Annealing-tabu PAES: a multi-objective hybrid meta-heuristic

A. Alcayde, R. Baños, C. Gil, F.G. Montoya, J. Moreno-García & J. Gómez

Department of Rural Engineering, University of Almería, Carretera Sacramento s/n, 04120 Almería, Spain
Department of Computer Architecture & Electronics, University of Almería, Carretera Sacramento s/n, 04120 Almería, Spain
Department of Technologies and Information Systems, University of Castilla-La Mancha, Av/Carlos III s/n, 45071 Toledo, Spain
Department of Languages and Computation, University of Almería, Carretera Sacramento s/n, 04120 Almería, Spain

Published online: 15 Nov 2010.

To cite this article: A. Alcayde, R. Baños, C. Gil, F.G. Montoya, J. Moreno-García & J. Gómez (2011): Annealing-tabu PAES: a multi-objective hybrid meta-heuristic, Optimization: A Journal of Mathematical Programming and Operations Research, 60:12, 1473-1491

To link to this article: http://dx.doi.org/10.1080/02331934.2010.505960

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings,
demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Annealing-tabu PAES: a multi-objective hybrid meta-heuristic

A. Alcayde\(^a\), R. Baños\(^b\)*, C. Gil\(^b\), F.G. Montoya\(^a\), J. Moreno-Garcia\(^c\) and J. Gómez\(^d\)

\(^a\)Department of Rural Engineering, University of Almería, Carretera Sacramento s/n, 04120 Almería, Spain; \(^b\)Department of Computer Architecture & Electronics, University of Almería, Carretera Sacramento s/n, 04120 Almería, Spain; \(^c\)Department of Technologies and Information Systems, University of Castilla-La Mancha, Av/Carlos III s/n, 45071 Toledo, Spain; \(^d\)Department of Languages and Computation, University of Almería, Carretera Sacramento s/n, 04120 Almería, Spain

(Received 16 September 2009; final version received 25 June 2010)

Most real-life optimization problems require taking into account not one, but multiple objectives simultaneously. In most cases these objectives are in conflict, i.e. the improvement of some objectives implies the deterioration of others. In single-objective optimization there exists a global optimum, while in the multi-objective case no optimal solution is clearly defined, but rather a set of solutions. In the last decade most papers dealing with multi-objective optimization use the concept of Pareto-optimality. The goal of Pareto-based multi-objective strategies is to generate a front (set) of non-dominated solutions as an approximation to the true Pareto-optimal front. However, this front is unknown for problems with large and highly complex search spaces, which is why meta-heuristic methods have become important tools for solving this kind of problem. Hybridization in the multi-objective context is nowadays an open research area. This article presents a novel extension of the well-known Pareto archived evolution strategy (PAES) which combines simulated annealing and tabu search. Experiments on several mathematical problems show that this hybridization allows an improvement in the quality of the non-dominated solutions in comparison with PAES, and also with its extension M-PAES.

**Keywords:** hybrid meta-heuristics; multi-objective optimization; PAES; simulated annealing; tabu search

**AMS Subject Classifications:** 68T20; 90C29; 90C59

1. Introduction

Most real optimization problems entail simultaneous optimization of distinct and usually conflicting objectives. The success of meta-heuristics is based on their ability to find efficient solutions in a reasonable runtime. In recent years, the number of multi-objective meta-heuristics (MOMHs) proposed to solve complex multi-objective optimization problems (MOPs) has increased considerably.

*Corresponding author. Email: rbanos@ual.es*
Although the main research interest in this field has been the development of multi-objective evolutionary algorithms (MOEAs), other non-evolutionary methods have also been successfully applied. One of them is the Pareto archived evolution strategy (PAES) [28], which uses a single solution during the optimization process, in combination with an external archive of solutions which stores the best solutions found in the search.

As PAES is a hill climbing-based method, it is not often able to escape from local optima. In order to overcome this drawback, we propose a novel extension of PAES which combines aspects of simulated annealing and tabu search with the aim of improving the quality of the solutions in terms of proximity to the Pareto-optimal front. The performance of this new method has been analysed in several multi-objective test problems, obtaining an improvement in the quality of the solutions in comparison with PAES and M-PAES [29], a memetic algorithm which also uses many ideas from PAES.

The remainder of this article is organized as follows. Section 2 describes some basic multi-objective optimization concepts, while Section 3 gives an overview of multi-objective methods found in the literature. Section 4 details the hybrid algorithm here proposed, while the results obtained when solving some test problems are commented in Section 5. Finally, Section 6 provides the conclusions to this work.

2. Multi-objective optimization: concepts and methods

Multi-objective optimization is an important notion in economics, engineering and social sciences. The goal of Pareto-based multi-objective optimization [18] is to obtain a well-distributed set of non-dominated solutions as close as possible to the Pareto-optimal set, which is unknown for MOPs with large and highly complex search spaces. It is therefore very difficult to describe exactly what a good approximation is in terms of a number of criteria such as closeness to the Pareto set, diversity, etc. [6,10].

2.1. Pareto-based multi-objective optimization: basic concepts

There follows a description of some multi-objective optimization concepts used in Pareto-based multi-objective optimization.

