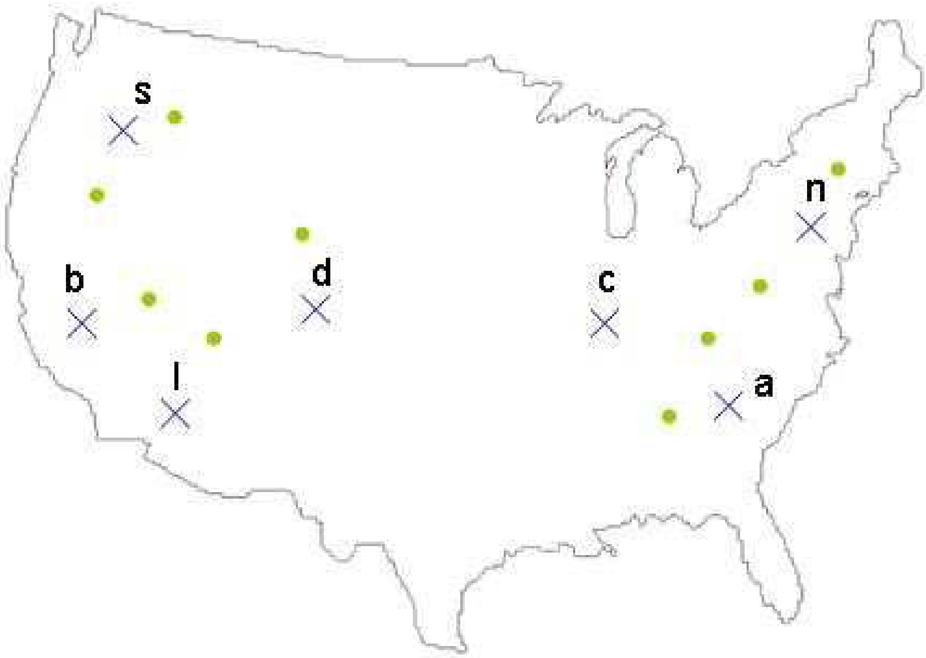
A New Greedy Approach for the Facility Location Problems

Kamal Jain
Microsoft Research

Joint work with:

- A. Saberi (Georgia Tech)
- M. Mahdian (MIT)

Facility Location Problem



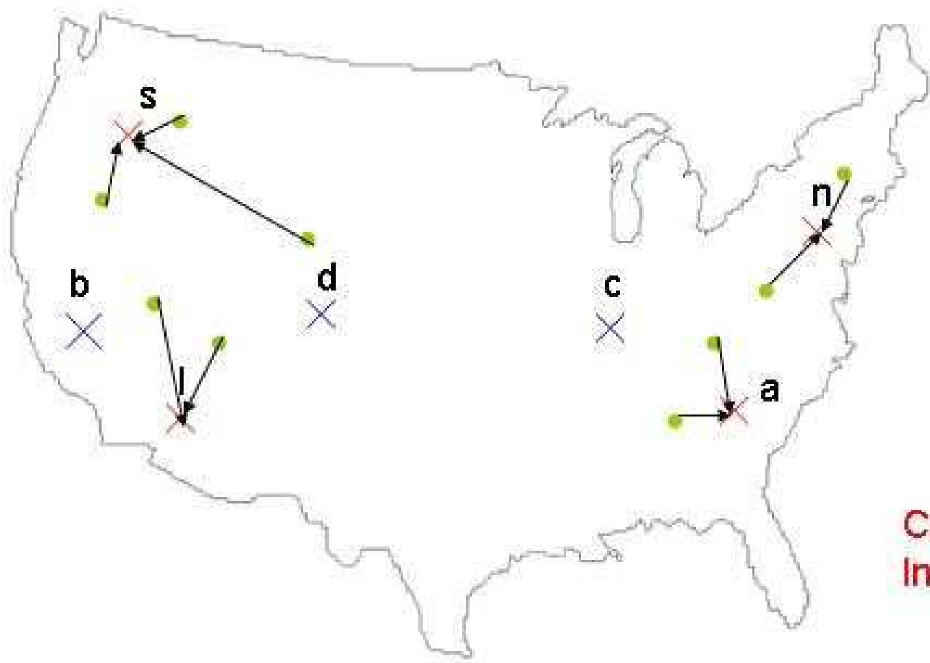
×: Server locations.

