# The Online Multicommodity Connected Facility Location Problem

Mário César San Felice\* Cristina G. Fernandes\* Carla Negri Lintzmayer<sup>†</sup>

\* IME-USP Institute of Mathematics and Statistics University of São Paulo <sup>†</sup>IC-UNICAMP Institute of Computing University of Campinas

< □ > < (四 > < (回 > ) < (u = 1) <

October 20th, 2017

Define and present a competitive algorithm for the Online Multicommodity Connected Facility Location problem.

But first ...

Maximization or minimization problems.

Algorithm receives an input.

Returns a solution with a cost.

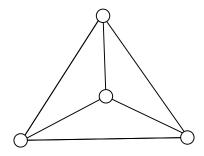
Some minimization problems are:

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

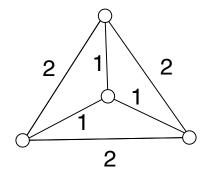
These problems are NP-hard with constant factor approximation algorithms known.



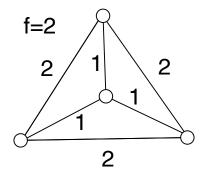
- < A

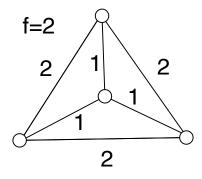


- 一司

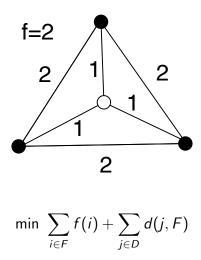


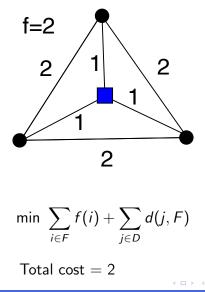
- 一司

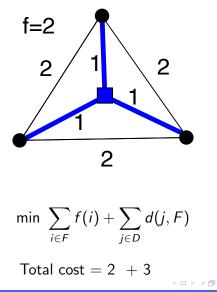


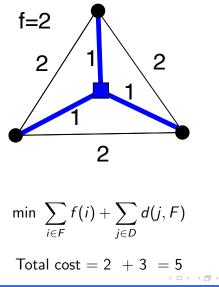


 $\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F)$ 





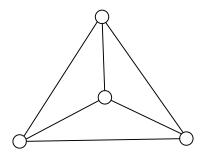






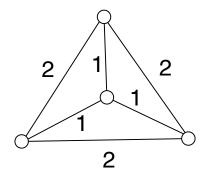
3

Image: A matrix



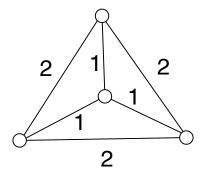
3

Image: A matrix



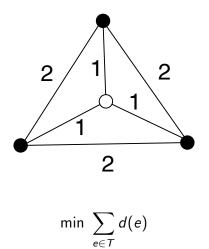
3

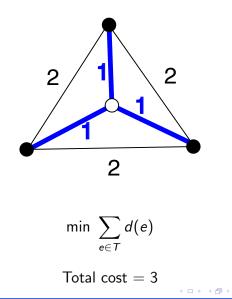
Image: A matrix



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

< 🗗 🕨





OMCFL

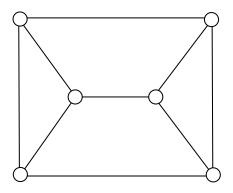
Combination of the Facility Location and the Steiner Tree problems through the rent-or-buy model.

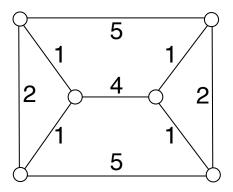
Motivation is to build a two-layer network.

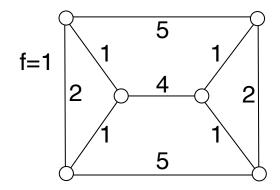
Algorithm receives a set of clients and connects each client to an opened facility.

