Multiple-Objective Genetic Algorithm Using The Multiple Criteria Decision Making Method TOPSIS

M. Méndez[1], B. Galván[2], D. Salazar[2], and D. Greiner[2]

[1] Computer Science Department, University of Las Palmas de Gran Canaria Edif. de Informática y Matemáticas. Campus de Tafira, CP 35017 Las Palmas, Spain
mmendez@dis.ulpgc.es

[2] IUSIANI, University of Las Palmas de Gran Canaria Edif. Polivalente. Campus de Tafira, CP 35017 Las Palmas, Spain
bgalvan@step.es, danielsalazaraponte@gmail.com,
dgreiner@iusiani.ulpgc.es

The so called second generation of Multi-Objective Evolutionary Algorithms (MOEAs) like NSGA-II, are highly efficient and obtain Pareto optimal fronts characterized mainly by a wider spread and visually distributed fronts. The subjacent idea is to provide the decision-makers (DM) with the most representative set of alternatives in terms of objective values, reserving the articulation of preferences to an a posteriori stage. Nevertheless, in many real discrete problems the number of solutions that belong the Pareto front is unknown and if the specified size of the non-dominated population in the MOEA is less than the number of solutions of the problem, the found front will be incomplete for a posteriori Making Decision. A possible strategy to overcome this difficulty is to promote those solutions placed in the region of interest while neglecting the others during the search, according to some DM's preferences. We propose TOPSISGA, that merges the second generation of MOEAs (we use NSGA-II) with the well known multiple criteria decision making technique TOPSIS whose main principle is to identify as preferred solutions those ones with the shortest distance to the positive ideal solution and the longest distance from the negative ideal solution. The method induces an ordered list of alternatives in accordance to the DM's preferences based on Similarity to the ideal point.

1 Introduction

Many well known and extendedly used Multi-Objective Evolutionary Algorithms (MOEAs) like NSGA-II [2] pursue to reach the efficient frontier and to sample it by a wide and even distributed set of non-dominated solutions.

Subsequently, the decision maker (DM) chooses one solution in accordance with his/her preferences. Nevertheless, this approach does not always turn out the most appropriate. For instance, in many real discrete problems, if the efficient set is numerous and the size of the non-dominated population is limited, the MOEA cannot contain the whole set of solutions, compelling the DM to lose potentially attractive alternatives. In order to solve the above-mentioned disadvantage, a possible strategy is to concentrate the search in a smaller set of Pareto optimal solutions, according to some DM's preferences. The incorporation of preferences into a MOEA is not new [1, 3]. Nonetheless, to the best of our knowledge there is no previous attempt at incorporating TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) method [6] into a MOEA. In that sense we propose TOPSISGA, combining the concept of TOPSIS (minimizing the distance to the ideal solution while maximizing the distance to the negative solution) with MOEA methodologies (we use NSGA-II). With TOPSIS, a DM needs input his/her preferences or weights that are used in the proposed method for guiding the search towards the region of interest. The method induces an ordering of the solutions based on Similarity to the ideal point.

2 TOPSIS method

The TOPSIS method was developed by Hwang and Yoon [6] for solving MCDM problems with a finite number of solutions. The TOPSIS method establishes that the chosen solution should have the shortest distance to the positive ideal solution (I^+) and the longest distance from the negative ideal solution (I^-), where the distances are calculated with a particular value of p ($1 \leq p \leq \infty$) of the Minkowski's metrics $L_p = \left\{ \sum_{i=1}^{k} w_i^p \left| f_i(\overrightarrow{x}) - f_i^* \right|^p \right\}^p$. Here, f_i^* ($i \in \{1, 2, ..., n\}$) is a vector whose coordinates corresponds to the coordinates of a reference point. With the TOPSIS method, that point is I^+ or I^-. The TOPSIS concept is rational and comprehensible. Since the Minkowski's metrics are weighted distances, the order strongly depends on the weights the DM assigns to each objective according to their preferences. The TOPSIS procedure consists of:

Step 1. Obtain a decision matrix, where a set of alternatives (solutions) $A=(a_j, j=1,2,...,k)$ is compared with respect to a set of criterion functions (objective functions) $C=(c_i, i=1,2,...,n)$, an element x_{ij} of the matrix, is a value indicating the performance rating of jth alternative with regard to the criterion c_i.

