



The Online Prize-Collecting Facility Location Problem

$\mathsf{DTC}/\mathsf{LOCo}$

Mário César San Felice*, Sin-Shuen Cheung[†], Orlando Lee* and David P. Williamson[†]

*IC-UNICAMP [†]ORIE-Cornell University

August 14, 2015

Problems in which an objective function needs to be minimized or maximized.

Problems in which an objective function needs to be minimized or maximized.

Minimization problems in which we are interested:

- Facility Location problem,
- Prize-Collecting Facility Location problem.

Problems in which an objective function needs to be minimized or maximized.

Minimization problems in which we are interested:

- Facility Location problem,
- Prize-Collecting Facility Location problem.

These problems are NP-hard and constant factor approximation algorithms are known for them.









Total cost
$$= 2$$



Total cost =
$$2 + 3$$



Total cost
$$= 2 + 3 = 5$$
.

San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell)

Parts of the input are revealed one at a time.

Parts of the input are revealed one at a time.

Each part must be served before the next one arrives.

Parts of the input are revealed one at a time.

Each part must be served before the next one arrives.

No decision can be changed in the future.

Competitive Analysis

Competitive Analysis

Worst case technique used to analyze online algorithms.

Competitive Analysis

Worst case technique used to analyze online algorithms.

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

$$\operatorname{ALG}(I) \leq c \cdot \operatorname{OPT}(I) + \kappa$$
,

for every input I and some constant κ .

Online Problems

Minimization problems in which we are interested:

Online Problems

Minimization problems in which we are interested:

• Online Facility Location (OFL),

Online Problems

Minimization problems in which we are interested:

- Online Facility Location (OFL),
- Online Prize-Collecting Facility Location (OPFL).



















The OFL has competitive ratio
$$\Theta\left(\frac{\log n}{\log\log n}\right)$$
 [Fotakis 2008].

The OFL has competitive ratio $\Theta\left(\frac{\log n}{\log \log n}\right)$ [Fotakis 2008].

There are randomized and deterministic $O(\log n)$ -competitive algorithms known for it [Meyerson 2001, Fotakis 2007].

The OFL has competitive ratio
$$\Theta\left(\frac{\log n}{\log \log n}\right)$$
 [Fotakis 2008].

There are randomized and deterministic $O(\log n)$ -competitive algorithms known for it [Meyerson 2001, Fotakis 2007].

[Nagarajan and Williamson 2013] give a dual-fitting analysis for the algorithm by [Fotakis 2007].

Online Facility Location LP Formulation

Online Facility Location LP Formulation

Linear programming relaxation

$$\begin{array}{ll} \min & \sum_{i \in F} f(i) y_i + \sum_{j \in D} \sum_{i \in F} d(j,i) x_{ji} \\ \text{s.t.} & x_{ji} \leq y_i & \text{for } j \in D \text{ and } i \in F, \\ & \sum_{i \in F} x_{ji} \geq 1 & \text{for } j \in D, \\ & y_i \geq 0, x_{ji} \geq 0 & \text{for } j \in D \text{ and } i \in F, \end{array}$$
Online Facility Location LP Formulation

Linear programming relaxation

$$\begin{array}{ll} \min & \sum_{i \in F} f(i) y_i + \sum_{j \in D} \sum_{i \in F} d(j,i) x_{ji} \\ \text{s.t.} & x_{ji} \leq y_i & \text{for } j \in D \text{ and } i \in F, \\ & \sum_{i \in F} x_{ji} \geq 1 & \text{for } j \in D, \\ & y_i \geq 0, x_{ji} \geq 0 & \text{for } j \in D \text{ and } i \in F, \end{array}$$

and its dual

$$\begin{array}{ll} \max & \sum_{j \in D} \alpha_j \\ \text{s.t.} & \sum_{j \in D} (\alpha_j - d(j, i))^+ \leq f(i) & \text{for } i \in F, \\ & \alpha_j \geq 0 & \text{for } j \in D. \end{array}$$

Online Facility Location Algorithm

Algorithm 1: OFL Algorithm.

