

Multiplex Anti-Asian Sentiment before and during the Pandemic: Introducing New Datasets from Twitter Mining

Pradeep Nalluri* Hao Lin[†] Lantian Li[‡] Yifan Sun[§] Yongjun Zhang[¶]

December 28, 2021

Abstract COVID-19 has unprecedentedly threatened minority communities in the U.S. But social scientists and policymakers lack critical data to capture the dynamics of the anti-Asian trend and evaluate its scale and scope. We introduce news datasets from Twitter mining related to anti-Asian hate sentiment before and in the pandemic. Relying on Twitter’s academic API, we retrieve hateful, counter-hate, and anti-Chinese politics tweets from the Twitter Historical Database. We then use the state-of-the-art hate speech classifiers to discern whether these tweets express hatred. These datasets can be used to study hate speech, general anti-Asian or Chinese sentiment, and hate linguistics by social scientists as well as to evaluate and build hate speech or sentiment analysis classifiers by computational scholars.

Keywords: Hate speech, Sinophobia, COVID19, Anti-Asian, Anti-China, Twitter mining

1 Introduction

The COVID-19 pandemic has unprecedentedly threatened minority communities in the U.S. (Chiriboga et al. 2020). In particular, Asians, Asian Americans, and Pacific Islanders (AAPI) are the primary victims of hate speech and hate crime due to systemic racism and xenophobia (Croucher, Nguyen and Rahmani 2020). Since the outbreak of COVID-19, there were over 4,000 hate incidents reported to stopaapihate.org. The growing anti-Asian attacks have led to the recent passage of the anti-Asian Hate Crimes Bill by the U.S. House after the mass shootings in Atlanta. Despite the troublesome surging COVID-hate incidents and crimes targeting AAPI communities, *social scientists and policymakers lack critical data and quantitative measures to capture the evolution of anti-AAPI trends in the U.S., and cannot evaluate the scale and scope of anti-AAPI hate incidents in the pandemic.*

*Pradeep Nalluri is a master student in the Department of Computer Science at Stony Brook University.

[†]Hao Lin is a doctoral student in the Department of Sociology and the Institute for Advanced Computational Science at Stony Brook University.

[‡]Lantian Li is a Ph.D. candidate in the Department of Sociology at Northwestern University.

[§]Yifan Sun is an Assistant Professor in the Department of Computer Science and the Institute for Advanced Computational Science at Stony Brook University.

[¶]Yongjun Zhang is an Assistant Professor in the Department of Sociology and the Institute for Advanced Computational Science at Stony Brook University. To whom correspondence should be addressed. E-mail: yongjun.zhang@stonybrook.edu. We wish to thank the Institute for Advanced Computational Science at Stony Brook University for access to high performance computing and its generous seed grant support for this project.

Recent scholarship has used social media data with machine learning techniques to track anti-Asian hate speech on social media platforms (Vidgen et al. 2020; Ziems et al. 2020; Jiang et al. 2020). For instance, Ziems et al. (2020) examined the evolution and spread of anti-Asian hate speech using 30 million tweets collected on Jan 15–Apr 17, 2020. Cook, Huang and Xie (2021) classified over 297 million tweets about China or COVID-19 up to June 2020 using a BERT model trained on 5000 tweets and found that the awareness of COVID-19 has led to a sharp rise in anti-China sentiments in the U.S. Although these studies provide training data sets to build hate speech classifiers and offer some insights about the spread of anti-Asian hate at the early stage of the outbreak, little is known about the enduring evolution of anti-Asian hate or counter hate before and in the continuing pandemic.

In this paper, we report on trends and patterns of anti-Asian sentiments and hate speech on Twitter by introducing new datasets. Twitter has been one of the most salient battlegrounds fighting against misinformation, fake news, hatred, and xenophobia in the COVID-19 pandemic. We use computational tools with natural language processing and machine learning methods to detect hate speech from Twitter before and in the pandemic. Our data sets contain 63.35 million tweets related to general hate terms targeting Asians and COVID19-specific racial slurs as well as anti-Chinese politics and counter hate, and they provides a comprehensive portrait of the dynamics of anti-Asian hate sentiment spanning from 2007 to 2021. Thus, we are able to address important questions related to the evolution of anti-Asian hate sentiments over time.

