An Empirical Study on the Influence of Social Interactions for the Acceptance of Answers in Stack Overflow

Zhang Zhang†⋆, Xinjun Mao†⋆∗, Yao Lu†⋆, Shangwen Wang†, Jinyu Lu†
†*National University of Defense Technology, Changsha, China*
⋆*Key Laboratory of Software Engineering for Complex Systems, Changsha, China*
{zhangzhang14, xjmao, luyao08, wangshangwen13}@nudt.edu.cn, jinyu_smile@foxmail.com

Abstract*—In knowledge-sharing communities like Stack Overflow (SO), users can post questions, give answers and choose one answer as an accepted answer. The accepted answers will be important references for users when they encounter similar questions. Essentially, posting questions and giving answers is an interactive process occurring among community users, and choosing accepted answers is actually a decision-making process involving multiple factors. Previous works examined the impact on this decision process from the user, question and answer viewpoints. Social interactions between the questioners and answerers, although being popular according to our pre-analysis, have never been considered as a factor that can influence the decisions. To fill this gap, this paper first proposes a comprehensive answer acceptance model that integrates the answer features established by social interactions as well as information of users, questions and answers. We then divide social interactions into two stages and propose a method to calculate the relationship between the questioner and the answerer by analyzing these social interactions. Finally, we investigate the influence of social interactions for the acceptance of answers by performing logistic regression analysis. The results reveal several findings: (1) social-based features explain 16.6% of the variance explained together, indicating that social interactions have significant and important effects on the acceptance of answers; (2) social interactions that occur after the answer is posted are more influential than these occur before the answer is posted. Based on the findings, we further conduct an online study of 132 SO users, and the respondents report that social interactions have a greater impact on the acceptance of answers than other judgments of answers such as upvotes, downvotes and not accepting answers.

***Index Terms*—answer acceptance model, Q&A community, Stack Overflow, social interaction**

I. INTRODUCTION

Software development and maintenance are complex activities that often involve many concepts and reference documents. Developers tend to spend much time finding solutions to their problems, as a result, the solutions from others are effective ways for them [1]. Recently, Question and Answer (Q&A) communities have emerged as popular media for online knowledge sharing, providing a large amount of information that can be referred by users when meeting problems [2] [3]. Among these Q&A websites, SO is one of the most famous and popular social Q&A websites [4] [5]. Users can post questions for help or help their peers by answering their questions.

The questioner can choose an answer as the accepted answer in SO, meaning that this answer helps him or her solve the problem. And when users encounter similar questions, they prefer to refer to these accepted answers to solve the problems [6] [7]. As a result, the choices of accepting answers are of great importance and can affect the users' impressions for the quality of answers. It is very necessary to study the factors that affect the acceptance of answers [8] and utilize these factors rationally to ensure the quality of the accepted answers. Previous studies have found that some factors can influence the likelihood of an answer being accepted, such as the content of the answer [9], score [10], response time [11], code snippet [12], etc.

Besides being a knowledge base, SO is also a community where users can interact with each other through a variety of behaviors including questioning, answering, commenting and voting. Users will gradually form an impression of each other through social interactions [13], and prior studies have pointed out that social interaction is likely to affect the decision in the programming-related platform [14] [15]. However, social interactions between questioner and answerer in SO have never been considered as factors affecting users' acceptance of answers in previous works. Motivated by this, we perform a pre-analysis including more than 74 million comments and 29 million answers, and we find the phenomenon that questioners and answerers have social interactions is common, indicating that studying the influence of social interactions for the acceptance of answers is necessary. Based on this finding, we propose to perform an empirical study to understand the influence for the acceptance of answers brought by social interactions comprehensively. In particular, we seek to answer the following research questions in this paper:

- RQ1: Do social interactions affect users accepting answers? If so, to what extent?
- RQ2: How do users perceive the impact of social interactions on acceptance of answers?

To answer RQ1, we propose a comprehensive answer acceptance model that integrates 16 answer features established by social interactions as well as information of users, questions

∗ Corresponding author.
This work was supported by National Science Foundation of China under granted number 61532004

and answers, then we propose a method to calculate the social relationship among users, and investigate the influence brought by social interactions for accepting answers by performing logistic regression analysis. Our results suggest that: (1) 40% of the top 10 features ranked by Sum sq. are social-based features, which show a significant and important impact on the acceptance of answers; (2) social interactions that occur after the answer is posted are more influential than these occur before the answer is posted.

To answer RQ2, we perform online survey by sending emails to 1000 respondents, and obtain 132 responses about (1) the impact of social interactions on four judgments of answers; (2) the effect of social-based features and other features on acceptance of answers; (3) the polarity of impact of social interactions. From the participants' opinions, social interactions between questioners and answerers have a greater impact on the acceptance of answers than other judging actions including upvoting, downvoting and not accepting answers. They select the social-based features as the least influential features. 47.3% of them think that the impact of social interactions is more positive, however, 14.86% of them think it is negative.

To summarize, we make the following contributions:

- we bring the attention to one important yet neglected factor that can affect questioners accepting answers in SO, the social interactions between the questioners and the answerers;
- we provide evidence that it is common for questioners and answerers to have social interactions in the past in SO;
- we corroborate the important role of the social interactions between the questioner and the answerer in accepting answers by performing logistic regression analysis with 0.99 million observations, and with an online survey of 132 SO users, and the answer features related to the social interactions explain about 16.6% of the variance explained in our logistic regression analysis.

II. Background and Related Work

The questioners post questions, look through a list of candidate answers and choose one answer as the accepted answer in SO. Many previous works have found that much information related to the content of question and answer can affect this process, such as structure of the answer [12], etc. Roy *et al.* [16] introduce a new tab called `promising answers tab` where answers are listed based on their usefulness, and several textual features of answers established by content of answers are used as features to predict the usefulness of the answers. The results are validated with good values of precision, recall, F1-score, area under the receiver operating characteristic curve (AUC) and root mean squared error. Sun *et al.* [17] propose an answer quality evaluation algorithm based on semantic. They firstly obtain seed keywords in a certain field that are firstly obtained by the topic model and expand the seed keyword in two ways: context and synonyms. The experimental result shows that the accuracy of answer quality classification can be effectively improved.

