OPTIMIZATION USING AN EVOLUTIONARY HYPERPLANE GUIDED APPROACH

Corina Rotar, Dan Dumitrescu and Rodica Lung

Abstract. A new evolutionary technique for multicriteria optimization called Target Hyperplane Evolutionary Algorithm (THEA) is proposed. The originality of the approach consists in the fact that the fitness assignment is realized by using a guiding hyperplane and a new non Pareto optimality concept. THEA has been tested on consecrated benchmarks.

Keywords: Evolutionary Computation, Multicriterial Optimization.

1. Introduction

In the last decades, multicriterial optimization has been a fertile ground in the field of evolutionary computation. Various evolutionary techniques have been successfully developed and applied on complex multicriterial optimization problems. Nevertheless, the subject is still a real challenge for the researchers in the field of evolutionary algorithms both with regard to developing new methods and with the analysis, comparison and evaluation of algorithms performances. A new evolutionary technique for multicriterial optimization called Target Hyperplane Evolutionary Algorithm (THEA) is proposed. THEA introduces a new non Pareto fitness assignment procedure based on guiding hyperplanes providing an intuitive form of guidance towards the desired front for the population.

The proposed approach proves to perform well in numerical experiments for static and dynamical multicriteria optimization problems. Two performance metrics are used to validate the results obtained for the considered test functions.

2. Target Hyperplane Evolutionary Algorithm

Two well-known issues arise regarding population dynamics within evolutionary algorithms: the first refers to preserving population diversity as an important factor for the success of the search process and the second one refers to the convergence towards the problem solutions. Moreover, for complex problems such as multicriteria optimization problems, the balance between the two mentioned components is compulsory. Our main goal is to produce an efficient population dynamics with respect to both diversity and convergence aspects.

Consider the following n objectives minimization problem:

$$\begin{cases} \text{minimize} F(c) = f_1(c), ..., f_n(c) \\ \qquad c \in D \end{cases} \tag{1}$$

where:$F : D \rightarrow R^n$and $D \in R^m$

The use of a hyperplane to guide the search in the objective space is suggested. The guiding hyperplane is built such that the most individuals in the current population are dominated by hyperplane points. Two characteristics of the current population are used to construct the hyperplane:

- An ideal vector in the objective space that is better with regard to every component than all individuals in the current population;

- The most important direction along which the population is spread out in the objective space.

This hyperplane is updated every generation according to the configuration of the current population.

Such a hyperplane would guide the population towards promising areas of the search space. The movement of the population guided by the hyperplane ultimately leads to convergence to the Pareto front.

The guiding hyperplane which is built for each generation is called the *target* of the current generation.

Individuals in the current population are not evaluated through a performance function in relation to the problem's objectives but through how close they are to the guiding hyperplane. Each generation the evolution of the current population towards the corresponding target hyperplane is realized by

focusing on the production of offspring closer to the target and not necessarily on the convergence to the Pareto front. However, the manner of building the target hyperplane for the current population leads to convergence to the Pareto front without using a performance function based explicitly on the Pareto dominance concept.

Regarding the diversity preservation, the target hyperplane can also be used to estimate the diversity degree of the current population. Thus, projecting the individuals of the present population on the target hyperplane could provide an image of the population's distribution. In order to maintain a good distribution during the search process THEA is supplemented with a mechanism that favors increased search in the areas poor in individuals of the search space.

Two measures are associated to each individual in the current population: the first one refers to the distance to the target hyperplane and the second refers to the agglomeration degree of its projection in the guiding hyperplane. Both measures are used to set the balance between exploring and exploiting, which is so necessary for obtaining a good population dynamics.

The mechanism of replacing parents with offspring is designed in such a way as to accelerate population convergence, i.e. to favor the survival of the best solutions found.

2.1. GUIDING HYPERPLANE CONSTRUCTION

The objective function values corresponding to the current population may be considered as a cluster $A = (\alpha^1, \alpha^2, ..., \alpha^d)$ of n-dimensional points. This cluster can be described by means of principal components analysis used in classification theory.

