How COVID-19 Conspiracy Theories Spread on Twitter

Anatoliy Gruzd
Toronto Metropolitan University
gruzd@torontomu.ca

Amira Ghenai
Toronto Metropolitan University
aghenai@torontomu.ca

Philip Mai
Toronto Metropolitan University
philip.mai@torontomu.ca

Abstract

Since the onset of the COVID-19 pandemic, conspiracy theories (CTs) related to the virus have been widely circulated on social media. The uncertainty surrounding the pandemic and available treatment options likely contributed to the wide dissemination of such theories on social media platforms like Twitter. This retrospective study examines the spread of CTs surrounding Bill Gates and COVID-19 vaccines on Twitter and identifies what accounts contributed to their dissemination. Based on the social network analysis of 100,601 Bill Gates and vaccine-related tweets shared by 71,364 users between March 1 and May 31, 2020, the study found that automated and suspended accounts had a significant impact on the spread of CTs around this topic. Their tweets were more likely to be reshared by others than by chance alone. This highlights the need for social media platforms to continue to act against harmful automated accounts, particularly considering recent trends to ease content moderation policies and debunking interventions by social media giants in the post-pandemic era.

COVID-19, Social Media, Conspiracy Theories, Vaccine Hesitancy, Twitter.
Keywords:

1. Introduction

The severe respiratory syndrome coronavirus 2 (SARS-CoV-2) was first reported in December 2019 (WHO, 2022), and has since rapidly spread around the world, resulting in the COVID-19 pandemic. A scientific consensus emerged in the first year of the pandemic that an effective vaccine is the most significant tool in curbing the spread of the virus and reducing severe cases of the disease. However, the anti-vaccination community has expressed strong opposition to COVID-19 vaccines from early on in the pandemic, likely contributing to vaccine hesitancy among the population (Johnson et al., 2020; Wiysonge et al., 2022). For example, one of the early polls in the U.S. conducted even before COVID-19 vaccines became available showed that only 42 percent intended to get vaccinated against the virus (Reinhart, 2022). In the past, vaccine hesitancy has found to be one of the leading causes of the 2019 measles outbreak in the U.S and other countries (P. Hotez, 2019; Olive et al., 2018).

In this retrospective study, we wanted to better understand what social media accounts were the main spreaders of conspiracy theories (CTs) linking Bill Gates to COVID-19 vaccines. These theories included claims that COVID-19 vaccines contained microchips to track the population or that they were designed to control the population growth. Our object is two-fold: 1) contribute to the existing literature on the negative aspects of social media in circulating harmful CTs, and 2) learn from the COVID-19 pandemic to better prepare for future outbreaks.

Health-related CTs in general, and specifically those pertaining to vaccines, raise particular concerns as they have been linked to negative individual and societal consequences. For example, prior studies found that people who believe in CTs are less likely to adhere to public health guidelines (Allington & Dhavan, 2020; Allington, Duffy, et al., 2020; Georgiou et al., 2020; Imhoff et al., 2020; Oleksy et al., 2021), less likely to participate in future treatments (Freeman et al., 2020), and are more likely to use violence (Imhoff et al., 2020; Jolley et al., 2020). Parents who believe in CTs are less likely to vaccinate their children (Chung, 2009; Jolley et al., 2014; Shapiro et al., 2016) due to mistrust in authorities, fear of side effects, and feelings of powerlessness (Jolley et al., 2014).

We employed exponential random graph modeling to analyze the retweet communication network and assess the factors that influence the spread of Bill Gates and COVID-19 vaccine-related CTs on Twitter. Our results indicate that Twitter's content moderation measures had only been partially successful in reducing the presence of accounts spreading harmful health-related CTs. We also discovered that probable bot accounts remained active during the study and were

URI: https://hdl.handle.net/10125/106682
978-0-9981331-7-1

among the most frequently retweeted sources of CTs on this topic.

2. Literature Review

2.1. Conspiracy theories

Douglas et al. (2019) defines CT as an attempt to explain social and political events with claims of secret plans by powerful individuals or groups. For example, the 9/11 attack CT implicated the George W. Bush administration, the Saudi government, and the Jewish people as responsible for the event. Similarly, CTs about the assassination of U.S. President John F. Kennedy contended that individuals other than Lee Harvey Oswald were involved in the assassination (McCauley et al., 1979). More recently, the “birther” CT questioned former U.S. president Barack Obama’s citizenship (Enders et al., 2020).

A term relevant to research on CT is “conspiracy belief” which refers to believing in one or many CTs. For example, during the 2016 Brexit referendum in the UK, about 46 per cent of those intending to vote “leave” believed in the presence of an electoral fraud (YouGov, 2016). Another term relevant to this research is “conspiracy thinking”, referring to conspiratorial mindset that may drive an individual to believe in a CT (Wood et al., 2012).

Surprisingly, believing in CTs may have potential benefits. This is because CTs encourage to challenge the actions of the government and those in power, which may lead to greater transparency (Clarke, 2002; S. Miller, 2002; Swami et al., 2010).

Although some research suggests potential benefits in believing CTs, a general consensus is that they have harmful social, health, and political consequences (Douglas et al., 2019). Furthermore, believing in CTs is also linked to feelings of isolation or helplessness (Abalakina-Paap et al., 1999).

