Performance Measurement for Interactive Multi-objective Evolutionary Algorithms

Long Nguyen
Department of Information Technology
National Defense Academy
Hanoi, Vietnam
Email: longit76@gmail.com

Hung Nguyen Xuan
Faculty of Information Technology
Le Quy Don Technical University
Hanoi, Vietnam
Email: nguyenxuanhung@outlook.com

Lam Thu Bui
Faculty of Information Technology
Le Quy Don Technical University
Hanoi, Vietnam
Email: lam.bui07@gmail.com

***Abstract*—This paper suggests to use a different metric for performance of multiple-point interactive evolutionary multi-objective algorithms. We defined a preferred region based on a set of user's reference points. Based on the preferred region, we also define a User based Front (UbF) which is generated from the preferred region. UbF is used in calculation of Generational Distance (GD) and Inverse Generational Distance (IGD). The usage of the metric in experiments indicated meaningful comparisons for interactive multi-objective evolutionary algorithms using multiple reference points.**

I. Introduction

We need to simultaneously optimize several objective functions in order to solve multi-objective optimization problems (MOPs)[1]. As a result, we usually obtain *trade-offs*, which are called *Pareto optimal solutions* or *Pareto optimal Front*. Methods for multi-objective optimization can be classified into several classes including *the Interactive methods*. With the interactive methods, Decision Maker (DM) iteratively directs the searching process by indicating her/his preference information over the set of solutions until DM satisfies or prefers to stop the process[2]. During the optimal process (iteratively) DM is able to learn about the underlying problem as well as her/his own preference. To date, many interactive techniques have been proposed [3], [4], [5], [6], [7], [8] for solving MOPs. It is worthwhile to note that the aim of the interactive method is to find the most suitable solution in several conflicting objectives regarding the DM's preference. It requires a mechanism to support DM in formulating her/his preferences and identifying preferred solutions in the set of Pareto optimal solutions.

The usage of multi-reference points is surveyed and implemented for MOEA/D[9] and DMEA-II[10] in recent proposals, which are described in the next section. In these methods, DM gives a set of reference points in the objective space at several generations. A point is aggregated from the set of reference points and it is used as primary DM's preference. In this way, DM has more flexibility to express her/his preference.

There are not many performance measurements for interactive user-preference based multi-objective evolutionary algorithms (MOEAs). In [11], the authors extended the standard Hypervolume (HV) metric in the manners for one preferred region in three steps: 1) Obtain the solution point closest to the ideal point; 2) Define a volume for HV calculation; 3) Filter solution points and calculate the HV. In [12] the authors combine the solution sets of the algorithms that are to be compared and extracted the non-dominated solutions into a composite front. The composite front is used to defined a preferred region based on the location of a user-supplied reference point in the objective space an a threshold r.

However, there are drawbacks of these metrics: In [11], a preferred region is determined from the location of the ideal point. This causes misleading results when reference point is biased towards one objective more than other objectives. In other case, many high quality solutions to fall outside the preferred region when DM chooses a bad ideal point. The proposed metric in [12] is highly competitive but it only works on single reference point interactive methods.

In this paper, we propose a new performance metric for interactive multi-objective evolutionary algorithms based on defining the preferred region from multiple reference points. In the remainder of the paper, section II gives an overview about multiple reference point interactive methods for MOEAs. Thereafter, in Section III, a new performance metric for user preference based MOEAs is introduced. In section IV, we use the proposed metric to compare two user preference based MOEAs. Finally, the conclusion of this paper is outlined in section V.

II. Multi-point interactive methods

The reference point interactive method is suggested by Wierzbicki[13], this method is known as classical reference point approach. The idea of the method is in order to control the search by reference points using *achievement functions*. Here the achievement function is constructed in such a way that if the reference point is dominated, the optimization will advance past the reference point to a non-dominated solution. A reference point z^* is given for an M-objective optimization problem of minimizing $(f_1(x), \ldots, f_k(x))$ with $x \in S$. Then single-objective optimization one as following:

$minimize$

$$max_{i=1}^{M}[w_i(f_i(x) - z_i^*)] \tag{1}$$

subject to $x \in S$.

