Why do many questions on Stack Exchange not have many answers? A preliminary agent-based model of answer contributions to Stack Exchange

Arpan Jani[1] and Naren Peddibhotla[2]

[1] University of Wisconsin – River Falls, River Falls, WI 54022, USA
[2] SUNY Polytechnic Institute, Utica, NY 13502, USA

Abstract. Many questions on community question-answering sites such as Stack Exchange either do not get answers or get very few answers. Using a simple agent-based model, we examine the role of the expertise of individuals who may potentially be interested in answering questions at such sites. Two contribution rules are hypothesized to determine whether potential contributors answer a question with the assumption that these potential contributors have expertise that follows a normal distribution. Under the condition that agents only contribute when the question difficulty closely matches with their expertise level, simulation runs show that questions that are too easy or too difficult do not get many answers. Under the condition that contributors answer questions as long as their ability exceeds the difficulty threshold, simulation runs show that easy questions get more answers but there is a steep decline in number of answers as complexity of questions increase. This model is preliminary, but we plan to extend this study by incorporating model parameters that reflect the structure of a site such as Stack Exchange.

Keywords: Agent-based Modeling (ABM), Community Question-Answering (CQA), Unanswered Questions, Contributor Expertise.

1 Introduction

1.1 Motivation

The value of many online platforms, such as social media sites (e.g., LinkedIn.com), question-answering forums (e.g., Stackexchange.com), review sites (e.g., Tripadvisor.com) is based almost entirely on user contributions. On any given platform, however, there is an uneven distribution of content among topics, whether it is the amount of content or the quality of the content. It is desirable for a site to provide high-quality and complete content on all the domains of relevance to that site. Stack Exchange is one such platform with a collection of 176 Q&A (question and answering) sites focused on various domains. Stack Overflow is the most popular Q&A site on Stack Exchange with more than 20 million questions posted so far. Stack Exchange provides easy access to their data through the Data Explorer tool (https://data.stackexchange.com/) and data dumps (https://archive.org/details/stackexchange). As Stack Exchange provides access to their data, many studies have looked at a variety of research questions. Chua and Banerjee (2015) had found that many questions on Stack Overflow did not receive any answers and that the quantity of unanswered questions had increased over time.

Using the Data Explorer tool, we analyzed the number of answers provided for the 20 million questions posted on Stack Overflow from years 2008 – 2020 to confirm and extend the analysis of Chua and Banerjee (2015). As shown in figures 1 and 2, a significant majority of questions posted between 2008 and 2020 (August) were found to have fewer than 3 answers. About half of the questions, i.e. about 10 million, had received only one answer whereas about 2.85 million (14%) questions had received no answer at all. As shown in figure 3, the trend in the number of questions that had not received an answer as identified by Chua and Banerjee (2015) has continued. While there has been a declining trend in the total number of questions posted on Stack Overflow in recent years, the proportion of questions that have not received an answer has grown.

1.2 Research Question

Chua and Banerjee (2015) attempted to find an explanation for many questions not receiving any answers on Stack Overflow. Their analysis indicated that there was a greater likelihood of a question receiving an answer when the question was posed by someone who had asked many questions already or when the question was complete, clear and had a shorter title and shorter description. It may be that more experienced question-askers were able to formulate better questions and question clarity and completeness reduced the cognitive effort involved in understanding the question. Saha et al. (2013) found that unanswered questions were longer on an average as compared to questions that received answers. This again could be due to the cognitive effort required in understanding longer questions. Yang et al. (2011) had found in their study on Yahoo! Answers that users with more experience were more likely to get answers to their questions and medium length questions were less likely to get answers as compared to short and long questions. Asaduzzaman, Mashiyat, Roy and Schneider (2013) found that the reputation of contributors influenced the number of questions that were unanswered: those with lower reputation had much higher number of questions unanswered than those with higher reputations. In their study of Stack Overflow, Hassan et al. (2018) classified the questions into basic, intermediate and advanced based on the difficulty level and the level of knowledge needed to answer the questions. They found that basic questions got more votes for accepted answers and had higher views, whereas advanced questions took longer to resolve and had much lower views.

