**Approximation Algorithm**

**Prof. Palash Dey**

**Department of Computer Science and Engineering**

**Indian Institute of Technology, Kharagpur**

**Week – 07**

**Lecture 31**

Lecture 31 : A 4 Factor Approximation Algorithm for Uncapacitated Facility Location Problem Contd.

Welcome. So, in the last class we have started doing four factor approximation algorithm for facility location problem. We have first written down the primal and dual LPs and the idea is to solve both of them to get optimal solutions, then we defined a concept of neighboring facilities of a client. Then we have seen that if somehow we can assign each client to its neighboring facility, then total assignment cost is at most opt and then we saw that if I am able to find clients $j_1, j_2, \ldots, j_t$ whose neighbours are disjoint they do not overlap and I open a cheapest facility from each that neighbourhood then total facility opening cost is also at most off. The idea is now to find out $j_1, j_2, \ldots, j_t$ in such a way we can open cheapest facility from each neighbor and then assign all the other clients to these facilities in such a way that the assignment cost does not shoot up by too much.

So, let us see how this idea is implemented in this lecture. uncapacitated facility location problem . So, for that we define double neighborhood of ah of clients. So, definition So, for client $j \in D$ define $N^2(j)$ to be the set of all clients such that k neighbors some facility in $N(j)$ pictorially here you have a client j and this is $N(j)$the neighbouring and you ask which other clients have a neighboring facility in this set.

So, this clients is $N^2(j)$. So, this is a subset of clients. So, let me first briefly discuss the idea of the algorithm, then we will see the pseudo code and then we will analyze that algorithm. Initially none of the facilities are open, it is a iterative algorithm in each iteration I will open a facility and assign some clients to some facilities. So, the first iteration I pick a client with certain property we will see what is the property, but for the time being let us take a pick a client and the property is that $v_{j_1}^*$ is the minimum among all clients who has not been assigned yet. Initially none of the clients are assigned. So, in the first iteration pick the client which has the minimum $v_j^*$ value. Why we are doing it? We will see it is useful in the analysis. It will be used crucially in bounding the approximation factor of our algorithm.

So, I pick a client which has minimum v star values v j star values among all unassigned clients ok. And then I look at $N(j_1)$ and which facility I will open? I will open the cheapest facility a cheapest facility in $N(j_1)$ that I open. and I look at $N^2(j_1)$ ok. So, this is the facility may be $i_k$ and all the clients $N^2(j_1)$ I assign it to $i_k$ and so on ok, this is the first iteration. So, you see in first iteration I have opened a facility and I have assigned some subset of clients to that open facility.

The remaining clients who is which does not belong to $N^2(j_1)$ they remain unassigned In the next iteration if there is any unassigned client left then again I pick a client $j_2$ which has minimum $v_j^*$ value among all unassigned clients. I look at the neighbour set of $j_2$ which is $N(j_2)$. as usual I open the cheapest facility in $j_2$ let us call it $i_2$ here let us call it $i_1$ and again I look at $N^2(j_2)$ all the clients in $N^2(j_2)$ I assign it to Now, is it the case that $N(j_1)$ and $N(j_2)$ are disjoint? Of course, yes you can take it as a homework or I would suggest you pause the video here and prove it yourself that $N(j_1)$ and $N(j_2)$ are disjoint. I hope you have pause the video and proved it yourself. So, here is the easy idea if they overlap if $N(j_1)$ and $N(j_2)$ overlap.

So, what will happen consider the facility which is common to both $N(j_1)$ and $N(j_2)$, but then that if $N(j_1)$ and $N(j_2)$ overlap. then you see this is $j_1$ and this is also $j_2$, then $j_2$ should be in $N^2(j_1)$. If $N(j_1)$ and $N(j_2)$ overlap then $j_2$ will be in $N^2(j_1)$, but $j_2$ is not in $N^2(j_1)$ why because all the clients in $N^2(j_1)$ we have assigned to facility facility $i_1$. So, if $N(j_2)$ is in $N^2(j_1)$ then in the beginning of second iteration $j_2$ would not have been unassigned hence they cannot overlap. If $N(j_1)$ intersection $N(j_2)$ is not empty, then $j_2$ is in $N^2(j_1)$which contradicts the fact that $j_2$ is unassigned after first iteration.

So, they must be disjoint and we continue this process until we assign all the facilities all the clients. That means, I have $j_1$ here is $N(j_1)$ and I open the cheapest facility and assign it to all the clients in. $N^2(j_1)$. So, $j_2$ I open the cheapest facility $i_2$ and assign all the facilities in $N^2(j_2)$ to $i_2$ and so on I continue this process until I assign all clients to some open facility. So, let us see the pseudo code of the algorithm.

Step 1, solve both the primal and dual LPs let $(x^*, y^*)$ be a primal optimal solution and $(v^*, w^*)$ a dual optimal solution. I will maintain a set C of clients which are unassigned initially all clients are unassigned and I will maintain a counter k which will keep track of number of facilities open till now. So, while C not equal to empty set that means, there exist some unassigned clients we continue what k is $k+1$. So, I will choose a unassigned client which has a lowest v star value. So, choose $j_k \in C$ that has smallest $v^*_{j_k}$ over all or $v^*_j$ over all $j \in C$.

