Control and Cybernetics

vol. **25** (1996) No. 1

A multiobjective metaheuristic approach to the location of petrol stations by the capital budgeting model[1]

by

Piotr Czyżak and **Andrzej Jaszkiewicz**

Institute of Computing Science
Poznań University of Technology
60-965 Poznań, Piotrowo 3a, Poland

Abstract: Localization of a chain of petrol stations within the area of Poland is considered. The problem is formulated in terms of capital budgeting model. To evaluate different possible solutions three conflicting criteria which express present and future profit as well as the environmental impact are taken into account. To solve the problem two stage procedure is proposed. In the first stage a multiobjective metaheuristic technique is applied. It generates an approximation of the Pareto solutions set. In the second stage, the decision maker selects the best compromise among these solutions guided by an interactive multiobjective procedure.

1. Introduction

A big, multinational company is going to initiate a chain of petrol stations located in Poland. After the preliminary analysis of the Polish petrol market the company board decided to invest some capital. Currently, 50 potential locations are considered. For each of them the expected investment costs are estimated. Moreover, some research has been performed resulting in evaluation of each location from the three, different points of view: medium expected profit in the first five years, expected profit in 15 years and the potential, negative environmental impact.

1 Work supported by State Committee of Scientific Research KBN grant 8 S503 016 06

The above described investment problem could be formulated as a multiobjective capital budgeting (MCB) problem and written down as follows:

$$\max/\min\{f_1(x), f_2(x), \ldots, f_k(x)\}$$

subject to

$$\sum_{i=1}^{n} a_i x_i \leq B$$

$$x_i \in \{0,1\}$$

where $\mathbf{x} = [x_1, \ldots, x_n]$ are decision variables equal to 1 if project (location) i is accepted and 0 otherwise, a_i $(i = 1, \ldots, n)$ is the first period expected cash outflow of project i, B is budget available for allocation to all projects in first period and f_1, $f_2, \ldots, f_k$ are criteria used to evaluate different feasible subsets of projects.

The MCB problem consists in finding the best compromise solution among the Pareto solution set. Many authors have dealt with such problem. However, in most cases only two objectives have been considered (Weingartner, 1966; Baum, Carlson and Jucker, 1978; McBridge, 1981). Moreover, usually the assumption about project dependencies and known covariances have been made (Baum, Carlson and Jucker, 1978). Other proposed solution methods, assuming independence among projects, still restrict the number of objectives to two (Rosenblatt and Sinuany-Stern, 1989).

The MCB problem, because of its combinatorial nature, may posses a large number of noninferior solutions. Moreover, the generation of the entire noninferior solution set for such problems is impractical, even if possible. In the paper a metaheuristic procedure called the Pareto Simulated Annealing (PSA) is used (Czyżak, Jaszkiewicz, 1994) to find a good approximation of the efficient solution set in the possibly shortest time. Then, the interactive Light Beam Search Discrete procedure (LBS-Discrete) proposed by Jaszkiewicz and Słowiński (1994) is applied. The LBS-Discrete allows the DM for both free scanning of the whole set of efficient solutions and iterative improvement of the currently considered solution.

The second section of the paper contains some basic definitions. In the third section the formulation and data for the problem concerning localization of a chain of petrol station are given. Fourth section presents the general scheme of the PSA and LBS methods. In the fifth section the solution process of the exemplar problem is described. The last section contains the conclusions drawn from the work.

2. Basic definitions

The general multiobjective combinatorial problem may be formulated as:

$$\max/\min\{f_1(\mathbf{x}) = \mathbf{c}_1^T\mathbf{x}, \ldots, f_k(\mathbf{x}) = \mathbf{c}_k^T\mathbf{x}\}$$

subject to

$$\mathbf{Ax} \leq \mathbf{b}$$

$$x_i \in \{0, 1\}$$

where $\mathbf{A}(k \times n)$, $\mathbf{c}_k(1 \times n)$, $\mathbf{b}(m \times 1)$, $\mathbf{x}(n \times 1)$.

Please note, that MCB problem is a special case of the multiobjective combinatorial problem.