Islam *et al.* [18] have utilized the textual features of the answers' comments with the other metadata of the answers to building the recommender system for predicting the accepted answer, and their system has achieved 89.7% accuracy to predict the accepted answer by utilizing the textual metadata as a feature. Bodke *et al.* [19] propose a system that finds the best answer from all the available answers based on answer content and the relationship between question and answer. They train the system using three different classifiers and note the accuracy results of each; the classifier which gave maximum accuracy is selected.

Instead of focusing on information about questions and answers, some researchers propose the assumption that authoritative users tend to produce high-quality answers and use it to rank answers [20] [21]. To identify the user criteria and data-driven features, both textual and non-textual, for assessing the quality of answers, Fu *et al.* [22] propose a total of 23 user criteria and 24 data features proposed and test them with high-quality answers obtained from four social Q&A sites in Stack Exchange. Findings indicate that content-related criteria and user and review features are the most frequently used in quality assessments, while the importance of user criteria and data features is variable across the knowledge domains. Elalfy *et al.* [23] propose a hybrid model for predicting the best answer. The proposed model is consisting of two modules. The first module is the content feature, and the second module uses the non-content feature to predict the best answers by using a novel reputation score function. The prediction accuracy of the whole model is very promising.

Users interact with each other in the process of participating in the community, and gradually form an impression of each other. However, social interactions between the questioner and the answerer have never been considered as factors affecting users' acceptance of answers in previous works.

III. Motivation and Research Questions

A. Motivation

In this section, we demonstrate the motivation of our study. The answers in SO can be marked as *accepted* by the questioner, indicating that he/she thinks this answer can solve the question. When users encounter similar problems, these *accepted* answers can serve as an important reference for them to solve the problems. Thus, the decision process of accepting answers is important for community development. Prior studies focus on factors that influence the decision of accepting answers from the perspective of questions, answers, users, such as user reputation [21], answer textual features [20], answer readability [24], response time [11], etc. By characterizing the influence of these factors, on the one hand, SO can personally recommend appropriate answers for users [18]; on the other hand, by controlling these factors, SO can make the questioner judge the answer more objectively and fairly, and decide whether to accept the answer based on the quality of the answer. Such a community atmosphere will not only help to highlight high-quality answers, but also motivate users with higher expertise to make more contributions [25]. So it is necessary to study the factors that can influence the decision process of accepting answers in SO.

Prior studies have pointed out that social interaction is likely to affect the decision in the programming-related platform. For instance, Tsay *et al.* [14] have observed this in the context of the evaluation of GitHub contributions, while Calefato *et al.* [15] related previous social interaction to trust. Due to the influence of social interactions for the decision in other communities, we aim to investigate whether social interactions between the questioner and the answerer are popular in SO. It is necessary to study the impact on the decision of accepting answers brought by this phenomenon if the answer is positive.

Due to the protection of user privacy by SO, we cannot obtain voting data of users, so the social interactions we study in this paper include commenting and answering. Firstly, we divide the time of the social interactions between the questioner and the answerer into two stages:

- **Stage 1.** From the time SO is established to the time point when the answer is posted.
- **Stage 2.** From the time point when the answer is posted to the time point when the questioner marks *accepted* answer.

The social interactions that occur in the **Stage 2** are often the communication between the questioner and the answerer through comments on the answer, which is popular according to works of [26]. Then we study the popularity of social interactions occur in **Stage 1** in the past ten years by analyzing data of SO from July 31, 2008 to December 31, 2019[1]. The dataset includes 19 million questions, 29 million answers, 74 million comments and 12 million users.

For an answer, if there is answering interaction or commenting interaction between the questioner and the answerer before the answer is posted, the answer is called *Type-1 answer*, otherwise, it is called *Type-2 answer*. The change in the proportion of *Type-1 answers* to all answers in the past ten years is shown in Fig.1. For each year, we only analyze the answers that are posted this year. As we can see, the proportion of *Type-1 answers* is showing an overall upward trend (from 0.125 in 2010 to 0.134 in 2019), meaning that there is about one *Type-1 answer* for every seven answers in 2019. Thus, we conclude that social interactions between the questioner and the answerer in **Stage 1** are also common.

Overall, for the answers in SO, social interactions between the questioner and the answerer are rather common. So it is of great value to study the impact of social interactions on the acceptance of answers considering the influence of social interactions on other programming-related platforms in decision-making.

B. Research Questions

Based on the finding of our pre-analysis, we propose to study the impact of social interactions on the acceptance of answers. In particular, we design the following two questions from the method of quantitative analysis and qualitative analysis:

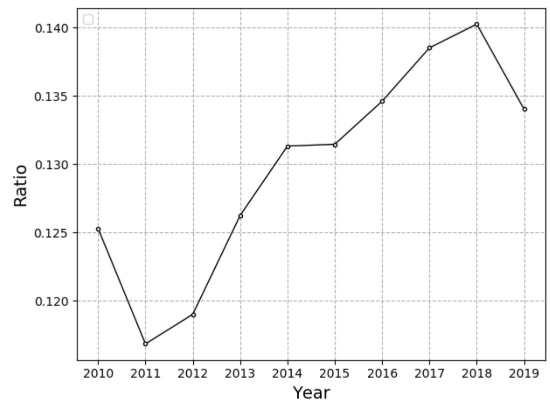

Fig. 1. The trend of proportion of *Type-1 answers*.

> ***RQ1:*** *Do social interactions affect users accepting answers? If so, to what extent?*

In prior studies [12] [20] [27] [28], some features have been found that they have a great impact on accepting answers. We first want to investigate whether social interactions affect accepting answers by comparing the features established by social interactions with features in prior studies, and the method of regression analysis is used in the process.

> ***RQ2:*** *How do users perceive the impact of social interactions on acceptance of answers?*

Users play an important role in the development of Q&A websites. We want to understand users' perceptions about effect of social interactions. We design a subset of questions in our survey to (1) compare the impact of social interactions on acceptance of answers with that on other judgments including upvoting, downvoting and not accepting answers. (2) compare the impact of answer features established by social interactions between questioners and answerers with that of other answer features on acceptance of answers. (3) understand users' attitudes towards the polarity of the impact of social interactions.

