Answer Quality Characteristics and Prediction on an Academic Q&A Site: A Case Study on ResearchGate

Lei Li
Nanjing University of Science and Technology
No. 200 Xiaolingwei, Nanjing, China
lli@pitt.edu

Daqing He, Wei Jeng, Spencer Goodwin
University of Pittsburgh
135 N Bellefield Ave
Pittsburgh PA
{dah44, wej9}@pitt.edu
spencer.j.goodwin@gmail.com

Chengzhi Zhang
Nanjing University of Science and Technology
No. 200 Xiaolingwei
Nanjing, China
zhangcz@njust.edu.cn

ABSTRACT

Despite various studies on examining and predicting answer quality on generic social Q&A sites such as Yahoo! Answers, little is known about why answers on academic Q&A sites are voted on by scholars who follow the discussion threads to be high quality answers. Using 1021 answers obtained from the Q&A part of an academic social network site ResearchGate (RG), we firstly explored whether various web-captured features and human-coded features can be the critical factors that influence the peer-judged answer quality. Then using the identified critical features, we constructed three classification models to predict the peer-judged rating. Our results identify four main findings. Firstly, responders' authority, shorter response time and greater answer length are the critical features that positively associate with the peer-judged answer quality. Secondly, answers containing social elements are very likely to harm the peer-judged answer quality. Thirdly, an optimized SVM algorithm has an overwhelming advantage over other models in terms of accuracy. Finally, the prediction based on web-captured features had better performance when comparing to prediction on human-coded features. We hope that these interesting insights on ResearchGate's answer quality can help the further design of academic Q&A sites.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval, Information filtering, Selection process

Keywords

User judgment, peer rating, social Q&A, academic social networking, Academic Q&A site, ResearchGate

1. INTRODUCTION

Along with the increasing popularity of social web, Internet users are provided with many social platforms for them to generate, share, and evaluate their own content. Therefore, social web users are both the consumers and producers of content, information, and knowledge.

Social question and answer sites (hereafter: social Q&A site) are one such platform, which allow users to interact with others through asking questions and providing answers, as well as evaluating others' contributions. Besides those well-known generic social Q&A sites such as Yahoo! Answers and Answers.com, we are also witnessing the growth of professional Q&A sites (e.g., Quora) and academic social networking sites featuring Q&A functions. A popular example of the latter is ResearchGate. Compared to generic social Q&A sites, academic social Q&A sites such as ResearchGate engage scholars and researchers in proposing academic-related questions and providing professional quality answers. Through their interactions in the Q&A sites, scholars form a virtual community for themselves and their peers.

Quality is a critical issue for user-generated content, so analyzing and predicting answer quality have been active research topics in studying generic social Q&A [2, 14]. There are also studies on answer quality from a more user-oriented angle, which argued that information seeker's satisfaction should be examined [1]. However, we have not seen studies on examining either answer quality or user's satisfaction associated with academic Q&A sites.

In this paper, we argue that it is imperative to study answer quality in academic Q&A sites. There are several reasons for our position. Firstly, on academic Q&A sites, scholars may ask questions only requested quick comments or answers, but more often their academic questions are complex with multiple facets [9], which require significant professional or disciplinary knowledge in order to provide quality answers. More importantly, scholars can engage in discussions that are exploratory in nature, where there are no widely accepted right answers to some questions. Therefore, academic Q&A sites can be considered as informal scholarly communication platforms for professionals, which is different than generic Q&A sites. Secondly, the responders at the academic Q&A sites are scholars and work in higher education, research institutions, or engage in professional work. They are different to the answerers at the generic Q&A site.

Because of these characteristics in academic Q&A sites, we think that the answer quality should be examined in a peer-based fashion, which resembles the peer-review of academic publications. This helps to cope with the exploratory nature of many questions discussed in these sites.

