Prioritizing Critical Challenges for the Quality of Crowdsourced Software Products

Zanab Safdar¹, Shahid Farid², Moiz Uddin Ahmed³

¹Department of Information Technology, Bahauddin Zakariya University, Multan, Pakistan
²Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan
³Department of Computer Science, Allama Iqbal Open University, Islamabad, Pakistan

ABSTRACT

Crowdsourcing is an online dispersed model for solving problems where the organizations utilize a vast number of commitments from internet users. Crowdsourcing platforms develop cheap labor ideas and administrations towards solving problems. However, sometimes crowdsourcing incorporates work from unqualified workers that subsequently brings up issues on the quality of crowdsourced developed products. Therefore, this study intends to identify various challenges in order to assure the quality of crowdsourced software products. Moreover, identified challenges have been prioritized following empirical investigation conducting a survey of experts. Different statistical models have been deployed in order to achieve the formulated objectives of the study. Investigations reflect that security, the privacy of data and lack of interaction between crowd and client have emerged as a most critical challenge in order to ensure the quality of crowdsourcing software product.

Keywords: Crowdsourcing, crowdsourced software products, quality of crowdsource product, crowdsource challenges.

INTRODUCTION

Crowdsourcing is a compelling way of receiving commitments from a vast number of internet users to produce unique ideas and solve multiple problems¹,². Crowdsourcing has been utilized as a part of different fields of workmanship, business, administration and arts and so forth. Crowdsourcing platforms are an online and open source to help crowd getting the opportunity to work and get paid. Crowdsourcing associations endeavor to raise their growth through developing products from cheap workers. However, software products quality is a subjective parameter as offering products alone is deficient. Hence, Crowdsourcing platforms develop low-quality products due to unqualified workers and wrong description of the task that implies less believable output³.

Latterly, several studies have conducted oneself crowdsourcing platforms yet the quality of developed products need to be focused [4-9]. Furthermore,
crowdsourcing platforms enable numerous unfit workers to work that result in the extensive amount of unusable and below quality output. This has been raised numerous questions about how to gage the quality and who shall gage the quality of software products developed in crowdsourcing environment. Therefore, this study aims to expound crucial issues that affect the quality of crowdsource developed products. Therefore, this study intends to identify crucial issues affecting the quality of a crowdsource software product. Furthermore, the study contributes in a way by prioritizing the identified quality issues facilitating the crowdsource organizations to enhance the quality of developed software products.

This study is worded as Section 2 presents background the crowdsourcing software products quality; Section 3 presents research methodology adopted by this study. Section 4 presents results and discussion of the study. Section 5 delineates identified issues that affect the quality of crowdsourced software products whereas Section 6 concludes this study.

**BACKGROUND**

Crowdsourcing approach was instituted in wired magazine (2016) by Jeff Howe encouraging a pool of cheap work suppositions and thoughts for critical thinking [10]. The term crowdsourcing is a combination of "crowd" and "outsourcing" where crowd covers diverse internet users that work to solve problems defined by organizations. Outsourcing is well known term for separating work between workers despite the fact that the work originates from defined sources whereas; crowdsourcing incorporates work from undefined sources. Various crowdsourcing platform includes TopCoder, uTest, Amazone Mechanical Turk, Crowdx, Bountify, Pay4Bugs, AppStori, eBay, Stack Overflow, Crowd flower and etc. are available to facilitate with crowdsource services [2, 11-18]. The crowdsourcing platforms consists of organizations and workers where organization (requester) request for work by submitting problem (task) on the platform and grasps discovering arrangement of problems from unidentified group of individuals as depicted in Fig. 1. The people who take a problem from submitted task are called workers. Requester is responsible for grant of reward on completion of each problem whereas reward can be in the form of money, cards, material etc. [10, 15, 19]. Requester analysis all the tasks and sometimes they allow workers to give feedback on each task. Sometimes requesting organization needs to remunerate all workers who have finished the task as indicated by the characterized directions. Crowdsourcing platforms have been utilized to depict task and administrations through which they get commitments from assorted specialists.

