



On the distribution approach to location problems

Włodzimierz Ogryczak*

Warsaw University of Technology, Institute of Control & Computation Engineering, 00-665 Warsaw, Poland

Abstract

While making location decisions, the distribution of travel distances among the service recipients (clients) is an important issue. It is usually tackled with the minimax (center) or the minisum (median) solution concepts. Both concepts minimize only simple scalar characteristics of the distribution: the maximal distance and the average distance, respectively. In this paper, all the distances for the individual clients are considered as a set of multiple uniform criteria to be minimized. This results in a multiple criteria model taking into account the entire distribution of distances. Our analysis of the multiple criteria problem focuses on the symmetrically efficient solutions which comply with minimization of distances as well as with impartial consideration of the clients. Various solution concepts generating symmetrically efficient location patterns are analyzed. Finally, the reference distribution approach is developed as an interactive technique which enables us to identify a satisfactory symmetrically efficient location pattern by evolving a reference (target) distribution of distances. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Location problems; Multiple criteria; Symmetric efficiency; Interactive methods; Reference point

1. Introduction

A host of operational models has been developed to deal with the facility location optimization (cf. [1–3]). Most classical location studies focus on some aspects of two major approaches: the minimax (center) or the minisum (median) solution concepts. Both concepts minimize only simple scalar characteristics of the distribution: the maximal distance and the average distance, respectively. In this paper all the distances for the individual clients are

* Tel.: +48-22-658-3165; fax: +48-22-658-3164.

E-mail address: ogryczak@mimuw.edu.pl (W. Ogryczak).

considered as the set of multiple uniform criteria to be minimized. This results in a multiple criteria model taking into account the entire distribution of distances.

The generic location problem that we consider may be stated as follows. There is given a set of m clients (service recipients). Each client is represented by a specific point. There is also given a set of n potential locations for the facilities. It may be, in particular, a subset (or the entire set) of points representing the clients. Further, the number (or the maximal number) p of facilities to be located is given ($p \leq n$). Thus, we limit our discussion to discrete location problems [4]. They can be viewed, however, as network location problems with possible locations restricted to some subset of the network vertices [5].

The main decisions to be made in the location problem can be described with the binary variables: x_j is equal to 1 if location j is to be used and equal to 0 otherwise ($j = 1, 2, \dots, n$). To meet the problem requirements, the decision variables x_j have to satisfy the following constraints:

$$\sum_{j=1}^n x_j = p, \quad x_j \in \{0, 1\}, \quad \text{for } j = 1, 2, \dots, n, \quad (1)$$

where the equation is replaced with the inequality (\leq) if p specifies the maximal number of facilities to be located. Note that constraints (1) take a very simple form of the binary knapsack problem with all the constraint coefficients equal to 1. However, for most location problems the feasible set has a more complex structure due to explicit consideration of allocation decisions. These decisions are usually modeled with the additional allocation variables: x'_{ij} is equal to 1 if location j is used to service client i and equal to 0 otherwise ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$).

The allocation variables have to satisfy the following constraints:

$$\sum_{j=1}^n x'_{ij} = 1, \quad \text{for } i = 1, 2, \dots, m. \quad (2)$$

$$x'_{ij} \leq x_j, \quad \text{for } i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n. \quad (3)$$

$$x'_{ij} \in \{0, 1\}, \quad \text{for } i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n. \quad (4)$$

In the capacitated location problem, for each client the demand for services is defined with the corresponding coefficient v_i and the capacities of the potential facilities are given as q_j (for $j = 1, 2, \dots, n$). This implies the additional constraints:

$$\sum_{i=1}^m v_i x'_{ij} \leq q_j, \quad \text{for } j = 1, 2, \dots, n.$$

Let us assume that for each client i ($i = 1, 2, \dots, m$) a function $f_i(\mathbf{x})$ of the location pattern \mathbf{x} has been defined. This function, called the individual objective function, measures the outcome (effect) of the location pattern for client i (cf. [6]). Individual objective functions f_i depend on effects of several allocations decisions. It means they depend on allocation effect coefficients

$d_{ij} > 0$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$), called hereafter simply distance coefficients or distances as they usually express the distance (or travel time) between client i and location j . For the standard uncapacitated location problem, it is assumed that all the potential facilities provide the same type of service and each client is serviced by the nearest located facility. The individual objective functions then take the following form:

$$f_i(\mathbf{x}) = \min_{j=1, \dots, n} \{d_{ij} : x_j = 1\}, \quad \text{for } i = 1, 2, \dots, m.$$

With the explicit use of the allocation variables and the corresponding constraints (2) and (3), the individual objective functions f_i can be written in the linear form:

$$f_i(\mathbf{x}) = \sum_{j=1}^n d_{ij} x'_{ij}, \quad \text{for } i = 1, 2, \dots, m. \quad (5)$$

These linear functions of the allocation variables are applicable for the uncapacitated as well as for the capacitated facility location problems. One may be interested in putting into formula (5) some additional factors related to the client attributes, like the service demand v_i , etc. This can be implemented without affecting formula (5) by an appropriate transformation of the distance coefficients d_{ij} (e.g., $d_{ij} = v_i d'_{ij}$, for $j = 1, 2, \dots, n$). Note that in such transformations, the set of potential locations and the set of clients are treated independently thus generating possible nonsymmetric distances if the same “geographical” point belongs to both categories.

In typical formulations of location problems related to desirable facilities, smaller value of the individual objective function means better effect (higher service quality or client satisfaction). This remains valid for location of obnoxious facilities if the distance coefficients are replaced with their complements to some large number: $d'_{ij} = d - d_{ij}$, where $d > d_{ij}$ for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. Therefore, without loss of generality, we can assume that each function f_i needs to be minimized. Hence, the generic location problem can be viewed as the following multiple criteria minimization problem:

$$\min\{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in Q\}, \quad (6)$$

where:

$\mathbf{f} = (f_1, \dots, f_m)$ is a vector function that maps the decision space $X = R^n$ into the criterion space $Y = R^m$,

$Q \subset X$ denotes the feasible set of location patterns,

$\mathbf{x} \in X$ denotes the vector of decision variables (the location pattern).

