To be Defined

Anonymous
Department of Computer Science
University of Southern California
Anonymous@usc.edu

Anonymous
Department of Computer Science
University of Southern California
Anonymous@usc.edu

***Abstract*—In open source software systems nowadays, we have various software metrics to evaluate software quality of those projects. Some of them are from the metadata of each revision while others are from calculation and analysis (need further info from tools — J.H.). Among all the metadata, commit messages can also be considered to be one kind of software quality metrics. In this research, we investigate how different kind of changes, based on an established categorization, are related to commit message, including its length and contents, and how is commit message related to other software quality metrics. Also, are the relations correlations, or causal? (perhaps using Tetrad.)**

***Index Terms*—Software Engineering, Software Maintenance, Software Quality, Open Source Software**

I. Introduction

A. TBD

The length of commit messages (should be manually written?). Longer commit messages may imply higher software quality. This research will focus on how different kinds of commits (represented by tags) impact the software quality metrics, including the commit messages and other metrics like security and vulnerability metrics.

B. TBD

What are the commit messages? We consider commit messages to be another kind of software quality metric. At the same time, it is used to convey the information of contributors of the commented revisions about what they did to the revision to others contributing to the projects, perhaps to other branches.

II. Related Works

[1]

III. Research Questions

A. Do different type of changes have different level of impact on software quality?

B. Do longer commit messages imply higher software quality?

C. Is there any relation between commit message lengths and commit types?

Are those relations causal relations or correlations?

IV. Data Collection

V. Approach

To be defined

VI. Results

To be defined

VII. Threats to Validity

To be defined

VIII. Conclusions

To be defined

Acknowledgment

This material is based upon work supported in part by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract No. HQ0034-13-D-0004 Research Task WRT 1016 — "Reducing Total Ownership Cost (TOC) and Schedule." SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

References

[1] A. Hindle, D. M. German, M. W. Godfrey, and R. C. Holt, "Automatic classication of large changes into maintenance categories," in *2009 IEEE 17th International Conference on Program Comprehension*, May 2009, pp. 30–39.