Identifying the Spread of COVID-19 Misinformation on Twitter: Network Properties and Community Detection

Md Ashraf Ahmed
Ph.D. Candidate
Department of Civil and Environmental Engineering
College of Engineering and Computing
Florida International University
10555 West Flagler Street, EC 2900, Miami, FL 33174
E-mail: mpave002@fiu.edu

Arif Mohaimin Sadri, Ph.D.
Assistant Professor
School of Civil Engineering & Environmental Science
University of Oklahoma
202 W. Boyd St., Norman, OK 73019-1024
E-mail: sadri@ou.edu
(Corresponding Author)

Word Count: 04 tables + 4,199 words = 5,199 word equivalents
Submitted for presentation in Transportation Research Board Conference

ABSTRACT

The spread of misinformation through social media during COVID-19 has created pandemic of information which is often called as "infodemic". The goal of this study is to understand how information is deviating and misinformation is spreading through social media during the COVID-19 pandemic, whereas defining the misinformation is crucial. Dynamics of information is captured by quantifying the network effects on such communication behavior and characterizing how information is exchanged among people who are socially connected online. Twitter data was collected using COVID-19 related keywords (corona, ncov and covid) and analyzed by complex networks (e.g., community detection algorithm), natural language processing techniques, and statistical modeling. A subgraph of Twitter users who follow each other is created to identify and track misinformation propagation. Findings of this study show that users with higher neighbor, more connection and closeness to other users in the network are likely to deviate from original information and users with higher degree and connection with influential users spread information quickly. Besides, users with similar sentiments tend to cluster in the same communities, whereas relatively negative communities are responsible for spread of misinformation. The proposed methodology for identifying misinformation will be useful to public health, transportation, and emergency management agencies for tailoring effective information dissemination policies for users, based on their social network characteristics, activities, and interactions. The proposed techniques, if combined, may also serve as an efficient misinformation detection tool for government agencies, enabling them to respond more quickly and confidently to provide correct information to their citizens.

Keywords: COVID-19, social media, Communication network, Information deviation, Misinformation, AEP35

INTRODUCTION

COVID-19 has been rapidly spreading over the world since December 2019, and there is an increasing amount of discussion about this in both online and traditional media. These conversations about COVID-19 are disseminating through social media, specifically Twitter, i.e., the themes of discussion, where the discussion is coming from, misinformation about the virus shared, and how much of it is connected to Twitter users. Existing literature suggests that there is a meaningful spatiotemporal relationship between information flow through social media and new cases of COVID-19. Besides, discussions about misinformation while connected to poor quality information, their existence is found less influential than other disaster-specific subjects [1]. One survey data (conducted on 21,000 individuals across the United States between Aug. 7 and 26, 2020) showed that 26% survey respondents were likely to believe a false claim about COVID-19. This experiment also stated those who believe COVID-19-related misinformation are also less likely to seek the COVID-19 vaccine as well as wearing mask [2].

Online social media platforms, and specifically microblogging sites such as Twitter, has recently overtaken traditional news broadcasting platforms [3, 4]. Social media is becoming a more prevalent platform for information sharing [5, 6], especially in COVID-19 situation. It is commonly noticed that news stories are first disseminated on Twitter, and then telecasted by the electronic and print media. The validity of information (tweet) is a key issue in Twitter due to its distributed structure and lack of fairness, as well as the compulsion of releasing a newsworthy topic early on Twitter. Due to this, information overload and spread of misinformation, rumor, hoax is also increasing. Rumor is defined as any information spreading in the Twitter world that contradicts information from a reliable source [7].

Misinformation is also defined as an honest mistake, or the intentions are not to blatantly mislead people. For an example, advising other people to eat garlic or gargle with salt water as protection against COVID-19 [8]. Besides, different emergency management and health care organizations (FEMA, CDC, WHO) listed several myths spreading through social media. Such as, one myth is *taking a hot bath both raises body temperature and prevents coronavirus infection*. But the fact is scientific evidence suggests that hot baths can minimally affect body temperature; Temperatures needed to deactivate coronavirus are typically >56°C, which exceed safe bath temperatures.