IV. Study Design

In this section, we will describe the design of our empirical study, which consists of the **dataset**, **answer acceptance model** and **implementation**.

A. Dataset

For RQ1, since the official data released by SO is as of May 31, 2020[2], we select all question-answer pairs from January 1, 2020 to May 31, 2020 for logistic regression analysis. For each question-answer pair, we remove the answer that is answered by the questioner and the answers that are posted

[1] https://archive.org/details/stackexchange

[2] https://archive.org/details/stackexchange

after the questioner mark *accepted* answer (from 1.08 million question-answer pairs to 0.99 million question-answer pairs). There are around 0.29 million *accepted* answers in the dataset. Considering that the data in a shorter period of time does not fully reflect the user's social interaction, and the data in a longer period of time may interfere with the timeliness of the reflected user's social interaction, we add the data for the two years before 2020 to calculate the social relationship among users. So the social relationship among users is calculated by analyzing the data from January 1, 2018 to December 31, 2019.

For RQ2, according to our experience and reference to prior study [29], the response rate of the questionnaire is usually between 10% and 30% and we think 100 responses can represent the views of community users to a certain extent. Thus, our target population comprises 1000 users. SO doesn't provide users' email addresses in their dataset, and active users in Github often post questions for help or help their peers by answering their questions in SO [30]. So we crawled the email addresses of active users in Github and sent 1,000 questionnaires to them through the SurveyMonkey platform[3].

B. Answer Acceptance Model

In this section, we introduce how we build the answer acceptance model in detail. The model is based on the information received by the questioner when browsing the answers and it integrates the answer features related to social interaction history and answer features not related to social interaction history. As is shown in Fig.2, when questioners look through answers in SO, the information they can get include content-related information (e.g., text and code), user-related information (e.g., reputation), competition-related information (e.g., the number of answers to the question), peer-judgment related information (e.g., comments and score), and they can also get the social-interaction-history related information: the history of social interactions between the questioners and the answerers, including commenting, questioning, answering and voting activities. We will build the model based on the above information.

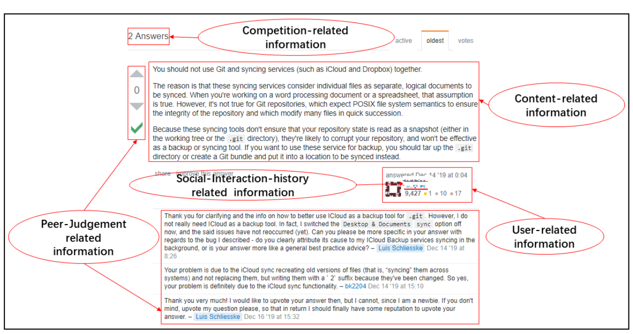

Fig. 2. An example of an answer in SO.

[3]https://www.surveymonkey.com/

As is shown in Fig.3, this model consists of social-based answer features and non-social-based answer features. The former features are extracted from social-interaction-history related information, and the latter features are extracted from user-related information, content-related information, peer-judgment related information and competition-related-related information. The details of all answer features are described in Table I and Table II. Each table shows for each feature (a) its name, (b) its description, and (c) a reference meaning this feature acted as an important feature related to the quality of the answer, or *New* if it's not used before.

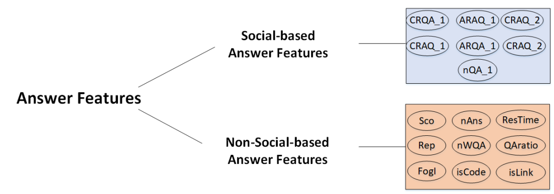

Fig. 3. Answer acceptance model.

Social-based features. For the first time, this paper considers applying answer features based on social interactions, called social-based features, to the study of affecting users choosing accepted answers. Considering the directions of social interactions, we extract features based on the social interaction between the questioner and the answerer at different stages. We consider social interactions of answering and commenting in **Stage 1**: by referring to the method of [31], we use the intimacy characteristics established by answering as two of social-based features (`ARAQ_1`, `ARQA_1`), considering the frequency and timestamp characteristics of social interaction; we use similar method to extract the intimacy characteristics established by commenting (`CRAQ_1`, `CRQA_1`). We consider comments on the answer for **Stage 2**: we use the intimacy characteristics established by these comments as two of social-based features (`CRAQ_2`, `CRQA_2`). In addition, the answers can be divided into *accepted* answers and *unaccepted* answers, and these two answers will give the questioners different impressions. Thus, we choose the number of *accepted* answers in **Stage 1** as one of social-based features (nQA_1). We describe the calculation methods of these features in **Extract social-based features** of Section C **Implementation**.

Non-social-based features. Previous studies have found that some answer features can influence the likelihood of an answer being accepted. In order to study the influence of social-based features on accepting the answers, we select some answer features that have a non-negligible influence on the decision process of users accepting answers in CQA from previous studies. The details and the calculation methods of these features are described in the corresponding reference in Table II.

C. Implementation

In this section, we describe the implementations of methods used to answer the research questions.

TABLE I
SOCIAL-BASED FEATURES

Feature	Description	Ref
nQA_1	the number of answers posted by the answerer that are accepted by the questioner in **Stage 1**	New
CRQA_1	relationship established by the questioner commenting on the answerer in **Stage 1**.	New
CRAQ_1	relationship established by the answerer commenting on the questioner in **Stage 1**.	New
ARAQ_1	relationship established by the answerer answering the questioner's question in **Stage 1**.	New
ARQA_1	relationship established by the answerer answering the questioner's question in **Stage 1**.	New
CRQA_2	relationship established by the questioner commenting on the answerer in **Stage 2**.	New
CRAQ_2	relationship established by the answerer commenting on the questioner in **Stage 2**.	New

TABLE II
NON-SOCIAL-BASED FEATURES.

