

# Bio-Inspired Algorithms and Preferences for Multi-objective Problems

Daniel Cinalli<sup>1</sup>(✉), Luis Martí<sup>1</sup>, Nayat Sanchez-Pi<sup>2</sup>,  
and Ana Cristina Bicharra Garcia<sup>1</sup>

<sup>1</sup> Universidade Federal Fluminense, Rio de Janeiro, Brazil  
{`dcinalli,lmarti,bicharra`}@ic.uff.br

<sup>2</sup> Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil  
nayat@ime.uerj.br

**Abstract.** Multi-objective optimization evolutionary algorithms have been applied to solve many real-life decision problems. Most of them require the management of trade-offs between multiple objectives. Reference point approaches highlight a preferred set of solutions in relevant areas of Pareto frontier and support the decision makers to take more confidence evaluation. This paper extends some well-known algorithms to work with collective preferences and interactive techniques. In order to analyse the results driven by the online reference points, two new performance indicators are introduced and tested against some synthetic problem.

**Keywords:** Collective intelligence · Preferences · Reference points · Evolutionary multi-objective optimization algorithms

## 1 Introduction

Multi-objective optimization problems (MOPs) simultaneously optimize a set of objective functions. Formally posed, a MOP can be defined as  $\min \mathbf{F}(\mathbf{x}) = \{f_1(\mathbf{x}), \dots, f_k(\mathbf{x})\}$ , where  $\mathbf{x} = \langle x_1, \dots, x_n \rangle \in \Omega$  is an  $n$ -dimensional vector of decision variables. The solution of a MOP is a (possibly infinite) Pareto-optimal set  $P_S = \{\mathbf{x} \in \Omega \mid \nexists \mathbf{y} \in \Omega : \mathbf{y} \prec \mathbf{x}\}$  that contains all the elements of  $\Omega$  that not Pareto-dominated ( $\prec$ ) by any other element [2]. Elements of  $P_S$  represent different trade-offs between the objective functions values. The projection of  $P_S$  through  $F()$  is known as the Pareto-optimal front,  $P_F$ .

The application of evolutionary algorithms to MOPs has prompted the creation of multi-objective optimization evolutionary algorithms (MOEAs) [2]. They approximate the  $P_S$  as a discrete set of points.

Having a solution to a MOP, a decision maker (DM) must identify which of those solutions are the ones that satisfy her/his preferences and would be realized in practice. This task can be rather complex and requires in-depth knowledge of the problem being solved, something that is impossible in many practical situations.

Reference points and interactive techniques can be used to mitigate these inconveniences and support the DM in reaching a proper specification. These approaches also allow the optimization algorithm to focus on areas of interest and thus reaching satisfactory solutions at a lower computational cost.

In practice, however, optimization problems pose difficulties in defining *a priori* reference points or preferences. In this regard, collective intelligence reference points obtained by the interaction and aggregation of multiple opinions can be used to produce an accurate representation of preferences and, hence, reference points. This approach can also eliminate the unilateral choice bias that can be introduced by a single DM. Collective environments can improve MOEAs with cognitive and subjective evaluation to find better solutions in more relevant regions of  $P_F$ .

This work incorporates online interactive collective preferences and interactive behaviour into three existing MOEAs: NSGA-II [6], SPEA2 [12] and SMS-EMOA [1] and, therefore, introduces a collective intelligence version of them. Similarly, while working in this task it became evident the lack of adequate performance indicators that take into account preferences. Therefore, as part of the paper are also introduced two new performance indicators that are used to evaluate the quality of the  $P_F$  approximation driven by the online collective preferences.

The paper proceeds with Sect. 2 that covers some required formal foundations regarding reference points and collective intelligence. Section 3 introduces the novel interactive MOEAs. Subsequently, Sect. 4 presents the new performance indicators that are to be used later on. After that, Sect. 5 analyzes the performance of the algorithms when faced with benchmark problem. Finally, conclusive remarks and future work directions are put forward.