In Fig.1: z_a, z_b are reference points, w is chosen weight vector used for scalarizing the objectives. The algorithm for this method is described in five following steps:

Step 1: Present information to the DM. Set h=1

Step 2: Ask the DM to specify a reference point z_*^h

Step 3: Minimize achievement function. Present z^h to the DM

Step 4: Calculate k other solutions with reference points $\overline{z}(i) = \overline{z}^h + d^h e^i$ where $d^h = ||z_*^h - z^h||$ and e_i is the i^{th} unit vector

978-1-4673-8013-3/15 $31.00 © 2015 IEEE
DOI 10.1109/KSE.2015.51

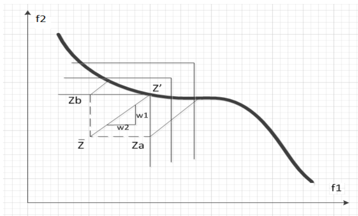

Fig. 1. Altering the reference point

Step 5: If the DM can select the final solution, stop. Otherwise, ask the DM to specify z_*^{h+1}. Set $h = h + 1$ and go to **Step 3**. Here h is the number that DM specifies a reference point during process. By the way of using the series of reference points, DM actually tries to evaluate the region of Pareto Optimality, instead of one particular Pareto-optimal point. However DM usually deals with two situations:

1) The reference point is feasible and not a Pareto-optimal solution, the DM is interested in knowing solutions which are Pareto-optimal and near the reference point.
2) The DM find Pareto-optimal solutions which near supplied reference point.

There are some surveys on the usage of multi-reference points for MOEAs such as interactive methods for MOEA/D and DMEA-II. These reports indicate that the usage of multiple reference points instead of using a single point for following reasons:

1) In some cases, DM does not have an exactly point of preference. So it is better to give them a facility for defining the region of interest.
2) The conventional interactive methods require DM giving several reference points via many iterations. Sometimes, they might be close in a region. Hence, it might be convenient for them to give these points at an one-go.

This paper used the proposed performance metric for two MOEAs: MOEA/D and DMEA-II with proposed interactive methods respectively.
In the interactive method for MOEA/D, the authors proposed an interactive method using multiple reference points with multi-objective optimization based on decomposition-based MOEA (MOEA/D). In the alternative method the authors use a set of reference points in objective space to represent for DM's preferred area. The aggregated point formed set of reference points is used in optimal process by two ways: replace or combine the current ideal point at the loop. In the experimental study the authors use ZDTs problems with two objectives. The ideal point replacement approach is used. In the interactive method for DMEA-II, the authors proposed an interactive method using multiple reference points with direction based multi-objective evolutionary algorithm-II (DMEA-II). In the alternative method a set of new rays is generated from reference points given by DM in objective space. These rays will replace corresponding the farthest rays to DM's preferred region. By applying a niching with new adjusted distribution of rays, the final solutions strongly converged to the DM's preferred region. It ensures convergence and spreading of population and concept to use two kinds of improvement directions. With the interactive method, DM can get the most preferred solutions and concept of using two kind of improvement directions: Spread direction and Convergence direction. The Ray replacement approach is used in this paper.

III. A NEW PERFORMANCE METRIC

A. *Conventional metrics for MOEAs*

In order to allow a quantitative comparison of results among different algorithms, there are two distinct goals that we pursue: (1) obtaining the solutions as close to the Pareto optimal solutions as possible (closer to the true Pareto front) and (2) obtaining the solutions as diverse as possible along the Pareto front (good distribution of solutions). Apparently, these two goals are independent from each other and there exist different performance measures to deal with one or both of the goals. Thus, it does not exist a single performance measure that can indicate the superiority of one algorithm over another in these two aspects. So, there is a clear need of having at least two performance measures for adequately evaluate both goals (convergence and diversity) of an MOEA.
An MOEA will be termed as a good multi-objective solver if both goals are properly satisfied. This is, it is expected to find solutions that are very close to the true Pareto front and, at the same time, are well spreaded along the Pareto front. Two typical performance metrics are commonly used for MOEA validation are listed:

- **The generational distance (GD)** [14] is defined as the average distance from a set of solutions, denoted *P*, found by evolution to the global POF. The first-norm equation is defined as

$$GD = \frac{\sum_{i=1}^{n} d_i}{n} \tag{2}$$

  where d_i is the Euclidean distance (in objective space) from solution *i* to the nearest solution in the POF, and n is the size of P. This measurement is considered for convergence aspect of performance. Therefore, it could happen that the set of solutions is very close to the POF, but it does not cover the entire the POF.

