Improved Metaheuristic Based on the R2 Indicator for Many-Objective Optimization

Raquel Hernández Gómez
Computer Science Department
CINVESTAV-IPN
México, D.F. 07360
rhernandez@computacion.cs.cinvestav.mx

Carlos A. Coello Coello
Computer Science Department
CINVESTAV-IPN
México, D.F. 07360
ccoello@cs.cinvestav.mx

ABSTRACT

In recent years, performance indicators were introduced as a selection mechanism in multi-objective evolutionary algorithms (MOEAs). A very attractive option is the $R2$ indicator due to its low computational cost and weak-Pareto compatibility. This indicator requires a set of utility functions, which map each objective to a single value. However, not all the utility functions available in the literature scale properly for more than four objectives and the diversity of the approximation sets is sensitive to the choice of the reference points during normalization. In this paper, we present an improved version of a MOEA based on the $R2$ indicator, which takes into account these two key aspects, using the achievement scalarizing function and statistical information about the population's proximity to the true Pareto optimal front. Moreover, we present a comparative study with respect to some other emerging approaches, such as NSGA-III (based on Pareto dominance), Δ_p-DDE (based on the Δ_p indicator) and some other MOEAs based on the $R2$ indicator, using the DTLZ and WFG test problems. Experimental results indicate that our approach outperforms the original algorithm as well as the other MOEAs in the majority of the test instances, making it a suitable alternative for solving many-objective optimization problems.

Categories and Subject Descriptors

I.2.8 [**Computing Methodologies**]: Artificial Intelligence—*Problem Solving, Control Methods, and Search.*

Keywords

Multi-objective optimization; performance measures; genetic algorithms.

1. INTRODUCTION

A wide variety of real-world problems require the simultaneous optimization of several (often conflicting) objectives, whose solution involves finding a set of trade-off solutions. These are called Multi-objective Optimization Problems (or MOPs) and for solving them, Multi-objective evolutionary algorithms (MOEAs) have been proven effective [2], especially when no gradient information is assumed. An important trend in MOEAs is the incorporation of performance indicators as a selection mechanism. This has been mainly motivated by the fact that Pareto-based selection schemes do not perform properly when dealing with MOPs having four or more objectives (the so called many-objective optimization problems [11]).

The indicator that has been most commonly adopted in the selection mechanism of a MOEA has been the hypervolume [22], being the only unary indicator known to be Pareto compliant. However, its drawback is that its computational cost grows exponentially as the number of objectives increases. An alternative is the $R2$ indicator [8], which is correlated with the hypervolume [19], but requires a much lower computational cost. Moreover, the use of the $R2$ indicator is recommended [13] for many-objective problems, since it simultaneously evaluates convergence and diversity of an approximation set. One of the first MOEAs that have been proposed based on this indicator is the Many Objective Metaheuristic Based on the $R2$ indicator (MOMBI) [9]. Nevertheless, in some test instances it experimented loss of diversity for high dimensionality. In this paper, we address this problem and propose an improved version, called MOMBI-II. Furthermore, we provide a comparative study with respect to other recent MOEAs, such as Δ_p - Differential Evolution (Δ_p-DDE) [16], which is based on the Δ_p indicator; the Nondominated Sorting Genetic Algorithm III (NSGA-III)[4], which relies on Pareto dominance and a niching strategy; and three MOEAs based on the $R2$ indicator: $R2$-IBEA [15], $R2$-MOGA [7] and the original MOMBI.

The remainder of this paper is organized as follows. Section 2 introduces basic concepts that will be used in the paper. Section 3 presents an overview of emerging MOEAs that can handle many-objective problems. Section 4 describes in detail our proposal. Then, we present the experimental results in Section 5. Our conclusions and some paths for future research are provided in Section 6.

2. PRELIMINARIES

The formal definition of a MOP is given by:

$$\text{Minimize} \quad \vec{f}(\vec{x}) := (f_1(\vec{x}), f_2(\vec{x}), \dots, f_m(\vec{x})) \qquad (1)$$

$$\text{subject to} \quad x \in \mathcal{S}, \qquad (2)$$

GECCO '15, July 11 - 15, 2015, Madrid, Spain
© 2015 ACM. ISBN 978-1-4503-3472-3/15/07…$15.00
DOI: http://dx.doi.org/10.1145/2739480.2754776

where $\vec{x}$ is the *decision variable vector*, $\mathcal{S} \subset \mathbb{R}^n$ is the *feasible region set* and $\vec{f}(\vec{x})$ is the vector of m (≥ 2) *objective functions* ($f_i : \mathbb{R}^n \rightarrow \mathbb{R}$).

The aim is to seek from among the set of all numbers which satisfy the constraints functions defined in equation (2) the particular set $\vec{x^*}$ which yields the optimum values of all the objective functions. In the following, we provide concepts related to Pareto optimality. Unless otherwise stated, let be $\vec{x}, \vec{y} \in \mathcal{S}$.

Definition 1. A decision vector $\vec{x}$ is *Pareto optimal* if there does not exist another decision vector $\vec{y}$ such that $f_i(\vec{y}) \leq f_i(\vec{x})$ for all $i \in \{1, \ldots, m\}$ and $f_j(\vec{y}) < f_j(\vec{x})$ for at least one index j.

Definition 2. The *Pareto Optimal Set* $\mathcal{P}^*$ is defined by:

$$\mathcal{P}^* := \{\vec{x} \mid \vec{x} \text{ is Pareto optimal}\}.$$

Definition 3. The *Pareto Optimal Front* $\mathcal{PF}^*$ is given by:

$$\mathcal{PF}^* := \{\vec{f}(\vec{x}) \in \mathbb{R}^m \mid \vec{x} \in \mathcal{P}^*\}.$$

Definition 4. A solution $\vec{x}$ is said to *weakly dominate* a solution $\vec{y}$ (denoted by $\vec{x} \preceq \vec{y}$), if and only if $\forall i \in \{1, \ldots, m\}$, $f_i(\vec{x}) \leq f_i(\vec{y})$.

