Dominating Global Best Selection for Multi-objective Particle Swarm Optimization

Heming Xu, Yinglin Wang, Xin Xu
Department of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
hemingqiubai@126.com, ylwang@sjtu.edu.cn, xuxin0336@sjtu.edu.cn

***Abstract*—For multi-objective particle swarm optimization, the selection of the global best becomes an interesting topic, because it balances convergence and diversity. But the global best selected by the existing strategies has a high probability of not dominating the particle. The flight towards the global best not dominating the particle is expected to cause some objectives to become worse, thus does not surely promote convergence. As the accumulation of the flights, the algorithm suffers from slow convergence. Therefore we propose the dominating strategy to accelerate the convergence by decreasing that probability. Experimental results show our strategy outperforms other strategies.**

Keywords-multi-objective optimization problem; particle swarm optimization; global best; pareto dominance; archive

I. INTRODUCTION

Particle swarm optimization (PSO) [1] is a relatively recent heuristic inspired by the choreography of a bird flock, a kind of swarm intelligence [2] developed in research on artificial life [3]. For its simple concept, easy implementation and fast convergence, it is now widely used to solve multi-objective optimization problems (MOPs).

For multi-objective particle swarm optimization (MOPSO), the selection of the global best becomes an interesting topic, because it balances convergence and diversity. Convergence requires solutions close to the Pareto front, whereas diversity requires uniform solutions along the Pareto front. Though great varieties of global best selection strategies have been proposed, the selected global best still has a high probability of not dominating the particle. The flight towards the global best not dominating the particle is expected to cause some objectives to become worse, thus does not surely promote convergence. As the accumulation of the flights, the algorithm suffers from slow convergence. Therefore we propose the dominating strategy to accelerate the convergence by decreasing that probability. We adopt the classical MOPSO – Coello's MOPSO [4] as the basic algorithm, and apply different global best selection strategies. Comparison on the WFG series [5] shows that our strategy provides the best performance.

The reminder of this paper is organized as follows. Section II defines basic concepts. Section III briefly reviews related work. Section IV elaborates the dominating global best selection. Section V presents experimental results. Section VI draws conclusions and discusses future research directions.

II. BASIC CONCEPTS

Definition 1 Multi-objective Optimization Problem

A MOP attempts to find the decision vector $\mathbf{x}$ in the domain Ω that will optimize the objective vector $\mathbf{f}(\mathbf{x})$. Without loss of generality, it can be described as minimizing the value for every objective.

$$\min \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_k(\mathbf{x})), \mathbf{x} \in \Omega \quad (1)$$

A decision vector $\mathbf{x}$ consists of n decision variables. An objective vector $\mathbf{f}(\mathbf{x})$ consists of k objectives. The objective function $\mathbf{f}$ maps the decision space into the objective space.

Definition 2 Pareto Dominance

$$\begin{aligned} &\forall i \in \{1, 2, ..., k\}, f_i(\mathbf{x}) \le f_i(\mathbf{y}), \wedge \exists i, \\ &f_i(\mathbf{x}) < f_i(\mathbf{y}) \Rightarrow \mathbf{f}(\mathbf{x}) \prec \mathbf{f}(\mathbf{y}) \end{aligned} \quad (2)$$

Given two decision vectors $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}$ dominates $\mathbf{y}$, or $\mathbf{y}$ is dominated by $\mathbf{x}$, iff $\mathbf{f}(\mathbf{x})$ is not larger than $\mathbf{f}(\mathbf{y})$ for any objective and is less for at least one objective; $\mathbf{x}$ is equivalent to $\mathbf{y}$, iff $\mathbf{f}(\mathbf{x})$ and $\mathbf{f}(\mathbf{y})$ are identical in all the objectives; $\mathbf{x}$ and $\mathbf{y}$ are incomparable, iff they are not equivalent, and neither dominates the other.

Definition 3 Pareto Optimal Set

$$\rho^* = \{\mathbf{x} \in \Omega \mid \neg \exists \mathbf{y} \in \Omega, \mathbf{f}(\mathbf{y}) \prec \mathbf{f}(\mathbf{x})\} \quad (3)$$

A decision vector $\mathbf{x}$ is Pareto optimal, iff it is not dominated by any decision vector in the decision space. The Pareto optimal set is the set of all the Pareto optimal decision vectors.