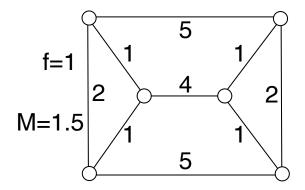
Also, it builds an expensive tree which connects all facilities.

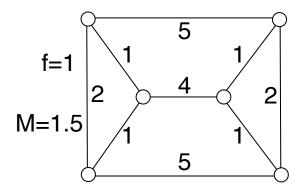




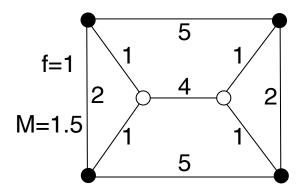




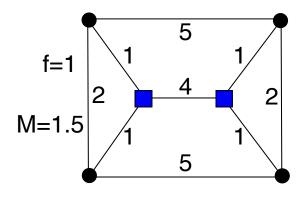




$$\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F) + M \sum_{e \in T} d(e)$$

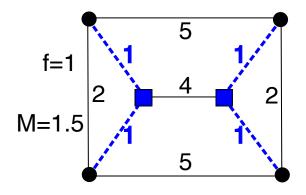


$$\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F) + M \sum_{e \in T} d(e)$$



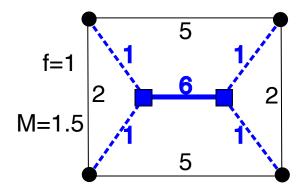
$$\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F) + M \sum_{e \in T} d(e)$$

Total cost = 2



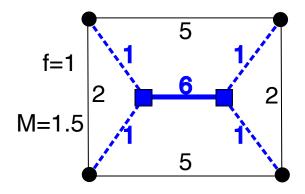
$$\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F) + M \sum_{e \in T} d(e)$$

Total cost = 2 + 4



$$\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F) + M \sum_{e \in T} d(e)$$

Total cost = 2 + 4 + 6 = 4



$$\min \sum_{i \in F} f(i) + \sum_{j \in D} d(j, F) + M \sum_{e \in T} d(e)$$

Total cost = 2 + 4 + 6 = 12

Generalization of the Connected Facility Location problem.

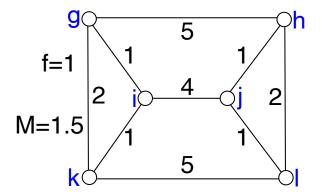
Proposed by Fabrizio Grandoni and Thomas Rothvoß, who presented a constant approximation for it.

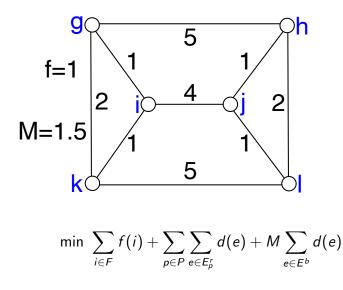
Algorithm receives a set of pairs to connect.

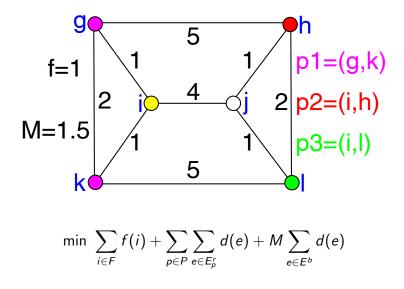
It may rent or buy edges and open facilities to connect each pair.

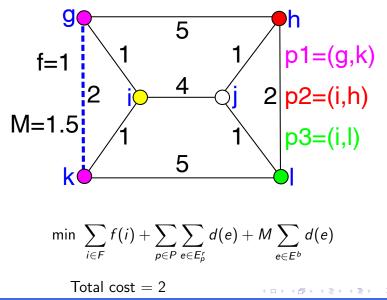
The path connecting a pair may only change between rented and bought edges at an opened facility.



