Approximating (r,p)-Centroid on a Path

Joachim Spoerhase and Hans-Christoph Wirth

University of Würzburg · Department of Computer Science

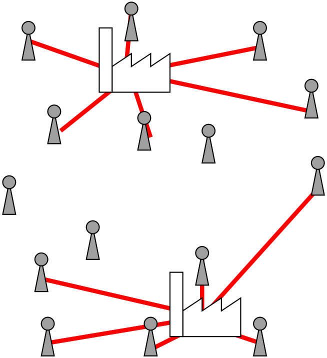
CTW · May 2008 · Gargnano

Contents

Contents

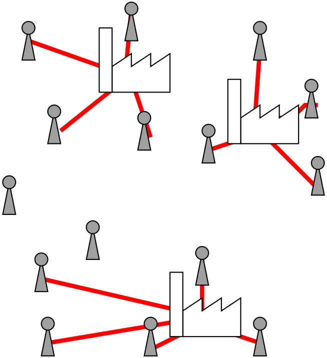
Location Problems

- **Input:** A Graph. Users located at nodes.
- **Output:** Placement of Facilities
- **Goal:** Serve customers, optimizing various costs



Location Problems

- **Input:** A Graph. Users located at nodes.
- **Output:** Placement of Facilities
- **Goal:** Serve customers, optimizing various costs



Location Problems (Examples)

Set X of facilities

Minimize Objective	Problem
$\sum_v d(v, X)$	*p*-Median
$\max_v d(v, X)$	*p*-Center
$\sum_{x \in X} f(x) + \sum_v d(v, X)$	Facility Location

Location Problems (Examples)

Set X of facilities

Minimize Objective	Problem
$\sum_v d(v, X)$	p-Median
$\max_v d(v, X)$	p-Center
$\sum_{x \in X} f(x) + \sum_v d(v, X)$	Facility Location

Location Problems (Examples)

Set X of facilities

Minimize Objective	Problem
$\sum_v d(v, X)$	p-Median
$\max_v d(v, X)$	p-Center
$\sum_{x \in X} f(x) + \sum_v d(v, X)$	Facility Location

Location Problems (Examples)

Set X of facilities

Minimize Objective	Problem
$\sum_v d(v, X)$	p-Median
$\max_v d(v, X)$	p-Center
$\sum_{x \in X} f(x) + \sum_v d(v, X)$	Facility Location

Prior: **one central planner**

Now: **two competitive providers**

Competitive Location

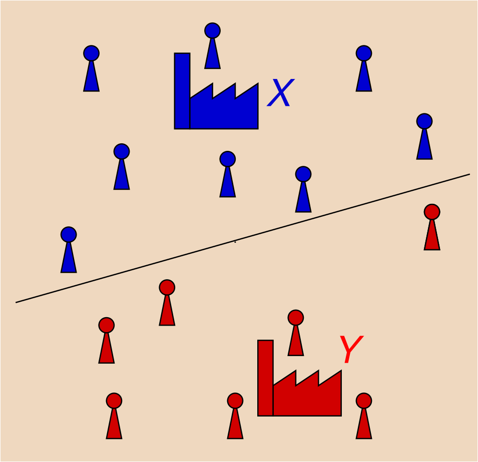
Competing providers locate facilities on a graph

- Graph $G = (V, E)$
- Edge lenghts $d\colon E \to \mathbb{R}_0^+$
- Node weights $w\colon V \to \mathbb{R}_0^+$
- Users are located at nodes
- Facilities may be placed along edges

User u prefers X over Y:

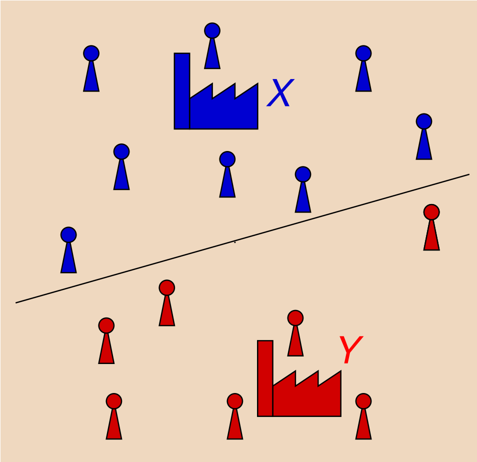
$$\begin{aligned} X \prec_u Y \quad &:\iff \quad d(u, X) < d(u, Y) \\ w(X \prec Y) \quad &: \text{weight of users preferring } X \end{aligned}$$

Example of Competitive Providers



- blue users prefer facility X
- red users prefer facility Y

Example of Competitive Providers



- blue users prefer facility X

  $w(X \prec Y) = 8$

- red users prefer facility Y

  $w(Y \prec X) = 7$

Game Rules

Providers place **sequentially** . . .

Game Rules

Providers place **sequentially** . . .

1. **Leader** places p facilities ($\rightsquigarrow$ leader problem)

Game Rules

Providers place **sequentially** . . .

