How Effective Developers Investigate Source Code: An Exploratory Study

Martin P. Robillard, Wesley Coelho, and Gail C. Murphy (IEEE)

Presentation
By
Naga Prasanthi Bobbili
Student ID - 101057215

Agenda

❑Idea

❑Methodology followed

❑Study & Data analysis

❑Observations

❑Validation

❑Discussion

Idea of the Research

✓ To understand the nature of program investigation behavior of a developer.

✓ Study of link between program investigation behavior and success at a software modification task.

✓ To prove that developer should understand the system's source code and its impact on the other modules.

Agenda

✓ Idea

❑ Methodology followed

❑ Study & Data analysis

❑ Observations

❑ Validation

❑ Discussion

Methodology

✓ Realism

    - The task given should be realistic

    - Challenging

    - Spends significant time

✓ Replication

    - Helps to contrast the difference between successful and unsuccessful subjects.

Agenda

- ✓ Idea
- ✓ Methodology followed
- ❑ Study & Data analysis
- ❑ Observations
- ❑ Validation
- ❑ Discussion

Study

Task

  - Autosave feature in jEdit text editor(Version 4.6)
  - Subjects to understand source code
  - Develop application to provide users to explicitly disable the Autosave feature.

Subject Selection

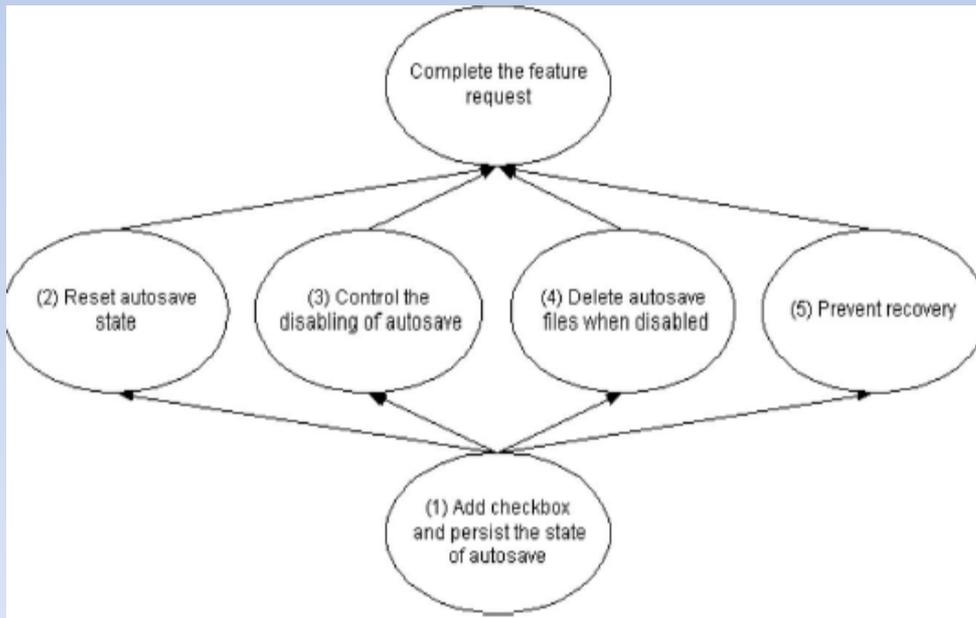 - Student, Co-ops with JAVA knowledge at BC university.

Study Setting

  - Eclipse Training phase
  - Program Investigation phase
  - Program Change phase

Data Analysis

❑ **Time Taken to Complete change**

Subject	1	2	3	4	5
Time(min)	125	62	72	114	120

❑**Quality of change**

Sub-Task/Subject	1	2	3	4	5
1-Check box	S	S	S	IE	S
2-State rest	NA	B	S	B	NA
3-Disabling	UW	S	S	S	UW
4-Deletion	UW	S	S	B	UW
5-Recovery	UW	S	S	S	NA

S - Success; NA - Not attempted
UW – Unworkable; B - Buggy
IE   - Inelegant

Data Analysis

☐Behavior of the developer

- Recorded videos to check behavior of subject
- Following were recorded for each event created (when a new method is see in the code editor)
    - Time
    - Method
    - Modification
    - Navigation of a method