Data obtained through the Stack Exchange repository could be insightful in identifying factors that influence the outcomes. However, using this available data it is difficult to investigate cognitive and behavioral factors and the dynamic process of interaction among individuals that give rise to the emergent patterns such as the one captured in figures 1-3[1]. Given the complexity of the phenomenon, it is challenging to study the dynamics of submission of answers by a large number of contributors for questions asked by many individuals over time. Agent-based modeling (ABM) enables investigation of how agent behavior at the micro-level, such as decisions being made by agents, could give rise to patterns at the macro-level. In the study reported in this paper, we used an agent-based modeling (ABM) approach implemented using NetLogo (Wilensky, 1999) to address our following research question.

RQ: *Why do some questions on Stack Exchange receive many answers whereas many questions receive no answers or very few answers?*

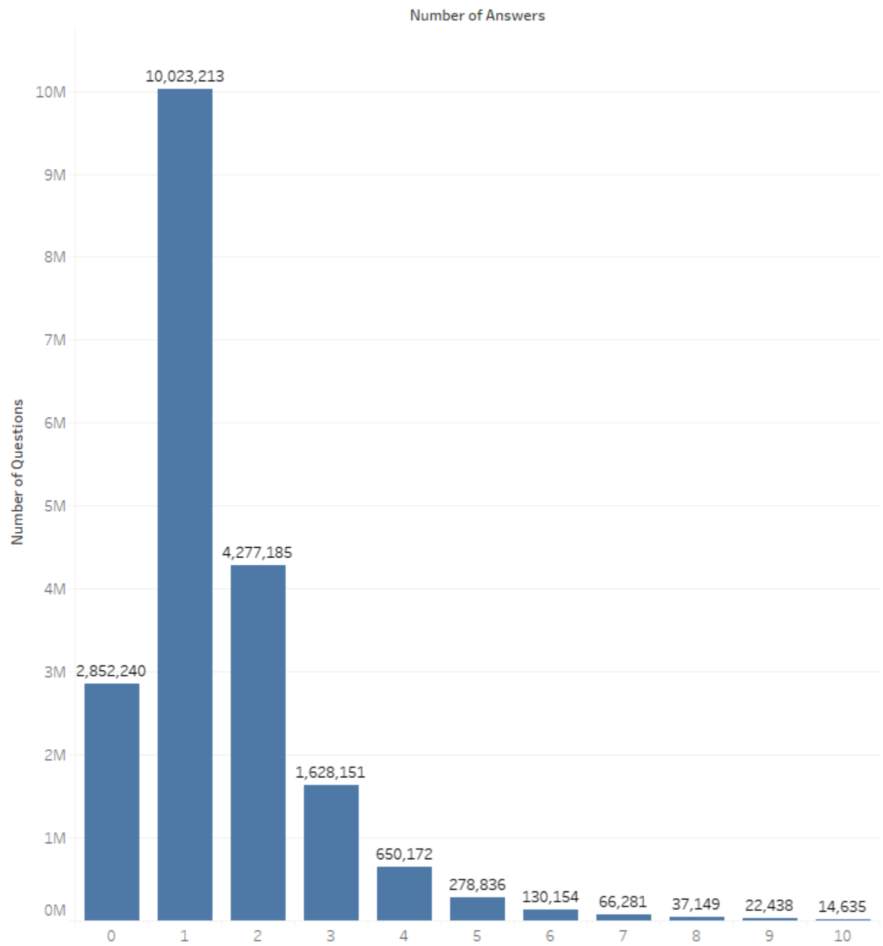

Fig. 1. Number of questions posted in years 2008-2020 with no answers or more than one answers
Note. Only questions receiving 10 or less answers were included since proportionally fewer questions had received more than 10 answers.