1. **Leader** places p facilities ($\rightsquigarrow$ leader problem)
2. **Follower** places r facilities ($\rightsquigarrow$ follower problem)

Game Rules

Providers place **sequentially** …

1. **Leader** places p facilities ($\rightsquigarrow$ leader problem)
2. **Follower** places r facilities ($\rightsquigarrow$ follower problem)

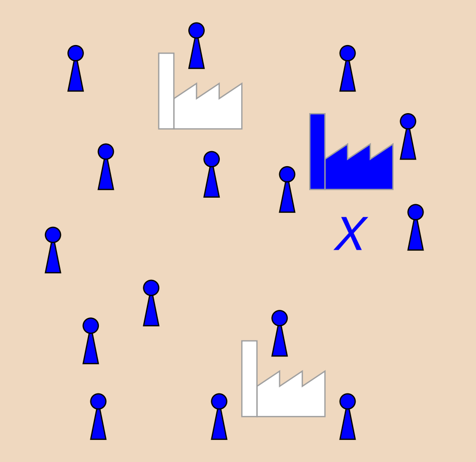
… each maximizes **own benefit** (sum of user demand)

Follower Problem

Let us start with the **follower problem**...

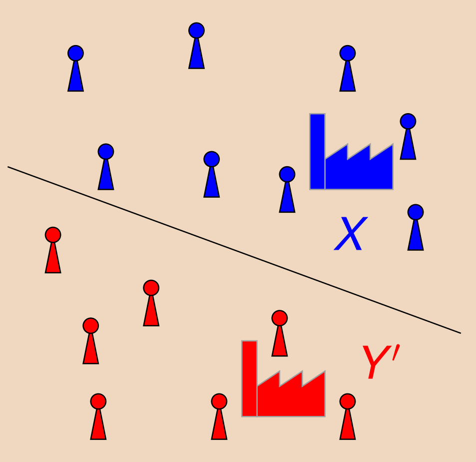
Example Follower Problem

Given X, what is the best choice for the follower?



Example Follower Problem

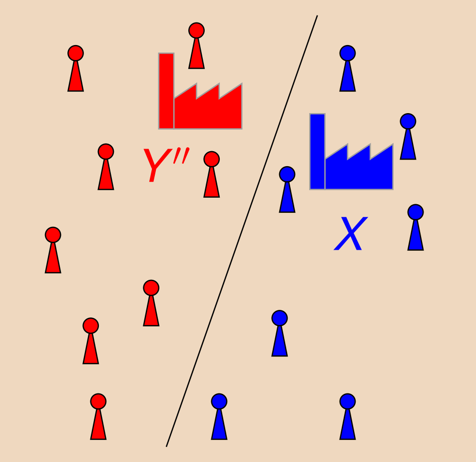
Given X, what is the best choice for the follower?



- $w(Y' \prec X) = 7$

Example Follower Problem

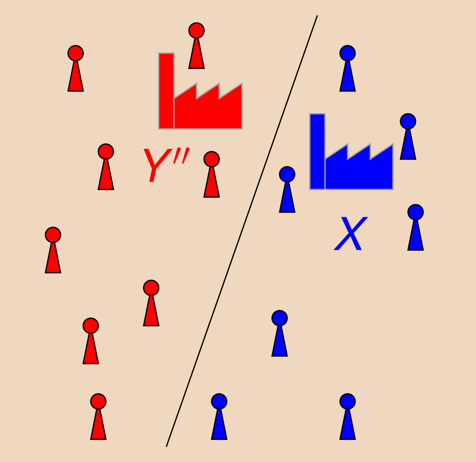
Given X, what is the best choice for the follower?



- $w(Y' \prec X) = 7$
- $w(Y'' \prec X) = 8$

Example Follower Problem

Given X, what is the best choice for the follower?



- $w(Y' \prec X) = 7$
- $w(Y'' \prec X) = 8$
- Max $w(\cdot \prec X)$

Definition Follower Problem

Definition (Hakimi '83)

Given p-element leader placement X_p, define

$$w_r(X_p) := \max_{\substack{Y_r \subseteq G \\ |Y_r| = r}} w(Y_r \prec X_p),$$

which denotes the **maximum gain** of **follower**.

Definition Follower Problem

Definition (Hakimi '83)

Given p-element leader placement X_p, define

$$w_r(X_p) := \max_{\substack{Y_r \subseteq G \\ |Y_r| = r}} w(Y_r \prec X_p),$$

which denotes the **maximum gain** of **follower**.

Definition (Hakimi '83)

We call an r-element follower placement Y_r with

$$w(Y_r \prec X_p) = w_r(X_p)$$

an (r, X_p)**-Medianoid**.

And the Leader Problem?

- Leader can **predict** the follower's reaction:

And the Leader Problem?

- Leader can **predict** the follower's reaction:
- $X_p \mapsto w_r(X_p)$

And the Leader Problem?

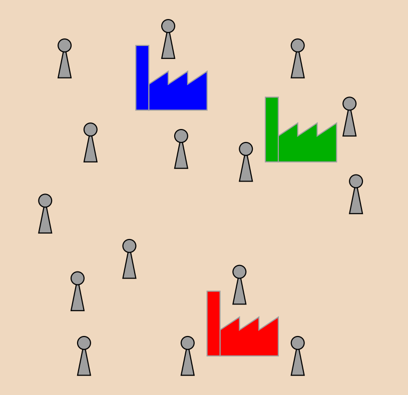
- Leader can **predict** the follower's reaction:
- $X_p \mapsto w_r(X_p)$
- ... the leader gets the rest

And the Leader Problem?

