Average-case analysis of some approximation algorithms for the metric uncapacitated facility location problem.

Abraham Flaxman, CMU
Alan Frieze, CMU
Juan Vera, CMU
Algorithmic goals in combinatorial optimization:

(1) In polynomial time,
(2) for every problem instance
(3) find an optimal solution.
Two out of three = research area:

Faster exponential time algorithms

(1) In polynomial time,
(2) for every problem instance
(3) find an optimal solution.

Average-case analysis of algorithms

(1) In polynomial time,
(2) for “most” problem instance
(3) find an optimal solution.
The most popular two out of three Approximation algorithms

(1) In polynomial time, (2) for every problem instance (3) find a solution within a known factor of optimal
This talk: Two two-out-of-threes

We will take approximation algorithms and see how good the apx ratio is in the average-case analysis setting.
Facility location problem

Set $C$ of cities
Set $F$ of facilities
opening costs: $f_i$ for facility $i$
connection costs: $d_{ij}$ for city $j$ to be served by facility $i$

Goal: Find set of facilities to open so that total cost is minimized.
Approx. fac. loc. state-of-the-art

- Connection costs unrestricted: $O(\log n)$ [Hochbaum]

- Connection costs form a metric

  1.52-approx exists [Mahdian, Ye, Zhang]
  1.46-approx does not [Guha, Khuller]

(unless $\text{NP} \subseteq \text{PTIME}(n^{O(\log n)})$)
Brief description of an alg.

\[(1.86 \text{- apx;} \text{ similar to } \& \text{ simpler than } 1.52 \text{- apx, } [\text{Jain, Mahdian, Markakis, Saleri, Vazirani}])\]

\[\text{Init: Time } t=0. \text{ Funds } \delta_i=0 \text{ for each city.} \]

Set of unconnected cities \(\mathcal{U} = \mathcal{G} \)
Brief description of an alg.

\textbf{Init:} Time } t = 0. \text{ Funds } S_i = 0 \text{ for each city.}
\text{Set of unconnected cities } U = \mathcal{G}

\textbf{While } U \neq \emptyset : \hfill \text{Take appropriate action.}
\text{For every } i \in U, \text{ increase } S_i \text{ simultaneously}
\text{until: } \begin{align*}
1) \quad & S_i = d_{ij} \quad \text{for some open facility } j. \\
2) \quad & \sum_{i \in U} (S_i - d_{ij})^+ = f_j \quad \text{for some unopened facility } j.
\end{align*}
Random instances (random how?)

Not \( d_{i,j} \in [0,1] \) uniformly at random

Geometric instances:
Place \( n \) cities u.a.r. in \([0,1]^{2}\).
Keep it simple:
- each city is a candidate facility
- opening costs all same value, \( f \)
- connection costs all given by \( l_{\infty}\)-norm
Random instances

The opening cost of the solution structure.

If $f = 0$ then every city opens as a facility.

If $f > n$ then just one city opens as the most central.
Random instances

The opening cost $f$ controls solution structure.

If $f = 0$ then every city opens as a facility.

If $f > n$ then just one city opens as a facility—the one most central.
Random instances

The opening cost $f$ controls solution structure.

If $f = 0$ then every city opens as a facil.

If $f > n$ then just one city opens as a facility—the one most central.
Random instances

The opening cost $f$ controls solution structure.

If $f = 0$ then every city opens as a fac.

If $f > n$ then just one city opens as a facility—the one most central.
Random instances

If \( \frac{(\log n)^{3/2}}{n} \ll f \ll n \)
then \( \text{OPT} \sim \frac{1}{2} \alpha n \)
where \( \alpha = 3 \sqrt{\frac{6f}{n}} \).
Random instances

If \( \frac{(\log n)^{3/2}}{n} \ll f \ll n \)

then \( \text{OPT} \approx \frac{1}{2} \alpha n \)

where \( \alpha = \sqrt[3]{\frac{6f}{n}} \).

**Pf:** Partition square into \( \alpha \times \alpha \) cells. Open facility near center of each cell. (Upper bound)

(Lower bound) Construct dual soln which is feasible w.h.p.
Main Theorem If \( \frac{(\log n)^{3/2}}{n} \ll f \ll n \)

then w.h.p., \( \text{APX} \geq (1+\varepsilon) \text{OPT} \)

Pf: 1) Show that in any near-optimal soln, most of the Voronoi cells of open facilities must be close to square, close to centered, and have area approximately \( \frac{1}{x^2} \).

2) Show \( \text{APX} \) does not do this.
Proof of main theorem, part 1

A sort of isoperimetric inequality says the connection cost per facility is minimized when the Voronoi cell is square.

We quantize to a fine grid and prove a quantitative version of this.
Proof of main theorem, part 2

Each approx will open facilities at (dependent) random locations in the square.

Greedy can't fix Voronoi cells after this happens, which is the case often.
Summary

Whp, $\text{APX} > (1+\varepsilon)\text{OPT}$

Future work

- Tighter bounds on $\frac{\text{APX}}{\text{OPT}}$
  (finding it exactly is probably hard)

- See how other approx. algs do
  (for starters, analyze “Solve LP-Relax and round independently”)