Slimium: Debloating the Chromium Browser with Feature Subsetting

CHENXIONG QIAN, HYUNGJOON (KEVIN) KOO, CHANGSEOK OH, TAESOO KIM, WENKE LEE
Background

• Chromium dominates Web browser market share.

• Ever-increasing Features:
  • 2300+ Html/Javascript properties
  • 700+ CSS properties
  • Hundreds of experimental features
Problem

- Not all features are used commonly.
- Attack surface is increasing.
Slimium

- Remove code of unused features.
- Given a set of websites, generate a slim version of Chromium.
Overview

Map code to features

Remove unused features

Profile websites

Entire Code

F₂

F₁
Feature Code Mapping

- **Challenge**
  - Large-scale & Complex
  - Code generation during compiling
Feature Code Mapping

• **Approach**
  ➢ Manual Analysis
  • Investigate source code and documents.
  • Create an initial mapping between features and source code (i.e., files).
Feature Code Mapping

• Approach
  ➢ Manual Analysis
    • Investigate source code and documents.
    • Create an initial mapping between features and source code (i.e., files).
  ➢ Static Analysis
    • Build the call graph
    • Compute a relation vector \( R = (r_c, r_s) \)
      • \( r_c \) -- Call Invocation Relation (0 ~ 1)
      • \( r_s \) -- File Name Similarity (0 ~ 1)
    • If \( r_c \) and \( r_s \) are greater than the thresholds, add the file to the feature’s mapping.
Webpage Profiling

- Challenge & Approach
  - Nondeterministic Code
    - Keep profiling until stable.
  - Performance
    - Adopt AFL’s approach.

Identified nondeterministic code from visiting the top 1000 Alexa websites
Removing Unused Features

• Keep nondeterministic code.

• Calculate a feature’s code coverage based on profiling results.
  • If the code coverage is greater than the threshold (i.e., $T$), keep the feature.
  • Otherwise, remove the feature’s unexecuted code.

• Rewrite the binary to remove code.
## Evaluation

### Feature Code Mapping

<table>
<thead>
<tr>
<th>Class</th>
<th>Features (#)</th>
<th>Functions (#)</th>
<th>Function Size (KB)</th>
<th>CVEs (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTML5</td>
<td>6</td>
<td>8,103</td>
<td>1,721</td>
<td>15</td>
</tr>
<tr>
<td>JS API</td>
<td>100</td>
<td>71,082</td>
<td>17,204</td>
<td>57</td>
</tr>
<tr>
<td>Non-web</td>
<td>57</td>
<td>62,594</td>
<td>21,303</td>
<td>77</td>
</tr>
<tr>
<td>Wasm</td>
<td>1</td>
<td>1,189</td>
<td>869</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>164</strong></td>
<td><strong>142,968</strong></td>
<td><strong>41,097</strong></td>
<td><strong>153</strong></td>
</tr>
</tbody>
</table>

**Manual Analysis**

**Static Analysis**
## Evaluation

➢ Code Reduction & Security Benefits

- Visit 40 websites from 10 different groups.

<table>
<thead>
<tr>
<th>Category</th>
<th>Websites</th>
<th>User Activities</th>
<th>Code Reduction Size (MB)</th>
<th>Code Reduction Rate (%)</th>
<th>Number of Removed CVEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airline</td>
<td>aa, delta, spirit, united</td>
<td>Login; search a flight; make a payment; cancel the flight; logout.</td>
<td>24.17</td>
<td>53.8</td>
<td>97</td>
</tr>
<tr>
<td>Email</td>
<td>gmail, icloud, outlook, yahoo</td>
<td>Login; read/delete/reply/send emails (with attachments); open attachments; logout.</td>
<td>23.75</td>
<td>52.9</td>
<td>97</td>
</tr>
<tr>
<td>Financial</td>
<td>americanexpress, chase, discover, paypal</td>
<td>Login; check a statement; pay a bill; transfer money; logout.</td>
<td>23.45</td>
<td>52.2</td>
<td>91</td>
</tr>
<tr>
<td>News</td>
<td>cnn, cnbc, nytines, washingtonpost</td>
<td>Read breaking news; watch videos; search other news.</td>
<td>24.19</td>
<td>53.9</td>
<td>98</td>
</tr>
<tr>
<td>Remote Working</td>
<td>bluejeans, slack, webex, zoom</td>
<td>Schedule a meeting; video/audio chat; share a screen; end the meeting.</td>
<td>18.57</td>
<td>41.4</td>
<td>81</td>
</tr>
<tr>
<td>Shopping</td>
<td>amazon, costco, ebay, walmart</td>
<td>Login; track a previous order; look for a product; add it to the cart; checkout; logout.</td>
<td>24.33</td>
<td>54.2</td>
<td>98</td>
</tr>
<tr>
<td>Social Media</td>
<td>instagram, facebook, twitter, whats app</td>
<td>Login; follow/unfollow a person; write a post and comment; like a post; send a message; logout.</td>
<td>23.30</td>
<td>51.9</td>
<td>93</td>
</tr>
<tr>
<td>Sports</td>
<td>bleacherreport, espn, nfl, nba</td>
<td>Check news, schedules, stats, and players.</td>
<td>24.39</td>
<td>54.3</td>
<td>98</td>
</tr>
<tr>
<td>Travel</td>
<td>booking, expedia, priceline, tripadvisor</td>
<td>Login; search hotels; reserve a room; make a payment; logout.</td>
<td>24.16</td>
<td>53.8</td>
<td>97</td>
</tr>
<tr>
<td>Video</td>
<td>amazon, disneyplus, netflix, youtube</td>
<td>Search a keyword; play a video (forward/pause/resume); switch screen modes (normal/theatre/full); adjust a volume</td>
<td>24.18</td>
<td>53.9</td>
<td>91</td>
</tr>
<tr>
<td>All</td>
<td>–</td>
<td>–</td>
<td>17.43</td>
<td>38.8%</td>
<td>73</td>
</tr>
</tbody>
</table>
Evaluation

➢ Feature Usages
Related Works

• Snyder’s work at CCS’17
  • “Most Websites Don’t Need to Vibrate: A Cost-Benefit Approach to Improving Browser Security”
  • Scope
  • API blocking vs Code removing
Limitations

• Rely on manual analysis.
• Not 100% guaranteed stable.
Questions?