Analysis of COVID-19 Misinformation: Origin and Cure Narratives

Anika Halappanavar
Richland High School, Richland, WA, USA
anika.halappana@gmail.com

ABSTRACT

The interplay between three critical aspects — biology of the virus, human behavior, and government policies — significantly determine the spread and impact of any given epidemic. With over 3 million deaths worldwide, COVID-19 has exposed the vulnerabilities of modern societies to not only an epidemic outbreak, but the corresponding spread of misinformation on social media platforms. In order to understand the human behavior that facilitated the rapid spread of COVID-19, we study the spread of misinformation through social media. While the biology governing the spread of a virus through a population is well understood, the spread of misinformation is relatively less understood, and therefore difficult to control.

By selecting different narratives on origin of COVID-19 and its fake cures, we collected data from Reddit posts and their corresponding comments. Using metrics from social media analysis, we characterize different aspects of content dynamics such as number of shares over time, lifetime of a topic, and speed of spread, for the chosen narratives. We associate the peaks in online activity to corresponding peaks in general media for a given topic to identify the sources driving the creation and spread of misinformation. We also compute graph-theoretic measures of the graphs representing user interactions to identify key users in a narrative. We then use the empirical evidence to design a set of recommendations to minimize the spread of misinformation in social media, with the ultimate goal of providing a fast and effective response to epidemic outbreaks in the future.

KEYWORDS

Misinformation, COVID-19, social media analytics, graph analytics

1 INTRODUCTION

People form their perceptions and opinions about several aspects of their life under the influence of people around them [17]. Influence can be measured in terms of the information flow that happens from one person to another. Misinformation can be simply defined as reconfigured content, where existing and often true information is spun, twisted, recontextualized, or reworked to be misleading, whether intentionally or unintentionally. Further, disinformation is deliberately deceptive misinformation, for example, propaganda meant to mislead people into believing outright lies [22, 23]. Misinformation has the power to influence large populations of people based on what they see, hear, or read. Misinformation is now enabled by heavily unregulated and ubiquitous online social media platforms, which are primarily driven by business needs to maximize user engagement [5].

In order to study the role and impact of misinformation in the ongoing COVID-19 pandemic, we studied several COVID-19 related origin and cure narratives that were demonstrated to be untrue. In particular, we studied the false or fake origin narratives involving Fauci, Soros, Bioweapon, Gates and 5G as the source of COVID-19 pandemic. We also studied narratives involving Hydroxychloroquine, Silver solution and Sputnik as the false or fake cures for COVID-19. We collected and analyzed data from a social media platform called Reddit, and collected data from various posts, comments, and shares from numerous users for the period beginning March 2020 to March 2021. Using several metrics from social media analytics and graph analytics, we present the insights on the data in support of a set of actions that can be taken towards controlling or minimizing the spread of misinformation.

We make the following contributions:

- Characterize the dynamics of information spread using several metrics such as lifetime, speed, Gini index and Palma ratio (§3.1). We demonstrate that there are several differences between narratives that can used to design strategies to control the viral spreading of misinformation (§4.1).
- Characterize user interaction using graph-theoretic modeling and metrics such as average path length, clustering coefficient and modularity (§3.2). We demonstrate that there are a few users that can be targeted for removal to break the spread of misinformation (§4.2).
- Supported by empirical evidence, we present several strategies for minimizing and controlling the spread of misinformation (§5).

To the best of our knowledge, this is the first study to explore data from Reddit and combine metrics from social media analytics and graph analytics to support strategies for minimizing the spread of misinformation. With $3,557,073$ deaths worldwide as of June 1, 2021, the impact of misinformation on the human society should not be underestimated.

2 RELATED WORK

Conroy et al. observed from a meta-analysis of more than 200 experiments that humans have just 4% better than chance probability of detecting misinformation [7]. The human susceptibility to recognize and pinpoint deceptive texts is low and thus indicating that humans are more likely to fall victim to the misinformation they see online. Scheufele and Krause observe: "Being misinformed is often conceptualized as believing in incorrect or counterfactual claims. However, the line between being misinformed or uninformed —

