Homework III, Approximation Algorithms and Hardness of Approximation 2013

Due on Tuesday April 30 at 16.15 (send an email to alantha.newman@epfl.ch). Solutions to many homework problems, including problems on this set, are available on the Internet, either in exactly the same formulation or with some minor perturbation. It is *not acceptable* to copy such solutions. It is hard to make strict rules on what information from the Internet you may use and hence whenever in doubt contact Ola Svensson. You are, however, allowed to discuss problems in groups with up to three students, but solutions should be handed in individually and please write with whom you have collaborated.

1 (30 pts) In class, we showed an approximation algorithm for the facility location problem assuming triangle inequality on the costs of connecting the facilities. Consider the facility location without triangle inequality. Give an approximation algorithm for this problem with the best guarantee you can. Do you think there exists a better approximation guarantee? Why or why not?

2 (35 pts, Problem 24.12 in book of Vazirani) This exercise shows that we can apply the method of *dual fitting* to obtain another factor-3 approximation algorithm for the metric uncapacitated facility location problem.

Consider the following modification to Algorithm 24.2 (the algorithm we saw in class). As before, dual variables, α_j, of all unconnected cities are raised uniformly. If edge (i, j) is tight, β_{ij} is raised. As soon as facility, say i, is paid for, it is declared open. Let S be the set of unconnected cities having tight edges to i (recall that an edge is *tight* if $\alpha_j = c_{ij} + \beta_{ij}$)). Each city $j \in S$ is declared connected and stops raising its α_j. So far, this algorithm is the same as Algorithm 24.2. The main difference appears now: Each city $j \in S$ withdraws its contribution from other facilities, i.e., for each facility $i' \neq i$, set $\beta_{i'j} = 0$. When all cities have been declared connected, the algorithm terminates. Observe that each city contributes towards the opening of at most one facility–the facility to which it is connected.

2a This algorithm has a simpler description as a greedy algorithm. Provide this description. **Hint:** Use the notion of cost-effectiveness defined for the greedy set cover algorithm.

2b Let i be an open facility and let $\{1, \ldots, k\}$ be the set of cities that contributed to opening i at some point in the algorithm. Assume wlog that $\alpha_1 \le \alpha_j$ for $j \le k$. Show that for $j \le k$, $\alpha_j - c_{ij} \le 2\alpha_1$. Also, show that

$$\sum_{j=1}^{k} \alpha_j \le 3 \sum_{j=1}^{k} c_{ij} + f_i. \tag{1}$$

Hint: Use the triangle inequality and the following inequality which is a consequence of the fact that at any point, the total amount contributed for opening facility i is at most f_i:

$$\sum_{j: c_{ij} \le \alpha_1} \alpha_1 - c_{ij} \le f_i. \tag{2}$$

2c Show that $\alpha/3$ is a dual feasible solution.

2d How can this analysis be improved? Give the best improvement you can.

3 (35 pts) Let $T = (V, E)$ be a tree on V and let $w : E \rightarrow \mathcal{R}^+$ be the edge lengths. Consider the shortest path distances on T and note that these distances form a metric d on V.

3a Prove that d is an ℓ_1 metric.

3b Suppose there is an embedding mapping any metric to a tree metric with $\alpha(n)$ distortion. Give a randomized α-approximation algorithm for the sparsest cut problem using this assumption.