A Memetic NSGA-II with EDA-Based Local Search for Fully Automated Multiobjective Web Service Composition

Chen Wang, Hui Ma, Gang Chen
Victoria University of Wellington
Wellington, New Zealand
chen.wang,hui.ma,aaron.chen@ecs.vuw.ac.nz

Sven Hartmann
Clausthal University of Technology
Clausthal-Zellerfeld, Germany
sven.hartmann@tu-clausthal.de

ABSTRACT

Web service composition aims to provide added values by loosely coupling web services to accommodate users' complex requirements. Evolutionary computation techniques have been used to efficiently find near-optimal composite services to satisfy users' requirements reasonably well. Often, the quality of a composite service is measured by two important quality criteria that are related to the non-functional quality (i.e., Quality of service, QoS for short) and function quality (i.e., Quality of semantic matchmaking, QoSM for short). One recent work [2] proposed a Hybrid method that combines NSGA-II and MOEA/D with swap-based local search to enhance the performance of NSGA-II. This Hybrid method handles two quality criteria in QoS as two trade-off objectives. However, the local search of this method is randomly applied to a predefined large number of subproblems without focusing on the most suitable candidate solutions. In this paper, we propose a memetic NSGA-II with EDA-based local search. Particular, EDA performs the local improvements of a few well-selected composite services in different regions of the Pareto front. We also aim to handle two practical trade-off objectives with respect to QoS and QoSM. Our experiments have shown that our proposed method outperforms the recent state-of-the-art algorithms and the baseline NSGA-II method with respect to effectiveness and efficiency.

CCS CONCEPTS

• **Mathematics of computing** → **Combinatorial optimization**; • **Information systems** → *Web services*; • **Theory of computation** → Evolutionary algorithms;

KEYWORDS

Web service composition, QoS, Combinatorial optimisation, EDA

ACM Reference Format:
Chen Wang, Hui Ma, Gang Chen and Sven Hartmann. 2019. A Memetic NSGA-II with EDA-Based Local Search for Fully Automated Multiobjective Web Service Composition. 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/10.1145/3319619.3321937

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

ACM ISBN 978-1-4503-6748-6/19/07…$15.00
https://doi.org/10.1145/3319619.3321937

1 INTRODUCTION

Web services can be composed to provided added values to accommodate users' requirements. Often, for a given service request, there are many possible composite services, different in both execution workflow and component services, as well as the quality. Evolutionary computation (EC) techniques have been widely used to find composite services efficiently with near-optimal non-functional quality (i.e., Quality of service, QoS for short) and function quality (i.e., Quality of semantic matchmaking, QoSM for short). One recent work (i.e., Hybrid [2]) combines NSGA-II and MOEA/D, where a simple local search, named one-point "swap", is randomly applied to a large number of decomposed subproblems (i.e., 500 subproblems). Although this method is more effective than NSGA-II, its local search method suffers from high level of randomness and ineffective swap operators. To address these two limitations, new memetic approaches must be developed to avoid pre-determining a large number of subproblems in advance for local search and to perform effective local search by explicitly exploiting knowledge embedded in good candidate solutions in each generation. To fulfill this idea, a clustering technique is used to determine some well-selected composite services in different regions of the Pareto front for local search and EDA is used to sample solutions with local improvements on the selected composite services.

2 PROPOSED APPROACH

To enhance the performance of NSGA-II in both effectiveness and efficiency, a memetic NSGA-II is proposed with EDA-based local search (this method is referred to as **MNSGA2-EDA**) for fully automated and multi-objective service composition. We use MNSGA2-EDA to handle two objectives related to QoSM and QoS respectively.

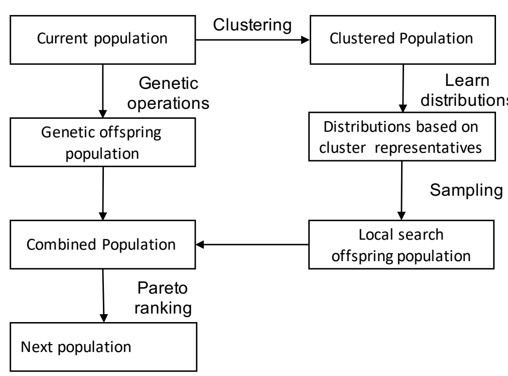

Figure 1: Generation updates in MNSGA2-EDA

Fig. 1 shows the generation updates in MNSGA2-EDA. These updates start with the current population, which is used to produce two offspring populations: genetic offspring population and local search offspring population. Particularly, the genetic operations (e.g., crossover and mutation) produce the genetic offspring population from the current population and the sampling procedure produces the local search offspring population from the distribution models. For each cluster, the model is learned based on its representative solution sampled from the current population. Later on, we combine these two produced offspring populations into one combined population, where the Pareto ranking technique is utilized to form the next generation from the combined population.

Representation. Let $\Pi = (\pi_0, \ldots, \pi_t, \ldots, \pi_{n-1})$ be a permutation-based composite solution of service indexes $\{0, \ldots, t, \ldots, n-1\}$ such that $\pi_i \neq \pi_j$ for all $i \neq j$. Usually, a decoding algorithm [4] is utilized to decode this permutation into a composite service represented in a form of a service workflow. Afterwards, this workflow is further encoded into another permutation Π' using Breadth-first search, see [4] for details.

Genetic operation. We employ two-point crossover and one one-point swap mutation [2] to produce the genetic offspring population, see [2] for the details.