Definition 4 Pareto Front

$$\rho f^* = \{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in \rho^*\} \quad (4)$$

The Pareto front is the image set of the Pareto optimal set in the objective space.

III. RELATED WORK

As the only information sharing mechanism proposed by PSO, the selection of the global best is vital to the performance of MOPSO. If the particles are separated into several swarms [6, 7], the global best is selected inside the swarm the particle residing in. If the particles are not fully connected [8-10], the global best is selected among the particle's neighbors. In this paper, we exclude these situations, and only discuss how to choose a global best for a particle when given a swarm of particles that are fully connected.

Common used strategies are summarized and analyzed as the following. The random strategy randomly chooses an archive member, but the sustained change of the global best which can also be far from the particle may cause erratic oscillations and lead to the chaotic search behavior. The sigma strategy [11] selects the archive member with the most similar sigma value, but the computational cost greatly increased when handling problems with high objective dimensions. The nearest strategy chooses the archive member that is the nearest to the particle in the objective space, but the particle has a great chance to select the same global best for many iterations, thus may fall into local optima. The grid strategy [4] uses the roulette wheel to select the hypercube with more archive members and then randomly choose one member from the hypercube, but needs the maintenance of the adaptive grid. The dominated tree strategy [12] constructs composite points to reorganize the archive to reduce the computational cost, thus allows a larger archive. But it does not suit the common used asynchrony model in which the archive is updated after each update of the particles, because it needs to frozen the archive to generate composite points. The stripe strategy [13] assumes the shape of the Pareto front for a bi-objective problem is similar to a line connecting two extreme solutions, so the nearest stripe center is chosen, but it can only deal with bi-objective problems, does not suit the asynchrony model, and the performance greatly deteriorates when the hypothesis is wrong.

Other strategies are also summarized. The average strategy [14] uses the average of the complete set of the global bests on each objective, but such formed global best may not be good and the same global best for all the particles apparently violates the demand for diversity. The uniform strategy [15] assigns each particle a serial number, sorts the archive and assigns each archive member a serial number, and then uniform distributes each particle by its order to an archive member, in order to promote uniformity, but does not suit the asynchrony model. The comprehensive learning strategy [16] selects different global bests for different decisional dimensions of a particle, and each global best may be a random archive member, its personal best, or the combination of the personal bests of some other particles. The artificial strategy [17] fractionally creates an artificial global best by collecting the best components. The two-lbests strategy [18] selects the archive member close to the personal best which is also an archive member to prevent the chaotic search behavior. The binary tournament selection [19] uses a binary tournament based on the crowding value of the non-dominated archive members to choose the global best. The crowding strategy [20] selects the one having the smaller crowding distance to its nearest neighbor in the archive between two random archive members.

Now we have analyzed the existing strategies, but little attention has been paid to the issue that the high probability of the selected global best not dominating the particle decreases the convergence, and that is what we concern.

IV. DOMINATING GLOBAL BEST SELECTION

Given a particle, a global best provides mixed information of convergence and diversity. Convergence guides the particle towards the Pareto front, whereas diversity guides the particle along the Pareto front. Only when the global best dominates the particle, the flight towards such a global best has the expectation of becoming not worse on all the objectives and better on some objectives, thus convergence is surely expected.

Our proposed dominating strategy chooses the archive member dominating the particle if possible. Therefore given a particle, the selection is as follows. We randomly select an archive member and record it as the start member, and then begin to traverse the archive members. If an archive member dominating the particle is found, we terminate the traverse and select it as the particle's global best. If we do not find any qualified member during a traverse, we simply select the start member.

Three reasons support the efficiency of the strategy. Firstly, if archive members dominating the particle exist, only such members can be selected, thus convergence is promoted. Secondly, such members are selected randomly, so the global best for the particle always changes, thus the randomness prevents premature convergence. Thirdly, if no such member exists, the archive member is selected randomly, thus diversity is also promoted.

