OSS Unsung Heroes: Crafting Productive Communities Invisibly

Zixuan Feng
dept. EECS
Oregon State University
Corvallis, USA
fengzi@oregonstate.edu

I. Introduction

Open source software (OSS) has become an integral part of the digital landscape and our society today. OSS communities are led by volunteers with diverse backgrounds, collaborating on open platforms to develop and maintain complex software. Existing studies have shown that diversity imbalances have persisted for years in OSS projects. For example, fewer women participate and reach leadership positions in OSS projects. It has been found that women contributors often disengage earlier than men [1]. Similarly, contributors from Non-western countries report facing more barriers than those from Western countries, and non-English speaking contributors report challenges when participating in discussions [2], [3].

Another barrier to participation is the lack of recognition of non-code contributions (e.g., discussions, mentoring), which are often considered "glue work". For example, women contributors are frequently seen as nurturers and often provide more mentoring comments that help with contributor retention and in creating higher quality contributions. Yet, these contributions frequently go unacknowledged, leading to feelings of under-appreciation [2], [4]. Similarly, non-English speakers often dedicate more time to assisting each other in addition to routine activities [3]. However, their efforts are often overlooked, potentially discouraging further engagement [4].

When contributors from minority groups are not encouraged and non-code contributions are unacknowledged, contributors do not feel a sense of belonging in the community. This can result in negative consequences, such as a toxic environment, high turnover, and hindering the community's sustainability. [1], [5], [6].

My work aims to foster OSS sustainability by promoting underrepresented contributors and acknowledging invisible contributions through two primary hypotheses: 1) providing actionable information to OSS project maintainers will enhance the retention of OSS contributors, and 2) inferring the often "invisible" non-code contributions and bringing them to the forefront will foster a more inclusive OSS community.

II. Sustaining diverse contributors

Diversity is a multifaceted concept including age, gender, seniority, ethnicity, role, and expertise [2]. There is a lack of understanding regarding the current state of diversity and where to focus efforts on promoting diversity in OSS communities. In previous research, we conducted a longitudinal analysis of two cross-sectional surveys in 2020 and 2022 to investigate the state of diversity in OSS across various attributes, including gender, English proficiency, seniority, and country of residence [2]. The results from over 1,000 participants show improvements in technical, social, and process challenges across diverse backgrounds. Yet, for underrepresented groups such as gender minorities, newcomers, Non-western contributors, or non-English speakers, significant gaps persist, along with challenges such as feeling underrepresented and difficulty in finding comfortable mentors to work with.

To further understand the role of diversity in sustaining OSS contributors, we conducted a preliminary analysis using "Kaplan-Meier curves" to empirically analyze the impact of diverse backgrounds on contributor retention in OSS. We selected the Flutter community as a sample due to its extensive history. When comparing historical data to the recent three years, the retention difference between man and woman contributors has decreased, resulting in improved retention of woman contributors. Non-western contributors tend to disengage earlier than Western contributors. We also found that affiliation impacts a contributor's retention, with differences between volunteers and company-affiliated contributors.

As we have outlined the challenges of diversity in sustaining OSS contributors, there's a pressing need to transform this understanding into actionable insights for OSS maintainers. We bridge this gap by designing a dashboard as a strategic tool for project maintainers. Our approach is guided by three primary research questions (RQs): **RQ1.** What approaches can project maintainers adopt to monitor contributor retention in OSS projects effectively? **RQ2.** How effective and useful is the designed dashboard in assisting project maintainers with monitoring contributors' retention? **RQ3.** Using the dashboard, what strategies can maintainers employ to improve contributor retention in OSS projects?

We first reviewed the existing literature and conducted interviews with OSS project maintainers. This helped us understand their specific needs for monitoring contributor retention and identify activity metrics that could affect it. We then designed the dashboard utilizing statistical methods, including Cox-regression [7] and Kaplan-Meier survival estimates [8] to provide a detailed view of the likelihood of contributor disengage-

ment, disaggregated by diverse attributes, including gender, region, and affiliations. The dashboard allows maintainers to compare continued engagement probabilities between diverse groups, like women vs. men or company-affiliated vs. volunteer contributors. It also highlights contributors at higher risk of disengagement, aiding maintainers to take early targeted actions to sustain contributors better.

III. Acknowledging invisible contributions

To promote a sustainable OSS community, it is crucial to infer the invisible contributions. Failing to acknowledge these efforts, the project's sustainability is at risk, as it can diminish motivation and hinder career growth for contributors, ultimately causing them to leave communities. For instance, the absence of acknowledging women's non-code contributions can lead to social challenges such as a lack of peer parity and non-inclusive communication [5].

