A Tight Bound on Approximating Arbitrary Metrics by Tree Metrics

Jittat Fakcheroenphol  Satish Rao  Kunal Talwar

STOC 2003, JCSS 2004

Presented by Jian XIA

for COMP670P: Topics in Theory: Metric Embeddings and Algorithms
Spring 2007, HKUST

May 8, 2007
Random Tree Embedding

Given a metric \((V, d)\). Let \(S\) be a family of metrics over \(V\), and let \(D\) be a distribution over \(S\). We say that \((S, D)\) \(\alpha\)-probabilistically approximates a metric \((V, d)\), if

- every metric in \(S\) dominates \(d\);
  \((d'(u, v) \geq d(u, v), \text{ for every } u, v \in V \text{ and every metric } d' \in S.)\)
- for every \(u, v \in V\),

\[
E_{d' \in (S, D)}[d'(u, v)] \leq \alpha \cdot d(u, v).
\]

We call \(\alpha\) the distortion.

Question
What is the distortion for probabilistic approximation by dominating trees?
Known Results

- Embedding $C_n$ (unit weight $n$-cycle) into a spanning tree requires distortion at least $n - 1$.
- Embedding $C_n$ into a tree requires $\Omega(n)$ distortion. [Rabinovich and Raz, 95]
- $C_n$ can be embedded into a distribution of dominating trees with distortion $2(1 - 1/n)$. [Karp, 89]
- $2^{O(\sqrt{\log n \log \log n})}$ distortion for graph metrics, using spanning trees. [Alon et al., 95]
- $O(\log^2 n)$ distortion; there exists a graph requiring $\Omega(\log n)$ distortion. [Bartal, 96]
  - Note: Tree metrics can be isometrically embedded into $\ell_1$
- $O(\log n \log \log n)$ distortion [Bartal, 98]
- This paper closes the gap!
- $O(\log^2 n \log \log n)$ distortion for graph metrics, using spanning trees. [Elkin et al., 05]
Hierarchical Cut Decomposition

- assumption: the smallest distance in the given \( n \)-point metric space \((V, d)\) is strictly more than 1; and the diameter of the metric is \( \Delta = 2^\delta \).

- A *hierarchical cut decomposition* of \((V, d)\) is a sequence of \( \delta + 1 \) nested cut decompositions \( D_0, D_1, \ldots, D_\delta \) such that
  - \( D_\delta = \{V\} \),
  - \( D_i \) is a \( 2^i \)-cut decomposition, and a refinement of \( D_{i+1} \).(that is, each set in \( D_{i+1} \) is a disjoint union of some sets of \( D_i \).)

where, given a parameter \( r \), an \( r \)-cut decomposition of \((V, d)\) is a partitioning of \( V \) into clusters, each centered around a vertex and having radius at most \( r \).

- Property
  - the diameter of each cluster in \( D_i \) (referred as *level i cluster*) is at most \( 2^{i+1} \)
  - each cluster in \( D_0 \) is a singleton vertex.
  - a hierarchical cut decomposition naturally corresponds to a rooted tree.
Corresponding tree

- The vertices of the tree have the form \((S, i)\), where \(S \in D_i\), and \(i = 0, 1, \ldots, \delta\).
- The root is \((V, \delta)\).
- The children of a vertex \((S, i)\) are \((T, i - 1)\) with \(T \in D_{i-1}\) and \(T \subseteq S\).
- The edge connecting \((S, i)\) to \((T, i - 1)\) has length \(2^i\).

The tree metric \(d_T\) is the shortest-path metric induced by this tree on the set of its leaves.

- \(d_T\) dominates \(d\).
- upper bound on \(d_T\): Let \(u\) and \(v\) be leaves and \(w\) be their LCA. Let \(l_w\) be the length of the edges from \(w\) to its children. Then, \(d_T(u, v) \leq 4l_w\).
- Steiner points don’t (really) help. (only introducing 4-distortion.) [Gupta, 01; Konjevod et al., 01]
Construct a random hierarchical cut decomposition, and let $T$ be the associated tree.

An edge $(u, v)$ is at level $i$ if $u$ and $v$ are first separated in the decomposition $D_i$.

Thus $d_T(u, v) \leq 4 \cdot 2^{i+1} = O(2^i)$

Since $d_T(u, v) \geq d(u, v)$, $(u, v)$ cannot be at a level $i$ less than roughly $\log d(u, v)$

For $i$ above, we’ll show that the probability $(u, v)$ is at level $i$ decreases geometrically with $i$.

$\mathbb{E}[d_T(u, v)] = \sum_i \Pr[(u, v) \text{ is at level } i] \cdot O(2^i)$
Decomposition Algorithm

**Algorithm** *Partition*(\(V, d\))

1. Choose a random permutation \(\pi\) on \(V\).
2. Choose \(R\) uniformly at random from \([\frac{1}{2}, 1]\).
3. Let \(D_\delta = \{V\}\).
4. for \(i = \delta - 1\) downto 0
5. Let \(R_i = 2^i R\).
6. for \(l = 1, 2, \ldots, n\)
7. for every cluster \(S \in D_{i+1}\)
8. Create a new cluster consisting of all unassigned vertices \(v\) in \(S\) satisfying \(d(\pi(l), v) \leq R_i\)
Illustration
Analysis

- We get a hierarchical cut decomposition
- Now we only need to prove that given an arbitrary edge \((u, v)\), the expected value of \(d_T(u, v)\) is bounded by 
  \(O(\log n) \cdot d(u, v)\)

- \(w\) settles the edge \((u, v)\) at level \(i\) if \(w\) is the first center to which at least one of \(u\) and \(v\) get assigned at level \(i\).
- Note: exactly one center settles any edge \((u, v)\) at any particular level
- \(w\) cuts the edge \(e = (u, v)\) at level \(i\) if it settles \(e\) at this level, and exactly one of \(u\) and \(v\) is assigned to \(w\) at level \(i\).