Step 2. Calculate the normalized decision matrix according to:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{k} x_{ij}^2}}$$

Step 3. Calculate the weighted normalized values as:

$$v_{ij} = w_i r_{ij}$$

w_i is the weight of the ith criterion set by the DM and $\sum_{i=1}^{n} w_i = 1$.

Step 4. Determine the positive ideal solution I^+ and the negative ideal solution I^- as:

$I^+ = (\max_j v_{1j}, \max_j v_{2j}, ..., \max_j v_{nj}) = (v_1^+, v_2^+, ..., v_n^+)$, see figure 1(a)

$I^- = (\min_j v_{1j}, \min_j v_{2j}, ..., \min_j v_{nj}) = (v_1^-, v_2^-, ..., v_n^-)$, see figure 1(a).

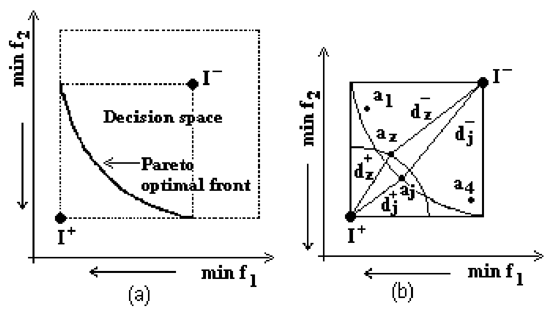

Fig. 1. (a) Positive and negative solutions, (b) TOPSIS distances.

Step 5. Calculate the Euclidean distances for each alternative from the positive ideal solution as:

$$d_j^+ = \sqrt{\sum_{i=1}^{n} (v_{ij} - v_i^+)^2} \qquad j=1,2,...k.$$

Similarly, the Euclidean distances from the negative ideal solution is given as:

$$d_j^- = \sqrt{\sum_{i=1}^{n} (v_{ij} - v_i^-)^2} \qquad j=1,2,...k.$$

Step 6. Calculate the relative closeness to the positive ideal solution (rating of Similarity to the ideal positive) as:

$$D_j^+ = \frac{d_j^-}{d_j^+ + d_j^-} \qquad D_j^+ \text{ value } [0,1].$$

Step 7. Sort the solutions in terms of similarity. The final (increasingly labelled) order is obtained sorting the set of alternatives decreasingly in terms of D_j^+, i.e. from the most similarity to the less. Figure 1(b) shows the basic principle, a_j is closer to the positive ideal and farther from the negative ideal than a_z because $d_j^+ < d_z^+$ and $d_j^- > d_z^-$; $D_j^+ > D_z^+$ and the alternative a_j is better than a_z.

3 Proposed TOPSISGA method

The present approach has a similar structure to other MOEAs, and it introduces two modifications to the original formulation of NSGA-II [2]: First the size of population and the archive of non-dominated solutions can differ; Second, TOPSISGA varies the crowding operator of NSGA-II [2] substituting the crowding distances with the relative distances D_j^+. We assume two populations: P^t, which represents the current population (size M) during generation t; and P_A^t, which consists of non-dominated solutions (archive size N). Initially, M individuals are randomly generated and the archive of non-dominated solutions is set empty. At each generation t, a combined population $R^t = P^t + P_A^t$ (size M+N) is formed (since all previous population members are included in R^t elitism is ensured). Then R^t is sorted based on dominance (figures 2(a) and 2(b)). The following population P_A^{t+1} is established with the non-dominated solutions of R^t starting with the set (F_1) of rank 1 followed by the set (F_2) with rank 2 until the last set (F_L) of rank L. If the count of solutions in all sets from F_1 to F_L is larger than the new population P_A^{t+1} (size N), we sort the last set F_L using the Similarity (D_j^+ in ascending order) (figures 2(a) and 2(c)). Afterwards a reproductive selection of individuals randomly selected from P_A^{t+1} is accomplished using a binary tournament and a mating pool (MP) is filled up, at this stage M new individuals are generated by applying recombination operators on MP.