My previous work defined a non-code contribution as implicit mentoring, where contributors guide each other through daily activities, such as discussion comments [9]. Through an empirical investigation of pull requests across OSS projects, we developed a classifier to identify instances of implicit mentoring. We found that nearly one in three pull request comments contain implicit mentoring. Through a large-scale survey of OSS contributors, we found that over 90% of the participants confirmed that this non-code contribution improves technical skills, enhances career progression, and expands their social network.

IV. Discussion and future work

A. Proposed evaluation of dashboard

I have summarized my preliminary analysis aimed at understanding and promoting diverse contributors. To continue, I plan to evaluate dashboard design with OSS project maintainers by conducting a one-month diary study to investigate strategies that contributors find beneficial for retention. The maintainers will be assigned to check the dashboard and monitor the retention data of their projects every week by using the dashboard. A survey will also be assigned for them to record retention data every week, including details about the state of retention, any concerns they have noticed, potential actions they plan to take, and any actions they have taken based on the retention data from previous weeks. At the conclusion of the study, project maintainer participants will be invited to a semi-structured interview to share their experience in using the dashboard. Questions may cover topics such as evaluating the usefulness and effectiveness of the dashboard.

B. Promoting contributions beyond code

Another thrust of my work is to recognize invisible contributions, such as implicit mentoring, as we have defined before. Such invisible contributions often serve as the "glue" that holds the OSS community together beyond the code. However, many other invisible contributions exist and remain unrecognized in OSS. As a continuation of my work, my goal is to answer the following research questions: **RQ1.** What are the different invisible contributions in OSS, and how prevalent are these contributions? **RQ2.** How do invisible contributions impact OSS contributors and the community? **RQ3.** What are potential approaches to better trace and acknowledge invisible contributions?

To answer these research questions, I plan to conduct interviews with OSS practitioners from diverse backgrounds, with a particular focus on underrepresented groups as their contributions are often overlooked [5], [9]. These interviews will help us identify and understand the invisible contributions of OSS. Additionally, I plan to survey the OSS community to validate invisible contributions and investigate their impacts on engagement, considering factors like contributors' activity duration, motivation, and sense of belonging within OSS. Moreover, I plan to develop a taxonomy that categorizes different types of invisible contributions. I will then investigate potential approaches for tracing and inferring these contributions through empirical analysis.

V. Conclusion

Through promoting diverse contributors and inferring invisible contributions, I hope to pave new paths for acknowledging unsung heroes in OSS, thereby nurturing a diverse, inclusive, and sustainable OSS community.

Acknowledgement

This work is supported by the National Science Foundation under Grant No.: 1901031. I am grateful to Dr. Anita Sarma for her insightful guidance throughout this work.

References

[1] H. S. Qiu, A. Nolte, A. Brown, A. Serebrenik, and B. Vasilescu, "Going farther together: The impact of social capital on sustained participation in open source," in *2019 ieee/acm 41st international conference on software engineering (icse)*. IEEE, 2019, pp. 688–699.

[2] Z. Feng, M. Guizani, M. A. Gerosa, and A. Sarma, "The state of diversity and inclusion in apache: A pulse check," in *2023 IEEE/ACM 16th International Conference on Cooperative and Human Aspects of Software Engineering (CHASE)*, 2023, pp. 150–160.

[3] S. Balali, I. Steinmacher, U. Annamalai, A. Sarma, and M. A. Gerosa, "Newcomers' barriers... is that all? an analysis of mentors' and newcomers' barriers in oss projects," *Computer Supported Cooperative Work (CSCW)*, vol. 27, pp. 679–714, 2018.

[4] Z. Feng, A. Chatterjee, A. Sarma, and I. Ahmed, "Implicit mentoring: The unacknowledged developer efforts in open source," *arXiv preprint arXiv:2202.11300*, 2022.

[5] B. Trinkenreich, I. Wiese, A. Sarma, M. Gerosa, and I. Steinmacher, "Women's participation in open source software: A survey of the literature," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 31, no. 4, pp. 1–37, 2022.

[6] B. Trinkenreich, K.-J. Stol, A. Sarma, D. M. German, M. A. Gerosa, and I. Steinmacher, "Do i belong? modeling sense of virtual community among linux kernel contributors," *arXiv preprint arXiv:2301.06437*, 2023.

[7] M. Lunn and D. McNeil, "Applying cox regression to competing risks," *Biometrics*, pp. 524–532, 1995.

[8] M. K. Goel, P. Khanna, and J. Kishore, "Understanding survival analysis: Kaplan-meier estimate," *International journal of Ayurveda research*, vol. 1, no. 4, p. 274, 2010.

[9] Z. Feng, A. Chatterjee, A. Sarma, and I. Ahmed, "A case study of implicit mentoring, its prevalence, and impact in apache," in *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2022, pp. 797–809.