Comparison of Analytical and Empirical Performance Models

A Case Study on Multigrid Systems

Christian Kaltenecker
University of Passau, Germany

Team

Christian Kaltenecker

Alexander Grebhahn

Sven Apel

Configurable Systems

Which configuration is performance-optimal?

Performance Models

- Mathematical formulas for estimating the performance

Analytical

- Created by domain experts

\+ Applicable on different hardware

\- Requires domain knowledge
\- Difficult for complex software

Empirical

- Created by tools with measurement results of a specific hardware

SPL Conqueror

\+ Applicable on complex systems

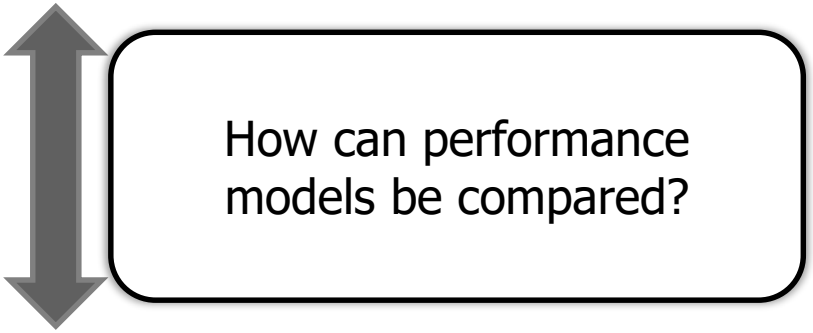
± Requires performance measurements

\- Only applicable on measured hardware

Performance Models

- Analytical:

$$23.37 * \log_2(px) + 23.37 * \log_2(nx) + 2.34 * 10^{-4} * nx$$



- Empirical:

$$-5 - 3 * px + 1.5 * px^2 + 10^{-5} * nx$$

Comparison Strategies

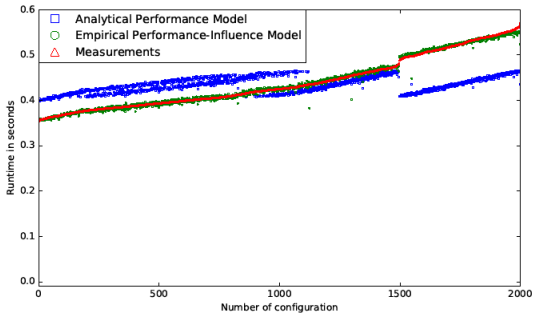
Syntactic

Compares the coefficients of the variables

	$[nx]$
$T_{analytical}$	$2.34 * 10^{-4} * nx$
$T_{empirical}$	$10^{-5} * nx$

Semantic

Compares the prediction results



Application of:

- Error Rate
- Jaccard Similarity
- Pearson Correlation

Hybrid

Uses both syntactic and semantic elements for the comparison

	$[nx]$
$T_{analytical}$	$2.34 * 10^{-4} * nx$
$T_{empirical}$	$10^{-5} * nx$

Calculate influence and use it as weight for Pearson correlation

Conversion of the Performance Models

- Sort each term into equivalence classes

$$23.37 * \log_2(px) + 23.37 * \log_2(nx) + 2.34 * 10^{-4} * nx$$

	$[px]$	$[nx]$
$T_{analytical}$	$23.37 * \log_2(px)$	$23.37 * \log_2(nx)\,;\, 2.34 * 10^{-4} * nx$

Syntactic Comparison

- Use the equivalence classes and compute score according to the formula:

$$\text{score}OfTerms(e,a) = \begin{cases} -1, & if\ the\ other\ model\ has\ no\ such\ equivalence\ class \\ 0, & if\ the\ equivalence\ class\ exists, but\ no\ such\ term \\ 1 + simValue(e,a), & if\ the\ term\ exists\ in\ the\ other\ model \end{cases}$$

