Duplicate Bug Report Detection

Tarun Chhabra

Computer Science

Department

North Carolina State University

Raleigh, NC, 27606, USA

Devika Desai

Computer Science Department

North Carolina State University

Raleigh, NC, 27606, USA

Sudipto Biswas

Computer Science Department

North Carolina State University

Raleigh, NC, 27606, USA

There are many users interacting with a system and reporting issues concerned with it in terms of a bug report. Bug reports are then used to guide software corrective maintenance activities so that it can result in more reliable software systems. Bug repositories are maintained as a collection of bug reports. Despite the advantages of a bug report system, it does cause some challenges. As bug reporting process is often ad-hoc, often the same bugs are reported by different users, or a different bug caused by the same potential software defect result in duplicate bug reports. A number of studies have attempted to address this issue by automating bug-report deduplication. The following review provides an overview of all the techniques used for duplicate bug detection before the concept of detection by information retrieval [1] and how the field has changed after that. A number of past studies have proposed a number of automated approaches to detect duplicate bug reports. We comment on how this approaches can be improved by integrating two or more techniques or considering even more possible factors for accurate detections.

INTRODUCTION

Due to system complexity and inadequate testing, many software systems are often released with defects or have some unknown bugs. To overcome such situations developers often need proper feedback on the bugs that are present in the systems. For this, they allow users to report such bugs using bug report systems such as Bugzilla, Jira or other propriety systems. Bug reporting is standard practice in both open source software development and closed source software development.

REFERENCES.

[1] Anh Tuan Nguyen, Tung Thanh Nguyen, Tien N. Nguyen, David Lo, Chengnian Sun. 2012. Duplicate Bug Report Detection with a Combination of Information Retrieval and Topic Modeling. Automated Software Engineering (ASE), 2012 Proceedings of the 27th IEEE/ACM International Conference.