15-854: Approximation Algorithms **Lecturer:** Anupam Gupta
Topic: Approximating Metrics by Tree Metrics **Date:** 10/19/2005
Scribe: Roy Liu

12.1 Introduction

In this lecture, we show how to embed metric weighted graphs into distributions of trees. Let $G = (V, E)$ be a metric weighted graph (where the distance function satisfies $d(i, k) \le d(i, j) + d(j, k)$ for all $i, j, k \in V$), and let $\mathcal{D}$ be a probability distribution over trees $\{T = (V, E(T))\}$. We say that a graph G α-probabilistically embeds into a distribution over trees $\mathcal{D}$ if:

- For every tree T in the sample space $\Omega(D)$, for all $u, v \in V$, $d_T(u, v) \ge d_G(u, v)$.
- For every tree $T \in \Omega(D)$, $V(T) = V$.
- $\mathbf{E}[d_T(u, v)] \le \alpha * d_G(u, v)$ for all $u, v \in V$. Expectation is taken over $\mathcal{D}$.

In other words, we want to reduce problems on arbitrary metrics to problems on tree metrics. The α-probabilistic embedding guarantees that our solution on some tree drawn from a distribution is "not too bad" (expected to be within a factor of α) when compared to its representation in the original graph.

12.2 Bartal's Algorithm

How low can we set the distortion factor α so that, for any metric weighted graph G, we can find a distribution on trees $\mathcal{D}$ that allows an α-probabilistic embedding of G? A theorem due to Bartal [1] proposes such an α and a way to construct $\mathcal{D}$ from any metric weighted graph G. Before proving the theorem, however, we introduce a procedure for finding a "Low Diameter Randomized Decomposition" (LDRD for short) of graphs. The randomized, recursive procedure LDRD takes as input a graph $G = (V, E)$ and a parameter $\delta > 0$ and outputs a partition $V_i \vdash V$ such that:

- $\text{diam}(G[V_i]) \le \delta$ for all i, where $\text{diam}(\cdot)$ is the diameter function on graphs.
- $\mathbf{Pr}[\text{edge } \{i, j\} \text{ not in any } E(G[V_k])] \le (d(i, j) * \log n)/\delta$. In other words, if the parts V_i were assigned distinct colors, this is an upper bound for the probability of an edge $\{i, j\}$ being bichromatic.

Assume for now we have such a procedure LDRD. We use it as a subroutine of the recursive algorithm Bartal, which takes as input a metric weighted graph G and outputs a tree T with root r and induced metric:

$\Delta \leftarrow \text{diam}(G)$;
$\mathfrak{G} \leftarrow \text{LDRD}(G, \Delta/2)$;

$\mathfrak{G} = \{G_1, \ldots, G_n\}$ is a partition of G returned by the decomposition.

$\{(T_i, r_i)\} \leftarrow \{\text{Bartal}(G_i)\}$;

Run Bartal on each of the parts G_i to get corresponding rooted trees (T_i, r_i).

$E \leftarrow \bigcup_i E(T_i)$;

$r \leftarrow r_1$;

for each $i \neq 1$ **do**

$E \leftarrow E \cup \{r_i, r\}$;

$d(r_i, r) \leftarrow \Delta$;

end for

$T \leftarrow (V, E)$;

Consolidate each of the T_i's by adjoining r_i to r. Set the newly formed edge to have distance Δ. See Figure 12.2.1.

return (T, r);

