



**HAL**  
open science

# The capacitated plant location problem with customers and suppliers matching

Zhanguo Zhu, Feng Chu, Linyan Sun

► **To cite this version:**

Zhanguo Zhu, Feng Chu, Linyan Sun. The capacitated plant location problem with customers and suppliers matching. *Transportation Research Part E: Logistics and Transportation Review*, 2010, 46 (3), pp.469-480. 10.1016/j.tre.2009.09.002 . hal-00645146

**HAL Id: hal-00645146**

**<https://hal.science/hal-00645146>**

Submitted on 30 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The capacitated plant location problem with customers and suppliers matching

Zhanguo Zhu<sup>a,b,\*</sup>, Feng Chu<sup>c</sup>, Linyan Sun<sup>a,d</sup>

<sup>a</sup>School of Management, Xi'an Jiaotong University, Xi'an 710049, China

<sup>b</sup>Université de technologie de Troyes, Institut Charles Delaunay, FRE CNRS 2848, Laboratoire d'optimisation des systèmes industriels (LOSI), 12 rue Marie Curie, BP 2060, 10010 Troyes Cedex, France

<sup>c</sup>Laboratoire d'informatique, biologie intégrative et systèmes complexes (IBISC), FRE CNRS 3190, Université d'Evry Val d'Essonne, 40 Rue du Pelvoux, CE1455 Courcouronnes, 91020 Evry Cedex, France

<sup>d</sup>State Key Laboratory for Mechanical Manufacturing Systems Engineering, Xi'an 710049, China

This paper introduces a new problem called the capacitated plant location problem with customer and supplier matching (CLCSM). The product distribution from plants to customers and the material supply from suppliers to plants are considered together. We merge a distribution trip and a supply trip into one triangular trip for saving allocation cost. Vehicles from plants visit a customer and a supplier for each trip. We provide a heuristic solution procedure based on Lagrangian relaxation. Computational results indicate that the proposed heuristic solution procedure is shown to be efficient yielding optimal or near-optimal solutions for randomly generated instances.

## 1. Introduction

A supply chain consists of suppliers, manufacturers, customers, the flow of material, information, and finances between them. The management of a supply chain system is divided into three levels: strategic, tactical, and operational, an efficient coordination of which can significantly reduce the total cost. Supply chain network design usually begins with the selection of the potential sites and the required capacity of the plants. Opening new plants requires a great deal of investment and the plants are expected to operate for a long time, as mentioned by Melo et al. (2009). The location of plants has a long-term effect on the operation of an entire supply chain and is one of the most important strategic decisions in the design of the supply chain network.

The plant location problem and its extensions have been widely studied. The early location problems studied include  $p$ -median problem, the set covering problem, the minimax problem amongst others. The first one involves selecting  $p$  plants in a finite set to minimize the total weighted distances between the customers and the plants which satisfy the demand (Hakimi, 1964). The second one finds the minimum number of plants needed to cover all demands with a maximum acceptable service distance (Toregas et al., 1971). The third one which is  $p$ -center problem in this context addresses the problem of minimizing the maximum distance between any given customer and their nearest plants, for details about these three problems please see Owen and Daskin (1998). In addition, there are other kinds of plant location problems such as the

---

\* Corresponding author. Address: School of Management, Xi'an Jiaotong University, Xi'an 710049, China. Tel.: +33 325 718091; fax: +33 325 715649. E-mail addresses: zhanguo.zhu@utt.fr (Z. Zhu), feng.chu@ibisc.univ-evry.fr (F. Chu), lysun@mail.xjtu.edu.cn (L. Sun).

capacitated plant location problem (CPLP), the uncapacitated plant location problem (UPLP), the single and multi source location problem, two-echelon location problem, the location-routing problem (LRP) and so on. Details of these problems are given in the following.

CPLP aims to find the set of open plants with limited capacity which satisfy the demands of all customers and whose objective is to minimize the total cost, including the opening cost of plants and the allocation cost of assigning products to customers (Balinski, 1965). Ignoring the capacity limitation of the plant, it is just the UPLP. If each customer is served by only a single plant, the problem is the single source capacitated plant location problem (SSCPLP). Two heuristic methods: greedy heuristic and interchange heuristic are used to solve a large scale SSCPLP in early research by Jacobsen (1983). Another heuristic method usually used to solve SSCPLP is the Lagrangian heuristic method proposed by Cornuejols et al. (1991).

Depending on the number of facility layers in which location decisions are made, the location problem is divided into single-echelon, two-echelon and multiple-echelons. The process of two-echelon facility location problem is to locate facilities, deliver products from plants to warehouses (hubs), then to customers. The objective is to minimize the opening cost of the facilities and the allocation cost in order to satisfy the demands of the customers. It includes capacitated and uncapacitated two-echelon facility location problem. Many researchers have studied the two-echelon uncapacitated facility location problem and developed algorithms to solve it, such as Marín (2007), Gao and Robinson (1992), Ro and Tcha (1984). Tragantalerngsak et al. (1997) have studied two-echelon capacitated facility location problem. In their model, each facility in the second echelon which has limited capacity can be delivered by only one facility in the first echelon, and each customer is served by only one facility in the second echelon. They proposed six heuristic methods based on Lagrangian relaxation and compared these methods. Amiri (2006) has studied a special two-echelon capacitated location problem which allows for multiple levels of capacities available to the warehouses and plants. Hinojosa et al. (2000) dealt with a multi-period multi-commodity two-echelon capacitated plant location problem in their paper. They developed a heuristic algorithm based on Lagrangian relaxation.

The location-routing problem (LRP) is a research direction combining a classic location problem with a vehicle routing problem. Generally LRP minimizes the opening cost of plants and routing cost to satisfy the demand of the customers by limited vehicle and plant capacities in a potential set of facilities. Webb (1968), Eilon et al. (1971) and Laporte (1989) involved this field early. Recently, Aksent and Altinkemer (2008) have studied a location-routing problem from the perspective of distribution logistics. An augmented Lagrangian relaxation method is used to solve the problem. Caballero et al. (2007) have presented a multi-objective location-routing problem, proposed a metaheuristic method based on tabu search and given a practical application from real life. For more details see the survey by Nagy and Salhi (2007) and Min et al. (1998).

