Emotion and Interaction Processes in a Collaborative Online Network

Deepak Rishi[1], Jesse Hoey[1], Mei Naggappan[1], Kimberley B. Rogers[2], and Tobias Schröder[3]

[1]David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada
[2]Department of Sociology, Dartmouth College, NH, USA
[3]Institute for Urban Futures, Potsdam University of Applied Sciences, Potsdam, Germany

keywords: emotion, collaborative network, interaction process, sentiment analysis, github

Understanding the social forces behind self-organized collaboration is increasingly important in today's society, where political problem-solving and the creation of economic value occur less and less in formal, hierarchical organizations. Instead, we live in what scholars have described as an emerging distributed economy and digital democracy, where technological and social innovations are increasingly generated through informal processes of collaboration in and across startups, civic laboratories, fabrication labs and the like, often enabled through cheap and ubiquitous information and communication technology (e.g., see Blowfield and Johnson [2013], Bogers and West [2012], Helbing and Pournaras [2015], Townsend [2013]). In the THEMIS.COG project [1], we study the open, collaborative development of software in online social coding communities like GitHub [2] as one key example of these economic changes. Exploring collaboration dynamics in communities like GitHub can further our understanding of the social and psychological mechanisms that drive the novel kind of human collaboration so central to the 21st century's economy and society.

Prior research suggests that people care at least as much about maintaining social relationships as they do about striving to maximize personal gains in their transactions with others. This makes intuitive sense, since maximizing one's gains depends on sustaining valuable relationships over time. Building on a long tradition of sociological theory and research, we hypothesise that identity dynamics explain how and why actors pursue each of these goals through interactions with others, and seek to develop a mathematically precise model that can be used to predict and test collaborative dynamics. The model will be based on the interaction process analysis (IPA) dynamics model of Bales [1999], implemented and simulated using the *BayesACT* sentiment and identity based model of human dyadic and group interactions [Schröder et al., 2016]. The general underlying assumption of both *BayesACT* and IPA is that humans strive for their social experiences to be coherent at a deep, emotional level with their sense of identity and general worldviews and values as constructed through culturally shared symbols. In pursuit of this goal, in this paper we examine the problem of classifying pull request comments on GitHub (one of the methods for contributing to a software project by a developer) into the categories of IPA and emotions.

Emotions and interaction processes play an important role in software collaborations. For example, Fredickson [2001] states that positive emotions like happiness help people to be more creative, which is essential for successful software design. On the other hand, Ambler [2002] states that negative emotions, fear, or absence of courage might make developers refrain from

[1]`themis-cog.github.io`
[2]`github.com`

changing/refactoring their code. De Choudhury and Counts [2013] also show that emotions affect task quality, productivity, creativity, group rapport and job satisfaction. While data on software discussions are openly available, analyses of sentiment and emotions in these data are challenging as discussions can be of a technical nature. Previous attempts include: Murgia et al. [2014], who perform a feasibility study of emotions mining using *Parrott's framework* on *Apache* issue reports; Guzman et al. [2014], who use lexical sentiment analysis to study emotions expressed in commit comments of open source projects; and Pletea et al. [2014], who use the Natural Language Text Processing (NLTK) tool to explore sentiment analysis of security related discussions on GitHub.

To our knowledge, no studies have attempted a theory-driven approach informed by social psychological research on group dynamics to guide analysis of mass data from software collaborations. Interaction Process Analysis (IPA) is such a theoretical framework that uses (among other things) a small set of behaviour categories (twelve) that were developed through many observational studies of groups by Bales [1999]. The categories are shown in Table 1. We also look at ten emotions (positive: *Thanks, Calm, Cautious, Happy* and negative: *Sorry, Nervous, Careless, Aggressive, Defensive, Angry*) that span a three dimensional emotional space as identified by and related to IPA categories by Heise [2013]. The association between group behavior taxonomies and emotions builds upon the insight that emotions are fundamental tools for the subtle coordination of human social interaction. In the following we discuss our efforts towards classifying interactions on GitHub into these IPA and emotion categories. The ability to make such classifications will allow us to build and test group process simulations as in Heise [2013] at a massive, unprecedented scale.