Input:
$$(G, d, f, F)$$

 $F^{a} \leftarrow \emptyset; D \leftarrow \emptyset;$
while a new client j' arrives do
increase $\alpha_{j'}$ until one of the following happens:
(a) $\alpha_{j'} = d(j', i)$ for some $i \in F^{a};$ /* connect only */
(b) $f(i) = (\alpha_{j'} - d(j', i)) + \sum_{j \in D} (d(j, F^{a}) - d(j, i))^{+}$ for some
 $i \in F \setminus F^{a};$ /* open and connect */
 $F^{a} \leftarrow F^{a} \cup \{i\}; D \leftarrow D \cup \{j'\}; a(j') \leftarrow i;$
end
return $(F^{a}, a);$



















OPFL Results

OPFL Results

Our contribution: we proposed the problem and showed a primal-dual $(6 \log n)$ -competitive algorithm for it, by extending the algorithm from [Fotakis 2007, Nagarajan and Williamson 2013].

OPFL Results

Our contribution: we proposed the problem and showed a primal-dual $(6 \log n)$ -competitive algorithm for it, by extending the algorithm from [Fotakis 2007, Nagarajan and Williamson 2013].

Since the OPFL is a generalization of the OFL, the lower bound of $\Omega\left(\frac{\log n}{\log \log n}\right)$ applies to it.

OPFL LP Formulation

OPFL LP Formulation

Linear programming relaxation

$$\begin{array}{ll} \min & \sum_{i \in F} f(i) y_i + \sum_{j \in D} \sum_{i \in F} d(j, i) x_{ji} + \sum_{j \in D} p(j) z_j \\ \text{s.t.} & x_{ji} \leq y_i & \text{for } j \in D \text{ and } i \in F, \\ & \sum_{i \in F} x_{ji} + z_j \geq 1 & \text{for } j \in D, \\ & y_i \geq 0, x_{ji} \geq 0, z_j \geq 0 & \text{for } j \in D \text{ and } i \in F, \end{array}$$

OPFL LP Formulation

Linear programming relaxation

$$\begin{array}{ll} \min & \sum_{i \in F} f(i) y_i + \sum_{j \in D} \sum_{i \in F} d(j, i) x_{ji} + \sum_{j \in D} p(j) z_j \\ \text{s.t.} & x_{ji} \leq y_i & \text{for } j \in D \text{ and } i \in F, \\ & \sum_{i \in F} x_{ji} + z_j \geq 1 & \text{for } j \in D, \\ & y_i \geq 0, x_{ji} \geq 0, z_j \geq 0 & \text{for } j \in D \text{ and } i \in F, \end{array}$$

and its dual

$$\begin{array}{ll} \max & \sum_{j \in D} \alpha_j \\ \text{s.t.} & \sum_{j \in D} (\alpha_j - d(j, i))^+ \leq f(i) & \text{for } i \in F, \\ & \alpha_j \leq p(j) & \text{for } j \in D, \\ & \alpha_j \geq 0 & \text{for } j \in D. \end{array}$$

San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell)

OPFL Algorithm

Algorithm 2: OPFL Algorithm.

Input:
$$(G, d, f, p, F)$$

 $D \leftarrow \emptyset; F^a \leftarrow \emptyset;$
while a new client j' arrives do
increase $\alpha_{j'}$ until one of the following happens:
(a) $\alpha_{j'} = d(j', i)$ for some $i \in F^a$; /* connect only */
(b) $f(i) = (\alpha_{j'} - d(j', i)) + \sum_{j \in D} (\min\{d(j, F^a), p(j)\} - d(j, i))^+$ for some $i \in F \setminus F^a$; /* open and connect */
(c) $\alpha_{j'} = p(j')$; /* pay the penalty */
(in this case *i* is choose to be null, i.e., $\{i\} = \emptyset$)
 $F^a \leftarrow F^a \cup \{i\}; D \leftarrow D \cup \{j'\}; a(j') \leftarrow i;$
end
return $(F^a, a);$

Analysis

Towards the $O(\log n)$ competitive ratio

• Resulting assignment is feasible.