The location problems introduced above have a common feature that empty runs exist during the backhaul from customers to facilities. In a supply chain, a large manufacturing company usually has upstream material/parts suppliers (suppliers), intermediate manufacturing/assembly plants, and downstream distribution centers (customers). The plants need to serve customers to satisfy their demand while also needing a mass of material/parts supply to manufacture/assemble products. The material/parts from suppliers to plants and the products from plants to customers are transported by vehicles. So the transportation process is divided into two parts in this system: the supply trip and the distribution trip. Generally the two trips are independent and there is an empty run for each trip. However, if the empty run can be used effectively, great cost-savings can be made. It means merging a supply trip with a distribution trip into one trigonal trip. This is possible if suppliers and customers are located and mixed in the same supply chain network. In the vehicle routing problem, Imai et al. (2007) has described a similar case in which empty runs are used efficiently. As an important variant of vehicle routing problems, the vehicle routing problem with backhauls (VRPB) contains both deliveries and pickups. The problem is to find a good set of routes by which vehicles make deliveries and pickups after leaving the unique depot and any pickup loads are carried on the backhaul after all deliveries have been made. Literature on the related vehicle routing problem with backhauls is very numerous such as Gendreau et al. (1997), Osman and Wassan (2002). Exact branch and bound method, tabu search algorithm, greedy heuristic method and so on are used to solve this kind of problem. The significant interest of combining the plant location problem with VRP is self-evident. Considering customer and supplier matching in the same trip as part of the plant location decision process will significantly influence the total location cost and future operational cost in a supply chain.

In this paper, we study an extension of the plant location problem: the capacity plant location problem with customer and supplier matching (CLCSM). To the best of our knowledge, the most closely related research on matching customers and suppliers is the uncapacitated facility location problem with client matching (LCM), is proposed by Gourdin et al. (2000) where two clients allocated to a facility can be matched. In the LCM, the allocation cost includes two aspects: the cost of a return trip between the facility and the client, and the cost of a trip containing the facility and two clients. A greedy heuristic and a branch and cut algorithm are developed to solve the problem. However, the limitation of facility capacity and material supply are not considered in this problem. In the CLCSM, customers, suppliers and a set of potential plants are mixed in a same network and each open plant holds its fleet of homogenous vehicles. It is assumed that the vehicles serve customers from a plant, the same vehicles visit suppliers from customers, and then the vehicles transport material/parts from suppliers back to the plant.

The remainder of the paper is organized as follows: Section 2 describes the characteristics of the problem. Section 3 presents the problem formulation. Lagrangian relaxation method and solution approaches to construct feasible solutions based on an optimal solution of the relaxation problem are described in Section 4. Computational results are reported in Section 5. Section 6 concludes the paper and gives the future research direction.

## 2. Problem description

In general, in plant location problems with customers like situation 1 in Fig. 1, the plants send vehicles loaded with products to serve customers and the empty vehicles come back from customers to plants. If we consider the transportation process from the point of view of plants and ignore the difference in cost between loaded runs and empty runs, the allocation cost of type 1 is proportional to the distances between customers and plants and is made up of two parts: the loaded

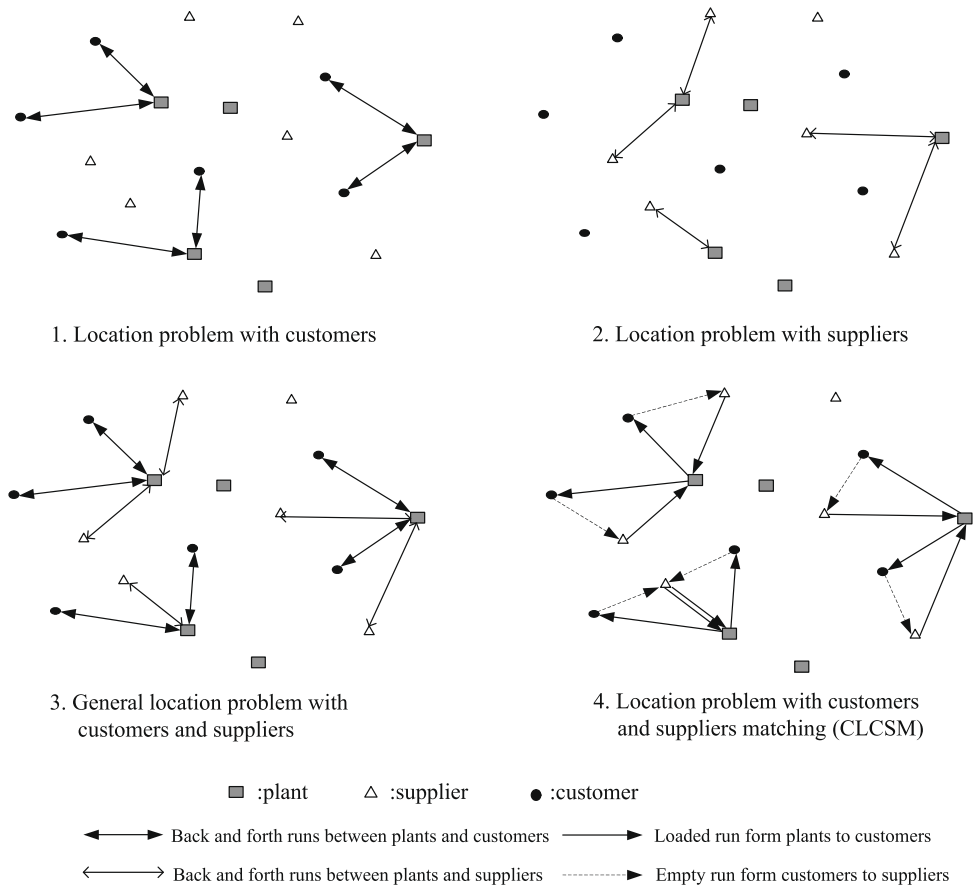


Fig. 1. The location problems.

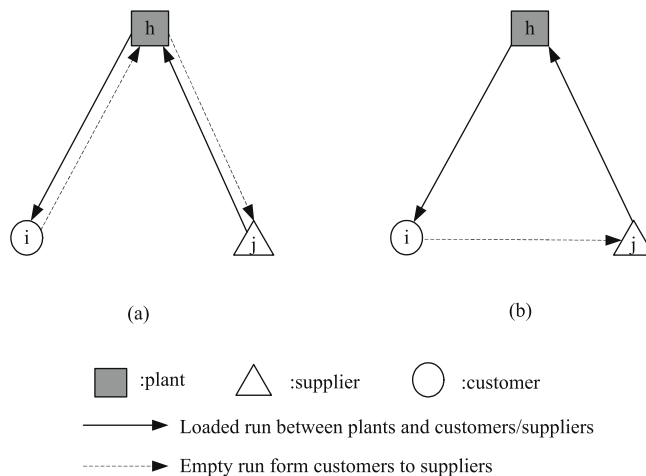


Fig. 2. Trigonal customer and supplier.

allocation runs and the empty backhaul runs. Yang et al. (2007) have described a location problem with suppliers like situation 2 in Fig. 1, where the upstream network suppliers are assigned to plants, and the allocation cost is proportional to the distances between plants and suppliers, this also consists of two parts like the situation 1. We can see that the products allocation to customers from plants or the material/parts transportation from the upstream suppliers to plants need two runs: a loaded run and an empty run.

