Analyzing COVID-19 on Online Social Media: Trends, Sentiments and Emotions

Xiaoya Li*, Mingxin Zhou*, Jiawei Wu*, Arianna Yuan*, Fei Wu* and Jiwei Li*
* Department of Computer Science and Technology, Zhejiang University
* Computer Science Department, Stanford University
* Shannon.AI
{xiaoya_li, mingxin_zhou, jiawei_wu, jiwei_li}@shannon.ai.com
xfyuan@stanford.edu, wufei@cs.zju.edu.cn

Abstract—At the time of writing, the ongoing pandemic of coronavirus disease (COVID-19) has caused severe impacts on society, economy and people’s daily lives. People constantly express their opinions on various aspects of the pandemic on social media, making user-generated content an important source for understanding public emotions and concerns.

In this paper, we perform a comprehensive analysis on the affective trajectories of the American people and the Chinese people based on Twitter and Weibo posts between January 20th, 2020 and May 11th 2020. Specifically, by identifying people’s sentiments, emotions (i.e., anger, disgust, fear, happiness, sadness, surprise) and the emotional triggers (e.g., what a user is angry/sad about) we are able to depict the dynamics of public affect in the time of COVID-19. By contrasting two very different countries, China and the Unites States, we reveal sharp differences in people’s views on COVID-19 in different cultures. Our study provides a computational approach to unveiling public emotions and concerns on the pandemic in real-time, which would potentially help policy-makers better understand people’s need and thus make optimal policy.

I. INTRODUCTION

The emergence of COVID-19 in early 2020 and its subsequent outbreak have affected and changed the world dramatically. According to the World Health Organization (WHO), by mid-May 2020, the number of confirmed COVID-19 cases has reached 5 millions with death toll over 300,000 world wide. Several mandatory rules have been introduced by the government to prevent the spread of the coronavirus, such as social distancing, bans on social gatherings, store closures and school closures. Despite their positive effects on slowing the spread of the pandemic, they nevertheless caused severe impacts on the society, the economy and people’s everyday life. There have been anti-lockdown and anti-social-distancing protests in many places around the world. Given these difficult situations, it is crucial for policy-makers to understand people’s opinions toward the pandemic so that they can (1) balance the concerns of stopping the pandemic on the one hand and keeping people in good spirits on the other hand and (2) anticipate people’s reactions to certain events and policy so that the policy-makers can prepare in advance. More generally, a close look at the public affect during the time of COVID-19 could help us understand people’s reaction and thoughts in the face of extreme crisis, which sheds light on humanity in moments of darkness.

People constantly post about the pandemic on social media such as Twitter, Weibo and Facebook. They express their attitudes and feelings regarding various aspects of the pandemic, such as the medical treatments, public policy, their worry, etc. Therefore, user-generated content on social media provides an important source for understanding public emotions and concerns.

In this paper, we provide a comprehensive analysis on the affective trajectories of American people and Chinese people based on Twitter and Weibo posts between January 20th, 2020 and May 11th 2020. We identify fine-grained emotions (including anger, disgust, fear, happiness, sadness, surprise) expressed on social media based on the user-generated content. Additionally, we build NLP taggers to extract the triggers of different emotions, e.g., why people are angry or surprised, what they are worried, etc. We also contrast public emotions between China and the Unites States, revealing sharp differences in public reactions towards COVID-19 related issues in different countries.

By tracking the change of public sentiment and emotion over time, our work sheds light on the evolution of public attitude towards this global crisis. This work contributes to the growing body of research on social media content in the time of COVID-19. Our study provides a way to extracting public emotion towards the pandemic in real-time, and could potentially lead to better decision-making and the development of wiser interventions to fight this global crisis.

The rest of this paper is organized as follows: we briefly go through some related work in Section 2. We then present the analyses on topic trends in Weibo and Twitter (section 3), the extracted emotion trajectories (section 4) and triggers of those emotions (section 5). We finally conclude this paper in Section 6.

