An Empirical Study on Factors Impacting Bug Fixing Time

Feng Zhang, Foutse Khomh, Ying Zou
and Ahmed E. Hassan

Queen's University
A Typical Process of Bug Fixing

New → Assign → Bug Fixing → Verified → Resolved
Bug Fixing Time

- Introduced
- New
- Resolved
Example: delay before change

Back to bug 162007

<table>
<thead>
<tr>
<th>Who</th>
<th>When</th>
<th>What</th>
<th>Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>mik.kersten</td>
<td>2006-10-23</td>
<td>Assignee</td>
<td>mik.kersten</td>
</tr>
<tr>
<td></td>
<td>17:57:16 EDT</td>
<td>Priority</td>
<td>P2</td>
</tr>
<tr>
<td>steffen.pingel</td>
<td>2011-05-21</td>
<td>Status</td>
<td>RESOLVED</td>
</tr>
<tr>
<td></td>
<td>14:19:36 EDT</td>
<td>Resolution</td>
<td>FIXED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assignee</td>
<td>steffen.pingel</td>
</tr>
</tbody>
</table>

Reported & Assigned: Oct-2006
Fixing started & Resolved: Mar-2011

almost 5 years
Example: delay after change

Reported & Assigned: September-2008
Fixing started & ended:
Resolved: February-2011

over 2 years
Intervals during Bug Fixing

- **DBR**: Delay Before Response
- **DBA**: Delay Before Assigned
- **DBC**: Delay Before Change
- **DBF**: Duration of Bug Fixing
- **DAC**: Delay After Change

New
First Response
Assign
Bug Fixing Start
Bug Fixing End
Resolved
Benefits of Studying Delays

Locate time-consuming steps
Understand factors affecting the delays

A Typical Process of Bug Fixing

New  Assign  Bug Fixing  Verified  Resolved

Improve the process of bug fixing
Subject Systems

<table>
<thead>
<tr>
<th>System</th>
<th>Bugs</th>
<th>Logs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mylyn</td>
<td>2,722</td>
<td>3,883</td>
</tr>
<tr>
<td>Eclipse Platform</td>
<td>606</td>
<td>793</td>
</tr>
<tr>
<td>PDE</td>
<td>524</td>
<td>638</td>
</tr>
</tbody>
</table>
Data Sources

- Bugzilla
  - Bug Reports
  - Mylyn Logs

- CVS
  - Static metrics
  - History metrics
  - Fine grained change types

Properties of bug
Change history of bug
Properties of file edits

Bug ID
Metrics from three Dimensions

- Title
- Summary
- Severity of Bug
- Operating System
- Comments

Properties of Bug Report

Properties of Source Code

Properties of Code Change

Size (LOC, NOM) Complexity (WMC)

Code Churn
Number of Changed Files
Fine Grained Change Types
Research Questions

*RQ1:* Do delays by developers exist during bug fixing process?

*RQ2:* Can we characterize delays incurred by developers before and after fixing bugs?

*RQ3:* What factors contribute to the delays most?
RQ1: Bug fixing time extracted from two data sources
RQ1: Intervals during bug fixing process
RQ2: Properties of Bug Report

Enhancement v.s. Defect

- Delay Before Change: 364
- Delay After Change: 56
- Delay Before Change: 186
- Delay After Change: 4

Low Severity v.s. High Severity

- Delay Before Change: 198
- Delay After Change: n.s
- Delay Before Change: 110
- Delay After Change: n.s
RQ2: Properties of Bug Report (cont')

Length of comment
- delay before change
- delay after change
RQ2: Properties of Source Code

Size of Source Code
- delay before change (sum)
- delay after change (avg/sum/max)

Complexity of Source Code
- delay before change (sum)
- delay after change (sum/max)
RQ2: Properties of Code Change

**Code Churn**
- Delay before change (avg/sum/max)
- Delay after change (avg/sum/max)

**Fine Grained Change Types**
- Delay before change
- Delay after change
RQ3: What factors contribute to the delays most?

Analysis Method: (Logistic Regression Model)

Predict the probability of an event’s occurrence.
Combines different factors.

Interpretation by Coefficient:
magnitude describes importance
sign describes direction

Events:
Long DBC : greater than median
Long DAC : greater than median
RQ3: Factors Impacting Delays Before Change

-0.5

-0.5

is bug with high severity

sum Code Churn

0.4

0.5

sum NOM

0.3

0.2

max length of comment

Coefficient
RQ3: Factors Impacting Delays After Change

- max length of comment: 0.7
- max WMC: 0.5
- sum WMC: -0.7
- is bug: -0.5
- is OS Linux: -0.5
- length of description: -0.3
- avg Code Churn: 0.3

Coefficient
Conclusions

Top Two Intervals in Bug Fixing Process

delay before and after change

Delay Before Change

level of severity, code churn (most influential factors)

Delay After Change

comments, complexity (most influential factors)