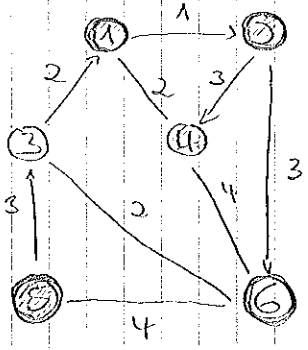
3. The Uncapacitated Facility Location Problem

3.1. The general Uncapacitated Fac. Loc. Prob.

3.1.1 Ex. (Schöbel & Schmidt [2003]) & Def. (Uncapacitated Fac. Loc. Problem):



$G = (V, E \cup A)$ network

$I \subseteq V$ customers (clients)

$J \subseteq V$ potential facilities

$d_{ij} \in \mathbb{R}_{\geq 0},\ ij \in A$ transportation costs $i \to j$

$f_i \in \mathbb{R}_{\geq 0},\ i \in I$ set-up costs

$A \subseteq I \times J$

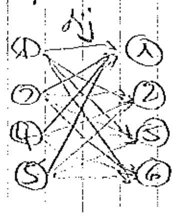
Uncapacitated Facility Loc. Prob.:

$$(UFL) \min_{I^* \subseteq I} f(I^*) + \sum_{j \in J} \min_{i \in I^*} d_{ij}$$

$I = \{1, 2, 5, 6\}$

$J = \{1, 2, 4, 5\}$

$$D = (d_{ij}) = \begin{array}{c|cccc|c|c|c} & 1 & 2 & 5 & 6 & \Sigma & \rho = & \rho = \\ \hline 1 & 0 & 1 & 8 & 4 & 13 & 18 & 15 \checkmark \\ 2 & 5 & 0 & 7 & 3 & 15 & 20 & \\ 4 & 2 & 3 & 8 & 4 & 17 & 22 & \\ 5 & 6 & 3 & 0 & 4 & 13 & 18 & \checkmark \end{array}$$



associated bipartite graph $G = (I \cup J, A)$

$f = (f_i) = (5, 5, 5, 5)$

3.1.2. Prop. (IP-Formulation for the UFL, Balinski [1965]).

$y_i \in \{0,1\},\ i \in I$ facility setup variables

$x_{ij} \in \{0,1\},\ j \in J$ client assignment variables

$$(UFL) \quad \min \sum_{i \in I} f_i y_i + \sum_{ij \in A} d_{ij} x_{ij}$$

(i) $\sum_{i \in I} x_{ij} \geq 1 \quad \forall j \in J$

(ii) $y_i \geq x_{ij} \quad \forall ij \in A$

(iii) $y_i \in \mathbb{Z}_+ \quad \forall j \in J$

(iv) $x_{ij} \in \mathbb{Z}_+ \quad \forall ij \in A$

a) (UFL) has an optimal 0/1-solution.

b) (UFL) $\Leftrightarrow$ (UFL)(i),(ii),(iii), $x_{ij} \geq 0\ \forall ij \in A$, i.e., the client assignment variables are automatically integral.

Peter Coup ← Proof: ex. sheet 11. □

3.1.3. Prop. (Set Covering Model for the UFL):

$J(i) := \{j \in J : ij \in A\}$ clients covered by fac. i

$\mathcal{J} := \{(i, J') : i \in I,\ \emptyset \neq J' \subseteq J(i)\}$

$c_{(i,J')} := f_i + \sum_{j \in J'} d_{ij}$

$$(SCP) \quad \min \sum_{(i,J') \in \mathcal{J}} c_{(i,J')} z_{(i,J')}$$

$$\sum_{J' \ni j} z_{(i,J')} \geq 1 \quad \forall j \in J$$

$$z_{(i,J')} \in \{0,1\} \quad \forall (i,J') \in \mathcal{J}$$

Then there is a one-to-one correspondence between optimal solutions of (SCP) and (UFL).

Proof: Ex. sheet 11. $\square$

Cor. 3.1.3: a) The greedy algorithm for UFL is $H(\max |J(i)|)$-approximate.

b) UFL is APX-hard.

Proof: a) Follows from Thm. 2.6.8

b) SCP is APX-hard. Given an SCP $\min c x, Mx \geq 1, x \in \{0,1\}^n$, construct an UFL (in polytime)

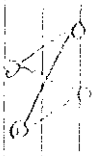
$I = \{1, \dots, n\}, \; J = \{1, \dots, m\}, \; A = \{ij : m_{ij} = 1\}, \; f_j = c_j$.

Then (UFL) $\Leftrightarrow$ (SCP) and the equivalence is approximation preserving. $\square$

Obs 3.1.4 (Relation to the p-median problem):

a) (UFL) $\Leftrightarrow$ $|I|/I/\cdot/d_{ij}/\Sigma$

b) $p/V/\cdot/d_{ij}/\Sigma \Rightarrow$ (UFL), (V) $\sum_{i \in I} y_i \leq p$.



3.2. The Metric Uncapacitated Facility Location Problem

3.2.1 Def. (Metric UFL): A UFL s.t.

i) $I, J \subseteq \mathbb{R}^k, \; k \in \mathbb{N}$, I, J are embedded in $\mathbb{R}^k$

ii) $A = I \times J$, all assignments are possible

iii) $d_{ij} = \|i,j\|$, distances with some norm

is a metric UFL (MUFL).

