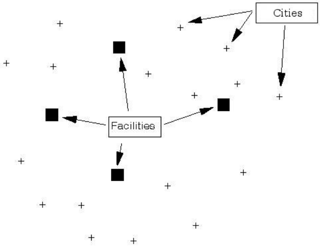
Approximation Algorithms

Facility Location

vissarion fisikopoulos

June 18, 2009

Problem Definition



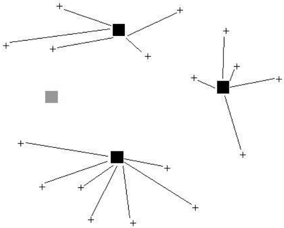
Metric uncapacitated facility location

- G: bipartite graph with bipartition (F, C)
- F: facilities, C: cities
- f_i: the cost of opening facility i
- c_{ij}: the cost of connecting city j to (opened) facility i
- the connection costs satisfy the triangle inequality

the problem is to find:

1. a subset $I \subseteq F$ of facilities that should be opened
2. a function $\phi : C \to I$ assigning cities to open facilities

s.t. minimize the total cost of opening facilities
and connecting cities to open facilities



The Integer Program

y_i: 1: facility i is open, 0: otherwise
x_{ij} 1: city j is connected to the facility i, 0:otherwise

minimize

$$\sum_{i\in F, j\in C} c_{ij}x_{ij} + \sum_{i\in F} f_i y_i$$

subject to

$$\sum_{i\in F} x_{ij} \geq 1 \qquad j \in C \tag{1}$$

$$y_i - x_{ij} \geq 0 \qquad i \in F, j \in C \tag{2}$$

$$x_{ij} \in \{0, 1\} \qquad i \in F, j \in C \tag{3}$$

$$y_i \in \{0, 1\} \qquad i \in F \tag{4}$$

- (1) each city is connected to at least one facility
- (2) facilities connected with cities must be open

LP-relaxation and Dual Program

minimize

$$\sum_{i\in F, j\in C} c_{ij}x_{ij} + \sum_{i\in F} f_i y_i$$

subject to

$$\sum_{i\in F} x_{ij} \geq 1 \qquad j \in C$$

$$y_i - x_{ij} \geq 0 \qquad i \in F, j \in C$$

$$x_{ij} \geq 0 \qquad i \in F, j \in C$$

$$y_i \geq 0 \qquad i \in F$$

maximize

$$\sum_{j\in C} a_j$$

subject to

$$a_j - \beta_{ij} \leq c_{ij} \qquad i \in F, j \in C$$

$$\sum_{j\in C} \beta_{ij} \leq f_i \qquad i \in F$$

$$a_j \geq 0 \qquad j \in C$$

$$\beta_{ij} \geq 0 \qquad i \in F, j \in C$$

Understanding the Dual

Slackness Conditions

optimal integral solution (I, ϕ) with conditions:

- $y_i = 1 \Leftrightarrow i \in I$
- $x_{ij} = 1 \Leftrightarrow i = \phi(j)$

primal slackness conditions:
i. $\forall i \in F, j \in C : x_{ij} > 0 \Rightarrow a_j - \beta_{ij} = c_{ij}$ a: total price of cities
ii. $\forall i \in F : y_i > 0 \Rightarrow \sum_{j \in C} \beta_{ij} = f_i$ each opened facility must fully paid for

dual slackness conditions:
iii. $\forall j \in C : a_j > 0 \Rightarrow \sum_{i \in F} x_{ij} = 1$ each city connected to exactly 1 facility
iv. $\forall i \in F, j \in C : \beta_{ij} > 0 \Rightarrow y_i = x_{ij}$ a city contribute only to the connected facility

Note: from weak duality theorem proof must be: $x_{ij} = 0 \vee a_j - \beta_{ij} = c_{ij}$ etc...

Understanding the Dual

$a_j = \beta_{ij} + c_{ij}$: total price of city j

- c_{ij} goes towards the use of edge (i, j)
- β_{ij} is the contribution of j towards opening facility i (from ii if $i \in I$ then $\sum_{j:\phi(j)=i} \beta_{ij} = f_i$)

Relaxing primal complementary slackness conditions

i $\forall i \in F, j \in C : \frac{1}{3} c_{\phi(j)j} \le a_j - \beta_{ij} \le c_{\phi(j)j}$

ii $\forall i \in F : \frac{1}{3} f_i \le \sum_{j \in C} \beta_{ij} \le f_i$

. . . we can relax in a way that dual must pay completely for open facilities

Solution:

- partition the cities into directly connected ($\beta_{ij} > 0 \vee \beta_{ij} = 0$) and indirectly connected ($\beta_{ij} = 0$)
- relax only the indirectly connected cities i.e. $\frac{1}{3} c_{\phi(j)j} \le a_j \le c_{\phi(j)j}$ (becuase $\beta_{ij} = 0$)

```
% Initialization
time = 0
∀i,j (city_j = unconnected, facility_i = closed,
        a_j = 0,  b_ij = 0)
% Phase 1
while (∃j, city_j = unconnected) do ∀j :
    if city_j = unconnected then a_j = a_j + 1
    if a_j = c_ij then (i,j) is tight

    if (i,j) is tight && a_j − β_ij ≤ c_ij then β_ij = β_ij + 1
    if β_ij > 0 then (i,j) is special

    if ∑_{j∈C} β_ij = f_i then facility_i = temporarily_open
    if city_j = unconnected
        && facility_i= temporarily_open && (i,j) is tight
    then
        facility_i is the connected witness of city j
        city_i = connected;   β_ij = 0
    time = time + 1
```

```
% Phase 2
F_t = {temporarily open facilities}
T = (V',E'), E'={e ∈ E| e is special}
T^2 = (V'',E''), E''={e=(u,v) ∈ E'| |path(u,v)∈T|≤2}
H: subgraph of T^2 induced on F_t
find a maximal independent set I in H
each facility in I is declared open

for each city j:
F_j = {i ∈ F_t | (i,j) is special}
if (∃i, facility_i ∈ F_j) then
    φ(j) = i; city_j = directly_connected (β_ij > 0)
else let i' the connecting witness for j, (i',j) is tight
    if (i' ∈ I) then
        φ(j) = i'; city_j = directly_connected (β_ij = 0)
    else
        let i'' be a any neighbour of i' in H
        φ(j) = i''; city_j = indirectly_connected (β_ij = 0)
```

Analysis

Definitions

$a_j = a_j^f + a_j^e$, a_j^f: opening facilities, a_j^e: connecting cities to facilities

- directly connected: $a_j = c_{ij} + \beta_{ij}$, $a_j^f = \beta_{ij}$, $a_j^e = c_{ij}$
- indirectly connected: $a_j^f = 0$ and $a_j = a_j^e$

Lemma

Let $i \in I$. Then

$$\sum_{j:\phi(j)=i} a_j^f = f_i$$

Proof Sketch

- i temporarily open, i.e. $\sum_{j:(i,j)\ \texttt{is special}} \beta_{ij} = f_i$ (Phase 1)
- each city j that has contributed to f_i must be directly connected to i ($a_j^f = \beta_{ij}$)
- for any other connected city $a_{j'}^f = 0$

Corollary

$$\sum_{i \in I} f_i = \sum_{j \in C} a_j^f$$

Lemma

For an indirectly connected city j, $c_{ij} \leq 3a_j^e$ *, where* $i = \phi(j)$*.*

Theorem

$$\sum_{i \in F,\ j \in C} c_{ij} x_{ij} + 3 \sum_{i \in F} f_i y_i \leq 3 \sum_{j \in C} a_j$$

Proof Sketch

- for a directly connected city j, $c_{ij} = a_j^e \leq 3a_j^e$, where $\phi(j) = i$
- previous lemma
- corollary

Implementation

$n_c = |C|,\ n_f = |F|,\ n = n_c + n_f,\ m = n_c \times n_f$

$\forall$ facility i:

- number of cities currently contributing (init: 0)
- anticipated time t_i: completely paid for if no other event happens (init: infinite)

binary heap for t_i: minimum: $O(1)$, update: $O(\log n_f)$

Events:

- edge (i, j) goes tight
  - **i not temporarily open:** 1. numofcities(i)++; 2. recompute t_i and update heap;
  - **i already temporarily open:** 1. j connected;
    for each i' counting j: 2. numofcities(i')−−;
    3. recompute t_i and update heap;
- facility i completely paid for: 1. i temporarily open 2. all cities contributing to i connected 3. steps 2,3 from previous event 2^{nd} case

Running Time

Theorem

The algorithm achieves approximation factor 3 for the facility location problem and has running time $O(m \log n_f)$.

Tight Example

$f_1 = \epsilon,\ f_2 = (n+1)\epsilon$

- OPT = $(n+1)\epsilon + n$
- SOL = $\epsilon + 1 + 3(n-1)$