

Track 1 Lightning Talk: Forking as a Tool for Software Sustainability—An Empirical Study.

SARAH ALHOZAIMY

ROBERT HAINES AND CAROLINE JAY

SCHOOL OF COMPUTER SCIENCE

UNIVERSITY OF MANCHESTER

SEPTEMBER 2017



INTRODUCTION

- Ensuring sustainability has been identified as one of the key challenges in the development of software and how to achieve it remains an open question.
- SSI defines 'sustainable software' as that which 'you use today will be available— and continue to be improved and supported—in the future'.

PURPOSE

Examines the link between the use of *forking*— the process of cloning a repository, such that development can progress separately to the original project—and the active life of the project.

SOFTWARE SUSTAINMENT DEFINITIONS

- Sustainment of the original project, S , is defined as the time period from the first commit of a repository through to its last commit, measured in days.

$$S = t_{\text{last-commit}} - t_{\text{initial-commit}}$$

SOFTWARE SUSTAINMENT DEFINITIONS

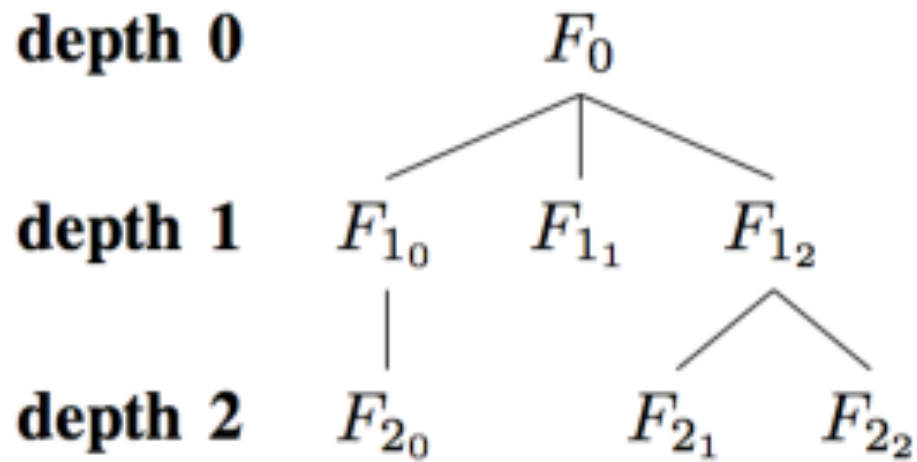
- Software sustainment with forking, SF, is defined as the number of days from the first commit of the original repository through to the last commit when considering all forks.

