Thursday, January 25, 2024 11:49 PM

Low-Stretch Trees and Low-Diameter Decompositions

Recall APSP Conjecture: no $n^{3-\varepsilon}$ time algo for APSP.

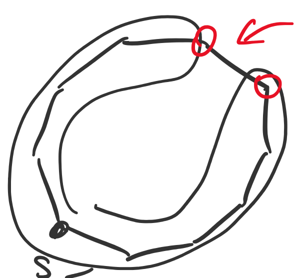
This lecture: $O(\log n)$-approx APSP on undir. graphs in $\tilde{O}(n^2)$ time

Main idea: approximate pairwise distances by <u>trees.</u> Then solve APSP on trees

Ideal statement: for any graph G, there is a tree T on V s.t.
$d_G(u,v) \approx d_T(u,v)$ for all $u,v \in V$.
(Then compute APSP on T in $O(n^2)$ time.)

Remark: true for single-source:
the SSSP tree satisfies $d_G(s,v) = d_T(s,v)$ $\forall v \in V$.

Consider a cycle:



$\leftarrow$ not preserved!

Shortest path tree looks like

<u>Idea</u>: <u>Randomized</u> low-stretch tree, preserve distances in <u>expectation</u>

Def: A (randomized) low-stretch tree of <u>stretch</u> α for a graph $G=(V,E)$ is a probability distribution $\mathcal{D}$ over spanning trees of G s.t.

$d_G(x,y) \le d_T(,)$ for all T in the support of $\mathcal{D}$

$\mathbb{E}\,[d_T(x,y)] \le \alpha\, d_G(x,y)$ for all $x,y \in V$. $(\Longleftrightarrow \forall (x,y) \in E)$

$\mathbb{E}_{T \sim \mathcal{D}} [d_T(x,y)] \le \alpha\, d_G(x,y)$ for all $x, y \in V$.
$(\iff \forall (x,y) \in E)$

Example: cycle. Sample T by dropping a random edge.
($\approx$ run SSSP from a random source)

Claim: is randomized low-stretch tree with stretch $\alpha = 2$.

Proof: $d_G(x,y) \le d_T(x,y)$ since $T \subseteq G$ (distances only increase)

Fix $(x,y) \in E$.

$$\mathbb{E}_T [d_T(x,y)] = \Pr[\text{don't delete edge } (x,y)] \cdot 1 + \Pr[\text{delete edge } (x,y)] \cdot (n-1)$$

$$= \underbrace{\frac{n-1}{n}}_{\le 1} \cdot 1 + \underbrace{\frac{1}{n} \cdot (n-1)}_{\le 1} \le 2.$$

So, $\mathbb{E}_T [d_T(x,y)] \le 2\, d_G(x,y) \quad \forall (u,v) \in E$.

Now consider general $x, y \in V$.

Let $x = v_0 - v_1 - v_2 - \cdots - v_k = y$ be shortest $x-y$ path.

Know $\mathbb{E}_T [d_T(v_i, v_{i+1})] \le 2\, d_G(v_i, v_{i+1}) \quad \forall i$

$$\mathbb{E}_T [d_T(x,y)] \le \mathbb{E}_T [d_T(v_0,v_1) + d_T(v_1,v_2) + \cdots + d_T(v_{k-1},v_k)]$$

$$= \mathbb{E}_T [d_T(v_0,v_1)] + \cdots + \mathbb{E}[d_T(v_{k-1},v_k)]$$

$$\le 2\, d_G(v_0,v_1) + \cdots + 2\, d_G(v_{k-1},v_k)$$

$$= 2\, d_G(x,y).$$

APSP algo: Sample $10 \log n$ trees $T \sim \mathcal{D}$: $T_1, T_2, \ldots, T_{10 \log n}$

Compute APSP on each tree T_i in $O(n^2)$ time.

Compute APSP on each tree T_i in $O(n^2)$ time.

For each x, y, output $\widetilde{dist}_G(x,y) = \min_i dist_{T_i}(x,y)$ as estimate of $dist_G(x,y)$.

Claim: $\widetilde{dist}(x,y) \geq dist_G(x,y)$

Proof: $dist_{T_i}(x,y) \geq dist_G(x,y)$ $\forall i$ by definition.

Claim: With prob $\geq 1 - n^{-8}$, $\widetilde{dist}(x,y) \leq 2\alpha\, dist_G(x,y)$ $\forall u,v \in V$.

Proof: Fix $u,v \in V$. $\mathbb{E}_{T \sim \mathcal{D}}\, dist_T(x,y) \leq \alpha\, dist_G(x,y)$ by definition

$$\Rightarrow \Pr[dist_T(x,y) > 2\alpha\, dist(x,y)] \leq 1/2.$$

The event $\widetilde{dist}(x,y) > 2\alpha\, dist_G(x,y)$ happens iff

$dist_{T_i}(x,y) > 2\alpha\, dist_G(x,y)$ $\forall i$,

happens w.p. $\leq (1/2)^{10 \log n} = n^{-10}$.

So $\widetilde{dist}(x,y) > 2\alpha\, dist_G(x,y)$ w.p. n^{-10}.

By union bound over all $x,y \in V$,

$\widetilde{dist}(x,y) > 2\alpha\, dist_G(x,y)$ for some $x,y \in V$ w.p. n^{-8}

Bartal's Low-Stretch Tree construction.