The plant location problem considering customers and suppliers similar to situations 1 and 2 is presented as situation 3 in Fig. 1. The total allocation cost is proportional to the sums of the distances between plants and customers and the distances between plants and suppliers. Besides, if the product assignment and material/parts transportation are matched, the total cost will obviously decrease according to the geometric property that the sum of the length of two edges is longer than the third one in a triangle, as shown in Fig. 2 based on the one by Imai et al. (2007). That is to say, vehicles with products leave plants to satisfy the customers, unload the products there, visit proper suppliers empty (instead of going back to the plants directly), then load material/parts there, transport the material/parts back to the plant. This is the capacitated plant location problem with customer and supplier matching (CLCSM) described in situation 4 in Fig. 1. The total allocation cost of this situation is proportional to the distance from plants to customers, from customers to suppliers and from suppliers to plants.

The problem discussed here is the location of intermediate plants which need mass supplies of material/parts and products distribution, so full truckloads become necessary to the transportation process in this strategy and each vehicle visits a customer and a supplier in each trip.

### 3. Problem formulation

In this section, a directed network  $G(N, E)$  comprised of a set of nodes  $N$  and a set of directed arcs  $E$  is considered. The set of nodes  $N$  is composed of three categories: a set of potential plants  $H$ , a set of customers  $I$ , a set of potential suppliers  $J$  and  $N = H \cup I \cup J$ . The trigonal distance  $d_{hij}$  is equal to  $d_{hi} + d_{ij} + d_{jh}$ . That means the distance from the plant  $h$  visiting customer  $i$ , from customer  $i$  to supplier  $j$ , from supplier  $j$  back to plant  $h$ . Each potential plant  $h \in H$  with a limited capacity  $b_h$  has a fixed cost  $f_h$ . The limited capacity of supplier  $j$  is  $w_j$ . The demand of customer  $i$  is  $v_i$ . We assume  $v_i \leq w_j$  for all  $i$  and  $j$ ,  $v_i \leq b_h$  for all  $i$  and  $h$ . As mentioned in Section 2, supplies of material/parts and products distribution are mass. Full truckloads are considered in the transportation process, so the demand and other abilities are multiple of truck capacity which means that the unit of  $v_i, w_j$  and  $b_h$  is measured by the quantity corresponding to the number of vehicles. For example,  $v_i = 5$  means that customer  $i$  needs five vehicles of products.  $c$  is the cost per unit distance per unit demand (vehicle). Klose and Drexl (2005) mentioned a model similar to this. However, the allocation variable is not binary and the authors did not give an approach to solve it. The problem can be formulated as a zero-one integer linear program as follows:

Model P:

$$Z_p = \min \sum_h f_h y_h + c \sum_h \sum_i \sum_j v_i d_{hij} x_{hij} \quad (1)$$

s.t.

$$\sum_h \sum_j x_{hij} = 1, \quad \forall i \in I \quad (2)$$

$$\sum_i \sum_j v_i x_{hij} \leq b_h y_h, \quad \forall h \in H \quad (3)$$

$$\sum_h \sum_i v_i x_{hij} \leq w_j, \quad \forall j \in J \quad (4)$$

$$x_{hij} \in \{0, 1\}, \quad \forall h \in H, \quad \forall i \in I, \quad \forall j \in J \quad (5)$$

$$y_h \in \{0, 1\}, \quad \forall h \in H \quad (6)$$

The objective function (1) is to minimize the total opening cost of plants and allocation cost. Constraints (2) ensure each customer is served by one plant and matched with one supplier. Constraints (3) guarantee that each open plant does not supply more than its capacity. Constraints (4) mean that suppliers' capacity must be respected, and a selected supplier can be associated with one or several customers. Constraints (5) and (6) are binary constraints.  $y_h$  is equal to 1 if and only if a plant is opened at site  $h$ ,  $x_{hij}$  is equal to 1 if and only if customer  $i$  and supplier  $j$  are assigned to open plant  $h$ , which means that customer  $i$ , supplier  $j$ , and plant  $h$  are in the same trigonal trip.

To explain the model clearly, we give the following example. Input  $v_1 = 8, v_2 = 9, w_1 = 25, w_2 = 30, b_1 = 50$ , output  $y_1 = 1, x_{111} = 1, x_{121} = 1$ , which means that the plant 1 transports eight vehicles of products to customer 1, unloads products there, then eight empty vehicles travel to supplier 1, loads eight vehicles of material there, then travel to plant 1. The demand of customer 2 is satisfied by plant 1 and supplier 1 in a similar way. So the supplier 1 is visited  $8+9=17$  times, the quantity of material transported from supplier 1 to plant 1 is 17 which keep the flow equilibrium of plant 1.

The CLCSM is NP-hard, which is easily proved. When there is only one supplier with enough capacity, the CLCSM corresponds to the particular case, the single source capacitated plant location problem (SSCPLP), which is known to be NP-hard mentioned by Cortinhal and Captivo (2003). So the CLCSM is NP-hard.

## 4. Solution approaches

### 4.1. The lower bound

As shown above, the problem is NP-hard. Commercial optimization software-packages can only solve small instances exactly, and the solving process is time-consuming with large scale size of instances (see Section 5 below). Besides, basic relaxations such as the linear relaxation usually provide a weak lower bound. So we are seeking an effective heuristic algorithm based on Lagrangian relaxation to solve this problem. Relaxing constraints (2) and (4) of problem  $P$  with multipliers  $\lambda_i$  and  $\mu_j, \lambda_i$  unrestricted,  $\mu_j \geq 0, \forall i \in I, \forall j \in J$ , the relaxed problem is presented as follows:

$P_{LR}$ :

$$\begin{aligned} Z_{LR}(\lambda, \mu) &= \min \left( \sum_h f_h y_h + c \sum_h \sum_i \sum_j v_i d_{hij} x_{hij} + \sum_i \lambda_i \left( \sum_h \sum_j x_{hij} - 1 \right) + \sum_j \mu_j \left( \sum_h \sum_i v_i x_{hij} - w_j \right) \right) \\ &= \min \sum_h f_h y_h + \sum_h \sum_i \sum_j (c v_i d_{hij} + \lambda_i + \mu_j v_i) x_{hij} - \sum_i \lambda_i - \sum_j \mu_j w_j \end{aligned}$$

subject to (3), (5) and (6).