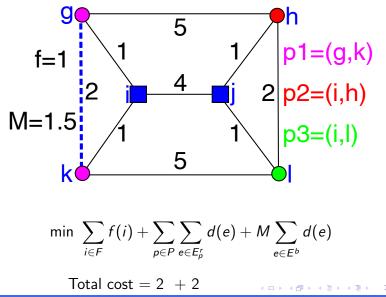


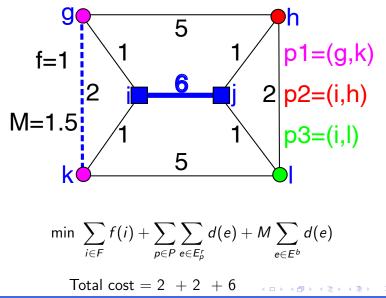


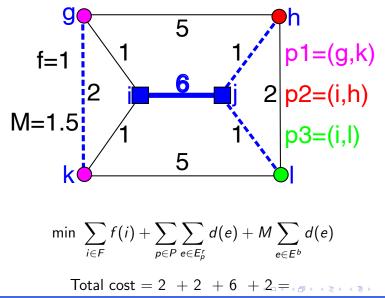


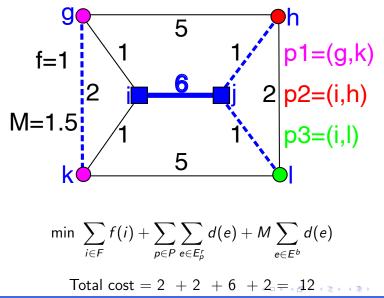


San Felice, Fernandes and Lintzmayer









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.

An online algorithm ALG is *c*-competitive if:

 $ALG(I) \leq c OPT(I)$ ,

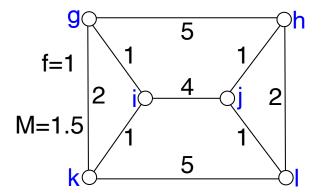
for every input *I*.

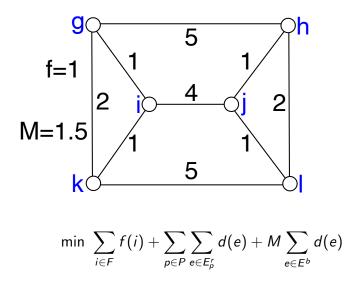
Competitive ratio is similar to approximation ratio.

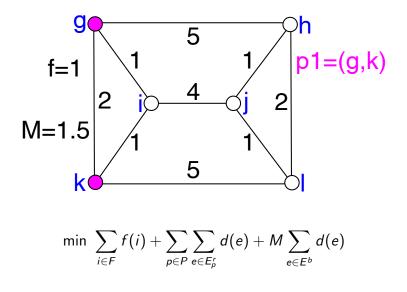
Online version of the Multicommodity Connected Facility Location problem.

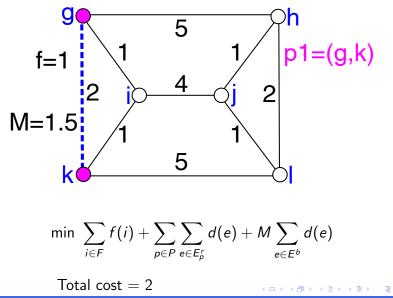
Pairs arrive one at a time and their nodes must be immediately connected to each other.

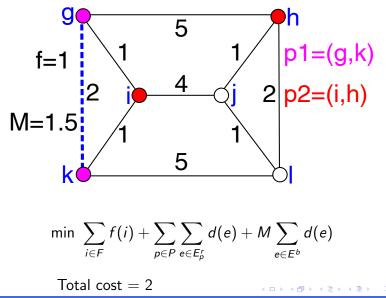
Opened facilities and rented or bought edges may not be removed in the future.