$$SF = t_{\text{last-commit-on-all-forks}} - t_{\text{initial-commit}}$$

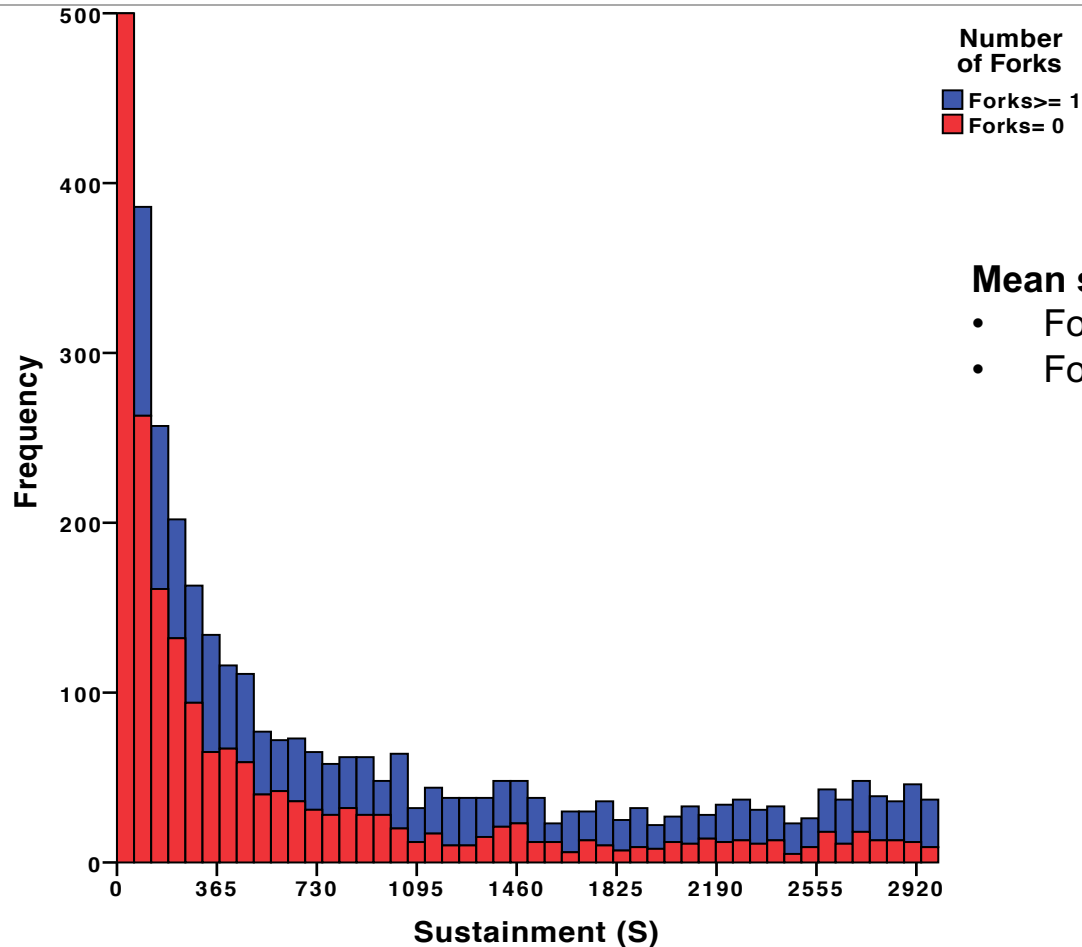
RESEARCH METHOD

- 9,118 projects retrieved from GitHub with 71 different languages.
- GitHub projects selection criteria:
 - The project was created between 1st January and 31st December 2009.
 - Five continuous days from each month were sampled.
 - The project had at least one commit.
 - The first commit occurred on or after 1st of January 2009.

