

# Facility Location Problem with Different Type of Clients

Lisheng Wang<sup>1</sup>, Rongheng Li<sup>1\*</sup>, Jingui Huang<sup>2</sup>

<sup>1</sup>Department of Mathematics, Hunan Normal University, Changsha, China

<sup>2</sup>Department of Computer Education, Hunan Normal University, Changsha, China

E-mail: ldwls@126.com, {lirongheng, hjg}@hunnu.edu.cn

Received December 25, 2010; revised January 5, 2011; accepted March 28, 2011

## Abstract

This paper proposes a new model of facility location problem referred to as  $k$ -product uncapacitated facility location problem with multi-type clients. The  $k$ -product uncapacitated facility location problem with multi-type clients consists of two set of sites, one is the set of demand points where clients are located and the other is the set of sites where facilities of unlimited capacities can be set up to serve the clients. Each facility can provide only one kind of products. Each client needs to be served by a set of facilities depending on which products it needs. Each facility can be set up only for one of the  $k$  products with a non-negative fixed cost determined by the product it is designated to provide. There is also a nonnegative cost of shipping goods between each pair of locations. The problem is to determine the set of facilities to be set up and to find an assignment of each client to a set of facilities so that the sum of the setup costs and the shipping costs is minimized. Under the assumption that the setting costs is zero and the shipping costs are in facilities centered metric space, it is shown that the problem with two kinds of clients is NP-complete. Furthermore a heuristic algorithm with worst case performance ratio not more than  $2-1/k$  is presented for any integer  $k$ .

**Keywords:** Heuristic Algorithm, Complexity, Facility Location

## 1. Introduction

In the last few years, a number of constant factor approximation algorithms have been proposed for facility location problem when the service cost is assumed to be in the metric space by Ageev [1], Chudak and Shmoys [2], Guha & Khuller [3], Mahdian *et al.* [4], Shmoys *et al.* [5] and Zhang [6]. That is, the service cost is assumed to be symmetric and satisfy the triangle inequality. The first heuristic algorithm with a performance guarantee of 3.157 was given by Shmoys *et al.* [5], which is based on a linear program rounding algorithm extended from the filter technique of Lin & Vitter [7].

In the classical simple uncapacitated facility location problem, each client only needs one kind of product. Recently  $k$ -product uncapacitated facility location problem which is proposed by Huang and Li [8] can be described as follow: there is a set of demand points where clients are located and a set of potential sites where facilities of unlimited capacities can be set up. There are  $k$  different kinds of products. Each client needs to be supplied with  $k$  kinds of products by a set of  $k$  different facilities and each facility can be set up to supply only a

distinct product with a non-negative fixed cost determined by the product it intends to supply.

But in practice, it is not the case that all of the clients' demands are the same. For example, let  $k = 2$ , some clients may only need product  $p_1$ , some clients may only need product  $p_2$  and some other clients need both product  $p_1$  and  $p_2$ . More concisely, let  $D$  be the set of clients and  $F$  be the set of potential facilities. There are  $k$  kinds of products,  $P = \{p_1, p_2, \dots, p_k\}$ . For each  $j \in D$ , there is a set  $P_j$  of products demanded by client  $j$ . Certainly  $P_j$  is a subset of  $P$  which contains  $k$  products. Each facility  $i \in F$  may be set up to provide at most one of the products. The cost of setting up a facility  $i$  to supply product  $p_l$  is  $f_i^l$ ,  $i \in F$ ,  $1 \leq l \leq k$ . The cost of shipping between any two points  $i, j \in F \cup D$  is equal to  $c_{ij}$ . Each client  $j \in D$  must be supplied with all of the products in  $P_j$  by a set of  $|P_j|$  facilities. In other words, split source is not allowed for a given product.

