

An $O(\log \log n)$ -Approximation for Submodular Facility Location

Fateme Abbasi¹, Marek Adamczyk¹, Miguel Bosch-Calvo²,
Jarosław Byrka¹, Fabrizio Grandoni², Krzysztof Sornat², Antoine Tinguely²

1 University of Wrocław, Poland

2 IDSIA, USI-SUPSI, Switzerland

Banf 20.09.2023

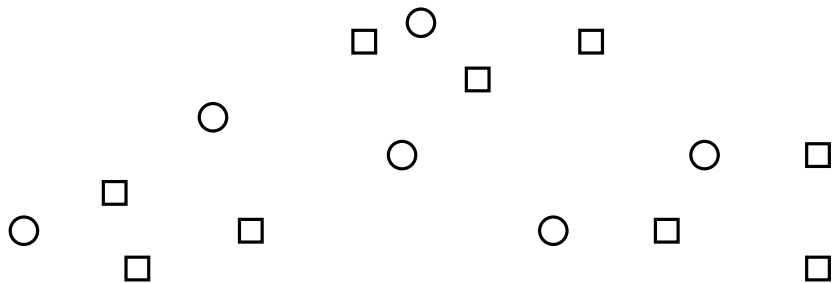
Facility Location

Given:

- | set of Clients C , set of Facilities F
- | opening facility cost o_f

Goal:

- | minimize $\sum_{c \in C} d(c, F(c)) + \sum_{f \in F} o_f$



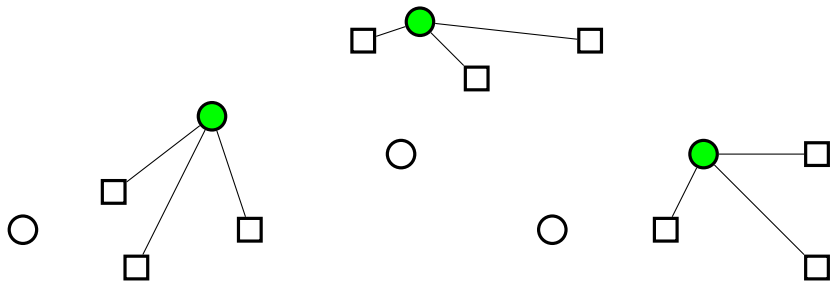
Facility Location

Given:

- | set of Clients C , set of Facilities F
- | opening facility cost o_f

Goal:

- | minimize $\sum_{c \in C} d(c, F(c)) + \sum_{f \in F} o_f$



Submodular Set Functions

$g : 2^V \rightarrow \mathbb{R}$ is **submodular** if:

$$A, B \subseteq V, \quad g(A) + g(B) \geq g(A \cup B) + g(A \cap B)$$

$$A \subseteq B \subseteq V, \quad x \in V \setminus B, \quad g(A \cup \{x\}) - g(A) \geq g(B \cup \{x\}) - g(B)$$

Submodular Facility Location

Given:

- | n clients C and m facilities F ,
- | $d : (C \times F) \times (C \times F) \rightarrow \mathbb{R}_0$
- | a monotone submodular opening cost $g(\cdot)$

Goal:

$$\text{Minimize} \sum_{c \in C} d(c, \pi(c)) + \sum_{f \in F} g(\pi^{-1}(f))$$

Where $\pi : C \rightarrow F$ is assignment of each client to some facility

Known Results

SFL is APX-hard [Guha, Khuller,1999]

Svitkina and Tardos show [2010] that:

- | There is $O(\log n)$ approximation for general SFL with multiple submodular function. The result is tight, because of reduction from Set Cover Problem.
- | $(4.237 + \epsilon)$ approximation for special case of SFL where submodular function $g(\cdot)$ is specified by a rooted cost tree T

Our Results

Main Contribution

- | There is a polynomial-time $O(\log \log n)$ -approximation algorithm for SFL.

Our Results

Main Contribution

- | There is a polynomial-time $O(\log \log n)$ -approximation algorithm for SFL.

Generalizations and Variants

- | There is a polynomial-time $O(\log \log n)$ -approximation algorithm for mul tSFL.
- | There is a polynomial-time $O(\log \log n)$ -approximation algorithm for addSFL.
- | There is a polynomial-time $O(\log \log \frac{n}{\min})$ -approximation algorithm for the Universal Stochastic Facility Location problem.

LP Relaxation

Conf-LP:

$$\begin{aligned}
 \min \quad & \sum_{f \in F, R \in C} g(R) \cdot x_R^f + \sum_{c \in C, f \in F, R \in C} d(c, f) \cdot x_R^f & (1) \\
 \text{s.t.} \quad & \sum_{f \in F, R \in C} x_R^f = 1 & c \in C; \\
 & \sum_{R \in C} x_R^f = 1 & f \in F; \\
 & x_R^f \geq 0 & R \in C, f \in F.
 \end{aligned}$$

Overview of the algorithm

1. Compute fractional solution to Conf-LP
2. Sample partial assignment S_1 to remove most of the clients from the remaining instance
3. Embed the remaining instance into a tree
4. Use filtering techniques to remove the connection cost from the picture and obtain Descendent-Leaf Assignment problem (DLA)
5. Approximately solve DLA via LP rounding

2) Reducing the connection cost by removing client

1. Let \hat{x} be a solution to Conf-LP
2. For $\ln \ln N$ times, sample random partial assignments by selecting configuration (f, R) indep. with probability \hat{x}_R^f
3. let C_1 be the covered clients and S_1 the partial assignment, then:

| Each client belongs to C_1 with **large enough** probability,

$$\mathbb{P}[c \in C_1] = e^{-\ln \ln N} = \frac{1}{\ln N}$$

| The expected cost of S_1 is **small enough**

$$\mathbb{E}[\text{cost}(S_1)] = \ln \ln N \cdot \text{cost}(\hat{x})$$

3) Embedding the remaining instance on an HST

Let \check{x} be \check{x} restricted to $C_2 = C \setminus C_1$ we have:

$$\begin{aligned} | \text{open}(\check{x}) &= \text{open}(\check{x}) \text{ and } \mathbb{E}[\text{conn}(\check{x})] &= \frac{1}{\ln N} \text{conn}(\check{x}). \end{aligned}$$