3.2.3. Alg. (LP Rounding):

Input: (MUFL) $I, J \subseteq \mathbb{R}^k$, $d_{ij} = \|i,j\| \; \forall ij \in I \times J$, $f_i \in \mathbb{R}_{\geq 0} \; \forall i \in I$

Output: (x,y) feasible for (MUFL)

1. LP Solving: $(\bar{x}, \bar{y}) \leftarrow \operatorname{argmin} (\text{MUFL})_{LP}$, where

$$(\text{MUFL})_{LP} \quad \sum_{i \in I} f_i y_i + \sum_{ij \in A} d_{ij} x_{ij}$$

$$\sum_{i\in I} x_{ij} \ge 1 \quad \forall j \in J$$
$$y_i \ge x_{ij} \quad \forall ij \in A$$
$$y_i \ge 0 \quad \forall i \in I$$
$$x_{ij} \ge 0 \quad \forall ij \in A$$

1. Filtering:

$D_j \leftarrow \sum_{i\in I} d_{ij}\bar{x}_{ij}$ "total fractional distance"

$N_j \leftarrow \{ i \in I : \underbrace{\bar{x}_{ij} > 0}_{\text{"fractionally assigned"}} \wedge \underbrace{d_{ij} \le 2D_j}_{\text{"near"}} \}$ "near frac. assigned fac."

3.0.4 Claim: $\sum_{i\in N_j} \bar{x}_{ij} \ge \frac{1}{2} \quad \forall j \in J$

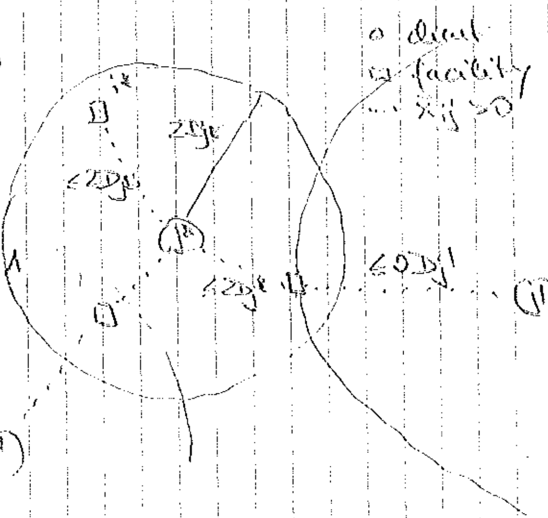
$$x'_{ij} \leftarrow \begin{cases} 0, & i \notin N_j \\ 2\bar{x}_{ij}, & \text{else} \end{cases} \quad \forall ij \in A$$

$$y'_i \leftarrow \max_{j\in J} x'_{ij}$$

3.0.5 Obs: (x', y') is feasible for (MUFL)$_{LP}$, and

i) $x'_{ij} > 0 \Rightarrow d_{ij} \le 2D_j$

ii) $c(x', y') \le 2c(\bar{x}, \bar{y})$



2. Rounding:

2a. $J^* = \emptyset$, $(x^0, y^0) \leftarrow (x', y')$, $k \leftarrow 1$

2b $j^k \leftarrow \arg\min_{j \in J\setminus J^*} D_j$

$i^k \leftarrow \arg\min_{i\in N_{j^k}} f_i$

$EN_{j^k} \leftarrow \{ j \in J\setminus J^* : i^k \in N_j \}$ "extended neighborhood of client j^k"

$$y^k_i \leftarrow \begin{cases} 1, & i = i^k \\ 0, & i \in N_{j^k}\setminus\{i^k\} \\ y^{k-1}_i, & \text{else} \end{cases}$$ select i^k, drop $N_{j^k}\setminus\{i^k\}$

$$x^k_{ij} \leftarrow \begin{cases} 1, & i = i^k, \ j \in EN_{j^k} \\ 0, & i \in N_{j^k}\setminus\{i^k\}, \ j \in J \\ 0, & i \notin N_{j^k}, \ j \in EN_{j^k} \\ x^{k-1}_{ij}, & \text{else} \end{cases}$$ assign all uncov. clients in EN_{j^k} to i^k

$J^* \leftarrow J^* \cup EN_{j^k}$

goto 2b

2c. $(x'', y'') \leftarrow (x^k, y^k)$, output (x'', y'').

3.2.6 Claim:

$D_{j^*} \le D_{j'}$

i) $j' \in N_{j^*} \Rightarrow d_{i^*j'} \le 2 \cdot 2D_{j^*} + 2D_{j'} \le 6D_{j'}$

ii) $f_{i^*} \le \sum_{i \in N_{j^*}} \underbrace{y'_i}_{\bar{x}_{ij^*}} f_i$, $\quad \underbrace{\sum_{i \in N_{j^*}} \bar{x}_{ij^*}}_{\le 1}$

iii) $$\underbrace{\sum_{i \in I} f_i y''_i}_{\overset{ii)}{\le} \sum_{i \in I} f_i y'_i \;\overset{3.2.5(ii)}{\le}\; 2\sum_{i \in I} f_i \bar{y}_i} + \underbrace{\sum_{ij \in A} d_{ij} x''_{ij}}_{\overset{i)}{\le} 6 \sum_{j \in J} D_j \,=\, 6\sum_{j \in J}\sum_{i \in I} \bar{x}_{ij} d_{ij} \,\le\, 6 \sum_{ij \in A} \bar{x}_{ij} d_{ij}} \le 6 \sum_{i \in I} f_i \bar{y}_i + \sum_{ij \in A} d_{ij} \bar{x}_{ij}.$$

