Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Impact of Group Norms in Eliciting Response in a Goal Driven Virtual Community

Sumeet JAIN[a*], Tanmay SINHA[b], Achal SHAH[c], Chandramouli SHARMA[d] & Carolyn ROSE[e]
[a] *Undergraduate student, Vellore Institute of Technology, India*
[b] *Undergraduate student, Vellore Institute of Technology, India*
[c] *Undergraduate student, IIT Guwahati, India*
[d] *Undergraduate student, NITK Surathkal, India*
[e] *Associate Professor, Carnegie Mellon University, USA*
* sumeet.jain44@gmail.com

Abstract: With the proliferation of social media into our daily lives, online communities have become an important platform for collaborative learning and education. To connect users with varying knowledge levels and increase the net learning throughput, these communities often follow a question-answer based approach. Understanding what drives attention to help-seeking questions can reduce the amount of questions that go unnoticed or remain unanswered by the community. In this paper we discuss an important feature that affects the activity of the community, namely the community norms. We present a machine learning based trigger-driven feedback model that functions by (i) differentiating between help-seeking questions and follow-up posts – i.e. posts that are part of an ongoing discussion, and (ii) a dynamic intervention scheme to help improve question formulation. Our findings show that adhering to the community norms significantly increases the chance of eliciting a response.

Keywords: Internet Relay Chat, Intervention, Group Norms, Text classification, Virtual Community, Intelligent Agent, Dynamic feedback

1. Introduction

Knowledge sharing has been the focus of research for a long time now and facilitates knowledge creation, transfer learning and development of knowledge management initiatives (He & Wei, 2009). In today's digitized world, technical platforms and discussion forums facilitate engagement, collaboration and dissemination of knowledge among the users. The objective is to connect users from varying demographical locations and help them to improve their technical skills. Most of these platforms take the shape of online communities, where a question-answer driven approach is followed to help beginners learn new concepts.

In such virtual communities built around a common goal, there is usually a certain code of behavior that people are expected to exhibit. These norms tell people how to behave, and importantly how not to misbehave (Danis & Alison, 2002). In these communities, however, it has been observed that newcomers who are not aware of the norms of the community often fail to elicit a response to their questions due to improper question formulation. As a result, the channel is effectively not able to meet its goal of delivering immediate solutions and learning may be inhibited. This can also result in a decrease in the participation of users in the community. To engage more users in goal-oriented communities, there is a growing need to understand the type of questions they are asking, whether it fits into the community norms, and possible ways of rephrasing their question so that it has a higher probability of getting answered.

To explore these and related questions, our paper sets out to study the following research question: How does an understanding of group norms impact learning of the community users? To achieve this, we implement a dynamic machine learning based intervention model that identifies questions that are less likely to elicit a response, and provides triggers (suggestions) help improve question formulation. Similar to the educational setting, where students are taught the importance of classroom or school etiquette, our model aims to improve collaborative learning processes that are a part of the online community.

2. Related Work

Prior work can be categorized into 3 parts: Virtual Communities, Social Norms, Intervention and feedback.

Virtual Communities: Successful communities have been found to share characteristics like common goal accomplishments, common practices for interaction and sharing information, major participant contribution with both producers and consumers of information. A critical factor that improves learning and promotes knowledge flow is asking questions (Svinicki, Marilla & McKeachie, 2011). However, it was observed that 40% of potential thread starting messages in Usenet groups received no response (Burke, Joyce, Kim, Anand & Kraut, 2007). The solutions so far have mainly focused on rhetorical strategies like group introductions, message length and other contextual factors as an aid to improve chances of interaction. Johnson (2001) points out the factors that impede the development of virtual communities include withdrawing, cultural differences, superficial discussion content, as well as lack of urgency in responding. However, majority of these approaches are community oriented and not from the user's perspective. To the best of our knowledge, these virtual communities have not been studied in the light of importance of group norms for response elicitation.