Regarding health-related CTs, studies have shown a strong correlation between believing in CTs and making health-related choices that may be harmful to one’s health. In 2014, Oliver et al. (2014) found that those who believe in health-related CTs were less likely to consult medical professionals and more likely to use alternative medicine. Goertzel (2010) reported that individuals who believed CTs have less trust in public health policies, science, and government in general.

2.2. Conspiracy theories related to COVID-19

Individuals tend to turn to CTs during societal crises due to uncertainty and anxiety, as it may help them make sense of a difficult situation (Desta et al., 2020; Van Prooijen et al., 2017). The pandemic was not an exception, as CTs concerning COVID-19 have emerged from its start (Romer et al., 2020). These theories include the belief that COVID-19 is a bioweapon created by China, that Bill Gates created the virus, that 5G networks spread the virus, or that the pandemic is a hoax. Although these theories may seem implausible, there are individuals open to them. For example, Dornan (2020) found that 25 per cent of Canadians believed that the virus was engineered in a Chinese laboratory, 23 per cent believed that hydroxychloroquine was an effective treatment, and 11 per cent believed the disease was being spread to cover up the effects of 5G radiation.

Several COVID-19 CTs have emerged due to political polarization in countries like the U.S. One such theory, as highlighted by J. M. Miller (2020), suggested that COVID-19 was exaggerated to undermine Trump’s presidency. Another politically-driven conspiracy theory (FilmYourHospital) encouraged people to take photos of empty hospitals in their area as a proof that the pandemic is not real. This theory was mostly spread by Trump supporters (Gruzd et al., 2020).

Other research have also found that believing in CTs about COVID-19 correlated with conspiratorial thinking more broadly (Georgiou et al., 2020; Goldberg et al., 2020; Klofstad et al., 2019). In addition, Cassese et al. (2020) showed that men are more likely to believe COVID-19 CTs than women.

2.3. Vaccine-related conspiracy theories

Resistance to vaccines has been around since the first documented vaccination attempt by Dorset farmer Benjamin Jesty, who vaccinated his family against smallpox in 1774 (Kaufman et al., 2018). In the early 1800s, Jenner demonstrated that cowpox could protect against smallpox (Dubé et al., 2015), but despite smallpox killing three in ten (CDC, 2022), many still resisted the use of the vaccine. It was not until mid-1800s that the UK made the smallpox vaccination mandatory, which led to opposition from those who refused to have their bodies “controlled” by the government (Colgrove et al., 2005). This marked the beginning of the anti-vaccination movement in the UK and worldwide (Dubé et al., 2015; Wolfe et al., 2002).

The anti-vaccine movement relies on several common claims to oppose vaccination, including claims that infectious diseases are natural and therefore benign, that vaccines are ineffective because they are not perfect, and that vaccines cause harm. Members of this movement often cite personal anecdotes rather than rely on statistically significant data. In addition, the anti-vaccine movement has embraced CTs about vaccines (e.g., vaccines contain a microchip) to support their opposition to vaccination (Kaufman et al., 2018).

Vaccine knowledge can moderate the impact of exposure to anti-vaccine CTs (Yang, Varol, et al., 2020), but correcting people's beliefs and behaviours after initial exposure is challenging (Jolley et al., 2014). This is especially true for deeply convicted believers (Uscinski, 2018).

2.4. Conspiracy theories, social media and social bots

In this work, we studied the role of Twitter in spreading CTs because of its prominence as a source of both credible and non-credible information related to the pandemic (Allington, Duffy, et al., 2020; Gruzd et al., 2020). Within the platform, we were especially interested in the role of social bots in disseminating COVID-19 vaccine related CTs. Social bots (or bots) are accounts run by scripts to automatically produce content and generate interactions with human accounts on social media (Davis et al., 2016). Prior research showed how bots promote different political ideologies and amplify controversial topics (Sayyadiharikandeh et al., 2020; Yang, Torres-Lugo, et al., 2020), including vaccine related content (Broniatowski et al., 2018) and COVID-19 CTs (Ferrara et al., 2020). We aimed to validate the prior work in this area.

Most bot detection methods rely on machine learning classifiers to differentiate between human-like accounts and bots (Sayyadiharikandeh et al., 2020; Yang, Torres-Lugo, et al., 2020). However, as the aim of this study is not to improve existing approaches, we utilized a widely-used bot detection service called Botometer (Davis et al., 2016). Botometer employs over 1,000 features to measure the probability of an account exhibiting bot-like behavior. It uses six different types of features, including account-specific features (such as language and location), friends features (such as information about the account's followers), temporal features (when the accounts tweets), content and sentiment features (what the accounts tweet), and network features to capture information diffusion patterns. According to the developers, the service achieves an accuracy of 0.95 AUC (Area Under the Curve) with ten-fold cross-validation.

2.5. Common methods to study the spread of conspiracy theories on Twitter

This section provides a brief overview of various methods used to examine the dissemination of CTs on social media. Most studies in this area relied on either content analysis (whether manual or automated) or social network analysis, or both.

Starbird (2017) studied CTs on Twitter related to mass shootings. The author used a mixed-method approach to analyze data collected in 2016. Results showed that alternative media outlets fueled CT content, which often had strong political agendas and attacks on mainstream media.