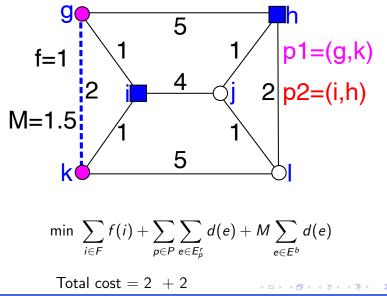


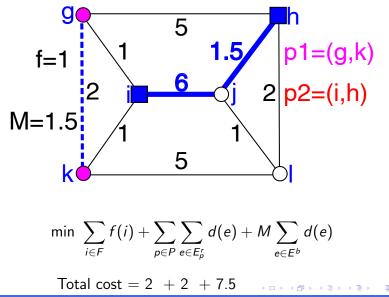


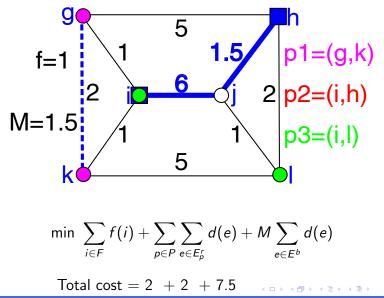


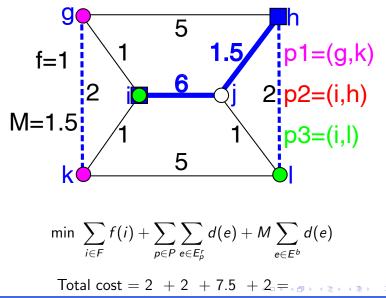


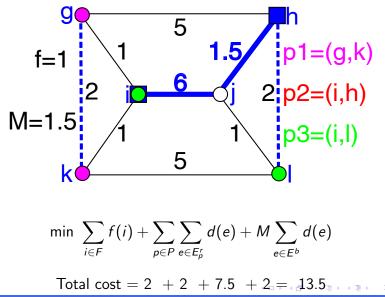












# Online Multicommodity CFL Algorithm

We present a sample-and-augment algorithm inspired on the algorithm for MCFL due to Grandoni and Rothvoß.

Sample-and-Augment is a technique, due to Gupta et al., to design randomized algorithms for rent-or-buy problems.

We highlight that the Online Multicommodity Connected Facility Location problem is not a typical rent-or-buy problem.

Because the constraints on rented edges are distinct from those on bought edges.

However, it still has a cost scaling factor which justify the use of this technique.

Algorithm 1: Draft of Algorithm for the OMCFL problem.

**Input**: (G, d, f, M)

while a new pair p = (s, t) arrives do

decide if and which facilities to open when serving *s* and *t*;  $\triangleright$  algorithm for the Online Prize-Collecting Facility Location mark *p* with probability  $\frac{1}{M}$ ;  $\triangleright$  balance cost scaling factor **if** *p* is marked **then** 

open facilities to which s and t are assigned and update  $F^a$ ; choose edges to connect these facilities and update  $E^b$ ;

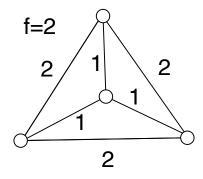
 $\triangleright$  algorithm for the Online Steiner Forest

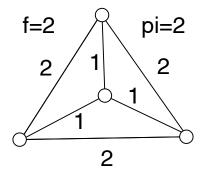
add zero cost edges connecting opened facilities which are in the same bought component;

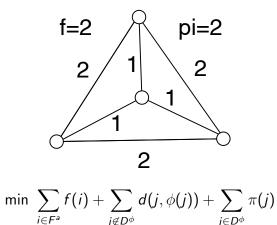
```
consider an (s, t)-shortest path in G;
```

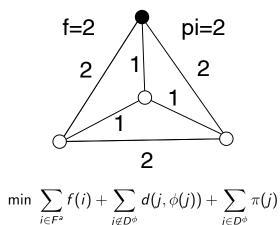
let  $E_p^r$  be the non zero cost edges of this path;

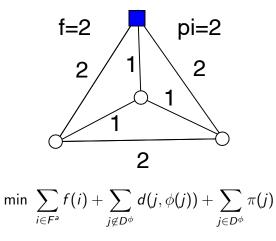
return  $(F^a, E^b, (E_p^r)_{p \in P});$ 



