Understanding the Socio-Economic Disruption in the United States during COVID-19's Early Days

Swaroop Gowdra Shanthakumar, Anand Seetharam, Arti Ramesh

Computer Science Department, SUNY Binghamton

(sgowdra1, aseethar, artir)@binghamton.edu

Abstract

In this paper, we collect and study Twitter communications to understand the socio-economic impact of COVID-19 in the United States during the early days of the pandemic. With infections soaring rapidly, users took to Twitter asking people to self isolate and quarantine themselves. Users also demanded closure of schools, bars, and restaurants as well as lockdown of cities and states. The communications reveal the ensuing panic buying and the unavailability of some essential goods, in particular toilet paper. We also observe users express their frustration in their communications as the virus spread continued. We methodically collect a total of 530,206 tweets by identifying and tracking trending COVID-related hashtags. We then group the hashtags into six main categories, namely *1)* General COVID, *2)* Quarantine, *3)* Panic Buying, *4)* School Closures, *5)* Lockdowns, and *6)* Frustration and Hope, and study the temporal evolution of tweets in these hashtags. We conduct a linguistic analysis of words common to all the hashtag groups and specific to each hashtag group. Our preliminary study presents a succinct and aggregated picture of people's response to the pandemic and lays the groundwork for future fine-grained linguistic and behavioral analysis.

1 Introduction

COVID-19 (also known as the novel coronavirus) is a truly global pandemic and has affected humans in all countries of the world. While humanity has seen numerous epidemics including a number of deadly ones over the last two decades (e.g., SARS, MERS, Ebola), the grief and disruption that COVID-19 will inflict is incomparable (and perhaps unimaginable). At the time of writing this paper, COVID-19 is still rapidly spreading around the world and projections for the next few months are grim and extremely disconcerting. The learnings from COVID-19 will also enable humankind to prevent such epidemics from transforming into global pandemics and minimize the socio-economic disruption.

In this preliminary work, our goal is to analyze the socio-economic disruption caused by COVID-19 in the United States of America, understand the chain of events that occurred during the spread of the infection, and draw meaningful conclusions so that similar mistakes can be avoided in the future. Though Twitter data has previously been shown to be biased [Morstatter and Liu, 2017], Twitter has emerged as the primary media for people to express their opinion especially during this time and our study offers a perspective into the impact as self-disclosed by people in a form that is easily understandable and can be acted upon. We summarize our main contributions below.

- We collect 530,206 tweets from Twitter between March 14th to March 24th, a time period when the virus significantly spread in the US and quantitatively demonstrate the socio-economic disruption and distress experienced by the people. Calls for closures started off with schools (e.g., #closenycschools), then moved on to bars and restaurants (e.g., #barsshut), and finally to entire cities and states (e.g., #lockdownusa). While these calls were initially mainly confined to the Seattle, Bay Area, and NY regions (e.g., #seattleshutdown, #shutdownnyc), they later expanded to include other parts of the country (e.g., #shutdownflorida, #vegasshutdown). Alongside, panic buying and hoarding escalated with essential items particularly toilet paper becoming unavailable in stores (e.g., #panicbuying, #toiletpapercrisis).

- We observe increased calls for social distancing, quarantining, and working from home to limit the spread of the disease (e.g., #socialdistancingnow, #workfromhome). To slow the exponential increase in the number of infections, people rallied for flattening the curve and staying at home for extended periods (e.g., #flattenthecurve). The challenges of working from home also surface in communications (e.g., #stayhomechallenge). With the passage of time, we see an increased fluctuation in emotions with some people expressing their anger at individuals flouting social distancing calls (e.g., #covidiots), while others rallying people to fight the disease (e.g., #fightback) and to save workers (e.g., #saveworkers).

- We group the hashtags into six main categories, namely *1)* General COVID, *2)* Quarantine, *3)* School Closures, *4)* Panic Buying, *5)* Lockdowns, and *6)* Frustration and Hope, to quantitatively and qualitatively understand the chain of events. We observe that general COVID and

quarantine related messages remain trending throughout the duration of our study. We observe calls for closing schools and universities peaking in the middle of March and then reducing when the closures go into effect. We observe a similar trend with panic buying. Lockdowns also have a significant number of tweets with calls initially being focused on closure of bars, followed by cities and then states. Tweets in the frustration and hope hashtag group have an overall increasing trend as the struggle with the virus mounts.