Nerghes et al. (2018) studied CTs on YouTube related to Zika virus. The authors analyzed 35 English YouTube videos with a minimum of 40,000 views related to the virus. They found that some videos promoted CTs, such as blaming specific organizations for the virus, or claiming it was a bio-weapon for depopulation. Authors used sentiment and content analysis to compare the differences between informative and CT videos (in terms of views, user activities, sentiment, and user comments content). They found no significant difference in user engagement between the two types of videos, suggesting that YouTube users respond similarly to both.

Kou et al. (2017) also examined the spread of Zika-related CTs but on Reddit. The researchers used qualitative analysis to analyze 156 top-commented posts and 47,551 associated comments to understand how these theories spread on the platform. They found that people tend to seek out potential explanations to cope with the outbreak, and that Reddit's design interface potentially facilitates the spread of conspiratorial thinking. For example, the platform allows for collective elaboration of conspiracy details between conspiracy theorists.

In addition to content analysis, social network analysis (SNA) is another common approach used to trace the spread of CTs on social media and identify those who are resharing such content. In an analysis of Twitter discussions surrounding Zika-related CTs, Wood (2018) examined a dataset of 25,162 tweets from September 1, 2015 to March 31, 2016. Using SNA, the researchers discovered that CTs were disseminated through a more decentralized network, in contrast to debunking messages. This indicates that multiple accounts likely created CTs, offering various claims to challenge mainstream media narratives. However, in the part of the network dedicated to debunking, users tended to follow one influential account (e.g., an authoritative source), resulting in a more centralized information diffusion network structure.

In another study that used SNA (in addition to other approaches), Bruns et al. (2020) examined the spread of the CT that suggested 5G mobile networks spread COVID-19 on Facebook (Bruns et al., 2020). The study found that the origin of the CT on Facebook came from pre-existing conspiracy theorists who usually target public figures, celebrities, sports stars, and media outlets to increase the chances of exposure. P. J. Hotez (2020) also used SNA to examine the same COVID-19 5G CT, but on Twitter. The authors found that reliable and trusted sources are absent when looking at the network of CT spread on the platform.

Considering the strengths of SNA in identifying how CTs spread from one account to another in online networks, we adopted this approach in our study. SNA is used to understand the attitudes and behaviours of people, both online and offline, by examining their social and communication connections with others (Alhajj et al., 2014; Tindall et al., 2001). On Twitter, retweeting is a feature that enables users to form and maintain social connections. By sharing content from other users, a new opportunity for connection can arise, as the original poster is notified of the retweet. This can lead to a stronger connection between users, creating a "latent tie" (Haythornthwaite, 2002). Thus, by studying communication practice based on retweets, we explored information and misinformation sharing across both existing and latent ties among Twitter accounts.

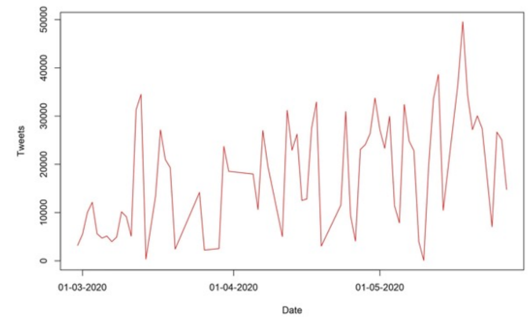


Figure 1: Total number of tweets per day

Table 1: Dataset statistics

Collection	Total tweets	Users	Tweets/User
COVID keywords	145,091,231	15,658,605	9.27
English language	81,000,234	12,477,542	6.5
Vaccine-related	1,279,959	740,479	1.73
Bill Gates CT	100,601	71,364	1.41

3. Method

3.1. Data collection

We used the Social Feed Manager open source platform to collect tweets by querying Twitter's Search API (v1.1) every 30 minutes for tweets containing the keywords "COVID19" and "COV", including variations of these terms. The resulting dataset included 145,091,231 tweets and 15,658,605 accounts from March 1 to May 31, 2020, which covered the peak of interest in COVID-19 when the pandemic was officially declared, and lockdown restrictions were implemented.

For this study, only English tweets were considered, which accounted for 81,000,234 tweets and 12,477,542 accounts. To collect tweets related to COVID-19 vaccines, we applied regular expressions over the text field to identify tweets containing words that start with either vaccin, vacin, vax (including hashtags) or hyphenated words with -vaccin, -vax, or -vacin. This search strategy was iteratively refined to minimize false positives, resulting in a final dataset of 1,279,962 tweets shared by 740,479 accounts. The number of tweets per day about COVID-19 vaccines during the studied period is shown in Figure 1.

In this paper we focused on CTs about the "Bill Gates" topic as it was one of the most popular topics associated with discussions about COVID-19. Figure 2 shows the number of tweets per day around Bill Gates CTs. To locate tweets concerning the Bill Gates-related CTs, we used Indri[1] to index all vaccine-related tweets, and then used search queries consisting of Boolean strings with relevant terms such as "patent 666", "microchip", "mark", "quantum dot", "rfid", and "id2020" (similarly to Qazvinian et al. (2011)). These terms are commonly associated with false claims that Bill Gates is involved in a patent concerning a microchip tattoo or an invisible mark that can track or manipulate human behavior[2]. Table 2 summarizes our retrieval strategy for Bill Gates-related tweets. Table 1 outlines the full data collection process.