Pareto-based multi-objective optimization is the process of searching for one or more decision variables which simultaneously satisfy all constraints, and optimize an objective function vector that maps the decision variables to \( K \geq 2 \) objectives. Mathematically, the goal is to minimize or maximize a certain mathematical function, called objective function, \( f: \text{minimize/maximize}(f_k(s)) \forall k \in [1,K] \).

Each decision vector or solution \( s = \{s_1,s_2,\ldots,s_m\} \) represents accurate numerical qualities for a MOP. The set of all decision vectors constitutes the decision space. The set of decision vectors that simultaneously satisfies all the constraints is called the feasible set (\( F \)). The objective function vector (\( f \)) maps the decision vectors from the decision space into a \( K \)-dimensional objective space \( Z = \mathbb{R}^K \), \( z = f(s), f(s) = \{f_1(s),f_2(s),\ldots,f_K(s)\}, z \in Z, s \in F \).
Given a MOP with \( K \geq 2 \) objectives, an Order relation between decision vectors is established as follows: Let \( s \) and \( s' \) be two decision vectors. The dominance relations in a minimization problem are:

- \( s \) dominates \( s' \) (\( s < s' \)) iff \( \forall k \in [1, K]: f_k(s) \leq f_k(s') \land \exists k' \in [1, K]: f_{k'}(s) < f_{k'}(s') \)
- \( s, s' \) are indifferent (\( s \sim s' \)) iff neither \( s < s' \) nor \( s' < s \)

Pareto-optimal solution: A solution \( s \) is Pareto-optimal if there is no other \( s' \in F \), such that \( s' < s \). All the Pareto-optimal solutions define the Pareto-optimal set.

Non-dominated solution: A solution \( s \in F \) is non-dominated with respect to a set \( S' \subseteq F \) iff \( \nexists s' \in S' \), verifying that \( s' < s \).

Non-dominated set: Given a set of solutions \( S' \), such that \( S' \subseteq F \), \( ND(S') \) returns the set of non-dominated solutions from \( S' \): \( ND(S') = \{ \forall s' \in S': \nexists s'' \in S': s'' < s' \} \).

Figure 1(b) graphically describes the Pareto-dominance concept for a minimization problem with two objectives (\( f_1 \) and \( f_2 \)), where solution \( s \) dominates the solution \( s' \). Figure 2(a) shows two fronts of non-dominated solutions, whose solutions are indifferent to the other solutions of the same non-dominated front.

### 2.2. Survey of MOMHs

To date, many MOMHs have been proposed to solve MOPs. These methods are usually classified according to whether they are MOEAs, multi-objective local-search algorithms (MOLSAs) or hybrid multi-objective meta-heuristics (hMOMHs). A brief survey of some of the best known approaches found in the literature now follows.

#### 2.2.1. Multi-objective evolutionary algorithms

Conceptually, evolutionary algorithms (EA) base their operation on maintaining a set of solutions (individuals) which evolve in successive iterations (generations) trying to optimize an objective function. In the multi-objective context, it is necessary to design strategies that obtain good solutions for all the objectives.

An intuitive idea to optimize several objectives simultaneously is to use aggregating functions [20], which combine all the objectives into a single function.
using a combination of mathematical operations. A typical approach is the sum of weights of the form:

\[
\text{minimize} \left( \sum_{k=1}^{K} (w_k \cdot f_k(s)) \right),
\]

where \( w_k \geq 0 \) are the coefficients that determine the relative importance of objectives, and often \( \sum_{k=1}^{K} w_k = 1 \).

Schaffer presented a multi-objective EA called the vector evaluated genetic algorithm (VEGA) [31]. VEGA is an adaptation of the single-objective genetic algorithm with a modified selection mechanism. An iteration of VEGA consists of generating \( K \) sub-populations of \( N/K \) individuals, which are responsible for optimizing different objectives. Subsequently, all the sub-populations are joined into a global population and recombination and mutation operators are applied. However, the performance of VEGA is poor due to the lack of diversity.

One of the earliest Pareto-based MOEAs was the multi-objective genetic algorithm (MOGA) [14]. It consists of assigning a rank to each individual of the population according to the number of individuals in the population by which it is dominated, i.e. all non-dominated solutions are assigned rank 1. These fitness values, which are later used in the selection process, are assigned by interpolating from the best to the worst rank. Although improved in recent years by other methods, MOGA has been a reference work in multi-objective optimization.

The niched-Pareto genetic algorithm (NPGA) [22] uses a tournament selection scheme based on Pareto-dominance. NPGA takes two random individuals and compares them against a subset of the entire population, using fitness sharing. The main disadvantage of this method is the necessity of setting the tournament size in addition to the sharing factor. An extension of NPGA, NPGA2, was proposed in [13], and its main novelty is the use of a Pareto-based ranking to classify the individuals of the population.

