Succinct Graph Structures and Their Applications	**Spring 2018**

Lecture 1: March 28

Lecturer: Merav Parter

Spanners

Graph spanners introduced by Peleg and Schäffer [PS89] are sparse subgraphs that faithfully preserve the distances in the original graph up to some bounded stretch. Formally, for a given undirected n-vertex graph $G = (V, E)$, a subgraph $H \subseteq G$ is an (α, β)-spanner for G if for every $u, v \in V$, it holds:

$$\texttt{dist}(u, v, H) \leq \alpha \cdot \texttt{dist}(u, v, G) + \beta \ .$$

When $\beta = 0$, the spanner is called *multiplicative* and when $\alpha = 1$, the spanner is *additive*. In this class, we mainly consider multiplicative spanners, which we usually denote by $(2k-1)$-spanners, i.e., spanners with multiplicative stretch $2k-1$. As a warm-up example, note that every n-clique has a 2-spanner with $n-1$ edges, simply by taking all edges incident to one vertex. There are, however, graphs for which any 2-spanner has $\Omega(n^2)$ edges (e.g., the complete bipartite graph).

Greedy Construction. We will now see a very simple greedy algorithm due to [ADD$^+$93] that constructs $(2k-1)$ spanners for (possibly) weighted graphs with $O(n^{1+1/k})$ edges. For simplicity assume that the edge weights in the graph are unique, $w(e_i) \neq w(e_j)$. For a weighted graph G and $u, v \in V(G)$, $\texttt{dist}(u, v, G)$ is the weight of the shortest-path between u and v in G. See Fig. 1.2 for the description of the algorithm. Next, we analyze this construction and begin by showing that H is a $(2k-1)$-spanner for G.

Algorithm $\mathsf{GreedySpanner}(G)$

1. Consider G edges in increasing weight ordering $w(e_1) < \ldots < w(e_m)$.
2. $H \leftarrow \emptyset$.
3. For $i = 1, ..., m$:
   (a) Add the edge $e_i = (u_i, v_i)$ to H only if: $\texttt{dist}(u_i, v_i, H) > (2k-1)\texttt{dist}(u_i, v_i, G)$.
4. Output H.

Figure 1.1: The Greedy Algorithm for Constructing $(2k-1)$ Spanners

Lemma 1.1 (Stretch) *For every $u, v \in V$, $\texttt{dist}(u, v, H) \leq (2k-1) \cdot \texttt{dist}(u, v, G)$.*

Proof: By construction, the lemma holds for every $(u, v) \in E$. Now, consider an arbitrary pair $x, y \in V$ and let P the x-y shortest path in G. Since for each edge $e = (u, v) \in P$, there is a u-v path in H of length at most $(2k-1)w(e)$, it holds that $\texttt{dist}(u, v, H) \leq (2k-1)\sum_{e \in P} w(e) = (2k-1)\texttt{dist}(u, v, G)$. The fact that it is sufficient to bound the stretch between neighboring vertex pairs is a very convenient property of multiplicative spanners. It does not hold for additive spanners, which are indeed much more intriguing structures as we will see in later sessions. ■

We will bound the size of the spanner in two steps. An important definition in this context is the *girth* of the graph, namely, the length of the shortest cycle in G. We first show that the output spanner of Alg. $\mathsf{GreedySpanner}(G)$ has a large girth and then show that every graph of large girth must be sparse.

Claim 1.2 *The girth of H is at least $2k+1$.*

Proof: Assume towards contradiction that H has a cycle C of length at most $2k$. Let $e = (u, v) \in C$ be the cycle edge that was added last to H (among all other C edges). By the addition of e, we have that $\sum_{e' \in C \setminus \{e\}} w(e') > (2k-1)w(e)$. On the other hand, by the fact that $C \setminus \{e\}$ contains at most $2k-1$ edges that are lighter than e, it holds that $\sum_{e' \in C \setminus \{e\}} w(e') < (2k-1)w(e)$, leading to contradiction. ■

Claim 1.3 *Every n-vertex graph G with girth $g(G) \geq 2k+1$ has $O(n^{1+1/k})$ edges.*

Proof Sketch: Assume towards contradiction that there exists an n-vertex graph G with at least $4n^{1+1/k}$ edges, and girth at least $2k+2$. We first compute a subgraph $G' \subseteq G$ with at least $2n^{1+1/k}$ edges and minimum degree $2n^{1/k}$. To compute G', we repeatedly omit from G vertices with degrees less than $2n^{1/k}$, one by one. Clearly, the girth of G' is also at least $2k+2$ (by the contradictory assumption and as $G' \subseteq G$). We now consider some vertex $u \in G'$ and bound the number of vertices at distance at most k from it in G'. The key observation is that since the girth of G' is at least $2k+1$, two vertices x, y at distance $i \leq k-1$ from u have no common neighbor. Hence, the number of vertices at distance i from u is at least $(2n^{1/k})^i$ for every $i \leq k-1$. Since pair of vertices at distance $k-1$ from u have no common neighbor, the number of vertices at distance k from u is at least $(2n^{1/k})^k > n$, leading to contradiction.

