

MANFRED KÖRKEL

**Discrete facility location with nonlinear facility costs**

*RAIRO. Recherche opérationnelle*, tome 25, n° 1 (1991), p. 31-43

[http://www.numdam.org/item?id=RO\\_1991\\_\\_25\\_1\\_31\\_0](http://www.numdam.org/item?id=RO_1991__25_1_31_0)

© AFCET, 1991, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## DISCRETE FACILITY LOCATION WITH NONLINEAR FACILITY COSTS (\*)

by Manfred KÖRKEL <sup>(1)</sup>

---

*Abstract. — We treat a discrete uncapacitated facility location problem where the installation cost per facility is a general nonlinear function of the number of established facilities. Assuming a convexity property, Mirchandani and Jagannathan developed a heuristic solution procedure for this problem. Here, we show how to modify their algorithm in order to get and prove exact solutions and, at the same time, to improve the computational efficiency.*

Keywords : Location; integer programming; Lagrange relaxation.

*Résumé. — On traite d'un problème discret lié au choix d'emplacements d'installations sans restriction de capacité où le coût d'établissement par installation est une fonction non linéaire générale du nombre total des installations existantes. En admettant une propriété de convexité, Mirchandani et Jagannathan ont trouvé une solution heuristique à ce problème. Ici, il est montré comment leur algorithme peut être modifié pour obtenir et prouver des solutions exactes et comment l'efficacité de calcul peut être améliorée.*

### 1. INTRODUCTION

Solving the discrete “uncapacitated facility” (or “simple plant”) location problem (UFLP), it is our intention to reach the minimum total cost by balancing two different cost types. The first of these types is caused by the individual supply of demands, while the second of them concerns the fixed charges of establishing facilities. In the usual case (see [13] for a survey), each fixed cost part depends only on the location of the corresponding facility. Mirchandani and Jagannathan [16] introduced to the standard problem a “fixed” cost component which is a convex function of the number of established facilities. To give an example, the authors referred to the situation where each “facility” coincides with a copy of a data file in a computer network. There, the cost of updating the data in the copies generally increases

---

(\*) Received April 1990.

(1) Forschungsinstitut der D.B.P.-Telekom, Postfach 10 00 03, D-6100 Darmstadt.

nonlinearly and rapidly as the number of copies grows. We here hint at an application with respect to the field of telecommunications. In a so-called Intelligent Network, each user must have access to a Service Control Point which contains customer data. The necessary administration of the customer data within the Service Control Points then creates the above-mentioned nonlinear cost component.

Mirchandani and Jagannathan attempted to tackle the case with convex and non-decreasing nonlinear facility costs by developing a bisection heuristic called DATALOC which decomposes the problem into a series of UFLPs. We agree to their main idea that it is useful to take advantage of the progress which, as far as the solution of the UFLP is concerned, has led to a significant increase of efficiency. A view on the literature [2, 7, 12] shows that the computer times necessary for determining exact UFLP solutions have been reduced by at least two orders of magnitude when large data sets with several hundred potential locations are treated. Indeed, we will show that, without changing this subproblem structure, we may construct an improved and exact algorithm which solves a general version of the nonlinear problem.

## 2. PROBLEM FORMULATION

We introduce the indices  $j \in J$  and  $i \in I$  for the demands and the potential facility sites, respectively. In the following, we will always omit the index sets to get a shorter notation. The decision variable  $y_i$  indicates whether a facility at site  $i$  exists ( $y_i = 1$ ) or not ( $y_i = 0$ ) while  $x_{ij}$  designates the point  $j$ 's demand being satisfied by facility  $i$  ( $x_{ij} = 1$ ) or not ( $x_{ij} = 0$ ). Furthermore, we define the feasible regions

$$R = \{ X, Y \mid \sum_i x_{ij} = 1; x_{ij} \leq y_i; x_{ij}, y_i \in \{0, 1\} \}, \quad (1)$$

$$Q(p) = R \cap \{ X, Y \mid \sum_i y_i = p \} \quad (2)$$

where  $X$  and  $Y$  are equivalent to the matrix  $(x_{ij})$  and the vector  $(y_i)$ , respectively.  $c_{ij} \in \mathbb{R}$  denotes a transportation cost coefficient,  $f_i \in \mathbb{R}$  corresponds to fixed charges for establishing a facility at site  $i$  and  $p \in \{1, \dots, m = |I|\}$  designates the number of installed facilities. With these definitions, we formulate the "generalized" [5, p. 111] or "fixed cost"  $p$ -median problem

$$\omega(p) = \min_{X, Y \in Q(p)} \sum_i \sum_j c_{ij} x_{ij} + \sum_i f_i y_i. \quad (3)$$

Finally, a nonlinear cost component for opening a facility is introduced by a given function  $g$  depending on the number  $p$  of facilities opened. Setting  $\sigma(p) = pg(p)$  enables us to state the location problem with nonlinear facility costs (NFCLP) in a concise way:

$$z_{\text{opt}} = \min_p \{ z(p) = \omega(p) + \sigma(p) \}. \quad (4)$$

### 3. LOWER AND UPPER BOUNDS TO THE OBJECTIVE

What we need for determining the true optimum of NFCLP are converging lower and upper bounds to the objective. Since  $\sigma(p)$  is given, we only have to develop bounds for  $\omega(p)$ .