## 2 Foundations

Preferences are user-defined parameters and denote values or subjective impressions regarding the trade-offs points. It transforms qualitative feelings into quantitative values to bias the search during the optimization phase and restrict the objective space. In this sense, a reliable preference vector improves the trade-off answers obtained.

The reference point approach concentrates the search of Pareto non-dominated solutions near a selected point. It is based on the achievement scalarizing function that uses a reference point to capture the desired values of the objective functions. Let  $\mathbf{z}^0$  be a reference point for an  $n$ -objective optimization problem of minimizing  $F(\mathbf{x}) = \{f_1(\mathbf{x}), \dots, f_k(\mathbf{x})\}$ , the reference point scalarizing function can be stated as follows:

$$\sigma(\mathbf{z}, \mathbf{z}^0, \boldsymbol{\lambda}, \rho) = \max_{i=1, \dots, k} \{\lambda_i(z_i - z_i^0)\} + \rho \sum_{i=1}^k \lambda_i(z_i - z_i^0), \tag{1}$$

where  $\mathbf{z} \in \mathcal{Z}$  is one objective vector,  $\mathbf{z}^0 = \langle z_1^0, \dots, z_k^0 \rangle$  is a reference point vector,  $\sigma$  is a mapping from  $\mathbb{R}^k$  onto  $\mathbb{R}$ ,  $\boldsymbol{\lambda} = \langle \lambda_1, \dots, \lambda_k \rangle$  is a scaling coefficients vector,

and  $\rho$  is an arbitrary small positive number. Therefore, the achievement problem can be rebuilt as:  $\min \sigma(\mathbf{z}, \mathbf{z}^0, \boldsymbol{\lambda}, \rho)$ .

A subset  $X$  of  $\mathbb{R}^n$  is convex if for any two pair of solutions  $x^1, x^2 \in X$  and  $\alpha \in [0, 1]$ , the following condition is true:  $\alpha x^1 + (1 - \alpha)x^2 \in X$ . The intersection of all the convex sets containing a given subset  $X$  of  $\mathbb{R}^n$  is called the convex hull of  $X$ . The convex hull of a set of points is the smallest convex set that contains the points.

The *convex hull of individual minima* (CHIM) is the set of points in  $\mathbb{R}^k$  that are convex combinations of  $F_i^* - F^*$ :

$$\mathcal{H} = \left\{ \Phi\beta : \beta \in \mathbb{R}^k, \sum_{i=1}^k \beta_i = 1, \beta_i \geq 0 \right\} \tag{2}$$

where  $x_i^*$  is the global minimizers of  $f_i(\mathbf{x})$ ,  $\forall i \in \{1, \dots, k\}$ . Let  $F_i^* = F(x_i^*)$ ,  $\forall i \in \{1, \dots, k\}$ ; and  $\Phi$  is a pay-off matrix  $k \times k$  whose the  $i^{th}$  column is  $F_i^* - F^*$ .

Since the beginning of 2000, the development of social network technologies and interactive online systems has promoted a broader understanding of the “intelligence” concept. A new phenomenon appeared based not only on the cognition of one individual, but also placed on a network of relationships with other people and the external world. The field known as collective intelligence (COIN) [9] is defined as the self-organized group intelligence arisen from participatory and collaboration actions of many individuals. Shared tasks or issues are handled by singular contributions in such a manner that their aggregation process creates better results and solves more problems than each particular contribution separately.

Inside collective environment, contributions come from different people. Clustering algorithms distinguish the users with similar preferences to perform a cooperative evolution or decision making choice. A mixture model is a probabilistic model to reveal distributions of observations in the overall population. Given a data set  $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_N\}$  where  $\mathbf{y}_i$  is a  $d$ -dimensional vector measurement with the points created from density  $p(\mathbf{y})$ , a finite mixture model is defined as:

$$p(\mathbf{y}|\Theta) = \sum_{k=1}^K \alpha_k p_k(\mathbf{y}|z_k, \theta_k) \tag{3}$$

