Emotion Impacts in Coding Efficiency

Jiaming Xu
jm.xu@mail.utoronto.ca
1007698831, Group 8, Department of ECE, University of Toronto
Toronto, Ontario, Canada

Teng Yue
teng.yue@mail.utoronto.ca
1007826792, Group 8, Department of ECE, University of Toronto
Toronto, Ontario, Canada

Wenrui Xu
wenrui.xu@mail.utoronto.ca
1008313228, Group 8, Department of ECE, University of Toronto
Toronto, Ontario, Canada

ABSTRACT

This report mainly discussed the impacts of different emotions may have on coding efficiency. We use both quantitative and qualitative analysis to find the results. For the quantitative analysis, we investigate the open-source repositories, mine the emotion hidden inside the commit messages and find the relationship between emotion and coding efficiency. For the qualitative analysis, we make some interviews with software engineers from dot-com companies, research teams and open-source community. Besides, some participants also shared their project emotional experience on other engineer-related approaches, which indicates some general clues of emotion impacts in computer engineering.

KEYWORDS

Emotion Mining, Coding Efficiency, Computer Engineering

1 INTRODUCTION

Brief introduction of the research topic, including the background information, the motivation and contribution of our research.

2 LITERATURE REVIEW

Detailed information of definition of emotions, categorization and emotional mining.
Major references are:[4],[3],[6],[10],[8],[9],[1],[7],[2],[5].

2.1 Definition & Categories of Emotions

The definition and features of emotions from a psychological aspect. Using a consistent categorization standard to analyze the emotions. Explain the categories and the words which belongs to these emotions. In this part, we will discuss definition from wikipedia[6], the Ekman emotion framework[3].

2.2 Emotion and productivity

In this part we will discuss how does emotion affect productivity. The papers are [5],[2],[7],[1],[8]

2.3 Emotional Mining

Introduce the emotional mining methods used in github. The papers discussed in this part are [9],[10],[4].

3 RESEARCH QUESTIONS

There are a number of research questions we plan to figure out through the research.

- How do we define and categorize different kinds of emotion?
- How do we know engineers' emotion from their codes?
- How do we measure the efficiency of coding?
- Does the emotion change have impact on the efficiency of coding?
- What effects will the emotion change cause on coding efficiency?

4 METHODOLOGY

We will collect and analyze our data in both qualitative and quantitative ways.

4.1 Quantitative Methods

We will use the emotional mining methods to mine the emotions in the selected repositories which are easy to analyze. Then we will calculate the coding efficiency of each mined commit messages accordingly.

4.2 Qualitative Methods

We designed an interview to make qualitative study on different groups of software engineers, including researchers in Universities, software engineers in large companies and open-source developers.

5 RESULTS

This part will show the results of both quantitative and qualitative research methods.

5.1 Quantitative Results

This part will show the results of emotional mining.

5.2 Qualitative Results

This part will show the results of interviews.

6 DISCUSSION

We will analyze the results of the study to show the relationship between different emotions and coding efficiency in both qualitative and quantitative ways.

7 CONCLUSION

Conclude the paper by summarizing the findings of the research questions and our contributions. Then we will consider future works to improve our researches.

REFERENCES

[1] Sigal G Barsade and Donald E Gibson. 2007. Why does affect matter in organizations? *Academy of management perspectives* 21, 1 (2007), 36–59.
[2] Decio Coviello, Erika Deserranno, Nicola Persico, and Paola Sapienza. 2020. *Effect of mood and worker incentives on workplace productivity.* Technical Report. Mimeo.
[3] Paul Ekman. 1999. Basic emotions. *Handbook of cognition and emotion* 98, 45-60 (1999), 16.

[4] Ramiro Gómez. 2012. Exploring Expressions of Emotions in GitHub Commit Messages. Website. https://geeksta.net/geeklog/exploring-expressions-emotions-github-commit-messages.
[5] Yoshihiko Kadoya, Mostafa Saidur Rahim Khan, Somtip Watanapongvanich, and Punjapol Binnagan. 2020. Emotional status and productivity: Evidence from the special economic zone in Laos. *Sustainability* 12, 4 (2020), 1544.
[6] NATALIE MENDES. 2017. How do emotions affect productivity? [New research]. Website.
[7] Muhammad Ali Pervez. 2010. Impact of emotions on employee's job performance: An evidence from organizations of Pakistan. *OIDA International Journal of Sustainable Development* 1, 5 (2010), 11–16.
[8] Phillip Shaver, Judith Schwartz, Donald Kirson, and Cary O'connor. 1987. Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology* 52, 6 (1987), 1061.
[9] Mohamed Yassine and Hazem Hajj. 2010. A Framework for Emotion Mining from Text in Online Social Networks. In *2010 IEEE International Conference on Data Mining Workshops*. 1136–1142. https://doi.org/10.1109/ICDMW.2010.75
[10] Jinxuan (Janice) Zhou, Vrushank Phadnis, and Alison Olechowski. 2020. Analysis of Designer Emotions in Collaborative and Traditional Computer-Aided Design. *Journal of Mechanical Design* 143, 2 (08 2020). https://doi.org/10.1115/1.4047685 arXiv:https://asmedigitalcollection.asme.org/mechanicaldesign/article-pdf/143/2/021401/6558698/md_143_2_021401.pdf 021401.

