Two simplified versions of Red-Blue Facility Location

Cristina G. Fernandes[1], **Rafael V. Pocai**[1]

[1]Instituto de Matemática e Estatística – Universidade de São Paulo
São Paulo – SP – Brazil

Abstract. *The Red-Blue Facility Location problem is a generalization of Facility Location where clients may have two types of demands and each open facility must provide for one of the types of demands. We present preliminary results for two special cases of this problem.*

1. Introduction

The Facility Location problem is a central topic in combinatorial optimization. It formalizes the problem of deciding where to place facilities on a map to better serve the clients, minimizing the cost of opening the facilities and the distance from clients to the closest open facility. The general form of the problem is as follows:

Problem 1. *Facility Location*
Instance: *a set F of facilities and a set C of clients; a function $d : C \times F \to \mathbb{Q}^+$; for every $f \in F$, an opening cost $k_f \in \mathbb{Q}^+$ and, for every $c \in C$, a demand $k_c \in \mathbb{Q}^+$.*
Solution: *a set $O \subseteq F$ that minimizes $\sum_{f \in O} k_f + \sum_{c \in C} k_c \min_{f \in O} d(c, f)$.*

The function d, representing the distances, is not necessarily metric. An $\mathrm{O}(\lg n)$-approximation is known for this general version, where $n = |C|$ [Hochbaum 1982]. For the metric case, in which d satisfies the triangle inequality, the best known approximation ratio for the problem is 1.488 [Li 2013], and there is an inapproximability result that says that, if there is an approximation with ratio better than 1.463, then $\mathrm{NP} \subseteq \mathrm{DTime}(n^{\lg \lg n})$ [Guha and Khuller 1999].

A more general version of the problem considers that clients have different types of demands, and each open facility serves a specific type of demand. In this more general setup, one has to decide not only where to open facilities, but also what type of demand each open facility will serve [Baev et al. 2008]. The case that allows for two types of demands was called in the literature the Red-Blue Facility Location (RBFL). The k-medians problem with more than one type of demand has also been addressed [Hajiaghayi et al. 2010, Swamy 2016].

There is a 4-approximation algorithm for the Metric RBFL under the assumption that there is no opening cost for the facilities [Arora et al. 2014]. This algorithm uses linear programming and rounding. While searching for a better approximation for this version of the problem, we managed to design a simple 2-approximation for the particular case in which the blue and red demands are the same for each client. The problem remains NP-hard in this case, as we show at the end of Section 2, after presenting the 2-approximation. In Section 3, we prove that the variant of the Metric RBFL in which each client has either blue demand or red demand is also NP-hard.

2. Red-Blue Facility Location with Equal Red-Blue Demands

We consider the Unitary Demands Metric RBFL without opening costs, stated as follows.

Problem 2. *Unitary-Demands Red-Blue Facility Location (UD-RBFL)*
Instance: *a set F of facilities, a set C of clients, and a metric $d : (C \cup F)^2 \to \mathbb{Q}^+$.*
Solution: *a partition $F = F_r \cup F_b$ that minimizes the sum $\sum_{c \in C} (d(c, F_r) + d(c, F_b))$.*

Next we describe a 2-approximation for UD-RBFL. The algorithm also works for instances of the Metric RBFL where the red and blue demands of each client are the same.

Let (F, C, d) be an instance of the UD-RBFL. If $|F| \leq 1$, any solution is optimal. So assume $|F| \geq 2$. Consider the directed graph D with set of vertices $F = \{f_1, \ldots, f_n\}$ and set of arcs $\{(f, f') : f, f' \in F$ and $f' \in \arg\min_{x \in F \setminus \{f\}} d(f, x)$ with smallest index$\}$. Since the outdegree of every vertex in D is one, we may define a function $F_D : F \to F$ that takes each facility to the head of the arc leaving it in D. We use the digraph D and the function F_D to design an approximation for UD-RBFL.