TABLE I. COMPARISON BETWEEN STRATEGIES OF GLOBAL BEST SELECTION

Name	Time Complexity	Optimality	Determinism	Asynchrony
Dominating	$O(kAN)$	Good	Weak	Yes
Grid	$O(N)$ *	Middle	Weak	Yes
Random	$O(N)$	Middle	Weak	Yes
Sigma	$O\left(k(k-1)AN/2\right)$	Middle	Middle	Yes
Nearest	$O(kAN)$	Middle	Strong	Yes
Dominated Tree	$O(A^2+AN)$	Middle	Middle	No
Stripe	$O(N), k=2$	Bad	Strong	No

We compare the dominating strategy with other common strategies in Table I. Time complexity describes the computation cost of the strategy. k is the objective dimension; A is the archive size; N is the particle number. "*" means the grid strategy needs to maintain the grid, and such time complexity is not included. Though the time complexity of the dominating strategy is not the smallest, such cost is acceptable, especially when handling problems with expensive evaluations. Optimality depicts the optimality of the global best provided by the strategy. The better it is, the faster the convergence. The global best from the archive that dominates the particle is good; the global best from the archive is middle; the global best even not selected from the archive is bad because it is usually dominated. Determinism indicates the probability of changing the global best in the following iterations. The weaker it is, the lower the risk of premature convergence. Random based strategies that always change the global best provide weak determinism, such as the dominating strategy, the grid strategy, and the random strategy; position based strategies that sometimes change the global best provide middle determinism, such as the sigma strategy and the dominated tree strategy; position based strategies that seldom change the global best provide strong determinism, such as the nearest strategy and the stripe strategy. Asynchrony demonstrates whether the strategy suits the asynchrony model, thus the dominated tree strategy and the stripe strategy are not recommended. To conclude, the dominating strategy obviously performs the best, because it suits the asynchrony model, and provides both good optimality and weak determinism with acceptable time complexity.

V. TEST FUNCTIONS AND PERFORMANCE ANALYSIS

We select all the WFG [5] test functions because this series provides the most abundant characteristics such as non-separable, deception, and degeneration [21]. The series uses $n = 24, k = 3$, with 4 position variables and 20 distance variables.

We propose the dominating ratio to measure the probability of the global best dominating the particle, which reflects the optimality of the strategy. We also adopt the hyper-volume [22, 23] metric to measure both convergence and diversity and provide an overall measure of the result. It is defined as the volume dominated by the archive members, but not dominated by the reference point – [10,10,10]. The larger the value, the better the performance is. We use box and whisker plots for visualization of the statistics of these metrics.

We use Coello's MOPSO [4] as the basic algorithm, and generate derived algorithms by applying the dominating strategy and some common global best selection strategies that suit the asynchrony model. The parameters are set as follows. The population size is 100; the maximum number of iterations is 250, i.e., 25000 evaluations are processed; the mutation rate is 0.5; the archive size is 300; the adaptive grid uses 10 divisions for each objective. Every algorithm runs 30 independent times on every test function.

Comparison of the obtained results is shown in Fig. 1. The label 'A' represents the dominating strategy; the label 'B' represents the grid strategy; the label 'C' represents the random strategy; the label 'D' represents the sigma strategy; the label 'E' represents the nearest strategy. For dominating ratio, the dominating strategy always performs the best, the nearest strategy the second, the sigma strategy the third, and the grid strategy and the random strategy the last, the same with the optimality sequence. For hyper-volume, the dominating strategy always performs the best, and only for WFG9, the difference is slight. Other strategies perform differently on different functions. Therefore the dominating strategy outperforms other strategies, no matter the function is separable or not, multi-modal or not, deceptive or not, and non-uniform or not.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have analyzed existing strategies related to the global best selection, found that the high probability of the global best not dominating the particle decreases the convergence, and suggested the dominating strategy that randomly selects an archive member dominating the particle if possible. Our strategy provides both good optimality and weak determinism. Experimental results demonstrate that our strategy outperforms other strategies.

As proposed future work, the current dominating strategy can be seen as mixed with the random strategy, so the future dominating strategy can be mixed with the grid strategy to increase the uniformity of the obtained solutions.

ACKNOWLEDGMENT

This paper was supported by the Science and Technology Innovation Action Plan (Grant Number:12511502902) of Shanghai Science and Technology Committee.

