

Analyses of software project characteristics on pull request acceptance in distributed software development

Haoyu Wang

School of Computing Science

Sir Alwyn Williams Building

University of Glasgow

G12 8RZ

A dissertation presented in part fulfillment of the requirements of the Degree of Master of Science at the University of Glasgow

<Date of submission placed here>

**Abstract**

With the continued growth of the web and the advent of distributed version control systems, distributed software development has become a mainstream development approach, especially as social coding tools such as GitHub [1] have changed the way software is developed collaboratively and publicly on the World Wide Web [2]. Instead of pushing changes to a central repository, development developers are pulling them from other repositories and merging them locally [3]. This work builds on the dataset of Xunhui Zhang, Ayushi Rastogi, Yue Yu et al.'s study [4] to investigate the impact of item features in pull requests on whether pull requests can be successfully merged. There are six main project features, namely Programming languages, Popularity of project, Age of project, Workload of a project, Activeness of project, and Openness of a project. Using the project characteristics provided in the dataset [4], data cleaning as well as data analysis and data visualisation were carried out to find out the relationship between project characteristics and the success rate of pull requests.

Education Use Consent

I hereby give my permission for this project to be shown to other University of Glasgow students and to be distributed in an electronic form.

<**Please note that you are under no obligation to sign this declaration, but doing so would help future students.>**

Name: Haoyu Wang Signature: Haoyu Wang

Acknowledgements

<Acknowledgements go here>

Contents

<Update the table of contents by right-clicking on it and selecting Update Field… and then select page numbers only.>

Chapter 1 Introduction <This is Heading 1> 1

Chapter 2 Survey 2

Chapter 3 Further Chapters 3

Chapter 4 Conclusion 4

Appendix A <Name of appendix> 1

Appendix B <Another appendix> 2

Bibliography 3

# Introduction

# Survey

<Each new chapter should appear on a new page.>

Background concepts (if required) and overview of relevant previous work (critically evaluate strengths and weaknesses).

# Further Chapters

The content of these chapters depends on the project and should be agreed with your supervisor (e.g. description of the solution, evaluation results, etc).

<Figure below is in style “figure” which continues to style “figure caption” when you press Enter and then back to “Normal” when you press Enter again.>

Figure 1: Some important shapes.

<If you wanted to show any code fragments, you could use the following style called code, which could then be followed by figure caption..>

*# This is a little bit of Python*

**for** i in range( 10 ):

**for** j in range( 10 ):

**print** i\*j,

**print**

Figure 2: A crucial algorithm for the project.

# Conclusion

Main conclusions of your project. Here you should also include suggestions for future work.

###### <Name of appendix>

<Use Heading 6 for the Appendix heading>

###### <Another appendix>

# Bibliography

[1] GitHub: Social Coding. [Online] Available:http://github.com/

[2] Khadke, N., Teh, M. H., & Shen, M. (2012). Predicting acceptance of github pull requests. Stanford–CS 229, Tech. Rep.

[3] Gousios, G., Pinzger, M., & Deursen, A. V. (2014, May). An exploratory study of the pull-based software development model. In Proceedings of the 36th International Conference on Software Engineering (pp. 345-355).

[4] Zhang, X., Rastogi, A., & Yu, Y. (2020, June). On the Shoulders of Giants: A New Dataset for Pull-based Development Research. In Proceedings of the 17th International Conference on Mining Software Repositories (pp. 543-547).