The Online Connected Facility Location Problem[2]

Mário César San Felice*
David P. Williamson†
Orlando Lee*

*UNICAMP
Institute of Computing
Laboratory of Optimization and Combinatorics

†Cornell University
School of Operations Research
and Information Engineering

UNICAMP

April 3rd, 2014

[2]grant No. 2012/06728-3, São Paulo Research Foundation (FAPESP)

Combinatorial Optimization Problems

- Maximization or minimization problems,
- Algorithm receives an input,
- Returns a solution with a cost.

Some minimization problems in which we are interested are:

- Facility Location problem,
- Steiner Tree problem,
- Connected Facility Location problem.

These problems are NP-hard with constant factor approximation algorithms known.

Combinatorial Optimization Problems

- Maximization or minimization problems,
- Algorithm receives an input,
- Returns a solution with a cost.

Some minimization problems in which we are interested are:

- Facility Location problem,
- Steiner Tree problem,
- Connected Facility Location problem.

These problems are NP-hard with constant factor approximation algorithms known.

Combinatorial Optimization Problems

- Maximization or minimization problems,
- Algorithm receives an input,
- Returns a solution with a cost.

Some minimization problems in which we are interested are:

- Facility Location problem,
- Steiner Tree problem,
- Connected Facility Location problem.

These problems are NP-hard with constant factor approximation algorithms known.

The Facility Location Problem

- Defined in a metric space with distance and facility costs,
- Algorithm receives a set of clients and connects each client to an opened facility,
- Goal: minimize total cost of opening facilities plus connecting clients.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F')$$

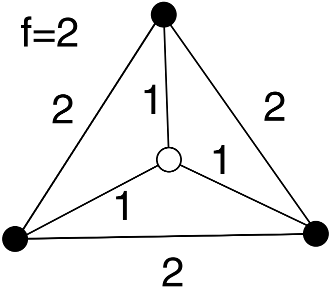
The Facility Location Problem

- Defined in a metric space with distance and facility costs,
- Algorithm receives a set of clients and connects each client to an opened facility,
- Goal: minimize total cost of opening facilities plus connecting clients.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F')$$

The Facility Location Problem (cont.)

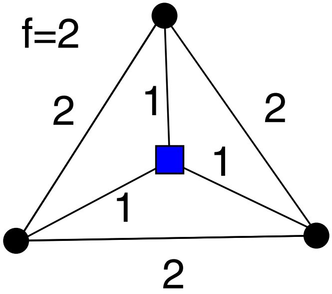
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F')$$



Total cost = 2 + 3 = 5.

The Facility Location Problem (cont.)

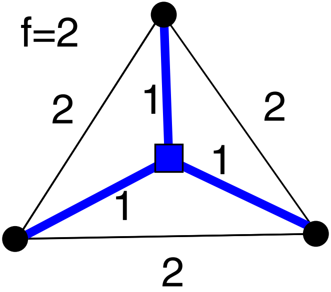
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F')$$



Total cost $= 2 + 3 = 5$.

The Facility Location Problem (cont.)

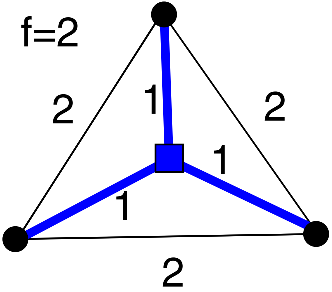
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F')$$



Total cost = 2 + 3 = 5.

The Facility Location Problem (cont.)

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F')$$



Total cost $= 2 + 3 = 5$.

The Steiner Tree Problem

- Defined in a graph with costs on the edges,
- Algorithm receives a set of terminal nodes and builds a tree that connects all terminals,
- Goal: minimize the total cost of the tree.

$$\sum_{e \in T} d(e)$$

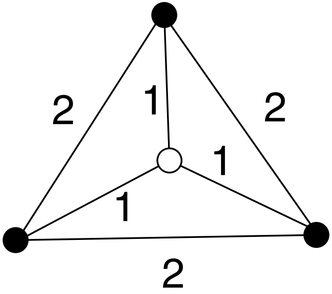
The Steiner Tree Problem

- Defined in a graph with costs on the edges,
- Algorithm receives a set of terminal nodes and builds a tree that connects all terminals,
- Goal: minimize the total cost of the tree.

$$\sum_{e \in T} d(e)$$

The Steiner Tree Problem (cont.)

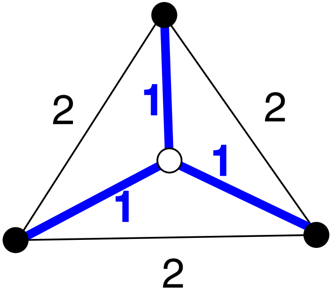
$$\sum_{e \in T} d(e)$$



Total cost = 3.

The Steiner Tree Problem (cont.)

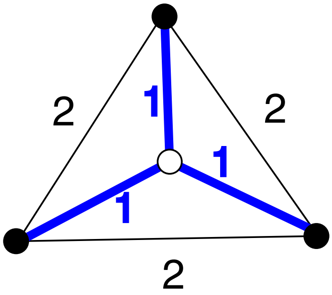
$$\sum_{e \in T} d(e)$$



Total cost = 3.

The Steiner Tree Problem (cont.)

$$\sum_{e \in T} d(e)$$



Total cost = 3.