Let m be the mean vector (the centroid) of set A. The most important directions along which cluster A is spread out are given by the eigenvectors of the scatter matrix S, given by

$$S = \sum_{i=1}^{d} \left(\alpha^i - m \right) \left(\alpha^i - m \right)^T \tag{2}$$

Eigenvectors of the matrix S are called the principal directions of the cluster A [2].

The eigenvector u corresponding to the largest eigenvalue gives the most important direction along which the cluster is spread out.

Denote by x^0 the vector of components:

$$x_i^0 = \min \left\{ a_i^k \,|\, k \in \{1, 2, ..., d\} \right\}, i \in \{1, 2, ..., n\} \tag{3}$$
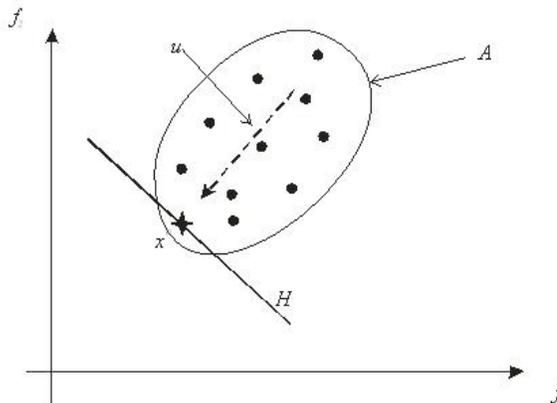

Figure 1: Construction of hyperplane H

Each component x_i^0 of vector x^0 is given by the minimum value of the components a_i^k of the vectors $a^k \in A$. Thus x^0 represents an 'ideal' vector formed with the best components from the vectors of A. The aim is to guide the search of population A toward what could be called the most promising region provided by it. Thus the hyperplane H orthonormal to direction u and passing though x^0 is constructed:

$$H : u^T \left(x - x^0 \right) = 0 \tag{4}$$

Hyperplane H characterizes the set A and it will be considered to be the target of the current population in the objective space. Hyperplane H will be used to evaluate individuals.

2.2. HYPERPLANE BASED FITNESS ASSIGNMENT

Each generation of the evolutionary process, the current target hyperplane H is constructed. Current solutions are evaluated using H. Two features characterize an individual:

- μ_1- measuring the distance to the target hyperplane;

52

- μ_2- measuring the crowding degree its neighborhood .

Consider an individual c in the current population $P(t)$ and $\alpha = (f_1(c), ..., f_n(c))$.

We denote $\mu_1 : D \to R$ the function computing the distance between individual c and the target hyperplane H:

$$\mu_1 (c) = d (H, \alpha) \tag{5}$$

It is essential to maintain the diversity of the cluster. The population will explore efficiently the search space if individuals from less crowded zones will be promoted to survive and evolve. A measure for the degree of crowdedness of a point into its vicinity is proposed. A simple function to evaluate the crowdedness is proposed.

We denote $\mu_2 : D \to R$ the function by which an individual c in the current population can be qualified regarding the crowdedness. Let c' be the projection of c into the target hyperplane H. Function μ_2 is given by:

$$\mu_2 (c) = \min \{d (c', c'_k) , c_k \in P (t)\} \tag{6}$$

Thus μ_2 represents the minimum of the distances between the projection of individual c into H and all other projections from the current population.

REMARK. *Euclidian distance is used.*

2.3. COMPUTING NEXT GENERATION. VERY WEAK OPTIMALITY CONCEPT.

THEA uses binary tournament selection and two genetic operators for variation: crossover and mutation. Offspring obtained by using crossover form an intermediary population P_1, while descendents obtained by mutation will be included into a second intermediary population P_2. Selection for recombination relies on the closeness to the hyperplane (μ_1), while the crowdedness into the target hyperplane (μ_2) gives the fitness values for the selection for mutation. A larger population Γ is obtained, comprising all parents and descendents. The best individuals from the united population will replace the population of parents constituting the next current population. Individuals in Γ are evaluated using a new concept of optimality described next.
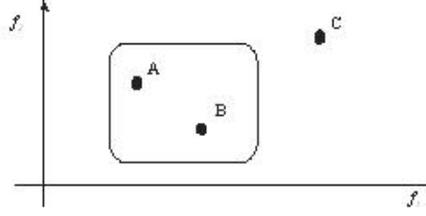
Figure 2: Solutions represented by the points A and B are weakly optimal. Solution C does not verify the weak optimality condition

A new optimality concept called very weak optimality condition for multi-criterial optimization problems is proposed.