- We then present a linguistic analysis of the tweets in the different hashtag groups and present the words that are representative of each group. We observe that words such as *family*, *life*, *health* and *death* are common across hashtag groups. We observe mentions to *mental health*, a possible consequence of social isolation. We also observe solidarity for essential workers and gratitude towards them (#saveworkers).

Our study unearths and summarizes the critical public responses surrounding COVID-19, paving the way for insightful fine-grained linguistic and graph analysis in the future.

2 Data and Methods

In this section, we discuss our methodology for data collection from Twitter to investigate the socio-economic distress and disruption in the United States caused by COVID-19 during its early days.

2.1 Data Collection

We collect data using the Twitter search API. The results presented in this paper are based on the data collected from March 14 to March 24, 2020. We track the trending COVID related hashtags every day and collect the tweets in those specific hashtags. We repeat this process to collect a total of 530,206 tweets during this time period.

Table 1: Number of Tweets by Category

Category	Number of Tweets
General Covid	4,81,398
Quarantine	142,297
Lockdowns	14,709
Frustration and Hope	13,084
Panic Buying	10,855
School Closures	2,133

2.2 Hashtag Categories

We group the hashtags into six main categories, namely *1)* General COVID, *2)* Quarantine, *3)* School Closures, *4)* Panic Buying, *5)* Lockdowns, and *6)* Frustration and Hope to quantitatively and qualitatively understand the chain of events. We collect data on per day basis for the different hashtags as and when they become trending. Tables 1 shows the number of tweets in each category, while Table 2 shows a few representative hashtags in each category. We observe that the total number of tweets as grouped by hashtags is 664,476, which is higher than the total number of tweets. This is because tweets
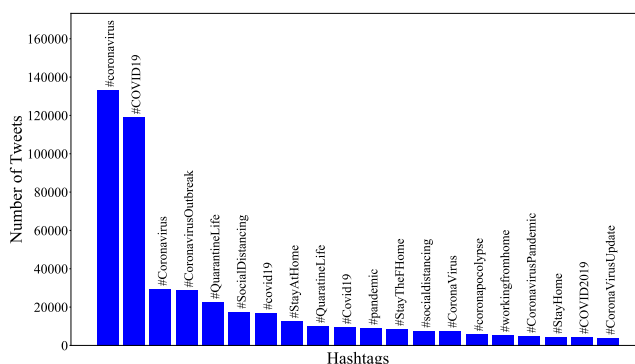

Figure 1: Number of tweets for the different hashtags

can contain multiple hashtags and thus the same tweet can be grouped into multiple categories.

1. General COVID: In this category, we group hashtags related to COVID related messages as it is the most discussed topic in conversations. This grouping is done by accumulating hashtags related to COVID-19.

2. Quarantine: Calls for social distancing and quarantines flooded Twitter during this outbreak. Communications centered around quarantines, working from home and flattening the curve to slow the spread of the virus.

3. School Closures: In this category, we collect data related to school closures. Before states decided to close schools, users on Twitter demanded the government to shut down public schools and universities. We collect data from a number of hashtags centered around this call for action.

4. Panic Buying: The spread of the virus also resulted in panic buying and hoarding. People rushed to shopping marts and there was a huge panic buying of sanitizers and toilet paper. This panic buying resulted in severe shortage of toilet papers around the middle of March, an issue that remained unresolved till the first week of April.

5. Lockdowns: With COVID-19 spreading unabated, lock downs of public stores, bars, restaurants, and cities began in many states. This resulted in a surge in lockdown related tweets.

6. Frustration and Hope: Emotions ran high during these times with people expressing anger and resentment towards those not abiding by social distancing and quarantine rules. Alongside, people also rallied to support workers working hard to keep essential services running. With the beginning of April approaching, many people started to worry about their next month's rent.

2.3 Gaps in Data Collection

Due to the data collection limits imposed by Twitter, we are able to only collect and analyze a portion of the tweets. Though we started collecting data as quickly as we conceived of this project, we were unable to collect data during the first week of March. Though we ran our script to collect data as far back as March 8, because of the way Twitter provides data, we obtained limited number of tweets from March 8 to March 13. Additionally, due to the rapidly evolving situation, it is likely that we have inadvertently missed some important hashtags, despite our best efforts. As is the case

