Using Collective Intelligence to Support Multi-Objective Decisions: collaborative and online preferences

Daniel Cinalli*, Luis Martí*, Nayat Sanchez-Pi†, Ana Cristina Bicharra Garcia*
*UFF - Universidade Federal Fluminense
{dcinalli,lmarti, bicharra}@ic.uff.br
†UERJ - Universidade Estadual do Rio de Janeiro
nayat@ime.uerj.br

***Abstract*—This research indicates a novel approach of evolutionary multi-objective optimization algorithms meant for integrating collective intelligence methods into the optimization process. The new algorithms allow groups of decision makers to improve the successive stages of evolution via users' preferences and collaboration in a direct crowdsourcing fashion. They can, also, highlight the regions of Pareto frontier that are more relevant to the group of decision makers as to focus the search process mainly on those areas. As part of this work we test the algorithms performance when face with some synthetic problem as well as a real-world case scenario.**

Keywords*—*Collective intelligence, preferences, reference points, evolutionary multi-objective optimization algorithms.

I. Introduction

Many real-life decision problems require managing trade-offs between multiple objectives. These problems can be posed as multi-objective optimization problems (MOPs) and must consider more than one criteria to be simultaneously optimized [1]. In the general case, there is not a single optimal solution which optimizes all the objectives at the same time, but a set of points that represent different trade-offs between the objectives known as Pareto-optimal set. According to some *a priori* high-level preferences, a decision maker (DM) has to select which of those solutions are the ones to satisfy its needs.

In some applications, the approximation of the entire Pareto-optimal set requires extensive time and computational resource. MOPs tend to be NP-hard or NP-complete. Therefore, deterministic search techniques are usually unsuitable for this task. Metaheuristic and stochastic approaches are a viable alternative to handle this difficulty. These circumstances have attracted attention inside the evolutionary computation community and have prompted the creation of what has been called multi-objective optimization evolutionary algorithms (MOEAs) [2].

MOEAs can use decision makers' preferences to improve the reference point or fitness function and drive the search process by focusing only on preferred solutions, instead of all possible ones. A global outcome built on the aggregation of vast and diverse masses of individual DM intelligence would be a helpful input parameter, rather than a unique or small group of decision makers provided with unilateral preferences. The interaction of group preferences explores a wider diversity of answer and enhance multi-objective results.

The present work proposes and compares new preference-based interactive MOEAs augmented by collective intelligence (COIN) [3]. It introduces a new useful connection between these two fields and extends some of the current state-of-the-art MOEAs. The hybrid algorithms aggregate consistent and online preferences from a collective environment to the optimization process. Built upon people collaboration and heterogeneity, the intelligence of participatory actions explores creative solutions on preferred areas.

The combination of COIN in MOEAs aims at improving the quality of the obtained Pareto frontier approximation. Their results are driven not by one decision maker, but a group of people that delimits their collective area of interest in the objective space. The new algorithms produce better solutions in the sense that they iteratively refine the search parameters and get users collaborations to generate more appropriated points in the final trade-off set.

The rest of this paper is organized as follows. Section II covers some required formal definitions of multi-objective optimization and collective intelligence field. Section III presents the new algorithms CI-NSGA-II, CI-SMS-EMOA and CI-SPEA2 based on interactive collective intelligence techniques. Some results from benchmark problems and a resource placement case study are analyzed in Section IV. Finally, Section V puts forward some conclusive remarks and future work plans.

II. Basic Concepts

MOP can be stated as follows:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{F}(\boldsymbol{x}) = \{f_1(\boldsymbol{x}), \dots, f_k(\boldsymbol{x})\}, \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, \\ & h_j(\boldsymbol{x}) = 0, \end{aligned} \tag{1}$$

where $\boldsymbol{x} = \langle x_1, \dots, x_n \rangle \in \Omega$ is an n-dimensional decision variable. Thus, a MOP consists of k objectives, $m + p$ constraints, n decision variables and an evaluation function $\boldsymbol{F} : \Omega \to \mathcal{Z}$ that maps from the vector $\boldsymbol{x}$ to output vectors $\boldsymbol{a} = \langle a_1, \dots, a_k \rangle$. The solution to this problem can be expressed by relying on the Pareto dominance relationship. An

978-1-4673-9775-9/15 $31.00 © 2015 IEEE
DOI 10.1109/ASEW.2015.12

$\boldsymbol{x}$ is said to dominate $\boldsymbol{v}$ (denoted as $\boldsymbol{x} \prec \boldsymbol{v}$) iff $\forall i \in \{1, \ldots, n\}$, $x_i \leq v_i \wedge \exists i \in \{1, \ldots, n\}$ such that $x_i < v_i$. A solution $\boldsymbol{x} \in \Omega$ is Pareto optimal if there does not exist another solution $\boldsymbol{x'} \in \Omega$ such that $F(\boldsymbol{x'}) \prec F(\boldsymbol{x})\}$.