A solution $\mathbf{y}$ *dominates* $\mathbf{x}$ if $f_j(\mathbf{y}) \geq f_j(\mathbf{x})\ \forall_j$ and $f_j(\mathbf{y}) > f_j(\mathbf{x})$ for at least one j.

A solution $\mathbf{x}$ is *efficient (Pareto optimal)* if there is no other feasible solution that dominates $\mathbf{x}$. The set of all efficient solutions is denoted by N.

The *ideal point* is the point in the objective space composed of the best attainable values of objectives.

The *nadir point* is the point in the objective space composed of the worst attainable values of objectives in the set of efficient solutions.

3. Formulation and data of the localization problem

The problem concerning a capital investment into a chain of petrol stations could be formulated in terms of the MCB problem as follows:

$$\begin{aligned} P_1 &= \max \sum_{i=1}^{50} P_{1i} x_i \\ P_2 &= \max \sum_{i=1}^{50} P_{2i} x_i \\ EI &= \max \sum_{i=1}^{50} E_i x_i \end{aligned}$$

subject to

$$\begin{aligned} \sum_{i=1}^{50} a_i x_i &\leq B \\ x_i &\in \{0, 1\} \end{aligned}$$

where P_{1i} $(i = 1, \ldots, 50)$ denotes medium profit obtained from i-th localization during the next five years, P_{2i} $(i = 1, \ldots, 50)$ is an expected profit from i-th petrol station in 15 years and E_i $(i = 1, \ldots, 50)$ means an expert evaluation of the potential negative environmental impact.

Such a specific definition of objective functions takes into account two factors. First one, is expected by the investor, stabilization of profit from the chain of petrol station for many years. Second one, are the plans concerning building of a system of highways crossing Poland from the north to the south and from

i	a_i	P_{1i}	P_{2i}	E_i	i	a_i	P_{1i}	P_{2i}	E_i
1	25	6	55	10	26	32	34	33	8
2	26	7	54	2	27	47	37	34	10
3	28	17	62	4	28	24	35	30	0
4	25	11	54	7	29	43	43	36	10
5	20	13	54	6	30	27	38	29	11
6	38	18	57	11	31	32	41	30	8
7	38	15	52	9	32	48	38	25	7
8	49	14	49	3	33	35	47	32	4
9	28	18	51	4	34	22	42	25	2
10	22	19	50	6	35	34	48	29	9
11	34	16	45	0	36	31	51	30	9
12	39	22	49	10	37	31	47	24	10
13	26	18	43	7	38	36	52	27	6
14	21	28	51	2	39	50	53	26	6
15	36	20	41	0	40	23	45	16	7
16	47	24	43	1	41	33	47	16	5
17	38	32	49	5	42	24	48	15	6
18	50	27	42	9	43	28	53	18	11
19	39	32	45	7	44	35	49	12	5
20	50	28	39	9	45	22	56	17	3
21	45	33	42	10	46	26	51	10	11
22	27	36	43	9	47	32	60	17	0
23	49	35	40	6	48	21	60	15	7
24	24	32	35	2	49	20	62	15	9
25	48	31	32	3	50	29	62	13	8

Table 1. Parameters of 50 potential locations of petrol-stations

the west to the east in the next 10-15 years. With the development of highways a deep reorganization of existing road system is connected. There will be a lot of variations in the transport flow during the process of rebuilding of the road system in Poland. The changes influence directly on expected profits of petrol station situated in different places.

Parameters of 50 potential locations of petrol station resulting from analysis are given in Tab.1. The budget B available for the investment is equal to 374.

4. Solution procedure

4.1. Stage I – the Pareto Simulated Annealing method

To solve the above described MCB problem a metaheuristic method for multiobjective combinatorial optimization, called Pareto Simulated Annealing (PSA), is used. The goal of this method is not to find a single solution, like in the single objective case, but a sample of solutions that is a good approximation of the set of efficient solutions. The general scheme of the PSA is similar to that of classical, single criterion Simulated Annealing.

The PSA method uses several concepts of two single objective metaheuristic procedures: genetic algorithms and simulated annealing. The main concepts of the method and their sources can be summarized as follows:

- the concept of neighborhood,
- acceptance of new solutions with some probability,
- dependence of the probability on a parameter called the temperature,
- the scheme of the temperature changes; these concepts are known from simulated annealing,
- the use of a sample (population) of solutions; concept coming from genetic algorithms.