Let  $K \geq 1$  be the number of components,  $p_k(\mathbf{y}|z_k, \theta_k)$  be the mixture components where each  $k$  is a density or distribution over  $p(\mathbf{y})$  and parameters  $\theta_k$ ,  $\mathbf{z} = \langle z_1, \dots, z_k \rangle$  be a  $K$ -ary random variable defining the identity of the mixture component that produced  $\mathbf{y}$  and  $\alpha_k = p_k(z_k)$  are the mixture weights representing the probability that  $\mathbf{y}$  was generated by component  $k$ . Hence, the parameters for a mixture model is  $\Theta = \{\alpha_1, \dots, \alpha_K, \theta_1, \dots, \theta_K\}$ ,  $1 \leq k \leq K$ .

The Central Limit Theorem [8], explains why many applications that are influenced by a large number of random factors have a probability density function that approximates a Gaussian distribution. Let  $Y$  be a sequence of random variables that are identically and independently distributed, with mean  $\mu$  and

variance  $\sigma^2$ . The distribution of the normalised sum  $S_n = \frac{1}{\sqrt{n}}(\mathbf{y}_1 + \dots + \mathbf{y}_N)$  approaches the Gaussian distribution,  $G(\mu, \sigma^2)$ , as  $n \rightarrow \infty$ .

In a Gaussian mixture model, each of the  $K$  components is a Gaussian density with parameters  $\theta = \{\boldsymbol{\mu}_k, \Sigma_k\}$ ,  $\mathbf{y} \in \mathbb{R}^d$  and function as:

$$p_k(\mathbf{y}|\theta_k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_k)^t \Sigma_k^{-1}(\mathbf{y}-\boldsymbol{\mu}_k)} \quad (4)$$

### 3 Interactive Algorithms

Interactive genetic algorithms (IGA) were successfully applied to get feedbacks of transitional results during the evolution process. IGA incorporates external evaluation to support problems whose optimization objectives are complex to be defined by *a-priori* exact functions. It employs users' subjectivities as fitness values to drive the search engine.

As the purpose of this study is the enhancement of MOEAs through the use of the collective preferences, interactive genetic algorithms are an appropriate technique to support this goal. The new algorithms are extensions of the classical MOEAs: NSGA-II, SPEA2 and SMS-EMOA. The main changes on the original methods are the incorporation of COIN into the selection procedure; the transformation of the continuous evolutionary process into an interactive one; and the adoption of reference points to drive the search towards relevant regions in Pareto-optimal front.

#### 3.1 CI-NSGA-II

One of the new algorithms is a variation of NSGA-II [6]. The NSGA-II is a non-domination based genetic algorithm for multi-objective optimization. It adopts two main concepts: a density information for diversity and a fast non-dominated sorting in the population. The crowding distance uses the size of the largest cuboid enclosing two neighboring solutions to estimate the density of points in the front. The non-dominated sorting places each individual into a specific front such that the first front  $\tau_1$  is a non-dominant set, the second front  $\tau_2$  is dominated only by the individuals in  $\tau_1$  and so on. Each solution inside the front  $\tau_n$  receives a rank equal to its non-domination level  $n$ .

The selection operator uses the rank ( $i_{rank}$ ) and crowding distance ( $i_{dist}$ ) in a binary tournament. The partial order  $\prec_c$  between two individuals  $i$  and  $j$ , for example, prefers the minor domination rank if they are from different fronts or otherwise, the one with higher values of crowding distance.

$$i \prec_c j := i_{rank} < j_{rank} \vee (i_{rank} = j_{rank} \wedge i_{dist} > j_{dist}) \quad (5)$$

In Algorithm 1, the new CI-NSGA-II converts the original NSGA-II into an interactive process. The subroutine *CollectiveContributions()* suspends the evolution progress and submits some individuals from population to the users'

evaluation. In this research, the individuals received can be analyzed in two different ways: (a) a pairwise comparison allows the selection of the best candidate between two or more individuals; (b) a dynamic game scenario stimulates the participant creativity to improve or produce new individuals. Both approaches discover online reference points.