•: Clients.

Install servers and connect each client with a server.

Aim: minimize $\sum$ installation cost + $\sum$ connection cost

Facility Location Problem



×: Server locations.

•: Clients.

Install servers and connect each client with a server.

Connection Cost satisfy triangle inequality.

Aim: minimize $\sum$ installation cost + $\sum$ connection cost

Previous Results

factor	reference	technique(s)/running time
$O(\ln n_c)$	Hochbaum	greedy/$O(n^3)$
3.16	Shmoys, Tardos, Aardal	LP rounding
2.41	Guha, Khuller	LP rounding, greedy augmentation
1.736	Chudak	LP rounding
$5+\epsilon$	Korupolu, Plaxton, Rajaraman	local search/$O(n^6 \log(n/\epsilon))$
3	J., Vazirani	primal-dual/$O(n^2 \log n)$
1.853	Charikar, Guha	primal-dual, greedy aug./$O(n^3)$
1.728	Charikar, Guha	LP r., primal-dual, greedy aug.
1.861	Mahdian, Markakis, Saberi, Vazirani	greedy/$O(n^2 \log n)$
1.61	J., Mahdian, Saberi	greedy/$O(n^3)$
1.582	Sviridenko	LP rounding
1.52	Mahdian, Ye, Zhang	greedy, greedy augmentation/$O(n^3)$

Lower bound: 1.463 (Guha, Khuller)

Versatality of JV's algorithm

An α factor approximation algorithm:

$$F + C \leq \alpha(F^* + C^*)$$

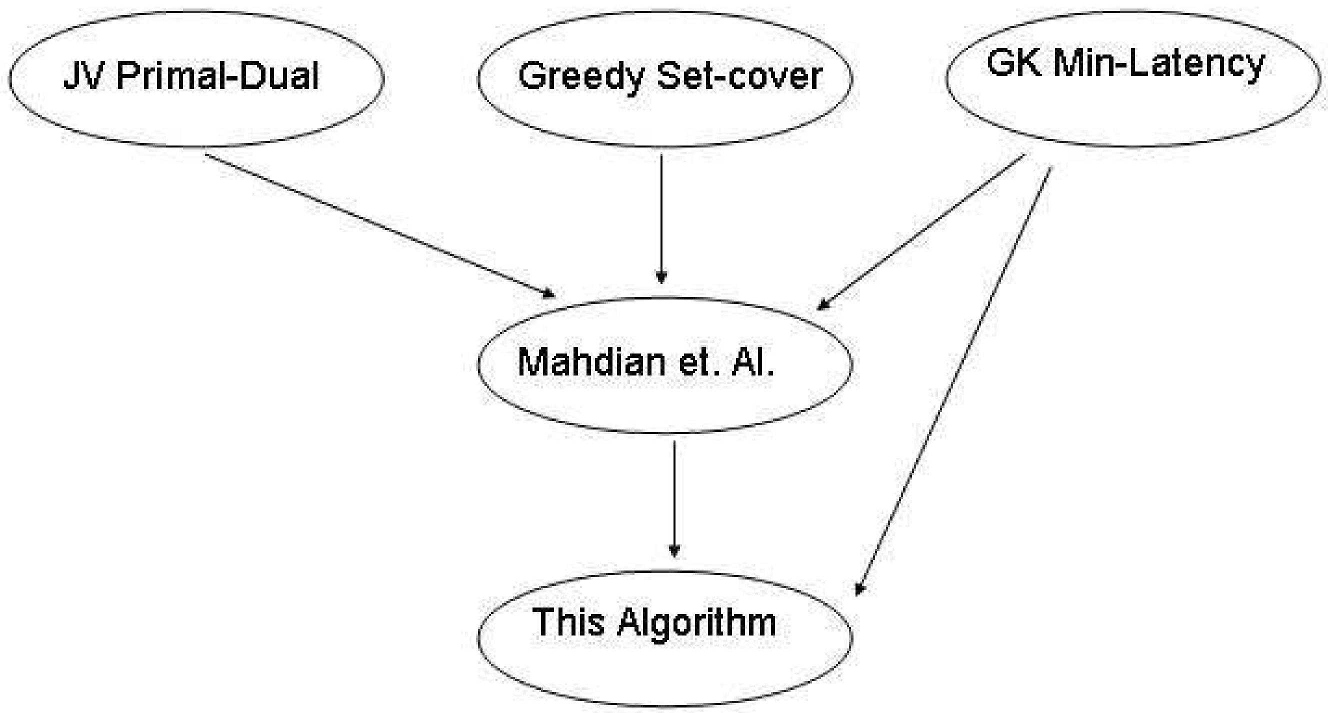
An LMP α factor approximation algorithm:

$$F + C \leq F^* + \alpha C^*$$

JV Algorithm is LMP 3 factor.