Let us assume the following notation:

D – the set of feasible solutions of a given multiple criteria combinatorial problem,

$\mathbf{x}, \mathbf{y} \in D$ – feasible solutions of the multiple criteria combinatorial problem,

$S \subset D$ – the generating sample of solutions, which replaces the current solution used in the single objective SA,

M – the set of *potentially efficient solutions*, i.e. the set composed of solutions efficient with respect to all solutions generated by the method,

T – the temperature,

T_o – the starting temperature,

$\Lambda = (\lambda_1, \ldots, \lambda_n)$ – vector of criteria weights,

$V(\mathbf{x})$ – the neighborhood of solution $\mathbf{x}$, i.e. the set of solutions that can be reached from $\mathbf{x}$ by a single basic move.

The basic move is a simple transformation which applied to a feasible solution $\mathbf{x}$ gives another feasible solution close to $\mathbf{x}$. The way of implementation of the basic move depends on a particular problem solved by the method.

The general scheme of the PSA procedure may be summarized as follows:

Select a starting sample of generating solutions $S \subset D$
for each $\mathbf{x} \in S$ **do**
    Update set M of potentially efficient solutions with $\mathbf{x}$
$T := T_o$
repeat
    for each $\mathbf{x} \in S$ **do**
        Construct $\mathbf{y} \in V(\mathbf{x})$

Update set M with $\mathbf{y}$ Select solution $\mathbf{x}' \in S$ closest to $\mathbf{x}$ and nondominated with respect to $\mathbf{x}$
if there is no such solution $\mathbf{x}'$ or it is the first iteration with $\mathbf{x}$ **then**
Set random weights such that:
$\forall_j \lambda_j \geq 0$ **i** $\sum_{\mathbf{j}} \lambda_j = 1$
else
for each objective f_j

$$\lambda_j = \begin{cases} \alpha\lambda_j^{\mathbf{x}}, & \text{if} \quad F_j(\mathbf{x}) \geq f_j(\mathbf{x}') \\ \lambda_j^{\mathbf{x}}/\alpha, & \text{if} \quad F_j(\mathbf{x}) < f_j(\mathbf{x}') \end{cases}$$

$\mathbf{x} := \mathbf{y}$ (accept $\mathbf{y}$) with probability $P(\mathbf{x}, \mathbf{y}, T, \Lambda)$
if the conditions of changing the temperature are fulfilled **then**
decrease T
until the stop conditions are fulfilled

where: $\alpha > 1$ is a constant close to one (e.g. $\alpha = 1.05$), $\Lambda^{\mathbf{x}} = [\lambda_1^{\mathbf{x}}, \ldots, \lambda_f^{\mathbf{x}}]$ is the weighting vector used in the previous iteration for solution $\mathbf{x}$.

The general scheme has to be customized for solving a given combinatorial problem. The customization consists in defining the neighborhood of a feasible solution, conditions of the temperature change and stop conditions as well as setting values of such parameters like: the size of generating sample and the starting temperature.

The general scheme of the method is similar to that of classical, single criterion Simulated Annealing. The presented method, however, uses a sample of generating solutions in spite of a single solution used by classical SA.

Unlike in the single criterion SA the outcome of PSA is not a single solution but a set M of potentially efficient solutions, i.e. the set composed of solutions efficient with respect to all generated solutions. The set is updated whenever a new solution is generated. Updating set M with a new solution $\mathbf{x}$ consists in:

- adding $\mathbf{x}$ to M if there is no other solution $\mathbf{v} \in M$ such that $\mathbf{v}$ dominates $\mathbf{x}$,
- removing from set M all solutions dominated by $\mathbf{x}$.

Another difference with the single criterion SA is in the way of calculating the probability of accepting a new solution, denoted by $P(\mathbf{x}, \mathbf{y}, T, \Lambda)$. In the case of single objective SA a new solution is accepted with probability equal to one if it is not worse than the current solution. Otherwise, it is accepted with probability less than one. In the case of multiple objectives one of the following three exclusive situations may occur while comparing a new solution $\mathbf{y}$ with the current one $\mathbf{x}$:

- $\mathbf{y}$ dominates or is equal to $\mathbf{x}$,
- $\mathbf{y}$ is dominated by $\mathbf{x}$,
- $\mathbf{y}$ is nondominated with respect to $\mathbf{x}$.

