Perfume
Resource-aware model inference

Tony Ohmann
Kevin Thai
Ivan Beschastnikh
Yuriy Brun

University of Massachusetts, Amherst
Facebook Inc.
University of British Columbia
University of Massachusetts, Amherst
Perfume: resource-aware model inference

Tony Ohmann
Kevin Thai
Ivan Beschastnikh
Yuriy Brun

University of Massachusetts, Amherst
Facebook Inc.
University of British Columbia
University of Massachusetts, Amherst
Motivation: system understanding

- A diagnostic system tests network connections

- A developer wants to understand:
  what causes the system to classify connections as problematic?

- The developer:
  - instruments the application
  - inspects the console log
Console log

- Complex
- Hard to parse
- Relevant information spread out

Perfume: resource-aware model inference
**Perfume: inferring a log model**

- **Perfume model**
  - separates types of behavior
    - broadband (left) and narrowband (right) separated
  - helps answer questions:

```
19.38.218.11 [31/May/2014:31200.0] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.1] "GET HTTP/1.1 /test bandwidth"
38.151.1.182 [31/May/2014:31200.2] "GET HTTP/1.1 /test bandwidth"
95.39.21.28 [31/May/2014:31200.3] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.8] "GET HTTP/1.1 /broadband"
38.151.1.182 [31/May/2014:31200.9] "GET HTTP/1.1 /broadband"
19.38.218.11 [31/May/2014:31202.0] "GET HTTP/1.1 /narrowband"
210.82.199.247 [31/May/2014:31202.1] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31202.2] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31202.3] "GET HTTP/1.1 /narrowband"
38.151.1.182 [31/May/2014:31203.6] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31204.1] "GET HTTP/1.1 /OK"
19.38.218.11 [31/May/2014:31205.7] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.0] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.8] "GET HTTP/1.1 /OK"
210.82.199.247 [31/May/2014:31208.3] "GET HTTP/1.1 /query"
210.82.199.247 [31/May/2014:31208.8] "GET HTTP/1.1 /problem"
19.38.218.11 [31/May/2014:31208.9] "GET HTTP/1.1 /query"
19.38.218.11 [31/May/2014:31209.7] "GET HTTP/1.1 /OK"
```

**Console log**

**Resource-aware model**

**Perfume**
Perfume: inferring a log model

- **Perfume** model
  - separates types of behavior
    - broadband (left) and narrowband (right) separated
  - visualizes problematic executions
  - helps answer questions:

```
19.38.218.11 [31/May/2014:31200.0] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.1] "GET HTTP/1.1 /test bandwidth"
38.151.1.182 [31/May/2014:31200.2] "GET HTTP/1.1 /test bandwidth"
95.39.21.28 [31/May/2014:31200.3] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.8] "GET HTTP/1.1 /broadband"
38.151.1.182 [31/May/2014:31200.9] "GET HTTP/1.1 /broadband"
19.38.218.11 [31/May/2014:31202.0] "GET HTTP/1.1 /narrowband"
210.82.199.247 [31/May/2014:31202.1] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31202.2] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31202.3] "GET HTTP/1.1 /narrowband"
38.151.1.182 [31/May/2014:31203.6] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31204.1] "GET HTTP/1.1 /OK"
19.38.218.11 [31/May/2014:31205.7] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.0] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.8] "GET HTTP/1.1 /OK"
210.82.199.247 [31/May/2014:31208.3] "GET HTTP/1.1 /query"
210.82.199.247 [31/May/2014:31208.8] "GET HTTP/1.1 /problem"
19.38.218.11 [31/May/2014:31208.9] "GET HTTP/1.1 /query"
19.38.218.11 [31/May/2014:31209.7] "GET HTTP/1.1 /OK"
```

Perfume: resource-aware model inference
Perfume: inferring a log model

- **Perfume** model
  - separates types of behavior
    - broadband (left) and narrowband (right) separated
  - visualizes problematic executions
  - helps answer questions:

```
19.38.218.11 [31/May/2014:31200.0] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.1] "GET HTTP/1.1 /test bandwidth"
38.151.1.182 [31/May/2014:31200.2] "GET HTTP/1.1 /test bandwidth"
95.39.21.28 [31/May/2014:31200.3] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.8] "GET HTTP/1.1 /broadband"
38.151.1.182 [31/May/2014:31200.9] "GET HTTP/1.1 /broadband"
19.38.218.11 [31/May/2014:31202.0] "GET HTTP/1.1 /narrowband"
210.82.199.247 [31/May/2014:31202.1] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31202.2] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31202.3] "GET HTTP/1.1 /narrowband"
38.151.1.182 [31/May/2014:31203.6] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31204.1] "GET HTTP/1.1 /OK"
19.38.218.11 [31/May/2014:31205.7] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.0] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.8] "GET HTTP/1.1 /OK"
210.82.199.247 [31/May/2014:31208.3] "GET HTTP/1.1 /query"
210.82.199.247 [31/May/2014:31208.8] "GET HTTP/1.1 /problem"
19.38.218.11 [31/May/2014:31208.9] "GET HTTP/1.1 /query"
19.38.218.11 [31/May/2014:31209.7] "GET HTTP/1.1 /OK"
```
Perfume: inferring a log model