3.2.7 Thm (Shmoys, Tardos, Aardal [1997]): Alg. 3.2.3 is 6-approximate.

3.2.8 Lemma (generalization of Claim 3.2.4): For $\beta > 0$ let

$$N_j(\beta) := \{ i \in I : \bar{x}_{ij} > 0 \wedge d_{ij} \le \beta D_j \}.$$

Then $\sum_{i \in N_j(\beta)} \bar{x}_{ij} \ge 1 - \frac{1}{\beta} = \frac{\beta - 1}{\beta}$.

Proof: Let

$$F_j(\beta) := \{ i \in I : \bar{x}_{ij} > 0 \wedge d_{ij} > \beta D_j \}$$

$$s_j := \sum_{i \in I} \bar{x}_{ij} \ge 1$$

Suppose $\sum_{i \in F_j(\beta)} \bar{x}_{ij} > \frac{s_j}{\beta}$. Then

$$D_j = \sum_{i \in I} d_{ij} \bar{x}_{ij} = \sum_{i \in F_j(\beta)} d_{ij} \bar{x}_{ij} + \sum_{i \in N_j(\beta)} d_{ij} \bar{x}_{ij}$$
$$\ge \sum_{i \in F_j(\beta)} d_{ij} \bar{x}_{ij}$$
$$\ge \beta D_j \sum_{i \in F_j(\beta)} \bar{x}_{ij}$$
$$> \cancel{\beta} D_j \frac{s_j}{\cancel{\beta}}$$
$$\ge D_j \quad ↯$$

$$\Rightarrow s_j = \underbrace{\sum_{i \in F_j(\beta)} \bar{x}_{ij}}_{\le \frac{s_j}{\beta}} + \sum_{i \in N_j(\beta)} \bar{x}_{ij} \quad \Rightarrow \sum \bar{x}_{ij}$$

$$\Rightarrow \sum_{i \in N_j(\beta)} \bar{x}_{ij} \ge s_j - \frac{s_j}{\beta} = s_j \left(1 - \frac{1}{\beta}\right) \ge 1 - \frac{1}{\beta}. \quad \square$$

25.06.12

(17)

Metric Uncapacitated Facility Location

I: set of facilities

J: set of clients

f_i: opening cost of facility i

d_{ij}: cost of assigning client j to facility i

Goal: Open a subset of facilities and assign each client to an open facility while incurring in a minimum cost

Primal-Dual Approximation Algorithm (Jain & Vazirani, 2001)

(LP): $\min \sum_{i\in I} f_i y_i + \sum_{i\in I}\sum_{j\in J} d_{ij}x_{ij}$

$$\text{s.t.} \quad \sum_{i\in F} x_{ij} \geq 1 \quad \forall j\in J \longrightarrow \alpha_j$$
$$x_{ij} \leq y_i \quad \forall i\in I, j\in J \longrightarrow \beta_{ij}$$
$$x_{ij}, y_i \geq 0 \quad \forall i\in I, j\in J$$

(D): $\max \sum_{j\in J} \alpha_j$

$$\text{s.t.} \quad \sum_{j\in J} \beta_{ij} \leq f_i \quad \forall i\in I \longrightarrow y_i$$
$$\alpha_j - \beta_{ij} \leq d_{ij} \quad \forall i\in I, j\in J \rightarrow x_{ij}$$
$$\alpha_j, \beta_{ij} \geq 0 \quad \forall i\in I, j\in J$$

Interpretation of dual:
- α_j: Total cost paid by j to get assigned to an open facility.
- β_{ij}: Cost paid by j to open facility i.

Why? Suppose (x,y) is integral and optimal for (LP). Let (α,β) be a corresponding dual solution.

By complementary slackness:

- $y_i > 0 \Rightarrow \sum_{j\in J} \beta_{ij} = f_i$ (To open i, f_i needs to be paid ~~for~~ by all clients)
- $x_{ij} > 0 \Rightarrow \alpha_j = d_{ij} + \beta_{ij}$ (If j is connected to i, j pays the assignment cost and its share for opening i)
- $\alpha_j > 0 \Rightarrow \sum_{i\in F} x_{ij} = 1$ (Not interesting)
- $\beta_{ij} > 0 \Rightarrow x_{ij} = y_i$ (Client j only pays its share of f_i if it uses facility i)

Algorithm: Phase 1

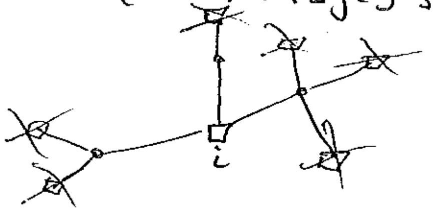
- $x, y, \alpha, \beta := 0$, $T := \{i \in I \mid \sum_{j \in J} \beta_{ij} = f_i\}$ (temporarily open facilities)

Definitions:
- j neighbors i if $\alpha_j \geq d_{ij}$
- j contributes to i if $\beta_{ij} > 0$

- while $\exists$ a client j not neighboring any facility in T.
  Raise α_j uniformly for all such j.
  - if j becomes neighbor of some $i \notin T \Rightarrow$ Increase β_{ij} at the same rate to maintain feasibility.
  - if j becomes neighbor of some $i \in T \Rightarrow$ Freeze variable α_j.