We present LMP 2 factor.

Leads

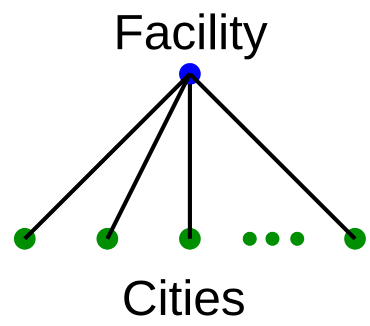

JV Primal-Dual
Greedy Set-cover
GK Min-Latency
Mahdian et. Al.
This Algorithm

LP Formulation

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in \mathcal{F}, j \in \mathcal{C}} c_{ij} x_{ij} + \sum_{i \in \mathcal{F}} f_i y_i \\
\text{subject to} \quad & \sum_{i \in \mathcal{F}} x_{ij} \geq 1 && \forall j \in \mathcal{C} \\
& x_{ij} \leq y_i && \forall i \in \mathcal{F}, j \in \mathcal{C} \\
& x_{ij}, y_i \in \{0, 1\} && \forall i \in \mathcal{F}, j \in \mathcal{C}
\end{aligned}
$$

Facility Location and Set Cover

A star consists of one facility and several cities.



The cost c_S of a star S is the sum of the opening cost of the facility and the connection costs between the facility and cities in S.

Let $\mathcal{R}$ be the collection of all stars.

We want to cover all cities with sets in $\mathcal{R}$.

Set Cover LP Formulation

set cover LP:

$$\begin{aligned} \text{minimize} \quad & \sum_{S \in \mathcal{R}} c_S x_S \\ \text{subject to} \quad & \forall j \in \mathcal{C}: \quad \sum_{S : j \in S} x_S \geq 1 \\ & \forall S \in \mathcal{R}: \quad x_S \geq 0 \end{aligned}$$

dual:

$$\begin{aligned} \text{maximize} \quad & \sum_{j \in \mathcal{C}} \alpha_j \\ \text{subject to} \quad & \forall S \in \mathcal{R}: \quad \sum_{j \in S \cap \mathcal{C}} \alpha_j \leq c_S \\ & \forall j \in \mathcal{C}: \quad \alpha_j \geq 0 \end{aligned}$$

JV Primal-Dual Algorithm

$$\text{maximize} \quad \sum_{j \in \mathcal{C}} \alpha_j$$

$$\text{subject to} \quad \forall S \in \mathcal{R}: \quad \sum_{j \in S \cap \mathcal{C}} \alpha_j \leq c_S$$

$$\forall j \in \mathcal{C}: \ \alpha_j \geq 0$$

Phase 1:

– Start at *time* $t = 0$. Set $\alpha_j = 0$ for every j.
(Think of α_j as the *budget* of city j.)

– Increase α_j for all *unconnected* cities j at the same rate until some Star goes tight.
Pick the Star and freeze the budget of every city in it.

Phase 2: Clean-up phase.

Paradigm

Primal-Dual Algorithm

– Feasibility of the Dual is the God given paradigm.

– Cost of the Dual may not be the cost of the Primal.

– A proof is needed that the Primal is still bounded.

Greedy + Dual Fitting

– Cost of the Dual must be the cost of the Primal.

– Dual may not be feasible.

– A proof is needed that infeasibility of each constraint is still bounded.

Algorithm - Greedy

- Start at *time* $t = 0$. Set $\alpha_j = 0$ for every j.
  - Increase α_j for all *unconnected* cities j at the same rate
    until some Star goes tight.
    Pick the Star.
  - From this iteration onwards, ignore
    the budget of every city and
    the cost of the facility
    in it.
  - Bounded within 1.86 factor. [Mahdian et. al. 2001]

Algorithm - Greedy + Local Improvement

– Start at *time* $t = 0$. Set $\alpha_j = 0$ for every j.