We can decompose this problem into  $|H|$  subproblems, one for each plant  $h$ :

$P_{LR}^h$ :

$$Z_{LR}^h(\lambda, \mu) = \min f_h y_h + \sum_i \sum_j (c v_i d_{hij} + \lambda_i + \mu_j v_i) x_{hij}$$

subject to

$$\begin{aligned} \sum_i \sum_j v_i x_{hij} &\leq b_h y_h \\ x_{hij} &\in \{0, 1\}, \quad \forall i \in I, \quad \forall j \in J \\ y_h &\in \{0, 1\} \end{aligned}$$

For each subproblem  $P_{LR}^h, y_h$  is 0 or 1. If  $y_h$  is equal to 0,  $x_{hij}$  is equal to 0 for all  $i, j$  and the objective function is equal to 0 too because of constraint  $\sum_i \sum_j v_i x_{hij} \leq b_h y_h$ . If  $y_h$  is equal to 1, the subproblem  $P_{LR}^h$  with  $y_h = 1$  can be formulated in the following way.

$$\min \sum_i \sum_j (c v_i d_{hij} + \lambda_i + \mu_j v_i) x_{hij}$$

subject to

$$\begin{aligned} \sum_i \sum_j v_i x_{hij} &\leq b_h \\ x_{hij} &\in \{0, 1\}, \quad \forall i \in I, \quad \forall j \in J. \end{aligned}$$

For each given plant  $h$ , each above subproblem can be transformed into a 0–1 knapsack problem. The transformed process is shown in Algorithm 1. We solve it with the minknap algorithm proposed by Pisinger (1997). So the optimal solution of  $P_{LR}, Z_{LR}(\lambda, \mu) = \sum_h Z_{LR}^h(\lambda, \mu) - \sum_i \lambda_i - \sum_j \mu_j w_j$ , is a lower bound on the optimal value of the original problem  $P$ . The algorithm which gives this lower bound is Algorithm 1.

#### Algorithm 1.

---

Fixed values for the multipliers and plant  $h$  are given.

**Step 1** : For a given  $h \in H$ , list all  $(c v_i d_{hij} + \lambda_i + \mu_j v_i), i \in I, j \in J$ , and set  $\theta_{hij} = -(c v_i d_{hij} + \lambda_i + \mu_j v_i)$ . For  $(c v_i d_{hij} + \lambda_i + \mu_j v_i) \geq 0$  which means  $\theta_{hij} \leq 0$ , we set  $x_{hij}^* = 0$ , because the profit ( $\theta_{hij}$ ) of those items of the knapsack problem below are non-positive.

**Step 2** : For each plant  $h$  listed, solve the following 0–1 knapsack problem using minknap and get the solution  $x_{hij}^*$ :

$$\begin{aligned} \max \sum_i \sum_j \theta_{hij} x_{hij} \\ \text{subject to} \\ \sum_i \sum_j v_i x_{hij} &\leq b_h \\ x_{hij} &\in \{0, 1\}, \forall i \in I, \forall j \in J \end{aligned}$$

**Step 3** : Calculate the objective function value of the subproblem for each plant site  $h$  with  $x_{hij}^*$ :  $Z_{LR}^h(\lambda, \mu) = \min\{f_h + \sum_i \sum_j (c v_i d_{hij} + \lambda_i + \mu_j v_i) x_{hij}^*, 0\}$  and if  $f_h + \sum_i \sum_j (c v_i d_{hij} + \lambda_i + \mu_j v_i) x_{hij}^* < 0, y_h^* = 1$ , otherwise  $y_h^* = 0$ .

**Step 4** : The function value of the relaxation problem is calculated by  $\sum_h Z_{LR}^h(\lambda_i, \mu_j) - \sum_i \lambda_i - \sum_j \mu_j w_j$

---

#### 4.2. The upper bound computation

We construct a feasible solution of CLCSM according to the optimal solution of the Lagrangian relaxation problem. There may be some customers that are not served by any plants at all or matched with many plants and suppliers which violate constraints (2) for the solutions of  $P_{LR}$ . The solutions are repaired to satisfy constraints (2) in the first three steps of the following algorithm. For the solutions which violate constraints (4), we deal with them in the following way.  $J'$  is the set of suppliers the capacities of which are violated.  $I_j$  is the set of customers that are matched with  $j \in J'$ . For each supplier  $j \in J'$ , we solve a subset-sum problem (SSP) described by Martello and Toth (1990) to obtain a subset  $I'_j$  of  $I_j$  the total demand of which is closest to and does not exceed the capacity of the supplier  $j \in J'$ . The customers  $i \in I_j/I'_j$  are ordered by decreasing order of their demand and assigned to other suppliers with enough (residual) capacity according to the lowest assignment cost. For details see the Algorithm 2.

The algorithm used to construct a feasible solution is as follows.

#### Algorithm 2.

*Step 1:* For any  $y_h^* = 0, h \in H$ , set  $x_{hij}^* = 0$ , for all  $i \in I, j \in J$ .

*Step 2:* For any  $i \in I$ , if  $\sum_h \sum_j x_{hij}^* > 1$ , then assign the customer  $i$  to the plant  $h (y_h^* = 1)$  and the supplier  $j$  with the lowest cost  $(c v_i d_{hij})$ , and set  $x_{hij}^* = 1$ , the others  $x_{hij}^* = 0$  (other matches are removed).

*Step 3:* For any  $i \in I$ , if  $\sum_h \sum_j x_{hij}^* = 0$ , order the customers  $i$ , by decreasing order of their demand  $v_i$ ; we denote the residual capacity of  $h$  with  $\sigma_h$ . In this order, assign  $i$  to the open plant  $h$  which has the enough capacity ( $\sigma_h - v_i \geq 0$ ), choose the plant  $h$  and supplier  $j$  which satisfies  $\min_{\sigma_h - v_i \geq 0} c v_i d_{hij}$ , and set  $x_{hij}^* = 1$ . We denote the set of open plants with  $H'$ . If there is no open plant  $h \in H'$  satisfying the condition  $\sigma_h - v_i \geq 0$ , a new plant  $h \in H \setminus H' (y_h^* = 0)$  must be open. The one with the least total cost ( $\min(f_h + c v_i d_{hij})$ ) relative to the customer  $i$  is chosen and set  $y_h^* = 1$ .

*Step 4:* For any  $j \in J', J' = \{j | \sum_h \sum_i v_i x_{hij}^* > w_j, \forall j \in J\}$ ,  $J'$  is the set of suppliers the capacities of which are violated,  $i \in I_j = \{i | x_{hij}^* = 1\}$ , solving the following special case of knapsack problem called the subset-sum problem. It is to find a subset of  $I_j, I_j = \{i | x_{hij}^* = 1, \text{ for a given } j \in J', \forall h \in H\}$ , which is the set of customers that are assigned to a supplier  $j \in J'$ :

$$(SSP) \max \sum_i v_i z_i$$

subject to

$$\sum_i v_i z_i \leq w_j, \forall j \in J'$$

$$z_i \in \{0, 1\}, \forall i \in I_j$$

where

$z_i = 1$ , customer  $i \in I_j$  is assigned to supplier  $j \in J'$  again, we denote the set with  $I'_j, z_i = 0$  otherwise. SSP can be solved with Pisinger's algorithm.

For any customer  $i \in I_j \setminus I'_j$  which means  $z_i = 0$ , order customer  $i$  by decreasing order of  $v_i$  and assign customer  $i$  to other suppliers  $\tilde{j} \in J \setminus \{j\}$  and  $\tilde{h} \in H'$  which have capacities with the lowest  $c v_i d_{\tilde{h}i\tilde{j}}$  and set  $x_{\tilde{h}i\tilde{j}}^* = 0, x_{hij}^* = 1$ .

