5-2013

Analyzing the friendliness of exchanges in an online software developer community

Brendan CLEARY

Margaret-Anne STOREY

Carlos GÓMEZ

Leif SINGER

Christoph TREUDE
Singapore Management University, ctreude@smu.edu.sg

Citation

Analyzing the Friendliness of Exchanges in an Online Software Developer Community

Brendan Cleary, Carlos Gómez,
Margaret-Anne Storey
University of Victoria
Victoria, BC, Canada
{bcleary,cgomez,mstorey}@uvic.ca

Leif Singer
Leibniz Universität Hannover
Hannover, Germany
leif.singer@inf.uni-hannover.de

Christoph Treude
McGill University
Montreal, QC, Canada
ctreude@gmail.com

Abstract—**Many online communities struggle with conflicts — e.g. between newcomers and elders — at some point. In July 2012, the Stack Exchange organization attempted to assess the overall "niceness" of the Stack Overflow community by rating the "friendliness" of 7,000 comments made on the site over a 4 year period. We performed a deeper examination of the comment dataset published by Stack Exchange. We find a high degree of comment repetition in the Stack Overflow database and suggest some simple heuristics that may help in automatically identifying unfriendly comments, providing managers of developer communities with simple means that could counter hostility.**

I. Introduction

Stack Overflow is an increasingly important part of the online resources relied on by millions of software developers [1]–[3]. Like most online communities, newcomer perceptions and future participation can be heavily influenced by their first interactions with the community, especially if they are negative or unfriendly. Being able to automatically classify comment friendliness would allow developer communities like Stack Overflow to preemptively warn users that attempt to post comments which may be received as unfriendly. This could help to reduce the number of unfriendly comments while increasing participation and ultimately the longevity of the community.

To investigate whether Stack Overflow was welcoming or hostile to newcomers, the Stack Exchange organization instigated a campaign in the summer of 2012 to study and promote "niceness" in the Stack Overflow community[1]. Stack Exchange sampled 7,000 of the more than 15 million comments in the Stack Overflow database. Using Mechanical Turk, these comments were manually rated for "friendliness". The Stack Exchange study found that Stack Overflow was becoming a friendlier place, with the percentage of friendly comments increasing over time. However, they also found that first time posters were more likely to receive negative comments[2].

In this position paper, we perform an analysis of the comment dataset published by Stack Exchange, focusing on the content of those comments and how frequently individual comments are repeated. This is a first step towards building tools that automatically identify potentially negative comments in developer communities.

II. Exploring the Stack Exchange Data Set

Our goal is to determine the viability of automatically classifying the "friendliness" of comments on Stack Overflow and other developer communities. To this end, we are interested in identifying the characteristics of comments that could be used for building a classifier. Specifically we are interested in:

1) Uniqueness — are comments on Stack Overflow unique, or is there repetition that could be utilized by a classification algorithm for comments?
2) Language — are there patterns to vocabulary use that a classification algorithm might use as heuristics?

The Stack Exchange dataset, by providing a pre-classified set of comments, represents a valuable opportunity to address these research questions. To investigate the pervasiveness of friendly and unfriendly comments, we used a simple text matching approach: we attempted to match the 7,000 comments in the published Stack Exchange dataset against the more than 13 million comments contained in the August 2012 Stack Overflow data dump (a snapshot of the Stack Overflow database published by Stack Exchange every 6 months).

We identified 3,561 comments out of the 7,000 that had been published and classified by Stack Exchange. The discrepancy results from comments that were modified or removed from Stack Overflow prior to August 2012.

III. Analysis & Findings

We performed a simple occurrence analysis to identify the number of times each of the 3,561 comments appeared in the August 2012 Stack Overflow data dump and the number of unique users who made the comment. Table I presents the results of this analysis along with a "niceness" classification based on the Stack Exchange study.

The majority of the most frequently repeated comments are classified as "friendly" with only one "unfriendly" comment being highly repeated. In total, we found over 20,000 occurrences of the 3,561 comments. However, only 212 comments were repeated more than once. If the Stack Exchange dataset is representative of Stack Overflow comments, the majority of comments are unique, but a small number are highly repeated.