- **Perfume** model
  - separates types of behavior
    - broadband (left) and narrowband (right) separated
  - visualizes problematic executions
  - helps answer questions:
    - problem = broadband clients with slow second query

```
19.38.218.11 [31/May/2014:31200.0] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.1] "GET HTTP/1.1 /test bandwidth"
38.151.1.182 [31/May/2014:31200.2] "GET HTTP/1.1 /test bandwidth"
95.39.21.28 [31/May/2014:31200.3] "GET HTTP/1.1 /test bandwidth"
210.82.199.247 [31/May/2014:31200.8] "GET HTTP/1.1 /broadband"
38.151.1.182 [31/May/2014:31200.9] "GET HTTP/1.1 /broadband"
19.38.218.11 [31/May/2014:31202.0] "GET HTTP/1.1 /narrowband"
210.82.199.247 [31/May/2014:31202.1] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31202.2] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31202.3] "GET HTTP/1.1 /narrowband"
38.151.1.182 [31/May/2014:31203.6] "GET HTTP/1.1 /query"
38.151.1.182 [31/May/2014:31204.1] "GET HTTP/1.1 /OK"
19.38.218.11 [31/May/2014:31205.7] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.0] "GET HTTP/1.1 /query"
95.39.21.28 [31/May/2014:31206.8] "GET HTTP/1.1 /OK"
210.82.199.247 [31/May/2014:31208.3] "GET HTTP/1.1 /query"
210.82.199.247 [31/May/2014:31208.8] "GET HTTP/1.1 /problem"
19.38.218.11 [31/May/2014:31208.9] "GET HTTP/1.1 /query"
19.38.218.11 [31/May/2014:31209.7] "GET HTTP/1.1 /OK"
```
Perfume motivation

Console logs

- rich, low-level descriptions of system behavior
- massive, difficult to interpret

**Our solution:** summarize log with a resource-aware model

Previous model-inference work

- Biermann IEEETC1972
- Walkinshaw ASE2008
- Lorenzoli ICSE2008
- Beschastnikh FSE2011
- Fahland ASE2013
- Ghezzi ICSE2014
Perfume motivation

Console logs
- rich, low-level descriptions of system behavior
- massive, difficult to interpret

Our solution: summarize log with a resource-aware model

Previous model-inference work
- Biermann IEEETC1972
- Walkinshaw ASE2008
- Lorenzoli ICSE2008
- Beschastnikh FSE2011
- Fahland ASE2013
- Ghezzi ICSE2014

Our contribution:
Improve inference precision and utility with resource information
Key insight

Observations

1) behavior depends on resource usage
   • caching   • timeouts   • network protocols

2) most runtime logs already contain resource usage data
   • time   • bytes transferred   • power/memory/CPU usage

Key insight: **Resource-aware inference ⇒ better models**

Key challenges: model precision, conciseness, usability
Addressing inference challenges

1) Precise models

2) Concise models

3) Usable models
Addressing inference challenges

1) Precise models
   - model must satisfy **observed** resource-based properties

2) Concise models
   - start with **minimal** model
   - expand model only to **satisfy observed properties**
   - **minimize** resulting model

3) Usable models
   - interactive and query-able models
   - an easy-to-use, cloud-based web interface
Resource-based properties

• To ensure model precision:
  1) mine observed properties from logged executions
     • approximate system's true properties
  2) ensure no model path can violate observed properties

• e.g., broadband always precedes problem in ≥ 8.7s
Evaluation summary

1) RQ1: Do resource-aware models increase system understanding?
   - user study: **Perfume** users 4-12% more correct, 5-12% faster

2) RQ2: Can **Perfume** model real network protocol behavior?
   - TCP case study: model revealed real TCP behavior

3) RQ3: Can **Perfume** model large-scale website behavior?
   - real estate website case study

**Perfume** model usability:
http://bestchai.bitbucket.org/perfume
Small-scale user study

RQ1: Do resource-aware models increase system understanding?

• 13 users

• asked questions about 3 systems

• measured:
  • response correctness
  • response speed

Compare Perfume to previous approaches
Small-scale user study

RQ1: Do resource-aware models increase system understanding?

• 13 users were shown:
  1) a console log,
  2) a console log + Synoptic\(^1\) model, or
  3) a console log + Perfume model

• asked questions about 3 systems

• measured:
  • response correctness
  • response speed

Compare Perfume to previous approaches

\(^1\) Beschastnikh FSE2011
Small-scale user study

RQ1: Do resource-aware models increase system understanding?