The non-dominated sorting genetic algorithm (NSGA) [35] is also based on a ranking of the population according to Pareto-dominance relationships. All the non-dominated solutions are assigned rank 1, and are then removed from
the population. The non-dominated solutions of the remaining are assigned rank 2 and are also removed from the population. The same process is repeated for the remaining solutions, until all the individuals are classified. The solutions of the first categories are more likely chosen in the selection process. This scheme gives a good robustness when the number of objectives to optimize increases, but it has some drawbacks, such as its lack of elitism, high-computational complexity, etc. NSGA-II [11] solves these deficiencies.

Other well-known MOEA is the strength Pareto evolutionary algorithm (SPEA) [42]. SPEA uses two sets of solutions: a main population of individuals where evolutionary operators are applied, and an external population consisting of a variable number of non-dominated solutions. In each generation, the best solutions of the main population are stored in the external population with the help of a clustering technique. Binary tournament selection is then performed between both sets. An extended version of SPEA, SPEA2 [41], incorporates density information to assign the fitness. In addition, the size of the external archive is greater or equal to the size of the main population.

The Pareto envelope-based selection algorithm (PESA) [8] is a multi-objective optimizer which uses a local search evolution strategy. PESA also uses an external population of solutions to store promising solutions found in the search process. The novelty of this scheme arises in the control of selection and diversity by using a hyper-grid-based scheme, which allows PESA to outperform NPGA and NSGA in some problems [8]. The same authors also proposed PESA-II [7], which uses a selection technique in which the unit of selection is a hyperbox in objective space. Instead of assigning a selective fitness to an individual, selective fitness is assigned to the hyperboxes which are currently occupied by at least one individual in the current approximation to the Pareto frontier.

2.2.2. Multi-objective local search algorithms

Local search-based methods (LS) try to improve the quality of the solutions by moving from the current solution to another one in the local neighbourhood. Given a solution $s$, its neighbouring solutions are those that are in some sense close to $s$, usually because they can be easily reached from $s$. This section briefly describes some of the LS-based MOMHs, here called MOLSAs, proposed to date.

One of the earliest MOLSAs was Serafini’s multi-objective simulated annealing (SMOSA) [33]. In single-objective simulated annealing (SA) [26] better neighbouring solutions are always accepted, whereas worse solutions are accepted with a certain probability, which is dependent on a parameter called temperature. In a multi-objective context a new solution ($s^*$) is accepted if it dominates the current one ($s$), and it is accepted with a certain probability if it is dominated by $s$. However, there is a special case to bear in mind when both solutions are indifferent ($s \sim s^*$). Serafini suggested several transition probabilities for this case.

Another important SA-based MOMH is Ulungu’s multi-objective simulated annealing (UMOSA) [37]. Unlike SMOSA, where weights in the scalarizing function can be dynamically adapted at runtime, UMOSA executes separate runs by using fixed weights. Thus, each run of UMOSA generates a set of non-dominated solutions. After the execution of the algorithm with different weights for each objective, all the non-dominated sets are joined together in a global set. The benefit
of this method is obtained when the number of executions with different combinations of weights is large enough. However, the runtime of UMOSA increases according to the number of separate runs.

A population-based version of SMOSA is Pareto simulated annealing (PSA) [9]. PSA is based on accepting neighbouring solutions with a certain probability, which, like SMOSA and UMOSA, depends on the temperature parameter. However, while SMOSA and UMOSA use only one solution in the optimization process, PSA uses several, which dynamically modify their weights in the objective function with the aim of assuring adequate dispersion of the non-dominated solutions.

Tabu search (TS) [17] is an optimization method which repeatedly moves from the current solution to the best one in the neighbourhood, while trying to avoid being trapped in a local optimum by maintaining a list of tabu movements. Some extensions of TS for multi-objective optimization are [15,21]. Hansen proposed Multi-objective tabu search (MOTS) [21]. It works with a set of current solutions which are simultaneously advanced towards the non-dominated front (like PSA). These solutions are upgraded using a TS acceptance criterion. In MOTS, the weighting values are adaptively modified during the search process, and the number of solutions changes according to the dominance rank among solutions.

The Pareto archived evolution strategy (PAES) [28] is a well-known MOLSA. The basic version of PAES, also called (1+1)PAES, which uses a single individual as the search agent, and an external set of solutions (archive) which is managed by a grid-based archiving strategy. It works by dividing the objective space occupied by the individuals of the archive into a constant number of rectangular areas, called grid regions. The goal is to obtain a set of non-dominated solutions so that the number of solutions in the same grid is minimized, which denotes that this set is well-spread. Some authors have parallelized PAES (pPAES) [1], where each processor computes the (1+1)PAES and maintains its own archive of non-dominated solutions. Periodically, a synchronous migration operation is applied in order to exchange solutions among processors.

2.2.3. Hybrid multi-objective meta-heuristics

The hybridization of meta-heuristics allows the cooperation of methods with complementary characteristics. Hybridization has several advantages, such as the ability to be close to the true Pareto-optimal front, especially when there are local optima. Interest in the design and implementation of hybrid meta-heuristics (hMOMHs) has increased remarkably in the last decade [36]. Most research effort has been made in single-objective optimization problems, mostly by combining population-based methods with local search strategies. However, hybridization has not yet been extensively applied in the multi-objective context. There follows a description of several effective hMOMHs found in the literature.