---

[1] Fig. 1-3 and Fig. 5 were created using Tableau visualization software

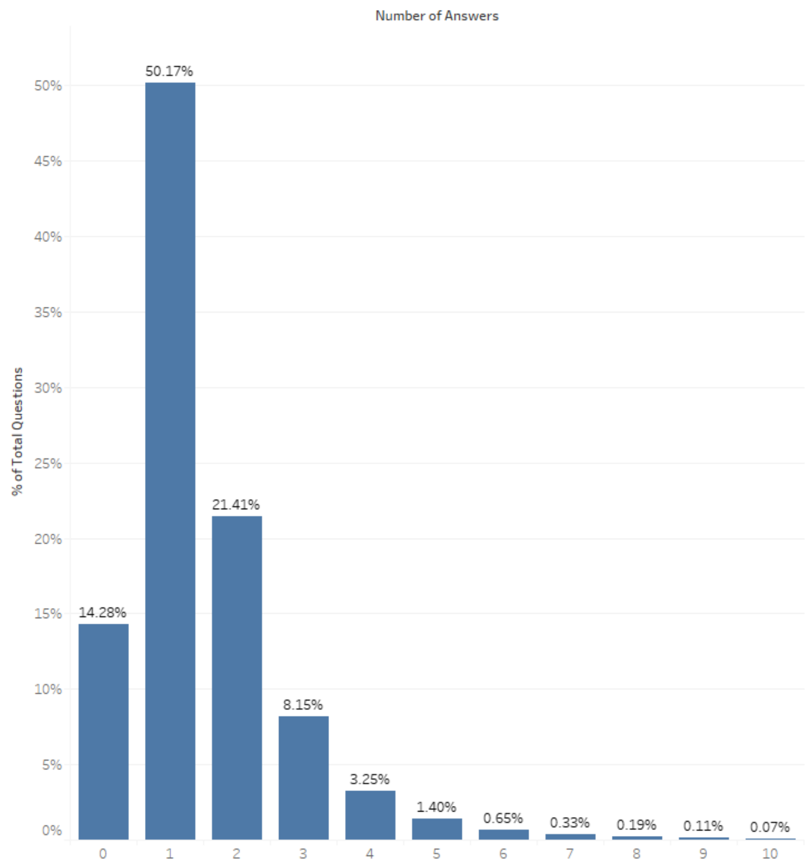

Fig. 2. Proportion of questions posted in years 2008-2020 with no answers or more than one answers
Note. Only questions receiving 10 or less answers were included since proportionally fewer questions had received more than 10 answers.

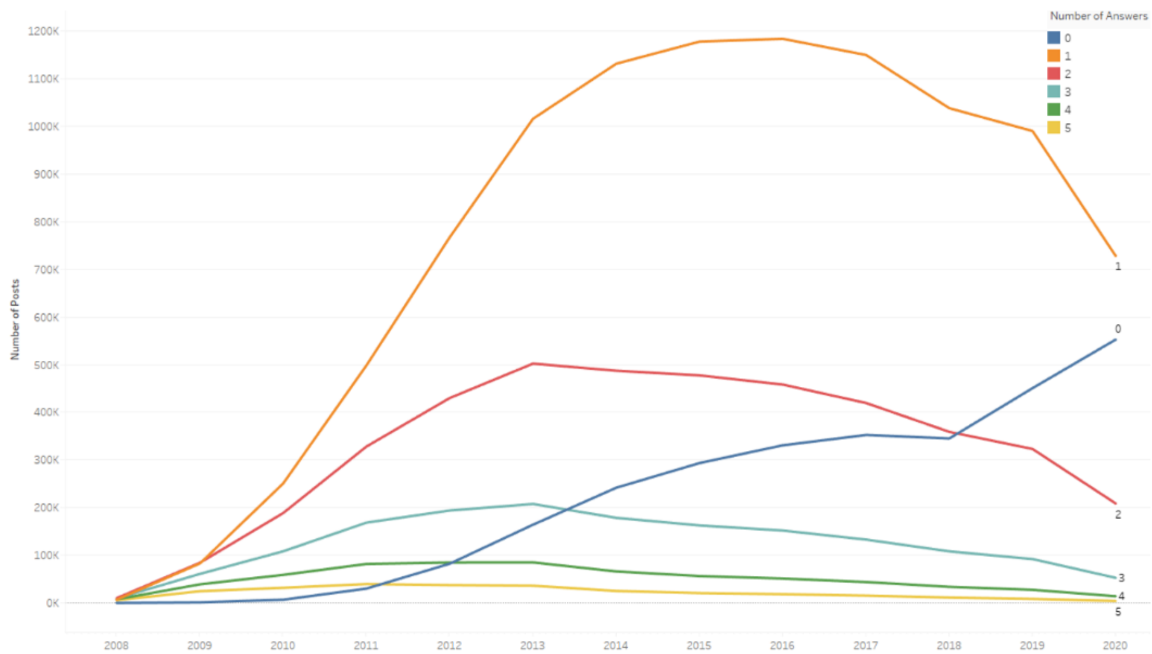

Fig. 3. Trend in number of questions without answers from years 2008 - 2020
Note. Only questions receiving 10 or less answers were included since proportionally fewer questions had received more than 10 answers.

