An Empirical Study of Using Source Code Snippets from Stack Overflow



Rabe Abdalkareem, Emad Shihab and Juergen Rilling rab_abdu@encs.concordia.ca, eshihab@cse.concordia.ca, juergen.rilling@concordia.ca Department of Computer Science and Software Engineering, Concordia University

Objectives

Study the impact of reusing source code from Stack Overflow (SO).

- Investigate how frequent the code has been reused from Stack Overflow.
- When was the reused code introduced in the project lifetime?
- Who modified the code after its introduction to the project?
- Defect-proneness of the Stack Overflow code.

Motivating Example



Description of a commit that reuse the code snippets

SHA1 ID: a6e3669c0101455e827e65c5c191fe7a0217d8bb Author: Dan Roundhill <dan@automattic.com> 2013-08-19 15:13:53

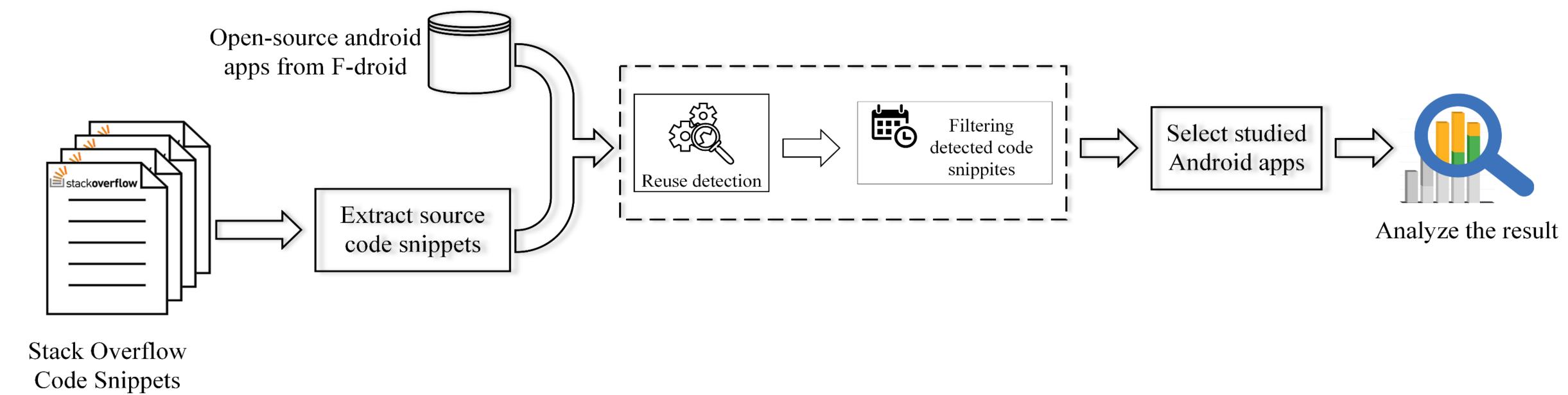
Message: ListView.getCheckedItemCount() is a v11 or higher api. Added solution from http://stackoverflow.com/questions/12330660/whats-the-

equivilent-of-getcheckeditemcount-for-api-level-11 that loops through the list to get the checked item count on older devices.

