Network Design foundation
Fall 2011
Lecture 5

Instructor: Mohammad T. Hajiaghayi
Scribe: M. Reza Khani

September 28, 2011

1 Overview

In this Lecture we review the ideas of the celebrated result of Jittat Fakcheroenphol, Kunal Talwar and Satish Rao (FRT) in their paper " A tight bound on approximating arbitrary metrics by tree metrics". They give an algorithm which for a given metric (V, d) finds a probabilistic embedding into tree metrics such that the distortion (stretch) is at most $O(\log n)$. Note that there are some graphs (such as diamond graphs) that have lower bound of $\Omega(\log n)$ for the probabilistic embedding, as a result the FRT bound is tight.

2 Definitions

(1) **r-cut decomposition**
An r-cut decomposition is a partitioning of V into some clusters. Each cluster is centered at a particular node called root such that the distance of each node in the cluster from the root is at most r (as you see later the root is not required to be a member of the cluster).

(2) **Laminar Family**
A laminar family is a collection of subsets $\mathcal{F} \subseteq 2^V$ such that:

$$\forall A, B \in \mathcal{F}, \quad A \subseteq B \text{ (or) } B \subseteq A \text{ (or) } A \cap B = \emptyset$$

(3) **Hierarchal Cut Decomposition**

A hierarchal cut decomposition of (V, d) is a sequence of $\delta + 1$ nested cut decompositions $D_0, \ldots, D_\delta$ such that:

- $D_\delta = V$
- D_i is a 2^i-cut decomposition and a refinement of D_{i+1} (*i.e.* each cluster in D_i is contained within a cluster in D_{i+1}). Note that $D_0 \cup \ldots \cup D_\delta$ makes a laminar family.

3 Intuition and Overview of the Algorithm

Remember from the previous lecture we are going to find a set of trees $T_1, \ldots, T_l$ along with the probability distribution $p_1, \ldots, p_l$, such that each T_i *dominates* metric d and the expected stretch of distance between each pair of nodes u, v is at most $O(\log n)$. *i.e.* $d(u, v) \leq O(\log n)\Sigma p_i d^{T_i}(u, v)$ where d denotes the distance in the original graph and d^{T_i} denotes the distance in the tree T_i.

Instead of giving trees T_is and probabilities p_i, we give a randomized algorithm that produces a tree T such that $d(u, v) \leq O(\log n)E[d^T(u, v)]$.

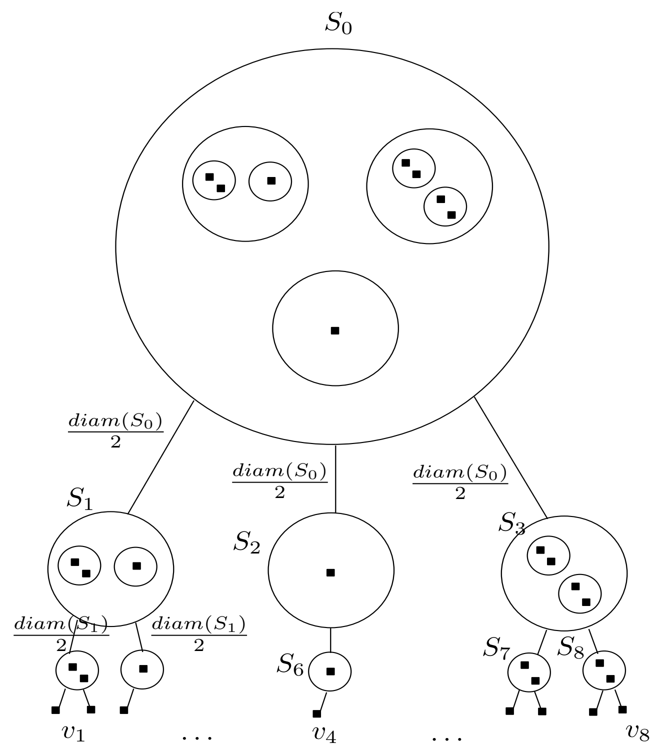

Figure 1: Hierarchical clustering

Figure 1 resembles the structure of T. Filled squares are the actual nodes of the graph and circles show the clustering. Note that T in addition to $v_1, \ldots v_8$ has 9 other nodes representing clusters $S_0, \ldots S_9$. For the sake of simplicity we can assume that the last level only contains the actual nodes of the graph, *e.g.* we add nodes S_2 and S_6 to keep v_4 in the last level.

Without loss of generality we can assume that the minimum distance is 1. Assume the maximum distance (or diameter of the graph) is Δ, and $2^{\delta-1} < \Delta \leq 2^{\delta}$. Note that T is actually a *hierarchical cut decomposition* with the first level being D_δ and the last level being D_0.