2 Simulation Model

In order to address the research question, an agent-based model was implemented in NetLogo (version 6.1.1). It was decided to build the model gradually to manage model complexity. A 101 x 101 torus grid was created. Ten patches represented questions such that each question belonged to a different topic indicated by the color assigned to that patch. Each question had a degree of complexity or difficulty level with a range of [0, 100] such that lower values indicated lower complexity of a question. 500 agents were created who were potential contributors to the questions. Agents (contributors) had heterogeneous expertise and heterogeneous interests. While collectively the agents were interested in many topics (ten topics in total), it was assumed that each agent was interested in a single topic represented by one of ten colors assigned to the agent. These were the same ten colors assigned to the patches. Agents' degree of expertise in the topic of their interest was indicated by a value drawn from a Gaussian (normal) distribution with a mean value of 50 and standard deviation of 10 with higher value indicating higher expertise. The assumption that the expertise of contributors was normally distributed can be considered reasonable and is similar to other studies e.g. Herbrich, Minka, and Graepel (2007) in the context of online game players and Pluchino, Biondo, and Rapisarda (2018) in a general context. The agents (contributors) moved around randomly on the grid. When a contributor agent came across a question that matched their interest, i.e. the agent was situated on a patch with the patch color matching its own color, it represented a potential contributor reading a question on Stack Overflow that matched their interest. The agent would contribute an answer to this question depending on the contribution rule discussed below.

2.1 Contribution Rules

It was assumed that when agents came across a question that matched their interest, the agents would assess the complexity of this question and compare it with their own expertise. Two rules were considered to determine how agents may contribute an answer to a question. The first rule determined that agents contributed an answer to a question if their own expertise was higher than the complexity of the question i.e., if agents were able to answer a question, they contributed an answer since answering questions would increase the reputation of agents. Building reputation is an important source of motivation for contributors (Anthony, Smith, and Williamson, 2009; Nam et al., 2009). Thus, it was assumed that agents would attempt to maximize the opportunity to build their reputation. Since Stack Exchange gives the same points for easy or difficult questions, answering easier questions is a quicker way to build reputation. The second rule assumed that agents contributed an answer only when their expertise closely matched with the complexity of a question i.e., agents will not contribute to questions that are either too easy or too difficult for them. This rule draws upon the flow theory (Csikszentmihalyi, 1982, 1988) which suggests that individuals experience a 'flow' state of optimal experience when there is a balance between the perceived challenge or difficulty of a situation (task demands) and perceived skills (ability to meet the task demands).

If a contributor answered a question, it was indicated by a patch with the same color as the patch that represented a question. Hence, each answer would "grow" or will be "stacked" on top of the previous answer and provide a bar-chart style visual indicator of how many questions had been answered at any point in time. Figure 4 below shows the result of a simulation run with 500 time steps with 10 questions each represented by a different patch color and the number of answers contributed to each question. It can be inferred from the height of the bars in figure 5 that some questions got more answers than others. This illustrates the role of randomness in how contributors find questions related to topics of interest and whether they provide answers to those questions.

Fig. 4. Visual pattern of questions answered

Chua and Banerjee (2015) had found that when questions were posted during business hours, they were more likely to be answered which indicates that there is a narrow window during which questions have greater likelihood of receiving answers. Since contributors on Stack Overflow may be interested in a topic for a limited duration, the simulation was run for 500 time steps (ticks) to mimic the limited time available to answer a question. Ten simulation runs were performed (N = 180 runs total) under each contribution rule and for each level of question complexity using the *nlrx* package (Salecker, Sciaini, Meyer, and Wiegand, 2019) for R statistical software (R Core Team, 2017) which allows running NetLogo simulations from R.