Definition 5. A solution $\vec{x}$ is said to *dominate* a solution $\vec{y}$ (denoted by $\vec{x} \prec \vec{y}$), if and only if $\vec{x} \preceq \vec{y} \wedge \vec{f}(\vec{x}) \neq \vec{f}(\vec{y})$.

Definition 6. A vector of decision variables $\vec{x} \in X$ is *nondominated* with respect to the set $X \subseteq \mathcal{S}$ if there does not exist another vector $\vec{y} \in X$ such that $\vec{f}(\vec{y}) \prec \vec{f}(\vec{x})$.

Definition 7. The set of *nondominated solutions* relative to the set X is expressed by $\{\vec{x} \in X \mid \vec{x} \text{ is nondominated}\}$.

The Pareto optimal front of a multi-objective optimization problem is bounded by two special vectors:

Definition 8. The *ideal objective vector* $\vec{z}^* \in \mathbb{R}^m$ minimizes all objective functions. Each i^{th}-component is defined as $z_i^* = \min_{\vec{x}} f_i(\vec{x})$.

Definition 9. The *nadir objective vector* $\vec{z}^{nad} \in \mathbb{R}^m$ is constructed using the worst values of $\mathcal{PF}^*$. Each i^{th}-component is defined as $z_i^{nad} = \max_{\vec{x} \in \mathcal{P}^*} f_i(\vec{x})$.

A utility function $u : \mathbb{R}^m \rightarrow \mathbb{R}$ is a model of the decision maker's preference that maps each objective vector into a scalar value. Given an approximation set A, and a set of utility functions U, the unary $R2$ indicator [1] is defined as:

$$R2(A, U) = \frac{1}{|U|} \sum_{u \in U} u^*(A), \tag{3}$$

where $u^*(A) = \min_{\vec{a} \in A}\{u(\vec{a})\}$ is the best utility value obtained in the set A.[1]

[1] For simplicity, we have applied the dual property: $\min \vec{z} = -\max(-\vec{z})$, and we also assume that the utility functions are non-negative.

3. PREVIOUS RELATED WORK

In this section, we review some recently created MOEAs that were designed for many-objective optimization problems.

Δ_p-DDE [16] uses the Δ_p indicator [18] as the selection mechanism of differential evolution. Since this semi-metric requires knowledge about the Pareto optimal front, Δ_p-DDE employs an echelon form of the nondominated individuals. The complexity of this MOEA is dominated by the building of the reference set, which is exponential with respect to the number of objectives and the quality of the outcome sets is influenced by a resolution parameter.

The Nondominated Sorting Genetic Algorithm III (NSGA-III) [4] is similar to the original NSGA-II algorithm [5], which ranks individuals using a nondomination criterion. But, the maintenance of diversity is aided by a set of weight vectors. The main advantage of this MOEA is the uniformity of its distributions. However, its implementation is not trivial, being sensitive to the construction of a hyper-plane.

The $R2$ Indicator Based Evolutionary Algorithm ($R2$-IBEA) [15] is an extension of IBEA [23], which eliminates Pareto dominance and performs a selection guided by an exponential amplification of a binary version of the $R2$ indicator. The set of normalized weight vectors is generated using a hypervolume-based approach. The reference point is dynamically updated according to the extent of current solutions in objective function space.

The $R2$ Multi-Objective Genetic Algorithm ($R2$-MOGA) and $R2$ Multi-Objective Differential Evolution ($R2$-MODE) [7] incorporate the $R2$ indicator to a modified version of the nondominated sorting method of NSGA-II [5], in order to separate individuals into layers. This approach is coupled to two different search algorithms, resulting in two MOEAs. The set of weight vectors is generated at each generation using a random approach. A reference point is updated per generation using the utopian point.[2]

The Many-Objective Metaheuristic Based on the $R2$ Indicator (MOMBI) [9] is another example of a MOEA based on the $R2$ indicator, which is easy to implement and does not require Pareto dominance (in the following sections we will discuss more about it).

These three $R2$ approaches adopt the weighted Tchebycheff functions as their utility functions.

4. OUR PROPOSED APPROACH

In this section we describe, in order of importance, the main limitations of the original MOMBI. First, we analyze the scalability of some utility functions. Then, we provide a mechanism to update the reference points during normalization. Finally, we introduce MOMBI-II, including a brief analysis of its computational complexity.

4.1 Scalability of Utility Functions

There are several possible utility functions, the majority of which have been borrowed from mathematical programming. The most popular is the Weighted Tchebycheff (WT) metric [12], which is of the form:

$$u_{tch}(\vec{v} : \vec{r}, \vec{w}) = \max_{i \in \{1, \cdots, m\}} \{w_i |v_i - r_i|\}, \tag{4}$$

where $r \in \mathbb{R}^m$ is a reference point, $\vec{w}$ is a weight vector such that $w_i \geq 0$ for all $i \in \{1, \cdots, m\}$ and $\sum_{i=1}^m w_i = 1$. The

[2] $z_i^{**} = z_i^* - \epsilon_i$ for all $i \in \{1, \cdots, m\}$, where $\epsilon_i > 0$ is a relatively small but computational significant scalar.

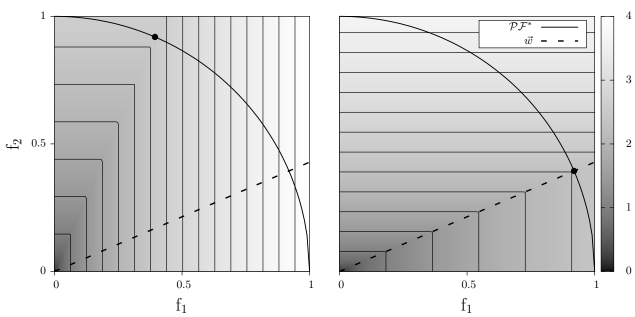

Figure 1: Contour lines of the WT (left) and ASF (right) metrics for the weight vector $\vec{w} = (0.7, 0.3)$.