To construct the lower bound to  $\omega(p)$ , we dualize the restriction on the number of facilities  $\sum_i y_i = p$  and introduce the corresponding Lagrange multiplier  $\gamma \in \mathbb{R}$ . With

$$\psi(\gamma) = \min_{X, Y \in R} \sum_i \sum_j c_{ij} x_{ij} + \sum_i (f_i + \gamma) y_i \quad (5)$$

we then get

$$\psi(\gamma) - p\gamma \leq \omega(p). \quad (6)$$

To calculate  $\psi(\gamma)$ , we have to solve an UFLP with the added "artificial" fixed cost component  $\gamma$ . The best lower bound to (4) which may be generated from  $\psi(\gamma)$  is then given by

$$\max_{\gamma} \min_p \psi(\gamma) - p\gamma + \sigma(p) \leq z_{\text{opt}}. \quad (7)$$

Let  $q(\gamma)$  denote the number of opened facilities which corresponds to one member of the possibly non-unique set of optimal solutions of  $\psi(\gamma)$ . From  $Q(q(\gamma)) \subseteq R$ , we get

$$\omega(q(\gamma)) = \psi(\gamma) - q(\gamma)\gamma \quad (8)$$

which immediately yields

$$z_{\text{opt}} \leq \min_{\gamma} \psi(\gamma) - q(\gamma)\gamma + \sigma(q(\gamma)) \quad (9)$$

as the –from the viewpoint of our UFLP approach– best available upper bound to the objective (4).

#### 4. SOLUTION APPROACH

In the following, it is our task to create a procedure that repeatedly solves  $\psi(\gamma)$  with the aim of optimizing (7) and (9) simultaneously. The technique of determining a solution with a prespecified number of established facilities (*i.e.* solving a “ $p$ -median” problem) by varying the fixed costs of UFLPs has been proposed several times in literature [3, 6, 11, 15, 17]. To obtain a certain  $p$  value, two papers [11, 17] independently suggested a scheme for updating  $\gamma$ . As it will be shown below, this scheme, which the publications [4, 10] already had treated in a more general context, also proves to be useful for NFCLP.

For solving NFCLP, we subsequently divide the closed interval  $[1, m]$  into subintervals and treat the latter separately. As subinterval bounds, we evaluate integer numbers  $a$  and  $b \geq a + 2$ . Assume that  $\omega(a)$  and  $\omega(b)$  have been generated by (8) with  $\gamma = \alpha$  and  $\gamma = \beta < \alpha$ , respectively. Then (6) provides a convex underestimation of  $\omega$  which consists of two line segments with slopes  $-\alpha$  and  $-\beta$  (see fig. 1).

The aim is now to improve the lower bounding function within the subinterval  $[a, b]$  as well as to create a new solution  $\omega(p)$  with  $a < p < b$ . With respect to this intention, the previously mentioned update scheme for  $\gamma$  proposes to solve  $\psi(\gamma)$  with

$$\gamma = (\omega(a) - \omega(b)) / (b - a) \quad (10)$$

Two situations may occur:

- If function values of  $\omega$  lie within the region limited by the lower bounding line segments and the line connecting  $\omega(a)$  and  $\omega(b)$  but excluding the latter (hatched region of figure 1), a new  $\omega(q(\gamma))$ ,  $a < q(\gamma) < b$ , will be determined. By correspondingly applying (10) to the two newly generated subintervals, further function values  $\omega$  may principally be detected if  $q(\gamma) - a \geq 2$  and/or  $b - q(\gamma) \geq 2$ . We may characterize this case also as having maximized the lower bound at  $p = q(\gamma)$ .