DEFINITION 1. *Let us consider problem (1) and denotea and b two solutions from $A \subseteq D, a, b \in A$. We say that a is weakly better than b and we write $a \rhd b$ if there is at least one objective f_i such that $f_i(a) < f_i(b)$.*

REMARK. *Note that $a \rhd b$ does not exclude the relation $a \lhd b$.*

DEFINITION 2. *Let a be a solution from $A \subseteq D$. We state that a is weakly optimal with regard to the set A if we have:*

$$\forall b \in A, \ a \rhd b,$$

which can be also written:

$$\forall b \in A, \exists i \in \{1, 2, ..., n\} \, | f_i(a) < f_i(b)$$

Figure 2 illustrates the introduced weak optimality concept. Several two dimensional points (individuals) are considered. The weakly optimal solutions are plotted in a box.

We denote by $P_1(t)$ the population of descendents obtained by crossover, and by $P_2(t)$ the population of descendents obtained by mutation. The three populations are combined to form a larger population $\Gamma(t) = P(t) \cup P_1(t) \cup P_2(t)$. Each individual c from $\Gamma(t)$ is evaluated using the following function of performance:

$$\Phi(c) = |\{x \in \Gamma(t) \, | c \rhd x\}| \, , c \in \Gamma(t) \tag{7}$$

Thus the performance of an individual is given by the number of individuals from the population $\Gamma(t)$, against which it is better in at least one objective.

The best individuals from $\Gamma(t)$ regarding the performance measure replace the current population $P(t)$, resulting the next generation $P(t+1)$.

Target Hyperplane Evolutionary Algorithm (THEA) is described in the following.

Target Hyperplane Evolutionary Algorithm

1. Randomly generate initial population $P(0)$, t=0.
2. Compute the target hyperplane H for the current population P(t).
3. Establish the population $P\prime(t)$ representing the projections of solutions from $P(t)$ onto the target hyperplane H.
4. Evaluate the population $P(t)$ regarding the target H.
5. $P_1(t) := Crossover(P(t))$
6. $P_2(t) := Mutation(P(t))$
7. Construct population $\Gamma(t) = P(t) \cup P_1(t) \cup P_2(t)$.
8. Evaluate the population $\Gamma(t)$ by using the proposed performance function Φ.
9. Replace the current population $P(t)$ with the best distinct individuals of the population $\Gamma(t)$; t := t+1.
10. If stopping condition is true then stop, else go to step 2.

In order to confer THEA a better generality a different mutation mechanism is proposed. The mutation operator may act in two ways: uniform mutation of a single gene or uniform mutation of all genes, depending on a randomly generated parameter that decides which type of mutation will be applied further.

3. NUMERICAL EXPERIMENTS

Some results obtained using THEA are presented. The well-known test suite of bi-criterial functions ZDT1-ZDT3, ZDT6 [4] is used in order to investigate the performance of the proposed algorithm.

The considered test functions share a particular feature: the set of decision variables is given a structure through which two subsets are defined. Thus each variable is written $X = (X_1, X_2)$, where $X_1 = x_1$ and $X_2 = (x_2, x_3, ..., x_m)$.

We shall treat the two subsets differently when applying variation operators. Thus the probability of selecting a gene also depends on the subset that includes it. More specific X_1 will be given a higher probability for mutation.