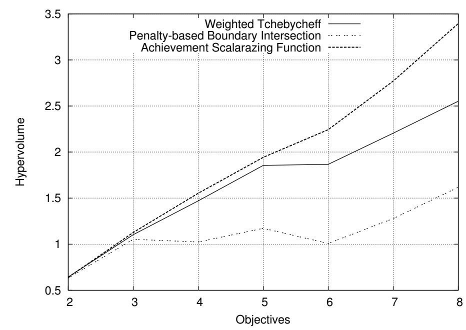

Figure 2: Average hypervolume for MOEA/D using three different scalarizing functions from 2 to 8 objectives on the DTLZ3 test problem.

minimization of this metric produces weakly Pareto optimal solutions. The geometric interpretation of equation (4) is presented in Figure 1. It is interesting to observe that the contour lines are incongruent with the weight vector.

Another important metric is the Penalty-based Boundary Intersection (PBI) [21]:

$$u_{pbi}(\vec{v} : \vec{r}, \vec{w}) = d_1 + \theta d_2, \tag{5}$$

$$d_1 = \frac{\left|\left|(\vec{v} - \vec{r})^T \vec{w}\right|\right|}{||\vec{w}||}, \tag{6}$$

$$d_2 = \left|\left|\vec{v} - \left(\vec{r} + d_1 \frac{\vec{w}}{||\vec{w}||}\right)\right|\right|, \tag{7}$$

where $\theta > 0$ is a penalty parameter. The minimization of the PBI metric produces Pareto optimal solutions much better uniformly distributed than those obtained by equation (4).

A metric that has been scarcely studied in MOEAs is the achievement scalarizing function (ASF) [20], defined as[3]:

$$u_{asf}(\vec{v} : \vec{r}, \vec{w}) = \max_{i \in \{1, \cdots, m\}} \left\{ \frac{|v_i - r_i|}{w_i} \right\}, \tag{8}$$

Its minimization produces weakly Pareto optimal solutions. Note that the weight vector agrees with the contour lines of Figure 1.

When these metrics are incorporated into a MOEA, a set of weight vectors is required in order to exploit its population-based nature. Usually, the Simplex-Lattice Design (SLD) method [17] is the most preferred. In this case, the distribution of the vectors is equally spaced over a simplex, producing more points at the boundary than at the center. This could explain the distinctive distribution generated by MOEAs based on the WT metric, such as MOEA/D [21] or MOMBI, where solutions are biased from the center of the Pareto front to a few boundary points, since this metric optimizes the opposite weight vectors.

In the following, we examine the scalability of these three metrics applied to the well-known MOEA/D on the DTLZ3 test problem [6]. We performed 30 independent runs of each instance from 2 to 8 objectives, adopting the same parameter settings used in [9]. To assess performance, we employed the hypervolume indicator.

In Figure 2, we present our experimental results. Here, for two and three objectives all the metrics gave very similar results. However, as the number of objectives increases, the performance of PBI and the WT metrics degrades. In contrast, the ASF metric obtained the maximum hypervolume for many-objective problems.

It is important to mention that the only difference between the WT and ASF utility functions is the weight parameter. Thus, the former can derive congruent results, if we change the vectors in such a way that each component is the reciprocal of those generated by the SLD method[4]. However, we decided to incorporate the ASF metric into MOMBI-II and left for future research the design of other weight vectors.

4.2 Update of Reference Points

In the original version of MOMBI each objective function was normalized using the following expression:

$$f_i'(\vec{x}) = \frac{f_i(\vec{x}) - z_i^{min}}{z_i^{max} - z_i^{min}} \quad \forall i \in \{1, \cdots, m\}, \tag{9}$$

where $\vec{z^{min}}$ and $\vec{z^{max}}$ are the ideal and nadir objective vectors, respectively, taken from the current population.

Some problems with this approach were that the true ideal vector of the feasible region was never retained, and in consequence, boundary solutions of the Pareto optimal front were missing. In multi-frontal problems, the nadir objective vector was always an outlier, causing a poor distribution in the approximation sets. In high dimensionality, it could happen that the ideal and the nadir vectors were close, flattening some objectives and thus causing MOMBI to lose diversity. Also, the constant update of these vectors made the population unstable in the final generations.

In order to overcome these drawbacks, we update the reference points in a smarter way, taking into account statistical information of previous generations. The idea is to monitor the nadir point of the parents population at each generation, determining how close are the individuals from the true Pareto front. A high variance means that solutions are far away, which strongly biases the location of the point. A low variance means that solutions are close, thus, small movements must be done. This mechanism can serve as a cut-off of objective space, since outliers are removed.

In Algorithm 1, we present the pseudo-code of this technique. Here, we use a data structure called *record* that stores

[3]We consider the absolute value from its original definition.

[4]The restriction of normalization is dropped without affecting weak Pareto dominance [12]. In order to avoid indeterminate form, a computational big scalar can be taken instead.

