SYM-OP-IS-2010
XXXVII Symposium on Operational Research
Ministry of Defence Republic of Serbia
Serbian Armed Forces

BILEVEL FACILITY LOCATION: DISCRETE MODELS AND COMPUTATIONAL METHODS

YURY KOCHETOV
Sobolev Institute of Mathematics, Novosibirsk, jkochet@math.nsc.ru

Abstract: *Discrete location theory is one of the dynamic areas in operations research. We present some mathematical models in this field concerning the hierarchical optimization. Main properties of the models and computational methods are reviewed.*

Keywords: *Hierarchical optimization, Stackelberg game, Matheuristics, Competitive locations.*

1. INTRODUCTION

Facility location constitutes a broad stratum of mathematical models, methods, and applications in operations research. This area is interesting for theoretical studies, experimental research and real-word applications. Examples include storage facilities, warehouses, police and fire stations, base stations for wireless services, and others.

In many applications decision are made in hierarchical order. The individual decision makers have no direct control on the decisions of the others. The decision makers at higher level of the hierarchy have the power to strongly influence the preferences and strategies of the decision makers at lower levels. In this paper we consider facility location models with a two-level hierarchy. An early model for a hierarchical decision process was proposed by von Stackelberg. In recent years hierarchical decision processes received an increasing interest in the mathematical programming and operations research. More details and examples can be found in [1].

Sure, we can not cover all aspects of this topic here. In Section 2, the basic single level uncapacitated facility location problem (UFLP) is presented. It is well-known model which closely related with the classical p-median problem. In Section 3, we introduce a bilevel facility location problem with user preferences and show some single level reformulations. In Section 4 the problem UFLP with partial external finance is discussed. An elegant reduction of the problem to a series of the UFLP is considered. In Section 5 we deal with the competitive facility location problem where two noncooperative decision makers, the leader and the follower, compete to attract clients from a given market. This Stackelberg game is more difficult than any problem in the class NP. We discuss complexity of the problem, heuristics and exact methods. Computational results for test instances from the benchmark library *Discrete Location Problems* (http://www.math.nsc.ru/AP/benchmarks/) are presented.

2. THE BASIC MODELS

In the most part of facility location models we assume that there is only one decision maker who selects the sites for facilities. For a given set of users $J = \{1, \dots, n\}$ he knows the production–transportation costs $c_{ij} \geq 0$ for servicing user j from facility if it will be opened in site i. The finite set of potential sites $I = \{1, \dots, m\}$ is given and the fixed cost $f_i \geq 0$ of opening facility for each site is known. The goal is to find a subset $S \subseteq I$ of opening facility in such a way that all users will be serviced with minimal total cost, i.e. to minimize the objective function

$$F(S) = \sum_{i \in S} f_i + \sum_{j \in J} \min_{i \in S} c_{ij}$$

The first item in the objective function specifies the fixed cost for opening facilities. The second item is the production–transportation cost for servicing all users.

The problem is commonly referred to as the uncapacitated facility location problem or, as mentioned in early researches, the simple plant location problem. It is NP-hard in the strong sense and hard to approximate. There is no constant-factor approximation for it unless P = NP. So, the problem does not belong to the class APX. For the metric case, when the matrix (c_{ij}) satisfies the triangle inequality, the problem is strongly NP-hard again and Max SNP-hard. The best approximation algorithm with guarantee performance ratio 1.52 is suggested and a 1.463 factor approximation algorithm would imply P = NP. For special case of the metric UFLP when facilities and users are points in the d-dimensional Euclidean space and the production–transportation costs are geometrical distances between the points, an approximation scheme is suggested, i.e. an ε-factor approximation algorithm for any $\varepsilon > 1$ with running time polynomial in n and m and exponential in d and ε [2].

Suppose now that the fixed costs are the same for all facilities and we open exactly p facilities. So, the first item in the objective function of the UFLP is a constant and we need to minimize the objective function

$$F(S) = \sum_{j\in J} \min_{i\in S} c_{ij}$$

$$s.t. \quad S \subset I, |S| = p.$$

This problem is known as discrete p-median problem. It is NP–hard in the strong sense and a $2^{q(n,m)}$ factor approximation algorithm for any polynomial q would imply P = NP. In other words, the problem does not belong to the class APX and good approximate solution is hard to find, as well as the optimal one.