The change in the genetic mutation operator positively affects the convergence of the population, drastically reducing the number of generations needed for the population to locate the Pareto front. For example: considering the test function ZDT6, defined for 100 decision variables, giving a probability of 0.1 to the first variable to be altered, as compared to 0.001, - the probability conferred by the standard option of the mutation - the convergence time improves sensibly. In order to avoid loading the algorithm with additional parameters, we chose random generation for each application of the mutation operator, of a probability that the selected gene corresponds to a variable from the first or second subset of decision variables. Figures 3-6 illustrate the results obtained by THEA against those provided by two other evolutionary techniques for multiobjective optimization: SPEA [3] and NSGA [5]. The following evaluation metrics were used: Coverage metric [7] and Spacing metric [8]. Solutions plotted in the figures 2-4 are obtained after a single run of the considered algorithms. The main parameters are:

THEA's main parameters	SPEA, NSGA's main parameters
Population size = 80	Population size = 100
Maximum number of generation =100	Maximum number of generation =250

Problem ZDT1 presents a convex Pareto front:

$$f_1(x) = x_1$$
$$g(x_2, ..., x_m) = 1 + 9 \cdot \sum_{i=2}^{n} \frac{x_i}{m-1}$$
$$h(f_1, g) = 1 - \sqrt{\frac{f_1}{g}}$$
$$f_2(x) = g(x2, ..., x_m) \cdot h(f_2(x), g(x2, ..., x_m))$$

where m=30 and $x_i \in [0, 1]$, $x = (x_1, ..., x_m)$
The true Pareto front is formed with $g(x) = 1$
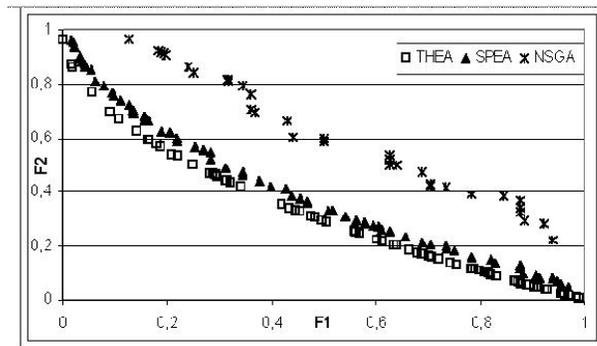Problem ZDT2 presents a concave Pareto front:

Figure 3: Solutions obtained by THEA, SPEA and NSGA for problem ZDT1.

$$f_1(x) = x_1$$
$$g(x_2, ..., x_m) = 1 + 9 \cdot \sum_{i=2}^{n} \frac{x_i}{m-1}$$
$$h(f_1, g) = 1 - \left(\frac{f_1}{g}\right)^2$$
$$f_2(x) = g(x2, ..., x_m) \cdot h(f_2(x), g(x2, ..., x_m))$$

where m=30 and $x_i \in [0,1], x = (x_1, ..., x_m)$
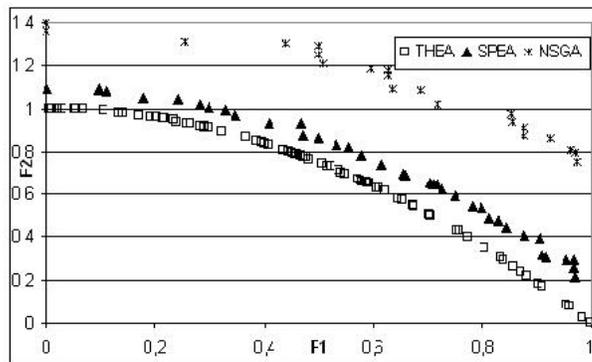The true Pareto front is formed with $g(x) = 1$

Figure 4: Solutions obtained by THEA, SPEA and NSGA for the test problem ZDT2. Entire final population of the THEA algorithm is close to the true Pareto front.

Problem ZDT3 has a discontinuous Pareto front.

