Friends and foes: Sinophobia was viral in Chinese language communities on Twitter during the early COVID-19 pandemic

Yongjun Zhang[1*], Hao Lin[1], Yi Wang[2], Xinguang Fan[3]

1 Department of Sociology and Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY, USA
2 Department of Asian and Asian American Studies, Stony Brook University, Stony Brook, NY, USA
3 Department of Sociology, Peking University, Beijing, China

* Corresponding Author: yongjun.zhang@stonybrook.edu

Abstract

COVID-19 has engendered a global surge of Sinophobia. Previous research using social media datasets shows that the pandemic triggered waves of negative attitudes toward China and Asian communities online, but limited research has examined how Chinese language users respond to COVID-19 on western social media. We address this gap by compiling a unique database (CNTweets) with over 25 million Chinese tweets mentioning any Chinese characters related to China, Chinese Communist Party (CCP), Chinese, and Asians from December 2019 to April 2021. Our analysis on Twitter users' self-reported geographic information shows that most Chinese language users on Twitter originated from Mainland China, Hong Kong, Taiwan, or United States. We then adopt the Robustly Optimized Bidirectional Encoder Representations from Transformers (RoBERTa) and structural topic modeling to further analyze the sentiments, contents, and topics of Chinese tweets during the COVID-19 pandemic. Our results suggest that the majority of tweets were negative toward China, and most of these tweets were only contributed by one percent of Twitter users. Despite the prevalence of anti-China sentiments, the target entity analysis shows that these negative sentiments were more likely to target the Chinese government and the communist party instead of Chinese people. Our findings also show that the most popular topics and themes are related to politics (e.g., Hong Kong protests and Taiwan issues), COVID-19, and United States (e.g., US-China relations and domestic issues). Anti-China users focus relatively more on political issues, while pro-China users mention more about cultural and economic topics. Our social network analysis reveals that these pro-China and anti-China Twitter users lack direct in-depth engagement. We conclude by discussing our contributions to China and social media studies as well as possible policy implications.

Introduction

COVID-19 has engendered xenophobia and racism against Asian communities (Lee & Huang, 2021; D. Zhang, 2021). The spread of misinformation and disinformation on the origin of COVID-19 has led to a global surge of Sinophobia (Cook, Huang, & Xie, 2021). Recent scholarship has begun to assess public sentiments toward China and the Chinese government on social media platforms. For instance, using English-language tweets, Cook and colleagues find that COVID-19 has caused a sharp spike in anti-China attitudes in the United States (Cook et al., 2021). But less attention has been directed to examine how Chinese language social media users respond to COVID-19 online. Lu and colleagues show that Chinese Sina Weibo users were more likely to be supportive rather than critical due to the effective COVID-19 responses by the Chinese government in the early pandemic (Lu, Pan, & Xu, 2021). Unlike Western social media platforms, Sina Weibo has censored various sensitive contents such as collective action potentials and political dissent (King, Pan, & Roberts, 2013, 2014). It remains unclear how Chinese language users discussed COVID-19 and their sentiments toward China on Western social media platforms like Twitter.

To fill the lacuna, this article examines how Chinese language users on Twitter engage in China-related discussions and the associated sentiments during the early COVID-19 pandemic. Unlike English Twitter users and Sino Weibo users, Chinese Twitter users are a distinctive group of accounts used by overseas Chinese, residents from Hong Kong, Taiwan, and Singapore, Mainland Chinese with VPN access, as well as other organizations and bots criticizing or supporting China.

In the Twitter-verse, who were those Chinese language users tweeting China-related issues during the pandemic? After the COVID-19 outbreak, how did Chinese language users on Twitter discuss the pandemic and China? What were the main public sentiments toward China? Were they targeting Chinese people or Chinese government? Were they part of the computational propaganda? Did those pro-China and anti-China users engage in each other's debate? To address these questions, we query the Twitter historical database using keywords related to China, Chinese, Chinese Community Party (CCP), and Asians in both simplified and traditional Chinese languages to generate our Chinese Tweets (CNTweets) analytic dataset with 25.30 million Tweets by 1.32 million Twitter users between December 2019 and December 2020. We then annotate a training dataset with 10,000 tweets to build a series of deep learning algorithms to classify the anti-China sentiment and topics in these tweets by fine-tuning pre-trained Chinese Robustly Optimized Bidirectional Encoder Representations from Transformers with the Whole Word Masking models (Chinese-RoBERTa-wwm-ext) (Devlin, Chang, Lee, & Toutanova, 2018; Liu et al., 2019; Cui, Che, Liu, Qin, & Yang, 2021).

Source of Chinese Tweets

Twitter has been blocked by the Chinese government since 2009 due to information control, so regular Mainland Chinese users have to rely on VPN services (Sullivan, 2012). As a result, Mainland Chinese users on Twitter could be a very selective group of individuals, such as lawyers, journalists, and human rights activists, seeking uncensored information and discussing sensitive topics that are not permitted in China (Song, Faris, & Kelly, 2015). These anti-Chinese state users are not the only Mainland users who circumvent the Great Fire Wall. Previous research also shows the prevalence of pro-Chinese state users, for instance, state-sponsored institutional accounts with free access to Twitter and regular pro-China internet users. China has initiated its own foreign propaganda program carried out mainly by state-run companies including China Central Television, China Daily, Global Times, Xinhua News, etc. Individual pro-state users could be paid 50-Cent party, government employees, and other regular nationalistic internet users (Bolsover & Howard, 2019; King, Pan, & Roberts, 2017). In addition to Mainland Chinese, Chinese language users on Twitter could stem from other countries and regions with a population of Chinese language speakers, overseas Chinese, or immigrants of Chinese descendant such as Hong Kong, Macao, Taiwan, Singapore, Thailand, U.S., Australia, and Canada. Twitter has been a battlefield for anti-Chinese state groups with few resources who are using Twitter to spread misinformation and disinformation on China and Chinese politics (Bolsover & Howard, 2019). The diversity of Chinese language users on Twitter motivates our first research question related to sources of Chinese tweets.

RQ1: Who are those Chinese Twitter users mentioning China-related issues during the early pandemic?

Sentiment of Chinese Tweets

A large body of literature has used Twitter to gauge public sentiments and the associated impacts on political, economic, and social outcomes, such as election (Tumasjan, Sprenger, Sandner, & Welpe, 2010; Bovet & Makse, 2019; Shmargad, 2018), stock market (Ranco, Aleksovski, Caldarelli, Grčar, & Mozetič, 2015), and public policies (Flores, 2017). Like other social media platforms (e.g., Weibo, Facebook), public sentiment is a mix of regular internet users, opinion leaders, organizations, and social bots, and it is part of the algorithmically infused societies co-shaped by algorithmic and human behaviour (Wagner et al., 2021).