REFERENCES

[1] Kennedy, J., and Eberhart, R., 1995, "Particle swarm optimization," IEEE, pp. 1942-1948 vol. 1944.

[2] Bonabeau, E., Dorigo, M., and Theraulaz, G., 1999, Swarm intelligence: from natural to artificial systems, Oxford University Press, USA.

[3] Adami, C., 1998, Introduction to artificial life, Telos Pr.

[4] Coello, C. A. C., Pulido, G. T., and Lechuga, M. S., 2004, "Handling multiple objectives with particle swarm optimization," Evolutionary Computation, IEEE Transactions on, 8(3), pp. 256-279.

[5] Huband, S., Barone, L., While, L., and Hingston, P., 2005, "A scalable multi-objective test problem toolkit," Springer, pp. 280-295.

[6] Blackwell, T., and Branke, J., 2006, "Multiswarms, exclusion, and anti-convergence in dynamic environments," Evolutionary Computation, IEEE Transactions on, 10(4), pp. 459-472.

[7] Yen, G. G., and Leong, W. F., 2009, "Dynamic multiple swarms in multiobjective particle swarm optimization," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 39(4), pp. 890-911.

[8] Kennedy, J., 1999, "Small worlds and mega-minds: effects of neighborhood topology on particle swarm performance," IEEE.

[9] Kennedy, J., and Mendes, R., 2002, "Population structure and particle swarm performance," IEEE, pp. 1671-1676.

[10] Hu, X., and Eberhart, R., 2002, "Multiobjective optimization using dynamic neighborhood particle swarm optimization," Ieee, pp. 1677-1681.

[11] Mostaghim, S., and Teich, J., 2003, "Strategies for finding good local guides in multi-objective particle swarm optimization (MOPSO)," IEEE, pp. 26-33.

[12] Fieldsend, J. E., Uk, E. Q., and Singh, S., 2002, "A Multi-Objective Algorithm based upon Particle Swarm Optimisation, an Efficient Data Structure and Turbulence."

[13] Villalobos-Arias, M. A., Pulido, G. T., and Coello, C. A. C., 2005, "A proposal to use stripes to maintain diversity in a multi-objective particle swarm optimizer," IEEE, pp. 22-29.

[14] Zhang, L. B., Zhou, C. G., Liu, X., Ma, Z., Ma, M., and Liang, Y., 2003, "Solving multi objective optimization problems using particle swarm optimization," Evolutionary Computation, 2003. CEC'03. The 2003 Congress on, IEEE, pp. 2400-2405.

[15] Hsieh, S. T., Sun, T. Y., Chiu, S. Y., Liu, C. C., and Lin, C. W., 2007, "Cluster based solution exploration strategy for multiobjective particle swarm optimization," ACTA Press, pp. 295-300.

[16] Huang, V., Suganthan, P., and Liang, J., 2006, "Comprehensive learning particle swarm optimizer for solving multiobjective optimization problems," International Journal of Intelligent Systems, 21(2), pp. 209-226.

[17] Kiranyaz, S., Ince, T., Yildirim, A., and Gabbouj, M., 2010, "Fractional particle swarm optimization in multidimensional search

space," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 40(2), pp. 298-319.

[18] Zhao, S. Z., and Suganthan, P., 2011, "Two-lbests based multi-objective particle swarm optimizer," Engineering Optimization, 43(1), pp. 1-17.

[19] Sierra, M., and Coello Coello, C., 2005, "Improving PSO-Based multi-objective optimization using crowding, mutation and ∈ -dominance," Evolutionary Multi-Criterion Optimization, Springer, pp. 505-519.

[20] Nebro, A., Durillo, J., Garcia-Nieto, J., Coello Coello, C., Luna, F., and Alba, E., 2009, "Smpso: A new pso-based metaheuristic for multi-objective optimization," IEEE, pp. 66-73.

[21] Huband, S., Hingston, P., Barone, L., and While, L., 2006, "A review of multiobjective test problems and a scalable test problem toolkit," Evolutionary Computation, IEEE Transactions on, 10(5), pp. 477-506.