3. FACILITY LOCATION WITH USER PREFERENCES

Up to now we have assumed that there was one decision maker only who tried to minimize the total cost of opening facilities and servicing the users. However, users may be free to choose the facility. They may have own preferences, for example, the travel time to a facility. They don't have to minimize the production and transportation costs of the firm. Hence, we should include the user preferences into the mathematical model [3].

Let the matrix (g_{ij}) define the user preferences on the set I. If $g_{i_1 j} < g_{i_2 j}$, then the user j prefers the facility i_1. We assume for simplicity that all elements are different in each column of the matrix. Otherwise, we have to consider cooperative and noncooperative strategies for the decision maker and users. So, the decision maker wishes to choose a subset $S \subseteq I$ of opening facilities in such a way that all users will be serviced with minimal total cost taking into account the user preferences. Let x_i=1 if facility i is opened and x_i=0 otherwise; x_{ij}= 1 is user j is serviced from facility i and x_{ij}= 0 otherwise. For this case, the mathematical model can be presented as the 0–1 bilevel linear programming problem: minimize

$$\sum_{i\in I} f_i x_i + \sum_{j\in J}\sum_{i\in I} c_{ij} x^*_{ij}(x_i)$$

$$s.t. \quad x_i \in \{0,1\},\ i\in I,$$

where $x^*_{ij}(x_i)$ is the optimal solution of the user problem:

$$\min_{x_{ij}} \sum_{j\in J}\sum_{i\in I} g_{ij} x_{ij}$$

$$s.t. \quad \sum_{i\in I} x_{ij} = 1, \quad j\in J,$$

$$x_{ij} \le x_i, \quad i\in I,\ j\in J,$$

$$x_{ij} \in \{0, 1\}, \quad i\in I,\ j\in J.$$

The objective function of the decision maker is, as before, the total cost for opening facilities and servicing all users. But now, the feasible domain is described by constraints and the auxiliary optimization problem. The values of variables x_i, $i\in I$ are given for the auxiliary problem. It is a new type of optimization problems which can be NP–hard even for continuous variables [1].

The uncapacitated facility location problem with user preferences (UFLPUP) can be reduced to a single level problem [3, 4]. Observe that only the ranking of the g_{ij} for each j is of importance and not their numerical values. Let $S_{ij} = \{\ l\in I \mid g_{lj} < g_{ij}\ \}$, $i\in I$, $j\in J$. For an optimal solution $x^*_{ij}(x_i)$ of the user problem we have

$$x^*_{ij} = 1 \Rightarrow x_l = 0, \quad l\in S_{ij}.$$

We can therefore rewrite the UFLPUP as follows:

$$\min \sum_{i\in I} f_i x_i + \sum_{j\in J}\sum_{i\in I} c_{ij} x_{ij}$$

$$s.t. \quad x_{ij} + x_l \le 1, \quad l\in S_{ij},\ i\in I,\ j\in J,$$

$$\sum_{i\in I} x_{ij} = 1, \quad j\in J,$$

$$x_{ij} \le x_i, \quad i\in I,\ j\in J,$$

$$x_i, x_{ij} \in \{0, 1\}, \quad i\in I,\ j\in J.$$

Indeed, in every optimal solution of the problem all constraints of UFLP will be satisfied and the first constraint will ensure that x_{ij} is an optimal solution for the user problem. The number of variables of the problem is the same as in the UFLP. However, while the UFLP already has the large number of constraints $n + nm$, the UFLPUP has $O(m^2 n)$ additional ones. This prohibits a direct resolution except for small instances. To avoid too numerous additional constraints we can rewrite them in the equivalent form:

$$\sum_{l\in S_{ij}} x_l \le |S_{ij}|(1 - x_{ij}), \quad i\in I,\ j\in J,$$

or

$$x_i \le x_{ij} + \sum_{l\in S_{ij}} x_l, \quad i\in I,\ j\in J,$$

or

$$x_i \le x_{ij} + \sum_{l\in S_{ij}} x_{lj}, \quad i\in I,\ j\in J.$$

It is easy to see that the last inequality produces better linear programming relaxation than three previous ones [3, 4].