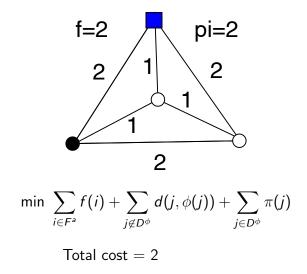


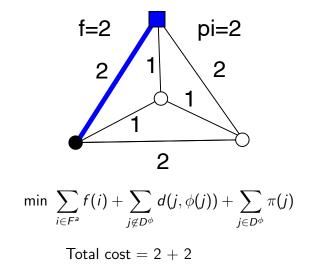


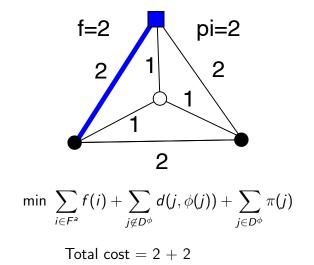


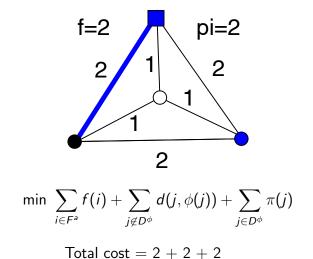


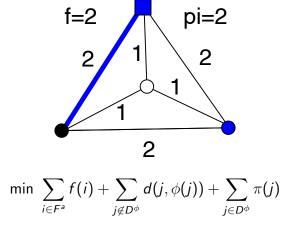
Total cost = 2











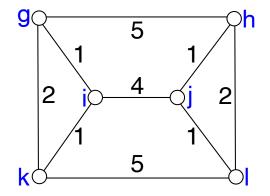
Total cost = 2 + 2 + 2 = 6

Elmachtoub and Levi, and San Felice et al. independently presented  $O(\log n)$ -competitive algorithms for the OPFL.

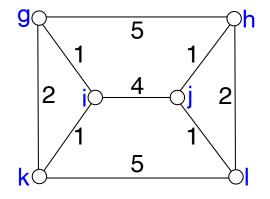
Since the OPFL is a generalization of the Online Facility Location problem, the  $\Omega\left(\frac{\log n}{\log \log n}\right)$  lower bound due to Fotakis applies to it.