3.2. Social Network Analysis (SNA)

As mentioned previously, in order to examine the spread of CTs about Bill Gates on Twitter, we used SNA, a theoretical and methodological framework that enabled us to visually and analytically examine user interactions on a large scale. Using SNA to analyze Twitter data requires the representation of interactions as a graph. We constructed a graph by identifying and connecting Twitter accounts (nodes) that retweeted one another (edges or ties). To generate this "retweet" network, we used a custom Python script based on the Network X[3] library (Araujo et al., 2017; Sanders et al., 2019; Schuchard et al., 2019; ten Thij

[1] http://www.lemurproject.org/indri.php

[2] https://archive.is/ox0k8,https://archive.is/teYst,https://www.snopes.com/fact-check/bill-gates-id2020/

[3] https://networkx.org/

Table 2: Conspiracy theories with corresponding Indri queries and sample tweets

Conspiracy theory topic	Query	Sample tweets
Bill gates	gates OR (bill gate) OR rfid OR microchip OR (micro chip) OR (patent 666) OR id2020 OR (quantum dot) OR mark	1. "Who could (or will) benefit from Covid-19, which (allegedly) came from a food market in China? Ultra wealthy 'elites' who will buy stocks during panic selling, then make a killing when markets rebound. Vaccine manufacturers and their large shareholders (including **Bill Gates**)" 2. "**Bill Gates** & the Cabal thought they could push vaccinations & **microchip** technology on us after releasing COVID-19 as a bio-weapon. But that power grab is about to fail. Soon everybody will know that Big Pharma and the Swamp Rats in DC have been covering up cures for DECADES."

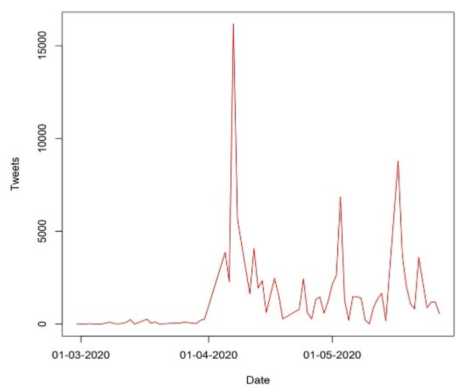


Figure 2: Total number of tweets per day related to Bill Gates CTs

et al., 2014). We excluded isolated accounts that have not retweeted anybody else or received retweets themselves. To facilitate a further analysis, we included the account-level metadata provided by Twitter API as node attributes. This included the username, the number of tweets in the user's timeline, the number of favourites, the number of followers, the number of followees, and whether or not the account is a verified.

In addition to including account-level attributes, we have also analyzed each account using Botometer[4]. As noted earlier, Botometer assigns a value between 0 to 1 to indicate how likely an account is automated based on a pre-trained machine learning classifier. Accounts with values closer to 1 are more likely to be automated.

6,314 accounts (8%) had no Botometer score since they were no longer available at the time of checking their status. At the time of the data collection, in March of 2020, Twitter was proactively suspending or temporarily restricting accounts if they were in violation of the Twitter community rules and standards, especially if they promoted COVID-19 misinformation that put others in harm. Some of the reasons for suspension include "artificially amplifying or suppressing information, interfering in elections, sharing synthetic/manipulated media which may cause harm, or promoting violence against, threatening, or harassing an individual or a group of people" (Twitter, 2022). These actions align with the behavior of automated accounts or bots, making it reasonable to assume that a considerable number of the unavailable accounts were suspended due to such activities. Therefore, we assigned the Botometer score of 1.1 to the accounts that were unavailable during the study.

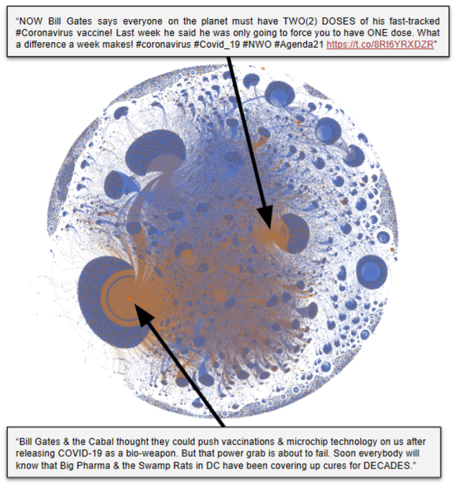
Figure 3 displays the resulting network consisting of 78,312 nodes and 93,567 edges. Based on a visual examination, the network exhibits a scale-free topology, characterized by the presence of few highly connected accounts surrounded by many less connected ones. The highly connected accounts act as hubs or sources of information (likely misinformation). Their central position in the network show their key role in driving the discussions and shaping the narratives on Twitter.

3.3. Exponential Random Graph Model (ERGM)

ERGM has emerged as a leading approach in SNA to examine and explain the formation of connections in social and communication networks. It considers both exogenous factors, such as node-level attributes, and endogenous factors, such as network-level factors (Sha et al., 2018). ERGM is applicable for statistical testing with network data where the assumption of independence of observations cannot be assumed. Using random network simulation through Markov chain Monte Carlo, the method estimates whether the observed node and network properties are likely to exist

[4]https://botometer.osome.iu.edu/api

Figure 3: “Bill Gates” Retweet Network (Color = Botometer Scale from 0-blue/less likely to be automated to 1-orange/more likely to be automated; and to 1.1 when an account is no longer available.)



Note: Two of the most frequently retweeted tweets (both are by suspended/deleted accounts: one by @Education4Libs, in the left low corner and another one by @BeachMilk, in the right top corner.

by chance alone (Morris et al., 2008).

In the area of social media research, ERGM has been successfully applied to studying social and communication networks in different communities and platforms. For instance, ERGM has been used to investigate friendship formation among politically motivated users in VK groups during the revolution of Dignity in Ukraine (Gruzd et al., 2015). Similarly, ERGM has been used to analyze the construction of relationships across different relief organizations on Twitter and Facebook during and after Typhoon Haiyan (Lai et al., 2017). Additionally, ERGM has been used to identify the commonalities shared by members of extremist groups in the darknet (Rashed et al., 2019).

In our case, we tried to uncover what explains a retweet behaviour in our network and what makes some accounts to be more likely retweeted than others. Using the regression analysis language, the dependent variable is the likelihood of A retweeting B, and the independent variables include both exogenous and endogenous factors. The exogenous factors that we tested are the account's number of followers, their number of tweets since they joined twitter, the age of their account, the verification status, and the botometer score. As for the endogenous factors, we only tested the reciprocity of retweets, in other words, how likely that if account A retweeted account B, the later would reciprocate. To perform ERGM, we used the ERGM

Table 3: Factors Underlying Tie Formation in the Twitter Retweet Network about Bill Gates CTs

Factors	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
edges	-11.81	0.01	-2248.0	<1e-04***
mutual	3.58	0.20	17.9	<1e-04***
nodeicov.botscore2	1.86	0.01	236.7	<1e-04***

Note: Significance codes: ‘***’ for p-values<0.001, ‘**’ for p-values<0.01, ‘*’ for p-values<0.05, ‘.’ for p-values<0.1, and ‘ ’ for p-value>=1. Null Deviance: 8.502e+09 on 6.133e+09 degrees of freedom. Residual Deviance: 2.215e+06 on 6.133e+09 degrees of freedom

library in R (Morris et al., 2008).

4. Results

Table 3 presents the results of our final ERGM model. It shows the estimates, standard errors, z-values, and p-values for each factor. The **edges** factor is the baseline factor. As expected, it has a negative estimate, indicating that as the number of edges increases, the likelihood of a connection (i.e., being retweeted) decreases. The most interesting result is that the **nodeicov.botscore2** factor has a positive estimate of 1.857348, indicating that bot-like accounts and those that are no longer available (potentially due to violations or automation) are more likely to be retweeted than accounts with lower botscore2 values. Furthermore, the fact that the **mutual** factor has a positive estimate of 3.583144 suggests that there is a tendency of accounts in this network to retweet each other. This may be a sign of coordinated sharing behavior among bot-like accounts.

All three factors in this model show statistical significance (with $p<1e\text{-}04$). The residual deviance is notably smaller than the null deviance, indicating a good fit of the model to the data. We also observed the reduction of the AIC and BIC values from the baseline model that only included the **edges** factor to the final model presented in Table 3. The number of followers, tweets since joining Twitter, account age, and verification status were found to be statistically not significant and therefore excluded from the final model. The goodness of fit test and MCMC diagnostics confirmed (based on 30,000 simulated networks) the final model generates networks that are structurally similar to the observed network (see Figure 4).

5. Discussion and Conclusion

The study reveals the significant role of automated accounts in disseminating CTs about Bill Gates amid the COVID-19 crisis. In particular, tweets pertaining