Prior studies show that both pro- and anti-Chinese state groups have used Twitter as a platform to serve their propaganda purposes (Bolsover & Howard, 2019), but these studies tend to focus on non-Chinese audiences and limited research has examined how these groups target Chinese language users on social media platforms. For instance, Bolsover and colleagues find no evidence of pro-Chinese-state computational propaganda on Twitter, but strong evidence of massive tweets associated with anti-Chinese-state perspectives published in simplified Mandarin (Bolsover & Howard, 2019). This is partly because China's foreign propaganda has been carried out by these traditional state-run media groups such as China Central Television and Global Times with massive human and monetary resources. However, these anti-Chinese state groups have used computational propaganda to promote and disseminate their messages targeting the Chinese government due to its lower operating costs. Thus, we might observe a lot of anti-Chinese state behavior on Twitter.

For pro-Chinese state groups, prior studies have shown the rise of Chinese digital nationalism (DeLisle, Goldstein, & Yang, 2016; Schneider, 2018). Cyber nationalists, especially young Chinese internet users, have defended China and the Chinese government on Western social media platforms without state blessings, such as Little Pinks (xiaofenhong) and Diba Expedition (diba chuzheng) (Han, 2019; Bi, 2021). These cyber nationalists tend to engage in the conversations with their opposing groups instead of posting comments like social bots. Previous research shows that government employees have played an important role in fabricating pro-Chinese messages online (King et al., 2017) and using clickbait strategy to gain visibility (Lu & Pan, 2021). In addition, in recent years, Beijing has initiated a series of campaigns via soft power messaging and COVID diplomacy to tell China's story well (Huang & Wang, 2019). Thus, the complexity and dynamics of pro- and anti-Chinese state groups lead us to the second set of research questions.

RQ2: What is overall pattern of public sentiments during the early pandemic?

RQ3: Who are the main targets of positive and negative sentiments?

RQ4: Is there any dialogue between pro-China and anti-China Twitter users?

Content of Chinese Tweets

Twitter has been a public sphere since its founding. After the COVID-19 outbreak, Twitter, like other social media platforms such as Facebook and Weibo, has been one of the major online spaces where individuals seek social support, tracking government announcements, and monitoring the spread of the coronavirus (Lu et al., 2021). We focus on any Chinese tweets mentioning China-related keywords during the pandemic. We would expect that Chinese Twitter users, such as overseas students and Chinese immigrants would use Twitter to share news and seek for help when COVID-19 emerged.

Twitter has also been a strong battlefield related to conspiracy theories, hate speech, misinformation, disinformation, and fake news. COVID-19 has led to a global surge of anti-Chinese sentiment (Cook et al., 2021), and racial slurs targeting Asian and Asian American communities have been widely spread on Twitter such as Chinese Virus and KungFlu (Ziems, He, Soni, & Kumar, 2020). Chinese Americans and overseas students might use Twitter as a platform to voice themselves and combat racism and anti-Asian attacks.

The increasing tension between U.S. and China such as trade wars and human rights issues related to Xinjiang and Tibet, and the Trump administration's tough policy on Chinese scientists might also spark overseas Chinese users to share concerns on the U.S.-Sino relationship, discuss immigration policies, and express angers or fears of uncertainty in the pandemic. Pro-democracy groups might use Twitter to discuss sensitive topics such as Xinjiang re-education camp, Uyghur, Falungong, etc, while pro-Chinese state users including state-sponsored organizations and paid 50 cents party might use Twitter to promote China's soft power and boost China's global image by tweeting Chinese culture, economic development, tourism, and so on.

The 2019-2020 protest cycles in Hong Kong have drawn great attention from Chinese societies. Protesters used Twitter as a platform to diffuse protest information, mobilize resources, and seek solidarity, while pro-Chinese state and anti-HK protesters might also strategically use Twitter for political propaganda by framing protests as conflicts and violence, disrupting social orders and economy, and destabilizing natural security (M. M. Zhang, Wang, & Hu, 2021). Twitter is also an online space that Chinese state-backed media and nationalists promote the reunification between Mainland China and Taiwan (Chang, Lai, Chang, & Lin, 2021). Similarly, Taiwan independence supporters use Twitter to seek and mobilize for support.

Due to the diversity of Chinese Twitter users and the confluence of COVID-19 and other political and social events, this leads to our third set of research questions.

RQ5: What is the content of these Chinese Tweets in the early pandemic?

RQ6: Is there any variation among different Twitter users?

Data and methods

CNTweets Data

We used Chinese keywords to retrieve all matched tweets posted in 2019-2020 from Twitter's historical database using academic Twitter API. Table 1 shows some descriptive statistics of Twitter data. We managed to obtain over 25 million tweets by 1.32 million users mentioning any keywords related to China, Chinese, and CCP. S1 Appendix documents the detailed keywords we used in data collection.

Table 1: Summary of Twitter Data

Data Type	Million
# of tweets	25.30
# of tweets mentioning China (中国)	16.63
# of tweets mentioning Asians or Chinese (亚裔/华裔)	0.28
# of tweets mentioning Chinese Communist Party (共产党)	7.46
# of Twitter users	1.32

Training Data

In order to extract sentiments and topics in CNTweets data, we annotated a training dataset with 10,000 tweets to build algorithms to classify CNTweets. S2 Appendix documents the detailed process of our training data construction, and here we briefly summarized our major steps. We started with those pro- and anti-China Twitter users and their followers or following accounts (e.g., PDChinese, dajiyuan). We scraped all their tweets posted in the past 2 years. We also used pro- and anti-China hashtags and keywords (e.g., against CCP) to extract potential tweets that either support or criticize the Chinese government or China. We then use a stratified sampling strategy to select 7,000 tweets from these potential positive or negative tweets targeting China. To add more potential neutral tweets in our training dataset, we then randomly selected 3,000 tweets from our CNTweets data to construct the final 10,000 tweets for human annotation. We hired both graduate and undergraduate research assistants to manually annotate the sentiment and topics in these tweets. Each tweet had been labelled by at least two annotators, if there is inconsistency, one of our authors then adjudicated the difference.

