

Viability of Crowd-Volunteered Open Research Reviews

Malay Bhattacharyya

Department of Information Technology
Indian Institute of Engineering Science and Technology, Shibpur
Howrah – 711103, India
E-mail: malaybhattacharyya@it.becs.ac.in

Abstract

In this study, we examine the feasibility of reviewing research publications through crowdsourcing. We propose a post-publication review model where the research materials are posted online for the crowd volunteers to review. One such platform, named PuGLi, is prepared and the data collected from it is analyzed. The analysis raises some interesting questions related to the challenges of attracting crowd volunteers, determining truthfulness of opinions, and obtaining the final judgment from multiple opinions received from the crowd.

Introduction

Crowdsourcing models have received considerable interests in the recent years due to its distributed yet united power of solving diverse problems (Kittur et al. 2013). These crowdsourcing models can be either collaborative, in which the problem is solved collectively by the crowd, or competitive, where the problem is solved by the crowd with competing interests. Collaborative crowdsourcing models have been successfully used earlier for several applications like collaborative research, collaborative learning, crowdfunding, etc. Recent criticisms of traditional publication models (Cross et al. 2014) demand for an alternative solution. We propose a collaborative crowdsourcing model for reviewing research publications publicly. We prepare an online crowdsourcing platform PuGLi, where the research materials are posted online as videos for the crowd volunteers to review. We study the behavior of crowd reviewers from this platform and analyze the data. Our analysis poses some open problems and some highlights to address these issues.

Related Works

Continuous criticisms of traditional publication models have advocated several alternative proposals. Some of these include open access publication (Beaudouin-Lafon 2012), post-publication peer-review (Terry 2014), public repositories of scientific research like Arxiv or CoRR, discussion forums like PubZone, etc. A recent research highlights some interesting psychological patterns of the scientific authors (Terry 2014). This leads to the suggestion of establishing

“high acceptance” publication venues for setting discussion platforms and make experimental data promptly available. We propose a novel crowd-volunteered review model that has been guided by these interesting highlights from the said literature. We study the proposed model through an experimental setting to understand its feasibility and limitation.

Experimental Design

The proposed post-publication review model was implemented as an online platform (accessible at: <http://www.pugli.grn.cc>), named as PuGLi, where the research materials were posted as video papers to be reviewed by crowd volunteers. We prepared these videos as short papers on emerging topics (unpublished in some cases). The review process was made public and accessible to all so that we can analyze a large-scale crowd activity. Author details were kept anonymous during the review process to obtain unbiased reports. Total seven decision options were given to the reviewers, namely ‘Strong Accept’, ‘Accept’, ‘Weak Accept’, ‘Borderline’, ‘Weak Reject’, ‘Reject’ and ‘Strong Reject’, in the decreasing order of superiority. All the video papers were posted on March 18, 2014 at midnight and a review period of 2 months (until May 16, 2014) was considered. In the first round of analysis, we kept the decisions given by the reviewers secret. In the second round, which started after 3 months (and lasted for a month), we made the decisions public and restarted taking reviews. This enabled us to compare closed (decisions are private) and open (decisions are public) review models. We analyzed the data for the same set of crowd volunteers. If no response is given in the second round then the earlier response is taken. After the user activity at PuGLi in the scheduled time, we downloaded the reviews that were submitted at the review portal (PuGLi) for the video papers.

Dataset Details

We collected the reviewers’ (crowd volunteers) name, email address, profession, as well as the relevant keywords, their comments and decisions during the review. The various details of the dataset are provided in Table 1. As can be seen from the table, we obtained a high average number of reviews per paper (min = 5, max = 9). Interestingly, there were common reviewers involved in reviewing multiple papers. This highlights their growing interest in the model.

Characteristics	Value
# Papers	05
# Papers to be considered for a journal	03
# Papers to be considered for a conference	02
# Reviews	35
# Reviewers	16
Average number of reviews per paper	7.00
Average number of reviewers per paper	3.20
Average number of reviews per reviewer	2.19

Table 1: Details of the data collected from PuGLi platform.

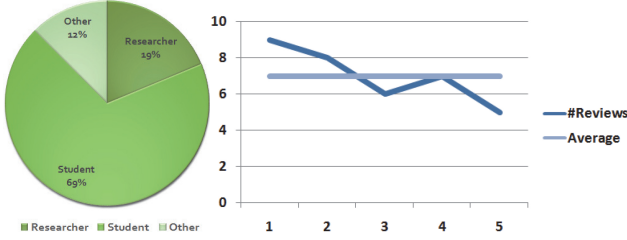


Figure 1: The distribution of reviewers' profession (left) and the number of reviews obtained for the papers in their order of appearance in PuGLi platform (right).

