Analysing Security and Privacy Threats in the Lockdown Periods of COVID-19 Pandemic: Twitter Dataset Case Study

BIBHAS SHARMA, The University of New South Wales, Australia
ISHAN KARUNANAYAKE, The University of New South Wales, Australia
RAHAT MASOOD, The University of New South Wales, Australia
MUHAMMAD IKRAM, Macquarie University, Australia

The COVID-19 pandemic era will be remembered as a uniquely disruptive period that altered the lives of billions of citizens globally, resulting in new-normal for the way people live and work. With the coronavirus pandemic, businesses, governments, and educational institutes adapted to the “work or study from home” operating model that has not only transformed our online lives but has also exponentially increased the use of cyberspace. Concurrently, there has been a huge spike in the usage of social media platforms such as Facebook and Twitter during the COVID-19 lockdown periods. These lockdown periods have resulted in a set of new cybercrimes, thereby allowing attackers to victimise users of social media platforms in times of fear, uncertainty, and doubt. The threats range from running phishing campaigns and malicious domains to extracting private information about victims for malicious purposes. To this end, it is vital to analyse the impact of drastic transformations that were taken during lockdown periods on the security and privacy of users.

This research paper performs a large-scale study to investigate the impact of lockdown periods during COVID-19 pandemic on the security and privacy of social media users. We analyse 10.6 Million COVID-related tweets from 533 days of data crawling and investigate users’ security and privacy behaviour in three different periods (i.e., before, during, and after lockdown). Our study shows that users unintentionally share more personal identifiable information when writing about the pandemic situation (e.g., sharing building name, nearby coronavirus testing location) in their tweets. The privacy risk reaches to 100% if a user posts three or more sensitive tweets about the pandemic. We investigate the number of suspicious domains shared in social media during different phases of the pandemic. Our analysis reveals an increase in the number of suspicious domains during the lockdown compared to other lockdown phases. We observe that IT, Search Engines, and Businesses are the top three categories that contain suspicious domains. Our analysis reveals that adversaries’ strategies to instigate malicious activities change with the country’s pandemic situation.

1 INTRODUCTION

COVID-19 was first discovered in Wuhan, China, in December 2019, which altered the lives of billions of people around the world [18]. The virus has spread all over the world with over 404 million confirmed cases and approximately 5.8 million deaths worldwide by February 2022 [24]. To contain the virus, countries worldwide have had to take multiple measures such as achieving a high vaccination rate, introducing state or area lockdowns, and border closures. As these restrictions were imposed, several workforces transitioned to work from home while schools and universities transitioned to online learning. A survey conducted by Gartner, Inc reveals that 74% of the companies want to shift their employees to remote work permanently [9]. Meanwhile, global internet traffic saw a significant surge during these lockdown periods across different continents. More specifically, a traffic surge of 86% in Asia, 78% in Europe, 65% in North America, and 70% in South America was measured during the lockdown periods [36]. This situation is an aftermath of the increase in web-based activities, including social media networking, online learning, virtual meetings, and online events during the height of the pandemic.
Apart from the devastating impact of COVID-19 on people’s personal, social, and professional lives, the pandemic also paved the way for a series of cybersecurity threats for online users. There has been a substantial increase in a range of different cyber attacks such as phishing, ransomware, spamming, and malicious messaging [17]. Employees were abruptly forced to work from home without proper training and arrangements when most companies did not have the necessary infrastructure and plans for such a drastic change. In addition, only a small percentage of companies had cybersecurity policies in place [26]. Moreover, school children and their parents became more frequent online users, and most of them were not aware of cybersecurity threats and their impact. These reasons considerably increased the playing field of cybercriminals by providing them with more attack vectors. For instance, German companies suffered around 53 billion Euros worth of damages due to Cyberattacks as a result of working from home [19].

In addition, the problem became more devastating with the increasing use of social media platforms to share public and personal information related to COVID-19. According to one of the reports [16], social media users increased by 13.2% (+490 million) in 2020 and by 10.1% (+424 million) in 2021. It has led to a massive increase in user generated-data that possess various privacy threats, thereby making social media platforms an appealing target for organizations to aggregate such information for legitimate or malicious intent [3]. According to a report by United State’s Federal Trade Commission [7], in 2020, there have been losses up to 258 million dollars as a result of social media scams, while that number rose up to a massive 770 million in 2021. Compared with the 134 million dollar losses overall in 2019, we can clearly see the impact of COVID-19 on these crimes. A recent example of exploiting social media for financial fraud is the use of Twitter bots to trick users to make payments to illegitimate accounts using Paypal or Venmo [11].

To this end, this research aims to investigate the impact of lockdown periods during COVID-19 pandemic on the security and privacy of social media users. Furthermore, we also intend to examine the influence of social media networks (in this case, Twitter) in managing the COVID-19 pandemic based on human behaviour and sentiments. We analyse 10 Million COVID-related tweets which were posted on Twitter from 01-Jan-2020 to 21-June-2021 (533 days). For most of our analysis, the dataset is sub-divided under 2 bases: the pandemic stage (i.e., before, during, and after the lockdown) and the country wise (Australia, India, UK, and the US). This classification allows us to carry out a systematic study to identify and analyse commonly discussed topics, public sentiments, privacy and security risks. We also investigate how these aspects vary across different countries compared with the global trends. Moreover, looking at the infection rates (IR) of the countries gives more insights in relating it to social posts and people sentiments. In essence, this paper makes the following four main contributions:

**Collection and characterisation of large-scale dataset.** We collect (cf. § 2) a large dataset consisting of 10 million tweets from four different geolocations spanning over 533 days of timeperiod. We classify this dataset into three phases of the pandemic (i.e., before, during, and after lockdown). We first perform Hashtag analysis to identify the topics that people are mostly discussing on social media platforms about the pandemic. Our analysis indicates that supporting businesses, politics, and latest news/updates have been more frequent topics during all the stages of the pandemic. We also perform URL analysis on the tweets and find that users share social media URLs from Twitter, Facebook, Instagram and news and media URLs to propagate information about the pandemic.

**Perception analysis toward COVID-19.** We perform (cf. § 3) Sentiment analysis on the tweets to explore the people perception (i.e., emotions and feelings of people) during the three phases of the pandemic. Our
study shows that COVID-19 restriction rules such as social distancing received a high positive sentiment of approximately 70% from the public. Similarly, staying home received a positive sentiment of approximately 45% from the community, respectively. On the contrary, political discussions and death tolls have a highly negative sentiment of approximately 50% for all three phases. Moreover, we observe that the user sentiments directly relate to the IR of a region. For example, people show negative sentiments on the death toll and positive sentiments on the social distancing topics when the IR of a country is higher.

**Privacy risks exposure.** We investigate (cf. § 4) the trend of sharing private information on social media platforms during the pandemic. For example, we investigate whether or not people are more inclined to share their personal information, such as their names, addresses, or locations, during a lockdown. We use a probabalistic framework that quantifies the privacy of user tweets based on three privacy probabilities, i.e., Uniqueness, Uniformity, and Linkability (explained later in the section). Our results indicate that users' average privacy risk reaches 100% after posting three sensitive tweets. Moreover, the average risk of predicting a user with just 1 sensitive tweet is 94% (0.94) before the lockdown, and 95% (0.95) during and after the lockdown, respectively.

**Exposure to suspicious content.** Finally, we perform (cf. § 4) a security analysis on the social media tweets. We investigate the number of suspicious domains shared in social media during different phases of the pandemic. Our analysis reveals an increase in the number of suspicious domains during the lockdown compared to other lockdown phases. We also observe that IT, Search Engines, and Businesses are the top three categories that contain suspicious domains. Moreover, we notice that adversaries strategies to instigate malicious activities change with the country’s pandemic situation. For example, if a government has imposed a lockdown, people are more likely to watch and hear news from government agencies, allowing adversaries to design government look-alike malicious websites.

## 2 DATA COLLECTION AND CHARACTERISATION

### 2.1 Data Collection Methodology

We begin by presenting our methodology to collect and analyse COVID-related data from Twitter.

**Data Collection:** We use a dataset provided by the Panacea Lab [1] to collect the COVID-related tweets. The Panacea Lab contains approximately 730 million COVID-related tweets, which can be used for scientific purposes. Since Panacea Lab only provides Tweet IDs, we need to hydrate (i.e., extracting the original content of the Tweet such as tweet text, geo-location, timestamp, likes, comments etc.) those tweet IDs. We use the Twitter API [30] and Twarc [29] python library for that purpose. We run our data-collection framework on High-Performance Cluster (HPC)–with over 4,000 CPU cores with multiple compute nodes, each having 1TB of memory–at our institute. Over the period of two months crawling, from June 2021 to August 2021, we collect tweets spanning over 533 days from January 1, 2020 to June 21, 2021. We further filter the collected tweets based on selected countries, English language, and lockdown periods. The filtering process is explained in detailed below.