Multi-objective genetic local search (MOGLS) consists of using scalar fitness functions with random weights for the selection of parents and local search for their offspring. Jaszkiewicz [25] improved the performance of Ishibuchi and Murate’s multi-objective genetic local search (IMMOGLS) [24] by modifying its selection mechanism of parents. Jaszkiewicz’s MOGLS randomly selects a pair of parents from a pre-specified number of the best solutions with respect to the scalar fitness function with the current weights. Taking into account the experimental results
obtained by Jaszkiewicz [25], his MOGLS algorithm performed significantly better than IMMOLGS in the case of multi-objective travelling salesperson problems.

*Genetic tabu search (GTS)* [4] combines the global nature of genetic search and the local nature of tabu search. Results obtained in the multi-objective knapsack problem demonstrate the good performance of GTS in comparison to NSGA and Jaszkiewicz’s MOGLS.

*Memetic-PAES (M-PAES)* [29] is an extension of PAES that uses a population of solutions and a crossover operator which is used to recombine solutions found by the PAES procedure. Promising solutions found in the search are also stored in external archives. Results obtained in the 0/1 knapsack problem outperformed PAES and were similar to those obtained by SPEA [42].

Hu et al. [23] combined sequential quadratic programming (SQP) with SPEA and NSGA-II. This hybridization allows an improvement in the quality of the solutions in several mathematical test functions. In addition, authors also conclude that this hybrid approach involves a better diversity of solutions.

*Mixed spreading between PESA and NSGA-II (msPESA)* [16] uses, such as PESA, a hypergrid-based scheme which is dynamically re-sized according to the extreme values of the non-dominated solutions found in each objective. However, while PESA divides the search space in a $K$-dimensional grid of boxes, msPESA divides it in a $K-1$ dimensional one. On the other hand, msPESA uses a variation of the selection scheme proposed in NSGA-II, where only the best front of the internal population is considered for inclusion in the external population.

*Multi-objective simulated annealing and tabu search (MOSATS)* [3] is a population-based hMOMH that combines SA and TS during the optimization process. The acceptance of new solutions is based on Pareto-dominance relations among solutions. MOSATS also uses an external archive which contains the promising solutions found in the main population. Results obtained [3] in the graph partitioning problem demonstrated the good performance of the algorithm.

Recently, some authors have hybridized SA with other meta-heuristics, such as, particle swarm optimization (PSO-SA) [5] or ant colony optimization (ACO-SA) [34]. TS has also been combined with memetic algorithms [30].

## 3. Annealing-tabu PAES

This section describes how to improve the quality of the PAES by using SA and TS. In contrast with population-based MOMHs [3,9,11,41], PAES performs the multi-objective search by using a single solution, which is improved by applying local movements. Non-dominated solutions found in the search are stored in a secondary population by the adaptive grid archiving (AGA) strategy [27]. This strategy works by dividing up the objective space occupied by the individuals of the population into $R \times R$ different rectangular grids, also called regions. The number of regions is constant during the search process. However, as the location of the solutions changes during this process, the space, location and size of the grid regions in the objective space also vary in runtime. The goal is to obtain a set of non-dominated solutions (ND) so that the number of solutions in the same grid (crowding degree) is minimized indicating that ND is well-spread. This grid-based
archiving technique helps to expand the non-dominated solutions in order to build an extensive set.

Knowles proposed in [27] that PAES could be extended in two different ways. One of these suggested extensions is based on using SA, and the other one on TS. Thus, the aim of this work is double: first we have implemented annealing-PAES (A-PAES) and tabu-PAES (T-PAES) to determine their improvement in comparison with PAES. On the other hand, we have also designed and implemented a hybrid version, Annealing-Tabu-PAES (AT-PAES), which combines both of them, and whose performance is compared with PAES, A-PAES and T-PAES.

As commented above, a common drawback when solving optimization problems is that hill-climbing methods, such as PAES, are not able to escape from local optima. SA has the ability to escape from local optima in function of the current temperature \( t \). We have therefore designed A-PAES with the aim of avoiding this disadvantage. Figure 1 compares the general operation of PAES and A-PAES. Figure 1 shows the case when PAES is applied, i.e. a neighbouring solution \( s^* \) substitutes the current solution \( s \) if \( s^* < s \). Figure 1 displays the criterion used by A-PAES. In this case, the neighbouring solution, \( s^* \), is accepted if it does not dominate \( s \), but \( s' \), which is calculated by the equation \( f(s)^*(1 + w*t) \), where \( w \) changes according to the percentage of use of the external archive (ND) and \( t \) is the current temperature. Both parameters, \( t \) and \( w \) oscillate in the interval [0.0, 1.0]. Then, when the temperature is \( t = 1.0 \), and the external archive is empty \( (w = 1.0) \) the function \( Annealing\_Test(t, s) \) returns a reference solution \( s' \) which has values equal to the double of \( s \) in both objectives (freedom degree is 100%). On contrary, when the temperature is \( t = 0.0 \) or ND is full \( (w = 0.0) \), A-PAES only accepts neighbouring solutions non-dominated by the current one (Figure 1a).