If vertices $v_1, v_2, \ldots, v_k \in F$ form a directed cycle in D, then, by the definition of D, $d(v_1, v_2) \leq d(v_k, v_1) \leq d(v_{k-1}, v_k) \leq \cdots \leq d(v_2, v_3) \leq d(v_1, v_2)$, and hence $d(v_i, F_D(v_i)) = d(v_1, v_2)$ for every i in $1, 2, \ldots, k$. Because the arcs leaving the vertices $v_1, v_2, \ldots, v_k$ will land on the vertex with minimum distance and smallest index, there is no cycle longer than two in D. Indeed, no vertex in a cycle would point to the one with the largest index in the cycle unless it was its only option.

Because the outdegree of D is one, there is exactly one (length-two) cycle in each connected component. Therefore, the digraph is bipartite. Algorithm 1 uses the bipartition to decide which facilities are red and which are blue. In it, we denote by DIGRAPH the procedure that constructs the digraph D from F and d.

Algorithm 1: UD-RBFL-2approx(F, C, d)

```
F_r ← ∅
F_b ← ∅
D ← DIGRAPH(F, d)
foreach {f, f'} ⊆ F that forms a cycle in D do
    F_r ← F_r ∪ {f}
    F_b ← F_r ∪ {f'}
while F ≠ F_r ∪ F_b do
    Let f ∈ F \ (F_r ∪ F_b) be such that F_D(f) ∈ F_r ∪ F_b
    if F_D(f) ∈ F_r then
        F_b ← F_b ∪ {f}
    else
        F_r ← F_r ∪ {f}
return F_r, F_b
```

Theorem 1. *Algorithm 1 is a 2-approximation for UD-RBFL.*

Proof. Let (F, C, d) be an instance of UD-RBFL with $|F| \geq 2$. For a client $c \in C$, we denote by $\text{cost}(c)$ the cost for client c in the solution given by the algorithm, that is, $d(c, F_r) + d(c, F_b)$. Let $f_1, f_2 \in F$ be the first and second closest facilities to c according to d and let $\text{optcost}(c) = d(c, f_1) + d(c, f_2)$ be a lower bound on the cost that c must pay to access a red and a blue facility in an arbitrary solution. This lower bound is achieved when one among f_1 and f_2 is red and the other is blue in the considered solution. In particular, $\text{cost}(c) = \text{optcost}(c)$ if Algorithm 1 assigns different colors to f_1 and f_2.

If Algorithm 1 makes both f_1 and f_2 red (blue), then, by construction of the coloring, $f = F_D(f_1)$ is blue (red), and the cost for c is

$$\begin{aligned}
\text{cost}(c) &\leq d(c, f_1) + d(c, f) \\
&\leq d(c, f_1) + d(c, f_1) + d(f_1, f) \\
&\leq d(c, f_1) + d(c, f_1) + d(f_1, f_2) \\
&\leq d(c, f_1) + d(c, f_1) + d(c, f_1) + d(c, f_2) \\
&\leq d(c, f_1) + d(c, f_2) + d(c, f_1) + d(c, f_2) \\
&\leq 2\, \text{optcost}(c).
\end{aligned}$$

Since the cost of the solution produced by Algorithm 1 is the sum of the costs for all clients, the solution found is a 2-approximation. □

We can apply Algorithm 1 also to instances in which the ratio between the blue and red demands of each client lies in $[1/c, c]$ for a constant $c \geq 1$. The approximation ratio achieved is $1 + c$, which is better than 4 if $c < 3$ and approaches 2 as c goes to 1.

Now we prove that UD-RBFL is NP-hard by presenting a reduction from the following classical decision problem:

Problem 3. *Splitting Set*
Instance: *a set S and a collection $\mathcal{C}$ of subsets of S.*
Question: *is there a partition $S = S_r \cup S_b$ such that $C \cap S_r \neq \emptyset$ and $C \cap S_b \neq \emptyset$ for every $C \in \mathcal{C}$?*

Theorem 2. *UD-RBFL is NP-hard.*

Proof. Let $(S, \mathcal{C})$ be an instance of Splitting Set. We construct an instance of UD-RBFL where $F = S$ are the facilities, $C = \mathcal{C}$ are the clients, and the distance between a client c and a facility f is defined as:

$$d(c, f) = \begin{cases} 1 & \text{if } f \in c \\ 3 & \text{otherwise.} \end{cases}$$

The construction of d clearly takes polynomial time. Besides, d may be extended to the shortest path metric induced on $F \cup C$. It remains to show the following.