Social Norms: Group norms and social identity have been identified as two major social determinants that impact virtual community participation (Bagozzi & Dholakia, 2002). They represent generally accepted and widely-sanctioned routines, values and perspectives that people follow (Gerber & Macionis, 2011). An interesting prior work by Dholakia, Richard & Lisa (2004) stressed that higher level of value perceptions leads to stronger group norms regarding the virtual community. This leads to increased participation and an increased willingness to accommodate and involve others in the discussion. Rfc (1855) also lays out guidelines for network etiquettes. It stresses on an understanding of the group culture before communicating, taking extreme care in what should be posted by conforming to the community norms, not going off topic and engaging in flame wars. People, especially newcomers who do not conform to these norms are often made to come back into line using some discretionary stimuli (Chong, 2000). These stimuli in our current study are dynamically designed and timely triggers, which help people in improving their way of asking questions, adhering to the community norms.

Intervention and feedback: Feedback Interventions (FIs) are "actions taken by external agents to provide information regarding some aspect of one’s task performance” (Kluger & Angelo, 1996). Proper intervention helps in encouraging peer dialog around learning and engaging people in more discussion. It provides opportunities to close the gap between current and desired performance, i.e. improvement in terms of knowledge and learning (Juwah, Macfarlane-Dick, Matthew, Nicol, Ross & Smith, 2004). Scruggs, Margo, Sheri & Janet (2010) in their analysis on the importance of educational support, pointed out that educational interventions like study aids, classroom learning strategies and computer-assisted instruction had a significant impact on learning. In their review article, Paul & Dylan (1998) drew together over 250 studies of formative assessment with feedback carried out since 1988 spanning all educational sectors. Their analysis of these studies showed that feedback resulted in positive benefits on learning and achievement across all content areas, knowledge and skill types and levels of education. While the inclusion of such instructional interventions has been mainly studied in educational settings, we connect the same to virtual communities.

3. Experiments

3.1 Workspace setting and dataset

The virtual community chosen for the purpose of study was "Ubuntu-beginners IRC channel". The major challenge in this community is that many people (mostly beginners) do not follow group norms. This leads to repeated manual interventions by experienced users and affects the efficiency of the community. We intend to automate this and inform users about the group norms, which benefit both the user and community.

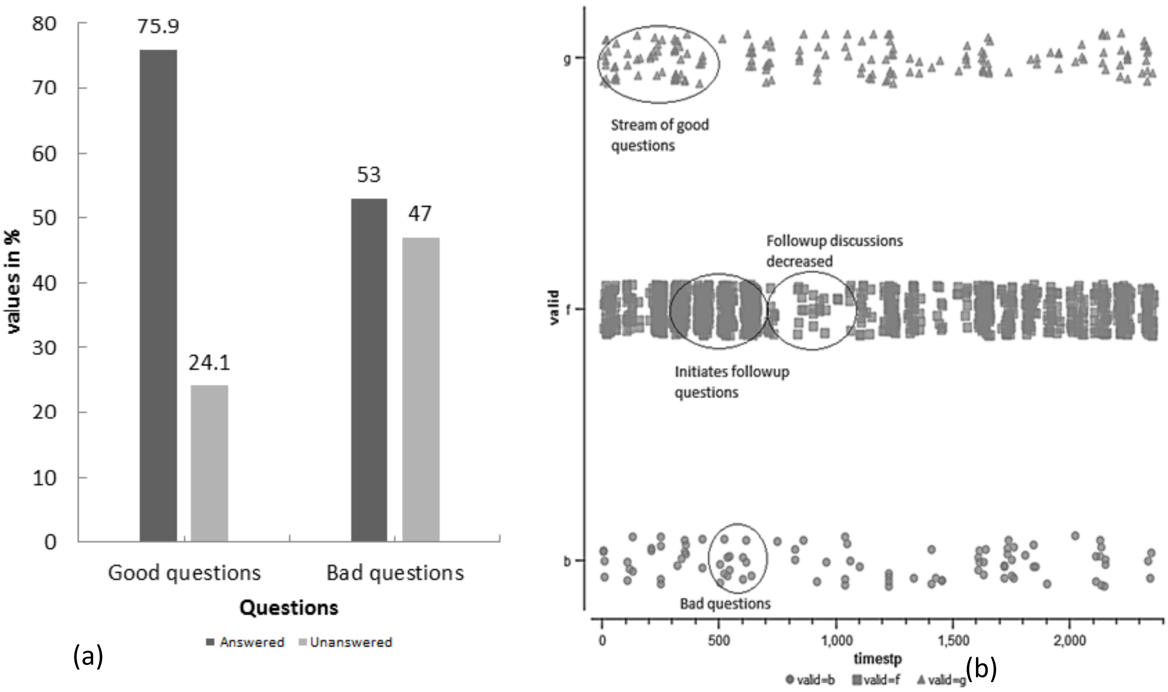
For this study, IRC logs of #ubuntu-beginner channel (a part of the online Ubuntu community) were used. We analyzed 1412 questions for this study, which were extracted from the logs of year 2011 from 1 month window. To extract the questions from the continuous stream of chat logs, pattern matching based on regular expressions was used.