## 2. The Formulation of the Problem

Let  $x_{ij}^l$  be equal to 1 if facility  $i$  supplies client  $j$  with product  $p^l$ , and 0 otherwise for any  $i \in F$ ,  $j \in D$

and  $l \in P_j$ . Let  $y_i^l$  be 1 if facility  $i$  is set up to supply product  $p_l$ , and 0 otherwise. Then the problem described above can be formulated as the following integer program:

$$(P1) \min \sum_{l=1}^k \sum_{i \in F} f_i^l y_i^l + \sum_{j \in D} \sum_{i \in F} \sum_{l \in P_j} c_{ij} x_{ij}^l \quad (1)$$

$$s.t. \sum_{i \in F} x_{ij}^l = 1 \quad \forall j \in D, l \in P_j \quad (2)$$

$$x_{ij}^l \leq y_i^l, \quad \forall i \in F, j \in D, l \in P_j \quad (3)$$

$$\sum_{l=1}^k y_i^l \leq 1, \quad i \in F \quad (4)$$

$$x_{ij}^l, y_i^l \in \{0,1\}, \quad \forall i \in F, j \in D, l \in P_j \quad (5)$$

In the above formulation, constraints (2) ensure that each client is supplied by precisely one facility for a product it demands. Constraints (3) ensure that facility  $i$  is set up to supply product  $l$  if client  $j$  receives product  $l$  from this facility. Constraints (4) ensure that each facility is set up to supply at most one kind of the  $k$  products.

Now we consider a special case of integer program (P1) in which  $|P_j|=1$  for any  $j \in D$ , i.e., each client need only one products. But different clients may need different products. Under such an assumption, we can get a partition of client set  $D = D_1 \cup D_2 \cup \dots \cup D_k$  such that all of the clients in  $D_l$  only need product  $p_l$ ,  $l = 1, 2, \dots, k$ . This special case can be formulated as the following integer program:

$$(P2) \min \sum_{l=1}^k \sum_{i \in F} f_i^l y_i^l + \sum_{j \in D} \sum_{i \in F} \sum_{l \in P_j} c_{ij} x_{ij}^l$$

$$s.t. \sum_{i \in F} x_{ij}^l = 1, \quad j \in D_l, l \in \{1, 2, \dots, k\};$$

$$x_{ij}^l \leq y_i^l, \quad i \in F, j \in D_l, l \in \{1, 2, \dots, k\};$$

$$\sum_{l=1}^k y_i^l \leq 1, \quad i \in F;$$

$$x_{ij}^l, y_i^l \in \{0,1\}, \quad i \in F, j \in D_l, l \in \{1, 2, \dots, k\}.$$

**Theorem 1.** To solve any instance of (P1) is equivalent to solve an instance of (P2).

**Proof.** For any instance of (P1), let  $j \in D$ . If  $|P_j| \neq 1$ , then we use  $|P_j|$  copies of client  $j$  to replace client  $j$  such that each copy need only one different product of  $P_j$ , respectively. Then the instance of (P1) is equivalently transformed into an instance of (P2).

Because of Theorem 1, we only consider (P2) in the following of this paper and refer to (P2) as  $k$ -product uncapacitated facility location problem with multi-type clients. When all of the setting up costs are zero, then all facilities can be set up. We use  $P-k-N$  to denote  $k$ -product uncapacitated facility location problem with multi-type

clients when the setting up costs are zero:

$$P-k-N \min \sum_{l=1}^k \sum_{i \in F} \sum_{j \in D_l} c_{ij} x_{ij}^l$$

$$s.t. \sum_{i \in F} x_{ij}^l = 1, \quad \forall j \in D_l, l \in \{1, 2, \dots, k\};$$

$$x_{ij}^l \leq y_i^l, \quad \forall i \in F, j \in D_l, l \in \{1, 2, \dots, k\};$$

$$\sum_{l=1}^k y_i^l = 1, \quad i \in F;$$

$$x_{ij}^l, y_i^l \in \{0,1\}, \quad \forall i \in F, j \in D_l, l \in \{1, 2, \dots, k\}.$$

Throughout this paper, we make the following assumptions on costs unless specially mentioned:

$$(1) \quad f_i^l \geq 0, \forall i \in F, l = 1, 2, \dots, k;$$

$$(2) \quad c_{ij} \geq 0, \forall i, j \in F \cup D;$$

$$(3) \quad c_{ij} = c_{ji}, \forall i, j \in F \cup D,$$

$$(4) \quad c_{ij} \leq c_{ih} + c_{hj}, \forall i, j, h \in F \cup D,$$

$$(5) \quad c_{ij} \leq \min\{c_{ih}, c_{hj}\}, \forall i, j \in F, h \in D$$

Conditions (2)-(4) mean that the ship costs are in metric space and condition (5) means that we only consider the problem with centered facilities. In the following of this paper we say the ship costs are in facilities centered metric space if they satisfy conditions (2)-(5) mentioned above.