The special case of the UFLPUP when $f_i = 0$ for all $i\in I$ is interesting too. For the UFLP this case is trivial, optimal solution can be computed in linear time. But for the UFLPUP this case is NP–hard and the integrality gap can be arbitrary close to 1. If $c_{ij} = g_{ij}$ we get the UFLP. If $c_{ij} = -g_{ij}$ we can solve the problem in polynomial time [3]. Other reformulations, valid inequalities, branch and cut methods and computational results can be found in [5-6].

4. FACILITY LOCATION WITH PARTIAL EXTERNAL FINANCE

Let us consider the UFLP for the case when we have to send a part of product to investor (UFLPI). The investor receives the product by reduced prices and tries to maximize own profit. In fact, we face with a bilevel mixed integer problem. The upper level is an uncapacitated facility location problem with an additional restriction for product distribution. The lower level is a knapsack problem with continuous variables. To present the mathematical formulation, let us introduce the total investment W and the following parameters for facility i:

$d_i > 0$ is the product cost of facility,
$p_{ij} \geq 0$ is the demand of user j if it is serviced by facility,
$\alpha_i \geq 0$ is the product weight for investor,
$\beta_i \geq 0$ is the reduced price for investor.

Additional decision variables:
$v_i \geq 0$ is amount of product of facility,
$w_i \geq 0$ is amount of product for investor.

Using the previous and new notations we can write the problem as follows [7]:

$$\min \sum_{i\in I} \left\{ f_i x_i + d_i v_i + \sum_{j\in J} c_{ij} x_{ij} \right\} \quad (1)$$

$$s.t. \quad \sum_{i\in I} x_{ij} = 1, \quad j \in J, \quad (2)$$

$$x_i \geq x_{ij} \geq 0, \quad i\in I,\ j\in J, \quad (3)$$

$$\sum_{j\in J} p_{ij} x_{ij} \leq v_i - w_i^*, \quad i \in I, \quad (4)$$

$$x_i \in \{0,1\}, \quad i\in I, \quad (5)$$

where w_i^* is the optimal solution of the investor problem:

$$\max \sum_{i\in I} \alpha_i w_i \quad (6)$$

$$s.t. \quad \sum_{i\in I} \beta_i w_i \leq W\ , \quad (7)$$

$$v_i \geq w_i \geq 0, \quad i\in I. \quad (8)$$

The objective function (1) is the total cost for opening facilities and servicing users taking into account the additional product for investor. Constraint (2) requires servicing all users. Constraint (3) allows servicing users from open facilities only. Constraint (4) guarantees that first of all the product is sent to investor and the remaining part is used by users. Objective function (6) expresses the goal of investor which wishes to maximize a weighted sum of product obtained. Constraints (7), (8) limit the product by the amount of investment and the total product of each facility.

Problem UFLPI is NP-hard in the strong sense. If $W = 0$ the problem can be reduced to UFLP. It is known [9] that UFPLI can be solved by a reduction to series of the UFLP. Assume for convenience that $\frac{\alpha_1}{\beta_1} > \frac{\alpha_2}{\beta_2} > ... > \frac{\alpha_m}{\beta_m}$. Let $I_k = \{k, k+1,\dots, m\}$ and P_k is the following problem:

$$\min \sum_{i\in I_k} \left(f_i x_i + \sum_{j\in J} (c_{ij} + d_i p_{ij}) x_{ij} \right) + d_k \frac{W}{\beta_k}$$

$$s.t. \quad \sum_{i\in I} x_{ij} = 1, \quad j\in J,$$

$$x_i \geq x_{ij} \geq 0, \quad i\in I,\ j\in J,$$

$$x_k = 1,\ x_i \in \{0,1\}, \quad i\in I_k.$$

For given optimization problem $(\cdot)$ we denote $V(\cdot)$ an optimal value of the objective function. It is easy to see that $V(\text{UFLPI}) = \min_{k\in I} V(P_k)$. In [8,9] exact and approximate methods for the problem UFLPI are developed based on the results for UFLP.