The Connected Facility Location Problem

- Combination of Facility Location and Steiner Tree.
- Motivation is to build a two level network.
- Algorithm receives a set of clients and connects each client to an opened facility,
- Also, builds an expensive tree that connects all facilities,
- Goal: minimize total cost of opening facilities, connecting clients and building the tree.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$

The Connected Facility Location Problem

- Combination of Facility Location and Steiner Tree.
- Motivation is to build a two level network.
- Algorithm receives a set of clients and connects each client to an opened facility,
- Also, builds an expensive tree that connects all facilities,
- Goal: minimize total cost of opening facilities, connecting clients and building the tree.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$

The Connected Facility Location Problem

- Combination of Facility Location and Steiner Tree.
- Motivation is to build a two level network.
- Algorithm receives a set of clients and connects each client to an opened facility,
- Also, builds an expensive tree that connects all facilities,
- Goal: minimize total cost of opening facilities, connecting clients and building the tree.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$

The Connected Facility Location Problem

- Combination of Facility Location and Steiner Tree.
- Motivation is to build a two level network.
- Algorithm receives a set of clients and connects each client to an opened facility,
- Also, builds an expensive tree that connects all facilities,
- Goal: minimize total cost of opening facilities, connecting clients and building the tree.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$

The Connected Facility Location Problem

- Combination of Facility Location and Steiner Tree.
- Motivation is to build a two level network.
- Algorithm receives a set of clients and connects each client to an opened facility,
- Also, builds an expensive tree that connects all facilities,
- Goal: minimize total cost of opening facilities, connecting clients and building the tree.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$

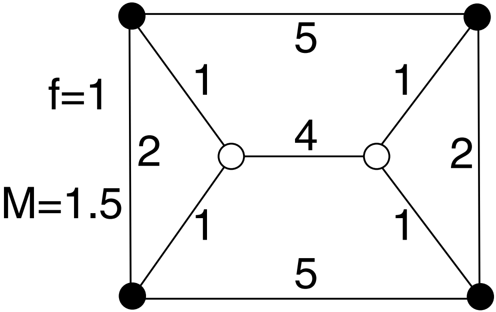
The Connected Facility Location Problem

- Combination of Facility Location and Steiner Tree.
- Motivation is to build a two level network.
- Algorithm receives a set of clients and connects each client to an opened facility,
- Also, builds an expensive tree that connects all facilities,
- Goal: minimize total cost of opening facilities, connecting clients and building the tree.

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$

Connected Facility Location Problem (cont.)

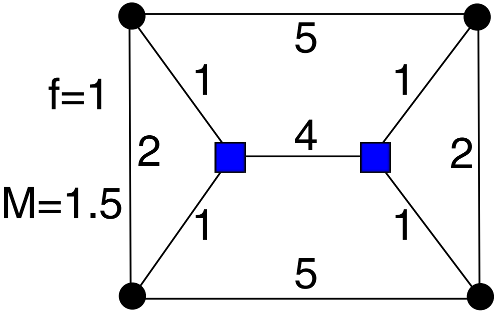
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$



Total cost = 2 + 4 + 6 = 12.

Connected Facility Location Problem (cont.)

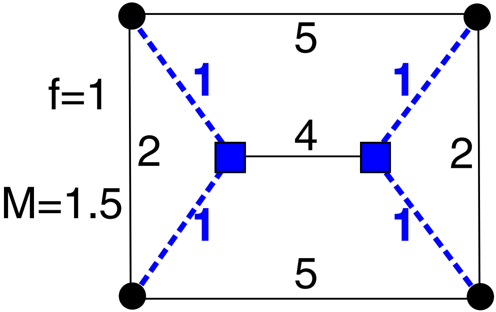
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$



Total cost = 2 + 4 + 6 = 12.

Connected Facility Location Problem (cont.)

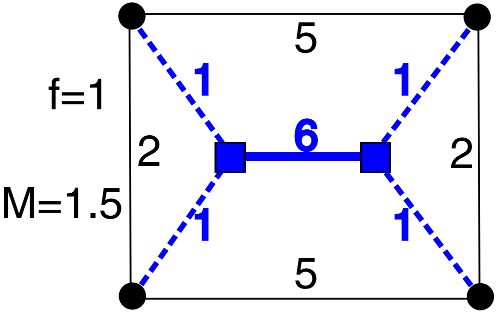
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$



Total cost $= 2 + 4 + 6 = 12$.

Connected Facility Location Problem (cont.)

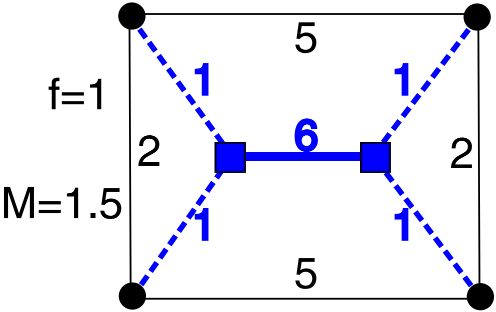
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$



Total cost = 2 + 4 + 6 = 12.

Connected Facility Location Problem (cont.)

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, F') + M \sum_{e \in T} d(e)$$



Total cost $= 2 + 4 + 6 = 12.$

Online Problems

- Parts of the input are revealed one at a time,
- Each part is served before the next one arrives,
- No decision made may be changed in the future.

Some online problems in which we are interested are:

- Online Facility Location problem,
- Online Steiner Tree problem,
- Online Connected Facility Location problem.