**Data Filtering.** Next, we filter tweets based on the Geolocation and Language. Specifically, we selected tweets from Australia, India, United States of America (US), and United Kingdom (UK) from the dataset. We selected India, US, and UK because they had the highest IRs and death tolls during the pandemic. On
Table 1. Breakdown of number (#) and percentage (%) of tweets collected for each country in three different stages of COVID-19 pandemic.

| Period | Australia | India | UK       | US       |
|--------|-----------|-------|----------|----------|
|        | #         | %     | #        | %        |
| Before | 95,416    | 17    | 350,638  | 20       |
|        | 1,030,106 | 21    | 1,260,757| 41       |
| During | 398,825   | 70    | 1,041,865| 58       |
|        | 2,441,744 | 51    | 925,964  | 30       |
| After  | 75,404    | 13    | 400,090  | 22       |
|        | 1,324,941 | 28    | 870,943  | 29       |
| Total  | 569,645   | 100   | 1,792,593| 100      |
|        | 4,796,791 | 100   | 3,057,664| 100      |

the contrary, we selected Australia because its strategy to eliminate/contain COVID was different from other countries. Australia imposed international and domestic border closure and state lockdown for a prolonged period of time. We assume this would provide some interesting trends in Australia as compared to the other countries. Next, we filtered out non-English tweets from the dataset of selected countries. That is because multilingual tweets can affect the accuracy of hashtag analysis (cf. § 2.2.3) and sentiment analysis (cf. § 3). For example, it has been previously shown by Boyd-Graber et al. [6] that the topics learned by the Latent Dirichlet Allocation (LDA) algorithm are language-specific when used with bilingual datasets. They used LDA against a dataset of English and German Tweets, and their results indicated that English and German tweets were clustered independently. After filtering non-English tweets from the selected countries, our dataset contained 10.22 Million tweets for further analysis.

2.1.3 Ethics Consideration: We obtain a publicly available dataset from Panacea Lab. Prior to data collection, we obtained ethics approval from our organisation's ethics board. Throughout data collection, we did not attempt to obtain the real identities of the participants via, for instance, a linkage study. We follow ethics guidelines [27] and do not use, track, or de-anonymise users from the collected dataset. The data collected was not released publicly. We did not store any identifying information other than the attributes such as user tweets, timestamps, and hashtags on our servers.

2.2 Data Characterisation

In this section, we discuss our findings after characterising COVID-19 tweets. We first discuss our findings on the hashtags analysis, followed by the discussion on the URL analysis. For hashtag analysis, we identify the most discussed topics related to COVID-19 during the pandemic and discern their variation with respect to IRs in certain locations during different stages of the lockdown. We examine the most widely shared URLs in tweets for the URL analysis and identify which types of websites people are frequently visiting. We also try to identify any relationships among user behaviour with the IRs of countries during different phases of the lockdown.

2.2.1 Data Statistics. Table 1 shows the number of tweets collected from each country in different lockdown periods. We collect a majority of the tweets from the UK and the US. Of the 10.22 Million tweets, 4.8 Million and 3.06 Million tweets are from the UK and USA, respectively. We collect 1.79 Million tweets from India, while a meagre 0.57 Million are from Australia. For Australia, 70% of the tweets are from during the lockdown period, while this number is 58%, 51%, and 30% for India, UK, and US, respectively. We have a lower number of US tweets because it has a shorter lockdown period as given by the stringency level. We have a similar distribution in the number of tweets during the three stages for all the other countries. Furthermore, the
hydrated tweets were stored as jsonl files at our HPC server. We use the following attributes from the jsonl file for our analysis: anonymised user IDs, timestamp of the Tweet, Tweet ID, Tweet Text, URLs, geo-location of the Tweet and hashtags. Figure 1 shows the distribution of tweets per user and the distribution of hashtags per tweet, respectively. Figure 1a clearly shows that most users have posted less than 20 tweets; however, a sheer amount of users have also posted more than 40 tweets. Approximately 0.7% of users have posted more than 100 Tweets, whereas approximately 50% of users in the dataset have posted exactly one tweet. Similarly, we see a 3,812,764 (37.3%) number of tweets with up to 20 hashtags in Figure 1b.

2.2.2 Data Classification: We classify the filtered dataset into three phases based on the country’s lockdown dates i.e., before lockdown, during lockdown, and after lockdown. This classification helps us analyse the variations in the public sentiments, security, and privacy risks across different pandemic stages. Moreover, it also helps us identify various trends across the pandemic and come to better conclusions on potential drivers behind them. We refer to these phases as “lockdown Periods” throughout the paper. To determine the lockdown phases (dates) for each country, we use the stringency level given by the COVID-19 stringency level dashboard [23]. The stringency index measures how strict the government restrictions have been in response to COVID-19. We select 65 (100 is the maximum) as our stringency index for determining the lockdown dates. We select this number by first checking the lockdown dates of the selected countries on various news articles. We then put these dates into the stringency website to get the stringency index. For most of the lockdown dates, the index was >=65, hence giving us a clear indication of the lockdown stringency index. Table 2 shows the lockdown periods of the selected four countries.

Table 2. Lockdown Periods and Infection Rates (IR)–the ratio of total number of COVID-19 cases to the number of days [34].

| Country | Before Lockdown | During Lockdown | After Lockdown |
|---------|-----------------|-----------------|---------------|
| Australia | 5 Mar - 20 Mar (2020) 63.5 | 21 Mar - 15 May (2020) 229.4 | 16 May - 1 Jun (2020) 10.4 |
| | 21 Jun - 6 Jul (2020) 80.1 | 7 Jul - 19 Oct (2020) 176.4 | 20 Oct - 5 Nov (2020) 11.8 |
| India | 24 Feb - 23 Mar (2020) 17.1 | 24 Mar - 31 May (2020) 2,754.7 | 1 Jun - 30 Jun (2020) 12,903.7 |
| US | 29 Feb - 28 Mar (2020) 4,392.8 | 29 Mar - 28 Apr (2020) 28,445.5 | 29 Apr - 27 May (2020) 22,688.7 |
| UK | 10 Mar - 25 Mar (2020) 1,059.6 | 26 Mar - 1 Jun (2020) 3,521.8 | 2 Jun - 17 Jun (2020) 1,033.6 |
| | 7 Oct - 22 Oct (2020) 17,941.1 | 23 Oct - 7 Nov (2020) 21,297.5 | 8 Nov - 23 Nov (2020) 22,281.1 |
| | 20 Dec (2020) - 5 Jan (2021) 43,326.4 | 6 Jan - 16 Mar (2021) 26,181.8 | 17 Mar - 1 Apr (2021) 5,146.5 |
We also calculate the average number of infections per day for each country and each period. We define this value as the **Infection Rate (IR)**, which is calculated by dividing the total number of COVID-19 cases for a specific period by the total number of days for that period. This value provides us with a high-level idea about the COVID-19 situation of a certain country before, during, and after a lockdown. The number of cases for each period was retrieved from the *Our World in Data* website[34]. We can observe a few interesting details in Table 2.

For example, let’s consider the first lockdown period between 21st March to 15th May in Australia. We can see that the IR is higher (229.38) during the lockdown than before the lockdown (63.5) and eventually reduces after the lockdown (10.41). It shows the impact of strict COVID containment strategies followed by Australia. The states imposed lockdown when the cases were rising and eased restrictions when the transmission was under control. In India, we can see the IR is still high even after the lockdown (2,754.68 during the lockdown and 12,903.7 after the lockdown), which indicates that they haven’t been able to get ahead of the virus and prevent community transmission. In the US, although we can observe a slight decline of the IR after the lockdown, the number itself (22,688.66) is very high and suggests that community transmission must have been happening. However, the restrictions seem to have slowed down the rate of transmission. In the UK, we can see some contradicting relationships between lockdown periods and IR. The first lockdown between 26th March and 1st June seems to have controlled the spread of the disease considerably, while the second lockdown seems to have only managed to slow the rate of transmission. After two lockdowns, it seems that the UK was a bit late to impose a third lockdown, resulting in an extremely high IR of 43,326.41 before the lockdown.