- If this region does not contain values of  $\omega$ , we will find  $\omega(q(\gamma))$ ,  $a \leq q(\gamma) \leq b$ , on the line connecting  $\omega(a)$  and  $\omega(b)$ . Such a result proves that no further  $\omega(p)$ ,  $a < p < b$ , can be determined by solving an UFLP. (Note that we always assume a deterministic behaviour of the used UFLP algorithm.) On

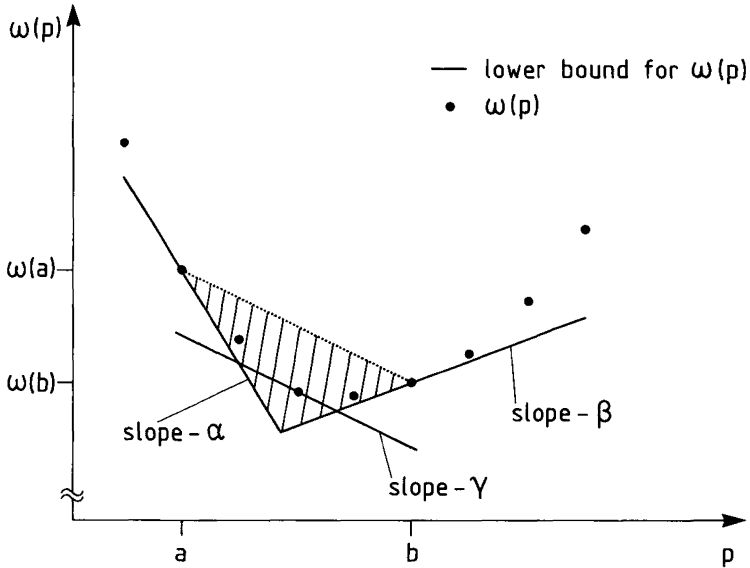


Figure 1.

the other hand, this situation implies that the maximum possible lower bound for the whole interval  $[a, b]$  has been calculated (the dotted line of figure 1).

Thus, we may state that, for determining the maximum piece-wise linear underestimation of  $\omega$ , at most  $(m-2)$  UFLPs have to be solved if we start with the (trivial) solutions  $\omega(1)$  and  $\omega(m)$  and apply the above  $\gamma$  update scheme subsequently. We define for subsequent use:

A scheme for calculating  $\gamma_1$  needed to generate a lower bounding function for  $\omega$  in the interval  $[a, b]$  by solving  $\psi(\gamma_1)$  has the *maximum lower bound property* if there exists a  $p$ ,  $a < p < b$ , with  $\psi(\gamma_1) - p\gamma_1 = \max_{\gamma} \psi(\gamma) - p\gamma$  or if, for each  $p$ ,  $a < p < b$ , it is  $\psi(\gamma_1) - p\gamma_1 \geq \tilde{\omega}(p)$  where  $\tilde{\omega}(p)$  is a given upper limit to  $\omega(p)$ .

### 5. BASIC SOLUTION PROCEDURE

Using the considered method for generating the subproblems  $\psi(\gamma)$ , we are now able to construct the algorithm which solves NFCLP. In order to formulate this procedure, here called NFCLOC, we define:

$l, p, p_+$ , variables for the number of opened facilities;  
 $\underline{z}(p)$ , lower bound to  $z(p)$ ;

$\bar{z}$ ,	upper bound to $z_{\text{opt}}$ ;
$r, s, t$ ,	subinterval indices;
$a_s, b_s$ ,	subinterval bounds;
$z_s$ ,	$\min_{a \leq p \leq b} z(p)$ ;
$S$ ,	index set of subintervals to be treated.

We, thus, state:

#### Procedure NFCLOC

- (a1)  $S := \{1\}$ ;  $t := 1$ ;  $z_t := -\infty$ ;  $a_t := 1$ ;  $b_t := m$   
(a2) for  $p = 1$  to  $m$  do:  $\underline{z}(p) = -\infty$   
(a3) calculate  $\omega(1)$ ;  $z(1) := \omega(1) + \sigma(1)$   
(a4) calculate  $\omega(m)$ ;  $\bar{z}(m) := \omega(m) + \sigma(m)$   
(a5)  $\bar{z} := \min\{z(1), \bar{z}(m)\}$   
(b1) while  $S \neq \emptyset$  do:  
(b2) get  $s$  with  $z_s = \min\{z_r | r \in S\}$ ;  $S := S \setminus \{s\}$   
(b3) if  $z_s \geq \bar{z}$  then: leave "while...do" loop  
(b4)  $\gamma := (\omega(a_s) - \omega(b_s)) / (b_s - a_s)$   
(b5) calculate  $\psi(\gamma)$ ;  $p_+ := q(\gamma)$ ;  $\omega(p_+) := \psi(\gamma) - p_+ \gamma$   
(b6) for  $p = a_s$  to  $b_s$  do:  $z(p) := \max\{z(p), \psi(\gamma) - p \gamma + \sigma(p)\}$   
(b7)  $\bar{z} := \min\{\bar{z}, z(p_+)\}$   
(b8) if  $a_s < p_+ < b_s$  then:  
(b9)  $t := t + 2$ ;  $S := S \cup \{t-1, t\}$   
(b10)  $z_{t-1} := \min\{z(p) | p = a_s \text{ to } p_+\}$   
(b11)  $z_t := \min\{z(p) | p = p_+ \text{ to } b_s\}$   
(b12)  $a_{t-1} := a_s$ ;  $b_{t-1} := p_+$ ;  $a_t := p_+$ ;  $b_t := b_s$   
(c1) get  $p$  with  $z(p) = \min\{z(l) | l = 1 \text{ to } m\}$   
(c2) while  $z(p) < \bar{z}$  do:  
(c3) calculate  $\omega(p)$ ;  $z(p) := \omega(p) + \sigma(p)$ ;  $\bar{z} := \min\{\bar{z}, z(p)\}$   
(c4) get  $p$  with  $\underline{z}(p) = \min\{\underline{z}(l) | l = 1 \text{ to } m\}$