- Leader can **predict** the follower's reaction:
- $X_p \mapsto w_r(X_p)$
- ... the leader gets the rest
- $\rightsquigarrow$ find X_p which **minimizes** $w_r(\cdot)$

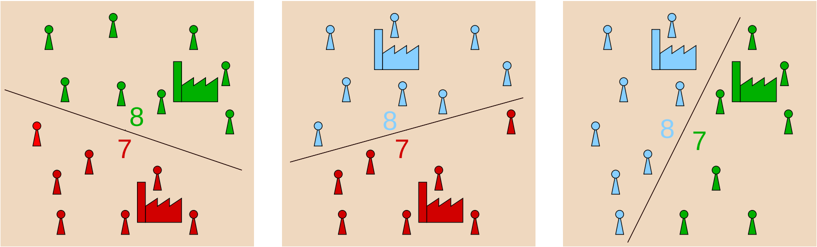
Example of Leader Problem

Open one of three facilities . . .



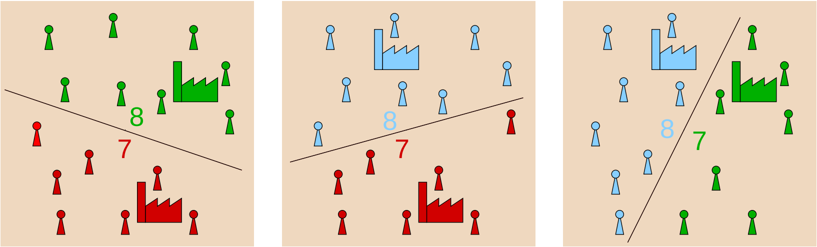
Example of Leader Problem

Open one of three facilities . . .



Example of Leader Problem

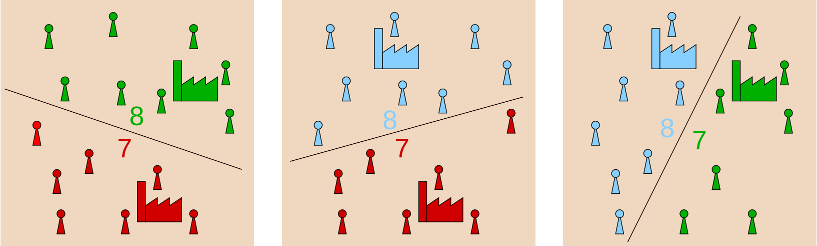
Open one of three facilities . . .



- $w_1(R) = \max\{w_1(G \prec R), w_1(B \prec R)\} = 8$

Example of Leader Problem

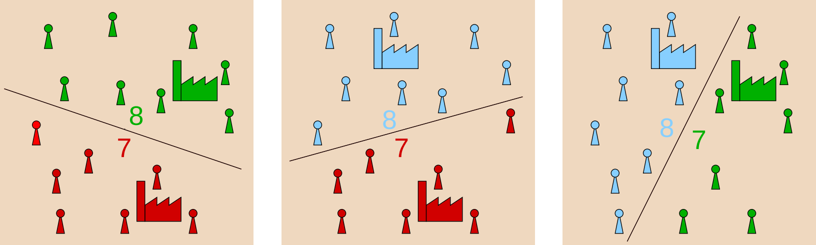
Open one of three facilities . . .



- $w_1(R) = \max\{w_1(G \prec R), w_1(B \prec R)\} = 8$
- $w_1(G) = \max\{w_1(B \prec G), w_1(R \prec G)\} = 8$

Example of Leader Problem

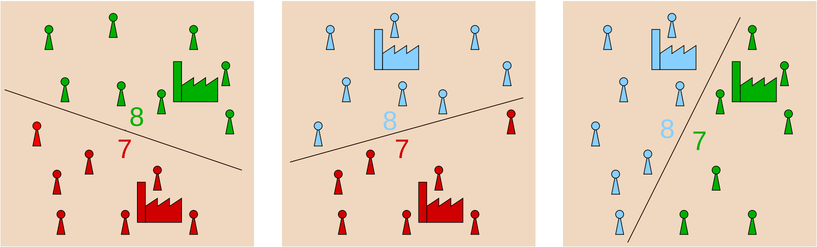
Open one of three facilities . . .



- $w_1(R) = \max\{w_1(G \prec R), w_1(B \prec R)\} = 8$
- $w_1(G) = \max\{w_1(B \prec G), w_1(R \prec G)\} = 8$
- $w_1(B) = \max\{w_1(G \prec B), w_1(R \prec B)\} = 7$

Example of Leader Problem

Open one of three facilities ...



- $w_1(R) = \max\{w_1(G \prec R), w_1(B \prec R)\} = 8$
- $w_1(G) = \max\{w_1(B \prec G), w_1(R \prec G)\} = 8$
- $w_1(B) = \max\{w_1(G \prec B), w_1(R \prec B)\} = 7$
- $\Rightarrow$ Blue is the optimal leader position

Definition Leader Problem

Definition (Hakimi '83)

Given r, p, we define

$$w_{r,p}(G) := \min_{\substack{X_p \subseteq G \\ |X_p| = p}} w_r(X_p),$$

Definition Leader Problem

Definition (Hakimi '83)

Given r, p, we define

$$w_{r,p}(G) := \min_{\substack{X_p \subseteq G \\ |X_p| = p}} w_r(X_p),$$

Definition (Hakimi '83)

Given r, p, we call a p-element leader placement X_p with

$$w_r(X_p) = w_{r,p}(G)$$

an (r, p)-**Centroid**.