5. COMPETITIVE LOCATION WITH FORESIGHT

Let us assume that two firms want to open facilities. The first firm, we call it the leader, opens its own set of facilities $X \subset I$, $|X| = p$. Later, the second firm, we call it the follower, observes the set X and opens its own set of facilities $Y \subset I$, $Y \cap X = \varnothing$, $|Y| = r$. Each user selects one facility from the union $X \cup Y$ according to its own preferences. We will assume that any firm will get the profit $w_j > 0$ if it services the user j. Each firm tries to maximize the own profit. The firms do not have the same rights. The leader makes a decision first. The follower makes a decision analyzing the set X. It is a Stakelberg game for two players, where we need to maximize the total profit of the leader [10]. Let us formulate this game as a linear 0–1 bilevel programming problem. Introduce decision variables:

$x_i = 1$ if facility i is opened by the leader, $x_i = 0$ otherwise;
$y_i = 1$ if facility i is opened by the follower, $y_i = 0$ otherwise;
$z_j = 1$ if user j is serviced by the leader, $z_j = 0$ if user j is serviced by the follower.

For a given solution x, we can define the set of facilities which allow capturing user j by the follower:

$$I_j(x) = \{i\in I \mid g_{ij} < \min_{l\in I}(g_{lj} \mid x_l = 1)\},\ j\in J.$$

Note that we consider *conservative* users. If a user has the same distances to the closest leader and the closest follower facilities, he prefers the leader facility. So, the follower never opens a facility at a site where the leader has a facility [10]. Now the model can be written as a linear 0–1 bilevel programming problem [11]:

$$\max_x \sum_{j\in J} w_j z_j^*(x) \quad (9)$$

$$s.t. \quad \sum_{i\in I} x_i = p, \quad (10)$$

$$x_i \in \{0,1\}, \quad i\in I, \quad (11)$$

where $z_j^*(x)$ is a component of the optimal solution of the follower problem:

$$\max_{y,z} \sum_{j\in J} w_j (1 - z_j) \quad (12)$$

$$s.t. \quad \sum_{i\in I} y_i = r, \quad (13)$$

$$1 - z_j \leq \sum_{i\in I_j(x)} y_i\ , \quad j\in J, \quad (14)$$

$$x_i + y_i \leq 1, \quad i \in I, \qquad (15)$$

$$y_i, z_j \in \{0,1\}, \quad i \in I, \ j \in J. \qquad (16)$$

The objective function (9) defines the total profit of the leader. Equation (10) guarantees that the leader opens exactly p facilities. The objective function (12) defines the total profit of the follower. Equation (13) requires opening exactly r facilities for the follower. Constraint (14) guarantees that user j is serviced by the leader if the follower has no facilities in the set $I_j(x)$. Constraint (15) allows opening a facility by at most one decision maker. As we have mentioned, the constraint (15) is redundant. Nevertheless, we use it to reduce the feasible domain of the follower problem.

It is known that the problem is Σ_2^P -hard [12]. So, we deal with the more difficult problem than any NP-complete problem. Polynomially solvable cases and complexity results can be found in [13]. In order to get an upper bound for this maximization problem, it can be rewritten as a single level mixed integer linear program with exponential number of constraints and variables. If we extract a subfamily of constraints and variables, we get the desired upper bound. In [11] a nonclassical column generation method is applied to find optimal solution for the bilevel problem. Computational experiments for the test instances from the benchmark library *Discrete Location Problems* indicate that the exact method allows to find the global optimum for $p = r = 5$, $n = m = 100$.

For higher dimension we may apply heuristics or metaheuristics. The simplest heuristic for the leader is to ignore the follower. The leader opens own facilities to minimize the total distance between users and his facilities. He wishes to service all users and solves the classical p-median problem. This strategy is not so bad despite ignoring the follower. Computational experiments show that this lower bound can be improved by a few percents only.

