Investigating the Use of Bayesian Networks in the Hora Approach for Component-based Online Failure Prediction

Teerat Pitakrat, André van Hoorn
University of Stuttgart
Institute of Software Technology (ISTE)
Reliable Software Systems (RSS) Group
Stuttgart, Germany

Nov 27, 2014 @ SOSP 2014, Stuttgart
Microsoft recovering from seven-hour long Outlook.com outage

By Tom Warren on August 14, 2013 05:32 pm

Sorry, there seems to be a problem with Outlook.

We're having a problem accessing email. You might not be able to send or receive email.

We identified a solution to the problem and have brought it up gradually. Aug 14 21:50

Show earlier updates
Microsoft recovering from seven-hour long Outlook.com outage

AWS Server Issues Take Down Instagram, Vine, Airbnb And IFTTT

Those of you looking to spend the rest of today watching people do other things on Instagram or Vine probably just had a rough time trying to do it. Both services went offline for over an hour, most likely because of issues with Amazon Web Services.

Instagram was the first of the two to publicly acknowledge its issues on Twitter, and Vine followed suit half an hour later.

The deluge of tweets that accompanied the services’ initial hiccups first started at around 4 p.m. Eastern time, and only increased in intensity as users found they couldn’t share pictures of their food or their meticulously crafted video
Service Failure

Motivation: Failure Management

Investigating the Use of Bayesian Networks in Hora for C-B OFP

T. Pitakrat, A. van Hoorn

Nov. 27, 2014 @ SOSP 2014 2 / 29
“A service failure, often abbreviated here to failure, is an event that occurs when the delivered service deviates from correct service.”

— Avizienis et al. [2004]
Reactive vs. Proactive Failure Mgmt.

Motivation: Failure Management

Reactive

Failure detected
Start recovery
System recovered

QoS
100%
0%
Reactive vs. Proactive Failure Mgmt.

Motivation: Failure Management

Reactive
- Failure detected
- Start recovery
- System recovered
- QoS: 100% 0%

Proactive
- Failure predicted
- Prepare recovery
- Failure
- System recovered
- QoS: 100% 0%
Agenda

1. Motivation: Failure Management

2. [Recap] Hora: Online Failure Prediction for CB Systems

3. Hora: Framework and Implementation

4. Evaluation

5. Conclusion
Related Approaches vs. Hora Approach

[Recap] Hora: Online Failure Prediction for CB Systems

Amin et al. [2012], Liang et al. [2007]

\[ S_1 \circ \]
\[ S_2 \circ \]
Related Approaches vs. Hora Approach

[Recap] Hora: Online Failure Prediction for CB Systems

Bielefeld [2012]
[Recap] Hora: Online Failure Prediction for CB Systems

Related Approaches vs. Hora Approach

Component Dependency
Component-level Prediction Models
System-level Prediction Model

Pitakrat [2013], Pitakrat et al. [2014b]
Related Approaches vs. Hora Approach

[Recap] Hora: Online Failure Prediction for CB Systems

Component-level Prediction Models

+ Component Dependency

System-level Prediction Model

Pitakrat [2013], Pitakrat et al. [2014b]
[Recap] Hora: Online Failure Prediction for CB Systems
[Recap] Hora: Online Failure Prediction for CB Systems

Hora: Component-level Prediction Models

Investigating the Use of Bayesian Networks in Hora for C-B OFP

T. Pitakrat, A. van Hoorn

Nov. 27, 2014 @ SOSP 2014
Hora: Framework Architecture

[Recap] Hora: Online Failure Prediction for CB Systems

Hora

System-level Predictor
Monitoring
Reader
!
!

Component-level Predictors

HDD Failure Predictor

PAD

Event Log Analyzer

System-level Predictor

PCM
SLAastic
...

CDT

Hora

Kieker, Weka, R, ESPER, ...

Questions:

1. What is a suitable model for System-level Prediction Model (SPM)?
2. How to transform architectural models to CDT and to SPM?
3. How does Hora improve online failure prediction?

Becker et al. [2009], Bielefeld [2012], Pitakrat et al. [2013; 2014a], van Hoorn [2014]
Questions:

1. What is a suitable model for System-level Prediction Model (SPM)?
2. How to transform architectural models to CDT and to SPM?
3. How does Hora improve online failure prediction?