Sources of Chinese Twitter Users

To tackle the first research question on sources of Chinese tweets, we rely on partial information provided by Twitter users' self-reported locations when they signed up for a Twitter account. To extract the major countries and regions, our location analysis first uses regular expressions to search country/region names and other abbreviations names and then searches major states/provinces/cities for a country or region. For instance, to identify whether a Twitter user is from U.S., we first search United States, U.S., or US, and then incorporate different states, cities, and their abbreviations like New York and NY. In addition, we also asked our annotators to identify whether a tweet is related to personal opinion, organizations, government announcements, and spams. This allows us to identify whether these tweets are from individual or organizational accounts.

Sentiment of Chinese Tweets

To answer the second question about the overall pattern of public sentiments, we fine-tune the pre-trained Robustly Optimized BERT Pretraining Approach (RoBERTa) with the Whole Word Masking models (Chinese-Roberta-wwm-ext) (Liu et al., 2019). The recent development in natural language processing with deep learning techniques has shown that BERT has outperformed other state-of-the-art language models (Vaswani et al., 2017; Devlin et al., 2018; Cui et al., 2021). We use the pre-trained Chinese-roberta-www-ext models and fine-tune the last classification layer and some hyper-parameters of the models such as learning rate and batch size. Table 2 shows our accuracy and F1 scores for each classifier. We will use different architectures to train our models as robustness check. Note that we broadly define China here. China can be a nation as whole, Chinese people, Chinese central/local government, CCP, State-sponsored enterprises and organizations, places, other entities related to China, etc. We group each tweet into positive, negative, or neutral.

Table 2: Model Performance on Topic Classification

Outcomes	F1 Score	Accuracy
Tweet type	0.91	0.91
COVID-19	0.93	0.97
Culture	0.16	0.98
Democracy	0.63	0.90
Economy	0.23	0.98
Politics	0.92	0.92
US Politics	0.70	0.96
Taiwan Politics	0.68	0.99
HK Politics	0.70	0.98
Religion	0.27	0.99
US	0.86	0.96
US-China Relation	0.45	0.96

To tackle the third research question, we build RoBERTa models to further discern the target entities: Chinese people, the Chinese government, and China in general. If a tweet mentions anything related to ordinary Chinese people, we label it as "Chinese people". If a tweet discusses the political system in China, we label it as "Chinese government". Examples of entities in the category include the Chinese central/local government or the communist party (also referred to as CCP), general politics in China, police departments, state media, state-sponsored companies, major political figures in China, and Beijing or Zhongnanhai when they are used to refer to the government. Sometimes people mention the Chinese government without using any specific term related to the government, and probably only using “China”, or “authoritarian regime.” In this case, it requires our annotators to use their own judgments to identify their targets and label those tweets. If a tweet talks about China, but it can’t be categorized as “Chinese people” or “Chinese government”, we label it as “China in general”. For example, Chinese traditional culture, festivals, traveling, food, etc. Table 3 shows our accuracy and F1 scores for each classifier.

Table 3: Model Performance on Sentiment and Target Classification

Outcomes	F1 Score	Accuracy
Sentiment	0.81	0.81
Target China in general	0.70	0.86
Target Chinese people	0.61	0.94
Target Chinese government	0.84	0.89

To answer the fourth research question on the dynamics between pro- and anti-Chinese state groups, we conduct a social network analysis. We use the conversation id from Twitter to construct a bipartite conversation network based on whether these pro- and anti-users classified by our BERT model engage in same conversations.

Content of Chinese Tweets

To address the fifth research question about the content of Chinese tweets, we train a series of classifiers to identify whether a tweet is related to *COVID-19, politics, economy, culture, religion, and Unite States*. Table 4 shows our accuracy and F1 scores for each classifier. We also supplement our topic classification results with structural topic models (Roberts, Stewart, & Tingley, 2019). Structural topic model, as an unsupervised text analysis tool, has been used to retrieve information from large-scale textual data and it allows researchers to flexibly estimate how document-level metadata shapes topic prevalence (Roberts et al., 2014). We run a series of structural topic models with 30 topics. We report the main themes embedded in these tweets.

Table 4: Descriptive Statistics of Prediction Results

Outcomes	Million of Tweets	Percent
Sentiments		
Negative	15.74	62.2
Neutral	5.54	21.9
Positive	4.02	15.9
Target Chinese Government	15.19	60.0
Target Chinese people	2.79	11.0
Target China in general	6.32	25.0

To address the six question, we focus on two types of accounts that either support or oppose China. We then analyze the differences in their posted tweets in our CNTweets database.

Results

The Sources and Types of Chinese Twitter users

We begin by describing the overall pattern on who produced these Chinese tweets. The descriptive analysis shows that 1% of Twitter users generated over 62% of Chinese tweets in the pandemic in our CNTweets database. The 1% rule on Internet culture suggests that only a tiny proportion of users produced the vast majority of content on a digital platform (Van Mierlo, 2014). Our CNTweets data is consistent with the 1% rule that 1% of Twitter users accounted for 61% of total Chinese tweets mentioning China, Chinese, or Asians from December 2019 to March 2021. 10% of Twitter users contributed to over 90% of total Chinese tweets in our CNTweets database. Thus, in the Twitter-verse of Chinese language users, the majority of Chinese tweets targeting China, CCP, and Asians in either positive or negative direction were driven by a handful of Twitter users ($\tilde{1}3$ thousands).

We then ran a geospatial analysis of Chinese Twitter users' self-reported locations. The majority of Chinese language Twitter users reported a location from Mainland China, U.S., Taiwan, or Hong Kong. Among 1.32 millions of Twitter users in our dataset, 0.58 million (43.83%) of users self-reported a location on their public profiles. Among those who reported certain information in the location part of the profile, we were able to identify 0.33 million (58%) users' countries/regions (e.g., Europe, Singapore, Indonesia, Japan). Among those users with identified countries/regions, the majority reported a location of Mainland China (31.62%), the United States (18.09%), Taiwan(8.95%), or Hong Kong (8.59%).

The majority of Chinese language tweets were associated with personal opinions, followed by news contents. We trained a RoBERTa classifier to discern the types of these tweets. Each tweet was classified into personal opinion (i.e., any personal expression such as personal opinion, comments, discussion or emotions about any topic), news content (e.g., news related to COVID, China, US, or other countries), government or any other institutions' announcements (e.g., announcements by government officials and World Health Organization's health advice), advertisements and spams, and others. We found that 68.4% of tweets were related to personal opinions, 27.6% were associated with news media, 0.71% of tweets were related to governments' or other institutions' announcements, and 2.16% were ads and spams. The disproportional concentration on personal opinions shows that Twitter has been a public space to express public opinions by Chinese language users.

