In the Eyes of the Beholder: Sentiment and Topic Analyses on Social Media Use of Neutral and Controversial Terms for COVID-19

Long Chen, Tongyu Yang, Jiebo Luo  
Department of Computer Science  
University of Rochester  
{lichen62,tyang20}@u.rochester.edu, jluo@cs.rochester.edu

Hanjia Lyu  
Goergen Institute for Data Science  
University of Rochester  
hlyu5@ur.rochester.edu

Yu Wang  
Political Science  
University of Rochester  
ywang176@ur.rochester.edu

Abstract—During the COVID-19 pandemic, “Chinese Virus” emerged as a controversial term for coronavirus. To some, it may seem like a neutral term referring to the physical origin of the virus. To many others, however, the term is in fact attaching ethnicity to the virus. In this paper, we attempt to shed light on the term’s real-world usage on Twitter. Using sentiment feature analysis and topic modeling, we reveal substantial differences between the use of the controversial terms such as “Chinese virus” and that of the non-controversial terms such as “COVID-19”. For example, tweets using controversial terms contain a higher percentage of anger as well as negative emotions. They also point to China more frequently. Our results suggest that while the term “Chinese virus” could be interpreted either as neutral or racist, its usage on social media leans strongly towards the latter.

Index Terms—COVID-19, Topic Modeling, LIWC2015, Twitter, Controversial Term, Social Media

I. INTRODUCTION

Starting in late 2019, the COVID-19 pandemic has rapidly impacted over 200 countries, areas and territories. As of April 19, according to the World Health Organization (WHO), 2,241,359 COVID-19 cases were confirmed worldwide, with 152,551 deaths. This disease has tremendous impacts on the world’s populations daily lives.

In light of the deteriorating situation in the United States, discussions of the pandemic on social media has drastically increased since March 2020. Within these discussions, an overwhelming trend is the use of controversial terms targeting Asian and, specifically, Chinese population, insinuating that the virus originated in China. On March 16, the President of United States, Donald Trump, posted on Twitter calling COVID-19 the Chinese Virus. Around March 18, media coverage of the term Chinese Flu also took off. Although most public figures who used the controversial terms claimed them to be non-discriminative, such terms have stimulated racism and discrimination against Asian-Americans in the US, as reported by New York Times, the Washington Post, and other mainstream news media.

A recent work was done with social media data to characterize users who used controversial or non-controversial terms associated with COVID-19 and found the associations between demographics, user-level features, political following status, and geo-location attributes with the use of controversial terms. In this study, we analyze from a language perspective crawled tweets (Twitter posts) with and without controversial terms associated with COVID-19. To operationalize this idea, we perform two investigations. First, latent Dirichlet allocation (LDA) is applied to extract the topics in controversial and non-controversial posts. Next, LIWC2015 (Linguistic Inquiry and Word Count 2015) is applied to build multi-dimensional profiles of the posts. We then make comparisons between the topics and profiles presented in both controversial and non-controversial posts, trying to investigate any association between the use of controversial terms and the underlying mindsets.

II. RELATED WORK

Our work is built upon previous works on text mining using data from social media during influential events.

Studies have been conducted using topic modeling, a process of identifying topics in a collection of documents. The commonly used model, Latent Dirichlet Allocation (LDA), provides a way to automatically detect hidden topics in a given number. Previous research has been conducted on inferring topics on social media. Kim et al. investigated the topic coverage and sentiment dynamics on Twitter regarding the issue of Ebola. Chen et al. found LDA-generated topics from e-cigarette related posts on
A large number of studies were performed with LIWC, an API to do linguistic analysis of documents. Tumasjan et al. [10] used LIWC to capture the political sentiment and predict elections with Twitter. The API was also used by Zhang et al. [13] to provide insights into the sentiment of the descriptions of crowdfunding campaigns. Our motivation is to combine qualitative analysis with LDA and quantitative analysis with LIWC, comparatively investigate discrepancies between the tweets that use controversial terms associated with COVID-19 and the tweets that use non-controversial terms.