Becker et al. [2009], Bielefeld [2012], Pitakrat et al. [2013; 2014a], van Hoorn [2014]
Motivation: Failure Management

[Recap] Hora: Online Failure Prediction for CB Systems

Hora: Framework and Implementation

Evaluation

Conclusion
Hora: Framework Architecture

Hora: Framework and Implementation

Questions:

1. What is a suitable model for System-level Prediction Model (SPM)?
2. How to transform architectural models to CDT and to SPM?
3. How does Hora improve online failure prediction?
Bayesian network: probabilistic graphical model

Bayesian network library used in Hora: https://github.com/kutschkem/Jayes
Bayesian network: probabilistic graphical model

<table>
<thead>
<tr>
<th>Rain</th>
<th>Sprinkler</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.01 0.99</td>
</tr>
<tr>
<td>F</td>
<td>0.4 0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sprinkler</th>
<th>Rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.8 0.2</td>
</tr>
<tr>
<td>F</td>
<td>0.0 1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sprinkler</th>
<th>Rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rain</th>
<th>Grass wet</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.99 0.01</td>
</tr>
<tr>
<td>F</td>
<td>0.9 0.1</td>
</tr>
<tr>
<td>T</td>
<td>0.8 0.2</td>
</tr>
<tr>
<td>F</td>
<td>0.0 1.0</td>
</tr>
</tbody>
</table>

Bayesian network library used in Hora: https://github.com/kutschkem/Jayes
Bayesian network: probabilistic graphical model

Bayesian network library used in Hora: https://github.com/kutschkem/Jayes
Questions:

1. What is a suitable model for System-level Prediction Model (SPM)?
2. How to transform architectural models to CDT and to SPM?
3. How does Hora improve online failure prediction?
Component Dependency Table
Component Dependencies
Hora: Framework and Implementation
<table>
<thead>
<tr>
<th></th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$H_1$</th>
<th>$H_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>$C_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>●</td>
</tr>
<tr>
<td>$H_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$C_1$</td>
<td>$C_2$</td>
<td>$C_3$</td>
<td>$H_1$</td>
<td>$H_2$</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>$C_1$</td>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td></td>
<td></td>
<td>0.5</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>$H_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Investigating the Use of Bayesian Networks in Hora for C-B OFP
System-level Prediction Model

<table>
<thead>
<tr>
<th></th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$H_1$</th>
<th>$H_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>0.5</td>
<td></td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td>0.5</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>$H_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>$H_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
System-level Prediction Model

Hora: Framework and Implementation

\[
\begin{array}{|c|c|c|c|c|}
\hline
 & C_1 & C_2 & C_3 & H_1 & H_2 \\
\hline
C_1 & 0.5 & 0.5 & 1.0 & & \\
C_2 & & 0.5 & 1.0 & & \\
C_3 & & & 1.0 & & \\
H_1 & & & & 1.0 & \\
H_2 & & & & & 0.8 & 0.2 \\
\hline
\end{array}
\]
System-level Prediction Model

Hora: Framework and Implementation

\[
\begin{array}{c|ccccc}
 & C_1 & C_2 & C_3 & H_1 & H_2 \\
\hline
C_1 & 0.5 & 0.5 & 1.0 & & \\
C_2 & & 0.5 & 1.0 & & \\
C_3 & & & 1.0 & & \\
H_1 & & & & & \\
H_2 & & & & & \\
\end{array}
\]

\[
\begin{array}{c|cc}
H_2 & C_3 \\
\hline
\checkmark & 0.9 & 0.1 \\
\times & 0.0 & 1.0 \\
\end{array}
\]
System-level Prediction Model

Hora: Framework and Implementation

<table>
<thead>
<tr>
<th></th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$H_1$</th>
<th>$H_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td></td>
<td>0.5</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_1$</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>$H_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H_1$</th>
<th>$C_3$</th>
<th>$C_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔</td>
<td>✗</td>
<td>0.9</td>
</tr>
<tr>
<td>✔</td>
<td>✗</td>
<td>0.5</td>
</tr>
<tr>
<td>✗</td>
<td>✔</td>
<td>0.0</td>
</tr>
<tr>
<td>✗</td>
<td>✔</td>
<td>0.0</td>
</tr>
</tbody>
</table>

T. Pitakrat, A. van Hoorn
Investigating the Use of Bayesian Networks in Hora for C-B OFP
Nov. 27, 2014 @ SOSP 2014
Hora: Framework Architecture

Hora: Framework and Implementation

Questions:

1. What is a suitable model for System-level Prediction Model (SPM)?
2. How to transform architectural models to CDT and to SPM?
3. How does Hora improve online failure prediction?
Evaluation

1. Motivation: Failure Management

2. [Recap] Hora: Online Failure Prediction for CB Systems

3. Hora: Framework and Implementation

4. Evaluation

5. Conclusion
Experiment Setup

Evaluation

- System under analysis
Experiment Setup

Evaluation

- System under analysis
- Fault injection
Experiment Setup

Evaluation

- System under analysis
- Fault injection
- Predicting response time violation using time series forecasting
SPM for JPetStore
Component Failure Prediction
CachingExecutor-1-update
Evaluation

CPM

SPM

CachingExecutor-1-update

Response time (ns)