II. RELATED WORK

A. Analyses on Social Media Content about COVID-19

At the time of writing, analyses on people’s discussions and behaviors on social media in the context of COVID-19 has attracted increasing attention. [1] analyzed tweets concerning
COVID-19 on Twitter by selecting important 1-grams based on rank-turbulence divergence and compare languages used in early 2020 with the ones used a year ago. The authors observed the first peak of public attention to COVID-19 around January 2020 with the first wave of infections in China, and the second peak later when the outbreak hit many western countries. They released the first COVID-19 Twitter dataset. Another study provided a ground truth corpus by annotating 5,000 texts (2,500 short + 2,500 long texts) in UK and showed people’s worries about their families and economic situations. They viewed emotions and sentiments on social media as indicators of mental health issues, which result from self-quarantining and social isolation. A follow-up study revealed increasing amount of hateful speech and conspiracy theories towards specific ethnic groups such as Chinese on Twitter and 4chan’s. Other researchers started looking at the spread of misinformation on social media and social isolation. They showed increasing amount of hatefu...
F1 scores for the three rounds of binary classification are 0.74, 0.82, 0.86 respectively. After the final round of bootstrapping, we collected a total number of 78 million English tweets concerning the topic of COVID-19. We used this strategy to retrieve COVID-related posts on Weibo and collected a total number of 16 million posts.

B. Analyses

We report the intensity scores for Weibo and Twitter in Figure 1. We split all tweets by date, where $X_t$ denotes all tweets published on day $t$. The value of intensity is the number of posts classified as COVID-related divided by the total number of retrieved posts, i.e., $|X_t|$. On Weibo, we observe a peak in late January and February, then a drop, followed by another rise in March, and a gradual decline afterwards. The trend on Chinese social media largely reflects the progress of the pandemic in China: the outbreak of COVID-19 and the spread from Wuhan to the rest of the country corresponds to the first peak. The subsequent drop reflects the promise in containing the virus, followed by a minor relapse in March. For Twitter, we observe a small peak that is aligned with the news from China about the virus. The subsequent drop reflects the decline of the attention to the outbreak in China. The curve progressively went up since March, corresponding to the outbreak in the US. Upon the writing of this paper, we have not observed a sign of drop in the intensity score of COVID19-related posts.

IV. THE EVOLUTION OF PUBLIC EMOTION

In this section, we present the analyses on the evolution of public emotion in the time of COVID-19. We first present the algorithms we used to identify the emotions expressed in a given post. Next we present the results of the analyses.
A. Emotion Classification

We adopted the well-established emotion theory by Paul Ekman [12], which groups human emotions into 6 major categories, i.e., anger, disgust, worry, happiness, sadness, and surprise. Given a post from a social network user, we assign one or multiple emotion labels to it [46, 47]. This setup is quite common in text classification [48, 49, 50, 51, 52]. For emotion classification of English tweets, we take the advantage of labeled datasets from the SemEval-2018 Task 1e [20], in which a tweet was associated with either the “neutral” label or with one or multiple emotion labels by human evaluators. The SemEval-2018 Task 1e contains eleven emotion categories in total, i.e., anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise and trust, and we only use the datasets for a six-way classification, i.e., anger, disgust, fear, happiness, sadness, and surprise. Given that the domain of the dataset used in [20] covers all kinds of tweets, and our domain of research covers only COVID-related tweets, there is a gap between the two domains. Therefore, we additionally labeled 15k COVID-related tweets following the guidelines in [20], where each tweet can take either the neutral label or one/multiple emotion labels. Since one tweet can take on multiple emotion labels, the task is formalized as a multi-label classification task, in which six binary (one vs. the rest) classifiers are trained. We used the description-based BERT model [53] as the backbone, which achieves current SOTA performances on a wide variety of text classification tasks. More formally, let us consider a to-be-classified tweet \( x = \{x_1, \ldots, x_L\} \), where \( L \) denotes the length of the text \( x \). Each \( x \) will be tagged with one or more class labels \( y \in Y = [1, N] \), where \( N = 6 \) denotes the number of the predefined emotion classes (the six emotion categories). To compute the probability \( p(y|x) \), each input text \( x \) is concatenated with the description \( q_y \) to generate \( \{[CLS]; q_y; [SEP]; x\} \), where \( [CLS] \) and \( [SEP] \) are special tokens. The description \( q_y \) is the Wikipedia description for each of the emotions. For example, \( q_y \) for the category anger is “Anger, also known as wrath or rage, is an intense emotional state involving a strong uncomfortable and hostile response to a perceived provocation, hurt or threat.” Next, the concatenated sequence is fed to the BERT model, from which we obtain the contextual representations \( h_{[CLS]} \), \( h_{[CLS]} \) is then transformed to a real value between 0 and 1 using the sigmoid function, representing the probability of assigning the emotion label \( y \) to the input tweet \( x \):