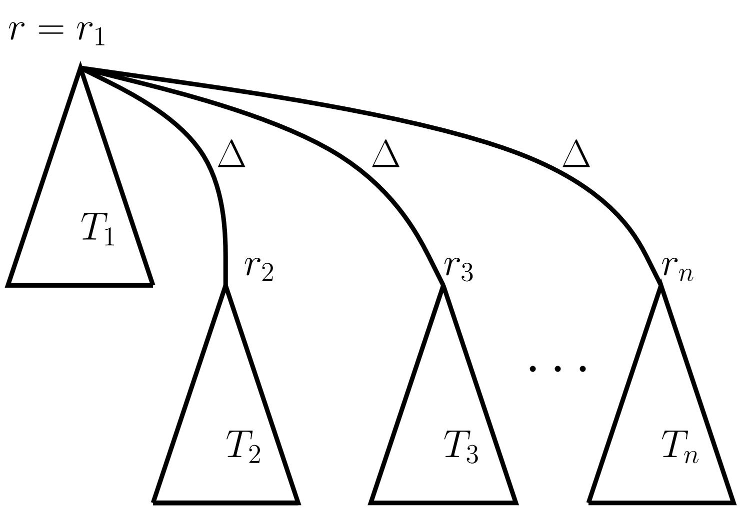

Figure 12.2.1: Consolidating the $T_i's$.

We now prove the main theorem.

Theorem 12.2.1 *Let G be any metric weighted graph. Bartal's Algorithm gives an $O(\log\Delta\log n)$-probabilistic embedding into a distribution $\mathcal{D}$ of trees, where $\Delta = diam(G)$.*

Proof: Let T be a tree returned by Bartal. We first show that the distances in T dominate those in G.

Claim 12.2.2 *For all $u, v \in V$, $d_T(u,v) \geq d_G(u,v)$.*

Proof: We proceed by induction; the base case clearly holds, because Bartal's recursion starts to unwind once it reaches a single vertex. In the inductive step, if u, v belong to the same T_i, then the claim holds by the inductive applied to each of the T_i's. If they don't, then they're separated by at least Δ, by the way T is constructed from the T_i's. Notice that Δ is the diameter of G, and so $d_T(u,v) \geq d_G(u,v)$. ∎

Before we show that Bartal is an $O(\log \Delta \log n)$-probabilistic embedding, we calculate some distances among nodes in the tree T.

Claim 12.2.3 *For all $x \in V(T)$, $d_T(r, x) \leq 2\Delta$. For all $u, v \in V(T)$, $d_T(u, v) \leq 4\Delta$.*

Proof: The claim follows trivially from induction; the base case holds because a single node has diameter 0. In the inductive step, apply the hypothesis to each of the T_i's, which are trees computed from G_i's, the parts resulting from LDRD$(G, \Delta/2)$, to get $d_{T_i}(r_i, x) \leq 2 * (\Delta/2) = \Delta$ for all $x \in V(T_i)$. Let $x \in V(T_i)$ be given, for some T_i. We get $d_T(x, r) = d_{T_i}(x, r_i) + d_T(r_i, r) \leq \Delta + \Delta = 2\Delta$. It follows that for all $u, v \in V$, we get $d(u, v) \leq 4\Delta$: Walk up to the root from u, incurring distance $\leq 2\Delta$, and walk down from the root to v, incurring another distance $\leq 2\Delta$. The total distance incurred is $\leq 4\Delta$. ■

Now we are ready to calculate $\mathbf{E}[d_T(u, v)]$, where expectation is over $\mathcal{D}$, the distribution on trees generated by the randomness of the LDRD subroutine. Suppose inductively that, for T_i's, $\mathbf{E}[d_{T_i}(u, v)] \leq 8 \log \operatorname{diam}(G[V(T_i)]) \log n * d_G(u, v)$. We see that

