

Anomaly Detection Approach

Matthias Rohr

BTC Business Technology Consulting AG, Germany,
Graduate School TrustSoft*, University of Oldenburg, Germany

March 21, 2011

Software Engineering Group, University of Kiel,
Kiel, Germany



* This work is supported by the German Research Foundation (DFG), grant GRK 1076/1

Primary thesis contribution

- Trace context sensitive timing behavior analysis (TracSTA)
- Workload intensity sensitive timing behavior analysis (WiSTA)

Secondary thesis contribution

- Application of TracSTA and WiSTA in anomaly-detection-based fault localization scenario ([Rohr, 2006; Rohr et al., 2007])

Primary contribution:
Timing Behavior Analysis Methods

Evaluation:
Quantitative empirical evaluation in industry studies



Secondary contribution: Fault localization approach

Evaluation: Proof-of-concept demonstration in lab-studies and application of the monitoring infrastructure in industry systems

1. Instrumentation and Monitoring

- Recording of:
 - **Response times** of software operation executions
 - **Execution sequences** corresponding to user requests
 - Host identifier
- Reconstruction of Traces and Dependency Graphs

2. Trace-Context-Sensitive Timing Behavior Analysis

3. Workload-Intensity-Sensitive Timing Behavior Analysis

4. Anomaly Detection

5. Anomaly Correlation and Fault Localization

1. Instrumentation and Monitoring

2. Trace-Context-Sensitive Timing Behavior Analysis

- 1 Identification of trace contexts
- 2 Defining classes of observations based on the trace context
- 3 (Re-)Merging classes

3. Workload-Intensity-Sensitive Timing Behavior Analysis

4. Anomaly Detection

5. Anomaly Correlation and Fault Localization

Thesis
overview

Fault
Localization
& Anomaly
Detection

References

1. Instrumentation and Monitoring

2. Trace-Context-Sensitive Timing Behavior Analysis

3. Workload-Intensity-Sensitive Timing Behavior Analysis

- 1 Definition of a workload metric by machine learning
- 2 Splitting observations according to workload-intensity

4. Anomaly Detection

5. Anomaly Correlation and Fault Localization

Thesis
overview

Fault
Localization
& Anomaly
Detection

References

1. Instrumentation and Monitoring

2. Trace-Context-Sensitive Timing Behavior Analysis

3. Workload-Intensity-Sensitive Timing Behavior Analysis

4. Anomaly Detection

- Evaluation of new observations in context of a profile.
- How normal is a new observation?

5. Anomaly Correlation and Fault Localization

Thesis
overview

Fault
Localization
& Anomaly
Detection

References

1. Instrumentation and Monitoring

2. Trace-Context-Sensitive Timing Behavior Analysis

3. Workload-Intensity-Sensitive Timing Behavior Analysis

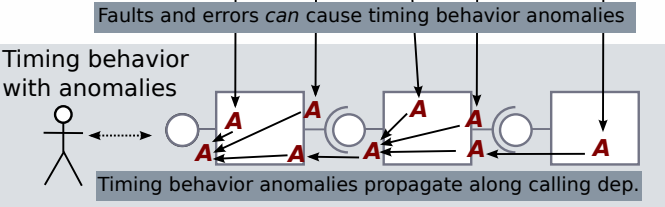
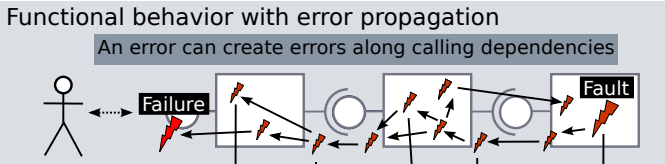
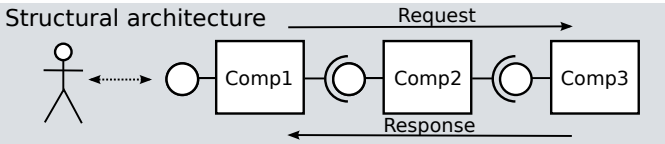
4. Anomaly Detection

5. Anomaly Correlation and Fault Localization

- Derivation of component ratings from execution ratings
- Derivation of causes from symptoms



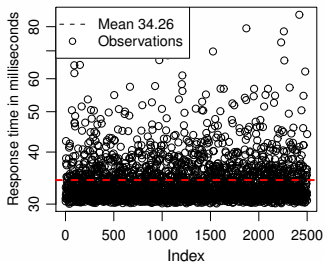
Assumptions: Anomaly propagation model



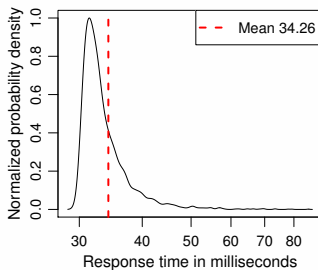
Thesis
overview

Fault
Localization
& Anomaly
Detection

References



(a)

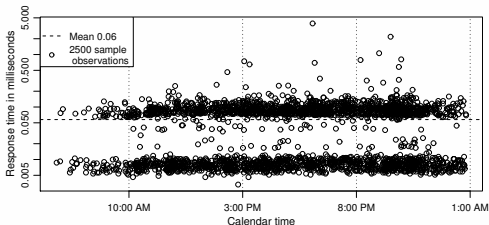


(b)

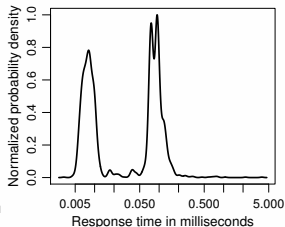
Software timing behavior anomaly detection requirements

- Heavy tails, right skewed (i.e. mode < median < mean), multi-modality
- Context information (parameters, workload, state ...)