### 3. The Complexity of the P-2-N

**Theorem 2.** The P-2-N is NP-complete even under the assumption that the ship costs are in facilities centered metric space.

**Proof.** We will establish the proof by reducing the max-cut problem, which is an NP-complete problem in Garey and Johnson [9], to the P-2-N. Consider the max-cut problem defined on an undirected graph  $G = (\mathbf{V}, \mathbf{E})$  with node set  $\mathbf{V}$  and arc set  $\mathbf{E}$ . Suppose  $\mathbf{V} = \{v_1, v_2, \dots, v_n\}$ , and  $\mathbf{E} = \{e_1, e_2, \dots, e_m\}$ . Let  $\mathbf{V} = S_1 \cup S_2$  be a partition of  $\mathbf{V}$ . The set of the total edges between  $S_1$  and  $S_2$  is defined as a cut of graph  $G$ . Let  $\text{CUT}(S_1, S_2)$  denote the set of such edges. The max-cut problem is to find a partition of  $\mathbf{V}$ ,  $\mathbf{V} = S_1 \cup S_2$ , such that  $|\text{CUT}(S_1, S_2)|$  is maximal. By graph  $G$ , we construct an instance of the P-2-N with the set of facilities  $F = \mathbf{V}$  and the set of clients  $D = E \cup \bar{E}$ , where  $\bar{E} = \{\bar{e}_1, \bar{e}_2, \dots, \bar{e}_m\}$  is a copy of the edge set  $E$ . We define the serving costs as follows:

$$c_{ij} := c(v_i, e_j) = \begin{cases} 1 & e_j \text{ is incident to } v_i, \\ 2 & \text{otherwise.} \end{cases}$$

$$\bar{c}_{ij} := c(v_i, \bar{e}_j) = c_{ij};$$

$$c(v_{i_1}, v_{i_2}) = c(e_{j_1}, e_{j_2}) = c(\bar{e}_{j_1}, e_{j_2}) = c(\bar{e}_{j_1}, \bar{e}_{j_2}) = 1,$$

$$\forall v_{i_1}, v_{i_2} \in V, v_{i_1} \neq v_{i_2}, \forall e_{j_1}, e_{j_2} \in \mathbf{E}, e_{j_1} \neq e_{j_2}.$$

It is easy to show that the ship costs are in facilities centered metric space. Let  $F = \mathbf{V} = S_1 \cup S_2$  be a partition of  $\mathbf{V}$ .  $S_1$  and  $S_2$  are set up to supply products  $p_1$  and  $p_2$ , respectively. Let  $C(S_1, S_2)$  denote the cost of the P- $k$ -N under the ordered partition  $F = \mathbf{V} = S_1 \cup S_2$ . Now we consider the servicing cost of each  $e_j \in \mathbf{E}$  and  $\bar{e}_j \in \bar{\mathbf{E}}$  under this partition of  $F$  and let  $C_j(S_1, S_2)$  denote the total servicing cost of  $e_j$  and  $\bar{e}_j$ . Let  $v_{i_1}$  and  $v_{i_2}$  be the two end points of  $e_j$ , i.e.  $e_j = (v_{i_1}, v_{i_2})$ .

**Case 1.**  $e_j \in \text{CUT}(S_1, S_2)$ .

In this case, the two end points of  $e_j$  fall into  $S_1$  and  $S_2$ , respectively. Without loss of generality, suppose  $v_{i_1}$  is in  $S_1$  and  $v_{i_2}$  is in  $S_2$ . We assign  $v_{i_1}$  to supply  $e_j$  with product  $p_1$  and  $v_{i_2}$  to supply  $\bar{e}_j$  with product  $p_2$ , respectively. It is easy to see that  $C_j(S_1, S_2) = 2$ .