As shown in Figure 1, ResearchGate has a feature for users to upvote an answer. An upvote in ResearchGate means to "vote to

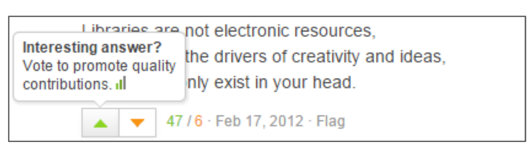

Figure 1. The upvote information on ResearchGate

WWW'15 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2742129

promote quality contributions." Therefore, the number of upvotes for a given answer can be seen as the members of ResearchGate community's peer judgment on that answer's quality.

Following what the literature suggested, we identified two groups of characteristics associated with each answer. The first group of characteristics can be obtained directly from the web content (hereafter called web-captured features). They include: the responder's participatory scores (named RG score on the site), responder's publication impacts (called impact points), responder' institution participatory scores and publication impacts, answer length, and response time. The second group of characteristics was obtained through a content analysis conducted by the authors (hereafter called human-coded features). These include social elements, consensus building (i.e., agreement), and providing factual information, resources, opinion, personal experience or reference to other researchers[19, 20].

One of the goals of this study was thus to study the influence of these two groups of characteristics on the answer's peer-judged quality in ResearchGate, stated as the first research question:

RQ1: What are the relationships between answer characteristics and answer quality based on peer judgments on ResearchGate?

Further, we also wanted to examine whether those identified characteristics can be used for prediction. Therefore, the second research question is:

RQ2: Is the answer quality based on peer judgments at the ResearchGate predictable? Which prediction models perform best? Which groups of characteristics can better predict the answer quality, web-captured or human-coded?

To the best of our knowledge, this is the first study of answer quality based on peer judgments on an academic Q&A site. In particular, we combine the answer web-captured features and human-coded features to predict the answer quality at the academic Q&A site.

2. RELATED WORK

The social web offers novel interactive possibilities to facilitate scholars' collaboration and information access. Researchers have paid increasing attention to the impact of academic social networking sites in disseminating publications, measuring scholarly impact, interacting with other researchers, and engaging non-professionals [11, 12]. Existing studies have explored the usage, motivations [7], utilities [8], and network patterns [9, 10] on academic social platforms such as ResearchGate (https://www.researchgate.net/), Academia.edu, and Mendeley (http://www.mendeley.com/).

Conventional evaluation methods for answer quality on Q&A sites, especially Yahoo! Answers, include responder's reputation tracking [15, 16] and user satisfaction [17] (i.e., peer satisfaction) rating. Harper et al determined how Q&A sites differ in the quality and characteristics of answers to questions, based on two coders coding answer correctness, asker confidence in an answer, helpfulness of an answer, progress towards receiving an answer, and monetary worth of the answer [2]. Researchers also outsourced these judgments by asking Amazon Mechanical Turkers to rate answer quality [3]. When Fichman compared answer quality among four Q&A sites, the concept of answer quality was determined by the reliability as constructed by three specific measures: accuracy, completeness, and verifiability [4]. Researchers also considered multiple variables in order to capture a better understanding of answer quality. A quality framework comprising social, textual, and content-appraisal features of answers was developed and tested by [5, 6].

However, similar inquiries have not yet been addressed in the academic social networking context. Existing research does not combine the answer's web-captured features and human-coded features to evaluate answer quality. Based on the above conclusions, this paper uses the judgment of peers, as evidenced by their votes on answers, to predict answer quality based on both web-captured and human-coded features.

3. RESEARCH METHODS

3.1 Study Site: ResearchGate Q&A

In this study, we selected ResearchGate.net as our testbed. ResearchGate (hereafter: RG) is one of the most well-known academic social network sites with around 5 million users, as of January 2015 [18].

Besides allowing users to disseminating their publications and build personal networks, the most apparent feature on RG is the Q&A platform. The Q&A platform on RG allows users to exchange information by asking and answering questions [9].

A typical question thread is composed of topic tags, asker's name, institution, and RG score. Those who read this question are able to provide an answer to the question. As Figure 1 shows, other users who do not contribute the answer can also rate each answer by using "upvote" or "downvote" buttons. This "Vote to Promote" design pattern is common and similar to other social platforms such as reddit and StackOverflow [12], which helps us to identify the answer quality by peer judgments through these votes on RG.