The second strategy is more sophisticated. The leader anticipates that the follower will react to his decision. So, the (p+r) facilities will be opened. According to the second heuristic, the leader solves the (p+r)-median problem and opens p most profitable facilities. Unfortunately, this strategy is weak.

The third strategy is alternate. It was suggested for continuous locations. This heuristic is iterative. For a given solution of one decision maker, we find the optimal solution for another one. In discrete case this strategy produces a cycle. The best solution in the cycle is the result of the approach. If we use the previous strategies to create a starting solution, we can improve the profit of the leader. Surely, it is a more time consuming procedure.

One of the most powerful approaches is a hybrid memetic algorithm where a tabu search is used to improve the elements of the population [11]. To evaluate neighboring solutions for the leader, the linear programming relaxation of the follower problem is solved by CPLEX software. To reduce the running time at each step of the tabu search, the idea of randomized neighborhoods is used. Other heuristics can be found in [14, 15].

REFERENCES

[1] Dempe, S., *Foundations of bilevel programming*. Kluwer Academic Publishers, Dordrecht, 2002.

[2] Korte, B., Vygen, J., *Combinatorial Optimization. Theory and Algorithms*. 3-rd Edition, Springer. 2005.

[3] Hansen, H., Kochetov, Y., Mladenovic, N., *Lower bounds for the uncapacitated facility location problem with user preferences*. Preprint G-2004-24, GERAD-HEC, Montreal, 2004.

[4] Vasil'ev, I., Klimentova, K., Kochetov, Y., “New lower bounds for the facility location problem with user preferences”. *Computational Mathematics and Mathematical Physics*, 49(6) (2009) 1055-1066.

[5] Canovas, L., Garcia, S., Labbe, M., Marin, A., “A strengthened formulation for the simple plant location problem with order”. *Operations Research Letters*, 35(2) (2007) 141-150.

[6] Alekseeva, E., Kochetov, Y., “Genetic local search for the *p*-median problem with user preferences”. *Discrete Analysis and Operations Research*, 14(1) series 2 (2007) 3-31 (in Russian).

[7] Kochetov, Y., Alexandrov, D., *Simple plant location problem with partial external finance: lower bounds, heuristics and exact solution.* Operations Research Proceedings 1996, (1997) 90-94.

[8] Kochetov, Y., Plyasunov, A., “The choice of tools collection problem with partial external finance”. *Discrete Analysis and Operations Research*, 9(2) series 2 (2002) 78-96 (in Russian).

[9] Ivanenko, D., Plyasunov, A., “ On Lagrangean relaxa tions for the choice of tools collection problem with partial external finance and capacity constraints”. *Discrete Analysis and Operations Research*, 11(2) series 2 (2004) 69-93 (in Russian).

[10] Hakimi, S.L., Locations with spatial interactions: competitive locations and games. In: P. Mirchandani, R. Francis (Eds.) *Discrete Location Theory*. Wiley, 1990, 439-478.

[11] Alekseeva, E.V., Kochetova, N.A., Kochetov, Y.A., Plyasunov, A.V., *A heuristic and exact methods for the discrete (r|p)-centroid problem.* LNCS 6022, Berlin: Springer (2010) 11-22.

[12] Noltemeier, H., Spoerhase, J., Wirth, H.C., “Multiple voting location and single voting location on trees”. *European J. Oper. Res*. 181 (2007) 654-667.

[13] Spoerhase, J., Wirth, H.C., *(r,p)-Centroid problems on paths and trees*. Tech. Report 441, Inst. Comp. Science, University of Wurzburg, 2008.

[14] Alekseeva, E., Kochetova, N., *Upper and lower bounds for the competitive p-median problem.* Proceedings of XIV Baikal International School-Seminar, 1 (2008) 563-569 (in Russian).

[15] Kochetov, Y., Kononov, A., Plyasunov, A., “Competitive facility location models”. *Computational Mathematics and Mathematical Physics*. 49 (6) (2009) 994-1009.