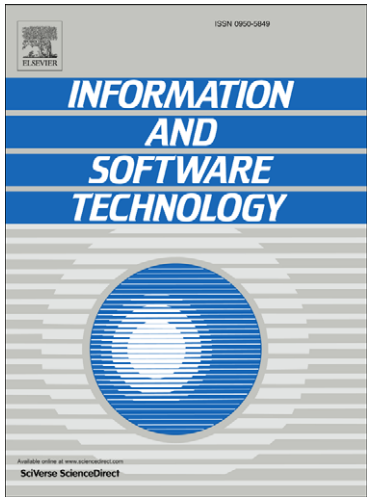

Software Fault Prediction Metrics:A Systematic Literature Review

Danijel Radjenović, Marjan Heričko, Richard Torkar, Aleš Živkovič

PII: S0950-5849(13)00042-6
DOI: http://dx.doi.org/10.1016/j.infsof.2013.02.009
Reference: INFSOF 5329

To appear in: *Information and Software Technology*

Received Date: 3 September 2012
Revised Date: 13 February 2013
Accepted Date: 14 February 2013

Please cite this article as: D. Radjenović, M. Heričko, R. Torkar, A. Živkovič, Software Fault Prediction Metrics:A Systematic Literature Review, *Information and Software Technology* (2013), doi: http://dx.doi.org/10.1016/j.infsof.2013.02.009

Software Fault Prediction Metrics: A Systematic Literature Review

Danijel Radjenović[a,b,*], Marjan Heričko[b], Richard Torkar[c,d], Aleš Živkovič[b]

[a]*Pro-bit programska oprema d.o.o., Stari trg 15, SI-3210 Slovenske Konjice, Slovenia*
[b]*University of Maribor, Faculty of Electrical Engineering and Computer Science, Smetanova ulica 17, SI-2000 Maribor, Slovenia*
[c]*Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden*
[d]*Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden*

Abstract

Context: Software metrics may be used in fault prediction models to improve software quality by predicting fault location.

Objective: This paper aims to identify software metrics and to assess their applicability in software fault prediction. We investigated the influence of context on metrics' selection and performance.

Method: This systematic literature review includes 106 papers published between 1991–2011. The selected papers are classified according to metrics and context properties.

Results: Object-oriented metrics (49%) were used nearly twice as often compared to traditional source code metrics (27%) or process metrics (24%). Chidamber and Kemerer's (CK) object-oriented metrics were most frequently used. According to the selected studies there are significant differences between the metrics used in fault prediction performance. Object-oriented and process metrics have been reported to be more successful in finding faults compared to traditional size and complexity metrics. Process metrics seem to be better at predicting post-release faults compared to any static code metrics.

Conclusion: More studies should be performed on large industrial software systems to find metrics more relevant for the industry and to answer the question as to which metrics should be used in a given context.

Keywords: Software metric, Software fault prediction, Systematic Literature Review

1. Introduction

Fault prediction models are used to improve software quality and to assist software inspection by locating possible faults[1]. Model performance is influenced by a modeling technique [9, 16, 20, 29, 30] and metrics [89, 138, 130, 88, 66]. The performance difference between modeling techniques appears to be moderate [38, 30, 60] and the choice of a modeling technique seems to have lesser impact on classification accuracy of a model than the choice of a metrics set [60]. To this end, we decided to investigate the metrics used in software fault prediction and to leave the modeling techniques aside.

In software fault prediction many software metrics have been proposed. The most frequently used ones are those of Abreu and Carapuca (MOOD metrics suite) [1, 51], Bansiya and Davis (QMOOD metrics suite) [3], Bieman and Kang [5], Briand et al. [70], Cartwright and Shepperd [74], Chidamber and Kemerer (CK metrics suite) [12, 13], Etzkorn et al. [17], Halstead [24], Henderson-Sellers [25], Hitz and Montazeri [26], Lee et al. [37], Li [41], Li and Henry [39, 40], Lorenz and Kidd [42], McCabe [44], Tegarden et al. [49]. Many of them have been validated only in a small number of studies. Some of them have been proposed but never used. Contradictory results across studies have often been reported. Even within a single study, different results have been obtained when different environments or methods have been used. Nevertheless, finding the appropriate set of metrics for a fault prediction model is still important, because of significant differences in metrics performance. This, however, can be a difficult task to accomplish when there is a large choice of metrics with no clear distinction regarding their usability.

The aim of this systematic literature review (SLR) is to depict current state-of-the-art software metrics in software fault prediction. We have searched in seven digital libraries, performed snowball sampling and consulted the authors of primary studies to identify 106 primary studies evaluating software metrics. The most commonly used metrics were identified and their fault prediction capabilities were assessed to answer the question of which metrics are appropriate for fault prediction. Ten properties were extracted from the primary studies to assess the context of metrics usage. We identified the most important studies according to study quality and industry relevance. We aggregated the dispersed and contradictive findings to reach new conclusions and to provide directions for practitioners and fu-

*Corresponding author. Tel.: +386 31 461 290
Email addresses: danijel.radjenovic@pro-bit.si (Danijel Radjenović), marjan.hericko@uni-mb.si (Marjan Heričko), richard.torkar@bth.se (Richard Torkar), ales.zivkovic@uni-mb.si (Aleš Živkovič)

[1]A correct service is delivered when the service implements the system function. A service failure is an event that occurs when the delivered service deviates from the correct/expected service. The deviation is called an error. The adjudged or hypothesized cause of an error is called a fault [2].