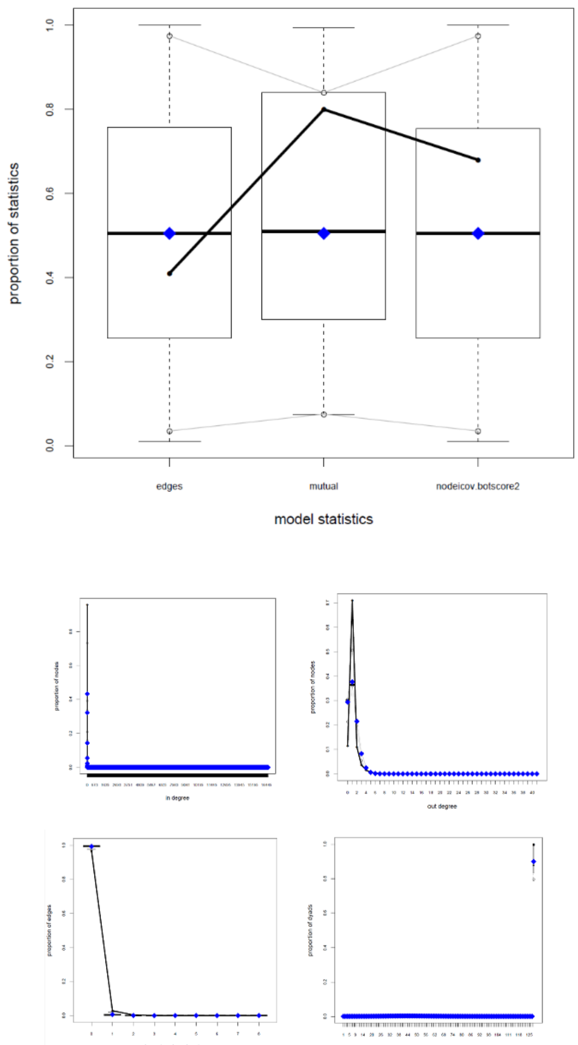


Figure 4: Goodness of fit diagnostics

to “Bill Gates” CTs, originally posted by probable bots, were reshared by a greater number of accounts compared to similar content shared by other accounts (more likely than by chance alone). Additionally, the existence of reciprocity within the network implies that these accounts may be more inclined to retweet each other in a possibly coordinated fashion.

Our findings confirm previous studies on the impact of bots in disseminating unreliable information about COVID-19 on social media. For instance, researchers found that 45% of tweets related to the topic of COVID-19 originated from bot accounts even though Twitter posed various rules to filter out such content (Allyn, 2020). Another study analyzing 43 million COVID-19 related tweets found that bots were responsible for spreading CTs, such as the virus originating in a lab or being a biological weapon (Ferrara, 2020). Our study not only validates the previous research, but also shows that CTs shared by bot-like accounts were frequently re-shared by other accounts, including other probable bots.

These results suggest that Twitter could potentially limit the spread of highly viral CTs by identifying automated accounts involved in their dissemination. However, our findings reveal that Twitter’s content moderation efforts were inadequate in preventing the spread of harmful CTs. As of December 2020, only approximately 8% of the 78k accounts that disseminated CTs related to Bill Gates were suspended or restricted. Ha et al. (2022) found similar results regarding the spread of CTs related to Bill Gates during the COVID-19 pandemic on YouTube, concluding that design and policy changes are necessary to counteract the dissemination of harmful CTs on social media.

To exacerbate the issue, Twitter is significantly reducing access to its API (as of May 2023), which is the primary means of data collection for this study. This move jeopardizes future studies similar to this one. Additionally, the termination of free access to Twitter’s API means that independent researchers will not have access to research tools such as Botometer.

The study has several limitations that suggest avenues for future research. First, the process of constructing queries to identify tweets related to “Bill Gates” CTs relied on a set of pre-determined keywords. In the future, we intend to rely on a topic modeling technique, such as LDA, to enhance the generalizability of the approach. Second, the dataset collected for this study only captures how CTs were disseminated on Twitter. However, CTs are not limited to a single platform; they can shift and migrate across different social media platforms. Conducting further research on various platforms will offer a more comprehensive understanding of this phenomenon.

6. Acknowledgements

This research is funded in part through grants from the Canadian Institutes of Health Research (PIs: Veletsianos, Hodson, Gruzd) and the Natural Sciences and Engineering Research Council of Canada (PIs: Gruzd, Ghenai).

References

Abalakina-Paap, M., Stephan, W. G., Craig, T., & Gregory, W. L. (1999). Beliefs in conspiracies. *Political Psychology*, *20*(3), 637–647.

Alhajj, R., & Rokne, J. (2014). Encyclopedia of social network analysis and mining.

Allington, D., & Dhavan, N. (2020). The relationship between conspiracy beliefs and compliance with public health guidance with regard to COVID-19. *Centre for Countering Digital Hate: London, UK*.

Allington, D., Duffy, B., Wessely, S., Dhavan, N., & Rubin, J. (2020). Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychological Medicine*, 1–7.

Allyn, B. (2020). *Researchers: Nearly half of accounts tweeting about coronavirus are likely bots*. https://www.npr.org/sections/coronavirus-live-updates/2020/05/20/859814085/researchers-nearly-half-of-accounts-tweeting-about-coronavirus-are-likely-bots (accessed: 04.05.2023)

Araujo, T., Neijens, P., & Vliegenthart, R. (2017). Getting the word out on twitter: The role of influentials, information brokers and strong ties in building word-of-mouth for brands. *International Journal of Advertising*, *36*(3), 496–513.

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized health communication: Twitter bots and russian trolls amplify the vaccine debate. *American Journal of Public Health*, *108*(10), 1378–1384.

Bruns, A., Harrington, S., & Hurcombe, E. (2020). 'corona? 5g? or both?': The dynamics of COVID-19/5g conspiracy theories on facebook. *Media International Australia*.

Cassese, E. C., Farhart, C. E., & Miller, J. M. (2020). Gender differences in COVID-19 conspiracy theory beliefs. *Politics & Gender*, 1–12.

CDC. (2022). *History of smallpox*. https://www.cdc.gov/smallpox/history/history.html (accessed: 04.10.2022)

Chung, K. (2009). The phenomenon of the conspiracy theory has contributed substantially to the belief that vaccination is the direct cause of autism. *The Journal of the American Osteopathic Association*, (7), 384–386.

Clarke, S. (2002, June). *Conspiracy theories and conspiracy theorizing*.

Colgrove, J., & Bayer, R. (2005). Could it happen here? vaccine risk controversies and the specter of derailment. *Health Affairs (Project Hope)*, *24*(3), 729–739.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016). BotOrNot: A system to evaluate social bots. *Proceedings of the 25th International Conference Companion on World Wide Web*, 273–274.

Desta, T. T., & Mulugeta, T. (2020). Living with COVID-19-triggered pseudoscience and conspiracies. *International Journal of Public Health*, *65*(6), 713–714.

Dornan, C. (2020). *Science disinformation in a time of pandemic*.

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding conspiracy theories. *Political Psychology*, *40*, 3–35.

Dubé, E., Vivion, M., & MacDonald, N. E. (2015). Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: Influence, impact and implications. *Expert Review of Vaccines*, *14*(1), 99–117.

Enders, A. M., Smallpage, S. M., & Lupton, R. N. (2020). Are all 'birthers' conspiracy theorists? on the relationship between conspiratorial thinking and political orientations. *British Journal of Political Science*, *50*(3), 849–866.

Ferrara, E. (2020). What types of covid-19 conspiracies are populated by twitter bots? *First Monday*.

Ferrara, E., Cresci, S., & Luceri, L. (2020). Misinformation, manipulation, and abuse on social media in the era of COVID-19. *Journal of Computational Social Science*, *3*(2), 271–277.

Freeman, D., Waite, F., Rosebrock, L., Petit, A., Causier, C., East, A., Jenner, L., Teale, A.-L., Carr, L., Mulhall, S., Bold, E., & Lambe, S. (2020). Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in england. *Psychological Medicine*, 1–13.

Georgiou, N., Delfabbro, P., & Balzan, R. (2020). COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. *Personality and Individual Differences*, *166*, 110201.

Goertzel, T. (2010). Conspiracy theories in science. *EMBO reports*, *11*(7), 493–499.

Goldberg, Z. J., & Richey, S. (2020). Anti-vaccination beliefs and unrelated conspiracy theories. *World Affairs*, *183*(2), 105–124.

Gruzd, A., & Mai, P. (2020). Going viral: How a single tweet spawned a COVID-19 conspiracy theory on twitter. *Big Data & Society*, *7*(2).

Gruzd, A., & Tsyganova, K. (2015). Information wars and online activism during the 2013/2014 crisis in ukraine: Examining the social structures of pro- and anti-maidan groups. *Policy & Internet*, *7*(2), 121–158.

Ha, L., Graham, T., & Grey, J. W. (2022). Where conspiracy theories flourish: A study of youtube comments and bill gates conspiracy theories. *Harvard Kennedy School Misinformation Review*, *3*(5), 1–12.

Haythornthwaite, C. (2002). Strong, weak, and latent ties and the impact of new media. *The Information Society*, *18*(5), 385–401.

Hotez, P. (2019). America and europe's new normal: The return of vaccine-preventable diseases. *Pediatric Research*, *85*(7), 912–914.

Hotez, P. J. (2020). COVID19 meets the antivaccine movement. *Microbes and Infection*, *22*(4), 162–164.

Imhoff, R., & Lamberty, P. (2020). A bioweapon or a hoax? the link between distinct conspiracy beliefs about the coronavirus disease (COVID-19) outbreak and pandemic behavior. *Social Psychological & Personality Science*.

Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S., & Lupu, Y. (2020). The online competition between pro- and anti-vaccination views. *Nature*, *582*(7811), 230–233.

Jolley, D., & Douglas, K. M. (2014). The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *British Journal of Psychology*, *105*(1), 35–56.

Jolley, D., & Paterson, J. L. (2020). Pylons ablaze: Examining the role of 5g COVID-19 conspiracy beliefs and support for violence. *British Journal of Social Psychology*, *59*(3), 628–640.

Kaufman, A. B., & Kaufman, J. C. (2018). *Pseudoscience: The conspiracy against science*. The MIT Press.

Klofstad, C. A., Uscinski, J. E., Connolly, J. M., & West, J. P. (2019). What drives people to believe in zika conspiracy theories? *Palgrave Communications*, *5*(1), 36.

Kou, Y., Gui, X., Chen, Y., & Pine, K. (2017). Conspiracy talk on social media: Collective sensemaking during a public health crisis. *Proceedings of the ACM on Human-Computer Interaction*, *1*, 61:1–61:21.

Lai, C.-H., She, B., & Tao, C.-C. (2017). Connecting the dots: A longitudinal observation of relief organizations' representational networks on social media. *Computers in Human Behavior*, *74*, 224–234.

McCauley, C., & Jacques, S. (1979). The popularity of conspiracy theories of presidential assassination: A bayesian analysis. *Journal of Personality and Social Psychology*, *37*(5), 637–644.

Miller, J. M. (2020). Psychological, political, and situational factors combine to boost COVID-19 conspiracy theory beliefs. *Canadian Journal of Political Science/Revue canadienne de science politique*, *53*(2), 327–334.

Miller, S. (2002). Conspiracy theories: Public arguments as coded social critiques: A rhetorical analysis of the twa flight 800 conspiracy theories. *Argumentation and Advocacy*, *39*(1), 40–56.

Morris, M., Handcock, M. S., & Hunter, D. R. (2008). Specification of exponential-family random graph models: Terms and computational aspects. *Journal of statistical software*, *24*(4), 1548–7660.

Nerghes, A., Kerkhof, P., & Hellsten, I. (2018). Early public responses to the zika-virus on YouTube: Prevalence of and differences between conspiracy theory and informational videos. *Proceedings of the 10th ACM Conference on Web Science*, 127–134.

Oleksy, T., Wnuk, A., Maison, D., & Łyś, A. (2021). Content matters. different predictors and social consequences of general and government-related conspiracy theories on COVID-19. *Personality and Individual Differences*, *168*.

Olive, J. K., Hotez, P. J., Damania, A., & Nolan, M. S. (2018). The state of the antivaccine movement in the united states: A focused examination of nonmedical exemptions in states and counties. *PLOS Medicine*, *15*(6).

Oliver, J. E., & Wood, T. (2014). Medical conspiracy theories and health behaviors in the united states. *JAMA internal medicine*, *174*(5), 817–818.

Qazvinian, V., Rosengren, E., Radev, D. R., & Mei, Q. (2011). Rumor has it: Identifying misinformation in microblogs. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 1589–1599.

Rashed, M., Piorkowski, J., & McCulloh, I. (2019). Evaluation of extremist cohesion in a darknet forum using ERGM and LDA. *Proceedings of the 2019 IEEE/ACM International Conference* *on Advances in Social Networks Analysis and Mining*, 899–902.

Reinhart, R. (2022). *More americans now willing to get covid-19 vaccine*. YouGov. (accessed: 02.05.2022).

Romer, D., & Jamieson, K. H. (2020). Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. *Social Science & Medicine*.

Sanders, W. S., Wang, Y. J., & Zheng, Q. (2019). Brand's social media presence as networks: The role of interactivity and network centrality on engagement. *Communication Research Reports*, *36*(2), 179–189.

Sayyadiharikandeh, M., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2020). Detection of novel social bots by ensembles of specialized classifiers. *Proceedings of the 29th ACM international conference on information & knowledge management*.

Schuchard, R., Crooks, A. T., Stefanidis, A., & Croitoru, A. (2019). Bot stamina: Examining the influence and staying power of bots in online social networks. *Applied Network Science*, *4*(1), 55.

Sha, Z., Huang, Y., Fu, J. S., Wang, M., Fu, Y., Contractor, N., & Chen, W. (2018). A network-based approach to modeling and predicting product coconsideration relations. *Complexity*, *2018*, 1–14.

Shapiro, G. K., Holding, A., Perez, S., Amsel, R., & Rosberger, Z. (2016). Validation of the vaccine conspiracy beliefs scale. *Papillomavirus Research*, *2*, 167–172.

Starbird, K. (2017). Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on twitter, 10.

Swami, V., & Coles, R. (2010). *The truth is out there: Belief in conspiracy theories*. https://www.bps.org.uk/psychologist/truth-out-there (accessed: 09.10.200)

ten Thij, M., Ouboter, T., Worm, D., Litvak, N., van den Berg, H., & Bhulai, S. (2014). Modelling of trends in twitter using retweet graph dynamics. In A. Bonato, F. C. Graham, & P. Prałat (Eds.), *Algorithms and models for the web graph* (pp. 132–147). Springer International Publishing.

Tindall, D. B., & Wellman, B. (2001). Canada as social structure: Social network analysis and canadian sociology. *Canadian Journal of Sociology/Cahiers canadiens de sociologie*, 265–308.

Twitter. (2022). *The twitter rules*. https://help.twitter.com/en/rules-and-policies/twitter-rules (accessed: 05.10.2022)

Uscinski, J. E. (2018, November 12). *Conspiracy theories and the people who believe them*. Oxford University Press.

Van Prooijen, J.-W., & Douglas, K. M. (2017). Conspiracy theories as part of history: The role of societal crisis situations. *Memory studies*, *10*(3), 323–333.

WHO. (2022). *Coronavirus disease (covid-19)*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19 (accessed: 04.05.2022)

Wiysonge, C. S., Ndwandwe, D., Ryan, J., Jaca, A., Batouré, O., Anya, B.-P. M., & Cooper, S. (2022). Vaccine hesitancy in the era of covid-19: Could lessons from the past help in divining the future? *Human vaccines & immunotherapeutics*, *18*(1), 1–3.

Wolfe, R. M., & Sharp, L. K. (2002). Anti-vaccinationists past and present. *BMJ (Clinical research ed.)*, *325*(7361), 430–432.

Wood, M. J. (2018). Propagating and debunking conspiracy theories on twitter during the 2015–2016 zika virus outbreak. *Cyberpsychology, Behavior, and Social Networking*, *21*(8), 485–490.

Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and alive: Beliefs in contradictory conspiracy theories. *Social Psychological and Personality Science*, *3*(6), 767–773.

Yang, K.-C., Torres-Lugo, C., & Menczer, F. (2020). Prevalence of low-credibility information on twitter during the COVID-19 outbreak. *arXiv:2004.14484 [cs]*.

Yang, K.-C., Varol, O., Hui, P.-M., & Menczer, F. (2020). Scalable and generalizable social bot detection through data selection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(1), 1096–1103.

YouGov. (2016). *Yougov survey results*. http://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/463g4e5e0e/LBCResults_160614_EUReferendum_W.pdf (accessed: 04.20.2022)