Algorithm 1 Update Reference Points

Require: $\vec{z}^{min}$, $\vec{z}^{max}$, population P, num. objectives m
Ensure: $\vec{z}^{min}$ and $\vec{z}^{max}$

1: Update vectors $\vec{z}^*$ and $\vec{z}^{nad}$
2: $z_i^{min} \leftarrow \min\{z_i^{min}, z_i^*\} \quad \forall i \in \{1, \cdots, m\}$
3: Store $\vec{z}^{nad}$ in record
4: Obtain vector of variances $\vec{v} \in \mathbb{R}^m$ for $\vec{z}^{nad}$ from *record*
5: **if** $\max_{j \in \{1,\cdots,m\}} v_j > \alpha$ **then**
6: $\quad z_i^{max} \leftarrow \max_{j \in \{1,\cdots,m\}} z_j^{nad} \quad \forall i \in \{1, \cdots, m\}$
7: **else**
8: $\quad$ **for all** $i \in \{1, \cdots, m\}$ **do**
9: $\quad\quad$ **if** $|z_i^{max} - z_i^{min}| < \epsilon$ **then**
10: $\quad\quad\quad z_i^{max} \leftarrow \max_{j \in \{1,\cdots,m\}} z_j^{max}$
11: $\quad\quad\quad$ Mark z_i^{max}
12: $\quad\quad$ **else if** $z_i^{nad} > z_i^{max}$ **then**
13: $\quad\quad\quad z_i^{max} \leftarrow 2z_i^{nad} - z_i^{max}$
14: $\quad\quad\quad$ Mark z_i^{max}
15: $\quad\quad$ **else if** $v_i = 0$ **and** z_i^{max} has not been marked recently **then**
16: $\quad\quad\quad$ Obtain the maximum value a for z_i^{nad} from record
17: $\quad\quad\quad z_i^{max} \leftarrow (z_i^{max} + a)/2$
18: $\quad\quad\quad$ Mark z_i^{max}
19: **return** $\vec{z}^{min}$ and $\vec{z}^{max}$

the nadir vector of a few generations. We also need the parameter α, which is the threshold of variances for the vector $\vec{z}^{nad}$, and the parameter ϵ, which is a tolerance threshold. In line 1, the ideal and nadir objective vectors are updated using Definitions 8 and 9. In line 2, if $\vec{z}^{min}$ is improved, then it is updated. In lines 3 and 4, the nadir vector is kept in *record* and its vector of variances, for each objective, is estimated. If the maximum variance of all objectives is high, then $\vec{z}^{max}$ is updated using the maximum objective of the nadir vector (lines 5 and 6). Otherwise, from lines 8 to 18, each component of $\vec{z}^{max}$ is examined. If the absolute difference between its i^{th} value and z_i^{min} is less than ϵ, then it is marked and updated, using the maximum objective of $\vec{z}^{max}$. Otherwise, if the i^{th} value of the nadir vector is greater than the component, then it is expanded and marked. Otherwise, if the variance is zero and z_i^{max} has not been marked yet, then it is marked and averaged between its previous value and the maximum value stored for it in *record*. In all cases, the mark lasts the same number of generations that *record* is kept. The total complexity of this algorithm can be done in $\mathcal{O}(|P|m)$, where $|P|$ denotes the population size.

4.3 MOMBI-II

In this subsection, we introduce MOMBI-II (Many-Objective Metaheuristic Based on the $R2$ Indicator II), which is inspired on a Genetic Algorithm.

The $R2$ ranking procedure (see Algorithm 2) remains almost with no changes. Solutions that optimize the set of weight vectors are chosen and placed on top such that they get the first rank (the best). Such points will then be removed and a second rank will be identified in the same manner. The process will continue until all the solutions had been ranked. When two individuals contribute with the same utility value, then we choose as tiebreaker the one with the lower Euclidean distance. This eliminates weak-Pareto solutions. The ASF metric is adopted as our utility function. The notation assumes that each individual p contains the vector of objective functions $p.\vec{f}$, the hierarchy of the individual $p.rank$ and $p.\mu$, the current utility value for a weight vector $\vec{w}$. The complexity of the algorithm is $\mathcal{O}(|W||P|(\log|P| + m))$.

Algorithm 2 $R2$ Ranking Algorithm

Require: Population P, set of weight vectors W
Ensure: Ranking of the population

$p.rank \leftarrow p.\mu \leftarrow \infty \quad \forall p \in P$
for all $\vec{w} \in W$ **do**
$\quad$ **for all** $p \in P$ **do**
$\quad\quad p.\mu \leftarrow u_{asf}(p.\vec{f} : \vec{0}, \vec{w})$
$\quad$ Sort P w.r.t. the fields μ and L_2 in increasing order
$\quad rank \leftarrow 1$
$\quad$ **for all** $p \in P$ **do**
$\quad\quad p.rank \leftarrow \min\{p.rank, rank\}$
$\quad\quad rank \leftarrow rank + 1$

Algorithm 3 Main Loop of MOMBI-II

Require: MOP, stopping criterion, set of weight vectors W
Ensure: Pareto set approximation

1: Initialize population P_i, $i \leftarrow 1$
2: Evaluate population P_i
3: Calculate the L_2-norm of objectives for P_i
4: Set $\vec{z}^{min} \leftarrow \vec{z}^*$ and $\vec{z}^{max} \leftarrow \vec{z}^{nad}$
5: **while** termination condition is not fulfilled **do**
6: $\quad$ Perform parent selection
7: $\quad$ Generate offspring P_i' using variation operators
8: $\quad$ Evaluate population P_i'
9: $\quad$ Calculate the L_2-norm of objectives for P_i'
10: $\quad$ Normalize objective functions for $P_i \bigcup P_i'$
11: $\quad$ Execute $R2$ ranking algorithm $(P_i \bigcup P_i', W)$
12: $\quad$ Reduce population $P_{i+1} \leftarrow \{P_i \bigcup P_i'\}$
13: $\quad$ Update reference points $(\vec{z}^{min}, \vec{z}^{max}, P_{i+1}, m)$
14: $\quad i \leftarrow$i + 1
15: **return** P_i

The main loop of MOMBI-II is presented in Algorithm 3. The normalization step in line 10 is performed using equation (9). The criteria for population reduction takes into account the rank of each individual and the Euclidean distance. Next, we determine the complexity of this approach. Parent selection is performed in $\mathcal{O}(|P|)$, as well as the offspring generation. The evaluation of the population, the calculation of the norm, the update of reference points and the normalization is done in $\mathcal{O}(|P|m)$ each. As seen before, the ranking procedure takes $\mathcal{O}(|W||P|(\log|P|+m))$ and the reduction can be performed in $\mathcal{O}(|P|\log|P|)$. Therefore, the overall complexity of MOMBI-II at each generation is $\mathcal{O}(|W||P|(\log|P| + m))$ and the storage is $\mathcal{O}(|P|m)$.

