Lecture 24 (Oct 31): Tree Metrics — Part 2

Lecturer: Zachary Friggstad *Scribe: Yifeng Zhang*

24.1 Tree Metrics

We finish the discussion of the tree embedding algorithm in this lecture. The basic definitions are found in the previous lectures' notes. We recall the algorithm here.

Algorithm 1 Partitioning Scheme for the Tree Metric

Sample r_0 uniformly from $[1/2, 1)$. Set $r_i = 2^i \cdot r_0$ for $1 \leq i \leq \Delta$. (Recall that $\log_2(\Delta) = \lceil \log_2(\text{diam}(V)) \rceil$)
Let $\pi : V \to V$ be a random permutation of V
$\mathcal{C}(\log_2(\Delta)) = \{V\}$
for i from $\log(\Delta)$ down to 1 **do**
  $\mathcal{C}(i-1) \leftarrow \emptyset$
  for each $\mathcal{S} \in \mathcal{C}(i)$ **do**
    $S' \leftarrow S$
    for each $v \in V$ in order of π **do**
      $S^* \leftarrow B(v, r_{i-1}) \cap S'$
      if $S^* \neq \emptyset$ **then**
        $\mathcal{C}(i-1) \leftarrow \mathcal{C}(i-1) \cup \{S^*\}$
        $S' \leftarrow S' - S^*$
        Add an edge with distance 2^i connecting the vertices corresponding to S^* and S

Note that for each level i, $\mathcal{C}(i)$ is a partition of V the leaves of the constructed tree. Furthermore, the leaf nodes are the singleton sets in $\mathcal{C}(1)$ and we identify $v \in V$ with the leaf node for set $\{v\} \in \mathcal{C}(1)$.

For any two $u, v \in V$, we say that their least common ancestor is at level i if there is some $S \in \mathcal{C}(i)$ such that $u, v \in S$ but u and v lie in different sets of $\mathcal{C}(i+1)$. We prove the following lemma last class, which we are restating for reference.

Lemma 1 *For all pairs $u, v \in V$, $d(u, v) \leq T(u, v) \leq 2^{i+2}$ where the least common ancestor of u and v is at level i.*

In this lecture, we complete the analysis of Algorithm 1 by proving the following.

Theorem 1 *For any pair $u, v \in V$, $\mathrm{E}[T(u, v)] \leq O(\log n) \cdot d(u, v)$*

From now on, we fix u and v. The following concepts will help us examine the expected stretch of (u, v).

Definition 1 *For a vertex $w \in V$, say w **settles** u, v at level i if w is the first vertex (with respect to π) such that $B(r_{i-1}, w) \cap \{u, v\} \neq \emptyset$.*

Note that for each level i, there is exactly one vertex w that settles u, v at level i.

Definition 2 *For a vertex $w \in V$, w **cuts** u, v at level i if $|B(r_{i-1}, w) \cap \{u, v\}| = 1$.*

In the following discussion, for a pair $u, v \in V$, we will use some $\{0,1\}$-random variables:

- $\mathbf{S}_{iw} = 1$ if w settles u, v at level i.
- $\mathbf{X}_{iw} = 1$ if w cuts u, v at level i.

Claim 1 *If the least common ancestor of u and v is at level i, then there is some $w \in V$ that both settles and cuts u, v at level i (i.e. $\mathbf{S}_{iw} = \mathbf{X}_{iw} = 1$).*

Proof. Let w be the vertex such that $\mathbf{S}_{iw} = 1$. Suppose, without loss of generality, that $u \in B(r_{i-1}, w)$. If $v \in B(r_{i-1}, w)$ as well then both u and v would be added to the same set in $\mathcal{C}(i-1)$ which contradicts the fact that i is the level of their least common ancestor. Therefore, $v \notin B(r_{i-1}, w)$ so $\mathbf{X}_{iw} = 1$ as well. ■

Using these tools, we complete the proof of the main result.

Proof of Theorem 1. If the least common ancestor of u and v is at level i, then $T(u, v) \le 2^{i+2}$ by Lemma 1. By Claim 1, there is some $w \in V$ such that $\mathbf{X}_{iw} = 1$ and $\mathbf{S}_{iw} = 1$. Thus,

$$T(u, v) \le \sum_{i=1}^{\log_2 \Delta} \sum_{w \in V} 2^{i+1} \cdot \mathbf{X}_{iw} \cdot \mathbf{S}_{iw}.$$

In other words, by Claim 1 the latter sum includes 2^{i+2} where i is the level of the least common ancestor of u and v.

Therefore,

$$\begin{aligned} \mathrm{E}[T(u, v)] &\le \sum_{i=1}^{\log_2 \Delta} \sum_{w \in V} 2^{i+1} \cdot \Pr[\mathbf{X}_{iw} = \mathbf{S}_{iw} = 1] \\ &\le \sum_{i=1}^{\log_2 \Delta} \sum_{w \in V} 2^{i+1} \cdot \Pr[\mathbf{S}_{iw} = 1 | \mathbf{X}_{iw} = 1] \cdot \Pr[\mathbf{X}_{iw} = 1]. \end{aligned} \tag{24.1}$$

We simplify this sum in two ways. First, we will show that $Pr[\mathbf{S}_{iw} = 1 | \mathbf{X}_{iw} = 1]$ is bounded by some value b_w that is independent of level. If so, then we can bound (24.1) by

$$\mathrm{E}[T(u, v)] \le \sum_{w \in V} b_w \sum_{i=1}^{\log \Delta} \Pr[\mathbf{X}_{iw} = 1] 2^{i+2}.$$

More explicitly:

Lemma 2 *$Pr[\mathbf{S}_{iw} = 1 | \mathbf{X}_{iw} = 1] \le b_w$ for some constant that is independent of level where $\sum_{w \in V} b_w = H_n = O(\log n)$.*

Proof. Sort V by distance to $\{u, v\}$ i.e. $V = \{w_1, w_2, ..., w_n\}$ such that $d(w_j, \{u, v\}) \le d(w_{j+1}, \{u, v\})$.