57

$$f_1(x) = x_1$$
$$g(x_2, ..., x_m) = 1 + 9 \cdot \sum_{i=2}^{n} \frac{x_i}{m-1}$$
$$h(f_1, g) = 1 - \sqrt{\frac{f_1}{g}} - \frac{f_1}{g} \cdot \sin(10\pi f_1)$$
$$f_2(x) = g(x2, ..., x_m) \cdot h(f_2(x), g(x2, ..., x_m))$$

where m=30 and $x_i \in [0,1]$, $x = (x_1, ..., x_m)$
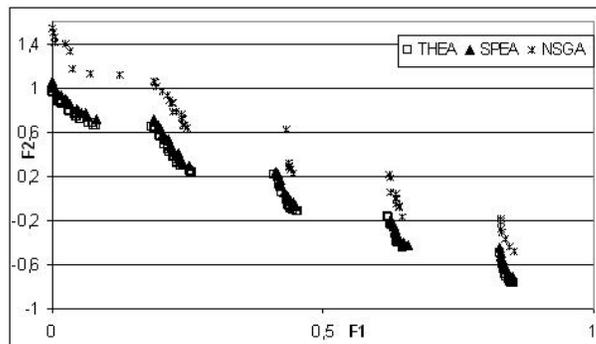The true Pareto front is formed with $g(x) = 1$.

Figure 5: Solutions obtained by THEA, SPEA and NSGA for the test function ZDT3

The values for the considered metrics are summarized next.

Coverage	Zdt1	Zdt2	Zdt3	Coverage	Zdt1	Zdt2	Zdt3
C(Spea,THEA)	0	0	0	C(Nsga,THEA)	0	0	0
C(THEA,Spea)	1	1	0.85	C(THEA,Nsga)	1	1	1

Table 1. Coverage metric - Comparison's results between THEA and other popular algorithms: SPEA and NSGA on the test function considered

Spacing	Zdt1	Zdt2	Zdt3
THEA	0,0163	0,0106	0,0098

Table 2. Spacing Metric for THEA algoritm. The sets of solutions plotted in Figures 2-4 are considered

58

Regarding the more recent and improved SPEA2 and NSGA2 [1], [6] a series of numeric experiments were made. We present next the results obtained for the ZDT6 test problem. For an objective view concerning the performances of the algorithms the following remarks are made:

- The algorithms NSGA2 and SPEA2 use a population size of 100 individuals. The maximum number of evaluations allowed is 100000.
- THEA uses a base population of 100 individuals for 100 generations.
- For each algorithm, solutions found in 30 runs were unified and Pareto dominated solutions were removed.

Figure 5 illustrates the Pareto nondominated solutions found by THEA, SPEA2 and NSGA2 after 30 runs. Without any further numerical comparison, it is obvious that THEA outperforms the other algorithms for the considered test function. THEA's number of fitness evaluation is smaller then 100000 evaluations. Even if the base population size is 100 individuals, a larger population of descendents and parents is evaluated regarding the very weak Pareto optimality concept, therefore the fitness evaluation number if larger than 100x100, but still less than 100.000 which is the maximum number of fitness evaluation allowed for the other two algorithms. Maximum number of fitness evaluations for a single run of the algorithm THEA is 30.000.

Problem ZDT6:

$$f_1(x) = 1 - \exp(-4x_1)\sin^6(6\pi x_1)$$
$$g(x_2, ..., x_m) = 1 + 9 \cdot \sum_{i=2}^{n} \frac{x_i}{m-1}$$
$$h(f_1, g) = 1 - \left(\frac{f_1}{g}\right)^2$$
$$f_2(x) = g(x2, ..., x_m) \cdot h(f_2(x), g(x2, ..., x_m))$$

where m=100 and $x_i \in [0, 1]$, $x = (x_1, ..., x_m)$
The true Pareto front is formed with $g(x) = 1$.