Contents

Complexity on General Graphs

- Leader and follower problem are difficult (Hakimi '83)

Complexity on General Graphs

- Leader and follower problem are difficult (Hakimi '83)
- ... but leader problem is much more difficult

Complexity on General Graphs

- Leader and follower problem are difficult (Hakimi ’83)
- ... but leader problem is much more difficult
- Hakimi (’90) conjectured: it is “exceedingly difficult”

Complexity on General Graphs

- Leader and follower problem are difficult (Hakimi '83)
- ... but leader problem is much more difficult
- Hakimi ('90) conjectured: it is "exceedingly difficult"
- $\rightsquigarrow$ Is it in NP?

Complexity on General Graphs

- Leader and follower problem are difficult (Hakimi ’83)
- ... but leader problem is much more difficult
- Hakimi (’90) conjectured: it is “exceedingly difficult”
- $\rightsquigarrow$ Is it in NP?
- (Probably) it is not: (r, p)**-Centroid is Σ_2^p-complete** (Noltemeier, Spoerhase, Wirth ’07)

Why is Centroid so hard?

Computing the follower problem

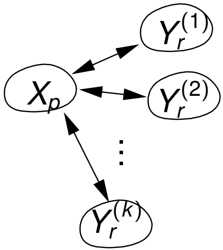
$$w_r(X_p) = \max_{Y_r} w(Y_r \prec X_p)$$

Why is Centroid so hard?

Computing the follower problem

$$w_r(X_p) = \max_{Y_r} w(Y_r \prec X_p)$$

- Compare X_p with **all** oppositions Y_r

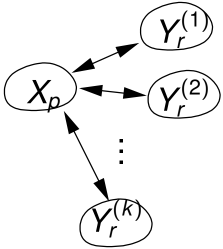


Why is Centroid so hard?

Computing the follower problem

$$w_r(X_p) = \max_{Y_r} w(Y_r \prec X_p)$$

- Compare X_p with **all** oppositions Y_r
- **One-stage** enumeration



Why is Centroid so hard?

Computing Centroid

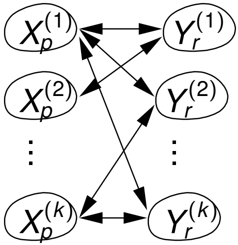
$$w_{r,p} = \min_{X_p} w_r(X_p) = \min_{X_p} \max_{Y_r} w(Y_r \prec X_p)$$

Why is Centroid so hard?

Computing Centroid

$$w_{r,p} = \min_{X_p} w_r(X_p) = \min_{X_p} \max_{Y_r} w(Y_r \prec X_p)$$

- Compare **all** X with **all** Y

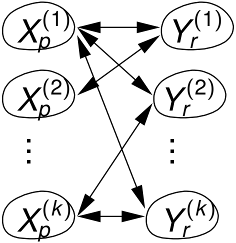


Why is Centroid so hard?

Computing Centroid

$$w_{r,p} = \min_{X_p} w_r(X_p) = \min_{X_p} \max_{Y_r} w(Y_r \prec X_p)$$

- Compare **all** X with **all** Y
- **Two-stage** optimization



Complexity on Trees

Follower Problem ((r, X_p)-Medianoid)

- Megiddo, Zemel, Hakimi '83: (r, X_p)**-Medianoid** is **computable** in $O(rn^2)$ on **trees**

Complexity on Trees

Follower Problem ((r, X_p)-Medianoid)

- Megiddo, Zemel, Hakimi '83: (r, X_p)**-Medianoid** is **computable** in $O(rn^2)$ on **trees**
- $\rightsquigarrow$ Bottom-up dynamic programming

Complexity on Trees

Follower Problem ((r, X_p)-Medianoid)

- Megiddo, Zemel, Hakimi '83: (r, X_p)**-Medianoid** is **computable** in $O(rn^2)$ on **trees**
- $\rightsquigarrow$ Bottom-up dynamic programming

Leader Problem ((r, p)-Centroid)

Complexity on Trees

Follower Problem ((r, X_p)-Medianoid)

- Megiddo, Zemel, Hakimi '83: (r, X_p)**-Medianoid** is **computable** in $O(rn^2)$ on **trees**
- $\rightsquigarrow$ Bottom-up dynamic programming

Leader Problem ((r, p)-Centroid)

- Is (r, p)**-Centroid easy on trees?**

Complexity on Trees

Follower Problem ((r, X_p)-Medianoid)

- Megiddo, Zemel, Hakimi '83: (r, X_p)**-Medianoid** is **computable** in $O(rn^2)$ on **trees**
- $\rightsquigarrow$ Bottom-up dynamic programming

Leader Problem ((r, p)-Centroid)

- Is (r, p)**-Centroid easy on trees?**
- ... a long standing open question (Hakimi '90, Eiselt,Laporte '96, Benati '00)

Complexity on Trees

Follower Problem ((r, X_p)-Medianoid)

- Megiddo, Zemel, Hakimi '83: (r, X_p)**-Medianoid** is **computable** in $O(rn^2)$ on **trees**
- $\rightsquigarrow$ Bottom-up dynamic programming

Leader Problem ((r, p)-Centroid)

- Is (r, p)**-Centroid easy on trees?**
- ... a long standing open question (Hakimi '90, Eiselt,Laporte '96, Benati '00)
- Answer: (r, p)**-Centroid is hard** even on a **path!** (Spoerhase, Wirth '08)

Approximability

Follower Problem ((r, X_p)-Medianoid)

- Tight bound $(1 - \frac{1}{e})$ of approximability (Noltemeier, Spoerhase, Wirth '07)

Approximability

Follower Problem ((r, X_p)-Medianoid)