When the search space is explored, the use of a limited neighbourhood can cause the appearance of cycles. An interesting way to reduce this drawback is the use of TS. In our implementation a tabu list, \( TL_s \), is used for each solution \( s \). Thus, if the solution \( s^* \) comes from \( s \), this movement is included in the the tabu list. In each iteration of T-PAES and AT-PAES, the movements included in the tabu list are forbidden. When \( TL_s \) is updated, tabu movements included in an iteration prior to the tabu memory are reset to non-tabu.

Algorithm 1: Annealing-Tabu PAES (AT-PAES)

\begin{verbatim}
Input: T_start; T_cr; TL_min; RT;
\end{verbatim}

\begin{verbatim}
s ← random_solution; ND ← s; t ← T_start;
Do{
    s* ← Neighbour(s);
    )While (movement(s, s*) ∈ tabu_list);
    s' ← Annealing_Test(t, s);
    If (s' < s*) { discard s*; }
    Else {
        If (s* < s') { s ← s*; ND ← re-sizeGrid(s*); ND ← ND \cup s*; }
        Else {
            If (\exists s'' ∈ ND : s'' < s*) { discard s*; }
            Else {
                s'' ← solution of ND with highest crowding degree;
            }
        }
    }
}\end{verbatim
Algorithm 1 describes the general procedure of AT-PAES. The initial solution is randomly obtained and included in the external archive of non-dominated solutions (ND), and the temperature is initialized according to the input parameters. While the stop criterion is not fulfilled, AT-PAES works as follows. A neighbouring solution is obtained by applying a movement not included in the tabu list (TL). Then, the tabu list is updated by including the last movement, which will be maintained in this list during a given number of iterations, according to the memory length (TL_{ml}). On the other hand, the movements included TL_{ml} iterations prior to TL are removed. Subsequently, the Annealing_Test(t, s) function returns a temporal solution s', which is dependent on the current temperature and the solution s. Then, while PAES compares solution s* with s in terms of Pareto-dominance, AT-PAES compares s* with solution s'. If s' < s*, s* is discarded. If not, there are three remaining possibilities to consider. The first possibility is that s* < s'. In this case, s* substitutes the current solution, s, the grid is re-sized if necessary and s* is included in ND. The second and third cases correspond to when s* ~ s'. If there is any solution in the archive that dominates s*, s* is discarded. Otherwise, if s* ~ s and to all the solutions located in ND, s* substitutes that solution in ND whose crowding degree is higher than the degree of s*. Finally, temperature is decreased according to the cooling rate, T_{cr}. When the stop condition is fulfilled, the set of non-dominated solutions (ND) is returned.

4. Experimental analysis

4.1. Test problem suite

The six bi-objective test problems [40] used in this study are now described. These test problems, which are often used to evaluate multi-objective meta-heuristics [19,40], are structured in the same manner and consist of three functions f_1, g, h:

Minimize F(x) = (f_1(x_1), f_2(x))
subject to: f_2(x) = g(x_2, \ldots, x_m) * h(f_1(x_1), g(x_2, \ldots, x_m))
where x = (x_1, \ldots, x_m)

The function f_1 is a function of x_1 only, g is a function of the remaining m-1 variables, and the parameters of h are the function values of f_1 and g. The test functions differ in these three functions. Each of the functions is a two-objective minimization problem on m parameters.

Test Problem ZDT1 (convex Pareto-optimal front)

\[
f_1(x_1) = x_1 \\
g(x_2, \ldots, x_m) = 1 + \frac{g}{(m-1)} \sum_{i=2}^{m} x_i
\]
Euclidean distance between extreme solutions, do not reveal the true distribution of the front. Other alternatives, such as the maximum spread metric, which measures the distance metric, etc. The second category, which includes spacing metrics, has the disadvantage that a set of non-dominated solutions with a good spacing metric does not necessarily mean a good distribution of solutions in the entire Pareto-optimal front. Several authors [12,38] have analysed a large variety of performance metrics for MOPs. In these studies, metrics are classified into two main categories, depending on whether they evaluate the closeness to the Pareto-optimal front or the diversity in the obtained front. In the first category, almost all metrics suggested for bi-objective MOPs can be applied to problems with more objectives. In this category we can find the error ratio metric, set coverage metric, hyper-volume metric, generational distance metric, etc. The second category, which includes spacing metrics, has the disadvantage that a set of non-dominated solutions with a good spacing metric does not necessarily mean a good distribution of solutions in the entire Pareto-optimal front. Other alternatives, such as the maximum spread metric, which measures the Euclidean distance between extreme solutions, do not reveal the true distribution of

\begin{align*}
h(f_1(x_1), g(x_2, \ldots, x_m)) &= 1 - \sqrt{\frac{g}{h}} \quad \text{where } m = 30, \ x_i \in [0,1].
\end{align*}