We map the input metric (M, d) into a metric on a Hierarchically well-Separated Tree (HST), (M, d_T) and obtain:

1. Every $a \in M$ is mapped to some leaf $v(a)$ of T
2. $\mathbb{E}[d_T(v(a), v(b))] \leq 8 \log |M| \cdot d(a, b)$;
3. T has depth $O(\log d_{\max})$.

Therefore,

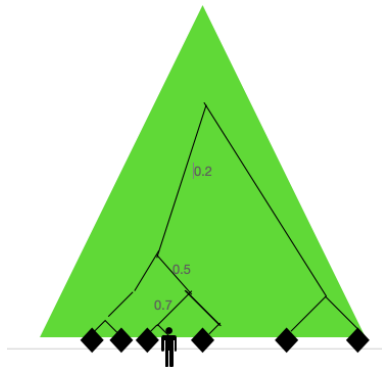
$$\mathbb{E}[\text{conn}_{d_t}(\check{x})] = O(1) \cdot \text{conn}(\check{x})$$

4) Filtering

Focus on a single client:

- | Think of fractional connection as of a unit flow.
- | Focus on the path from the client to the root of the tree.
- | Find the lowest edge on the path on which the flow has value < 0.5 .

Figure 1: A client fractionally connected to facilities.



4) Filtering and Descendent-Leaf Assignment (DLA)

DLA

Goal: Find an assignment of each $c \in \tilde{C}$ to some $f \in \tilde{F}_c$ so that the total opening cost is minimized.

Figure 2: A client restricted to a subtree.

Convex-programming (CP) relaxation for DLA:

$$\begin{aligned} \min \quad & \sum_{f \in \tilde{F}} \hat{h}(z^f) \\ \text{s.t.} \quad & \sum_{f \in \tilde{F}_c} z_c^f = 1 \quad c \in \tilde{C}. \\ & z_c^f \in \{0, 1\} \quad c \in \tilde{C}, f \in \tilde{F}. \end{aligned}$$

5) Solving DLA via LP rounding

We adapt method of [Bosman, Olver 2020]

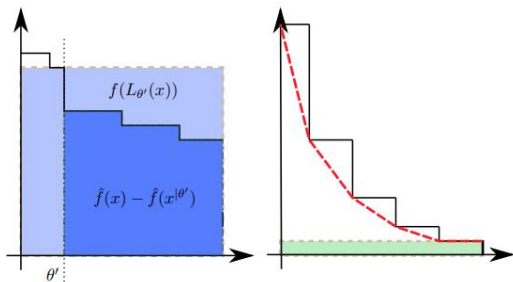
- | proceed by levels bottom-up.
- | feature of the relaxation:
in extreme solutions,
subsets of clients served
by a facility form a chain.
- | when
processing a node: select
a subset of clients (from
the chain) to be integrally
served via threshold.
- | merge nodes at the bottom.

Figure 3: A client restricted to a subtree.

Lemma (Bosman, Olver, 2020)

Given $x \in [0, 1]^V$ and $\theta \in (0, 1]$, at least one of the following holds:

1. there exists $\theta' \in [0, 1]$, which can be computed in polynomial time, such that $L_\theta(x)$ is $\frac{\theta'}{32}$ -supported;
2. $2^{1/\theta'} f(L_\theta(x)) \leq \hat{f}(x)$.



Summary: $(\log \log N)$ approximation for SFL

- | Compute a random partial assignment S_1 ,
 $\mathbb{E}(\text{cost}(S_1)) = O(\log \log N) \cdot \text{cost}(\dot{x})$
- | Obtain residual fractional solution \ddot{x} restricted to $S_2 = C \setminus S_1$ and embed it on the HST-type instance, we have:

$$\begin{aligned}\mathbb{E}[\text{cost}_T(\ddot{x})] &= \text{open}(\ddot{x}) + \mathbb{E}[\text{conn}_T(\ddot{x})] \\ &= \text{open}(\dot{x}) + O(\log N) \cdot \mathbb{E}[\text{conn}(\ddot{x})] \\ &= O(\text{cost}(\dot{x})).\end{aligned}$$

- | Randomly round \ddot{x} to an assignment of S_2 via a red. to DLA
- | Obtain S_2 of cost at most $O(\log \log N) \text{cost}(\dot{x})$
- | Return $S_1 + S_2$ as a feasible solution to SFL so that:

$$\begin{aligned}\mathbb{E}(S_1 + S_2) &= O(\log \log N) \cdot \text{cost}(\dot{x}) \\ &= O(\log \log N) \cdot \text{cost}(opt)\end{aligned}$$

Open Questions

- | Is there any constant approximating for SFL Problem?
- | Is there any constant factor approximation over tree instance for SFL problem?
- | Is there $O(\log \log N)$ approximation for AFFINE SFL problem over tree instance, where the opening costs are submodular functions of form $g_f(S^f) = o_f + w_f \cdot g(S^f)$?

IPCO'24 in Wrocław, Poland

ipco2024.ii.uni.wroc.pl

6 November 2023: submission deadline

1-2 July 2024: summer school

3-5 July 2024: IPCO conference

PC chair: Jens Vygen

Local chair: Jarek Byrka

Summer school speakers:

- Sophie Huiberts
- Neil Olver
- Vera Traub