

# Paper Two Summary

Matthew Neal, Joseph Sankar, and Alexander Sobran

September 7, 2015

## Reference

Erik Arisholm, Lionel C. Briand, and Eivind B. Johannessen. 2010. A systematic and comprehensive investigation of methods to build and evaluate fault prediction models. *J. Syst. Softw.* 83, 1 (January 2010), 2-17. DOI=10.1016/j.jss.2009.06.055 <http://dx.doi.org/10.1016/j.jss.2009.06.055>

## Important Keywords

**Fault-proneness** The number of defects detected in a software-component. A measurement of the likelihood of software or software component error.

**Object-oriented (OO) code measures** Measurement of the structural properties of the source code.

**Delta measures** Measurement of the amount of change (churn) in a file between two successive releases.

**Process measures** Measures from a configuration management system. Includes developer experience, the number of developers that have edited a file, the number of faults in the previous releases, the number of line added, the number of lines removed.

## Feature Extraction

**Motivational statements** The authors mentioned that most previous studies only examined the structure of the code when constructing fault-proneness prediction models. In the few studies that did, they did not systematically measure the cost-effectiveness of the models with factors with different data collection costs. Very few studies have looked at the methods and criteria for evaluating the models. Finally, it is very difficult to draw general conclusions using confusion matrices since they do not provide any insight into determining the cost-effectiveness of the models under test.

**Future Work** The authors used the default parameters for the statistical techniques they employed in the study. They have stated that they hope to tweak the parameters and observe the results to see if there are optimizations that can be made while at the same time try to avoid overfitting. The authors also stated they are going to work on a large-scale evaluation of the costs and benefits of various prediction models in the COS project.

## Possible Improvements

- In the Introduction section, the authors reference a paper published in 1989 to cite a percentage of effort spent on testing. Is there a more recent paper the authors could have chosen to cite? If not, why quote a figure that is 26 years old and quite possibly outdated?