Approximation Algorithm

Prof. Palash Dey

Department of Computer Science and Engineering

Indian Institute of Technology, Kharagpur

Week – 06

Lecture 30

Lecture 30 : A 4 Factor Approximation Algorithm for Uncapacitated Facility Location Problem contd.

Welcome. So, in the last class we have started seeing a four factor approximation algorithm for uncapacitated facility location problem. We have written down the linear programming relaxation and worked out the dual. So, let us continue from there. facility location problem.

So, let us write down the primal LP. minimize $\sum f_i y_i$, y_i is a variable for every facility $i \in F$ it is 1 if the facility is open otherwise it is 0 plus $\sum \sum c_{ij} x_{ij}$ subject to each demand $j \in D$ is connected to exactly one and each facility each client or demand can be assigned to a facility only if it is open. x_{ij} can take value 1 only if y_i takes value 1, if y_i takes value 0 x_{ij} should take value 0. So, this is encoded using this constraint that x_{ij} is less than equal to y_i

And of course, we have all the variables x_{ij} greater than equal to 0 for all $i \in F$ for all $j \in D$ and y_i is greater than equal to 0 for all $i \in F$. So, this is the primal LP let me write down the dual LP. the variables for these constraints corresponding to this constraints let us call it v_j and these constraints the variables corresponding let us call it w_{ij}. So, maximize summation v_j, $j \in D$ subject to for all facility $i \in F$ $\sum w_{ij} \leq f_i$ and for all facility $i \in F$ for all client $j \in D$ we have $v_j - w_{ij} \leq c_{ij}$ and the variables w_{ij} can take only non-negative recall because $v_j$corresponds to equalities v_j can take negative value this is the place where is stopped in the last class. So, this algorithm is based on deterministic rounding.

So, we solve both primal and dual and get hold of a primal and dual optimal solutions that is the first step solve both primal and dual LP. Let (x^*, y^*) be a primal optimal solution and (v^*, w^*) dual optimal solution. So, recall we have variables x_{ij} if it takes value 1, if it if client j is assigned to facility i otherwise it takes value 0 that is in the integer linear programming formulation. In this relaxed LP they can take fractional values. But, if x_{ij} is text value 0 still it this LP solution indicating us to not assign client j to facility i if x_{ij} equal to 0.

So, using that we define a concept called neighborhood, neighborhood of clients. with respect to this optimal solutions, we say that facility i is a neighbor of client j if $x_{ij}^{*}>0$. We define set of neighbors of a client j to be the set of facilities $i \in F$ such that $x_{ij}^{*}>0$. This is a subset of F, j belongs to D. Why do we care about neighboring facilities? So, here is the lemma.

So, again this lemma with is with respect to this optimal solutions of primal and dual. So, if $x_{ij}^{*}>0$ for any $i \in F, j \in D$, then the assignment cost c_{ij} is less than equal to v_j^{*}. So, if $x_{ij}^{*}>0$ then in some sense the cost of assigning demand j to facility i is upper bounded by v_j^{*}. First let us see the proof and then we will see why we care about this what is the use of this lemma proof. it follows directly from complementary slackness.

So, let us recall complementary slackness with respect to this primal dual LPs, it says that whenever a primal variable is non-negative then the corresponding dual constraint is tight and vice versa. Whenever the dual variable is non-zero then the primal constraint is tight if it holds this conditions hold if and only if the primal and dual assignments are optimal assignments. So, in particular in one direction of complementary slackness says that if I start with optimal solutions $\left(x^{*}, y^{*}\right)$ and $\left(v^{*}, w^{*}\right)$ of primal dual LPs, whenever a primal variable x_{ij}^{*} is positive then the corresponding dual constraint is tight. So, this dual constraint $v_j - w_{ij} \leq c_j$ this holds with equality that is the meaning of saying that a constraint or inequality is tight. So, right from complementary slackness $x_{ij}^{*}>0$ implies that the corresponding dual constraint is tight $v_j - w_{ij} = c_j$

Now see that w_{ij}'s are non-negative. So, from here we conclude that c_{ij} is less than equal to v_j^{*} all these are star this holds with respect to optimal solutions of primal and dual LPs. Now you see why do we care? So, if by some reason we are able to connect each client to its neighbouring facility then the total assignment cost is at most opt let us write if we are able to connect each client to its to one of its . neighboring facility, then total assignment cost. You see if client j is assigned to a facility i which is neighbouring that is same as saying $x_{ij}^{*}>0$, then the assignment cost for client j is at most v_j^{*}.

So, the total assignment cost and if this holds for all the clients in the total assignment cost of the clients. at most $\sum v_j^{*}$. Now, what is the connection between $\sum v_j^{*}$ and opt? For here you see $\sum v_j^{*}$ is optimum value of dual LP and because we have started with optimal solutions of primal LP and dual LP from weak duality theorem, the any solution to dual LP the corresponding value is a lower bound on primal opt. So, this is less than equal to LP opt which is less than equal to opt. So, the assignment costs are same are at most of total assignment cost.

