Predicting Bug-Fixing Time using Bug Change History

Alessandro Murgia, Javier Pérez, Serge Demeyer
Coen De Roover, Christophe Scholliers, Angela Lozano, Viviane Jonckers

Benevol 2013
Introduction

- Bug fixing time \(\rightarrow\) allocate resources

- Fixing time prediction:
  Today: snapshot-based
  CHAQ: snapshot-based + Change History
Goals

Identify:

- features useful for predicting bug fixing time
- industry-requirements for useful estimator
Bug 12345 - [DOGFOOD] Unable to Forward a message received as an Inline

**Status:** VERIFIED FIXED  
**Whiteboard:** [PDT+][PR1]  
**Keywords:**  
**Product:** MailNews Core (show info)  
**Component:** Backend (show info)  
**Version:** Trunk  
**Platform:** x86 Windows NT  
**Importance:** P1 normal (vote)  
**Target Milestone:** M11  
**Assigned To:** jefft  
**QA Contact:** Ichiang  
**URL:**  
**Depends on:** #5069  
**Blocks:** #1094 +7926  

Show dependency tree / graph

**Reported:** 1999-08-23 16:34 PDT by marina  
**Modified:** 2013-11-15 06:54 PST (History)  
**CC List:** 8 users (show)

**See Also:**

**Crash Signature:**

**Tracking Flags:**
Bug 12345 - [DOGFOOD] Unable to Forward a message received as an Inline

Status: VERIFIED FIXED
Whiteboard: [PDT+][PR1]
Keywords:

Product: MailNews Core (show info)
Component: Backend (show info)
Version: Trunk
Platform: x86 Windows NT
Importance: P1 normal (vote)
Target Milestone: M11
Assigned To: jefft
QA Contact: Ichiang

URL:

Depends on: #5069
Blocks: #1091 #7926
Show dependency tree / graph

Reported: 1999-08-23 16:34 PDT by marina
Modified: 2013-11-15 06:54 PST (History)
CC List: 8 users (show)

See Also:

Crash Signature:
Tracking Flags:
## Change History 2/2

<table>
<thead>
<tr>
<th>Who</th>
<th>When</th>
<th>What</th>
<th>Removed</th>
<th>Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>jeff</td>
<td>1999-08-29 10:43:15 PDT</td>
<td>Status</td>
<td>NEW</td>
<td>ASSIGNED</td>
</tr>
<tr>
<td>ichiang</td>
<td>1999-09-20 15:58:39 PDT</td>
<td>Blocks</td>
<td></td>
<td>++1091</td>
</tr>
<tr>
<td>jeff</td>
<td>1999-10-29 17:59:57 PDT</td>
<td>Depends on</td>
<td></td>
<td>++5069</td>
</tr>
<tr>
<td>ichiang</td>
<td>1999-11-02 11:13:05 PST</td>
<td>Summary</td>
<td>Unable to Forward a message received as an inline</td>
<td>[DOGFOOD] Unable to Forward a message received as an Inline</td>
</tr>
<tr>
<td>chofmann</td>
<td>1999-11-04 06:05:18 PST</td>
<td>Blocks</td>
<td></td>
<td>++7976</td>
</tr>
<tr>
<td>jeff</td>
<td>1999-11-04 11:50:21 PST</td>
<td>Status</td>
<td>ASSIGNED</td>
<td>RESOLVED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resolution</td>
<td>---</td>
<td>FIXED</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008-07-31 04:30:00 PDT</td>
<td>Product</td>
<td>Core</td>
<td>MailNews Core</td>
</tr>
<tr>
<td>nobody</td>
<td>2010-04-18 10:08:26 PDT</td>
<td>CC</td>
<td></td>
<td>tymerkaev</td>
</tr>
<tr>
<td>tymerkaev</td>
<td>2013-11-15 06:45:30 PST</td>
<td>Depends on</td>
<td></td>
<td>93918+</td>
</tr>
<tr>
<td>ysakham</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experimental Setup

Dataset:
- Mozilla,
- Industrial Project

Size:
- 4000 bug reports

Performance:
- Accuracy
  (Precision & Recall)

Classifier:
- Input
  - Bug Closed, Resolved
- Algorithms
  - Decision Tree,
  - Multinomial Naive Bayse
  - K-nearest-neighbor
  - Support Vector Machine
- Output
  - Mozilla (16 Classes)
  - Industrial Proj. (14 Classes)
Dataset

Mozilla

IndustrialCase

# report

0 2 4 6 8 10 12 14 16

time to fix

0 2 4 6 8 10 12 14 16

time to fix

Benevol 2013

Thursday, December
Results
Results

Mozilla

IndustrialCase
Results

Mozilla

IndustrialCase
Top 30 features - Multinomial Naive Bayes

Mozilla

Project: 11960

IndustrialCase

Benevol 2013
Conclusion

Bug Fixing time prediction:
• Snapshot-based features (e.g. platform, title)
• Change-related features
  + number of re-assignment

Industry requirements
• 4 hours time slots (fast bug fixing)
• incremental time window (long bug fixing)
Set Up

Input

- description
- priority
- #reassignment

Output

- Machine Learning Classifier
- time to fix:
  - 0> 0-4 hours
  - 1> 4-8 hours
  - 2> 8-16 hours

Dataset

Mozilla

IndustrialCase

Top 30 features - Multinomial Naive Bayes

- `text:mozilla`
- `bug_severity: normal`
- `text:window`
- `op_sys: All`
- `text:bug`
- `nAssignee`
- `complexity...impact: null`
- `text:companyName`
- `nAssignee`
- `nAssignee`
- `text:crash`
- `project:11960`
- `text:crash`

Benevol 2013
Thanks
Any questions?

alessandro.murgia@uantwerpen.be
Welcome to IEEE CSMR-WCRE 2014 Software Evolution Week

Software Evolution Week joins The Working Conference on Reverse Engineering (WCRE), the premier research conference on the theory and practice of recovering information from existing software and systems, with The European Conference on Software Maintenance and Reengineering (CSMR), the premier European conference on the theory and practice of maintenance, reengineering and evolution of software systems.

The first joint meeting will promote discussion and interaction among researchers and practitioners about the development of maintainable systems, and the evolution, migration and reengineering of the existing ones. In addition explore innovative methods of extracting the many kinds of information that can be recovered from software, software engineering documents, and systems artifacts, and examines innovative ways of using this information in system renovation and program...