximation schemes for NP- and PSPACE-hard problems for geometric graphs. *J. Algorithms*, 26(2):238–274, 1998.

[11] X.-Y. Li, G. Calinescu, and P.-J. Wan. Distributed Construction of a Planar Spanner and Routing for Ad Hoc Wireless Networks. In *IEEE INFOCOM 2002*, New York, NY, June 23–27 2002.

[12] C. A. Mead and L. Conway. *Introduction to VLSI Systems*. Addison-Wesley, 1980.

[13] G. L. Miller, S.-H. Teng, and S. A. Vavasis. An unified geometric approach to graph separators. In *Proc. 32nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 538–547, 1991.

[14] J. Tenenbaum, V. de Silva, and J. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:22, December 2000.

[15] M. Thorup. Compact oracles for reachability and approximate distances in planar digraphs. In *Proc. 42nd IEEE Symposium on Foundations of Computer Science*, pages 242–251, 2001.