3.2 Data labeling and analysis

The 1412 questions extracted were manually labeled by four independent annotators. It was ensured that the annotators reached an agreement, by asking them to independently label a sample of 200 questions from the dataset and then cross validate with other annotators. To avoid confusion, a set of labeling rules were formed through general observation and interaction with the frequent channel users. These rules are mentioned below. The following three categories of questions had a negative impact on the efficiency of the community and that did not adhere to the norms of the community. These are – i) Social questions (such as 'Hi, how is everyone?'), ii) Meta-questions (these were questions that asked to ask, such as 'Can anyone help me?') and iii) Indirect questions (such as 'Does anyone know about Ubuntu?').

- Rule 1: Questions belonging to any of these three categories were marked as 'Off-norm' or 'Bad' questions.
- Rule 2: All help-seeking questions that were relevant to the channel and that did not belong any of the above-mentioned categories were labeled as 'On-norm' or 'Good' questions.
- Rule 3*:* Questions that were part of an ongoing discussion, directed to a specific user, or that were in response to another question were labeled as "Follow-up" questions.

After the labeling, 171 questions were classified as 'Good', 93 as 'Bad' and the remaining 1148 were 'Follow-up' questions. It is observed that 'Follow-up' questions are larger in number as compared to 'Good' and 'Bad' questions, the reason being large number of questions interchanged between the asker and the responder. Figure 1(a) shows the distribution of answered questions among 'Good' and 'Bad' categories. It is evident that a higher proportion of 'Good' questions are answered as compared to 'Bad' questions. This supports our claim that the chances of questions being answered increase when they adhere to the group norms. Figure 1(b) represents the question category on the vertical axis and time on horizontal axis. An important observation is that a stream of 'Good' questions (denoted by an oval boundary) is followed by a large number of 'Follow-up' questions (indicating a healthy discussion), whereas a stream of 'Bad' questions results in a decrease in the number of 'Follow-up' questions. This further strengthens our claim that users respond (in terms of follow up questions) more to 'Good' questions.



<u>Figure 18: (a) Answered/Unanswered question distribution; (b) Scatter plot of question category vs. timestamp.</u>

3.3 Feature Engineering

Understanding what drives attention in virtual communities involves defining a collection of features and then filtering out the important ones. The following text/post features were considered to identify attributes that could be used to categorize questions in one of the above mentioned classes.

- *Complexity (Unigrams and Bigrams)*: To measure the complexity of a post we use N-grams with N=1 and N=2. N-grams are a continuous sequence of n lexical resources like words, letters and syllables collected from text or speech corpus.
- *Style (Part of speech (POS) tagging)*: Various posts have unique style of grammar. To measure this, we use POS tagging. Stanford's POS tagger library was used for this purpose.
- *Informativeness*: The novelty of the post's terms with respect to other posts. We derive this measure using the Term Frequency-Inverse Document Frequency (TF-IDF) measure.

3.4 Classifier training and results

We performed the question category identification using Sequential Minimal Optimization (SMO) (Platt, 1998) method of Support Vector Machines (SVM) (Hearst, Dumais, Osman, Platt & Scholkopf, 1998). The kernel used for SVM was Polynomial with order 3. SVMs have been shown to outperform other existing methods (naïve Bayes, k-NN, and decision trees) in text categorization (Joachims, 1998). Their advantages are robustness and elimination of the need for feature selection and parameter tuning. To guarantee that the model performs well in practice and generalize to an independent data set, k-*fold* cross validation technique was used with $k = 10$. To assess the model's performance we measured accuracy (correctly classified instances), macro averaged F score (computed globally over all category decisions) and Cohen's kappa/kappa measure. To gain further insights into which features contributes to each class, we use precision and kappa value measure for individual feature. After the training the accuracy obtained was 94.75%, macro-averaged F score .957 and kappa value 0.832.