**Insight.** The restrictions seem to have helped control the transmission as the number has fallen to 26,181.82 during the lockdown. It has kept falling, which suggests the possibility of herd immunity—*takes place when a substantial population of a community becomes immune to a disease*—in the UK.

2.2.3 Hashtag Analysis. A hashtag is a metadata tag prefaced by a hash sign (#) and is used on microblogging and other photo-sharing websites to identify digital content on a specific topic. In Twitter, the hashtag indicates the topic associated with the tweet. The hashtags are also used to index keywords and help users follow a specific topic they are interested in. We analyse hashtags in our dataset to identify the widely discussed COVID-related topics during the lockdown periods and their relation with IRs in specific countries.

**Methodology:** We use K-Means clustering to group the tweets that contain similar hashtags. This algorithm identifies $k$ number of centroids for a given dataset and assigns every data point to the nearest cluster. For our analysis, we first selected tweets that contained hashtags. Next, we duplicated each tweet by the number of hashtags it contained. For example, if a tweet had 3 hashtags, we created two additional duplicates of that tweet. This approach gave us a dataset of 10.6 million tweets. We then processed this new dataset by removing URLs, mentions, punctuations, and stopwords\(^1\) using NLTK library [4]. Generally, a single hashtag is a concatenated text of multiple words with different semantics, e.g. *stayHomeSayNoToDrugs*. Therefore, we cannot extract lexical features of hashtags as they have significant noise. We argue that since a hashtag represents the main content of a tweet, we can use the main content to represent the hashtag. Therefore, we remove the hashtags from the selected tweets for this analysis and only consider the main content.

After the initial processing, we tokenised and lemmatised the tweets to extract numerical features from the text. Tokenisation is used to protect private data, while lemmatisation removes redundancy and converts the words into their lemma (the root word in vocabulary). We used Sklearn TF-IDF vectorizer [25] for this

\(^1\)Since all the tweets in the dataset are covid related, we removed words such as COVID, COVID-19, coronavirus to avoid redundancy.
task. Finally, we used the Sklearn library [25] to fit the features into a K-Means model with 15 clusters. We experimented with the number of clusters ranging from 5 to 35 and found 15 to produce the best accuracy. The cluster names were assigned manually after inspecting the top-10 words in each cluster. Some of the clusters having words with similar semantics were merged to eliminate redundancy, which left us with 13 hashtag clusters in total.

**Insight 1.** Our hashtag analysis reveals that people mostly talk about supporting businesses during the three stages of the pandemic, as shown in Figure 2a. We find approximately 7.2 Million hashtags related to supporting businesses. For instance, people are frequently using hashtags such as #fundraising, #charities, #our_work_is_our_identity in their tweets. It indicates that the economic disaster was also immense during the pandemic apart from the death and sickness. Tourism, which was one of the most profitable industries before 2020, was almost brought to its knees. People from small roadside sellers at tourist attractions to commercial airline pilots lost their jobs and main sources of income. Almost every business that thrived on close human interactions or large numbers of people, including salons, massage parlours, pubs, nightclubs, gyms, restaurants, and cafes, had to be shut down during lockdowns. This situation affected a lot of livelihoods and directly impacted basic human needs, making supporting businesses the most frequent topic of discussion during the pandemic.

**Insight 2.** We also observe that topics such as politics, latest updates, PCR testing, and lockdown, have been quite frequently discussed on Twitter. We found approximately 285K tweets related to politics with hashtags such as #BorisJohnson, #PM, #Trump during the lockdown. Similarly, we observe that the topic latest updates (with hashtags such as #LIVE, #WATCH, #currentaffairs) was tweeted approximately 54K, 250K, and 148K times before, during and after the lockdown, respectively. Another noteworthy insight is that the topic “Mask Wearing” has been discussed less before the lockdown. However, the proportion increases by 84.1%, i.e. from 2,547 to 15,986 during the lockdown for face covering.

**Insight 3.** Another interesting insight from hashtag analysis is that the frequency for most of the topics increased during the lockdown and slightly decreased after the lockdown, as shown in Figure 2a. This pattern is only different for the vaccination, and the face mask topics, which had kept rising even after lockdowns.
The reasons can be that the governments kept pushing people to get vaccinated and mandated face masks most of the time, even after lockdowns. On the contrary, staying home related tweets had declined considerably after the lockdown. That is quite reasonable as staying home is not a relevant topic after a lockdown.

**Insight 4.** From Figure 2, we can observe that the general trend of Hashtags compared to the frequency based on lockdown periods is quite consistent. An interesting insight is that hashtags relating to wearing masks have been discussed more during early 2021 than in 2020. This could be because people realised that wearing a mask is a very effective preventive measure to contain the spread of COVID-19. We can see a similar trend for staying home, where people have been posting these hashtags more during lockdown compared to other periods. The volumetric trend seems to be consistent for the rest of the hashtags.

**Insight 5.** When observing these trends for each of the countries considered in our paper (refer Appendix A), we notice that supporting businesses are the main topic in each of the countries during all the lockdown periods. Before the lockdown, common discussion topics are more or less the same for Australia, the UK, and the US. However, India seemed to have more discussions related to COVID-19 prevention (e.g., topics such as PCR testing, preventing spread, staying home). Considering India’s low IR (17.1) before the lockdown, we can assume that the people were extremely concerned about the virus, which may be due to the devastating news they were receiving from other countries. During the lockdown period, we notice that topics such as front-line workers and death toll are highly discussed in countries such as USA and UK. The extremely high IRs in these countries (Table 2), which subsequently caused an increase in death rates, must be the reason for this surge in topics. The trends we see in Australia are more consistent than other countries, resulting from its low infection rates throughout the pandemic. For further analysis, we refer readers to Table 5.

![Fig. 3. General Trend of Domain Categories over time and the top 10 domain categories](image)

2.2.4 **URL Analysis.** People use Twitter as a medium to share articles and resources from other websites. The restricted character length in tweets encourages someone to write a short text with a piece of news or an opinion and share some supporting material. As a result, we can find a large number of URLs in tweets. An analysis of these URLs, their domains and categories can give important insights on widely discussed topics at a particular period among the Twitter community. With the objective of further investigating global trends, we did an URL analysis on the Twitter dataset.
Methodology: To perform URL analysis, we first extract all the URLs from the tweets. The Twitter API provides an attribute `URL` which can be used to extract URLs from the tweets while hydrating. Twitter usually shortens URLs using its URL shortening tool, which causes all URLs to have a Twitter domain. Nevertheless, Twitter API also provides the `expanded_url` attribute to extract the original URLs. Using `expanded_url`, we collected a corpus of 6.95 Million fully resolved URLs for our study.

We then used the python `tld` library [2] to extract the domains of these URLs. For each domain, we use Fortiguard [8] to classify them to a specific category. Figure 3a depicts the top 10 domain categories we identified using the above technique.

Insight 1. The URL analysis provides a slightly different perspective on the global trends. According to Figure 3a, social networking related URLs from domains such as twitter.com, instagram.com and facebook.com have been mostly shared on Twitter. Users shared approximately 1.43 Million social urls during the lockdown and 793K after the lockdown. This suggests that overall social media usage across multiple platforms increased during the pandemic. However, as we only consider the domain category of the URL for our analysis, we do not examine the underlying content in those articles or posts. The topics of the shared articles can be anything, although we can assume that they are more or less similar to the results of our hashtag analysis.

The same limitation applies to the news and media related URLs. There are 953K and 500K news related URLs during and after the lockdown, respectively. For example, some of the most widely shared news related URL domains are https://subscribe.theepochtimes.com/ (during lockdown: 1307, after lockdown: 1309), https://theconversation.com/ (during lockdown: 1445, after lockdown: 965), and https://ncbi.nlm.nih.gov/ (during lockdown: 1441, after lockdown:963).

Insight 2. We also notice that URLs related to information technology (IT) were also frequently shared on Twitter, especially during the lockdown period (approximately 330K number of times). That suggests that COVID-19 has significantly transformed business operations by forcing organisations to switch to remote working, increasing the load on IT equipment and the network traffic. This transformation hence forced people to share IT related URLs frequently. Some of the IT-related URLs include https://apps.apple.com/ (during lockdown: 3110, after lockdown: 95), https://play.google.com/store/apps/ (during lockdown: 3469, after lockdown: 99), and https://dailym.ai/ios (during lockdown: 1436, after lockdown: 586). We observe that approximately 329K IT-related URLs were shared during the lockdown, followed by 200K after the lockdown. Other most frequent URLs’ categories include: business, government, & legal organisations, streaming media, health & wellness, entertainment, and education.