San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell)

Analysis

Towards the $O(\log n)$ competitive ratio

- Resulting assignment is feasible.
- **2** Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.

Towards the $O(\log n)$ competitive ratio

- Resulting assignment is feasible.
- **2** Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{i}$.

3
$$\left\{\frac{\alpha_j}{3H_n}\right\}_j$$
 is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT}$.

Towards the $O(\log n)$ competitive ratio

- Resulting assignment is feasible.
- **2** Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{i}$.

3
$$\left\{\frac{\alpha_j}{3H_n}\right\}_j$$
 is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT}$.

Analysis

2 Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.

- **2** Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.
 - total cost for opening facilities $\leq \sum_{j} \alpha_{j}$.

- **2** Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.
 - total cost for opening facilities $\leq \sum_{i} \alpha_{i}$.
 - total cost for connections and penalties $\leq \sum_{i} \alpha_{j}$.

- 2 Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.
 - total cost for opening facilities $\leq \sum_{i} \alpha_{j}$.
 - total cost for connections and penalties $\leq \sum_{i} \alpha_{j}$.

Towards $O(\log n)$ competitive ratio

- $\textcircled{0} \quad \text{Resulting assignment is feasible. } \checkmark$
- **2** Total cost of the assignment is bounded by $2 \cdot \sum_{j} \alpha_{j}$.

$$\left\{ \frac{\alpha_j}{3H_n} \right\}_j \text{ is feasible to the dual problem, so } \sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT.}$$

$$\left\{ \frac{\alpha_j}{3H_n} \right\}_j \text{ is feasible to the dual problem, so } \sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT.}$$

•
$$\left\{\frac{\alpha_j}{3H_n}\right\}_j$$
 is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT}$.
 $D^c := \{\text{connected clients}\}$ $D^p := \{\text{penalized clients}\}$

•
$$\left\{\frac{\alpha_j}{3H_n}\right\}_j$$
 is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT}$.
 $D^c := \{\text{connected clients}\}$ $D^p := \{\text{penalized clients}\}$
(i) $\sum_{j \in D^c} \left(\frac{\alpha_j}{3H_n} - d(j, i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$.

•
$$\left\{\frac{\alpha_j}{3H_n}\right\}_j$$
 is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT}$.
 $D^c := \{\text{connected clients}\}$ $D^p := \{\text{penalized clients}\}$
(i) $\sum_{j \in D^c} \left(\frac{\alpha_j}{3H_n} - d(j,i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$.
(ii) $\sum_{j \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j,i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$.

San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell)

Analysis: (i) $\sum_{i \in D^c} \left(\frac{\alpha_i}{3H_r} - d(j, i) \right)^+ \leq \frac{f_i}{2}$ for each $i \in F$ Techniques in [NW, 2013] For each $i \in F$. $f(i) \ge (\alpha_{[k]} - d(j_{[k]}, i)) + \sum (\min\{d(j, F^a_{[k]}), p(j)\} - d(j, i))^+$ $j \in D_{[k-1]}^c$ $(j \text{ connected} \Rightarrow d(j, F^a_{[k]}) \text{ is smaller})$ $= (\alpha_{[k]} - d(j_{[k]}, i)) + \sum (d(j, F^{a}_{[k]}) - d(j, i))^{+}$ $j \in D_{[k-1]}^c$ (triangle inequality) $\geq (\alpha_{[k]} - d(j_{[k]}, i)) + \sum (\alpha_{[k]} - d(j_{[k]}, i) - 2d(j, i))^+$ $j \in D_{tk-1}^c$ $\Rightarrow f(i) \ge (1 + \underbrace{(k-1)}) \cdot (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \sum_{i \in \mathcal{N}} d(j, i)$ $j \in D_{[k-1]}^c$ $=|D_{[k-1]}^{c}|$