The main goal of this study is to understand the misinformation propagation in social media through user's network. The objectives are to identify how information is deviating from the original content through social media, how misinformation dissemination is related to information deviation and what are the social network metrics which influence the users to change their perception in a timely manner (quickly or slowly). To achieve these objectives, following research questions are developed to answer through this study-

- *What is the best way to identify the information deviation (change in sentiment, topic) in social media?*
- *Are there significant temporal changes existing in deviated information?*

- *Do we observe misinformation from the information deviation?*
- *Does followers' network (1st degree, 2nd degree neighbor) of social media users influence the spread of misinformation?*

This study uses twitter data during COVID-19 pandemic to understand the characteristics of misinformation propagation through different users. The spread of misinformation may also depend on the social network criterions (retweet, follower etc.) [3, 9] which is also captured by implementing machine learning approach (community detection). This approach also identifies the users who are responsible for the misinformation dissemination. The proposed methodology to identify misinformation may help policy makers, emergency management and health care organizations to develop effective information dissemination strategies to recover from similar public health hazards.

LITERATURE REVIEW

Researchers have been studying the dissemination of incorrect information, or so-called fake news, for several years. Researchers define fake news as "completely manufactured and typically biased content portrayed as factual." Understanding why people share disinformation and how it spreads leads to recommended remedies, a goal that grows in importance as people spend more time on social media platforms [10]. Findings suggest that false information propagates faster than the authentic information [11]. Besides, the person who spread misinformation is most likely distracted rather than biased and the people who perform more analytical thinking are more likely to perceive truth [12]. Some false information disseminates from politicians which might assist them to get more support from mass people [13].

Civil rights groups have been warned of online surveillance of social media chatter by city officers and police departments since 2016. Online chats are analyzed by services like Media Sonar, Social Sentinel, and Geofeedia, which alert police and local leaders to what hundreds of thousands of people are saying. According to law enforcement experts, an AI based tool called Zencity can assist them in combating misinformation. It might be used for mass spying, according to civil liberties organizations. Cities such as Phoenix, New Orleans, and Pittsburgh claim to use the service to counteract misinformation and assess public opinion on issues such as social distancing enforcement and traffic rules [14].

Bursztyn et al. explored the effects of news coverage of the novel coronavirus by the two most widely viewed cable news shows (Hannity and Tucker Carlson Tonight, both on Fox News) on viewers' behavior and downstream health outcomes. Authors found that greater exposure to Hannity relative to Carlson is associated with a greater number of county-level cases and deaths [15]. Zaman et al. developed a systematic study of the problem of finding the source of a rumor in a network. Authors found surprising threshold phenomenon: on trees which grow faster than a line, the estimator always has nontrivial detection probability. On trees that grow like a line, the detection probability will go to zero as the network grows [16].

Kouzy et al. analyzed and rated individual tweets for authenticity against verified and peer-reviewed sources (CDC, WHO) to track misinformation. To compare keywords and hashtags, as well as to identify individual tweets and account attributes, descriptive statistics were used. Informal individual/group accounts had a higher rate of misinformation. More disinformation was spread by tweets from unauthorized Twitter accounts. The lowest percentage of unverifiable information was found in tweets from healthcare/public health accounts [17]. By learning to predict accuracy ratings in two credibility-focused Twitter datasets, a research provides a strategy for automating fake news detection on Twitter. Models trained against crowdsourced workers surpassed models based on journalist ratings and models trained on a combined dataset of both crowdsourcing workers and journalists, according to the study [18].