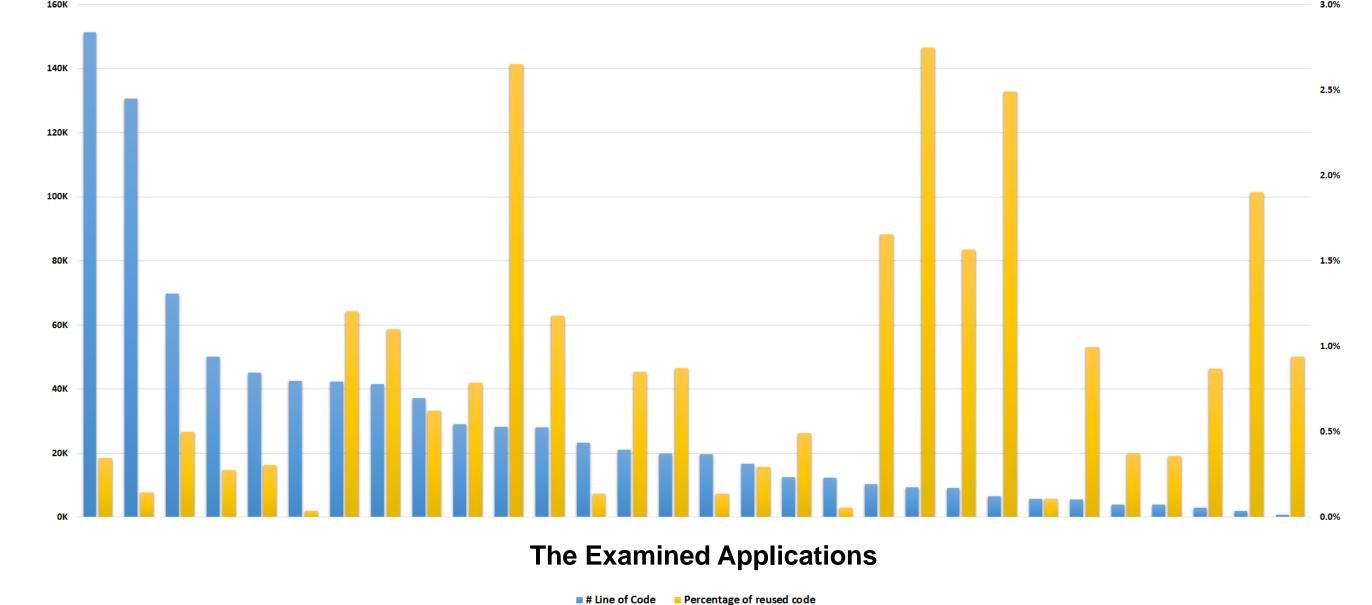
@@-392,4 +393,16 @@ public class SelectCategoriesActivity

```
extends SherlockListActivity {
   private int getCheckedItemCount(ListView listView)
      if (Build.VERSION.SDK_INT >= 11) return
listView.getCheckedItemCount();
      else
        int count = 0;
        for (int i = listView.getCount() - 1; i >= 0; i--)
          if (listView.isItemChecked(i)) count++;
        return count;
```

Approach

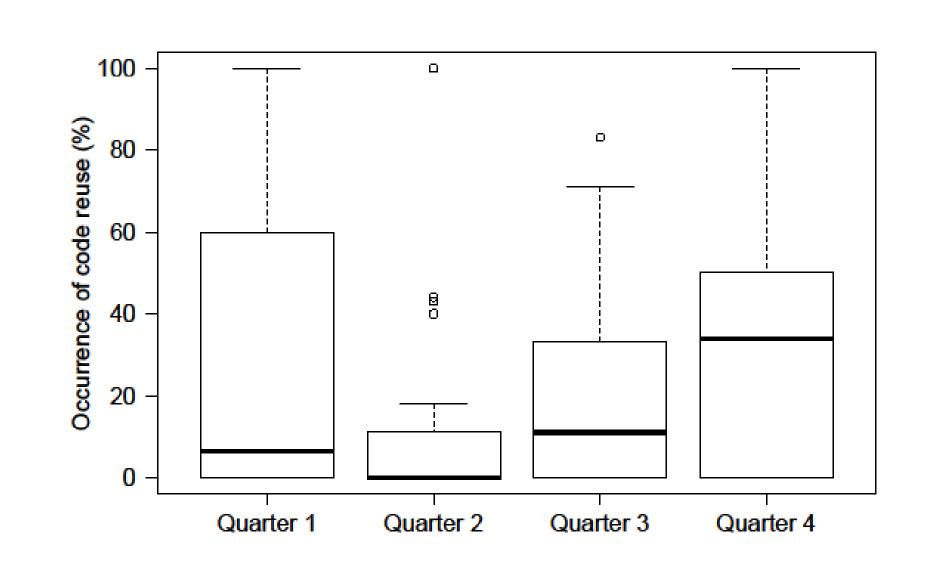


RQ1: Amount of Reused Code



0.87% of the examined apps' source code comes from Stack Overflow. Additionally, small apps have higher percentage of reused code.