*Step 5:* The value of the feasible solution (FS) can be obtained with the formula:  $\sum_h f_h y_h^* + \sum_h \sum_i \sum_j c v_i d_{hij} x_{h(ij)}^*$ .

#### 4.3. The overall algorithm

In this section, a subgradient optimization method is adapted to solve the Lagrangian dual problem  $Z_{LR}(\gamma^*) = \text{Max}_{\gamma} (Z_{LR}(\gamma))$  to find the optimal multiplier vector  $\gamma^*$ . This subgradient method is introduced by Fisher (1981) and applied widely in plant

**Table 1**

Notation used in numerical results.

$CT_L$	Computation time of the Lagrangian heuristic algorithm (s)
$CT_{opt}$	Computation time of small instances by the commercial solver Lingo 9.0 (s)
$CT_{LN}$	Computation time of linear relaxation (s)
NOpt	Number of optimal solutions found by the Lagrangian heuristic algorithm
U_gap	100 * (the upper bound - the optimal value) / the optimal value
L_gap	100 * (the lower bound obtained by Lagrangian relaxation - the lower bound obtained by linear relaxation) / the lower bound obtained by linear relaxation
Gap	100 * (the upper bound - the lower bound obtained by Lagrangian relaxation) / the lower bound obtained by Lagrangian relaxation
$ I $	Number of customers
$ J $	Number of suppliers
$ H $	Number of potential plants

location problems such as Nozick (2001) and Li et al. (2009). In each iteration, for a given Lagrangian vector  $\gamma = \{\lambda_i, \mu_j\}$ , the corresponding problem  $P_{LR}$  is solved with Algorithm 1, then a lower bound of the original problem  $Z_{LR}(\lambda, \mu)$  can be obtained. Based on the solution of problem  $P_{LR}$ , a feasible solution is given by the Algorithm 2. The Lagrangian multipliers  $\lambda_i, \mu_j$  are updated with (7) and (8) in the Algorithm 3. In order to prevent constraints (4) dominating the calculation of the updates

**Table 2**  
Comparison of heuristic feasible solutions to optimal solutions in small instances.

$ I $	$ J $	$ H $	U_gap (%)	Gap (%)	NOpt	$CT_L$	$CT_{opt}$
20	5	5	0.68	4.03	3	33	1
20	10	5	0.25	7.57	4	31	2
30	10	5	0.35	4.97	2	35	6
40	10	5	0.23	3.97	1	35	4
50	15	15	0.47	4.32	0	107	43
80	20	20	1.85	6.74	0	160	970

**Table 3**  
Comparison of Lagrangian relaxation and linear relaxation.

$ I $	$ J $	$ H $	L_gap (%)	Gap (%)	$CT_L$	$CT_{LN}$
100	15	10	3.22	6.24	103	11
100	20	10	1.81	5.39	102	17
100	50	10	3.82	5.44	134	47
100	50	15	0.65	8.27	193	93
200	50	20	2.75	4.79	429	758

**Table 4**  
Performance of the overall solution procedure.

$ I $	$ J $	$ H $	Gapavg (%)	Gapmin (%)	Gapmax (%)	$CT_L$
100	20	10	5.39	3.24	7.13	102
100	25	10	4.25	0.98	6.16	97
100	30	10	5.78	2.52	9.49	110
100	35	10	4.82	3.22	6.47	111
100	40	10	4.67	2.92	7.35	115
100	45	10	5.04	1.87	8.18	124
100	50	10	4.93	1.59	8.86	131
100	55	10	4.69	3.38	6.39	138
100	60	10	7.10	4.48	8.49	144
100	65	10	4.28	2.77	5.88	154
100	30	15	6.55	5.39	8.18	155
100	40	15	6.42	3.57	8.86	182
100	50	15	6.87	4.16	8.75	197
100	60	15	6.91	5.07	8.71	214
100	70	15	6.65	5.08	7.85	272
200	30	20	6.10	3.75	8.08	323
200	40	20	6.18	4.65	9.33	374
200	50	20	4.79	2.96	5.66	430
200	60	20	5.86	5.09	8.07	496
200	70	20	7.16	5.58	9.68	563
200	100	20	6.22	4.95	8.69	740
200	30	30	7.09	5.58	9.32	499
200	40	30	5.95	4.34	7.77	538
200	50	30	6.77	4.32	8.34	621
200	60	30	7.19	5.02	9.13	718
200	70	30	6.84	4.42	9.79	812
300	40	30	6.76	3.11	8.74	775
300	50	30	6.78	4.98	7.70	1020
300	60	30	5.92	4.02	7.57	1061
300	70	30	5.94	4.39	6.93	1201
300	100	30	7.25	6.25	9.79	1650
300	40	45	6.78	5.20	8.38	1474
300	50	45	7.17	5.61	9.46	1579
300	60	45	7.12	5.07	8.93	1575
300	70	45	6.33	5.20	6.98	2052
400	60	40	6.06	4.75	7.38	1842



**Table 5**Performance of the overall solution  $c = 0.005, w_j \in [200, 500]$ .

$ I $	$ J $	$ H $	Gapavg (%)	Gapmin (%)	Gapmax (%)	$CT_L$
100	20	10	5.31	3.11	9.46	115
100	25	10	5.56	3.24	9.77	120
100	30	10	4.61	1.78	7.63	131
100	35	10	6.09	3.52	7.97	148
100	40	10	4.42	1.35	8.98	155
100	45	10	8.27	6.78	9.93	176
100	50	10	6.10	3.52	8.25	185
100	55	10	6.76	5.45	9.41	196
100	60	10	5.98	1.82	8.64	202
100	65	10	3.82	1.05	5.23	219
100	30	15	3.90	2.08	8.05	202
100	40	15	4.04	2.08	7.84	220
100	50	15	6.23	4.44	8.58	253
100	60	15	5.95	2.26	10.45	315
100	70	15	5.37	2.97	7.40	347
200	30	20	5.89	3.99	8.79	434
200	40	20	3.40	1.63	6.81	519
200	50	20	5.11	1.98	7.77	621
200	60	20	7.53	4.40	10.18	711
200	70	20	3.89	1.39	6.66	811
200	100	20	4.76	2.83	8.91	1166
200	30	30	5.19	3.17	8.03	620
200	40	30	6.53	3.13	9.78	776
200	50	30	4.15	2.82	6.02	920
200	60	30	4.58	3.75	5.49	1018
200	70	30	5.20	2.20	7.36	1239
300	40	30	6.51	2.86	9.23	1091
300	50	30	5.45	3.20	8.70	1346
300	60	30	5.25	3.04	7.36	1539
300	70	30	4.28	2.77	6.41	2015
300	100	30	3.77	1.87	6.18	2423
300	40	45	4.09	2.58	6.26	1649
300	50	45	5.01	3.70	6.51	1911
300	60	45	4.16	3.11	5.47	2267
300	70	45	4.21	2.86	6.04	2581

of the multipliers we also deal with constraints (4) by dividing the equation by  $w_j$  just like Nozick (2001) did in his paper, because the updates are based on the sum of the square of the differences between the both sides of the relaxed constraints. In Algorithm 3, two termination criteria are set. One is *Maxit* which means the maximum number of iterations, the other is  $\epsilon$  which denotes the allowable gap between the best upper bound *UB* and the best lower bound *LB*. We set *Maxit* = 700 and  $\epsilon = 0.001$ . If one of the termination criteria is met, Algorithm 3 stops.  $\alpha$  is a positive parameter with  $0 \leq \alpha \leq 2$ . We set  $\alpha = 2$  initially and halved its value when the best lower bound did not improve over 50 consecutive iterations. We also initialize *LB*, *UB*, and multipliers.