Feature	Description	Ref
Sco	the score of answer	[10]
Rep	the reputation of user	[21]
nWQA	number of non-stop word overlap between question text and answer text	[20]
QAratio	ratio of question length to answer length	[20]
nAns	the number of answers to the question	[27]
FogI	the Gunning Fog Index of answer text	[24]
isCode	Whether there is a code snippet in the answer (exist as 1, otherwise 0)	[12]
ResTime	Interval between the time the question was posted and the time the answer was posted	[11]
isLink	Whether there is a link in the answer (exist as 1, otherwise 0)	[28]

1) RQ1: We first introduce the procedure of **logistic regression analysis**: extract social-based features, extract non-social-based features and establish the logistic regression model.

Extract social-based features. For `nQA_1`, we can easily obtain the value of it from answering social interactions between the questioner and the answerer in **Stage 1**.

The methods of calculating the values of other social-based features are based on the method of Yu *et al.* [31]. Relationship established by U_i commenting on U_j's posts (`CRQA_1, CRAQ_1, CRQA_2, CRAQ_2`) can be calculated by Equation 1:

$$cr_{i,j} = \sum_{r=1}^{k} t(i,j,r) \tag{1}$$

where k is the total number of commenting behaviour, the element *t(i,j,r)* is a time-sensitive factor of corresponding comment which can be calculated by Equation 2:

$$t(i,j,r) = \frac{timestamp_{(i,j,r)} - baseline}{deadline - baseline} \in (0,1] \tag{2}$$

where *timestamp(i,j,r)* is the datetime of `r-th` comments posted by U_i on the posts that belongs to U_j. The baseline and deadline are highly related to selection of trainset.

Relationship established by U_i answering U_j's questions (`ARQA_1, ARAQ_1`) can be calculated by Equation 3:

$$ar_{i,j} = \sum_{r=1}^{k} t(i,j,r) \tag{3}$$

where k is the number of answers posted by U_i to questions posted by U_j, and the meaning of *t(i,j,r)* is similar to that of Equation 1, which is calculated by Equation 4:

$$t_{(i,j,r)} = \frac{timestamp_{(i,j,r)} - baseline}{deadline - baseline} \in (0,1] \tag{4}$$

where *timestamp(i,j,r)* is the datetime of the answer posted by U_i to the question `r` poster by U_j.

In our experiment, the social interactions between the questioner and the answerer in **Stage 1** are the social interactions that occurred within two years before the answer was posted (we explained it in the first subsection of Section IV). And the social interactions in **Stage 2** are the social interactions that occur after the answer is posted and before the questioner marks *accepted* answer. For a question, if the questioner marks *accepted* answer, the end time of **Stage 2** of social interactions between the questioner and each answerer of the question is the acceptance time, otherwise, the end time for the questioner and each answerer is the datetime of the most recent activity of the answer posted by the answerer (it is the value of `LastActivityDate` in the table `posts` of the official data set).

Then the calculation of other social-based features is as follows: (*baseline* is two years before the time the answer is posted, and *deadline* is the time the answer is posted for **Stage 1**; *baseline* is the time the answer is posted, and *deadline* is the accpetance time for **Stage 2**)

- `CRQA_1 / CRAQ_1` is the value of relationship established by the questioner / answerer commenting on the answerer / questioner in **Stage 1** according to Eq.1, Eq.2.
- `ARQA_1 / ARAQ_1` is the value of relationship established by the questioner / answerer answering the the answerer / questioner's questions in **Stage 1** according to Eq.3, Eq.4.
- `CRQA_2 / CRAQ_2` is the value of relationship established by the questioner / answerer commenting on the answerer / questioner in **Stage 2** according to Eq.1, Eq.2.

The method of calculating the relationship between the questioner and the answerer has several desirable qualities.

- The time-sensitive factor t is introduced to guarantee that recent comments/answers are more valuable for the relation than old those and the frequency of social interactions is considered.
- Due to the different roles of users in social interaction, we set the relationship to be directional.
- Real-time update. For each questioner and answerer, we analyze all social interactions between them from *baseline* to *deadline*.

Extract Non-social-based Features. For each question and answer pair, we extract `nWQA`, `QAratio`, `FogI`, `isCode`, ResTime and isLink from information of text and code, Rep from information of the answerer, and Sco, nAns from score of the answer and the number of answers to this question respectively.

Establish the Logistic Regression Model. Finally, we get the 16 answer features. These values of answer features are numerical or binary, and the state of the answer is binary (accepted or not accepted). The characteristics of the data are suitable for analyzing the relationship among these variables using logistic regression [32]. We use answer features to act as explanatory variables, and isAccepted (the state of the answer: 1 means that the answer is accepted by the questioner, 0 otherwise) is the response variable.

In regression analysis, few data points may have disproportionate effects on the slope of the regression equation, which can lead to the overfitting problem [33]. We log transform explanatory variables to stabilize the variance and improve model fit [34]. We also normalized the variables before running the regression. To avoid multicollinearity between explanatory variables, we consider the variance inflation factor (VIF) of the set of explanatory variables, comparing against the recommended maximum of 5 (in our case all remained well below 4, indicating the absence of multicollinearity).

2) RQ2: We conduct an online survey about the impact of social interactions on the acceptance of answers in SO by sending emails to users.

Survey Design. We use three multiple-choice questions to collect demographic information, one Likert scale question to compare the impact of social interactions on accepting answers with the impact on other judgments (upvoting, downvoting, not accepting answers), one Likert scale question to compare the impact of social-based answer features with the impact of non-social-based answer features on accepting answers, and one open-ended questions to ask respondents about their opinions about the polarity of this impact. Participation is voluntary and the estimated time to complete the survey is 5-10 minutes. The settings of our survey refer to the works of [29], and the following is our questionnaire:

Q1: Do you consent to participate in this survey? Choices: {Yes, No (Stop filling out this form).}

Q2: How long have you participated in Stack Overflow? Choices: {Less than one year, 1∼3 years, 3∼5 years, more than 5 years.}

Q3: How often do you contribute to Stack Overflow? Choices: {Always, Usually, Sometimes, Rarely, Never.}

Q4: Do you agree that the social relationship influenced your upvoting/ downvoting/ accepting answer/ not accepting answer actions? Choices: {Upvoting actions, Downvoting actions, Accepting answer actions, Not accepting answer actions}; Rating scale: {Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree.}

Q5: When you look through answers, what factors influenced your acceptance? Choices: {Answer length, Answer readability, Code snippet, The relevance of answer and question, Whether the answer solves the problem, Author information (reputation, badges), Peer judgment (score, comments), Social relationship between you and the answerer.}; Rating scale: {Extremly influential, Very influential, Somewhat influential, Not so influential, Not at all influential.}

Q6: Do you think that the influence of social interactions for Stack Overflow is more positive or negative? And please state the reasons.