Erdős' Girth Conjecture. A well-known conjecture by Erdős' states that for every $k \geq 1$ and for sufficiently large n, there are n-vertex graphs with girth at least $2k+2$ and $\Omega(n^{1+1/k})$ edges. The conjecture was verified for $k = 1, 2, 3, 5$. Consider the n-vertex graph G with girth at least $2k+2$ and $\Omega(n^{1+1/k})$ edges which is promised to exist by the conjecture. A removal of any edge (u, v) from G increases the distance between u, v from 1 to $2k+1$ and hence any $(2k-1)$-spanner for G must include all edges of G. This in turn implies that under the girth conjecture, the greedy construction gives a spanner of optimal size (up to constant factors).

3-Spanners via a Clustering Method

We will now see a different approach for constructing 3-spanners by Baswana and Sen [BS07]. The algorithm works also for weighted graphs (up to minor modifications). For demonstrating the clustering ideas in a clean manner, we will focus on unweighted graphs. Unlike the greedy construction, this algorithm is *randomized* and it guarantees to construct a 3-spanner with probability $\geq 1 - 1/n^c$ for some constant c[1]. For a vertex $u \in V$, let $\Gamma(u)$ be the neighbors of u in G and let $\deg(u)$ denotes its degree in G. Call a vertex v *high-degree* if $\deg(v, G) \geq 5\sqrt{n}\log n$, otherwise, the vertex is *low-degree*. Let V_h be the collection of high-degree vertices and let V_ℓ be the set of all low-degree vertices. First, we add to the spanner H all edges incident to low-degree vertices. To handle the high-degree vertices, we sample a set of $O(\sqrt{n})$ *centers* $S \subseteq V$ by sampling each vertex $v \in V$ independently with probability $1/\sqrt{n}$ and adding it to S. Each center $s \in S$ would correspond to a cluster that consists of a subset of its high-degree neighbors. These clusters are defined by letting each high-degree vertex joins the cluster of one of its neighboring centers, as a result, we have $|S| = O(\sqrt{n})$ star-clusters in H. Finally, for each $v \in V_h$ and for each cluster $C(s)$ centered at s, we add one edge between v and $C(s)$ (if exists) to H. We next analyze the construction.

Observation 1.4 *Each high-degree vertex v has a neighbor in S, i.e., $\Gamma(v) \cap S \neq \emptyset$, with probability $1 - 1/n^4$.*

Proof: Fix a high-degree vertex $v \in V$. The probability that none of v's neighbors is sampled into S is $(1 - 1/\sqrt{n})^{\deg(v)} \leq (1 - 1/\sqrt{n})^{5\log n/\sqrt{n}} \leq 1/n^5$. By doing a union bound over all high-degree vertices, the observation follows. ■

Lemma 1.5 *$H \subseteq G$ is 3-spanner with $O(n^{3/2}\log n)$ edges, with high probability.*

Proof: Recall that for multiplicative spanners, it is sufficient to bound the stretch for neighboring pairs. Since all edges incident to low-degree vertices are in H, it remains to consider an edge (u, v) where $u, v \in V_h$. If $c(u) = c(v)$, both neighbors belong to the same star-cluster and hence $\texttt{dist}(u, v, H) \leq 2$. Otherwise, if $(u, v) \notin H$, u must have another neighbor $w \in V_h$ that belongs to the same cluster of v and such that (u, w) was added to the spanner. Letting s be the cluster center of w and v, there is a u-v path $u \to w \to s \to v$ in

[1] We call such a success probablity, *high probability*.

Algorithm $\mathsf{3Spanner}(G)$

1. $E_1 = (V_\ell \times V) \cap E(G)$.
2. $S \leftarrow \mathsf{Sample}(V, 1/\sqrt{n})$.
3. For each $v \in V_h$, let $c(v)$ be an arbitrary center vertex in $S \cap \Gamma(v)$.
4. $E_2 = \{(v, c(v)) \mid v \in V_h\}$.
5. For each $s \in S$, let $C(s) = \{v \in V_h \mid c(v) = s\}$.
6. For each $v \in V_h$, and $s \in S$, let $e(v, s)$ be one edge connecting v and $C(s)$ if exists.
7. $E_3 = \bigcup_{v \in V_h} \bigcup_{s \in S} \{e(v, s)\}$.
8. Output $H \leftarrow E_1 \cup E_2 \cup E_3$.

Figure 1.2: The Clustering-Based Algorithm for Constructing 3-Spanners

H of length 3. We next bound the size of H. The set E_1 has $O(n^{3/2} \log n)$ edges. The set E_2 contains $O(n)$ edges and since each high-degree vertex adds at most one each for each of the $O(\sqrt{n})$ clusters, $|E_3| = O(n^{3/2})$. ■

References

[ADD+93] Ingo Althöfer, Gautam Das, David Dobkin, Deborah Joseph, and José Soares. On sparse spanners of weighted graphs. *Discrete & Computational Geometry*, 9(1):81–100, 1993.

[BS07] Surender Baswana and Sandeep Sen. A simple and linear time randomized algorithm for computing sparse spanners in weighted graphs. *Random Structures and Algorithms*, 30(4):532–563, 2007.

[PS89] David Peleg and Alejandro A Schäffer. Graph spanners. *Journal of graph theory*, 13(1):99–116, 1989.