---

**Algorithm 1.** The Collective Intelligence NSGA-II.

---

```

1: generation ← numgeneration
2: block ← subsetgeneration           ▷ iteration interval
3: while i < generation do
4:   while block do
5:     offspring ← Tournament(pop)
6:     offspring ← Crossover(offspring)
7:     offspring ← Mutation(offspring)
8:     pop ← COIN Selection(offspring)
9:     i ++
10:   end while
11:   contributions ← CollectiveContributions(front)
12:   pop ← contributions
13:    $\Theta$  ← ExpectationMaximization(contributions)
14:   pop ← ReferencePoint Distance(pop,  $\Theta$ )
15: end while

```

---

Inside collective environment, the contributions come from different people. Assuming the Central Limit Theorem [8], the inputs have a distribution that is approximately Gaussian. Therefore, after each collective interaction, the subroutine *ExpectationMaximization()* gets the users' collaboration as a Gaussian Mixture model to emulate the evaluation landscape of all participants' preferences.

The expectation maximization approach creates online reference points ( $\Theta$ ) for search optimization. Whether the user's collaboration is a simple vote on the best individual presented to him (pairwise comparison) or a complete re-edited individual, the clustering algorithm distinguishes the users with similar preferences to perform a cooperative evolution and a decision making choice through the collective reference points.

The procedure *ReferencePointDistance()* calculates the minimum distance from each point in the population to the nearest collective reference points in  $\Theta$ . This way, the point near the reference point is favoured and stored in the new population. CI-NSGA-II develops a partial order similar to the NSGA-II procedure, but replaces the crowding distance operator by the distance to collective reference points ( $i_{ref}$ ).

$$i \prec_c j := i_{rank} < j_{rank} \vee (i_{rank} = j_{rank} \wedge i_{ref} < j_{ref}) \quad (6)$$

CI-NSGA-II prioritizes the points close to the online collective reference point. The algorithm consumes preference information to explore satisfactory solutions for DMs.

### 3.2 CI-SMS-EMOA

The SMS-EMOA [1] is a steady-state algorithm that applies the non-dominated sorting as a ranking criterion and the hypervolume measure ( $S$ ) as a selection operator.

After the non-domination ranking, the next step is to update the last front population,  $P_{worst}$ . It replaces the member with the minimum contribution to  $P_{worst}$  hypervolume by a new individual that increases the hypervolume covered by the population.

The new algorithm CI-SMS-EMOA converts the original SMS-EMOA into an interactive process. The *CollectiveContributions()* and *ExpectationMaximization()* subroutines have the same purpose and work as the CI-NSGA-II.

The selection operation, performed by the *COIN Selection()* procedure, prefers individuals with minor domination rank ( $i_{rank}$ ). If they belong to the same front, the one with the maximum contribution to the hypervolume of the set and the closest reference point distance ( $i_{ref}$ ) is selected.

A procedure *Hype-RefPoint Distance()* gets the hypervolume contribution ( $S$ ) and calculates the minimum distance from each solution in the population to the nearest collective reference points in  $\Theta$ . This way, the point with high hypervolume value and short reference point distance is favoured and stored in the new population.

### 3.3 CI-SPEA2

The strength pareto evolutionary algorithm 2 (SPEA2) [12] developed a fitness assignment strategy based on the number of individuals that one solution dominates and it is dominated by. SPEA2 implements elitism by keeping an external population (archive) of size  $N$ . The archive preserves the best solutions since the beginning of the evolution.

The strength  $ST(i)$  for each individual  $i$  is the number of population members it dominates:  $ST(i) = |\{j : j \in P_t \oplus \bar{P}_t \wedge i \prec j\}|$ ; where  $\oplus$  is the multiset union,  $P_t$  and  $\bar{P}_t$  are the population and archive population at generation  $t$ , respectively. The fitness  $F(i)$  for an individual  $i$  is given by the strength of its dominators:  $F(i) = \sum ST(j)$ ; where  $j \in P_t \vee \bar{P}_t$ ,  $j \prec i$ . High values of  $F(i)$  means the individual  $i$  is dominated by many others and  $F(i) = 0$  corresponds to a non-dominated individual.