Claim 1. *The instance $(S, \mathcal{C})$ of Splitting Set has a positive answer if and only if there is a solution for (F, C, d) with cost at most $2\,|C|$.*

Proof of the claim. Assume there is a solution $S = S_r \cup S_b$ for the instance of Splitting Set. Then, for every client $c \in C$, there are two facilities $f_1 \in S_r$ and $f_2 \in S_b$ with distance one to c, and $F = S = S_r \cup S_b$ is a solution for (F, C, d) with cost $2\,|C|$. Now assume there is a solution $F = S = S_r \cup S_b$ for (F, C, d) with cost at most $2\,|C|$. Then there must be facilities $f_1 \in S_r$ and $f_2 \in S_b$ such that $d(c, f_1) = 1$ and $d(c, f_2) = 1$ for every $c \in C$. This is the same as saying that $c \cap S_r \neq \emptyset$ and $c \cap S_b \neq \emptyset$, hence there is a solution to problem Splitting Set. □

3. Red-Blue Facility Location with Exclusive Demands

On the other extreme, we consider the case of RBFL in which each client has either red demand or blue demand, and we show that this case is also NP-hard. We keep studying the unitary version, since the weighted one, in this case, is equivalent to the general RBFL.

Problem 4. *Unitary-Demand Exclusive Red-Blue Facility Location (UDX-RBFL)*
Instance: *set F of facilities, sets R and B of clients, and metric $d : (F \cup R \cup B)^2 \to \mathbb{Q}^+$.*
Solution: *partition $F = F_r \cup F_b$ that minimizes the sum $\sum_{c \in R} d(c, F_r) + \sum_{c \in B} d(c, F_b)$.*

Theorem 3. *UDX-RBFL is NP-hard.*

Proof. The reduction is from the following classical NP-complete problem.

Problem 5. *Exact Cover by 3-Sets (X3C)*
Instance: *a 3-regular bipartite graph $G = ((X, Y), E)$ such that $|X| = |Y| = 3k$.*
Question: *is there a subset of A with k vertices that dominates B?*

For an instance $((X, Y), E)$ of X3C, consider the following instance of UDX-RBFL. The set $F = X \cup X'$, where X' is a copy of X. The set $B = Y$ and the set $R = X''$, where X'' is another copy of X. For every facility $x \in X$ and blue client $y \in Y$ with $\{x, y\} \in E$, set $\ell(x, y) = 1$. For every red client $x'' \in X''$, set $\ell(x'', x) = 1$ and $\ell(x'', x') = 3$, where x and x' are the copies of x'' in X and X' respectively. Take d as the shortest path metric induced by ℓ on $F \cup B \cup R$. The construction of this instance is polynomial, so Theorem 3 follows from the next claim, whose proof we omit in this abstract.

Claim 2. *Instance $((A, B), E)$ of X3C has a positive answer if and only if there is a solution of cost at most $8k$ for the (F, B, R, d) instance of UDX-RBFL.*

□

References

Arora, S., Gupta, N., Khuller, S., Sabharwal, Y., and Singhal, S. (2014). Facility location with red-blue demands. *Oper. Res. Lett.*, 42(6):462–465.

Baev, I., Rajaraman, R., and Swamy, C. (2008). Approximation algorithms for data placement problems. *SIAM J. Comput.*, 38(4):1411–1429.

Guha, S. and Khuller, S. (1999). Greedy strikes back: Improved facility location algorithms. *Journal of Algorithms*, 31(1):228 – 248.

Hajiaghayi, M., Khandekar, R., and Kortsarz, G. (2010). Budgeted red-blue median and its generalizations. In *Proceedings of the 18th Annual European Conference on Algorithms: Part I*, ESA'10, pages 314–325, Berlin, Heidelberg. Springer-Verlag.

Hochbaum, D. S. (1982). Heuristics for the fixed cost median problem. *Math. Program.*, 22(1):148–162.

Li, S. (2013). A 1.488 approximation algorithm for the uncapacitated facility location problem. *Information and Computation*, 222:45–58.

Swamy, C. (2016). Improved approximation algorithms for matroid and knapsack median problems and applications. *ACM Transactions on Algorithms*, 12(4). Article 49.