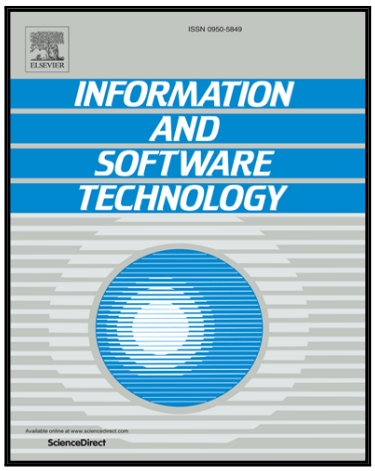
Accepted Manuscript



Reproducibility and Replicability of Software Defect Prediction Studies

Zaheed Mahmood, David Bowes, Tracy Hall, Peter C.R. Lane, Jean Petrić

PII: S0950-5849(17)30420-2
DOI: 10.1016/j.infsof.2018.02.003
Reference: INFSOF 5957

To appear in: *Information and Software Technology*

Received date: 29 May 2017
Revised date: 6 February 2018
Accepted date: 7 February 2018

Please cite this article as: Zaheed Mahmood, David Bowes, Tracy Hall, Peter C.R. Lane, Jean Petrić, Reproducibility and Replicability of Software Defect Prediction Studies, *Information and Software Technology* (2018), doi: 10.1016/j.infsof.2018.02.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Reproducibility and Replicability of Software Defect Prediction Studies

Zaheed Mahmood[a], David Bowes[a], Tracy Hall[b], Peter C. R. Lane[a], Jean Petrić[a]

[a] *University of Hertfordshire*
[b] *Brunel University London*

Abstract

Context: Replications are an important part of scientific disciplines. Replications test the credibility of original studies and can separate true results from those that are unreliable.

Objective: In this paper we investigate the replication of defect prediction studies and identify the characteristics of replicated studies. We further assess how defect prediction replications are performed and the consistency of replication findings.

Method: Our analysis is based on tracking the replication of 208 defect prediction studies identified by a highly cited Systematic Literature Review (SLR) [1]. We identify how often each of these 208 studies has been replicated and determine the type of replication carried out. We identify quality, citation counts, publication venue, impact factor, and data availability from all 208 SLR defect prediction papers to see if any of these factors are associated with the frequency with which they are replicated.

Results: Only 13 (6%) of the 208 studies are replicated. Replication seems related to original papers appearing in the Transactions of Software Engineering (TSE) journal. The number of citations an original paper had was also an indicator of replications. In addition, studies conducted using closed source data seems to have more replications than those based on open source data. Where a paper has been replicated, 11 (38%) out of 29 studies revealed different results to the original study.

Conclusion: Very few defect prediction studies are replicated. The lack of replication means that it remains unclear how reliable defect prediction is. We provide practical steps for improving the state of replication.

Keywords: Replication, Reproducibility, Software Defect Prediction

1. Introduction

Defect prediction is a very active area of research in software engineering. However the quality of defect prediction modelling is regularly criticised [2, 3]. Replications are

Email addresses: z.mahmood4@herts.ac.uk (Zaheed Mahmood), d.h.bowes@herts.ac.uk (David Bowes), tracy.hall@brunel.ac.uk (Tracy Hall), peter.lane@bcs.org.uk (Peter C. R. Lane), j.petric@herts.ac.uk (Jean Petrić)

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/6948046

Download Persian Version:

https://daneshyari.com/article/6948046

Daneshyari.com