The Impact of Test Ownership and Team Structure on the Reliability and Effectiveness of Quality Test Runs

Kim Herzig, Nachiappan Nagappan

Microsoft Research
Cambridge UK, Redmond USA

kimh@microsoft.com
nachin@microsoft.com

Abstract: *This submission presents work submitted and accepted at the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement in 2014* [Hn2014]. Software testing is a crucial step in most software development processes. Testing software is a key component to manage and assess the risk of shipping quality products to customers. But testing is also an expensive process and changes to the system need to be tested thoroughly which may take time. Thus, the quality of a software product depends on the quality of its underlying testing process and on the effectiveness and reliability of individual test cases. In this paper, we investigate the impact of the organizational structure of test owners on the reliability and effectiveness of the corresponding test cases. Prior empirical research on organizational structure has focused only on developer activity. We expand the scope of empirical knowledge by assessing the impact of organizational structure on testing activities. We performed an empirical study on the Windows build verification test suites (BVT) and relate efficiency and reliability measures of each test run to the complexity and size of the organizational sub-structure that enclose all owners of test cases executed. Our results show, that organizational structure impacts both test efficiency and test execution reliability. We are also able to predict effectiveness and reliability with fairly high precision and recall values.

1 Talk Summary

Testing is part of a software engineer's daily development process. Changes applied to a system must be tested to ensure that the modified code base matches specifications and constraints, e.g. security, privacy, performance, or functional correctness. However, testing may slow down product development: running hundreds of tests on large and complex software systems can easily take hours and thus may impact development agility. Therefore, frequently executed tests should be of high quality. Ineffective or unreliable tests may use expensive resources ineffectively and slow down product development without adding essential benefit—or may even harm the product by letting code issues slip into the final product. In this study, we investigated the impact of

organizational structure on the effectiveness and reliability of test executions. Do test suites owned by a smaller part of the development organization perform better when compared to test suites whose owners are distributed over a wider range of the organizational structure? To answer such questions, we performed an empirical study on Microsoft Windows build verification (BVT) test suites. We related effectiveness and reliability measures of each test suite to the complexity and size of the organizational sub-structure that enclose all owners of test cases executed by the corresponding test suite.

Results from correlating individual metrics against test suite effectiveness and reliability show that the size and structure of the team owning a test suite can have significant impact both test quality measures. Test suites owned by a larger organizational subgroups with short communication paths tend to be more effective and more reliable. Not surprisingly, test suites owned by engineers that left the organization seem to be less effective.

To investigate how well metrics capturing organizational structure can be used to predict effectiveness and reliability of test suites we used actual classification models to identify test suites associated with above-median effectiveness and above-median reliability values. Using our organizational metrics, we were able to build classification models that identify the effectiveness of test suites with precision and recall values around 0.7. We further showed, that organizational structure metrics are excellent predictors (precision around 0.8 and recall around 0.9) for test suite reliability issues and should be considered as strong indicators.

The results of this study are aligned with results of previous Microsoft studies investigating the effect of code ownership on code quality showing that distributed code ownership can have severe impact on code quality. These results on production code and our results on test suites are surprisingly aligned and show the same trend: code and test ownership are important properties and should be considered as a driving factor for code and test issues.

The original published paper this talk is covering can be found on the publisher's website: http://dl.acm.org/citation.cfm?doid=2652524.2652535

References

[Hn2014] Herzig K. und Nagappan N. (2014): The Impact of Test Ownership and Team Structure on the Reliability and Effectiveness of Quality Test Runs. In: *Proceedings of the 8th International Symposium on Empirical Software Engineering and Measurement*. ACM. S. 2:1--2:10.