3 Results and Discussion

The model creates two different predictions on the pattern that can be expected to arise based on the assumptions of when contributors provide answers. If contributors contribute when the difficulty of a question is lower than their ability, it can be expected that easier questions will have many answers and difficult questions will have fewer answers.

Figure 5 shows the aggregate results of the simulation runs under each contribution rule and varying levels of question complexity. Under the assumption that agents only contribute when the question difficulty matches with their expertise level, questions that are too easy or too difficult do not get many answers. When it is assumed that contributors answer questions as long as their ability exceeds the difficulty, it can be seen that easy questions get more answers and there is a steep decline for complex questions.

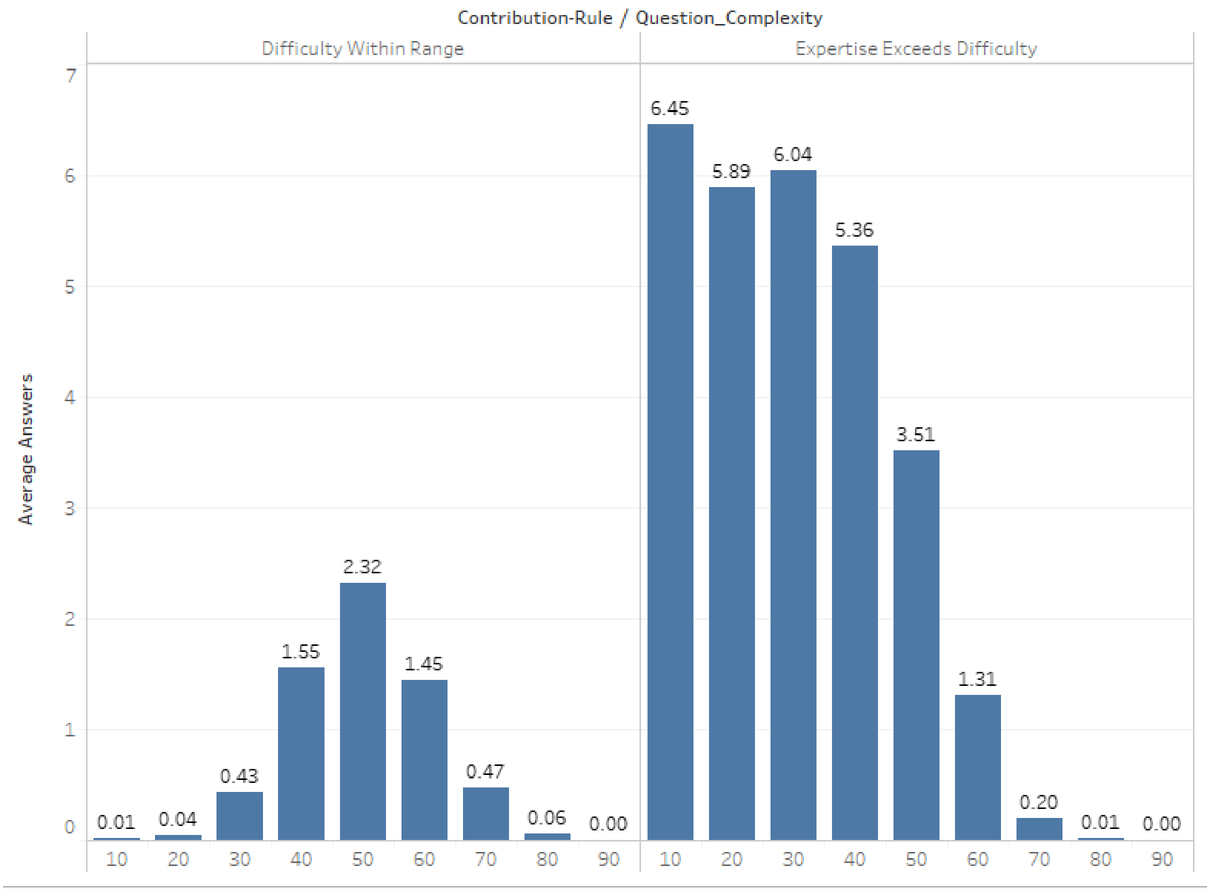

Fig. 5. Average number of answers received for varying complexity of questions and under two contribution rules

3.1 Limitations and future work

Our initial model of contribution behavior at Stack Overflow is based on two simple rules based respectively on the reputation motivation and flow theory. We plan to develop this model further to incorporate aspects of Stack Overflow (e.g., number of views received by a question, number of answers already received, and votes received by questions and answers) based on user behavior that has been observed in prior empirical research. To advance knowledge in this domain, we also need to test rival models of contribution behavior and compare their predictions with observed patterns at sites such as Stack Overflow.