2 Literature Review

In the past decade, computational scholars have made great efforts to detect hate speech on social media platforms (Davidson et al. 2017). Although there is no clear and formal definition of hate speech, scholars tend to define hate speech as language or speech “used to express hatred towards a targeted group or is intended to be derogatory, to humiliate, or to insult the members of the group” Davidson et al. (2017). Anti-Asian hate speech, especially Sinophobia, has attracted numerous attention from both computer scientists and social scientists (Vidgen et al. 2020; Cook, Huang and Xie 2021; Lee 2021; Ziems et al. 2020). Anti-Asian hate speech could be understood as any speech that targets Asians and people of Asian descendants in a way that promotes hatred, violence, or social disorder.

Anti-Asian is not a unique thing only in the COVID-19 pandemic. Scapegoating immigrants in public crises has a long tradition in American history (Daniels et al. 2021). For instance, tuberculosis was called “the Jewish disease” in the 1890s and Italian immigrants were blamed for the 1916 polio epidemic. Asian Immigrants were also the frequent targets of medical scapegoating. Chinese immigrants were blamed for the spread of smallpox in the later 1870s in San Francisco and the SARS epidemic in 2003 (Daniels et al. 2021). Now Asian immigrants, specifically Chinese immigrants, were scapegoated for the origin and spread of COVID-19. Since the pandemic, around 2 million Asian American adults have experienced some form of anti-Asian hate incident such as being beaten, spit on, and harassed based on AAPI Data reports (Lee 2021). Federal-level hate crime data also shows that anti-Asian hate crimes increased by 149% in 16 major cities while hate crimes, in general, decreased by 7%. When former President Trump singled out China for blaming the origin of the pandemic and called the novel coronavirus “Chinese Virus” and “Kung Flu,” Asian Americans, particularly immigrants of Chinese descent, became the primary target of anti-Asian hate crimes. But even before the COVID-19 outbreak, hate crimes against Asians increased 31% in

2016-2018 under the Trump administration (Lee 2021). Anti-Asian bigotry, violence, and misogyny were ingrained in American history following the passage of the Page Act of 1875 and the 1882 Chinese Exclusion Act which prohibited the importation of Chinese laborers and women. Asian Americans were often perceived as “disease, unassimilable aliens, and economic, cultural, and moral threats to a free White republic” (Lee 2021). Even today, Asian Americans, particularly women of Asian descendants, are still the victims of anti-Asian misogyny and hatred (e.g., the Atlanta mass shooting in March 2021).

3 Data

To understand the anti-Asian hate sentiment before and in the pandemic, we use keywords related to COVID specific hate terms, general anti-Asian hate terms, anti-Chinese politics terms, and counter-hate terms to extract all relevant tweets from Twitter’s historical database. This allows us to obtain a holistic view of multiplex anti-China or Asian hate sentiment. Here we briefly describe how we built our datasets but you can get more information in the Appendix. Since online hate speech mainly targets China, we focus on these anti-China related keywords in the present study. We are incorporating more keywords related to other Asian countries.

COVID-Specific Hate Data. We use keywords, such as ‘*chinese virus*’, ‘*china virus*’, ‘*wuhan virus*’, ‘*kungflu*’, and their variants, to extract all relevant tweets that target AAPI communities. We use Ziems et al.’s classifier to exclude all these tweets related to counter-hate.

General Anti-AAPI Hate Data. We use keywords, such as ‘*bamboo coon*’, ‘*chigger*’, ‘*chinese wetback*’, ‘*ching chong*’, ‘*chonky*’, ‘*chunky*’, ‘*slant eye*’, ‘*slopehead*’, ‘*bat eater*’, ‘*chink*’, ‘*ling ling*’, and ‘*commies*’ to extract all non-COVID19 related tweets that target AAPI communities. Some of the keywords can be used in a multitude of scenarios, in such cases we removed using a few filters to get the tweets that are just targeted towards the AAPI community.