We do not assume any special form of the feasible set while analyzing properties of the solution concepts. We rather consider that the feasible set is a general discrete (nonconvex) set. Therefore, the results of our analysis apply to various discrete location problems. However, for the solution procedure of the reference distribution approach we will assume that the feasible set includes constraints (1)–(4) and possible side constraints. Similarly, we do not assume any special form of the individual objective functions nor their special properties (like convexity) while analyzing properties of the solution concepts. However, for the solution procedure of the

reference distribution approach we will assume that the individual objective functions are defined in terms of formula (5).

The elements of the criterion space we refer to as achievement vectors. An achievement vector $\mathbf{y} \in Y$ is attainable if it expresses outcomes of a feasible solution $\mathbf{x} \in Q$ ($\mathbf{y} = \mathbf{f}(\mathbf{x})$). The set of all the attainable achievement vectors is denoted by Y_a , i.e. $Y_a = \{\mathbf{y} \in Y: \mathbf{y} = \mathbf{f}(\mathbf{x}), \mathbf{x} \in Q\}$.

Model (6) only specifies that we are interested in the minimization of all objective functions f_i for $i \in I = \{1, 2, \dots, m\}$. The individual objective functions f_i are usually conflicting when minimized. Therefore, the location problem (6) is a true multiple criteria decision problem and one needs to decide for a compromise among the individual outcomes. There is, however, a specificity of problem (6) related to the location decision circumstances. In typical multiple criteria problems values of the individual objective functions are assumed to be incomparable [7]. The individual objective functions in our multiple criteria location model express the same quantity (usually the distance) for various clients. Thus the functions are uniform in the sense of the scale used and their values are directly comparable. Moreover, especially while locating public facilities, the clients should be considered impartially. Thus the distribution of distances among the clients is more important than the assignment of several distances to the specific clients. In other words, a location pattern generating individual distances: 4, 3 and 0 for clients 1, 2 and 3, respectively, should be considered equally good as a solution generating distances 0, 3 and 4. Our approach will take into account this specificity of the multiple criteria location model (6).

2. Symmetric efficiency

It is clear, or rather commonly accepted, that an achievement vector is better than another if all its individual outcomes are better or at least one individual outcome is better whereas no other one is worse. In fact, it is the most general assumption about the preference model underlying the multiple criteria optimization [8]. This is mathematically formalized with the domination relation defined on the criterion space Y .

Definition 1. We say that achievement vector $\mathbf{y}' \in Y$ dominates $\mathbf{y}'' \in Y$, or \mathbf{y}'' is dominated by \mathbf{y}' , if $y'_i \leq y''_i$ for all $i \in I$ and for at least one index i_0 strict inequality holds (i.e., $y'_{i_0} < y''_{i_0}$).

Unfortunately, there usually does not exist an attainable achievement vector that dominates all the others with respect to all the criteria. Thus, in terms of the domination relation, we cannot distinguish the best attainable achievement vector. We can only distinguish the attainable achievement vectors which are not dominated by the others.

Definition 2. We say that achievement vector $\mathbf{y} \in Y_a$ is nondominated, if there does not exist a vector $\mathbf{y}' \in Y_a$ such that \mathbf{y}' dominates \mathbf{y} .

Definition 3. We say that feasible solution $\mathbf{x} \in Q$ is an efficient (Pareto optimal) solution to the multiple criteria problem (6), if $\mathbf{y} = \mathbf{f}(\mathbf{x})$ is a nondominated achievement vector.

Recall that in the multiple criteria location problem (6) all the individual objective functions are uniform and equally important. Moreover, we want to consider all the clients, and thereby all the individual objective functions, impartially. Thus we are interested in comparison rather distributions of outcomes of the achievement vectors. Note that having two possible location patterns generating achievement vectors $\mathbf{y}' = (5, 0, 5)$ and $\mathbf{y}'' = (0, 1, 0)$, respectively, we recognize both the location patterns as efficient. In fact, neither \mathbf{y}' dominates \mathbf{y}'' nor \mathbf{y}'' dominates \mathbf{y}' . However, the first location pattern generates two outcomes (distances) equal to 5 and one outcome equal to 0, whereas the second pattern generates one outcome equal to 1 and two outcomes equal to 0. Thus, the second location pattern is clearly better.

For multiple criteria problems with uniform and equally important objective functions, we need to introduce an efficiency concept based rather on the set of outcomes than on the achievement vectors. For this purpose, we assume that the preference model satisfies the principle of impartiality (anonymity)

$$(y_{\tau(1)}, y_{\tau(2)}, \dots, y_{\tau(m)}) \cong (y_1, y_2, \dots, y_m), \quad \text{for any permutation } \tau \text{ of } I. \tag{7}$$

Condition (7) reads that any permutation of the achievement vector is equally good (indifferent) as the original achievement vector. Adding the principle of impartiality to the domination relation leads us to the concept of symmetric domination which is not affected by any permutation of the achievement vector coefficients.

Definition 4. We say that achievement vector $\mathbf{y}' \in Y$ symmetrically dominates $\mathbf{y}'' \in Y$, or \mathbf{y}'' is symmetrically dominated by \mathbf{y}' , if there exist permutations τ' and τ'' such that $y'_{\tau'(i)} \leq y''_{\tau''(i)}$ for all $i \in I$ and for at least one index i_0 strict inequality holds (i.e., $y'_{\tau'(i_0)} < y''_{\tau''(i_0)}$).

Definition 5. We say that feasible solution $\mathbf{x} \in Q$ is a symmetrically efficient solution to the multiple criteria problem (6), if $\mathbf{y} = \mathbf{f}(\mathbf{x})$ is symmetrically nondominated, i.e. does not exist $\mathbf{x}' \in Q$ such that $\mathbf{y}' = \mathbf{f}(\mathbf{x}')$ symmetrically dominates $\mathbf{y} = \mathbf{f}(\mathbf{x})$.

Note that an achievement vector \mathbf{y} symmetrically dominates all achievement vectors which are dominated in the standard sense as well as vectors dominated by any permutation of \mathbf{y} (see Fig. Fig. 1). Hence, the concept of symmetric efficiency is a refinement of the standard efficiency and the symmetrically efficient set is a subset of the standard efficient set.