- Tight bound $(1 - \frac{1}{e})$ of approximability (Noltemeier, Spoerhase, Wirth '07)

Leader Problem ((r, p)-Centroid)

Approximability

Follower Problem ((r, X_p)-Medianoid)

- Tight bound $(1 - \frac{1}{e})$ of approximability (Noltemeier, Spoerhase, Wirth '07)

Leader Problem ((r, p)-Centroid)

- Not approximable within $n^{1-\varepsilon}$ on planar graphs (Noltemeier, Spoerhase, Wirth '07)

Approximability

Follower Problem ((r, X_p)-Medianoid)

- Tight bound $(1 - \frac{1}{e})$ of approximability (Noltemeier, Spoerhase, Wirth '07)

Leader Problem ((r, p)-Centroid)

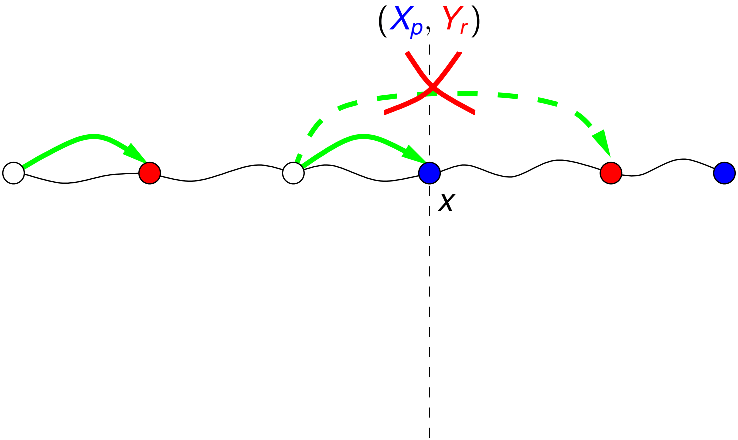
- Not approximable within $n^{1-\varepsilon}$ on planar graphs (Noltemeier, Spoerhase, Wirth '07)
- $\rightsquigarrow$ **Approximability on paths and trees?**

Contents

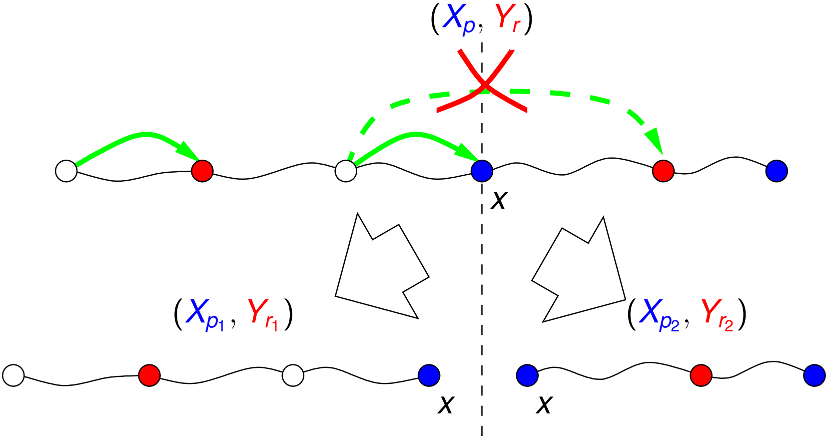
Dynamic Programming

- Typical Approach: pseudopolynomial algorithm and scaling
- Dynamic Programming requires **optimal substructure property**

Optimal Substructure Paradoxon

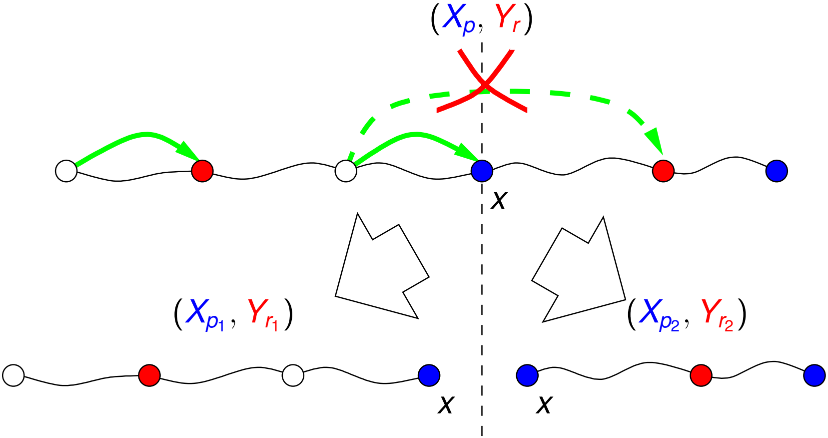


Optimal Substructure Paradoxon



$$w_r(X_p) = w_{r_1}(X_{p_1}) + w_{r_2}(X_{p_2})$$

Optimal Substructure Paradoxon

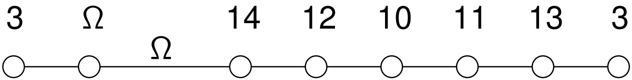


$$w_r(X_p) = w_{r_1}(X_{p_1}) + w_{r_2}(X_{p_2})$$

However the problem is not splittable!

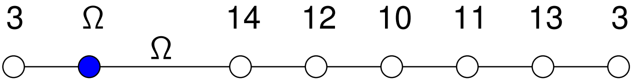
Optimal Substructure Paradoxon

Determine $(2, 2)$-Centroid ...