SPEA2 uses a nearest density estimation technique, adapted from the kNN method, to distinguish individuals having the same fitness values. This density function is a function of the distance to the  $k$ -th nearest data point and it is added to the fitness function  $F$ .

In the new algorithm CI-SPEA2, subroutine *COIN Selection()* computes the strength of all individuals and the non-dominated members are copied to the archive  $\bar{P}_t$ . The  $k$ -th nearest data point used to calculate the original density function in SPEA2 was substituted by the collective reference points  $\Theta$ . If the archive  $|\bar{P}_t| \leq N$ , the algorithm chooses the nearest individuals to the collective reference point until the archive size is reached. Otherwise, if  $|\bar{P}_t| > N$ , it removes the more distant ones proportionally to the number of individuals in each reference point cluster. This way, the archive keeps the same distribution of points around its reference points.

## 4 Performance Indicators

Several performance indicators are used to evaluate the outcome sets of MOEAs. They measure the quality of the Pareto front approximation and allow comparison between different algorithms. There exist a variety of approaches that analyse the distribution of points in objective function space and the accuracy in terms of convergence.

The Pareto-optimal front coverage indicator,  $D_{S \rightarrow P_F}$ , is a proximity indicator that defines the distance between an achieved approximation set  $S$  and their closest counterpart in the current Pareto-optimal front:

$$D_{S \rightarrow P_F}(S) = \frac{1}{|S|} \sum_{\mathbf{x} \in S} \min_{\mathbf{x}' \in P_S} \{d(\mathbf{x}, \mathbf{x}')\}, \quad (7)$$

where  $d$  is the Euclidean distance between two points. Small values of  $D_{S \rightarrow P_F}$  indicate proximity to the  $P_F$ .

However, some others indicators like: coverage of two sets, diversity and the hypervolume could not be employed in this study. Because their values depend on the spread of solutions in the whole Pareto front and, on contrary, the proposed algorithms aim to obtain subsets of solutions close to the collective reference points. There is a lack of performance indicators that focus only on the proportion of occupied area in  $P_F$ . For that reason, the following subsection presents two new ones.

### 4.1 Referential Cluster Variance Indicator

Instead of a good spread of solutions along  $P_F$ , the method proposed in this work wants to obtain subsets of solutions close to the collective reference point. In this context, a small cluster variance means the individuals from the sample  $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_N\}$  of size  $N$  are clustered closely around the population mean ( $\mu$ ) or the reference point ( $\mathbf{z}^0$ ). A low dispersion for a group of preferred points in  $P_F$  denotes a better efficiency of the approach tested. The referential cluster variance indicator  $\kappa$  is represented as follows:

$$\kappa = \frac{1}{N} \sum_{i=1}^N (\mathbf{y}_i - \mu)^2 \quad (8)$$

In cases with more than one collective reference point ( $\mathbf{z}^j$ ), the points are clustered based on the closest distance to one of the reference points:  $C_j = \{\mathbf{a} \in \mathbb{R}^k : \|\mathbf{a} - \mathbf{z}^j\| \leq \|\mathbf{a} - \mathbf{z}^i\|, \forall i\}$ . Cluster  $C_j$  consists of all points for which  $\mathbf{z}^j$  is the closest. The referential cluster variance is calculated to each cluster separately.

### 4.2 Hull Volume Indicator

Convex hull is a well-known geometric object widely used in various fields such as: shape analysis, pattern recognition, geographical information systems, image

processing, etc. It has been also applied in the multi-objective optimization scenario.

The Normal Boundary Intersection method (NBI) [4] projects elements of the CHIM towards the boundary  $\partial\mathcal{Z}$  of the objective space  $\mathcal{Z}$  through a normal vector  $N$ . The intersection point between  $\partial\mathcal{Z}$  and  $N$  the normal pointing is a Pareto optimal point, if the  $P_F$  surface is convex. Martínez and Coello [10] introduced an archiving strategy based on the CHIM to find evenly distributed points along the  $P_F$ . Their convex hull multi-objective evolutionary algorithm (CH-MOEA) uses an archiving mechanism that stores non-dominated solutions which are orthogonal to each point of CHIM ( $h \in \mathcal{H}$ ). Likewise, Shan-Fan et al. [11] presented a MOEA where the non-dominated solutions are picked out from dominated solutions by the quick convex hulls algorithm.