Another study presented an analysis of the impact of automated accounts, or bots, on opinions in a social network. The opinions were modeled using a variant of the famous DeGroot model, which connects opinions with network structure. Authors found a strong correlation between opinions, based on the network model and tweets of Twitter users discussing the 2016 U.S. presidential election between Hillary Clinton and Donald Trump. The Clinton bots produced almost twice as large a change as the Trump bots, despite being fewer in number [19]. Zaman et al. mentioned that online social networks are often subject to influence campaigns by malicious actors using automated accounts known as bots. Analysis of the behavior of bots in social networks identified that they exhibit heterophily, meaning they interact with humans more than other bots. This property is used to develop a detection algorithm based on the Ising model from statistical physics. The bots are identified by solving a minimum cut problem. The Ising model algorithm can identify bots with higher accuracy while utilizing much less data than other state of the art methods. Findings showed that a limited number of bots can cause significant shifts in the population opinions [20].

METHODOLOGY

People are connected through online social media and interact with each other in many ways. These virtual connections create a complex network of users within social media platforms. This phenomenon is depicted in Figure 1 where the connected people exchange information among themselves and tend to create different groups with similar intention. Over the time, different types of information (positive/negative) begin to spread among the users and information turnover or deviation start to take place. Some of the users tend to influence by others perspective and some of them add new thoughts on existing information as well as some other remain inactive. These activities result in different groups of people who are bonded by trust and similarity in their sentiments. Along with time, theses connections become more stronger as the clustered people have similar types of opinion, eventually creates communities. These communities often represent the people who believe and spread misinformation (hoax community) as well as the people who disseminate information based on facts (authentic community). Besides, there is another group of people who do not express their thoughts and remain quiet (inactive community).