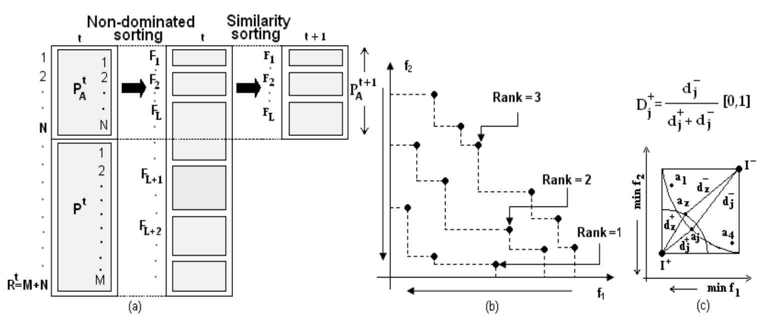

Fig. 2. Rank (a)(b) and similarity (c) concepts used by TOPSISGA.

Notice, that the selection operator uses a binary tournament and the criterion is: 1) non-dominated rank -smaller rank- 2) similarity -bigger similarity-.

4 Experimental results

4.1 The 0-1 multiobjective knapsack problem

In this section the TOPSISGA method is applied to two problems. The first is the 0-1 multiobjective knapsack problem (0-1 MOKP), which has been widely studied in the multiobjective community. The second application example is a real world engineering problem in the domain of reliability. Independently of the test problem the R^t population (size M+N) was set to 200 individuals, the crossover probability to 0.8, the mutation rate to 0.01, the p value to 2 and the archive size of non-dominated solutions was changed progressively following the sequence N=10,20,30,40,50,100 individuals. The maximum generation number (G) for the 0-1 MOKP problem was G=500 and for the reliability problem was G=50, G=100 and G=250.

Description

The 0-1 MOKP problem is well known and has been the subject of in-depth studies in the multiobjective domain. It is easy to implement it, but because of its NP-hard nature, it becomes a very difficult problem to be solved in practice. The 0-1 MOKP can be used to model many real problems and it possesses a high number of applications in finance particularly. Various evolutionary algorithms have been used to solve the 0-1 MOKP, e.g. [7, 8].
The 0-1 MOKP consists of to find a subset of items (weights and profits are associated to each item) maximizing a multiobjective function -expressed as a function of the profit values- and considering the constraints of capacity of each knapsack (maximum weight). The 0-1 MOKP can be defined formally by (1):

$$\begin{cases} max. \qquad f_i(x) = \sum_{j=1}^{m} c_{ij} x_j \qquad i = 1, 2, ..., n \\ Such\ that \qquad \sum_{j=1}^{m} w_{ij} x_j \le b_i \qquad x_j \epsilon \{0, 1\} \end{cases} \tag{1}$$

where:

m = number of items
x_j = a decision variable
n = number of objectives
c_{ij} = profit of item j according to knapsack i
w_{ij} = weight of item j according to knapsack i
b_i = capacity of knapsack i

The data adopted have been: two objectives and 100 items, the true Pareto frontier is known (figure 3), for more details see:
http://www.tik.ee.ethz.ch/%7ezitzler/testdata.html#testproblems.

Results

Figure 4 shows the results with TOPSISGA and NSGAII (for G=500 and N=10). The labels correspond to the TOPSIS classification of the final front

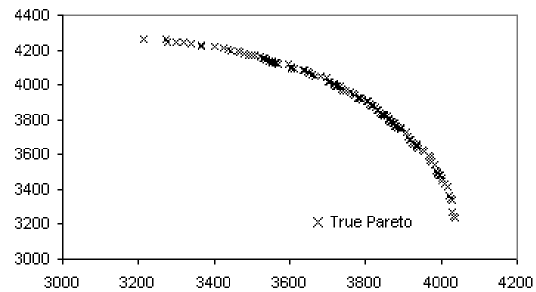

Fig. 3. True Pareto frontier knapsack problem.

when the weights are $w_1 = w_2 = 0.5$. Notice also that, TOPSISGA focuses upon a particular region of the efficient frontier while NSGA-II finds an even final set.