Monitoring data

Failure threshold

Failure probability

Time

0e+00 1e+06 2e+06 3e+06 4e+06 5e+06


15:03 15:31 15:59 16:27

Monitoring data

Failure threshold

Failure probability

Time

0e+00 1e+06 2e+06 3e+06 4e+06 5e+06


15:03 15:31 15:59 16:27

Monitoring data

Failure threshold

Failure probability
Investigating the Use of Bayesian Networks in Hora for C-B OFP

Component Failure Prediction

Evaluation

CPM

SPM

CachingExecutor-1-update

CachingExecutor-1-update

Response time (ns)

Failure probability

Time

Failure threshold

Monitoring data

Failure probability

Monitoring data

Failure threshold
Component Failure Prediction

CachingExecutor-1-update

Evaluation

CPM

SPM

CachingExecutor–1–update ROC Curve

False positive rate

True positive rate

CachingExecutor–1–update ROC Curve

False positive rate

True positive rate
Component vs. System Failure Prediction

Evaluation

OrderMapper-7-insertOrderStatus

CPM

SPM

OrderMapper-7-insertOrderStatus

Response time (ns)

0e+00 1e+06 2e+06 3e+06 4e+06 5e+06 6e+06

Time


Monitoring data

Failure threshold

Failure probability
Evaluation
Component vs. System Failure Prediction
OrderService-9-insertOrder

Evaluation

CPM

SPM

OrderService-9-insertOrder

Monitoring data
Failure threshold
Failure probability

Time

Response time (ns)
Component vs. System Failure Prediction

Comparison of CPM and SPM for OrderService-9-insertOrder

CPM

SPM

Evaluation

T. Pitakrat, A. van Hoorn
Investigating the Use of Bayesian Networks in Hora for C-B OFP
Discussion of Preliminary Results

Evaluation

- CPMs perform quite well but can still be improved
- SPM shows potential to predict failure propagation using component dependency
- Although the failure propagates to other components, the predictions are not very good as the failure thresholds are not set properly
Agenda

1. Motivation: Failure Management
2. [Recap] Hora: Online Failure Prediction for CB Systems
3. Hora: Framework and Implementation
4. Evaluation
5. Conclusion

T. Pitakrat, A. van Hoorn
Investigating the Use of Bayesian Networks in Hora for C-B OFP
Nov. 27, 2014 @ SOSP 2014
Component-level Prediction Models

\[ C_1 \rightarrow C_2 \rightarrow C_3 \]

System-level Prediction Model

\[ H_1 \rightarrow C_1 \rightarrow H_2 \]

Component Dependency

$S_1, S_2 \rightarrow C_1 \rightarrow C_2 \rightarrow C_3 \rightarrow H_2 \rightarrow H_1$
Summary

Conclusion

Component-level Prediction Models

+ Component Dependency

System-level Prediction Model

<table>
<thead>
<tr>
<th></th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$H_1$</th>
<th>$H_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td></td>
<td>0.5</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_1$</td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>$H_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>
Component-level Prediction Models

\[ C_1 \] \[ C_2 \] \[ C_3 \] \[ H_1 \] \[ H_2 \]

- \[ C_1 \]: 0.5 \[ C_2 \]: 0.5 \[ C_3 \]: 1.0 \[ H_1 \]: 1.0 \[ H_2 \]: 0.5
Component Dependency

Component-level Prediction Models

\[ C_1 \quad C_2 \quad C_3 \quad H_1 \quad H_2 \]

\[
\begin{array}{c|cccc}
  & C_1 & C_2 & C_3 & H_1 & H_2 \\
\hline
C_1 & 0.5 & 0.5 & 1.0 & & \\
C_2 & 0.5 & 1.0 & & & \\
C_3 & 1.0 & & & & \\
H_1 & & & 1.0 & & \\
H_2 & & & & 1.0 & \\
\end{array}
\]

System-level Prediction Model

\[
\begin{array}{cccc}
H_1 & C_3 & & C_2 \\
\hline
\checkmark & \checkmark & 0.9 & 0.1 \\
\checkmark & \times & 0.5 & 0.5 \\
\times & \checkmark & 0.0 & 1.0 \\
\times & \times & 0.0 & 1.0 \\
\end{array}
\]

T. Pitakrat, A. van Hoorn

Investigating the Use of Bayesian Networks in Hora for C-B OFP

Nov. 27, 2014 @ SOSP 2014
Next Steps

Conclusion

- Improve prediction of CPMs
  - Apply different prediction techniques
- Calibrate the configuration parameters of Hora
  - Failure threshold
  - Lead time
  - …
- Extend evaluation settings
  - Evaluate with distributed lab study
  - Evaluate with large-scale production systems
- Develop a failure injection framework/benchmark