The Overall Sentiments and Main Targets

Our BERT sentiment classifier shows that the sentiments in the Chinese tweets were predominantly negative. In our CNTweets database, tweets sharing negative, positive, and neutral sentiments toward China accounted for 62%, 22%, and 16%, respectively during the early pandemic. Fig 1 shows the time series of positive, negative, and neutral tweets. It suggests a robust pattern over time that the Chinese Twitter community was consistently negative about China during the early pandemic.

Keywords analysis shows that China and CCP were more likely to be mentioned than people of Asian or Chinese descendants. Fig 2 shows the daily trends of China, CCP, and people of Asian or Chinese descendants (亚裔/华裔). It clearly shows that Chinese language

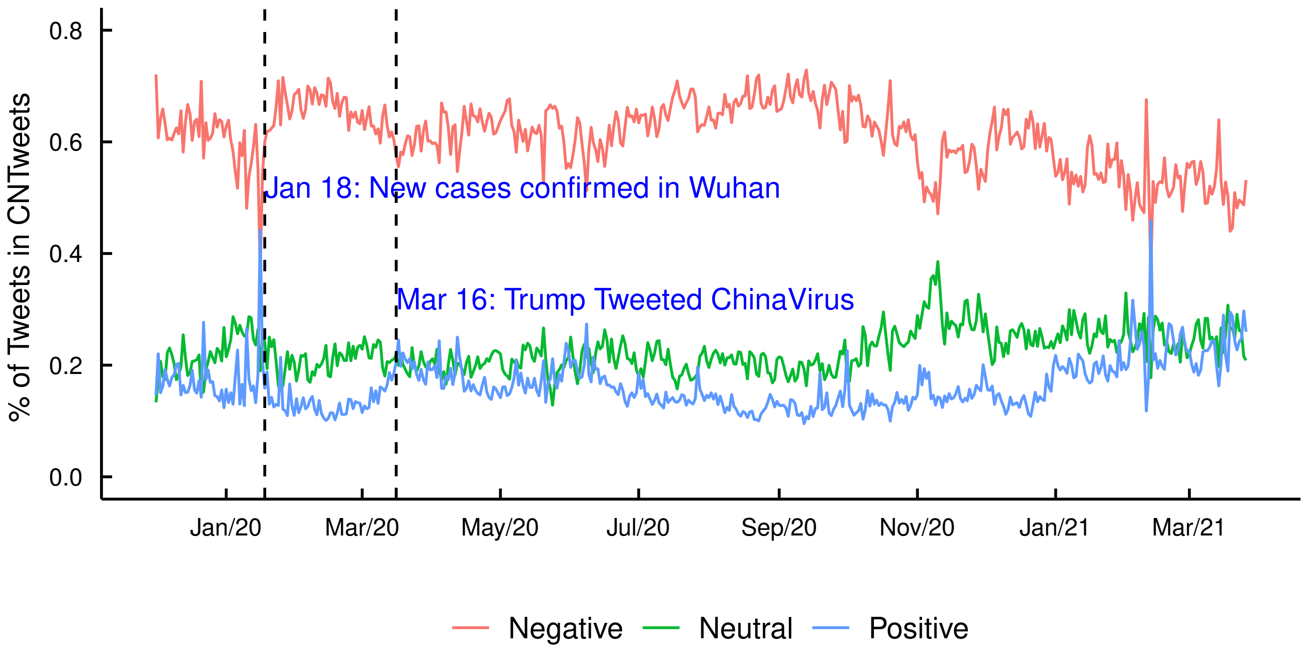

Figure 1: Sentiment analysis based on Chinese-RoBERTa-www-ext-large

Twitter users mentioned China and CCP more often than Asian or Chinese descendants. They focused on Chinese issues over their Asian/Chinese communities overseas. Fig 2 also shows that China/CCP keywords surged during the early pandemic, peaked after former U.S. President Donald Trump tweeted "China Virus" on March 16, 2020, and then remained relatively steady. For Asian related keywords, we have the similar pattern during the early pandemic, but these keywords also surged after March 2021 because of the tragic Atlanta SPA mass shootings.

Our sentiment target analysis shows that most negative tweets were targeting Chinese government or China in general instead of Chinese people. Fig 3 shows the daily trends of tweets targeting different China-related entities. The majority of sentiments in the CNTweets database were directed toward the Chinese government. During the early pandemic, around 60% of tweets were targeting Chinese government, around 11% were targeting Chinese as a ethnic group, and around 25% were targeting China in general.

For those tweets with negative sentiments, as shown in Table 5, 80% were targeting Chinese government, 11% Chinese people, and 19% China in general. For those tweets with positive sentiment, the proportions associated with Chinese government, people, and broad China were 20%, 34%, and 46%, respectively. Clearly, it shows that positive tweets were more likely to support the Chinese government.