Anti-Chinese Politics Data We use keywords and hashtags, such as ‘*BoycottChina*’, ‘*MakeChinaPay*’, ‘*StandWithHongKong*’, ‘*FreeTibet*’, ‘*FuckChina*’, ‘*CCP_is_terrorist*’, and ‘*Chinazi*’ to extract all tweets that target china politically.

Counter-Hate Data. We collect counter hate data from two sources. First, we use keywords related to counter hate terms, such as *Racisimisvirus*, *StopAsianHate*, and *StopAAPIHate*. We then use Ziems et al.’s classifier to extract those counter-hate tweets from the dataset compiled using hate terms.

Table 1 shows basic statistics for our four main datasets compiled using Twitter academic API. Note that these datasets span across different time periods. Researchers can use these datasets for different purposes. For instance, we can use these datasets to test the following hypotheses related to the overall anti-Asian or Chinese hate sentiment in the COVID-19 pandemic or the persistence of anti-Asian sentiment.

Hypothesis 1. The overall anti-Asian hate sentiment should be consistent before and after the outbreak of the COVID-19 pandemic.

Hypothesis 2. The pandemic threat has engendered the rise of COVID-19 specific hate sentiment on Twitter.

TABLE 1. Summary of Twitter Data (in Million)

	COVID-Hate	AAPi-Hate	Anti-Chinese Politics	Counter-Hate
# of tweets	12.93	12.92	32.6	4.9
# of unique tweets	3.29	7.24	6.36	0.68
# of retweets	9.64	5.68	26.29	4.2
# of Twitter users	3.15	4.58	2.85	1.97
Time range	2019.12-2021.3	2008.1-2021.3	2019.12-2020.12	2019.12-2021.3

4 Results

Next, we report some major trends and patterns from our four main datasets. We start with COVID19-Specific hate data.

4.1 COVID-Specific Hate Data

In the early stage of the COVID-19 pandemic, Twitter users used COVID-19 specific terms such as ‘chinese virus’, ‘china virus’, ‘wuhan virus’, and ‘kungflu’ to describe the novel coronavirus. On Feb 2020, WHO named the disease caused by the novel coronavirus as COVID-19, but still these racial slurs remained popular on Twitter, especially after for U.S. President Trump tweeted Chinese Virus multiple times in three consecutive days on Mid March, 2020.

The United States will be powerfully supporting those industries, like Airlines and others, that are particularly affected by the Chinese Virus. We will be stronger than ever before!
Mar 16, 2020

Cuomo wants “all states to be treated the same.” But all states aren’t the same. Some are being hit hard by the Chinese Virus, some are being hit practically not at all. New York is a very big “hotspot”, West Virginia has, thus far, zero cases. Andrew, keep politics out of it.... Mar 17, 2020

I always treated the Chinese Virus very seriously, and have done a very good job from the beginning, including my very early decision to close the “borders” from China-against the wishes of almost all. Many lives were saved. The Fake News new narrative is disgraceful false! Mar 18, 2020

Figure 1 clearly shows the pattern. The blue line shows the overall trend of tweets mentioning any COVID-19 racial slurs and it peaked around Mid March after Trump tweets Chinese Virus.

We also used those state-of-the-art hate speech detection algorithms to classify whether these racial slurs count as hate speech. In general, all hate speech detectors have some degree of noise and subjectivity; in this case, running the model provided in Davidson et al. (2017) on the dataset provided in Ziems et al. (2020) gives a 42% false positive rate and 21% false negative rate. We provide our potential users with three sets of labels classified by Ziems et al. (2020), Davidson et al. (2017), and Vidgen et al. (2020). We observe some consistency between Ziems classifier and Vidgen classifier. These are both trained on COVID-19 specific tweets. Davidson classifier is less likely to classify these racial slurs as hate speech, given that it focuses more on general hatred.