Table 2: Hashtags by Category

| Category	Hashtags
General COVID	#covid19, #COVID19, #Covid19, #covid_19, #covid, #corona, #coronavirus, #Coronavirus,
Quarantine	#QuarantineLife, #quarantined, #staysafestayhome, #staytheFhome, #staythefuckhome,
School Closures	#schoolclosures, #closenypublicschools, #closenycschools, #closenyschools, #suny, #cuny,
Panic Buying	#panicbuying, #PanicBuying, #panicshopping, #PanicShopping, #panicbuyers, #toiletpaper,
Lockdowns	#lockdown, #Shutdown, #seattleshutdown, #ShutItDown, #shutdownnyc, #lalockdown, #sflockdown,
Frustration and Hope	#fightback,#saveyourlife, #COVIDIOTS, #COVIDIDIOTS, #WhenThisIsAllOver, #coronafighters,
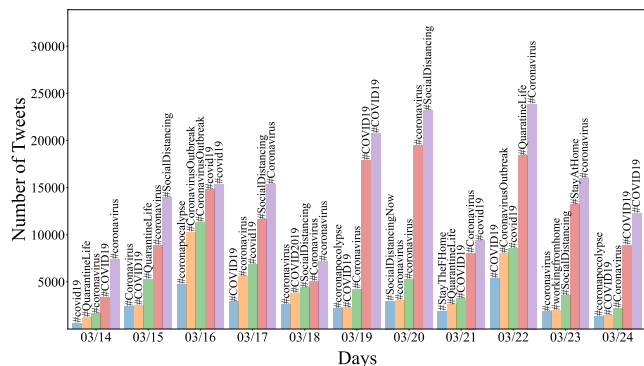
Figure 2: Most trending hashtags by day

with most studies based on Twitter data, we also acknowledge the presence of bias in data collection [Morstatter and Liu, 2017]. Having said that, the goal of this study is to provide a panoramic summarized view of the impact of the pandemic on people's lives and aggregate public opinion as expressed by them. Due to the nature of this study, we are confident that the results presented here help in appreciating the sequence of events that transpired and better prepare ourselves from a possible second outbreak of COVID-19 or another pandemic.

3 Data Analysis

In this section, we present observations and results based on our preliminary analysis of the tweets. We study the popularity of individual hashtags and investigate how the number of tweets in particular hashtag groups evolve over time. We also explore the term frequency for each hashtag group to understand the main points of discussion. Our analysis summarizes the critical public responses surrounding COVID-19 and paves the way for more insightful and fine-grained linguistic analysis in the future.

3.1 Temporal Evolution of Hashtag Groups

Figure 1 shows the top 20 hashtags observed in our data. As expected, we see that hashtags corresponding directly to COVID or coronavirus are the most popular hashtags as most communications are centered around them. We observe that hashtags around social isolation, staying at home, and quarantining are also popular. Figure 2 shows the most popular hashtags by date. Similar to Figure 1, we observe that hashtags related directly to COVID and social distancing trend most on Twitter. The figures and the number of tweets highlight how the pandemic gripped the United States.

We investigate the evolution of the number of tweets in various hashtag groups over time. To calculate the number of

tweets in each hashtag group, we count the number of mentions of hashtags in that group across all the tweets. If the tweet contains more than one hashtag, it is counted as part of all the hashtags mentioned in it. As the number of tweets for hashtag groups vary significantly, we plot the groups that have similar number of tweets together. Similar to Figures 1 and 2, we observe from Figure 3a that the total number of tweets in the General COVID and Quarantine categories are relatively high throughout the time period of the study.

Interestingly, from Figure 3b, we observe that panic buying and calls for school closures peak around the middle of March and then decrease as school closures and rationing of many essential goods such as toilet paper, cereal, and milk take effect. From Figure 3c, we see that calls for lock downs related to schools, bars, and cities peak in the middle of March. With the virus spreading unabated, we observe intense calls for lock downs of cities and entire states around the beginning of the fourth week of March, resulting in an increased number of tweets in this category. With passage of time, we observe people increasingly expressing their frustration and distress in communications, while some hashtags attempt to inject a more positive outlook.

3.2 Linguistic Word-Usage Analysis

In this section, we present results from a linguistic word usage analysis across the different hashtag groups. First, we identify and present the most commonly used words across all the hashtags. To construct the first group of common words across all hashtags, we remove the words that are same or similar to the hashtags mentioned in Table 2 as those words are redundant and tend to also be high in frequency. We also remove the names of places and governors such as New York, Massachusetts and Andrew Cuomo. After filtering out these words, we then rank the words based on their occurrence in multiple groups and their combined frequency across all the groups. We observe words such as *family*, *health*, *death*, *life*, *work*, *help*, *thank*, *need*, *time*, *love*, *crisis*. In Table 3, we present some notable example tweets containing the common words. While one may think that *health* refers to the virus-related health issues, we notice that many people also refer to *mental health* in their tweets as a possible consequence to social distancing and anxiety caused by the virus. We observe the usage of words such as *death* and *crisis* to indicate the seriousness of the situation. Supporting workers and showing gratitude toward them is another common tweet pattern that is worth mentioning. We plan to use these observations to guide our future exploration and fine-grained analysis.