[22] Zitzler, E., 1999, Evolutionary algorithms for multiobjective optimization: Methods and applications, Shaker.

[23] Zitzler, E., Brockhoff, D., and Thiele, L., 2007, "The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration," Springer, pp. 862-876.

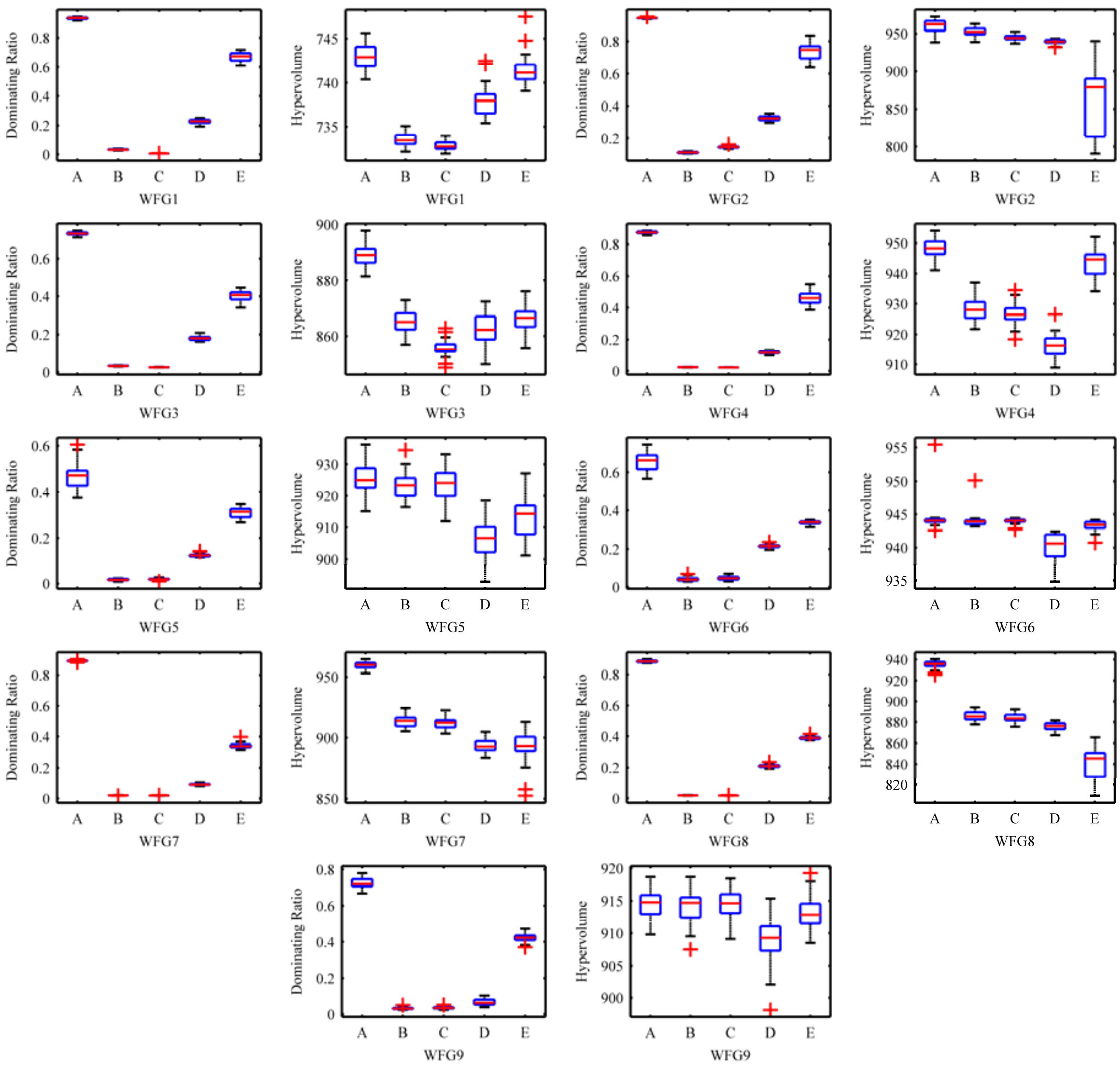

Figure 1. Comparison of the obtained results.