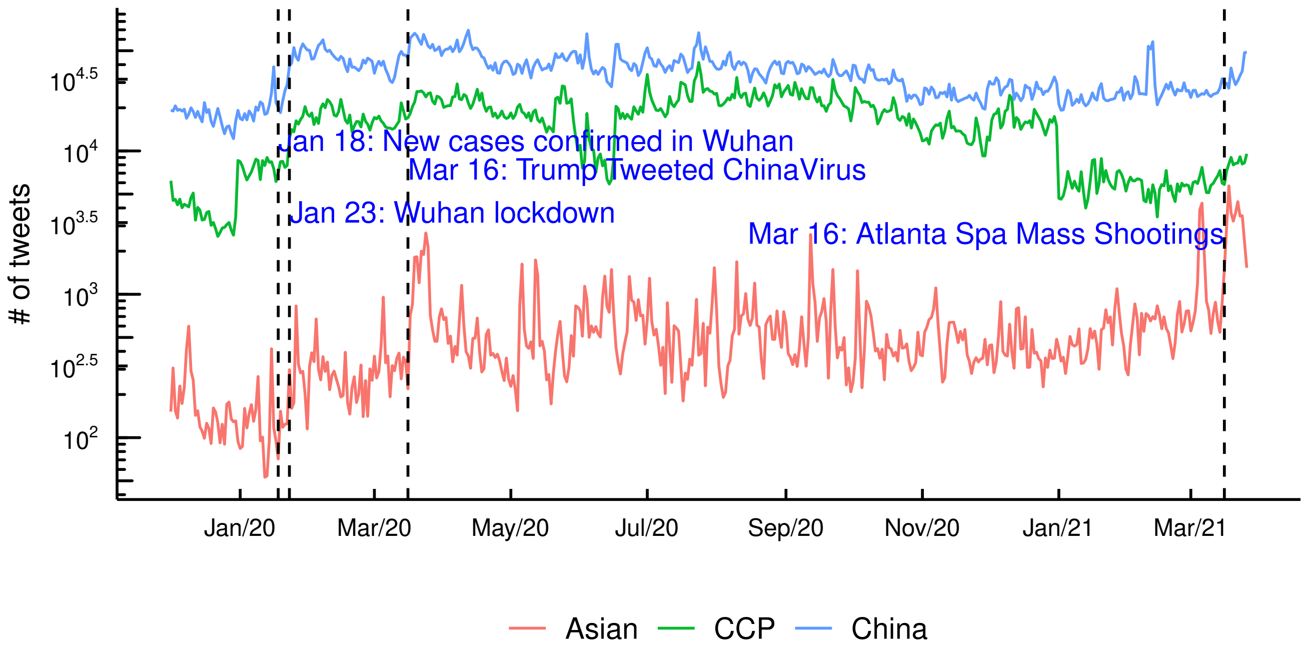

Figure 2: Daily Trend of Chinese Tweets mentioning China, Asians/Chinese, and CCP.

Table 5: Proportion of Target entities by different sentiments

Sentiment	Chinese People	Government	Broad China
Negative	0.11	0.80	0.19
Neutral	0.06	0.22	0.27
Positive	0.20	0.34	0.46

A Network Analysis of Pro- and Anti-China Twitter Users

We use the results from sentiment analysis to classify Twitter users (at least having 10 tweets in our database) into pro-China and anti-China users based on the rate of positive tweets. If a user's positive rate is larger than 0.6, we label it as pro-China user; if it is less than 0.4, we label it as anti-China user. We have 80,901 anti-China users and 29,395 pro-China users.

Then we constructed a conversation network for these pro- and anti-China users in our database based on whether these users engaged in the same conversations using Twitter's conversation_id. Twitter assigns a unique conversation id to each tweet if they engage in the same conversation sequence. Typically, the conversation id should be identical to the tweet id posted by the first user. For these identified pro- or anti- China users, we observed 18.2 million unique conversations in our database. Among these conversations, 1.04 million conversations were solely posted by one pro-China user, while 16.51 conversations had solely one anti-China user, 0.83 million conversations had at least one pro and anti-China users. Thus,

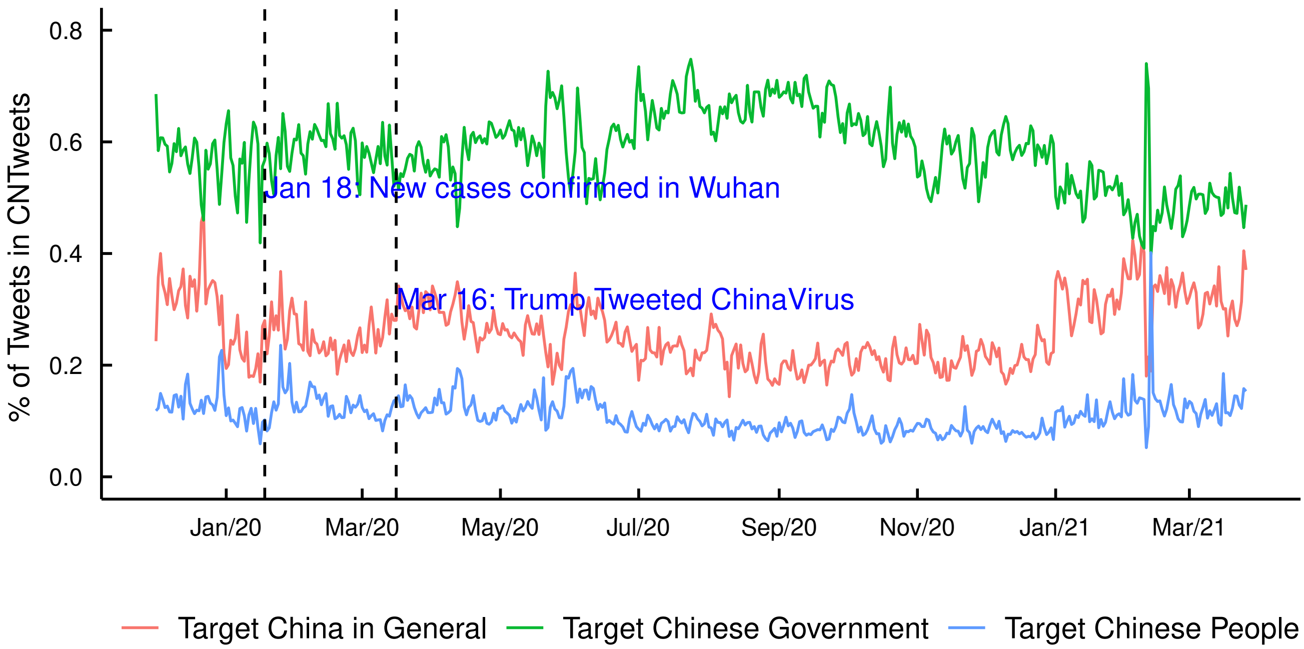

Figure 3: Daily Trend of Chinese Tweets Targeting Different Entities.

conversations between pro- and anti-China users accounted for 4.5% of total conversations occurred among identified pro- or anti-China users in our database. Fig 4 visualized the conversation network among Twitter users (with at least 10 conversations). Red dots indicate pro-China users, while blue dots denote anti-China users. It clearly shows the polarized pattern, but pro- and anti-China users did engage in some dialogues that might support or criticize China. For 173,130 conversations with at least one pro and anti-China user, we find that 22% only had one pro-China and one anti-China participant and the majority (73%) of these conversations had less than 10 pro- or anti-China users. This clearly shows that pro- and anti-China users lack direct in-depth engagement.

The Content of Chinese Tweets

The majority of tweets were related to politics, followed by democracy and freedom, U.S. issues, and COVID-19 topics. Our BERT topic classifiers show that 73% of tweets were broadly related to politics. More specifically, 31% were associated with discussions on democracy and freedom, 22% were discussing US politics, 9% were discussing Hong Kong protest issues, and 6% were mentioning Taiwan politics. 27% of these tweets were related to United States topics. Note that 14% were related to US-China relation. This is reasonable as the trade war between China and US. 20% of tweets were discussing COVID-19 related issues, while Culture, economy, and religion related topics only accounted for 6%, 5%, and 2%, respectively.

