Network Design and Game Theory
Spring 2008
Lecture 4

Instructor: Mohammad T. Hajiaghayi
Scribe: Asif Jamshed

February 18, 2008

1 Probabilistic Embedding into Trees

In this lecture we introduce the technique of probabilistic embedding of arbitrary graphs into trees and describe some of its applications. We give various definitions, theorems and present approximation algorithms making use of probabilistic embedding. The main benefit that we get out of embedding graphs into trees is that we can then solve the given graph problem on the corresponding tree, which might otherwise have been difficult to solve on general graphs.

2 Definitions

Definition 1 *An* n***-point Metric Space*** (X, d)*is defined by a distance function* $d : X \times X \rightarrow \mathcal{R}^+ \cup \{0\}$ *with the following properties:*

(i) $d(i, i) = 0$ *for all* $i \in X$

(ii) $d(i, j) = d(j, i)$ *for all* $i, j \in X$ *(symmetry)*

(iii) $d(i, j) \leq d(i, k) + d(k, j)$ *for all* $i, j, k \in X$ *(triangle inequality)*

where $|X| = n$.

Two ways of representing an n**-point Metric Space** (X, d) are:

1. An $n \times n$ symmetric matrix M in which each of its entry $a_{ij} = d(i, j)$.
2. A weighted graph $G(X, E)$ where the edges in E represent the distances between each pair of points.

A natural metric on a graph may be considered where the distance between two points (vertices) is the length of the shortest path between the pair. However, a problem with this "natural metric" is that it may not obey the triangle inequality. In that case we simply replace the length of the offending side of the triangle by the sum of the other two sides. For a given adjacency matix M_G we can **complete** the matrix by entering the smallest distance between each pair of vertices. This operation is known as **matrix completion**.

Definition 2 *Given an undirected graph* $G(V, E)$ *and its completed adjacency matrix* M'_G*, the* ***corresponding graph metric of*** G *is* (V, d_G) *such that* $d_G(v_i, v_j) = a_{ij}$*, where* a_{ij} *represents the entry on the* i*-th row and* j*-th column of* M'_G*.*

We can also talk about metrics on other settings. The vertices may be points in:

- L_2^2 (the Euclidean metric) where

$$d((X_1, X_2), (X'_1, X'_2)) = \sqrt{(|X_1 - X'_1|)^2 + (|X_2 - X'_2|)^2}$$

- L_1^2 (the Manhattan metric) where

$$d((X_1, X_2), (X'_1, X'_2)) = |X_1 - X'_1| + |X_2 - X'_2|$$

- L_∞ (the infinity metric) where

$$d((X_1, X_2), (X'_1, X'_2)) = max(|X_1 - X'_1|, |X_2 - X'_2|)$$

In this lecture, however, we would essentially be talking about graphs metrics.

Definition 3 *Given metric spaces* (X, d) *and* (X', d')*, a map* $f : X \to X'$ *is called an* ***embedding****. Furthermore, if this map is distance-preserving (i.e.* $\forall x, y \in X, d(x, y) = d'(f(x), f(y))$*), then it is called an* ***isometric embedding****.*

Why is the study of embeddings useful? It might be the case that problems are easier to solve in a different domain. For example, if a difficult problem on general graphs is considered on a tree into which the graph has been embedded, it might be much simpler to get to a slightly suboptimal solution. It might not always be possible to come up with isometric embeddings of graphs and metrics into trees and this fact contributes to the suboptimality of the solution found on the trees.