We randomly selected 834 pull requests and a total of 3000 pull request comments from GHTorrent's GitHub dump up to February 2017. Out of the 834 pull requests, 41 were open, 343 were closed without being merged and 450 were merged. The comments were filtered to remove sections of code, then annotated by four coders using the twelve IPA labels and ten emotions described above. One annotator was the first author of the study, while three were hired on Amazon Mechanical Turk (MTurk) who had experience in programming and had heard of Github. Further, they were screened according to their ratings on an initial set of 50 pull request comments according to their ratings agreement with two IPA raters from our lab (one the first author of the study). Detailed instructions on how to annotate a particular pull request comment were provided, and each comment could be annotated into a maximum of three IPA categories and a maximum of three emotions. The participants were also asked to filter out any unnecessary sections of code. Majority voting was used to threshold all the ratings (see Table 1 for examples).

We used a linear Support Vector Machine (SVM) with Term Frequency–Inverse Document Frequency (TF-IDF) vectors for each comment, and show the F1-scores (evenly weighted precision and recall) for a one-vs-all classification task of all IPA categories and all emotions in Table 2(a) and (b), resp. Parameters were set by maximizing F1-score in a grid search and results are for a 5-fold cross validation. We also examined aggregated IPA and emotion categories by grouping IPA categories into positive vs. negative reactions and questions vs. attempted answers, and grouped emotions into positive and negative categories. Using linear SVM, and averaged TF-IDF weighted Google word vectors, the results in 2(c) show that this task is much simpler, and F1-scores over 0.73 can be achieved [3].

The results of this preliminary data analysis show that the task of sentiment and interaction analysis is a major challenge in cases with more objective conversations than what is usually attempted. This is perhaps not a surprise, given that raters that classify the behavior of human groups typically undergo substantial training to ensure they understand the theoretically-based category systems well [Bales, 1999]. And yet, it is known that subjective emotional and so-

[3]Logistic regression, metric learning, and a deep learning methods yielded similar results (see Rishi [2017]).

cial interactions play a significant role in the online software development process. We have therefore exposed a significant gap for research in this area. Our current work is aimed at more fine-grained (sentence or word-level) sentiment analysis [Alhothali and Hoey, 2015], and further group process analysis that may provide top-down information which can improve the overall effectiveness of the analysis. Longer term goals include the development of artificial agents to assist in software development by catalyzing more effective group processes online.

IP group	IP Category	Example pull request comment	Emotions
positive reactions	Shows solidarity	*im sure youll recover somehow*	Calm
	Shows tension release	*ooops sorry my mistake*	Sorry, Careless
	Agrees	*allright will do thanks for the feedback*	Thanks, Calm
attempted answers	Gives suggestion	*needs a metric tonne of docs*	Cautious
	Gives opinion	*love it*	Happy
	Gives orientation	*fucking hell im hungry now*	Aggressive, Angry
questions	Asks for orientation	*what if the file does not exist*	Nervous, Cautious
	Asks for opinion	*what about filtering by type and tag*	Cautious
	Asks for suggestion	*how could i show the name of the fighter that wins the turn*	Calm, Cautious
negative reactions	Disagrees	*for me just says linux which is not very useful at all*	Aggressive
	Shows tension	*um i dont know i dont remember changing that and probably did it by accident*	Nervous, Defensive
	Shows antagonism	*Kill this method with an axe and then burn its body*	Defensive, Aggressive

Table 1: IP categories used in the study, along with example comments and emotion ratings.

