

Chapter 23

Important Issues in Software Fault Prediction: A Road Map

Golnoush Abaei

University Technology Malaysia, Malaysia

Ali Selamat

University Technology Malaysia, Malaysia

ABSTRACT

Quality assurance tasks such as testing, verification and validation, fault tolerance, and fault prediction play a major role in software engineering activities. Fault prediction approaches are used when a software company needs to deliver a finished product while it has limited time and budget for testing it. In such cases, identifying and testing parts of the system that are more defect prone is reasonable. In fact, prediction models are mainly used for improving software quality and exploiting available resources. Software fault prediction is studied in this chapter based on different criteria that matters in this research field. Usually, there are certain issues that need to be taken care of such as different machine-learning techniques, artificial intelligence classifiers, variety of software metrics, distinctive performance evaluation metrics, and some statistical analysis. In this chapter, the authors present a roadmap for those researchers who are interested in working in this area. They illustrate problems along with objectives related to each mentioned criterion, which could assist researchers to build the finest software fault prediction model.

1. INTRODUCTION

As today's software grow rapidly in size and complexity, the prediction of software reliability plays a crucial role in a software development process (Catal, 2011). Software fault¹ is an error situation of the software system. The explicit and potential

violation of security policies at runtime causes these faults because of wrong specification and inappropriate development of configuration (Dowd, MC Donald, & Schuh, 2007). Early detection of fault-prone software components could enable verification experts to concentrate their time and resources on the problematic areas of the system

DOI: 10.4018/978-1-4666-6026-7.ch023

under development. In fact, when limited budget and time do not allow for complete testing of an entire system, software managers can use predictors for software quality to focus on testing parts of the system that seems to be defect prone. So the defect prone trouble spots can then be examined in more detail by model checking and intensive testing. In fault prediction approaches, previous reported faulty data with the help of distinct metrics could identify the fault-prone modules. In addition, important information about location, number of faults and distribution of defects are extracted to improve test efficiency and software quality of the next version of the program. In other word, fault prone modules can be identified prior to system test by using these models. Application of defect predictors in software developments helps the managers allocate the resources such as time and effort more efficiently. It would also be cost effectively to test some certain section of the code (Calikli, Tosun, Bener, & Celik, 2009).

Area of software fault prediction still poses many challenges and unfortunately, none of the techniques proposed within the last decade has achieved widespread applicability in the software industry due to several reasons including the lack of software tools to automate this prediction process, the unwillingness to collect the fault data, and other practical problems (Catal, Sevim, & Diri, 2009).

According to (Catal & Diri, 2009b), analyzing and predicting defects are needed for three main purposes, firstly, for assessing project progress and plan defect detection activities for the project manager. Secondly, for evaluating product quality in order to assess the finished product, and thirdly, for improving capability and assessing process performance for process management.

Any fault prediction model could be built based on variety of learning methods, distinctive software metrics, and different performance evaluation metrics; each of them has its own challenges. We should mention that the type of applications that software metrics extracted from them could also

be varied. In this chapter, we reviewed all related parameters to build a finest prediction model. Although literature review paper (Catal, 2011) and many systematic literature review papers (Arisholm, Briand, & Johannessen, 2010; Catal & Diri, 2009b; T. Hall, Beecham, Bowes, & Counsell, 2011; Radjenović, Heričko, Torkar, & Živković, 2013) are available, but to our knowledge none of them has classified the fault prediction issues and challenges as we have done here. Figure 1 shows the overall structure of building any software fault prediction model. All problems and objectives related to each part of this diagram are discussed and reviewed throughout this chapter.

Often, fault prediction approaches are most suitable in less critical domains such as software packages and telecommunication companies. Anyhow, in limited time and budget and specially for testing different versions of the software that have been completely tested before, using software prediction model would be a good choice. One of the practical examples could be found in Chapter 24 entitled “Building Defect Prediction Models in Practice”. Key questions and consequences arising when establishing defect prediction in a large software development project are answered in the mentioned chapter.

This chapter is organized as follows: section 2 presents background of the problem. Section 3 describes different fault prediction issues. Related works are presented in section 4. Section 5 presents fault prediction challenges and finally section 6 is conclusion.

2. BACKGROUND OF THE PROBLEM

Testing is the most popular method for defect detection in most of the software projects. However, when project's sizes grow in terms of both lines of code and effort spent, the task of testing becomes more difficult and costly. According to Bohem (1988), fault removal is considerably less costly when performed in the design phase

28 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/important-issues-in-software-fault-prediction/108634

Related Content

Requirements Engineering Process Improvement and Related Models

Badariah Solemon, Shamsul Sahibuddin and Abdul Azim Abd Ghani (2012). *Software Process Improvement and Management: Approaches and Tools for Practical Development* (pp. 18-33).

www.irma-international.org/chapter/requirements-engineering-process-improvement-related/61208

Research on Measuring Semantic Correlation Based on the Wikipedia Hyperlink Network

Feiyue Ye and Feng Zhang (2013). *International Journal of Software Innovation* (pp. 1-11).

www.irma-international.org/article/research-on-measuring-semantic-correlation-based-on-the-wikipedia-hyperlink-network/103278

A Novel Approach to Parkinson's Disease Progression Evaluation Using Convolutional Neural Networks

Mhamed Zineddine (2023). *International Journal of Software Innovation* (pp. 1-26).

www.irma-international.org/article/a-novel-approach-to-parkinsons-disease-progression-evaluation-using-convolutional-neural-networks/315655

A Formal Semantics of Kermeta

Moussa Amrani (2013). *Formal and Practical Aspects of Domain-Specific Languages: Recent Developments* (pp. 270-309).

www.irma-international.org/chapter/formal-semantics-kermeta/71823

Unsupervised Estimation of Facial Expression Intensity for Emotional Scene Retrieval in Lifelog Videos

Shota Sakaue, Hiroki Nomiya and Teruhisa Hochin (2018). *International Journal of Software Innovation* (pp. 30-45).

www.irma-international.org/article/unsupervised-estimation-of-facial-expression-intensity-for-emotional-scene-retrieval-in-lifelog-videos/210453