Online Problems

- Parts of the input are revealed one at a time,
- Each part is served before the next one arrives,
- No decision made may be changed in the future.

Some online problems in which we are interested are:

- Online Facility Location problem,
- Online Steiner Tree problem,
- Online Connected Facility Location problem.

Competitive Analysis

- Technique used to analyse online algorithms,
- An online algorithm ALG is c-competitive if:

$$\mathrm{ALG}(I) \leq c\mathrm{OPT}(I) + \alpha,$$

  for every input I and some α constant.
- Competitive ratio is similar to approximation ratio.

Competitive Analysis

- Technique used to analyse online algorithms,
- An online algorithm ALG is c-competitive if:

$$\mathrm{ALG}(I) \leq c\,\mathrm{OPT}(I) + \alpha,$$

for every input I and some α constant.
- Competitive ratio is similar to approximation ratio.

Competitive Analysis

- Technique used to analyse online algorithms,
- An online algorithm ALG is c-competitive if:

$$\mathrm{ALG}(I) \leq c\,\mathrm{OPT}(I) + \alpha,$$

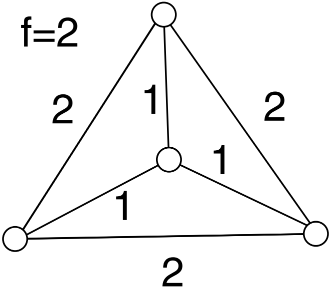
  for every input I and some α constant.
- Competitive ratio is similar to approximation ratio.

The Online Facility Location Problem

- Defined similarly to the Facility Location problem,
- Clients arrive one at a time and each one must be immediately connected to some facility,
- Opened facilities and clients connections may not change.

The Online Facility Location Problem (cont.)

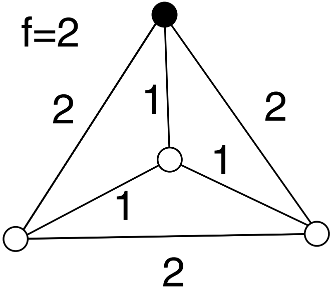
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost $= 2 + 4 = 6$.

The Online Facility Location Problem (cont.)

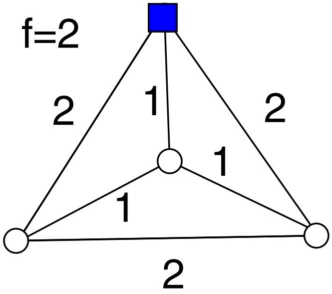
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost = 2 + 4 = 6.

The Online Facility Location Problem (cont.)

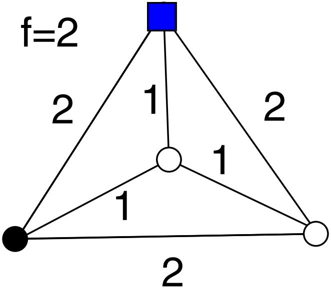
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost = 2 + 4 = 6.

The Online Facility Location Problem (cont.)

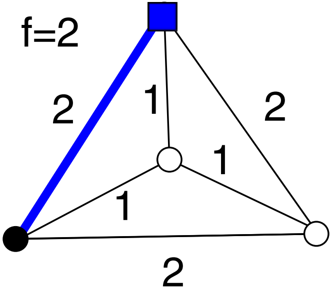
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost $= 2 + 4 = 6$.

The Online Facility Location Problem (cont.)

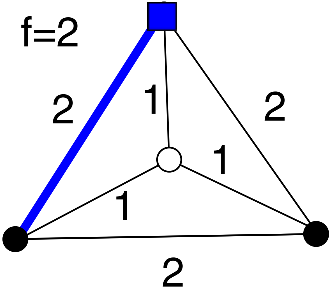
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost = 2 + 4 = 6.

The Online Facility Location Problem (cont.)

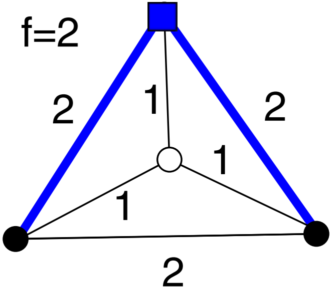
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost $= 2 + 4 = 6$.

The Online Facility Location Problem (cont.)

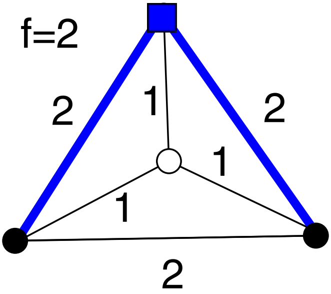
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost = 2 + 4 = 6.

The Online Facility Location Problem (cont.)

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(j))$$



Total cost $= 2 + 4 = 6$.

The Online Facility Location Problem (cont.)

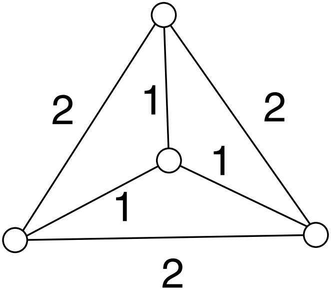
- There are $O(\log n)$-competitive algorithms known,
- In particular, a primal-dual algorithm due to Fotakis that is used in our result,
- Also, it is known a $\Omega(\frac{\log n}{\log \log n})$ lower bound to the competitive ratio.