- 13 users were shown:
  1) a console log,
  2) a console log + Synoptic\(^1\) model, or
  3) a console log + Perfume model
- asked questions about 3 systems
- measured:
  - response correctness
  - response speed

Compare **Perfume** to previous approaches

\(^1\) Beschastnikh FSE2011
User study results

- Participants using Perfume:
  - on average, answered more questions correctly

<table>
<thead>
<tr>
<th>Questions answered correctly</th>
<th>Log</th>
<th>Log+Synoptic</th>
<th>Log+Perfume</th>
</tr>
</thead>
<tbody>
<tr>
<td>RADIUS$^1$ protocol</td>
<td>62.5%</td>
<td>70.8%</td>
<td>81.7%</td>
</tr>
<tr>
<td>Caching web browser</td>
<td>77.8%</td>
<td><strong>80.6%</strong></td>
<td>60.0%</td>
</tr>
<tr>
<td>Connection tester</td>
<td>78.3%</td>
<td>83.3%</td>
<td><strong>97.2%</strong></td>
</tr>
<tr>
<td>Average</td>
<td>72.4%</td>
<td>78.0%</td>
<td><strong>81.4%</strong></td>
</tr>
</tbody>
</table>

$^1$ Rigney, RFC 2865, 2000
User study results

- Participants using Perfume:
  - on average, answered more questions correctly
  - on average, answered faster

<table>
<thead>
<tr>
<th>Questions answered correctly</th>
<th>Log</th>
<th>Log+Synoptic</th>
<th>Log+Perfume</th>
</tr>
</thead>
<tbody>
<tr>
<td>RADIUS protocol</td>
<td>62.5%</td>
<td>70.8%</td>
<td><strong>81.7%</strong></td>
</tr>
<tr>
<td>Caching web browser</td>
<td>77.8%</td>
<td><strong>80.6%</strong></td>
<td>60.0%</td>
</tr>
<tr>
<td>Connection tester</td>
<td>78.3%</td>
<td>83.3%</td>
<td><strong>97.2%</strong></td>
</tr>
<tr>
<td>Average</td>
<td>72.4%</td>
<td>78.0%</td>
<td><strong>81.4%</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time to answer questions for one system (minutes)</th>
<th>Log</th>
<th>Log+Synoptic</th>
<th>Log+Perfume</th>
</tr>
</thead>
<tbody>
<tr>
<td>RADIUS protocol</td>
<td>11.5</td>
<td>13.0</td>
<td><strong>9.1</strong></td>
</tr>
<tr>
<td>Caching web browser</td>
<td>21.5</td>
<td><strong>8.7</strong></td>
<td>10.5</td>
</tr>
<tr>
<td>Connection tester</td>
<td><strong>7.3</strong></td>
<td>13.5</td>
<td>16.6</td>
</tr>
<tr>
<td>Average</td>
<td>13.0</td>
<td>11.9</td>
<td><strong>11.4</strong></td>
</tr>
</tbody>
</table>

On average, Perfume models supported 4% more correct, 5% faster comprehension

---

1 Rigney, RFC 2865, 2000
RQ2: Can Perfume model real network protocol behavior?

- **Perfume**-inferred model of web browser TCP traffic illustrated:
  1) Timeouts
  2) Buffer pushes (push packets)
  3) Connection resets (reset packets)

- server events shaded
- slow ack ⇒ server reset

- Actual TCP properties visible without prior TCP knowledge
Contributions

- Defined resource-based system properties
- Developed a resource-aware inference algorithm
- Developed cloud-based user interface
- Evaluated Perfume in a small user study
  - 12% more correct, 12% faster than logs
  - 4% more correct, 5% faster than Synoptic
- Evaluated Perfume in two case studies
  - Revealed real TCP behavior
  - Revealed website navigation bug

http://cs.umass.edu/~ohmann/perfume
Backup: Approach

**Input:** console log

Parse log → Initial, minimal model

Refine until properties satisfied → Coarsen model

**Output:** Resource-aware model
Backup: behavioral model

- “What type of clients experience problems?”
  - Behavioral model
    - problems only in broadband clients
Backup: user study models

Caching web browser

Connection Tester

RADIUS protocol

Perfume: resource-aware model inference

6/3/14
Problem

Debugging

- consumes 50% of programmers' time\(^1\)
- costs $300 billion annually\(^1\)
- bugs often caused by misunderstanding system behavior

Console logs

- rich but low-level description of system behavior
- massive, difficult to interpret
- our solution: summarize with a model

Goals

1) Predictive models
   - predict unobserved executions

2) Precise models
   - predicted executions are likely to be possible

3) Concise models
   - human-readable
   - generalizing
Addressing goals

1) Predictive models
   • allow observed executions to form unobserved paths

2) Precise models
   • resourced-based properties

3) Concise models
   • combine model states unless property is violated