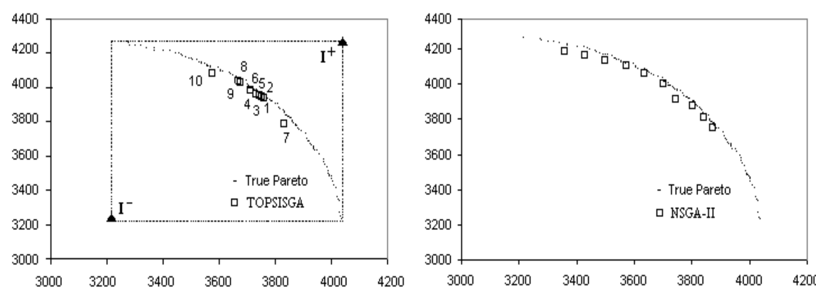

Fig. 4. Non-dominated front found by TOPSISGA and NSGAII.

TOPSISGA was compared with NSGA-II based on the C metric [10] (the lower the better), using the efficient frontier as a reference set R. Table 1 reports the percentage (average after ten runs) of the final outcomes (labelled A) dominated by the true Pareto frontier. Figure 5 shows graphically the results of table 1.

Table 1. Metric C(R,A) values for 500 generations

Method	N=10	N=20	N=30	N=40	N=50	N=100
TOPSISGA	74	74.5	86.27	87.65	84.5	90.95
NSGA-II	99	86	86.33	83.97	87.68	97.83

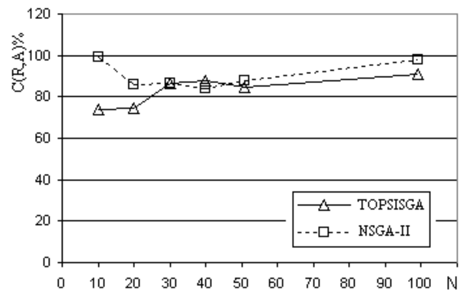

Fig. 5. Graphic view of table 1.

4.2 Safety systems design optimisation

Description

As an application example we will use here a well known dependability problem, the design optimization of a Safety System (SS) -as practical test we use the Containment Spray Injection System (CSIS) of a Nuclear Power Plant (NPP)-. The problem is combinatorial in nature and NP-hard and it has been widely studied before [4, 5]. In figure 6(a) the CSIS layout design is depicted and table 2 shows the Unavailability and Cost of the different market available components (valves and pumps) for the system, being the optimization purpose to obtain the best design. For each combination of pumps and valves the system unavailability and the system cost are computed, the former using a fault tree with design alternatives and the later using a single aggregating formula. Both objectives are in conflict so a multiobjective optimisation is the appropriate methodology. When any number of objective function evaluations can be made during the optimization, the true Pareto front of non dominated solutions can be obtained using ad hoc multiobjective methods like the NSGA-II [2]. Figure 6(b) shows the true Pareto frontier [5].

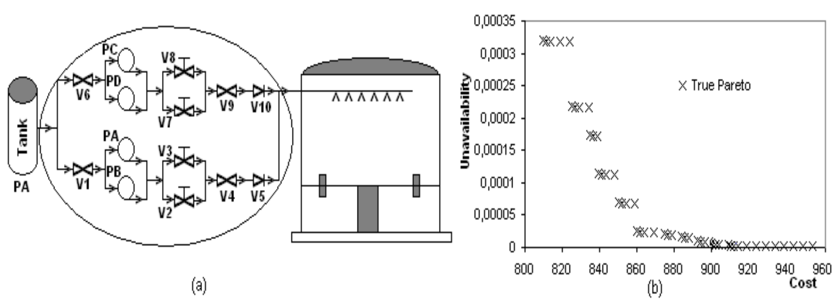

Fig. 6. (a) CSIS system for NPP, (b) The true Pareto frontier SS problem.

Table 2. Component models available on the market.

