Evolution of Near-Miss Clones

Saman Bazrafshan

12th IEEE International Working Conference on Source Code Analysis and Manipulation

September 23th 2012
String nameParts[] = author[1].split(" ");
int i = 0;
for (i = 0; i < nameParts.length; i++) {
    c = nameParts[i].charAt(0);
    sb.append(c + ".");
}
String nameParts[] = author[1].split(" ");
int i = 0;
for (i = 0; i < nameParts.length; i++) {
    c = nameParts[i].charAt(0);
    sb.append(c + ".");
}

String nameParts[] = author[1].split(" ");
int y = 10;
for (i = 0; i < nameParts.length; i++) {
    c = nameParts[i].charAt(0);
    sb.append(c + ".");
}

String nameParts[] = author[1].split(" ");
String s = in.readLine();
for (i = 0; i < nameParts.length; i++) {
    c = nameParts[i].charAt(0);
    sb.append(c + ".");
}
“Irregularities in clone evolution can help to differentiate the potentially harmful from the less interesting clones.”
Göde's study

“Evolution of Type-1 Clones” (SCAM 2009)

“Irregularities in clone evolution can help to differentiate the potentially harmful from the less interesting clones.”

What about near-miss clones?

• Analyzed 7 subject systems
• 3 programming languages
• 4 year period (daily snapshots)
1. Which clone type has more impact on the clone ratio?

The bar chart compares the mean clone ratio of different software tools. The x-axis represents the software tools: ArgoUML, Ant, JabRef, httpd, Nautilus, FileZilla, and Umbrello. The y-axis shows the mean clone ratio in percentage. The blue bars represent Identical Cloned Code, and the red bars represent Near-Miss Cloned Code.
2. Which clone type is more persistent?

Total Number Versions: 1212 1181 660 1222 986 936 1401
3. How frequent do identical clones change into near-miss clones and vice-versa?

<table>
<thead>
<tr>
<th></th>
<th>ArgoUML</th>
<th>Ant</th>
<th>JabRef</th>
<th>httpd</th>
<th>Nautilus</th>
<th>FileZilla</th>
<th>Umbrello</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Near-Miss to Identical</strong></td>
<td>44</td>
<td>3</td>
<td>8</td>
<td>18</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td><strong>Identical to Near-Miss</strong></td>
<td>19</td>
<td>8</td>
<td>13</td>
<td>29</td>
<td>4</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>
3. How frequent do identical clones change into near-miss clones and vice-versa?

<table>
<thead>
<tr>
<th></th>
<th>ArgoUML</th>
<th>Ant</th>
<th>JabRef</th>
<th>httpd</th>
<th>Nautilus</th>
<th>FileZilla</th>
<th>Umbrello</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near-Miss to Identical</td>
<td>44</td>
<td>3</td>
<td>8</td>
<td>18</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Identical to Near-Miss</td>
<td>19</td>
<td>8</td>
<td>13</td>
<td>29</td>
<td>4</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

```java
if (nameParts.length > 0)
    nameParts[i] = sb + ".";
if (nameParts.length > 0) {
    nameParts[i] = sb + ".";
}
if (nameParts.length > 0) {
    nameParts[i] = sb + ".";
}
if (nameParts.length > 0) {
    nameParts[i] = sb + ".";
}
```
4. Which clone type is changed more frequently (consistently / inconsistently)?
Questions?