Sa

Analysis: (i)
$$\sum_{j \in D^{c}} \left(\frac{\alpha_{j}}{3H_{n}} - d(j,i)\right)^{+} \leq \frac{f_{i}}{2}$$
 for each $i \in F$
Techniques in [NW, 2013]
For each $i \in F$,
 $f(i) \geq k \cdot (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \sum_{j \in D_{[k-1]}^{c}} d(j, i)$
 $\Rightarrow \frac{f(i)}{k} \geq (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \frac{\sum_{j \in D_{[k-1]}^{c}} d(j, i)}{k}$
 $\Rightarrow H_{|D^{c}|} \cdot f(i) \geq \sum_{k=1}^{|D^{c}|} (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \sum_{k=1}^{|D^{c}|} \frac{1}{k} \cdot \sum_{j \in D_{[k-1]}^{c}} d(j, i)$
 $\Rightarrow H_{|D^{c}|} \cdot f(i) \geq \sum_{k=1}^{|D^{c}|} (\alpha_{[k]} - d(j_{[k]}, i)) - (2H_{|D^{c}|} - 1) \cdot \sum_{j \in D^{c}} d(j, i)$
 $\Rightarrow H_{|D^{c}|} \cdot f(i) \geq \sum_{k=1}^{|D^{c}|} (\alpha_{[k]} - 2H_{|D^{c}|} d(j_{[k]}, i))$

26

Analysis: (i) $\sum_{i \in D^c} \left(\frac{\alpha_j}{3H_c} - d(j, i) \right)^+ \leq \frac{f_i}{2}$ for each $i \in F\checkmark$ Techniques in [NW, 2013] For each $i \in F$. $f(i) \geq \mathbf{k} \cdot (\alpha_{[k]} - \mathbf{d}(j_{[k]}, i)) - 2 \sum \mathbf{d}(j, i)$ $j \in D_{[k-1]}^c$ $\Rightarrow \frac{f(i)}{t} \ge (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \frac{\sum_{j \in D_{[k-1]}^c} d(j, i)}{t}$ $\Rightarrow \quad H_{|D^{c}|} \cdot f(i) \geq \sum_{k=1}^{|D^{c}|} (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \sum_{k=1}^{|D^{c}|} \frac{1}{k} \cdot \sum_{j \in D_{[k-1]}^{c}} d(j, i)$ $\Rightarrow \quad \boldsymbol{H}_{|\boldsymbol{D}^{c}|} \cdot f(i) \geq \sum_{i=1}^{j-1} (\alpha_{[k]} - d(j_{[k]}, i)) - (2\boldsymbol{H}_{|\boldsymbol{D}^{c}|} - 1) \cdot \sum_{i=1}^{j-1} d(j, i)$ i∈D° $\Rightarrow \quad H_{|D^c|} \cdot f(i) \geq \sum (\alpha_{[k]} - 2H_{|D^c|}d(j_{[k]}, i))$ San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell) August 14, 2015 19 / 26

Analysis: (ii) $\sum_{i \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j,i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$ Same trick for D^{p} ? For each $i \in F$, $f(i) \ge (\alpha_{[k]} - d(j_{[k]}, i)) + \left(\sum_{j \in \mathcal{I}} \min\{d(j, F^*_{[k]}), \underline{p(j)}\} - d(j, i)\right)^+$ $j \in D_{[k-1]}^p$ (j not connected $\neq d(j, F_{[k]}^a)$ is smaller) $\overline{D}_{[k-1]}^p := \{j \in D_{[k-1]}^p : \alpha_j \ge d(j, F_{[k]}^a)\}$ $\geq (\alpha_{[k]} - d(j_{[k]}, i)) + \sum (d(j, F^{a}_{[k]}) - d(j, i))^{+}$ $j \in \overline{D}_{lk-1}^p$ (triangle inequality) $\geq (\alpha_{[k]} - d(j_{[k]}, i) + \sum (\alpha_{[k]} - d(j_{[k]}, i) - 2d(j, i))$ $j \in \overline{D}_{[k-1]}^p$ $\Rightarrow f(i) \ge (1 + \left\lfloor \overline{D}_{[k-1]}^p \right\rfloor) \cdot (\alpha_{[k]} - d(j_{[k]}, i)) - 2 \sum_{i \in \mathbb{N}^n} d(j, i)$ $j \in D_{[k-1]}^c$ could <(k-1)

San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell)

Analysis: (ii) $\sum_{j \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j, i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$

How to fix?