\[
p(y|x) = \text{sigmoid}(W_2 \text{ReLU}(W_1 h_{[CLS]} + b_1) + b_2) \quad (1)
\]

where \( W_1, W_2, b_1, b_2 \) are some parameters to optimize. Classification performances for different models are presented in Table 3.

B. Analyses

For emotion \( y \), its intensity score \( S(t, y) \) for day \( t \) is the average probability (denoted by \( P(y|x) \)) of assigning label \( y \) to all the texts in that day \( X_t \). For non COVID-related texts, \( P(y|x) \) is automatically set to 0. We thus have:

\[
S(t, y) = \frac{1}{|X_t|} \sum_{x \in X_t} p(y|x) \quad (2)
\]

For Chinese emotion classification, we used the labeled dataset in [54], which contains 15k labeled microblogs from weibo. In addition to the dataset provided by [54], we labeled COVID-related 20k microblogs. The combined dataset is then used to train a multi-label classification model based on the description-BERT model [53]. Everyday emotion scores for Weibo are computed in the same way as for Twitter.

The time series of intensity scores of six different emotions, i.e., sadness, anger, disgust, worry, happiness, surprise, for Weibo and Twitter are shown in Figures 2 and 3 respectively. For Weibo, as can be seen, the trend of worry is largely in line with the trend of the general intensity of the COVID-related posts. It reached a peak in late January, and then gradually went down, followed by a small relapse in mid-March. For anger, the intensity first went up steeply at the initial stage of the outbreak, staying high for two weeks, and then had another sharp increase around February 8th. The peak on February 8th was due to the death of Wenliang Li, a Chinese ophthalmologist who issued the warnings about the virus. The intensity for anger then gradually decreased, with no relapse afterwards. The intensity for disgust remained relatively low across time. For sadness, the intensity reached the peak at the early stage of the outbreak, then gradually died out with no relapse. For surprise, it went up first, mostly because of the fact that the public was surprised by the new virus and the unexpected outbreak, but then gradually went down. The intensity for happiness remained relatively low across time, with a small peak in late April, mostly because the countrywide lockdown was over.

For Twitter, the intensity for worry went up shortly in late January, followed by a drop. The intensity then went up steeply in mid-March in response to the pandemic breakout in the States, reaching a peak around March 20th, then decreased a little bit and remained steady afterwards. The intensity for anger kept going up after the outbreak in mid-March, with no drop observed. The trend for sadness is mostly similar to that of the overall intensity. For surprise, the curve went up first after the breakout in early March, reaching a peak around Mar 20th, then dropped, and remained steady afterwards. For happiness, the intensity remained low over time.

1 The original dataset contains 7 categories of emotions, and we used only six of them.

| Model                | Acc | micro F1 | macro F1 |
|----------------------|-----|----------|----------|
| SVM diagram          | 51.4| 63.0     | 52.7     |
| BERT [41]            | 65.0| 75.2     | 66.1     |
| BERT-description [53]| 66.8| 77.0     | 68.3     |

TABLE I: Results for the multi-label emotion classification for English tweets.
TABLE II: Top 10 extracted trigger spans regarding different emotions on Twitter.