In the initialization phase (a), we set up the interval within which  $\omega$  is defined, compute the trivial solutions  $\omega(1)$  and  $\omega(m)$  and calculate the initial lower and upper bounds. In the main loop (b), the intervals stored in  $S$  are treated subsequently. As interval selection rule, we suggest the minimum lower bound strategy (b2) with the intention of accelerating the convergence of the lower and the upper bound to the objective. For each interval, the procedure computes a new Lagrangian multiplier  $\gamma$  (b4) and solves the corresponding UFLP (b5) using any exact algorithm, e. g. [7, 12]. Steps (b6)-(b7) update the bounds. If the UFLP algorithm has determined a solution with a new  $p$  value, the generated subintervals are stored for further treatment in (b8)-(b12). [Note: If we prove for any  $\gamma$  and  $p$  with  $p \neq q(\gamma)$  that  $\omega(p) = \psi(\gamma) - p \gamma$ , it is not necessary to store the subinterval with the bounds  $p$  and  $q(\gamma)$  within  $S$ . To simplify the procedure, this collinearity test has been omitted.] NFCLOC terminates loop (b) if either  $S$  is empty (b1) or all lower bounds which correspond to the intervals contained in  $S$  exceed or equal the current upper bound  $\bar{z}$  (b3). The latter step also avoids that subintervals of length 1 are processed unnecessarily.

It is easy to show that, if  $\omega$  can be embedded within a strictly convex function, phase (b) ends up with the exact optimum, the function value of which is given by  $\bar{z}$ . If, on the other hand, this property does not hold, it depends on  $\sigma$  whether, in line (c1), there remain unresolved subproblems  $z(p)$  with  $z(p) \leq \bar{z}$ . In the latter case, we have to solve  $\omega(p)$  directly in (c2)-(c4). For this problem, several branch-and-bound algorithms have been proposed in literature: [9] is explicitly addressed to a fixed cost  $p$ -median problem, while the solution methods [1, 6, 8, 11] have to be modified slightly by introducing the facility costs  $f_i$ . However, the reported experience [11, 15, 17] shows that the non-convex case occurs rather seldom when practical data is processed. So we can expect that, in the most situations, NFCLP will already be solved during phase (b).

## 6. MODIFICATIONS OF THE BASIC PROCEDURE

For partitioning the search intervals, the basic version of NFCLOC uses only information on  $\omega(a)$  and  $\omega(b)$ . Especially at the first iteration, this may lead to poor bounds on the objective while, at subsequent iterations, the algorithm may unnecessarily fail to compute better bounds. To improve NFCLOC with respect to these shortcomings, we here propose

- to calculate a more “intelligent”  $\gamma$  at the first iteration by using additional information on  $\omega$  and  $\sigma$  and
- to apply a modified  $\gamma$  update scheme at subsequent iterations, which is able to exploit previously computed small gaps between the lower and the upper bounds.

### 6.1. Calculating the initial $\gamma$

We first state two useful properties. Note that here we refer to a discrete function as being “convex” if, for each point within the domain of definition, there exists a global subgradient [17, p. 129].

PROPERTY 1: If  $\gamma = \min_p g(p)$  and  $g(q(\gamma)) = \gamma$ , we get  $\underline{z}(p) \geq \psi(\gamma) = z_{\text{opt}}$  for  $p = 1, \dots, m$ .

PROPERTY 2: If  $\omega$  is strictly convex,  $g_{\min} = \min_p g(p)$  and  $g_{\max} = \max_p g(p)$ , there exists always an  $\varepsilon > 0$  such that, for all functions  $g$  with  $g_{\max} - g_{\min} \leq \varepsilon$ , the optimum of NFCLP may be proved by solving at most  $\psi(\gamma)$  with  $\gamma = g_{\min}$ ,  $g_{\min} + m\varepsilon$  and  $g_{\min} - m\varepsilon$ . For convex functions  $\sigma$  with  $\sigma(p) = pg(p)$  at most one of the two latter problems has to be solved.