epiDAMIK 2021, Aug 15, 2021, Virtual

ACM ISBN 978-1-xxxx-XXXX-X…$15.00
https://doi.org/xx.xxxx/xxxxxxxxx.xxxxxxx

that is, simply not knowing — has long been blurry in different literature" [21]. To fully understand the effects of misinformation on society, researchers have successfully conducted studies and formulated the potential risks associated with it. Hopf et al. observe that false information can not only alter attitudes and policies on critical ecological, social and political issues, but in some extreme cases, it can render entire populations at local, national, and even global scales at risk of severe harm [13]. With over 3.5 million deaths worldwide as of June 2021, COVID-19 misinformation can be considered one such disaster of massive proportions. Researchers have found the influence of misinformation to be so profound that people are willing to die for a blatantly false cause. A real-world example is COVID-19 in which misinformation is so persuasive that patients dying from the disease still claim that COVID-19 is a hoax [1]. Another example of a viral misinformation spread is the belief that 5G wireless technology was the origin of COVID-19. Across the United Kingdom, about seven 5G mobile towers were burned and forty employees of one UK carrier were physically or verbally attacked [18]. The correlational and causal evidence found in several studies suggests that emotion increases belief in misinformation. Studies have shown that people who engage in rational thinking are less likely to fall victim to misinformation [15]. We build on the observations from these fundamental studies to hypothesize and select COVID-19 topics for exploration in this paper.

A closely related topic to this work is machine learning, which is one of the most rapidly developing fields in computer science at solving diverse real-world problems [19]. With successful application of machine learning methods in a wide variety of impactful problems, there is a hope of putting an end to the intake of fake news. Ghenai and Mejova studied Twitter tweets posting misinformation on the Zeka Virus using automated Latent Dirichlet Allocation (LDA) for topic discovery as well as a high-precision expert-led Information Retrieval approach to identify relevant tweets in the stream. Using crowdsourcing, they distinguished between rumor and clarification tweets, which they then used to build automatic classifiers [10]. This is one particular example of machine learning-based approach to identify and address misinformation found in social media on the narrative of Zika virus outbreak. In a study analyzing reliable and unreliable health-related articles from multiple Chinese online social media sites it was found there were differences in writing style, text topic, and feature distribution by both intuitive and statistical analysis. A Health-related Misinformation Detection framework (HMD) that included a feature-based method and a text-based method for detecting unreliable health-related information was proposed to help combat misinformation found in healthcare [14]. Additionally, there is a possibility of the use of news feed algorithms to down-rank content from sources that users rate as untrustworthy. However, there is a strong justification that laypeople will be handicapped by reasoning and lack of expertise, and therefore, unable to identify misinformation content [9]. It is evident that machine learning models significantly help to identify, categorize, and tackle misinformation. However, the issue lies in the lack of adoption of these technologies by social media platforms.

To the best of our knowledge, this is the first study on data from Reddit for the specific topics on origin and cure. It is also the first study to combine graph-theoretic modeling of social dynamics with content dynamics for the same dataset.

3 METHODS

COVID-19 has exposed the vulnerabilities of modern societies to not only an epidemic outbreak, but also to the corresponding spread of misinformation on social media platforms. Misinformation is often triggered, conceptualized, and spread through social media platforms and significant political events. As Allen observed "Since the 2016 U.S. presidential election, the deliberate spread of online misinformation, in particular on social media platforms such as Twitter and Facebook, has generated extraordinary interest across several disciplines" [2]. With the monumental economic and human loss, the spread of misinformation on social media platforms has become an urgent problem of significant social interest. In order to hold responsible parties accountable and recommend effective strategies to contain the spread of misinformation, we proposed the following research questions for this study: (*i*) Can we quantify the spread of misinformation? (*ii*) Can we quantitatively characterize the users who spread misinformation? (*iii*) Can we study the differences in the spread of different narratives? (*iv*) Can we observe correlation between the narratives and the people who spread the narratives?

In this section, we provide a brief overview of the methods we use in this work to address the research questions that we posed. We divide the methods into two categories: (*i*) Content dynamics, to quantify different aspects of the content of the narratives, and (*ii*) Social dynamics, to quantify the human interactions that govern the content.

Data: We collect the data from Reddit using the `PushshiftAPI`() interface and search for specific queries using a given origin or cure narrative [4]. We filter the queries using specific terms such as `create, created, engineered, made, developed, from, cause, causes`, or `caused` for origin, and `cures, cured, curing` for the cure narratives. We combine several variations for `Gates, Soros, Fauci, 5G` for origin, and `Hydroxychloroquine, silver, and Sputnik` for cure narratives. We collect the data from 03/2020 to 03/2021 time period.

3.1 Content Dynamics

We process the data from Reddit using a Python-based social media analytics tool named SocialSim [11, 20]. In particular, we compute the folllowing metrics for all the origin and cure narratives studied in this work, including the generation of graph-theoretic models from the raw data:

- *Shares*: total number of spread actions for a specific topic.
- *Audience/Users*: total number of users who spread the information.
- *Lifetime*: how long the spreading activity continues for using a given metric of time.
- *Speed*: how quickly the number of shares grows over time.
- *Gini Index or Gini ratio*: represents inequality in a given population. For our study, it measures the imbalance in shares per user (for both initiation and participation).
- *Palma Ratio*: is defined as the ratio of the share of the top 10% to the bottom 40% of users in the population. For our study, it

measures the imbalance in shares per user (for both initiation and participation).

