Comparing Random Forest and Gaussian Process Modeling in the Gp-demo Algorithm

Miha Mlakar, Tea Tušar, Bogdan Filipic
Department of Intelligent Systems, Jozef Stefan Institute and
Jozef Stefan International Postgraduate School
Jamova cesta 39, SI-1000 Ljubljana, Slovenia
{miha.mlakar, tea.tusar, bogdan.filipic}@ijs.si

ABSTRACT: *In surrogate-model-based optimization, the selection of an appropriate surrogate model is very important. If solution approximations returned by a surrogate model are accurate and with narrow confidence intervals, an algorithm using this surrogate model needs less exact solution evaluations to obtain results comparable to an algorithm without surrogate models. In this paper we compare two well known modeling techniques, random forest (RF) and Gaussian process (GP) modeling. The comparison includes the approximation accuracy and confidence in the approximations (expressed as the confidence interval width). The results show that GP outperforms RF and that it is more suitable for use in a surrogate-model-based multiobjective evolutionary algorithm.*

1. Introduction

One of the most effective ways to solve problems with multiple objectives is to use multiobjective evolutionary algorithms (MOEAs). The MOEAs draw inspiration from optimization processes occuring in nature and perform many solution evaluations to find high-quality solutions. Due to the high number of solution evaluations the MOEAs are not very suitable for computationally expensive optimization problems where exact solution evaluation takes a lot of time. In order to obtain the results of such problem more quickly, we usually use surrogate models to approximate the objective functions of the problem.

But due to inaccurate approximations, the solution comparisons can be incorrect, which can result in very good solutions being discarded. In order to minimize the impact of incorrect comparisons, we defined the relations under uncertainty ([5]) for comparing approximated solutions presented with an approximated value and a confidence interval. By including the confidence interval in the comparison we were able to consider this additional information and minimize the number of incorrect comparisons.

We used these relations under uncertainty in the algorithm called Differential Evolution for Multiobjective Optimization based

on Gaussian Process modeling (GP-DEMO) [4]. We discovered that the quality of the gained result depends greatly on the surrogate model. If the surrogate model is accurate, GP-DEMO finds high-quality results with a low number of exact solution evaluations, while if it is not, GPDEMO needs more exact solution evaluations to achieve similar results.

Since the accuracy of the surrogate model in surrogate model-based optimization is crucial, we decided to apply two different modeling techniques and compare their approximations to determine which one is more suitable for use in a surrogate-model-based algorithm. In addition to Gaussian process (GP) modeling that is used in GP-DEMO, we used random forest (RF) for comparison. The reason for choosing RF was the fact that the methodology is well-known and that the solutions approximated with this method in addition to approximated values return also confidence intervals.

The structure of this paper is as follows. In Section 2, we present how the comparison of RF and GP modeling techniqueswas carried out. In Section 3, we discus the results gained with both techniques, compare them and determine which technique performs better. Section 4 concludes the paper with an overview of the work done.

2. Comparison of Rf and Gp Surrogate Models

In this section we compare random forest and Gaussian process modeling techniques used for solution approximations. The aim of the comparison is to determine which of the two techniques is more suitable for use in surrogate-model-based optimization.

To test the two techniques, we used relations under uncertainty to compare their approximated solutions. If two solution approximations had overlapping confidence intervals, we, in order to determine their relation, exactly evaluated one solution and compared the solutions again. Together with the number of these additional exact evaluations, we measured also the number of incorrect solution comparisons and the width of the confidence intervals.

In addition to using relations under uncertainty, we also compared the approximated solutions with Pareto dominance relations and measured the number of incorrect comparisons. With Pareto dominance relations the confidence intervals are not included in the comparisons, so in general, the number of incorrect comparisons hints at the accuracy of the approximations.

Relation type	Solutions used for surrogate model	Number of comparisons	Incorrect comparisons	Number of comparisons with confidence interval reductions	Proportion of confidence interval reductions [%]	Confidence interval width
Relations under uncertainty	20 30 50 100 200	3,940,200	1;515 682 138 65 13	3,635,805 3,152,124 1,218,337 672,384 549,380	92 80 31 17 14	26.25 15.41 1.29 0.012 0.002
Pareto dominance relations	20 30 50 100 200	3,940,200	367,684 159,945 22,032 2,309 1,219	/ / / / /	/ / / / /	26.25 15.41 1.29 0.012 0.002

Table 1. Comparison of the relations under uncertainty and Pareto dominance relations for GP modeling on the Poloni problem

The solutions selected for testing were not generated randomly, but rather produced by the well-known NSGA-II algorithm [3]. This ensured that the solution comparisons were similar to the comparisons performed in evolutionary multiobjective algorithms and thus provided relevant results.