Property 1 follows directly from (6) and (8). To prove property 2, we have to show that values of  $\omega$  which may improve the resulting upper bound  $z(q(g_{\min}))$  lie within the region between two lines of the slope  $-g_{\min}$  and that the distance between these lines is bounded by  $m\varepsilon$ . If we set  $k = q(g_{\min})$ , this can be done by using the conditions  $\omega(p) \geq \omega(k) + kg_{\min} - pg_{\min}$  and  $\omega(p) \leq \omega(k) + kg(k) - pg(p)$  which are derived from (6) and (8), respectively.

The properties illustrate that initializing  $\gamma$  with  $g_{\min}$  may be efficient in the sense that only few iterations may be necessary to prove the optimum. If  $g(p)$  is constant, property 1 clearly applies and we have to solve only one UFLP (the problem then simply reduces to an UFLP).

To decide whether or not starting with  $g_{\min}$  is useful if  $g(p)$  varies with  $p$ , we have to know more about  $\omega$ . Estimated values of  $\omega$  may quickly be gained



from upper bounds  $\bar{\omega}(p) \geq \omega(p)$  which may e. g. be calculated by the well-known ADD heuristic ([14], see [18] for a fast variant).  $\bar{\omega}(p)$  immediately yields an upper bound  $\bar{z}(p) = \bar{\omega}(p) + \sigma(p)$  to  $z(p)$ . Let  $\bar{z}(p') = \min_p \bar{z}(p)$  and assume for a moment  $1 < p' < m$ . We then denote  $\alpha = \bar{\omega}(p') - \bar{\omega}(p'+1)$  and  $\beta = \bar{\omega}(p'-1) - \bar{\omega}(p')$ . Note that  $\bar{\omega}$  is convex if it is derived from the ADD heuristic, i. e.  $\alpha \leq \beta$ . We propose to initialize  $\gamma$  with  $g_{\min}$  if

$$2p'(g_{\max} - g_{\min}) \leq \beta - \alpha \quad \text{or} \quad g_{\min} \in [\alpha, \beta]. \quad (11)$$

The first test attempts to detect whether or not  $g$  influences significantly the position of the optimum, while the latter test estimates whether  $g_{\min}$  is likely to yield  $q(g_{\min}) = p'$ .

Let us demonstrate the usefulness of the above test by an example with  $\sigma(p) = 0$ ,  $1 \leq p \leq \hat{p}$ , and  $\sigma(p) = L$ ,  $\hat{p} < p \leq m$ , where  $L$  is sufficiently high to exclude the region  $\hat{p} < p \leq m$ . If  $q(0) \leq \hat{p}$ , we are very likely to get  $g(p') = 0$  which means that (11) will be satisfied. We, then, will immediately prove the optimum by solving  $\psi(0)$ . If, however,  $q(0) \gg \hat{p}$ , it is almost sure that (11) will be violated. In this case solving  $\psi(0)$ , in fact, is not a good choice because the resulting upper bound  $\psi(0) + L$  is very poor.

In the case where (11) is violated, we suggest to initialize  $\gamma$  with an estimation of the value which computes  $\omega$  at the point  $p'$  of the minimum upper bound (see figure 2):

$$\gamma' = (\alpha + \beta)/2 \quad (12)$$

Finally, we compute the closed interval  $[\gamma_{\min}, \gamma_{\max}]$  where

$$\gamma_{\min} = \min_{1 \leq p \leq m-1} (\bar{\omega}(p) - \bar{\omega}(m))/(m-p), \quad (13)$$

$$\gamma_{\max} = \max_{2 \leq p \leq m} (\bar{\omega}(1) - \bar{\omega}(p))/(p-1). \quad (14)$$

If we project the chosen value of  $\gamma$  (either  $g_{\min}$  or  $\gamma'$ ) into this interval [i. e.  $\gamma := \max(\min(\gamma, \gamma_{\max}), \gamma_{\min})$ ], the maximum lower bound property is guaranteed for this initial iteration (see fig. 2). If  $p' = 1$  or  $p' = m$ , we set  $\gamma$  to  $\gamma_{\max}$  or  $\gamma_{\min}$ , respectively.

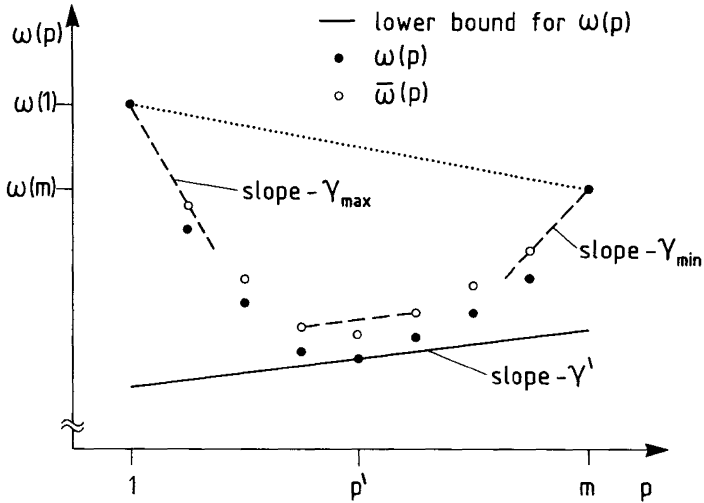


Figure 2.