**Case 2.**  $e_j \notin \text{CUT}(S_1, S_2)$ .

In this case, both  $v_{i_1}$  and  $v_{i_2}$  belong to either  $S_1$  or  $S_2$ . Without loss of generality, suppose both  $v_{i_1}$  and  $v_{i_2}$  belong to  $S_1$ . We set  $v_{i_1}$  to supply  $e_j$  with products  $p_1$ . Because  $c_{i_1 j} = 1$  by our definition, the servicing cost of  $e_j$  for  $p_1$  is 1. Furthermore we must select a  $v_{i'} \in S_2$  to supply  $\bar{e}_j$  with  $p_2$ . For any  $v_i \in S_2$ ,  $\bar{c}_{ij} = 2$  because  $e_j$  is not incident to  $v_i$ . Thus we have  $C_j(S_1, S_2) = 3$ .

Because there are exactly  $m - |\text{CUT}(S_1, S_2)|$  edges which are not in  $\text{CUT}(S_1, S_2)$ , from **Case 1** and **Case 2** discussed above, we can conclude that

$$\begin{aligned} C(S_1, S_2) &= \sum_{e_j \in D} C_j(S_1, S_2) \\ &= \sum_{e_j \in \text{CUT}(S_1, S_2)} C_j(S_1, S_2) + \sum_{e_j \notin \text{CUT}(S_1, S_2)} C_j(S_1, S_2) \\ &= 3m - |\text{CUT}(S_1, S_2)|. \end{aligned}$$

Because  $m$ , the number of edges in  $\mathbf{E}$ , is fixed,  $C(S_1, S_2)$  gets its minimal value when  $|\text{CUT}(S_1, S_2)|$  reaches its maximal value, and vice versa. This means an optimal solution of the P-2-N provides an answer to the max-cut problem. Thus we have proved that the P-2-N is NP-complete.

#### 4. Heuristic algorithm for P- $k$ -N

If we remove the conditions that each facility can only supply one product, then (P- $k$ -N) becomes the following problem (P- $k$ -N)':

$$\begin{aligned} (\text{P-}k\text{-N})' \quad & \min \sum_{l=1}^k \sum_{i \in F} \sum_{j \in D_l} c_{ij} x_{ij}^l \\ \text{s.t.} \quad & \sum_{i \in F} x_{ij}^l = 1, \quad \forall j \in D_l, l \in \{1, 2, \dots, k\}; \\ & x_{ij}^l \in \{0, 1\}, \quad \forall i \in F, j \in D_l, l \in \{1, 2, \dots, k\}. \end{aligned}$$

In the following we refer to the optimal solution of (P- $k$ -N)' as the overlap optimal solution of (P- $k$ -N). It is obvious that the value of overlap optimal solution is a lower bound of the value of the optimal solution for (P- $k$ -N). It is easy to get an overlap optimal solution by letting each client be supplied by its closest facility. Now we give an approximation algorithm for (P- $k$ -N):

**Algorithm A.**

**Step 1.** Find an overlap optimal solution by letting each client be supplied by its closest facility. Set  $S_i := \emptyset$ ,  $i = 1, 2, \dots, k$  and  $\bar{F} := F$ .

**Step 2.**  $\forall i \in \bar{F}$ , let

$$B_i^l = \{j \in D_l \mid j \text{ is supplied by } i \text{ in the overlap optimal solution}\};$$

$$a_i^l = \sum_{j \in B_i^l} c_{ij}, \quad l = 1, 2, \dots, k;$$

$$a_i^s = \max\{a_i^l \mid l = 1, 2, \dots, k\}.$$

Let  $a_i^l = 0$  if  $B_i^l = \Phi$ . Set

$$S_s := S_s \cup \{i\}, S_l := S_l, l \neq s, \bar{F} := \bar{F} \setminus \{i\}$$

We iterate this process until  $\bar{F} = \Phi$ .

**Step 3.** We use the facilities in  $S_l$  to supply products  $p_l$  and for each client  $j \in D_l$ , we select its closest facility in  $S_l$  as its supplier,  $l = 1, 2, \dots, k$ .

**Theorem 3.** For the P- $k$ -N with the assumption that the ship costs are in facilities centered metric space, the algorithm A has worst case performance ratio not more than  $2 - \frac{1}{k}$ .