San Felice, Cheung, Lee and Williamson (UNICAMP and Cornell)
Analysis: (ii) $\sum_{j \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j, i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$

How to fix? Goal: harmonic series as coefficients, e.g. $f(i) \ge \mathbf{k} \cdot (\alpha_k - \mathbf{d}(j_k, i))$

for some ordering $\{j_k\}$ over D^p .

(*)

Analysis: (ii) $\sum_{j \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j, i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$

How to fix?

Goal: harmonic series as coefficients, e.g.

 $f(i) \ge \mathbf{k} \cdot (\alpha_k - \mathbf{d}(j_k, i)) \tag{(*)}$

for some ordering $\{j_k\}$ over D^p . Observe: For larger $\alpha_k - d(j_k, i)$ value, smaller coefficient k should be assigned.

Analysis: (ii) $\sum_{j \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j, i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F$

How to fix?

Goal: harmonic series as coefficients, e.g.

 $f(i) \ge \mathbf{k} \cdot (\alpha_k - \mathbf{d}(j_k, i)) \tag{(*)}$

for some ordering $\{j_k\}$ over D^p . Observe: For larger $\alpha_k - d(j_k, i)$ value, smaller coefficient k should be assigned.

New argument: For each $i \in F$, order $\alpha_k \in D^p$ in nonincreasing order of $\alpha_k - d(j_k, i)$ and show that inequality (*) holds.

Analysis: (ii) $\sum_{j \in D^p} \left(\frac{\alpha_j}{3H_n} - d(j,i)\right)^+ \leq \frac{f_i}{2}$ for each $i \in F\checkmark$

How to fix?

Goal: harmonic series as coefficients, e.g.

 $f(i) \ge \mathbf{k} \cdot (\alpha_k - \mathbf{d}(j_k, i)) \tag{(*)}$

for some ordering $\{j_k\}$ over D^p . Observe: For larger $\alpha_k - d(j_k, i)$ value, smaller coefficient k should be assigned.

New argument: For each $i \in F$, order $\alpha_k \in D^p$ in nonincreasing order of $\alpha_k - d(j_k, i)$ and show that inequality (*) holds.

Analysis

Towards $O(\log n)$ competitive ratio

- Resulting assignment is feasible.
- 2 Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.
- $\left\{\frac{\alpha_j}{3H_n}\right\}_j$ is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT}$. \checkmark

Analysis

Towards $O(\log n)$ competitive ratio

- Resulting assignment is feasible.
- 2 Total cost of the assignment is bounded by $2 \cdot \sum_{i} \alpha_{j}$.

$$\left\{\frac{\alpha_j}{3H_n}\right\}_j$$
 is feasible to the dual problem, so $\sum_j \frac{\alpha_j}{3H_n} \leq \text{OPT.} \checkmark$

Conclusion

3

Our algorithm has $(6 \log n)$ -competitive ratio.

Future Research:

Future Research:

Other variants of Facility Location, like:

- Online Robust Facility Location,
- Online Multicommodity Facility Location,
- Online Prize-Collecting Facility Leasing.

Acknowledgements

Thank you!

Questions?

References



A. Meyerson.

Online Facility Location.

Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science, Page 426, 2001.

D. Fotakis.

On the Competitive Ratio for Online Facility Location. Algorithmica, Volume 50, Pages 1–57, 2008.

References (cont.)

D. Fotakis.

A Primal-Dual Algorithm for Online Non-Uniform Facility Location.

Journal of Discrete Algorithms, Volume 5, Pages 141–148, Elsevier, 2007.

 C. Nagarajan and D.P. Williamson. Offline and Online Facility Leasing. Discrete Optimization, Volume 10, Number 4, Pages 361–370, 2013.