Optimal Substructure Paradoxon

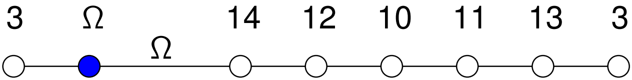
Determine $(2, 2)$-Centroid …



- Leader always locates at Ω node $\rightsquigarrow$ split

Optimal Substructure Paradoxon

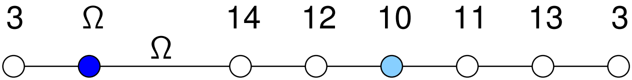
Determine $(2, 2)$-Centroid . . .



- Leader always locates at Ω node $\rightsquigarrow$ split
- Locate the second facility at the right subpath

Optimal Substructure Paradoxon

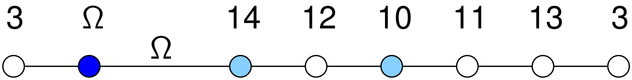
Determine $(2,2)$-Centroid ...



- Leader always locates at Ω node $\rightsquigarrow$ split
- Locate the second facility at the right subpath
- Local $(1,1)$-centroid is at 10 (split equally)

Optimal Substructure Paradoxon

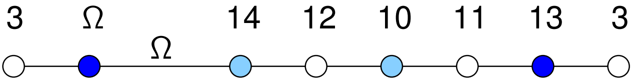
Determine $(2, 2)$-Centroid . . .



- Leader always locates at Ω node $\rightsquigarrow$ split
- Locate the second facility at the right subpath
- Local $(1, 1)$-centroid is at 10 (split equally)
- Local $(2, 1)$-centroid is at 14 (heaviest node)

Optimal Substructure Paradoxon

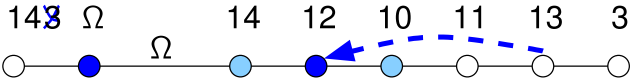
Determine $(2,2)$-Centroid . . .



- Leader always locates at Ω node $\rightsquigarrow$ split
- Locate the second facility at the right subpath
- Local $(1,1)$-centroid is at 10 (split equally)
- Local $(2,1)$-centroid is at 14 (heaviest node)
- But 13 is the best choice!

Optimal Substructure Paradoxon

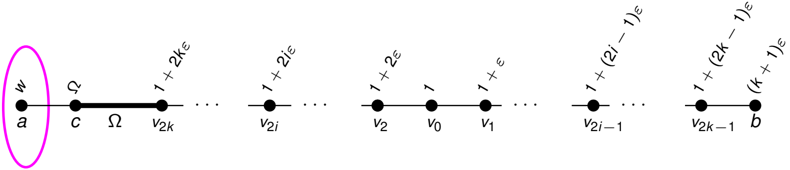
Determine $(2,2)$-Centroid . . .



- Leader always locates at Ω node $\rightsquigarrow$ split
- Locate the second facility at the right subpath
- Local $(1,1)$-centroid is at 10 (split equally)
- Local $(2,1)$-centroid is at 14 (heaviest node)
- But 13 is the best choice!
- If we change the weight of the leftmost node. . .
- . . . the optimum switches to 12

Optimal Substructure Paradoxon

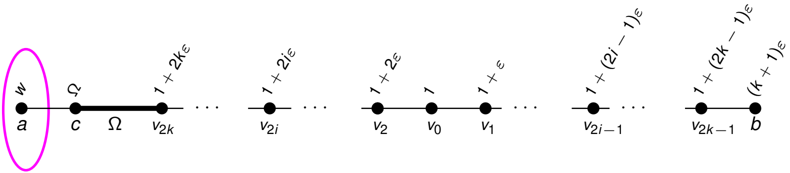
Generalization . . .



- **Every node** v_j can be part of a $(2,2)$-centroid . . .

Optimal Substructure Paradoxon

Generalization . . .



- **Every node** v_j can be part of a $(2, 2)$-centroid . . .
- . . . depending on the weight w

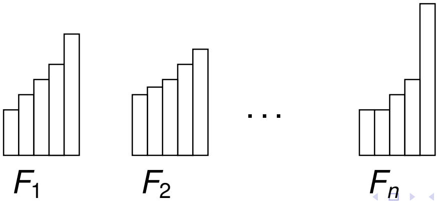
Contents

k-Sum Optimization

k-Sum Optimization Problem (Punnen '96)

Given:

- Set $E = \{e_1, \ldots, e_m\}$ of weighted ground elements
- Family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ with $F_i \subseteq E$



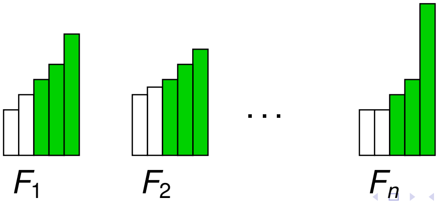
k-Sum Optimization

k-Sum Optimization Problem (Punnen '96)

Given:

- Set $E = \{e_1, \dots, e_m\}$ of weighted ground elements
- Family $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$ with $F_i \subseteq E$

- $c_k(F_i) :=$ sum of the k **heaviest** elements in F_i
- $c(F_i) :=$ sum of **all** elements $\rightsquigarrow$ **minisum** problem



k-Sum Optimization

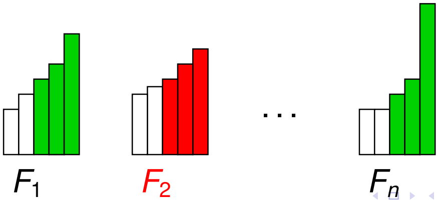
k-Sum Optimization Problem (Punnen '96)