4. CONCLUSIONS

A new approach for evolutionary optimization called Target Hyperplane Evolutionary Algorithm (THEA) is proposed. THEA is based on creating an ideal target hyperplane that guides the search toward the Pareto front without using the traditional Pareto dominance concept.
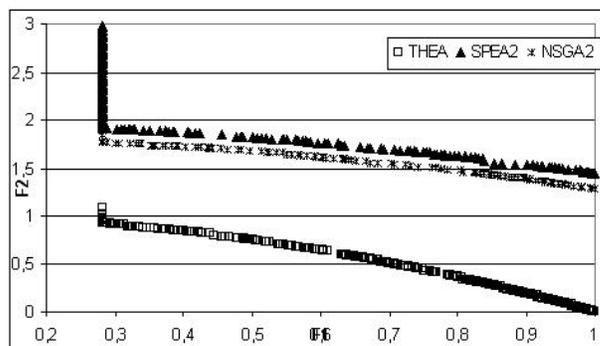
Figure 6: ZDT6. THEA versus SPEA2 and NSGA2 - solutions found by each algorithm in 30 runs. Only Pareto nondominated solutions are plotted

As far as the main features of a good approximation of the Pareto front are concerned, i.e. diversity and convergence, two evaluation functions are used to facilitate the evaluation of each solution regarding the closeness to the current target hyperplane and regarding the crowdedness of the solutions' projections into the hyperplane. The target hyperplane guides the search gradually until the Pareto front is found.

Experimental results show that the approach performs well in various static optimization problems. The target hyperplane, which guides the current population toward the Pareto front together with the new concept of optimality, form an efficient procedure for solving the multicriterial optimization problem.

The proposed algorithm performs better than well-known algorithms for the selected set of test problems. It outperforms SPEA and NSGA for the test suite ZDT1-ZDT3, ZDT6 regarding the closeness to the Pareto front. Generally, SPEA algorithm provides a better-distributed set of solutions.

THEA algorithm also outperforms improved algorithms SPEA2 and NSGA2 for problem ZDT6. Consequently, we can assert that THEA is efficient for solving multicriterial optimization problems.

References

[1] K. Deb, S. Agrawal, A. Pratab, and T. Meyarivan, *A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II*. KanGAL report 200001, Indian Institute of Technology, Kanpur, India, 2000.

[2] R. Duda, Hart, P., *Pattern recognition and scene analysis*,Wiley- Interscience, New York, 1974.

[3] N. Srinivas, K. Deb, *Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms*, Evolutionary Computation 2(3): 221-248 (1994)

[4] E. Zitzler, K. Deb, and L. Thiele, *Comparison of multiobjective evolutionary algorithms: Empirical results*, Tech. Rep. 70, Comput. Eng. Networks Lab. (TIK), Swiss Federal Inst. Technol. (ETH) Zurich, Switzerland, Feb. 1999.

[5] E. Zitzler and L. Thiele, *An Evolutionary Algorithm for Multiobjective Optimization: The Strength Pareto Approach*, Technical Report 43, Computer Engineering and Communication Networks Lab (TIK), Swiss Federal Institute of Technology (ETH) Zurich, Gloriastrasse 35, CH-8092 Zurich, May 1998.

[6] E. Zitzler, M. Laumanns, and L. Thiele, *SPEA2: Improving the Strength Pareto Evolutionary Algorithm* Technical Report 103, Computer Engineering and Communication Networks Lab (TIK), Swiss Federal Institute of Technology (ETH) Zurich, Gloriastrasse 35, CH-8092 Zurich, May 2001.

[7] E. Zitzler, K. Deb, and L. Thiele, *Comparison of Multiobjective Evolutionary Algorithms: Empirical Results*, Evolutionary Computation, 8(2), pages 173-195, 2000.

[8] J. R. Schott, *Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization*, Master's thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, May 1995.

Authors:

Corina Rotar
Department of Computer Science
"1 Decembrie 1918" University of Alba Iulia
N. Iorga 11-13 Alba Iulia
e-mail: crotar@uab.ro

Dan Dumitrescu
Department of Computer Science
"Babes-Bolyai" University of Cluj Napoca
e-mail: ddumitr@cs.ubbcluj.ro

Rodica Lung
Department of Computer Science
"Babes-Bolyai" University of Cluj Napoca
e-mail: srodica@cs.ubbcluj.ro