Fall 2006, CS 598CC: Approximation Algorithms
Homework 3
Due: 11/28/2006 in class

Instructions and Policy: You are not allowed to consult any material outside of the textbook and class notes in solving these problems. Unless otherwise noted you are to solve the problems on your own. Even when collaboration is allowed, each person should write up their own solutions independently. You need to indicate the names of the people you discussed a problem with. Ideally you should discuss with no more than two other people.

Please write clearly and concisely - clarity and brevity will be rewarded. Refer to known facts as necessary. Your job is to convince me that you know the solution, as quickly as possible.

For this homework collaboration is allowed on all problems.

Embedding a finite metrics into tree: For Problems 1 and 2 we will use probabilistic tree embeddings. You can assume that there is a randomized algorithm that given a n point finite metric on a set V gives a tree T on V such that

- $d_T(uv) \geq d_G(uv)$ for all $u, v \in V$.
- $E[d_T(uv)] \leq \alpha(n) d_G(uv)$ for all $u, v \in V$.

We showed $\alpha(n) = O(\log^2 n)$ in class but the best known result gives $\alpha(n) = O(\log n)$. However we can leave it as a a parameter in the two problems.

Problem 1 [20 pts] Consider the (uniform) buy-at-bulk problem. In this problem we are given an undirected edge-weighted graph $G = (V, E)$ and k pairs of nodes $s_1t_1, s_2t_2, \ldots, s_kt_k$. Each pair has a non-negative demand d_i. Further we are given some p cable types; cable type i has a capacity u_i and cost c_i. Assume that $u_1 < u_2 < \ldots < u_p$ and $c_1 < c_2 < \ldots < c_p$. These cable types exhibit economies of scale, the higher capacity cables are cheaper per unit capacity; formally, $c_i/u_i < c_{i-1}/u_{i-1}$ for $1 < i \leq p$. The goal is to select, for each pair s_it_i, a path P_i that connects them and route the pair's demand d_i on P_i. The total flow routed on an edge e is $f(e) = \sum_{i:e\in P_i} d_i$. To support this flow we need to install cables on the edges; on edge e the total capacity of the cables installed should be at least $f(e)$. Note that multiple copies of a cable might need to be installed for this purpose. To install a cable of type i on an edge e it costs $\ell(e)c_i$ where $\ell(e)$ is the length of the edges e. The problem involves finding the paths P_i and installing the necessary cables to minimize total cost.

- Suppose you are given the paths P_i and the only goal is to install cables of cheapest cost to support the flow on the paths.. Give a constant factor approximation to minimize the cost of installing cables. Note that you can do this separately for each edge e. (Why is this problem hard?).

- Use above and the tree embedding result to obtain a randomized $O(\alpha(n))$ approximation to the problem. Carefully argue why you can use the tree embedding result.

Problem 2 [20 pts] Let $T = (V, E)$ be a tree on V and let $w : E \to \mathcal{R}^+$ denote edge lengths on T. The shortest path distances on T induce a metric d_T on V. Prove that d_T is an ℓ_1 metric. Use this and the tree embedding result to show that any n point metric can be embedded into ℓ_1 with $\alpha(n)$ distortion. Is the resulting embedding a contraction or an expansion? Give a randomized approximation algorithm with ratio $\alpha(n)$ for the sparsest cut problem in general graphs using this observation: use the LP relaxation we discussed in class and instead of using Bourgain's embedding, use the tree embedding. Why do you not need to compute the embedding explicitly?

Problem 3 [10 pts] Recall the Set Cover problem and the Maximum Coverage problem. Assume that for every fixed $\epsilon > 0$, it is NP-hard to obtain a $(1 - \epsilon) \ln n$ approximation for Set Cover where n is the number of elements in the given instance. Prove that for every fixed $\epsilon > 0$, it is NP-hard to obtain a $(1 - 1/e + \epsilon)$ approximation for Maximum Coverage.