Respondents. We obtained 132 responses (13.2% response rate): 85.61% of respondents have contributed to SO more than 5 years, 8.33% of respondents have contributed to SO in 3-5 years, a few respondents have experience of fewer than 3 years, and only 1 respondent has never contributed to SO. So users participating in the survey are very familiar with the development of this community and their shared opinions about the impact of social interactions are very persuasive.

V. Results

In this section, we show the results of the empirical study and analyze the results to answer the research questions.

A. RQ1: Do social interactions affect users accepting answers? If so, to what extent?

We first only use non-social-based features to establish logistic regression model, then use social-based-features and non-social-based features to perform logistic regression analysis. And the results are shown in Table III and Table IV: the pseudo R^2 of the former is reported to be 0.343, while the pseudo R^2 of the latter is 0.374, which means about 37.4% of the variability in the data can be explained by this model. And the AUC of the latter is 83.0%, which indicates an acceptable model fit. We can infer from that the social-based features can improve the fitness of the model in explaining accepting answers, and social interactions between the questioner and the answerer affect users accepting answers. The Sum.sq. of all social-based answer features in logistic regression model can also illustrate the impact of social interactions. Then, we compare the effect of social-based and non-social-based answer features from the perspective of Sum.sq.

TABLE III
LOGISTIC REGRESSION MODEL FOR NON-SOCIAL-BASED FEATURES: R^2 = 0.343, AUC = 0.814.

Variable	Coef (S.E.)	Sum sq.
Intercept	-23.1163 (0.0959)***	
Rep	1.3582 (0.0136)***	10008.04***
QAratio	0.3460 (0.0233)***	220.43***
nWQA	2.6045 (0.0290)***	8051.91***
nAns	-8.1730 (0.0333)***	60351.53***
FogI	-0.3038 (0.0273)***	123.69***
isCode	0.2717 (0.0082)***	1100.81***
isLink	0.0492 (0.0056)***	77.15***
ResTime	-1.9870 (0.0161)***	15255.81***
Sco	41.8848 (0.1885)***	49349.96***

***p<0.001,**p<0.01,*p<0.5

We sort the variables in descending order of Sum sq, and Fig.4 shows top 10 features. 40% of them are social-based features: CRQA_2, nQA, ARAQ_1, and CRAQ_1, and all social-based features explained 16.6% of the variance

TABLE IV
LOGISTIC REGRESSION MODEL FOR SOCIAL-BASED FEATURES AND NON-SOCIAL-BASED FEATURES: R^2 = 0.374, AUC = 0.830.

Variable	Coef (S.E.)	Sum sq.
Intercept	-23.0313 (0.0974)***	
nQA_1	4.9708 (0.0785)***	4013.33***
ARAQ_1	-4.3383 (0.0831)***	2728.03***
ARQA_1	2.6807 (0.2651)***	102.24***
CRAQ_1	1.5899 (0.0369)***	1855.32***
CRQA_1	0.8903 (0.0596)***	223.31***
CRAQ_2	1.8585 (0.1926)***	93.13***
CRQA_2	4.7957 (0.0367)***	17039.05***
Rep	1.1903 (0.0140)***	7240.99***
QAratio	0.2416 (0.0238)***	103.16***
nWQA	2.1364 (0.0296)***	5200.01***
nAns	-7.9514 (0.0339)***	55078.91***
FogI	-0.4275 (0.0278)***	235.89***
isCode	0.2529 (0.0083)***	924.22***
isLink	0.0308 (0.0057)***	28.88***
ResTime	-1.9615 (0.0166)***	13963.17***
Sco	42.0411 (0.1916)***	48136.69***

***p<0.001,**p<0.01,*p<0.5

explained together, which shows that *The social interactions between the questioner and the answerer plays an important and non-negligible role in affecting the questioners accepting the answers when compared with the answer features that can influence the decision-making process of accepting answers in previous studies.* As expected, `nQA_1` has a positive effect on the acceptance of answers (2.56% of the variance explained), indicating that *the more the answerer's answers have been accepted by the questioner in the past, the more likely his answer is to be accepted.* However, `ARAQ_1` has a negative effect on the acceptance of answers (1.74% of the variance explained). This effect suggests that *a good relationship established by the answerer answering the questioner's questions is not conducive to the acceptance of the answer*, which is not consistent with our expectations. We speculate that *utilizing the frequency and timestamp to establish the relationship without distinguishing the quality of the answers could make the measurement inaccurate, for example, high-quality answers enhance the questioner' trust in the answerer, while low-quality answers do the opposite*.

As for social-based features related to commenting interactions, `CRQA_1` and `CRAQ_1` have a significant, positive effect on the acceptance of answers, explaining about 1.32% of the variance explained. We infer from the result that *the relationship established by commenting interactions between the questioner and the answerer is positively related to the likelihood of the answer being accepted.* One possible reason is that commenting social interactions between the questioner and the answerer in **Stage 1** are helpful to strengthen the questioner's understanding of the answerer's expertise about specific tags, which can be a reference for the decision of accepting answers.

For features related to commenting interactions in **Stage 2**, `CRQA_2` is the most influential social-based features, explaining about 10.86% of the variance explained. One possible reason is that *the questioner can express his confusion about the answer in the comments*. However, `CRAQ_1` has a weak impact (less than 1% of variance explained), which is not consistent with our expectations. We speculate that we may have errors in selecting the comments of the answerer to the questioner: if the answerer @ the questioner's username in the comment of the answer, the comment is considered to be the answerer's comment on the questioner, otherwise, it will not be considered as the answerer's comment on the questioner.