5. EXPERIMENTAL RESULTS

In this section, we first describe the parameters used to compare the performance on 3, 5 and 10 objectives of Δ_p-DDE, NSGA-III, $R2$-MOGA, $R2$-IBEA[5], MOMBI-II and MOMBI. With the aim of evaluating the impact of only one improvement in the proposed MOEA, we considered another version, named MOMBI-ASF, which incorporates the ASF metric and the normalization is done as in the original algo-

[5]Ten objectives were excluded from the study, since this MOEA is scalable up to 5 objectives [15].

Table 1: **Parameters adopted in the study.**

m	Weights		$\|P\|$	Δ_p-DDE	WFG	
	H	H_{ref}		r	n	k
3	12	66	92	10	24	4
5	5	30	126	4	47	8
10	3	10	220	3	105	18

rithm. Then, we establish the adopted performance metric in the experiments, the test problems and finalize with the discussion of the results.

5.1 Parameters Settings

The parameters were identical for all the algorithms (see Table 1). The set of weight vectors was generated using the SLD method. Its cardinality is given by the combinatorial number:

$$|W| = \binom{H+m-1}{m-1}, \tag{10}$$

where m represents the number of objectives and H is a proportionality parameter. The population size $|P|$ was the same as the number of weight vectors, except for 3 objectives (in this case, it was increased by 1 to fulfill the requirement of adopting even numbers in the binary tournament selection). For Δ_p-DDE, the p-norm was set to 1, the parameters F and C were both established to 0.5 and the resolution of the reference set r varied according to the objectives as shown in Table 1. In the genetic algorithms, the variation operators were simulated binary crossover (SBX) and polynomial-based mutation [3]. As suggested in [4], the crossover rate and its distribution index were set to 1.0 and 30, respectively. The mutation rate was set to $1/n$ (here, n represents the number of variables) and its distributed index was set to 20. All these MOEAs were implemented using real-numbers encoding and their source code was provided by their authors, except for NSGA-III.

For MOMBI-II, the parameters ϵ and α were set to $1e-3$ and 0.5, respectively. These values were determined from an analysis of several samplings. We observed that they are independent of the dimensionality and the problem to be solved. As an example, we show in Figure 3(a) the median values of the Δ_2 indicator (see next subsection for its definition) for 30 independent runs of DTLZ1 with 5 objectives. Similarly, the *record* size (used for tracing the variance of the nadir vector) was set to 5 generations. It can be appreciated from Figure 3(b), that low values of the indicator are achievable using this value and it is computationally less expensive than if we perform the normalization at every generation.

5.2 Performance Measure

We selected the Δ_p [18] as a performance assessment measure. This indicator simultaneously evaluate proximity to the Pareto optimal front and spread of solutions along it. Given an approximation set A and a discretized Pareto optimal front $\mathcal{PF}$ of a MOP, the Δ_p indicator is defined as:

$$\Delta_p(A, \mathcal{PF}) = \max\left(\left(\frac{\sum_{i=1}^{|A|} d_i^p}{|A|}\right)^{\frac{1}{p}}, \left(\frac{\sum_{i=1}^{|\mathcal{PF}|} e_i^p}{|\mathcal{PF}|}\right)^{\frac{1}{p}}\right), \tag{11}$$

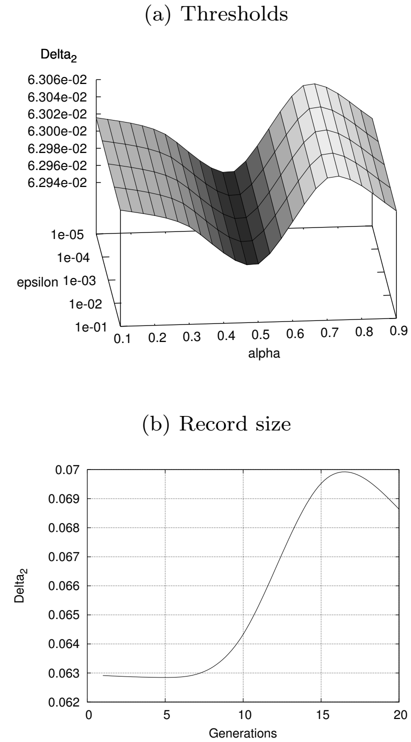

Figure 3: Parameter sampling in MOMBI-II for the DTLZ1 with 5 objectives.

where d_i is the Euclidean distance from $a_i \in A$ to its nearest member in $\mathcal{PF}$ and e_i is the Euclidean distance from $pf_i \in \mathcal{PF}$ to its nearest member in A. Small values of Δ_p are preferred.

Since this indicator requires that $\mathcal{PF}$ is known, we sampled it for each type of geometry in the adopted test problems by generating a set of weight vectors using the SLD method and then finding the intersection point of $\mathcal{PF}$ with the lines defined between the origin and the weight vectors. The number of points in $\mathcal{PF}$ is defined by equation (10), using the parameter H_{ref} of Table 1.

5.3 Test Problems

For comparison purposes, we select four normalized problems of the Deb-Thiele-Laumanns-Zitzler (DTLZ) set [6]. These are DTLZ1, with a linear and multi-frontal Pareto front and the others with a concave geometry: DTLZ2, DTLZ3 (multi-frontal) and DTLZ4 (biased). We also include in the study two problems of the Walking-Fish-Group (WFG) set [10], which are concave with different scale in each objective. Such instances are WFG6 (non-separable) and WFG7 (biased). Their properties make these two problems much harder to solve for MOEAs. The number of decision variables and position-related (k) parameters are specified in Table 1.

