A study of similarity measure between tasks for multifactorial evolutionary algorithm

Lei Zhou, Liang Feng*
College of Computer Science, Chongqing Univ.
China
liangf@cqu.edu.cn

Jinghui Zhong[1], Zexuan Zhu[2]
College of Computer Science, South China Univ. of Technology[1]
College of Computer Science and Software Engineering, Shenzhen Univ.[2] China

Bingshui Da[3], Zhou Wu[4]
School of Computer Engineering, Nanyang Technological Univ.[3]
Singapore
School of Automation, Chongqing Univ.[4] China

ABSTRACT

In contrast to the traditional single-task evolutionary algorithms, multi-factorial evolutionary algorithm (MFEA) has been proposed recently to conduct evolutionary search on multiple tasks simultaneously. It aims to improve convergence characteristics of the tasks to be tackled by seamlessly transferring knowledge among them. Towards superior multitasking performance, the evaluation of task relationship plays an important role for grouping the related tasks, and solve them at the same time. However, in the literature, only a little work has been conducted to provide deeper insights in the measure of task relationship in MFEA. In this paper, we thus present a study of similarity measure between tasks for MFEA from three different perspectives. 21 multitasking problem sets are developed to investigate and analyze the effectiveness of the three similarity measures with MFEA for evolutionary multitasking.

CCS CONCEPTS

• **Theory of computation → Evolutionary algorithm**; • **Computing methodologies → Search methodologies**;

KEYWORDS

Evolutionary multitasking, multifactorial evolutionary algorithm, task relationship, similarity measure

ACM Reference Format:
Lei Zhou, Liang Feng, Jinghui Zhong[1], Zexuan Zhu[2], and Bingshui Da[3], Zhou Wu[4]. 2018. A study of similarity measure between tasks for multifactorial evolutionary algorithm. In *GECCO '18 Companion: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan.* ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3205651.3205736

1 INTRODUCTION

Recently, in contrast to traditional EAs, the multifactorial evolutionary algorithm (MFEA) has been proposed for evolutionary multitasking, which solves multiple optimization problems (tasks) concurrently [4]. MFEA is capable of exploiting the latent synergies between distinct (but possibly similar) optimization problems and has been demonstrated superior search performance over the traditional single-task EAs, in terms of solution quality and convergence speed, on many complex optimization tasks [2, 4, 6].

In spite of the success obtained by MFEA, it is worth noting here that the performance of MFEA is greatly affected by the similarity between tasks [1]. In the literature, there is only a little work having been proposed to study the task similarity for MFEA. In particular, in [3], the authors presented a synergy metric ξ to measure the correlation between the optimization functions. However, the calculation of ξ requires the information of global optimal solution and function gradients which is unpractical in real world applications. Further, in [1], the Spearman's rank correlation coefficient (SRCC) is employed as a measure for task similarity evaluation. However, as only a few multitasking sets have been investigated in [1], the generalization capability of SRCC for task similarity evaluation requires further investigation.

In this paper, we embark a study to measure the similarity between tasks for MFEA from three different perspectives, i.e., the distance between best solutions, the fitness rank correlation and the fitness landscape analysis. Further, new multitasking benchmarks, which contain 21 multitasking problem sets, are developed to investigate and analyze the effectiveness of the three similarity measures with MFEA.

2 MEASURING THE TASK SIMILARITY FROM DIFFERENT PERSPECTIVES

2.1 Similarity Measure based on distance between best solutions

An intuitive way of measuring the similarity between tasks for MFEA is based on the distance or discrepancy between the best solutions of different tasks. In this paper, we propose to approximate the distance between tasks via maximum mean discrepancy (MMD), which can be calculated as follows[1]:

$$MMD = \left\| \frac{1}{m} \sum_{i=1}^{m} g(\mathbf{Sb}_1^i) - \frac{1}{m} \sum_{i=1}^{m} g(\mathbf{Sb}_2^i) \right\|^2 \quad (1)$$

where $||\ .\ ||$ denotes the Euclidean norm, $g(s)$ extracts the first D_{min}=min(D_1, D_2) dimensions of a solution s. $\mathbf{Sb}_1$ and $\mathbf{Sb}_2$ contain the first m best solutions selected from two sets of solutions which are uniformly and independently sampled on task T_1 and T_2, respectively.

*Liang Feng is the corresponding author

[1]In this study, the linear mapping is considered in the calculation of MMD.

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

ACM ISBN 978-1-4503-5764-7/18/07.
https://doi.org/10.1145/3205651.3205736

Table 1: Results of the true performance obtained by MFEA over SEA, and the predicted performance of MFEA based on task similarity, on the 21 multitasking sets.