Phase 2 (Cleanup)

- while $\exists$ a client contributing to at least two facilities in T:
  - Pick an arbitrary such i
  - $T := T - \{h \in T \setminus \{i\} \mid \exists j \in J$ s.t. $\beta_{ij} > 0$ and $\beta_{hj} > 0\}$



Phase 3

- Set $y_i = 1 \quad \forall i \in T$

For all $j \in J$:
- Case 1: j contributes to a unique $i \in T \Rightarrow x_{ij} = 1$
- Case 2: j contributes to no $i \in T$ ~~but~~ but neighbors some $i \in T$
  $\Rightarrow x_{ij} := 1$ for arbitrary such i.
- Case 3: j has no neighbors in $T \Rightarrow$ Let i be closest facility in T, $x_{ij} := 1$.

Analysis:

Lemma: If at the end of Phase 2 there is a client j with no neighbors in T, there exists $i \in T$ s.t. $d_{ij} \leq 3\alpha_j$.

Proof: Later

Theorem: The algorithm gives a 3-approximation.

Proof: For every i, let $A(i) = \{$neighbours of i s.t. $x_{ij} = 1\}$

$$\sum_{i \in T} \Big(f_i + \sum_{j \in A(i)} d_{ij}\Big) = \sum_{i \in T} \sum_{j \in A(i)} (\beta_{ij} + d_{ij}) = \sum_{i \in T} \sum_{j \in A(i)} \alpha_j$$

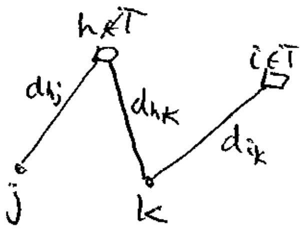
Let $Z := J \setminus \bigcup_{i \in T} A(i)$ (clients with no neighbors in T)

$$\sum_{i \in T} \sum_{j \in Z} d_{ij} x_{ij} = \sum_{j \in Z} \sum_{i \in T} d_{ij} x_{ij} \underset{\text{Lemma}}{\leq} \sum_{j \in Z} 3\alpha_j.$$

$$\Rightarrow \sum_{i \in T} \Big(f_i + \sum_{j \in J} d_{ij} x_{ij}\Big) = \sum_{i \in T} \sum_{j \in A(i)} \alpha_j + \sum_{j \in Z} 3\alpha_j \leq 3 \sum_{j \in J} \alpha_j \leq 3\, OPT_D \leq 3\, OPT_{LP}. \quad \square$$

Proof of Lemma:

Claim 1: Given such j, there exists h, k, i s.t.:



Why?
- We stopped increasing α_j when j neighboured some $h \in T$.
- h was removed from T in Phase 2 because k was contributing both to h and to i (which remains in T).

Claim 2: $d_{ij} \leq 3\alpha_j$

Pf: By triangle inequality.

1. $d_{hj} \leq \alpha_j$ because h and j are neighbours.
2. k contributes to h $\Rightarrow \beta_{hk} > 0 \Rightarrow \alpha_k > d_{hk}$.
   When α_j stops growing, h is in T. k contributed to h $\Rightarrow$ At that time, k already neighbours h
   $\Rightarrow \alpha_k$ doesn't ~~keep~~ growing after α_j stops $\Rightarrow \alpha_j \geq \alpha_k \Rightarrow \alpha_j \geq d_{hk}$.
3. $\beta_{ik} > 0 \Rightarrow \alpha_k \geq d_{ik} \Rightarrow \alpha_j \geq d_{ik}$.

So, $d_{ij} \leq d_{hj} + d_{hk} + d_{ik} \leq 3\alpha_j$. $\square$

Best approximation known: 1.52-approx. by Mahdian, Ye, Zhang.

Hardness: No approximation better than 1.46 possible (unless $P = NP$)

References:
- Jain, Vazirani: Approx. algorithms for metric facility location and k-median problems (2001)
- Shmoys, Williamson: The Design of Approximation Algorithms (2010)
- Vazirani: Approximation Algorithms (2001)

28.06.12

3.2.9 Alg. (Modified greedy Algorithm for the metric UFL):

Input: $I, J \subseteq \mathbb{R}^k$, $d_{ij} = \|i,j\|$ $\forall ij \in I \times J$, $f_i \in \mathbb{R}_{\geq 0}$ $\forall i \in I$

Output: $I^* \subseteq I$ opened facilities

1. $I^* \leftarrow \emptyset$
2. while $J \neq \emptyset$ do

2a $\quad (\underset{\in I}{i}, \underset{\subseteq J}{A}) \leftarrow \text{argmin} \frac{f_i + \sum_{j \in A} d_{ij}}{|A|}$

2b $\quad I^* \leftarrow I^* \cup \{i\}$

2c $\quad f_i \leftarrow 0$ (reuse open facility for free)

2d $\quad J \leftarrow J \setminus A$

3. Output I^*.

3.2.10 Rem.: The difference between the standard and the modified greedy is step 2c, which allows an open facility to be reused at 0 cost.