IPA Category	F1
Shows Solidarity	56.8
Shows Tension Release	10.0
Agrees	**64.0**
Gives Suggestion	33.4
Gives Opinion	51.4
Gives Orientation	58.6
Asks for Orientation	36.2
Asks for Opinion	22.9
Asks for Suggestion	10.6
Disagrees	56.6
Shows Tension	30.0
Shows Antagonism	13.2

(a)

Emotion	F1
Thanks	54.7
Sorry	58.7
Calm	69.3
Nervous	23.6
Careless	15.7
Cautious	**69.8**
Aggressive	25.2
Defensive	16.7
Happy	2.5
Angry	0

(b)

Aggregated sets	F1-score
positive vs negative reactions	73.2
questions vs. attempted answers	81.0
positive vs negative emotions	80.5

(c)

Table 2: One vs. All classifications: (a) IP categories; (b) Emotions; (c) aggregated classes

Acknowledgments This work was supported in the framework of the TransAtlantic Platform by NSERC, SSHRC, the DFG (Germany) and the NSF (USA). The study was reviewed by and received full ethics clearance by the Universisty of Waterloo Office of Research Ethics Review Board.

References

Areej Alhothali and Jesse Hoey. Good news or bad news: Using affect control theory to analyze readers' reaction towards news articles. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1548–1558, Denver, Colorado, May–June 2015. Association for Computational Linguistics. URL `http://www.aclweb.org/anthology/N15-1178`.

Scott Ambler. *Agile modeling: effective practices for extreme programming and the unified process*. John Wiley & Sons, 2002.

Robert Freed Bales. *Social Interaction Systems: Theory and Measurement*. Transaction Publishers, New Brunswick, NJ, 1999.

M. Blowfield and L. Johnson. *Turnaround challenge: Business and the city of the future*. OUP, 2013.

M. Bogers and J. West. Managing distributed innovation: Strategic utilization of open and user innovation. *Creat. Innov. Manage.*, 21:61–75, 2012.

Munmun De Choudhury and Scott Counts. Understanding affect in the workplace via social media. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 303–316, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1331-5. doi: 10.1145/2441776.2441812. URL `http://doi.acm.org/10.1145/2441776.2441812`.

BL Fredickson. The role of positive emotions in positive psychology. *American psychologist*, 56(3):218–226, 2001.

Emitza Guzman, David Azócar, and Yang Li. Sentiment analysis of commit comments in github: An empirical study. In *Proceedings of the 11th Working Conference on Mining Software Repositories*, MSR 2014, pages 352–355, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2863-0. doi: 10.1145/2597073.2597118. URL `http://doi.acm.org/10.1145/2597073.2597118`.

David R. Heise. Modeling interactions in small groups. *Social Psychology Quarterly*, 76:52–72, 2013.

Dirk Helbing and E. Pournaras. Build digital democracy. *Nature*, 527:33–34, 2015.

Alessandro Murgia, Parastou Tourani, Bram Adams, and Marco Ortu. Do developers feel emotions? an exploratory analysis of emotions in software artifacts. In *Proceedings of the 11th Working Conference on Mining Software Repositories*, MSR 2014, pages 262–271, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2863-0. doi: 10.1145/2597073.2597086. URL `http://doi.acm.org/10.1145/2597073.2597086`.

Daniel Pletea, Bogdan Vasilescu, and Alexander Serebrenik. Security and emotion: sentiment analysis of security discussions on github. In *Proceedings of the 11th working conference on mining software repositories*, pages 348–351. ACM, 2014.

Deepak Rishi. Affective sentiment and emotional analysis of pull request comments on github. Master's thesis, University of Waterloo, 2017.

Tobias Schröder, Jesse Hoey, and Kimberly B. Rogers. Modeling dynamic identities and uncertainty in social interactions: Bayesian affect control theory. *American Sociological Review*, 81(4), 2016.

A.M. Townsend. *Smart cities: Big data, civic hackers, and the quest for a new utopia*. Norton, 2013.