



Emoji use in China: popularity patterns and changes due to COVID-19

Chuchu Liu¹ · Xu Tan² · Tao Zhou³ · Wei Zhang⁴ · Jianguo Liu⁵ · Xin Lu¹

Accepted: 30 December 2021 / Published online: 21 March 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Emojis are small pictograms that are frequently embedded within micro-texts to more directly express emotional meanings. To understand the changes in the emoji usage of internet users during the COVID-19 outbreak, we analysed a large dataset collected from Weibo, the most popular Twitter-like social media platform in China, from December 1, 2019, to March 20, 2020. The data contained 38,183,194 microblog posts published by 2,239,472 unique users in Wuhan. We calculated the basic statistics of users' usage of emojis, topics, and sentiments and analysed the temporal patterns of emoji occurrence. After examining the emoji co-occurrence structure, we finally explored other factors that may affect individual emoji usage. We found that the COVID-19 outbreak greatly changed the pattern of emoji usage; i.e., both the proportion of posts containing emojis and the ratio of users using emojis declined substantially, while the number of posts remained the same. The daily proportion of Happy emojis significantly declined to approximately 32%, but the proportions of Sad- and Encouraging-related emojis rose to 24% and 34%, respectively. Despite a significant decrease in the number of nodes and edges in the emoji co-occurrence network, the average degree of the network increased from 34 to 39.8, indicating that the diversity of emoji usage increased. Most interestingly, we found that male users were more inclined towards using regular textual language with fewer emojis after the pandemic, suggesting that during public crises, male groups appeared to control their emotional display. In summary, the COVID-19 pandemic remarkably impacted individual sentiments, and the normal pattern of emoji usage tends to change significantly following a public emergency.

Keywords Emoji · COVID-19 · Sentiment analysis · Emotion evolution · Weibo

1 Introduction

Emojis, graphic symbols carrying specific meanings, which have been quickly adopted into online conversations and widely used to express emotions, have become a universal language used across apps, platforms, and

cultures [1–3]. Compared with words, emojis can express individual emotions more accurately and directly [4]. In analysing short informal texts, such as tweets, blogs, or comments, emojis provide crucial information on user sentiment [5–8]. The emergence of emojis provides new opportunities to analyse sentiment expressions in a textual context. Moreover, it has been proven that emerging events, such as natural disasters, may lead to unrepresentative user moods and behaviours and affect the usage of emojis [9]. Therefore, the changes in emoji use patterns can reflect the impact of public health emergencies on personal moods. Despite the popularity of emojis, however, few studies have analysed the usage of emoji characters in social networks at scale and in-depth, especially regarding the emoji use pattern after a public emergency.

The coronavirus disease 2019 (COVID-19) pandemic [10–12] had caused over 242 million cumulative confirmed cases and more than 4.93 million deaths worldwide as of October 22, 2021 [13]. Some studies have indicated that the pandemic not only threatens physical health but also affects individuals' mental state and emotions [14]. According to

✉ Xin Lu
xin.lu@flowminder.org

¹ College of Systems Engineering, National University of Defense Technology, Changsha 410073, China

² Shenzhen Institute of Information Technology, Shenzhen 518172, China

³ Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 610054, China

⁴ West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu 610041, China

⁵ Institute of Accounting and Finance, Shanghai University of Finance and Economics, Shanghai 200433, China

behavioural immune system theory [15], people are more likely to develop negative emotions (such as aversion and anxiety) [16, 17] for self-protection when faced with a public crisis. In line with this theory, stress theory [18] and perceived risk theory [19] also suggest that public health emergencies trigger more negative emotions. These negative emotions keep people away from potential pathogens when referring to the disease. However, long-term negative emotions may reduce the immune function of people and destroy the balance of their normal physiological mechanisms [14, 20]. Therefore, it is essential to promptly understand the potential psychological changes caused by COVID-19.

Here, we present an effort to collect and analyse all emojis embedded by online users in posts during the COVID-19 outbreak on Sina Weibo, the most popular public social media site in China. To the best of our knowledge, this is the first study to examine a large and fine-grained dataset of over 38 million social media posts to understand the evolution of emoji usage before and after COVID-19 and accordingly uncover the impact of this pandemic on individual behaviour and moods. There are three main innovations in this paper:

- (1) We compared emoji usage with the evolution of individual sentiment during normal times and after the COVID-19 pandemic outbreak and revealed the emoji use patterns of Chinese social media users after public health emergencies.
- (2) We built an emoji co-occurrence network and a maximum frequency co-occurrence network between emojis on the Sina Weibo platform, which directly disclosed the co-occurring relationship of different emojis in posts.
- (3) We comprehensively analysed the relationship between the development trend of the pandemic, different post topics, user types (user level, user gender) and individual emoji use habits.

2 Related work

The unprecedented magnitude and transmission speed of COVID-19 has intensified massive social media activities as people are isolated at home to break infection chains [21–26]. The large amount of social data generated by online users can help us understand the topics of concern and emotional fluctuations of the epidemic-related population in real time, which is necessary for timely detection of people's needs and emotional counselling. Recent research has shown that public data on social networks can be employed to analyse the sentiments of internet users.

Nemes and Kiss [27] analysed users' sentiments on the Twitter social media platform by using a recurrent neural network model to determine what emotional manifestations arose

surrounding a given topic. Manguri et al. [28] collected a seven-day dataset from Twitter based on two specified hashtag keywords ("COVID-19, coronavirus") and analysed the sentiment polarity and subjectivity of these COVID-19-related texts. They found that neutral emotions regarding both coronavirus and COVID-19 keywords were significantly high (50%). Currently, studies have attempted to use public data on Chinese social platforms to explore the emotional changes of Chinese citizens during the COVID-19 period. Wang et al. [29] randomly selected 999,978 COVID-19-related Weibo posts from 1 January 2020 to 18 February 2020 and adopted the bidirectional encoder representations from transformers (BERT) model to conduct sentiment classification. Tan et al. [30] conducted a sentiment analysis based on social media posts to examine the long-term negative effects of the COVID-19 pandemic on public sentiments.

However, almost all studies extracted the text content posted by users online as clues for sentiment analysis and focused on classifying sentences into three sentiment categories (positive, neutral, and negative). In addition, due to the expression diversity of Chinese textual content, the variability of Chinese syntax and semantics, and the rapid update of internet slang, it is still challenging to identify users' sentiments from Chinese plain texts with high accuracy [31–34]. In recent years, due to their simplicity, convenience and conduciveness to emotional expression, emojis have been widely used in internet communication. Like nonverbal clues in face-to-face communication, emojis are able to help clarify intentions in ambiguous contexts [35], express emotions [36, 37] and improve the efficiency of communication [38]. Studies in the field of computer science have focused on using emojis as an effective means for sentiment analysis on UGC (user-generated content) data [39, 40].

Li et al. [41] constructed a sentiment model with automatically annotated training sets using an emoji approach, which performs well in classifying tweets into five different sentiment classes. Kaye [42] proved that as a contextual cue, emojis can help users establish an emotional tone, reduce the ambiguity of semantic expressions and improve appropriateness relative to context. Settanni and Marengo [43] investigated the relationship between user-generated textual content shared on Facebook and emotional well-being, and they examined the correlation between emoji usage behaviour and users' stress level. In addition, Barbieri et al. [44] compared the meaning and usage of emojis across two Spanish cities and suggested that cultural differences between different regions can reflect how they communicate in social networks. Cheng et al. [45] found that the different cultural backgrounds of message senders could influence their emoji usage and emotional expressivity.

A number of studies have confirmed the effective performance of emojis in sentiment analysis. However, currently,

there is little research on emoji usage during the COVID-19 period [46, 47], as well as on how public emergencies affect people's emoji usage patterns (emotional expressions). In particular, to the best of our knowledge, this is the first study to examine the emoji usage of Chinese citizens before and after the COVID-19 outbreak.

3 Data and methods

3.1 Data description

Weibo is a popular Twitter-like social media platform in China, making it a rich publicly available data source for opinion mining and sentiment analysis. We collected all data from Weibo that were posted publicly by users located in Wuhan (the capital of Hubei Province in China), including microblog text, posting time, author ID, and gender, from December 1, 2019, to March 20, 2020. By extracting emoji tags from the post texts, we compared the changes in emoji usage patterns and observed the sentiment evolution of Weibo users before and after the COVID-19 outbreak to explore the impact of COVID-19 on users' online behaviour. Overall, 38,183,194 microblog posts posted by 2,239,472 unique users were collected.

We found that emotion tokens, i.e., emoji characters, were commonly used in Weibo posts. There were 15,609,843 posts containing emoji symbols, accounting for 40.88% of all posts. In addition, 1,279,828 users used emojis at least once, accounting for 57.15% of all unique users. Surprisingly, while no substantial change was observed in the daily number of posts over time (Fig. 1A), both the proportion of posts with emojis and the ratio of users using emojis declined substantially after January 18, 2020. Around this time, a high-level epidemiological team led by the well-known Chinese respiratory expert Zhong

Nanshan arrived in Wuhan from Beijing and later confirmed human-to-human transmission of severe acute respiratory syndrome coronavirus 2 (the causative virus for COVID-19) with infections among medical staff (Fig. 1B). The public began to recognize the high risk of COVID-19 and pay more attention to the epidemic. At this time, the average number of emojis used by each user dropped dramatically; in contrast, the number of posts published by each user increased slightly (Fig. 1C). Similarly, we calculated the average number of emojis employed in each post as well as the number of emojis used by each user in each post and found that they both decreased remarkably after January 18, 2020. In addition, the topic keywords discussed on Weibo in the early period of our dataset were analysed, and it was found that "pneumonia" ("肺炎" in Chinese) began to appear among the popular topics on January 5, 2020 (see Fig. S2 in Supplementary Information), even before COVID-19 was discovered and announced.

3.2 Analytical framework

To comprehensively measure the impact of COVID-19 on users' online social behaviour and public sentiment, changes in the emoji usage of online users during the COVID-19 outbreak were compared from multiple perspectives. First, by determining the most frequently used emojis daily, we selected two sets of popular emojis on Weibo before and after the COVID-19 outbreak and compared the differences. By utilizing the official annotations of emoji characters, which provide their textual descriptions by "translating" emojis back into words, we manually classified all popular emoji characters into five groups and examined the daily use frequency of different emoji groups. In analysing the correlation of emoji usage with COVID-19 infections, the Pearson correlation coefficient (r) was calculated. In addition, we compared the

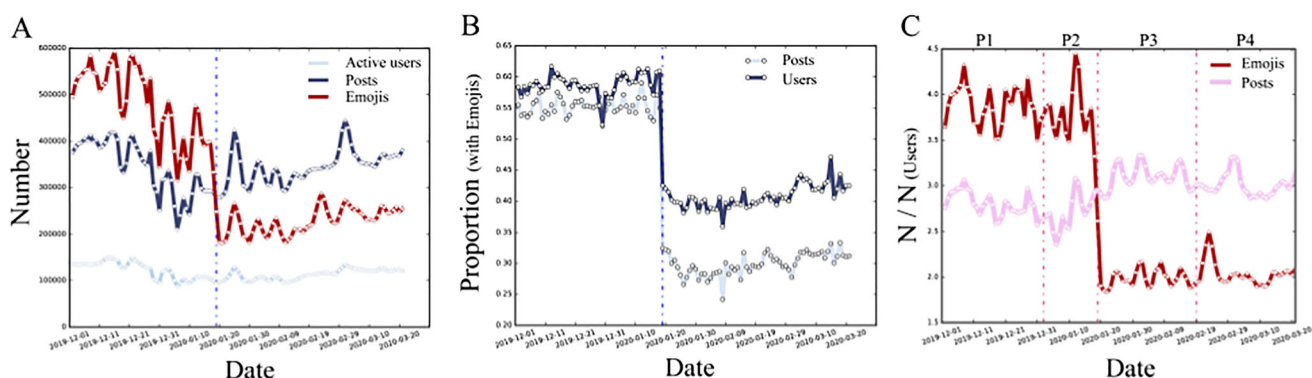


Fig. 1 (A) Number of active users, posts, and emojis; (B) The proportion of Weibo posts with embedded emojis and the proportion of users who used emojis; (C) Average number of posts and emojis per user

temporal patterns of emoji occurrence over 111 days to explore the impact of COVID-19 on the diurnal moods of Wuhan residents. By considering emojis as nodes and the co-occurrence relationship between any two emojis as edges, we also analysed the emoji co-occurrence structure from the perspective of complex networks. Finally, we examined the difference in emoji usage between different topics, user levels, and gender groups to explore other factors that may affect emoji usage. In this study, all text segmentation work was carried out with Jieba [48, 49], a commonly used Chinese word segmentation package. Topics discussed by online users were examined by employing the Latent Dirichlet Allocation (LDA) model [50, 51], a three-level hierarchical Bayesian model, which aims to understand and extract hidden topics from large volumes of text documents [52]. The overall analytical framework is illustrated in Fig. 2.

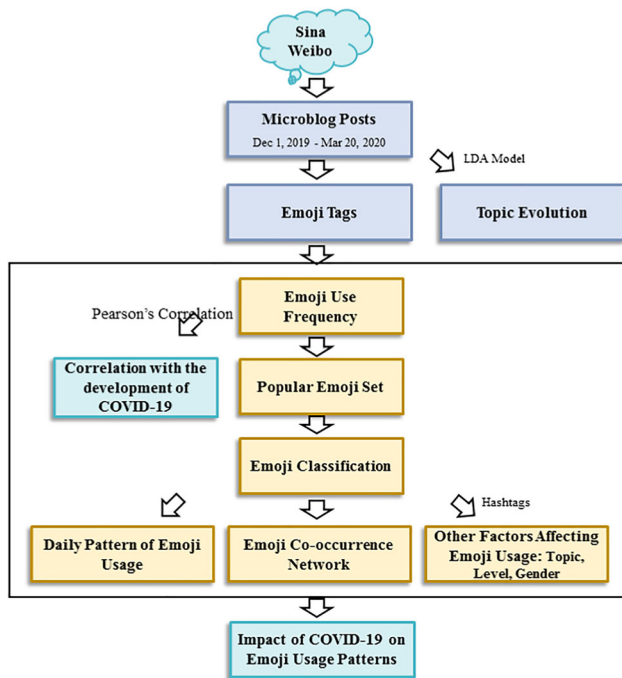


Fig. 2 The workflow of analysis in this study

4 Results

4.1 Basic statistical analysis

Popular emojis Similar to word frequency, which usually conforms to Zipf’s law or Pareto’s law [53, 54], there is also a large heterogeneity in the popularity of different emojis; i.e., the top 100 popular emojis account for approximately 96% of all emojis used daily, and the top 10 popular emojis account for approximately 50.4% (Fig. S3). Among the 1,840 unique emojis, we determined the top 100 most frequently used emojis anytime during the collection period as *popular emojis*, which yielded 154 unique emojis overall. In determining the daily top 10 popular emojis on Weibo before and after the COVID-19 outbreak (Fig. 3), we found that after the pandemic, most of the newly added popular emojis were related to negative emotions, such as sadness, anger, and disappointment. A few positive emojis, such as *Cheer for Wuhan* (🇨🇳) and *Peace Lantern* (🏮), were also posted frequently, expressing support for anti-epidemic work and encouragement and cheering for medical staff. In addition, due to the shortage of living materials in quarantined communities, the number of food-related emojis increased significantly, such as *Chicken Leg*.

Emoji classification According to the emotion expressed by the emoji, we classified the 154 popular emojis into 5 categories, i.e., Happy, Encouraging, Questioned, Sad, and Angry (Table S1). By analysing the evolution in the usage of different emoji groups over time, we found that after January 18, 2020, the number of emojis in the category Happy decreased rapidly, while the number of emojis relevant to Sad increased. Furthermore, the daily proportion of Happy emojis declined significantly and remained low (approximately 0.32) for nearly 29 days after the COVID-19 outbreak, while the daily proportion of emojis in the Sad and Encouraging categories rose to 0.24 and 0.34, respectively (Fig. 4). On January 31, 2020, Li Wenliang [55], an ophthalmologist at Wuhan Central Hospital, posted his experience of being summoned by

Fig. 3 Comparison of the daily top 10 popular emoji sets before and after the COVID-19 outbreak

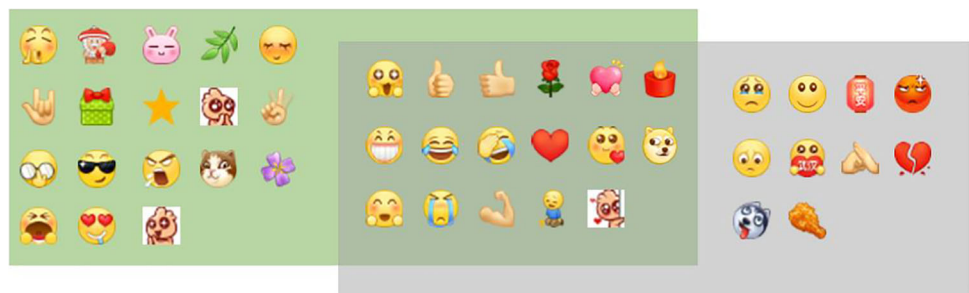
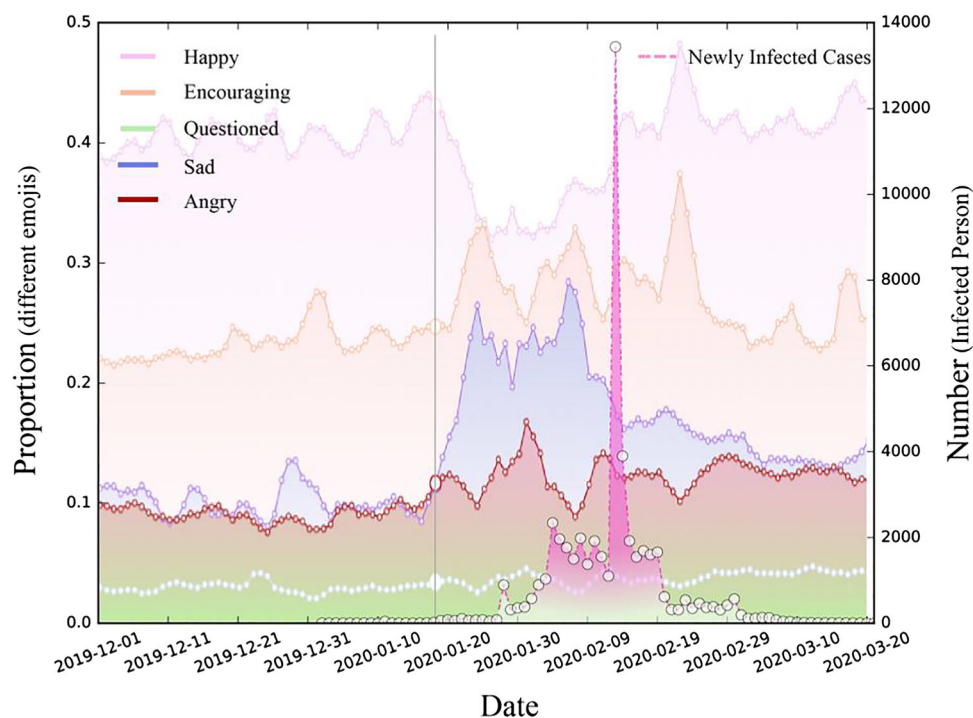


Fig. 4 The use probability of emojis in different emotion categories and the newly infected cases of COVID-19 in Wuhan



the police for "making false comments on the internet about unconfirmed SARS outbreaks" on Weibo. At this time, the use of anger-related emojis reached its peak. These results indicate that the COVID-19 epidemic significantly impacted individuals' moods, and accordingly, people tended to express more sadness and encouragement after the COVID-19 outbreak.

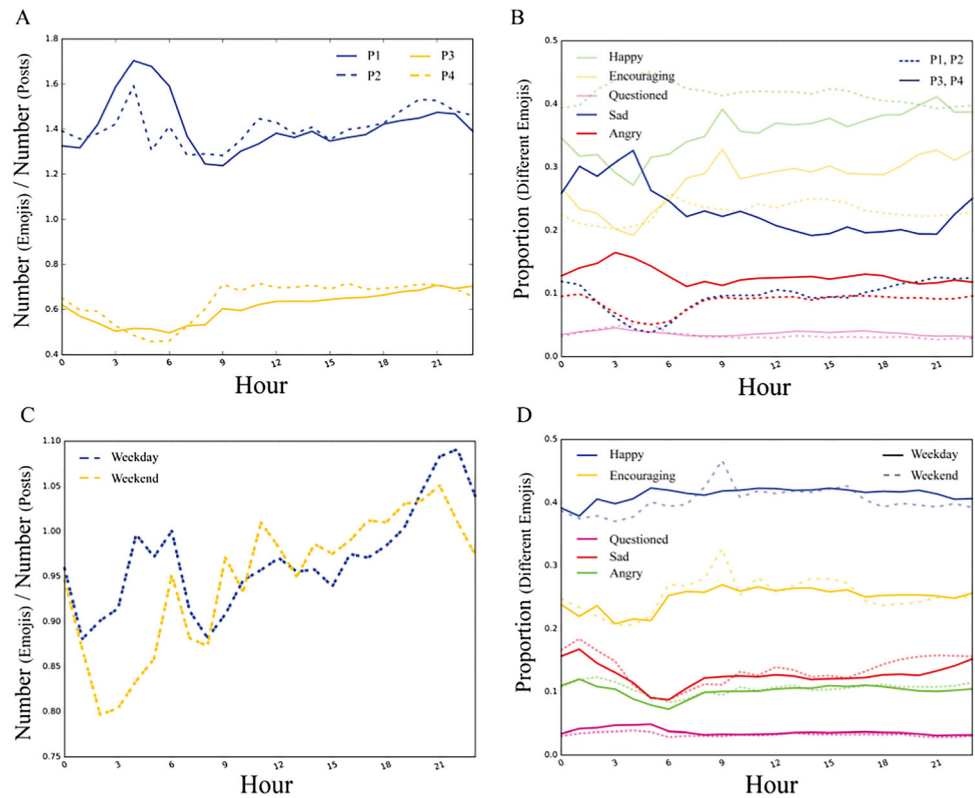
Correlation of emoji usage with COVID-19 We found that the average number of emojis used in each post during the period was negatively correlated with the proportion of pneumonia-related posts ($r = -0.698, p < 0.001$); i.e., the more pneumonia-related posts were published, the fewer emojis were utilized. This suggests that Weibo users were likely to post fewer emojis when mentioning COVID-19-related topics. Moreover, the rate of emojis relevant to Happy was negatively correlated with the ratio of pneumonia-related posts ($r = -0.808, p < 0.001$ $r = -0.808, p$). In contrast, the proportion of Sad emojis and the rate of posts mentioning pneumonia showed a significant positive correlation ($r = 0.813, p < 0.001$ $r = 0.813, p$). Surprisingly, the proportion of Encouraging emojis was also positively correlated with the proportion of pneumonia-related posts ($r = 0.604, p < 0.001$ $r = 0.604, p$), indicating that although the COVID-19 outbreak negatively affected individuals' moods, the positive emotions regarding encouragement and cheering also increased, which to some extent demonstrated the confidence and support of Wuhan residents for anti-epidemic work.

4.2 Temporal pattern of emoji usage

Hourly pattern In the temporal analysis of emoji usage, we first divided the study time into 4 periods in accordance with the development of COVID-19. The first period (P1) covers the time from December 01, 2019, to January 01, 2020; the second period (P2) is from January 01, 2020, to January 18, 2020; the third period (P3) is from January 18, 2020, to February 18, 2020; and the fourth period (P4) is from February 18, 2020, to March 20, 2020.

We found that the hourly changes in the number of emojis used during the whole period were consistent with the number of posts; i.e., the number of emojis and posts both peaked at 10:00 pm and were minimal at approximately 5:00 am. However, the hourly emoji use rate (average number of emojis used in each post) after the COVID-19 outbreak (P3 and P4) was clearly different from the pattern before the outbreak (P1 and P2), as shown in Fig. 5A. The peak time of the emoji use rate before the pandemic was at midnight or in the early morning, when the rate during P3 and P4 was at the lowest. In addition, users became more negative at midnight (approximately 3:00 am); i.e., the proportion of emojis related to happiness and encouragement decreased significantly, while emojis in the Sad and Angry categories increased dramatically (Fig. 5B). This proves that COVID-19 negatively impacted individuals' moods. Furthermore, online users were more psychologically vulnerable at midnight,

Fig. 5 (A) Daily pattern of the emoji use frequency in 4 periods (standardized with the number of posts per hour); (B) Daily pattern of the proportion of emojis in different emotion categories (Standardized with the total number of emojis per hour) before and after the COVID-19 outbreak; (C) The hourly number of emojis used in each post on weekdays and weekends; (D) The proportion of different types of emojis used at each hour on weekdays and weekends



often communicating the pessimistic part of their emotions.

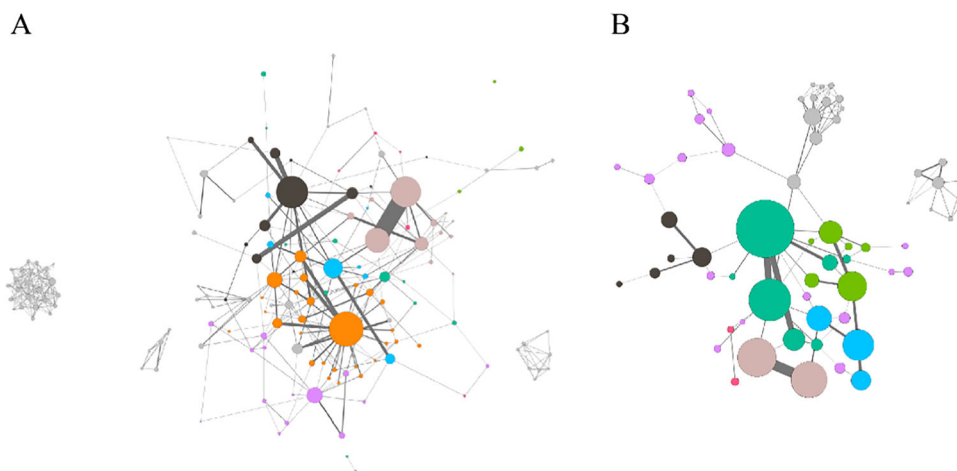
Weekly pattern On weekdays, the number of emojis used in each post was the highest (emoji use rate) during the evening (approximately 10:00 pm) and early morning (4:00–6:00 am). However, the peak period of the emoji use rate on weekends was not obvious, and the emoji use pattern daily showed a more even distribution (Fig. 5C). In addition, the diversity of emojis used on weekends was much higher than that on weekdays, even if there were not more emojis used. On weekends, the proportion of Encouraging and Happy emojis used was significantly higher in the morning (approximately 9:00 am) than at other times (Fig. 5D), indicating that people tended to be more positive and optimistic in the morning on weekends. However, whether on weekdays or weekends, the use of emojis related to negative emotions (Sad and Angry) reached its peak at midnight. This also proves that people were more likely to reveal their psychologically vulnerable and negative side late at night. We found that the difference in emoji usage between weekdays and weekends varied significantly with the development of COVID-19. During P1 and P2, there was a clear peak in the emoji use rate in the early morning (approximately 4:00 am) on working days, while after the COVID-19 outbreak, the daily trends on weekdays and weekends both

remained relatively stable throughout the day (Fig. S6). This could be attributed to the stay-home quarantine for controlling the spread of COVID-19; afterwards, residents had a more even time distribution.

4.3 Emoji co-occurrence network

Network construction To understand the collocation pattern among emojis, we constructed a co-occurrence network with nodes standing for the unique emojis and the weight of edges being the number of times the two emojis have been used together. The network contains 1,711 nodes and 31,878 edges, with an average degree of 37.3; i.e., each emoji is used, on average, with approximately 37 other emojis together in posts. ❤️ is the most connected node, with a degree of 591, while there are more than 45% nodes with a degree less than 10 and approximately 87% nodes with a degree less than 100. We found that the average weighted degree of the network is 298.4, indicating that, on average, each emoji appears with other emojis approximately 298 times. The node with the highest weighted degree (12,970) is 🎄, a consequence of the great number of Christmas greetings during the data collection period, and the emoji pair (edge) [❤️🤔] has the highest co-occurrence frequency in posts (weight of edge, 66,653); however, the weight of approximately 91% of the

Fig. 6 (A) Net1 (node's degree > 1, edge's weight > 100); (B) Net2 (node's degree > 1, edge's weight > 100). Nodes represent emojis, the size of the node represents the node's weighted degree, and the colour of the node distinguishes different communities. The thickness of the edge represents the edge's weight, i.e., the co-occurrence frequency between two nodes



edges in the network is less than 100, and the weight of more than 74% edges is less than 10. In addition, we found that there was a clear community structure in the co-occurrence network.

Evolution of the emoji co-occurrence network We extracted the emojis in all posts and constructed the emoji co-occurrence networks before and after the COVID-19 outbreak (Net1 and Net2, for short), as shown in Fig. 6A and B. We found that the number of nodes and edges in the network

greatly decreased after the COVID-19 outbreak. This is consistent with the above findings, indicating that due to the epidemic, users reduced the use frequency of emojis. However, the average degree of this network increased from 34 to 39.8, which means that although the number of emojis used decreased, the diversity of emoji usage increased. After the COVID-19 outbreak, the average clustering coefficient of the emoji co-occurrence network also increased from 0.23 to 0.32, indicating that more triangular structures appeared in the emoji connections. The top 3 nodes with the largest weighted

Table 1 Comparison between Net1 and Net2

	Before the COVID-19 outbreak	After the COVID-19 outbreak
Network measurements	<p>Number of nodes: 1,659</p> <p>Number of edges: 28,212</p> <p>Average degree: 34.01</p> <p>Average weighted degree: 298.95</p> <p>Network diameter: 7</p> <p>Modularity: 0.48</p> <p>Clustering coefficient: 0.23</p> <p>Average shortest path length: 2.53</p>	<p>Number of nodes: 867</p> <p>Number of edges: 17,263</p> <p>Average degree: 39.82</p> <p>Average weighted degree: 273.68</p> <p>Network diameter: 8</p> <p>Modularity: 0.49</p> <p>Clustering coefficient: 0.32</p> <p>Average shortest path length: 2.72</p>
Top 5 emoji pairs with the highest frequency	<p>❤️😂 (52072)</p> <p>❤️😱 (44672)</p> <p>❤️💕 (40910)</p> <p>😂😂 (38859)</p> <p>❤️😱 (37422)</p>	<p>😂😂 (22228)</p> <p>❤️🌹 (20277)</p> <p>❤️😱 (20221)</p> <p>😂😂 (19737)</p> <p>❤️💕 (18458)</p>
Node with the highest degree	❤️ (578)	❤️ (200)
Top 5 nodes with the highest weighted degree	<p>😂 (15849)</p> <p>🌿 (14545)</p> <p>🍷 (13700)</p> <p>😱 (11175)</p> <p>😱 (8754)</p>	<p>😂 (9909)</p> <p>😱 (7103)</p> <p>😂 (6529)</p> <p>😂 (6139)</p> <p>🍷 (4913)</p>

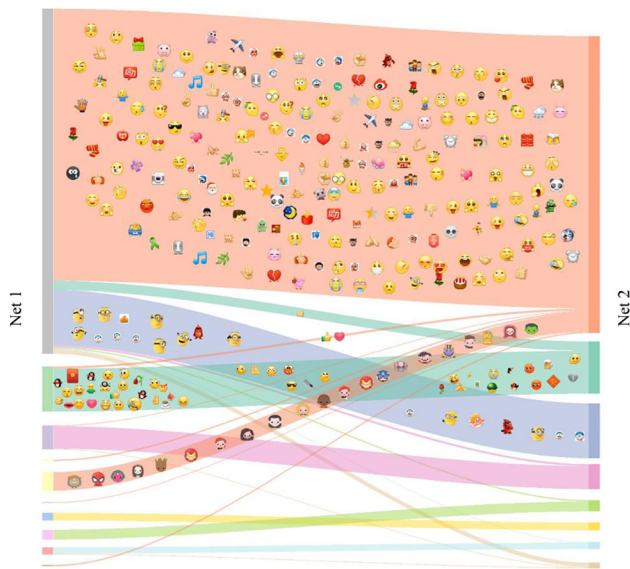


Fig. 7 Member flow in the TOP-10 communities of Net1 and Net2

degree in Net2 are 🙄🙄🙄 (Table 1), related to emotions about grievance, sadness and doubt, which are significantly different from the most connected nodes in Net1. The emoji pairs with the highest co-occurrence frequency after the pandemic also changed considerably, and the use frequency of emojis relevant to sadness (e.g., 😞😞) increased. By comparing the community structure of Net1 and Net2, we also described the flow of members (nodes) in the TOP-10 communities before and after the pandemic (Fig. 7).

Maximum co-occurrence frequency network To discover the typical emoji collocation on Weibo, we built the maximum co-occurrence frequency network by retaining the node pair (edge) with the largest co-occurrence probability. The weight of the edge represents the co-occurrence probability between nodes (from 0 to 1). There were mainly 4 hub nodes (❤️😞😞😞) in the network (Fig. 8A) during P1 and P2. After the COVID-19 outbreak, *Flower* (🌺) replaced *Longing* (🥺) as a new hub node, and more discrete emoji pairs, as well as

small subclusters, appeared, e.g., the emoji pairs connected by *Angry* (😡) and a series of emojis about gestures (Fig. 8B). Emojis and emoji pairs related to sadness, gratitude, and encouragement were used more frequently. In addition, the cluster connected by *Love* (❤️) decreased, but the size of the community linked by *Facepalm* (🤦) grew significantly, and there were more emojis relevant to negative emotions, such as *Dizzy* (🤡), *Scared* (😱), and *Wearing Mask* (😷). These results indicate that users’ habits regarding emoji usage indeed changed after the COVID-19 outbreak. Due to the emergence of newly popular emojis and the evolution of the moods users wanted to express, the fixed emoji co-occurrence patterns were significantly altered.

4.4 Topic, user level and gender

Correlation between emoji usage and post topics To examine whether the topic affected emoji usage in the post, we extracted all posts containing hashtags and analysed the correlation between different topics and emoji use patterns. All hashtags (##) on Weibo from December 1, 2019, to March 20, 2020, with over 500 users and 500 microblog posts, were manually divided into 4 groups according to the topic, i.e., stars, daily life, games or e-sports, and social events (including COVID-19). These 4 topics covered a total of 3,830,100 posts. It was observed that before the COVID-19 outbreak, most topics discussed online were involved with stars and daily life. After the COVID-19 outbreak, discussions on these topics diminished dramatically, while the number of posts related to public social events rapidly increased (Fig. S10).

We then compared the evolution of the proportion of posts containing emojis on these four topics (Fig. 9A). Users posted emojis most frequently when discussing topics related to stars and entertainment, while the emoji use in posts related to social events was the lowest, suggesting that users tend to be more cautious and objective when discussing public social events. After the COVID-19 outbreak, the proportion of posts containing emojis in topics on social events moderately

Fig. 8 The maximum co-occurrence frequency network of the top 100 most frequent emojis (A) before and (B) after the COVID-19 outbreak. The size of the node represents the node’s degree, and the thickness of the edge represents the probability of co-occurrence

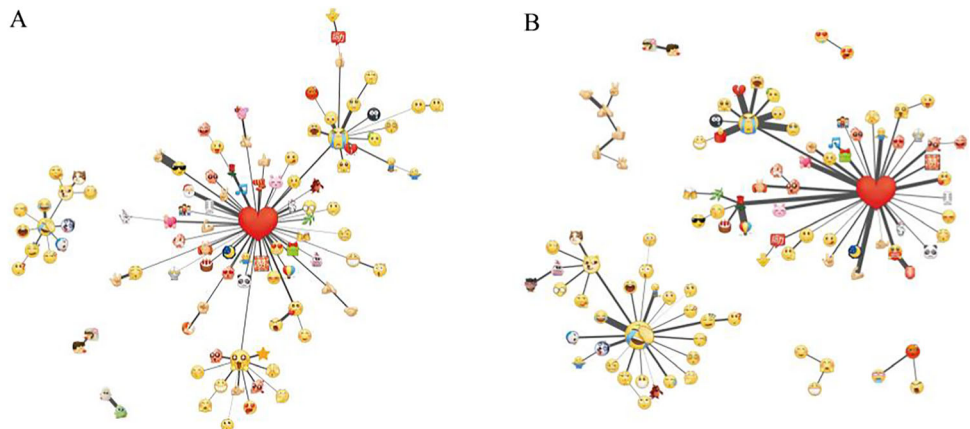
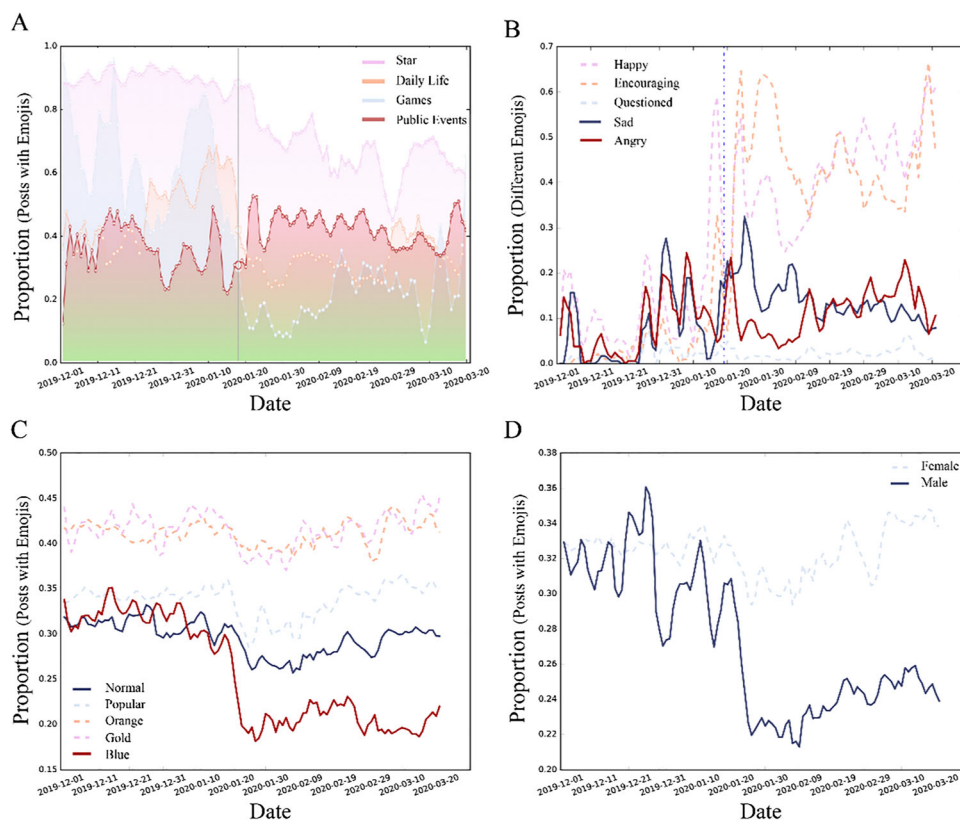


Fig. 9 (A) The proportion of posts containing emojis in different topics; (B) The proportion of different emojis used in posts related to social events; (C) The percentage of posts containing emojis posted by users at different levels; (D) The percentage of posts containing emojis posted by users of different genders



increased, while the proportion regarding stars, daily life, and games dropped greatly. In addition, we found that most emojis used in these topics were related to Happy and Encouraging, while more Sad emojis were posted in posts related to social events. Using the total number of emojis used daily to standardize the number of emojis in different categories, we found that after the COVID-19 outbreak, users' emotions showed greater fluctuations than before when discussing public social events, and they expressed more incentives, sadness, and anger (Fig. 9B).

Differences between individual users and official users There are a total of five user levels on Weibo, i.e., *Ordinary User*, *Popular User*, *Orange V*, *Gold V*, and *Blue V*. The first four levels all represent user accounts, while *Orange V* as well as *Gold V* are personal authentication accounts, standing for public figures with a certain social influence (e.g., actors). *Blue V* stands for officially certified enterprise or government accounts. We found that the proportion of posts containing emojis published by official users dropped dramatically after the COVID-19 outbreak, while the proportion of posts published by personal users changed much more moderately (Fig. 9C). This indicates that due to the public health emergency, official users decreased their emoji use frequency, and the texts posted by the government and enterprises tended to be more formal and serious than usual.

Emoji usage by user gender Comparing the evolution of emoji usage between male and female users, we found that the average number of emojis posted by male users decreased significantly after the COVID-19 outbreak, while that posted by female users slightly increased (Fig. S12). Comparing the proportion of posts containing emojis between these two groups, it is also obvious that after the epidemic, the emoji use frequency of male users when posting online dramatically decreased (Fig. 9D). Male users were more inclined to utilize fewer emotional symbols and more formal language, which to some extent suggests that when facing public risks, men tend to conceal or control their emotional display, while women are more emotionally expressive [56–58].

5 Conclusion & discussion

As the COVID-19 pandemic sweeps across the world, it is causing widespread concern, fear, and stress. This study aims to explore the impacts of this public health emergency on individuals' social behaviour and sentiment by comparing the emoji usage of online users on Weibo before and after the COVID-19 outbreak.

We found that the number of emojis used decreased dramatically after the COVID-19 outbreak, and more emojis related to sadness and encouragement were posted. However, the use probability of Happy emojis significantly decreased. In

addition, the proportion of Happy emojis showed a negative correlation with the rate of COVID-19-related posts ($r = -0.808$), while the proportion of Sad emojis was positively correlated with the proportion of COVID-19-related posts ($r = 0.813$). Surprisingly, the rate of Encouraging emojis was also positively correlated with the rate of posts related to COVID ($r = 0.604$). Although the COVID-19 outbreak negatively affected individual sentiments, some positive emotions also increased, implying the confidence and support of Wuhan residents for the anti-epidemic work. On the other hand, we found that at midnight, users tend to be more psychologically vulnerable and more willing to communicate the negative side of their moods. In addition, after the pandemic, the government and enterprises were inclined to post fewer emojis, preferring regular textual language. Similarly, male users used fewer emotional symbols, indicating that when facing public risks, compared with women, men tend to conceal or control their emotional display.

This study reveals the pattern of emoji use on social networks after a public emergency and contributes to the knowledge gap of online behavioural changes and the sentiment evolution of social platform users due to the COVID-19 outbreak. These results can provide effective references for policymakers to plan reasonable interventions and psychological counselling for populations whose sentiments are strongly affected by COVID-19. Although there may be slight deviations in the analysis due to the inability to obtain deleted posts, this has no significant impact on the overall conclusions. As different use habits of emojis may lead to different sentiment semantics [59, 60], in future work, we tend to further explore the emotional meanings conveyed by emojis by introducing detailed emoji use habits of online users, and analyse the usage patterns of more massive emojis across platforms and contexts.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10489-022-03195-y>.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10489-022-03195-y>.

Acknowledgements The authors would like to thank Ziqiang Cao, Suoyi Tan and Jayson Jia for helpful discussions on data analysis and visualization.

Authors' contributions Study Design, X.L.; Methodology and experiment, C.L.; Data collection, J.L.; Funding acquisition, X.L., W.Z. & X.T.; Writing—original draft, C.L.; Writing—review and editing, X.L., X.T., W.Z., J.L., & T.Z.

Funding This study is supported by the National Natural Science Foundation of China [72025405, 91846301, 82041020, 71771213, 72171150, 61773248, 11975071], the Hunan Science and Technology Plan Project [2020JJ4673, 2020TP1013], the National Social Science Foundation of China [20ZDA060], and the Shenzhen Basic Research Project for Development of Science and Technology [202008291726500001, JCYJ20200109141218676].

Availability of data and material All data analysed in this study are publicly available, and all posts in the datasets can be collected on the website of Sina Weibo (<https://weibo.com/>).

Declarations

Ethics approval and consent to participate This study uses only publicly available data and does not involve any physical, social or legal risks to users, whose personal identifying information is either unknown to the public or removed if they choose to post online.

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

References

- Novak P K, Smailović J, Sluban B, et al (2015) Sentiment of emojis. *PLoS one* 10(12)
- Shiha M, Ayvaz S (2017) The effects of emoji in sentiment analysis. *Int J Comput Electr Eng (IJCEE)* 9(1):360–369
- Al RF (2018) Functions of emojis in WhatsApp interaction among Omanis. *Discourse, context & media* 26:117–126
- Hogenboom A, Bal D, Frasinca F, Bal M, de Jong F, Kaymak U (2015) Exploiting emoticons in polarity classification of text. *Journal of Web Engineering* 14(1–2):22–40
- Wall HJ, Kaye LK, Malone SA (2016) An exploration of psychological factors on emoticon usage and implications for judgement accuracy. *Comput Hum Behav* 62:70–78
- Alshenqeeti H (2016) Are emojis creating a new or old visual language for new generations? A socio-semiotic study. *Advances in Language and Literary Studies* 7(6)
- Djenouri Y, Comuzzi M (2017) Combining Apriori heuristic and bio-inspired algorithms for solving the frequent itemsets mining problem. *Inf Sci* 420:1–15
- Chen X, Siu KWM (2017) Exploring user behaviour of emoticon use among Chinese youth. *Behaviour & Information Technology* 36(6):637–649
- Dodds PS, Harris KD, Kloumann IM, et al (2011) Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS one* 6(12): e26752
- Niehus R, De Salazar PM, Taylor AR, et al (2020) Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. *The Lancet Infectious Diseases*
- Wang C, Horby PW, Hayden FG et al (2020) A novel coronavirus outbreak of global health concern. *The Lancet* 395(10223):470–473
- Phelan AL, Katz R, Gostin LO (2020) The novel coronavirus originating in Wuhan, China: challenges for global health governance. *JAMA* 323(8):709–710
- WHO. Novel coronavirus (2019-nCoV), <https://covid19.who.int/>, last accessed 2021/10/22.
- Li S, Wang Y, Xue J et al (2020) The impact of COVID-19 epidemic declaration on psychological consequences: a study on active Weibo users. *Int J Environ Res Public Health* 17(6):2032
- John ATJ, Natalie JS, Michael AM (2013) The behavioral immune system and social conservatism: A meta-analysis. *Evol Hum Behav* 34:99–108
- Mortensen CR, Becker DV, Ackerman JM, Neuberg SL, Kenrick DT (2010) Infection breeds reticence: The effects of disease salience on self-perceptions of personality and behavioral avoidance tendencies. *Psychol Sci* 21:440–447

17. Lu X, Brelsford C (2014) Network structure and community evolution on twitter: human behavior change in response to the 2011 Japanese earthquake and tsunami. *Sci Rep* 4(1):1–11
18. Nadal KL (2018) Microaggressions and traumatic stress: Theory, research, and clinical treatment. American Psychological Association
19. Khosravi M (2020) Perceived risk of COVID-19 pandemic: The role of public worry and trust. *Electron J Gen Med* 17(4):em203. 2020.
20. Ziarko M, Jasielska A, Stanisławska-Kubiak M, et al (2021) Mental Health Outcomes Associated with COVID-19 Pandemic in a Group of Health Care Professionals. *The Journal of Behavioral Health Services & Research* 1–10
21. Zhou P, Yang XL, Wang XG et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798):270–273
22. Li Q, Guan X, Wu P, et al (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*
23. Zhou F, Yu T, Du R, et al (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*
24. Huang C, Wang Y, Li X et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan. *China The Lancet* 395(10223):497–506
25. Jia J S, Lu X, Yuan Y, et al (2020) Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 1–5
26. Guo S, Fang F, Zhou T et al (2021) Improving Google flu trends for COVID-19 estimates using Weibo posts. *Data Science and Management* 3:13–21
27. Nemes L, Kiss A (2021) Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication* 5(1): 1–15
28. Manguri KH, Ramadhan RN, Amin PRM (2020) Twitter sentiment analysis on worldwide COVID-19 outbreaks. *Kurdistan Journal of Applied Research* 54–65
29. Wang T, Lu K, Chow K P, et al. COVID-19 sensing: negative sentiment analysis on social media in China via BERT model. *Ieee Access*, 2020, 8: 138162–138169.1
30. Tan H, Peng S L, Zhu C P, et al (2021) Long-term Effects of the COVID-19 Pandemic on Public Sentiments in Mainland China: Sentiment Analysis of Social Media Posts. *Journal of Medical Internet Research* 23(8): e29150
31. Kiritchenko S, Zhu X, Mohammad SM (2014) Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research* 50:723–762
32. Wang Z, Yu Z, Guo B et al (2015) Sentiment Analysis of Chinese Micro Blog Based on Lexicon and Rule Set. *Comput Eng Appl* 51(8):218–225
33. Dun X, Zhang Y, Yang K (2017) Fine-grained sentiment analysis based on Weibo. *Data Analysis and Knowledge Discovery* 1(7): 61–72
34. Xiao J, Ding X, He R (2015) Analysis of Chinese Micro-blog Emotion Which Based on Field of Emotional Dictionary. *Electronic Design Engineering* 23(12):18–21
35. Thompson D, Mackenzie IG, Leuthold H et al (2016) Emotional responses to irony and emoticons in written language: evidence from EDA and facial EMG. *Psychophysiology* 53(7):1054–1062
36. Aldunate N, González-Ibáñez R (2017) An integrated review of emoticons in computer-mediated communication. *Front Psychol* 7:2061
37. Kaye LK, Malone SA, Wall HJ (2017) Emojis: Insights, affordances, and possibilities for psychological science. *Trends Cogn Sci* 21(2):66–68
38. Tang Y, Hew KF (2019) Emoticon, emoji, and sticker use in computer-mediated communication: A review of theories and research findings. *Int J Commun* 13:27
39. Bai Q, Dan Q, Mu Z et al (2019) A systematic review of emoji: Current research and future perspectives. *Front Psychol* 10:2221
40. Cappallo S, Svetlichnaya S, Garrigues P et al (2018) New modality: Emoji challenges in prediction, anticipation, and retrieval. *IEEE Trans Multimedia* 21(2):402–415
41. Li M, Ch'ng E, Chong AYL, et al (2018) Multi-class Twitter sentiment classification with emojis. *Industrial Management & Data Systems*
42. Kaye LK, Wall HJ, Malone SA (2016) “Turn that frown upside-down”: A contextual account of emoticon usage on different virtual platforms. *Comput Hum Behav* 60:463–467
43. Settanni M, Marengo D (2015) Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. *Front Psychol* 6:1045
44. Barbieri F, Espinosa-Anke L, Saggion H (2016) Revealing patterns of Twitter emoji usage in Barcelona and Madrid. *Artif Intell Res Dev* 288:239–244
45. Cheng L (2017) Do I mean what I say and say what I mean? A cross cultural approach to the use of emoticons & emojis in CMC messages. *Fonseca* 15:199
46. Al-Rawi A, Siddiqi M, Morgan R, et al (2020) COVID-19 and the gendered use of emojis on Twitter: infodemiology study. *Journal of medical Internet research* 22(11):e21646
47. Zappavigna M, Logi L (2021) Emoji in social media discourse about working from home. *Discourse, Context & Media* 44:100543
48. Wu X, Du Z, Guo Y et al (2019) Hierarchical attention based long short-term memory for Chinese lyric generation. *Appl Intell* 49(1): 44–52
49. Project description of jieba [pypi], <https://pypi.org/project/jieba/>, last accessed 2021/02/20.
50. Na L, Ming-xia L, Hai-yang Q, et al (2021) A hybrid user-based collaborative filtering algorithm with topic model. *Applied Intelligence* 1–14
51. Wan C, Peng Y, Xiao K et al (2020) An association-constrained LDA model for joint extraction of product aspects and opinions. *Inf Sci* 519:243–259
52. Wang W, Guo B, Shen Y et al (2020) Twin labeled LDA: a supervised topic model for document classification. *Appl Intell* 50(12): 4602–4615
53. Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. *Contemp Phys* 46(5):323–351
54. Aoki S, Nirei M (2017) Zipf's law, Pareto's law, and the evolution of top incomes in the United States. *Am Econ J Macroecon* 9(3):36–71
55. Introduction of Li Wenliang, https://en.wikipedia.org/wiki/Li_Wenliang, last accessed 2021/03/12.
56. Prada M, Rodrigues DL, Garrido MV et al (2018) Motives, frequency and attitudes toward emoji and emoticon use. *Telematics Inform* 35(7):1925–1934
57. Rodrigues D, Prada M, Gaspar R et al (2018) Lisbon Emoji and Emoticon Database (LEED): Norms for emoji and emoticons in seven evaluative dimensions. *Behav Res Methods* 50(1):392–405
58. Butterworth SE, Giuliano TA, White J et al (2019) Sender gender influences emoji interpretation in text messages. *Front Psychol* 10:784
59. Liu C, Fang F, Lin X et al (2021) Improving sentiment analysis accuracy with emoji embedding. *Journal of Safety Science and Resilience* 2(4):246–252
60. Fang F, Wang T, Tan S, et al. Network Structure and Community Evolution Online: Behavioral and Emotional Changes in Response to COVID-19. *Frontiers in Public Health*, 2274.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.