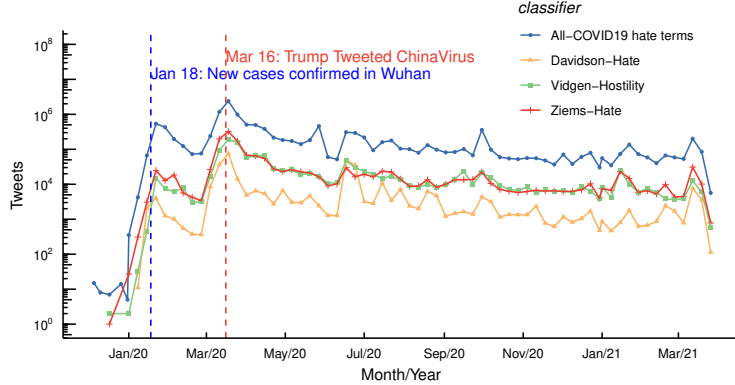


FIGURE 1. WEEKLY TREND OF COVID-19 HATE TERMS, AS CLASSIFIED BY THREE DIFFERENT HATE-SPEECH DETECTORS.

4.2 General Anti-AAPI Hate Data

COVID19-specific hate terms can only give us the anti-China or Asian hate sentiment after the novel coronavirus outbreak. We are not sure whether these hate sentiment against Chinese or Asians is consistent spanning across different time periods. For comparison, we built the general anti-AAPI hate data using these anti-Chinese or Asian hate terms.

Figure 2 shows the monthly trends of different anti-AAPI hate terms in our database. Note that our dataset only captures those tweets that still remained in Twitter historical database and might underestimate the total volume of hateful tweets because some tweets might get deleted by Twitter or users.

The top blue line in Figure 2 shows the overall trend of all general hate terms we used to obtain our Twitter data in Jan 2008 to March 2021. We see a rapid increase in tweets using anti-AAPI racial slurs from the founding of Twitter in 2007 to early 2013, and this growth may be attributed to the exponential growth of Twitter users at the same time. But after that, we see a decline pattern in the Obama administration before 2017. After Trump took over the Oval office, we see a clear increase in these hateful tweets, given that the U.S.-Sino relation worsened due to trade war and other sensitive issues like Taiwan and Hong Kong protest. We also present major general hate terms used by Twitter users, including Chink, Coolie, Sideways, Chinaman, Chunk, Chonk, etc. One interesting pattern is that we see a huge increase in using Chunk or chonk after 2018.

4.3 Anti-Chinese Politics Data

COVID19-specific and general anti-AAPI hate terms provides us with a general understanding of hate sentiment against Chinese and Asian communities. Still, they might just one tip of the iceberg. To examine the relationship between the general anti-China sentiment on Twitter, we used a series of anti-Chinese politics terms to retrieve potential anti-Chinese state tweets in 2019-2020 for comparison. We also provide our hate speech labels with different classifiers.

Figure 3 shows the weekly trend of these anti-Chinese politics terms on Twitter from Jan 2019 to December 2020. We observe a increasing in Tweets mentioning any anti-Chinese politics such as BoycottChina, MakeChinaPay, and Uyghur before the outbreak of COVID-19. But since then, the total number of anti-Chinese politics tweets bounced back and fluctuated in the early stage of