The relation of symmetric domination can be expressed as domination of the achievement vectors with coefficients ordered in nonincreasing order. This can be mathematically formalized with the ordering map $\Theta: R^m \rightarrow R^m$ such that $\Theta(\mathbf{y}) = (\theta_1(\mathbf{y}), \theta_2(\mathbf{y}), \dots, \theta_m(\mathbf{y}))$, where $\theta_1(\mathbf{y}) \geq \theta_2(\mathbf{y}) \geq \dots \geq \theta_m(\mathbf{y})$ and there exists a permutation τ of set I such that $\theta_i(\mathbf{y}) = y_{\tau(i)}$ for $i = 1, 2, \dots, m$. The following proposition is valid [9].

Proposition 1. An achievement vector $\mathbf{y}' \in Y$ symmetrically dominates $\mathbf{y}'' \in Y$, if and only if $\bar{\mathbf{y}}' = \Theta(\mathbf{y}')$ dominates $\bar{\mathbf{y}}'' = \Theta(\mathbf{y}'')$, i.e. $\theta_i(\mathbf{y}') \leq \theta_i(\mathbf{y}'')$ for all $i \in I$ and for at least one index i_0 strict inequality holds (i.e., $\theta_{i_0}(\mathbf{y}') < \theta_{i_0}(\mathbf{y}'')$).

Proposition 1 permits one to express symmetric efficiency for problem (6) in terms of the standard efficiency for the multiple criteria problem with objectives $\Theta(\mathbf{f}(\mathbf{x}))$:

$$\min\{(\theta_1(\mathbf{f}(\mathbf{x})), \theta_2(\mathbf{f}(\mathbf{x})), \dots, \theta_m(\mathbf{f}(\mathbf{x}))) : \mathbf{x} \in Q\}. \quad (8)$$

Corollary 1. *A feasible solution $\mathbf{x} \in Q$ is a symmetrically efficient solution to the multiple criteria problem (6), if and only if it is an efficient solution to the multiple criteria problem (8).*

Example 1. In order to illustrate the concept of symmetric domination, let us consider a problem of locating two facilities among 10 points where each point represents one client and it can be considered as a potential location for a facility. We assume that the facilities have unlimited capacities and each client is served by the nearest facility. Thus, the problem takes the form (1)–(6) with $m = n = 10$ and $p = 2$. To make possible an easy analysis of the problem without complex computations, we consider points P1, P2, ..., P10 on a line, say X -axis, with coordinates: 0, 4, 5, 6, 8, 17, 18, 19, 20 and 28, respectively.

Table 1 contains four various solutions to the location problem. The first one corresponds to the lexicographic minimax (center) solution concept [10] where in addition to the largest outcome, the second largest outcome is also minimized (provided that the largest one remains as small as possible), the third largest outcome is minimized (provided that the two largest remain as small as possible), and so on. This solution depends on locating facilities at points P2 and P9. In the second row of Table 1, there are distances for another, in our opinion the worst, minimax solution. It is based on locating facilities at points P1 and P9. Further, we have included the minisum (median) solution and the solution minimizing the Gini coefficient, which is perhaps the most popular inequality measure in economics [6]. The minisum solution

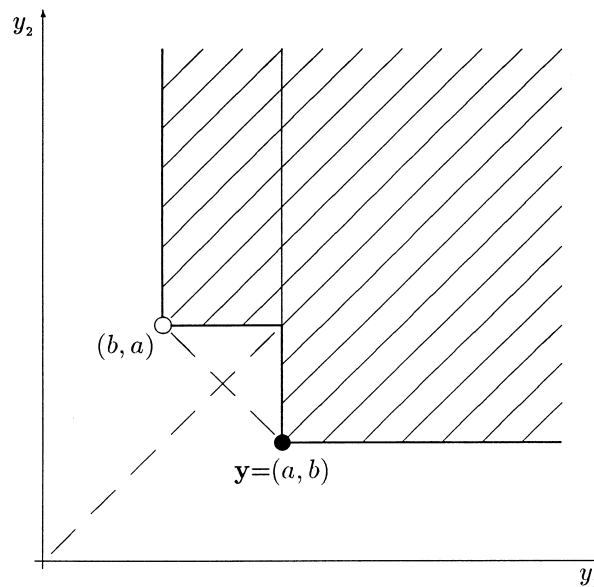


Fig. 1. Achievement vectors symmetrically dominated by $\mathbf{y} \in R^2$.

is based on locating facilities at points P3 and P8. The Gini solution uses locations P1 and P10.

Note that among the four solutions (achievement vectors) presented in Table 1 no one is dominated by any other. In fact, all these solutions are efficient as, due to the problem specificity, each feasible solution is efficient. However, one can easily observe that the ordered achievement vector of the second solution is dominated by that of the first one. The ordered achievement vector of the fourth solution is dominated by each of the other three vectors. Thus, both the second and the fourth solutions are not symmetrically efficient.

3. Generation techniques

Efficient solutions of the multiple criteria problem (6) can be generated with simple scalarizations of the problem. Most of them are based on the minisum approach:

$$\min \left\{ \sum_{i=1}^m f_i(\mathbf{x}) : \mathbf{x} \in Q \right\}, \tag{9}$$

or on the minimax approach:

$$\min \left\{ \max_{i=1, \dots, m} f_i(\mathbf{x}) : \mathbf{x} \in Q \right\}. \tag{10}$$

However, the latter generates an efficient solution only in the case of a unique optimal solution. In the general case, the optimal set of (10) includes an efficient solution and some additional refinement (regularization) is necessary to select the optimal solution which is efficient. Therefore, the minimax scalarization is usually regularized by the additional minisum term thus generating the augmented minimax problem [7]:

$$\min \left\{ \max_{i=1, \dots, m} f_i(\mathbf{x}) + \varepsilon \sum_{i=1}^m f_i(\mathbf{x}) : \mathbf{x} \in Q \right\}, \tag{11}$$

where ε is an arbitrarily small positive number.