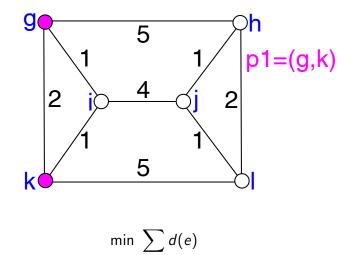
3



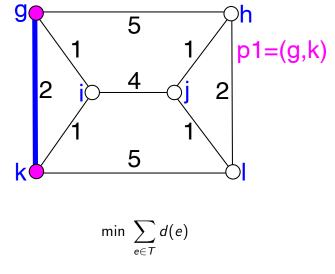
< A



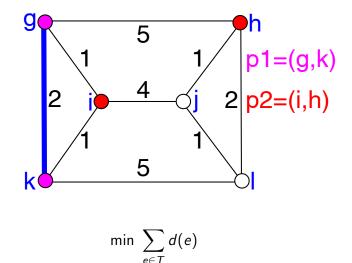
 $\min \sum_{e \in \mathcal{T}} d(e)$ 



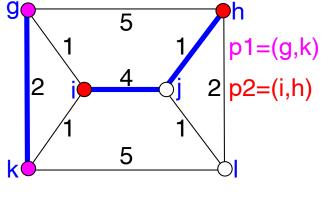
 $e \in T$ 



Total cost 
$$= 2$$

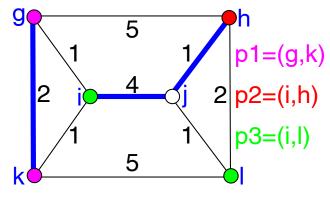


Total cost = 2



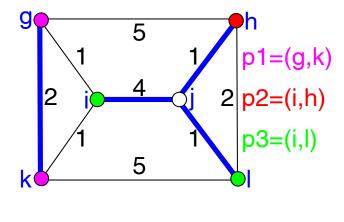
 $\min \sum_{e \in T} d(e)$ 

Total cost = 
$$2 + 5$$

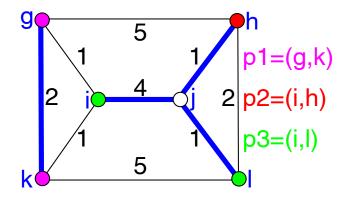


 $\min \sum_{e \in T} d(e)$ 

Total cost = 
$$2 + 5$$



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



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

Total cost = 2 + 5 + 1 
$$\pm$$
 8  $\oplus$   $\pm$   $\pm$   $2$   $\pm$   $2$ 

- Berman and Coulston presented a deterministic  $O(\log n)$ -competitive algorithm for the OSF.
- Also, a  $\Omega(\log n)$  lower bound to the OST due to Imase and Waxman applies to the OSF.

Algorithm 2: Algorithm for the OMCFL problem.

**Input**: (G, d, f, M)while a new pair p = (s, t) arrives do  $\pi_p \leftarrow \operatorname{dist}(G, d', s, t)/2; \quad \triangleright \text{ decide if and which facilities}$ send  $(s, \pi_p)$  and  $(t, \pi_p)$  to ALG<sub>OPFL</sub> obtaining  $\phi(s)$  and  $\phi(t)$ ; if  $\phi(s) \neq \text{null and } \phi(t) \neq \text{null then}$ mark p with probability 1/M;  $\triangleright$  balance cost scaling factor if p is marked then send  $(\phi(s), \phi(t))$  to ALG<sub>OSF</sub> obtaining an edge set  $E_n^b$ ;  $F^a \leftarrow F^a \cup \{\phi(s), \phi(t)\}; E^b \leftarrow E^b \cup E_p^b;$ for x,  $y \in F^a$  in the same component of  $G[E^b]$  do  $d'(x, y) \leftarrow 0; \quad E' \leftarrow E' \cup \{xy\};$ consider an (s, t)-shortest path in G with costs d'; let  $E_{\rho}^{r}$  be the edges of this path except for those in E'; return  $(F^a, E^b, (E_p^r)_{p \in P});$ 

## Analysis of the OMCFL Algorithm

Cost of Algorithm for OMCFL is divided between facilities opening cost(O), edges buying cost(B) and edges renting cost(R):

 $ALG_{OMCFL}(P) = O(P) + B(P) + R(P)$ .

And the edges renting cost (R) is divided according to the pairs in  $P^{\pi}$ ,  $P^{m}$  and  $P^{u}$ :

$$R(P) = R^{\pi}(P) + R^{m}(P) + R^{u}(P)$$

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

$$OPT_{MCFL}(P) = O^{*}(P) + B^{*}(P) + R^{*}(P)$$

#### Lemma

 $OPT_{PFL}(D) \leq OPT_{MCFL}(P).$ 

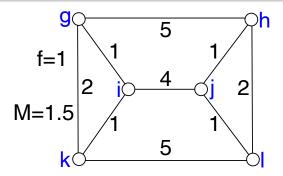
E 5 4

Image: A matrix of the second seco

3

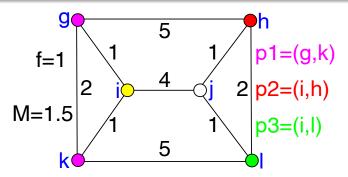
#### Lemma

 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



#### Lemma

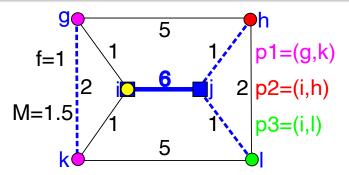
 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