Test problem ZDT2 (non-convex Pareto-optimal front):
\begin{align*}
f_1(x_1) &= x_1 \\
g(x_2, \ldots, x_m) &= 1 + \frac{g}{(m-1)} \sum_{i=2}^{m} x_i \\
h(f_1(x_1), g(x_2, \ldots, x_m)) &= 1 - \left(\frac{x_1}{g}\right)^2 \quad \text{where } m = 30, \ x_i \in [0,1].
\end{align*}

Test problem ZDT3 (discontinuous Pareto-optimal fronts):
\begin{align*}
f_1(x_1) &= x_1 \\
g(x_2, \ldots, x_m) &= 1 + \frac{g}{(m-1)} \sum_{i=2}^{m} x_i \\
h(f_1(x_1), g(x_2, \ldots, x_m)) &= 1 - \sqrt{\frac{g}{h}} - \frac{g}{h} \sin(10\pi f_1) \\
\text{where } m = 30, \ x_i \in [0,1].
\end{align*}

Test problem ZDT4 (multi-modal problem):
\begin{align*}
f_1(x_1) &= x_1 \\
g(x_2, \ldots, x_m) &= 1 + 10(m-1) + \sum_{i=2}^{m} (x_i^2 - 10 \cos(4\pi x_i)) \\
h(f_1(x_1), g(x_2, \ldots, x_m)) &= 1 - \sqrt{\frac{g}{h}} - \frac{g}{h} \sin(10\pi f_1) \\
\text{where } m = 10, \ x_1 \in [0,1], \ x_2, \ldots, x_m \in [-5,5].
\end{align*}

Test problem ZDT5 (deceptive problem):
\begin{align*}
f_1(x_1) &= 1 + u(x_1) \\
g(x_2, \ldots, x_m) &= \sum_{i=2}^{m} v(u(x_i)) \\
(f_1, g) &= \frac{1}{h} \\
\text{where } u(x_i) \text{ gives the number of ones in the bit vector } x_i, \\
v(u(x_i)) &= 2 + u(x_i) \text{ if } u(x_i) < 5, \\
v(u(x_i)) &= 1 \text{ if } u(x_i) = 5, \\
m &= 11, \ x_1 \in \{0,1\}^{50}, \ x_2, \ldots, x_m \in \{0,1\}^5.
\end{align*}

Test problem ZDT6 (non-uniformity problem):
\begin{align*}
f_1(x_1) &= 1 - \exp^{-4x_1} \sin^6(6\pi x_1) \\
g(x_2, \ldots, x_m) &= 1 + 9 \left(\sum_{i=2}^{m} \frac{x_i}{m-1}\right)^{0.25} \\
h(f_1(x_1), g(x_2, \ldots, x_m)) &= 1 - \left(\frac{f_1}{g}\right)^2 \\
\text{where } m = 10, \ x_i \in [0,1].
\end{align*}

4.2. Performance measures

Several authors [12,38] have analysed a large variety of performance metrics for MOPs. In these studies, metrics are classified into two main categories, depending on whether they evaluate the closeness to the Pareto-optimal front or the diversity in the obtained front. In the first category, almost all metrics suggested for bi-objective MOPs can be applied to problems with more objectives. In this category we can find the error ratio metric, set coverage metric, hyper-volume metric, generational distance metric, etc. The second category, which includes spacing metrics, has the disadvantage that a set of non-dominated solutions with a good spacing metric does not necessarily mean a good distribution of solutions in the entire Pareto-optimal front. Other alternatives, such as the maximum spread metric, which measures the Euclidean distance between extreme solutions, do not reveal the true distribution of
intermediate solutions. Zitzler and Thiele [42] proposed applying the hyper-volume metric in combination with the relative set coverage metric to solve the 2/3/4-objective knapsack problem. Both performance measures have been used in our empirical study, in addition to Schott’s spacing metric [32]. All of these are often used to compare multi-objective algorithms [19] and they are described below.

**Set coverage (SC):** Given two sets of non-dominated solutions, $X$, $X'$, function $SC$ maps the ordered pair $(X, X')$ to the interval $[0, 1]$. The value $SC(X, X') = 1$ means that all solutions in $X'$ are dominated by at least one solution of $X$. Figure 2(a) shows that set $X'$ covers most of the solutions of $X$ in a minimization problem.

\[
SC(X, X') = \frac{|s' \in X'; \exists s \in X: s < s'|}{|X'|}.
\]

**Hyper-volume (HV):** Another metric often used to determine the quality of non-dominated sets is hyper-volume. Let \(X = (s_1, s_2, \ldots, s_n)\) be a set of non-dominated solutions, then the function $HV(X)$ returns the area of objective space that lies between a chosen reference point ($x_r$) that acts as a corner to a hyper-cube, and the set $X$. For a minimization problem, the coordinates of $x_r$ are those corresponding to the maximum values in each objective of all the non-dominated solutions of the fronts to be compared. This metric was first proposed by Zitzler and Thiele [42], and called size of the space covered. Later, Coello et al. [6] described this metric using the Lebesgue measure $\Lambda$ [39]. Thus, non-dominated fronts with solutions close to the (unknown) Pareto-optimal front, and having extreme solutions will obtain a higher hyper-volume. Figure 2(b) and (c) shows the hyper-volume of $X$ and $X'$. In this article, the hyper-volume values obtained by the non-dominated fronts have been normalized taking as reference the area of the hyper-cube (i.e. the rectangle enclosed by the points $(0, 0)$ and $(x_{r1}, x_{r2})$):

\[
HV(X) = \frac{1}{|X|} \sum_{i=1}^{|X|} (d_i - d)\]