Insight 3. The number of Tweets containing News related URLs seems to have been the highest from August 2020 to October 2020, which falls during the lockdown phase (see Table 2). Other times Social Network related URLs has been shared the most, which can be observed in Figure 3b. That can be because people relied more on the News to get updated information about COVID-19 during lockdown phases. The other clusters have similar trends and are consistent with Figure 2a.

Insight 4. The URL domain categorisation for individual countries (see Table 5 in Appendix A) is mostly consistent with Figure 3a. For example, in the UK, the top 5 domain categories before, during, and after the lockdown are the same as the order of categories in Figure 3a. In India and USA, streaming related URLs seems to be more popular than government and law related URLs. Meanwhile, in Australia, we can see that News and Media and government related URLs have been shared more times than social media and business URLs. That is because Australia adopted a “zero COVID-19 strategy” and strictly implemented lockdowns, forcing people
to share government announcements and news alerts to keep up with the changes in restrictions. Another important observation is that Australians share Health and Wellness related URLs after the lockdown as compared to other countries.

3 PERCEPTION ANALYSIS TOWARD COVID-19

Next, by topic modelling and sentiment analysis of people tweets’ text, we illuminate people perceptions (i.e., feeling and emotion) the during different stages of the COVID-19 pandemic.

3.1 Topic Modelling

Topic modelling is the first step towards sentiment analysis. It is a clustering approach that helps in discovering some abstract topics in the dataset. For hashtag analysis, we only considered tweets with hashtags and duplicated them to represent multiple hashtags. However, for the sentiment analysis, we consider all the tweets from our dataset. We applied Latent Dirichlet Allocation (LDA) to generate 15 prominent topics. After obtaining the top 15 clusters, we manually inspected the top 15 words in each cluster and labelled them with a suitable name. We merged two similar topics related to politics hence ending up with 14 topics for our sentiment analysis.

As shown in Figure 4, most of these topics are consistent with the clusters we obtained in our hashtag analysis (e.g., vaccination, stay home, political, lockdown, and PCR testing). Some interesting topics identified additionally are sports, human rights, economic crisis, stock prices, and closing schools. Sports and stock prices are commonly discussed topics in twitter regardless of the pandemic.

Insight. At the start of the pandemic in 2020, almost all sporting events were cancelled. However, gradually they resumed in controlled environments (e.g. bio-secure bubbles). Moreover, the pandemic caused major changes in the business world, collapsing many businesses, meanwhile the valuation of some businesses skyrocketed. For instance, Video conferencing tool Zoom, pharmaceutical company Pfizer Inc., which developed an effective vaccine, and the e-commerce giant Amazon are some of the companies that had considerable increases in their stock prices as a result of the pandemic. Two other important topics we can observe in our topic modelling results are Human rights and Economic crisis. Due to some border restrictions, families were separated for prolonged periods in some countries. At the same time, many world leaders directly or indirectly mandated people to take vaccination to enjoy their freedom out of lockdown periods. Unvaccinated people even had to resign from their jobs in certain situations. Some people believe these actions violate human rights. Moreover, the closing of businesses due to lockdown periods, lack of seasonal and migrant workers due to border restrictions, and the huge decline in tourism have forced many countries into an economic crisis.
3.2 Public Sentiment during the Pandemic

For each topic, we performed sentiment analysis of the tweets using the VADER sentiment library [15] in Python. The main goal of sentiment analysis is to evaluate a body of text and comprehend its viewpoint. Usually, we measure this feeling by assigning the text a positive or negative number known as polarity. The sign of the polarity score is then used to determine whether the prevailing emotion is positive, neutral, or negative. Finally, we normalized the count of each sentiment in each topic to produce results shown in Figure 4.

**Insight 1.** We observe that the topic with the highest positive sentiments are social distancing, support businesses, and stay home. Social distancing has a positive sentiment of 69.85%, 70.5%, and 71.65% before, during, and after the lockdown periods, respectively. Similarly, support businesses has a positive polarity score of more than 50% for all periods, while this number is greater than 45% for stay home. It shows that people were happy with preventative methods and restrictions even though those measures limited their freedoms to some extent. The most negative comments seems to be towards the death toll (>50% for during and after lockdown) and politics (>50% for all periods). The increase in the number of cases and the resulting deaths were very upsetting to everyone worldwide. Moreover, the pandemic is a challenge for politicians as they have to implement strategies that were not welcoming from the public. For example, closing the borders affected families and businesses, making people angry with the governments. However, if open borders increase the number of cases and deaths in a country, citizens become angry with the government for allowing COVID-19 and its variants into the country. This suggests that these kinds of situations lead to negative sentiment for political tweets. Meanwhile, latest updates seems to have a highly neutral sentiment (~50% for all periods) along with closing schools (>45% for all periods). Topics such as vaccination, sports, lockdown, human rights, economic crisis, and stock prices seem to have more or less balanced between the positive and the negative sentiments.

**Insight 2.** When inspecting the sentiments of individual countries (please see Figure 11 in Appendix A), we observe similar trends with some quite noticeable results: The negative sentiment for the death toll in India is significantly lower (31.06%), while it is comparatively higher (57.31%) in the UK. One reason is that the IR is considerably lower in India than in the UK for the lockdown dates we considered in our analysis. Moreover, it can be the same reason for the UK’s highly positive sentiment (74.91%) for social distancing compared to other countries.

4 PRIVACY RISKS EXPOSURE ANALYSIS

In this section, we discuss our findings on privacy risks associated with COVID tweets that could lead to privacy leakages, such as sensitive information disclosure and user identification and tracking. Previous studies have shown that users share an immense amount of private information on the Internet through their web actions that mainly include web searches, social media posts with location details, photos and video sharing, and forum comments [21, 28, 33].

To analyse the impact of the COVID pandemic on the privacy of social media users, we quantify the privacy risks of tweets using the methodology given in [20]. In this study, the authors proposed a probabilistic framework that quantifies privacy risks based on three key aspects: uniqueness, uniformity, and linkability of the web data. Considering the results of this study and the generic nature of the model, we apply this framework to the Twitter dataset for our privacy risk analysis.
4.1 Privacy Threat Model

Our privacy risk quantification and estimation is based on a defined threat model. The model considers an anonymised dataset of tweets that do not contain any user identification, i.e., all the user identity attributes have been removed from the dataset. We assume an adversary as a third-party who has been given access to the dataset for non-malicious purposes (e.g., checking aggregated statistics). However, the adversary can analyse the tweets and identify the user based on their tweets. We assume that an adversary has sufficient resources to execute the privacy attack on the dataset.

The user identification from an anonymised dataset is possible using three different scenarios. 1. **Uniformity in Tweets** refers to a set of similar tweets posted by a user (e.g., posting about home quarantine after testing positive for the coronavirus), 2. **Uniqueness in Tweets** refers to a unique sequence of tweets posted by a user (e.g., continuously posting about air travel from one country to another during a pandemic), and 3. **Linkability in Tweets** refers to mentioning Personal Identifiable Information (PII) in the tweets (e.g., giving the location of COVID vaccination clinic). Our proposed threat model assumes that the continuous flow of information in the form of the above three scenarios could lead to user tracking and identification, even if the data is anonymised. In the next section, we define our privacy risk quantification method based on the threat model discussed above.

4.2 Privacy Risk Quantification Method

For our work, we define privacy risk as the probability of identifying social media users by learning their private or sensitive information through their tweets. The three key probabilities that are involved in risk quantification are: (1) **Probability of Uniqueness**: is measured as the non-likelihood of user’s tweets sequence being similar to tweets of other users such that the sequence is unique or distinguished to reveal the user’s identity. (2) **Probability of Uniformness**: is measured as the likelihood of a user entering the specific tweet (and thereby interested in the specific topic) based on the user’s previous tweet history. The more the user has entered a certain type of tweet, the more confidence in the inference that the user is interested in this topic. (3) **Probability of Linkability**: is based on how much PII available from user’s tweet data. PII could reveal the identity of a user and therefore allows linking the corresponding data to the user.

The overall privacy risk is measured as the joint probability of identifiability (uniqueness and uniformity) and linkability probabilities. In order to measure the above probabilities, we use the Hidden Markov Model (HMM) that represents probability distributions over sequences of observations. Let’s consider \( u_t \) represents a user and \( X_t \) represents a tweet at a time \( t \). Also, assume a sequence of events (i.e., tweets) by a user at time \( t \) is \( X_1, X_2, ..., X_t \), respectively. We train the HMM model using previous tweets of a user in order to predict the privacy risk of his current tweet. The tweets entered by a user becomes a node, and the probabilities of uniqueness, uniformity, and linkability are modelled in the HMM.