5.4 Discussion of Results

We performed 30 independent runs of each of the 7 MOEAs compared on all the test instances adopted. With the aim of comparing the performance of all algorithms among themselves in a pairwise fashion, the Wilcoxon rank sum test (one-tailed) with the Bonferroni correction [14] was applied

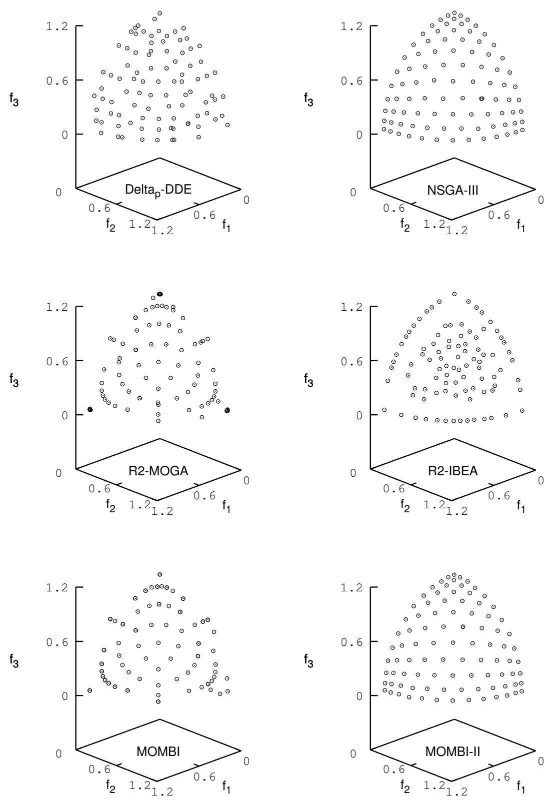

Figure 4: Pareto fronts produced by MOEAs on DTLZ3 for 3 objectives.

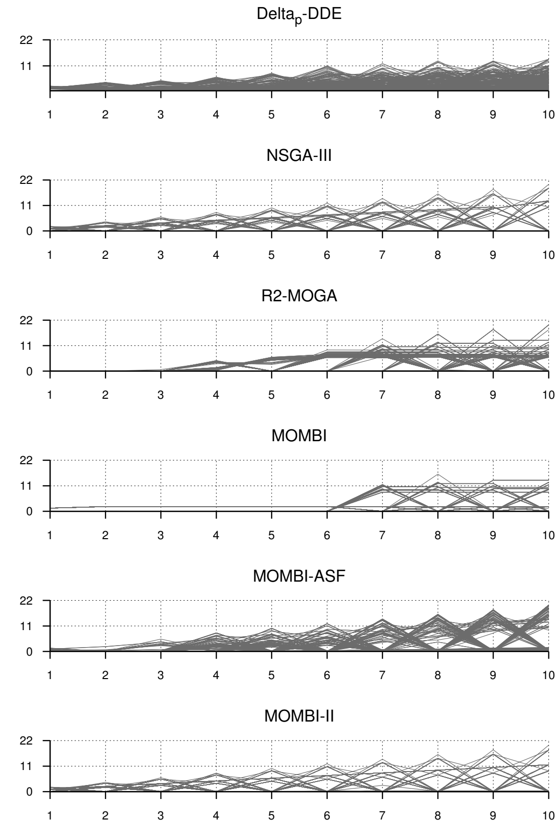

Figure 5: Parallel coordinates of the approximations obtained by MOEAs on WFG7 for 10 objectives.

to the Δ_2 indicator values. Experimental results appear in Tables 2 and 3. Examples of the Pareto fronts, corresponding to the median values, are depicted in Figures 4 and 5.

The best algorithm was MOMBI-II, outperforming the other MOEAs in 96.1% of the tests. It was followed by NSGA-III with 65.7%, both Δ_p-DDE and MOMBI-ASF with 52%, $R2$-IBEA with 26.4%, $R2$-MOGA with 18.6% and in last place is MOMBI with 5.9%.

MOMBI-II tied on two instances for 3 objectives with MOMBI-ASF and NSGA-III, producing almost identical results. On WFG6 and WFG7 for 10 objectives, MOMBI-II lost against Δ_p-DDE. As can be seen in Figure 4, the distributions of the Pareto fronts produced by MOMBI-II are more uniform than those generated by Δ_p-DDE and the other MOEAs based on the $R2$ indicator.

NSGA-III faced difficulties on multi-frontal problems for 10 objectives, being unable to reach the true Pareto front in some cases. In general, this algorithm produced similar results to those generated by MOMBI-II.

Δ_p-DDE experimented the same difficulties as NSGA-III on multi-frontal problems. Surprisingly, it won on WFG6 and WFG7 for ten objectives. Here, we suspect that differential evolution is able to produce better solutions than SBX and polynomial-based mutation. An advantage of Δ_p-DDE is that it explores more regions for many-objective problems than those algorithms based on weight vectors.

The performance of MOMBI-ASF was poor, being unable to beat NSGA-III. This indicates that the mechanism for updating the reference points plays an important role in MOMBI-II, specially for many-objective problems.

The solutions generated by $R2$-IBEA have a strong bias to the center of the Pareto front than to its edges, being similar to those generated by the hypervolume-based algorithms.

In $R2$-MOGA and MOMBI the WT metric presented difficulties in convergence in high dimensionality problems.

6. CONCLUSIONS AND FUTURE WORK

In this work, we have introduced an improved version of a MOEA based on the $R2$ indicator (called MOMBI-II), which substitutes the weighted Tchebycheff metric by the achievement scalarizing function, since the former presents difficulties for many-objective problems. Moreover, we developed a mechanism to update the reference points, required during the normalization process of the objective functions, using statistical information of the population proximity to the true Pareto optimal front. This is an important aspect since the diversity is sensitive to the choice of such points.

Experimental results indicate that the new version outperformed MOMBI, Δ_p-DDE, NSGA-III, $R2$-MOGA and $R2$-IBEA in more than the 96% of the test instances. Therefore, we believe that MOMBI-II is a suitable alternative for solving many-objective optimization problems.