A CRITIQUES AND SOLUTIONS

Critique: "one concern that comes to my mind is that developers are often required to write commit messages in a certain format, which may not be able to convey emotions. For example, some short messages like '[merge] merge from main branch' are neural and could be difficult to analyze. How would you cleanse the dataset while remaining unbiased is a challenge. Overall, the proposal was well structured and the research methods make sense to me."
Solution: We make an innovative design to select some open-source GitHub repositories which contains relatively more valid data. Besides, we will pick some of the commits from which are easy to analyze emotions hidden inside, and abandon the others by using some text filters. We will reduce the bias by calculating the efficiency with the time between the selected commit logs and its previous logs. So we think this does not add much bias to our results.

Critique: "Suggestions: - Your research topic is really interesting! Lots of great work is being done around the emotions of engineers/designers/developers. If you need some inspiration for how to analyze and process data related to emotions, here is a paper that investigates the emotions during mechanical design, that I think could be relevant: https://doi.org/10.1115/1.4047685 - At times, the writing is vague or unclear. It may be more impactful for the statements/claims you make if you use examples and provide sources. - Try to reconsider using lines of code as your measure of working efficiency, considering it was used quite a few times in class as an example of a flawed metric. - Try to add more to your motivation for the next deliverables and really prove that what you are studying is important and impactful!"
Solution: We read the paper and try to improve our writing by adding more examples, and we notice that you made a good point about the measurement. We are considering using the coding value metrics as a backup option. And we will use more examples to prove our study's significance.

Critique: "It is worth noting that generally speaking, the wording of the common commit logs is usually quite neutral. How to analyze the emotions of the programmers based on these commit messages accurately deserves further refinement. "
Solution: We have refined our mining methods and try to mine the emotions from the commit messages using Regular Expression. We will pick the messages which contains the words expressing their feelings.

Critique: "I only have one concern that the fourth research question, 'Does the emotion change have impact on the efficiency of coding?' is a close-end question, maybe some further supplement to find some exact relationship or modifying it to be a 'what' question would be better? And I would suggest that combining qualitatively analyzing the coding efficiency via something like self-report of developers may also provide some help."
Solution: We added a new RQ to ask 'What impact the emotion change will have on the efficiency of coding?'. And we have changed part of our plans to include interviews of engineers/open-source developers.

Critique: "First, the use of interviews may generate biases, such as people who are willing to be interviewed tend to show a more positive mood when they are chosen and tracked, etc. Also, you intend to analyze the programmer's commit messages to analyze the mood, but in reality, in many cases, the commit messages are probably just a very concise but objective description of the content of this code commit, not mixed with subjective emotions, so I think more considerations are needed to choose the quantitative analysis object. "
Solution: We carefully select the interviewees instead of just recruit them with a reward. The interviewees are selected by us, not recruited with an advertisement. And we select them considering different factors including working experience, ages, job positions, etc. So I don't think there exist such problems. As to the second point, we have mentioned it before.

Critique: "Probably need more literature reviews on previous studies for comparison"
Solution: We will read more papers and references to add more literature reviews. Thanks for your suggestion.

Critique: "Firstly, the research questions are specific, but the first three questions seem more about the background information instead of the analysis result of collected data. If it is possible to add more questions about the connection between emotion and code efficiency, like will engineers (students) try to switch their emotions when they try to promote efficiency? Is that useful? Do engineers realize which emotion is most optimal for their code efficiency? etc. Secondly, for the interview part, whether keeping a record of mood and code efficiency at the beginning of working will have an impact on their work efficiency? Since they may be more productive if they record their progress."
Solution: For the first problem, we think emotion is mostly an unconsciously psychological activity, so most of time, people will not try to change it consciously. So we add some questions in our interviews about the emotion changes and the coding efficiency changes to solve this problem. For the second point, we revised our interview methods. We will not record their performance any more for two reasons. First, the time of research is too short to

complete some projects. So it's meaningless to record their coding efficiency in a month. Second, as you said, if recorded, they will realize that they are monitored and increase their working efficiency unconsciously. So we just use an interview to ask them recall their emotion change, working efficiency and evaluation during the project.

Critique: "For the research questions, I think most of them are description questions, while other types of questions are relatively few. I suggest that more frequency and descriptive-process questions can be supplemented to the research questions. "
Solution: We revised the RQ to add more descriptive-process questions. It's a very good point, thanks for your comment.

B TODOS FOR THE NEXT 2 WEEKS

Here is our plan of work for the next 2 weeks.

- Keep interviewing more software engineers, including open-source coders, junior/senior engineers, PhD. students and researchers.
- Get to know more about the emotion mining from the commit messages, and get the valid commit logs as the raw data.
- Select some appropriate open-source repositories which are easy to analyze as the study case.
- Record and translate the interview into English version to analyze the cases.