We end this section by proving that d^T dominates d.

Lemma 1 *For all* $x, y \in V$, $d^T(x,y) \geq d(x,y)$.

Proof: Suppose $\frac{\Delta}{2^j} \leq d(x,y) \leq \frac{\Delta}{2^{j-1}}$, so x and y cannot be in a component with level less than j. As a result they are separated at a level $j' \geq j$ which implies that at least two edges with weight at least $\frac{\Delta}{2^{j'}}$ are in between them. Thus their distance in T is at least $2\frac{\Delta}{2^{j'}} \geq \frac{\Delta}{2^j} \geq \frac{\Delta}{2^{j-1}}$. ■

4 FRT Algorithm

In this section we provide FRT algorithm which builds a tree T as described in the previous section.

1. Choose a random permutation π of $v_1, \ldots, v_n$.
2. Choose β in the range $[1,2]$ randomly from the distribution $p(x) = \frac{1}{x \ln 2}$.
3. Let $D_\delta \leftarrow \{V\}$ and $i \leftarrow \delta - 1$.
4. while D_{i+1} has non-singleton clusters do
   (a) $\beta_i \leftarrow 2^{i-1}\beta$
   (b) For $l \leftarrow 1, \ldots, n$ do
      i. For every cluster S in D_{i+1}
         A. Create a new cluster consisting of all unassigned vertices in S with distance less than β_i to $\pi(l)$.
   (c) $i \leftarrow i - 1$

5 Bounding the Stretch

In this section we bound the maximum stretch between all pairs of nodes, *i.e.* we prove that for all $u, v \in V$, $E[D^T(u,v)] \leq O(\log n) d(u,v)$.

Definition 1 *An edge* (u,v) *is at level* i *if* u *and* v *are first separated at level* i.

Lemma 2 *If* (u, v) *is at level* i, $d^T(u, v) = 2\Sigma_{j=1}^{i} 2^j = 2^{i+2}$.

Following lemma is a property related to the selection of β.

Lemma 3 *For any* $x \geq 1$, $P\big[$*some* B_i *lies in* $[x, x + dx]\big] \leq \frac{1}{x \ln 2} dx$.

Definition 2 *A node* w *settles an edge* (u, v) *if one of* u *or* v *first assigned to* w*. We say* w *cuts* (u, v) *if* w *settles* u *and* v*, and exactly one of them is assigned to* w*.*

If w cuts u and v at level i then the distance between u and v (we show it by $d^T_w(u, v)$) is $d^T_w(u, v) = d^T(u, v) \leq 2^{i+2}$ (**fact 1**).

Note that $E[d^T(u, v)] = \Sigma_w P[w \text{ cuts } (u, v)] d^T_w(u, v)$.

In the following we prove that $\Sigma_w P[w \text{ cuts } (u, v)] d^T_w(u, v) \leq O(\log n) d(u, v)$ which completes our proof.

Sort vertices according to their distance to either of u or v, let the sorted list be $w_1, \ldots, w_n$. Now we want to calculate $P[w_s \text{ cuts } (u, v)]$, without loss of generality assume u is closer to w_s than v.

There are two conditions in order to w_s cuts (u, v):

(1) $d(w_s, u) \leq \beta_i < d(w_s, v)$.

(2) w_s settles u and v at level i.

As we select a random permutation for the nodes condition (2) happens with probability $\frac{1}{s}$. By Lemma 3 we know that condition (1) happens with probability $\frac{1}{x \ln 2} dx$ for the interval $[x, x + dx]$. Thus, we have:

$$\begin{aligned}
P[w_s \text{ cuts } (u, v)] d^T_{w_s}(u, v) &= \int_{d(w_s,u)}^{d(w_s,v)} \frac{1}{x \ln 2} \frac{1}{s} d^T_{w_s}(u, v) \\
&\leq \int_{d(w_s,u)}^{d(w_s,v)} \frac{1}{x \ln 2} \frac{1}{s} 8x && \text{from the fact that } \beta \text{ is at most 2 and fact 1} \\
&= \frac{8}{s \ln 2} (d(w_s, v) - d(w_s, u)) \\
&\leq \frac{8}{s \ln 2} d(u, v) && \text{triangle property}
\end{aligned}$$

As a result we have:

$$d^T(u, v) \leq \Sigma_w P[w \text{ cuts } (u, v)] d^T_w(u, v) \leq \frac{8}{\ln 2} d(u, v) \Sigma_{s=1}^{n} \frac{1}{s} = O(\ln n) d(u, v)$$