**Uniqueness** is modelled as transition probabilities in the HMM. Transition probability is a conditional probability of a tweet by all users given previous tweets sequence from all users. In HMM, edges contain the transition probabilities between nodes \( p(X_t|X_{t-1}) \). These transition probabilities are weighted by their confidence in terms of how many transitions have occurred, which is \( w_T = 1/count(X_t|X_{t-1}) \). Hence, the weighted transition probabilities are considered as, \( w_T \times p(X_t|X_{t-1}) \).
Uniformity is modelled as observation probabilities in the HMM. Observation probability is a probability of a tweet found in previous tweets history of different users ($u_i$), including the user whose risk is to be predicted (if available). In HMM, we model the observation probabilities as different users’ probabilities of the given tweet, $X_t$, found in previous tweet entries ($p(u_i | X_t)$). Again these probabilities are weighted by $w_O = 1/count(u_i | X_t)$ and then inversed (as more uniform a user is higher the privacy risk is and therefore lower privacy probability), i.e., $(1 - w_O \times p(u_i | X_t))$.

Linkability is measured from the prior probabilities of a user based on previous tweets that include PII (names, locations, and organisations). The privacy risks of user tweets that include PII are modelled in a separate HMM. For a given user $u_i$, the prior risk probability is calculated by getting the minimum privacy probability (maximum privacy risk) from all the paths in the PII HMM, which include nodes $X_t$ that contain an observation probability for the user, i.e., $p(u_i | X_t) > 0$.

The overall privacy probability of a user $u_i$ along a sequence of tweets $X_1 \rightarrow X_2 \rightarrow \ldots \rightarrow X_t$ is calculated as:

$$
p(X_1, \ldots, X_t | u_i) = \min(HMM_{PII} | u_i) \times w_T \times p(X_1) \times (1 - w_O \times p(u_i | X_t)) \times \prod_{x=2}^{t} w_T \times p(X_x | X_{x-1}) \times (1 - w_O \times p(u_i | X_x)),
$$

where $HMM_{PII} | u_i$ returns a list of privacy probabilities calculated from the PII HMM for all paths that include nodes where the user has an observation probability of $> 0.0$.

### 4.3 Privacy Analysis

We apply the above privacy risk quantification methodology on the Twitter dataset and analyze the results from the three aspects of uniqueness, uniformity, and linkability, and also present overall risk prediction results combining all three.
Before applying the quantification method, we first split the data into a 20-80 testing approach where 20% of the tweets dataset were used for testing, while 80% were used to train the HMM model. Furthermore, to reduce training time, we applied k-means clustering that partitions the training data into k clusters and then used a multi-processing technique to run each training cluster simultaneously. As mentioned earlier, the k-means algorithm helps group similar tweets based on the nearest mean (centroid). For our datasets, we selected 14 clusters based on our topic modelling (see Section 3.1). Results from each multi-processed cluster are then combined to create one training model. We use cosine similarity to find similar tweets.

**Insight 1.** Our results indicate that an average privacy risk reaches 100% (1.0 privacy risk) when a user enters 3 tweets for all three lockdown periods. Surprisingly, the above result holds true for most of the COVID-related topics. For instance, vaccination and lockdown topics reach 100% identification rate after posting just 3 tweets. Figure 10 in Appendix A shows the average privacy risk when users post 40 tweets on 14 different topics. We also illustrate some specific examples of tweets where the risk becomes 100% after entering 3 queries in Table 6 of Appendix A. Moreover, the average risk of predicting a user with just 1 sensitive tweet is 94% (0.94) before the lockdown and 95% (0.95) during and after the lockdown, respectively. Comparing our results with [20], we observe that COVID-related tweets have 70% higher privacy risk than normal web data. The higher privacy risk is perhaps because the quantification framework calculates risks based on three aspects, i.e., uniformity, uniqueness, and linkability. Even if a user does not have uniformity in his tweets, he might be identified through the unique pattern of tweets and vice versa. For instance, we can predict after 4 tweets of the user in Table 6 of Appendix A with user ID ‘168973’ that a person has a 6-year-old daughter and is currently stuck in the UK without her parents. Similarly, we observe that another user (with user ID ‘666231’) has a blood clotting condition and cannot have vaccination because of a pre-medical condition. We find similar cases for all the topics and observe that users can be identified through their unique tweet patterns. For instance, we discover that the user with ID ‘905643’ is a male whose wife is 33 weeks pregnant and is concerned about giving coronavirus to an unborn baby. Likewise, the user with ID ‘369225’ informs on Twitter that her daughter named Miha is coming to Dubai-UAE after getting a travel exemption.

**Insight 2.** Figure 5 shows the CDF of users with their predicted privacy risks in three lockdown periods. Before the lockdown, topics such as vaccination, lockdown, and PCR Testing, have a risk higher than 0.85 for more than 50% of users, while Stock Price, Human Right, and Economic Crisis has a prediction rate of 0.8 for more than 50% of users. During the lockdown, we observe that Death Toll and Economic Crisis have an average privacy risk of 0.95 for more than 50% of users, followed by Support Business and School Close topics with a 0.85 prediction rate for 50% of users. This data clearly indicate that people share more information regarding their personal situation during the lockdown. After the lockdown, we see that topics
such as Politics, Stay Home, and Death Toll have highest privacy risk with 0.95 prediction rate for 50% of users.

Insight 3. on Uniformity: We now discuss our results on the uniformity of users’ tweets during different lockdown periods. Before the lockdown, people are consistently discussing vaccination and PCR Testing, which results in an average privacy risk of 0.97 with just 1 tweet. For instance, we observe that a user enters the tweet ‘my son should be returning to #school today but @stocktoncouncil have withdrawn his transport with no plans to restart whilst there is #COVID19 whats the plan from’ twice, which makes her 97% identifiable. Similarly, during the lockdown, topics such as death toll and economic crises have been discussed consistently by the users making them 97% identifiable with just 1 tweet. After the lockdown, politics and vaccination have an identification rate of 97% with just 1 tweet. Figure 6 shows the average risk for uniform queries. Overall, our results indicate that users are 100% identifiable after posting 5 uniform tweets for all the topics and all the lockdown periods.

Insight 4. on Uniqueness: Figure 7 shows the results of posting unique tweets each time. Our analysis shows that around 95% of tweet sequences are unique and can lead to 100% privacy risk for all the topics before the lockdown periods. For example, we observe that out of 82,069 unique sequences during the lockdown, 81,859 unique Tweets are 100% identifiable. Finally, 48,927 unique sequences are 100% identifiable out of 49,011 after the lockdown state.

Insight 5. on Linkability: We now investigate the linkability of users’ tweets using their PII. We found few users who have PII information available in their tweets. For instance, a user in a Lockdown topic shared a tweet ‘As fate would have it, I was scheduled to fly to Cairo this evening (tickets were cancelled weeks ago). I haven’t been in one place for this long in over six years. Today, of all days, I wish the skies were fully open and I could go and see my family.’. Another user in Social Distancing topic entered PII query ‘Great family day out to the Chester Zoo today - great outdoors walk with socially-distanced measures in place all the way around the park. Congrats @chesterzoo for making it work so well within these COVID-19 times’. Figure 8 shows the average privacy risk for the queries having PII available for the three lockdown periods. We also present results without linkability information, i.e., we remove PII and evaluate the privacy risk for the same set of entries. Our results indicate that linking tweets with PII has a higher privacy risk compared to the tweets with no PII. For instance, before the lockdown, Vaccination topic has the minimum average risk of 98% for linkability, which reduces to 94% if we remove PII. Similarly, during the lockdown, Death Toll has 98% minimum privacy risk with PII and 95% without PII. After the lockdown, the Support Business topic, for example, has a 96% of minimum average risk with PII but reduces to 95% without PII. However, we found that tweets with or without PII can eventually reach up to 100% identifiability (uniqueness and uniformity) for all the topics and all the stages of lockdown, respectively.
5 EXPOSURE TO SUSPICIOUS CONTENT