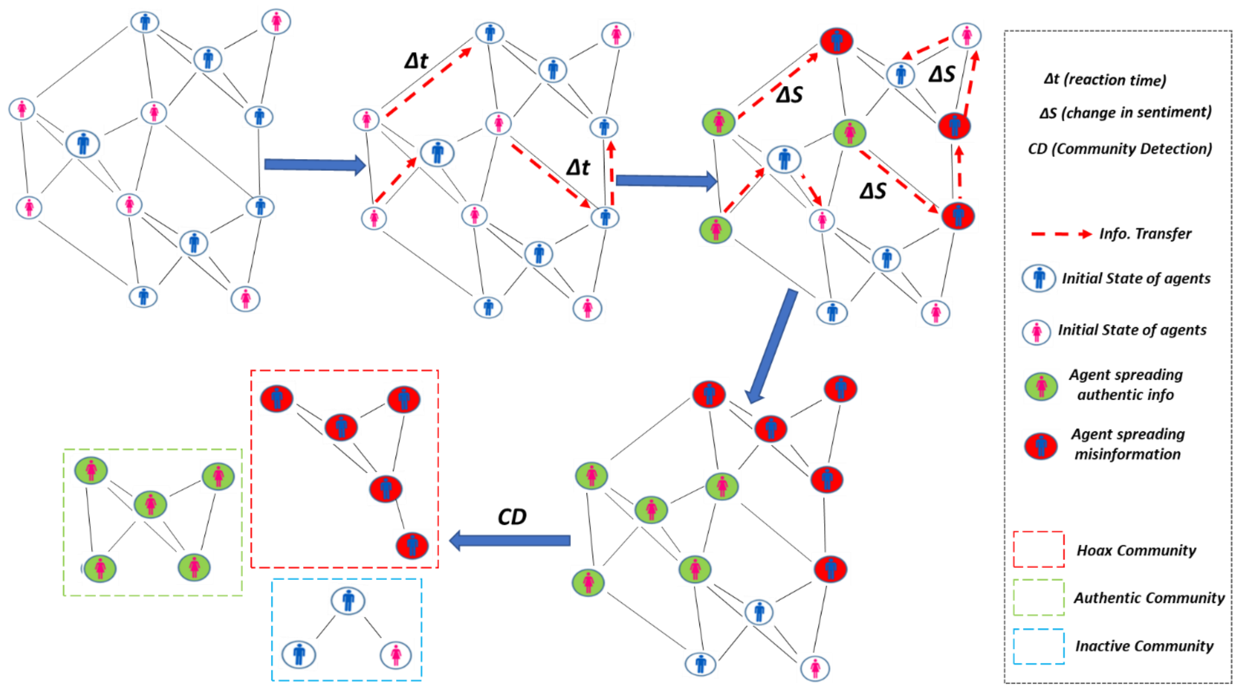

Figure 1. Conceptual framework for identifying misinformation on social media

DATA DESCRIPTION

In this study, social media data (Twitter) was collected by using streaming API which fetched real time tweets from Twitter. The data was collected for one hour, on 5th May 2020 (as it was during the 1st phase of lockdown and the number of cases and deaths due to COVID-19 were also high), using different keywords related to the COVID-19 pandemic (covid, ncov, corona). To understand how the spread of misinformation propagates through online social media platforms, a sample dataset of Twitter consists of 5,376 tweets (number of users: 4,834; number of retweets with quote: 1,371: number of original tweets: 424, number of retweets and replies: 3,581) are analyzed. To capture the information deviation (change in sentiment and topic), retweets with quote (reshared tweet with additional comments) are extracted from the sample dataset. Because retweets with quote have additional comments from Twitter users along with the original tweet (the tweets being retweeted with quotes), which results in different sentiments and topics (as compared to the original tweet). An example of original tweet and retweet with quote is given below from the sample dataset used in this study-

Original Tweet: *We've learned so much during this #Coronavirus crisis; the power of community, which really makes a difference in our life.*
Retweet with quote: *Well said... But no animal age is globally sustainable...Have you read Dr. Sailesh Rao's whitepaper? (We've learned so much during this #Coronavirus crisis; the power of community, who really makes a difference in our life.)*

After filtering the retweets with quote, the individual sentiment values of these as well as corresponding original tweets are computed using sentiment analyzer [21, 22] library in python. Then the change in sentiment of these retweets with quotes and original tweets have identified. The left side of the following Figure 2 is illustrating these steps sequentially.