$$\begin{aligned}
\mathbf{E}[d_T(u, v)] &= \mathbf{E}[d_T(u, v) \,|\, u, v \text{ don't lie in the same } T_i] * \mathbf{Pr}[u, v \text{ don't lie in the same } T_i] \\
&\quad + \mathbf{E}[d_T(u, v) \,|\, u, v \text{ lie in the same } T_i] * \mathbf{Pr}[u, v \text{ lie in the same } T_i] \\
&\leq \mathbf{E}[d_T(u, v) \,|\, u, v \text{ don't lie in the same } T_i] * \mathbf{Pr}[u, v \text{ don't lie in the same } T_i] \\
&\quad + \mathbf{E}[d_T(u, v) \,|\, u, v \text{ lie in the same } T_i] \\
&\leq \frac{d_G(u, v)}{\Delta/2} \log n * 4\Delta + \mathbf{E}[d_{T_i}(u, v)] \\
&= 8 \log n * d_G(u, v) + 8 \log \operatorname{diam}(G[V(T_i)]) \log n * d_G(u, v) \\
&\leq 8 \log n * d_G(u, v) + 8 \log(\Delta/2) \log n * d_G(u, v) \\
&= 8 \log \Delta \log n * d_G(u, v)
\end{aligned}$$

Thus, we have shown that Bartal's algorithm gives an $O(\log \Delta \log n)$-probabilistic embedding for any metric weighted graph G. ■

12.3 Low Diameter Randomized Decomposition

All that remains is to implement the LDRD procedure and prove its important key properties. The procedure goes as follows:

```
R ← random real in [δ/2, δ/4];
π ← random permutation over [n];
{v_1, ... v_n} ← V;
for each v ∈ V do
    c(v) ← NULL;
end for
```

Set all initial colors to NULL.

for each $i \in [n]$ **do**
  for each $v \in B_R(v_{\pi(i)})$ **do**
    if $c(v) = \mathsf{NULL}$ **then**
      $c(v) \leftarrow i$;
    end if

Let $B_r(x) := \{v \in V \mid d(x,v) \le r\}$. It can be thought of as the ball of radius r around x. If the color of v is not assigned, assign it to i.

  end for
end for
return $\{V_i\} \leftarrow \{\{v \in V \mid c(v) = i)\}\}$;

Output each color class as a V_i.

We follow up with a proof.

Claim 12.3.1 *The decomposition algorithm achieves the bound* $\mathbf{Pr}[\textit{edge } e = \{u,v\} \textit{ is bichromatic}] \le (d(u,v) * \log n)/\delta$.

Proof:

We see that e is "first cut by x" if one of e's bichromatic endpoints is colored $c(x)$ *and* $c(x)$ is the first of two colors to be applied to the endpoints. See Figure 12.3.2 for an illustration.

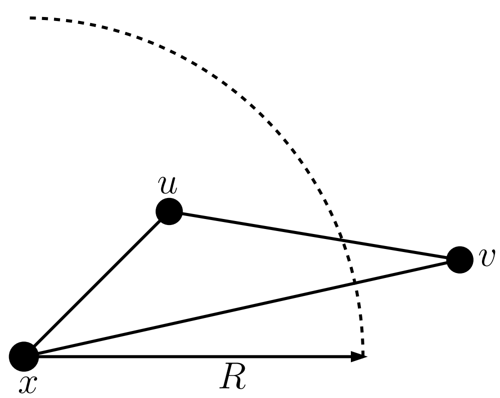

Figure 12.3.2: Suppose $e = \{u,v\}$ is first cut by x.

Define $L_x = \min\{d(x,u), d(x,v)\}$ and $U_x = \max\{d(x,u), d(x,v)\}$. We make a claim about L_x and U_x telling us something about R.

Claim 12.3.2 *If e is first cut by x, then $R \in [L_x, U_x)$ and $R \in [L_x, L_x + d(u,v)]$.*

Proof: $R \ge L_x$ because it must cut an endpoint of e. On the other hand, $R < U_x$ because otherwise e would not be bichromatic – both endpoints would be colored $c(x)$. The second statement follows from the observation that $[L_x, U_x) \subseteq [L_x, L_x + d(u,v)]$. ■

With the claim in mind, fix x. To reduce notational complexity, define $\widehat{U}_x = L_x + d(u,v)$. Sort the vertices in V by non-decreasing L_v value as in Figure 12.3.3.