**Schott’s spacing (SS):** Given a set of solutions, $X$, the function $SS$ measures the variance of the distance of each solution located in $X$ to its closest non-dominated solution in $X$. The following equation describes this metric, where $d_i$ is the minimum euclidean distance between solution $i$ and any other included in $X$, $\bar{d}$ is the average of all $d_i$ and $|X|$ is the cardinal of $X$. Therefore, the ideal distribution of solutions in $X$ is obtained when $SS = 0$:

\[
SS(X) = \sqrt{\frac{1}{|X| - 1} \sum_{i=1}^{|X|} (d_i - \bar{d})^2}.
\]

### 4.3. Parameter settings

As commented above, the aim of this empirical analysis is to evaluate the advantages provided by the simultaneous application of SA and TS in comparison with the use of hill climbing, as in PAES. Furthermore, the memetic version of PAES, M-PAES, is also included in this study. To compare the results of different executions, the stop criterion in the experiments cannot be fixed to a number of iterations, since each MOMH has particular characteristics (e.g. M-PAES uses a population of agents, while PAES, Annealing-PAES, Tabu-PAES, AT-PAES use a single solution) that could result in differences in their runtimes. Given this circumstance, one way to
guarantee the equality of conditions in the comparison is to maintain the same runtime for all the methods. Therefore, the stop condition for all the algorithms analysed here is established when the runtime is 30 s ($RT = 30$). Regarding the representation of the solutions, each one of the $m$ parameters of the test problems are encoded using real numbers. A new solution $s^*$ is obtained from the current one $s$ by applying a mutation to one of these $m$ parameters, i.e. a random parameter $p$ ($p \in [1, m]$) is selected and it is increased or decreased. On the other hand, the tabu list ($TL$) is an array of $m$ integers such that $0 \leq TL[p] \leq TL_{ml} \forall p \in [1, m]$. If, after applying a mutation, the new solution $s^*$ is accepted, the tabu list is then updated by adding this movement, i.e. the parameter $p$ has increased/decreased, which means that this parameter cannot be increased/decreased again during the following $TL_{ml}$ iterations. The memory length of the tabu list in our experiments is $TL_{ml} = 5$ iterations. The initial temperature is set to $T_{start} = 1.0$, and it is decreased using a cooling rate $T_{cr} = 0.995$. Previous studies have demonstrated the advantages of using cooling rate values close to one [2]. The number of agents in M-PAES is set to 50, while the size of the external archive for all the methods has been set to $ND_{size} = 500$. The adaptive grid archiving strategy used in PAES and its extensions presented here use the same grid dimensions: $R \times R = 10 \times 10$ regions. Taking into account the stochastic nature of the algorithms implemented here, 15 independent runs of each method have been performed using the same parameter described above. Thus, the median non-dominated front of each method, in terms of hyper-volume, has been chosen as the representative front of each algorithm and used in the comparison among them.

4.4. Results and discussion

The first aspect we analyse is the quality of the solutions obtained by PAES, A-PAES, T-PAES, AT-PAES and M-PAES considering the hyper-volume. The results displayed in Table 1 show that the normalized hyper-volume is similar in some test functions, as in ZDT1, ZDT2, ZDT3 and ZDT6. However, in ZDT4, AT-PAES obtains the best results; while in ZDT5, M-PAES is slightly better than the other methods.

The hyper-volume metric is able to include both the closeness of the solutions to the Pareto-optimal set and, to some extent, the spread of the solutions across objective space in a single scalar value [39]. However, this metric strongly penalizes those sets that, having good solutions in terms of closeness to the Pareto-optimal set, do not have good extreme solutions. For instance, Figure 2(a) shows that, though
most of the solutions of the non-dominated set $X$ are dominated by the solutions of set $X'$ while none of the solutions of $X'$ is dominated by the solutions of $X$, the hyper-volume of $X$ (Figure 2b) is higher than the one obtained by $X'$ (Figure 2c). This is why set coverage is used to directly establish a comparison between two fronts in terms of Pareto-dominance. Table 2 shows the set coverage relationships among the algorithms for all the test functions detailed above. Table 3 shows the average results obtained in all test functions. The first conclusion we obtain is that A-PAES and T-PAES both outperform PAES. On average, 21.4% of the non-dominated solutions obtained by PAES are dominated by solutions of A-PAES and T-PAES, while PAES is able to dominate only 4.5% and 3.8% of the solutions of these methods, respectively. The second conclusion obtained is that the simultaneous use of SA and TS outperforms the quality of the solutions when they are applied independently. On average, 19.3% and 20.3% of the non-dominated solutions obtained by A-PAES and T-PAES are dominated by solutions of AT-PAES, while these methods are able to dominate only 5.0% and 4.5% of the solutions of the latter, respectively. On the other hand, AT-PAES also outperforms M-PAES in
terms of set coverage, which reinforces the conclusions about the good performance of this hybridization.