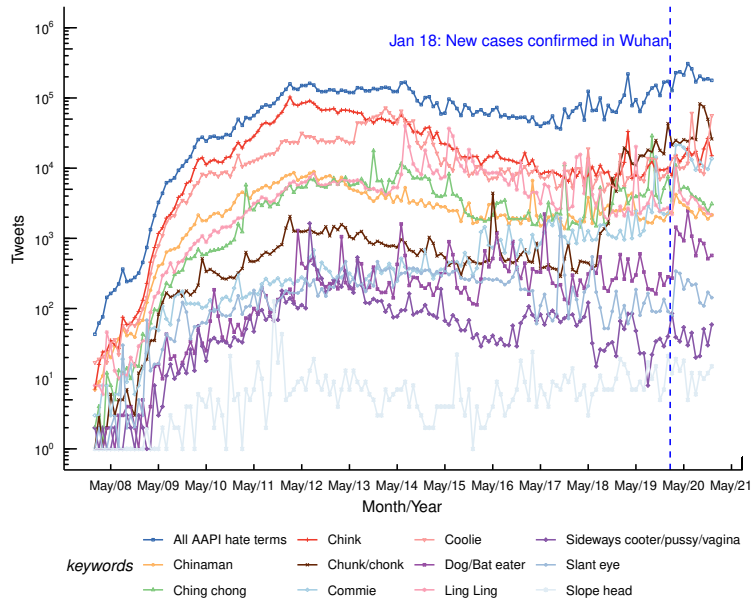


FIGURE 2. MONTHLY TREND FOR ANTI-AAPI HATE TERMS, AS CLASSIFIED BY THREE DIFFERENT HATE-SPEECH DETECTORS.

the pandemic. We suspect Twitter users' attention has shifted from anti-Chinese politics to these COVID-19 specific issues.

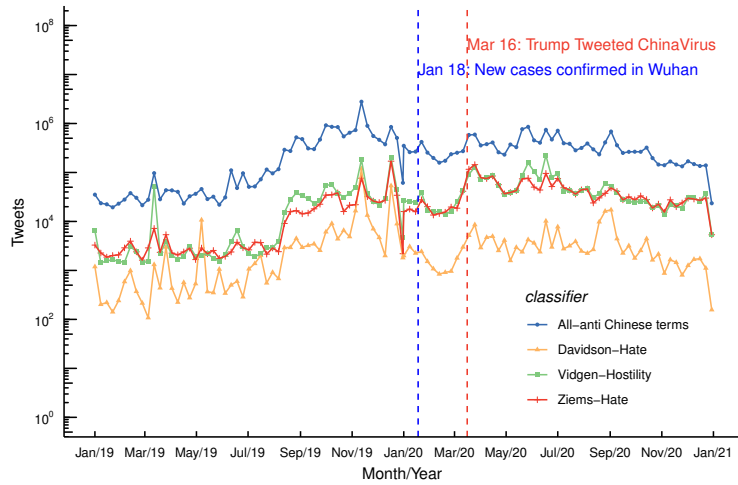


FIGURE 3. WEEKLY TREND OF ANTI-CHINESE POLITICS TWEETS, AS CLASSIFIED BY THREE DIFFERENT HATE-SPEECH DETECTORS.

4.4 COVID Counter-Hate Data

In addition to hate datasets, we also built a counter-hate dataset to assess the dynamics between pro- and anti-Chinese or Asian groups. Figure 4 shows the overall counter-hate weekly tweets after the outbreak of COVID-19.

Since the pandemic, we have seen a troublesome surge of anti-Asian attacks. This raises substantial concerns within the AAPI communities. We see a rapid increase in tweets that counter anti-AAPI hate speech such as *RacismIsVirus* and *StopAAPIHate*. The counter-hate tweets peaked after Trump tweets Chinese Virus on March and then declined. The *StopAAPIHate* movement took off after the early 2021 and peaked after the tragedy of Atlanta Spa Mass Shootings on March 16 2021. Our dataset provides unique author ids and conservation ids which allow researchers to assess the interaction among Twitter users.

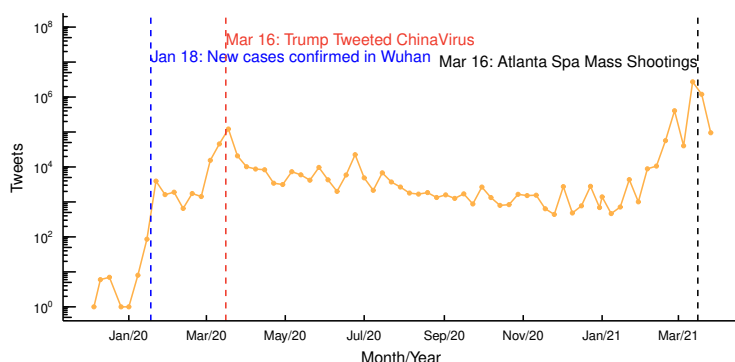


FIGURE 4. WEEKLY TREND OF COUNTER-HATE TWEETS

4.5 Hashtag Analysis

Here we provide some basic hashtag analysis in our four main datasets. What are the most popular hashtags in our datasets?