The overall algorithm is presented as follows:

### Algorithm 3.

#### Initialization

Set the current iteration ( $k$ ) to 0.

Set the Lagrangian multipliers  $\lambda^k = \lambda^* = 0, \mu^k = \mu^* = 0$ .

Set the lower and upper bound,  $LB = 0, UB = 10^8$ .

Set the step length  $\alpha = 2$ .

Set the maximum number of iterations to be *Maxit*=700.

**While** not (termination criterion)

Solve  $Z_{LR}^k(\lambda^k, \mu^k)$  with Algorithm 1 and denote the optimal function value as  $\bar{Z}_{LR}^k(\lambda^k, \mu^k)$ .

If ( $\bar{Z}_{LR}^k(\lambda^k, \mu^k) > LB$ ), set  $LB = \bar{Z}_{LR}^k(\lambda^k, \mu^k), \lambda^* := \lambda^k, \mu^* := \mu^k$ .

Find a feasible solution to the original problem and get *FS* with Algorithm 2.

If ( $FS < UB$ ), set  $UB := FS$ .

Update the multipliers:

$$\lambda_i^{k+1} = \lambda_i^k + \frac{\alpha(UB - \bar{Z}_{LR}^k(\lambda^k, \mu^k))(\sum_h \sum_j x_{hij}^k - 1)}{\sum_i (\sum_h \sum_j x_{hij}^k - 1)^2 + \sum_j (\sum_h \sum_i v_i x_{hij}^k / w_j - 1)^2} \quad (7)$$

$$\mu_j^{k+1} = \mu_j^k + \frac{\alpha(UB - \bar{Z}_{LR}^k(\lambda^k, \mu^k))(\sum_h \sum_i v_i x_{hij}^k / w_j - 1)}{\sum_i (\sum_h \sum_j x_{hij}^k - 1)^2 + \sum_j (\sum_h \sum_i v_i x_{hij}^k / w_j - 1)^2} \quad (8)$$

If (there is no improvement in the best lower bound over 50 consecutive iterations)

Half the step length  $\alpha := \alpha/2$ , reset the multipliers  $\lambda := \lambda^*$ ,  $\mu := \mu^*$ .

Set  $k:=k+1$ .

**End While**

## 5. Numerical experiments

In this section the performance of the proposed algorithm is evaluated by using randomly generated instances with different problem sizes and other parameters. The basic parameters in the following way: the coordinates of locations are generated as  $U(10, 60)$ . Here the notation  $U(a,b)$  means a uniform distribution in the interval  $[a,b]$ . The demand  $v_i$  is generated randomly from  $U(10, 50)$ . The capacities  $b_h$  of the plant range from 200 to 600 and the fixed costs range from 300 to 700, which were based on paper presented by [Holmberg et al. \(1999\)](#). The cost per unit distance per unit demand  $c$  is set to 0.1. Thirty-six problem sets and five instances for each set are generated in order to evaluate the performance of the algo-

**Table 6**

Performance of the overall solution for procedure with  $c = 0.1, w_j \in [100, 400]$ .

$ I $	$ J $	$ H $	Gapavg (%)	Gapmin (%)	Gapmax (%)	$CT_L$
100	20	10	4.57	1.69	7.73	96
100	25	10	5.90	5.05	7.82	104
100	30	10	4.67	2.53	7.69	107
100	35	10	4.46	1.93	7.31	114
100	40	10	7.14	4.93	10.59	125
100	45	10	4.48	2.68	7.85	140
100	50	10	6.63	2.92	9.46	135
100	55	10	4.37	3.72	5.02	137
100	60	10	6.11	3.61	7.91	144
100	65	10	6.78	4.26	10.41	157
100	30	15	6.21	4.85	8.43	148
100	40	15	6.09	5.12	7.36	169
100	50	15	6.87	4.16	8.75	197
100	60	15	8.01	6.44	9.35	217
100	70	15	6.06	3.06	8.24	239
200	30	20	4.96	3.80	6.45	319
200	40	20	4.23	3.13	5.86	374
200	50	20	5.52	3.94	7.33	431
200	60	20	5.53	3.33	9.39	529
200	70	20	5.27	2.85	8.22	559
200	100	20	5.79	3.69	7.45	768
200	30	30	6.60	5.55	7.91	461
200	40	30	7.31	6.00	8.77	532
200	50	30	7.68	6.07	8.96	617
200	60	30	7.18	6.32	7.95	698
200	70	30	7.16	5.53	9.22	792
300	40	30	6.74	5.89	8.49	774
300	50	30	6.49	4.73	7.62	937
300	60	30	6.51	4.69	8.01	1085
300	70	30	6.04	4.92	7.97	1218
300	100	30	5.73	4.53	6.92	1696
300	40	45	8.33	7.21	10.65	1108
300	50	45	7.32	5.94	8.49	1417
300	60	45	8.50	6.17	10.18	1636
300	70	45	7.72	6.27	9.40	1789

rithm. So a total of 180 problem instances are tested for basic parameters. We divide the instances into two kinds: small scale (less than 100 customers), medium and large scale (more than 100 customers).

The algorithm is coded in Visual C++ and has been performed on a CPU of 1.66 GHZ and Windows XP.  $|I|$  is the number of customers,  $|J|$  is the number of suppliers, and  $|H|$  is the number of potential plants. *Gap* reflects the quality of our proposed heuristic algorithm. *Gapavg*, *Gapmin* and *Gapmax* are corresponding to the average value, minimum and maximum of *Gap* for each problem set, respectively. *U\_gap* denotes the gap between the best upper bound of the heuristic algorithm and the optimal objective value solved by Lingo 9.0 in small scale instances. By comparing *Gap* with *U\_gap* we can find where the main part of *Gap* exists: the lower bound or the upper bound in small scale instances, which may provide significant implications for the quality of our proposed heuristic algorithm in large scale instances. The  $CT_L$  and  $CT_{opt}$  are computation times of the heuristic algorithm and Lingo 9.0 in seconds, respectively. With the notations in Table 1, the results of the instances are given in Tables 2–7.