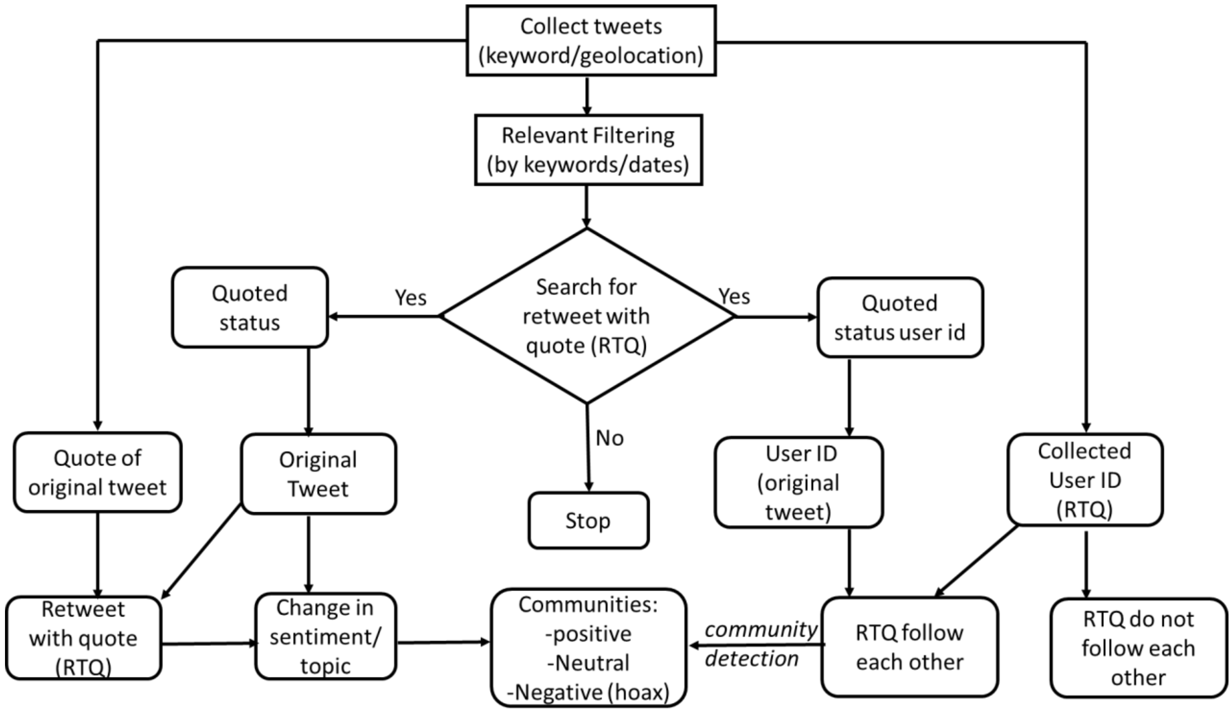

Figure 2. Flowchart of data analysis for identifying misinformation (hoax) on Twitter

Then, the Twitter users who reacted (by retweeting with quotes) to a same original tweet are identified which clustered in different tweet contagion. Community detection algorithm (based on modularity) is applied on these specific tweets contagion to identify different communities where Twitter users follow each other. Finally, observing the discussion topics of these small communities and comparing with the original tweet content may lead to identify the misinformation.

RESULTS

At first the subset of retweets with quote from the sample dataset is extracted. From the sample dataset of 5,376 tweets, 1,371 retweets with quote are found from 1,270 users. Then, a followers' subgraph is created which consists of 1,694 users, where they are following each other or not is checked.

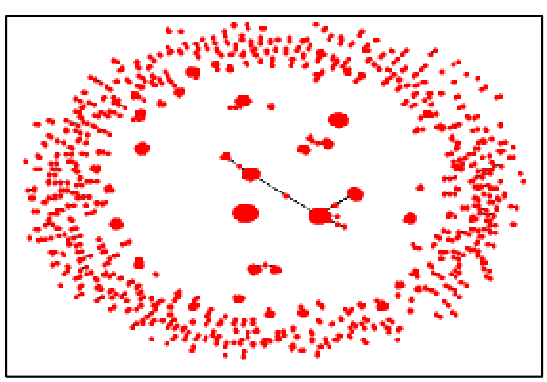

Figure 3. Network of Twitter users-followers who retweeted with quotes

Figure 3 is showing how these twitter users (who retweeted with quotes) are connected with the other users (who tweeted the original tweet) in a followers' network diagram. The people who generated the original tweets are found with more follower than the other users in the network which defines their influential characteristics. The graph properties of follower's subgraph who retweeted with quotes are listed as below:

Table 1 Graph Properties of user-follower network

Number of Nodes	1694 (1270 RTQ, 424 user of original tweet)
Number of Links	1293
Density	0.00045
Average in degree	0.7633
Average out degree	0.7633

Statistical Analysis

Then, to capture the network effects on information deviation (change in sentiment, Δ sentiment) and contagion (how quickly information is changing, Δ time), different network metrics (i.e., degree, centrality) are calculated for all the users (N=1,694) as listed below (Table 2). In following Table 2, descriptive statistics of the undirected unweighted follower's graph (retweeted with quotes) is summarized.

Table 2 Descriptive Statistics

Network Metrics	N	Minimum	Maximum	Mean	Std. Deviation
Δ Sentiment	1694	-0.3182	0.8845	0.2093	0.3751
Δ time	1694	0	110.1437	1.6578	8.1343
degree	1694	1	144.0000	1.5300	5.1260
Indegree centrality	1694	0	0.0851	0.0005	0.0031
Outdegree centrality	1694	0	0.0018	0.0005	0.0003
Average neighbor degree	1694	1	144.0000	27.7644	44.8842
Degree centrality	1694	0.0006	0.0851	0.0009	0.0030
Closeness centrality	1694	0	0.0851	0.0105	0.0154
Betweenness centrality	1694	0	0.0136	0.0000	0.0005
Eigenvector centrality	1694	0	0.7072	0.0054	0.0237
Harmonicon centrality	1694	0	144.0000	19.1339	28.3631
Load centrality	1694	0	0.0136	0.0000	0.0005

To identify the network metrics which influence information deviation, Tobit regression is applied as the change in sentiment values have a range from -2 to 2. From the analysis (Table 3) it is found that the users with higher closeness centrality are less likely to deviate from original information. Besides, users with higher average neighbor degree and harmonic centrality (more connection and closer to other nodes) are more likely to deviate from original information.