Given:

- Set $E = \{e_1, \ldots, e_m\}$ of weighted ground elements
- Family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ with $F_i \subseteq E$

Task: Compute F_i minimizing $c_k(\cdot)$

- $c_k(F_i) :=$ sum of the k **heaviest** elements in F_i
- $c(F_i) :=$ sum of **all** elements $\rightsquigarrow$ **minisum** problem



k-Sum Optimization

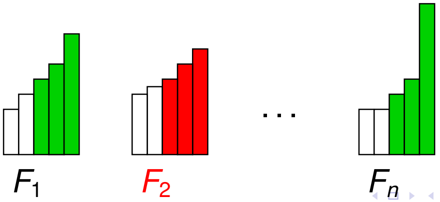
k-Sum Optimization Problem (Punnen '96)

Given:

- Set $E = \{e_1, \ldots, e_m\}$ of weighted ground elements
- Family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ with $F_i \subseteq E$

Task: Compute F_i minimizing $c_k(\cdot)$

- $c_k(F_i) :=$ sum of the k **heaviest** elements in F_i
- $c(F_i) :=$ sum of **all** elements $\rightsquigarrow$ **minisum** problem
- Examples: k-sum shortest path, k-sum MST



k-Sum Optimization

Theorem (Punnen '96)

A k**-sum optimization problem** is **easy** if the corresponding **minisum problem** is tractable

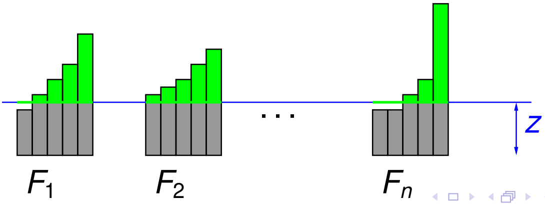
- Introduce "Water Level" z

k-Sum Optimization

Theorem (Punnen '96)

A k**-sum optimization problem** is **easy** if the corresponding **minisum problem** is tractable

- Introduce "Water Level" z
- $w^{(z)}(e_i) := w(e_i) \dot{-} z := \max\{w(e_i) - z, 0\}$

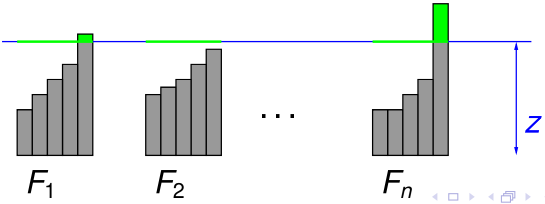


k-Sum Optimization

Theorem (Punnen '96)

A k**-sum optimization problem** is **easy** if the corresponding **minisum problem** is tractable

- Introduce "Water Level" z
- $w^{(z)}(e_i) := w(e_i) \dot{-} z := \max\{w(e_i) - z, 0\}$
- Consider minisum problem under $w^{(z)}(\cdot)$

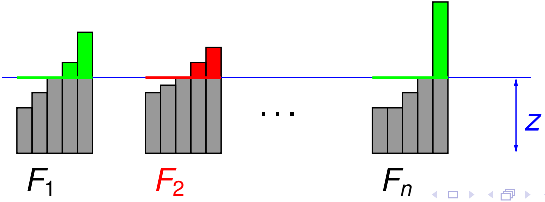


k-Sum Optimization

Theorem (Punnen '96)

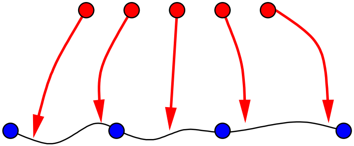
A k**-sum optimization problem** is **easy** if the corresponding **minisum problem** is tractable

- Introduce "Water Level" z
- $w^{(z)}(e_i) := w(e_i) \dot{-} z := \max\{w(e_i) - z, 0\}$
- Consider minisum problem under $w^{(z)}(\cdot)$
- For some z, **minisum** under $w^{(z)}$ and k**-sum** under w become **equivalent**

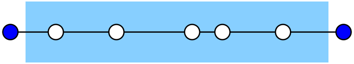


Formulation as *k*-sum problem

Observation: Follower distributes its facilities among leader intervals

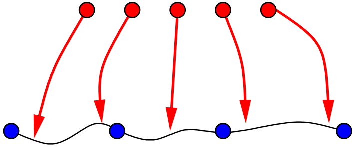


Within a leader interval the follower places either. . .

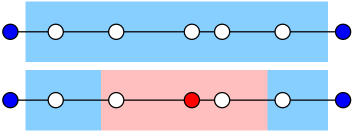
- 0 Servers 

Formulation as k-sum problem

Observation: Follower distributes its facilities among leader intervals



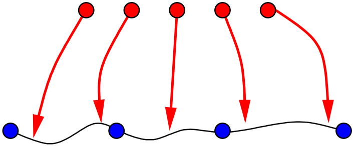
Within a leader interval the follower places either...



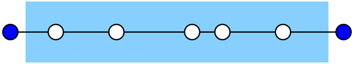
- 0 Servers
- 1 Servers
  $\rightsquigarrow$ incremental gain δ'

Formulation as k-sum problem

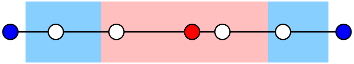
Observation: Follower distributes its facilities among leader intervals



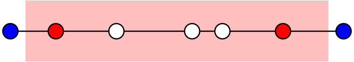
Within a leader interval the follower places either. . .