Definition 4 *For an embedding* $f : X \to X'$*, the* ***contraction*** *of* f *is the maximum factor by which the distances are shrunk, i.e.*

$$\max_{x,y \in X} \frac{d(x, y)}{d'(f(x), f(y))}$$

Definition 5 *For an embedding $f : X \to X'$, the **expansion** or **stretch** of f is the maximum factor by which the distances are stretched, i.e.*

$$\max_{x,y \in X} \frac{d'(f(x), f(y))}{d(x, y)}$$

Definition 6 *The **distortion** of f, $\alpha(f)$, is the product of the contraction and the expansion of f.*

Distortion is a measure of the "distance" of any embedding from an *isometric embedding.* A large distortion in an embedding is undesirable since we end up paying for it with worse upperbounds and approximation guarantees. Note that distortion is invariant under scaling and by scaling an embedding by a factor of $\frac{1}{\text{contraction}}$ one can get rid of contraction. This latter fact allows us to assume that f is non-contractive and thus $\alpha(f)$ is just the expansion i.e.

$$d(x, y) \leq d'(f(x), f(y)) \leq \alpha d(x, y)$$

where α is the distortion. (When the embedding is isometric, $\alpha = 1$.)

Network problems may be modeled as graphs with constraints such as length, cost, capacity etc. However, it is not always possible to embed graphs into trees isometrically. For example, consider a cycle, C_n, of length n. This cycle cannot be embedded into any tree, T_n, on n vertices with a distortion better than $\Omega(n)$. However, better results can be obtained if one wants to embed the given metric (X, d) into a distribution of trees, instead of one tree, so that the *expected distances* are preserved.

Definition 7 *A metric space (X, d') **dominates** the metric space (X, d) if for evey $u, v \in X$, $d'(u, v) \geq d(u, v)$.*

Definition 8 *The **support** of a probability distribution $\mathcal{D}$ is the closure of the set of all points x, such that $\mathcal{D}(x) > 0$.*

We say that a metric (X, d) embeds probabilistically into a distribution $\mathcal{D}$ of trees with distortion α if:

- every tree $T = (V_T, E_T)$ in the support of $\mathcal{D}$ contains the points of the metric i.e. $X \subseteq V_T$. Furthermore, the distances in T dominate those in d. (Support of $\mathcal{D} = \{T_1, T_2, \ldots, T_k\}$, $\mathcal{D}(T_i) = p_i$ where $k = O(n \log n)$.)

- given a pair of vertices $x, y \in X$, the expected distance is not too much larger than $d(x, y)$, i.e. $E[d_T(x, y)] = \sum_{i=1}^{k} p_i d_{T_i}(x, y) \leq \alpha d(x, y)$

For example, consider a cycle, C_n, with n edges, each of unit length. $\mathcal{D}$ can be obtained from C_n by giving the n paths of length $n-1$, obtained by deleting a different edge every time, an equal probability. Then:

$$\alpha = 1 \times \frac{n-1}{n} + (n-1) \times \frac{1}{n} = 2(1 - \frac{1}{n})$$

Bartal first showed that α is in $O(\log^2 n)$ and then showed that it lies in $O(\log n \log\log n)$. Fakcharoenphol, Rao & Talwar later improved the above result to $O(\log n)$.

Theorem 1 (FRT'03) *Any n-point metric space can be embedded into a distribution over dominationg tree metrics such that the expected expansion of any edge is $O(\log n)$.*

This bound is tight even for simple planar graphs (e.g. the recursive diamond construction.) Note that X may be a strict subset of V_T but by a result of Anupam Gupta '01, the Steiner vertices $(V_T \setminus X)$ may be removed by introducing a constant factor of 8 in the distortion. The edges sets and the lengths may still vary greatly from those in G. However, Elkin, Emek, Spielman & Teng have shown that:

Theorem 2 (EEST'05) *Any n-point metric space can be embedded into a distribution over dominating* ***spanning*** *tree metrics such that the expected expansion of any edge is $O(\log^2 n \log\log n)$.*

Note that in this case domination is trivial. Again the lowerbound is $\Omega(\log n)$ but closing this gap is a very important open problem (the conjectured value is $\Theta(\log n)$.)

Working with **EEST'05** is easier but we can usually get better bounds by using **FRT'03**. Furthermore, the techniques used to prove **FRT'03** are quite simple and useful and we will be discussing its proof in the next session. We now present some applications of embedding.