< 67 ▶

#### Lemma

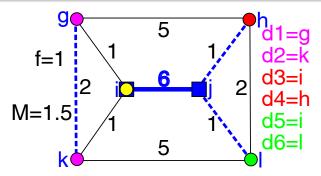
 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



- 一司

#### Lemma

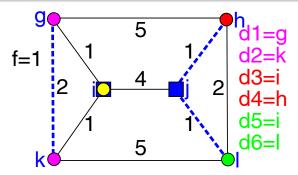
 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



- 一司

#### Lemma

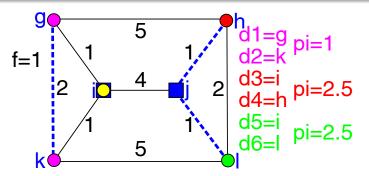
 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



- 一司

#### Lemma

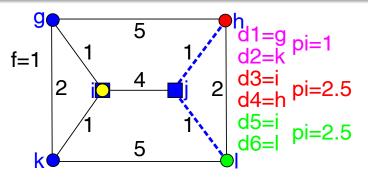
 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



< 67 ▶

#### Lemma

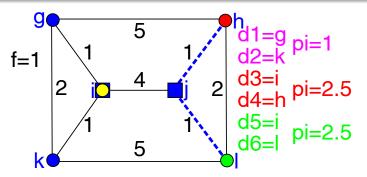
 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



< 67 ▶

#### Lemma

 $\operatorname{OPT}_{\operatorname{PFL}}(D) \leq \operatorname{OPT}_{\operatorname{MCFL}}(P).$ 



 $\operatorname{ALG}_{\operatorname{OPFL}}(D) \leq \operatorname{O}(\log n)\operatorname{OPT}_{\operatorname{PFL}}(P) \leq \operatorname{O}(\log n)\operatorname{OPT}_{\operatorname{MCFL}}(P)$ 

# Some Simple Lemmas

Cost of Algorithm for OPFL is divided between facilities opening cost (O'), clients penalty cost ( $\Pi$ ) and clients connection cost (C'): ALG<sub>OPFL</sub>(D) = O'(D) +  $\Pi(D)$  + C'(D).

Lemma (Facility Opening Cost)

 $O(P) \leq O'(D)$ . ALG<sub>OMCFL</sub> opens a subset of ALG<sub>OPFL</sub> facilities.

Lemma (Close Pairs Renting Cost)

 $R^{\pi}(P) \leq 2\Pi(D)$ . At least one node of each pair paid penalty.

Lemma (Marked Pairs Renting Cost)

 $R^{m}(P) \leq C'(D)$ . For every marked pair, its renting edges correspond to its nodes connections.

# Central Lemma

#### Lemma (Buying Cost)

 $\mathbf{E}[B(P)] = \mathrm{O}(\log^2 n) \operatorname{OPT}_{\mathrm{MCFL}}(P).$ 

 $B(P) \leq M \operatorname{ALG}_{OSF}(Q) = M \operatorname{O}(\log n) \operatorname{OPT}_{SF}(Q)$ .

 $\mathbf{E}[\operatorname{OPT}_{\operatorname{SF}}(Q)] \leq \left(B^*(P) + R^*(P) + C'(D)\right)/M \ .$ 

$$\begin{split} \mathbf{E}[B(P)] &= \mathrm{O}(\log n) \left( B^*(P) + R^*(P) + C'(D) \right) \\ &= \mathrm{O}(\log n) \left( B^*(P) + R^*(P) + \mathrm{ALG}_{\mathrm{OPFL}}(D) \right) \\ &= \mathrm{O}(\log n) \left( \mathrm{OPT}_{\mathrm{MCFL}}(P) + \mathrm{O}(\log n) \mathrm{OPT}_{\mathrm{PFL}}(D) \right) \\ &= \mathrm{O}(\log^2 n) \operatorname{OPT}_{\mathrm{MCFL}}(P) \ . \end{split}$$