→ Scrolling

→ Browsing

→ Cross Reference

→ Recall

→ Keyword Search
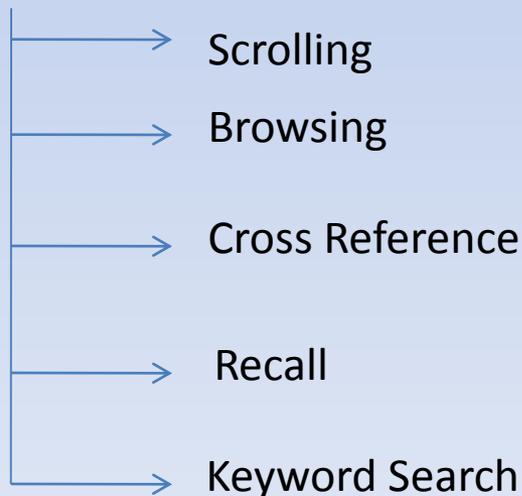
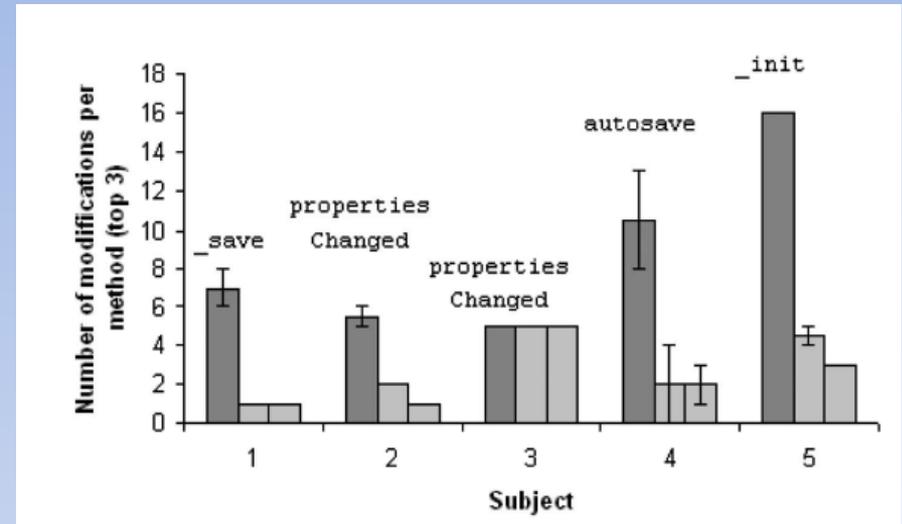| Transcript Excerpt		
0:26:51	B89	Keyword
0:27:55	A4	Recall
0:28:05	L3	Recall
0:28:09	L2	Scrolling

Example transcript of subject 1

Agenda

✓ Idea

✓ Methodology followed

✓ Study & Data analysis

❏ Observations

❏ Validation

❏ Discussion

Results (Observation 1)

Locality of changes!!!

✓Single method modification => Inadequate Investigation prior to modification

✓Except for subject 3, rest all relied mostly on unstructured techniques for task.

✓Unsuccessful subjects changed only single method more than successful subjects

✓ Subject 3 showed even pattern of modification (unlike other subjects).

Takeaway : Developer should have a broad spectrum of study while making a change to make a good code

Characterization of the Investigation Phase

Subject	Number of methods	Ratio of cross-reference and keyword events
1	8.5 ± 1.5	$2.0\% \pm 0.7$
2	27.5 ± 0.5	$23.3\% \pm 2.0$
3	34 ± 2	$31.7\% \pm 1.4$
4	38 ± 9	$30.8\% \pm 6.2$
5	17.5 ± 4.5	$10.7\% \pm 3.6$

Results(Observation 2)

Inattention/Blindness!!

✓Attention to details is very important

✓Unsuccessful subjects, either did not acknowledge or accidentally acknowledged the required code changes but did not actually change.

Takeaway: Search should be explicit and specific while doing the changes

Results(Observation 3)

Planning!!

✓ Detailed plan was made by successful subjects rather than unsuccessful subjects.

✓ Only subjects 2 & 3 made a proper plan and followed it.

✓ Proper plan ahead ensures:
- Assessment of the change made
- More focus on the program search
- eases the cognitive load on developer

Takeaway : Developers should use tools to make explicit plan for their designs before hand

Results(Observation 4)

Reinvestigation Frequency

✓ Unsuccessful subjects frequently revisits modified methods unlike successful subjects.

✓ Linear Traversal of the code vs Opportunistic approach.

✓ Video transcripts were used to calculate the mean of n number of reinvestigations happened on a particular method in a given time frame.

Takeaway: Developer should carefully assess the potential relevance of each method investigated. Should understand thoroughly before making a change to avoid repetition.

Results(Observation 5)

Structured Guided Searches!!

✓Structurally guided searches:
   - Keyword search
   - Cross-reference search

✓Developer should have a detailed knowledge of the implementation of a system rather than trial and error methods like scrolling.

Takeaway : Developer should minimize scrolling and practice methodical investigation methods.

Agenda

- ✓ Idea

- ✓ Methodology followed

- ✓ Study & Data analysis

- ✓ Observations

- ❑ Validation

- ❑ Discussion

Validation strategies

✓ **Construct validity** is tough in empirical studies

      - Evaluation of developer's solution.

      - Analysis of the behavior of the developers.

✓ **Internal Validity**(Possibilities)

     - Subject expertise in jEdit.

     - Experimental Setup

     - Evaluation of the best solution

✓ **External Validity**

     - Generalization of the study

Agenda

✓ Idea

✓ Methodology followed

✓ Study & Data analysis

✓ Observations

✓ Validation

❑ Discussion

Discussion

- How do you think the author could have better validated this work?

- How do we think we can use this work in an organization?

- Alternate best process to find a best solution for evaluating the developer's code.

- How do you think this research can be extended?