Crowdsourcing clients endeavor to raise their profit through developing products from cheap workers. However, software products quality is a subjective parameter where quality is defined by its customers and that must be fit for their use although crowdsourcing platforms develop low quality products due to unqualified workers; wrong description of task that implies less believable output [3]. One of the boundless objectives of crowdsourcing is to grow high quality products within extension, time and spending plan. Quality is imperative and urgent for the survival of the software product item as reported in the literature. In recent years, many studies have discussed quality in various computing fields including data and information quality, software products and services quality. Some literatures have discussed quality control in crowdsourcing however; a few have discussed the quality of developed products [5, 20-24]. Software products quality contains various attributes including controllability, compactness, unwavering quality, usefulness, execution, effectiveness, security, practicality, ease of use and so on [20, 25]. In this respect, the eccentricity of crowdsourcing require considering novel issues that are rising and affecting the quality of developed products [21].
Multiple factors like people, software, process and etc. may influence the quality of developed products that need to be addressed in order to produce quality output. State of the art literature is deficient to address critical issues that encountered by the quality of developed crowdsource software products. Therefore, this study has been devoted to identifying the crucial issues encountered by the crowdsource organizations. Moreover, the identified issues have been prioritized deploying different statistical methods which have never done before for the quality crowdsourcing products. Furthermore, prioritization of the identified issues may lead our effort to facilitate crowdsource organization in enhancing the quality of the developed software product.

Selection of an appropriate research approach is one of the critical tasks for any research. Exploratory mode of research was utilized to explore state of the art literature intensively. One of the reasons behind adopting exploratory approach is that confined material is reported regarding the identification of critical challenges confronted by the quality of crowdsource products. As the exploratory method is significant in the situation where either the targeted issue has never been addressed or inadequate information is achievable and investigator intends to probe the research arena [26, 27]. Additionally, an exploratory research commences constructing observations and penetrating for a pattern. The researcher puts forward an idea about why this pattern occurs. This method is occasionally known as the inductive approach. Therefore, this mode of research facilitates by furnishing an appropriate manner to alleviate investigator with the basic work for afterward studies [26, 28].

a) Sample

The targeted population comprises of experts in the arena, from various higher educational institutions and software industry. These experts were selected on the basis of their experience in the area of quality of conventional and crowdsourcing products. It is crucial to highlight that targeted sample was having more than 5 years of experience in the domain. Snowball sampling method was adopted while selecting experts from the population. The rationale behind utilizing snowball sampling method was limited availability of experts in the crowdsource arena. Therefore, recommendations were sought from the experts in order to enhance the sample size. The experts were requested to rank each of the challenge using five-point Likert Scale from not crucial (1) to most crucial (5).

b) Criteria

An instrument was developed a pertaining list of identified challenges after extensively reviewing of state of the art literature reported by well-known journals and conference articles. 45 questionnaires were distributed amongst the targeted population. However, only 33 responses were obtained at the rate of 73%. Table 1 demonstrates the demographic detail of the respondents of the study.

<table>
<thead>
<tr>
<th>Table 1: Demographic Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
</tr>
<tr>
<td>Gender:</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Qualification:</td>
</tr>
<tr>
<td>Ph.D.</td>
</tr>
<tr>
<td>MS</td>
</tr>
<tr>
<td>Master</td>
</tr>
<tr>
<td>Designation:</td>
</tr>
<tr>
<td>Academia</td>
</tr>
<tr>
<td>Industry</td>
</tr>
</tbody>
</table>

In order to analyze the collected data through empirical investigation, SPSS software was utilized in terms of Cronbach’s Alpha and Principal Component Analysis (PCA) respectively.

a) Cronbach’s Alpha

Cronbach's alpha has been used to gauge the internal consistency and reliability of the collected data. The range of Cronbach’s Alpha begins from 0.4 to 0.9 [29] as illustrated in Table 2. Hence, values higher than 0.7 demonstrates the higher reliability of the data [30]. The Cronbach’s Alpha of this study reaches to 0.91 reflecting that data is reliable and there exist strong consistency among the items of the instrument.