3 Applications

<u>First Example: Steiner Tree</u>

Given a set of vertices (terminals) $t_1, t_2, \ldots, t_n$, connect them to a root s in an undirected graph $G(V, E)$ in which every edge e has a buying cost $b(e)$.

Algorithm

1. (V, b) should be a metric, otherwise we perform the matrix completion operation to make it a metric.
2. Apply **EEST'05**: we get a distribution $\mathcal{D} = \{p_1, p_2, \ldots, p_k\}$ over the set of spanning trees $T_1, T_2, \ldots, T_k$ of G, where $k = poly(n)$.
3. In each tree there is a unique path from each t_i to s. Buy all the edges on these paths.
4. Output the best solution C_{EEST} with the minimum total cost c found on the tree $T_{c_{EEST}}$.

Since the edge lengths in $T_{c_{EEST}}$ dominate those in G, the solution is the best possible that we can get out of spanning subtrees of G (**FRT'03**, however, may give a better solution.) We now prove that $c \leq (\log^2 n \log\log n)\mathsf{OPT}_G$.

Proof: Let OPT_G be the solution (and its cost) in G. Now for each edge $e = (u, v)$ of OPT_G, we map the edge to the unique path between u and v in $T_{c_{EEST}}$ and any tree T_i where $i \neq c$. Say we obtain a solution OPT_i in T_i where $\mathsf{OPT}_i \geq c_i \geq c$.

$$(\log^2 n \log\log n)\mathsf{OPT}_G \geq \alpha \mathsf{OPT}_G \geq \sum_{i=1}^{k} p_i c_i \geq \sum_{i=1}^{k} p_i c = c \sum_{i=1}^{k} p_i = c \tag{1}$$

■

If we had used **FRT'03** instead, C_{FRT} might not have been a solution in G since $T_{c_{FRT}}$ is allowed to choose non-edges in G. However, again we can make use of the fact that the embedding dominates (V, b) and so the cost of the solution can only go down when we replace some non-edge (u, v) by the shortest path between u and v in G. This then gives us $O(\log n)$-approximation by the above argument. Indeed, we can often use **EEST'05** and **FRT'03** interchangably as long as we can compare OPT_G with C_{FRT}

The above argument also works in a more general setting known as the **generalized Steiner tree** or the **Steiner forest** problem where instead of just one root, s, we are given pairs (t_i, s_i) which we want to connect. The same argument goes through to give the desired approximation factors. It should be kept in mind, however, that there exist 1.55- and $(2 - \frac{1}{n})$-approximations for the Steiner tree and the generalized Steiner tree problems, respectively, that do not employ the above methods.

The techinques that we used above cannot be used with directed graphs and the best approximation for the directed Steiner tree problem is n^{ϵ}. Similarly, the directed generalized Steiner tree problem has a hardness of $\Omega(2^{\log^{1-\epsilon} n})$.

Second Example: Group Steiner Tree

Given a collection of groups $X = \{S_1, S_2, \ldots, S_p\}$, $S_i \subseteq V$, find a minimum cost tree that connects at least one vertex of each S_i in the undirected graph $G(V, E)$ in which every edge e has a buying cost $b(e)$.
We again use probabilistic embedding along with the following non-trivial result:

Theorem 3 *Group Steiner tree problem on trees may be approximated within a factor of* $\Theta(\log^2 n)$.

We proceed with the algorithm in the same manner as before. We can see that it gives a solution with cost c that is at most $O(\log n)$ larger in expectation on each tree (as in the case for **FRT'03**, the optimal solution on a tree can be embedded back into the original graph with at most the same cost.) This optimal solution, in turn, is only a $\Theta(\log^2 n)$-approximation by the above theorem.

Thus, by similar argument as before, we get the following inequality which is analogous to 1:

$$(\log^2 n \times \log n)\mathsf{OPT}_G = (\log^3 n)\mathsf{OPT}_G \geq c \tag{2}$$

This time, however, it is the best known result for the group Steiner problem.