4 Conclusion

A preliminary model of Stack Overflow, a popular community question answering system was created that focused on understanding why many questions remain unanswered on Stack Overflow. This model does not aim to be comprehensive to include all the important factors that lead to questions not being answered. It shows how randomness, distribution of expertise and varied interests in topics may play a role in the outcome: whether questions attract many answers or none. In our analysis of Stack Overflow data, we also found that questions that had received fewer answers also had received fewer views whereas questions with many answers have received many views. We plan to develop this model further to incorporate aspects such as number of views received by a question, number of answers already received, and votes received by questions and answers.

References

Anthony, Denise & Smith, Sean & Williamson, Timothy. (2009). Reputation and Reliability in Collective Goods: The Case of the Online Encyclopedia Wikipedia. Rationality and Society. 21. 281. 10.1177/1043463109336804.

Asaduzzaman, M., Ahmed Shah Mashiyat, Chanchal K. Roy, and Kevin A. Schneider. 2013. Answering questions about unanswered questions of stack overflow. In Proceedings of the 10th Working Conference on Mining Software Repositories (MSR '13). IEEE Press, 97–100.

Chua, A. Y. K., & Banerjee, S. (2015). Answers or no answers: Studying question answerability in Stack Overflow. Journal of Information Science, 41(5), 720–731. https://doi.org/10.1177/0165551515590096

Csikszentmihalyi, M. (1982). "Towards a psychology of optimal experience," in Annual Review of Personality and Social Psychology, Vol. 3, ed. L. Wheeler (Beverly Hills, CA: Sage), 13–36.

Csikszentmihalyi, M. (1988). "The flow experience and its significance for human psychology," in Optimal Experience: Psychological Studies of Flow in Consciousness, eds M. Csikszentmihalyi and I. Csikszentmihalyi (Cambridge: Cambridge University Press), 15–35.

Hassan S. A., Dipto Das, Anindya Iqbal, Amiangshu Bosu, Rifat Shahriyar and Toufique Ahmed, SOQDE: A Supervised Learning based Question Difficulty Estimation Model for Stack Overflow. In Proceedings of the 25th Asia Pacific Software Engineering Conference, APSEC'2018, Nara, Japan.

Nam, K. K., Ackerman, M. S., & Adamic, L. A. (2009). Questions in, Knowledge iN? A study of Naver's Question Answering Community. ACM Conference on Human Factors in Computing Systems (CHI'09).

Pluchino, A., Biondo, A. E. and Rapisarda, A. (2018). Talent versus luck: The role of randomness in success and failure. Advances in Complex Systems, 21, 3-4, 1850014. https://doi.org/10.1142/S0219525918500145

Herbrich, Ralf, Tom Minka, and Thore Graepel. TrueSkill(TM): A Bayesian Skill Rating System. In Advances in Neural Information Processing Systems, pages 569–576. MIT Press, January 2007. URL: https://www.microsoft.com/en-us/research/publication/trueskilltm-a-bayesian-skill-rating-system/.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Saha RK, Saha AK and Perry DE. Toward understanding the causes of unanswered questions in software information sites: A case study of Stack Overflow. In: Proceedings of the joint meeting on foundations of software engineering. New York: ACM, 2013, pp. 663–666.

Salecker J, Sciaini M, Meyer KM, Wiegand K. The nlrx R package: A next-generation framework for reproducible NetLogo model analyses. Methods Ecol Evol. 2019;2041-210X. https://doi.org/10.1111/2041-210X.13286.

Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Yang L, Bao S, Lin Q et al. (2011). Analyzing and predicting not-answered questions in community-based question answering services. In: Proceedings of the Association for the Advancement of Artificial Intelligence conference. Palo Alto, CA: AAAI, pp. 1273–1278.