b) Principal Component Analysis
PCA was deployed to visually gauge which of the factor (challenges in this case) have most variance in data. Table 1: Guidelines to assess the reliability

<table>
<thead>
<tr>
<th>Cronbach’s Alpha Coefficient</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.8</td>
<td>Good</td>
</tr>
<tr>
<td>0.7</td>
<td>Acceptable</td>
</tr>
<tr>
<td>0.6</td>
<td>Questionable</td>
</tr>
<tr>
<td>0.5</td>
<td>Poor</td>
</tr>
<tr>
<td>0.4</td>
<td>Unacceptable</td>
</tr>
</tbody>
</table>

PCA is one of well-known and widespread multivariate statistical technique to prioritize the factor with respect to their severity. Table 3 illustrates the proportion of each of the identified challenges. Security, privacy and data protection emerges as the most crucial challenge which has higher big proportional value i.e. 0.55 or 55%. Whereas lack of interaction between crowd and client, lack of collaborative tools and lack of qualified people are having proportion equal to 0.20, 0.15 and 0.05 respectively. The challenges with too small proportion (say≤0.03) can be negligible [31]. Therefore, challenges like lack of monetary motivation and task design are having too small proportion i.e. 0.03 and 0.02 respectively.

**Identified Issues:**

This study identifies the issues that evolve around and affect the quality of crowdsourced software products. As crowdsourcing allows any person to participate which sometimes result in a large number of unusable contributions. The ideas of workers get funded by requester yet those ideas quality is still a large argument in crowdsourcing. For making a plan to transform into a reality right off the associations require funds. The number of funds can range from a couple of hundreds to thousands that depend on scope or complexity of the crowdsourcing projects. These issues in crowdsourcing affect the quality of outputs. Following are the identified issues that should be resolved to improve the quality of crowdsourcing projects.

**Security, Privacy and Data Protection:**

In crowdsourcing privacy issues are linked to personal data and information about the workers or individuals [32, 33]. Sometimes contributors do not want to publicize their personal data such as their phone number, name, gender, address and so on. Different crowdsourcing platforms have a different method of interaction so security and privacy risks are linked with contributors. The issues related to data protection in crowdsourcing are important to be considered. Data protection is a key component of general crowdsourcing towards business crowdsourcing; in every crowdsourcing initiatives, data protection is important for saving personal data from misuse. Relational perspective shows that data privacy and security both are linked with data protection. As crowdsourcing is becoming wider different people started using it for different purposes [34].

**Lack of Crowd and Client Interaction:**

Another issue that effect crowdsourcing products quality is due to lack of interaction between the participating crowd and the crowdsourcing organizations (client or requester). Organizations interaction with workers is essential part of the design process of product quality [35]. When workers know little information about the desired product due to less interaction with the client organization this leads to the development of low quality products. Delivering quality products means to meet customer expectations. Customers perceived value of product quality is perhaps most critical determinants for customers willingness to buy products [36]. Involvement programs helps to improve quality of products by training the crowd to improve their skills, communication ways, multiskilling and improve flexibility, responsibility and measuring user’s satisfaction [37].

**Lack of Collaboration Tools:**

Another cause of the decrease in product quality in crowdsourcing is due to deficiency of collaboration tools. Typical business workplaces organize its co-workers in a...
way so that they can work with each other and share each other’s knowledge or ideas [37, 38]. Using crowdsourcing as an approach for evaluating software products is surely gaining foothold furthermore, the business workplaces provide its employees necessary information, guidance, procedures, and tools frequently to fulfill their responsibilities. However, in crowdsourcing, crowd workers depend on their own knowledge and means for completing tasks. Many crowdsourcing platforms deficient to provide better information about tasks, mostly fillers are used to define tasks that contain out-of-vocabulary words that are proved to be troublesome [39]. Currently, there are no techniques for iterative correctness that can increase the performance of overall tasks.