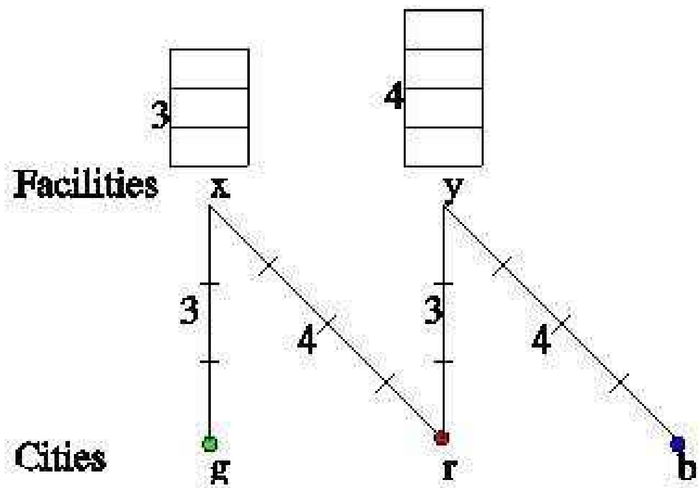
– Increase α_j for all *unconnected* cities j at the same rate
until some Star goes tight.
Pick the Star.

– From this iteration onwards, ignore
only the part of the budgets contributed towards the facility cost, and
the cost of the facility.

– Bounded within 1.61 factor. [This paper]

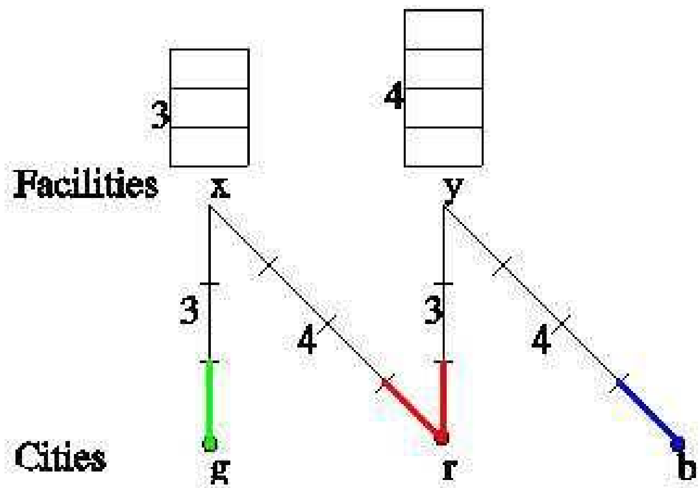
– Is a 2 LMP approximation algorithm. [This paper]

An Instance



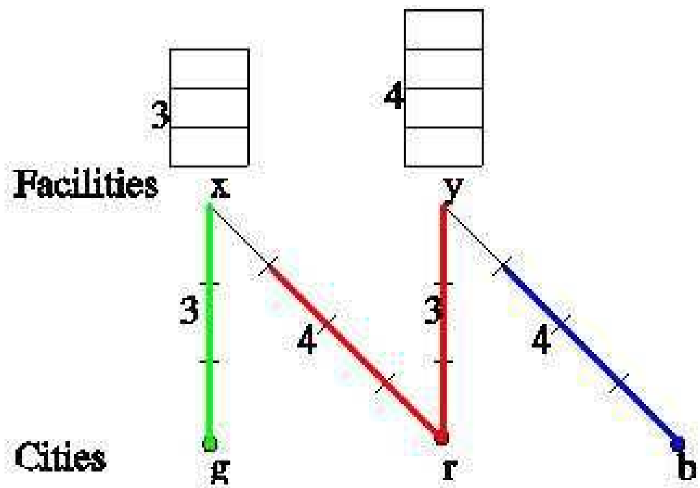
Time = 0

Time = 1

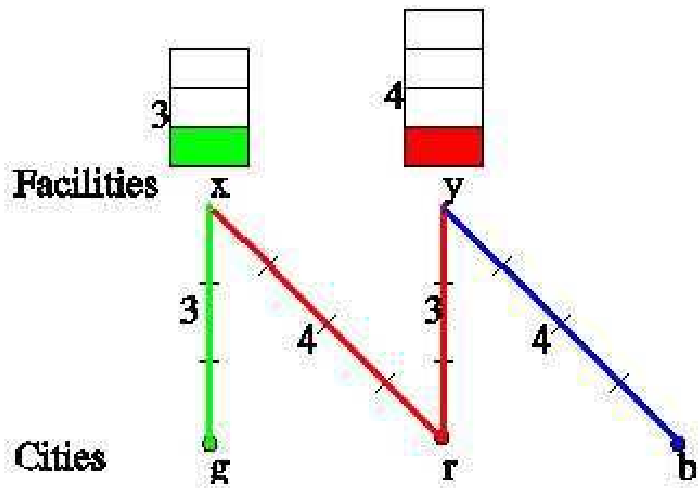


Cities start increasing their budgets.