Table XIII: Top features table for each category

Category	Top Features (By precision)
Bad questions	anyon_know , anyon_help, help_me, anyon, MD_NN (POS) , can_anyon, doe_anyon, someon
Follow-up questions	did_you, oh , you_mean, you_need, hav_you, do_that , you_try, username
Good questions	and_i, CC_VBP (POS), comput, gui, emapthi, read, after

Table I shows the top features (stemmed) for each class. An observation is that predictors of 'Bad' questions are words generally used in Meta-questions, such as 'Does anyone know how to install Ubuntu?'. This question would be better if it were more specific, and framed as 'How to install ubuntu 12.04 on dell 1555?'. Other features like 'help_me' indicate that although the user is asking for help, he is not clear and is asking indirect questions. Features predicting 'Follow-up' questions seem logical and valid. These features are words used in conversation as a reply to someone's query, such as 'did_you', 'do_that', 'you_try'. An important observation is that the features of good questions are specific to community and these may change if we analyze a different community. Features like 'comput' and 'gui' are specific to computer related fields. In summary, the above results suggest that the features predicting the Bad or Follow-up category are independent of the channel (goal directed) or topic, whereas those belonging to the 'Good' category depend on the topic and context of the questions posted in the IRC channel.

3.5 Intervention and Feedback

In this section, we discuss our intervention and feedback model, based on principals of good feedback practice as identified by Juwah, Macfarlane-Dick, Matthew, Nicol, Ross & Smith (2004) from the conceptual model and research literature on formative assessment. Our model seeks to provide users with effective and timely feedback, to help them understand how their question could be better framed to increase the chance of eliciting a response from the community. The trigger advice and direction helps them avoid abortive work. 'Follow-up' questions do not require an automated intervention, as they indicate an on-going discussion amongst users in the community. 'On-norm' questions include those framed adhering to the norms of the community, and require no automated intervention. On the other hand, "Off-norm" questions can be further subcategorised as "Social questions", "Metaquestions" or "Indirect questions". Questions so framed are less likely to attract the attention of the community and may remain ignored. For the development of timely and correct triggers, we used the features which were strong predictors of different categories of bad questions (Table I) based on weights. Then, we used regular expressions to create various combinations of such features, to form a baseline for comparison. An important aspect of the model is that the feedback delivered to the user is in the form of a personal message, rather than a public message in the channel. The purpose this serves is twofold – i) It facilitates the development of self-assessment in learning, without affecting motivational beliefs and self-esteem, ii) It helps in maintain the efficiency of the community.

Table XIV: Sample intervention based on questions identified

Question Type	**Example**	**Sample Intervention**
Good	Which is latest ubuntu?	No intervention
Follow-up	sam: You used sudo?	No intervention
Bad (Social)	Hey all, watsup?	"Such questions are not apt for this community. You might want to ask questions without being too social."
Bad (Indirect)	Does anyone know git?	"You might want to rephrase your question. Indirect questions generally do not get answered.
Bad (Metaquestion)	Hey, can anyone help me?	"Hey this community would appreciate if you directly asked your question. Do not ask to ask"

For the implementation, we used an open source Java framework for writing IRC bots called "Pircbot". It includes an event-driven architecture to handle common IRC events. To facilitate proper communication between IRC bot (which delivers personalized message to end user) and the machine learning model, we used Bazaar (Adamson & Rosé, 2012), an open source architecture for collaborative conversational agents. The detailed flow of events in our implementation is shown in Figure 2. The personalized interventions that are given by our model based on question type are shown in Table II.