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 is defined as the self-organized group intelligence arisen from participatory and collaboration actions of many individuals. Shared tasks are handled by singular contributions, but their aggregation process creates better results and solves more problems than each particular contribution separately [3]. This phenomenon develops a *sui generis* intelligence. It raises a global experience of collective attitudes without centralized control, bigger than its isolated pieces and sub-product of their combination.

COIN involves groups of individuals collaborating to create synergy and augment the intellectual processes of human beings. A decision-making process over the Internet has to manage users' interactions. It must get valuable knowledge concealed or dispersed in the group, even when the participants are not specialized in the subject. This environment includes large and heterogeneous audiences that are mostly independent among each others. Therefore, the problem must be decomposed in tasks that sustain diversity and transient members' attendances to align the interest of crowds.

III. Algorithms

DMs must choose one solution from a potentially large —or perhaps, infinite— optimal frontier (P_F). This work proposes the application of collective intelligence as a reference to drive the optimization process of MOEAs towards relevant regions in Pareto-optimal set. Consequently, instead of the entire front, it reaches a smaller sub-set of the front and use the collective preferences to support decisions upon multi-objective situations.

In addition, this method manages the collaboration of people to carry out a local search for new creative solutions. Following the Find-Fix-Verify method, some individuals from population are given to the users in order to get their update and feedback. This approach encourages the interaction of multiple participants and takes rational collaborations to improve the overall quality of EA population.

Complex scenarios that are hard for computer might be easier or natural for human's mind. People are able to improve the multi-objective algorithms with cognitive and subjective evaluation to find better solutions. Characteristics such as perception, strategy, weighting factors, agility, among others subjectivities might be introduced into the algorithm to generate a better pool of answers and enhance the optimization.

A. CI-NSGA-II

The new algorithm CI-NSGA-II converts the original NSGA-II [4] into an interactive process. It suspends the evolution progress and submits some individuals from population to the users' evaluation. A participant can update the received individual and send it back to the evolution process.

After each collective interaction, a Gaussian Mixture model is used to emulate the evaluation landscape of all participants' preferences. As a result, online reference points are created with the expectation maximization approach and several rational improved individuals are inserted into the population.

The CI-NSGA-II develops a partial order similar to the NSGA-II procedure, but replaces the original crowding distance operator by the distance to collective reference points (i_{ref}). The selection operator is based on this new partial order. Like NSGA-II, individuals with minor domination rank are preferred. But if they belong to the same front, the one with the closest reference point distance is used instead.

$$i \prec_c j := i_{rank} < j_{rank} \vee (i_{rank} = j_{rank} \wedge i_{ref} < j_{ref}) \quad (2)$$

B. CI-SMS-EMOA

The new algorithm CI-SMS-EMOA converts the original SMS-EMOA [5] into an interactive process. A Gaussian Mixture model is used to create online reference points based on the participants' collaboration.

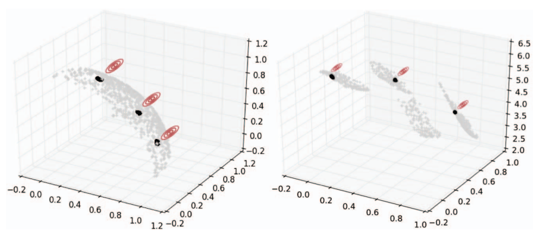
The difference between the original and CI-SMS-EMOA algorithm is the selection operator. Both apply the non-dominated sorting as a ranking criterion and prefer individuals with minor domination rank (i_{rank}). But if the individuals belong to the same front, CI-SMS-EMOA chooses not only the one with the maximum contribution to the hypervolume [6] of the set, but also the one with the closest reference point distance (i_{ref}).

C. CI-SPEA2

The new algorithm CI-SPEA2 converts the original SPEA2 [7] into an interactive process. In the selection operator of CI-SPEA2, the strength of all individuals are computed 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 Θ. If the archive $\mid \bar{P}_t \mid \leq N$, the algorithm chooses the nearest individuals to the collective reference point until the archive size is reached. Otherwise, if $\mid \bar{P}_t \mid > N$, 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.

IV. Results

This section presents some results of CI-NSGA-II, CI-SMS-EMOA and CI-SPEA2. The multi-objective test problems ZDT [8] and DTLZ [9] have a known optimal front and can be used to benchmark the outcome of the algorithms. A real-world case is formally introduced afterwards and submitted to a COIN experiment.



(a) DTLZ2 (b) DTLZ7

Fig. 1: CI-NSGA-II results for DTLZ2 and DTLZ7 problems.