In this section, we aim to investigate people’s exposure to suspicious content. We analyse the suspicious domains and the associated security risks for four individual countries (Australia, India, the US, and the UK) and three different periods of the COVID-19 pandemic (before, during and after lockdown periods). We use the URLs shared in tweets and utilise VirusTotal [31] to determine whether or not the second-level domains of those URLs are involved in any malicious activities. We also use the URL categories to analyse the most suspicious categories of second-level domains.

| Time Period | Total URLs | Unique URLs | Suspicious Domains | Unique Suspicious Domains |
|-------------|------------|-------------|--------------------|---------------------------|
| Before      | 1,061,853  | 586,787     | 15,411             | 140                       |
| During      | 3,771,312  | 2,219,095   | 66,441             | 259                       |
| After       | 2,116,832  | 1,254,057   | 35,139             | 196                       |

Table 3. Number of suspicious urls and domains in our dataset (VT Score >= 3)

**Methodology:** After extracting the expanded URLs as explained in Section 2.2.4, we removed duplicate URLs. It left us with 4.06 million URLs out of the 6.95 million total URLs. Next, we queried VirusTotal to get reports on each domain in our dataset. VirusTotal is an information aggregator, which presents a combined output of different antivirus products, file and website characterization tools, website scanning engines, etc. For a URL or a domain, we can obtain a report from VirusTotal. This report provides a number (positives), which indicates the number of tools that find the URL or the domain suspicious. Using this information, we calculate a parameter called VirusTotal Score (VTScore) for our analysis.

For every unique domain, we query VirusTotal to get all the reports between Jan 1st 2020 and Nov 6th 2021. Then, for each domain with positives >= 1, we take the sum of positives in all the reports and divide it by the total number of reports. We use the VT Score as a metric to identify how suspicious a particular domain is. The higher the VT Score, a domain is deemed more suspicious. Table 3 shows the number of suspicious domains we obtained for a VT Score greater than or equal to three. We draw the following insights from our analysis.

**Insight 1.** Firstly, we observe that the number of URLs shared during the lockdown periods is higher than before or after lockdown periods. It has caused a proportionate increase in the number of malicious domains. We found 345 unique suspicious domains overall (some domains are found in more than one period). For example, docsquiffy.com and peoples.it are two suspicious domains flagged during the lockdown periods, while ccp.it and comapncsr.com are flagged before and after lockdown, respectively. Meanwhile, buzzsawpolicies.com is flagged before and in lockdown periods. Table 4 is complementary to Table 3, where we show the number of suspicious domains and the number of unique domains out of them based on different VT Scores. We can observe that as the VT Score increases, the number of suspicious domains decreases considerably. For a VT Score greater than and equal to 55, which means the domains are extremely suspicious, we obtained 9 unique domains. cjso.org and ccp.it are flagged before the lockdown, and dudmc.com, geitpl.com, vietnam.travel, and itcslimited.com are flagged during the lockdown. Moreover, india.org is flagged during and after lockdown periods, while begadistrictnews.com.au and grantuk.com are flagged during all three periods.

We also analysed the domain categories of suspicious URLs and their distribution among Australia, India, the UK, and the US. We present our results in Figure 9, and we can notice some interesting traits there. The
most significant point is that we can see a domain category called Search Engines representing a considerable number of suspicious domains. We used the same methods described in §2.2.4 for the domain categorisation here. However, we do not see the Search Engines category as one of the widely shared domain categories in Figure 3a.

**Insight 2.** It indicates that URLs with a search engine related domain have a higher chance of being malicious compared to other domain categories. Moreover, we can see that Social Network related domains are not included in the top 6 most suspicious domain categories, even though it was the most widely shared domain categories according to our URL analysis. We can assume that the main reason for this is that people mostly share URLs from major social media networks such as facebook.com, instagram.com, and twitter.com, which are legitimate domains. IT and Business related domains contribute to a majority of suspicious domains. Since both these categories can be work-related most of the time, people tend to click URLs with these domains without much hesitancy. This behaviour can encourage malicious actors to use such domains to distribute malicious URLs.

**Insight 3.** When observing the distribution of suspicious URLs in individual countries, we can observe several interesting facts. One of the most significant observations is the high number of suspicious domains related to Government and Legal Organisations during the lockdown periods in Australia. As discussed earlier, due to how the Australian government handled the pandemic, individuals in Australia had to continuously rely on announcements from the government and legal authorities. This situation must have motivated malicious entities to act upon the government-related domains. If we consider India, we can notice two unique features in the distribution. First, we can deduce that the domain category of Blogs contains a noticeable portion of suspicious domains before the lockdown. We can only assume that blog related domains which are suspicious are widely shared in India during regular times, while COVID-19 has shifted people’s focus to other topics of interest. Second, we can see a significant increase in search Engine related domains during lockdown for India. During lockdown periods, people are mostly confined to their homes which can increase Internet usage, which could have caused this peculiarity. However, it is difficult to understand why it has not happened in other countries as well. For the UK, we can observe that News and Media have taken prominent places in the charts, while it is not the case with other countries. Being a country with high infection rates and many lockdown periods, we can assume that the people in the UK mostly relied on news and media related domains to get updated about the situation in the country. This situation may have resulted in attackers sharing more suspicious URLs in Twitter, which belong under that category. For the US, we cannot observe any significant traits in Figure 9.
6 RELATED WORK

Since the pandemic, several research works have focused on identifying user behaviour and perceptions using Twitter data. Boot-Itt et al. [5] identified three main topics of concern in Twitter users during the pandemic using topic modelling. The topics were COVID-19 emergency, COVID-19 control mechanisms and reports on COVID-19. Their sentiment analysis confirmed the common notion that people had a negative outlook toward COVID-19. Meanwhile, Hussain et al. [14] used Twitter data to assess the public opinion regarding the COVID-19 vaccine. Their research revealed more than 50% positive sentiments from the people in the UK and the US. Huang et al. [13] proposed Twitter data analysis as an efficient, cost-effective, and privacy-preserving method to assess human mobility dynamics during the pandemic. Their results suggested that Twitter data is capable of quantifying mobility dynamics in various geographical scales. Guntuku et al. [10] used Twitter data to analyse mental health and symptoms. Additionally, Visentin et al. [32] tried to identify the relationship of words, linguistics, styles, and emotions with privacy concerns and conspiracy theories in Twitter and how those elements contributed to the spread of such theories. To this end, they analysed tweets related to an Italian tracing app called “Immuni”.

A number of studies have also been conducted to analyse the privacy risk on social media platforms. Hoeisini et al. [12] analysed approximately 351K URLs on Twitter by modelling them based on different topics and performed a content-based analysis to determine differences between group messaging platforms such as WhatsApp, Telegram, and Discord. The study also analysed the level of PII exposure on the three platforms and collected over 34,000 phone numbers. Narayan et al. [22] discussed the possibilities of adversaries to de-anonymise social media datasets using different strategies. A more comprehensive study was performed by Masood et al. [20] that used Hidden Markov Model (HMM) to predict the privacy risk based on the probabilities of uniqueness, uniformity and linkability of user’s web data. Authors conducted experiments using AOL search queries dataset and Android application reviews dataset. The results show that with a minimum of only 10 sensitive web queries, a user’s privacy risk reaches 100%.

The security of social media users can be analysed with respect to different types of emerging security threats such as malware, phishing, spam email, ransomware, etc. Xia et al. [35] conducted research to identify and characterise COVID-19 themed malicious domains. Authors aggregated a dataset containing 4,500 malicious COVID-19 themed domains from a number of different sources. They differentiated the COVID-19 malicious campaigns based on the underlying network infrastructure such as subnet distribution, cloud IPs, geolocation, domain registration, WHOIS records and so on. They then constructed a network knowledge graph followed by clustering the nodes based on the relations in the graph (related IPs, name servers, etc.). The study concludes that the adversaries are rapidly exploiting COVID-19 to facilitate cyber-attacks.
In this paper, we carry out large-scale study on the impact of COVID-19 pandemic on the security and privacy of social media users. This study is first of its kind to comprehensively investigate the 10 Million tweets from various aspects that mainly include characterization, sentiment analysis, security analysis, and privacy analysis, respectively. Our study complements all the prior research and delivers new contributions to the knowledge on privacy and security in social media.

7 CONCLUSION AND FUTURE RESEARCH

COVID-19 brought many changes to the lives of each and every person on the planet. As most of these changes were unprecedented, the impact of most such changes is still unknown. However, this situation encouraged many researchers to go beyond their established bounds and engage in impactful research. Under these circumstances, we decided to conduct a comprehensive study on user behaviour on social media with the major objective of understanding privacy and security risks. We try to identify the main topics of discussion during the pandemic related to COVID-19 and the generic user sentiment towards them. In addition, we try to extend our analysis to examine the impact of different phases during the pandemic and different countries, their infection rates and COVID-19 related policies on our results. Hence our study consists of statistical, sentiment, privacy, and security analyses. All analyses are based on the three lockdown periods (before, during, and after) and consider Australia, India, the UK, and the US.

