Fall 2006, CS 598CC: Approximation Algorithms
Homework 4/Take Home Final
Due: 12/12/2006

Instructions and Policy: You are not allowed to consult any material outside of the textbook and class notes in solving these problems. No collaboration is allowed for this homework, this is a take home final.

Please write clearly and concisely - clarity and brevity will be rewarded. Refer to known facts as necessary. Partial credit will be given for your ideas so write them down even if you cannot figure out the full solution.

The test is for 50 pts. The intention is that one of the two problems, 4, 5 is optional. However you can choose to do both for extra credit or do them both in place of other problems.

Problem 1 [10 pts] In the k-MST problem you are given an undirected edge-weighted graph $G = (V, E)$ with edge weightes $c : E \to \mathcal{R}^+$ and an integer k. The goal is to find a tree $T = (V_T, E_T)$ in G of smallest edge weight ($\sum_{e \in E_T} c(e)$) such that $|V_T| \geq k$. Show that if there is an α-approximation for k-MST then there is an α-approximation for the Steiner tree problem. Recall the Steiner tree problem. The input is an edge-weighted graph $G = (V, E)$ and a set of terminals $S \subseteq V$ and the goal is to find a tree T of minimum edge-weight that connects (contains) all the terminals S. (Hint: reduce the Steiner tree problem to the k-MST problem by using many dummy vertices at each Steiner vertex and choosing k appropriately large).

Problem 2 [10 pts] We consider a multi-dimensional generalization of bin-packing. We are given n non-negative vectors $v_1, v_2, \ldots, v_n$ in R^d. Let $v_i(j)$ be the jth coordinate of v_i. The vectors satisfy the property that $v_i(j) \leq 1$ for $1 \leq i \leq n, 1 \leq j \leq d$. The goal is to pack these vectors into the least possible number of d-dimensional bins where a bin is a d-dimensional vector v with $v(j) = 1$ for $1 \leq j \leq d$. A set of vectors S fit into a bin v if the sum of vectors in S is at most 1 (in each coordinate). Obtain an $O(d)$ approximation for this problem. (Hint: Analyze the simple greedy algorithm.)

Problem 3 [10 pts] Given a graph $G = (V, E)$ with edge-weights $c : E \to \mathcal{R}^+$, you wish to partition G into $G_1 = G[V_1], G_2 = G[V_2], G_3 = G[V_3]$ such that $|V_i| \leq \lceil |V|/3 \rceil$ for $1 \leq i \leq 3$, and the cost of the edges between the partitions is minimized. Note that $|V_i| \geq \lfloor |V|/3 \rfloor$. Using an α-approximation for the sparsest cut problem, give a pseudo-approximation for this problem where you partition the graph into 3 pieces $G[V_1'], G[V_2'], G[V_3']$ such that $|V|/c_2 \leq |V_i'| \leq |V|/c_1$ for some constants $1 < c_1 < c_2$ and the cost of the edges between the partitions is $O(\alpha)\textsc{opt}$. What constants c_1, c_2 can you guarantee? Note that c_1 and c_2 should be independent of the graph size. (Hint: this problem is similar to the one on partitioning into two pieces that we covered in class and is in the book on applications of sparsest cut (Section 21.6.3).)

Problem 4 [20pts] Consider the multicut problem when the underlying graph is a tree. That is, we are given an edge-weighted tree $T = (V, E)$ (edge e has weight/cost c_e) and k node pairs $s_1t_1, s_2t_2, \ldots, s_kt_k$. The goal is find a minimum weight set of edges whose removal separates each pair s_it_i. Consider the LP relaxation for this problem that we discussed in class (also see Chapters 18 and 19 in Vazirani's book) which has variables d_e for each edge and constraints that state that the distance between s_i and t_i is at least 1 in the metric imposed by d.

- Suppose the tree is rooted at some vertex r and in this rooting, for each pair s_it_i, either s_i is an ancestor of t_i or t_i is an ancestor of s_i. Prove that in this case the LP gives an optimum integral solution by using the fact that the matrix obtained from the LP is a network matrix and hence totally unimodular (see class notes of Lecture 11). **You can do the second part even if you cannot figure this part out.**

- Use above to obtain a 2 approximation for the tree case as follows. Root the tree arbitrarily. For each pair s_it_i let v_i be the least common ancestor of s_i and t_i (v_i could be one of s_i, t_i). Solve the LP to obtain an optimum feasible solution d^*. Obtain a new solution d' where $d'_e = \min\{1, 2d^*_e\}$. Argue that one of s_i or t_i is separated fractionally from v_i in the new solution d'. Now apply the first part to obtain a 2 approximation for the orignal problem.

Problem 5 [20 pts] We consider a problem that combines aspects of the Steiner tree problem and Set Cover. We are given an undirected graph $G = (V, E)$ with edge weights given by $c : E \rightarrow \mathcal{R}^+$. We are also given groups of vertices $S_1, S_2, \ldots, S_k$ where each $S_i \subseteq V$. Assume without loss of generality that the S_i are *disjoint* (that is $S_i \cap S_j = \emptyset$ for all $i \neq j$). The goal is to find a minimum edge-cost tree $T = (V_T, E_T)$ such that each T contains at least one vertex from each S_i; that is $V_T \cap S_i \neq \emptyset$ for each $1 \leq i \leq k$. To simplify the problem we can consider the rooted version. We have a root vertex r that needs to be in the tree T. Let $S = \cup_i S_i$ denote the terminals.

- Write an LP formulation for the above problem using two types of variables; x_e for whether an edge e is in the tree or not, and a variable y_v for each $v \in S$. Does your LP have exponential number of constraints? Can you show a separation oracle for it? (Hint: use the ideas behind the Steiner tree/forest LP except now you have additional variables y_v. You only need to consider cuts that separate root from a vertex $v \in S$). **If you cannot do this first part, you can come to me and obtain the LP formulation to be used for the second part. You will forfeit 10 pts.**

- We now want to use the above LP formulation to obtain an approximation algorithm. Let $f = \max_i |S_i|$. Obtain a $2f$ approximation as follows. Obtain an optimum fractional solution x^*, y^* to the above LP. Obtain a new LP solution x', y' by scaling as follows; set $x'_e = \min\{1, fx*_e\}$ and $y'_v = \min\{1, fy_v\}$. Use the information in the solution x', y' and the fact that the Steiner forest (which generalizes Steiner tree) LP has an integrality gap of 2 to obtain the desired $2f$ approximation. (Hint: For each group S_i there should be some $v \in S_i$ such that $y'_v = 1$.)