Unfortunately, it may not be possible to assign each client to its neighboring facility because for some facilities the facility opening cost may be high. But, nevertheless we will see that we will use the metric property and triangle inequality to cleverly open some subset of facilities and assign the clients to them in a such a way that the assignment cost does not shoot up much. That idea we will implement next to get a constant factor approximation algorithm. So, what is the idea for that here is the next idea to open low cost facilities. First let us see the idea and then we will see how to implement them.

So, suppose we can partition some subset $F' \subseteq F$. of the facilities into sets F_k where F_k is the neighbouring facility of some client j_k. So, here is a set of facilities F. Now, if I have some clients say j_1. and its neighbour is this is $N(j_1)$, I have j_2 and its neighbour $N(j_2)$. and so on. So, important part is that $N(j_1)$, $N(j_2)$, $N(j_3)$ they forms a partition of some facilities some subset of facilities f prime in particular they should not overlap. $N(j_1)$ should be disjoint with $N(j_2)$, $N(j_3)$ and so on. If this thing if this holds then if we open the cheapest facility or a cheapest facility if it is not unique $i_k \in N(j_k)$. So, for every neighbouring sets $N(j_1)$, $N(j_2)$ and so on from each set I open the cheapest facility, then we can bound the cost of i_k as what is the cost of i_k that is f_{i_k}, again f_{i_k} I write it as f_{i_k} times 1 and this is f_{i_k}.

Now, there is a primal constraint that $\sum x_{ij} = 1$ this I use in place of 1 this $\sum_{i \in N(j_k)} x^*_{i j_k}$. ok. And because i_k is the cheapest facility that in $N(j_k)$ that means, among all f_{i_k} s f_i among all facilities the f value the facility opening cost is cheapest in $N(j_k)$ that is the i_k this is less than equal to $\sum_{i \in N(j_k)} f_i x^*_{i j_k}$. Why this is the case? Because for all $i \in N(j_k)$, i_k is the cheapest facility that means, f_{i_k} is less than equal to f_i.

So, what have we achieved then? So, if we only open facilities in this disjoint neighborhoods, then we will see that the total facility opening cost is small. So, total facility opening cost. If we only open facilities from this disjoint neighborhood suppose this is up to $N(j_t)$ is there, then the total facility opening cost is $\sum_{k=1}^{t} f_{i_k}$. Now, for each f_{i_k}, f_{i_k} is less than this. So, this inequality I use this is less than equal to $\sum_{i \in N(j_k)} f_i x^*_{i j_k}$ ok.

Till now we have not used the fact that this $N(j_1)$, $N(j_2)$ up to $N(j_k)$ they are non overlapping. Now, we are going to use it because they are non overlapping this is less than equal to or this is equal to $\sum_{i \in F'} f_i x^*$. What is F'? $F' = \cup_{k=1}^{t} N(j_k)$. Now, here we need to do one more step first is $N(j_k), k=1,\ldots,t$ because I am I need to get rid of this k here.

So, for that what I do is x this inequality I use that $x_{ij} \leq y_i$. So, means $x_{ij}^* \leq y_i$. Now, this terms does not depend on k, now I write $F' = \cup_{k=1}^{t} N(j_k)$ and use the fact that $N(j_k)$ s are disjoint. So, this is then $\sum_{i \in F'} f_i y_i^*$ ok. So, here because $N(j_k)$ s are disjoint ok.

So, this is less than equal to $\sum_{i \in F'} f_i y_i^*$. I include the other facilities also f other facilities in f which are not in f prime this holds because y_i^* they take value non negative y_i^* is greater than equal to 0. Now, what is this value? Again you look at the primal because x_{ij} is non negative this is at most primal opt. So, this is less than equal to LP opt which is less than equal to opt. So, you see if by somehow I get hold of some clients $j_1, j_2, \ldots, j_t$ in such a way that their neighbors neighboring facilities do not overlap and I ensure that or the I open facilities only from those neighboring facilities, then my total facility opening cost is at most So, let us write it down.

So, if we can find clients $j_1, j_2, \ldots$ such that $N(j_1), N(j_2), \ldots, N(j_t)$ are disjoint and we open cheapest facility ah cheapest facility from each $N(j_1), N(j_2), \ldots, N(j_t)$ that means, in total I am opening t facilities from each $N(j_1), N(j_2), \ldots, N(j_t)$ I open one facility which is the chip which is our cheapest facility. Then the total facility opening cost is at most. So, in the next lecture we will implement this idea, we will see how we can open the facilities in such a way in this way while ensuring that the assignment cost does not shoot up much ok. So, let us stop here. Thank you.