3.2.11 Rem. (Primal and Dual LP for UFL):

$$(P)\ \min \sum_{i \in I} f_i y_i + \sum_{ij \in I \times J} d_{ij} x_{ij} \quad \overset{(D)}{=} \quad \max \sum_{j \in J} \alpha_j$$

$$\sum_{i \in I} x_{ij} \geq 1 \quad \forall j \quad (\alpha_j) \qquad \sum_{j \in J} \beta_{ij} \leq f_i \quad \forall i$$

$$y_i - x_{ij} \geq 0 \quad \forall ij \quad (\beta_{ij}) \qquad \alpha_j - \beta_{ij} \leq d_{ij} \quad \forall ij$$

$$y_i \geq 0 \quad \forall i \qquad \alpha_j \geq 0$$

$$x_{ij} \geq 0 \quad \forall ij \qquad \beta_{ij} \geq 0$$

$\alpha_j \hat{=}$ transport contribution of client j

$\beta_{ij} \hat{=}$ contribution of client j to facility i

3.2.12 Alg. (Primal-Dual Alg. for the metric UFL):

Input: $I, J \subseteq \mathbb{R}^k$, $d_{ij} = \|i,j\|$ $\forall ij$, $f_i \in \mathbb{R}_{\geq 0}$ $\forall i$

Output: $I^* \subseteq I$, $A_i \subseteq J$, $i \in I \cup \{0\}$

Data Structures:

- $I^* \subseteq I$ opened facilities
- $J' \subseteq J$ unassigned clients (with active duals)
- $(\alpha, \beta) \geq 0$ infeasible solution for (D)
- $A_i \subseteq J$ clients assigned to facility $i \in I$ when i is opened
- A_0 " " " opened facility

1. $I^* \leftarrow \emptyset$, $J' \leftarrow J$, $\alpha \leftarrow 0$, $\beta \leftarrow 0$, $A_i \leftarrow \emptyset$ $\forall i \in I \cup \{0\}$
2. while $J' \neq \emptyset$ do

   uniformly raise

   α_j $\forall j \in J'$

   β_{ij} $\forall j \in J', i \in I \setminus I^*$: $\alpha_j \geq d_{ij}$

   until

   a) $\alpha_j = d_{ij}$ for $j \in J'$, $i \in I^*$; then

   $J' \leftarrow J' \setminus \{j\}$ (assign j to opened facility i)

   $\beta_{ij} \leftarrow 0$ $\forall i \in I$

   $A_0 \leftarrow A_0 \cup \{j\}$

   b) $\sum_{j \in J} \beta_{ij} = f_i$ for $i \in I \setminus I^*$; then

   $A_i \leftarrow \{j \in J' : \beta_{ij} > 0\}$ (open facility i and assign clients in A to i)

   $\beta_{i'j} \leftarrow 0$ $\forall i' \in I \setminus \{i\}, j \in A_i$

   $I^* \leftarrow I^* \cup \{i\}$

   $J' \leftarrow J' \setminus A_i$

3.2.13 Prop.: Alg. 3.2.12 computes $I^*, A_i, i \in I \cup \{0\}, \alpha, \beta$ such that

$$f(I^*) + \sum_{i \in I^*} d(i, A_i) = \alpha(J),$$

i.e., the objective values of the primal solution $(x,y) = (\chi_{I^*}, \chi_{A_i})$ for (P) and the dual (infeasible) solution (α, β) for (D) coincide.

Proof:

(1) $|\{i \in I : \beta_{ij} > 0\}| \leq 1$ $\forall j \in J$, i.e., $J = \dot{\bigcup} A_i$

Subst: $|\cdot| = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$, if j is assigned to i in $\begin{Bmatrix} 2a \\ 2b \end{Bmatrix}$.

(2) $f_i = \sum_{j \in A_i} \beta_{ij} = \sum_{j \in A_i} (\alpha_j - d_{ij})$ $\forall i \in I^*$

Subst: By construction.

(3) $\alpha_j = d_{ij}$ $\forall j \in A_0$

Subst: By construction.

Hence

$$\sum_{i \in I^*} f_i + \sum_{j \in J} d(I^*, j) \overset{(2)}{=} \sum_{i \in I^*} \sum_{j \in A_i} (\alpha_j - d_{ij}) + \sum_{j \in J} d(I^*, j)$$
$$= \sum_{i \in I^*} \sum_{j \in A_i} \alpha_j + \sum_{j \in A_0} d(I^*, j)$$

$$\overset{(5)}{=} \sum_{j \in \bigcup_{i \in \sigma} A_i} \alpha_j + \sum_{j \in A_0} \alpha_j$$

$$\overset{(4)}{=} \sum_{j \in J} \alpha_j . \quad \square$$

Prop. 3.2.14: $(\alpha/3, \beta)$ is feasible for (D).

Cor 3.2.15: The modified greedy is a 3-approx. alg. for metric UFL.