III. Data and Methodology

The related tweets (Twitter posts) were crawled with the Tweepy API using keyword filtering in reference to [6]. Simultaneous streams were collected to build the controversial dataset (CD) and the non-controversial dataset (ND) from March 23 April 5. The controversial keywords consist of “Chinese virus” and “#ChineseVirus, whereas non-controversial keywords include “corona, “covid-19, “covid19, “coronavirus, “#Corona, “#Covid 19. In total, 2,607,753 tweets for CD and 69,627,062 tweets for ND were collected. We then randomly sampled 2 million tweets from the two datasets, respectively. For preprocessing, we removed all URLs, emails and newlines, as they are not informative for language or textual analysis.

### Table 1: Topics generated by the LDA model for both controversial and non-controversial tweets.

| Classification | Topics                  | Top 10 Topic Words                                                                 |
|----------------|-------------------------|----------------------------------------------------------------------------------|
| Controversial  | Racism, Anecdote, Conspicacy, Work in hospital, Blame the lie | call chinese virus stop racist people let kill pandemic infect virus chinese people say world covid would death call take chinese spread virus be send full conspiracy exactly theory vid get good chinese test fight virus hospital need work last chinese virus must people lie blame tweet fact need die |
| Non-controversial | Test cases, Anecdote, Trump, Health workers, Stay home | case test covid death new day positive patient break number tell virus covid man story spread free say try government say people know get trump could make would help need covid crisis health worker work government pandemic people virus home stay corona take get die |

Note: All appearances of “Chinese virus” related keywords were removed prior to the LDA process. Bigrams and trigrams were included in the LDA model. None of them appears in the top 10 topic words due to infrequency. Some topics contain less than 10 topic words due to deletion of some short (less than 3 characters) words.

Next, LIWC2015 was applied to extract the sentiment of the tweets of CD and ND. LIWC2015 is a dictionary-based linguistic analysis tool that can count the percentage of words that reflect different emotions, thinking styles, social concerns, and capture people’s psychological states. We focused on 4 summary linguistic variables and 12 more detailed variables that reflect psychological states, cognition, drives, time orientation, and personal concerns of the Twitter users of CD and ND. We followed the similar methodology used by Yu et al. [12] by concatenating all tweets posted by users of CD and ND, respectively. One text sample was composed of all the tweets from the aforementioned sampled dataset of CD, and the other was composed of all the tweets from that of ND. Then we applied LIWC2015 to analyze these two text samples. In the end, there are 16 linguistic variables for CD and ND, respectively.

IV. RESULTS

A. Topic Modeling Results

In Table 1, 5 topics generated by LDA on CD and ND are reported, respectively, together with the top 10 topic words. We manually assigned each topic a topic name to generalize what would most likely be discussed under the topic.

Comparing across CD and ND, we observe that the topics in CD contain more opinions. 3 out of 5 topics have very strong opinion-related topic words, which are highlighted in Table 1. In the topic “blame the lie, a very strong signal about lie and blame is present, as well as indication of some correlation between “lie and people “die. In contrast, the generated topics in ND are more related to fighting the pandemic (n=3), stories (n=1) and government (n=1), which are all related to COVID-19 in the US. No strong opinion-related keywords could be found.

One finding is that all 5 topics in CD contain the topic word “chinese, even though we have removed all keywords related

---

7https://liwc.wpengine.com/

8https://liwc.wpengine.com/how-it-works/
to Chinese virus in the documents for LDA. This suggests that discussions in CD are closely related to China or the Chinese people/government. In addition, all 5 topics in CD contain topic word “virus, whereas only 2 of 5 topics in ND contain this topic word. Such a difference suggests that in CD, discussions are more related to the virus, while in ND discussions are more related to the pandemic caused by this virus. In fact, only one topic in CD is about how people work towards containing the pandemic (“work in hospital”), whereas 3 topics in ND are discussing measures to relieve this pandemic (“test cases,” “health workers” and “stay home”).

These discrepancies in the topic modeling result contradict the claim of “only referring to the geo-locational origin of the pandemic” by some public figures who employed the use of “Chinese virus” when referring to COVID-19. Nevertheless, such words have provoked, to certain degree, racist or xenophobic opinions and hate speeches towards China or people with Chinese ethnicity on social media. Furthermore, hate speeches can spread extremely fast on online social media platforms and can stay online for a long time [3]. Gagliardone et al. [3] found that such speeches are also itinerant, meaning that despite forcefully removed by the platforms, one can still find related expression elsewhere on the Internet and even offline.