From the time period of social interactions, the social interactions that occur in **Stage 2** are more influential than those that occur in **Stage 1** in affecting decision of accepting answers. This may be because the social interactions in **Stage 1** tend to form the expertise impression between the questioner and the answerer, while the social interactions in **Stage 2** express the views on the answer to improve the quality of the answer, so that the questioner can get a satisfactory answer. The latter is more focused on the answer itself, so it is reasonable that it has more influence on decision-making of accepting answers.

From the perspective of the types of social interactions, we can see that the variance explained by `nQA_1`, `ARAQ_1` and `ARAQ_2` is about three times as much as that explained by `CRQA_1` and `CRAQ_1`. One possible reason is that answering interactions reflect the experise about specific tags of the answerer more than commenting interactions.

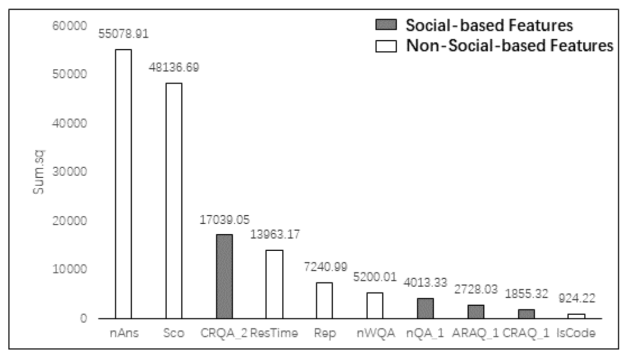

Fig. 4. Top 10 features ranked by Sum.sq.

RQ1: *40% of the top10 features ranked by Sum sq. are social-based features, and they explain 16.6% of the variance explained together, indicating that social interactions have significant and important effects on acceptance of answers. Social interactions that occur in* ***Stage 2*** *are more influential than those that occur in* ***Stage 1*** *in affecting the decision of accepting answers.*

B. RQ2: How do users perceive the impact of social interactions on acceptance of answers?

We compare the impact of social interactions on acceptance of answers with the impact on other judgments of answers (upvoting, downvoting, not accepting answers). And the results are shown in Fig.5. 14% of respondents (8 for *strongly agree* and 10 for *agree*) approve the impact of social interactions between the questioner and the answerer on the acceptance of answers, the percentages for downvoting, upvoting, not accepting answer are 4% (5 for *agree*), 12% (5 for *strongly agree* and 11 for *agree*), 5% (3 for *strongly agree* and 4 for *agree*) respectively. So social interactions between the questioner and the answerer have a greater impact on the acceptance of answers than other judging actions. The judgments of the answers can be divided into positive judgments (accepting the answer, upvoting) and negative judgments (downvoting, not accepting the answer). From this perspective, *more users approve the impact of social interactions on positive judgments than users who approve the impact on negative judgments*. In addition, about 56%-73% of respondents deny the impact of social interactions on the judgments of answers, and about 1/3 of them are neutral.

Then, we compare the impact of social-based answer features with the impact of non-social-based answer features on the acceptance of the answers. The results are shown in Table V. We can find that users consider the relevance of the answer to the question, whether the answer can solve the question, and the readability of the answer as the top 3 features that affect the acceptance of the answer. On the one hand, the first two features are in line with our conclusions of **RQ1**, but the readability of answers is contrary to our results in **RQ1**. This may be because `FogI` is more suitable for measuring text readability, and acts as a poor measure of text and code readability in a programming environment. On the other hand, the social relationship between the questioner and the answerer has the lowest grade, suggesting that the users think that they are less affected by social interactions when they choose to accept the answers. This is inconsistent with our findings of **RQ1**. We speculate that this may be because *users tend to behave as fair, and being influenced by social relationships can make them appear not objective, or users are not aware of the impact of social interactions on them, for example, users can know each other's professionalism and enthusiasm through questioning, answering and commenting social interactions in the past, however, they do not see this perception as a reference provided by social interactions among users.*

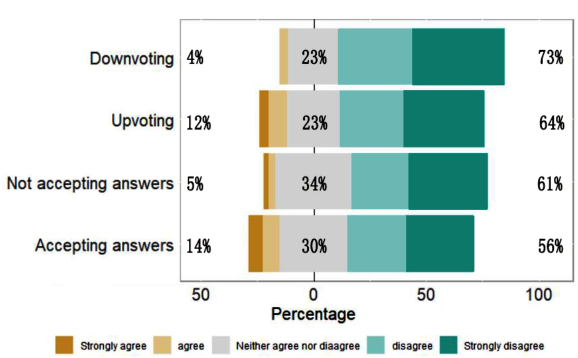

Fig. 5. Responses about the impact of social interaction on four judgments of answers.

Finally, we get 74 valid answers from users about their opinions about the polarity of this impact, and these answers can be divided into 3 types: positive (35 respondents), negative (11 respondents), and neutral (28 respondents). 47.3% of them think that social interactions play a more positive role in community development, for example, it can provide extra information, *"get background information on level of experience and credibility"*, and make answer more perfect, *"Comment discussion is helpful to get clarity on answers"*, and attract more contributors, *"they will appreciate my contribution and would ask me to answer their questions"*, and improve problem-solving skills, *"I learn how to discuss efficiently"*. However, 14.86% of them think that the negative effect is bigger, some respondents think that it can reduce the quality of answers, *"many people posting low effort answers just to get more points to climb higher on the social ladder",* others think it can hurt the fairness of the judgment, *"It's good overall but some users can abuse the system by downvoting interactions of strangers",* and some respondents think that user behaviors in social interactions are terrible, *"Negative because toxic comments/users continue to highlight that the Internet is an unfriendly place"* 37.84% of them are neutral, some respondents think that the impact of social interactions is minimal, *"I don't really chat. I just look at answers"* others think the impact varies with the users, *"Depends strongly on the character of the person",* and some respondents think both positive and negative effects cannot be ignored, *"it's engaging because it's addicting. which is both a positive and a negative."*

> ***RQ2:*** *Respondents think that social interactions have a greater impact on the acceptance of answers than other judgments. They select the social-based features as the least influential features, maybe because they want to behave as fair or they are not aware of the impact of social interactions. 47.3% of them think that the impact of social interactions is more positive, however, 14.86% of them think it is negative. The rest of them are neutral.*