| Emotion | Trigger Spans |
|---------|---------------|
| sadness | passed away, died, anger |
| disgust | covid / covid-19, Chinese, virus |
| worry   | job, bill, Chinese, virus |
| happiness | healthy, help |
| surprise | covid / covid-19, human-to-human, transmission |

TABLE III: Top mentions of different subcategories for anger on Twitter.

| Trigger Spans   | Families | Finance and Economy | Jobs and Food | Increasing Cases and Deaths |
|-----------------|----------|---------------------|---------------|-----------------------------|
| China lockdown  | quarantined, stay, lockdown |
| Trump hospitals | realdonaldtrump, hospital, patients, test |
| Increasing cases and deaths | increasing, confirmed, case |

TABLE IV: Top mentions of different subcategories for worry on Twitter.

| Trigger Spans   | Families | Finance and Economy | Jobs and Food | Increasing Cases and Deaths |
|-----------------|----------|---------------------|---------------|-----------------------------|
| China lockdown  | quarantined, stay, lockdown |
| Trump hospitals | realdonaldtrump, hospital, patients, test |
| Increasing cases and deaths | increasing, confirmed, case |

V. EMOTIONAL TRIGGERS

For a given emotion, we would like to dive deeper into its different subcategories. For example, for worry, we wish to know what the public is worried about, and how these triggers of worry change over time. In this section, we first present our methods for extracting triggers/subcategories for different emotions, followed by some analyses and visualization on the
Twitter data.

A. Extracting the Triggers for Different Emotions

In order to extract the emotional triggers from Twitter’s noisy text, we first annotate a corpus of tweets. For the ease of annotation, each emotion is associated with only a single trigger: the person/entity/event that a user has a specific emotion towards/with/about. A few examples are shown as follows with target triggers surrounded by brackets.

- **Angry** protesters are traveling 100’s of miles to join organized rallies over COVID-19 [lockdown]attr_anger.
- Feeling awfully tired after a 5.30am start for work today. Worried too about [the early return to school]attr_worry my grandchildren are so very dear to me. I could not bear to lose them to covid.
- All Americans are very angry with [@realDonaldTrump]attr_anger 81,647 dead Americans would be very angry as well. If they weren’t dead.
- Fucking [bat-eating chinks]attr_anger, go die in a hole, far away from us.
- Well, I am ANGRY as hell at Trumpattr_anger.
- With great sadness we report the sad loss of two dear Friendsattr_sadness.
- The lockdownattr_anger was implemented when there were hardly any cases and now it is above lakhs and people are acting so carelessly. Making me so angry.

In order to build an emotional trigger tagger, we annotated 2,000 tweets in total, and split them into training, development and test sets with ratio 8:1:1. We treat the problem as a sequence labeling task, using Conditional Random Fields for learning and inference with BERT-MRC features [55]. Comparing with vanilla BERT tagger [41], the BERT-MRC tagger has the strength of encoding the description of the to-be-extracted entities, e.g., *what they are worried about*. As this description provides the prior knowledge about the entities, it has been shown to outperform vanilla BERT even when less training data is used. In addition to the representation features from BERT-MRC, we also considered the Twitter-tuned POS features [56], the dependency features from a Twitter-tuned dependency parsing model [57] and the Twitter event features [58]. The precision and recall for segmenting emotional triggers on English tweets are reported in Table V.

The precision and recall for segmenting triggering event phrases are reported in Table 3. We observe a significant performance boost with linguistic features such as POS and dependency features. This is mainly due to the small size of the labeled dataset. The best model achieves an F1 score of 0.66.

B. Clustering Trigger Mentions

Since different extracted tokens may refer to the same concept or topic, we would like to cluster the extracted trigger mentions. The method of supervised classification is unsuitable for this purpose since (1) it is hard to predefine a list of potential triggers to people’s anger or worry; (2) it is extremely labor-intensive to annotate tweets with worry types or anger types and (3) these types may change over time. For these reasons, we decided to use semi-supervised approaches that will automatically induce worry/anger types that match the data. We adopt an approach based on LDA [59]. It was inspired by work on unsupervised information extraction [60], [58], [61].