Proof (of Prop. 3.2.14):

(1) $\alpha_j \le d_{ij} + 2\alpha_{j'} \quad \forall i \in I, j, j' \in J : \alpha_j \ge d_{ij}, \ \overbrace{\alpha_{j'} \ge d_{ij'}}^{(i)}$.

If $\alpha_{j'} \ge \alpha_j$ ✓. So let $\alpha_{j'} < \alpha_j$.

$\alpha_j \hat{=}$ time at which j is assigned to some facility

$\alpha_{j'} < \alpha_j \Rightarrow j' \to i', \ j \to i'$ later $\Rightarrow \overbrace{\alpha_j \le d_{i'j}}^{(ii)}$

$\Rightarrow d_{i'j} \overset{(iii)}{\le} \underbrace{d_{i'j'}}_{\le \alpha_{j'}} + \underbrace{d_{ij'}}_{\overset{(i)}{\le} \alpha_{j'}} + d_{ij} \le 2\alpha_{j'} + d_{ij}$

$\alpha_{j'} + \beta_{i'j'} = d_{i'j'}$

$\Rightarrow \alpha_j \overset{(ii)}{\le} d_{i'j} \overset{(iii)}{\le} 2\alpha_{j'} + d_{ij}$.

(2) $f_i \ge \sum_{j \in J} \max\{0, \frac{\alpha_j}{3} - d_{ij}\} \quad \forall i \in I$

$C_i = \{ j \in J : \underbrace{\frac{\alpha_j}{3} \ge d_{ij}}_{\alpha_j \ge 3 d_{ij}, \text{ clients that pay a lot for fac } i} \} = \{1, \dots, k\}$ s.t. $\alpha_1 \le \dots \le \alpha_k$

$\Rightarrow \underbrace{\sum_{j \in J} \max\{0, \frac{\alpha_j}{3} - d_{ij}\} = \sum_{j \in C_i} \left(\frac{\alpha_j}{3} - d_{ij}\right)}_{(iv)} = \sum_{j=1}^{k} \left(\frac{\alpha_j}{3} - d_{ij}\right)$

Claim: $\sum_{j=1}^{k} \max\{0, \alpha_1 - d_{ij}\} \le f_i$.

Subst.: At time α_1, $\beta_{ij}(\alpha_1) = \max\{0, \alpha_1 - d_{ij}\}$.

$\alpha_j / 3 \ge d_{ij} \Rightarrow \alpha_j \ge d_{ij} \quad \forall j \in C_i$

$\overset{(1)}{\Rightarrow} \alpha_j \le d_{ij} + 2\alpha_1$

$\Rightarrow \sum_{j=1}^{k} \alpha_j \le \sum_{j=1}^{k} (d_{ij} + 2\alpha_1) = 2\sum_{j=1}^{k} (\alpha_1 - d_{ij}) + 3\sum_{j=1}^{k} d_{ij}$

$\le 2 f_i + 3 \sum_{j=1}^{k} d_{ij}$

$\Rightarrow f_i \ge \frac{2}{3} f_i \ge \sum_{j=1}^{k} \left(\frac{\alpha_j}{3} - d_{ij}\right) = \sum_{j \in J} \max\{0, \frac{\alpha_j}{3} - d_{ij}\}$

(3) $(\bar{\alpha}, \bar{\beta}) = \left(\frac{\alpha}{3}, \max\{\frac{\alpha_j}{3} - d_{ij}, 0\}\right)$ is feasible for (D)

$\sum_{j \in J} \bar{\beta}_{ij} = \sum_{j \in J} \max\{\frac{\alpha_j}{3} - d_{ij}, 0\} \overset{(2)}{\le} f_i$.

$$\overline{\alpha_j} - \overline{\beta_{ij}} = \frac{\alpha_j}{3} - \max\left\{\frac{\alpha_j}{3} - d_{ij}, 0\right\} \le d_{ij}$$

$\overline{\alpha}, \overline{\beta} \ge 0$. $\square$

3.2.16 Prop. Alg. 3.2.9 and 3.2.12 produce the same solution.

Proof: Consider some point in time t in Alg. 3.2.12. Then

$\alpha_j = t \quad \forall j \in J'$.

If 2a) doesn't hold $\alpha_j = t \le d_{ij} \quad \forall i \in I^*, j \in J'$.

If 2b) " $\sum_{j \in J'} \max\{0, t - d_{ij}\} \le f_i \quad \forall i \in I \setminus I^*$

$$\sum_{j \in S} (t' - d_{ij}) = f_i \Leftrightarrow t' = \frac{f_i + \sum_{j \in S} d_{ij}}{|S|}$$

$\therefore$ The next decision in Alg. 3.2.12 is at time

$$t' = \min\Bigg\{ \underbrace{\min_{i \in I^*, j \in J'} d_{ij}}_{= \min\limits_{\substack{i \in I^* \\ S \subseteq J'}} \frac{0 + \sum_{j \in S} d_{ij}}{|S|}}, \min_{\substack{i \in I \setminus I^* \\ S \subseteq J'}} \frac{f_i + \sum_{j \in S} d_{ij}}{|S|} \Bigg\}$$

This is exactly the criterion used by the greedy algorithm 3.2.9. $\square$

Thm. 3.2.17: The modified greedy algorithm 3.2.9 for the metric UFL is 3-approximative.