**6.2. Updating  $\gamma$  in subsequent iterations**

The rule (10) does not always lead to an improved lower bound to the objective whenever this would be possible. Figure 3 gives an illustrative example of such a situation.  $\tilde{\omega}(p) = \min(\bar{\omega}(p), \bar{z} - \sigma(p))$  denotes the upper bound to values of  $\omega$  which are able to improve  $\bar{z}$ . We see that the minimum lower bound is located at  $p^*$ . The multiplier value  $\gamma$  derived from (10) yields  $\omega(p^* + 1) > \tilde{\omega}(p^* + 1)$  which improves neither the lower nor the upper bound. However, choosing  $\gamma^*$  immediately generates the improved objective value  $z(p^*)$  since  $\omega(p^*) < \tilde{\omega}(p^*)$ .

In the modified procedure, we start each iteration by calculating the interval  $[\gamma_{min}, \gamma_{max}]$  within which  $\gamma$  may vary without violating the maximum lower bound property. Define  $\underline{\omega}(p) = \underline{z}(p) - \sigma(p)$ . Let  $\tilde{P} = \{p \in \mathbb{N} \mid a < p < b, \tilde{\omega}(p) > \underline{\omega}(p)\}$  be the domain of the search interval  $[a, b]$  where improved function values of  $z$  may still be discovered. We then compute

$$\gamma_{min} = \max_{p \in \tilde{P}} (\tilde{\omega}(p) - \omega(b)) / (b - p), \tag{15}$$

$$\gamma_{max} = \min_{p \in \tilde{P}} (\omega(a) - \tilde{\omega}(p)) / (p - a). \tag{16}$$

If  $\gamma_{min} \geq \gamma_{max}$  [i. e. the interval is empty since no value of  $\tilde{\omega}$  remains under the line connecting  $\omega(a)$  and  $\omega(b)$ ], we set  $\gamma$  to the standard value (10).

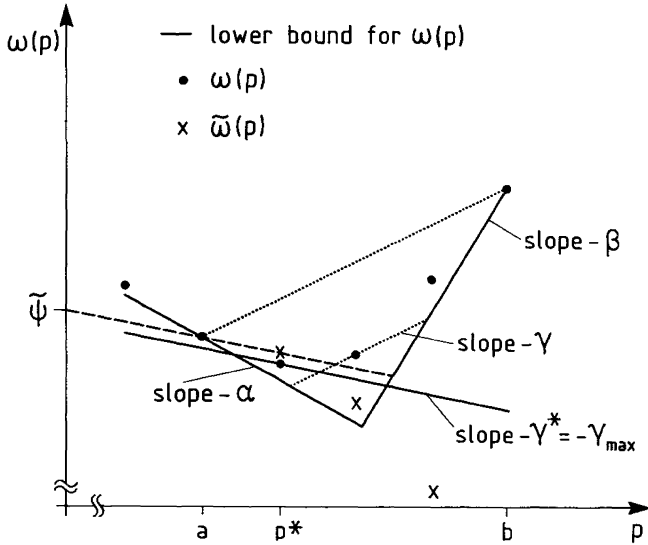


Figure 3.

If  $\gamma_{\min} < \gamma_{\max}$ , we compute the negative slope  $(\omega(p^* - 1) - \omega(p^* + 1))/2$  of the current lower bounding function  $\omega$  at  $p^*$  where  $\underline{z}(p^*)$  is the minimum lower bound to the objective. Projecting this slope into the interval  $[\gamma_{\min}, \gamma_{\max}]$  yields the modified multiplier  $\gamma^*$  and ensures the maximum lower bound property (see figure 3 where  $\gamma^* = \gamma_{\max}$ ).

With these measures, we try to reduce the size of the search region and, thus, to avoid unnecessary iterations. An acceleration of the convergence may especially be expected if the search interval is large while the reached gap between the lower and the upper bounds to the objective is small. We, finally, remark that, for the objective of the subproblems  $\psi(\gamma)$ , it is possible to compute an *a priori* upper bound  $\tilde{\psi}(\gamma)$  above which an improvement of  $\bar{z}$  is impossible (see also figure 3):

$$\tilde{\psi}(\gamma) = \max_{p \in \bar{P}} \min(\tilde{\omega}(p) + p\gamma, \omega(a) + \alpha\gamma, \omega(b) + b\gamma) \quad (17)$$

The called UFLP algorithm may internally use this upper bound to reduce the computational effort of determining the UFLP solution by facilitating variable fixations or diminishing the size of the branch-and-bound tree.

