Synchronizing software variants: A two-dimensional approach

Christoph König, Kamil Rosiak, Lukas Linsbauer, Ina Schaefer
Software Product Lines

Domain Engineering

Feature model
- Coffee machine
  - Coffee
    - Latte macchiato
    - Espresso
    - Cappuccino
    - Filter coffee
  - Temperature control

Artifact selection

Feature selection
- Coffee machine
- Coffee
  - Latte macchiato
  - Espresso
  - Cappuccino
  - Filter coffee
  - Temperature control

Application Engineering

Generator

Product
Horizontal cloning

Basic
+ Filter coffee
+ Espresso
+ Temperature control

Pro 100
+ Filter coffee
+ Espresso
+ Cappuccino
+ Temperature control

Pro 250
+ Filter coffee
+ Espresso
+ Latte Macchiato
+ Cappuccino
+ Temperature control
Vertical cloning

1. Heat water
2. Heat milk
3. Grind coffee
4. Brew espresso
5. Foam milk
6. ...

Pro 250

1. Heat water
2. Heat milk
3. Grind coffee
4. Foam milk
5. Brew espresso
6. ...

Cappuccino

Latte Macchiato
Synchronizing software variants
Overview

1) Intra-clone detection

2) Inter-clone detection

3a) Revision-based synchronization

3b) Model-based synchronization

4) Clone analysis

Revision control system

Family model

Clustering

Folder

Class

Method

Body

While

Method call

Return

Variable var1

Variable constant 0

Argument 1: var1

Argument 2: var2

Software variants
Intra-clone detection

1) AST transformation and extraction

2) Clustering and matching

3) Derivation of multiset model
Inter-clone detection

1) AST transformation

2) Matching

3) Derivation of family model
Cloning model

Multiset model

Latte Macchiato

Clone model

Family model
RQ1: How relevant is two-dimensional clone synchronization?

RQ2: How correctly can we synchronize intra- and inter-clones?

RQ3: How does synchronization scale in time and space with increasing number of variants?
Evaluation

■ Experimental subject: Five revisions of MobileMedia SPL¹

■ Metrics

■ Co-occurrence: Relationship between two inter-clones c₁ and c₂, where c₁ has at least one intra-clone c₁a

■ Subsequent revisions Rᵢ and Rⱼ (i < j)
  ■ True positives (TP): Clones manually synchronized in Rⱼ and two-dimensionally in Rᵢ
  ■ False negatives (FN): Clones manually synchronized in Rⱼ but not two-dimensionally in Rᵢ
  ■ False positives (FP): Clones not manually synchronized in Rⱼ but two-dimensionally in Rᵢ

■ Precision (PR): \( \frac{TP}{TP + FP} \)

■ Recall (RC): \( \frac{TP}{TP + FN} \)

¹ Eduardo Figueiredo, Nelio Cacho, Claudio Sant’Anna, Mario Monteiro, Uira Kulesza, Alessandro Garcia, Sergio Soares, Fabiano Ferraz, Safoua Khan, Fernando Castor Filho, and Francisco Dantas. Evolving software product lines with aspects
## RQ1: Relevance

<table>
<thead>
<tr>
<th>Revision</th>
<th>#Variants</th>
<th>$T_I$</th>
<th>$T_{II}$</th>
<th>$T_{III}$</th>
<th>$T_I$</th>
<th>$T_{II}$</th>
<th>$T_{III}$</th>
<th>(\Phi) Intra-group size</th>
<th>(\Phi) Inter-group size</th>
<th>#Co-occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$R_2$</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>86</td>
<td>0</td>
<td>5</td>
<td>4.00</td>
<td>2.00</td>
<td>24</td>
</tr>
<tr>
<td>$R_3$</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>88</td>
<td>2</td>
<td>6</td>
<td>4.00</td>
<td>3.92</td>
<td>48</td>
</tr>
<tr>
<td>$R_4$</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>52</td>
<td>104</td>
<td>2</td>
<td>7</td>
<td>3.19</td>
<td>7.61</td>
<td>172</td>
</tr>
<tr>
<td>$R_5$</td>
<td>16</td>
<td>16</td>
<td>44</td>
<td>100</td>
<td>134</td>
<td>2</td>
<td>10</td>
<td>3.30</td>
<td>13.62</td>
<td>424</td>
</tr>
</tbody>
</table>

The number of intra- and inter-clones per clone type, the average group size and the number of co-occurrences in each revision of MobileMedia.
RQ2: Correctness

<table>
<thead>
<tr>
<th></th>
<th>Intra-clones</th>
<th></th>
<th></th>
<th>Inter-clones</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TP</td>
<td>FP</td>
<td>FN</td>
<td>PR</td>
<td>RC</td>
<td>TP</td>
</tr>
<tr>
<td>$R_1$</td>
<td>92</td>
<td>3</td>
<td>6</td>
<td>0.97</td>
<td>0.93</td>
<td>94</td>
</tr>
<tr>
<td>$R_2$</td>
<td>146</td>
<td>5</td>
<td>26</td>
<td>0.97</td>
<td>0.85</td>
<td>211</td>
</tr>
<tr>
<td>$R_3$</td>
<td>419</td>
<td>31</td>
<td>153</td>
<td>0.93</td>
<td>0.73</td>
<td>1308</td>
</tr>
<tr>
<td>$R_4$</td>
<td>404</td>
<td>111</td>
<td>28</td>
<td>0.79</td>
<td>0.94</td>
<td>2040</td>
</tr>
</tbody>
</table>

Precision and recall of variant synchronization between subsequent revisions of MobileMedia
Threats to validity

- Bias of matching algorithm
- MobileMedia as synthetic clone-and-own experimental subject
- Java source code only
Summary

- Approach: Synchronizing intra- and inter-clones simultaneously
- High precision and recall in change propagation
- Relevance of two-dimensional software variant synchronization