Model	Valves 1,4,6,9	Valves 2,3,7,8	Valves 5,10	Pumps a,b,c,d
Mod. 1	U=2,9E-03 C=50	U=3,0E-03 C=65	U=5,0E-04 C=37	U=3,5E-03 C=90
Mod. 2	U=8,7E-03 U=35	U=1,0E-03 C=70	U=6,0E-04 C=35	U=3,8E-03 C=85
Mod. 3	U=4,0E-04 C=60			

Results

Figures 7(a) and 7(b) show the results when the unavailability and the cost are weighted 0.5, 0.5 and 0.8, 0.2 respectively (for G=100 and N=10). The labels correspond to the TOPSIS classification of the final front.

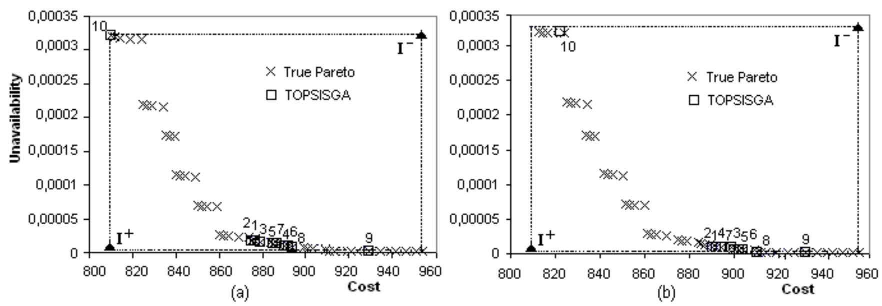

Fig. 7. Non-dominated front found by TOPSISGA.

Notice that the final ordering changes with the weights as expected, providing the DM with a final pre-order according to their preferences. Notice also that, while NSGA-II finds an even final set (figure 8 -for G=100 and N=10-), the TOPSISGA focuses upon a particular region of the efficient frontier (figures 7(a) and 7(b)) and the final result is far different from the one reached by NSGA-II despite of the fact that the objectives were equally weighted. On the other hand, it is evident that it is impossible to obtain a similar classification to the one obtained by TOPSISGA from the final set presented by NSGA-II and vice versa. It raises the question of where is the right moment to introduce preferences and under what criterion?
Finally, the proposed approach was compared with NSGA-II based on the C metric [10], using the efficient frontier as a reference set R. Table 3 reports the percentage (average after ten runs) of the final outcomes (labelled A) dominated by the true Pareto frontier. Figures 9(a) and 9(b) show graphically the results of table 3 for 50 and 100 generations.

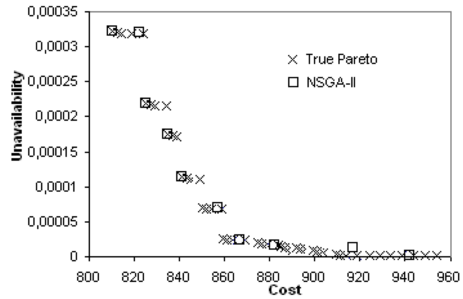

Fig. 8. Non-dominated frontier found by NSGA-II.

Table 3. Metric C(R,A) values for 50, 100 and 250 generations.

Method	G	N=10	N=20	N=30	N=40	N=50	N=100
TOPSISGA	50	21	10	10.33	8.25	6.38	8.15
NSGA-II		46	29.5	10.33	7.25	7.93	9.45
TOPSISGA	100	18	9.5	9.66	6.5	6.97	4.58
NSGA-II		47	28.55	9.66	4.25	6.96	4.97
TOPSISGA	250	18	6	7	4.5	3.31	2.7
NSGA-II		34	23	4.7	4	4.6	2.47

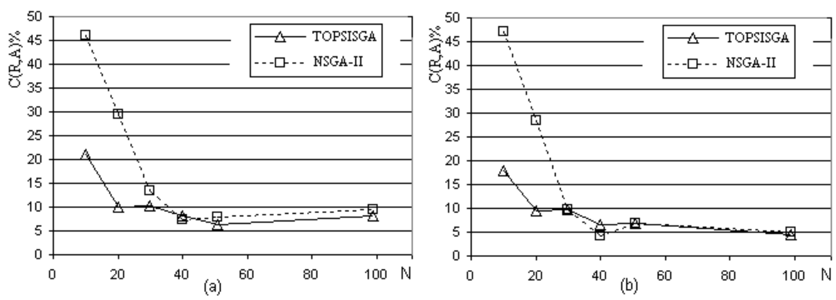

Fig. 9. Graphic view of table 3: (a) 50G, (b) 100G.