Figure 5 shows the hashtags used by Twitter users co-occurring with other keywords. Panel A shows that the most popular hashtags in anti-Chinese politics dataset are *#StandwithHongKong* and *#HongKong*. Panel B shows that the most popular hashtags used by counter hate users are related to *#StopAsianHate* and *#AsiansAreHuman*. Panel C shows that the most popular hashtags in COVID19-specific hate dataset are *#ChineseVirus*, *#CoronaVirus*, and *#WuhanVirus*. Panel D shows that the most popular hashtags used in general anti-AAPI hate dataset are *#boycottChina*, *#China*, and *#CCP*.

4.6 Overlapping Analysis

We conduct an extra analysis to examine the overlapping between COVID-Hate and AAPI-Hate data as well as between COVID-Hate and anti-Chinese politics data. We suspect that Twitter users who expressed general anti-AAPI hate and anti-Chinese politics were also more likely to show COVID-specific hatred in the pandemic. For those who posted COVID-19 specific hate terms in the pandemic, there are 741,802 Twitter users from the general anti-AAPI hate dataset and contributed 7.2 million tweets. These twitter users accounted for 23.57% of total users and 55.71%

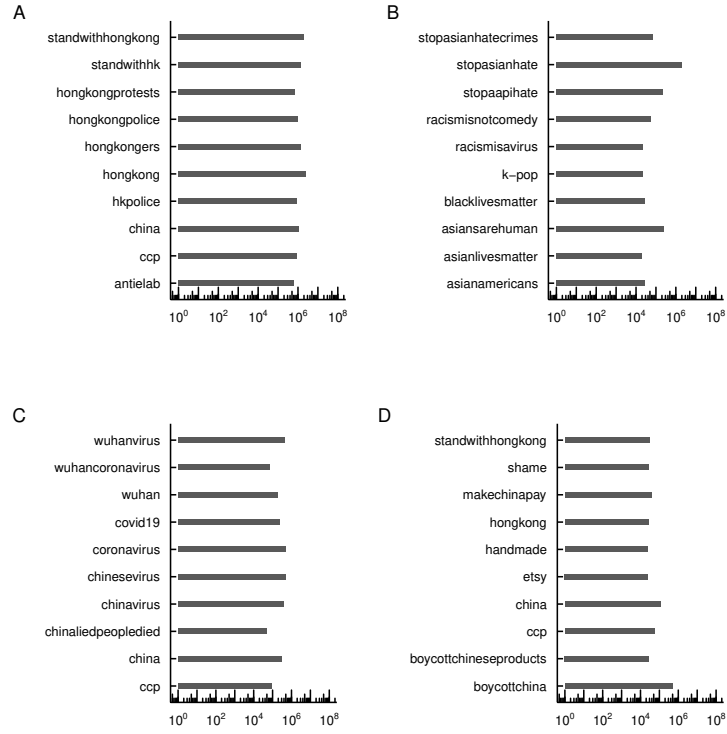


FIGURE 5. TOP HASHTAGS IN FOUR MAIN DATASETS. PANEL A: ANTI-CHINESE POLITICS; PANEL B: COUNTER HATE; PANEL C: COVID-SPECIFIC HATE; PANEL D: GENERAL ANTI-AAPI HATE.

of total tweets in COVID-Hate dataset. There are also 864,287 Twitter users from anti-Chinese politics dataset overlapping with COVID specific hate data and contributed 7.86 million tweets. These twitter users accounted for 27.46% of total users and 60.77% of total tweets in COVID-Hate dataset. Figure 6 shows the monthly or weekly trends of these overlapping Twitter users.