Table 1
Location solutions for Example 1

Solution	Locations		Achievement vector										Ordered achievement vector									
			1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
1	P2	P9	4	0	1	2	4	3	2	1	0	8	8	4	4	3	2	2	1	1	0	0
2	P1	P9	0	4	5	6	8	3	2	1	0	8	8	8	6	5	4	3	2	1	0	0
3	P3	P8	5	1	0	1	3	2	1	0	1	9	9	5	3	2	1	1	1	1	0	0
4	P1	P10	0	4	5	6	8	11	10	9	8	0	11	10	9	8	8	6	5	4	0	0

For location problems, the minimum and the minimax approaches represent the (unweighted) median and the center solution concepts, respectively. Halpern [11] has introduced a parametric solution concept of λ -cent-dian as the optimal solution of problem:

$$\min \left\{ \lambda \max_{i=1, \dots, m} f_i(\mathbf{x}) + (1 - \lambda) \sum_{i=1}^m f_i(\mathbf{x}) : \mathbf{x} \in Q \right\}. \quad (12)$$

The λ -cent-dian covers as special cases the center ($\lambda = 1$) and the median ($\lambda = 0$) solution concepts. For $0 < \lambda < 1$, the λ -cent-dian minimizes a convex combination of the average and maximum distance and it is an extension of the augmented minimax approach (11) for not necessarily small ε . Note that:

$$\max_{i=1, \dots, m} \theta_i(\mathbf{f}(\mathbf{x})) = \max_{i=1, \dots, m} f_i(\mathbf{x}) \quad \text{and} \quad \sum_{i=1}^m \theta_i(\mathbf{f}(\mathbf{x})) = \sum_{i=1}^m f_i(\mathbf{x}).$$

Hence, all the above scalarizations can also be considered as applied to the ordered problem (8). Thus, due to the theory of multiple criteria optimization [7] and Corollary 1, the following proposition is valid.

Proposition 2. *For any $0 \leq \lambda < 1$, each optimal solution to the λ -cent-dian problem (12) is a symmetrically efficient solution to the multiple criteria location problem (6).*

Usually there exist many symmetrically nondominated achievement vectors and they are incomparable with each other on the basis of the specified set of objective functions. Therefore, usually there exist many symmetrically efficient solutions and they are different not only in the decision space but also in the criteria space. So, there arises a need for further analysis, or rather decision support, to help the decision maker (DM) in the selection of one solution for implementation. Of course, the original objective functions do not allow one to select any symmetrically efficient solution as better than any other one. Therefore, this analysis depends usually on additional information about the DMs preferences. The DM, working interactively with a decision support system (DSS), specifies the preferences in terms of some control parameters and the DSS provides the DM with a symmetrically efficient solution which is the best according to the specified control parameters. For such an analysis, there is no need to identify the entire symmetrically efficient set prior to the analysis, since contemporary optimization software is powerful enough to be used on-line for direct computations at each interactive step. Thus, the DSS can generate at each interactive step only one solution that meets the current preferences. Such a DSS can be used for the analysis of decision problems with finite as well as infinite symmetrically efficient sets. In order to allow the DSS to meet various DMs preferences, it is important, however, that the control parameters provide the completeness of the control (cf. [12]), i.e., by varying the control parameters, the DM can identify every symmetrically nondominated achievement vector.

For an interactive DSS dealing with multiple uniform criteria, we need parametric solution concepts generating symmetrically efficient solutions. In the case of the standard efficiency, one may consider weighting of the objective functions. In the case of impartial criteria, we cannot

assign various weights to individual objective functions. Due to Corollary 1, the weights should be assigned rather to the specific coefficients of the ordered achievement vectors. Such an ordered weighting approach was proposed by Yager [13] in the so-called Ordered Weighted Averaging (OWA) aggregation. Applying the OWA aggregation operator to the multiple criteria problem (6), we get the following single objective problem:

$$\min \left\{ \sum_{i=1}^m w_i \theta_i(\mathbf{f}(\mathbf{x})) : \mathbf{x} \in Q \right\}. \quad (13)$$

Due to Corollary 1, the following proposition is valid.

Proposition 3. *In the case of positive weights $w_i > 0$ ($i = 1, \dots, m$), every optimal solution to problem (13) is a symmetrically efficient solution to the multiple criteria problem (6).*

Note that the OWA scalarization (13) with positive weights covers as its special cases: the minimum approach (9) ($w_1 = \dots = w_m = 1$), the augmented minimax approach (11) ($w_1 = 1 + \varepsilon$, $w_2 = \dots = w_m = \varepsilon$) and the λ -cent-dian (12) for $0 \leq \lambda < 1$ ($w_1 = 1$, $w_2 = \dots = w_m = 1 - \lambda$), whereas the pure minimax approach (10) is represented with $w_1 = 1$ and $w_2 = \dots = w_m = 0$. As a limiting case of the OWA problem (13), when the differences among the weights increase to infinity (i.e., $w_1 \gg w_2 \gg \dots \gg w_m$), we get the lexicographic problem:

$$\text{lexmin} \{ (\theta_1(\mathbf{f}(\mathbf{x})), \theta_2(\mathbf{f}(\mathbf{x})), \dots, \theta_m(\mathbf{f}(\mathbf{x}))) : \mathbf{x} \in Q, \} \quad (14)$$

which represents the lexicographic minimax approach [10] to the original multiple criteria problem (6). Problem (14) is a regularization of the standard minimax scalarization (10). In the former, having minimized the largest outcome, we minimize also the second largest outcome (provided that the largest one remains as small as possible). Further, we minimize the third largest outcome (provided that the two largest remain as small as possible), and so on. Since the lexicographic optimization generates efficient solutions, thus due to Corollary 1, we get the following corollary.

Corollary 2. *The optimal solution to the lexicographic minimax problem (14) is a symmetrically efficient solution to the multiple criteria problem (6).*

Unfortunately, the ordered weighting does not provide us with a complete parameterization of the entire symmetrically efficient set. This is due to the specificity of the linear weighting approach. In the case when the multiple criteria problem is a discrete one (like the location problem (6)), there exist symmetrically efficient solutions that cannot be generated as optimal solutions to problem (13) with any set of positive weights. We illustrate this with a small example.