- 0 Servers



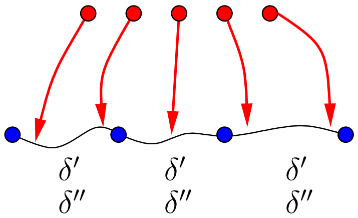
- 1 Servers
  $\rightsquigarrow$ incremental gain δ'



- 2 Servers
  $\rightsquigarrow$ incremental gain $\delta'' \leq \delta'$

Formulation as k-sum problem

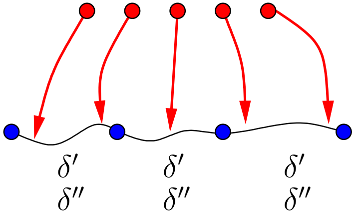
Hence the follower chooses the r largest incremental gains



- We consider X_p as sorted vector of $2p$ incremental gains $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_{2p}$

Formulation as k-sum problem

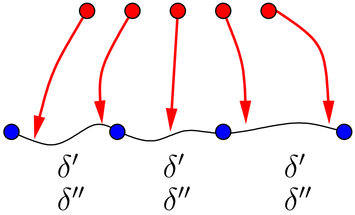
Hence the follower chooses the r largest incremental gains



- We consider X_p as sorted vector of $2p$ incremental gains $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_{2p}$
- Observe that $w_r(X_p) = \sum_{i=1}^{r} \delta_i$

Formulation as k-sum problem

Hence the follower chooses the r largest incremental gains

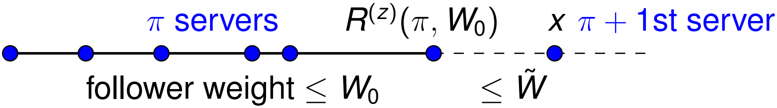


- We consider X_p as sorted vector of $2p$ incremental gains $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_{2p}$
- Observe that $w_r(X_p) = \sum_{i=1}^{r} \delta_i$
- According to Punnen: need to solve the minisum problem under reduced cost $w^{(z)}(X_p) := \sum_{i=1}^{2p} (\delta_i \dot{-} z)$

Solving the minisum problem

This new problem **is** splittable.

$$R^{(z)}(\pi+1, W) = \max \left\{ x \;\middle|\; \begin{array}{l} \exists(W_0, \tilde{W}) : \; W_0 + \tilde{W} = W \\ \text{and} \\ w^{(z)}\left(R^{(z)}(\pi, W_0), x\right) \leq \tilde{W} \end{array} \right\}$$



- $R^{(z)}$ rightmost leader position of placing π servers such that $w^{(z)}$ restricted to the interval $[0, R^{(z)}]$ does not exceed W
- $w^{(z)}(a, b)$ reduced gain in interval $[a, b]$

Approximating the (r, p)-Centroid

Computing (r, p)-centroid

An (r, p)-centroid of a path P can be computed in pseudo-polynomial running time $O(p \cdot w(P)^2 \cdot n^2)$.

and by scaling we obtain. . .

Approximating (r, p)-centroid

There is a FPTAS for (r, p)-centroid on a path.

	(r, p)-centroid *(leader problem)*		(r, X_p)-medianoid
	absolute	discrete	*(follower problem)*
arb. r arb. p	NP-hard on path [Spoerhase, Wirth 08] FPTAS on path	$O(pn^4)$ on path [Spoerhase, Wirth 08] NP-hard on spider [Spoerhase, Wirth 08] Σ_2^p-complete on graph [Noltemeier, Spoerhase, Wirth 07]	$O(n)$ on path [Megiddo et al. 83] $O(rn^2)$ on tree [Megiddo et al. 83] NP-hard on graph [Megiddo et al. 83]
$r = 1$ arb. p	$O(n^3 \log W \log D)$ on tree [Spoerhase, Wirth 08]	$O(n^2 (\log n)^2 \log W)$ on tree [Spoerhase, Wirth 08] NP-hard on pathwidth bounded graph [Spoerhase, Wirth 08]	$O(n (\log n)^2 / \log \log n)$ on tree [Spoerhase, Wirth 07] $O(n^2 \log n + nm)$ on graph [by enumeration]
$r = 1$ $p = 1$	$O(n^4 m^2 \log mn \log W)$ on graph [Hansen, Labbe 88]	$O(n^3)$ on graph [Campos, Moreno 03]	

	(r,p)-centroid		(r,X_p)-medianoid
	hardness	approximability	approximability
Graph	Σ_2^p-complete [NSW 07]	Lower bound $n^{1-\varepsilon}$ [NSW 07]	$(1-\frac{1}{e})$ [NSW 07]
Tree	NP-hard [Spoerhase, Wirth 08]	r	1 [Megiddo et al.]
Path	NP-hard [Spoerhase, Wirth 08]	FPTAS	

Open Questions

- There is a simple r-approximation on trees
- Are better approximations factors possible?
- Sublinear factors on general graphs?

	(r,p)-centroid		(r,X_p)-medianoid
	hardness	approximability	approximability
Graph	Σ_2^p-complete [NSW 07]	Lower bound $n^{1-\varepsilon}$ [NSW 07]	$(1-\frac{1}{e})$ [NSW 07]
Tree	NP-hard [Spoerhase, Wirth 08]	r	1 [Megiddo et al.]
Path	NP-hard [Spoerhase, Wirth 08]	FPTAS	