It is worth noticing that the solutions produced by MOMBI-II are uniformly distributed in objective space, being similar to those generated by NSGA-III. However, MOMBI-II requires much less computational effort and the source code is in the public domain.

Much more work is still required. We are interested, for example, in studying the scalability of some other utility functions available in the literature and in incorporating a mechanism to handle constraints, in order to solve real-world problems.

Acknowledgments

The first author acknowledges support from CONACyT and CINVESTAV-IPN to pursue graduate studies in Computer Science. The second author gratefully acknowledges support from CONACyT project no. 221551.

7. REFERENCES

[1] D. Brockhoff, T. Wagner, and H. Trautmann. On the Properties of the $R2$ Indicator. In *2012 Genetic and Evolutionary Computation Conference (GECCO'2012)*, pages 465–472, Philadelphia, USA, July 2012. ACM Press. ISBN: 978-1-4503-1177-9.

[2] C. A. Coello Coello, G. B. Lamont, and D. A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York, second edition, September 2007. ISBN 978-0-387-33254-3.

[3] K. Deb and R. B. Agrawal. Simulated Binary Crossover for Continuous Search Space. *Complex Systems*, 9:115–148, 1995.

[4] K. Deb and H. Jain. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems with Box Constraints. *Evolutionary Computation, IEEE Transactions on*, 18(4):577–601, Aug 2014.

[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA–II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, April 2002.

[6] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable Test Problems for Evolutionary Multiobjective Optimization. In A. Abraham, L. Jain, and R. Goldberg, editors, *Evolutionary Multiobjective Optimization. Theoretical Advances and Applications*, pages 105–145. Springer, USA, 2005.

[7] A. Diaz-Manriquez, G. Toscano-Pulido, C. Coello, and R. Landa-Becerra. A Ranking Method Based on the $R2$ Indicator for Many-Objective Optimization. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 1523–1530, 2013.

[8] M. P. Hansen and A. Jaszkiewicz. Evaluating the Quality of Approximations to the Non-dominated Set. Technical Report IMM-REP-1998-7, Technical University of Denmark, March 1998.

[9] R. Hernandez Gomez and C. Coello Coello. MOMBI: A New Metaheuristic for Many-Objective Optimization Based on the $R2$ Indicator. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 2488–2495, 2013.

[10] S. Huband, L. Barone, L. While, and P. Hingston. A Scalable Multi-objective Test Problem Toolkit. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005*, pages 280–295, Guanajuato, México, March 2005. Springer. Lecture Notes in Computer Science Vol. 3410.

[11] H. Ishibuchi, N. Tsukamoto, and Y. Nojima. Evolutionary Many-Objective Optimization: A Short Review. In *2008 Congress on Evolutionary Computation (CEC'2008)*, pages 2424–2431, Hong Kong, June 2008. IEEE Service Center.

[12] M. Kaisa. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston, 1st edition, 1999.

[13] J. Knowles and D. Corne. On Metrics for Comparing Nondominated Sets. In *Congress on Evolutionary Computation (CEC'2002)*, volume 1, pages 711–716, Piscataway, New Jersey, May 2002. IEEE Service Center.

[14] R. G. Miller. *Simultaneous Statistical Inference*. Springer-Verlag, New York, 1981.

[15] D. H. Phan and J. Suzuki. R2-IBEA: R2 Indicator Based Evolutionary Algorithm for Multiobjective Optimization. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 1836–1845, 2013.

[16] C. A. Rodríguez Villalobos and C. A. Coello Coello. A New Multi-Objective Evolutionary Algorithm Based on a Performance Assessment Indicator. In *2012 Genetic and Evolutionary Computation Conference (GECCO'2012)*, pages 505–512, Philadelphia, USA, July 2012. ACM Press. ISBN: 978-1-4503-1177-9.

[17] H. Scheffé. Experiments with Mixtures. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 20:344–360, 1958.

[18] O. Schütze, X. Esquivel, A. Lara, and C. A. Coello Coello. Using the Averaged Hausdorff Distance as a Performance Measure in Evolutionary Multiobjective Optimization. *IEEE Transactions on Evolutionary Computation*, 16(4):504–522, August 2012.

[19] S. Wessing and B. Naujoks. Sequential Parameter Optimization for Multi-Objective Problems. In *Evolutionary Computation (CEC), 2010 IEEE Congress on*, pages 1–8, July 2010.

[20] A. Wierzbicki. The Use of Reference Objectives in Multiobjective Optimization. In G. Fandel and T. Gal, editors, *Multiple Criteria Decision Making Theory and Application*, volume 177 of *Lecture Notes in Economics and Mathematical Systems*, pages 468–486. Springer Berlin Heidelberg, 1980.

[21] Q. Zhang and H. Li. MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731, December 2007.

[22] E. Zitzler. *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*. PhD thesis, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, November 1999.

[23] E. Zitzler and S. Künzli. Indicator-based Selection in Multiobjective Search. In X. Y. et al., editor, *Parallel Problem Solving from Nature - PPSN VIII*, pages 832–842, Birmingham, UK, September 2004. Springer-Verlag. Lecture Notes in Computer Science Vol. 3242.

Table 2: Median and standard deviation of the Δ_2 indicator for 3 and 5 dimensions. In each case, the outperformance relation among algorithms is shown, using a significance level of $\alpha = 0.5$ (for example, Δ_p-DDE performs significantly better than $R2$-MOGA and MOMBI on DTLZ1 for 3 objectives). The two best values are shown in gray scale, where the darker tone corresponds to the best value. The number of generations appears within parentheses in the first column.