Example 2. Let us consider a simple single facility location problem with two clients (C1 and

C2) and three potential locations (P1, P2 and P3). The distances between the clients and potential locations are given as follows: $d_{11} = 15$, $d_{12} = 14$, $d_{13} = 12$, $d_{21} = 10$, $d_{22} = 11$, $d_{23} = 12$.

Note that all three feasible solutions are efficient in the standard and symmetric sense. One can easily verify that while dealing with the ordered weighting approach, location P2 cannot be selected for any set of positive weights. If $3w_1 < 2w_2$, then location P1 is the unique optimal solution to problem (13). If $3w_1 > 2w_2$, then location P3 is the unique optimal solution to problem (13). Finally, if $3w_1 = 2w_2$, then both locations P1 and P3 are optimal. Location P2 is never an optimal solution to problem (13) with the data of this example.

4. The reference point approach

In the case of multiple criteria problems with discrete (or more general nonconvex) feasible sets, the entire efficient set can be parameterized with the weighted augmented minimax approach (cf. [7]). This is used as the basis of the reference point method [14]. Due to Corollary 1, we can apply the reference point method to the ordered problem (8) to parameterize the entire symmetrically efficient set of the original multiple criteria problem (6). The reference point method is an interactive technique for an open search for a satisfying efficient solution. The basic concept of the interactive scheme is as follows. The DM specifies requirements in terms of aspiration levels for individual objective functions. Depending on the specified aspiration levels a special scalarizing achievement function is built which when minimized generates an efficient solution to the problem. The computed efficient solution is presented to the DM as the current solution allowing comparison with previous solutions and modifications of the aspiration levels if necessary.

The scalarizing achievement function not only guarantees efficiency of the solution but also reflects the DMs expectation as specified by the aspiration levels. In building the function the following assumption regarding the DMs expectations is made: the DM prefers outcomes that satisfy all the aspiration levels to any outcome that does not reach one or more of the aspiration levels. In other words, the aspiration levels are considered to be the desired values for the corresponding outcomes. However, if there is an opportunity to reach all the aspiration levels and to improve further some outcomes, then the reference point method will do it. One of the simplest scalarizing achievement functions takes the following form (cf. [7]):

$$s(\mathbf{x}) = \max_{i=1, \dots, m} \{ \lambda_i (f_i(\mathbf{x}) - a_i) \} + \varepsilon \sum_{i=1}^m \lambda_i (f_i(\mathbf{x}) - a_i), \quad (15)$$

where:

- a denotes the vector of aspiration levels,
- λ is a scaling vector, $\lambda_i > 0$,
- ε is an arbitrarily small positive number.

Minimization of the scalarizing achievement function (15) over the feasible set generates an efficient solution. The selection of the solution within the efficient set depends on two vector

parameters: an aspiration vector \mathbf{a} and a scaling vector λ . In practical implementations the former is usually designated as a control tool for direct use by the DM during the interactive analysis. The latter is automatically calculated on the basis of some predecision analysis or it is adjusted during the interactive process depending on values of the so-called reservation levels used as additional control parameters expressing the worst accepted values of outcomes (cf. [15]). The small scalar ε is introduced only to guarantee efficiency in the case of a nonunique optimal solution. It can be replaced by two level lexicographic minimization of the corresponding terms [15]. The reference point approach was successfully implemented in many DSS (cf. [16]) with real-life applications including multiple criteria location problems (e.g. [17]).

One may apply the scalarizing achievement function (15) to the ordered problem (8). It results in the scalarizing achievement function:

$$\bar{s}(\mathbf{x}) = \max_{i=1, \dots, m} \{ \lambda_i (\theta_i(\mathbf{f}(\mathbf{x})) - \bar{a}_i) \} + \varepsilon \sum_{i=1}^m \lambda_i (\theta_i(\mathbf{f}(\mathbf{x})) - \bar{a}_i), \tag{16}$$

where $\bar{\mathbf{a}} = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_m)$ is an ordered aspiration vector ($\bar{\mathbf{a}} = \Theta(\mathbf{a})$). Applying function (16) to the multiple criteria problem (6) we get the following parameterized single objective problem generating symmetrically efficient solutions:

$$\min \{ \bar{s}(\mathbf{x}) : \mathbf{x} \in Q \}. \tag{17}$$

Due to Corollary 1, the parametric problem (17) provides us with a complete parameterization for the symmetrically efficient set of the multiple criteria problem (6). This means that any optimal solution to problem (17) is a symmetrically efficient solution of (6), and any symmetrically efficient solution of the multiple criteria problem (6) can be found as an optimal solution to problem (17) for some aspiration vector $\bar{\mathbf{a}}$.

The ordering operator Θ used in the definition of the scalarizing achievement function (16), in general, makes the corresponding problem (17) very difficult to implement. However, even the unweighted scalarizing achievement function (16) with all $\lambda_i = 1$ provides us with a complete parameterization of the entire symmetrically efficient set. If we decide to use such an unweighted scalarizing achievement function, we can form the corresponding scalarized problem (17) without the ordering operator in the following form:

$$\begin{aligned} &\text{minimize} && \max_{i=1, \dots, m} z_i + \varepsilon \sum_{i=1}^m z_i \\ &\text{subject to} && \mathbf{x} \in Q \\ &&& z_i = f_i(\mathbf{x}) - \sum_{l=1}^m a_l u_{il}, \quad \sum_{l=1}^m u_{il} = 1, \quad \text{for } i = 1, 2, \dots, m \\ &&& \sum_{i=1}^m u_{il} = 1, \quad \text{for } l = 1, 2, \dots, m \\ &&& u_{il} \in \{0, 1\}, \quad \text{for } i = 1, 2, \dots, m; l = 1, 2, \dots, m, \end{aligned}$$

where the ordering of outcomes is implemented with additional assignment constraints.

Note that the aspiration vector $\bar{\mathbf{a}}$ in the scalarizing achievement function (16) is used in the ordered form ($\bar{\mathbf{a}} = \Theta(\bar{\mathbf{a}})$). Thus, it is rather an aspiration set of outcomes than a vector. For problems with large number of objectives, like location problem (6) with large number of clients, we can consider it as an aspiration distribution of outcomes.