In the first situation the new solution may be considered as not worse than the current one and accepted with probability equal to one. In the second situation the new solution may be considered as worse than the current one and accepted with probability less than one. Serafini (1992) and Fortemps, Teghem and Ulungu (1994) have proposed several multiple objective rules for acceptance probability which in different way treat the third situation. In the PSA method the following rule for acceptance probability is used:

$$P(\mathbf{x}, \mathbf{y}, T, \Lambda) = \min\{1, \exp(\max_j\{\lambda_j(f_j(\mathbf{y}) - f_j(\mathbf{x}))/T\})\}.$$

Please note, that the higher is the weight associated with a given objective the lower is the probability of accepting moves that decrease the value on this objective and the greater is the probability of improving value on this objective. So, controlling the weights one can increase or decrease the probability of improving values of the particular objectives.

In each iteration of the procedure a sample of solutions, called generating sample, is used. The main idea of PSA is to assure a tendency for approaching the set of efficient solutions as well as an inclination for dispersing the solution constituting the generating sample over the whole set N. In result each solution tends to investigate a specific region of set N.

The tendency for approaching the set of efficient solutions is assured by using one of the mentioned above multiple objective rules for acceptance probability. The inclination for dispersing the solutions from the generating sample over the whole set N is obtained by controlling the weights of particular objectives used in these rules. For a given solution $\mathbf{x} \in S$ the weights are changed in order to increase the probability of moving it away from its closest neighbor in S denoted by $\mathbf{x}'$. This is obtained by increasing weights of the objectives on which $\mathbf{x}$ is better than $\mathbf{x}'$ and decreasing weights of the objectives on which $\mathbf{x}$ is worse than $\mathbf{x}'$.

Please note that the algorithm of PSA is essentially parallel because calculations required for each solution from S, i.e. construction of a new solution form its neighborhood, setting the weights and accepting the new solution, may be done on different processors.

4.2. Stage II - the LBS-discrete method

Because the set of potentially efficient solutions of a MCB problem may posses a large number of solutions the DM needs a further support in order to select the best solution - the best compromise among the objectives. An interactive procedure Light Beam Search-Discrete (LBS-Discrete) (Jaszkiewicz, Słowiński, 1994) for problems with explicitly known set of solutions was used. Interactive procedures are characterized by phases of decision alternating with phases of computation. In each computation phase, a solution, or a sample of solutions, is selected for examination in the decision phase. As a result of the examination, the DM inputs some preferential information which intends to improve the solution(s) selected in the next computation phase. The interactive analysis for such a large set of solutions has the following advantages:

- in a single iteration the DM is required to supply relatively simple preference information,
- the DM is involved in the decision process which increases his/her confidence to the final solution,
- the DM is supported in learning of his/her preferences and of the possible trade-offs.

The general scheme of the LBS-Discrete method is as follows:

Step 1. Present to the DM the ideal and nadir point, i.e. the points composed of the best and worst values of the objective functions in set M, respectively.
Step 2. Ask the DM to specify a reference point or make the ideal point the first reference point.
Step 3. Find a starting current solution by projecting the reference point onto set M using the achievement scalarizing function.
Step 4. Ask the DM to specify the preferential information defining the size of a vicinity of the current solution in the set M
Step 5. Present to the DM's sample of solutions coming from the vicinity.
Step 6. Allow the DM to see other solutions in the vicinity. Terminate the procedure at this step if the best compromise solution has been found, or give the DM the possibility to move the current solution and return to step 4.