**Proof.**  $\forall j \in D_l$ ,  $l = 1, 2, \dots, k$ , let  $i_l(j)$  is the closest facility of  $j$  in  $S_l$ . For each facility  $i \in F$ , let  $B_i = \bigcup_{l=1}^k B_i^l$  be the set of the clients which use facility  $i$  as its supplier in overlap optimal solution. Now we consider the total costs  $A_i$  produced by the clients of  $B_i$  in the algorithm A. Suppose  $i \in S_{l_0}$ . Then we have

$$a_i^{l_0} = \max\{a_i^l \mid l = 1, 2, \dots, k\}.$$

Therefore we get

$$\begin{aligned} A_i &= \sum_{l=1}^k \sum_{j \in B_i^l} c_{i_l(j)j} = a_i^{l_0} + \sum_{l=1, l \neq l_0}^k \sum_{j \in B_i^l} c_{i_l(j)j} \\ &\leq a_i^{l_0} + \sum_{l=1, l \neq l_0}^k \sum_{j \in B_i^l} (c_{i_l(j)i} + c_{ij}) \\ &\leq a_i^{l_0} + \sum_{l=1, l \neq l_0}^k \sum_{j \in B_i^l} 2c_{ij} \\ &= a_i^{l_0} + 2 \sum_{l=1, l \neq l_0}^k a_i^l \\ &\leq \left(2 - \frac{1}{k}\right) \sum_{l=1}^k a_i^l, \end{aligned}$$

where the first inequality follows from the triangle inequality, the second follows from (5) of the assumption on ship costs and the last follows from

$$a_i^{l_0} = \max \{a_i^l \mid l = 1, 2, \dots, k\}.$$

## 5. Discussion

In this paper we consider the k-product uncapacitated facility location problem with multi-type clients and suggest an approximation algorithm for P-k-N under the assumption that the ship costs are in facilities centered metric space. One interesting question that remains open is whether there exist approximation algorithms with constant worst case performance ratio for the P-k-N when we only suppose that the ship costs are in metric space.

## 6. Acknowledgements

The authors would like to express their thanks to the National Natural Science Foundation of China for financially supporting under Grant No. 10771060 and No. 60872039.

## 7. References

- [1] A. A. Ageev, "Improved Approximation Algorithms for Multilevel Facility Location Problems," *Operations Research Letters*, Vol. 30, No. 5, 2002, pp. 327-332.
- [2] F. A. Chudak and D. B. Shmoys, "Improved Approximation Algorithms for the Uncapacitated Facility Location Problem," *SIAM Journal on Computing*, Vol. 33, No. 1, 2003, pp. 1-25. [doi:10.1016/S0167-6377\(02\)00162-1](https://doi.org/10.1016/S0167-6377(02)00162-1)
- [3] S. Guha and S. Khuller, "Greedy Strikes Back: Improved Facility Location Algorithms," *Journal of algorithm*, Vol. 31, 1999, pp. 228-248.
- [4] M. Mahdian, Y. Y. Ye and J. W. Zhang, "Approximation Algorithms for Metric Facility Location Problems," *SIAM Journal on Computing*, Vol. 36, No. 2, 2006, pp. 411-432. [doi:10.1137/S0097539703435716](https://doi.org/10.1137/S0097539703435716)
- [5] D. B. Shmoys, E. Tardos and K. I. Aardal, "Approximation Algorithms for Facility Location Problems," In the *Proceedings of the 29<sup>th</sup> Annual ACM Symposium on Theory of Computing*, 1997, pp. 265-274.
- [6] J. W. Zhang, "Approximating the Two-Level Facility Location Problem via a Quasi-Greedy Approach," *The 15th ACM-SIAM Symposium on Discrete Algorithms SODA*, 2004, pp. 808-817.
- [7] J. H. Lin and J. S. Vitter, "Approximation Algorithms for Geometric Median Problems," *Information Processing Letters*, Vol. 44, No. 5, 1992, pp. 643-666.
- [8] H. C. Huang and R. H. Li, "A k-Product Uncapacitated Facility Location Problem," *European Journal of Operations Research*, Vol. 185, No. 2, 2008, pp. 552-562. [doi:10.1016/j.ejor.2007.01.010](https://doi.org/10.1016/j.ejor.2007.01.010)
- [9] M. R. Garey and D. S. Johnson (eds.), "Computers and Intractability - a Guide to the Theory of NP-Completeness," W. H. Freeman and Company, San Francisco, 1979.