Anomaly detection (2/3)



(c) Scatter plot of response times.



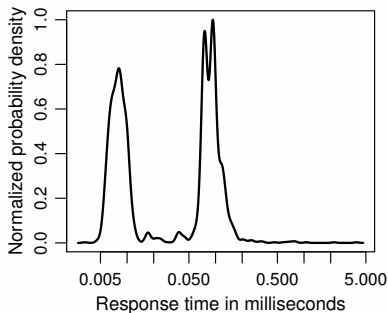
(d) Normalized probability density of the response times.

Thesis
overview

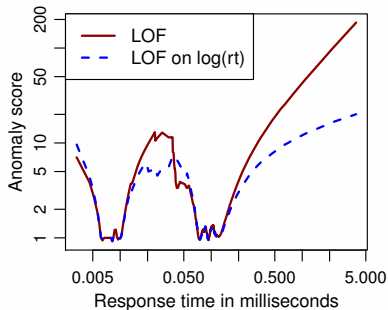
Fault
Localization
& Anomaly
Detection

References

Anomaly detection (3/3)



(e) Normalized probability density of the response times.



(f) LOF anomaly scores for response times and log-transformed response times.

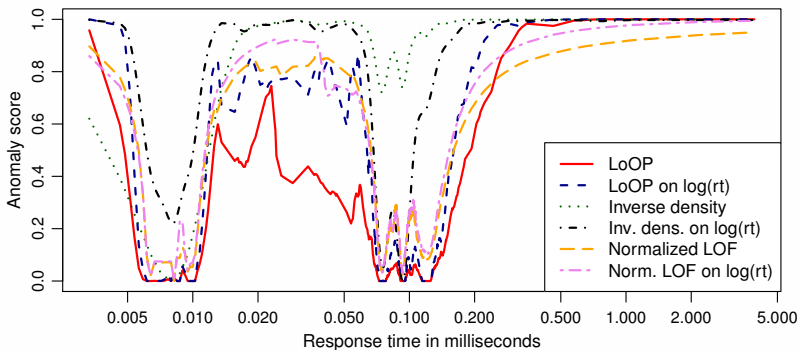
LOF [Breunig et al., 2000] and LoOP [Kriegel et al., 2009] algorithms provided by ELKI Framework [Achtert et al., 2010].

Thesis
overview

Fault
Localization
& Anomaly
Detection

References

Anomaly detection (4/4)



(g) Anomaly scores for response times using different anomaly detection functions.

LOF [Breunig et al., 2000] and LoOP [Kriegel et al., 2009] algorithms provided by ELKI Framework [Achtert et al., 2010].

Thesis
overview

Fault
Localization
& Anomaly
Detection

References

Integration into Kieker or our other ongoing work

- Sharing of anomaly detection algorithms
- Stream-oriented wrappers for non-stream-processing algorithms
- Standardized interfaces and patterns (e.g. annotation?) for Kieker integration

References I

- E. Aichert, H.-P. Kriegel, L. Reichert, E. Schubert, R. Wojdanowski, and A. Zimek. Visual evaluation of outlier detection models. In *Database Systems for Advanced Applications, 15th International Conference, DASFAA 2010, Tsukuba, Japan, April 1-4, 2010, Proceedings, Part II*, volume 5982 of *Lecture Notes in Computer Science*, pages 396–399. Springer, 2010. ISBN 978-3-642-12097-8. doi: 10.1007/978-3-642-12098-5_34.
- M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander. Lof: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data, SIGMOD '00*, pages 93–104. ACM, 2000. ISBN 1-58113-217-4. doi: 10.1145/342009.335388.
- H.-P. Kriegel, P. Kröger, E. Schubert, and A. Zimek. LoOP: local outlier probabilities. In D. W.-L. Cheung, I.-Y. Song, W. W. Chu, X. Hu, and J. J. Lin, editors, *Proceedings of the 18th ACM Conference on Information and Knowledge Management, CIKM 2009, Hong Kong, China, November 2-6, 2009*, pages 1649–1652. ACM, 2009. doi: 10.1145/1645953.1646195.
- M. Rohr. Timing Behavior Anomaly Detection for Fault Localization. In J. Happe, H. Koziolk, M. Rohr, C. Strom, and T. Warns, editors, *Proceedings of the International Research Training Groups Workshop, Dagstuhl*, page 20, Nov. 2006. ISBN 3-936771-871.
- M. Rohr, S. Giesecke, and W. Hasselbring. Timing Behavior Anomaly Detection in Enterprise Information Systems. In J. Cardoso, J. Cordeiro, and J. Filipe, editors, *Proceedings of the Ninth International Conference on Enterprise Information Systems (ICEIS'07)*, volume DISI, pages 494–497. INSTICC Press, June 2007. ISBN 978-972-8865-88-7.

Thesis
overview

Fault
Localization
& Anomaly
Detection

References