### TABLE II

| Variables          | Controversial | Non-Controversial |
|--------------------|---------------|-------------------|
| i                  | 0.96          | 1.04              |
| we                 | 1.25          | 1.0               |
| she/he             | 0.69          | 0.7               |
| they               | 1.05          | 0.71              |
| present orientation| 9.37          | 9.22              |

### B. LIWC Sentiment Features

Fig. 1 shows 4 summary variables for CD and ND. We observe that the clout scores for CD and ND are similar. A high clout score suggests that the author is speaking from the perspective of high expertise [8]. At the same time, analytical thinking, authentic and emotional tones scores for ND are higher than those for CD. The analytical thinking score reflects the degree of hierarchical thinking. Higher numbers indicate a more logical and formal thinking [8]. A higher authentic score suggests that the content of the text is more honest, personal and disclosing [8]. The emotional tones scores for CD and ND are both lower than 50, indicating that the overall emotions for CD and ND are negative. This is consistent with our expectation. However, the emotional tone score for ND is higher than that for CD, indicating that the Twitter users in ND are expressing a relatively more positive emotion.

Fig. 2 shows 12 more detailed linguistic variables of tweets of CD and ND. The scores of “future-oriented” and “past-oriented” reflect the temporal focus of attention of the Twitter users by analyzing the verb tense used in the tweets [9]. The tweets of ND are more future-oriented, while those of CD are more past-oriented. To better understand this difference, we conducted a similar analysis as Gunsch et al. [4]. We extract five more linguistic variables including four pronouns score and one time orientation score. The scores of “i”, “we”, “she/he”, “they”, and present-orientation are shown in Table 1. The tweets of CD show more other-references (“they”), whereas more self-references (“i”, “we”) are present in the tweets of ND. The scores for “she/he” of CD and ND are close. The score of present orientation of CD is higher than that of ND. From this similar observation to the findings of Gunsch et al. [4], we can infer that the tweets of CD focus on the past and present actions of the others, and the tweets of ND focus more on the future acts of themselves. Research shows that LIWC can identify the emotion in language use [9]. From the aforementioned discussion, the tweets of both CD and ND are expressing a negative emotion, and the emotion expressed by the Twitter users of ND is relatively more positive. This is consistent with the positive emotions score and negative emotions score. However, there are nuanced differences across the sadness, anxiety and anger scores. When referring to COVID-19, the tweets of ND express more sadness and anxiety than those of CD do. More anger is expressed through the tweets of CD. The certainty and tentativeness scores reveal the extent to which the event the author is going through may have been established or is still being formed [9].
A higher percentage of words like “always” or “never” results in a higher score for certainty, and a higher percentage of words like “maybe” or “perhaps” leads to a higher score for tentativeness [8]. We observe a higher tentative score and a higher certainty score for the tweets of CD, while these two scores for the tweets of ND are both lower. We have an interesting hypothesis for this subtle difference. Since 1986, Pennebaker et al. [8] have been collecting text samples from a variety of studies including blogs, expressive writing, novels, natural speech, New York Times, and Twitter to get a sense of the degree to which language varies across settings. Of all the studies, the tentative and certainty scores for the text of New York Times are the lowest. However, these two scores for expressive writing, blog, and natural speech are relatively higher. This observation leads to our hypothesis that the tweets of CD are more like blog, expressive writing, or natural speech that focus on expressing ideas, whereas the tweets of ND are more like newspaper articles which focus more on one describing facts. As for the score of “achievement”, McClelland [7] found that the stories people told in response to drawings of people could provide important clues to their needs for achievement. We hypothesize that the higher value of the “achievement” score for the tweets of ND reflects the need of these Twitter users to succeed in fighting against COVID-19. As for the personal concerns, the scores for “work” and “money” of ND are both higher than those of CD which shows that the Twitter users of ND focus more on the work and money issue (e.g. working from home, unemployment). According to the reports of the U.S. Department of Labor, the advance seasonally adjusted insured unemployment rate was 8.2% for the week ending April 4. The previous high was 7.0% in May of 1975 [10].