A. Multi-Objective Test Problems

ZDTs and DTLZs are a set of well established scalable multi-objective test problems. Extensively used in MOEA studies, these benchmark problems were selected to analyse the behaviour of the proposed COIN MOEAs algorithms in the first moment. Each of these test functions knows *a priori* the exact shape and location of Pareto-optimal front.

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 create the collective reference points. Figure 1 shows the relevant regions found in Pareto front to the DTLZ2 and DTLZ7 problems.

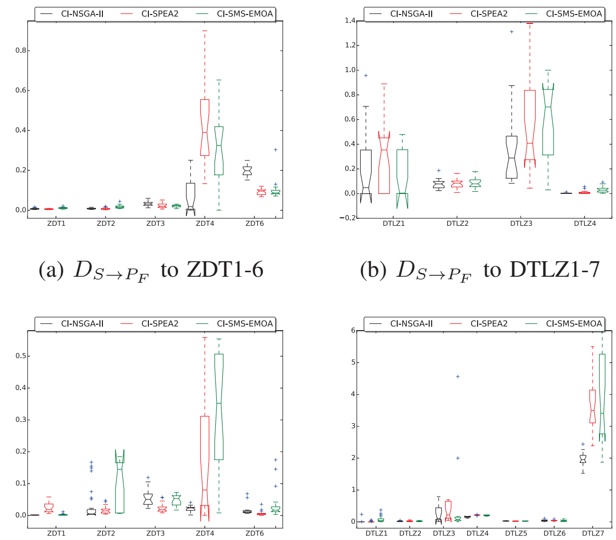
In addition to the Gaussian Mixture model, the K-means algorithm was implemented to bring a different clustering technique into the analysis of the algorithms. But the performance of Gaussian Mixture in these cases was consistently better than K-means. The front coverage ($D_{S \to P_F}$) and the variance (σ^2) indicators were used to measure the quality of the results [10].

After 30 independent executions per EA on each test problem, figure 2 reports the distribution of the front coverage and dispersion indicators in the form of box plots, respectively. The CI-NSGA-II with Gaussian Mixture model consistently outperformed the CI-SMS-EMOA and CI-SPEA2 in these benchmarks. Concerning the convergence and dispersion measures, it was ranked best in all functions except for ZDT3 and ZDT6. CI-NSGA-II and its collective reference points proved to be well matched for the range of these synthetic problems.

B. Resource Distribution problem

Many companies face problems of resource placement and assignment. A mining industry is one of the domain contexts where these problems are present. Those companies must extract valuable minerals or other geological materials from resource areas and allocate warehouses in such a way that optimizes its operational costs and production of collected resources. This general idea transforms the resource management into a multi-objective problem where one have to operate in an economic way and, at the same time, prioritize the performance or production.

The problem —to put it in simple terms— has to find a good solution for positioning the processing units according



(a) $D_{S \to P_F}$ to ZDT1-6 (b) $D_{S \to P_F}$ to DTLZ1-7

(c) σ^2 to ZDT1-6 (d) σ^2 to DTLZ1-7

Fig. 2: Distribution of the front coverage and dispersion indicators in the form of box plots.

the resource area. It is formally represented as:

$$\min \sum_{i=1}^{N} \sum_{j=1}^{M} \sigma_{ij} d_{ij} + \sum_{j=1}^{M} c_j \mu \,, \tag{3}$$

$$\max \sum_{i=1}^{N} \sum_{j=1}^{M} \sigma_{ij} v_j \,. \tag{4}$$

Let μ be the cost of one processing unit, v the productive capacity of one processing unit linked to one resource area, M a set of available positions to production units, N a set of available positions to resource area and D a distance matrix $(d_{ef})_{nxm}$, where $n \in N$ and $m \in M$. The decision variables are the processing unit c_j ($j \in M$) that assumes 1 if it is placed at position j or 0 otherwise and σ_{ij} that assumes 1 if there is a link between the resource area at position $i \in N$ and the processing unit at position $j \in M$.

Different constraints from real life and several new interdependencies among the variables might increase the search complexity of this MOP. Progressive articulation of preferences and collective intelligence can implement a dynamism not managed by *a priori* methods and enhance its efficiency. Therefore, the problem described is a candidate for this experiment due to some reasons: a) the objectives and decision variables are meaningful to the group, the problem is intuitive and allows an interaction with the crowd's cognition; b) incentive engines and gamification can be used to retain the users' interest on the interaction during the optimization; c) the problem can be decomposed in small blocks to be presented to the participants; d) the users' feed-backs can be parallelized in synchrony with the evolution of individuals in a MOEA.

TABLE I: The performance of the algorithms.