First, we compare the upper bound of the proposed heuristic algorithm with the optimal solution of the commercial software Lingo 9.0 in small scale instances. Table 2 shows that our proposed heuristic provides very good feasible solutions. *U\_gap* changes from 0.23% to 1.85% and most best feasible solutions are very near to optimal solutions. In addition, the best feasible solutions obtained from our proposed heuristic algorithm are the optimal solutions in 10 out of all over 30 instances. The *Gap* varies from 3.97% to 7.57%. So the *Gaps* are mainly caused by the relative weak lower bound. From Table 2 we can also see that in very small instances the proposed heuristics algorithm needs more time than lingo 9.0. However, with the increase of problem size the computation time of Lingo 9.0 increases more quickly. Large scale instances, such as 200 customers, 40 suppliers and 20 plants, could not be solved optimally by using Lingo 9.0 software after 72,000 s of computation.

The lower bounds obtained by the Lagrangian relaxation in this paper and these obtained by the linear relaxation are compared as shown in Table 3. The average value of  $L_{gap}$  for all instances is 2.45% which means that the Lagrangian relaxation generates tighter lower bounds than linear relaxation. We can also find that in relative small size of instances the computational time required by Lagrangian relaxation is more than that of linear relaxation. However, in larger instances, the linear relaxation needs more time. With the increase of the size of the instances the bounds of linear relaxation may not be obtained because of the enormous consumption of memory size and CPU time. For example, the Lingo runs out of memory

**Table 7**  
Performance of the overall solution  $c = 0.005, w_j \in [100, 400]$ .

$ I $	$ J $	$ H $	Gapavg (%)	Gapmin (%)	Gapmax (%)	$CT_L$
100	20	10	5.66	2.34	9.52	121
100	25	10	6.52	4.15	9.84	138
100	30	10	4.33	1.82	6.88	153
100	35	10	7.46	5.12	10.56	153
100	40	10	6.31	2.84	9.67	169
100	45	10	8.34	5.75	10.71	182
100	50	10	5.55	2.91	8.86	194
100	55	10	5.59	2.50	9.02	209
100	60	10	5.34	3.88	7.85	226
100	65	10	7.79	4.73	10.58	230
100	30	15	6.06	2.66	9.07	217
100	40	15	6.65	1.89	9.12	232
100	50	15	5.91	3.54	9.29	282
100	60	15	4.41	1.78	6.15	315
100	70	15	5.20	4.31	6.36	362
200	30	20	8.12	5.49	9.58	454
200	40	20	6.13	4.01	8.15	577
200	50	20	6.17	3.45	9.92	631
200	60	20	5.64	3.63	8.92	741
200	70	20	5.84	4.45	9.54	857
200	100	20	6.30	3.90	9.41	1159
200	30	30	6.15	5.58	6.99	758
200	40	30	4.12	3.70	5.10	955
200	50	30	5.42	3.97	8.11	1160
200	60	30	4.01	2.16	5.99	1243
200	70	30	5.65	3.58	7.46	1285
300	40	30	5.59	3.43	8.62	1128
300	50	30	5.64	3.38	7.45	1514
300	60	30	5.66	2.89	9.16	1578
300	70	30	6.19	3.61	8.27	1883
300	100	30	4.78	3.07	7.12	2605
300	40	45	5.36	2.55	9.78	1718
300	50	45	4.44	2.83	6.46	2240
300	60	45	4.11	3.42	5.57	2327
300	70	45	3.78	2.58	5.32	2677

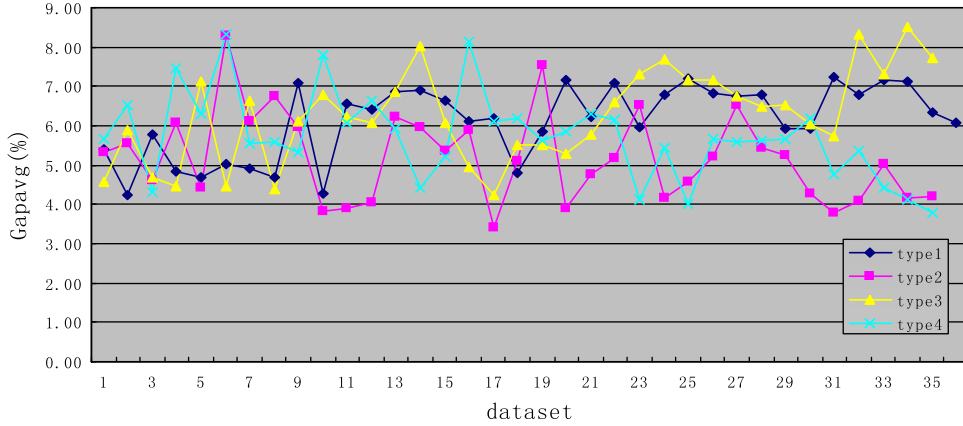


Fig. 3. The Gapavg in all types of instances.

on the instance (400 customers, 60 suppliers, and 40 plants). So it is necessary to propose a heuristic algorithm to solve the large size instances of the problem.

Table 4 gives the computation results for medium and large scale instances. The *Gapavg* for each set changes from 4.25% to 7.25% and the largest *Gapmax* is 9.79%. The average value of *Gapavg* and computation time for all 36 sets in Table 3 are 6.13% and 627 s, respectively. From Table 4, we can also see that the number of customers, suppliers and plants has little influence on the computation result of the proposed heuristic algorithm.

To test the stability of the proposed heuristic algorithm, parameter  $c$  and  $w_j$  are changed based on basic parameters  $c = 0.1$  and  $w_j \in [200, 500]$  keeping other parameters unchanged. Three other types of instances are evaluated. We tested 35 sets for each of them. The numerical results for  $c = 0.005, w_j \in [200, 500]$ ,  $c = 0.1, w_j \in [100, 400]$  and  $c = 0.005, w_j \in [100, 400]$ , are presented in Tables 5–7.

The average values of *Gapavg* for all instances in Tables 5–7 are 5.18%, 6.26% and 5.72%, respectively. The average value of *Gapavg* for all instances of these four types is 5.82%. The scopes of *Gapavg* are 3.40–8.27%, 4.23–8.50% and 3.78–8.34%, respectively. The comparison of the *Gapavg* in all types of instances is shown in Fig. 3. Type 1–4 are corresponding to Tables 4–7, respectively. We can find that the Gaps vary little with changes of the problem parameters and problem size, so the heuristic approach is not only efficient but also robust. The computation time of all instances is reasonable. Even for a large-scale instance with 400 customers, 60 suppliers, 40 plants the time is within (less than) 1900 s.