3.2 The Overall Research Framework

Data analyzed in this paper are a random sample of question and answers sets from RG Q&A and associated data; the answers serve as the initial data set in the answer quality analysis. Three subsets were extracted to form answer web-captured features, answer human-coded features and answer votes information data sets. After this, an analysis of answer quality characteristics was conducted. A prediction model for answer quality was then developed. Naive Bayes model, support vector machine (SVM) and multiple regression models were all tested, and the optimal one can be selected upon the evaluation of their performance in accordance with the test data. The basic framework of answer quality analysis is as shown in Figure 2.

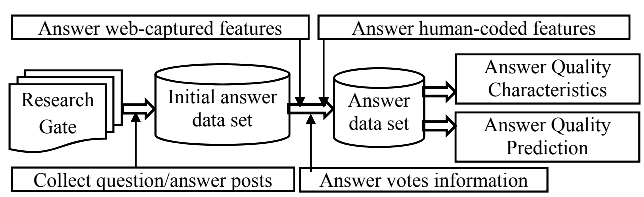

Figure 2. Research framework of answer quality analysis

3.3 Key Steps

3.3.1 Data Collection

In this paper, for analyzing the answer quality on ResearchGate Q&A we collected the answer data and conducted a content analysis in order to scan for human-coded features.

We started with three disciplines to set up an initial data set. We collected the URLs of 38 Q&A threads in the LIS discipline, totaling 413 posts, including posts ranging from September 2009 to November 2013. A detailed documentation of data collection

Table 1. The detailed introduction for each feature

	Features	Descriptions	References
Web-captured features	①RG scores	The RG Score takes all responder research and turns it into a source of reputation, including responder's publications, questions, answers, followers. RG Score is calculated based on how other researchers interact with this responder's content, how often, and who they are. The higher their score, the more this responder's RG scores will increase.	[1] [3] [5] [15, 16]
	②Impact points	The impact points in ResearchGate are auto-generated by adding the responder publications' impact points.	[1] [3] [5] [15, 16]
	③Inst. RG score	Responder's institutional RG score refers to the combined RG scores of all users in this institution based on entirety of ResearchGate data.	[1] [3] [5] [15, 16]
	④Inst. impact points	Responder's institution impact point refers to the combined impact points of all users in this institution based on entirety of ResearchGate data.	[1] [3] [5] [15, 16]
	⑤Response time	Response time refers to the difference between time of answer given and question posed.	[6]
	⑥Answer length	Number of words in an answer.	[1] [3] [5]
Human-coded features	①Social elements	In the answer, a post may contain some social elements that have no direct informational content. These elements may be offering comfort to another user, being polite (such as saying "Hello!" or "Thank you!"), or leaving contact information for offsite discussion.	[19,20]
	②Consensus building	A post has to explicitly state an agreement or a disagreement with an initiator or respondent has been reached through language such as "I agree" or "I disagree".	
	③Factual information	The characteristics of an answer can be specified as providing information to clarify the current knowledge of the question initiator.	
	④Provide resources	The answer provided a URL, citation, or other form of information that is informative to the question.	
	⑤Refer to other researchers	If an answer mentioned other researchers' theory or studies in the post without providing a direct link, we categorized this as referring to other researchers.	
	⑥Provide opinion	An answer may reveal the opinion of the responder.	
	⑦Provide personal experience	An answer may reveal the personal experience of the responder relating to the question.	

can be found in [9]. To extend our dataset, we also selected "History of Art" (311 posts in 33 question sets) and "Astrophysics" (404 posts in 36 question sets), with ranges from November 2012 to August 2014, and from March 2013 to October 2014, respectively. As described in Table 1, for each post we captured responder's name, date and time, RG scores, impact points, as well as the institution's RG score and impact points they belong to.