5. The reference distribution approach

For the discrete location problem (1)–(6) there exists a finite set of all possible outcomes of the individual objective functions. Therefore, we can directly deal with the distribution of outcomes described by frequencies of several outcomes. Let $V = \{v_1, v_2, \dots, v_r\}$ (where $v_1 > v_2 > \dots > v_r$) denote the set of all attainable outcomes (all possible values of the individual objective functions f_i for $\mathbf{x} \in Q$). We introduce integer functions $h_k(\mathbf{x})$ ($k = 1, 2, \dots, r$) expressing the number of values v_k taken in the achievement vector $\mathbf{f}(\mathbf{x})$. Analytically, the functions h_k can be introduced into any discrete model by auxiliary assignment (binary) variables u_{ik} with the following formulas:

$$h_k(\mathbf{x}) = \sum_{i=1}^m u_{ik}, \quad \text{for } k = 1, 2, \dots, r.$$

$$f_i(\mathbf{x}) = \sum_{k=1}^r v_k u_{ik}, \quad \sum_{k=1}^r u_{ik} = 1, \quad \text{for } i = 1, 2, \dots, m.$$

$$u_{ik} \in \{0, 1\}, \quad \text{for } i = 1, 2, \dots, m; k = 1, 2, \dots, r.$$

Note that in our discrete location problem with explicit allocation variables (2)–(4), the functions h_k can be introduced directly to the model without auxiliary variables u_{ik} . Let us partition the set of all the allocation indices (i, j) $i = 1, \dots, m; j = 1, \dots, n$ into classes C_k ($k = 1, \dots, r$) defined by equal distance coefficients d_{ij} . Let $d(C_k)$ ($k = 1, \dots, r$) denote the value of the distance coefficient for the class C_k and $d(C_1) > d(C_2) > \dots > d(C_r)$. Then, the functions h_k can be defined with the simple linear formula:

$$h_k(\mathbf{x}) = \sum_{(i,j) \in C_k} x'_{ij}, \quad \text{for } k = 1, 2, \dots, r. \quad (18)$$

Having defined the functions h_k , we can introduce cumulative distribution functions:

$$\bar{h}_k(\mathbf{x}) = \sum_{l=1}^k h_l(\mathbf{x}), \quad \text{for } k = 1, 2, \dots, r. \quad (19)$$

The function \bar{h}_k expresses the number of outcomes greater or equal to v_k . Certainly, for the discrete location problem (1)–(6) with functions h_k defined by formula (18), the corresponding functions \bar{h}_k are linear. Since we want to minimize all the outcomes, we are interested in the minimization of all the functions \bar{h}_k . This leads us to the following multiple criteria problem:

$$\min\{(\bar{h}_1(\mathbf{x}), \bar{h}_2(\mathbf{x}), \dots, \bar{h}_r(\mathbf{x})) : \mathbf{x} \in Q\}. \tag{20}$$

The following proposition (the proof is given in Appendix) permits one to express symmetric efficiency for problem (6) in terms of the standard efficiency for the multiple criteria problem (20).

Proposition 4. *A feasible solution $\mathbf{x} \in Q$ is a symmetrically efficient solution to the multiple criteria problem (6), if and only if it is an efficient solution to the multiple criteria problem (20).*

For the discrete location problem (1)–(6), the number of objective functions \bar{h}_k in problem (20) is given by the number of classes C_k which is related to the accuracy of distances. While dealing with discrete location problems we have a discrete set of feasible alternatives (location patterns). However, the alternatives are evaluated and compared on the basis of outcomes (distances) which may take real values. This leads us to the fundamental problem of the accuracy of distances, or rather the accuracy of distinguishing different distances. We argue that the DM is usually not interested in a reduction of some distance by an arbitrarily small quantity. In our opinion, the DM rather wishes to distinguish some categories of distances expressed with fuzzy terms such as: extremely close, very close, ..., far away, very far away, etc. In the simplest model, the distance categories and the corresponding classes C_k can be defined by appropriate intervals of distances without any need to incorporate the fuzzy set theory (cf. [18]). If one wishes to take advantages of the fuzzy set theory by modeling each distance category with the membership function μ_k , then the corresponding functions h_k are defined by the linear formula:

$$h_k(\mathbf{x}) = \sum_{i=1}^m \sum_{j=1}^n \mu_k(d_{ij})x'_{ij}, \quad \text{for } k = 1, 2, \dots, r.$$

We believe that most decision makers distinguish a quite small number of the distance categories. With such a modeling approach, there is a very limited number of the corresponding objective functions in problem (20). Thus, we argue that problem (20) has not only simpler objective functions than problem (8), but also it has a rather small number of objective functions independently of the number of clients.

Due to Proposition 4, we can apply the standard reference point method to the multiple criteria problem (20) for an interactive analysis of the symmetrically efficient set of problem (6). Note that function \bar{h}_r represents the number of outcomes greater than or equal to the smallest possible outcome v_r . Hence, function \bar{h}_r is constant ($\bar{h}_r(\mathbf{x}) = m$ for all $\mathbf{x} \in Q$) and it can be dropped from the analysis. Thus, the corresponding scalarizing achievement function takes the following form:

$$s(\mathbf{x}) = \max_{k=1, \dots, r-1} \{\lambda_k(\bar{h}_k(\mathbf{x}) - \bar{q}_k)\} + \varepsilon \sum_{k=1}^{r-1} \lambda_k(\bar{h}_k(\mathbf{x}) - \bar{q}_k), \tag{21}$$

where:

\bar{q} denotes the vector of aspiration levels for the cumulative distribution of outcomes,

- λ is a scaling vector, $\lambda_k > 0$,
 ε is an arbitrarily small positive number.

