Feedback-based integrated prediction: Defect prediction based on feedback from software testing process

Peng Xiao, Bin Liu, Shihai Wang

PII: S0164-1212(18)30106-7
DOI: 10.1016/j.jss.2018.05.029
Reference: JSS 10166

To appear in: *The Journal of Systems & Software*

Received date: 11 October 2017
Revised date: 13 May 2018
Accepted date: 18 May 2018

Please cite this article as: Peng Xiao, Bin Liu, Shihai Wang, Feedback-based integrated prediction: Defect prediction based on feedback from software testing process, *The Journal of Systems & Software* (2018), doi: 10.1016/j.jss.2018.05.029

Available online at www.sciencedirect.com

ScienceDirect

The Journal of System and Software 00 (2018) 1–??

Journal Logo

Feedback-based integrated prediction: Defect prediction based on feedback from software testing process

Peng Xiao[a,b], Bin Liu[a,b], Shihai Wang [a,b,*],

[a]School of Reliability and System Engineering, Beihang University, 100191, Beijing, China
[b]Science and Technology on Reliability and Environmental Engineering Laboratory, Beihang University, 100191, Beijing, China

Abstract

Test resource constraints is a common phenomenon in software testing. Using defect prediction to guide the resource allocation can significantly improve the efficiency and effectiveness of available test resources. However, traditional defect prediction (t-DP) is a static strategy, where the predictor cannot be dynamically adjusted during the software testing process (STP). This paper combines defect prediction with feedback control in STP and proposes a feedback-based defect prediction model, where the test results generated during STP is used as feedback information for on-line adjustment of predictor to optimize the prediction result. In addition, a novel approach called feedback-based integrated prediction (FIP) is proposed to improve the prediction accuracy, where a global predictor and a local predictor are employed to make an integrated prediction using the weight to adjust the effects of predictors at different test stages. A systematic experiment is conducted to investigate the performance of the FIP over 10 public data sets. Results show that FIP has better prediction efficiency and better robustness for external data than the t-DP, especially when the percentage of the test modules is 40%.

Keywords: Test resource constraints, Software testing, Defect prediction, Feedback control, Integrated prediction.

1. Introduction

Software testing is an indispensable activity in software development and is widely used in industry for quality assurance [1], since it can effectively eliminate defects before software delivery. However, software testing is also a time-consuming and expensive work, during which generally, software engineers need to spend 50%-60% development costs on software testing activities [2, 3]. In fact, the situation of test resource constraints is a common phenomenon during software testing process (STP) due to the limitation of deadline or development cost in software development [4]. In addition, for high-risk software parts (modules, components, files, or classes), more test resources are required to ensure their qualities. However, the distribution of software defects is generally non-equilibrium, that is, the majority of software defects are contained in a few software parts [5, 6, 7]. If the priorities of software parts can be sorted according to defect-proneness, available resources will be used in a more reasonable and effective way. As a result, the product's quality can be improved and the development cost can be reduced.

Defect prediction provides an effective way to reveal the likely defective parts before testing [8]. Its result is often used to guide the testing strategy. Therefore, the performance of defect prediction is a key factor in determining the efficiency of software testing. In order to improve the accuracy of defect prediction, many defect prediction algorithms have been proposed, such as defect prediction based on support vector machine [9] and Bayesian-based defect prediction [10], etc. However, sometimes the effectiveness of traditional defect prediction (t-DP) is far from being satisfying since the defect prediction and software testing are two independent processes. The defect prediction model trained by other similar project data cannot be dynamically adjusted according to the characteristics of the software under testing (SUT) during STP. Once the defect prediction model is established, the prediction results cannot be modified. If there are deviations between the characteristics of the training data and those of the SUT, the test efficiencywill be deteriorated.

To address this problem, this paper introduces the feedback control strategy into defect prediction, and proposes an novel feedback-based defect prediction strategy combining STP, so

*Corresponding author, Shihai Wang, a lecturer at Science and Technology on Reliability and Environmental Engineering Laboratory, and School of Reliability and Systems Engineering, Beihang University, Beijing, China. His research interests include machine learning, defect prediction, and software testing. E-mail: wangshihai@buaa.edu.cn.