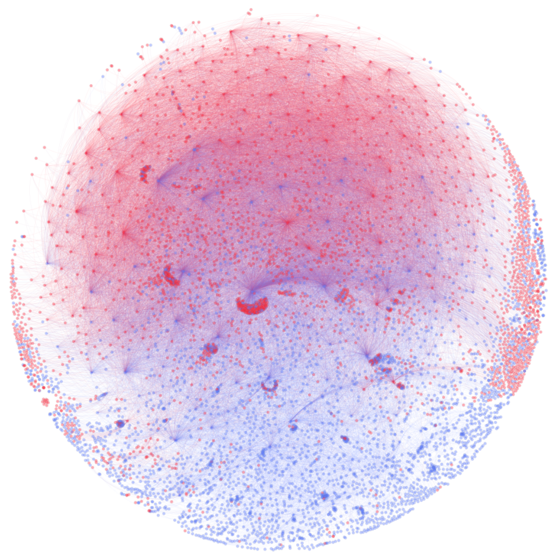

Figure 4: Network Visualization of Pro- and Anti-China Users. Nodes are pro (Red) and anti (Blue) China Twitter users and edges are conversations (at least 10).

Table 6: Proportion of main content

Outcomes	Proportion
Politics	0.73
Democracy	0.31
US	0.27
US Politics	0.22
COVID19	0.20
US-China Relation	0.14
HK Politics	0.09
Taiwan Politics	0.06
Culture	0.06
Economy	0.05
Religion	0.02

The keyword analysis shows that COVID-related keywords were frequently mentioned in the Chinese language community after the outbreak, but U.S. and Hong Kong related topics prevailed during the early pandemic. Fig 5 shows the daily trend of some keywords of interest, including COVID, Taiwan, USA, Hong Kong, Tibet, and Xinjiang. Unsurprisingly, COVID related Chinese keywords increased rapidly in the twitter community after outbreak and peaked after March but declined after April 2020. However, the U.S. and Hong Kong related topics were often discussed in the community as the U.S. trade wars and Hong Kong protests were dominating the issue attention cycle, followed by Taiwan, Xinjiang, and Tibet issues.

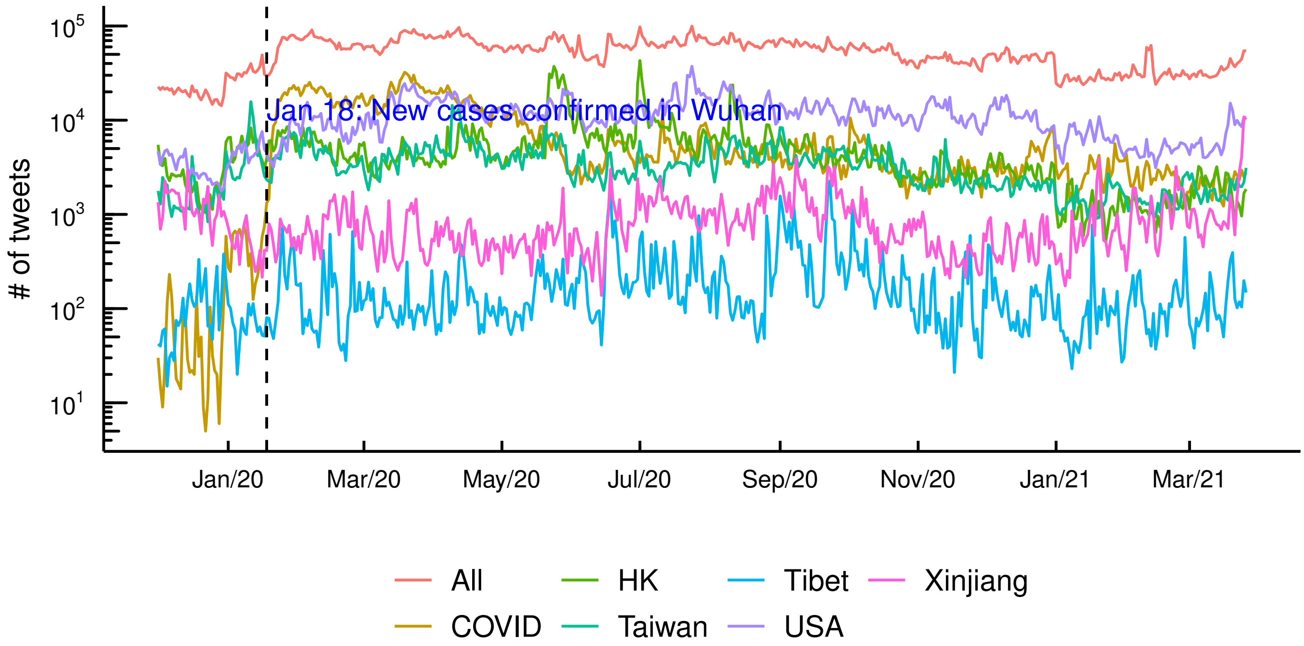

Figure 5: Daily Trend of Chinese Tweets mentioning Hong Kong, Taiwan, Xinjiang, Tibet, USA, and COVID19.

Structural topic modeling also shows that the most popular themes in CNTweets were China's domestic politics, COVID-19, US politics, and Hong Kong and Taiwan issue. Fig 6 plots the distribution of themes extracted from our CNTweets data. We estimated 30 topics using structural topic models. Results suggest that democracy-freedom (8%), U.S. election (6.9%), global issues (6%), 50 cents (i.e., supporting CCP,5.4%), culture-education (5.1%), COVID-19 (4.9), Hong Kong-National Security Law (4.8%), Wuhan outbreak(4.8%), human rights (e.g., Xinjiang, 3.7%), and U.S. China Initiative (3.6%) were the most top 10 themes during the early pandemic on Twitter. Other prevalent topics include COVID origin (made in a Wuhan lab), Huawei Ban, Chinese policing, Chinese economy, anti-CCP, etc.

The Topic Variation between Pro- and Anti-China Users

To examine whether different types of users engaged in distinct topics, we ran an additional analysis to compare topic proportions between 80,901 anti-China users and 29,395 pro-China users. Table 7 reports the average number of tweets and overall proportions for each topic within all tweets posted by these pro- or anti-China users. Both sides were heavily engaged in the topics including politics, U.S. topics, and COVID-19 issues. Over 30% of pro- or anti-China users' tweets were involved in some aspects of politics.

