The Emergence of Sparse Spanners and Greedy Well-Separated Pair Decomposition

Jie Gao* Dengpan Zhou*

* Department of Computer Science, Stony Brook University. {dpzhou, jgao}@cs.sunysb.edu

I. Introduction

A geometric graph G defined on a set of points $P \subseteq \mathbb{R}^d$ with all edges as straight line segments of weight equal to the length is called *Euclidean spanner* [4], if for any two points $p, q \in P$ the shortest path between them in G has length at most $\lambda \cdot |pq|$, where $|pq|$ is the Euclidean distance. The factor λ is called the *stretch factor* of G and the graph G is called a λ-spanner. Spanners with a sparse set of edges provide good approximations for the pairwise Euclidean distances and are good candidates for network backbones. Thus, there has been a lot of work on the construction of Euclidean spanners in both the centralized setting and the distributed setting. In this paper we are interested in the emergence of good Euclidean spanners formed by uncoordinated agents. The work in this paper initiates the study of the emergence of good spanners in the setting when there is little coordination between the peers and the users only need a modest amount of incomplete information of the current overlay topology.

Our contribution We consider in this paper the following model. There are n points in the plane. Each point represents a separate agent and may build edges from itself to some other points by the strategy to be explained later. The edges in the final graph is the collection of edges built by all the agents. The agents may decide to build edges at different point in time. When an agent p plans on whether an edge from itself to another point q should be built or not, p checks to see whether there is already an edge from some points p' to q' such that $|pp'|$ and $|qq'|$ are both within $\frac{1}{4(1+1/\varepsilon)} \cdot |p'q'|$ from p and q respectively. If not, the edge pq is built, otherwise it is not. This strategy is very intuitive — if there is already a cross-country highway from Washington D.C to San Francisco, it does not make economical sense to build a highway from New York to Los Angeles. We assume that each agent will eventually check on each possible edge from itself to all the other points, but the order on who checks which edge can be *completely arbitrary*. With this strategy, the agents only make decisions with limited information and no agent has full control over how and what graph will be constructed. It is not obvious that this strategy will end up with a sparse spanner on all points. It is even not clear that the graph will be connected.

The main result in this paper is to show that with the above strategy executed in *any* arbitrary order, the graph built at the end of the process is a sparse spanner graph with the following properties:

- Between any two points p, q, there is a path with stretch $1+\varepsilon$ and $O(|pq|^{1/(1+2/\varepsilon)})$ hops.
- The number of edges is O(n).
- The total edge length of the spanner is $O(|\text{MST}| \cdot \log \alpha)$, where α is the *aspect ratio*, i.e., the ration of the distance between the furthest pair and the closest pair, and $|\text{MST}|$ is the total edge length of the minimum spanning tree of the point set. Clearly $|\text{MST}|$ is a lower bound on the total edge length of any constant stretch spanner.

II. Uncoordinated spanner construction algorithm

Given n points in $\mathbb{R}^d$, each point is represented by an agent. We consider the following algorithm for constructing a sparse spanner with stretch factor s in an uncoordinated way. For any point p, denote by $B_r(p)$ the collection of points that are within distance r from point p, i.e., inside the ball with radius r centered at p.

Uncoordinated spanner construction. Each point/agent p will check to see whether an edge from itself to another point q should be constructed or not. At this point there might be some edges already constructed by other agents. The order of which agent checks on which edge is completely arbitrary. Specifically, p performs the following operation:

Check where there is already an edge $p'q'$ such that p and q are within distance $\frac{|p'q'|}{2(s+1)}$ from p', q' respectively. If so, p does not build the edge to q. Otherwise, p will build an edge to q.

This incremental construction of edges is executed by different agents in a completely uncoordinated manner. Any agent makes its decision only based on local information and the current network already constructed. The algorithm terminates when all agents finish checking the edges from themselves to all other points. To show the above algorithm output a good spanner, we first show the connection of G with the notion *well-separated pair decomposition*.

Definition 2.1 (Well-separated pair). *Let $s > 0$ be a constant, and a pair of sets of points A, B is s-separated, if $d(A, B) \geq s \cdot \max(\text{diam}(A), \text{diam}(B))$, where $\text{diam}(A)$ is the diameter of the point set A, $\text{diam}(A) = \max_{p,q \in A} |pq|$, and $d(A, B) = \min_{p \in A, q \in B} |pq|$.*

Definition 2.2 (Well-separated pair decomposition). *Let $s > 0$ be a constant,and $\mathcal{P}$ be a point set. An s-well-*

separated pair decomposition (WSPD) of $\mathcal{P}$ is a set of pairs $\mathcal{W} = \{(A_1, B_1), \ldots, (A_m, B_m)\}$, s.t.

1) *$A_i, B_i \subseteq P$, and the pair sets A_i and B_i are s-separated for every i.*
2) *For any two points $p, q \in \mathcal{P}$, there is at least one pair (A_i, B_i) such that $p \in A_i$ and $q \in B_i$.*

Here m is called the size of the WSPD.

A greedy algorithm for well-separated pair decomposition The above theorem shows the connection of the uncoordinated graph G with a WSPD W. In fact, the way to compute the WSPD W via the construction of G is equivalent to the following algorithm that computes an s-WSPD, in a greedy fashion, with $s > 1$.

1) Choose an arbitrary pair (p, q), not yet covered by existing well-separated pairs in W.
2) Include the pair of point sets $B_r(p)$ and $B_r(q)$ in the WSPD W, with $r = |pq|/(2 + 2s)$.
3) Label every pair (p_i, q_i) with $p_i \in B_r(p)$ and $q_i \in B_r(q)$ as being covered.
4) Repeat the above steps until every pair of points is covered.

