Predicting WCET Trends in Long-lived Real-time Applications

22nd International Conference on Reliable Software Technologies
Ada-Europe 2017

Xiaotian Dai and Alan Burns

Real-time Systems Group
University of York, UK
Contents

• I. Motivation
• II. Adaptive Scheduling Framework
• III. Trend Identification Methods
• VI. Evaluation
• V. Conclusion
Background | Motivation

• Worst-case execution time (WCET)
  o is important in **timing analysis** (DO-178 and ISO26262)
  o static and measurement-based
  o pWCET

Background | Motivation

- Current understanding of WCETs:
  - A theoretical boundary exists, if designed and programmed with constrained models.
  - Known as a static, upper-bound value of execution times

Issues | Motivation

- Data accessing time ↑
  - relevant data growth
  - hard disk fault/fragmented
Issues | Motivation

- Hardware ageing: computer systems age just like humans
  - CPU transistor ageing: fundamental speed ↓
  - Thermal performance decreased: lacking maintenance
Issues | Motivation

- Emerging systems
  - Self-adaptive systems: increased software complexity
  - Machine that learns and evolves, e.g., autonomous robots
Issues (continue) | Motivation

- Contribute negative and non-deterministic effects on WCETs.
- Subtle in a short period, but noticeable in long-term.
- Traditional WCET analysis could solve this by giving a very pessimistic boundary.

- A new perspective on WCET: a dynamic view of WCET (dWCET), as an extension of traditional WCET analysis.
dWCET | Motivation

• Run-time modelling of WCET.
• Enhanced Parametric WCET: $WCET = f(t, \text{system changes[, mode, state, input, ...]})$).

• Pro 1: Early detection of potential timing errors, and achieve graceful degradation.
• Pro 2: Utilize resources better (with feedback scheduling).
Adaptive Feedback Scheduling

• A variation of Feedback Control Scheduling (FBS)
• Adaptive
  o ability to handle unexpected events
  o understanding of the system increases

In Practice | A-FBS

- A-FBS uses with an adaptive control system:
Advantages | A-FBS

• Explicitly monitoring and modelling the system.
• Handling uncertainties in run-time executions.
• Increase system resilience: automated the process of (proactive) fault tolerance.
• Dynamic resource allocation: run-time optimization of scheduling.
What’s Next?

• The activation of system changes/degrades will be propagated in the system and reflects on WCETs.

• There are many ways we can model dynamics in WCETs.

• In this initial study, we consider one of these: trends in WCET.

• Use a linear model to describe trend.
Trend Identification

• Many techniques in the literature:
  o AR / ARMAX
  o Regression Analysis
  o Non-parametric
  o EVT
  o Neural Network
  o Decision Tree Regression
  o ...

• but not all of them fit our case:
  o data points are execution times
  o distribution is not known
  o few prior knowledge
  o need a long-term prediction
Methods | Trend Identification

• Non-parametric Methods
  o TSE: Theil-Sen Estimator

• Regression Analysis
  o OLS: Ordinary least-squares regression (OLS-regression)

• Extreme-value Theory
  o EVD: Generalized Extreme-value distribution

• Machine Learning Methods
  o SVR: Support Vector Regression

• These methods have never been used to analysis trends in WCETs. How to evaluate?

Dataset | Evaluation

- Use synthetic data to make it evaluable.
- One observation represents a high watermark of run-time executions.
- Markov model with multiple dominated paths.
- An increasing trend only in the worst-case path.