3.2 Social Dynamics

Connections between people, either online or in-person, can be found throughout the world, regardless of where one lives. Vicario observed that "The wide availability of user-provided content easily available through online social media facilitates the aggregation of people around common interests, worldviews, and narratives" [8]. For our study, that common interest lies with COVID-19 misinformation. The intake of misinformation is universal and as users from all over the globe partake in the spreading of false narratives it raises the severity of the issue. Thus, studying and characterizing the social dynamics of the spread becomes critically important.

Reddit data is collected using the `PushshiftAPI()` library. Users are assigned unique tokens for anonymity by Reddit. In the graph representation of the data, $G = (V, E)$, the vertex set V represents unique users, and the edge set E represents binary relations between two users if one of the users comments on the post (or a comment) by the other user. The edges are directed based on the directions of the action. For the metrics presented in §4, we treat the graph as undirected. We collect the following graph-theoretic measures using Gephi [3]. We also use Gephi for the two renderings presented in §4.

- *Degree distribution*: the distribution of vertex-degrees in a graph $G = (V, E)$. We denote the average degree as δ_{avg}.
- *Components*: the number of connected components in G. Since we model G as undirected, we report weakly connected components (where a path between any two vertices in a component exists).
- *Modularity score*: measures how well the nodes cluster with other vertices in their own cluster (or community) for a given clustering of a graph. The value ranges from 0 to 1, with higher values indicating the existence of a good community structure.
- *Number of communities (C)*: the number of communities in G.
- *Average path length* : Average shortest path length between all pairs of vertices in G.
- *Diameter (Φ)*: The longest shortest path between any two vertices in G.
- *Clustering coefficient*: of a vertex is the ratio of the actual number of edges among its neighbors to the total possible number of edges between them. Average clustering coefficient is the average of clustering coefficients all the vertices in G.
- *Triangles*: the number of triangles in G, $\{(a, b), (b, c), (c, a)\}$ exist in E for any three vertices a, b, and c in V.

4 RESULTS & ANALYSIS

We present the key results from our experiments in this section. Details of different metrics used are provided in §3.

4.1 Content Dynamics

The key insights we gain from the analysis of content revels the amount of information disseminated, the number of people sharing the information, and how quickly information gets shared. These are important metrics that are used in the state-of-the-art research and analysis of social media [11]. We summarize the information representing content dynamics in Table 1 for the five origin narratives, and the three cure narratives. We observe that all the topics are discussed for roughly the same lifetime, but the number of shares and user reach are different. Topics getting different levels of attention with respect to each other. Further insight is provided in Figure 2 for four of these topics. Of particular importance is the rapid drop-off for the `Silver` narrative, which coincided with activities in the real world. Speed — the average number of shares per day — for `Hydroxy` (short for Hydroxychloroquine) is significantly higher than `Silver`. This coincided with the multiple news outbreaks for `Hydroxy` in the news media. In contrast, the interest in `Silver` had a sudden drop-off. Similarly, there is significant difference between `Fauci` and `Bioweapon` — marking the stark differences between a human subject versus non-human subject and the coinciding narratives in news media. Topic `Fauci` has the maximum number of shares and `Hydroxy` has the maximum number of unique users over time. We refer the reader to a brief video by the author with overlaid pictures of news clips that correspond to the peaks in the user activity [12].

Table 1: Content Dynamics for Origin and Cure Narratives

Topic	Lifetime	#Shares	#Users	Speed
Origin Narratives				
Fauci	389.26	33876	15231	**87.03**
Soros	389	12326	5768	31.69
Bioweapon	365.75	819	422	**2.24**
Gates	357.86	4338	2055	12.12
5g	296.32	4316	2952	14.57
Cure Narratives				
Sputnik	390.77	15463	7880	39.57
Silver	382.37	2415	1573	**6.32**
Hydroxy	373.05	44668	22324	**119.74**

Palma score [6], which captures the imbalances in user engagement, demonstrate that by and large inequality increases, i.e., more people leave after one or few comments. However, `5G` is rather interesting since there is an equal number for initialization and participation, as illustrated in Figure 1. A maximum difference in contrast to `5G` are that of `Hydroxy` and `Silver`.