Figure 3 shows the results obtained by all the methods in the optimization of the test function ZDT3. It can be observed that all the methods are very close to the Pareto-optimal front, although PAES is not able to obtain good extreme solutions in both objectives, as given in Figure 4. Figure 5 shows the non-dominated sets obtained by all the algorithms in the test function ZDT4. This function is very difficult to optimize as it has a large number of local Pareto-optimal fronts. It can be seen that the non-dominated solutions obtained by AT-PAES dominate most of those of the other algorithms. On the other hand, Figure 6 shows the same comparison in the test function ZDT5. In this case all the approaches are close to the Pareto-optimal front.

Table 4 shows information about the distribution of solutions using Schott’s spacing. It can be seen that in easy test problems, such as ZDT1 and ZDT2, all the methods are well distributed. However, when the problem complexity and the objective space grow, the solutions are often more disperse. Table 4 shows that AT-PAES is often well distributed in comparison to the other methods, while M-PAES obtains low-quality distributions. In some cases, poor distributions occur because the non-dominated fronts contain very extreme solutions, which means that dispersion is worse, such as for example M-PAES in ZDT5, as shown in Figure 6.

Finally, the evolution of the five MOMHs has been analysed according to the runtime. Due to the sequential (non-parallel) implementation of the five algorithms, the metrics cannot be directly applied in runtime. In particular, the set coverage metric compares two non-dominated fronts, while the hyper-volume metric uses a common reference point \( (x_r) \) for all the non-dominated fronts to be compared.
Figure 4. Solutions of each algorithm in ZDT3.

Figure 5. Global comparison in ZDT4.
which is why in both cases the metrics can only be applied after execution is completed. For this reason, the quality of the non-dominated fronts of the five MOMHs have been analysed using different runtimes. In particular, Figure 7 shows the hyper-volume obtained by the five algorithms when the algorithm is applied during $RT = \{2, 5, 10, 20, 30\}$ seconds in ZDT1 and ZDT4. While in the first problem all the methods obtain similar results, their performance differs in the complex ZDT4 problem. In both cases it is observed that, in general, the higher the runtime is, the higher the hyper-volume. It is also observed that annealing-based approaches (A-PAES and AT-PAES) obtain poor results when $RT$ is low, which is because the temperature is still far from zero, and therefore the probability of accepting dominated solutions is still high. On the contrary, when the runtime increases, the temperature reaches values close to zero and therefore the probability of rejecting dominated solutions increases.

Table 4. Schott’s spacing (metric $SS$) values for PAES, A-PAES, T-PAES, AT-PAES and M-PAES.

<table>
<thead>
<tr>
<th></th>
<th>ZDT1</th>
<th>ZDT2</th>
<th>ZDT3</th>
<th>ZDT4</th>
<th>ZDT5</th>
<th>ZDT6</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAES</td>
<td>0.0027</td>
<td>0.0022</td>
<td>0.0055</td>
<td>0.0169</td>
<td>0.6346</td>
<td>0.0292</td>
</tr>
<tr>
<td>A-PAES</td>
<td>0.0026</td>
<td>0.0017</td>
<td>0.0046</td>
<td>0.0365</td>
<td>0.6671</td>
<td>0.0568</td>
</tr>
<tr>
<td>T-PAES</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0041</td>
<td>0.1209</td>
<td>0.6672</td>
<td>0.0120</td>
</tr>
<tr>
<td>AT-PAES</td>
<td>0.0043</td>
<td>0.0022</td>
<td>0.0045</td>
<td>0.0596</td>
<td>0.2672</td>
<td>0.0112</td>
</tr>
<tr>
<td>M-PAES</td>
<td>0.0022</td>
<td>0.0019</td>
<td>0.0050</td>
<td>0.2476</td>
<td>1.8241</td>
<td>0.1135</td>
</tr>
</tbody>
</table>

Figure 6. Global comparison in ZDT5.
5. Conclusions

This article has analysed how to improve the quality of the PAES for multi-objective optimization using SA and TS. To this end, we have implemented three extensions of PAES. The first two are A-PAES and T-PAES proposed by Knowles but not implemented until now. A hybrid procedure combining both extensions (AT-PAES) has also been implemented. Although an order of merit between different algorithms is very difficult to establish, using accurate metrics provides interesting conclusions on their performance. Results obtained in a test suite of six bi-objective test problems indicate that A-PAES and T-PAES outperform PAES in most cases. In general, it can be observed that hybridizing SA and TS (AT-PAES) improves on the result obtained by their separate use. AT-PAES also obtains solutions of equal or higher quality than M-PAES, particularly in test function ZDT4. These conclusions reinforce the perception that hybridization is a good tool to improve Pareto-based multi-objective optimization algorithms. The new hybrid procedure presented here can also be applied to other optimization problems.

Acknowledgements

This work was supported by grants P2007-TIC-02988 (CICE, Junta de Andalucia) and TIN2008-01117 (Spanish Ministry of Science and Innovation).

References