In every generation NSGA-II creates a new set of solutions, adds them to the current ones and then performs selection on the union to identify the most promising ones. The selection procedure includes comparing every solution with all other solutions to determine its dominance status. These were the comparisons used in our study.

Relation type	Solutions used for surrogate model	Number of comparisons	Incorrect comparisons	Number of comparisons with confidence interval reductions	Proportion of confidence interval reductions [%]	Confidence interval width
Relations under uncertainty	20	3,940,200	74,181	2,289,682	58	42.81
	30		21,861	1,934,212	49	25.98
	50		19,342	1,426,775	36	25.05
	100		144	712,298	18	0.07
	200		152	271,821	7	0.03
Pareto dominance relations	20	3,940,200	336,049	/	/	42.81
	30		136,357	/	/	25.98
	50		49,790	/	/	25.05
	100		1,736	/	/	0.07
	200		1,453	/	/	0.03

Table 2. Comparison of the relations under uncertainty and Pareto dominance relations for GP modeling on the OSY problem

Relation type	Solutions used for surrogate model	Number of comparisons	Incorrect comparisons	Number of comparisons with confidence interval reductions	Proportion of confidence interval reductions [%]	Confidence interval width
Relations under uncertainty	20	3,940,200	7,407	2,703,783	69	50.03
	30		16	2,338,535	59	0.074
	50		2	749,258	19	0.099
	100		3	359,952	9	0.022
	200		11	183,625	5	0.009
Pareto dominance relations	20	3,940,200	188,401	/	/	50.03
	30		161	/	/	0.074
	50		543	/	/	0.099
	100		645	/	/	0.022
	200		648	/	/	0.009

Table 3. Comparison of the relations under uncertainty and Pareto dominance relations for the GP modeling on the SRN problem

Relation type	Solutions used for surrogate model	Number of comparisons	Incorrect comparisons	Number of comparisons with confidence interval reductions	Proportion of confidence interval reductions [%]	Confidence interval width
Relations under uncertainty	20	3,940,200	22,497	3,906,474	99	32.89
	30		5,206	3,937,230	99	31.83
	50		2,180	3,935,723	99	28.42
	100		125	3,930,277	99	23.97
	200		4	3,909,386	99	19.76
	1,000		2	3,619,402	92	12.11
Pareto dominance relations	20	3,940,200	1,021,750	/	/	32.89
	30		965,491	/	/	31.83
	50		1,043,216	/	/	28.42
	100		894,889	/	/	23.97
	200		733,044	/	/	19.76
	1,000		379,928	/	/	12.11

Table 4. Comparison of the relations under uncertainty and Pareto dominance relations for RF modeling on the Poloni problem

Relation type	Solutions used for surrogate model	Number of comparisons	Incorrect comparisons	Number of comparisons with confidence interval reductions	Proportion of confidence interval reductions [%]	Confidence interval width
Relations under uncertainty	20	3,940,200	1	2,663,597	68	842.41
	30		0	2,663,597	68	789.04
	50		0	2,663,597	68	767.92
	100		0	2,663,597	68	720.44
	200		0	2,663,597	68	677.79
	1,000		0	2,663,597	68	548.19
Pareto dominance relations	20	3,940,200	885,416	/	/	842.41
	30		770,439	/	/	789.04
	50		810,251	/	/	767.92
	100		683,578	/	/	720.44
	200		661,919	/	/	677.79
	1,000		555,983	/	/	548.19

Table 5. Comparison of the relations under uncertainty and Pareto dominance relations for RF modeling on the OSY problem

Relation type	Solutions used for surrogate model	Number of comparisons	Incorrect comparisons	Number of comparisons with confidence interval reductions	Proportion of confidence interval reductions [%]	Confidence interval width
Relations under uncertainty	20	3,940,200	18	3,384,351	86	359.51
	30		0	3,385,285	86	350.55
	50		0	3,385,242	86	308.94
	100		0	3,384,910	86	266.55
	200		0	3,378,456	86	224.77
	1,000		0	3,133,626	79	139.89
Pareto dominance relations	20	3,940,200	387,854	/	/	359.51
	30		425,691	/	/	350.55
	50		365,606	/	/	308.94
	100		288,611	/	/	266.55
	200		216,634	/	/	224.77
	1,000		136,656	/	/	139.89

Table 6. Comparison of the relations under uncertainty and Pareto dominance relations for the RF modeling on the SRN problem

The experiments were performed on three benchmark multiobjective optimization problems. One is the Poloni optimization problem [6] and two are from [2], called OSY and SRN. All of them are two-objective problems.