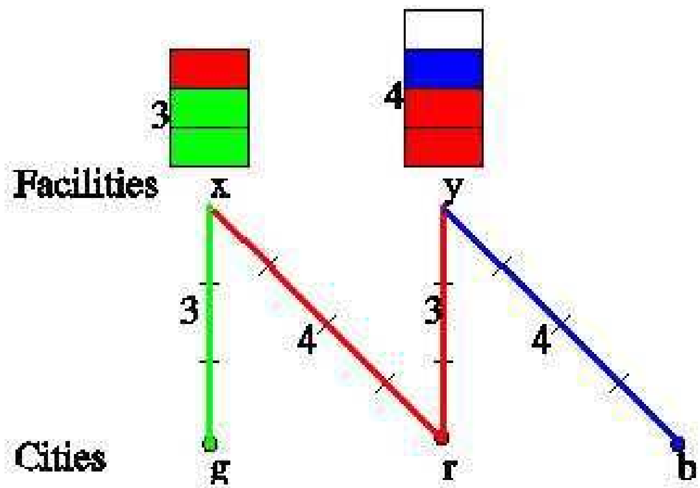
Time = 3



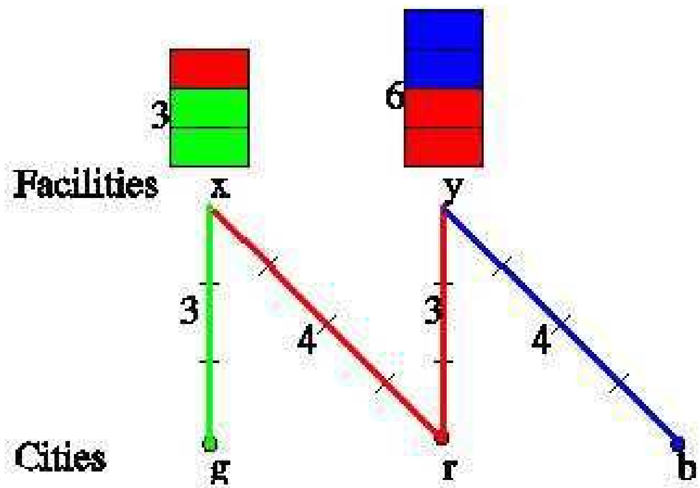
Time = 4



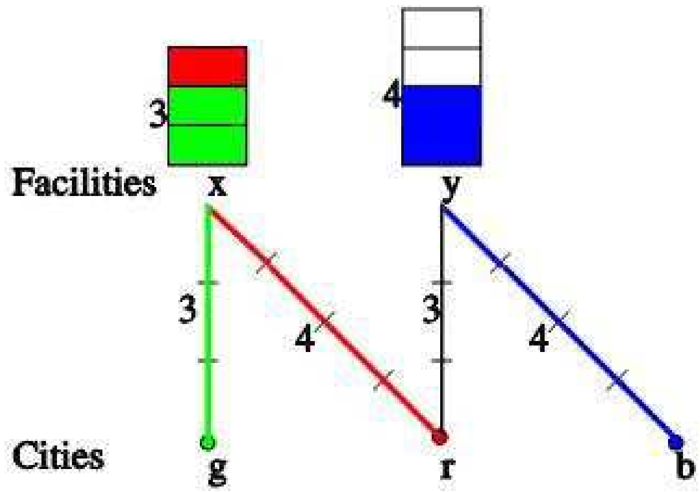
Time = 5



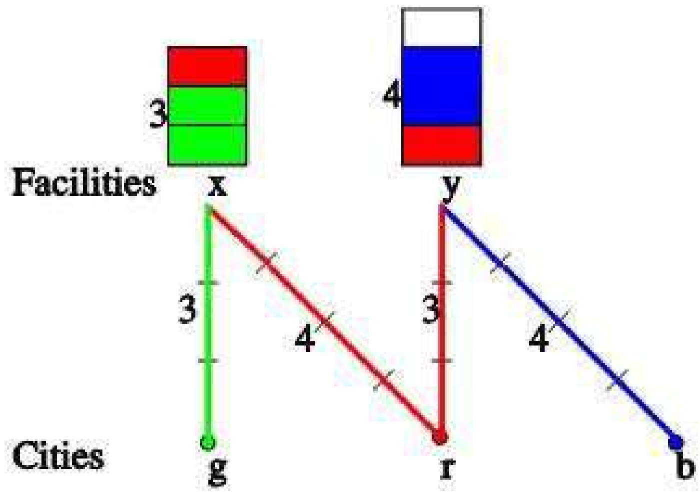
Time = 6
(Primal-Dual)



Time = 6
(Greedy)



Time = 6
(Greedy + Local Improvement)



Properties

- 1.61 factor - analysis is simpler.
- 2 LMP factor - analysis is actually simple.
- Almost as versatile as JV Primal-Dual algorithm.
- Invariant under local improvements.
- Variant under scaling. A scaling invariant version is 1.52 approximation factor. [Mahdian et. al. 2002].

Dual-Fitting

Recall the dual LP:

$$\begin{array}{ll} \text{maximize} & \sum_{j \in \mathcal{C}} \alpha_j \\ \text{subject to} & \forall S \in \mathcal{R}: \sum_{j \in S \cap \mathcal{C}} \alpha_j \leq c_S \\ & \forall j \in \mathcal{C}: \ \alpha_j \geq 0 \end{array}$$

The α_j's computed by Algorithm 1 are not a feasible dual solution.

However,

if we can prove that for every star S, $\sum_{j \in S \cap \mathcal{C}} \alpha_j \leq \gamma c_S$, then α_j / γ is a feasible dual solution.

An Optimization Problem

– Fix an algorithm. Find the worst instance of the problem.

– Find an instance with maximum possible γ.

– Computation of the maximum γ is a mathematical program, where variables represent an instance and a run of the algorithm. In this case it is a linear program called *Factor Revealing LP*.

– Any feasible dual solution of the Factor Revealing LP is an upper bound on γ.

Toy Problem–Set Cover

Given:

- set U of *elements*,
- *subsets* $S_1, S_2, \cdots S_m$ with costs $C(S_1), C(S_2), \cdots, C(S_m)$.

find:

- a collection of *sets*, with minimum cost, covering U.

Greedy Algorithm:

- Start at *time* $t = 0$. Set $\alpha_j = 0$ for every element j.
- Increase α_j for all *uncovered* elements j at the same rate until some Set goes tight.
  Pick the Set and ignore all its elements.

Approximation Factor

– We want to find out how much is the following dual constraint violated.

$$\sum_{i \in S} \alpha_i \leq C(S)$$

i.e., what is the maximum value of

$$\frac{\sum_{i \in S} \alpha_i}{C(S)}$$

– w.l.o.g assume $C(S) = 1$, $S = \{1, 2, \cdots, k\}$, and $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_k$.

Observation

$$(k - i + 1)\alpha_i \leq C(S) = 1$$

Factor-Revealing LP

$$
\begin{array}{lll}
\text{maximize} & \displaystyle\sum_{i=1}^{k} \alpha_i & \\
\text{subject to} & \alpha_i \le \alpha_{i+1} & 1 \le i \le k-1 \\
 & (k-i+1)\alpha_i \le 1 & 1 \le i \le k \\
 & \alpha_i \ge 0 & 1 \le i \le k
\end{array}
$$

Dual of Factor-Revealing LP

minimize $\sum_{i=1}^{k} y_i$

subject to $z_1 + k y_1 \geq 1$

$z_i - z_{i-1} + (k-i+1) y_i \geq 1 \quad 2 \leq i \leq k$

$z_i \geq 0 \quad 1 \leq i \leq k$

$y_i \geq 0 \quad \forall 1 \leq i \leq k$

A feasible solution: $z_i = 0$, $y_i = 1/(k-i+1)$.

– A greedy algorithm as versatile as primal-dual algorithms.

– Some hardness results.

– Furthering a computer aided proof technique of Dual-Fitting with Factor-Revealing LP.