---

[1] http://blog.stackoverflow.com/?p=11931 (Jul 20, 2012)
[2] https://www.dropbox.com/l/NECSweKlIYKvMBRE (Jul 23, 2012)

CHASE 2013, San Francisco, CA, USA

TABLE I
TOP 10 COMMENT OCCURRENCES

Comment	Users	Occurrences	Classification
what have you tried?	1,472	3,570	FRIENDLY
what have you tried so far?	981	1,834	FRIENDLY
is this homework?	939	1,530	UNFRIENDLY
you're welcome!	603	1,001	FRIENDLY
what is your question?	605	892	FRIENDLY
what's the question?	364	477	FRIENDLY
what is the error?	373	433	FRIENDLY
what error do you get?	300	382	FRIENDLY
thanks very much!	263	323	FRIENDLY
what error are you getting?	277	312	FRIENDLY

TABLE II
COMMENT REPEATERS

Comment	Users	Occurrences
please learn java naming …	1	150
needs 'homework' tag?	1	73
you can get help from facebook …	1	40
i am doing something similar …	1	26
can you answer your own …	1	24

Another notable finding were comments that occurred frequently, but were made by only a small number of users. For example, Table II lists a selection from the 100 most frequently occurring comments that each originated from a single user (comments have been abbreviated). Table III shows the top 10 most frequently occurring "unfriendly" comments. Notably, the "homework" keyword is not only present in the most repeated "unfriendly" comment, but is also present in half of the top 10 "unfriendly" comments. This constrained vocabulary presents a simple heuristic that could be used to flag "unfriendly" comments.

Table IV shows the most frequently occurring ambiguously classified comments. In comparison to the most frequently occurring "friendly" or "unfriendly" comments, ambiguous comments tend to be repeated far less often and appear to have a richer vocabulary.

IV. LIMITATIONS

As we examined only comments that remained in the Stack Overflow database in August 2012, there is a possibility that extremely negative comments captured in the Stack Exchange study might have been deleted before this date. Because the

TABLE III
TOP 10 OCCURRING UNFRIENDLY COMMENTS

Comment	Users	Occurrences
is this homework?	939	1,530
and the question is?	148	180
belongs on superuser.com	29	108
is this a homework?	56	74
needs 'homework' tag?	1	73
smells like homework.	50	69
what are you asking?	33	65
not programming related	30	54
so what is the question?	50	50
homework question?	45	48

TABLE IV
TOP 10 OCCURRING AMBIGUOUS COMMENTS

Comment	Users	Occurrences
why the downvotes?	85	89
this question is not a good fit to our q&a format …	9	21
belongs on superuser	17	20
what flavor of sql?	10	11
belongs on programmers.stackexchange.com	6	8
needs to be cw?	1	7
your question is better suited for serverfault.com	1	7
why is this off topic?	4	4
community wiki …?	1	3
exact duplicate	3	3

Stack Exchange dataset contained comments sampled at specific times over four years, it is possible that some frequently occurring comments were not captured in that dataset, and therefore would not be considered in our analysis. Finally, the comment classifications provided in the Stack Exchange dataset were collected using Mechanical Turk: participants were not explicitly trained in comment classification and did not have access to the context in which a comment appeared.

V. CONCLUSIONS & FUTURE WORK

Using data published by Stack Exchange that rated the "friendliness" of Stack Overflow comments, we analyzed how frequently those comments were repeated on Stack Overflow and the vocabulary used in different types of comments. We found that the most frequently occurring "unfriendly" comments use a constrained vocabulary and that some very frequently repeated comments are created by small numbers of users. This qualitative analysis adds further insights to the quantitative analysis conducted by Stack Exchange so far. Our findings suggest possible features that could be implemented on Stack Overflow and other developer communities that would warn users that are about to post a comment that might be received as negative. In future work, we will investigate the automatic classification of comments and posts.

ACKNOWLEDGMENTS

REFERENCES

[1] B. Dagenais and M. Robillard, "Creating and evolving developer documentation: understanding the decisions of open source contributors," in *Proceedings of the eighteenth ACM SIGSOFT international symposium on Foundations of software engineering*. ACM, 2010, pp. 127–136.

[2] S. M. Nasehi, J. Sillito, F. Maurer, and C. Burns, "What makes a good code example?: A study of programming q&a in stackoverflow," in *Software Maintenance (ICSM), 2012 28th IEEE International Conference on*, sept. 2012, pp. 25–34.

[3] C. Parnin and C. Treude, "Measuring api documentation on the web," in *Proceedings of the 2nd International Workshop on Web 2.0 for Software Engineering*, ser. Web2SE '11. New York, NY, USA: ACM, 2011, pp. 25–30. [Online]. Available: http://doi.acm.org/10.1145/1984701.1984706