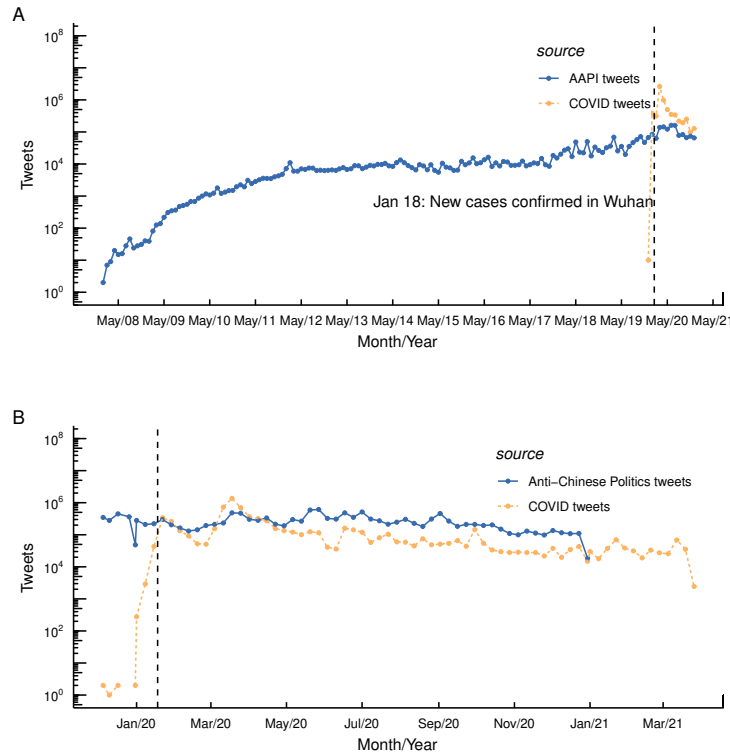


FIGURE 6. MONTHLY OR WEEKLY TREND OF OVERLAPPING TWITTER USERS. PANEL A: BETWEEN COVID-HATE AND AAPI-HATE; PANEL B: BETWEEN COVID-HATE AND ANTI-CHINESE POLITICS

5 Discussion

This paper introduces new datasets to study anti-Asian hate speech and sentiment on Twitter before and in the pandemic. We show that the overall anti-Chinese/Asian hate sentiment were consistent before and after the outbreak of the COVID-19 pandemic, but the pandemic threat has engendered the rise of COVID-19 specific hate sentiment on Twitter.

Hate speech online is a multiplex phenomenon. We built our datasets using keywords related to COVID-19 specific hate terms, general anti-AAPI hate terms, and anti-Chinese politics terms as well as counter hate terms. Instead of training our own deep learning algorithms to classify whether or not a tweet is hate speech, we provide a series of labels using state-of-the-art hate speech detectors. As shown in our main analysis, we demonstrate that we can use these datasets to illustrate the overall trends and patterns of anti-Asian hate speech online.

Researchers can also use these datasets to study questions of how Twitter users are radicalized by engaging into controversial conversations or what the linguistic features of hate speech on Twitter. Researchers can also use the specific author id and conversation id to conduct conversation network analysis. But future users should be cautious when using these datasets. Although our datasets provide a relatively comprehensive picture of anti-Asian tweets, we definitely underestimate the total volume of these tweets given that we use keywords to build our datasets. Still, our novel datasets can contribute to research in the areas of computational social science, machine learning, and hate speech detection.

6 Code and Data Availability

Codes and aggregated data used to replicate main figures are available via <https://osf.io/xtw4c/>. Dis-aggregated data are available for academic use upon request. You can email the corresponding author for the data sharing agreement form.

7 Appendix

Here we provide a detailed list of keywords and hashtags we used to extract all tweets.

COVID-Specific Hate Data. We use keywords, including *'chinese virus', 'china virus', 'wuhan virus', 'wuhan coronavirus', 'kungflu', 'china coronavirus', 'chinese coronavirus', 'chinavirus', 'chinesevirus', 'wuhanvirus', 'Kung flu'*.

