The price performance of performance models



Marcus Ritter, Technical University of Darmstadt

2022 SIAM Conference on Parallel Processing for Scientific Computing



2/25/2022 | Technical University of Darmstadt, Germany | Marcus Ritter | 1

Photo: Alex Becker / TU Darmstadt

Acknowledgement



TU Darmstadt

•

- Alexander Geiß
- Benedikt Naumann
- Thorsten Reimann
- Sebastian Rinke
- Marcus Ritter
- Felix Wolf



ETH Zurich

- Alexandru Calotoiu
- Torsten Hoefler



Performance model



Formula that expresses relevant performance metric as a function of one or more execution parameters



Automatic empirical performance modeling with Extra-P





Source: A. Calotoiu, D. Beckingsale, C. W. Earl, T. Hoefler, I. Karlin, M. Schulz, F. Wolf: Fast Multi-Parameter Performance Modeling. In Proc. of the 2016 IEEE International Conference on Cluster Computing (CLUSTER), Taipei, Taiwan, September 2016

Challenges





Cost of the required experiments

Money Bag Image Source: https://publicdomainvectors.org/en/free-clipart/Money-bag-vector/2806.html Applications

Run-to-run variation / noise



Experiments can be expensive

Need $5^{(m+1)}$ experiments, m = # parameters





How many data points do we really need?



TECHNISCHE

UNIVERSITÄT DARMSTADT

Learning cost-effective sampling strategies [Ritter et al., IPDPS'20]



TECHNISCHE UNIVERSITÄT

DARMSTADT

2/25/2022 | Technical University of Darmstadt, Germany | Marcus Ritter | 9

Heuristic parameter-value selection strategy





Synthetic data evaluation





Synthetic evaluation results

3 parameters, 5% noise





Case studies





Applicatio	n	#Parameters	Extra points	Cost savings [%]	Prediction error [%]
FASTEST		2	0	70	2
Kripke		3	3	99	39
Relearn		2	0	85	11

2/25/2022 | Technical University of Darmstadt, Germany | Marcus Ritter | 13

Neural Network Image Source: https://freesvg.org/img/Neural-Network-2.png



DNNs often better at guessing models in the presence of noise





2/25/2022 | Technical University of Darmstadt, Germany | Marcus Ritter | 15

Gaussian Process Regression

Optimized Measurement Point Selection

Goal: better tradeoff between accuracy and cost for specific models



GPR Image Source: https://commons.wikimedia.org/wiki/File:Gaussian_Process_Regression.png

TECHNISCHE UNIVERSITÄT DARMSTADT

Conclusion



Торіс	Learning cost-effective sampling strategies	Noise-resilient modeling with DNNs	
Improvements	 Cheaper performance models Less measurements required for modeling More flexible experiment design 	 Increased model accuracy and predictive power for noisy measurements Adaptive modeler automatically selects optimal modeling approach 	
Bibliography	Marcus Ritter, Alexandru Calotoiu, Sebastian Rinke, Thorsten Reimann, Torsten Hoefler, Felix Wolf: Learning Cost-Effective Sampling Strategies for Empirical Performance Modeling. IPDPS 2020.	Marcus Ritter, Alexander Geiß, Johannes Wehrstein, Alexandru Calotoiu, Thorsten Reimann, Torsten Hoefler, Felix Wolf. Noise-Resilient Empirical Performance Modeling with Deep Neural Networks. IPDPS 2021.	

Extra-P 4.0

- Includes the new sparse modeler
- Available as a Python package
- Available at: https://github.com/extra-p/extrap
- Support for Windows, Linux (Ubuntu), and MacOS
- Easy installation via pip
- BSD 3-Clause License