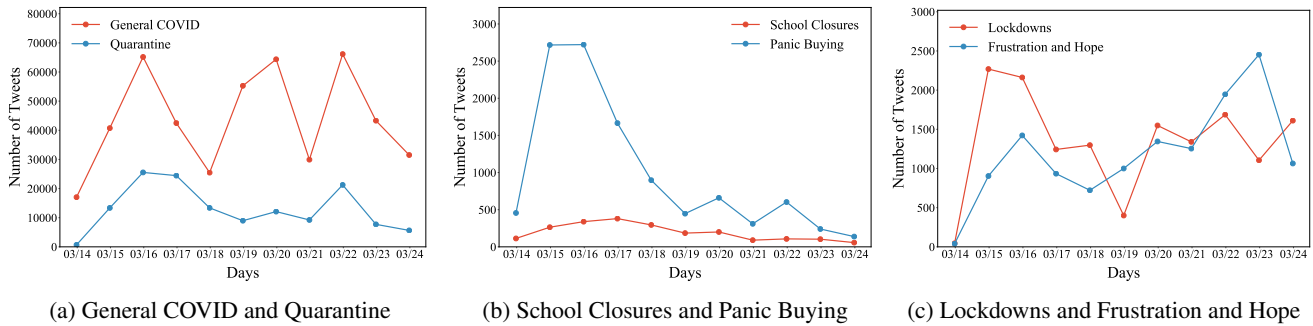Second, we present the most semantically meaningful and

(a) General COVID and Quarantine (b) School Closures and Panic Buying (c) Lockdowns and Frustration and Hope

Figure 3: Temporal Evolution of Tweets in Hashtag Groups

(a) General COVID word frequency

(b) School Closures word frequency

(c) Panic Buying word frequency
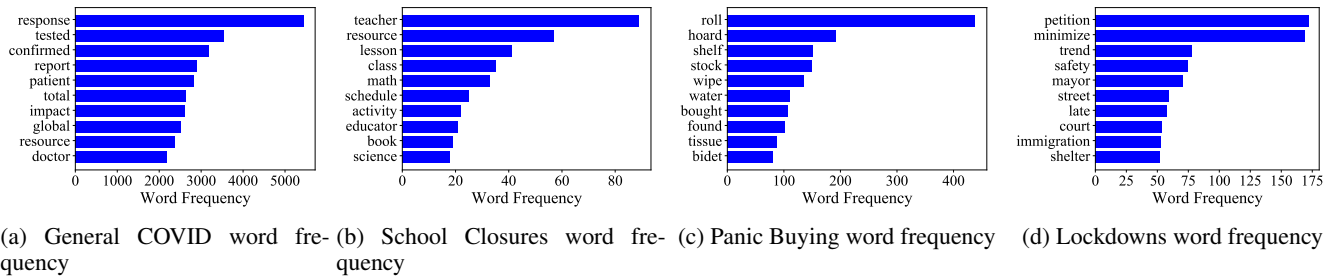
(d) Lockdowns word frequency

Figure 4: Word Frequencies

Table 3: Number of Tweets by Category

Example of Tweets
It's more important than ever to prioritize your mental **health**. You are not in this alone.
Has 13th century returned back to 20th century? Black **Death**. We must act very fast.
First, we take care of the **health** care and emergency **workers**. Then, we take care of whoever is in charge of keeping Netflix and Hulu running or it's going to get ugly #distancesocializing #coronavirus

uniquely identifying words in each hashtag group. To do this, we remove the common words calculated in the above step from each group. From the obtained list of words after the filtering, we then select the top 10 words. Due to lack of space, we only present results for four hashtag groups. Figure 4 gives us the uniquely identifying and semantically meaningful words in each hashtag group. In the General COVID group, we find words such as impact, response, resource, doctor. Similarly, for School Closures, we find words such as teacher, schedule, educator, book, class. The Panic Buying top words mostly resonates the shortages experienced by people during this time such as roll and tissue (referring to toilet paper), hoard, bidet (as an alternative to toilet paper), wipe, and water. Top words in the Lockdowns group include immigration, shelter, safety, court, and petition, signifying the different issues surrounding lockdown.