Cluster representative and its distribution model. We cluster the current population into a set of clusters using K-means++ [1]. Afterwards, we identify one promising solution of each cluster by randomly selecting a solution from the non-dominant composite services in each cluster. Consequently, we learn a group of distributions based on each representative and its belonging cluster. Particularly, a *node histogram matrix* (NHM) for the cl^{th} cluster with the cluster representative Rep^{cl} in generation g is denoted as $\mathcal{NHM}^g_{cl}$, which is an $n \times n$-matrix with entries $e_{i,j}$ as follows:

$$e_{i,j} = \sum_{k=0}^{m-1} \delta_{i,j}(\Pi'^g_k) + \varepsilon \tag{1}$$

$$\delta_{i,j}(\Pi'^g_k) = \begin{cases} w(\Pi'^g_k) & \text{if } \pi_i = \text{j} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$w(\Pi'^g_k) = 1 - ||\vec{f}(\Pi'^g_k) - \vec{f}(Rep^g_{cl})||_2 \tag{3}$$

Eventually, we can use node histogram-based sampling [3] to sample new permutations to form a local search offspring population.

Fitness function. Two trade-off objectives that reflect QoSM and QoS are defined in Eq. (4) and Eq. (5) respectively are to be minimized, see [4] for calculations of each quality criteria.

$$f_1(\Pi) = w_1(1 - \hat{MT}) + w_2(1 - \hat{SIM}) \tag{4}$$

$$f_2(\Pi) = w_3(1 - \hat{A}) + w_4(1 - \hat{R}) + w_5\hat{T} + w_6\hat{CT} \tag{5}$$

3 EXPERIMENTS

We experimentally compare MNSGA2-EDA with NSGA-II and Hybrid [2] that are recently used to solve a similar problem. The benchmark dataset utilized for the comparison is obtained from [2] with doubled service repository size. The population size is set to 500 with a maximum generation of 51. Crossover, mutation, and reproduction rates are 0.8, 0.1 and 0.1. ε is calculated referring to [4].

Table 1: Comparison results of Mean IGD (Note: the lower the IGD the better)

Task	MNSGA2-EDA	NSGA-II	Hybrid [2]
WSC09-1	0.0701 ± 0.0132	$0.0731 \pm 6e-04$	0.0654 ± 0.0199
WSC09-2	0.0055 ± 0.001	0.0065 ± 0.0011	$0.0061 \pm 9e-04$
WSC09-3	$0.002 \pm 9e-04$	0.0126 ± 0.0085	0.0107 ± 0.0076
WSC09-4	0.0025 ± 0.001	$0.0061 \pm 7e-04$	0.0056 ± 0.0012
WSC09-5	0.0025 ± 0.0014	0.0052 ± 0.0011	$0.0045 \pm 7e-04$

Table 2: Comparison results of Mean Hypervolume (Note: the higher the Hypervolume the better)

Task	MNSGA2-EDA	NSGA-II	Hybrid [2]
WSC09-1	0.4435 ± 0.0028	$0.4424 \pm 9e-04$	0.4434 ± 0.0031
WSC09-2	$0.2751 \pm 1e-04$	0.2742 ± 0.0016	$0.2747 \pm 7e-04$
WSC09-3	$0.3693 \pm 1e-04$	0.361 ± 0.0064	0.3618 ± 0.0054
WSC09-4	0.239 ± 0.0014	$0.2346 \pm 9e-04$	0.2355 ± 0.0017
WSC09-5	0.2376 ± 0.001	$0.235 \pm 5e-04$	$0.2353 \pm 5e-04$

Table 3: Comparison results of Mean execution time (in s) (Note: the shorter the time the better)

Task	MNSGA2-EDA	NSGA-II	Hybrid [2]
WSC09-1	198 ± 67	155 ± 76	327 ± 90
WSC09-2	5634 ± 679	6139 ± 1678	14634 ± 2816
WSC09-3	2968 ± 301	2820 ± 714	6527 ± 2403
WSC09-4	269207 ± 23542	255195 ± 28813	646897 ± 117538
WSC09-5	39370 ± 5125	35338 ± 8350	86281 ± 19944

The weights in the fitness function Eq. (4) and Eq. (5) are set to balance quality criteria in both QoSM and QoS, i.e., w_1 and w_2 are set to 0.5, and w_3, w_4, w_5 and w_6 to 0.25. We run the experiment with 30 independent repetitions with the results shown in Tables 1, 2 and 3. We can observe that MNSGA2-EDA achieves significantly better values of IGD and hypervolume for all tasks. Meanwhile, the mean execution time for MNSGA2-EDA and NSGA-II are very comparable to each other. However, Hybrid consistently takes much more execution time.

4 CONCLUSION

In this paper, we proposed a novel NSGA2-EDA method for handling two trade-off objectives in fully automated service composition. Particularly, EDA is used to make local improvements of well-selected candidate composite services in the different regions of the Pareto front. This method outperforms one recent work and a baseline method with respect to both effectiveness and efficiency. Future works can investigate more effective sampling techniques to produce local search offspring by considering problem-specific templates for permutation-based solutions.

REFERENCES

[1] David Arthur and Sergei Vassilvitskii. 2007. K-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 1027–1035.

[2] Alexandre Sawczuk Da Silva, Hui Ma, Yi Mei, and Mengjie Zhang. 2018. A Hybrid Memetic Approach for Fully Automated Multi-Objective Web Service Composition. In *2018 IEEE International Conference on Web Services (ICWS)*. IEEE, 26–33.

[3] Shigeyoshi Tsutsui. 2006. A comparative study of sampling methods in node histogram models with probabilistic model-building genetic algorithms. In *Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on*, Vol. 4. IEEE, 3132–3137.

[4] Chen Wang, Hui Ma, Gang Chen, and Sven Hartmann. 2018. Knowledge-Driven Automated Web Service Composition — An EDA-Based Approach. In *International Conference on Web Information Systems Engineering*. Springer.