Due to Proposition 4, the scalarizing achievement function (21) provides us with a complete parameterization for the symmetrically efficient set of the multiple criteria problem (6). This means that minimization of (21) over the feasible set generates a symmetrically efficient solution of (6) and any symmetrically efficient solution of the multiple criteria problem (6) minimizes (21) for some aspiration vector \bar{q} . The aspiration distribution vector \bar{q} is the main control tool for direct use by the DM during an interactive analysis. The scaling factors λ_k can be used as auxiliary control parameters and be modified by the DM during the interactive process in order to emphasize the importance of some criteria. Note that, in the case of larger r , the DM does not need to deal with all the aspiration coefficients \bar{q}_k . As \bar{q} represents the reference cumulative distribution, it can be specified with only a few coefficients \bar{q}_k and automatic interpolation of values for the remaining coefficients. Thus, during the initial stage of the interactive search the DM may only roughly specify the aspiration distribution taking into account larger classes of distances.

Example 3. In order to illustrate the reference distribution approach, we show the sample interactive analysis of a randomly generated discrete location problem (1)–(6). We consider location of two facilities ($p = 2$) among 10 given potential locations ($n = 10$) to serve the set of 50 clients ($m = 50$). To define the data, we have randomly generated (from a uniform distribution) 60 points with integer coordinates within the interval 0–100. The distances have been defined by calculating first the Euclidean distances, and next by rounding the results to the assumed distance step 10. Thus, we have received the multiple criteria problem (6) with 50 uniform criteria for which the values belong to the set $V = \{150, 140, \dots, 10, 0\}$, where $v_1 = 150, v_2 = 140, \dots, v_{15} = 10$ and $v_{16} = 0$. We have analyzed the problem with the reference distribution approach based on the scalarizing achievement function (21) with all $\lambda_k = 1$.

Table 2
Interactive analysis of the random location problem

k	Cumulative distribution of distances \bar{h}_k														
	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
v_k	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
Iteration 1															
Aspiration	45	43	31	22	18	11	6	3	0	0	0	0	0	0	0
Solution	48	44	31	22	21	12	7	5	0	0	0	0	0	0	0
Iteration 2															
Aspiration	50	50	40	30	18	11	6	3	0	0	0	0	0	0	0
Solution	50	49	38	27	18	11	6	3	1	0	0	0	0	0	0
Iteration 3															
Aspiration	50	50	48	32	22	14	6	3	0	0	0	0	0	0	0
Solution	50	49	38	27	18	11	6	3	1	0	0	0	0	0	0

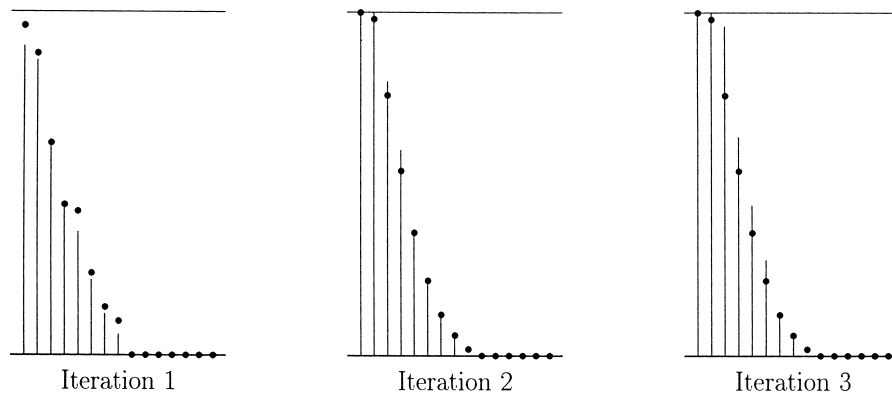


Fig. 2. Interactive analysis of the random location problem.

The interactive analysis is reported in Table 2 and Fig. 2. As the first aspiration, we have taken the utopia distribution, i.e. the vector $\bar{\mathbf{q}}$ consisted of the smallest possible values of the functions \bar{h}_k ($k = 1, 2, \dots, 15$), when minimized independently. This has resulted in the solution with a significant number of large distances (5 distances equal to 80, 7 distances greater than or equal to 70 and 12 distances greater than or equal to 60). To find a solution with a smaller number of large distances, in iteration 2 we relax (increase) the aspiration levels corresponding to the small distances ($\bar{q}_{15} = \bar{q}_{14} = 50$, $\bar{q}_{13} = 40$, $\bar{q}_{12} = 30$) leaving unchanged the aspiration levels corresponding to larger distances. As a result, we get the solution with a remarkable lower number of large distances (only 3 distances greater than or equal to 80, 6 distances greater than or equal to 70, and 11 distances greater than or equal to 60), although one distance is equal to 90. This distribution of distances seems to be a satisfying solution as we have achieved the smallest possible numbers of distances greater than or equal to 40, 50, 60 and 70, respectively.

To verify if there exists a similar solution without any distance greater than 80, in iteration 3 we further relax the aspiration levels for small distances. It turns out that we get again the same solution. There does not exist another solution with a similarly low number of larger distances. This is clear from Fig. 2, where the solution distributions are presented with dots on the bars representing the corresponding reference distributions (aspirations). Thus, we accept the current solution which generates: 1 distance of 90, 2 distances of 80, 3 of 70, 5 of 60, 7 of 50, 9 of 40, 11 of 30, 11 of 20 and one of 10. The selection of the final solution depends, certainly, on the DMs preferences. Our sample session shows, however, how easily the DM can learn the decision problem and control the analysis.

6. Concluding remarks

While making location decisions, the distribution of travel distances among the service

recipients (clients) is an important issue. It is usually tackled with the minimax (center) or the minisum (median) solution concepts. Both concepts minimize only simple scalar characteristics of the distribution: the maximal distance and the average distance, respectively. The entire distribution of distances can be taken into account in the multiple criteria model where all the distances for the individual clients are considered as the set of uniform criteria to be minimized. In order to comply with the minimization of distances as well as with an impartial consideration of the clients, the concept of symmetric efficiency must be used for this multiple criteria model. The concept is based on the principle of impartiality, i.e., on the assumption that any permutation of the achievement vector is equally good as the original achievement vector.