4 Related Work

In this section, we outline existing research related to modeling and analyzing twitter and web data to understand social, political, psychological and economic impacts of a variety of different events. Due to the recent nature of the outbreak, as far as we are aware there are no published research results related to COVID-19. Due to the space limitations, we only discuss social-media analysis work that are closely related to our work. Twitter has been used to study political events and related stance [Johnson and Goldwasser, 2016; Le *et al.*, 2017], human trafficking [Tomkins *et al.*, 2018a], and public health [Pérez-Pérez *et al.*, 2019; Dredze, 2012; Balani and De Choudhury, 2015; De Choudhury *et al.*, 2013; McIver *et al.*, 2015]. Several work perform fine-grained linguistic analysis on social media data [Zhang and Ramesh, 2018; Tomkins *et al.*, 2018b; Raisi and Huang, 2018].

5 Discussion and Concluding Remarks

In this paper, we studied Twitter communications in the United States during the early days of the COVID-19 outbreak. As the disease continued to spread, we observed panic buying as well as calls for closures of schools, bars, cities, social distancing and quarantining. We conducted a linguistic word-usage analysis and distinguished between words that are common to multiple hashtag groups and words that uniquely and semantically identify each hashtag group. In addition to words that represent the group, our analysis unearths many words related to emotion in the tweets. The research presented in this paper is preliminary and we plan to expand on our current research by performing topic modeling and expanding our linguistic analysis to unearth the main topics being discussed in the tweets. We also plan to conduct sentiment analysis to understand the extent of positive and negative sentiments in the tweets.

References

[Balani and De Choudhury, 2015] Sairam Balani and Munmun De Choudhury. Detecting and characterizing mental health related self-disclosure in social media. In *Proceedings of the Annual ACM Conference on Human Factors in Computing Systems*, 2015.

[De Choudhury *et al.*, 2013] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*, 2013.

[Dredze, 2012] M. Dredze. How social media will change public health. *Journal of IEEE Intelligent Systems*, 27:81–84, 2012.

[Johnson and Goldwasser, 2016] Kristen Johnson and Dan Goldwasser. All i know about politics is what i read in twitter: Weakly supervised models for extracting politicians' stances from twitter. In *International Conference on Computational Linguistics (COLING)*, 2016.

[Le *et al.*, 2017] Huyen Le, GR Boynton, Yelena Mejova, Zubair Shafiq, and Padmini Srinivasan. Bumps and bruises: Mining presidential campaign announcements on twitter. In *Proceedings of the Conference on Hypertext and Social Media*, 2017.

[McIver *et al.*, 2015] David J McIver, Jared B Hawkins, Rumi Chunara, Arnaub K Chatterjee, Aman Bhandari, Timothy P Fitzgerald, Sachin H Jain, and John S Brownstein. Characterizing sleep issues using twitter. *Journal of Medical Internet Research*, 17, 2015.

[Morstatter and Liu, 2017] Fred Morstatter and Huan Liu. Discovering, assessing, and mitigating data bias in social media. *Journal of Online Social Networks and Media*, 1:1–13, 2017.

[Pérez-Pérez *et al.*, 2019] Martín Pérez-Pérez, Gael Pérez-Rodríguez, Florentino Fdez-Riverola, and Anália Lourenço. Using twitter to understand the human bowel disease community: Exploratory analysis of key topics. *Journal of medical Internet research (JMIR)*, 21(8):e12610, 2019.

[Raisi and Huang, 2018] Elaheh Raisi and Bert Huang. Weakly supervised cyberbullying detection using co-trained ensembles of embedding models. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2018.

[Tomkins *et al.*, 2018a] Sabina Tomkins, Golnoosh Farnadi, Brian Amanatullah, Lise Getoor, and Steven Minton. The impact of environmental stressors on human trafficking. In *2018 IEEE International Conference on Data Mining (ICDM)*, 2018.

[Tomkins *et al.*, 2018b] Sabina Tomkins, Lise Getoor, Yunfei Chen, and Yi Zhang. A socio-linguistic model for cyberbullying detection. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 53–60. IEEE, 2018.

[Zhang and Ramesh, 2018] Yue Zhang and Arti Ramesh. Fine-grained analysis of cyberbullying using weakly-supervised topic models. In *International Conference on Data Science and Advanced Analytics (DSAA)*, 2018.