General Anti-AAPI Hate Data. *'bamboo coon', 'bamboo coons', 'celestial', 'celestials', 'chigger', 'chiggers', 'chinese wetback', 'chinese wetbacks', 'ching chong', 'ching chongs', 'chinig', 'chinigs', 'chink a billies', 'chink a billy', 'chonkies', 'chonky', 'chunkies', 'chunky', 'coolie', 'coolies', 'sideways cooter', 'sideways cooters', 'sideways pussies', 'sideways pussy', 'sideways vagina', 'sideways vaginas', 'slant eye', 'slant eyes', 'slopehead', 'slopeheads', 'aseng', 'bat eater', 'boycottchina', 'chinadidthis', 'chinaman', 'chinamen', 'chink', 'chinky', 'cokin, dog eater', 'fuckchina', 'ling ling', 'makechinapay', 'niakoué', 'pastel de flango', 'slant-eye', 'ting tong', 'idiot chink', 'chinky bat, commie', 'commies'.*

Anti-Chinese Politics Data. *'BoycottChina', 'MakeChinaPay', 'BoycottChineseProducts', 'boycottchina', 'StandWithHongKong', 'BoyCottChina', 'Uyghur', 'ReplaceIT', 'BoycottMadeInChina', 'FreeUyghur', 'boycottChina', 'antichinazi', 'CCPChina', 'BoycottChineseProduct', 'FreeTibet', 'FuckChina', 'CCP_is_terrorist', 'FreeHongKong', 'StopChina', 'BOYCOTTChina', 'StandwithHK', 'fuckchina', 'Chinazi', 'Tibet', 'Genocide', 'AnywherebutChina', 'ABC_challenge', 'Uyghurs', 'China_is_terrorist', 'HongKongers', 'BOYCOTTCHINA', 'XiJinping', 'MadeInChina', 'Boycottchina', 'TakeDownTheCCP', 'AntiChinazi', 'FreeHK', 'ChineseProductsInDustbin', 'SOSHK', 'BoycottChineseApp', 'FUCKCHINA', 'SanctionChina', 'RemoveChinaApps', 'chinazism', 'fuckchina', 'SaveUygur', 'Chinamustfall', 'HKPoliceState', 'HoldChinaAccountable', 'StandWithHK', 'Xitler', 'CCPChina', 'HongKongPolice', 'Communist', 'BoycottCh', 'antitotalitarianism', 'ChinaBacksTerror', 'antiELAB', 'FreedomHK', 'TaiwanIsNotChina', 'Hongkongprotest', 'boycottchinaproducts', 'fuckChina'.*

Counter-Hate Data. *Racisimivirus, StopAsianHate, and StopAAPIHate.*

References

- Cook, Gavin, Junming Huang and Yu Xie. 2021. “How COVID-19 has Impacted American Attitudes Toward China: A Study on Twitter.” *arXiv preprint arXiv:2108.11040* .
- Croucher, Stephen M, Thao Nguyen and Diyako Rahmani. 2020. “Prejudice toward Asian Americans in the COVID-19 pandemic: the effects of social media use in the United States.” *Frontiers in Communication* 5:39.
- Daniels, Chelsea, Paul DiMaggio, G Cristina Mora and Hana Shepherd. 2021. Has Pandemic Threat Stoked Xenophobia? How COVID-19 Influences California Voters’ Attitudes toward Diversity and Immigration. In *Sociological Forum*. Wiley Online Library.
- Davidson, Thomas, Dana Warmley, Michael Macy and Ingmar Weber. 2017. Automated hate speech detection and the problem of offensive language. In *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 11.
- Jiang, Julie, Emily Chen, Shen Yan, Kristina Lerman and Emilio Ferrara. 2020. “Political polarization drives online conversations about COVID-19 in the United States.” *Human Behavior and Emerging Technologies* 2(3):200–211.
- Lee, Jennifer. 2021. Reckoning with Asian America and the New Culture War on Affirmative Action. In *Sociological Forum*. Wiley Online Library.
- Vidgen, Bertie, Austin Botelho, David Broniatowski, Ella Guest, Matthew Hall, Helen Margetts, Rebekah Tromble, Zeerak Waseem and Scott Hale. 2020. “Detecting East Asian prejudice on social media.” *arXiv preprint arXiv:2005.03909* .
- Ziems, Caleb, Bing He, Sandeep Soni and Srijan Kumar. 2020. “Racism is a virus: Anti-asian hate and counterhate in social media during the covid-19 crisis.” *arXiv preprint arXiv:2005.12423* .