The idea of convex hull can be borrowed and applied as a performance indicator to measure the quality of the non-dominated points around the collective reference points. Connecting the closest final points to each reference point will produce a facet representation of the  $P_F$ . The volume of the convex hull can be used as a scalar indicator for the distribution of points in  $P_F$ . Small values of the hull volume ( $\Psi$ ) indicate concentrate points.

Non-convex problems can use alpha shapes to determine a concave hull of their points in  $P_F$ . The alpha shape is a subgraph of the Delaunay triangulation. The value of alpha ( $\alpha$ ) controls the geometric design of the shape. For large  $\alpha$  values the shape approaches to the boundary of the convex hull. On the other hand, as  $\alpha$  decreases the shape shows more cavities.

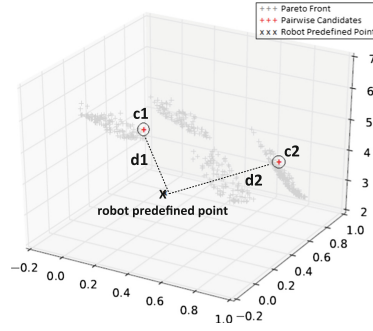
## 5 Experimental Results

This section presents some results of CI-NSGA-II, CI-SPEA2 and CI-SMS-EMOA. The scalable multi-objective test problems from the ZDT, DTLZ and WFG problem sets [2] have a known optimal front and can be used to benchmark the outcome of the algorithms. Their features cover different classes of MOPs: convex  $P_F$ , non-contiguous convex parts, non-convex, multi-modal, etc. For those reasons, the test problems subject the new algorithms to distinct optimization difficulties and compare their results.

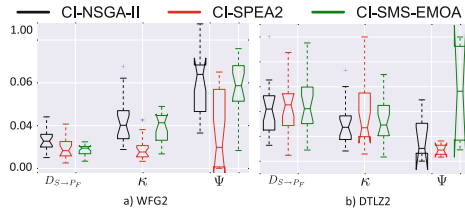
The experiment emulates the collectivity by developing some virtual DMs (robots). Each robot has a predefined point in the objective space which will be used to direct robots' votes in a pairwise comparison between two individuals from Pareto front. Therefore, the robot votes on a solution according to the closest distance between its predefined point and each of the two candidates. Figure 1 illustrates candidates  $c1$  and  $c2$  with their respectively distances to the predefined point,  $d1$  and  $d2$ . As  $d1 < d2$ , the robot would vote on  $c1$ . The robots create the collective reference points with a better reasoning strategy than simply random choice. It is important to notice that the collective reference point is built on the similarity of answers (votes) after the Gaussian Mixture model and cannot be confused with the robots' predefined points.

Very few MOEAs consider more than one user for reference point selection or evolutionary interaction. They neglect a collective scenario where many users





**Fig. 1.** The robot’s predefined point will choose candidate  $c1$  because the distance  $d1 < d2$ .



**Fig. 2.** Indicators for WFG2 and DTLZ2 tests.

could actively interact and take part of the decision process throughout the optimization. Furthermore, the new algorithms choose the reference points interactively. The references are not defined *a-priori*, like the R-NSGA-II from Deb [7], nor indicated by the DM as the middle point in the Light Beam approach [5]. Rather, all the references are discovered online with the support of a genuine collective intelligence of many users.

In this experiment, the robots abstract the collectivity within a controlled environment. So the algorithms can be tested, compared and better understood in their working principles. The quantity of online reference points is directly related to the number of  $k$  clusters in the Gaussian Mixture model.