4.2 Social Dynamics

Graphs capture the complex relationships in social interactions on how people post and comment for a given narrative. User interaction is modeled based on the users who post and those who comment on posts and comments by other users, and thus forming a complex interaction network. As summarized in Table 2, we observe that graph structure remains fairly similar for different topics, but with some distinctions such as largest diameter for `Soros`, where the graph is roughly a third of the size of the graph representing `Fauci`. It can also be noted that the graphs have a large number of connected components due to the lack of users that span across different posts within a given narrative, and the wide spread of the lifetime of a given topic. This can be connected back to the Palma ration presented in §4.1.

The primary goal of using graph-theoretic analysis is to identify the existence of a few key users that can be targeted for control if necessary. The different centrality measure distributions (not

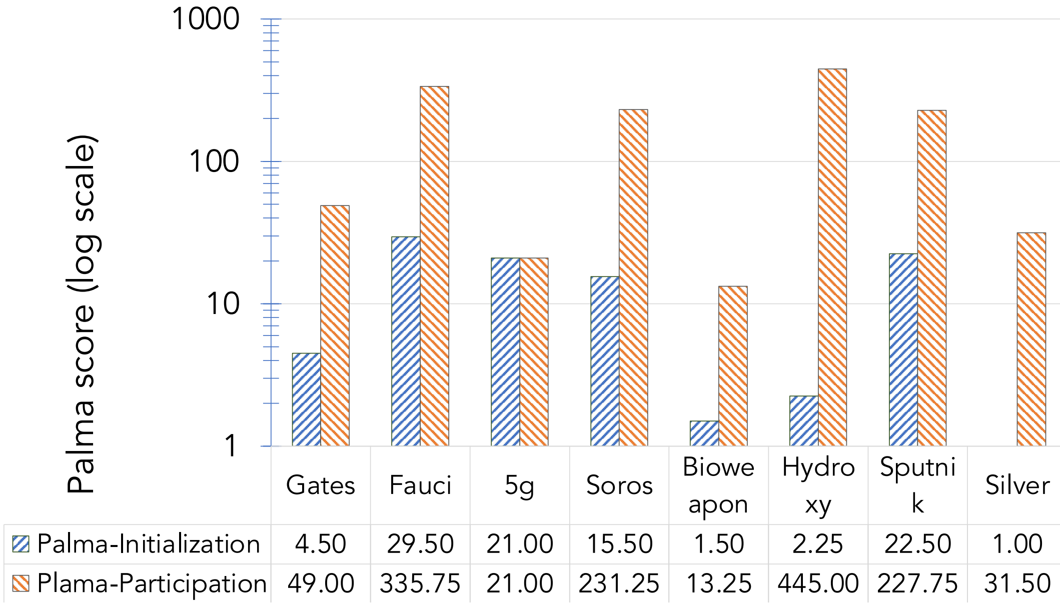


Figure 1: Palma scores for different origin and cure narratives, for initialization and participation activities (posts and comments). 5G offers a drastic difference from Hydroxy and Silver narratives, for initialization and participation activities.

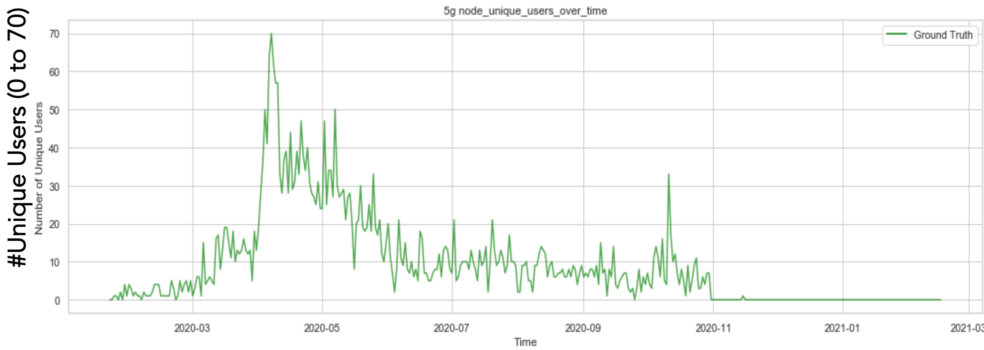
Table 2: Social Dynamics of Origin and Cure Narratives (metrics for 5G are omitted due to parsing errors in the data).