Algorithm	$D_{S \to P_F}$	σ^2
NSGA-II	0,7	189
CI-NSGA-II	**0,5**	**112**
SMS-EMOA	0,8	205
CI-SMS-EMOA	0,6	132
SPEA2	0,8	180
CI-SPEA2	0,7	135

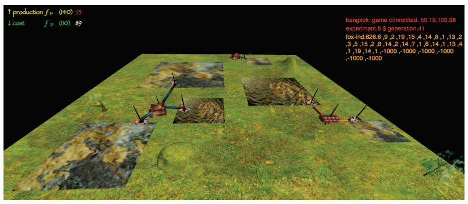


Fig. 3: Evolutionary algorithm solution for six areas.

In this context, the resource distribution problem was designed as a game where every player compete among themselves to obtain points and recognition of success. The game was implemented in a web-based platform and is open to all public. Some individuals from population are distributed to the players who have to fix and change their position arrangement. Following the problem definition and constraints, the dynamic scenario (figure 3) allows the creation of objects like trucks or warehouses, changing their arrangements and rebuilding their connections.

The experiment was applied in two different computer labs: a Brazilian undergraduate institution (SENAI) with more than 30 students' attendance and a private company training room. The front coverage indicator, $D_{S \to P_F}$, measures the distance between the current approximation set S and the Pareto-optimal front. The proximity to the Pareto-optimal front $D_{S \to P_F} = 0.5$ is the criteria to stop the evolution and compare the algorithms. The values in table I represent the mean of all completed games.

All the new algorithms iteratively refine the search parameters and adopt players collaborations to achieve more appropriated points in the final trade-off set. According to the table I, the CI-NSGA had a lower dispersion σ^2, which means the points are clustered closely around the collective reference point. In 250 generations, it consistently reached the convergence $D_{S \to P_F} = 0.5$, while the others algorithms would need more generations to succeed on that.

At the end of each game, only one scenario is presented to the players. Figure 3 shows the final solution from a single experiment game. After the collective reference points and users contributions in a game composed of six resource areas and two types of processing units, this candidate was progressively created with the support of users subjectivity and perception. From the group's point of view, it is the best alternative (winner candidate) and overcomes many others optimal points in the front.

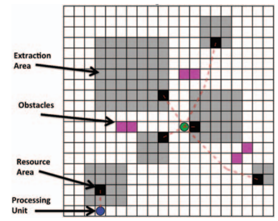


Fig. 4: Solution for six areas with obstacles.

Further studies on more complex scenarios intent to analyse the performance of CI-NSGA-II. Obstacles and zigzag routes were already implemented to the resource distribution problem (figure 4). In this context, new experiments are planned to be executed in collaborative environments.

V. Final Remarks

In this work we have introduced an interactive approach in multi-objective optimization evolutionary algorithms. The new algorithms use dynamic group collaboration to guide the search through relevant regions of Pareto-optimal front and discover creative resolutions. It is an opportunity to handle collective subjectivity, social creativity and cognition into the MOEAs optimization process. Results outlined the benefits of collective collaborations to unfold solutions designed by a group of people that is more intelligent when is working together.

In the near future, we plan to explore more complex scenarios with many constraints and non-explicit objectives hidden in the problem. It is important to validate if the complexity of the environment will favour even more the integration of COIN in MOEAs.

References

[1] M. Ehrgott, *Multicriteria optimization*. Springer, 2005, vol. 2.

[2] C. C. Coello, G. B. Lamont, and D. A. Van Veldhuizen, *Evolutionary algorithms for solving multi-objective problems*. Springer, 2007.

[3] J. Surowiecki, *The wisdom of crowds*. Random House LLC, 2005.

[4] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *Evolutionary Computation, IEEE Transactions on*, vol. 6, no. 2, pp. 182–197, 2002.

[5] N. Beume, B. Naujoks, and M. Emmerich, "Sms-emoa: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.

[6] M. Emmerich, N. Beume, and B. Naujoks, "An emo algorithm using the hypervolume measure as selection criterion," in *Evolutionary Multi-Criterion Optimization*. Springer, 2005, pp. 62–76.

[7] E. Zitzler, M. Laumanns, L. Thiele, E. Zitzler, E. Zitzler, L. Thiele, and L. Thiele, "Spea2: Improving the strength pareto evolutionary algorithm," 2001.

[8] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary computation*, vol. 8, no. 2, pp. 173–195, 2000.

[9] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, *Scalable test problems for evolutionary multiobjective optimization*. Springer, 2005.

[10] P. A. Bosman and D. Thierens, "The balance between proximity and diversity in multiobjective evolutionary algorithms," *Evolutionary Computation, IEEE Transactions on*, vol. 7, no. 2, pp. 174–188, 2003.