The Online Steiner Tree Problem

- Defined similarly to the Steiner Tree problem,
- Terminal nodes arrive one at a time,
- At all times terminals must be connected by a tree,
- No edge used may be removed in the future.

The Online Steiner Tree Problem (cont.)

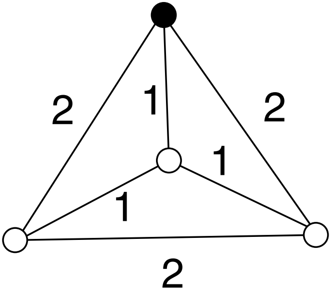
$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

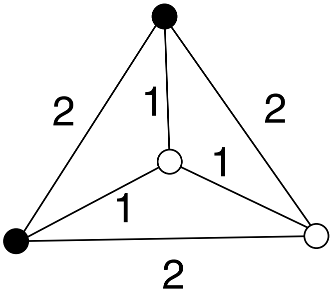
$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

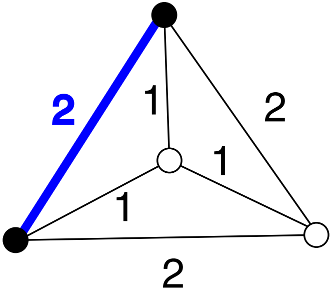
$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

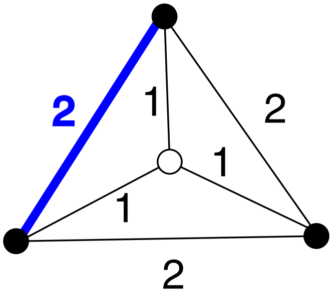
$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

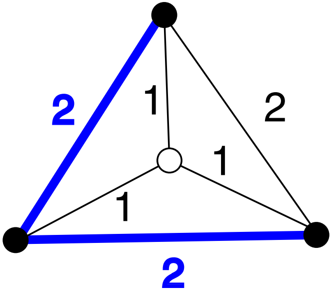
$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

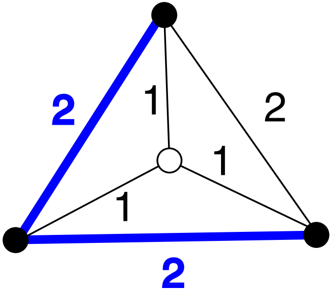
$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

$$\sum_{e \in T} d(e)$$



Total cost = 4.

The Online Steiner Tree Problem (cont.)

- There are $\mathrm{O}(\log n)$-competitive algorithms known,
- In particular, a greedy algorithm due to Imase and Waxman that is used in our result,
- Also, it is known a $\Omega(\log n)$ lower bound to the competitive ratio.

Online Connected Facility Location Problem

- Combination of Online Facility Location and Online Steiner Tree, is similar to the Connected Facility Location,
- Clients arrive one at a time and each one must be immediately connected to some facility,
- Also, there is a special facility always opened called **root**,
- At all times opened facilities must be connected by a tree,
- Opened facilities and clients connections may not change, and no edge used may be removed in the future.

Online Connected Facility Location Problem

- Combination of Online Facility Location and Online Steiner Tree, is similar to the Connected Facility Location,
- Clients arrive one at a time and each one must be immediately connected to some facility,
- Also, there is a special facility always opened called **root**,
- At all times opened facilities must be connected by a tree,
- Opened facilities and clients connections may not change, and no edge used may be removed in the future.

Online Connected Facility Location Problem

- Combination of Online Facility Location and Online Steiner Tree, is similar to the Connected Facility Location,
- Clients arrive one at a time and each one must be immediately connected to some facility,
- Also, there is a special facility always opened called **root**,
- At all times opened facilities must be connected by a tree,
- Opened facilities and clients connections may not change, and no edge used may be removed in the future.

Online Connected Facility Location Problem

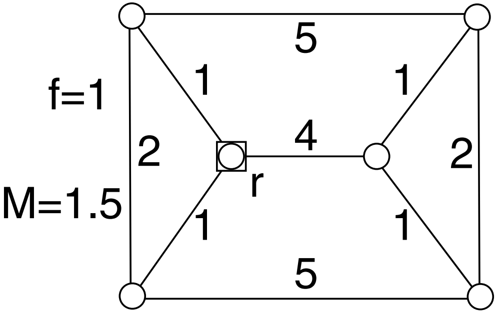
- Combination of Online Facility Location and Online Steiner Tree, is similar to the Connected Facility Location,
- Clients arrive one at a time and each one must be immediately connected to some facility,
- Also, there is a special facility always opened called **root**,
- At all times opened facilities must be connected by a tree,
- Opened facilities and clients connections may not change, and no edge used may be removed in the future.

Online Connected Facility Location Problem

- Combination of Online Facility Location and Online Steiner Tree, is similar to the Connected Facility Location,
- Clients arrive one at a time and each one must be immediately connected to some facility,
- Also, there is a special facility always opened called **root**,
- At all times opened facilities must be connected by a tree,
- Opened facilities and clients connections may not change, and no edge used may be removed in the future.

Online Connected Facility Location (cont.)

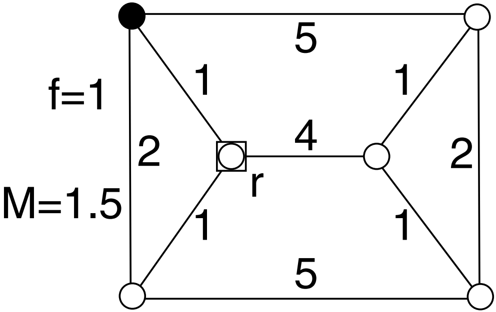
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost = 1 + 4 + 7.5 = 12.5.