Topic	$\|V\|$	$\|E\|$	δ_{avg}	#Comp	Modularity	#C	Φ	Avg Path	AvgCC	#Tringles
Origin Narratives										
Bioweapon	403	440	2.18	15	0.842	31	11	4.68	0.181	29
Gates	2023	2443	2.42	32	0.877	63	16	4.73	0.199	212
Soros	5544	7114	2.57	76	0.879	118	22	4.79	0.226	888
Fauci	14925	19649	2.63	192	0.911	271	18	5.1	0.186	2354
Cure Narratives										
Silver	1545	1699	2.21	42	0.849	72	13	4.45	0.143	80
Sputnik	3506	3933	2.24	280	0.911	317	23	6.22	0.123	304
Hydroxy	16143	17350	2.15	1401	0.936	1508	25	6.43	0.132	1185

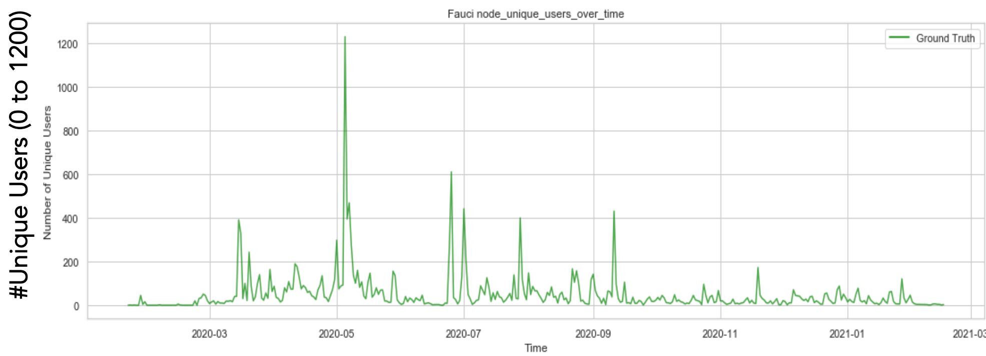
presented in this paper) provide ample evidence of such structure in the data we studied. We now provide empirical evidence in a visual manner using two examples: `Fauci` and `Silver`, as illustrated in Figure 3. The existence of a key central vertices in `Silver` (in Blue) is evident. For several topics, different communities are connected through a few vertices. Thus, supporting the recommendations we make in §5.

5 RECOMMENDATIONS FOR COMBATING MISINFORMATION:

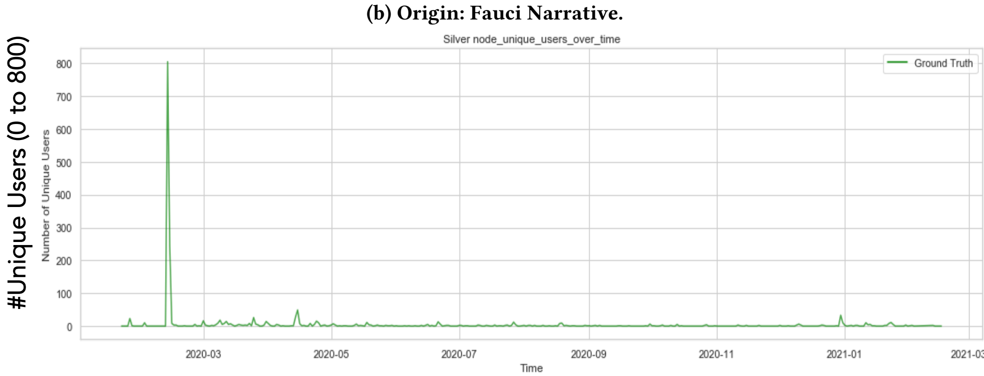
With the amount of misinformation that is spreading on media outlets, social media, and communities and influencing people all over the world, regulations and solutions must be implemented. We hypothesize that a large fraction of misinformation is propagated inadvertently due to a lack of understanding and judgment in the general public. Therefore, we propose that with the assistance of machine learning techniques we can develop software to identify and flag misinformation posts. This will not only help in educating people on how to differentiate facts from fiction, but it will also alert them and thus filter out deceptive and manipulative texts from getting propagated. Further, we propose that social media platforms limit the number of comments and shares of non-factual or suspicious posts that spread misinformation. The platforms must also introduce speed-breaking techniques, such as limiting the total number of shares for a post in a day and number of comments on a post, to slow the viral spread of misinformation as well as reduce its recurrence and longevity. These techniques can help minimize the increasingly large number of manipulative posts and the spread of misinformation over multiple social media platforms. Voluntary action by social media platforms and the general public are largely effective. However, when commercial interests promote misinformation, government regulation becomes necessary. The government can undertake a broad swath of actions including education through mandatory classes on media literacy at elementary,