V. CONCLUSION AND FUTURE WORK

We have presented a study of the topic preference related to the use of controversial and non-controversial terms associated with COVID-19 on Twitter during the ongoing COVID-19 pandemic. We first use LDA to extract topics from the controversial and non-controversial posts crawled from Twitter, and then qualitatively compare them across the two sets of posts. We find that topics in the controversial posts are more related to China, even after the keywords related to “Chinese virus” were removed before the analysis, whereas discussions in non-controversial posts are more related to fighting the pandemic in the US. We also find differences across the sentiment of the tweets posted by the users using controversial terms and the users using non-controversial terms. Both groups express a negative emotion, yet the tweets of ND are relatively more positive. The tweets of ND also show more analytical thinking and are expressed in a more truthful manner. The tweets of CD focus more on the past and present action of others, while the tweets of ND focus more on the future acts of the authors themselves. More anger is present in the tweets of CD, while more anxiety and sadness are observed in the tweets of ND. More tentativeness and certainty are observed in the tweets of CD, which is not contradictory since these two scores are both higher in the text samples from blogs and expressive writings that focus on expressing ideas and opinions. These two scores are both lower for the tweets of ND which is similar to the case of newspaper articles like New York Times. Tweets of ND reflect a strong need for achievement. As for the personal concerns, users of ND focus more on work and money issues.

It is reported that the widespread use of controversial terms associated with COVID-19 has induced hate speeches and, to some degree, racism and xenophobia on social media. Therefore, such content on social media should be closely monitored to prevent a further escalation in the situation, which could potentially lead to social unrest. Our next step is to use textual, demographic and account-level features to detect the use of hate speeches on social media, in an effort of predicting and analyzing such behaviors for uses including but not limited to social media content management and policy-making.

REFERENCES

[1] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation.  Journal of machine Learning research, 3(Jan):993–1022, 2003.
[2] Long Chen, Xinyi Lu, Jianbo Yuan, Joyce Luo, Jiebo Luo, Zidian Xie, and Dongmei Li. A social media study on associations of flavored e-cigarette with health symptoms: Observational study (preprint). December 2019.
[3] Iginio Gagliardi, Danit Gal, Thiago Alves, and Gabriela Martinez. Countering online hate speech. Unesco Publishing, 2015.
[4] Mark A Gunsch, Sheila Brownlow, Sarah E Haynes, and Zachary Mabe. Differential forms linguistic content of various of political advertising. Journal of Broadcasting & Electronic Media, 44(1):27–42, 2000.
[5] Erin Hea-Jin Kim, Yoo Kyung Jeong, Yuyoung Kim, Keun Young Kang, and Min Song. Topic-based content and sentiment analysis of ebola virus on twitter and in the news. Journal of Information Science, 42(6):763–781, 2016.
[6] Hanjia Lyu, Long Chen, Yu Wang, and Jiebo Luo. Sense and sensibility: Characterizing social media users regarding the use of controversial terms for covid-19, 2020.
[7] David C McClelland. Inhibited power motivation and high blood pressure in men. Journal of Abnormal Psychology, 88(2):182, 1979.
[8] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. The development and psychometric properties of liwc2015. Technical report, 2015.
[9] Yia R Tausczik and James W Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. Journal of language and social psychology, 29(1):24–54, 2010.
[10] Andranik Tumasjan, Timm O Sprenger, Philipp G Sandner, and Isabella M Welpe. Predicting elections with twitter: What 140 characters reveal about political sentiment. In Fourth international AAAI conference on weblogs and social media, 2010.
[11] Yu Wang, Jiebo Luo, Richard Niemi, Yuncheng Li, and Tianran Hu. Catching fire via “likes”: Inferring topic preferences of trump followers on twitter. In Tenth International AAAI Conference on Web and Social Media, 2016.
[12] Bei Yu, Stefan Kaufmann, and Daniel Diermeier. Exploring the characteristics of opinion expressions for political opinion classification. In Proceedings of the 2008 international conference on Digital government research, pages 82–91. Digital Government Society of North America, 2008.
[13] Xupin Zhang, Hanjia Lyu, and Jiebo Luo. What contributes to a crowdfunding campaign’s success? evidence and analyses from gofundme data. arXiv preprint arXiv:2001.05446, 2020.