Fix some w_j. Given that $\mathbf{X}_{iw_j} = 1$, then surely $B(r_{i-1}, w_j) \cap \{u, v\} \ne \emptyset$. But then $B(r_{i-1}, w_{j'}) \cap \{u, v\} \ne \emptyset$ for every $1 \le j' < j$. So if w_j both settles and cuts u, v at level i then π must order w_j before $w_{j'}, 1 \le j' < j$. The probability that w_j is ordered before each $w_{j'}, j' < j$ is exactly $1/j$.

Note that this also true in the conditional distribution (conditioned on the event $\mathbf{X}_{iw_j} = 1$) because the event that w_j is ordered before al $w_{j'}$ is independent of the random choice of radius. Therefore,

$$b_{w_j} := \Pr[\mathbf{S}_{iw_j} = 1 | \mathbf{X}_{iw_j} = 1] \leq 1/j.$$

Finally, we also note that

$$\sum_{w \in V} b_w = \sum_{j=1}^{n} b_{w_j} = \sum_{j=1}^{n} \frac{1}{j} = H_n.$$

■

Lemma 3 *For any vertex $w \in V$, $\sum_{i=1}^{\log \Delta} Pr[\mathbf{X}_{iw} = 1] \cdot 2^{i+2} \leq 16 \cdot d(u,v)$.*

Proof. We assume $d(u,w) \leq d(v,w)$, otherwise we can exchange the roles of u and v in the proof.

Observe $\mathbf{X}_{iw} = 1$ if and only $r_{i-1} = r_0 2^{i-1}$ lies in the half-open interval $[d(u,w), d(v,w))$. Since r_0 is sampled uniformly from $[\frac{1}{2}, 1)$, then

$$\Pr[\mathbf{X}_{iw} = 1] = \frac{|[2^{i-2}, 2^{i-1}) \cap [d(u,w), d(v,w))|}{|[2^{i-2}, 2^{i-1}]|}.$$

$$= \frac{|[2^{i-2}, 2^{i-1}) \cap [d(u,w), d(v,w))|}{2^{i-2}}$$

Therefore,

$$\begin{aligned} \Pr[\mathbf{X}_{iw} = 1] \cdot 2^{i+2} &= 2^{i+2} \cdot \frac{|[2^{i-2}, 2^{i-1}) \cap [d(u,w), d(v,w))|}{2^{i-2}} \\ &= 16 \cdot |[2^{i-2}, 2^{i-1}) \cap [d(u,w), d(v,w))|. \end{aligned}$$

This shows

$$\sum_{i=1}^{\log \Delta} \Pr[\mathbf{X}_{iw} = 1] 2^{i+2} = 16 \sum_{i=1}^{\log \Delta} |[2^{i-2}, 2^{i-1}) \cap [d(u,w), d(v,w))|$$

The union of the disjoint intervals $[2^{i-2}, 2^{i-1})$ covers $[0, \Delta)$, so the last sum is

$$16 \cdot |[d(u,w), d(w,v))| = 16 \cdot (d(w,v) - d(u,w)) \leq 16 \cdot d(u,v).$$

The last step is justified by the triangle inequality $d(w,v) \leq d(u,v) + d(u,w)$. ■

To wrap up, Lemma 2 allows us to bound (24.1) by $\sum_{w \in V} b_w \sum_{i=1}^{\log \Delta} Pr[\mathbf{X}_{iw} = 1] \cdot 2^{i+2}$. By Lemma 3, this is at most

$$16 \cdot d(u,v) \cdot \sum_{w \in V} b_w = 16 \cdot d(u,v) \cdot H_n = O(\log n) \cdot d(u,v).$$

This is what we wanted to show. ■

This tree embedding algorithm was proven by Fakcharoenphol, Rao, and Talwar [FRT04].

An interesting related topic is the following. Given a (not necessarily metric or even complete) graph $G = (V, E)$, let d denote the shortest path metric of G. The goal here is to find a distribution over *spanning trees* T of G such that $\mathrm{E}[T(u,v)] \leq \alpha \cdot d(u,v)$ for the smallest possible α. Abraham, Bartal, and Neiman show a nearly tight bound with $\alpha = O(\log n \cdot \log\log n \cdot (\log\log\log n)^3)$ [ABN08] (nearly tight with the known lower bound of $\Omega(\log n)$).

References

ABN08 I. Abraham, Y. Bartal, and O. Neiman, Nearly tight low stretch spanning trees, In Proceedings of FOCS, 2008.

FRT04 J. Fakcharoenphol, S. Rao, and K. Talwar, A tight bound on ppproximating arbitrary metrics by tree metrics, *Journal of Computer and System Sciences*, 69:485–497, 2004.