The overall sample size across three disciplines that we used in the current study was 1128 posts in 107 question threads.

3.3.2 The Establishment of the Answer Quality analysis Data Set

Data sets of answer quality analysis are developed in following steps: first, we obtained question threads as hereinbefore stated to set up the initial data sets for answer quality analysis; then we captured the feature value of all answers according to the web-captured feature of each post collected and the human-coded features of code, and figure out the quality value based on the vote information of each post collected; finally the analysis data sets in regard to answer quality can be developed after the comprehensive consideration of the feature value and quality value of each post. The process of developing analysis data sets of answer quality is shown in Figure 3 where ①, ②, ③, ④, ⑤, ⑥, ⑦stand for features of each post. The following part of this chapter will specifically introduce each feature.

(1) Quality Score

As we described above, the quality value of each answer is defined by the votes it received from other scholars. Posts with 3 upvotes or more were marked as a *high* quality answer, whereas answers receiving 1 and 2 votes were marked *medium*. Answers which did not receive any upvotes were categorized as low quality, because answers which do not receive any promotion may be unjudged or not good enough to give upvotes. We define this group as *low* quality answer. There are 505 answers with defined value of *low*, 438 with defined value of *medium* and 78 with defined value of *high* out of the 1128 posts collected. Figure 1 shows the upvote information appearing to ResearchGate users below each post. ResearchGate also provides a downvote ability, but the downvote information collected was too small, and thus not considered in this paper.

(2) Coding Answer Features Value

Answer features include web-captured features and human-coded features. Answer web-captured features include responder's RG scores and impact points, responder's institution RG score and institution impact points, response time and length of the answer. Answer human-coded features include social elements, consensus building, adding information, provide resources, refer to other researchers, provide opinion, provide personal experience. In our binary coding scheme "1" was marked if the element was present while "0" was marked when the element was absent. The detailed introduction for each feature is reported in Table 1.

Human-coded features were coded in two stages. In the first stage, three coders independently coded an initial set of 100 answers randomly drawn from the data set according to the features detailed in Table 1. The mean pair-wise Cohen's kappa among the coders was found to be 0.83, indicating non-chance level of agreement in human-coded features scores [21]. In the second stage, the entire 1021 answers were coded separately by the three coders in a non-overlapping fashion.

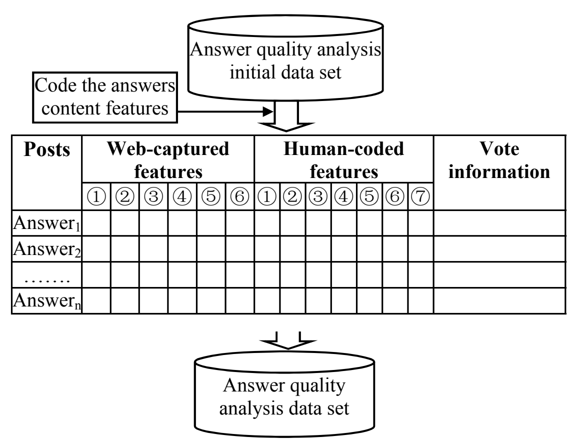

Figure 3. Establishment answer quality analysis dataset

3.3.3 Construction and Selection of Prediction Models for Answer Quality

(1) Construction of Prediction Models for Answer Quality

A prediction model for answer quality depends on all features of each answer. Naive Bayes model, SVM model and multiple regression were selected as candidate prediction models in this study.

Naive Bayes and SVM: Answer quality can be predicted as a problem of classification via naive Bayes model or SVM model. Naive Bayes classifier, which is a very simple and fast, is often a surprisingly effective method to quickly investigate the success of classification problem. Due to the high reported accuracy, support vector machines are the classifier of choice for many tasks.

Multiple Regression: Multiple regression model uses each feature of each answer as the independent variable for the dependent variable of answer quality during regression.