05.07.12

(01)

3.3 Local Search

3.3.1 Def. (Neighborhood, swap for the metric p-median problem): given $V = J \subseteq \mathbb{R}^k$, consider the metric p-median problem $p/I/\cdot/\|\cdot\|/\Sigma$:

a) $\mathcal{S} = \{S \subseteq I : |S| = p\}$ set of p-medians (feasible solutions)

b) $c(S) := \sum_{j \in J} \|S, j\| = \sum_{j \in J} \min_{i \in S} \|i, j\|$ cost of a solution

c) $\Gamma(S) := \{S' \in \mathcal{S} : |S' \Delta S| = 2\}$ neighborhood of a solution

d) $i \to i' : \{S \in \mathcal{S} : i \in S, i' \notin S\} \to \mathcal{S}$, $S \mapsto S \setminus \{i\} \cup \{i'\}$, swap

e) $S \in \mathcal{S}$ local optimum $:\Leftrightarrow \{S' \in \Gamma(S) : c(S') < c(S)\} = \emptyset$.

3.3.2 Alg. (Local Search for the metric p-median problem):

Input: $p/I/\cdot/\|\cdot\|/\Sigma$, $V = J$, $S \in \mathcal{S}$

Output: $\bar{S} \in \mathcal{S}$ local optimum

1. while $\exists S' \in \Gamma(S) : c(S') < c(S)$

   $S \leftarrow S'$

2. output S.

3.3.2 Rem. $p/I/\cdot/\|\cdot\|/\Sigma \in \mathcal{NP}$ (MMIL), $\sum_{i \in I} y_i = p$, $f \equiv$ const.

3.3.4 Goal: Let $\bar{S}$ be a local, S^* a global optimum. We want to know:

a) How large can $c(\bar{S}) / c(S^*)$ be?

b) How fast does Alg. 3.3.2 converge?

Arya, Garg, Khandekar, Meyerson, Munagala & Pandit [2004] showed

a) $c(\bar{S}) \leq 5\, c(S^*)$

b) $\forall S \in \mathcal{S} : \exists S' \in \Gamma(S) : \underbrace{c(S) - c(S')}_{\text{decrease per swap}} \geq \frac{c(S) - 5c(S^*)}{n^2}$

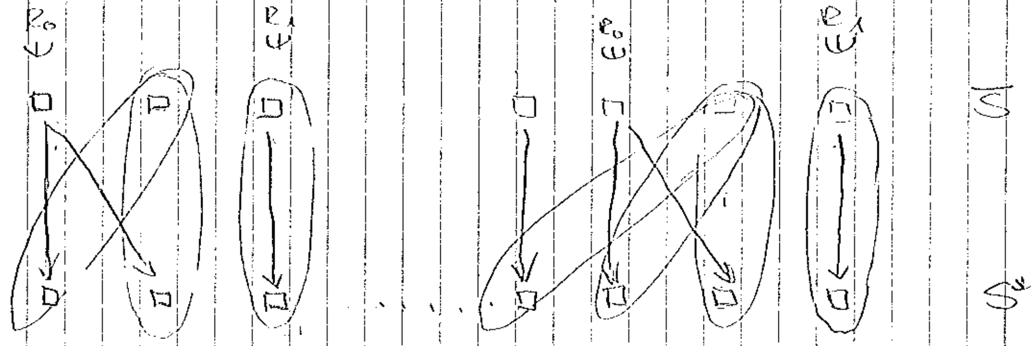
and that this yields an $5 + O(1)$-approximate polynomial time local search algorithm.

3.3.5 Thm (5-approximation ratio for local search): Alg. 3.3.3 yields a local optimum $\bar{S}$ such that $c(\bar{S}) \leq 5c(S^*)$.

Proof: Let $\bar{S}$ be the output of Alg. 3.3.3, S^* an optimal solution. Define a map

$$\eta : S^* \to \bar{S}, \quad i^* \mapsto \operatorname{argmin}_{i \in \bar{S}} \|i, i^*\|,$$

that maps each facility in S^* to the closest facility in $\bar{S}$, b.l.o.g. let



$R_k := \{ i \in \bar S : \eta(i^*) = i$ for exactly k facilities $i^* \in S^* \}$, $k = 0, 1$.

Construct a set P of k swaps, <u>one for each $i^* \in S^*$</u>, as follows:

a) $i \in R_1 \Rightarrow i \to \eta^{-1}(i) \in P$

b) $|R_0| + \underbrace{|\bar S \setminus (R_0 \cup R_1)|}_{\le \frac{1}{2}|S^* \setminus \eta^{-1}(R_1)|} = |S^* \setminus \eta^{-1}(R_1)| \ge \frac{1}{2}|S^* \setminus \eta^{-1}(R_1)|$

$\Rightarrow 2|R_0| \ge |S^* \setminus \eta^{-1}(R_1)|$

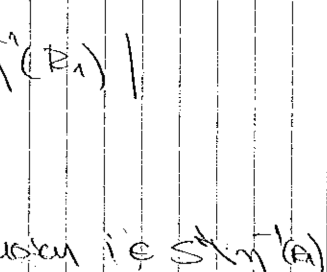
$i \in R_1 \Rightarrow i \to i^* \in P$ for at most 2 arbitrarily chosen $i \in S^* \setminus \eta^{-1}(R_1)$

<u>Rem.</u>: a) $i \in R_1$ is close to $\eta^{-1}(i)$, all other $i^* \in S^*$ are far away

$\Rightarrow$ $i \to \eta^{-1}(i)$ can be handled by assigning all of i's clients to $\eta^{-1}(i)$

b) $i \in \bar S \setminus (R_0 \cup R_1)$ is close to several facilities $\eta^{-1}(i)$

$\Rightarrow$ $i \to i^* \in \eta^{-1}(i)$ and assigning all of i's clients to i^* can cost much, so we consider only swaps $i \to i^* \in P$.