## 7. COMPARISON OF NFCLOC WITH DATALOC

Mirchandani and Jagannathan assumed that  $g$  is convex and  $\omega$  is strictly convex. The authors, moreover, required that  $g$  is non-decreasing and  $g(0) = 0$ .

If we analyze the influence of the convexity assumptions on NFCLOC's behaviour, we find that NFCLOC always provides  $a_s < p_+ < b_s$  in step (b8) and  $z_{t-1} \geq \bar{z}$  or  $z_t \geq \bar{z}$  in lines (b10)-(b11). We see that, in this case, NFCLOC operates like an interval reduction method where, at each iteration, exactly one UFLP is solved which always leads to a new candidate solution  $z(p)$ . For ensuring this behaviour, it is even sufficient that  $\sigma$  instead of  $g$  is convex.

To discuss some aspects of the behaviour of DATALOC, we state the procedure, using our notation ( $[u] = \max \{ n \in \mathbb{N} \mid n \leq u \}$ ):

Procedure DATALOC

- (a)  $a := 1$ ; calculate  $\psi(0)$ ;  $b := q(0)$
- (b) if  $b = a$  then:  $\bar{z} = \psi(0)$ ; stop
- (c1) while  $b - a \neq 1$  do:
- (c2)  $p_0 := \lfloor (a + b) / 2 \rfloor$ ;  $\gamma := g(p_0)$
- (c3) calculate  $\psi(\gamma)$ ;  $p_+ := q(\gamma)$
- (c4) if  $p_+ = p_0$  then:  $z = \psi(\gamma)$ ; stop
- (c5) else if  $p_+ < p_0$  then:  $b := p_0$
- (c6) else:
- (c7) calculate  $\omega(p_0)$  and  $\omega(p_0 + 1)$
- (c8) if  $\omega(p_0) + \sigma(p_0) \geq \omega(p_0 + 1) + \sigma(p_0 + 1)$  then:  $a := p_0$
- (c9) else:  $b := p_0 + 1$
- (d) calculate  $\omega(a)$  and  $\omega(b)$ ;  $\bar{z} := \min(\omega(a) + \sigma(a), \omega(b) + \sigma(b))$

We have to note here that DATALOC applies the previously mentioned technique [11, 17] of computing  $\omega(p)$  by decomposing (3) into a series of UFLPs. Contrary to this, NFCLOC deals directly with problem (3) only if a duality gap occurs which implies that, in this case, the DATALOC approach to computing  $\omega(p)$  would fail.

The underlying principle of DATALOC is a bisection method. At each iteration within the main loop, the considered interval  $[a, b]$  is halved. Exploiting the convexity assumption, DATALOC then determines the half interval containing the optimum number of facilities. Additionally, two termination tests (b), (c4) attempt to stop the program during the search process.

A comparison of DATALOC with NFCLOC immediately reveals some drawbacks of the first procedure:

- It cannot be excluded that, in step (c3), DATALOC computes the same value of  $p_+$  several times because different values of  $\gamma$  may produce the same number of facilities. Also it seems to be possible that  $p_+ < a$  or  $p_+ > b$  which, in a certain sense, provides "irrelevant" information. The case  $g(p) = 0$ ,  $1 \leq p < \bar{p}$  and  $g(p) = (p - \bar{p})L$ ,  $\bar{p} < p \leq m$ , yields such an example where DATALOC always computes either  $p_+ = q(0)$  or  $p_+ = 1$  if  $L$  is sufficiently high. Note that, for this example, the enhanced version of NFCLOC may prove the exact optimum by solving only one UFLP if  $g(q(0)) = 0$ .

- The calculated function values  $\psi(\gamma)$  are not used explicitly, neither for creating bounds to the optimal objective value of NFCLP nor for directly reducing the size of the interval  $[a, b]$ . If e.g.  $a < p_+ < b$ , always either the lower bound  $a$  or the upper bound  $b$  to the optimal number of facilities could immediately be set to  $p_+$ .

- If  $p_+ > p_0$ , DATALOC has to solve several UFLPs for just one interval reduction when it calculates function values of  $\omega$  in step (c7). For weakening to some extent the negative consequences of this method for the algorithmic efficiency, the authors only mentioned a possible reuse of previously calculated function values of  $\omega$  if DATALOC again has to evaluate  $\omega$  at prespecified points. On the other hand, the approach which DATALOC uses for calculating

$\omega(p_0)$  and  $\omega(p_0+1)$  is applied by NFCLOC in step (b4) to solve the complete problem. Therefore, NFCLOC, in total, needs roughly the computing time which DATALOC spends on only one iteration passing step (c7).