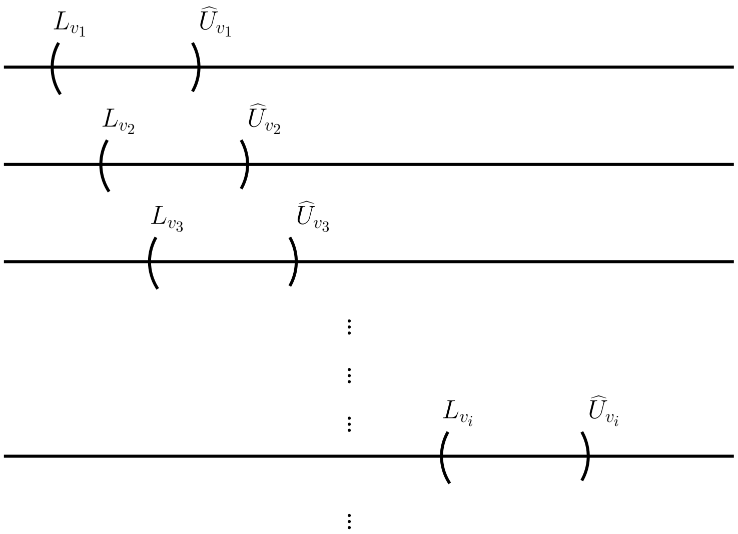

Figure 12.3.3: Sorting the L_v values.

Assign subscripts to the vertices so that L_{v_k} represents the kth smallest value. Well, we know that $x = v_i$, for some i. Since the edge e is cut by x, the vertices $V_{<i} = \{v_1, v_2, \ldots v_{i-1}\}$ must be ordered *after* x by the permutation π (otherwise, they would be first cutting e).

Now we start deriving some probabilities. Define INTERVAL(x) to be the event that $R \in [L_x, U_x)$. Define PRECEDE(x) to be the event that $x = v_i$ precedes everything in $V_{<i}$ under π.

$$\mathbf{Pr}[\text{INTERVAL}(x)] \leq \mathbf{Pr}\Big[R \in [L_x, \widehat{U}_x]\Big] \leq \frac{d(u,v)}{\delta/4}$$

Since R was chosen in $[\delta/4, \delta/2]$, we are simply upper bounding the probability R also lies in $[L_x, L_x + d(u,v)]$. We also have that

$$\mathbf{Pr}[\text{PRECEDE}(x)] = \frac{1}{i}$$

Noting that INTERVAL(x) and PRECEDE(x) are independent events, we get

$$\mathbf{Pr}[e \text{ first cut by } x] = \mathbf{Pr}[\text{PRECEDE}(x)] * \mathbf{Pr}[\text{INTERVAL}(x)] \leq \frac{4d(u,v)}{\delta i}$$

One should check that for e to be first cut by x, the events INTERVAL(x) and PRECEDE(x) are necessary and sufficient. Using a union bound, we get

$$\begin{aligned}\mathbf{Pr}[e \text{ bichromatic}] &\leq \sum_{v \in V} \mathbf{Pr}[e \text{ first cut by } v] = \sum_{i=1}^{n} \mathbf{Pr}[e \text{ first cut by } v_i] \\ &= \sum_{i=1}^{n} \frac{4d(u,v)}{\delta i} = \frac{O(\ln n) * d(u,v)}{\delta}\end{aligned}$$

■

12.4 Conclusions and Related Work

Embeddings into tree metrics are extremely versatile. The existence of an α-approximation scheme implies that without any effort, one has a quick and dirty approximation guarantee of α for any graph problem with metric weights. Bartal's algorithm gets a reasonable value of α, but $\log \Delta \log n$ isn't the best result of type. The optimal algorithm is given in [2], where the authors use ideas from Bartal's paper to get the coveted $O(\log n)$ approximation factor, which is shown to be tight over certain classes of input graphs.

References

[1] Yair Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *FOCS*, pages 184–193, 1996.

[2] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics, 2003.