Online Connected Facility Location (cont.)

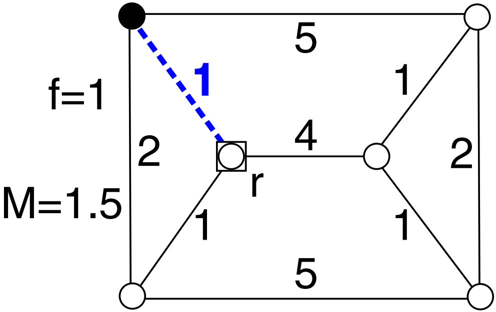
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost $= 1 + 4 + 7.5 = 12.5$.

Online Connected Facility Location (cont.)

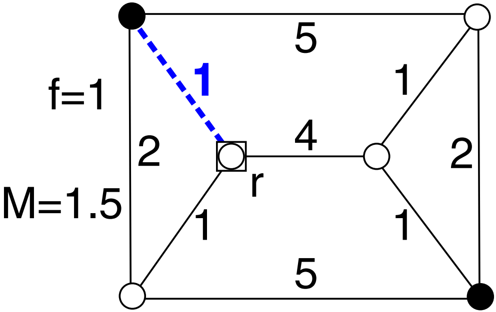
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost $= 1 + 4 + 7.5 = 12.5$.

Online Connected Facility Location (cont.)

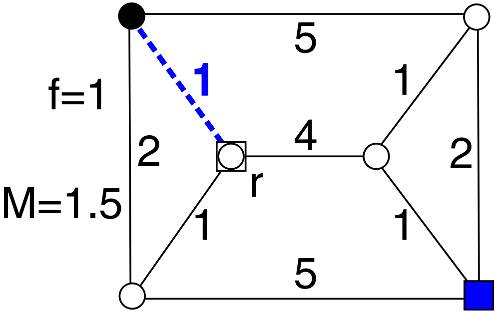
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost $= 1 + 4 + 7.5 = 12.5$.

Online Connected Facility Location (cont.)

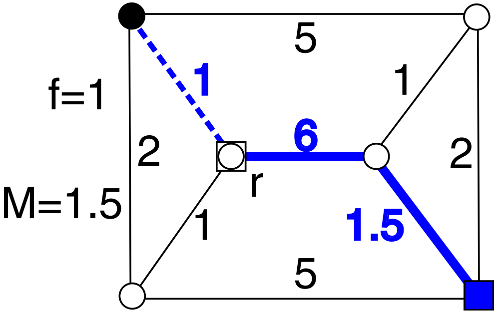
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost = 1 + 4 + 7.5 = 12.5.

Online Connected Facility Location (cont.)

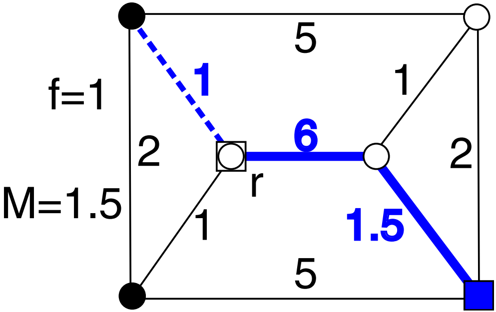
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost = 1 + 4 + 7.5 = 12.5.

Online Connected Facility Location (cont.)

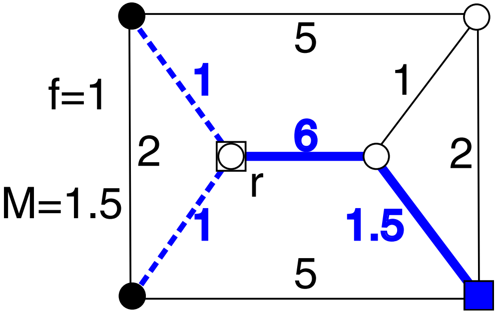
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost = 1 + 4 + 7.5 = 12.5.

Online Connected Facility Location (cont.)

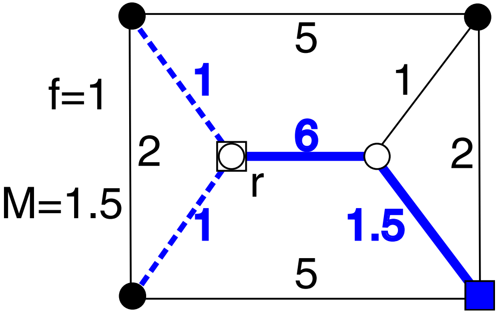
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost $= 1 + 4 + 7.5 = 12.5$.

Online Connected Facility Location (cont.)

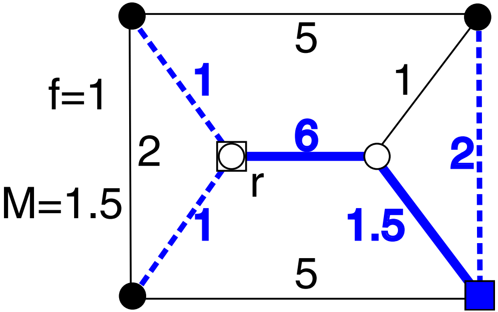
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost = 1 + 4 + 7.5 = 12.5.