III. Properties of the Greedy Spanner

Deformable spanner. Given a set of point $\mathcal{P}$ in the plane, a set of *discrete centers* with radius r is defined to be the maximal set $S \in \mathcal{P}$ that satisfies the *covering* property and the *separation* property: any point $p \in \mathcal{P}$ is within distance r to some point $p' \in S$; and every two points in S are of distance at least r away from each other. We now define a hierarchy of discrete centers in a recursive way. S_0 is the original point set $\mathcal{P}$. S_i is the discrete center set of S_{i-1} with radius 2^i. The deformable spanner is based on the hierarchy, with all edges between two points u and v in S_i if $|uv| \le c \cdot 2^i$, where c is a constant equal to $4 + 16/\varepsilon$.

Lemma 3.1 (Deformable spanner properties [1]). *For a set of n points in $\mathbb{R}^d$ with aspect ratio α,*

1) *For any point $p \in S_0$, its ancestor $P^{(i)}(p) \in S_i$ is of distance 2^{i+1} away from p.*
2) *Any point $p \in S_i$ has at most $(1 + 2c)^d - 1$ edges with other points of S_i.*
3) *The deformable spanner $\hat{G}$ is a $(1 + \varepsilon)$-spanner with $O(n/\varepsilon^d)$ edges.*
4) *$\hat{G}$ has total weight $O(|MST| \cdot \lg \alpha/\varepsilon^{d+1})$, where $|MST|$ is the weight of the minimal spanning tree of the point set S.*

With the WSPD $\hat{\mathcal{W}}$ constructed by the deformable spanner, we can prove several important properties of our greedy WSPD $\mathcal{W}$. The basic idea is to map the pairs in $\mathcal{W}$ to the pairs in $\hat{\mathcal{W}}$ and show that at most a constant number of pairs in $\mathcal{W}$ map to the same pair in $\hat{W}$.

Theorem 3.2. *The uncoordinated spanner G with parameter s is a spanner with stretch factor $(s+1)/(s-1)$ and has $O(ns^d)$ number of edges, with maximal degree of $O(\lg \alpha \cdot s^d)$ and average degree $O(s^d)$, and the total weight is $O(\lg \alpha \cdot |MST| \cdot s^{d+1})$.*

Theorem 3.3. *For any two point p and q in G, there is a path with stretch $(s + 1)/(s - 1)$ between p and q with at most $2|pq|^{1/(1+\lg s)}$ hops.*

IV. Application of the Greedy Spanner

The uncoordinated spanner construction has application in P2P network overlay design [3]. For that, we will first extend our spanner results on a more general metric, metric with constant doubling dimension [2]. The doubling dimension of a metric space (X, d) is the smallest value γ such that each ball of radius R can be covered by at most 2^γ balls of radius $R/2$.

Theorem 4.1. *For n points and a metric space defined on them with constant doubling dimension γ, the uncoordinated spanner construction outputs a spanner G with stretch factor $(s+1)/(s-1)$, has total weight $O(\gamma^9 \cdot \lg \alpha \cdot |MST| \cdot s^{O(\gamma)})$ and has $O(\gamma^4 \cdot n \cdot s^{O(\gamma)})$ number of edges. Also it has a maximal degree of $O(\gamma^4 \cdot \lg \alpha \cdot s^{O(\gamma)})$ and average degree $O(\gamma^4 \cdot s^{O(\gamma)})$.*

The spanner edges are recorded in a distributed fashion so that no node has the entire picture of the spanner topology. After each edge pq in G is constructed, the peers p, q will inform their neighboring nodes(those in $B_r(p)$ and $B_r(q)$ with $r = |pq|/(2s + 2)$) that such an edge pq exists so that they will not try to connect to one another. The total communication cost for the above construction is $O(n \log \alpha)$, and each node's storage cost is bonded by $O(\log \alpha)$. Although the spanner topology is implicitly stored on the nodes with each node only knows some piece of it, we are actually able to do a distributed and local routing on the spanner with only information available at the nodes such that the path discovered has maximum stretch $(s+1)/(s-1)$. In particular, for any node p who has a message to send to node q, it is guaranteed that (p, q) is covered by well-separated pair $(B_r(p'), B_r(q'))$ with $p \in B_r(p')$ and $q \in B_r(q')$. By the construction algorithm, all the nodes in $B_r(p') \cup B_r(q')$ will be informed the edge pq. Thus p includes in the packet a partial route with $\{p \rightsquigarrow p', p'q', q' \rightsquigarrow q\}$. By the same induction as used in the proof of spanner stretch, the final path is going to have stretch at most (s+1)/(s-1) and at most $2|pq|^{1/(1+\lg s)}$ hops. The constructed spanner can also be used to look for nearest peer in the P2P network, and it nicely supports node insertion and deletion as well.

References

[1] J. Gao, L. Guibas, and A. Nguyen. Deformable spanners and their applications. *Computational Geometry: Theory and Applications*, 35(1-2):2–19, 2006.

[2] A. Gupta, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *FOCS '03: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 534–543, 2003.

[3] K. Lua, J. Crowcroft, M. Pias, R. Sharma, and S. Lim. A survey and comparison of peer-to-peer overlay network schemes. *Communications Surveys & Tutorials, IEEE*, pages 72–93, 2005.

[4] G. Narasimhan and M. Smid. *Geometric Spanner Networks*. Cambridge University Press, 2007.