For testing purposes we used GP modeling as proposed by [7] and RF modeling as proposed in [1]. For the confidence interval width of the approximation we used two standard deviations (2), which corresponds to about 95% of the normal distribution of the approximations. The number of trees used for building RF was 10,000 and the minimum number of elements in the leaves was set to 1.

To test the correlation between the surrogate model accuracy and the incorrect comparisons, different models of increasing accuracy were built—each on larger number of solutions. Since building an RF surrogate model is faster than building a GP surrogate model, we, in addition to building surrogate models from 20, 30, 50, 100 and 200 exactly evaluated solutions, also built an RF surrogate model from 1000 exactly evaluated solutions. We tested how much the larger RF surrogate model built from 1000 exactly evaluated solutions increases the accuracy of the approximations.

The NSGA-II parameter values used in the experiments were the same for both modeling techniques and for all three problems.

They were set as follows:

• Population size: 100,

• Number of generations: 100,

• Number of runs: 30.

The results averaged over 30 runs are presented in Tables 1–3 (for GP modeling) and in Tables 4–6 (for RF modeling).

3. Discussion

The results gained with both modeling techniques show that, irrespectively of the accuracy of a surrogate model, using relations under uncertainty reduces the number of incorrect comparisons.

The comparison of the results gained with RF and GP reveals certain differences between the techniques. The main difference is in the width of the confidence intervals. RF surrogate models produce wider confidence intervals. Consequently, the number of comparisons with confidence interval reductions for RF is much higher than for GP.

In addition to yielding wider confidence intervals, the RF surrogate models are also less accurate. Comparing the number of incorrect comparisons performed with Pareto dominance relations where the confidence intervals are not considered, we can see that the number of incorrect comparisons is higher with the RF surrogate models.

Another difference is in the correlation between the number of solutions used for building the surrogate model and the accuracy of the surrogate model. By increasing the number of solutions used, the RF surrogate models do not improve as quickly as the GP models. Even in the cases where 1000 exactly evaluated solutions were used for building the RF surrogate models the confidence interval widths were not greatly reduced and the intervals were still much wider than the confidence intervals gained with GP models built from 200 solutions.

Looking at the number of incorrect comparisons, we can see that by using relations under uncertainty with RF the results are slightly better than with GP. The reason for that is in the fact, that the approximations with RF have relatively wide confidence intervals which results in high number of confidence interval reductions. Therefore, most solutions have to be exactly evaluated in order to perform the comparisons. So the reason for a lower number of incorrect comparisons is not the higher quality of the surrogate models, but in the fact that more solutions are exactly evaluated and are therefore without uncertainty. Since in surrogate-model-based optimization exactly evaluated solutions are typically computationally expensive, a modeling technique that exactly evaluates most of the solutions is not very useful.

4. Conclusion

In this paper we compared random forest and Gaussian process modeling techniques in the context of surrogate-modelbased multiobjective optimization. We compared their approximation accuracy and width of the confidence intervals.

The results show that surrogate models built with GP modeling produce more accurate approximations with narrower confidence intervals. Due to narrower confidence intervals the comparisons of solutions approximated with GP modeling require less additional exact solution evaluations. As a result, we can conclude that GP modeling is more appropriate for use in a surrogate-model-based algorithm than RF.

References

[1] Breiman, L. (2001). Random forests. Machine Learning, 45 (1) 5–32.

[2] Deb, K. (2001). Multi-Objective Optimization Using Evolutionary Algorithms. Wiley, New York.

[3] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGAII. *IEEE Transactions on Evolutionary Computation*, 6 (2) 182–197.

[4] Mlakar, M., Petelin, D., Tušar, T., Filipic, B. (2014). GPDEMO: Differential evolution for multiobjective optimization based on

Gaussian process models. *European Journal of Operational Research*, 2014, doi: 10.1016/j.ejor.2014.04.011.

[5] Mlakar, M., Tušar, T., Filipic, B. (2014). Comparing solutions under uncertainty in multiobjective optimization. *Mathematical Problems in Engineering*, 2014, doi: 10.1155/2014/817964.

[6] Poloni, C., Giurgevich, A., Onesti, L., Pediroda, V. (2000). Hybridization of a multi-objective genetic algorithm, a neural network and a classical optimizer for a complex design problem in fluid dynamics. *Computer Methods in Applied Mechanics and Engineering*, 186 (2) 403–420.

[7] Rasmussen, C. E., Williams, C. (2006). Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA.