Universiteit Leiden

Master Computer Science

The geographical distribution of Covid (mis)information on Twitter

Name: Ricardo L. Michels
Student ID: s1516744

Date: 09/10/2023

1st supervisor: Suzan Verberne
2nd supervisor: Frank Takes

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Abstract

The Covid-19 pandemic was accompanied by a large amount of misinformation. We attempt to find patterns in the geographical distribution of false news on twitter. We use an existing dataset of tweets about Covid news, labeled as true or false. We investigate geographical differences between widely discussed topics, and topics that were not widely discussed, as well as differences between true and false information. Location inference was used to greatly increase the amount of usable data. We find that the geographical distribution of marginally discussed false news is correlated with the geographical distribution of true news, whereas the geographical distribution of widely discussed false news is not correlated with the distribution of any other type of news. This suggests that widely discussed false news spreads in a different manner compared to other news.

1 Introduction

On the 31st of December 2019 the first confirmed case of Covid-19 was found in the Wuhan province of China [1]. On January 20 2020, the World Health Organisation declared it a “Public Health Emergency of International Concern” [2]. Since that time the virus has spread across the world, infecting people across the globe. As of April 4 2023, there have been 761 million confirmed cases, and 6.9 million deaths [3]. These numbers are likely an underestimation.

The pandemic was accompanied by a large amount of misinformation and disinformation that was spread online [4, 5]. Some people denied the existence of the pandemic, or stated that the severity of the disease was greatly exaggerated. Others spread misinformation about possible treatments, claiming for example that hydroxychloroquine could be used to treat the disease. Some claimed that a vaccine already existed at the start of the pandemic. And some conspiracy theorists stated that the pandemic did not exist and was caused by the 5G mobile network. The misinformation has likely worsened the pandemic, leading to increased numbers of infections. Many who believed in such conspiracies refused to cooperate with public health measures such as mask mandates and stay at home orders. A small percentage of people even refused to get vaccinated, citing misinformation as the cause of their concern [6]. Some of this misinformation is also directly dangerous to human health. For example, more than 100 people have died as a result of taking hydroxychloroquine to attempt to cure their Covid infections [7].

Understanding the ways in which this false news spreads is extremely useful in stopping the spreading of misinformation, and reducing its impact. This is true not only for misinformation relating to the pandemic, but for all misinformation.

A large amount of open source datasets became available as the pandemic started to become a global concern [8, 9, 10]. These datasets relate to a wide variety of topics, with applications related to diagnosis and treatment, surveys of ongoing research, and data on the social media response to the pandemic.

A lot of work has been done on understanding the way that misinformation spreads on online platforms. A lot of this research uses the concept of an information cascade. There are several ways of formally defining these cascades [11, 12], but in general these cascades contain one or more individuals who start, or first share a rumor or piece of information, and all individuals that are directly, or indirectly through others, influenced by that piece of information. On Twitter this is often slightly modified to be the user publishing the first tweet on a given subject, and any tweets that are directly or indirectly retweeting or replying to it. Alternatively it could be defined as the set of all tweets linking to the same URL, usually a news article.

Vosoughi *et al.* [13] found that misinformation spreads further on Twitter than true information, and that this misinformation also spreads more quickly. The authors also found that false rumours tended to be more novel compared to true rumours, i.e., the false rumours were more likely to be about topics not frequently encountered by the users in the two months before retweeting the false rumour. It was found that false covid information spreads more rapidly than partially true information on Twitter [14].

Some research has been done regarding the geographical spread of misinformation. Misinformation cascades have been shown to cross national borders [15]. The highest amounts of misinformation in the US was shown to occur in rural and conservative state [16, 17].

Our main research question is 'How can we detect differences between the geographical distributions of false and true news cascades on twitter?' We are also interested in differences between the geographical distributions of influential and uninfluential cascades. In order to answer this question we will utilize the CoAID dataset [8]. This dataset contains 283,531 tweets related to specific news articles surrounding the Covid epidemic, including both true and false news and information. These tweets were collected because they included a link to one of the articles, or because they replied to these first messages. All of these tweets are labeled true or false based on the veracity of the news articles. Geolocation data is gathered from the tweets, using coordinates for tweets where this information is present. We also attempt to infer the approximate location from the user profile for the large majority of tweets that contained no coordinates.

The number of tweets containing geolocation data is however only a small percentage (2–3%) of the total number of tweets [18]. Location inference may be used in order to increase the amount of available geographical data [19].

We analyse the geographical distributions of the true and false tweets, and plot these on the world map, and on the map of the contiguous United States. Finally within the US we also investigate whether the most frequently discussed topics are distributed in the same way geographically as the topics that were discussed less often for both true and false topics.

In order to answer our main research question, we will address five sub questions:

1. How quickly does Covid (mis)information spread?
2. What geographical differences exist in the spreading of misinformation?
3. How can we identify influential misinformation?
4. What differences exist between the geographical distributions of influential and uninfluential news?
5. Are the distributions of false and true news correlated with the US election results?

This thesis is structured as follows. In Section 2 we discuss the background and related work. Section 3 presents the data that we will use during this thesis. Section 4 presents the methods used to analyse the data. The results of the analysis are reported in Section 5. Section 6 presents a discussion of limitations of this thesis. Finally we present our conclusions in Section 7.

2 Background

In this section background and related work is presented about misinformation in general, misinformation cascades, location inference, geographical differences in the spreading of misinformation, and finally the CoAID dataset created by Cui *et al.* [8].

2.1 Misinformation

There are several different terms all relating broadly to the term false or fake news [13, 14]. Misinformation is generally any false information that is shared accidentally, while disinformation refers specifically to false information that is spread deliberately [14]. In this thesis we will use the term misinformation or false news for all of the false and questionable information. We make no claims as to the intent of the users spreading this misinformation.

Misinformation can be collected and categorised in several different ways. One method is to manually label statements or headlines. This method however does not allow for collection of large datasets which limits the use for large scale analysis. Another common method is to identify a number of reliable and unreliable sources of news and information, and collecting posts that link to one of these sources. The reliability of the source is then used to label the post as true/reliable, or false/unreliable [8, 4]. A method that is very similar to this is the use of fact-checking websites. When a statement is investigated by these websites, it is possible to search for posts containing that statement. Labeling these posts as true or false is then a simple task because these statements have been fact-checked before being included in the dataset [13, 14]. A similar process can be used to search for keywords or tags associated with factual discourse or misinformation [20, 17, 5, 15].

The large volume of misinformation related to Covid, even at the start of the pandemic, was quickly called the Covid-19 Infodemic [21, 22]. Cinelli *et al.* analysed the spreading of misinformation on several online platforms between January 1 and February 15, 2020 [4]. In general the number of interactions on true news grew at the same rate as the number of interactions on false news within this time frame. It was suggested that an increase in the amount of false news was likely a result of the increase in the amount of discussion surrounding Covid. The proportion of false news did differ between platforms however. The mainstream platforms contained relatively little false information. Of these Twitter contained the highest proportion, with 10% of posts and interactions regarding Covid containing misinformation. The far-right platform Gab contained a large proportion of misinformation. 41% of the posts relating to Covid were found to link to unreliable news sources [4]. On top of this, posts that linked to unreliable sources also received much more interactions from users compared to posts linking reliable sources.

2.2 Misinformation cascades

A common model used to investigate how quickly misinformation can spread on social media is a disease spreading through a population [4]. The simplest model consists of a Susceptible, and an Infected group, often called the SI model. Within the context of the spreading of misinformation users spreading false news are analogous to infected people exposing the uninfected users, i.e., those that have not yet accepted, or spread the misinformation [23]. For infectious diseases one useful tool is contact tracing: Tracing the origin of an infection, and determining when and how the disease was transmitted. Similarly one can imagine tracing where and how a false piece of information is shared, and who came into contact with that false information. This leads to the concept of an information cascade. On Twitter information cascades are defined as 'An unbroken re-tweet chain with a common origin.' [13] A graphical representation of a cascade can be seen in Figure Some authors only consider cascades with a singular origin, other definitions allow multiple origins within a single cascade. These cascades can be defined as a directed network, where each directed edge represents one user retweeting a tweet with false information by another.

By investigating the structure of these misinformation cascades, one can find interesting information about how this misinformation is spread online.

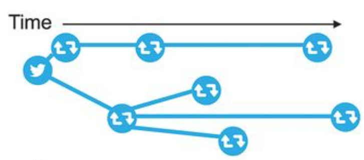


Figure 1: Example of a retweet cascade shown in Vosoughi *et al.* [13]. There is a single origin, and other users are added to the graph when they retweet or share the message.

Pierri *et al.* [24] were able to effectively classify and identify true and false information cascades on Twitter by utilising the differences in the structure between the true and false information cascades. The false information spreads more broadly and deeper than the true information cascades, and the users involved with the false news cascades were more closely connected to each other.

The spreading of misinformation was also investigated on Facebook. On Facebook the spreading of misinformation mostly occurred within 'echochambers' where misinformation is mostly shared between users that like and share similar posts [25]. Posts debunking false information are spread mostly within the science echo chambers [26]. A backfire effect was also observed with these debunking posts. When a user, who is mostly active in the false news echo chamber, comes into contact with a debunking post, then afterwards they tend to engage more actively with posts spreading misinformation.

Echochambers were also found on Twitter regarding Covid-19 (mis)information [9]. Users who were misinformed interacted with other misinformed users more frequently compared to the average amount of interaction between all users. The same is true for users who were not misinformed. Informed users were more likely to interact with other informed users, compared to the global average, but this effect was stronger with the misinformed users. Within the misinformation echochamber there was also a higher proportion of bots. A large proportion of the misinformed users were identified as being anti-vax, based on tweets posted before the pandemic.

Juul *et al.* [23] have found that structural differences between true and false cascades, as described in Vosoughi *et al.* [13], might be explained mostly by the difference in the size of the cascade. As an information cascade becomes larger, the other structural properties such as the maximum width, and the maximum depth also become larger. The conclusions in this paper were based in part on modelling the spreading of misinformation using several different infection models. No structural differences were found when comparing true and false cascades with the same size. When comparing the results between different models however, some structural differences remained.

In this thesis we have a lot of different cascades referring to the same news items or rumors. The collection of all cascades referring to the same news item or rumor, or the set of cascades referring to the same topic, is defined as the topic set.

2.3 Location inference

Several methods exist to obtain an (approximate) location from a tweet, and several types of location information exist within the data provided by Twitter [19, 18]. The most accurate locations can be obtained if the user has enabled geotags, these provide the coordinates of the user at the time a given tweet is posted. A distinction has to be made here between the location of the tweet, and the users home address as these are not necessarily the same. A user can also mention a place in a tweet. When a place is mentioned structured data with information about that place is also added to the data of the tweet. This data can include geographical coordinates or an address for example. These mentioned places again do not necessarily correspond to the actual location of a user. Tweets with location data in the form of coordinates represent only 2–3% of all tweets [18]. A user can also submit their location in a text field on their profile. These can provide a lot of useful information where the locations submitted can be parsed manually or programmatically. In their research Dredze *et al.* found that 56% of tweets had a non-empty user location [18]. There are however a number of users that submit nonsensical and joke locations that can not be traced to real world locations. Examples of these locations found in our dataset include 'Biebertown', 'Outer space', and 'In deep cover'.

In order to help extract the user location from the free form text field we use a so-called gazetteer [19]. A gazetteer uses a large dictionary to label pieces of text. Within the context of geographical locations, this is a large list of the names of locations, and for each location additional information may be provided such as the country, state or province, etc. The gazetteer can also contain multiple synonyms for the same location (e.g. 'The big apple' is labelled as New York

City). This is a method that is easy to implement. It also conveniently helps to filter joke or nonsense locations such as those mentioned above, as these will not be included in the dictionary of the gazetteer. A gazetteer does have the problem that some locations around the globe share the same name, e.g. London in Great Britain, Canada, and several locations in the USA, among others. It also relies on a complete list of synonyms, and if a synonym has not yet been included in the gazetteer, then locations might be missed.

It is also possible to extract information about a user based on the locations of users that they frequently interact with, and the followers of a user. A user is more likely to interact with users from the same geographical are [19]. This method requires that data is collected about the followers and/or interactions of a user. It might therefore not be suited for all purposes. Locations can also be based on context clues found within the text of tweets, using for example methods like a gazetteer. A location referenced in a tweet is not necessarily the location of the user itself however, and this method might be more suitable to investigate conversation about certain locations.

2.4 Geographical differences in the spreading of misinformation

Research has also been done on some aspects of the geographical spread of misinformation. Sharma *et al.* [15] were able to identify misinformation cascades spreading beyond national borders on Twitter, showing the geographical location of all tweets in a few examples. Their work was concerned mostly with identifying different types and topics of covid misinformation. They analysed for example differences in the sentiment between different topics, and changes in these sentiments through time, as well as determining what topics received the most attention in different countries. The volume of tweets mentioning countries in relation to Covid was found to be correlated with the case numbers [5]. An increase in the volume of tweets mentioning a given country would precede an increase in case numbers by 2–4 days, suggesting that the tweet volume could be predictive of an increase in confirmed cases.

When observing the geographical spread of misinformation on Twitter within the United States of America the highest volumes of misinformation per capita were found in rural states [16]. The authors identified the demographic most likely to spread misinformation by investigating the demographics of the counties where the highest number of misinformation tweets per capita were recorded. The number of tweets discussing topics related to Covid-19 were again shown to be correlated with case numbers, even when controlling for the population size in a state.

Erdemandi *et al.* investigated how different political and religious indicators in tweets was correlated to the presence of misinformation in those tweets [17]. Conservative users were found to be more likely to spread Covid misinformation, whereas Liberal users spread less misinformation. Religious users were also more likely to spread misinformation, although that effect was not as strong compared to the political ideology of the user. Misinformation was also shown to be concentrated in US states that skew more conservative.

All of these papers use geolocation data attached to the tweet. Some have elected to only include tweets containing these geolocations for some or all of the geospatial analysis [16, 5]. Others have also used user reported data in their analysis [17, 15].

2.5 CoAID dataset

The authors of CoAID [8] use an automated method for classifying tweets as true or false news based on the reliability of the original source. A number of reliable news websites, media outlets, and fact checking organisations were scraped to obtain the titles, content, and keywords for 200 false news articles, and 3500 true news articles. Tweets were then collected by searching for tweets containing the title of one of these articles. If the article was false, then the tweet is labelled as spreading false information, and conversely if a tweet mentions a true news article, then the tweet is labelled as true. This method allows for collecting a large amount of tweets, and the labeling of true and false is unambiguous. This method however may mislabel some tweets, as for example any tweets citing news articles with the intent of correcting or fact checking false information will still be labelled as spreading this false information. The CoAID dataset has been used in several instances in order to train classifiers to identify false news[27, 28, 29].

3 Data

The data used was the CoAID dataset [8]. This dataset contains the TweetID of tweets that were automatically collected and classified as true or false news based on news articles referenced in these tweets. The authors of the CoAID dataset obtained news articles from reliable news websites, fact checked by separate organisations. They then searched for any tweets sharing these reliable news articles. These tweets, and the replies to them were given the label true news. Links to unreliable news articles were found by using fact checking websites, and any tweets sharing these false news articles, and replies to them, were labelled as false news. The complete set of all tweets and replies about a given topic will be called a topic set in this thesis. This is similar to, but distinct from a cascade as described in Vosoughi *et al.* [13] A cascade is 'a rumor-spreading pattern that exhibits an unbroken re-tweet chain with a common, singular origin.' The topic set consists of multiple cascades about a given news article or claim. It contains all origins of the cascades, and all replies to these origins. It does not contain information about the cascade beyond the first replies. In other words, each cascade is stored up to a depth of 1. As such, this dataset gives an indication of the volume of tweets related to certain topics. It can not provide a lot of information about the spreading of a single cascade, as the data is only collected about the first link in the chain.

We collected all tweets still available from Twitter with the Twarc [30] package using the TweetIDs in the CoAID dataset. In Table 1 we show the number of TweetIDs in the dataset, and the number of tweets that were actually retrieved, as well as the percentage of tweets missing. We find that there are more tweets missing in the false news category, i.e. tweets spreading misinformation, compared to the tweets about true news. Tweets concerning false news also represent only a small fraction of the total dataset.

Figure 2 shows how many tweets were available for each of the 20 countries with most available data. The country with the most data available by far is the USA.

Table 1: Number of TweetIDs in the CoAID dataset, number of tweets retrieved during this study, and the percentage of tweets that were no longer available at the time the tweets were retrieved.

	Tweet IDs	Retrieved	Missing
True news	264,758	243,192	8.1%
False news	18,781	16,418	12.8%
Total	283,531	259,610	8.4%

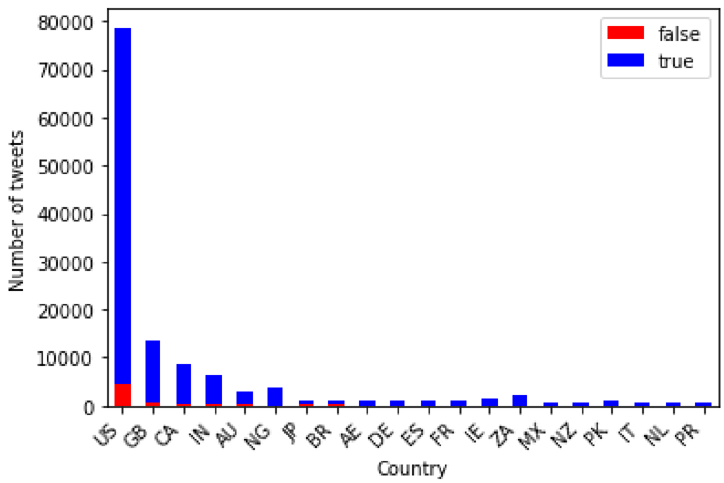


Figure 2: Number of tweets in the most prevalent countries in the CoAID dataset. This figure includes the locations obtained during the geocoding process described in section 4.

4 Methods

In this section we will cover the methods used to analyse the data described in Section 3. This follows roughly the order of the sub questions posed at the end of Section 1. First we cover the analysis of the speed at which (mis)information spreads, followed by the method used for location inference, the visualisation of geographical distribution of true and false news, analysing variables associated with each topic in order to identify more influential topics, visualising geographical differences in the distribution of influential and uninfluential news, and the correlation of these differences with the results of the US presidential election.

4.1 How quickly does Covid (mis)information spread?

In order to analyse the spread of the different topics in the dataset we first normalise the timestamp of each tweet to the number of days since the first tweet in the same topic set. The length of time that the topic is discussed is analogous to the cascade length described by Vosoughi *et al.* [13].

We analyse the differences between the spread of true and false news by counting the number of tweets and replies per day for the largest topic sets.

4.2 Location inference

Obtaining the location of a tweet was done using two different methods. A small number of tweets contained a location tag themselves in the form of coordinates.

Table 2: Number of tweets containing location data, separated by the method used to obtain these locations. For each source of location data the number of tweets using that method is shown. We list the percentage of the total dataset that used that source, and the share of each method in the obtained location data.

Origin	Tweets	% of all tweets	% of location data
coordinates in tweet	5,523	2.1%	3.5%
location text field	150,413	57.9%	96.5%

tweets with location	155,936
tweets without location	103,674

In this case, the location of the tweet was taken to be the the true location. This provided the location for 5,523 tweets.

Most tweets did not contain a location. In this case the user profile of the author of the tweet was used to obtain more information. Twitter provides a text field for users to enter their location. However, this text does not necessarily refer to real world locations, as a user can enter any text they wish. In our data for example, we find the locations ‘In deep cover’, ‘Vampire Castle’, and ‘State of Confusion’, among many others.

Most users have some text in this field, 106,159 out of 144,872 users, which is 73.3%. In order to obtain the geolocation for these users we compare the location text field to the Geonames database, and if this field is recognized as a place name, then that place is assumed to be the user’s actual location. If the text field is not a valid place, then no location is returned.

This process yielded a usable location for 86,585 users, 81.6% of the users with a non empty location field. The Geonames database was accessed using the Geocoder package [31]. This step increased the amount of available data by nearly two orders of magnitude. The number of tweets with location data, and with what method this data was obtained can be found in table 2.

Based on the location inferred from either the tweets or the user profile the tweets are separated based on the country of the user. Countries with fewer than 100 tweets were not considered for this analysis.

Visualising the geographical differences in the prevalence of false news. The ratio of false to true news is a direct measurement of the prevalence of false news in a country, compared to the total number of tweets observed. The ratio of false and true news tweets is calculated as $\frac{false}{true+1}$. This method is however very sensitive to small differences in countries where the amount of data is limited. We designed a separate metric called the false impact score in order to limit this sensitivity, and to give a better estimation of the global distribution of false news. For this metric the ratio of false news described previously is multiplied by the $\log_{10}$ of the total number of tweets in a country:

$$\frac{\mathit{false}}{\mathit{true}+1}\log_{10}(\mathit{total}) \tag{1}$$

Both of these metrics were plotted on the world map, with the colour of each country corresponding to the false news ratio or false impact score. The data was also plotted over time, with a map for each month represented in the dataset. The same techniques were used to generate a map of the contiguous United States. Here the metrics are calculated for each individual state, instead of for a whole country.

An influential cascade of true or false news is a cascade that spreads further, faster, and longer than other cascades. Identifying cascades or topics that have the potential to spread more virulently compared to other cascades will be very useful in halting the infodemic. The largest day to day increase in the number of tweets and replies are plotted against the total number of tweets in the topic set. The time between the first and last tweets in the dataset is also plotted against the total number of tweets. For all of these graphs the true and false topics are coloured differently.

A pearson correlation was performed on the variables shown in these graphs. The resulting r value will be between 1 and -1. Values close to 0 indicate that the variables are uncorrelated. Values close to 1 indicate a perfect positive correlation, and values near -1 indicate a negative correlation.

4.3 Spread of false and true news

In order to analyse the differences in how true and false news spread in this dataset, we gave each individual news story its own topic id, and then split the tweets according to the topic the tweet was associated with. Each topic set has one unique truth value, either true or false, according to the reputability of the news organisation that published the article shared in the first tweets of the cascades. All tweets within a given topic set share the same truth value, regardless of the contents of the tweets, e.g. a tweet debunking misinformation would still be classified as false. The timestamp of each tweet was then normalised to be the number of days after the first tweet in the topic set, in order to compare the virality of the news more directly. The largest topic set in the dataset was 510 tweets, and was in the true news category, the largest false news topic set was 209 tweets.

4.4 Geographical differences in influential topics

The overall goal of this thesis is to find whether there exist geographical differences in influential misinformation topics. In the previous section we have attempted to find a metric to label a misinformation topic as influential. We elect to use the total number of tweets in a topic set for the rest of this paper.

Our approach to identify geographical differences in influential misinformation topics consists of several steps.

1. separating the dataset into influential and uninfluential topics.
2. counting the number of tweets present in the influential and uninfluential topics per state.
3. transforming the counts of the tweets per state to a measure of relative abundance.
4. comparing the relative abundance of influential and uninfluential tweets in each state.

Topic sets were ranked based on the number of tweets in that topic in order to separate influential and uninfluential topics. To assemble the set of the most influential topics we iteratively add all tweets from the largest topic until at least 40% of tweets are included in the set of tweets belonging to the most influential topics. Selecting the tweets belonging to the least influential topics is done in a very similar way, where iteratively all tweets of the least influential topics are added until 40% of tweets are included in the set of tweets belonging to the least influential topics. Including the top and bottom 40% of data was done in order to have a decent amount of data for most states. If more data is available then a smaller subset of tweets could be used here in order to separate the two different groups more clearly. For true and false news these sets were generated separately. This resulted in four different categories: Influential true news, Uninfluential true news, Influential false news, and Uninfluential false news.

In order to analyse differences in the geographical distributions of these four categories, we count the number of tweets originating from each state for all four categories. The number of tweets found in a state varies by some orders of magnitude, so instead of raw counts, we use the we represent the volume of tweets in a state as $V_s = log_{10}(n_s)$ for our analysis. Here n_s is the number of tweets and replies in the state s for any given category of news (e.g. influential true news). The distribution of these $\log_{10}$ values resemble a normal distribution more closely making further analysis simpler.

Each of these four logarithmic distributions are then transformed to a standard distribution with a mean of 0, and a standard deviation of 1. This is done with the formula $(V_s - \bar{V_s})/\sigma_{V_s}$ where V_s is the aforementioned representation of the volume of tweets in a given state. $\bar{V}$ is the mean of that value across all states within that specific category, and σ_V is the standard deviation of that value across all states within a category. Transforming the distributions in this manner allows us to directly compare the distributions, even between the true and false categories, where a large difference in the amount of available data exists.

Each value calculated in the previous step now represents a relative abundance of tweets from a given state in one of the four categories.

The relative abundance is highly correlated to the population in a state. States with a high population have more tweets in each of the four categories compared to a state with a low population. Because of this, the difference between two different distributions is more informative for our purposes.

If the tweets are distributed in the same way for both categories, then the difference between values will be close to 0 for all states. A positive or negative value shows that there is a difference in the abundance of tweets between the two categories in a state.

Subtracting the distribution of the least influential news from the distributions with the most influential news results shows where the influential news has spread more relative to the uninfluential news. A positive value for a given state means that influential news was more abundant compared to uninfluential news. This is done with the formula $\delta_s = IT_s - UT_s$ where IT is the distribution of important true news, UT is the distribution of unimportant true news, and s is the index of a state. This is also done for false news: $\delta_s = IF_s - UF_s$. where IF and UF represent the important and uninfluential false news respectively. The only difference here is that we subtract the false news distributions from each other.

Subtracting the distribution of influential false news from influential true news $IT_s - IF_s$ indicates where false news spreads more readily compared to true news.

4.5 Correlating abundance of false and true news with the results of the US presidential election.

The Covid pandemic was an important topic in US politics, with a number of Republican politicians actively downplaying the severity of the pandemic, and spreading different types of misinformation. Because of this we were interested if the presence of misinformation was correlated with the political leaning of a state. We correlated the relative abundance of all of the previously mentioned categories to the results of the 2020 US presidential election, specifically to the percentage of votes cast for the Republican party. The correlation was performed using a standard function in the NumPy library. A 95% confidence interval was also plotted. This area contains the true best fit model with 95% certainty. This was achieved by methods described in [32, 33]

5 Results

In this section we will present the results of applying methods described in section 4 to the data presented in section 3. The first four subsections will cover the research questions posed in the introduction. In the last subsection we show how the presence of (mis)information is correlated with the US election results.

5.1 How quickly does Covid (mis)information spread?

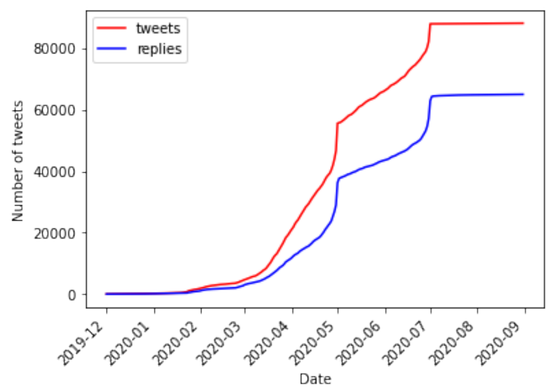
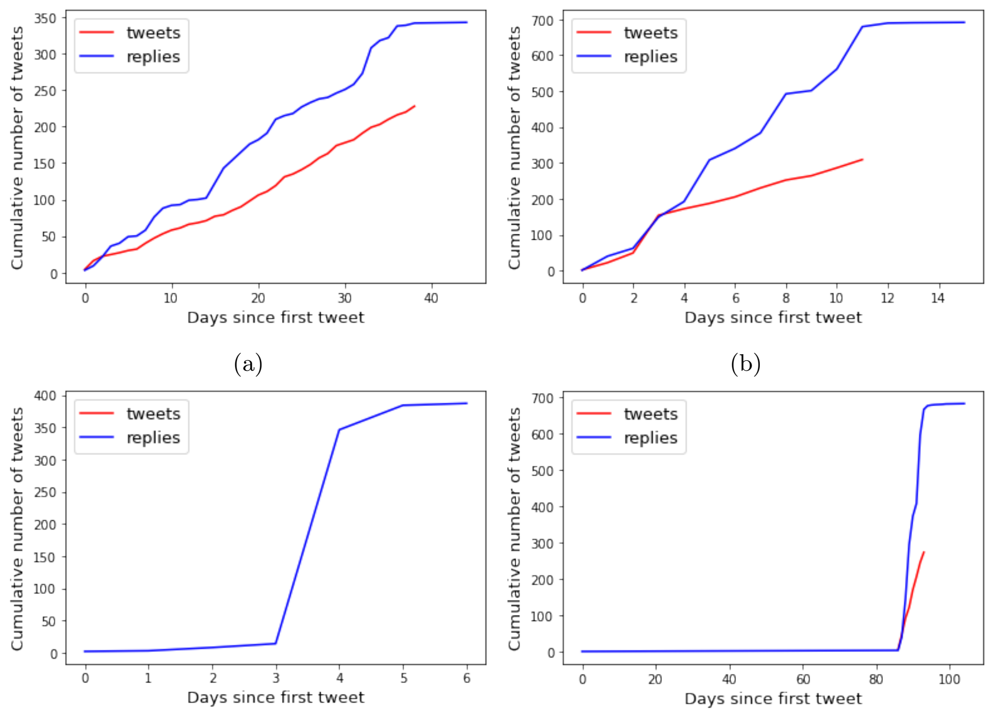


Figure 3: Cumulative number of tweets and replies over time

In Figure 3 the total number of tweets and replies over all topics is shown. The time is in days since the first tweet in the dataset. The number of tweets and replies per day is quite variable. The greatest number of tweets are recorded at the end of May and June. The single day with the most tweets and replies is the 1st of May with 8989 tweets, and 7239 replies recorded. The abrupt stop in the number of tweets and replies at the end of July is likely an artifact of the way the dataset is collected.

In Figure 4 we show the development over time of number of tweets and replies in the largest topic sets of true news, as determined by the number of tweet ids per topic set in the original CoAID dataset. In Figures 4c, and 4d it is clear to see that the quality of the data is lacking, and that a lot of information is missing in the dataset. A large amount of information about the original tweets are missing from these graphs. In the topic shown in Figure 4c all of the original tweets are missing. We do find that the tweets and the replies do follow similar trajectories. For the two topics where the data appears mostly complete we find that the number of new tweets and replies per day is mostly constant. The other two topics have most of their tweets and replies in a short amount of time, with one receiving more than 300 replies in a single day.



(a) (b) (c) (d)

Figure 4: Largest topic sets of true news messages

5.2 What geographical differences exist in the spreading of misinformation?

In Figure 5 we see the ratio of false news in every country in the dataset. A lot of countries in Africa and Eastern Europe are missing from the dataset. The highest ratio of false news is found in Poland. Japan, Greece, Brazil, Russia, and Mexico are well represented in the dataset, and show a large ratio of false news tweets compared to other countries.

In Figure 6 we see the ratio of false news in the states of the contiguous United States. All of the states were sufficiently represented in the dataset. The highest rate of false news was observed in New Mexico. In general, the highest amount of false news is found in the centre of the US, while eastern and western areas tend to have lower ratios of false news.

5.3 To what extent can we identify influential misinformation topics?

The Figure 7 shows three different scatter plots visualising all of the topic sets in the data. As shown in Figure 7a, The duration of a topic set concerning

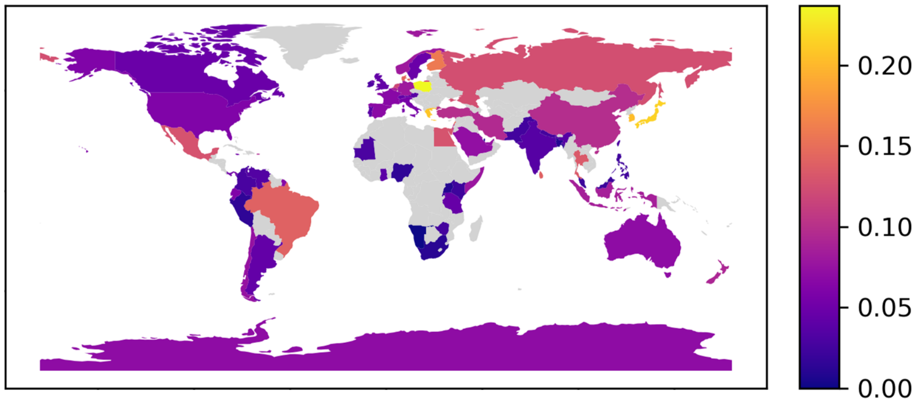


Figure 5: Ratio of false news per country

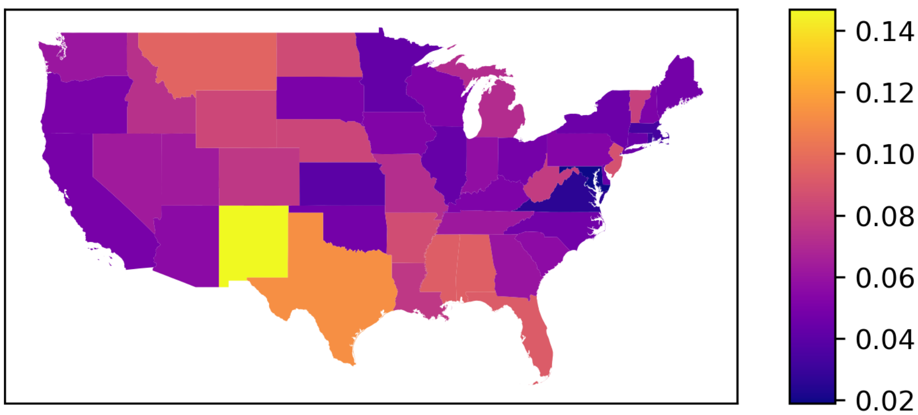
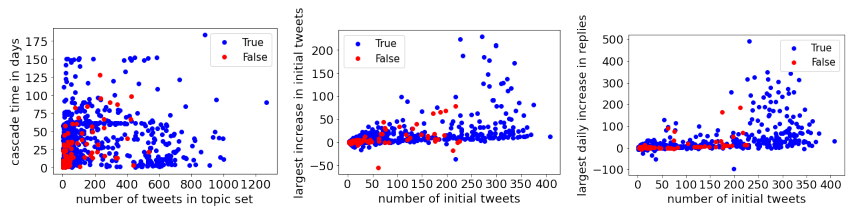


Figure 6: Ratio of false news per US state

true news is not correlated strongly with the number of initial tweets in that topic set (r=0.387). false news is more strongly correlated (r=0.605). Figure 7b shows a few topics where the largest change was negative, i.e., the largest change observed was a decrease in the number of tweets from one day to the next. In most cases however this was an increase in the number of tweets. There does not appear to be a difference in how influential the false and true news topics are. For a given number of initial tweets, the false topics do not seem to spread more rapidly or slowly compared to true news. The same holds for the number of replies and the time that the topic is discussed.

When the goal is to find highly influential topics, then the most interesting data points are the outliers. These represent topics that are spread much more rapidly, or much longer, compared to most other topics. The cascade time shows a few outliers where the topic is discussed for much longer compared to the number of tweets. The number of replies does not seem to have outliers.



(a) Cascade duration (b) Largest increase in number of initial tweets (c) Largest increase in number of replies

Figure 7: Visualising influential topic sets. Each point is a single topic. Blue points represent topics with the label true news, red points are topics with the label false news. Different variables are plotted as a function of the number of initial tweets in each topic set.

This plot does seem to split into two regions between 100 and 300 initial tweets. A small number of topics, both true and false, have a larger increase in the number of replies compared to most of the other topics. The significance of this is as of yet unknown, and further research should be done to ascertain whether or not this effect is still found in larger datasets.

To analyse more of the variables we plot the correlation of six different properties of the topics in Figure 8. The histograms on the diagonal in Figure 8 show how the individual variables are distributed with some interesting results. In social networks, a lot of variables tend have a lognormal distribution. The likelihood of a value x occurring is inversely proportional to e^x. In the log plot of the distributions it would be expected that there is a roughly linear decrease in the occurrences as the values increase. The histograms of the cascade time, and the max delta of both the initial tweets and replies all seem to follow this pattern.

The histograms of the number of initial tweets, replies, and the total number of tweets in the topics all show a much higher rate of occurrence at higher values. This is certainly the case for the topics involving true news. The larger availability of data makes it clear to see. The false news topics however also seem to follow this trend. This effect is most evident in the number of initial tweets. It should be noted that the graphs do not show the number of tweets in an individual topic, but rather a value related to the total volume of discussion surrounding a given topic. As such, the number of tweets in individual topics might still follow this power law distribution. The reason for the abnormal distribution is not known. It could be the case that a topic that is widely discussed could attract more discussions and cascades surrounding the same topic. Even if these new cascades are much smaller, they would still have been included in this dataset. Another possibility is that this abnormal distribution is an artifact of the way that this data was collected. The authors have collected topics from different (reliable or unreliable) news websites and from fact checking websites. Then they searched for tweets containing the title of the article. Some news websites obviously receive more traffic than others, meaning that these topics might naturally see more discussion for example.

Since we do not have information about the total number of tweets in any individual cascade It would certainly be interesting to compare the number of tweets in single cascades, and the number of tweets in all cascades surrounding a given news article or topic.

5.4 Geographic differences in the spreading of influential misinformation

In order to analyse how influential and uninfluential news is spread geographically, we plot each of the different distributions on the map in Figure 9. We see that states in the mid-west region of the US have a small number of tweets in all categories compared to other states. The west coast, and the east of the US have more tweets. This is to be expected, as the availability of the data is highly correlated to the amount of people living in that state. We do see some small differences between the distributions, however none of these are especially striking. California appears to have slightly less influential false news compared to other categories. Texas and Florida appear to have slightly more influential false news.

The differences between the distributions in Figure 10 show more interesting patterns. There is little difference in the distributions of the influential and uninfluential true news as shown in Figure 10a. No state appears to have any strong coloration at all. This means that the spatial distribution of the influential true news is nearly identical to the distribution of true news.

Looking at the differences between the influential and uninfluential false news in Figure 10b shows that there is more variation between these two distributions. Several states have more significant coloration compared to Figure 10a. States colored blue have a higher abundance of Influential false news, compared to the Uninfluential false news. The states with more Influential false news are found mostly in the central and southern US.

Figure 10c has the most intense coloration out of all the plots. There are relatively large differences in how the influential true and false news are spread spatially within the USA. States coloured blue have relatively more influential true news, and states coloured red have relatively more influential false news. The Influential false news seems to be over represented in the southern and mid-western states, although some states in these regions do not follow this pattern, most notably Kansas.

Although there are still relatively large differences between the uninfluential false and true news in Figure 10d, these differences are overall not as large compared to the differences found between the influential news. Interestingly, there seems to be a split between the eastern and western sides of the USA, where false news seems slightly more prevalent in the western parts of the USA. The spatial pattern does differ somewhat from the differences in the Influential news.

More interesting patterns arise when viewing the same data as correlation instead of a geographical map, as shown in Figure 11. The relative abundance of important and unimportant true news (subfigure a) is highly correlated ($r^2 = 0.92$). This is what we expected to find, especially considering the map in Figure 10a.

Similarly, when looking at the abundance of important true and false news in Figure 11c, there appears to be a slight correlation, although not significant ($r^2 = 0.02$). It seems that the presence of influential true news has little influence on the presence of influential false news.

What is interesting is the fact that there is no correlation between the presence of influential and uninfluential false news in Figure 11b, as the r^2 has a value of only 0.004. This is a very unexpected result. More variation is to be expected because of the smaller sample size in the categories of false news, but no correlation at all is unexpected. This suggests that the volume of discussion in a state about viral false topics is not related to the volume of discussion about false topics that have less conversation surrounding them.

There is however a correlation ($r^2 = 0.79$) between the presence of unimportant true and false news as shown in Figure 11d. This means that a state with more unimportant true messages also has more unimportant false messages. This correlation can be explained relatively easily as an effect of the different population sizes, similar to the results found in Figure 11a.

The previous two paragraphs then lead to the conclusion that the important, or viral false news is spread in a different manner compared to all other types.

5.5 Correlating the abundance of false news with US election results

Figure 12 shows the four different categories (influential and uninfluential true and false news) correlated with the results of the 2020 US presidential election. The x axis was chosen arbitrarily to be the percentage of votes for the republican party in a given state. Other variables like the percentage of Democrat votes, or the margin by which a state was won or lost by the Republican party were also tested, but these did not influence the observed results. In order to assess the significance of the observed differences we have also plotted a regression line, and the 95% Confidence interval for each of the different categories. The confidence interval represents the fact that there is some error due to randomness within the data. Given a regression line, there is a 95% chance that the true regression line lies somewhere within the Confidence interval.

The r^2 values are shown in Table 3. None of the different categories of news were positively correlated with the number of votes cast for the Republican party. The abundance of True news appears negatively correlated to the percentage of republican votes. This does not mean that republican states have less true news. The abundance is correlated with the population in a state, and a number of smaller states are well known to favour the Republican party. The correlation of the influential and uninfluential true news are virtually identical, to the extent that the difference between the two is hard to see. This is not entirely unexpected, considering the high degree of correlation found between these two categories themselves as shown in Figure 11a.

Although there is a difference between the true news and the uninfluential false news, the regression of the latter does fall comfortably within the 95% confidence interval of the true news. This is again not unexpected considering the correlation found between these categories in the previous section.

The most interesting is the regression of the influential false news. This regression is at the very edge of the 95% confidence interval for the true news, and the regression for the true news falls just within the 95% CI of the influential false news. As such we can not conclusively state these two categories are significantly different, but this is nevertheless an interesting result that can be a topic for further study. The abundance of Influential False news seems to be less correlated with the percentage of votes cast for the Republican party compared to true news.

The plot in Figure 12 does have an outlier on the x-axis, namely Washington DC, where the vast majority of voters did not vote for the republican party. In order to ascertain the influence of this outlier, the same regression was done excluding the data from this area. The result of this can be seen in Figure 13. Here we find that all the regressions now fall well within the 95% confidence interval of the others. Because of this we are not able to confidently state if the differences between the influential false news and the other categories are a real effect, or simply a coincidence within the dataset. The overall pattern does remain however, with the influential false news having the most significant differences with the true news, and the uninfluential false news falling approximately halfway between the true news and the influential false news. Although this is not conclusive proof by any means, it is again an interesting result, and does invite further study.

Table 3: r^2 values of the correlation of the abundance of each category of news with the US election results

News category	r^2
Influential True news	0.1803
Uninfluential True news	0.1905
Influential False news	0.0275
Uninfluential False news	0.0896

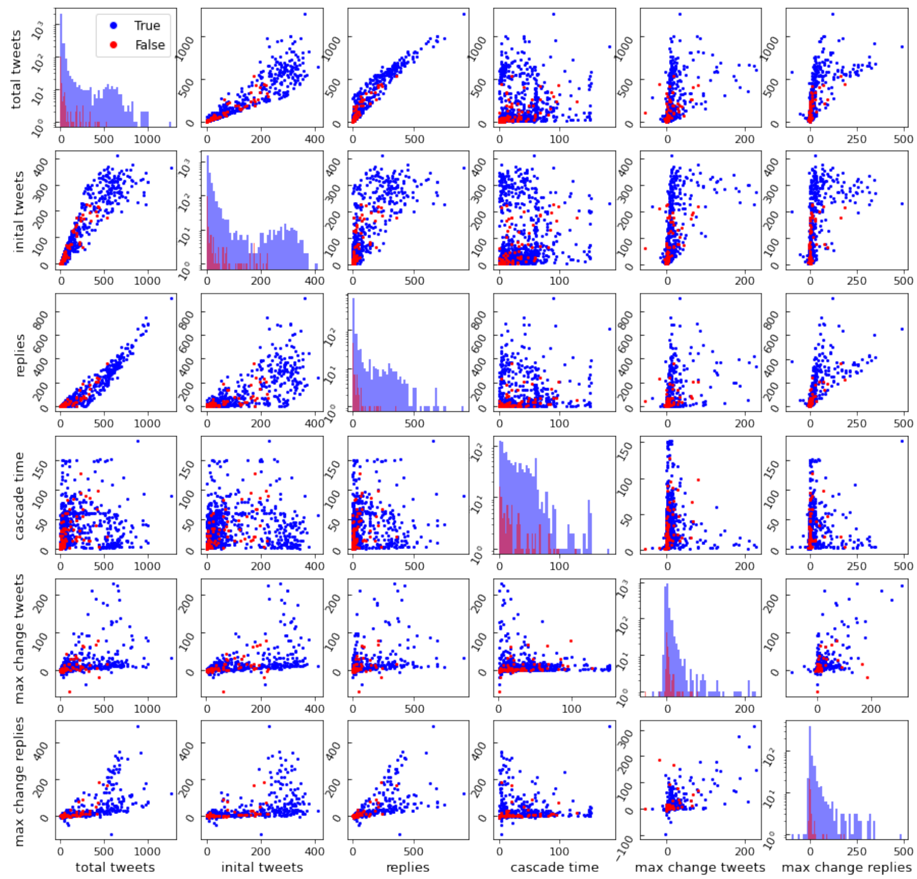


Figure 8: Visualising influential topics. Each scatter plot shows two different variables of the topics depending on the row and column it is in. From top to bottom and left to right; total number of tweets and replies combined, The number of initial tweets in the topic, the number of replies in the topic, the time between the first and last tweet in the topic in days, the largest difference in the number of initial tweets between two consecutive days, and the largest difference in the number of replies between two consecutive days. Each dot represents the values found for one topic. The colour of the dots represent whether the topic is true or false news. The diagonal contains histograms showing how that variable is distributed. The vertical axes of the histograms are logarithmically scaled. The red histogram shows how the false news is distributed for that variable, and the blue shows the distribution of the true news.

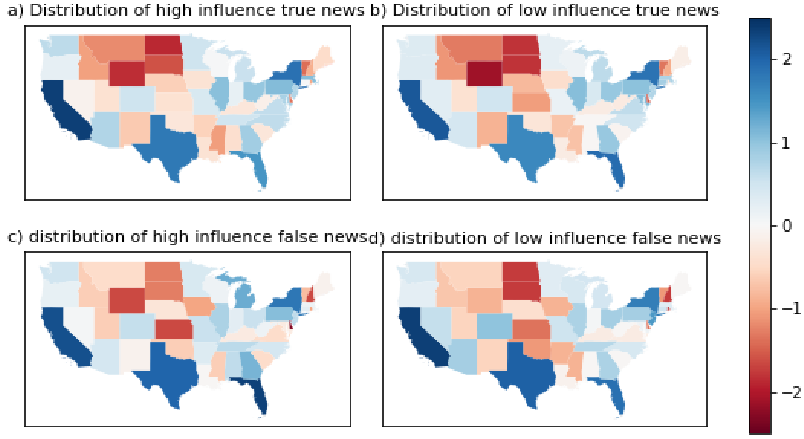


Figure 9: Distribution of tweets from influential and uninfluential information topics, both true and false. The data in these plots has been normalised to a mean value of 0, and a standard deviation of 1. A large positive or negative value in a given state means that the number of tweets is much larger or smaller than the mean number of tweets respectively.

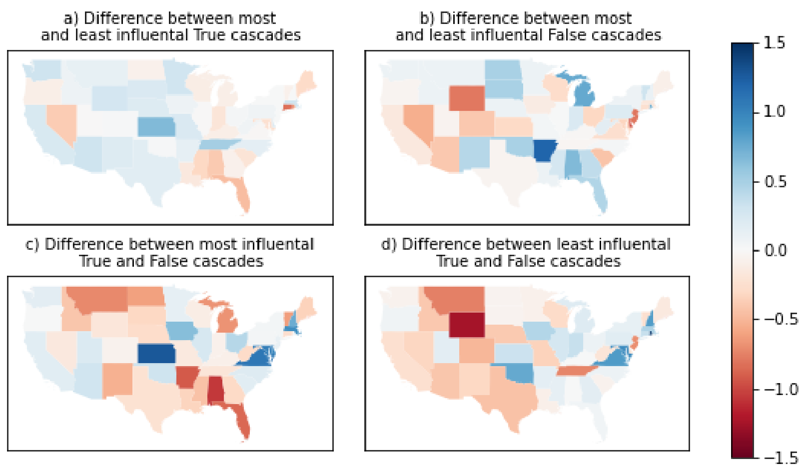
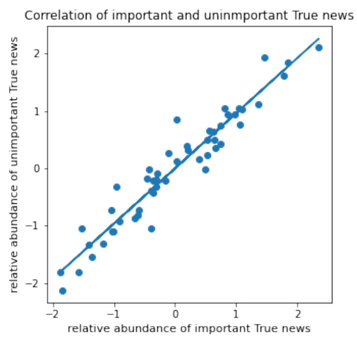
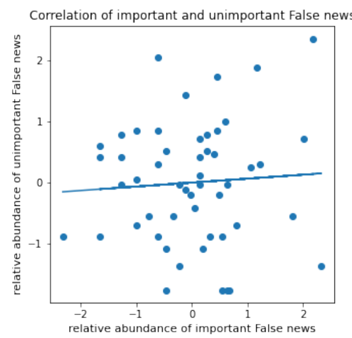


Figure 10: Differences between the distributions of tweets from influential and uninfluential information topics, both true and false. The different distributions have been subtracted from each other in order to visualise large differences between the distribution. In a: uninfluential true news was subtracted from influential true news. In b, the uninfluential false news was subtracted from influential false news. In c, influential false news was subtracted from influential true news. Finally, in d, uninfluential false news was subtracted from uninfluential true news.

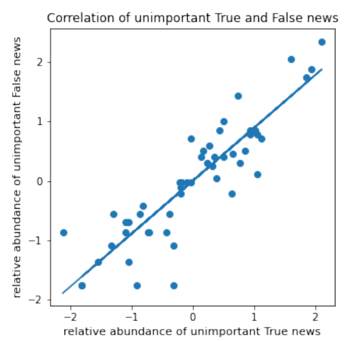


(a) Correlation of influential and uninfluential true news



(b) Correlation of influential and uninfluential false news

(c) Correlation of influential true and false news



(d) Correlation of uninfluential true and false news

Figure 11: Correlation between different categories. Each subplot is directly related to the same subplot in Figure 10. The top row shows the correlations between the abundance of important and unimportant topics for true and false news respectively. The bottom row shows the differences in abundance of true and false news for important and unimportant topics respectively.

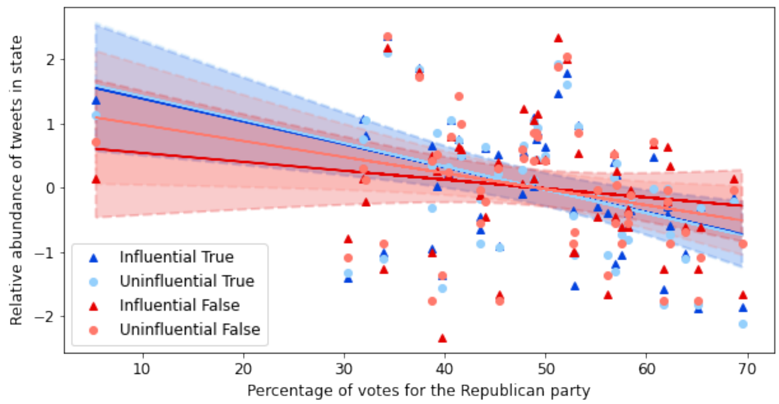


Figure 12: Correlation of the relative abundance of the different categories with the voting results in the USA. The solid line is the line of best fit for a given category. The shaded area around these lines, bounded by a dashed line, represents the 95% confidence interval for the given class.

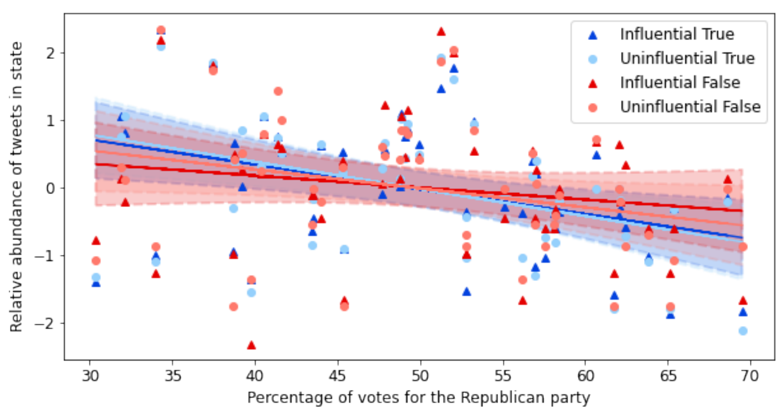


Figure 13: Correlation of the relative abundance of the different categories with the voting results in the USA, with the data for Washington DC removed. The solid line is the line of best fit for a given category. The shaded area around these lines, bounded by a dashed line, represents the 95% confidence interval for the given class.

6 Discussion

We encountered several limitations of the CoAID dataset, especially with regards to the analysis of the spreading of misinformation.

The cascades in the CoAID dataset only contain initial tweets, sharing the news, and the direct replies to that tweet. Any replies to these first replies were not collected, and as such a large part of the information cascade is not available. Any efforts to collect such data now would also be restricted due to the efforts of Twitter to reduce the amount of misinformation on their platform.

The dataset was several months old when we collected the tweets, and in these months a lot of tweets and user profiles containing misinformation, particularly pertaining to Covid, had been removed by Twitter. As much as removing potentially harmful misinformation is a good thing to have happen, it has resulted in numerous tweets missing in the dataset. The effects of this can be seen in Figure 4. Several of the topics are missing some of the initial tweets, and one topic is missing all initial tweets with only retweets remaining.

There is also a large difference in availability of data in different nations. The user base of Twitter is not uniformly distributed around the globe, with a large amount of users in North America and Europe. As such we should expect more data to be available in these countries.

However, half of the tweets in the dataset are from the United States. This is likely an over representation with regards to the actual distribution of twitter users. The over representation of the US might be explained as the result of two different factors related to the construction of the dataset. Firstly the authors of the CoAID paper only used English language websites and news outlets to construct the dataset. This already reduces the amount of tweets collected from non-English speaking countries, as the users in these countries are less likely to encounter and read these websites. Several of the websites used in the CoAID paper are websites primarily aimed at a US demographic, e.g. the websites for the NIH and the CDC, as well as politifact.com. Using these websites as the basis to construct a dataset will introduce a lot of selection bias for English speaking countries in general, and especially the US. This is not a fault of the authors of the CoAID paper, as they did not necessarily intend to construct a dataset to represent the distribution of false news around the globe. It is however a factor that should be considered when constructing or analysing data that aims to investigate geographical distribution on a global scale. Collecting a dataset of tweets surrounding one or several topics across languages and geographical regions is not a trivial task, and it is not entirely clear what the best way is to achieve this. In order to reduce the bias towards the US specifically one could use websites from a larger number of news outlets and fact checking organisations, spread across the globe.

The geotagging API is relatively robust, however there are still some limitations. Geographical gazetteers have a bias within their datasets towards more developed countries. This means that the results returned have a bias towards these countries as well. This bias is likely not a large issue in this thesis due to the geographical bias introduced in the collection of the data, as discussed previously. In general the results given by the gazetteer will not be completely accurate. Multiple cities and towns will have the same name, even if these are in completely different countries or continents, e.g. London in Ontario, Canada. This problem is also exacerbated by the fact that the amount of information in the location field of the user profiles is quite limited, and context clues might not be available.

Locating a user based on geotags in their tweets is another possible option. This approach was not very useful in this research, as only a small number of tweets was available for each individual user.

The biggest problem is again a relative lack of data. In order to investigate these regional differences one would ideally have a lot of data within each region, or state in this specific case, in order to reduce the variance and uncertainty present. The large difference in sample size between the true and false news also is a factor. There is much more data available for the true news. After splitting the false news into influential and uninfluential news, there are a number of states where less than 10 tweets are found within both of the false news categories. This increases the variance and the uncertainty of the results for the false news categories, as small differences in these states will have a larger effect on the calculated relative abundance of the tweets in that state. This effect is also not helped by the use of the log10 of the number of tweets when calculating the relative abundance because the logarithmic function is most sensitive near 0. Using the log10 was very useful in lowering the influence that large states have, however it does not reduce the variance in the case where tweets are very scarce.

The small number of topics represented in the influential false news category should also be noted. There were 8 different topics, that together accounted for 40% of all tweets concerning false news. We can by no means assume that this small set of topics is representative of the entirety of the false news, even when only Covid misinformation is considered. Any differences found in the spreading of these influential false cascades could be related to these topics, and where these might have been more or less relevant. We unfortunately were unable to assess how important differences in these topics were for the final results, as the amount of data over all these topics combined was already very sparse for a large portion of the USA.

The extremely strong correlation between the influential and uninfluential true news does support the idea that working with the relative abundance is a useful tool for investigating whether two geographical distributions are the same, or whether there might be an unexplained difference between the two. That extremely strong correlation also does lend slightly more weight to the idea that there is a real difference between the geographical distributions of the influential and uninfluential false news.

It might however also be a result of the fact that there was simply more data available, leading to a more accurate estimation of the true distribution of the tweets in the different categories.

The way the dataset was collected only gives limited insights into the spreading of false news unfortunately, as only the start of misinformation cascades is captured. The start of an information cascade does not provide much insight on the rest of the cascade. Including the rest of the cascade in the dataset would not only increase the amount of data available, but also allow for different types of analysis, for example based on the cascades themselves, or the retweet networks of the true and false information.

7 Conclusions

In this study we have examined the geographical distributions of false and true news cascades on twitter. We used the existing CoAID dataset to obtain tweets about topics surrounding covid, labeled as false or true news. We compared the geographical distributions of influential and uninfluential true and false news. The CoAID dataset is not well suited to determine the speed at which misinformation spreads. We find a mostly constant volume of tweets and replies for topics where enough data is present to provide a full picture, but a large number of topics seem to be missing a lot of data. We find much less tweets and replies related to false news, but this is largely due to the large number of true news topics included in the dataset compared to false news. Globally misinformation seems to have more prevalence in eastern Europe, eastern Asia, and South America. Data is however sparse, and a lot of countries do not have enough data to make meaningful conclusions. Within the US misinformation seems concentrated in the mid-western and rural states. We have tried to identify influential topics by analysing the total volume of conversation about a topic, and the number of tweets mentioning a specific news item. Better and more sophisticated methods for identifying influential topics might be possible, but were unfeasible due to limitations in the data. We have found that the geographical distribution of influential false news is different from uninfluential false news and all true news, while uninfluential false news is correlated with true news. This seems to indicate that the influential false news might spread through a different mechanism compared to other types of news, although further study is necessary to confirm these findings. The distribution of true news is negatively correlated with the percentage of votes cast for the republican party. The distribution of influential false news appears to be uncorrelated. Future work could try to confirm whether these patterns hold on a global scale using a more extensive dataset. Another interesting idea is to identify if this pattern holds for all types of misinformation, or if different types of misinformation like misinformation surrounding elections and conspiracy theories like flat-earth are distributed in a different way compared to each other.

References

[1] "Novel Coronavirus (2019-nCoV) SITUATION REPORT - 1," World Health Organisation, Tech. Rep., Jan. 2020.

[2] "Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)," World Health Organisation, Tech. Rep., Jan. 2020.

[3] World Health Organisation, "WHO Coronavirus dashboard," https://covid19.who.int/, 2020, accessed: 2023-04-04.

[4] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. Scala, "The covid-19 social media infodemic," *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.

[5] L. Singh, S. Bansal, L. Bode, C. Budak, G. Chi, K. Kawintiranon, C. Padden, R. Vanarsdall, E. Vraga, and Y. Wang, "A first look at covid-19 information and misinformation sharing on twitter," *arXiv preprint arXiv:2003.13907*, 2020.

[6] W. Jennings, G. Stoker, H. Bunting, V. O. Valgarsson, J. Gaskell, D. Devine, L. McKay, and M. C. Mills, "Lack of trust, conspiracy beliefs, and social media use predict covid-19 vaccine hesitancy," *Vaccines*, vol. 9, no. 6, p. 593, 2021.

[7] J. Fauber and D. Chen, "Trump's praise of hydroxychloroquine led to a boom in prescriptions. then came reports of thousands of patients experiencing adverse events." *Milwaukee Journal Sentinel*, Aug 2020. [Online]. Available: https://eu.jsonline.com/story/news/investigations/2020/08/06/trump-praised-hydroxychloroquine-possible-side-effects-reported/5538296002/

[8] L. Cui and D. Lee, "CoAID: COVID-19 healthcare misinformation dataset," *CoRR*, vol. abs/2006.00885, 2020.

[9] S. A. Memon and K. M. Carley, "Characterizing covid-19 misinformation communities using a novel twitter dataset," *arXiv preprint arXiv:2008.00791*, 2020.

[10] S. Latif, M. Usman, S. Manzoor, W. Iqbal, J. Qadir, G. Tyson, I. Castro, A. Razi, M. N. K. Boulos, A. Weller *et al.*, "Leveraging data science to combat covid-19: A comprehensive review," *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 1, pp. 85–103, 2020.

[11] S. Goel, D. J. Watts, and D. G. Goldstein, "The structure of online diffusion networks," in *Proceedings of the 13th ACM Conference on Electronic Commerce*, ser. EC '12. Association for Computing Machinery, 2012, p. 623–638.

[12] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Transactions on the Web (TWEB)*, vol. 1, no. 1, pp. 5–es, 2007.

[13] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[14] G. K. Shahi, A. Dirkson, and T. A. Majchrzak, "An exploratory study of covid-19 misinformation on twitter," *Online social networks and media*, vol. 22, p. 100104, 2021.

[15] K. Sharma, S. Seo, C. Meng, S. Rambhatla, and Y. Liu, "Covid-19 on social media: Analyzing misinformation in twitter conversations," *arXiv preprint arXiv:2003.12309*, 2020.

[16] A. M. Forati and R. Ghose, "Geospatial analysis of misinformation in covid-19 related tweets," *Applied Geography*, vol. 133, p. 102473, 2021.

[17] M. Erdemandi and J. Leach, "'Masks don't work:'ideology associations and the geospatial propagation of covid-19 disinformation on twitter," *Available at SSRN 3898097*, 2021.

[18] M. Dredze, M. J. Paul, S. Bergsma, and H. Tran, "Carmen: A twitter geolocation system with applications to public health," in *Workshops at the twenty-seventh AAAI conference on artificial intelligence*, 2013.

[19] O. Ajao, J. Hong, and W. Liu, "A survey of location inference techniques on twitter," *Journal of Information Science*, vol. 41, no. 6, pp. 855–864, 2015.

[20] Z. Zhao, J. Zhao, Y. Sano, O. Levy, H. Takayasu, M. Takayasu, D. Li, J. Wu, and S. Havlin, "Fake news propagates differently from real news even at early stages of spreading," *EPJ data science*, vol. 9, no. 1, p. 7, 2020.

[21] "Novel Coronavirus (2019-nCoV) SITUATION REPORT - 13," World Health Organisation, Tech. Rep., Feb. 2020.

[22] J. Zarocostas, "How to fight an infodemic," *The Lancet*, vol. 395, no. 10225, p. 676, 2020.

[23] J. L. Juul and J. Ugander, "Comparing information diffusion mechanisms by matching on cascade size," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, 11 2021.

[24] F. Pierri, C. Piccardi, and S. Ceri, "Topology comparison of twitter diffusion networks effectively reveals misleading information," *Scientific reports*, vol. 10, no. 1, pp. 1–9, 2020.

[25] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "The spreading of misinformation online," *Proceedings of the National Academy of Sciences*, vol. 113, no. 3, pp. 554–559, 2016.

[26] F. Zollo and W. Quattrociocchi, "Misinformation spreading on facebook," in *Complex spreading phenomena in social systems.* Springer, 2018, pp. 177–196.

[27] D. S. Abdelminaam, F. H. Ismail, M. Taha, A. Taha, E. H. Houssein, and A. Nabil, "CoAID-DEEP: An Optimized Intelligent Framework for Automated Detecting COVID-19 Misleading Information on Twitter," *IEEE Access*, vol. 9, pp. 27 840–27 867, 2021.

[28] N. L. Kolluri and D. Murthy, "CoVerifi: A COVID-19 news verification system," *Online Social Networks and Media*, vol. 22, p. 100123, 2021.

[29] A. Glazkova, M. Glazkov, and T. Trifonov, "g2tmn at constraint@aaai2021: Exploiting ct-bert and ensembling learning for covid-19 fake news detection," in *Combating Online Hostile Posts in Regional Languages during Emergency Situation*, T. Chakraborty, K. Shu, H. R. Bernard, H. Liu, and M. S. Akhtar, Eds. Cham: Springer International Publishing, 2021, pp. 116–127.

[30] "Twarc, A command line tool (and Python library) for archiving Twitter JSON," accessed: 2023-04-04. [Online]. Available: https://github.com/DocNow/twarc

[31] D. Carriere, "Geocoder: Simple, Consistent," accessed: 2023-04-04. [Online]. Available: https://geocoder.readthedocs.io/

[32] C. Zaiontz, "Confidence and prediction intervals for forecasted values." accessed: 2023-05-07. [Online]. Available: https://real-statistics.com/regression/confidence-and-prediction-intervals/

[33] rowan nicholls, "Statistics in Python: Simple Linear Regression," accessed: 2023-05-07. [Online]. Available: https://rowannicholls.github.io/python/statistics/hypothesis$_testing/simple_linear_regression.html$