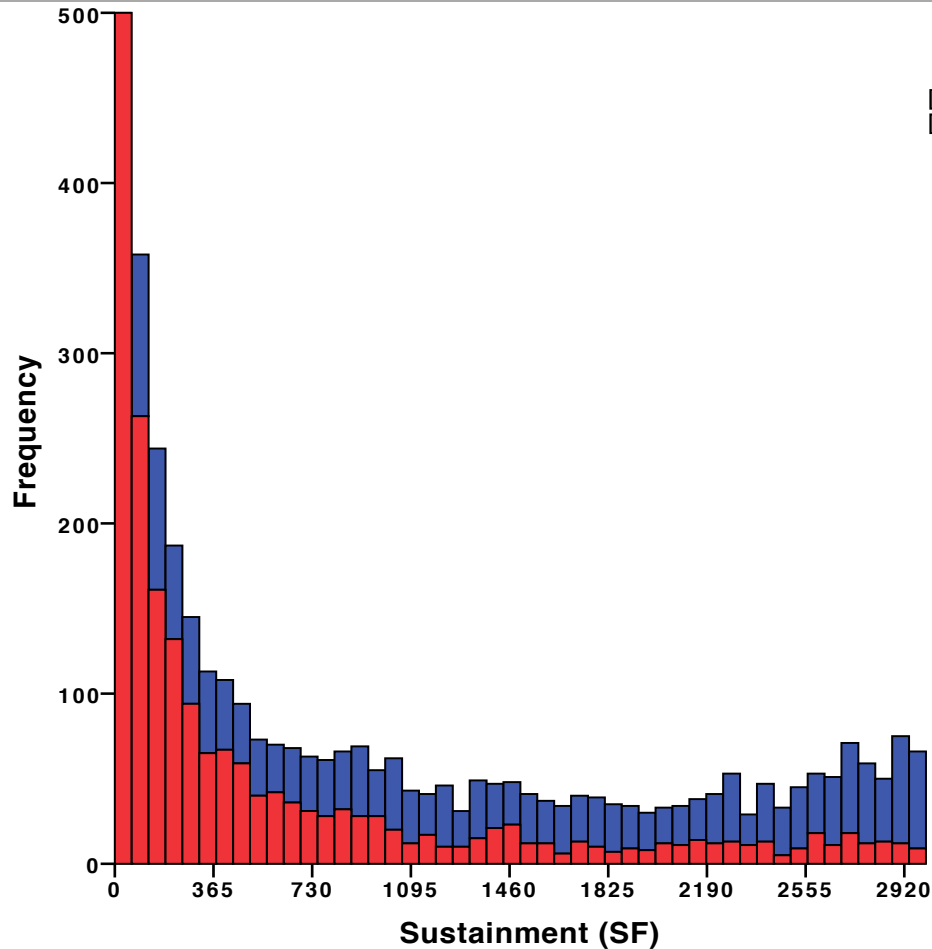
FORK TREE



Is forking related to sustainment of the original project?



Does forking sustain software beyond the original project?



Number of Forks

- Forks >= 1
- Forks = 0

Mean sustainment SF:

- Forks $\geq 1 \rightarrow 1,111$ days
- Forks = 0 $\rightarrow 303$ days

Conclusion

- Forking is positively related to software sustainment.
- Forking help to extend the active life of the original project and beyond the original project.
- Open source repositories such as GitHub that support forking may play a key role in supporting the sustainability of software.

Thank You !

Reference List

- [1] M. de Souza et al, “Defining Sustainability through Developers’ Eyes: Recommendations from an Interview Study,” in WSSSPE 2, 2014.
- [2] S. Betz and T. Caporale, “Sustainable Software System Engineering,” in 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, 2014.
- [3] C. Venters et al., “The Blind Men and the Elephant: Towards an Empirical Evaluation Framework for Software Sustainability,” JORS, vol. 2, no. 1, 2014. [4] K. Fogel, Producing Open Source Software How to Run a Successful Free Software Project, 2nd ed. LLC, 2005. [Online]. Available: <http://producingoss.com/en/producingoss.pdf>
- [5] A. Rastogi, “Forking and the Sustainability of the Developer Community Participation - An Empirical Investigation on Outcomes and Reasons,” SANR, 2016.
- [6] L. Meir, “Programs, Life Cycles, and Laws of Software Evolution,” IEEE, vol. 68, no. 9, Sep. 1980.

Reference List (Cont')

- [7] H. Kuusirati, "Forks in Open Source Software Projects," University of Oulu, 2012. [Online]. Available: https://wiki.oulu.fi/download/attachments/28092087/ossed_2012_kuusirati_seppanen.pdf?version=1&modificationDate=1353314930000
- [8] L. Nyman and T. Mikkonen, "To Fork or Not to Fork : Fork Motivations in SourceForge Projects," Springer, 2011.
- [9] R. Viseur, "Forks impacts and motivations in free and open source projects," IJACSA, vol. 3, no. 2, pp. 117–122, 2012.
- [10] L. Nyman, J. Lindman, and G. Moody, "Code Forking , Governance , and Sustainability in Open Source Software," TIM, pp. 7–12, Jan. 2013.
- [11] A. Aldabjan et al, "How should we measure the relationship between code quality and software sustainability?" WSSSPE 4, CEUR Workshop Proceedings, 2016.