Table 3 Tobit regression for information deviation (Δ Sentiment)

					Number of obs. = 1694	
					LR $chi^2(3)$ = 25.33	
					Prob > chi^2 = 0.0000	
Log likelihood = -729.46641					Pseudo R^2 = 0.0171	
Variables	Coefficient	Standard Error	t	P>|t|	[95% Conf. Interval]	
Closeness centrality	-43.3223	11.3735	-3.81	0	-65.6299	-21.0146
Average neighbor degree	0.0014	0.0009	1.5	0.134	-0.0004	0.0032
Harmonic centrality	0.0210	0.0052	4.06	0	0.0109	0.0311
Load centrality	-19.0191	22.0434	-0.86	0.388	-62.2543	24.2161
Constant	0.2247	0.0110	20.44	0	0.2031	0.2463

To model the information contagion, Tobit regression is used again as some of the retweets with quote (56) were generated after more than 15 days of the original tweet, which are not considered for the analysis. Results (Table 4) showed that the users with higher degree and eigenvector centrality (connection with influential nodes) are more likely to spread information quickly. Besides, users with higher average neighbor degree and indegree centrality are less likely to spread information quickly.

Table 4 Tobit regression for information contagion (Δ Time)

					Number of obs. = 1694	
					LR $chi^2(3)$ = 145.14	
					Prob > chi^2 = 0.0000	
Log likelihood = -4129.8156					Pseudo R^2 = 0.0173	
Variables	Coefficient	Standard Error	t	P>|t|	[95% Conf. Interval]	
Degree	1.7248	0.1617	10.67	0	1.4076	2.0419
Average neighbor degree	-0.0159	0.0021	-7.69	0	-0.0199	-0.0118
Indegree centrality	-2938.299	269.8211	-10.89	0	-3467.5180	-2409.0800
Eigenvector centrality	6.2513	4.2529	1.47	0.142	-2.0903	14.5929
Constant	0.0115	0.1402	0.08	0.934	-0.2634	0.2864

Community Detection

From the developed follower's subgraph, individual tweets' information deviation is tracked by gathering retweets with quote for specific original tweets. One negative tweet was tracked from a user where the tweet (*the experts in epidemiology repeated a vaccine #coronavirus may not be ready within 6 months)* discussed about the uncertainty of the vaccine availability. From the sample dataset, 29 retweets with quote are identified for this one tweet and the sentiment value of the original tweet is computed as -0.296 (negative). The user has 616,700 followers and follows only 892 people in Twitter. The information deviation (change in sentiment) for this tweet by retweets with quote is shown below in Figure 4. The original negative tweet was retweeted more negatively by most of the users as well as without changing the sentiment by some other users.