Problem	Δ_p-DDE 1	NSGA-III 2	$R2$-MOGA 3	$R2$-IBEA 4	MOMBI 5	MOMBI-ASF 6	MOMBI-II 7
3 Objectives							
DTLZ1 (400)	0.0252±8.60e-4 3,5	0.0220±2.25e-1 1,3,4,5	0.0352±7.00e-4 5	0.0239±2.01e-4 1,3,5	0.0361±1.32e-4 –	0.0225±6.10e-4 1,3,4,5	0.0212±9.48e-5 1,2,3,4,5,6
DTLZ2 (250)	0.0680±2.31e-3 3,4,5	0.0585±1.67e-4 1,3,4,5,6	0.0804±9.67e-4 4,5	0.0821±2.42e-3 –	0.0824±1.83e-2 –	0.0589±6.00e-4 1,3,4,5	0.0583±5.37e-4 1,3,4,5,6
DTLZ3 (1000)	0.0672±2.86e-1 3,4,5	0.0588±6.27e-4 1,3,4,5,6	0.0800±7.72e-1 4	0.0853±3.36e-3 –	0.0806±1.29e-2 4	0.0594±9.25e-3 1,3,4,5	0.0579±1.86e-4 1,2,3,4,5,6
DTLZ4 (600)	0.0685±1.87e-3 3,4,5	0.0584±3.42e-1 1,3,4,5,6	0.0829±1.50e-1 –	0.0833±2.40e-1 –	0.0827±1.50e-1 –	0.0947±1.58e-1 –	0.0578±1.12e-1 1,2,3,4,5,6
WFG6 (400)	0.3474±2.35e-2 3,4,5	0.2545±3.23e-3 1,3,4,5	0.3836±9.37e-3 5	0.3900±2.69e-2 5	0.4117±6.37e-3 –	0.2497±3.41e-3 1,2,3,4,5	0.2487±2.85e-2 1,2,3,4,5
WFG7 (400)	0.2799±1.15e-2 3,4,5	0.2420±9.81e-4 1,3,4,5	0.3878±1.11e-2 5	0.3831±2.14e-2 5	0.4073±6.03e-3 –	0.2418±2.21e-3 1,3,4,5	0.2402±1.18e-3 1,2,3,4,5,6
5 Objectives							
DTLZ1 (600)	0.0677±2.86e-3 3,5,6	0.0668±1.34e-3 1,3,4,5,6	0.1190±6.22e-4 5	0.0686±4.98e-4 3,5,6	0.1199±2.12e-2 –	0.0730±5.03e-3 3,5	0.0629±2.62e-4 1,2,3,4,5,6
DTLZ2 (350)	0.2113±2.69e-3 3,4,5	0.2092±1.32e-4 1,3,4,5,6	0.3376±5.32e-4 5	0.2271±2.16e-3 3,5	0.3397±2.51e-4 –	0.2107±3.69e-3 3,4,5	0.2051±7.87e-4 1,2,3,4,5,6
DTLZ3 (1000)	0.2117±7.44e+1 3,4,5,6	0.2098±6.89e-3 1,3,4,5,6	0.3379±1.02e-3 5	0.2252±3.65e-3 3,5	0.3402±8.10e-2 –	0.2136±4.30e-3 3,4,5	0.2049±9.29e-4 1,2,3,4,5,6
DTLZ4 (1000)	0.2146±2.97e-3 3,4,5,6	0.2091±1.91e-1 5	0.3368±7.38e-2 5	0.2281±7.33e-2 3,5,6	0.3396±8.80e-2 –	0.2644±1.00e-1 3,5	0.2038±1.18e-1 1,2,3,4,5,6
WFG6 (750)	1.3321±3.05e-2 3,4,5	1.2564±2.46e-3 1,3,4,5,6	3.3935±1.19e-1 –	1.4415±3.73e-2 3,5	2.3444±1.90e-2 3	1.2619±1.42e-2 1,3,4,5	1.2416±3.18e-3 1,2,3,4,5,6
WFG7 (750)	1.3953±1.92e-2 3,4,5	1.2657±1.66e-2 1,3,4,5,6	2.3404±1.60e-2 5	1.4255±2.67e-2 3,5	2.4201±7.19e-2 –	1.3384±2.46e-2 1,3,4,5	1.2477±4.76e-3 1,2,3,4,5,6

Table 3: Median and standard deviation of the Δ_2 indicator for 10 objectives (continuation).

Problem	Δ_p-DDE 1	NSGA-III 2	$R2$-MOGA 3	MOMBI 5	MOMBI-ASF 6	MOMBI-II 7
DTLZ1 (1000)	9.9311±7.45e+0 –	0.2113±2.78e+0 1	0.1569±1.32e-1 1,5,6	0.1647±9.54e-3 1,6	0.2087±2.74e-2 1	0.1235±1.08e-2 1,2,3,5,6
DTLZ2 (750)	0.4208±3.92e-3 2,3,5,6	0.4290±2.59e-1 3,5,6	0.6355±1.84e-2 5	0.6633±3.37e-2 –	0.4703±1.70e-2 3,5	0.4156±3.25e-4 1,2,3,5,6
DTLZ3 (1500)	512.7180±1.02e+2 –	0.9053±2.20e+0 1	0.6714±1.97e+0 1	0.6885±3.66e-2 1,2	0.4777±6.93e-2 1,2,3,5	0.4151±4.65e-2 1,2,3,5,6
DTLZ4 (2000)	0.4417±3.76e-3 3,5	0.4280±1.29e-1 1,3,5	0.6419±2.47e-2 5	0.6804±1.53e-2 –	0.4329±2.26e-2 1,3,5	0.4148±4.98e-2 1,2,3,5,6
WFG6 (2000)	4.6238±8.11e-2 2,3,5,6,7	4.8552±5.55e-3 3,5,6	7.3007±4.91e-1 5	7.6871±3.62e-1 –	5.2036±4.38e-2 3,5	4.7527±6.55e-3 2,3,5,6
WFG7 (2000)	4.5544±8.74e-2 2,3,5,6,7	4.9397±9.26e-1 3,5,6	7.7848±3.87e-1 5	7.8535±3.91e-1 –	5.8264±4.89e-1 3,5	4.7688±3.13e-2 2,3,5,6