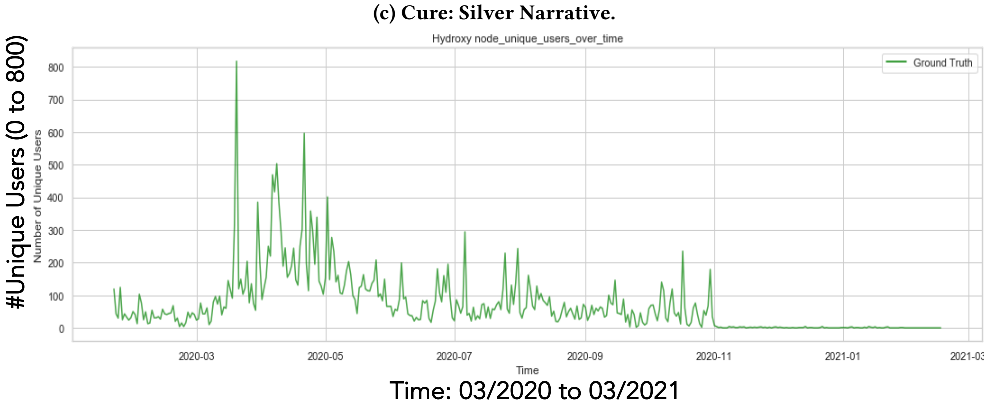
(a) Origin: 5G Narrative.



(b) Origin: Fauci Narrative.

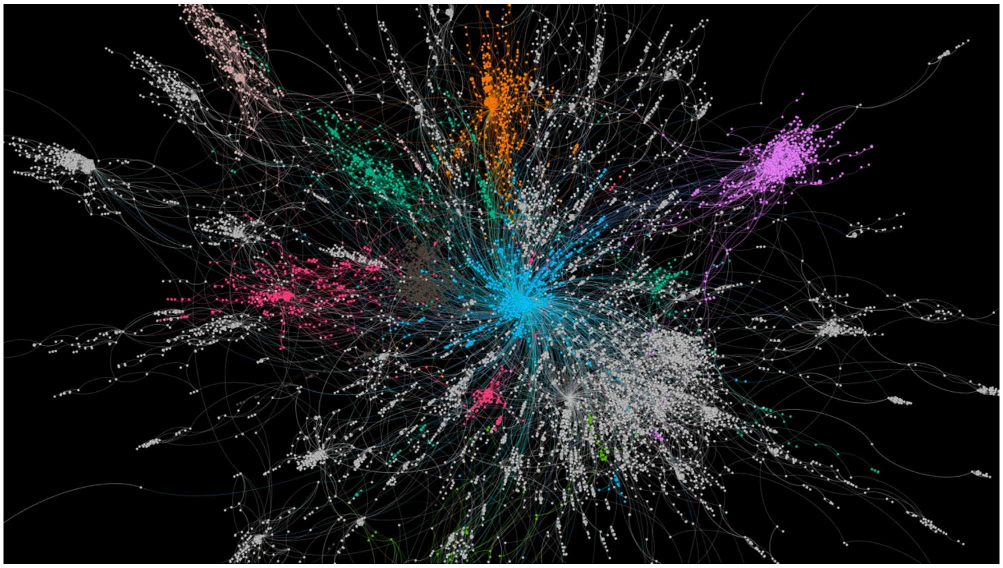


(c) Cure: Silver Narrative.

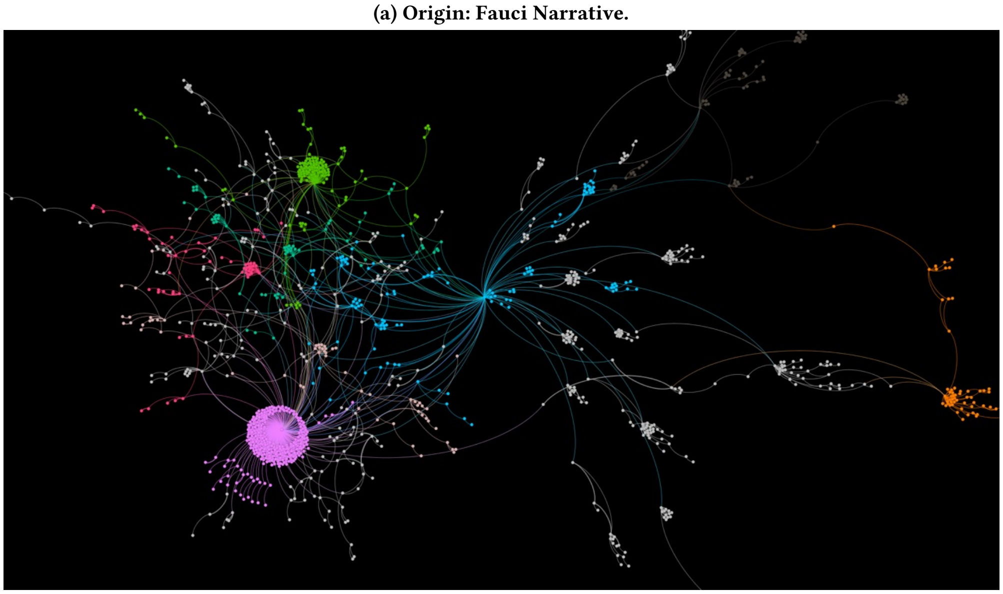


(d) Cure: Hyroxychloroquine Narrative.