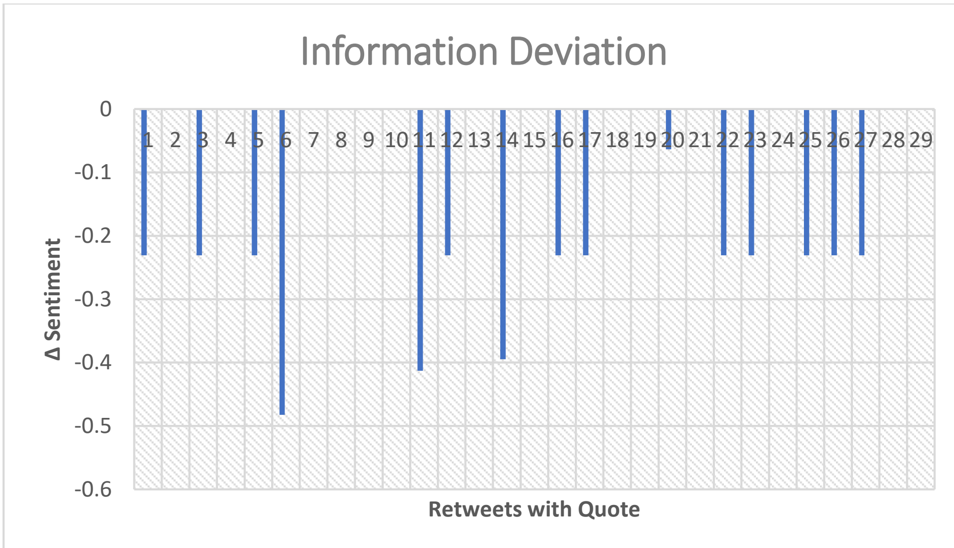

Figure 4. Information deviation (change in sentiment) of retweets with quote of an original tweet

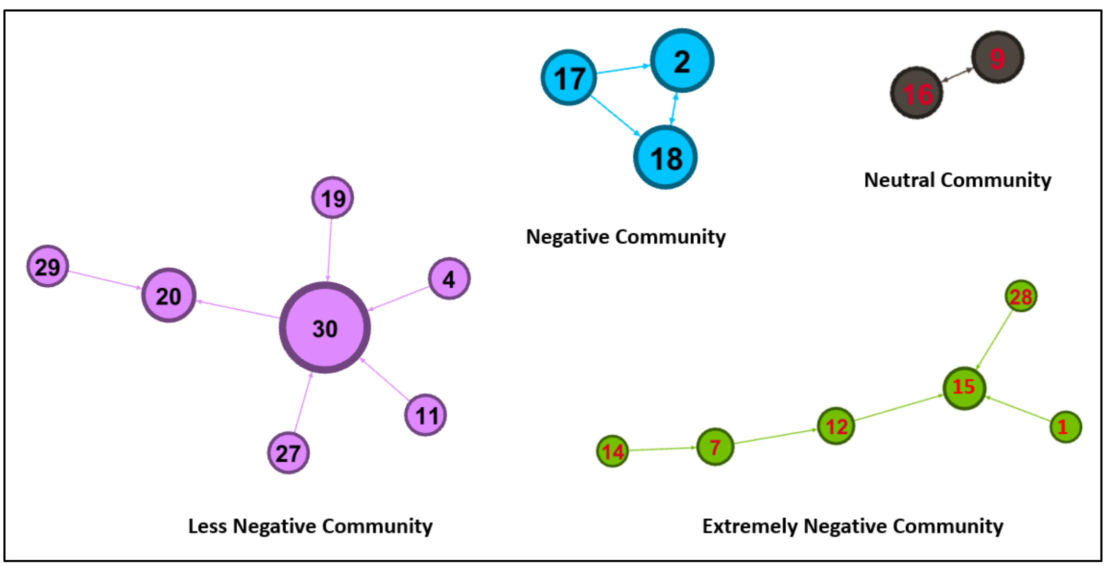

Figure 5. Communities of Twitter users who retweeted a particular negative tweet with quotes

From these 29 users who retweeted with quote, a followers' network is extracted where the users follow each other or not is checked. Then, different communities are generated from this follower's network by applying community detection algorithm based on modularity [23] (showed in different colors in Figure 5). Four discrete communities (Figure 5) are found from this single contagion where the users follow each other within each community. The sentiment values of these retweets with quote specified that most of the users of each community possesses either only negative or positive sentiments. Hence, the communities are labelled as extremely negative, negative, less negative, and neutral community. Then, the discussion topics from each community is observed and irrelevant discussion (in compared with the original tweet) patterns have found from extremely negative and negative communities. For an example, users from extremely negative communities were saying that the experts are not making enough effort to develop vaccine, without any conclusive reference or statistics. This indicates that the extremely negative and negative communities may possess the potential misinformation. These communities could be labelled as hoax community as well.