Online Connected Facility Location (cont.)

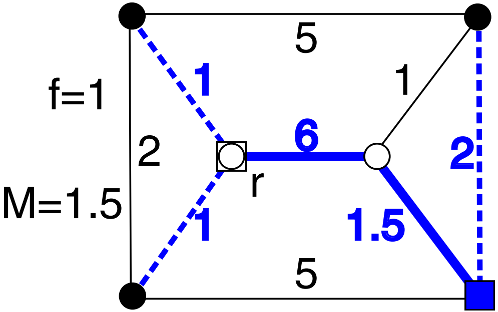
$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost = 1 + 4 + 7.5 = 12.5.

Online Connected Facility Location (cont.)

$$\sum_{i \in F'} f(i) + \sum_{j \in D} d(j, a(f)) + M \sum_{e \in T} d(e)$$



Total cost $= 1 + 4 + 7.5 = 12.5.$

Online CFL Algorithm

- Sample-and-augment algorithm based on algorithm for CFL due to Eisenbrand et al.
- Sample-and-Augment is a randomized technique, due to Gupta et al., to design algorithms for rent-or-buy problems.
- Online Connected Facility Location problem is not a rent-or-buy problem,
- However, has cost scaling characteristics that allow to use this technique on algorithms for it.

Online CFL Algorithm

- Sample-and-augment algorithm based on algorithm for CFL due to Eisenbrand et al.
- Sample-and-Augment is a randomized technique, due to Gupta et al., to design algorithms for rent-or-buy problems.
- Online Connected Facility Location problem is not a rent-or-buy problem,
- However, has cost scaling characteristics that allow to use this technique on algorithms for it.

Online CFL Algorithm

- Sample-and-augment algorithm based on algorithm for CFL due to Eisenbrand et al.
- Sample-and-Augment is a randomized technique, due to Gupta et al., to design algorithms for rent-or-buy problems.
- Online Connected Facility Location problem is not a rent-or-buy problem,
- However, has cost scaling characteristics that allow to use this technique on algorithms for it.

Online CFL Algorithm

- Sample-and-augment algorithm based on algorithm for CFL due to Eisenbrand et al.
- Sample-and-Augment is a randomized technique, due to Gupta et al., to design algorithms for rent-or-buy problems.
- Online Connected Facility Location problem is not a rent-or-buy problem,
- However, has cost scaling characteristics that allow to use this technique on algorithms for it.

Algorithm 1: The Online CFL algorithm.

Data: $G = (V, E)$, d, f, F, root r and M

Initializing auxiliar sets and compFL;

while *a new client j arrives* **do**

  send j to compFL; /* Update virtual OFL solution */

  mark j with probability $p = \frac{1}{M}$;

  if *j is marked and connected to a facility i that is not open* **then**

    $F' \leftarrow F' \cup \{i\}$; /* Open new facility */

    $T \leftarrow T \cup \{(i, j)\} \cup \{path(j, V(T))\}$;

    /* Connect new facility to the tree */

  end

  let i be the closest open facility to j;

  $D \leftarrow D \cup \{j\}$; $a(j) \leftarrow i$; /* Connect the client */

end

return $(F' \setminus \{r\}, T, a)$;

Analysis of the Online CFL Algorithm

Algorithm cost is divided between facilities opening cost (O), clients connection cost (C) and Steiner tree cost (S):

$$ALG_{OCFL}(D) = O + C + S.$$

The cost of the offline optimal solution is also divided in this way:

$$OPT_{CFL}(D) = O^* + C^* + S^*.$$

Analysis of the Online CFL Algorithm

Algorithm cost is divided between facilities opening cost (O), clients connection cost (C) and Steiner tree cost (S):

$$ALG_{OCFL}(D) = O + C + S.$$

The cost of the offline optimal solution is also divided in this way:

$$OPT_{CFL}(D) = O^* + C^* + S^*.$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Opening Cost)

$O \leq c_{\mathrm{OFL}}(O^* + C^*)$.

- Let O_{compFL} be the facility opening cost of compFL,
- Our algorithm opens a subset of the facilities opened by compFL,
- An optimal solution for CFL is a feasible solution for FL,
- So we have that:

$$O \leq O_{\mathrm{compFL}} \leq c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \leq c_{\mathrm{OFL}}(O^* + C^*) .$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Opening Cost)

$O \leq c_{\mathrm{OFL}}(O^* + C^*)$.

- Let O_{compFL} be the facility opening cost of compFL,
- Our algorithm opens a subset of the facilities opened by compFL,
- An optimal solution for CFL is a feasible solution for FL,
- So we have that:

$$O \leq O_{\mathrm{compFL}} \leq c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \leq c_{\mathrm{OFL}}(O^* + C^*) .$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Opening Cost)

$O \le c_{\mathrm{OFL}}(O^* + C^*)$.

- Let O_{compFL} be the facility opening cost of compFL,
- Our algorithm opens a subset of the facilities opened by compFL,
- An optimal solution for CFL is a feasible solution for FL,
- So we have that:

$$O \le O_{\mathrm{compFL}} \le c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \le c_{\mathrm{OFL}}(O^* + C^*) .$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Opening Cost)

$O \leq c_{\mathrm{OFL}}(O^* + C^*)$.