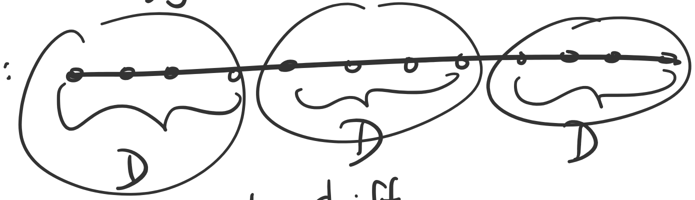
Def: A Low-Diameter Decomposition scheme with <u>approximation β</u> and <u>diameter D</u> on graph G is a randomized algorithm that partitions the vertex set V into clusters $V_1, \ldots, V_k$ s.t.

① $\forall i = 1, \ldots, k$, the diameter of $G[V_i]$ is $\leq D$

② $\forall x,y \in V$, $\Pr[x,y \text{ in different clusters}] \leq \beta \cdot \frac{d(x,y)}{D}$.

② $\forall x,y \in V$, $\Pr[x,y$ in different clusters$]$ ... D

Example: Path graph:



randomly shift

Algorithm: $p \leftarrow \min(1, \frac{4\log n}{D})$

while $\exists$ unclustered vertex $v \in V$:

sample $R_v \sim \text{Geometric}(p)$

take cluster $C_v = \{\text{unclustered } u : d(v,u) \le R_v\}$

Claim: with prob $1 - 1/n$, diam of each cluster is $\le D$.

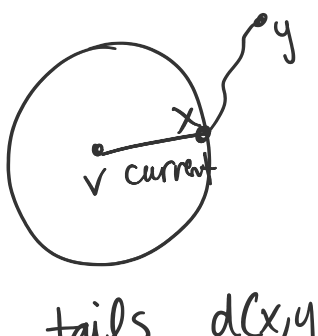
Proof: $\Pr[R_v \le D/2] \le \left(1 - \frac{4\log n}{D}\right)^{D/2} \le e^{-\frac{4\log n}{D}\cdot\frac{D}{2}} = n^{-2}$.

Union bound over all selected v.

Claim: any pair $x,y \in V$ separated w.p. $\le p\,d(x,y)$.

Proof: View clustering C_v as: start at v, flip coin w.p. p, if tails grow ball by 1, else stop. <u>Memoryless.</u>

Consider first time x or y enters current ball.



If flip tails $d(x,y)$ more times, then cluster y:

$d(v,y) \le d(v,x) + d(x,y)$.

In order to separate x,y, must flip a head in next $d(x,y)$ times. Union bound: $p \cdot d(x,y)$.

Assume that G is a <u>metric</u>: complete graph with $w_G(x,y) = dist_G(x,y)$ $\forall x,y \in V$

Algorithm $LST(G=(V,E), 2^\delta)$

Invariant: $diameter(G) \leq 2^\delta$

If $|U| = 1$, return empty graph on U.

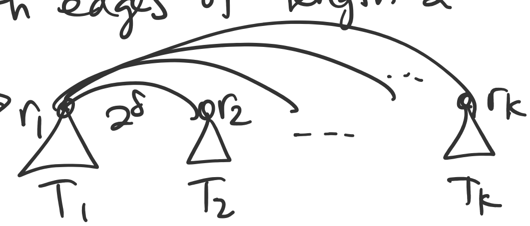
Run LDD on G with $D = 2^{\delta-1}$, get clusters $V_1 \dots V_k$

For each $i = 1, \dots, k$

$T_i \leftarrow LST(G[V_i])$

Connect the roots of T_i with edges of length 2^δ:

Return tree rooted at $r_1 \rightarrow$



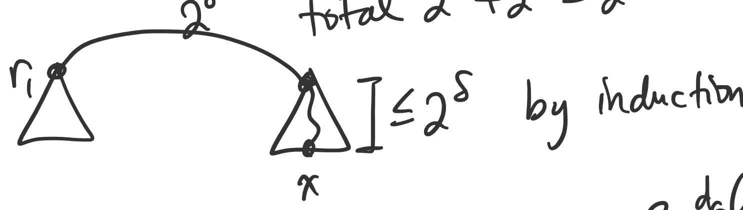
Lemma: If output $T \leftarrow LST(G, 2^\delta)$ has root r, then

① every vertex x in T has $d_T(x,r) \leq 2^{\delta+1}$

② $\mathbb{E}_T[d_T(x,y)] \leq 8\delta\beta\, d_G(x,y)$.

Proof: by induction on δ. Base case trivial.

① total $2^\delta + 2^\delta = 2^{\delta+1}$



② x,y separated by LDD w.p. $\beta \cdot \frac{d_G(x,y)}{2^{\delta-1}}$, in which case distance $\leq d_T(x,r) + d_T(r,y) \leq 2 \cdot 2^{\delta+1} = 4 \cdot 2^\delta$

Else, they lie in same cluster V_i, so by induction,

$d_{T_i}(x,y) \leq 8(\delta-1)\, d_G(x,y)$

$$d_{T_i}(x,y) \le 8(\delta-1)\, d_G(x,y)$$

So $d_T(x,y) = d_{T_i}(x,y) = 8(\delta-1)\, d_G(x,y)$

Overall, $\mathop{\mathbb{E}}_{T}\left[d_T(x,y)\right] \le \underbrace{\beta \cdot \frac{d_G(x,y)}{2^{\delta-1}}}_{\text{separated}} \cdot 4 \cdot 2^{\delta} + \underbrace{1}_{\text{not sep.}} \cdot 8(\delta-1) d_G(x,y)$

$$= 8\delta\, d_G(x,y).$$

So, stretch $\alpha = 8\delta\beta$ where $\delta = \log(\text{diameter}(G)) = O(\log n)$ and $\beta = O(\log n)$.

[FRT'04]: Low-stretch trees with $\alpha = O(\log n)$, also optimal.