Define \(\mathbf{E}[d^w_T(u, v)] = \sum_i \mathbf{1}(w\) cuts \((u, v)\) at level \(i) \cdot O(2^i)\)

Note:

\[ \mathbf{E}[d_T(u, v)] \leq \sum_i \Pr[(u, v)\ is\ at\ level\ i] \cdot O(2^i) \leq \sum_w \mathbf{E}[d^w_T(u, v)]. \]
Analysis cont.

- arrange the points $w_1, w_2, \ldots, w_k, \ldots$ in $V$ in increasing order of $\min\{d(u, w_k), d(v, w_k)\}$.
- For $w_k$ to cut $(u, v)$,
  - condition A: $R_i$ must fall in $[d(u, w_k), d(v, w_k)]$ for some $i$.
    (assume $d(u, w_k) \leq d(v, w_k)$)
  - condition B: $w_k$ settles $(u, v)$ at level $i$.

Consider an $x \in [d(u, w_k), d(v, w_k)]$,

$$\Pr[R_i \text{ falls in } [x, x + dx]] \leq \frac{dx}{2^{i-1}} \leq \frac{2}{x} \cdot dx$$

When $A$ is satisfied, any of $w_1, w_2, \ldots, w_k$ can settle $(u, v)$ at level $i$. Therefore, $\Pr[B|A] \leq 1/k$

$$\mathbb{E}[d_{T}^{w_k}(u, v)] \leq \int_{d(u, w_k)}^{d(v, w_k)} 2 \cdot O(x) \cdot \frac{1}{k} \cdot dx = O\left(\frac{d(v, w_k) - d(u, w_k)}{k}\right) \leq O\left(\frac{d(u, v)}{k}\right)$$

Using linearity of expectation, we have

$$\mathbb{E}[d_{T}(u, v)] \leq \sum_w \mathbb{E}[d_{T}^{w}(u, v)] = \sum_k O(d(u, v)/k) = O(\log n) \cdot d(u, v)$$
Second Analysis

Lemma

Given a vertex $u$ and a radius $\rho$, the probability that the ball $B(u, \rho)$ is cut at level $i$ is at most $(\rho/2^{i-2}) \cdot \log n$.

- A set $S$ is cut if there are two clusters in the partition such that vertices from $S$ lie in both these components.
- Given an edge $e = (u, v)$, consider the ball of radius $d(e)$ around $u$. Any partition that cuts the edge $e$ also cuts the ball $B(u, d(e))$. 
Proof of Lemma

Proof:

- arrange the points $v_1, v_2, \ldots$ in $V$ in order of increasing distance from $u$.
- $v_k$ intersects the ball $B(u, \rho)$ if $R_i \in [d(u, v_k) - \rho, d(u, v_k) + \rho]$.
- $v_k$ protects the ball if $R_i > d(u, v_k) + \rho$.
- $v_k$ cuts the ball first at level $i$ if,
  - condition A: $v_k$ intersects the ball — $\Pr[A] \leq 2\rho/2^{i-1}$
  - condition B: no node prior to $v_k$ in the permutation $\pi$ intersects or protects the ball — $\Pr[B|A] \leq 1/k$

\[
\Pr[B(u, \rho) \text{ is cut at level } i] \leq \sum_k \Pr[v_k \text{ cuts } B(u, \rho) \text{ first at level } i] \\
\leq \sum_k \frac{2\rho}{2^{i-1}} \cdot \frac{1}{k} \\
\leq (\rho/2^{i-2}) \cdot \log n
\]
Improvement

Observation

- Since $R_i \in [2^{i-1}, 2^i]$, a node that is closer to $u$ than $2^{i-1} - \rho$ or farther than $2^i + \rho$ cannot cut the ball $B(u, \rho)$ at all.
- We can assume $\rho \leq 2^{i-2}$

\[
\Pr[B(u, \rho) \text{ is cut at level } i] \leq \sum_{k=|B(u,2^{i-1}-2^{i-2})|}^{\frac{|B(u,2^{i}+r^{i-2})|}{|B(u,2^{i-1})|}} \Pr[v_k \text{ cuts } B(u, \rho) \text{ first at level } i] \\
\leq \sum_{k=|B(u,2^{i-2})|}^{\frac{|B(u,2^{i+1})|}{|B(u,2^i)|}} \Pr[v_k \text{ cuts } B(u, \rho) \text{ first at level } i] \\
\leq (\rho/2^{i-2}) \cdot O \left( \log \left( \frac{|B(u, 2^{i+1})|}{|B(u, 2^{i-2})|} \right) \right)
\]
\[ \mathbb{E}[d_T(u, v)] \leq \sum_{i} \Pr[(u, v) \text{ is at level } i] \cdot O(2^i) \]

\[ \leq \sum_{i=0}^{\delta-1} O(2^i) \cdot \Pr[(u, v) \text{ is cut at level } i] \]

\[ \leq \sum_{i=0}^{\delta-1} O(2^i) \cdot \Pr[B(u, d(u, v)) \text{ is cut at level } i] \]

\[ \leq \sum_{i=0}^{\delta-1} O(2^i) \cdot \frac{d(u, v)}{2^{i-2}} \cdot O \left( \log \left( \frac{|B(u, 2^{i+1})|}{|B(u, 2^{i-2})|} \right) \right) \]

\[ = O(\log n) \cdot d(u, v) \]