RQ2: Developer Experience

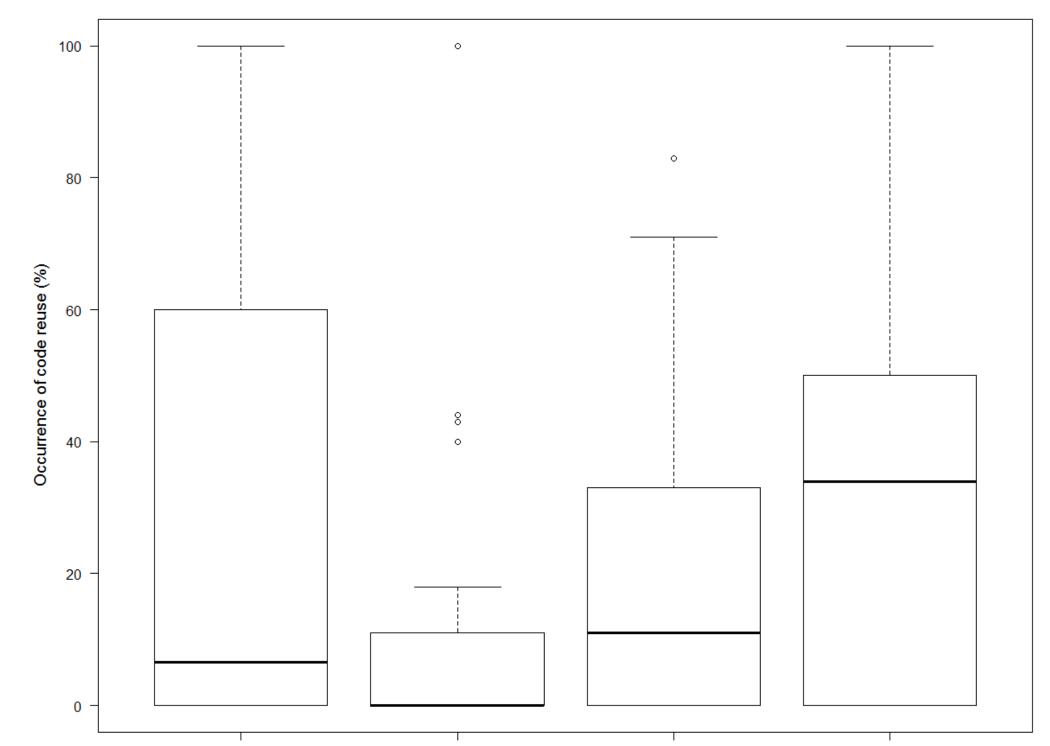


More experienced developers reuse Stack Overflow code in small apps, and less experienced developers reuse SO code in larger teams/apps.

RQ3: When is the Code Reused?

App	QTR 1	QTR 2	QTR 3	QTR 4
Web Opac	4%	4%	17%	74%
AdAway	0%	100%	0%	0%
ForPDA	83%	5%	12%	0%
WordPress	16%	16%	39%	29%

. . .



Percentage of reused code in all the projects per quarter

Software applications tend to reuse code form Stack Overflow later and early on in the project's life

Future Research

- Why did the developer introduce SO code?
- Who modified the code after its introduction to the project?
- What is the change- and defect-proneness of the SO code vs. non-SO code?