

A Replication and Reproduction

RoSE 6 ICSE 2020

A Replication Study: Where and When Should Defects be Re-Assigned

Tamer Abdou¹ and Ayse Bener²

¹ Github ID: @Tamer-Abdou

² Github ID:

{tamer.abdou, ayse.bener}@ryerson.ca

ABSTRACT

Links: A link to the **original** papers^{3 4}:

<https://ieeexplore.ieee.org/document/6227112>

<https://ieeexplore.ieee.org/document/6062117>

A link to the **replicated** paper⁵:

<https://ieeexplore.ieee.org/document/7592820>

What: The original studied the reopening and fixing of bug reports in the context of two versions of Microsoft Windows operating system project. The original study set to analyze multiple factors that have been seen to influence reopened bugs, and extend the results by analyzing the relationships between these factors and the impact on fixed-when-reopened and fixed bugs. The analysis demonstrated in the replication was based on data from a defect tracking database which aims to collect data from two open-source software (OSS) projects, namely Mozilla and Eclipse.

Why: One of the main values of a replication is the comparison of its results with the results of the original study. Through comparison the consistency between original study and the replication can be identified. The original study was conducted at a specific organization for a specific project (Microsoft operating system). We wanted to replicate the original study to test whether the findings are supported in the context of OSS projects.

How: We designed and conducted a controlled empirical study to evaluate the validity of some of the original study conclusions. Similar to the original quantitative study, the replication was divided into three sub research categories: (i) the probability that a bug will be reopened, (ii) The probability that a bug will be fixed after the bug has been reopened, and (iii) the probability that a bug will be fixed.

Where: In the original study, all variables were taken from the Windows bug database and the Microsoft employee personnel database. The database of the original study was based on two main data sets: i) The pre- and post-release bug reports for Windows Vista in July 2009. This means that data have been mined 2.5 years after Vista's release data. 2) Bug reports for Windows 7, representing the development period (3 years). The dataset used in the replication study comes from four products, mainly Mozilla Core (44.2%), Mozilla Firefox (41.6%), Mozilla Thunderbird (11.4%), and Mozilla Bugzilla (2.7%).

Discussion: On the basis of findings obtained from both studies, we made a number of general recommendations for improving the process of reassigning bugs. We analyzed and highlighted conclusions about the relationships between different factors in the maintenance process and re-assigned bugs. The original and replication studies each used a different dataset and combining results have been found useful to generalize evidences and draw conclusions.

³ T. Zimmermann, N. Nagappan, P. J. Guo, and B. Murphy, "Characterizing and Predicting Which Bugs Get Reopened," in Proceedings of the 34th International Conference on Software Engineering, ser. ICSE '12. Piscataway, NJ, USA: IEEE Press, 2012, pp. 1074–1083.

⁴ P. J. Guo, T. Zimmermann, N. Nagappan, and B. Murphy, "Characterizing and Predicting Which Bugs Get Fixed: An Empirical Study of Microsoft Windows," in Proceedings of the 32th International Conference on Software Engineering (ICSE). Association for Computing Machinery, Inc., may 2010.

⁵ Abdou, T., & Bener, A. (2016). A Replication Study: Where and When Should Defects be Re-Assigned. Software Engineering and Advanced Applications (SEAA), 2016 42th Euromicro Conference On, 368–371.