<u>Claim 1</u>: For $i \in \bar S$, $i^*, i^{*\prime} \in S^*$, and $i \to \overset{R_0, R_1}{i^*} \in P$ it holds $\eta(i^{*\prime}) \neq i$.

Let

$\bar\varphi : J \to \bar S$, $j \mapsto \arg\min \|\bar S, j\|$

$\varphi^* : J \to S^*$, $j \mapsto \arg\min \|S^*, j\|$

functions mapping j to the closest facility in $\bar S$ and S^*, respectively,

$\bar N(i) = \bar\varphi^{-1}(i)$, $i \in \bar S$

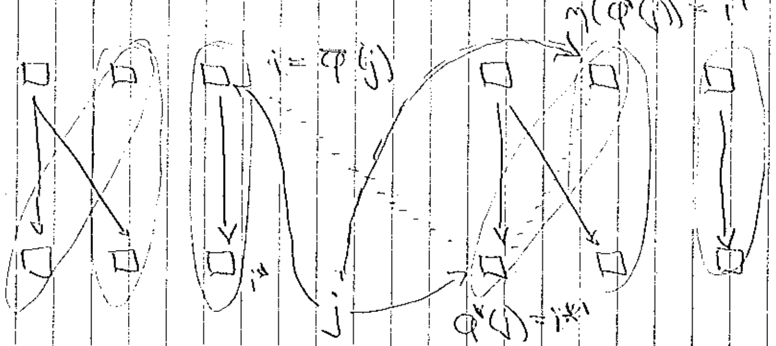
$N^*(i) = \varphi^{*-1}(i)$, $i \in S^*$

be the sets of clients assigned to $i \in \bar S$ and $i \in S^*$, respectively.

<u>Claim 2</u>: For each swap $i \to i^* \in P$ it holds

$$0 \overset{(1)}{\le} c(\bar S \setminus \{i\} \cup \{i^*\}) - c(\bar S) \overset{(2)}{\le} \sum_{j \in N^*(i^*)} \left(\|S^*, j\| - \|\bar S, j\|\right) + \sum_{j \in \bar N(i)} 2\|S^*, j\|.$$

<u>Subst.</u>: (1) is because $\bar S$ is a local optimum.



Consider the following assignment $\varphi' : J \to \bar{S} \setminus \{i\} \cup \{i^*\}$

$$\varphi'(j) = \begin{cases} i^*, & j \in N^*(i^*) \\ i' = \eta(\varphi^*(j)), & j \in \bar{N}(i) \setminus N^*(i^*) \\ \bar{\varphi}(j), & j \notin N^*(i^*), \bar{N}(i) \end{cases}$$

Because of claim 1, $i \neq i'$, so φ' is well defined. Further

$$c(\bar{S} \setminus \{i\} \cup \{i^*\}) - c(\bar{S}) \leq \sum_{j \in J} \left(\| \bar{S} \setminus \{i\} \cup \{i^*\}, j \| - \| \bar{S}, j \| \right)$$

$$\leq \sum_{j \in N^*(i^*)} (\underbrace{\| i^*, j \|}_{= \| S^*, j \|} - \| \bar{S}, j \|) + \sum_{j \in \bar{N}(i) \setminus N^*(i^*)} \underbrace{\left(\| i', j \| - \| i, j \| \right)}$$

$$\leq \| i^{*\prime}, j \| + \underbrace{\underbrace{\| i^{*\prime}, i' \|}_{\leq \| i^{*\prime}, i \|} - \| i, j \|}_{\leq \| i^{*\prime}, j \|}$$

$$\leq 2 \| i^{*\prime}, j \|$$

$$= 2 \| S^*, j \|$$

$$= \sum_{j \in N^*(i^*)} \left(\| S^*, j \| - \| \bar{S}, j \| \right) + \sum_{j \in \bar{N}(i) \setminus N^*(i^*)} 2 \| S^*, j \|$$

$$\leq \qquad \text{''} \qquad + \sum_{j \in \bar{N}(i)} 2 \| S^*, j \|$$

Summing over all $i \to i^* \in P$ gives

$$0 \leq \sum_{i^* \in S^*} \sum_{j \in N^*(i^*)} \left(\| S^*, j \| - \| \bar{S}, j \| \right) + 2 \sum_{i \in \bar{S}} \sum_{j \in \bar{N}(i)} 2 \| S^*, j \|$$

appears at most 2 times

$$= 5\, c(S^*) - c(\bar{S})$$

$$\Rightarrow c(\bar{S}) \leq 5\, c(S^*).$$

09.07.12

14