## 6. Conclusion

In this research, a new type of plant location problem: the capacitated plant location problem with customer and supplier matching (CLCSM) is studied. By matching a customer and a supplier in a triangle trip, transportation costs can be significantly reduced which also influence the location of the plants. The contributions of the paper to the literature are not only considering the matching in the plant location but also integrating the limitation of facility capacity and material supply process in the problem. An integer linear programming model is proposed first and a heuristic algorithm based on Lagrangian relaxation is developed. The Lagrangian relaxation problem which provides a lower bound is decomposed into series of 0–1 knapsack problems and the solutions obtained from the relaxed problem is used to construct a feasible solution of (CLCSM) where series of subset–sum problems are solved. The proposed algorithm was tested on different size of random instances and demonstrated that it is very efficient in solving the problems with the average gap about 6% in reasonable computation times, even for large problems with up to 400 customers, 60 suppliers and 40 potential plant sites. The algorithm is promising for practical applications to the location problem.

In future research, another kind of matching in the plant location problem will be studied. Each potential plant has a maximal matching distance based on travel time limitation of the vehicles which is a criterion to match a customer and a supplier. It is not necessary for every customer to be matched. Whether customers and suppliers are matched depends on the optimization process. Those which are not matched are transported separately. Besides, the ratio of loading of the vehicles will also be concerned.

## Acknowledgements

The authors would like to thank the Editor-in-Chief and the anonymous reviewer for their insightful comments and helpful suggestions on improving the quality of this work. This work was supported partly by National Natural Science Foundation of China (Grant number: 70433003 and 70701029) and State-Sponsored Study Abroad Program of China (CSC).

## References

- Aksen, D., Altinkemer, K., 2008. A location-routing problem for the conversion to the click-and-mortar retailing: the static case. *European Journal of Operational Research* 186, 554–575.
- Amiri, A., 2006. Designing a distribution network in a supply chain system: formulation and efficient solution procedure. *European Journal of Operational Research* 171, 567–576.
- Balinski, M.L., 1965. Integer programming: methods, uses, computation. *Management Science* 12, 253–313.
- Caballero, R., González, M., Guerrero, F.M., Molina, J., Paralara, C., 2007. Solving a multiobjective location routing problem with a metaheuristic based on tabu search. Application to a real case in Andalusia. *European Journal of Operational Research* 177, 1751–1763.
- Cornuejols, G., Sridharan, R., Thizy, J.M., 1991. A comparison of heuristics and relaxations for the capacitated plant location problem. *European Journal of Operational Research* 50, 280–297.
- Cortinhal, M.J., Captivo, M.E., 2003. Upper and lower bounds for the single source capacitated location problem. *European Journal of Operational Research* 151, 333–351.
- Eilon, S., Watson-Gandy, C.D.T., Christofides, N., 1971. *Distribution Management Mathematical, Modeling and Practical Analysis*. Griffin, London.
- Fisher, M.L., 1981. The Lagrangian relaxation method for solving integer programming problem. *Management Science* 27, 1–18.
- Gao, L.L., Robinson, E.P., 1992. A dual-based optimization procedure for the two-echelon uncapacitated facility location problem. *Naval Research Logistics* 39, 191–212.
- Gendreau, M., Laporte, G., Hertz, A., 1997. An approximation algorithm for the traveling salesman problem with backhauls. *Operations Research* 45, 639–641.
- Gourdin, E., Labbe, M., Lapporte, G., 2000. The uncapacitated facility location problem with client matching. *Operations Research* 48, 671–685.
- Hakimi, S., 1964. Optimum location of switching centers and the absolute centers and medians of a graph. *Operations Research* 12, 450–459.
- Hinojosa, Y., Puerto, J., Fernández, F.R., 2000. A multiperiod two-echelon multicommodity capacitated plant location problem. *European Journal of Operational Research* 123, 271–291.
- Holmberg, K., Rönnqvist, M., Yuan, D., 1999. An exact algorithm for the capacitated facility location problems with single sourcing. *European Journal of Operational Research* 113, 544–599.
- Imai, A., Nishimura, E., Current, J., 2007. A Lagrangian relaxation-based heuristic for the vehicle routing with full container load. *European Journal of Operational Research* 176, 17–105.
- Jacobsen, S.K., 1983. Heuristics for the capacitated plant location model. *European Journal of Operational Research* 12, 253–261.
- Klose, A., Drexl, A., 2005. Facility location models for distribution system design. *European Journal of Operational Research* 162, 4–29.
- Laporte, G., 1989. A survey of algorithms for location-routing problems. *Investigación Operativa* 1, 93–123.
- Li, J., Chu, F., Prins, C., 2009. Lower and upper bounds for a capacitated plant location problem with multi commodity flow. *Computers and Operations Research* 36, 3019–3030.
- Marín, A., 2007. Lower bounds for the two-stage uncapacitated facility location problem. *European Journal of Operational Research* 179, 1126–1142.
- Martello, S., Toth, P., 1990. *Knapsack Problems: Algorithm and Computer Implementations*. John Wiley & Son Ltd., New York.
- Melo, M.T., Nickel, S., Saldanha-da-Gama, F., 2009. Facility location and supply chain management – a review. *European Journal of Operational Research* 196, 401–412.
- Min, H., Jayaraman, V., Srivastava, R., 1998. Combined location-routing problems: a research directions synthesis and future research direction. *European Journal of Operational Research* 108, 1–15.
- Nagy, G., Salhi, S., 2007. Location-routing: issues, models and methods. *European Journal of Operational Research* 177, 649–672.
- Nozick, L.K., 2001. The fixed charge facility location problem with coverage restrictions. *Transportation Research Part E* 37, 281–296.
- Osman, I.H., Wassan, N.A., 2002. A reactive tabu search meta-heuristic for the vehicle routing problem with back-hauls. *Journal of Scheduling* 5, 263–285.
- Owen, S.H., Daskin, M.S., 1998. Strategic facility location: a review. *European Journal of Operational Research* 111, 423–447.
- Pisinger, D., 1997. A minimal algorithm for the 0–1 knapsack problem. *Operations Research* 45, 758–767.
- Ro, H., Tcha, D., 1984. A branch and bound algorithm for the two-level uncapacitated facility location problem with some side constraints. *European Journal of Operational Research* 18, 349–358.
- Toregas, C., Swain, R., ReVelle, C., Bergman, L., 1971. The location of emergency service facilities. *Operations Research* 19, 1363–1373.
- Tragantalerngsak, S., Holt, J., Rönnqvist, M., 1997. Lagrangian heuristics for the two-echelon, single-source, capacitated facility location problem. *European Journal of Operational Research* 102, 611–625.
- Webb, M.H.J., 1968. Cost functions in the location of depots for multiple-delivery journeys. *Operational Research Quarterly* 19, 311C320.
- Yang, Z., Chen, H., Chu, F., 2007. Facility location problem with minimum capacity constraints. In: *International Conference on Industrial Engineering and Systems Management*, IESM.