Figure 2: Number of unique users over time for origin and cure narratives.



(a) Origin: Fauci Narrative.



(b) Cure: Silver Narrative.

Figure 3: Graph renderings for two user interaction graphs – origin narrative on Fauci, and cure narrative on Silver. Different colors indicate different clusters (or communities), and vertices belonging to the same cluster have the same color.

middle, high school, and college levels. It is important to educate the youth so that they grow into educated adults who can identify and differentiate between truth and carefully constructed lies. Regulation of social media platforms through legal consequences is also an effective strategy. For example, categorizing social media as news media can introduce rigor and control.

Although effective techniques can be proposed in theory, the implementation of these methods will be challenging. For instance, machine learning models suffer from significant challenges such as fairness and bias that originates from the training data that is used, as well as the biases inadvertently introduced in the algorithms by human programmers with their own sets of beliefs [16]. Enforcement of existing technologies by media platforms is limited due to financial concerns and loss of users. Even when some social media platforms implement regulations, users can find loopholes in the system. Such as, creating multiple accounts or making private accounts. Due to the rapidly changing nature of technology, governmental policies to regulate misinformation often lag in regulation. Therefore, it is important that the government takes an active role in implementing new policies to solve these ethical and legal dilemmas. It should also be noted that the solutions proposed are based on scientific evidence found by researchers and psychologists. However, to understand the larger effects of implementing regulations, diverse perspectives need to be considered. Albeit the limitations, it cannot be stressed enough that the menace of misinformation has to be addressed effectively for the benefit of society. Understanding and regulating the spread of misinformation are necessary to reduce its adverse impact on the safety and health of the general public. While the government, social media platforms, and other organizations implement their own measures, a vast amount of power can be exercised by the general public by identifying misinformation and preventing its spread. And the time to act is now.

The key recommendations can be summarized as follows:

- *Containing the spread:* Cut-off strategies such as limiting the number of shares for a given post can be used to contain the spread.
- *Containing the discussion of a post:* Limiting the number of comments for a post can minimize its viral impact, measured both in terms of the speed and lifetime of a topic.
- *Minimizing the peaks:* Introducing speed breaking techniques such as delays and (automated) fact checks can impede the speed at with information spreads.
- *Minimizing lifetime of a narrative:* Implementing techniques, including some of the ones discussed above, to reduce the rate of recurrence can minimize the overall lifetime of a narrative.
- *Counter measures:* Effectively spreading positive and accurate messages can counteract the overall speed and lifetime of a narrative. Legal actions can also be a significant deterrent.

6 SUMMARY & FUTURE WORK

The spread of misinformation is a complex problem with adverse impact on the society. Using data from a social media platform on several COVID-19 fake origin and cure narratives, we provided empirical evidence in support of a set of recommendations to minimize the spread of misinformation on social media. In particular, we presented the metrics and results on two categories — content dynamics and social dynamics — for the same dataset. Supported by the empirical evidence, it becomes increasingly clear that social media platforms should incorporate regulations to minimize and prevent the spread of misinformation. These actions are vital to lessen the adverse impact on the general health of human society.

In the near future, we plan to advance the study through careful examination of the content of the posts and comments for specific links to news media and facts. In particular, actual user ids can be correlated with the topic of posts and the exact position in the graph structure, including the centrality metrics. We will also extend the study to include other COVID-19 narratives, and data from other social media platforms.

ACKNOWLEDGEMENTS

The work was performed for the 2021 regional, state and international science fair events under the guidance of Drs. Svitlana Volkova, Maria Glenski and Mahantesh Halappanavar. I sincerely thank my three mentors and numerous science fair judges who provided me valuable feedback to improve the work and its presentation. I completed this work during my 10th grade at Richland High School, WA, USA.

REFERENCES

[1] Zara Abrams. 2021. Controlling the spread of misinformation. *Monitor on Psychology* 52, 2 (2021), 44. http://www.apa.org/monitor/2021/03/controlling-misinformation

[2] Jennifer Allen, Baird Howland, Markus Mobius, David Rothschild, and Duncan J. Watts. 2020. Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances* 6, 14 (2020). https://doi.org/10.1126/sciadv.aay3539

[3] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media* 3, 1 (2009), 361–362.