• The fact that the condition for the termination test in (c4) is satisfied does not guarantee that the optimal solution has been found. This may be proved by the following counterexample:

$p$	$\omega(p)$	$(\gamma_1(p), \gamma_2(p))$	$g(p)$	$z(p)$
1	35	$(15, \infty)$	0	35
2	20	$(10, 15)$	2	24
3	10	$(6, 10)$	7	31
4	4	$(2, 6)$	13	56
5	2	$(-\infty, 2)$	20	102

The open interval  $(\gamma_1(p), \gamma_2(p))$  represents the domain for  $\gamma$  which yields exactly  $p$  installed facilities. In the above table, we obtain the optimal objective value for NFCLP at  $p=2$ , while  $g(g(3))=3$ . The stopping rule (c4) can be regarded as an infeasible generalization of property 1.

## 8. CONCLUSIONS

We have developed an improved exact algorithm, called NFCLOC, for solving NFCLP. Like the procedure DATALOC published in [16], NFCLOC decomposes the problem into a series of UFLPs. Compared to the first algorithm, NFCLOC possesses the following merits:

- The restriction on the facility costs to be a convex and non-decreasing function is dropped. NFCLOC is able to process any facility cost function.
- In the convex case (to which DATALOC is restricted), NFCLOC operates more efficiently than DATALOC.
- The new procedure determines the true optimum as far as the subproblems are exactly solved. Even when the program terminates prematurely, a lower and an upper bound to the objective is always provided if, at least, one UFLP has been solved.

Furthermore, we made several suggestions how to improve the basic version of NFCLOC in order to get a better initialization and a quicker convergence of the lower and the upper bounds to the objective.

## REFERENCES

1. J. E. BEASLEY, A Note on Solving Large  $p$ -Median Problems, *European Journal of Operational Research*, 1985, 21, pp. 270-273.
2. O. BILDE and J. KRARUP, Sharp Lower Bounds and Efficient Algorithms for the Simple Plant Location Problem, *Annals of Discrete Mathematics*, 1977, 1, pp. 79-97.

3. T. B. BOFFEY and J. KARKAZIS,  $p$ -Medians and Multi-Medians, *Journal of the Operations Research Society*, 1984, 35, pp. 57-64.
4. E. W. CHENEY, Introduction to approximation theory, *McGraw-Hill*, New York, 1966.
5. N. CHRISTOFIDES, Graph Theory – an Algorithmic Approach, *Academic Press*, London, 1975.
6. N. CHRISTOFIDES and J. E. BEASLEY, A Tree Search Algorithm for the  $p$ -Median Problem, *European Journal of Operational Research*, 1982, 10, pp. 196-204.
7. D. ERLKOTTER, A Dual-Based Procedure for Uncapacitated Facility Location, *Operations Research*, 1978, 26, pp. 992-1009.
8. R. D. GALVÃO, A Dual-Bounded Algorithm for the  $p$ -Median Problem, *Operations Research*, 1980, 28, pp. 1112-1121.
9. R. D. GALVÃO and L. A. RAGGI, A Method for Solving to Optimality Uncapacitated Location Problems, *Annals of Operations Research*, 1989, 18, pp. 225-244.
10. H. J. GREENBERG, The One-Dimensional Generalized Lagrange Multiplier Problem, *Operations Research*, 1977, 25, pp. 338-345.
11. P. HANJOUL and D. PEETERS, A Comparison of Two Dual-Based Procedures for Solving the  $p$ -Median Problem, *European Journal of Operational Research*, 1985, 20, pp. 387-396.
12. M. KÖRKEL, On the Exact Solution of Large-Scale Simple Plant Location Problems, *European Journal of Operational Research*, 1989, 39, pp. 157-173.
13. J. KRARUP and P. M. PRUZAN, The Simple Plant Location Problem: Survey and Synthesis, *European Journal of Operational Research*, 1983, 12, pp. 36-81.
14. A. A. KUEHN and M. J. HAMBURGER, A Heuristic Program for Locating Warehouses, *Management Science*, 1963, 9, pp. 643-666.
15. L. P. MAVRIDES, An Indirect Method for the Generalized  $k$ -Median Problem Applied to Lock-Box Location, *Management Science*, 25, pp. 990-996.
16. P. B. MIRCHANDANI and R. JAGANNATHAN, Discrete Facility Location with Non-Linear Diseconomies in Fixed Costs, *Annals of Operations Research*, 1989, 18, pp. 213-224.
17. P. B. MIRCHANDANI, A. OUDJIT and R. T. WONG, "Multidimensional" Extensions and a Nested Dual Approach for the  $m$ -Median Problem, *European Journal of Operational Research*, 1985, 21, pp. 121-137.
18. R. A. WHITAKER, A Fast Algorithm for the Greedy Interchange for Large-Scale Clustering and Median Location Problems, *Canadian Journal of Operations Research and Information Processing*, 1983, 21, pp. 95-108.