Pro-China users were more likely than anti-China users to tweet about economy, culture, COVID-19, and U.S. issues, compared to topics like politics. For an average pro-China user

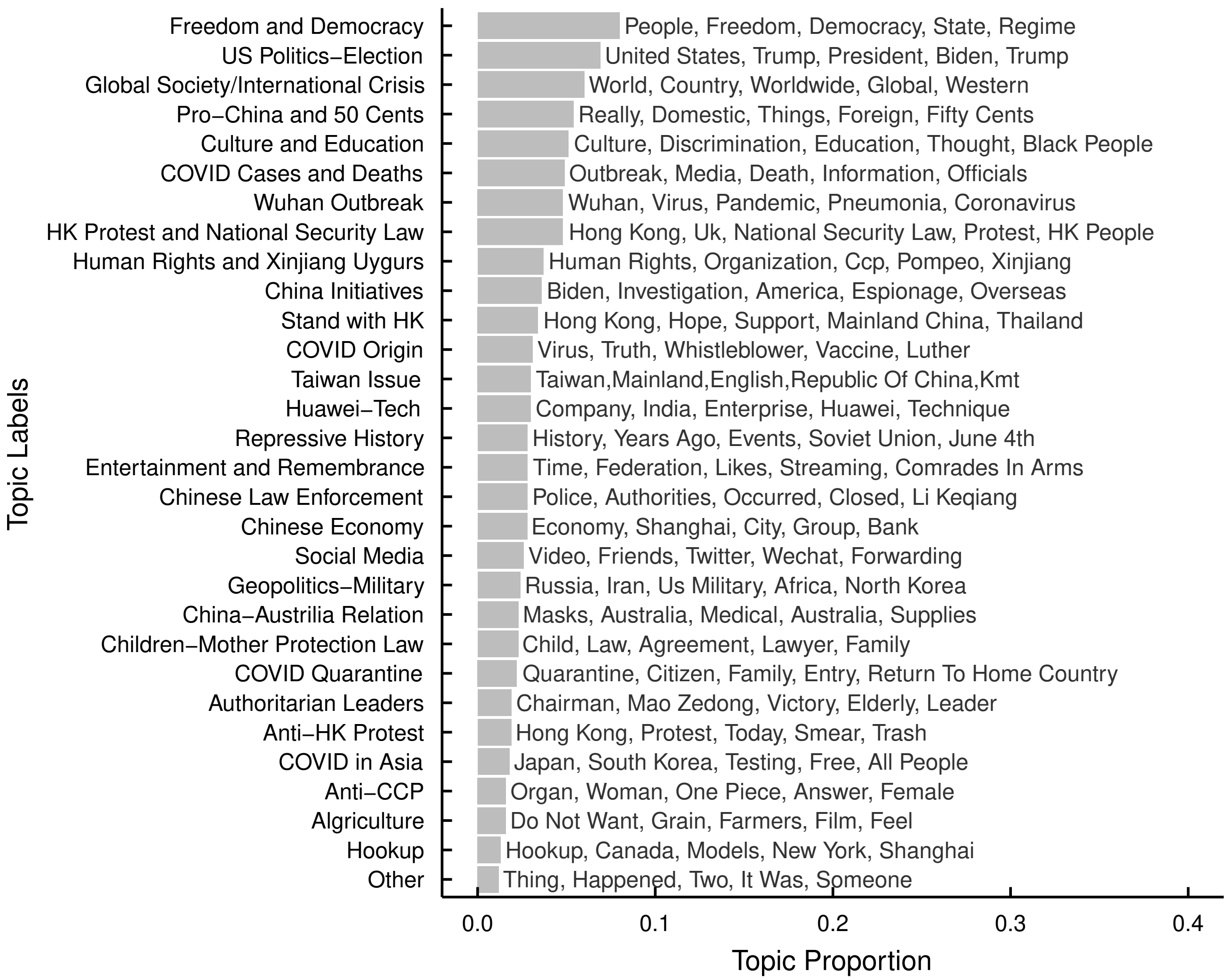

Figure 6: Structural topic model output, K=30.

in our CNTweets database, as shown in Table 7, they tend to be inactive in terms of the number of posts. For instance, a pro-China user had 38 tweets discussing politics, while an anti-China user had 191 tweets. But in terms of the topic shares for all tweets made by these users, pro-China users focused more on economy, culture, COVID-19 topic, and U.S. issues, while anti-China users focused more on politics, democracy and freedom, and Hong Kong politics. The variation in topics reflects the different agenda by these pro and anti-China users on Twitter.

Table 7: Topic Shares for Pro- and Anti-China Users

Outcomes	Pro-China (prop.)	Anti-China (prop.)
Politics	37.82(0.33)	190.73(0.34)
Democracy	6.86(0.06)	87.06(0.16)
US	17.56(0.15)	70.2(0.13)
US politics	12.9(0.11)	57.29(0.1)
COVID-19	12.05(0.11)	51.31(0.09)
US-China relation	6.62(0.06)	37.59(0.07)
HK politics	3.8(0.03)	24.44(0.04)
Taiwan politics	4.4(0.04)	15.34(0.03)
Economy	4.94(0.04)	11.21(0.02)
Culture	6.43(0.06)	10.61(0.02)
Religion	0.83(0.01)	5.47(0.01)

Discussion and Conclusion

This paper used multi-modal supervised and unsupervised machine learning tools to examine anti-China sentiments and topics in the Chinese language community on Twitter during the early COVID-19 pandemic. Since the outbreak, scholars have shown the global surge of anti-China sentiments. Our work was the first to systematically understand the sentiment dynamics targeting Chinese language communities on a major Western social media platform.

Based on the analysis of over 25 million Chinese tweets from December 2019 to April 2021, we find that the majority of these China-related tweets were generated by only 1% of Twitter users. These Chinese language users, who are most likely to report a location of Mainland China, U.S., Hong Kong, and Taiwan, tend to mention more about China or CCP instead of keywords related to people of Asian or Chinese descendants. The majority of these tweets are personal opinion oriented, followed by news-like contents and government or institutional announcements. These results suggest that tweets targeting Chinese communities might be a very selective group of users as a handful of Twitter users contributed to the majority of contents related to China topics.

We also find that the majority of tweets in our CNTweets database were negative toward China, although these sentiments were more likely to target the Chinese government or China in general instead of Chinese people. These pro-China and anti-China Twitter users were predominantly engaging with conversations on their own side, but we did observe a moderate size of Twitter users engaged in conversations on the other side. These results suggest that Twitter has been used as a major platform by anti-China users to disseminate negativity toward the Chinese government and CCP. These findings are consistent with previous literature on the lack of evidence related to computational propaganda by CCP but strong evidence of computational propaganda by anti-China groups on Twitter (Bolsover & Howard, 2019). Given that we focus solely on Chinese tweets, we cannot extend this conclusion to the entire universe of Twitter as CCP might target English communities instead of Chinese communities.