In the case of the standard multiple criteria optimization problems, the reference point method is a very convenient technique for an interactive analysis. It provides the DM with a tool for an open analysis of the efficient frontier. The interactive analysis is navigated with the commonly accepted control parameters expressing aspiration levels for the individual objective functions. The standard reference point method cannot be directly applied to search for symmetrically efficient solutions. In this paper we have developed, as an analogue of the reference point method, the reference distribution method taking into account the principle of impartiality. All the solutions generated during the interactive process belong to the symmetrically efficient set. The interactive analysis of the symmetrically efficient set is controlled with the aspiration cumulative distribution of outcomes. In the case of discrete location problems modeled with explicit allocation variables, the method does not complicate significantly the original problem, as all the introduced artificial objective functions are linear. The illustrative example with a randomly generated discrete location problem shows how easily the interactive analysis can be navigated with the aspiration cumulative distribution of distances.

This paper focuses on location problems. However, the location decisions are analyzed from the perspective of their effects for individual clients. Therefore, the general concept of the proposed distribution approach can be used for optimization of various systems which serve many users. Moreover, uniform individual objectives may be associated with some events rather than the physical users, like in many dynamic optimization problems where uniform individual criteria represent the same outcome for various periods.

Acknowledgements

The author gratefully acknowledges the comments and suggestions made by the anonymous reviewer which resulted in considerable improvement in the presentation of the material in this paper.

Appendix. Proof of Proposition 4

Let $\mathbf{x} \in Q$ be a symmetrically efficient solution to problem (6). Suppose that \mathbf{x} is not an efficient solution to the corresponding problem (20). This means that there exists $\mathbf{x}^0 \in Q$ such

that $\bar{h}_k(\mathbf{x}^0) \leq \bar{h}_k(\mathbf{x})$ for $k = 1, 2, \dots, r$ where for at least one index k_0 strict inequality holds (i.e., $\bar{h}_{k_0}(\mathbf{x}^0) < \bar{h}_{k_0}(\mathbf{x})$). Then, obviously, $\Theta(\mathbf{f}(\mathbf{x}^0))$ dominates $\Theta(\mathbf{f}(\mathbf{x}))$ which contradicts symmetric efficiency of \mathbf{x} for problem (6). Thus, symmetric efficiency of vector $\mathbf{x} \in Q$ for problem (6) implies its efficiency for problem (20).

Now, let $\mathbf{x} \in Q$ be an efficient solution to problem (20). Suppose that \mathbf{x} is not a symmetrically efficient solution to problem (6). This means that there exists $\mathbf{x}^0 \in Q$ such that $\bar{\mathbf{y}}^0 = \Theta(\mathbf{f}(\mathbf{x}^0))$ dominates $\bar{\mathbf{y}} = \Theta(\mathbf{f}(\mathbf{x}))$. Note that $\bar{h}_k(\mathbf{x}^0) = \bar{h}_k(\mathbf{x}) = 0$ if $v_k > \bar{y}_1^0$ and $v_k > \bar{y}_1$ as well as $\bar{h}_k(\mathbf{x}^0) = \bar{h}_k(\mathbf{x}) = m$ if $v_k < \bar{y}_m^0$ and $v_k < \bar{y}_m$. Moreover, for any $i \in I$ $\bar{y}_i^0 = v_{k'} \leq \bar{y}_i = v_{k''}$ implies $\bar{h}_k(\mathbf{x}^0) \leq \bar{h}_k(\mathbf{x})$ for $k' \leq k \leq k''$. Hence, the achievement vector $\bar{\mathbf{h}}(\mathbf{x}^0)$ dominates (in the standard sense) achievement vector $\bar{\mathbf{h}}(\mathbf{x})$ which contradicts efficiency of \mathbf{x} for problem (20). Thus, efficiency of vector $\mathbf{x} \in Q$ for problem (20) implies its symmetric efficiency for problem (6).

References

- [1] Love RF, Morris JG, Wesolowsky GO. Facilities location: models and methods. New York, NY: North-Holland, 1988.
- [2] Francis RL, McGinnis LF, White JA. Facility layout and location: an analytical approach. Englewood Cliffs, MA: Prentice-Hall, 1992.
- [3] Current J, Min H, Schilling D. Multiobjective analysis of facility location decisions. *European Journal of Operational Research* 1990;49:295–307.
- [4] Mirchandani PB, Francis RL, editors. Discrete location theory. New York, NY: Wiley, 1990.
- [5] Labbé M, Peeters D, Thisse J-F. Location on networks. In: Ball MO, Magnanti TL, Monma CL, Nemhauser GL, editors. Handbook in operations research and management science: network routing. Amsterdam, The Netherlands: North-Holland, 1996. p. 551–624.
- [6] Marsh MT, Schilling DA. Equity measurement in facility location analysis: a review and framework. *European Journal of Operational Research* 1994;74:1–17.
- [7] Steuer RE. Multiple criteria optimization — theory, computation and applications. New York, NY: Wiley, 1986.
- [8] Vincke Ph. Multicriteria decision-aid. New York, NY: Wiley, 1992.
- [9] Podinovskii VV. Multi-criterion problems with uniform equivalent criteria. *USSR Computational Mathematics and Mathematical Physics* 1975;15:47–60.
- [10] Ogryczak W. On the lexicographic minimax approach to location problems. *European Journal of Operational Research* 1997;100:566–85.
- [11] Halpern J. Finding minimal center-median convex combination (cent-dian) of a graph. *Management Science* 1978;24:534–44.
- [12] Wierzbicki AP. On completeness and constructiveness of parametric characterizations to vector optimization problems. *OR Spektrum* 1986;8:73–87.
- [13] Yager RR. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics* 1988;18:183–90.
- [14] Wierzbicki AP. A mathematical basis for satisficing decision making. *Mathematical Modelling* 1982;3:391–405.
- [15] Ogryczak W, Lahoda S. Aspiration/reservation decision support — a step beyond goal programming. *Journal of Multi-Criteria Decision Analysis* 1992;1:101–17.
- [16] Lewandowski A, Wierzbicki AP, editors. Aspiration based decision support systems — theory, software and applications. Berlin, Germany: Springer-Verlag, 1989.

- [17] Malczewski J, Ogryczak W. An interactive approach to the central facility location problem: locating pediatric hospitals in Warsaw. *Geographical Analysis* 1990;22:244–58.
- [18] Zimmermann H-J. *Fuzzy sets theory and its applications*, 3rd ed. Boston, MA: Kluwer Academic Publishers, 1996.