TABLE V
RESPONSES ABOUT THE IMPACT OF SOCIAL-BASED ANSWER FEATURES AND THE IMPACT OF NON-SOCIAL-BASED ANSWER FEATURES ON THE ACCEPTANCE OF THE ANSWERS

	EI	VI	SI	NotSI	NI	**Grade**
Answer length	6	16	67	33	10	2.81
Answer readability	36	68	25	3	0	4.04
Code snippet	33	60	30	7	2	3.87
The relevance of question and answer	82	43	5	2	0	4.55
Whether the answer solves the problem	95	27	11	1	0	4.06
Author information	2	7	34	49	40	2.11
Peer judgement	8	28	42	34	21	2.78
Social relationship between you and answerer	4	1	8	42	77	1.58

EI: Extremely influential.
VI: Very influential.
SI: Somewhat influentila
NotSI: Not so influential.
NI: Not at all influentila

VI. Discussion

A. Threats to Validity

We discuss our limitations in terms of internal and external threats to validity.

Internal Validity. Although the pseudo R^2 of the logistic regression model is 0.374, which is not a perfect fitting degree, what we built is an explanatory model rather than a prediction model, and we just want to study the coefficients' effect and don't explain the full phenomenon. This fitting degree is not a problem for our study by referring to similar studies [32].

In addition, we calculate the relationship between the questioner and the answerer by analyzing two years of user activity history, considering that both too long time periods and too short time periods of social interaction may not be conducive to reflecting the user's social interaction. However, we cannot ensure that 2-years is the best time period for constructing the social network.

As for our online survey, although most participants have contributed to SO over five years and they have a deeper understanding of the community, views from users evenly distributed in each stage of participation may be more representative of the opinions of the entire community. In addition, directly asking about the importance of social features might introduce bias in the respondents. It is much better to list multiple social types specifically, or else create a pool of multiple choices, including the one about social features.

External Validity. Our results only apply to the acceptance of answers in SO, they do not cover users in other social knowledge-sharing communities, like Yahoo! Answers. Our results are limited by mechanism design of SO, and different incentives, voting mechanisms or the questioning and answering mechanism will cause unpredictable changes in the degree of the impact. Still in the future, we would like to extend this study to include more social knowledge-sharing communities to reduce the threat even further.

B. Implications

From our results, we can distill several implications for different stakeholders:

Q&A Websites: According to the results of our study, questioners are influenced by the social interactions among them and answerers when they accept the answers. And the effects are both positive and negative: social interaction can be used as a reference for users' expertise and improve the quality of answers, but it may also affect the fairness of judgment. For the community designers, we provide some suggestions to refer to: anonymous answer (*anonymizing answerers can reduce the impact of social interaction history on questioners' accepting decisions*), give the reason for voting while voting (*Down votes MUST come with the useful comment. Otherwise MUST not allowed*), etc. And these suggestions will not affect the communication between the questioner and the answerer regarding the answer. For the community users, we suggest that developers evaluate the answers to the principles of CQA, that is, to use only the quality of the answers as the evaluation criterion, and to minimize the impact of subjective emotions on this decision-making process. In addition, for the answerer, he should communicate with the questioner after the answer is posted to help the questioner understand the answer or improve the answer.

Researchers: We find that among the options we offer, users choose social relationships as the least influential factor (***RQ2***), which is inconsistent with the conclusion we obtain from ***RQ1***. The reason we speculate is that users tend to behave fairly and being influenced by social relationships makes them appear not objective, or they are not aware of the impact of social interactions on their decisions. It is valuable for researchers to analyze this phenomenon and mine from the reasons why respondents don't choose social relationships.

SE Community: SO is a platform where developers of the SE community can seek help and guidance in programming. And we propose an answer acceptance model to explain the role of answer features in the decision-making behaviors of developers in SO when choosing the best knowledge they think. As a result, the SE community can refer to our model to analyze the decision-making patterns for users to choose the best knowledge they think and recommend the most satisfactory knowledge for them, which is helpful for users in the SE community to quickly acquire the knowledge they need.

VII. Conclusions

The *accepted* answers are important references for users in SO. So the choices of accepting answers are of great importance and can affect the users' impressions for the quality of answers in the community. Essentially, choosing *accepted* answers for questions is a decision-making process involving multiple factors. Prior studies examined the impact on this decision process from the user, question and answer viewpoints. Social interactions between the questioners and answerers, although being popular according to our pre-analysis, have never been considered as a factor that can influence the decisions.

In this study, (1) we first propose a comprehensive answer acceptance model that integrates the social-based features and non-social-based features; (2) we propose a method to calculate the social relationship among users in SO and investigate the influence of social interactions for the acceptance of answers by performing logistic regression analysis; (3) we further conduct an online study of 132 SO users to get their opinions about the impact of social interaction on the acceptance of answers. And we observe the following findings:

- Social interactions have a great impact on the acceptance of the answers, and `CRQA_2` is the most influential social-based feature.
- Social interactions between the questioner and answerer can help the questioner understand the answerer's expertise and the advantages and disadvantages of the answer.
- 47.3% of respondents have positive attitudes towards the impact of social interaction, 30% have negative attitudes, and the rest are neutral.

Based on our findings, in future work, we plan to propose a more accurate way to measure user social relationships using data of voting, commenting, questioning and answering. It can recommend potential answerers and reviewers, helping improve the quality of answers and solve problems. And the answer acceptance model can be improved by integrating more features that influence the process of accepting answers and adjusting the weight of different features, then we can transform our explanatory model to a prediction model that is used to get the result of the decision-making process. In addition, this paper only studied the impact of social interactions between the questioner and the answerer on the acceptance of the answers. We can consider the answerer's sociality on other posts in the future (besides previous social interactions with the same questioner, but also with anybody else). That is, if somebody helped others in the past, she/he also is likely to help me.

References

[1] Chen C, Yang Y, Yang L, et al. A Human-as-Sensors Approach to API Documentation Integration and Its Effects on Novice Programmers[C]//2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2019: 197-206.

[2] Ahasanuzzaman M, Asaduzzaman M, Roy C K, et al. Classifying stack overflow posts on API issues[C]//2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2018: 244-254.

