An Analysis of Control Parameters of Copula-based EDA Algorithm with Model Migration

Martin Hyrš
Brno University of Technology, Faculty of Information Technology, IT4 Innovations Centre of Excellence
Brno, Czech Republic
ihyrs@fit.vutbr.cz

Josef Schwarz
Brno University of Technology, Faculty of Information Technology, IT4 Innovations Centre of Excellence
Brno, Czech Republic
schwarz@fit.vutbr.cz

ABSTRACT

The copula-based EDA algorithms nowadays represent a promising technique for problem optimization in the continuous domain. This paper provides a detailed analysis on how six key parameters of the parallel copula-based EDA with model migration (mCEDA) influence the quality of optimization. In order to improve the performance of that kind of algorithm the most suitable setting of these control parameters is evaluated on the well known CEC 2013 benchmark using inferential statistics.

CCS CONCEPTS

• **Mathematics of computing → Probabilistic algorithms**; *Multivariate statistics*; • **Computing methodologies → Continuous space search**; Parallel algorithms;

KEYWORDS

EDA; copulas; parallelization; model migration; optimization

ACM Reference Format:
Martin Hyrš and Josef Schwarz. 2019. An Analysis of Control Parameters of Copula-based EDA Algorithm with Model Migration. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion), July 13–17, 2019, Prague, Czech Republic.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3319619.3321910

1 INTRODUCTION

Estimation of distribution algorithms (EDAs) belong to a class of evolutionary optimization methods that explore the search space by estimating and sampling an explicit probabilistic model of promising solutions. A new approach to building an efficient probabilistic model that is based on copula theory was proposed in [8]. Copulas can be used to model correlations within multivariate problems in which the joint distribution is separated into the univariate marginal distributions and the correlation structure is expressed by the copula function. This approach was later extended by introducing a parallel implementation and a new concept of migration of the probability model parameters [4].

GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic

ACM ISBN 978-1-4503-6748-6/19/07…$15.00
https://doi.org/10.1145/3319619.3321910

In this paper, the impact of six control parameters (copula type, population size, selection rate, the number of islands, migration rate and topology) of the mCEDA algorithm is analyzed.

2 EXPERIMENTAL WORKS

The mCEDA algorithm (with the initial setting: Gaussian copula, 500 individuals, selection rate 20 %, 10 islands, migration rate 20 generations and bi-directional ring topology) [4] is chosen as a reference algorithm (RA) in the systematical tuning of the control parameters. Tested variants of mCEDA algorithm with a particular setup of control parameters are evaluated using CEC 2013 benchmarks [5] for 10 dimension. As suggested by the authors of the benchmarks, the 51 independent runs on each function are performed, with maximally 100000 fitness evaluations.

To achieve the well performed variant of RA for the particular control parameters we used the two-sided Wilcoxon rank-sum test as a proper statistical test. It is directed to compare the performance of two algorithms – in our case RA and its algorithmic variant for each of 28 functions of CEC 2013 benchmarks. The result of comparison is expressed in the form: statistically better, worse or indifferent (on the significant level $\alpha = 0.05$).

Fig. 1–6 contain stacked histograms for all experiments devoted to parameters tuning: green (top stack) means that RA was statistically better (B); red (bottom stack) means that RA was worse (W); yellow (middle stack) means that RA is statistically indifferent (I) according to the tested variant of RA. The values associated with the color stacks give the number of benchmark functions for which a given statistical outcome holds. The coefficient $PB = \frac{B}{B+W}$ expresses the relative value of B. The minimal value of PB determines the best tested variant.

3 RESULTS

Type of Copula:

The candidate set of copulas contains six types: Gauss, Student, Gumbel, Clayton, Frank and product copula. From Fig. 1 it is evident that the **Gaussian copula** is significantly better than the others.

Population Size:

The tested size lies in the range from 125 to 2000 individuals. The most suitable value is **250 individuals**, see Fig. 2.

Selection Rate:

The tested selection rate lies in the range from 5 % to 60 %. Small rates are significantly better – the best value is **20 %**, see Fig. 3.

Number of Islands:

The entire population population consists of several subpopulations, each of them is assigned to one island. Variants with a small

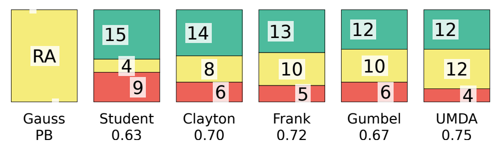

Figure 1: Test results for different copula types.

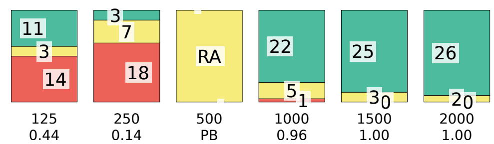

Figure 2: Test results for different population sizes.

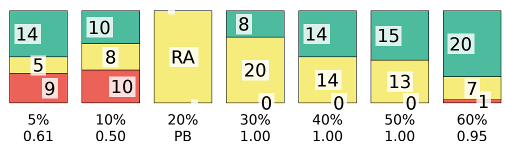

Figure 3: Test results for different selections rates.

number of islands perform better (see Fig. 4), **6 islands** is the most suitable setting.

Migration Rate:

The migration rate expresses the communication intensity among the associated islands. The promising value is in the range from 2 to 10 generations elapsed between consequent migrations; **10 generations** led to the best results, see Fig. 5.

