Association for Information Systems
AIS Electronic Library (AISeL)

MWAIS 2011 Proceedings   Midwest (MWAIS)

5-20-2011

Participation in OSS Projects: Does It Support Release Early Release Often?

Bryan Hosack
Illinois State University, bhosack@ilstu.edu

Glen Sagers
Illinois State University, gsagers@ilstu.edu

Follow this and additional works at: http://aisel.aisnet.org/mwais2011

Participation in OSS Projects: Does It Support Release Early Release Often?

Bryan Hosack, PhD
Illinois State University
bhosack@ilstu.edu

Glen Sagers
Illinois State University
gsagers@ilstu.edu

ABSTRACT

This research-in-progress paper explores the role participation in Open Source Software (OSS) projects has on the time between releases to begin to empirically test if "release early, release often" (Raymond, 1999) holds true in a practical sense. A model examining the effect that the number of project forum posts, as a proxy for participation, has on the average time between releases for that project. While the research question proposed expected a decrease in the average time as the number of post increases, in reality there is a significant, but small increase in the time. Further studies can build on this work by increasing the number of variables studied and working with project teams to qualitatively examine what factors drive participation.

Keywords

Open Source Software Projects, OSS Participation, Software Development.

INTRODUCTION

In Open Source Software (OSS) development, many volunteers collaborate to write software. Many different studies have addressed many topics in OSS, such as the quality of the software produced, (Mockus, Fielding & Herbsleb, 2002; Kuan, 2004), the organization of projects (Crowston & Howison, 2006; Kuk, 2006), and the motivation of individuals to participate in OSS development (Bagozzi & Dholakia, 2006).

One area that remains unexplored is whether high levels of participation influence software releases. It has been suggested that a philosophy of OSS is to "release early, release often" (Raymond, 1999, p. 7), and that a larger community of users allows quicker development (Raymond, 1999). Although there are many ways to determine whether a project is, in fact, developing software rapidly, the number of releases in a given time period indicates that sufficient changes have taken place in the software to warrant a new release.

THEORETICAL BACKGROUND

Since OSS developers are largely volunteers (Ghosh, Glott, Krieger & Robles, 2002; Lakhani, Wolf, Bates & DiBona, 2002), and no contractual relationship requires them to spend their time writing software, it may be difficult for the OSS project to meet a specific release schedule. In many cases, OSS projects have instead adopted a philosophy first formulated by Eric Raymond of "release early, release often" (1999, p. 7). This philosophy dictates that the project members release iterative improvements of the software, which can then be tested by users. Based on feedback from users, bugs can be fixed and new features added. This process allows very wide beta testing, instead of only developers and a few select beta testers giving feedback. This method has been claimed to result in higher quality software (Raymond, 1999). Interestingly, Raymond does not differentiate between users and developers (all are "co-developers"), as many other subsequent studies have. In this study, we also do not differentiate, all participants are counted, as any may provide feedback that helps the project, even though they may never submit code. The deadlines and feature lists for a specific program revision are largely determined by users of the software (Mockus, Fielding & Herbsleb, 2002). By allowing users to be co-developers, the time between software releases can be reduced, and improvements desired by the users can be made quickly.

Users who act as co-developers must communicate with developers in some fashion in order to pass along suggestions for improvement of the software. Given the distributed nature of OSS development, this communication is mostly electronic in nature, consisting of email, electronic forums, and similar technologies. The feedback from users is a valuable resource to allow developers to implement new features (Crowston, Anabi, & Howison, 2003), but the volume of communication can be overwhelming (Levine & Moreland, 1990). The question addressed by this research is whether high participation by users results in a "release early, release often" schedule for a project.

This research develops a model of participation which includes all users who post to the discussion forums, as well as those named as developers for the project as active users in the community. These users and their expertise in writing and testing

the software represent the resources available to the project for software development and debugging (Butler, 2001). By counting the number of posts made by active users to the project discussion forums, as well as the number of software releases over the same period, it can be determined whether there is a relationship between participation and the release schedule. This model of participation influencing the release schedule is given in Figure 1. This model does not attempt to discover all possible determinants of success in OSS projects, only whether communication activity in the forums of an OSS project enables the project to develop more quickly. Yielding the research question:

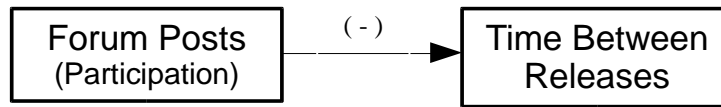**RQ: Do projects with a higher level of participation have a decreased time between releases?**

Figure 1. Theoretical Model

DATA SET AND ANALYSIS

Data from SourceForge.net from February 2011 was collected via SourceForge Research Data Archive for the most recent snapshot of project forum posts. The data collected included the total number of posts per project as the independent variable and the average time between releases as the dependent variable measured as seconds, see Table 1. The number of releases was used as a covariate, since the higher number of releases could lead to a higher number of posts. A project was dropped if it only had one release or if the number of forum posts was less than 10. The overall number of projects analyzed was 8547. SPSS 18 was used to analyze the model.