- Let O_{compFL} be the facility opening cost of compFL,
- Our algorithm opens a subset of the facilities opened by compFL,
- An optimal solution for CFL is a feasible solution for FL,
- So we have that:

$$O \leq O_{\mathrm{compFL}} \leq c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \leq c_{\mathrm{OFL}}(O^* + C^*) .$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Opening Cost)

$O \leq c_{\mathrm{OFL}}(O^* + C^*)$.

- Let O_{compFL} be the facility opening cost of compFL,
- Our algorithm opens a subset of the facilities opened by compFL,
- An optimal solution for CFL is a feasible solution for FL,
- So we have that:

$$O \leq O_{\mathrm{compFL}} \leq c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \leq c_{\mathrm{OFL}}(O^* + C^*) \ .$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Opening Cost)

$O \leq c_{\mathrm{OFL}}(O^* + C^*)$.

- Let O_{compFL} be the facility opening cost of compFL,
- Our algorithm opens a subset of the facilities opened by compFL,
- An optimal solution for CFL is a feasible solution for FL,
- So we have that:

$$O \;\leq\; O_{\mathrm{compFL}} \leq c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \leq c_{\mathrm{OFL}}(O^* + C^*)\ .$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Steiner Cost)

$E[S] \leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*)$.

- Our algorithm builds a tree connecting only marked clients (D'),
- And augments it connecting each client j to facility $a(j)$,
- The idea is to bound these costs by OPT_{ST} and OPT_{FL}.

$$\begin{aligned} E[S] &\leq E[M\mathrm{compST}(D')] + E\left[M \sum_{j \in D'} d(j, a(j))\right] \\ &\leq E[M c_{\mathrm{OST}} \mathrm{OPT}_{\mathrm{ST}}(D')] + c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \\ &\leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*) \ . \end{aligned}$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Steiner Cost)

$E[S] \leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*)$.

- Our algorithm builds a tree connecting only marked clients (D'),
- And augments it connecting each client j to facility $a(j)$,
- The idea is to bound these costs by OPT_{ST} and OPT_{FL}.

$$
\begin{aligned}
E[S] &\leq E[M\mathrm{compST}(D')] + E\left[M\sum_{j\in D'} d(j,a(j))\right] \\
&\leq E[Mc_{\mathrm{OST}}\mathrm{OPT}_{\mathrm{ST}}(D')] + c_{\mathrm{OFL}}\mathrm{OPT}_{\mathrm{FL}}(D) \\
&\leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*)\ .
\end{aligned}
$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Steiner Cost)

$E[S] \leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*)$.

- Our algorithm builds a tree connecting only marked clients (D'),
- And augments it connecting each client j to facility $a(j)$,
- The idea is to bound these costs by OPT_{ST} and OPT_{FL}.

$$
\begin{aligned}
E[S] &\leq E[M \mathrm{compST}(D')] + E\left[M \sum_{j \in D'} d(j, a(j))\right] \\
&\leq E[M c_{\mathrm{OST}} \mathrm{OPT}_{\mathrm{ST}}(D')] + c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \\
&\leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*) \ .
\end{aligned}
$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Steiner Cost)

$E[S] \le c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*).$

- Our algorithm builds a tree connecting only marked clients (D'),
- And augments it connecting each client j to facility $a(j)$,
- The idea is to bound these costs by OPT_{ST} and OPT_{FL}.

$$
\begin{aligned}
E[S] &\le E[M\mathrm{compST}(D')] + E\left[M\sum_{j\in D'} d(j,a(j))\right] \\
&\le E[Mc_{\mathrm{OST}}\mathrm{OPT}_{\mathrm{ST}}(D')] + c_{\mathrm{OFL}}\mathrm{OPT}_{\mathrm{FL}}(D) \\
&\le c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*) \ .
\end{aligned}
$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Steiner Cost)

$E[S] \leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*)$.

- Our algorithm builds a tree connecting only marked clients (D'),
- And augments it connecting each client j to facility $a(j)$,
- The idea is to bound these costs by OPT_{ST} and OPT_{FL}.

$$
\begin{aligned}
E[S] &\leq E[M\mathrm{compST}(D')] + E\left[M \sum_{j \in D'} d(j, a(j))\right] \\
&\leq E[M c_{\mathrm{OST}} \mathrm{OPT}_{\mathrm{ST}}(D')] + c_{\mathrm{OFL}} \mathrm{OPT}_{\mathrm{FL}}(D) \\
&\leq c_{\mathrm{OST}}(S^* + C^*) + c_{\mathrm{OFL}}(O^* + C^*) \ .
\end{aligned}
$$

Analysis of the Online CFL Algorithm (cont.)

Lemma (Connection Cost)

$E[C] \leq c_{\mathrm{OFL}}(O^* + C^*) + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*)).$

- This proof uses similar ideas to the previous one,
- Together with a conditional probability argument,
- Notice that the cost is bounded by a product of two $O(\log n)$ factors.

Analysis of the Online CFL Algorithm (cont.)

Lemma (Connection Cost)

$E[C] \leq c_{\mathrm{OFL}}(O^* + C^*) + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*)).$

- This proof uses similar ideas to the previous one,
- Together with a conditional probability argument,
- Notice that the cost is bounded by a product of two O(log n) factors.

Analysis of the Online CFL Algorithm (cont.)

Lemma (Connection Cost)

$E[C] \leq c_{\mathrm{OFL}}(O^* + C^*) + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*))$.

- This proof uses similar ideas to the previous one,
- Together with a conditional probability argument,
- Notice that the cost is bounded by a product of two O(log n) factors.