Preliminary Results

We studied what type of people is basically reviewing the papers at PuGLi. We observed that students are the major reviewers (see Fig. 1), not the general people. Even though the portal was public, the non-expert people might not have taken the risk to give feedback on the papers that deal with expert studies like research on emerging areas. We purposefully plagiarized in a paper and interestingly it was pointed out by a reviewer. This highlights the power of the crowd-powered reviewing. To show the deviation of the number of reviews received for the papers, we have plotted the review counts against the paper Ids in Fig. 1(b). It clearly shows that the succeeding papers (as they appear sequentially in the PuGLi platform) have obtained reduced attention from the reviewers. Therefore, the number of reviews has gradually decreased as the papers appear one after another.

Our initial observations suggested that open reviews might be biased by the preceding review reports or decisions. To verify this, we compared the review scores obtained from the same set of reviewers for the same video papers in the closed and open models. The decisions given in open reviews are apparently more close to each other in comparison with the closed case (see Fig. 2), even though the reviewers are same. The outliers in the righthand side of Fig. 2 either denote spam reviews or very stringent reviews (done by strict experts). To be sure about this, we looked into the review reports and found one of them to be just "What is this jaar? Everything went tangentially over my head!! :(" ('jaar' means 'friend' in an Asian language).

Challenges

After analyzing the individual review comments, we observed that many of the reviews are very short, no more

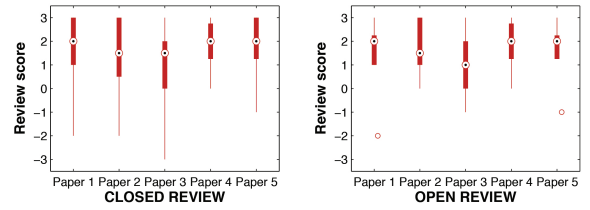


Figure 2: The comparative boxplots of review scores obtained through closed and open review process. Here, the scores corresponding to 'Strong Accept', 'Accept', 'Weak Accept', 'Borderline', 'Weak Reject', 'Reject' and 'Strong Reject' are given by 3, 2, 1, 0, -1, -2, -3, respectively.

than a sentence. Again, some of them are irrelevant and insignificant. Therefore, detecting spam reviews appears to be an interesting task of further study. Leveraging the distinction between novice and expert contribution is a challenging task. It is also interesting to analyze the linguistic patterns of the crowd-powered reviews. How the openness of decisions might affect the final review result is also an encouraging direction of further research. Another interesting problem appears to be motivating/incentivizing the crowd volunteers. This requires effective mechanism design and one possible solution could be the prerequisite of submitting reviews on others' papers before submitting a paper (of yours). Obtaining final judgment, where reviews are open, is also an interesting problem. As the succeeding reviewers get biased by the predecessors, a way of computing the final review score could be $S = \sum_{i=1}^n \frac{S_i}{2^i}$, where S_i denotes the review score given by the i^{th} reviewer. Note that, $S = S^*$ if $S_i = S^*$, $\forall i$.

Conclusions

In this article, we studied an online platform for realizing the viability of a model for publicly reviewing video papers with the help of crowd volunteers. Overall, our approach showed that crowdsourcing can be used as a great tool for providing quality and prompt reviews. Emerging approaches of crowd-powered video editing like VidWiki might enable such models to revise the papers (like the video papers in PuGLi) online, as a part of crowd-volunteered reviewing. But many of the drawbacks and challenges were also observed.

References

- Beaudouin-Lafon, M. 2012. Open access to scientific publications. *Communications of the ACM* 53(2):32–34.
- Cross, A.; Bayyapunedi, M.; Ravindran, D.; Cutrell, E.; and Thies, W. 2014. VidWiki: Enabling the Crowd to Improve the Legibility of Online Educational Videos. In *Proceedings of the CSCW*, 1167–1175. ACM Press.
- Kittur, A.; Nickerson, J. V.; Bernstein, M. S.; Gerber, E. M.; Shaw, A.; Zimmerman, J.; Lease, M.; and Horton, J. J. 2013. The Future of Crowd Work. In *Proceedings of the CSCW*, 1301–1318. ACM Press.
- Terry, D. 2014. Publish Now, Judge Later. *Communications of the ACM* 57(1):44–46.