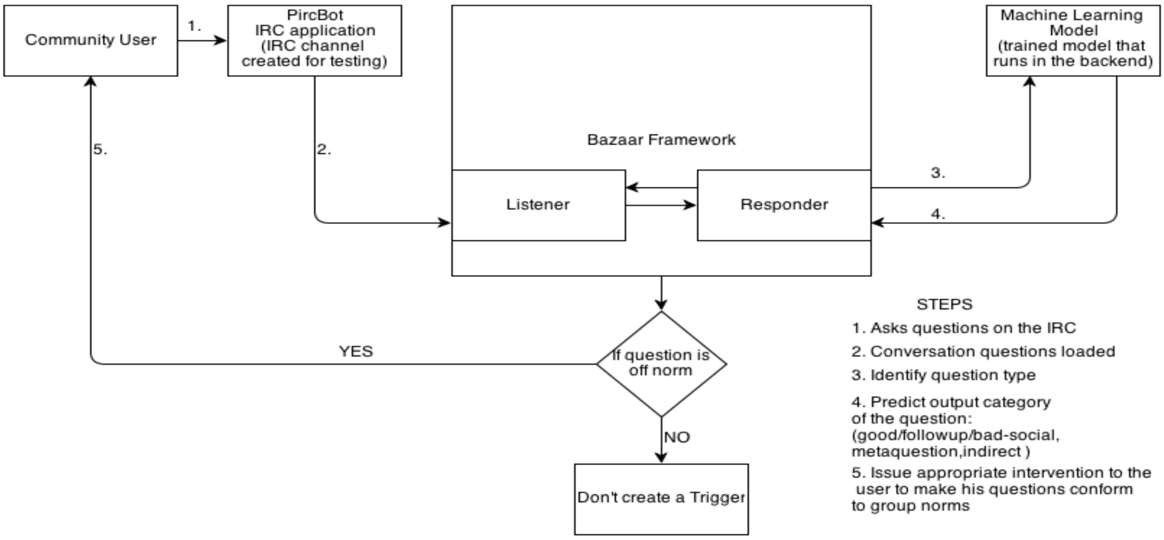


Figure 19: Implementation details of the intervention and feedback system

4. Conclusion and Future work

In this paper, we presented a new support mechanism for the users of IRC virtual community. The objective was to identify questions that did not adhere to the norms of the community and automate the intervention and support pipeline to help improve question formulation.

In this study, we focused only on the Ubuntu IRC community, instead of considering other open source discussion forums. As a result, the classification accuracy may not be a true reflector of the kind of questions asked in such virtual communities and it may have generalization errors.

Future work would involve extensive user studies on a large and diverse user base to understand the effectiveness of the model. We also aim to test our system using other IRC channels. We plan to study other relevant factors like context of the questions, timing of the query, number of active users etc., which lead to answered or unanswered questions. Based on these factors, we could then analyse and predict the specific user questions' attention span and the probability of eliciting an answer in the virtual community.

Acknowledgements

This research was started as a part of IPTSE Winter School by Carnegie Mellon University (CMU) and was later continued by authors. The authors thank Dr. Carolyn Rose (Language Technologies Institute, CMU) for her valued suggestions and guidance. The authors would also like to thank Mr. David Adamson (PhD Student, CMU) for his guidance and help in implementing IRC bot plugin for BAZAAR framework.

References

Adamson, D., & Rosé, C. P. (2012, January). Coordinating multi-dimensional support in collaborative conversational agents. In Intelligent Tutoring Systems (pp. 346-351). Springer Berlin Heidelberg.

Bagozzi, R. P., & Dholakia, U. M. (2002). Intentional social action in virtual communities. Journal of interactive marketing, 16(2), 2-21.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Granada Learning.

Burke, M., Joyce, E., Kim, T., Anand, V., & Kraut, R. (2007). Introductions and requests: Rhetorical strategies that elicit response in online communities. In Communities and Technologies 2007 (pp. 21-39). Springer London.

Chong, D. (2000). Rational lives: Norms and values in politics and society. University of Chicago Press.

Danis, C., & Lee, A. (2002). The Negotiation of Norms in an Online Community.

Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. (2004). A social influence model of consumer participation in network-and small-group-based virtual communities. International journal of research in marketing, 21(3), 241-263.

Gerber, L., & Macionis, J., Sociology: Seventh Canadian Edition., 2011.

He, W., & Wei, K. K. (2009). What drives continued knowledge sharing? An investigation of knowledge-contribution and-seeking beliefs. Decision Support Systems, 46(4), 826-838.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features (pp. 137-142). Springer Berlin Heidelberg.

Johnson, C. M. (2001). A survey of current research on online communities of practice. The internet and higher education, 4(1), 45-60.

Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., & Smith, B. (2004). Enhancing student learning through effective formative feedback. Higher Education Academy (Generic Centre).

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological bulletin, 119(2), 254.

Platt, J. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines.

RFC 1855, Based on: http://www.ietf.org/rfc/rfc1855.txt

Svinicki, M. D., & McKeachie, W. J. (Eds.). (2011). McKeachie's Teaching Tips: Strategies. Cengage Learning.