(2) Selection of Prediction Models for Answer Quality

A best prediction model for answer quality can be selected using our quantified data sets. A test on the data sets via N-fold cross-validation over a combination of validation of precision and recall rate (harmonic mean F_1) and AUC scores of the predictions, is able to select the model with highest F_1 and AUC scores as the best model for prediction among tested models.

4. ANALYSIS AND FINDINGS

4.1 Answer Characteristics and Answer Quality

We collected web-captured features for 1021 answers and further conducted content analysis on the answers. The descriptive results for these answers are summarized in Table 2, Table 3 and Table 4.

Table 2. Mean of web-captured features

Features	Low	Medium	High	Total
RG scores (points)	19.566	23.376	31.676	22.123
Impact points (points)	60.565	57.439	86.989	61.243
Inst. RG score (thousand)	10.341	10.893	16.645	11059.494
Inst. impact points (thousand)	14.264	14.257	25.211	15097.179
Response time (normalized hours)	0.493	0.368	0.358	0.429
Length of the answer (words)	75.140	92.296	133.462	86.913

Table 3. Ratio of human-coded features

Features	Low(505)	Medium(438)	High(78)	Total(1021)
Social elements	0.370	0.265	0.192	0.311
Consensus building	0.218	0.265	0.269	0.242
Provide resources	0.271	0.365	0.256	0.310
Factual information	0.368	0.445	0.487	0.410
Provide opinion	0.442	0.596	0.769	0.533
Refer to other researchers	0.109	0.146	0.205	0.132
Provide personal experience	0.089	0.135	0.115	0.111

Compared to medium and low quality answers, high quality answers (N=78) are more likely to have higher RG score (M=31.676, SD=26), more publication impact points (M=87.00, SD=208.38), and be composed by people from institutions that have higher RG score (M=16645). The average normalized response time (normalized response time means response time divided by the maximum value of each post) of answer among high quality answers is 0.358, comparing to low quality answers' 0.493, which suggests that users perceived an answer posted earlier as higher quality than a later answer. However, we did not find a significant correlation between responders' publication impacts points and their answer quality.

Based on our human-coded features (see Table 3 and Table 4), we observed that an answer was more likely to be rated as high quality if it had any of the following features: resource providing, factual information, opinion, referring to others, and personal experience. We also found that more than one-third of low quality answers (N=187 of 505, 37.0%) featured social elements, However, we did not observe the same phenomenon in the group of high quality answers (N=15 of 78, 19.2%). The correlation analysis also suggests that if an answer has more social features, such as greeting or affective languages, less users would click the upvote to promote it.

In summary, we found several important variables which are associated with the peer-judged answer quality. Longer answers, answers which were provided earlier, responders with higher RG score, or researchers from an institution with a higher RG score are more likely to be promoted for high quality content. In addition, users also preferred an answer with richer content, such as providing factual information, opinions, resource, referring to others, or experiences.

Table 4. Correlation between features and answer quality

Pearson/Spearman correlation	Answer quality based on peers' judgment	Sig.
Provide opinion	0.197**	0.000
Respond time	-0.159**	0.000
Length of the answer	0.147**	0.000
Social elements	-0.132**	0.000
RG scores	0.126**	0.000
Adding factual information	0.087**	0.005
Refer to other researchers	0.077*	0.014
Provide resources	0.066*	0.034
Responder's inst. RG score	0.062*	0.048
Provide personal experience	0.062*	0.048
Consensus building	0.055	0.081
Responder's inst. impact points	0.054	0.087
Impact points	0.019	0.541

*. p= .05 (2-tailed); **.01 (2-tailed);

4.2 Prediction of answer quality

(1) Prediction Result of SVM Classification Algorithm

We used LibSVM to apply the SVM classification algorithm, and compared the experimental results of 4 kernel functions on classification according to various measures. The experimental results of 10-fold cross-validation on SVM classification model is shown in Table 5.

Table 5. Result of SVM prediction

Kernel Function		Linear	RBF	Polynomial	Sigmoid
Accuracy		57.689%	58.178%	49.461%	56.024%
AUC scores	Low quality	0.604	0.611	0.500	0.587
	Medium quality	0.576	0.575	0.500	0.563
	High quality	0.500	0.500	0.500	0.500
Precision	Low quality	0.625	0.629	0.495	0.602
	Medium quality	0.529	0.534	0.000	0.514
Recall	Low quality	0.636	0.644	1.000	0.638
	Medium quality	0.612	0.614	0.000	0.571
F1	Low quality	0.630	0.636	0.662	0.619
	Medium quality	0.567	0.571	0.000	0.541

Note: High Quality is not shown for precision, recall or F1 because of the insufficient quantity of answers with high quality. The result is 0.

The RBF kernel enjoys advantages over the remaining three kernel functions. However, all kernel functions failed to predict the category of high quality answers which may be a result of an insufficient quantity of answers with high quality. The prediction in the paper has extracted 13 features that can be effectively mapped into the relatively high-dimensional feature space by RBF Kernel, allowing a good classification to be achieved. The parameter optimization is conducted for the RBF kernel that does better in classification, a better prediction of answer quality can be achieved via parameter optimization through which the optimal one can be selected after interactional verification. The experimental result is shown in Table 6.

Table 6. The result of optimizing SVM prediction

Kernel Function		RBF	Optimizing RBF
Accuracy		58.178%	62.684%
AUC scores	Low quality	0.611	0.613
	Medium quality	0.575	0.581
	High quality	0.500	0.511
Precision	Low quality	0.629	0.672
	Medium quality	0.534	0.575
	High quality	0.000	1.000
Recall	Low quality	0.644	0.691
	Medium quality	0.614	0.653
	High quality	0.000	0.064
F1	Low quality	0.636	0.681
	Medium quality	0.571	0.612
	High quality	0.000	0.120

As we can see from Table 6, the accuracy, precision, recall rate and F_1 value of the SVM model were all improved by parameter adjustment. The number of correctly predicted answers increased from 594 to 640 and the accuracy rose to 62.684% from 58.178%, indicating the improvement in prediction ability of answer quality after the parameter optimization.

(2) **Contrastive Analysis of Three Prediction Models**

The optimized SVM algorithm has an overwhelming advantage in accuracy over naïve Bayes or multiple regression as shown in Table 7. The performance of optimized SVM algorithms for prediction of low quality answers, medium quality answers and the precision rate of high quality answers is much better, while the naive Bayes analysis does better in predicting the recall rate and F1 value of high quality answers.

Table 7. The results of three prediction models

Classifier		Logistic regression	Naive Bayes	SVM (Optimizing RBF)
Accuracy		55.926%	52.204%	62.684%
AUC scores	Low quality	0.658	0.618	0.613
	Medium quality	0.616	0.594	0.581
	High quality	0.687	0.645	0.511
Precision	Low quality	0.593	0.581	0.672
	Medium quality	0.520	0.511	0.575
	High quality	0.429	0.164	1.000
Recall	Low quality	0.655	0.604	0.691
	Medium quality	0.541	0.493	0.653
	High quality	0.038	0.154	0.064
F1	Low quality	0.623	0.592	0.681
	Medium quality	0.530	0.502	0.612
	High quality	0.071	0.159	0.120

(3) **Contrastive Analysis of Features**

Three feature portfolios were developed to determine which features were most predictive of answer quality: human-coded features alone, web-captured features alone, and web-captured features + human-coded features. Because of the outstanding performance of the SVM classification algorithm, these feature portfolios were tested only in the optimized SVM algorithm. The experimental results are shown in Table 8.

The answer quality prediction based on the combination of web-captured features and human-coded features enjoys better accuracy than those based on the human-coded features or of web-captured features alone. That in prediction from a single source alone, that based on web-captured features does better than one based only on human-coded features.

Table 8. Contrastive analysis of feature portfolios

Classifier		Human-coded features	Web-captured features	Human-coded features+ Web-captured features
Accuracy		53.183%	59.158%	62.684%
AUC scores	Low quality	0.590	0.618	0.613
	Medium quality	0.555	0.585	0.581
	High quality	0.500	0.500	0.511
Precision	Low quality	0.591	0.595	0.672
	Medium quality	0.480	0.585	0.575
	High quality	0.000	0.000	1.000
Recall	Low quality	0.558	0.800	0.691
	Medium quality	0.596	0.457	0.653
	High quality	0.000	0.000	0.064
F1	Low quality	0.574	0.682	0.681
	Medium quality	0.532	0.513	0.612
	High quality	0.000	0.000	0.120

5. DISCUSSION

Our results found out that peer-judged answer quality on ResearchGate (RG) was associated with web-captured features such as responder's RG score, responder's institutional RG score, answer length, and response time. These findings are consistent with previous studies, where responders' authority and response time have influences on the answer quality [5, 6]. It is true that Responder's RG scores and institution RG scores, which stand for responders' academic authority, only exist on the ResearchGate platform, but other platforms also have similar users' reputation score, which could potentially be used to replace the RG scores. For example, both Stack Overflow and Yahoo! Answers have user reputation scores. Based on the results, we also found that quick answers have higher quality than later answers, which are consistent the previous work on generic Q&A sites [6]. However, we do notice that ResearchGate displays the most recent answer first, which could have negative impacts to the earlier answers. We need further research to know the influence of this setting to peer-judged quality. Meanwhile, contrary to previous literature [5, p.6] in which the authors suggested lengthy answers "may in fact turn out to be of poor quality" on generic Q&A sites, our findings suggest that the longer answers on RG are more likely to be promoted as a quality answer. Possible explanations can be that RG users preferred content–rich information. The difference between an academic Q&A site and generic Q&A site is worth studying in future works.

As for the human-coded features, we found that characteristics such as providing resources, adding factual information, providing opinions, referring to other researchers, and providing personal experience were positively correlated with the peer-judged answer quality. However, answers containing social elements such as greetings or other affective words actually harm the answer quality. A possible explanation can be that RG users are more concerned about the content relevance and knowledge volume in an answer. This result further helps us to conclude that academic Q&A users perceived long, information-rich content as better quality, also indicating that academic social platforms are different to generic social platforms, and the social components in academic social platforms should probably be kept separate from the information related components.

For the answer quality prediction, we learned that the optimized SVM algorithm has an overwhelming advantage over other models in terms of the accuracy. Meanwhile, we found that the prediction based on web-captured features had a better performance when comparing to the human-coded features. This phenomenon is worth our further exploration. If we can predict high quality answers accurately, users are expected to receive better answer recommendations. However, difficulties in predicting high quality answer should be noted due to the insufficient quantity of answers with high quality and a limited numer of answer features. In the future we would like to enlarge the quantity and types of data sets and propose more answer features, such as features related to the field of science and expertise levels of contributors as measured by their scientific contribution to more fully predict answer quality.

6. CONCLUSION AND FUTURE WORK

We studied answer quality of scholars answers in a Q&A setting based on the web-captured features (e.g., RG score or the length of answers) and human-coded features (e.g., if the answer provided factual information) on the academic Q&A site ResearchGate (RG). We found that the differing degrees of answer quality on RG are characterized differently from general Q&A sites. We also learned that the optimized SVM algorithm has an overwhelming advantage over other models in terms of accuracy and finally that prediction based on web-captured features had a better performance than prediction based on human-coded features.

In the future we need to perform a user study to explore whether the features proposed here as important are considered by users to determine whether an answer is high quality. We plan to enlarge the datasets for a better prediction accuracy rate. In addition, we plan to test our research model on other professional Q&A sites such as Quora or StackOverflow, allowing the study to probe into more features of social platforms.

7. ACKNOWLEDGMENTS

This work was supported by Major Projects of National Social Science Fund (13&ZD174), National Natural Science Foundation of China (No.71273126), National Social Science Fund Project (No.14BTQ033), and Project of the Education Ministry's Humanities and Social Science (No. 13YJA870020).This work was done while Lei Li was at the School of Information Sciences, University of Pittsburgh.

8. REFERENCES

[1] Liu, Y., Bian, J., & Agichtein, E. (2008, July). Predicting information seeker satisfaction in community question answering. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (pp. 483-490). ACM.DOI=10.1145/1390334.1390417

[2] Harper, F. M., Raban, D., Rafaeli, S., & Konstan, J. A. (2008, April). Predictors of answer quality in online Q&A sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*(pp. 865-874). ACM. DOI= 10.1145/1357054.1357191.

[3] Shah, C., & Pomerantz, J. (2010, July). Evaluating and predicting answer quality in community QA. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 411-418). ACM. DOI= 10.1145/1835449.1835518

[4] Fichman, P. (2011). A comparative assessment of answer quality on four question answering sites. *Journal of Information Science*,*37*(5), 476-486.DOI=10.1177/0165551511415584

[5] John, B. M., Chua, A. Y., & Goh, D. H. L. (2011). What makes a high-quality user-generated answer? *Internet Computing, IEEE*, *15*(1), 66-71. DOI=10.1109/MIC.2011.23

[6] Chua, A. Y., & Banerjee, S. (2013). So fast so good: An analysis of answer quality and answer speed in community Question - answering sites. *Journal of the American Society for Information Science and Technology*,*64*(10), 2058-2068. DOI=10.1002/asi.22902

[7] Jeng, W., He, D., & Jiang, J. (2014), User participation in an academic social networking service: A survey of open group users on Mendeley. Journal of the Association for Information Science and Technology. doi: 10.1002/asi.23225

[8] Nández, G., & Borrego, Á. (2013). Use of social networks for academic purposes: a case study. Electronic Library, The, 31(6), 781-791.

[9] Goodwin, S., Jeng, W. & He, D. (2014). Changing Communication on ResearchGate Through Interface Updates. ASIS&T 2014 Annual Meeting.

[10] Jordan, K. (2014). Academics and their online networks: Exploring the role of academic social networking sites. First Monday, 19(11).

[11] Thelwall, M., & Kousha, K. (2014). Academia. edu: Social Network or Academic Network?. Journal of the Association for Information Science and Technology, 65(4), 721-731.

[12] Thelwall, M., & Kousha, K. (2014). ResearchGate: Disseminating, communicating, and measuring Scholarship?. Journal of the Association for Information Science and Technology.

[13] Yahoo Design Patterns. Vote to Promote. https://developer.yahoo.com/ypatterns/social/objects/feedback/votetopromote.html

[14] Kim, Soojung, and Sanghee Oh. "Users' relevance criteria for evaluating answers in a social Q&A site." Journal of the American Society for Information Science and Technology 60, no. 4 (2009): 716-727.

[15] Suryanto, M. A., Lim, E. P., Sun, A., & Chiang, R. H. (2009, February). Quality-aware collaborative question answering: methods and evaluation. In Proceedings of the second ACM international conference on web search and data mining (pp. 142-151). ACM. DOI=10.1145/1498759.1498820

[16] Jurczyk, P., & Agichtein, E. (2007, November). Discovering authorities in question answer communities by using link analysis. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management (pp. 919-922). ACM. DOI=10.1145/1321440.1321575

[17] Ong, C. S., Day, M. Y., & Hsu, W. L. (2009). The measurement of user satisfaction with question answering systems. Information & Management,46(7), 397-403. DOI=10.1016/j.im.2009.07.004.

[18] About us. (n.d.). ResearchGate. Retrieved January 25, 2015, from https://www.researchgate.net/about

[19] Henri, F. (1992). Computer conferencing and content analysis. In *Collaborative learning through computer conferencing* (pp. 117-136). Springer Berlin Heidelberg. DOI=10.1007/978-3-642-77684-7_8

[20] Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, *42*(3), 243-265. DOI=10.1016/j.compedu.2003.08.003

[21] Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam Med*, *37*(5), 360-363.