[3] Yang X L, Lo D, Xia X, et al. What security questions do developers ask? a large-scale study of stack overflow posts[J]. Journal of Computer Science and Technology, 2016, 31(5): 910-924.

[4] Liu X, Zhong H. Mining stackoverflow for program repair[C]//2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2018: 118-129.

[5] Silva R F G, Paixao K, de Almeida Maia M. Duplicate question detection in stack overflow: A reproducibility study[C]//2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2018: 572-581.

[6] Guo C, Huang D, Dong N, et al. Deep Review Sharing[C]//2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2019: 61-72.

[7] Zhang Y, Lo D, Xia X, et al. Multi-factor duplicate question detection in stack overflow[J]. Journal of Computer Science and Technology, 2015, 30(5): 981-997.

[8] Suryanto M A, Lim E P, Sun A, et al. Quality-aware collaborative question answering: methods and evaluation[C]//Proceedings of the second ACM international conference on web search and data mining. ACM, 2009: 142-151.

[9] Gkotsis G, Stepanyan K, Pedrinaci C, et al. It's all in the content: state of the art best answer prediction based on discretisation of shallow linguistic features[C]//Proceedings of the 2014 ACM conference on Web science. ACM, 2014: 202-210.

[10] Gantayat N, Dhoolia P, Padhye R, et al. The synergy between voting and acceptance of answers on stackoverflow, or the lack thereof[C]//Proceedings of the 12th Working Conference on Mining Software Repositories. IEEE Press, 2015: 406-409.

[11] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec. 2012. Discovering value from community activity on focused question answering sites: a case study of stack overflow. Proc. of KDD '12. ACM, 850-858.

[12] Dalip D H, Goncalves M A, Cristo M, et al. Exploiting user feedback to learn to rank answers in q&a forums: a case study with stack overflow[C]//Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval. ACM, 2013: 543-552.

[13] Zhang, Y., Wang, H., Yin, G. et al. Social media in GitHub: the role of @-mention in assisting software development. Sci. China Inf. Sci. 60, 032102 (2017).

[14] Tsay J, Dabbish L, Herbsleb J. Influence of social and technical factors for evaluating contribution in GitHub[C]//Proceedings of the 36th international conference on Software engineering. ACM, 2014: 356-366.

[15] Calefato F, Lanubile F. Establishing personal trust?based connections in distributed teams[J]. Internet Technology Letters, 2018, 1(4): e6.

[16] Roy P K, Ahmad Z, Singh J P, et al. Finding and ranking high-quality answers in community question answering sites[J]. Global Journal of Flexible Systems Management, 2018, 19(1): 53-68.

[17] Sun M, Liu L, Chen H. An Answer Quality Evaluation Algorithm Based on Semantic in Community-Based Question Answering[C]//2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). IEEE, 2018: 971-978.

[18] Islam M M, Arafat S S I, Hossain M S, et al. RAiTA: Recommending Accepted Answer Using Textual Metadata[M]//Emerging Technologies in Data Mining and Information Security. Springer, Singapore, 2019: 119-131.

[19] Bodke S, Meher A, Shirsat K. Evaluating Answer Qualities on Q&A Community Sites (StackOverFlow)[J]. Available at SSRN 3367706, 2019.

[20] Agichtein E, Castillo C, Donato D, et al. Finding high-quality content in social media[C]//Proceedings of the 2008 international conference on web search and data mining. ACM, 2008: 183-194.

[21] Bosu A, Corley C S, Heaton D, et al. Building reputation in stackoverflow: an empirical investigation[C]//2013 10th Working Conference on Mining Software Repositories (MSR). IEEE, 2013: 89-92.

[22] Fu H, Oh S. Quality assessment of answers with user-identified criteria and data-driven features in social Q&A[J]. Information Processing & Management, 2019, 56(1): 14-28.

[23] Elalfy D, Gad W, Ismail R. A hybrid model to predict best answers in question answering communities[J]. Egyptian informatics journal, 2018, 19(1): 21-31.

[24] R. Gunning. The technique of clear writing. McGraw-Hill, 1952.

[25] Cavusoglu H, Li Z, Huang K W. Can gamification motivate voluntary contributions? The case of StackOverflow Q&A community[C]//Proceedings of the 18th ACM conference companion on computer supported cooperative work & social computing. 2015: 171-174.

[26] Soni A, Nadi S. Analyzing comment-induced updates on stack overflow[C]//2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR). IEEE, 2019: 220-224.

[27] Tian Q, Zhang P, Li B. Towards predicting the best answers in community-based question-answering services[C]//Seventh International AAAI Conference on Weblogs and Social Media. 2013.

[28] Ponzanelli L, Mocci A, Bacchelli A, et al. Improving low quality stack overflow post detection[C]//2014 IEEE International Conference on Software Maintenance and Evolution. IEEE, 2014: 541-544.

[29] Gousios G, Zaidman A, Storey M A, et al. Work practices and challenges in pull-based development: the integrator's perspective[C]//Proceedings of the 37th International Conference on Software Engineering-Volume 1. IEEE Press, 2015: 358-368.

[30] Badashian A S, Esteki A, Gholipour A, et al. Involvement, contribution and influence in GitHub and stack overflow[C]//Proceedings of 24th Annual International Conference on Computer Science and Software Engineering. IBM Corp., 2014: 19-33.

[31] Yu Y, Wang H, Yin G, et al. Reviewer recommendation for pull-requests in GitHub: What can we learn from code review and bug assignment?[J]. Information and Software Technology, 2016, 74: 204-218.

[32] Zhang Y, Vasilescu B, Wang H, et al. One size does not fit all: an empirical study of containerized continuous deployment workflows[C]//Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. 2018: 295-306.

[33] Wang X T, Shen D R, Bai M, et al. An efficient algorithm for distributed outlier detection in large multi-dimensional datasets[J]. Journal of Computer Science and Technology, 2015, 30(6): 1233-1248.

[34] Bogdan Vasilescu, Yu Yue, Huaimin Wang, Premkumar Devanbu, and Vladimir Filkov. 2015. Quality and productivity outcomes relating to continuous integration in GitHub. In Joint Meeting on Foundations of Softw. Eng.