$$simValue(e,a) = \max(0, 1 - \frac{|e-a|}{\max(e,a)})$$

$$23.37 * \log_2(px) + 23.37 * \log_2(nx) + 2.34 * 10^{-4} * nx$$

$$-5 - 3 * px + 10^{-5} * nx + 1.5 * px^2$$

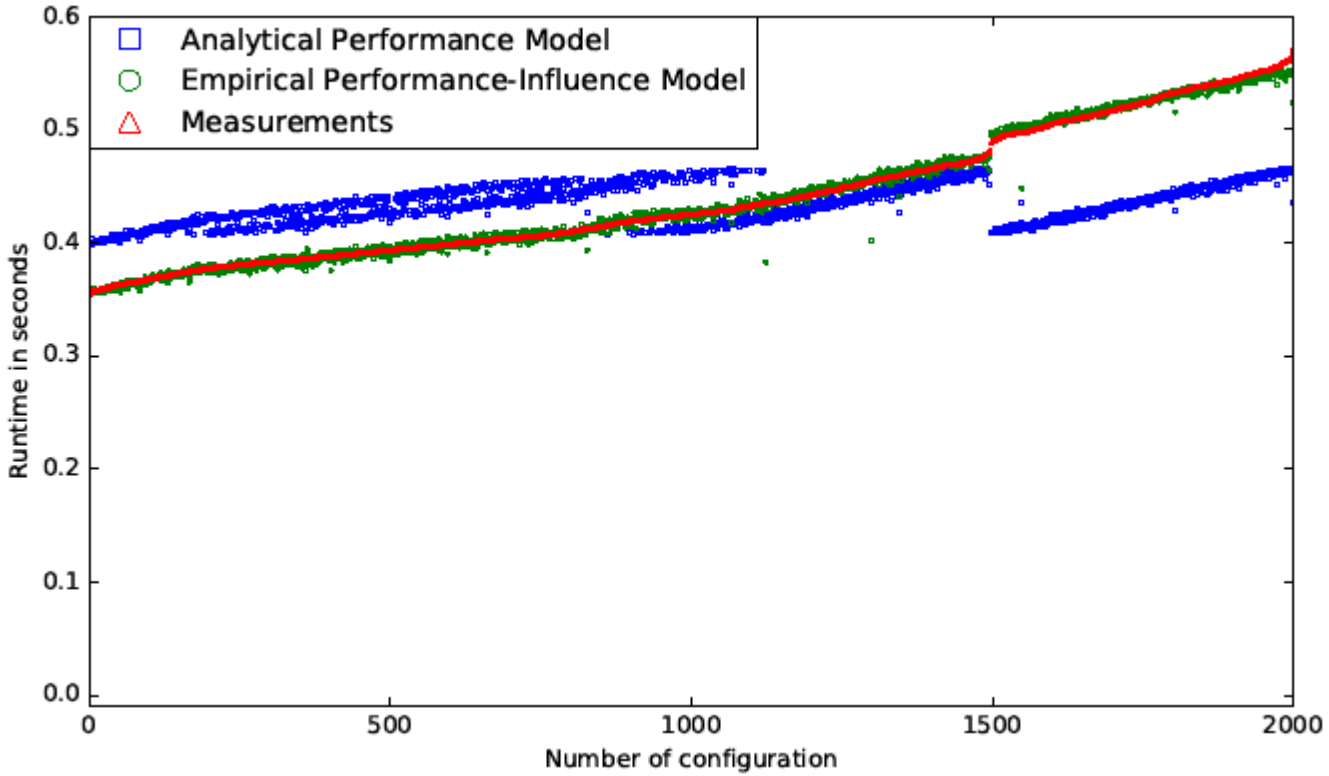
	$[constant]$	$[px]$	$[nx]$
$T_{analytical}$		$23.37 * \log_2(px)$	$23.37 * \log_2(nx); 2.34 * 10^{-4} * nx$
$T_{empirical}$	-5	$-3 * px; 1.5 * px^2$	$10^{-5} * nx$
$Score$	-1	0	$1 + 0.43$

Is this good or bad?

$$Score = 0.43$$

Semantic Comparison

- Compute the results of the performance models



Semantic Comparison

- Apply different measures on the computed results

Name	Result
Error Rate	10.1%
Jaccard Similarity	90%
Manhattan Distance	0.046
Euclidean Distance	0.0029
Pearson Correlation	0.41 (medium)

Good or bad?

Hybrid Comparison

- Calculate the influence of the equivalence classes:

	$[constant]$	$[px]$	$[nx]$
$T_{analytical}$		$23.37 * \log_2(px)$	$23.37 * \log_2(nx); 2.34 * 10^{-4} * nx$
$Influence_{analytical}$	0%	1.1%	98.9%
$T_{empirical}$	-5	$-3 * px; 1.5 * px^2$	$10^{-5} * nx$
$Influence_{empirical}$	0%	71.2%	28.8%

- Use the similarity of the influences as a weight for the Pearson correlation between the equivalence classes
  Result: 0.3

Conclusion

Performance Models

- Mathematical formulas for estimating the performance

Analytical

- Created by domain experts

\+ Applicable on different hardware

\- Requires domain knowledge
\- Difficult for complex sofware

Empirical

- Created by tools with measurement results of a specific hardware

\+ Applicable on complex systems

± Requires performance measurements

\- Only applicable on measured hardware

Syntactic Comparison

- Use the equivalence classes and compute score according to the formula:

$$scoreOfTerms(e,a) = \begin{cases} -1 & \text{if the other model has no such equivalence class} \\ 0 & \text{if the equivalence class exists,but no such term} \\ 1 + simValue(e,a) & \text{if the term exists in the other model} \end{cases}$$

$$simValue(e,a) = \max(0, 1 - \frac{|e-a|}{a})$$

$$23.37 * \log_2(px) + 23.37 * \log_2(nx) + 2.34 * 10^{-4} * nx$$

$$-5 - 3 * px + 10^{-5} * nx + 1.5 * px^2$$

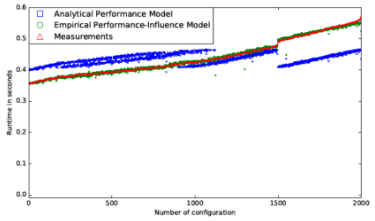
	$[constant]$	$[px]$	$[nx]$
$T_{analytical}$		$23.37 * \log_2(px)$	$23.37 * \log_2(nx); 2.34 * 10^{-4} * nx$
$T_{empirical}$	-5,	$-3 * px; 1.5 * px^2$	$10^{-5} * nx$
$Score$	-1	0	$1 + 0.43$

Is this good or bad?

$Score = 0.43$

Semantic Comparison

- Compute the results of the performance models



Hybrid Comparison

- Calculate the influence of the equivalence classes:

	$[constant]$	$[px]$	$[nx]$
$T_{analytical}$		$23.37 * \log_2(px)$	$23.37 * \log_2(nx); 2.34 * 10^{-4} * nx$
$Influence_{analytical}$	0%	27.7%	72.3%
$T_{empirical}$	-5	$-3 * px; 1.5 * px^2$	$10^{-5} * nx$
$Influence_{empirical}$	0%	73%	27%

- Use the influence as a weight for the Pearson correlation between the equivalence classes
  Result: 0.15 $(weak)$

Thank you for your attention!