- **The inverse generational distance (IGD)** [14]: The measurement takes into account both convergence and spread to all parts of the POF. The first-norm equation for IGD is as follows:

$$IGD = \frac{\sum_{i=1}^{\overline{N}} \overline{d_i}}{\overline{N}} \tag{3}$$

  where $\overline{d_i}$ is the Euclidean distance (in objective space) from solution *i* in the POF to the nearest solution in P, and $\overline{N}$ is the size of the POF. In order to get a good value for IGD (ideally zero), P needs to cover all parts of the POF. However, this method only focuses on the solution that is closest to the solution in the POF indicating that a solution in P might not take part in this calculation.

Based on GD, IGD we suggest a new performance metric for interactive MOEAs.

B. A new performance metric for interactive MOEAs

In interactive MOEAs, during generations, user preference is given to evolutionary process in order to drive the population towards user's preferred region. When stopping conditions are met, obtained solutions are converged to the preferred region. Based on the set of final solutions, we need to measure how good is user preference based MOEAs' performance. The criteria for the measurement is: solutions convergence and diversity with respect to the user's preferred region?
For the multiple-point interaction, we defined a preferred region as a boundary of reference points, a User based Front (UbF) is determined by a set of points which is generated from preferred region as the following:
From the boundary, a grid with size $M \times M$ and $M = \sqrt{m}$, m is the population size. Each node of the grid is a point of UbF. In case of single reference point, we can build the grid from the reference points as middle point with a specified size. The illustration of determining UbF is shown in Fig 2.
We define two measurements: iGD and iIGD which use UbF

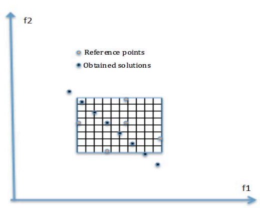

Fig. 2. Determining User based Front (UbF) in an objective space

as bellow:

$$iGD = \frac{\sum_{i=1}^{m} d_i}{m} \tag{4}$$

$$iIGD = \frac{\sum_{i=1}^{M} \overline{d_i}}{M} \tag{5}$$

Here, d_i is the shortest distance from solution i in the obtained set to a point in UbF. $\overline{d_i}$ is the shortest distance from point i in UbF to a solution in the obtained population. The metrics are simulated in Fig 3. Not change the properties of the

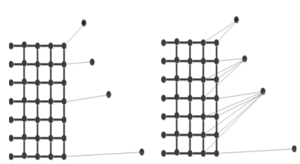

Fig. 3. Illustration of iGD (Left) and iIGD (Right).

comparison, we normalize the results in range of $[0, 1]$ as bellow:

$$iGD_k = \frac{iGD_k^*}{\max(iGD)} \tag{6}$$

Here, on a test problem, iGD_k is the iGD value for algorithm k, iGD_k^* is the original value of algorithm k which is calculated by 4, N is the number of algorithms in comparison.

$$iIGD_k = \frac{iIGD_k^*}{\max(iIGD)} \tag{7}$$

Here, on a test problem, $iIGD_k$ is the $iIGD$ value for algorithm k, $iIGD_k^*$ is the original value of algorithm k which is calculated by 5, N is the number of algorithms in comparison. The iGD for convergence measurement and $iIGD$ for both convergence and diversity measurements.
The difference between our proposed metric and the proposals in [12], [11] are: In [12] a composite front is generated from non-dominated solutions of a merged set which is collected from all algorithms. The preferred region is determined as a circle which is created from the closest solution to reference points, the radius of the circle is given by user as a parameter. The composite is used as POF in IGD and HV metrics. The disadvantages of this proposal: the metric results are dependent on the radius parameter which is given by user. It is not useful that user has to set many parameters in case of multiple reference points. In [11], the metric results are dependent on the choice of ideal point, in this case, many hight quality solutions fall outside the preferred region.
To resolve these issues, we suggest to use new user based performance metric. In our proposal, the preferred region is determined based on location of reference points only, this is the main criterion of the metric.

IV. Experiments

A. Benchmark sets and parameters

In our experiments, we use 5 popular test problems designed by Zitzler, Deb and Thiele:[15] and 5 unconstraint problems by Liu, Zou and Wu [16] with two objectives that POFs are *convex, non-convex, convex and disconnected, non-convex, non-uniformly spaced*. These test problems are used for two selected interactive MOEAs: MOEA/D and DMEA-II. The common testing parameters for these problems are: number objectives (2); number variables (30); population size (100); number of generations (1000). For MOEA/D and DMEA-II experiments, the mutation rate was kept at the same small rate of 0.01, and the perturbation rate was 0.4.

B. Results and Discussion

With the the same parameters for iterative DMEA-II and iterative MOEA/D and the same reference points are given during optimal process, example snapshots are shown in Fig 4, 5.

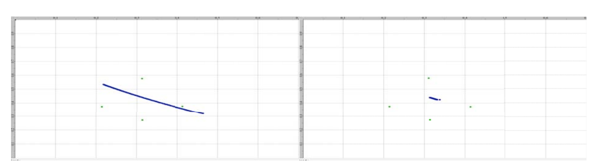

Fig. 4. The obtained solutions for DMEA-II (Left) and MOEA/D(Right) on ZDT1.

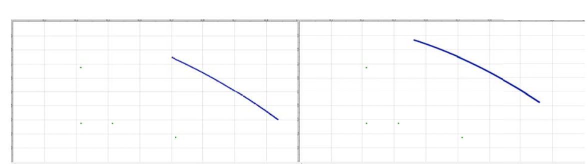

Fig. 5. The obtained solutions for DMEA-II (Left) and MOEA/D(Right) on ZDT6.

The metric values for DMEA-II and MOEA/D on 10 test problems are reported on Table I.

Problems		DMEA-II	MOEA/D		DMEA-II	MOEA/D
ZDT1	iGD	1.0	0.1	iIGD	0.05	1.0
ZDT2		0.04	1.0		0.02	1.0
ZDT3		1.0	0.06		1.0	0.9
ZDT4		0.6	1.0		1.0	0.2
ZDT6		1.0	0.7		0.5	1.0
UF1		0.58	1.0		0.87	1.0
UF2		0.35	1.0		1.0	0.65
UF3		1.0	0.88		0.32	1.0
UF4		0.7	1.0		0.09	1.0
UF5		0.52	1.0		1.0	0.78

TABLE I: The metric values for DMEA-II and MOEA/D.

Through the experiments, based on iGD and iIGD values we found that, for ZDT1: MOEA/D strongly converged to the user's preferred region (iGD value of 0.1), but DMEA-II was better in keeping the balance between convergence and diversity (iIGD value of 0.05). For ZDT2: DMEA-II was better than MOEA/D when it got 0.04 on iGD and 0.02 on iIGD. For ZDT3, MOEA/D was better than DMEA-II when it got 0.06 on iGD and 0.9 on iIGD. However, for ZDT4, DMEA-II converged to the user's preferred region when it got 0.6 on iGD, but MOEA/D was better than DMEA-II in keeping the balance between convergence and diversity when it got 0.2 on iIGD. For ZDT6, MOEA/D converged to the user's preferred region, but DMEA-II was better than MOEA/D in the balance between convergence and diversity when it got 0.5 on iIGD.

For UF1, DMEA-II was better than MOEA/D when it got 0.58 on iGD and 0.87 on iIGD. For UF2, DMEA-II was better than MOEA/D on iGD but MOEA/D was better on iIGD when it got 0.65. For UF3, MOEA/D was better than DMEA-II on iGD when it got 0.88 but it worse than DMEA-II on iIGD when DMEA-II got 0.32 on iIGD. For UF4, DMEA-II was better than MOEA/D when it got 0.7 on iGD and 0.09 on iIGD. Finally, DMEA-II was better than MOEA/D on iGD but it worse than MOEA/D on iIGD when it got 0.52 on iGD and 1.0 on iIGD.

Overall, the metric results indicated that, for convergence DMEA-II is better than MOEA/D on ZDT3, ZDT6, UF1, UF4. For keeping the balance of convergence and diversity, it is also better than MOEA/D on ZDT1, ZDT2, ZDT6, UF1, UF3, UF6. While MOEA/D works better than DMEA-II on ZDT3, ZDT4 in keeping the balance of convergence and diversity.

V. Conclusions

In this paper, we have suggested to use a new performance metric for interactive multi-objective evolutionary algorithms which use multiple reference points. In our proposal, a User based Front (UbF) is defined based on the preferred region. A preferred region is determined on a grid which is generated from the set of reference points. For convergence and diversity criteria, we use UbF as Pareto Optimal Front (POF) in two popular metrics: GD and IGD. We called the modified metrics for interactive MOEAs are iGD and iIGD. To simulate the effect of proposed metrics, we used these metrics for two interactive MOEAs: DMEA-II and MOEA/D on ZDT and UF benchmark sets.
Through experimental results, it indicated that the proposed metric is useful to compare the effectiveness and efficiency of interactive MOEAs on user's preference which are given during the search.

Acknowledgments

This research is supported by Software Engineering Laboratory (SELab), Faculty of Information Technology, Le Quy Don Technical University.

References

[1] K. Deb, *Multi-objective optimization using evolutionary algorithms*. New York: John Wiley & Sons, 2001. 1

[2] L. Thiele, K. Miettinen, P. J. Korhonen, and J. Molina, "A preference based evolutionary algorithm for multi-objective optimization," pp. 411–436, 2009. 1

[3] K. Deb and A. Kumar, "Interactive evolutionary multi-objective optimization and decision-making using reference direction method," in *GECCO '07*, 2007, pp. 781–788. 1

[4] K. Deb, A. Sinha, P. J. Korhonen, and J. Wallenius., "An interactive evolutionary multi-objective optimization method based on progressively approximated value functions," 2010. 1

[5] K. Deb and J. Sundar, "Reference point based multi-objective optimization using evolutionary algorithms," in *GECCO '06: Proceedings of the 8th annual conference on Genetic and Evolutionary Computation*. New York, NY, USA: ACM Press, 2006, pp. 635–642. 1

[6] M. Gong, F. Liu, W. Zhang, L. Jiao, and Q. Zhang, "Interactive moea/d for multi-objective decision making," in *GECCO' 2011*, 2011, pp. 721–728. 1

[7] E. B. J., D. K., M. K., S. R., and Berlin, "Consideration of partial user preferences in evolutionary multi-objective optimization. multi-objective optimization: interactive and evolutionary approaches," *OR Spectrum*, 2008. 1

[8] E. Petri and M. Kaisa, "Trade-off analysis approach for interactive nonlinear multiobjective optimization," in *OR Spectrum*, 2011, pp. 1–14. 1

[9] Q. F. Zhang and H. Li., "Moea/d: A multi-objective evolutionary algorithm based on decomposition," 2007. 1

[10] L. Nguyen, L. T. Bui, and H. A. Abbass, "Dmea-ii: the direction-based multi-objective evolutionary algorithm-ii," *Soft Computing*, vol. 18, no. 11, pp. 2119–2134, 2014. 1

[11] U. K. Wickramasinghe, R. Carrese, and X. Li, "Designing airfoils using a reference point based evolutionary many-objective particle swarm optimization algorithm." in *IEEE Congress on Evolutionary Computation*, 2010, pp. 1–8. 1, 3

[12] A. Mohammadi, M. Omidvar, and X. Li, "A new performance metric for user-preference based multi-objective evolutionary algorithms," in *Evolutionary Computation (CEC), 2013 IEEE Congress on*, June 2013, pp. 2825–2832. 1, 3

[13] W. A., "The use of reference objectives in multi-objective optimisation," *MCDMtheory and Application, Proceedings. No. 177 in Lecture notes in economics and mathematical systems*, pp. 468–486, 1980. 1

[14] D. Veldhuizen, "Multiobjective evolutionary algorithms: Classifications, analyses, and new innovation," Ph.D. dissertation, Department of Electrical Engineering and Computer Engineering, Airforce Institue of Technology, Ohio, 1999. 2

[15] E. Zitzler, L. Thiele, and K. Deb, "Comparision of multiobjective evolutionary algorithms: Emprical results," *Evol. Comp*, vol. 8, no. 1, pp. 173–195, 2000. 3

[16] Y. C. Minzhong Liu, Xiufen Zou and Z. Wu, "Performance assessment of dmoea-dd with cec 2009 moea competition test instances," *Proceeding CEC'09 Proceedings of the 11th conference on Congress on Evol. Comp*, pp. 2913–2918, 2009. 3