CO453: Network Design – Winter 2007

Instructor: Chaitanya Swamy

Assignment 5

Due: Monday, March 26, 2007 after class

You must give a proof of correctness of any algorithm you design, and argue briefly why it runs in polynomial time. You may use any proof or algorithm covered in class directly.

Q1: Consider the uncapacitated facility location (UFL) problem: we have a set $\mathcal{F}$ of facilities with opening costs f_i, a set $\mathcal{C}$ of clients, and assignment costs c_{ij} for every facility-client pair (i, j). We phrased this problem as one of deciding which facilities to open, and how to assign clients to open facilities. But observe that once we have decided to open a set $F' \subseteq \mathcal{F}$ of facilities, it is easy to assign clients to facilities — each client $j \in \mathcal{C}$ simply gets assigned to the facility $i \in F'$ for which c_{ij} is minimum.

Now suppose that we have a fractional solution to the LP-relaxation of UFL and we only know the values y_i^* of the facility-opening variables. Show how one can compute efficiently the values x_{ij}^* of the client-assignment variables and get a ("complete") fractional solution (x^*, y^*) such that, (total cost of (x^*, y^*)) $= \min_{x:(x,y^*) \text{ is a feasible fractional solution}}$ (total cost of (x, y^*)). **(10 marks)**

Q2: Recall the LP-relaxation of UFL. Throughout, we use i to index the facilities in $\mathcal{F}$ and j to index the clients in $\mathcal{C}$. The distances c_{ij} form a metric.

$$\min \quad \sum_i f_i y_i + \sum_{j,i} c_{ij} x_{ij} \qquad \text{(UFL-LP)}$$

$$\text{s.t.} \quad \sum_i x_{ij} \geq 1 \qquad \text{for all } j \qquad (1)$$

$$x_{ij} \leq y_i \qquad \text{for all } i, j$$

$$x_{ij}, y_i \geq 0 \qquad \text{for all } i, j.$$

In class, we gave an LP-rounding 4-approximation algorithm that used the optimal solution to the dual of (UFL-LP). Here, we will give a primal LP-rounding approximation algorithm with the same approximation guarantee. Let (x^*, y^*) be an optimal solution to (UFL-LP). Define $F_j^* = \{i : x_{ij}^* > 0\}$, and let $C_j^* = \sum_i c_{ij} x_{ij}^*$ denote the cost that the LP pays for assigning client j to an open facility.

Let $\alpha \in (0, 1)$ be a parameter that we will set later. For each j, let the facilities in F_j^* be sorted in increasing order of c_{ij} (assume that all c_{ij} are distinct for convenience). Let $i_j \in F_j^*$ be the first facility (according to the ordering) such that $\sum_{i \in F_j^*: i \leq i_j} x_{ij}^* \geq \alpha$. Let $N_j^\alpha = \{i \in F_j^* : i \leq i_j\}$ and $R_j^\alpha = c_{i_j j}$. Note that $R_j^\alpha = \max_{i \in N_j^\alpha} c_{ij}$.

(a) Prove that $R_j^\alpha \leq C_j^*/(1 - \alpha)$. **(5 marks)**

(Hint: If this is false, use the fact that $\sum_{i \geq i_j} x_{ij}^* > 1 - \alpha$ to obtain a contradiction.)

(b) Consider the solution $(\tilde{x}, \tilde{y})$ where $\tilde{y} = y^*/\alpha$, and $\tilde{x}_{ij} = x_{ij}^*/\alpha$ if $i \in N_j^\alpha$ and 0 otherwise. Argue that $(\tilde{x}, \tilde{y})$ is a feasible solution to (UFL-LP) (constraints (1) may be satisfied at strict inequality). **(3 marks)**

Clearly, we have $\sum_i f_i\tilde{y}_i = \left(\sum_i f_i y_i^*\right)/\alpha$. We now perform the same clustering step as in the LP-rounding algorithm covered in class, but we now consider clients in increasing order of R_j^α. Let L be the list of all clients ordered by increasing R_j^α. Initialize $\mathcal{C}' \leftarrow \emptyset$. While $L \neq \emptyset$, we do the following: (i) let j be the first client in L; add j to $\mathcal{C}'$ and remove j from L. (ii) For every client $k \in L$ such that $N_k^\alpha \cap N_j^\alpha \neq \emptyset$, set $\mathsf{nbr}(k) \leftarrow j$ and remove j from L.

Notice that the sets N_j^α are pairwise-disjoint for clients $j \in \mathcal{C}'$.

(c) Prove that if $k \in \mathcal{C} \setminus \mathcal{C}'$ and $j = \mathsf{nbr}(k)$ (which is in $\mathcal{C}'$) then $R_j^\alpha \leq R_k^\alpha$ and $c_{jk} \leq R_j^\alpha + R_k^\alpha \leq 2R_k^\alpha$. **(4 marks)**

Now for every $j \in \mathcal{C}'$, we open the cheapest facility in N_j^α, and assign j and all clients $k \notin \mathcal{C}'$ for which $\mathsf{nbr}(k) = j$ to this facility.

(d) Prove that the facility-opening cost incurred is at most $\sum_i f_i\tilde{y}_i$. Prove that the assignment cost of a client j is at most R_j^α if $j \in \mathcal{C}'$, and at most $3R_j^\alpha$ otherwise. Using these facts, and the relationship between $\sum_i f_i\tilde{y}_i$ and $\sum_i f_i y_i^*$, and R_j^α and C_j^* proved earlier, argue that the algorithm is a $\max\left(\frac{1}{\alpha}, \frac{3}{1-\alpha}\right)$-approximation algorithm. Thus, setting $\alpha = \frac{1}{4}$ gives a 4-approximation primal-rounding algorithm. **(8 marks)**

Q3 (Bonus question): Adapt the Jain-Vazirani primal-dual algorithm to give a 3-approximation algorithm for metric UFL with penalties. Recall that in this problem, we are allowed to not assign a client j to a facility by incurring a penalty p_j. Thus, a solution needs to decide which facilities are opened, which clients are assigned to open facilities and hence, which ones incur penalty, and how to assign the non-penalty clients to open facilities. **(10 marks)**