[4] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The Pushshift Reddit Dataset. *Proceedings of the International AAAI Conference on Web and Social Media* 14, 1 (May 2020), 830–839.

[5] J. Scott Brennen, Felix Simon, Philip N. Howard, and Rasmus Kleis Nielsen. 2020. Types, sources, and claims of COVID-19 misinformation. *Reuters Institute for the Study of Journalism* (2020).

[6] Alex Cobham, Andy Sumner, Andrea Cornia, Stefan Dercon, Lars Engberg-pedersen, Martin Evans, Nick Lea, Nora Lustig, Richard Manning, Branko Milanovic, Nuria Molina, Chris De Neubourg, and Gabriel Palma. 2013. PUTTING THE GINI BACK IN THE BOTTLE? 'THE PALMA ' AS A POLICY-RELEVANT MEASURE OF INEQUALITY.

[7] Nadia K. Conroy, Victoria L. Rubin, and Yimin Chen. 2015. Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–4. https://doi.org/10.1002/pra2.2015.145052010082

[8] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113, 3 (2016), 554–559. https://doi.org/10.1073/pnas.1517441113

[9] Ziv Epstein, Gordon Pennycook, and David G Rand. 2019. Will the crowd game the algorithm? Using layperson judgments to combat misinformation on social media by downranking distrusted sources. https://doi.org/10.1145/3313831.3376232

[10] Amira Ghenai and Yelena Mejova. 2017. Catching Zika Fever: Application of Crowdsourcing and Machine Learning for Tracking Health Misinformation on Twitter. In *2017 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, Park City, UT, USA, 518–518. https://doi.org/10.1109/ICHI.2017.58

[11] Maria Glenski, Emily Saldanha, and Svitlana Volkova. 2019. Characterizing Speed and Scale of Cryptocurrency Discussion Spread on Reddit. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 560–570. https://doi.org/10.1145/3308558.3313702

[12] Anika Halappanavar. 2021. *Analysis of COVID-19 Misinformation: Origin and Cure Narratives*. https://youtu.be/iZZBNm-t8o8

[13] Henning Hopf, Alain Krief, Goverdhan Mehta, and Stephen A. Matlin. 2019. Fake science and the knowledge crisis: ignorance can be fatal. *Royal Society Open Science* 6, 5 (2019), 190161. https://doi.org/10.1098/rsos.190161

[14] Yue Liu, Ke Yu, Xiaofei Wu, Linbo Qing, and Yonghong Peng. 2019. Analysis and Detection of Health-Related Misinformation on Chinese Social Media. *IEEE Access* 7 (2019), 154480–154489. https://doi.org/10.1109/ACCESS.2019.2946624

[15] Cameron Martel, Gordon Pennycook, and David G Rand. 2019. Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications* 5, 47 (Sep 2019). https://doi.org/10.1186/s41235-020-00252-3

[16] Cathy O'Neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy* (first edition ed.). Crown, New York.

[17] Sancheng Peng, Yongmei Zhou, Lihong Cao, Shui Yu, Jianwei Niu, and Weijia Jia. 2018. Influence analysis in social networks: A survey. *Journal of Network and Computer Applications* 106 (2018), 17–32. https://doi.org/10.1016/j.jnca.2018.01.005

[18] Corinne Reichert. 2021. *5G coronavirus conspiracy theory leads to 77 mobile towers burned in UK*. CNET. https://www.cnet.com/health/5g-coronavirus-conspiracy-theory-sees-77-mobile-towers-burned-report-says/

[19] Stuart Russell and Peter Norvig. 2009. *Artificial Intelligence: A Modern Approach* (3rd ed.). Prentice Hall Press, USA.

[20] Emily Saldanha, Maria Glenski, and Svitlana Volkova. 2021. *Measuring Information Spread Within and Across Social Platforms*. The 13th International AAAI Conference on Web and Social Media (ICWSM-2019). https://sites.google.com/alumni.nd.edu/icwsm19t3/

[21] Dietram A. Scheufele and Nicole M. Krause. 2019. Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences* 116, 16 (2019), 7662–7669. https://doi.org/10.1073/pnas.1805871115

[22] Li Zeng, Kate Starbird, and Emma S. Spiro. 2016. Rumors at the Speed of Light? Modeling the Rate of Rumor Transmission During Crisis. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*. IEEE, 1969–1978. https://doi.org/10.1109/HICSS.2016.248

[23] Xinyi Zhou and Reza Zafarani. 2020. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *Comput. Surveys* 53, 5 (28 Sept. 2020). https://doi.org/10.1145/3395046 Publisher Copyright: © 2020 ACM.