The front coverage ( $D_{S \rightarrow P_F}$ ), the referential cluster variance ( $\kappa$ ) and the hull volume ( $\Psi$ ) indicators were used to measure the quality of the algorithms. In addition to the Gaussian Mixture model, the K-means algorithm was implemented to bring a different clustering technique into the analysis of the algorithms. But the performance of Gaussian Mixture for these benchmarking cases was consistently better.

After 30 independent executions per algorithm on each test problem, the box plots were used to represent and support a valid judgment of the quality of the solutions and how different algorithms compare with each other. Figure 2 shows the distribution of the performance indicators for WFG2 and DTLZ2 problems in the form of box plots.

**Table 1.** Results of the Conover-Inman statistical hypothesis tests based on the mean values. Green cells (+) denote cases where the algorithm in the row statistically was better than the one in the column. Cells marked in red (−) are cases where the method in the column yielded statistically better results when compared to the method in the row.

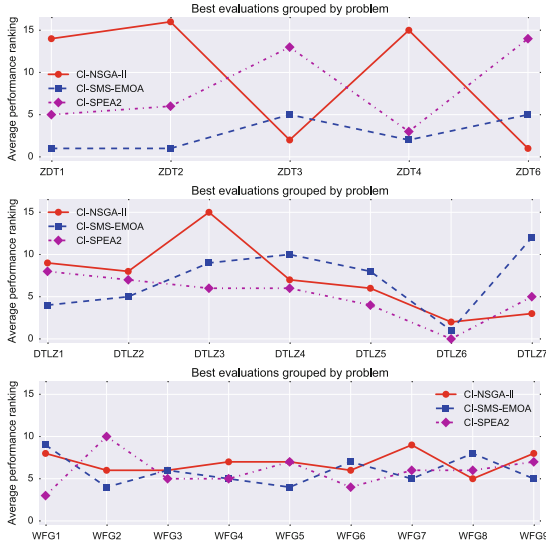
ZDT1			ZDT2			ZDT3			ZDT4		
CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
ZDT6			DTLZ1			DTLZ2			DTLZ3		
CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
DTLZ4			DTLZ5			DTLZ6			DTLZ7		
CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
WFG1			WFG2			WFG3			WFG4		
CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
WFG5			WFG6			WFG7			WFG8		
CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA	CI-NSGA-II	CI-SPEA2	CI-SMS-EMOA
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+

Although box plots allow a visual comparison of the results, it is necessary to go beyond reporting the descriptive statistics of the performance indicators and apply a statistical hypothesis test. The Conover-Inman procedure [3] is a non-parametric method for testing equality of population medians. It can be implemented in a pairwise manner to determine if the results of one algorithm were significantly better than those of the other. A significance level,  $\alpha$ , of 0.05 was used for all tests. Table 1 contains the results of the statistical analysis for all the test problems based on the mean values.

The process of discovering the best algorithm is rather difficult as it implies cross-examining and comparing the results of their performance indicators. Figure 3 present a more integrative representation by grouping their indicators. A higher value of average performance ranking implies that the algorithm consistently achieved lower values of the indicators being assessed. In this case, lower values mean better convergence to  $P_F$  and higher concentration around the collective reference points. For a given set of algorithms  $A_1, \dots, A_K$ , a set of  $P$  test problem instances  $\Phi_1, \dots, \Phi_P$ , the function  $\delta$  is defined as:

$$\delta(A_i, A_j, \Phi_p) = \begin{cases} 1 & \text{if } A_i \gg A_j \text{ solving } \Phi_p \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where the relation  $A_i \gg A_j$  defines if  $A_i$  is better than  $A_j$  when solving the problem instance  $\Phi_p$  in terms of the performance indicators:  $D_{S \rightarrow P_F}$ ,  $\kappa$  and  $\Psi$ . Relying on  $\delta$ , the performance index  $P_p(A_i)$  of a given algorithm  $A_i$  when solving  $\Phi_p$  is then computed as:  $P_p(A_i) = \sum_{j=1, j \neq i}^K \delta(A_i, A_j, \Phi_p)$ .