Analysis of the Online CFL Algorithm (cont.)

Lemma (Connection Cost)

$E[C] \leq c_{\mathrm{OFL}}(O^* + C^*) + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*)).$

- This proof uses similar ideas to the previous one,
- Together with a conditional probability argument,
- Notice that the cost is bounded by a product of two $O(\log n)$ factors.

Analysis of the Online CFL Algorithm (cont.)

Lemma (Connection Cost)

$E[C] \leq c_{\mathrm{OFL}}(O^* + C^*) + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*)).$

- This proof uses similar ideas to the previous one,
- Together with a conditional probability argument,
- Notice that the cost is bounded by a product of two $\mathrm{O}(\log n)$ factors.

Analysis of the Online CFL Algorithm (cont.)

Theorem

$E[\mathrm{ALG}_{\mathrm{OCFL}}(D)] \in \mathrm{O}(\log^2 n \mathrm{OPT}_{\mathrm{CFL}}(D))$.

$$\begin{aligned} E[\mathrm{ALG}_{\mathrm{OCFL}}(D)] &= E[O + S + C] \\ &\leq c_{\mathrm{OFL}}(O^* + C^*) + (c_{\mathrm{OST}}(S^* + C^*) \\ &\quad + c_{\mathrm{OFL}}(O^* + C^*)) + (c_{\mathrm{OFL}}(O^* + C^*) \\ &\quad + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*))) \\ &= \mathrm{O}(\log^2 n)\mathrm{OPT}_{\mathrm{CFL}}(D) \ . \end{aligned}$$

knowing that $c_{\mathrm{OFL}} \leq 4 \log n$ and $c_{\mathrm{OST}} \leq \log n$.

Analysis of the Online CFL Algorithm (cont.)

Theorem

$E[\mathrm{ALG}_{\mathrm{OCFL}}(D)] \in \mathrm{O}(\log^2 n \mathrm{OPT}_{\mathrm{CFL}}(D))$.

$$
\begin{aligned}
E[\mathrm{ALG}_{\mathrm{OCFL}}(D)] &= E[O + S + C] \\
&\leq c_{\mathrm{OFL}}(O^* + C^*) + (c_{\mathrm{OST}}(S^* + C^*) \\
&\quad + c_{\mathrm{OFL}}(O^* + C^*)) + (c_{\mathrm{OFL}}(O^* + C^*) \\
&\quad + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*))) \\
&= \mathrm{O}(\log^2 n)\mathrm{OPT}_{\mathrm{CFL}}(D) \ .
\end{aligned}
$$

knowing that $c_{\mathrm{OFL}} \leq 4 \log n$ and $c_{\mathrm{OST}} \leq \log n$.

Analysis of the Online CFL Algorithm (cont.)

Theorem

$E[\mathrm{ALG}_{\mathrm{OCFL}}(D)] \in \mathrm{O}(\log^2 n \mathrm{OPT}_{\mathrm{CFL}}(D)).$

$$\begin{aligned}
E[\mathrm{ALG}_{\mathrm{OCFL}}(D)] &= E[O + S + C] \\
&\leq c_{\mathrm{OFL}}(O^* + C^*) + (c_{\mathrm{OST}}(S^* + C^*) \\
&\quad + c_{\mathrm{OFL}}(O^* + C^*)) + (c_{\mathrm{OFL}}(O^* + C^*) \\
&\quad + c_{\mathrm{OST}}(S^* + C^* + c_{\mathrm{OFL}}(O^* + C^*))) \\
&= \mathrm{O}(\log^2 n)\mathrm{OPT}_{\mathrm{CFL}}(D) \ .
\end{aligned}$$

knowing that $c_{\mathrm{OFL}} \leq 4 \log n$ and $c_{\mathrm{OST}} \leq \log n$.

Lower Bound for the Online CFL Problem

- Online Steiner Tree problem can be reduced to Online Connected Facility Location problem,
- So, $\Omega(\log n)$ is a lower bound to the competitive ratio of any algorithm to the Online Connected Facility Location problem.

Lower Bound for the Online CFL Problem

- Online Steiner Tree problem can be reduced to Online Connected Facility Location problem,
- So, $\Omega(\log n)$ is a lower bound to the competitive ratio of any algorithm to the Online Connected Facility Location problem.

References

M. Imase and M.B. Waxman.
Dynamic Steiner Tree Problem.
SIAM Journal on Discrete Mathematics, Volume 4, Pages 369–384, 1991.

A. Gupta, A. Kumar, M. Pál and T. Roughgarden.
Approximation via cost sharing: Simpler and better approximation algorithms for network design.
Journal of the ACM, Volume 54, Article 11, ACM, 2007.

D. Fotakis.
A Primal-Dual Algorithm for Online Non-Uniform Facility Location.
Journal of Discrete Algorithms, Volume 5, Pages 141–148, Elsevier, 2007.

References (cont.)

C. Nagarajan and D.P. Williamson.
Offline and Online Facility Leasing.
Integer Programming and Combinatorial Optimization, Volume 5035 in LNCS, Pages 303–315, Springer-Verlag, 2008.

F. Eisenbrand, F. Grandoni, T. Rothvoß and G. Schäfer.
Connected Facility Location Via Random Facility Sampling and Core Detouring
Journal of Computer and System Sciences, Volume 76, Pages 709–726, Elsevier Inc., 2010.

Acknowledgements

Thank you!

Questions?

Acknowledgements

Thank you!

Questions?