**Lack of Qualified People:**

Skilled crowd and expertise are required to successfully achieve the desired results but oftentimes the scenario is that crowd is less qualified and only working for reward and money [40]. This scenario is more problematic for complex tasks, lack of qualified people impact the quality of products such as engineering product design and validation of products. Lack of information in the crowd can lead to the low-quality products due to that workers perform tasks very poorly. Due to the lack of submitted contributions verification techniques, many unqualified workers take advantage and produce products of low quality [37, 41]. In some cases, it becomes impossible for organizations to find qualified people from the crowd. Incorrect crowd members affect the reputation of the organization and impact the quality of products. Inaccurate crowd evaluations lead to the unqualified workers as evaluating experts exist along with the wrong crowd [40].

**Lack of Monetary Motivation:**

Mostly crowdsourcing software products failed due to the lack of monetary motivation or sometimes too few workers participate in problem-solving. In current crowdsourcing platforms, they lack all these features for motivating crowd through more coordination, influencing crowd and communicating with them. When crowdsourcing platforms belief that workers are only working for a hobby that has need of only minimal payment it is a hallucination. The platforms have to admit that majority of workers do not think about it as a hobby but as a main source of income and it is a labor marketplace for them. Many kinds of literature have discussed that there is no correlation between the reward and product quality in crowdsourcing platforms [39, 42]. High developed countries have low workers to participate because of the low amount of pay. Lack of monetary support stops qualified people to contribute to crowdsourcing platforms [42, 43]. There is also no guarantee for the workers that they will get paid for their work. Crowdsourcing platforms do not follow first in first out (FIFO) queue in case if tasks are not completed quickly so they might be over and done by search procedures or searching filters so workers will not see them again. Yet sometimes tasks are completed on time but it doesn’t produce quality results always.

**Task Design:**

Quality planning is an integral part of overall business planning; poor designing of tasks also impact the quality of crowdsourcing products. The existing practices of modeling crowdsourcing data collection unnecessarily restrict quality [15, 19, 41]. Workers provide more accurate information when phenomena are classified and defined at more general level [3]. Greater accuracy is expected from workers when they could provide more accurate outcomes as requested by the organization and compare tasks with the description of the tasks. Existing organizations do not focus on the task designing and planning strategies deployment that leads to the poor quality. Results exhibit that accuracy is depending on the module of specified information that is used for modeling domains. Planning strategies, deployment plans will help to resolve this issue.

**CONCLUSION**

Crowdsourcing is an effective approach for accepting commitments from vast number of internet users to generate ideas and solve problems. Multiple online and open source crowdsourcing platforms have been used for solving problems that includes TopCoder, uTest, Mturk, CrowdX, Bountify, Pay4Bugs and AppStori and the rest. In these platforms the requesting organization submits their task for completion that is completed by numerous workers. Without qualified group of workers, crowdsourcing platforms remains unappealing for requesters such as tasks are finished inaccurate and
gradually that result in effecting the quality of software products. Crowdsourcing enables any person to work that occasionally results in substantial amount of unusable assurances. In this study most crucial issues have been identified experienced by the organizations embrace in crowdsourced improvement. Result shows that these issues need to be focused for more accurate and reliable outcomes. Crowdsourcing is a subjective term; there are numerous issues and difficulties identified with quality of programming products created through crowdsourcing that require new quality control systems. In future work we will contribute by giving quality control mechanisms that will help crowdsourcing platforms to keep up their procedures and quality in control. Quality control mechanism will likewise help to keep up group and undertaking configuration and additionally keeping up quality guidelines for organization.

REFERENCES

2. Stol, K.-J. and B. Fitzgerald. Two’s company, three’s a crowd: a case study of crowdsourcing software development. in Proceedings of the 36th International Conference on Software Engineering. 2014. ACM.
6. He, Z., J. Cao, and X. Liu. High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility. in Computer Communications (INFOCOM), 2015 IEEE Conference on. 2015. IEEE.