**Fig. 3.** Average performance ranking across ZDT, DTLZ and WFG test problems.

The CI-NSGA-II with Gaussian Mixture model consistently outperformed the others algorithms in most cases. Concerning the convergence indicator, CI-NSGA-II was ranked best in 16 of 21 test functions, except for ZDT1, ZDTL6, DTLZ1, WFG6 and WFG8. The CI-SPEA2 and CI-SMS-EMOA have a similar performance on WFG tests. It is worth noticing that ZDT4 experiment demonstrated a premature convergence around the online reference points. A better control of the extent of obtained solutions must be investigated to avoid this behaviour.

In summary, the interactive MOEAs and their collective reference points proved to be well matched for the range of ZDT, DTLZ and WFG test problems.

## 6 Final Remarks

In this work we have introduced new algorithms to improve the successive stages of evolution via dynamic group preferences. The interactive algorithms can apprehend people’s heterogeneity and common sense to guide the search through relevant regions of Pareto-optimal front.

The new approaches have been tested successfully in benchmarking problems. Two different performance indicators were presented with the intention to measure the proportion of occupied area in  $P_F$ .

Currently, we have been exploring different features of the evolutionary process, such as: the usage of COIN as a local search for new individuals and opening the population for users’ update to augment its quality. Replacing the robots by human collaborators, this research has a simulated environment<sup>1</sup> that

<sup>1</sup> <http://playcanv.as/p/1ARj738G>.

represents the facility location problem. In this context, individuals from the evolutionary algorithm population are distributed to the participants who have to fix and change the position arrangement or the resource distribution in the scenario. This simulated environment promotes the collaboration and supports rational improvements in the quality of the population. The algorithms (CI-NSGA-II, CI-SMS-EMOA, CI-SPEA2) introduced and compared here are used to iteratively refine the search parameters with collective preferences.

In the near future, we plan to work on creative solutions unfolded by the participants and possibly explore non-explicit objectives hidden in more complex scenario. Furthermore, we intent to apply directional information from the collective reference points during the evolution process. This way, the technique can extract the intelligence of the crowds and, at the same time, minimize the interruptions of the algorithm.

**Acknowledgments.** This work was partially funded by CNPq BJT Project 407851/2012-7, FAPERJ APQ1 Project 211.500/2015, FAPERJ APQ1 Project 211.451/2015.

## References

1. Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: Multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* **181**(3), 1653–1669 (2007)
2. Coello, C.C., Lamont, G.B., Van Veldhuizen, D.A.: *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York (2007)
3. Conover, W.: *Practical Nonparametric Statistics*. Wiley, New York (1999)
4. Das, I., Dennis, J.E.: Normal-boundary intersection: a new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM J. Optim.* **8**(3), 631–657 (1998)
5. Deb, K., Kumar, A.: Light beam search based multi-objective optimization using evolutionary algorithms. In: *IEEE Congress on Evolutionary Computation, 2007, CEC 2007*, pp. 2125–2132. IEEE (2007)
6. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **6**(2), 182–197 (2002)
7. Deb, K., Sundar, J., Udaya Bhaskara Rao, N., Chaudhuri, S.: Reference point based multi-objective optimization using evolutionary algorithms. *Int. J. Comput. Intell. Res.* **2**(3), 273–286 (2006)
8. Grinstead, C.M., Snell, J.L.: *Introduction to Probability*. American Mathematical Soc., Providence (2012)
9. Malone, T.W., Laubacher, R., Dellarocas, C.: *Harnessing crowds: mapping the genome of collective intelligence* (2009)
10. Martinez, S.Z., Coello, C.A.C.: An archiving strategy based on the convex hull of individual minima for MOEAs (2010)
11. Shan-Fan, J., Xiong, S.W., Zhuo-Wang, J.: The multi-objective differential evolution algorithm based on quick convex hull algorithms. In: *Fifth International Conference on Natural Computation, 2009, ICNC 2009*, vol. 4, pp. 469–473. IEEE (2009)
12. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm (2001)