5 Conclusions

In many real discrete problems the number of solutions that belong to the Pareto front is unknown. If the specified size of the non-dominated population in the MOEA is less than the number of solutions of the problem, the found front will be incomplete for a posteriori Making Decision. In this work we introduce the MOEA structure TOPSISGA that combines the second generation of MOEAs (we use NSGA-II) with the multiple criteria decision making technique TOPSIS. The conducted experiments show that the proportion of efficient frontier reached by the algorithms is larger using TOPSISGA when the archive size of non-dominated solutions is small, but this difference seems to disappear when the archive size of non-dominated solutions increases. Besides, TOPSISGA focuses the search on the region of interest, giving an order list of alternatives in accordance to the DM's preferences. Nevertheless, it could be convenient to find a balance between the spread over the whole front produced by NSGA-II with the identification and the exploitation of the zone of interest realised by TOPSISGA. Kwangsun Yoon [9] measures the credibility of d_p distance function and obtains: the distance function becomes less specific or less credible as parameter p increases. He recommends the use of d_1 for obtaining the most credible compromise solution from the purely mathematical viewpoint. In TOPSISGA we use the p=2 metric, its influences hasn't been checked so far, it is left for future research.

References

1. Cvetkovic, D., and Coello C. (2005) Human Preferences and Their Applications in Evolutionary Multi-Objective Optimization. In: Yaochu, J. (Ed.): Knowledge Incorporation in Evolutionary Computation (Studies in Fuzziness and Soft Computing, vol. 167), Springer, pages 479-502.
2. Deb, K., Pratap, A., Agrawal, S. and Meyarivan, T. (2002) A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. In: IEEE Transactions on Evolutionary Computation, vol. 6, n. 2, pages 182-197.
3. Deb, K., Sundar, J., Udaya Bhaskara Rao, N., and Chaudhuri, S. (2006) Reference Point Based Multi-Objective Optimization Using Evolutionary Algorithms. In: International Journal of Computational Intelligence Research, vol. 2, n. 3, pages 273-286.
4. Galván B. (1999) Contributions to Fault Tree Quantitative Evaluation. Ph.D. Thesis, Physics Dep., Las Palmas de Gran Canaria University, Canary Islands, Spain, (In Spanish).
5. Greiner D, Galván B, Winter G. (2003) Safety Systems Optimum Design by Multicriteria Evolutionary Algorithms. In: Evolutionary Multi-Criterion Optimization, LNCS 2632, Springer-Verlag. pages 722-736.
6. Hwang C.L., Yoon K. (1981) Multiple Attribute Decision Making: Methods and applications. Springer, Heidelberg.
7. Ishibuchi H. and Narukawa K. (2004) Performance Evaluation of Simple Multiobjective Genetic Local Search Algorithms on Multiobjective 0/1 Knapsack Problems. Proc. 2004 Congress on Evolutionary Computation (CEC'2004), IEEE Service Center, vol. 1, pages 441-448.
8. Jaszkiewicz A. (2002) On the Performance of Multiple Objective Genetic Local Search on the 0/1 Knapsack Problem. A Comparative Experiment. IEEE Transactions on Evolutionary Computation, vol. 6, n. 4, pages 402-412.
9. Kwangsun Yoon (1987) A Reconciliation among Discrete Compromise Solutions. In: Journal of the Operational Research Society, vol. 38, n. 3, pages 277-286.
10. Zitzler E., Deb K. and Thiele L. (2000) Comparison of multiobjective evolutionary algorithms: Empirical results. In: Evolutionary Computation Journal, vol. 8, n. 2, pages 125-148.