So, among all unassigned clients pick a client which has smallest v star value. then choose a facility in the neighborhood of $j_k$ which is $N(j_k)$ to open a cheapest facility $i_k$ in $N(j_k)$ assign all unassigned clients. of $N^2(j_k)$ to $i_k$ and update the list of unassigned clients. I do not need to write $j_k$ explicitly because $j_k$ belongs to $N^2(j_k)$ and that is it. So, these are the set of open facilities and these are the assignments.

So, by design we have observe that total facility opening cost is at most opt ok, because this $N(j_1)$, $N(j_2)$ these are disjoint and we have opened the cheapest facility among these sets ok good. till now we have not used metric property and we also have not used why I am picking a unassigned facility unassigned client which has the lowest $v^*_j$ value this we will see now. So, and that will finish the proof that it is a 4 factor approximation algorithm. our algorithm has an approximation factor of at most 4 proof. We have argued that total facility opening cost is at most opt to show that total cost of our solution is at most 4 opt it is enough to show that total assignment cost is at most 3 opt.

it is enough to show that total assignment cost is at most 3 options. So, for that let us consider any arbitrary iteration k, let k be any iteration. In k-th iteration what is happening? I am opening a I am considering a client $j_k$ which has least $v^*_j$ value among all unassigned clients. I am looking at the set $N(j_k)$ the neighbors of j k opening a cheapest facility here. and for all clients in $N^2(j_k)$ suppose this is some j prime I am assigning this to $i_k$ that the open facility that the facility we have opened.

So, we will prove that if $j' \in D$ is a client which is assigned to $i_k$ then it is assignment cost $c_{i_k j'}$ this is the assignment cost is less than equal to $3v^*_{j'}$. and that will be enough to show that this total assignment cost is at most 3 opt. So, let us show that this. So, if I k belongs to $N(j')$ if $i_k$ is a neighbor of $j'$ then this cost is by complementary slackness we have shown that if $i_k$ is a neighbor of j prime then $c_{i_k j'}$ is less than equal to $v^*_j$ even better. So, if $i_k$ belongs to $N(j')$ then is the lemma we have ah shown in the last lecture, then the assignment cost is less than equal to $v^*_{j'}$ very good otherwise.

So, let us assume that $i_k$ does not belong to $N(j')$. very good. So, what is the situation then? I have picked $j_k$ which has the least $v^*_j$ value among all unassigned clients. Here is the neighbouring set $N(j_k)$, I open a cheapest facility here $i_k$ and here is $j'$ this $N(j')$

overlaps with $N(j_k)$ that is why $j^{'} \in N(j_k)$. So, if I look at the neighborhood of $N(j^{'})$ this is the neighborhood of $j_k$ and this is the neighborhood of $N(j^{'})$.

So, there is at least one common facility let us call it h. So, till now we have not used triangle inequality. So, let us use, but let us first make this picture less clumsy ok yeah. So, I know this cost cost of assigning $j^{'}$ to h this is at most $v_j^{'*}$ this is the lemma we have already proved. What is the cost of assigning $j_k$ to h? This is less than equal to $v_{j_k}^*$ because h belongs to $N(j_k)$, h is a neighbor of $j_k$ and what is the cost of assigning $j_k$ to $i_k$ this is less than equal to again $v_{j_k}^*$.

So, and I want to bound the cost of assigning $i_k$ to $j^{'}$. So, by triangle inequality here we are using the metric property now, what we have is $c_{i_k j^{'}}$ this is less than equal to $c_{i_k j_k} + c_{i_k h} + c_{h j_k}$ this is the sum of this green edges and $c_{h j^{'}}$ again because this is the lemma we have proved before ah by complementary slackness $c_{i_k j_k}$ is less than equal to $v_{j_k}^* + c_{h j_k}$ is less than equal to $v_{j_k}^* + c_{h j^{'}}$ is less than equal to $v_{j^{'}}^*$. Till now we have not used the fact that we have picked $j_k$ because it has the minimum $v_{j_k}^*$ , $v_j^*$ value now we are using it. Now, because $v_j^*$ is minimum among all unassigned clients and $j^{'}$ was also unassigned in the beginning of k-th iteration $v_{j_{k^{'}}}^*$ is less than equal to $v_j^{'*}$. So, what we have is less than equal to this is also$v_{j_{k^{'}}}^*$ this $v_{j_{k^{'}}}$ is also less than equal to $v_{j^{'}}^*$.

$v_{j^{'}}^*$ which is $3 v_{j^{'}}^*$ which is exactly what we need to show that this $c_{i_k j^{'}}$ is less than equal to $v_{j^{'}}^*$. So, what is the total assignment cost is less than equal to for $j^{'}$ the cost is less than equal to $v_{j^{'}}^*$. So, this is for all clients $j \in D$ 3 times $v_{j^{'}}^*$. Now, again summation $v_j^*$ it is a dual opt and hence this is less than equal to LP opt. which is less than equal to 3 opt.

Hence, total assignment cost is at most 3 times opt, total facility opening cost is at most opt, hence total cost is at most 4 times first one is less than equal to opt and second one is less than equal to 3 opt. This concludes the proof that our algorithm has a 4 factor approximation ok. So, let us stop here. Thank you.