3

E + 4 E +

# Final Lemma

#### Lemma (Unmarked Pairs Renting Cost)

 $\mathbf{E}[R^u(P)] = \mathrm{O}(\log^2 n) \operatorname{OPT}_{\mathrm{MCFL}}(P).$ 

 $S_p = E_p^u$  if  $p \notin P^m$  and  $Z_p = E_p^b$  if  $p \in P^m$ .

$$\mathbf{E} \left[ \sum_{e \in E_p^u} d(e) \mid d(S_p), d(Z_p) \right] = \frac{M-1}{M} d(S_p) \leq d(S_p) , \\ \mathbf{E} \left[ \sum_{e \in E_p^b} Md(e) \mid d(S_p), d(Z_p) \right] = \frac{1}{M} M d(Z_p) = d(Z_p) .$$

$$\mathsf{E}\left[\sum_{e\in E_{\rho}^{u}}d(e)
ight] \leq \mathsf{E}\left[\sum_{e\in E_{\rho}^{b}}Md(e)
ight]+d(s,\phi(s))+d(t,\phi(t))$$
 .

 $\mathbf{E}[R^{u}(P)] \leq \mathbf{E}[B(P)] + C'(D) = O(\log^{2} n) \operatorname{OPT}_{\operatorname{MCFL}}(P) \ .$ 

24 / 30

## Main Result

#### Theorem

$$\mathbf{E}[\mathrm{ALG}_{\mathrm{OMCFL}}(P)] = \mathrm{O}(\log^2 n) \operatorname{OPT}_{\mathrm{MCFL}}(P).$$

# $$\begin{split} \mathbf{E}[\mathrm{ALG}_{\mathrm{OMCFL}}(P)] &= \mathbf{E}[O(P)] + \mathbf{E}[B(P)] + \mathbf{E}[R(P)] \\ &= \mathbf{E}[O(P)] + \mathbf{E}[B(P)] \\ &+ \mathbf{E}[R^{\pi}(P) + R^{m}(P) + R^{u}(P)] \\ &\leq O'(D) + \mathrm{O}(\log^{2} n) \operatorname{OPT}_{\mathrm{MCFL}}(P) \\ &+ 2\Pi(D) + C'(D) + \mathrm{O}(\log^{2} n) \operatorname{OPT}_{\mathrm{MCFL}}(P) \\ &= \mathrm{O}(\log^{2} n) \operatorname{OPT}_{\mathrm{MCFL}}(P) \ . \end{split}$$

With a small change in the algorithm we are able to achieve a logarithmic bound on the expected buying cost (B(P)). Thus, we have:

#### Theorem

In the special case of OMCFL in which M = 1, we have

$$ALG2_{OMCFL}(P) = O(\log n) OPT_{MCFL}(P)$$
.

However, we are still working to improve the bound on the expected renting cost of unmarked clients  $(R^u(P))$ .

## References

M. Imase and M.B. Waxman.

Dynamic Steiner Tree Problem. SIAM Journal on Discrete Mathematics, Volume 4, Pages 369–384, 1991.

 P. Berman and C. Coulston.
 On-line Algorithms for Steiner Tree Problems.
 Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing (STOC), Pages 344–353, 1997.

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.

F. Grandoni and T Rothvoß.

Approximation Algorithms for Single and Multi-Commodity Connected Facility Location.

Integer Programming and Combinatorial Optimization (IPCO), Pages 248–260, Springer Berlin Heidelberg, 2011.

#### A.N. Elmachtoub and R. Levi.

From Cost Sharing Mechanisms to Online Selection Problems. Mathematics of Operations Research, Volume 40, Issue 3, Pages 542–557, 2015.

 M.C. San Felice, S.S. Cheung, O. Lee and D.P. Williamson. *The Online Prize-Collecting Facility Location Problem.* Electronic Notes in Discrete Mathematics, Volume 50, Pages 151–156, 2015. That's all!

Questions?



- 一司

3