Our statistical analysis revealed that supporting businesses and politics are the most widely discussed topics on Twitter. At the same time, URLs related to social networks and news and media domains have been widely shared. At the same time, the sentiment analysis shows that people have a highly positive sentiment for COVID-19 preventative methods, while they display highly negative sentiments towards discussions on politics and death tolls. These sentiments seem to be impacted by the infection rates in certain countries as well. Meanwhile, the privacy analysis revealed that how people share more information about their personal circumstances on social media networks. Users who posted just 3 sensitive tweets become 100% identifiable. Finally, the security analysis showed that a major portion of suspicious URL domains belonged to IT, business, or search engines.

As for future work, there are a few aspects that we would like to extend this work. For example, we can extend the VirusTotal analysis to individual URLs instead of domains. It will provide an in-depth URL analysis and a better characterisation of suspicious URLs. In addition, we can utilise Gaussian distribution, maximum entropy Markov model, etc., to quantify the privacy risk instead of using the basic HMM model. We also aim to extend our work on identifying, characterizing, and analysing the impact of spreading rumors and misinformation on social networks about the pandemic.

REFERENCES

[1] Juan M. Banda, Ramya Tekumalla, Guanyu Wang, Jingyuan Yu, Tuo Liu, Yuning Ding, Ekaterina Artemova, Elena Tutubalina, and Gerardo Chowell. 2021. A Large-Scale COVID-19 Twitter Chatter Dataset for Open Scientific Research—An International Collaboration. Epidemiologia (2021).
[2] Artur Barseghyan. 2021. tld — tld 0.12.7 documentation. https://tld.readthedocs.io/en/latest/
[3] Ghazaleh Beigi and Huan Liu. 2018. Privacy in social media: Identification, mitigation and applications. arXiv preprint arXiv:1808.02191 (2018).
[4] Steven Bird, Ewan Klein, and Edward Loper. 2009. Natural Language Processing with Python (1st ed.). O’Reilly Media, Inc.
[5] Sakun Boon-Itt and Yukolpat Skunkan. 2020. Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study. JMIR Public Health Surveill (2020).
[6] Jordan Boyd-Graber and David M. Blei. 2009. Multilingual Topic Models for Unaligned Text. In CIKM. ACM, New York, NY, USA.
[7] Emma Fletcher. 2022. Social media a gold mine for scammers in 2021. https://www.ftc.gov/news-events/blogs/data-spotlight/2022/01/social-media-gold-mine-scammers-2021
[8] FortiGuard. 2021. FortiGuard Labs: Global threat research and response. https://www.fortiguard.com/
[9] Gartner. 2020. Gartner CFO Survey Reveals 74% Intend to Shift Some Employees to Remote Work Permanently. https://www.gartner.com/en/newsroom/press-releases/2020-04-03-gartner-cfo-survey-reveals-74-percent-of-organizations-to-shift-some-employees-to-remote-work-permanently
[10] Sharath Chandra Guntuku, Garrick Sherman, Daniel C Stokes, Anish K Agarwal, Emily Seltzer, Raina M Merchant, and Lyle H Ungar. 2020. Tracking Mental Health and Symptom Mentions on Twitter During COVID-19.
[11] Jessica Haworth. 2021. Social media scam: Twitter bots are tricking users into making PayPal and Venmo payments into fraudsters’ accounts. https://portswigger.net/daily-swig/social-media-scam-twitter-bots-are-tricking-users-into-making-paypal-and-venmo-payments-into-fraudsters-accounts
[12] Mohamad Hoseini, Philippe Melo, Manoel Júnior, Fabrikcio Benevenuto, Balakrishnan Chandrasekaran, Anja Feldmann, and Savvas Zannettou. 2020. Demystifying the Messaging Platforms’ Ecosystem Through the Lens of Twitter. In IMC.
[13] Xiao Huang, Zhenlong Li, Yuqin Jiang, Xiaoming Li, and Dwayne Porter. 2020. Twitter reveals human mobility dynamics during the COVID-19 pandemic. PLOS ONE (2020).
[14] Amir Hussain, Ahsen Tahir, Zain Hussain, Zakariya Sheikh, Mandar Gogate, Kay Dastphtopur, Azhar Ali, and Aziz Sheikh. 2021. Artificial Intelligence–Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study. J Med Internet Res (2021).
[15] Clayton J. Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text.
[16] Simon Kemp. 2021. Global overview Report. https://datareportal.com/reports/?tag=Global+Overview
[17] Navid Ali Khan, Sarfraz Nawaz Brohi, and Noor Zaman. 2020. Ten deadly cyber security threats amid COVID-19 pandemic. (2020).
[18] Yu-Chin Liu, Rei-Lin Kuo, and Shin-Ru Shih. 2020. COVID-19: The first documented coronavirus pandemic in history. Biomedical Journal (2020).
[19] Carolynn Look. 2021. Cyber Attacks Are Making Work-From-Home Expensive for Businesses. https://www.bloomberg.com/news/articles/2021-08-23/cyber-attacks-are-making-work-from-home-expensive-for-businesses
[20] Rahat Masood, Dinusha Vatsalan, Muhammad Ikram, and Mohamed Ali Kaafar. 2018. Incognito: A method for obfuscating web data. In WWW.
[21] Xuying Meng, Suhang Wang, Kai Shu, Jundong Li, Bo Chen, Huan Liu, and Yujuan Zhang. 2019. Towards privacy preserving social recommendation under personalized privacy settings. WWW (2019).
[22] Arvind Narayanan and Vitaly Shmatikov. 2008. Robust de-anonymization of large sparse datasets. In S&P.
[23] University of Oxford. [n.d.]. Relationship between number of COVID-19 cases and government response. https://covidtracker.bsg.ox.ac.uk/stringency-scatter
[24] World Health Organisation. 2021. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/
[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. ML Research (2011).
[26] Bernardi Pranggono and Abdulhussab Abdullah. 2021. COVID-19 pandemic cybersecurity issues. Internet Technology Letters (2021).
[27] Caitlin M Rivers and Bryan L Lewis. 2014. Ethical research standards in a world of big data. F1000Research 3 (2014).
[28] Oleksii Starov and Nick Nikiforakis. 2017. Extended tracking powers: Measuring the privacy diffusion enabled by browser extensions. In WWW.
[29] twarc. 2021. twarc. https://twarc-project.readthedocs.io/en/latest/
[30] Twitter. 2021. Twitter API. https://developer.twitter.com/en/docs/twitter-api
[31] VirusTotal. 2021. VirusTotal. https://support.virustotal.com/hc/en-us
[32] Marco Visentin, Annamaria Tuan, and Giandomenico Di Domenico. 2021. Words matter: How privacy concerns and conspiracy theories spread on twitter. Psychology and Marketing (2021).
[33] Hua Wang, Yongzhi Wang, Tarek Taleb, and Xiaohong Jiang. 2020. Special issue on security and privacy in network computing. WWW (2020).
[34] Our World. 2021. Our World in Data. https://ourworldindata.org/
[35] Pengcheng Xia, Mohamed Nabeel, Ilsa Khalil, Haoyu Wang, and Ting Yu. 2021. Identifying and Characterizing COVID-19 Themed Malicious Domain Campaigns. In CODASPY.
[36] Yiting Xia, Ying Zhang, Zhizhen Zhong, Guanqian Yan, Chiulim Lin, Satyajeet Singh Ahuja, Sosnath Bali, Alexander Nikolaidis, Kimia Ghobadi, and Manya Ghobadi. 2021. A Social Network Under Social Distancing: Risk-Driven Backbone Management During COVID-19 and Beyond. In NSDI. 217–231.
Table 5. Top categories of discussion in each country in different stages of the pandemic