We use the emotion *anger* to illustrate how trigger mentions are clustered. Each extracted trigger mentions for *anger* is modeled as a mixture of *anger* types. Here we use subcategory, type, and topic interchangeably, all referring to the cluster of similar mentions. Each topic is characterized by a distribution over triggers, in addition to a distribution over dates on which a user talks about the topic. Taking dates into account encourages triggers that are mentioned on the same date to be assigned to the same topic. We used collapsed Gibbs
### TABLE V: Performances of different models on emotional trigger extraction from Tweets.

| Model                                                      | Pre  | Rec  | F1   |
|------------------------------------------------------------|------|------|------|
| BERT                                                       | 0.41 | 0.60 | 0.48 |
| BERT-MRC                                                   | 0.54 | 0.66 | 0.59 |
| CRF with BERT-MRC features                                 | 0.53 | 0.68 | 0.60 |
| CRF with BERT-MRC features, POS, event and parse tree      | 0.58 | 0.78 | 0.66 |

C. Analyses

We report the top triggers for different emotions in Table III. We adopt a simple strategy of reporting the most frequent triggers for different emotions. For sadness, the most frequent triggering events and topics are being test positive, and the death of families and friends. For anger, the top triggers are shutdown, quarantine and other mandatory rules. People also express their anger towards public figures such as President Donald Trump, Mike Pence, along with China and Chinese. For worry, the top triggers include jobs, getting the virus, payments and families. For happiness, the top triggers are recovering from the disease, city reopening and returning to
work. For surprise, the public are mostly surprised by the virus itself, its spread and the mass deaths it caused.

Next we report the results of the mention clustering for anger and worry in Tables 3 and 4 respectively. The unsupervised clustering reveals clearer patterns in the triggering events: top subcategories for anger include China with racist words such as chink and chingchong; lockdown and social distancing; public figures like President Donal Trump and Mike Pence; treatments in hospitals, and the increasing cases and deaths; Table 4 displays the change of intensity scores for the subcategories of anger. We observe a sharp increase in public anger toward China and Chinese around March 20th, in coincidence with President Donald Trump calling coronavirus ‘Chinese virus’ in his tweets. Public anger towards the lockdown sharply escalated in mid-March, but decreased a bit after late April when some places started to reopen.

Top subcategories for worry include syndromes for COVID-19, finance and economy, families, jobs and food, and increasing cases and deaths. Table 5 displays the change of intensity scores for subcategories of worry. People increasingly worried about families over time. It is interesting to see that the worry about finance and economy started going up in mid-February, earlier than other subcategories.

VI. CONCLUSION

In this paper, we perform analyses on topic trends, sentiments and emotions of the public in the time of COVID-19 on social media. By tracking the change of public emotions over time, our work reveals how the general public reacts to different events and government policy. Our study provides a computational approach to understanding public affect towards the pandemic in real-time, and could help create better solutions and interventions for fighting this global crisis.

REFERENCES

[1] T. Alshaabi, J. Minot, M. Arnold, J. L. Adams, D. R. Dewhurst, A. J. Reagan, R. Muhamad, C. M. Danforth, and P. S. Dodds, “How the world’s collective attention is being paid to a pandemic: Covid-19 related 1-gram time series for 24 languages on twitter,” arXiv preprint arXiv:2003.12614, 2020.

[2] E. Chen, K. Lerman, and E. Ferrara, “Covid-19: The first public coronavirus twitter dataset,” arXiv preprint arXiv:2003.07372, 2020.

[3] B. Kleinberg, I. van der Vegt, and M. Mozes, “Measuring emotions in the covid-19 real world worry dataset,” arXiv preprint arXiv:2004.04225, 2020.

[4] J. Gao, P. Zheng, Y. Jia, H. Chen, Y. Mao, S. Chen, Y. Wang, H. Fu, and J. Dai, “Mental health problems and social media exposure during covid-19 outbreak,” Plos one, vol. 15, no. 4, p. e0231924, 2020.

[5] L. Schild, C. Ling, J. Blackburn, G. Stringhini, Y. Zhang, and S. Zannettou, “‘Go eat a bat, chang!’: An early look on the emergence of sinophobic behavior on web communities in the face of covid-19,” arXiv preprint arXiv:2004.04046, 2020.

[6] R. Gallotti, F. Vallec, N. Castaldo, P. Sacco, and M. De Domenico, “Assessing the risks of infodemics” in response to covid-19 epidemics,” arXiv preprint arXiv:2004.03997, 2020.

[7] E. Ferrara, “# covid-19 on twitter: Bots, conspiracies, and social media activism,” arXiv preprint arXiv:2004.05637, 2020.

[8] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Britagnoi, A. L. Schmidt, P. Zoka, F. Zollo, and A. Scala, “The covid-19 social media infodemic,” arXiv preprint arXiv:2003.05004, 2020.

[9] R. Plutchik, “A general psychoevolutionary theory of emotion,” in Theories of emotion. Elsevier, 1980, pp. 3–33.

[10] N. H. Frijda, “The laws of emotion.” American psychologist, vol. 43, no. 5, p. 349, 1988.

[11] W. G. Parrott, Emotions in social psychology: Essential readings. Psychology Press, 2001.

[12] P. Ekman, “An argument for basic emotions,” Cognition & emotion, vol. 6, no. 3-4, pp. 169–200, 1992.

[13] Rag, Number Of Rasa-S. Adyar Library Research Centre, 1975, vol. 23.

[14] C. Yang, K.-H.-Y. Lin, and H.-H. Chen, “Emotion classification using web blog corpora,” in IEEE/WIC/ACM International Conference on Web Intelligence (WI’07). IEEE, 2007, pp. 275–278.

[15] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” Foundations and trends in information retrieval, vol. 2, no. 1-2, pp. 1–135, 2008.

[16] S. M. Mohammad, “Sentiment analysis: Detecting valence, emotions, and other affectual states from text,” in Emotion measurement. Elsevier, 2016, pp. 201–237.

[17] D. Tang, B. Qin, and T. Liu, “Document modeling with gated recurrent neural network for sentiment classification,” in Proceedings of the 2015 conference on empirical methods in natural language processing, 2015, pp. 1422–1432.

[18] S. M. Mohammad and F. Bravo-Marquez, “Wessa-2017 shared task on emotion intensity,” arXiv preprint arXiv:1708.03700, 2017.

[19] J. J. Louviere and G. G. Woodworth, “Best-worst scaling: A model for the largest difference judgments,” University of Alberta: Working Paper, 1991.

[20] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, “Semeval-2018 task 1: Affect in tweets,” in Proceedings of the 12th international workshop on semantic evaluation, 2018, pp. 1–17.

[21] G. Mishne, M. De Rijke et al., “Capturing global mood levels using blog posts,” in AAAI spring symposium: computational approaches to analyzing weblogs, vol. 6, 2006, pp. 145–152.

[22] Y. Liu, X. Huang, A. An, and X. Yu, “Arsa: a sentiment-aware model for predicting sales performance using blogs,” in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, 2007, pp. 607–614.

[23] P. S. Dodds and C. M. Danforth, “Measuring the happiness of large-scale written expression: Songs, blogs, and presidents,” Journal of happiness studies, vol. 11, no. 4, pp. 441–456, 2010.

[24] L. Mitchell, M. R. Frank, K. D. Harris, P. S. Dodds, and C. M. Danforth, “The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place,” PloS one, vol. 8, no. 5, 2013.

[25] J. Li, X. Wang, and E. Hovy, “What a nasty day: Exploring mood-weather relationship from twitter,” in proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, 2014, pp. 1309–1318.

[26] P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss, and C. M. Danforth, “Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter,” PloS one, vol. 6, no. 12, 2011.

[27] T. Rao and S. Srivastava, “Tweetsmart: Hedging in markets through twitter,” in 2012 Third International Conference on Emerging Applications of Information Technology. IEEE, 2012, pp. 193–196.