Multitasking Sets	True Performance	Predicted Performance			Multitasking Sets	True Performance	Predicted Performance		
		MMD	SRCC	CTFDC			MMD	SRCC	CTFDC
Griewank+Rastrigin	1	1	1	1	Griewank+Ackley	1	1	1	1
Griewank+Schwefe	1	**-1**	**-1**	1	Griewank+Sphere	1	1	1	1
Griewank+Rosenbrock	1	1	1	1	Griewank+Weierstrass	1	1	1	1
Rastrigin+Ackley	1	1	1	1	Rastrigin+Schwefel	1	**-1**	**-1**	1
Rastrigin+Sphere	1	1	1	1	Rastrigin+Rosenbrock	1	1	1	1
Rastrigin+Weierstrass	1	1	1	1	Ackley+Schwefel	-1	-1	-1	-1
Ackley+Sphere	1	1	1	1	Ackley+Rosenbrock	1	1	1	1
Ackley+Weierstrass	-1	**1**	-1	-1	Schwefel+Sphere	1	**-1**	**-1**	1
Schwefel+Rosenbrock	1	**-1**	**-1**	1	Schwefel+Weierstrass	-1	-1	-1	-1
Sphere+Rosenbrock	1	1	1	1	Sphere+Weierstrass	1	1	1	1
Rosenbrock+Weierstrass	1	1	1	1	Error rate	–	23.8%	19.0%	0%

2.2 Similarity Measure based on fitness rank correlation

According to [1], the Spearman rank correlation coefficient (SRCC) is adopted to estimate the degree of similarity between two tasks based on the rationale that if the solutions which perform well on one task are also good on the other, the improvement on one task is probable to be beneficial to the other. For more details of the calculation of SRCC, readers are referred to [1].

2.3 Similarity Measure based on fitness landscape analysis

Fitness landscape analysis (FLA) provides a global vision to measure the similarity based on tasks' search spaces [5]. In this paper, the commonly used fitness distance correlation (FDC) is adopted as the measure for FLA. As the original FDC evaluates the fitness-distance correlation of a single task, we propose a variant of FDC, called cross-task FDC (CTFDC), for the similarity measure of tasks in evolutionary multitasking, which can be calculated as follows:

$$CTFDC_{T_1 \to T_2} = \frac{1/n \sum_{i=1}^{n} (\mathbf{f_1^i} - \bar{\mathbf{f_1}})(\mathbf{d_2^i} - \bar{\mathbf{d_2}})}{\sigma(\mathbf{f_1})\sigma(\mathbf{d_2})} \qquad (2)$$

where $\bar{\mathbf{f_1}}$ and $\bar{\mathbf{d_2}}$, $\sigma(\mathbf{f_1})$ and $\sigma(\mathbf{d_2})$ are the means and standard deviations of $\mathbf{f_1}$ and $\mathbf{d_2}$, respectively. $\mathbf{f_1}$ is the fitness vector obtained by evaluating a set of uniformly and independently sampled solutions S, of size n, on task T_1. $\mathbf{d_2}$ is the distance vector contains the distances between each solution $s_i \in$ S and the best solution for task T_2 in S.

3 RESULTS AND DISCUSSION

In Table. 1, the column "True Performance" gives the comparison obtained by MFEA over the single-task EA (SEA), while the column "Predicted Performance" gives the predicted performance of MFEA indicated by the corresponding similarity measure. In particular, "1" denotes MFEA obtained superior performance, while "-1" indicates the deteriorated performance has been achieved by MFEA.

From Table. 1, it can be observed that on all the 21 bechmark sets, the prediction error rates of MMD, SRCC and CTFDC are 23.8%, 19% and 0%, respectively. The results indicate that, in contrast to MMD and SRCC, the CTFDC is able to provide a more reasonable measure of similarity between tasks for evolutionary multitasking with MFEA.

4 CONCLUSION

In this paper, we have presented a study of the similarity measure between tasks for MFEA from three different views, i.e., the distance between best solutions, the fitness rank correlation and the fitness landscape analysis, which are approximated by the maximum mean discrepancy (MMD), Spearman rank correlation coefficient (SRCC), and a variant of fitness distance correlation (FDC) called cross-task FDC (CTFDC), respectively. The experimental results showed that, among the three views, the CTFDC-based measure is more appropriate in evaluating the similarity between tasks for evolutionary multitasking. For future works, we would like to further study the generality of the CTFDC-based similarity measure on more multi-task optimization problems.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61603064, Frontier Interdisciplinary Research Fund for the Central Universities under Grant 106112017CDJQJ188828, and Chongqing Application Foundation and Research in Cuttingedge Technologies under Grant No. cstc2017jcyjAX0319.

REFERENCES

[1] Bingshui Da, Yew-Soon Ong, Liang Feng, AK Qin, Abhishek Gupta, Zexuan Zhu, Chuan-Kang Ting, Ke Tang, and Xin Yao. 2017. Evolutionary multitasking for single-objective continuous optimization: Benchmark problems, performance metric, and baseline results. *arXiv preprint arXiv:1706.03470* (2017).

[2] Abhishek Gupta, Jacek Mańdziuk, and Yew-Soon Ong. 2015. Evolutionary multitasking in bi-level optimization. *Complex & Intelligent Systems* 1, 1-4 (2015), 83–95.

[3] Abhishek Gupta, Yew-Soon Ong, B Da, L Feng, and Stephanus Daniel Handoko. 2016. Landscape synergy in evolutionary multitasking. In *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 3076–3083.

[4] Abhishek Gupta, Yew Soon Ong, and Liang Feng. 2016. Multifactorial Evolution: Toward Evolutionary Multitasking. *IEEE Transactions on Evolutionary Computation* 20, 3 (2016), 343–357.

[5] Erik Pitzer and Michael Affenzeller. 2012. A comprehensive survey on fitness landscape analysis. In *Recent Advances in Intelligent Engineering Systems*. Springer, 161–191.

[6] Lei Zhou, Liang Feng, Jinghui Zhong, Yew-Soon Ong, Zexuan Zhu, and Edwin Sha. 2016. Evolutionary multitasking in combinatorial search spaces: A case study in capacitated vehicle routing problem. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 1–8.