| Country | Hashtag | # | URL | Before | Hashtag | # | URL | During | Hashtag | # | URL | After | Hashtag | # | URL |
|---------|---------|---|-----|--------|---------|---|-----|--------|---------|---|-----|--------|---------|---|-----|
| Australia | Support Business | 70,066 | News & Media | 17,940 | Support Business | 420,189 | News & Media | 119,728 | Support Business | 91,748 | News & Media | 22,941 |
| | Political | 4,246 | Social Networks | 15,156 | Political | 28,396 | Social Networks | 99,852 | Political | 5,108 | Social Networks | 19,216 |
| | Latest Updates | 3,230 | IT | 3,275 | Latest Updates | 19,506 | IT | 19,551 | Latest Updates | 3,903 | IT | 4,051 |
| | Community Cases | 2,013 | Business | 1,965 | Community Cases | 102,268 | Business | 12,642 | Community Cases | 1,987 | IT | 2,066 |
| | PCR Testing | 2,928 | Gov & Legal Org | 2,257 | PCR Testing | 11,194 | Gov & Legal Org | 14,899 | Latest Updates | 2,089 | Gov & Legal Org | 4,032 |
| | Latest Updates | 2,013 | Business | 1,965 | Community Cases | 102,268 | Business | 12,642 | Community Cases | 1,987 | IT | 2,066 |
| | Support Business | 208,252 | Social Networks | 27,489 | Support Business | 1,145,364 | Social Networks | 230,476 | Support Business | 453,118 | Social Networks | 74,828 |
| | PCR Testing | 17,026 | News & Media | 22,388 | PCR Testing | 118,381 | News & Media | 64,939 | Support Business | 3,903 | News & Media | 74,270 |
| | Latest Updates | 2,013 | Business | 1,965 | Latest Updates | 102,268 | Business | 12,642 | Community Cases | 1,987 | IT | 2,066 |
| | Community Cases | 12,383 | Gov & Legal Org | 1,905 | Community Cases | 118,381 | News & Media | 186,274 | Community Cases | 64,939 | News & Media | 74,270 |
| | Support Business | 522,433 | Social Networks | 201,185 | Support Business | 1,845,705 | Social Networks | 762,958 | Support Business | 1,011,203 | Social Networks | 404,736 |
| | Political | 36,245 | News & Media | 107,457 | Political | 125,763 | News & Media | 458,109 | Political | 86,625 | News & Media | 196,675 |
| | Latest Updates | 29,165 | IT | 45,078 | Latest Updates | 110,149 | IT | 180,600 | Latest Updates | 75,306 | IT | 107,818 |
| | PCR Testing | 23,542 | Business | 24,447 | PCR Testing | 104,482 | Business | 103,528 | Political | 65,568 | Business | 65,349 |
| | Prevent Spread | 18,024 | Gov & Legal Org | 18,438 | Prevent Spread | 71,975 | Gov & Legal Org | 44,717 | Prevent Spread | 46,578 | Gov & Legal Org | 44,439 |
| | Community Cases | 12,383 | Gov & Legal Org | 1,905 | Community Cases | 118,381 | News & Media | 186,274 | Community Cases | 64,939 | News & Media | 74,270 |
| | Support Business | 599,658 | Social Networks | 177,730 | Support Business | 256,671 | Social Networks | 340,683 | Support Business | 256,671 | Social Networks | 294,266 |
| | Political | 35,219 | News & Media | 104,524 | Political | 58,666 | News & Media | 191,389 | Political | 50,181 | News & Media | 195,409 |
| | PCR Testing | 24,924 | IT | 25,062 | PCR Testing | 34,068 | IT | 54,082 | Latest Updates | 36,058 | IT | 57,521 |
| | Prevent Spread | 14,478 | Business | 7,972 | Prevent Spread | 22,577 | Business | 20,762 | Death Toll | 24,549 | Business | 27,333 |
| | Latest Updates | 13,885 | Streaming | 6,830 | Latest Updates | 19,262 | Streaming | 18,293 | News & Media | 22,136 | Streaming | 14,793 |

A.1 Data Characterisation

In Section 2.2, we perform data characterisation by discussing the trends in the URLs and hashtags of tweets. We identify the top most discussed COVID-related topics in hashtags and URLs during all the three lockdown periods. In Table 5, we provide more detailed view on data characterisation by breaking down into country level. We can clearly observe that Support Business is the most discussed hash-tagged topic among all the four countries, followed by Politics and Latest Updates topics. These topics are also common across the lockdown periods. Similarly, URLs related to News & Media and Social Networks are mostly shared among all the countries and all the periods.

A.2 Perception Analysis Toward COVID-19

In Section 3, we analyse the relation of people sentiments with infection rates (IR) and COVID restrictions in a region. In general, we try to identify if there is an impact of social media tweets in managing the pandemic. In regards to this, Figure 11 illustrates the trend in people sentiments across four countries, for all the lockdown periods. Clearly, Death Toll has received the highest negative sentiments from UK and US. In general, we observe similar trends for the topics across all the countries. For instance, topic Social Distancing has received approximately 70% positive sentiments from all the countries.

A.3 Privacy Analysis

In Section 4, we quantify privacy risks against social media tweets and reveal interesting findings about user identification from just 3 sensitive tweets. In Figure 10, we show an average privacy risk across various topics and increasing number of tweets. It is clear from the figure that COVID-19 related tweets are capable of re-identifying users with at least 94% of privacy risk. Similarly, in Table 6, we provide few examples where users are mentioning some personal identifiable information (PII) in their tweets.
Fig. 10. Average privacy risk with the increasing number of web entries (i.e., tweets) in three different periods: Before, During, and After lockdown.

Fig. 11. User Sentiment for each Topic in Australia, India, and UK, and US, for all the lockdown periods.
Table 6. Examples of tweets’s tweet from all lockdown periods.

| User Anonymized ID | Tweets                                                                                                                                                                                                                                           | Topic                  |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| 345767             | A friend of mine was issued a $1,652 on-the-spot coronavirus fine this morning for entering Costco in Docklands as it was beyond the 5 km radius of her house. Sad! Be careful out there, folks."                                                                 | Stay Home (1 Tweet)    |
| 168953             | @JamesMelville Is your point that they should be worried about catching Covid-19 here? According to ONS data, 24.9% of people who died of covid-19 died in a care home. Versus 26.3% of people who died of covid-19 who had diabetes. It would seem a comparable blind spot. (Of course some of those may have been diabetic &amp; also in a care home but the point still stands.) " #COVID19 But I can go to my parents with my kids whilst Im sick with coronavirus right?" @richardhyland @montepen Selfishness vs altruism. I was tutted at last weekend for trying to avoid three people walking side by side! (My 6 year old daughter yelled coronavirus! at them after they passed I) | Stay Home (4 Tweets)   |
| 789654             | Im helping to fight COVID19. We only need 164 more people on the app to get a COVID estimate for Jefferson County. Please help by taking 1 min daily to report how you feel. You also get an estimate of COVID in your area.                                                                 | DeathToll (1 Tweet)    |
| 904365             | #covisham up 12 #COVID19 confirmed cases. #Bromley up 46 confirmed #COVID19 cases as of 22 Oct. #Bromley up 41 confirmed #COVID19 cases as of 22 Oct. Another positive #COVID19 cases in my son’s year at school. #Lewisham up 26 confirmed #COVID19 cases as of 3 Nov. Sadly, an increase of 1 death registered to 23 Oct. | DeathToll (4 Tweets)   |
| 167843             | We got a call from a UAE-based airlines in Frankfurt that our daughter Miha has been accepted on the flight and they are bringing her to Dubai’                                                                                                                  | LockDown (1 Tweet)     |
| 666231             | My blood naturally wants to clot. (Factor V Leiden is a genetic mutation. Yes, I am *lucky* enough to have blood clotting as my mutant powers.) In addition to being old and fat, now I learn another reason ‘Rona REALLY wants to kill me.’                                                                 | Vaccination (1 Tweet)  |
| 541209             | OK! Now I’m really angry. The current evidence from Israel is that the single dose is only about 30% effective at best. I know you think that’s plenty but you’re playing with our lives and not very successfully. So listen to the real science: 2 doses. ‘ And I am still feeling the effects of *suspected* Covid19 - 6 weeks later. Very very mild. I am 99% better. But this virus lingers people ( and I have an excellent immune system - despite other health issues), Genetic quirk could explain how pangolins can tolerate coronaviruses’ | Vaccination (2 Tweets) |
| 905643             | @NHSBSoCC, my wife is currently 33 weeks pregnant and we are now going to hospital on a weekly basis for check and monitoring. I am worried about giving my unborn child coronavirus due to these frequent visits. Would I now be able to get the vaccine as they are at high risk? ‘Obstetricians and Cynaecologists are recommending that women who are offered a Covid vaccine have if before they get pregnant. There is no need to delay pregnancy after the vaccine.’ | Vaccination (2 Tweets) |