	N	Mean	Std. Deviation	Minimum	Maximum
number_of_posts	8547	544.48	5466.049	10	322715
avg_time_between_rel	8547	7.5343E6	8.57754E6	6.00	1.24E8

Table 1. Descriptive Statistics

Table 2 indicates that this preliminary model has some success in explaining the effect of participation on time between releases but has such a small R-squared that it is of little practical importance. It is also interesting to note that as the number of posts increases the overall time (in seconds) between releases.

Dependent Variable: Average Time Between Releases

Parameter					95% Confidence Interval	
	B	Std. Error	t	Sig.	Lower Bound	Upper Bound
number_of_posts	129.884	22.439	5.788	.000	85.898	173.870

R Squared = .004 (Adjusted R Squared = .004)

Table 2. Parameter Estimates for the model

CONCLUSION

While the results of this preliminary study are encouraging, there are limitations. Additional data needs to be collected on other factors that could influence the time between releases to yield a model that has better predictive capabilities. Examples of these factors include but are not limited to: characteristics of the participants (e.g., user to developer ratio, do users mostly ask questions or suggest fixes), and characteristics of the project such as type of software. Similarly, additional control data should be collected on the size and complexity of the projects being studied to see what impact they have on the model. Further, qualitative data from project teams and users could be collected to give a richer understanding of what forms participation in OSS take and how best to study this phenomenon. Future research can build on these limitations to help elaborate on the overall dynamics of this complex problem.

This preliminary presentation of a research-in-progress findings indicate that this is a problem worth examining in more

detail. Participation and its effects on OSS project release time can provide insight on the life cycle of these projects but also empirically validate if "release early, release often" (Raymond, 1999) can be seen in practice.

REFERENCES

1. Bagozzi, R. P. and Dholakia, U.M. "Open Source Software User Communities: A Study Of Participation In Linux User Groups," Management Science, (52:7), 2006, pp. 1099-1115.

2. Butler, B. S. "Membership Size, Communication Activity, And Sustainability: A Resource-Based Model Of Online Social Structures," Information Systems Research, (12:4), 2001, pp. 346- 362.

3. Crowston, K., Annabi, H. and Howison, J. "Defining Open Source Project Success," in 2003 International Conference On Information Systems, S. T. March, A. Massey and J. I. DeGross (Eds.), Seattle, Washington, 2003, pp. 327-340.

4. Crowston, K. and Howison, J. "Hierarchy And Centralization In Free And Open Source Software Team Communications," Knowledge, Technology and Policy, (18:4), 2006, pp. 65-85.

5. Ghosh, R. A., Glott, R., Krieger, B. and Robles, G. "Free/Libre And Open Source Software: Survey And Study (A.K.A. Floss Survey)," International Institute of Infonomics, University of Maastricht, 2002, (available online at http://www.infonomics.nl/FLOSS/report/, accessed Jan 2011)

6. Kuan, J. W. "Is Open Source Software Better Than Closed Source Software? Using Bug-Fix Rates To Compare Software Quality," Unpublished Paper, Stanford Institute for Economic Policy Research; Software Industry Center, Carnegie Mellon University, Irvine, CA, 2004.

7. Kuk, G. "Strategic Interaction And Knowledge Sharing In The KDE Developer Mailing List," Management Science, (52:7), 2006, pp. 1031-1042.

8. Lakhani, K. R., Wolf, B., Bates, J. and DiBona, C. "The Boston Consulting Group/OSDN Hacker Survey," Boston Consulting Group, 2002, (available online at http://www.osdn.com/bcg/BCGHACKERSURVEY-0.73.pdf, accessed June 2003)

9. Mockus, A., Fielding, R. T. and Herbsleb, J.D. "Two Case Studies Of Open Source Software Development: Apache And Mozilla," ACM Transactions on Software Engineering Methodology, (11:3), 2002, pp. 309-346.

10. Levine, J. M. and Moreland, R.L. "Progress In Small Group Research," Annual Review of Psychology, (41:none), 1990, pp. 585-634.

11. Raymond, E. S. The Cathedral And The Bazaar: Musings On Linux And Open Source By An Accidental Revolutionary, O'Reilly and Associates, Inc, Sebastopol, CA, 1999.

12. SourceForge Research Data Archive (SRDA): A Repository of FLOSS Research Data, Hosted by Notre Dame University at http://srda.cse.nd.edu/mediawiki/index.php?title=Main_Page.