In step 4, the DM is asked to specify the preferential information defining the size of the vicinity of the above solution. The preferential information consists in specifying indifference $q_j(f_j(\mathbf{x}))$ and preference thresholds $p_j(f_j(\mathbf{x}))$ on each objective f_j. By specifying the value of indifference threshold $q_j(f_j(\mathbf{x}))$ the DM states that any difference on this objective lower or equal to this value is insignificant to him/her. In other words, he/she considers two solutions $\mathbf{x}$ and $\mathbf{y}$ indifferent with respect to f_j if $|f_j(\mathbf{x}) - f_j(\mathbf{y})| \leq q_j(f_j(\mathbf{x}))$. By specifying the value of preference threshold $p_j(f_j(\mathbf{x}))$ the DM states that he/she strictly prefers $\mathbf{x}$ to $\mathbf{y}$ if $f_j(\mathbf{x}) - f_j(\mathbf{y}) > p_j(f_j(\mathbf{x}))$. The region between $q_j(f_j(\mathbf{x}))$ and $p_j(f_j(\mathbf{x}))$ corresponds to the hesitation of the DM between indifference and strict preference. The preference information is used to build an outranking relation defining a vicinity of the current solution (see Jaszkiewicz, Słowiński, 1994 for detailed description of this step).

In step 6, the method offers to the DM two possibilities to move the vicinity over the set M. The first one consists in specifying a new reference point in the criterion space which is then projected onto set M giving a new current solution. The second possibility consists in making the current solution a selected solution in its vicinity. The first possibility allows the DM for free scanning of the whole set M and learning of the possible trade-offs. By presenting samples of solutions in each iteration the method facilitates and speeds up the learning process. The second possibility allows for iterative improvement of the current solution.

5. An application of the Pareto-SA method to the localization of a chain of petrol stations problem

5.1. Generation of the potentially efficient solutions

In order to customize the PSA method to the localization of a chain of petrol stations problem the described below decisions were made.

One of the most important issues in the application of the metaheuristic method is an appropriate definition of the neighborhood $V(\mathbf{x})$ of the given solution $\mathbf{x}$. In general, the definition of the neighborhood should fulfil the following conditions:

- The solutions belonging to the neighborhood are "similar" to the given solution. It means that values of criteria evaluating solution $\mathbf{y} \in V(\mathbf{x})$ should be not too much different from evaluations of $\mathbf{x}$.
- Each solution from $V(\mathbf{x})$ can be calculated in a relatively short time.
- Each efficient solution should be accessible, i.e. there should be a possibility of moving to any efficient solution from any feasible solution in a finite number of basic moves.

In the considered problem solution $\mathbf{x}$ denotes set of accepted locations of petrol stations. The sum of investment costs for all of them should not exceed the available budget B. So, the neighborhood solutions of solution $\mathbf{x}$ were generated by the following algorithm:

repeat
remove a randomly selected location from the set of accepted projects
until there is a free budget for the most expensive location that is not accepted
repeat
accept a randomly selected location
until there is no free budget for any location that is not accepted

The temperature was decreased after making a given number of moves (accepted or not). The starting temperature was set equal to 50 which assured that above 80% of moves were accepted. The procedure was stopped after obtaining the temperature below 1 at which less than 5% of moves were accepted.

The procedure resulted in 1017 different potentially efficient solutions.

5.2. Interactive analysis of the set of potentially efficient solutions

The ideal and the nadir are the points composed of the worst and the best values of the criteria, respectively. In this case the ideal point is as follow:

$P_1 = 727, \quad P_2 = 691 \quad EI = 36,$

while the nadir point is:

$P_1 = 273 \quad P_2 = 263 \quad EI = 95.$

The above two points give the DM information about ranges of particular criteria in set M.

The DM decides to specify a reference point composed of desired values on particular objectives:

$P_1 = 650 \quad P_2 = 650 \quad EI = 40.$

The above point is unfeasible and the method projects it onto the set M to find the feasible solution:

$P_1 = 527 \quad P_2 = 534 \quad EI = 55.$

Then, the DM was asked to specify the preferential information defining the size of the vicinity of the above solution. He inputs the following values of the indifference and preference thresholds:

Objective (f_j)	$q_j(f_j(\mathbf{x}))$	$p_j(f_j(\mathbf{x}))$
$P1$	15	45
$P2$	14	42
EI	2	6

Then, the method presents to the DM the following sample of solutions belonging to the vicinity of the above solution:

$P_1 = 484 \quad P_2 = 565 \quad EI = 54,$
$P_1 = 521 \quad P_2 = 493 \quad EI = 51,$
$P_1 = 512 \quad P_2 = 557 \quad EI = 60.$

The DM thinks that values of the objectives P_1 and P_2 are too low and decides to specify the new reference point, relaxing a little the first two objectives and much more the third one:

$P_1 = 600 \quad P_2 = 600 \quad EI = 60.$

Projection of this point results in the following feasible solution:

$P_1 = 541 \quad P_2 = 546 \quad EI = 67.$

The DM accepts the same values of the thresholds for this solutions. The sample of solutions belonging to the vicinity of the above one are presented to the DM:

$P_1 = 496 \quad P_2 = 577 \quad EI = 68,$
$P_1 = 567 \quad P_2 = 504 \quad EI = 63,$
$P_1 = 546 \quad P_2 = 537 \quad EI = 72.$

The DM observes, however, that relatively small improvement on P_1 and P_2 corresponds to significant worsening of EI and decides to select as the best compromise the solution belonging to the first vicinity and defined as follows:

$P_1 = 512 \quad P_2 = 557 \quad EI = 60.$

6. Summary and conclusions

In the paper the problem of localization of a chain of petrol stations was formulated in terms of multiobjective capital budgeting. While it is a multiobjective combinatorial problem it posses a huge number of efficient solutions. To generate them the multiobjective metaheuristic PSA method was used. In result an approximation of efficient solutions set counting over 1000 solutions was obtained. To find the best compromise the DM's search over this set was guided by the interactive LBS-Discrete procedure. Finally, the DM found the efficient solution best fitting his preferences.

References

BAUM, S., CARLSON, R.C. and JUCKER, V. (1978) Some problems in applying the continuous portfolio selection model to the discrete budgeting problem. *J. Finan. Quant. Analysis*, 333-344.

COHON, J.L. (1978) *Multiobjective programming and planning.* Academic Press.

CZYŻAK and P., JASZKIEWICZ, A. (1994) Metaheuristic procedure for multiple criteria combinatorial optimization problems. *Proceedings of the Conference on Decision Analysis, Expert Systems and Applications of Computers Systems*, Warsaw 25-27 May 1994 (in Polish).

FORTEMPS, P., TEGHEM, J. and ULUNGU, B. (1994) Heuristics for multiobjective combinatorial optimization by simulated annealing. *Proceedings of the XI-th International Conference on Multiple Criteria Decision Making*, Coimbra 1-6.08.94.

GAREY, M. and JOHNSON, D. (1979) *Computers and intractability: A guide to the theory of NP-completeness.* Freeman, San Francisco, Calif.

GLOVER, F. (1989) Tabu Search - Part I. *ORSA J.Comput.*, **1**, 190-206.

GOLDBERG D.E. (1989) *Genetic Algorithm in Search, Optimization and Machine Learning.* Addison-Wesley.

JASZKIEWICZ, A., SŁOWIŃSKI, R. (1994) The LBS-Discrete Interactive Procedure for Multiple-Criteria Analysis of Decision Problems. *Proc. XIth International Conference on Multiple Criteria Decision Making*, Coimbra, Portugal, 1-6.08.94.

KIRKPATRICK, S., GELATT, C.D. and VECCHI, M.P. (1983) Optimization by simulated annealing. *Science*, **220**, 671-680.

LAARHOVEN, P.J.M. and AARTS, E.H.L. (1987) *Simulated Annealing: Theory and Applications.* Reidel, Dordrecht.

MCBRIDGE, R.D (1981) Finding the integer efficient frontier for quadratic capital budgeting. *J. Finan. Quant. Analysis*, **16**, 247-253.

ROSENBLATT, M.J and SINUANY-STERN, Z. (1989) Generating the discrete frontier to the capital budgeting problem. *Operations Research*, **37**, 3, 384-395.

SERAFINI, P. (1992) Simulated annealing for multiple objective optimization problems. *Proceedings of the Tenth International Conference on Multiple Criteria Decision Making*, Taipei 19-24.07.92, Vol. 1, 87-96.

WEINGARTNER, H.M. (1966) Criteria for programming investment project selection. *J. Indust. Economics*, 65-76.