<table>
<thead>
<tr>
<th>Group</th>
<th>Subgroup</th>
<th>Dataset Index</th>
<th>Data Size</th>
<th>Increasing Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A1</td>
<td>1 - 10</td>
<td>5,000</td>
<td>0%</td>
</tr>
<tr>
<td>B</td>
<td>B1</td>
<td>11 - 20</td>
<td>5,000</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>21 - 30</td>
<td>2,500</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>B3</td>
<td>31 - 40</td>
<td>1,667</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>B4</td>
<td>41 - 50</td>
<td>1,250</td>
<td>4%</td>
</tr>
</tbody>
</table>
Evaluation

• The Evaluation Framework
Evaluation

• The Evaluation Framework
Evaluation

• The Evaluation Framework
Evaluation

• The Evaluation Framework
Evaluation

• The Evaluation Framework
Evaluation

- The Evaluation Framework
Evaluation

• The Evaluation Framework
Evaluation

• The Evaluation Framework
Pre-processing | Evaluation

- Evaluated with raw, block maxima and r-largest
- Mean (absolute) error of trend magnitude
Pre-processing | Evaluation

- Evaluated with raw, block maxima and r-largest
- Mean (absolute) error of trend magnitude
Dataset Sensitivity | Evaluation

- All methods use block maxima
- Subgroups are separated by dashed lines
Trend Error | Evaluation

- Evaluate Trend error (= actual k - predicted k):

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Median</th>
<th>Mean</th>
<th>Maximum</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>olr-max</td>
<td>-1.91</td>
<td>4.16</td>
<td>6.31</td>
<td>27.64</td>
<td>7.21</td>
</tr>
<tr>
<td>olr-r</td>
<td>-2.07</td>
<td>7.68</td>
<td>10.59</td>
<td>32.10</td>
<td>9.31</td>
</tr>
<tr>
<td>tse-max</td>
<td>-1.12</td>
<td>2.23</td>
<td>3.07</td>
<td>17.45</td>
<td><strong>3.27</strong></td>
</tr>
<tr>
<td>tse-r</td>
<td>-1.15</td>
<td>9.14</td>
<td>9.91</td>
<td>26.00</td>
<td>7.72</td>
</tr>
<tr>
<td>svr-max</td>
<td>-5.24</td>
<td><strong>0.15</strong></td>
<td><strong>1.60</strong></td>
<td>25.65</td>
<td>5.71</td>
</tr>
<tr>
<td>svr-r</td>
<td>-1.00</td>
<td>9.65</td>
<td>12.72</td>
<td>44.36</td>
<td>12.74</td>
</tr>
<tr>
<td>evd-max</td>
<td>-1.46</td>
<td>1.60</td>
<td>3.40</td>
<td>23.47</td>
<td>4.75</td>
</tr>
<tr>
<td>evd-r</td>
<td>-0.45</td>
<td>5.34</td>
<td>6.86</td>
<td>30.20</td>
<td>6.77</td>
</tr>
</tbody>
</table>
Normalized Performance | Evaluation
Normalized Performance | Evaluation

Valid Estimations

Invalid Estimations

True Positives

False Positives

False Negatives

$k = 0$  $k = 1\%$  $k = 2\%$  $k = 3\%$  $k = 4\%$

IDP-MAX  LR-MAX  TSE-MAX  SVR-MAX  EVD-MAX
Mean Penalties | Evaluation

Table 4. Mean penalties over all datasets for each prediction method

<table>
<thead>
<tr>
<th></th>
<th>OLR</th>
<th>TSE</th>
<th>SVR</th>
<th>EVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>raw</td>
<td>62</td>
<td>62</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>maxima</td>
<td>58.28</td>
<td>29.02</td>
<td>42.26</td>
<td>49.68</td>
</tr>
<tr>
<td>r-largest</td>
<td>58.2</td>
<td>53.82</td>
<td>77.58</td>
<td>55.76</td>
</tr>
</tbody>
</table>
Conclusion

• Introduced dWCET and A-FBS
• Evaluated data pre-processing methods
• Result is sensitive to datasets
• Best two methods: svr-max and tse-max

• Future work
  o More dedicated dataset: e.g., with non-linear trend
  o Other analysis: anomaly detection, pattern recognition
  o Multiple variables + PCA
  o Evaluate with real-world data
Thank You for your attention!

Any Question/Comment?