The most common topics discussed by these anti-China Twitter users were politics, such as democracy and freedom, Hong Kong protests, Taiwan politics, Xinjiang, and Tibet issues. Even though both pro- and anti-China users were heavily engaged in the discussions of politics, pro-China users were more likely to engage with topics related to economy, COVID-19, U.S. issues, and culture, while anti-China users were more likely to focus on topics of democracy and HK politics. These findings also echo that pro-democracy activists tend to take advantage of these social media platforms to promote democracy and criticize the Chinese government, while pro-China Twitter users tend to use economy and culture topics to boost China's international image.

Taken together, our findings show that Sinophobia was viral among the Chinese communities on Twitter during the early pandemic, and Twitter-verse is a battlefield that contains misinformation and disinformation on China. Previous studies often focus on the English language communities on social media platforms and overlooked non-English communities. The misinformation on the origin of COVID19 and hate speech targeting Chinese ethnic groups have negative consequences in the community. As many social media platforms have developed policies and tools to mitigate these negative consequence such as blocking hateful terms and suspending controversial accounts, but very little resources have been devoted to communities of minorities.

Readers should note that our research has some limitations. For instance, some classifiers have a relatively low F1 score (e.g., culture, religion, and economy). One of the future directions is to use semi-supervised machine learning methods to improve predictive power by adding more positive cases. In addition, we only obtained tweets during the early pandemic using keywords instead of the whole universe. We leave these to future research.

Supporting information

S1 Appendix. Data Collection Process.

S2 Appendix. Pro or Anti-China Training Dataset Collection Strategies.

References

Bi, W. (2021). Playing politics digitally: young chinese people's political feelings on social media platforms. *Cultural Studies*, 1–20.

Bolsover, G., & Howard, P. (2019). Chinese computational propaganda: automation, algorithms and the manipulation of information about chinese politics on twitter and weibo. *Information, communication & society*, *22*(14), 2063–2080.

Bovet, A., & Makse, H. A. (2019). Influence of fake news in twitter during the 2016 us presidential election. *Nature communications*, *10*(1), 1–14.

Chang, R., Lai, C., Chang, K., & Lin, C. (2021). Dataset of propaganda techniques of the state-sponsored information operation of the people's republic of china. *CoRR*, *abs/2106.07544*.

Cook, G., Huang, J., & Xie, Y. (2021). How covid-19 has impacted american attitudes toward china: A study on twitter. *arXiv preprint arXiv:2108.11040*.

Cui, Y., Che, W., Liu, T., Qin, B., & Yang, Z. (2021). Pre-training with whole word masking for chinese bert. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *29*, 3504–3514. Retrieved from `http://dx.doi.org/10.1109/TASLP.2021.3124365` doi: 10.1109/taslp.2021.3124365

DeLisle, J., Goldstein, A., & Yang, G. (2016). *The internet, social media, and a changing china*. University of Pennsylvania Press.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, *abs/1810.04805*. Retrieved from `http://arxiv.org/abs/1810.04805`

Flores, R. D. (2017). Do anti-immigrant laws shape public sentiment? a study of arizona's sb 1070 using twitter data. *American Journal of Sociology*, *123*(2), 333–384.

Han, R. (2019). Patriotism without state blessing: Chinese cyber nationalists in a predicament. In *Handbook of protest and resistance in china*. Edward Elgar Publishing.

Huang, Z. A., & Wang, R. (2019). Building a network to "tell china stories well": Chinese diplomatic communication strategies on twitter. *International Journal of Communication*, *13*, 24.

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in china allows government criticism but silences collective expression. *American political science Review*, *107*(2), 326–343.

King, G., Pan, J., & Roberts, M. E. (2014). Reverse-engineering censorship in china: Randomized experimentation and participant observation. *Science*, *345*(6199), 1251722.

King, G., Pan, J., & Roberts, M. E. (2017). How the chinese government fabricates social media posts for strategic distraction, not engaged argument. *American political science review*, *111*(3), 484–501.

Lee, J., & Huang, T. J. (2021). *Reckoning with asian america* (Vol. 372) (No. 6537). American Association for the Advancement of Science.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., … Stoyanov, V. (2019). Roberta: A robustly optimized BERT pretraining approach. *CoRR*, *abs/1907.11692*.

Lu, Y., & Pan, J. (2021). Capturing clicks: How the chinese government uses clickbait to compete for visibility. *Political Communication*, *38*(1-2), 23–54.

Lu, Y., Pan, J., & Xu, Y. (2021). Public sentiment on chinese social media during the emergence of covid19. *Journal of Quantitative Description: Digital Media*, *1*.

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of twitter sentiment on stock price returns. *PloS one*, *10*(9), e0138441.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An r package for structural topic models. *Journal of Statistical Software*, *91*(1), 1–40.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., … Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064–1082.

Schneider, F. (2018). *China's digital nationalism*. Oxford University Press.

Shmargad, Y. (2018). Twitter influencers in the 2016 us congressional races. *Journal of Political Marketing*, 1–18.

Song, S. Y., Faris, R., & Kelly, J. (2015). Beyond the wall: Mapping twitter in china. *Berkman Center Research Publication*(2015-14).

Sullivan, J. (2012). A tale of two microblogs in china. *Media, culture & society*, *34*(6), 773–783.

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the international aaai conference on web and social media* (Vol. 4).

Van Mierlo, T. (2014). The 1% rule in four digital health social networks: an observational study. *Journal of medical Internet research*, *16*(2), e2966.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. In I. Guyon et al. (Eds.), *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc.

Wagner, C., Strohmaier, M., Olteanu, A., Kıcıman, E., Contractor, N., & Eliassi-Rad, T. (2021). Measuring algorithmically infused societies. *Nature*, *595*(7866), 197–204.

Zhang, D. (2021). Sinophobic epidemics in america: Historical discontinuity in disease-related yellow peril imaginaries of the past and present. *Journal of Medical Humanities*, *42*(1), 63–80.

Zhang, M. M., Wang, X., & Hu, Y. (2021). Strategic framing matters but varies: A structural topic modeling approach to analyzing china's foreign propaganda about the 2019 hong kong protests on twitter. *Social Science Computer Review*, 08944393211042575.

Ziems, C., He, B., Soni, S., & Kumar, S. (2020). Racism is a virus: Anti-asian hate and counterhate in social media during the covid-19 crisis. *arXiv preprint arXiv:2005.12423*.