Topology:

The role of used topology is less significant. We report the best results for the **bi-directional ring**, see Fig. 6. The other topologies (uni-directional ring, random, star = island containing the best individual is selected as the center, zero = no migration used) provide similar results.

Comparison with other published algorithms:

In Fig. 7 the comparison of final version mCEDAf with the initial mCEDA and with other published algorithms – namely IPOP-CMA-ES [7], CMAES-RIS [1], PSO [9], LaF [2], SPAM-AOS [3] is reported.[1] The pairwise t-test is used instead of Wilcoxon ranked-sum test.[2]

4 CONCLUSION

As the most important parameters influencing the mCEDAf performance we identified: the type of copula, the number of islands, the population size and the selection rate. The role of migration rate and topology are less important. The Gaussian copula was confirmed as the best variant. The most suitable parameters are: 250 individuals in the population, selection rate 20 % and 6 islands.

The final version mCEDAf utilizing the best-identified set of parameters outperforms the initial version of mCEDA, and it is comparable with PSO, LaF and SPAM-AOS algorithms. The IPOP-CMA-ES and CMA-ES-RIS outperform the mCEDAf algorithm.

In the future, the advanced adaptation of the significant control parameters will be investigated (see also [6]).

[1]The stacks are of the same meaning as in Fig. 1–6. – e.g. green stack gives the number of benchmarks for which mCEDAf was significantly better than the other algorithm.
[2]From the other published algorithms, only mean, median and standard deviation values of the tested functions are available.

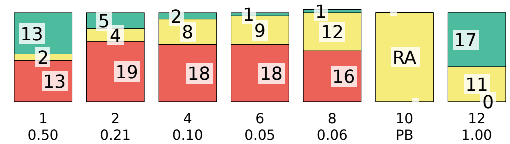

Figure 4: Test results for different number of islands.

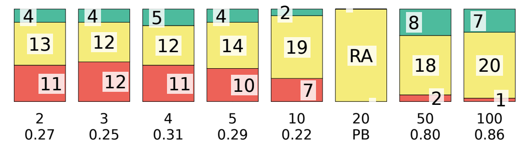

Figure 5: Test results for different migration rate.

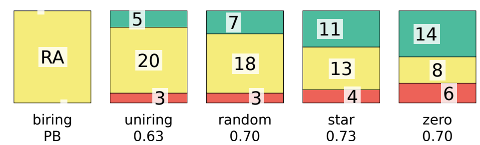

Figure 6: Test results for different topologies.

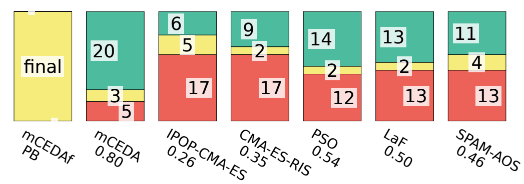

Figure 7: Comparison of mCEDAf with other published algorithms [1–4, 7, 9] using the pairwise t-test.

REFERENCES

[1] F. Caraffini, G. Iacca, F. Neri, L. Picinali, and E. Mininno. 2013. A CMA-ES super-fit scheme for the re-sampled inheritance search. In *2013 IEEE Congress on Evolutionary Computation*. 1123–1130. https://doi.org/10.1109/CEC.2013.6557692

[2] Stephen Chen and Yasser Gonzalez-Fernandez. 2015. *Leaders and Followers on the CEC2013 Real-Parameter Optimization Benchmark Functionss.* Technical Report. https://doi.org/10.13140/RG.2.1.3366.1927 Technical Report School of Information Technology York University.

[3] M. G. Epitropakis, F. Caraffini, F. Neri, and E. K. Burke. 2014. A Separability Prototype for Automatic Memes with Adaptive Operator Selection. In *Foundations of Computational Intelligence (FOCI), 2014 IEEE Symposium on.* 70–77. https://doi.org/10.1109/FOCI.2014.7007809

[4] Martin Hyrš and Josef Schwarz. 2015. Elliptical and Archimedean Copulas in Estimation of Distribution Algorithm with Model Migration. In *Proceedings of the 7th International Joint Conference on Computational Intelligence (IJCCI 2015).* SciTePress - Science and Technology Publications, 212–219.

[5] JJ Liang, BY Qu, PN Suganthan, and Alfredo G Hernández-Díaz. 2013. Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization. *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore, Technical Report* 201212 (2013).

[6] Manuel López-Ibáñez, Jérémie Dubois-Lacoste, Leslie Pérez Cáceres, Thomas Stützle, and Mauro Birattari. 2016. The irace package: Iterated Racing for Automatic Algorithm Configuration. *Operations Research Perspectives* 3 (2016), 43–58. https://doi.org/10.1016/j.orp.2016.09.002

[7] I. Loshchilov. 2013. CMA-ES with restarts for solving CEC 2013 benchmark problems. In *2013 IEEE Congress on Evolutionary Computation*. 369–376. https://doi.org/10.1109/CEC.2013.6557593

[8] Marta Soto, Yasser González-Fernández, and Alberto Ochoa. 2012. Modeling with Copulas and Vines in Estimation of Distribution Algorithms. *CoRR* abs/1210.5500 (2012).

[9] M. Zambrano-Bigiarini, M. Clerc, and R. Rojas. 2013. Standard Particle Swarm Optimisation 2011 at CEC-2013: A baseline for future PSO improvements. In *2013 IEEE Congress on Evolutionary Computation*. 2337–2344. https://doi.org/10.1109/CEC.2013.6557848

Acknowledgements. This work was supported by the Brno University of Technology project FIT-S-17-3994.