Another positive tweets' information deviation is also analyzed in similar way to confirm the phenomena described above. This original tweet *("I will respectfully challenge the Government - I want our country to succeed. However, I will not 'watch my tone")* was emphasizing on a debate between government official and healthcare worker about the pandemic situation of a specific country affected by COVID-19 severely. The user had 187,300 followers and followed only 7,130 people in Twitter. The sentiment of the original tweet is calculated as 0.7579 (highly positive) *and* 104 retweets with quote from followers are identified.

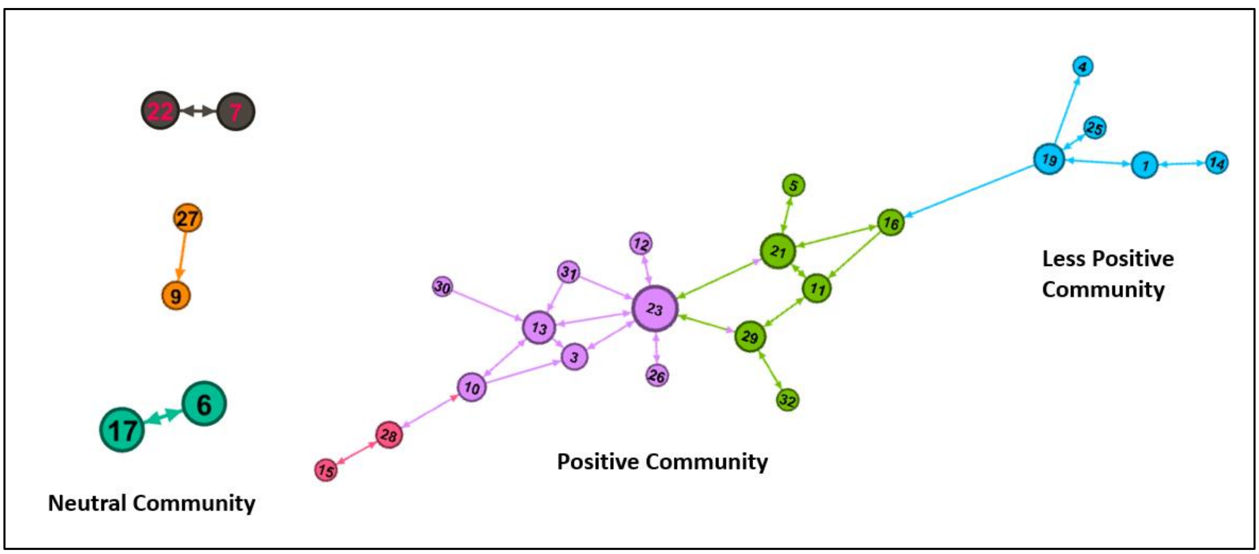

Figure 6. Communities of Twitter users who retweeted a particular positive tweet with quotes

After applying community detection algorithm on the followers' network from 104 users who retweeted the original tweet with quotes, seven communities (Figure 6) are identified based on modularity values (depicted in different colors). Two connected communities (positive and less positive) and three discrete (neutral) communities are clustered based on their followers. The users from the less positive communities were discussing irrelevant topics (cases of COVID-19 of other countries were much more than their country) without any evidence. Hence, this less positive community could be labeled as hoax community.

DISCUSSION AND CONCLUSION

The spread of misinformation is becoming a major problem for mass people and different agencies as both are being affected by its consequences. Small amount of misinformation can make a significant impact on society as there is no universal understanding of misinformation yet. Along with COVID-19 pandemic, this misinformation problem is creating an *Infodemic* which is much more contagious than coronavirus and affecting the society severely. This study has developed a new framework to quantify misinformation in social media by tracking the information deviation. This information deviation is captured by calculating the differences in sentiment values of tweets and retweets with quote from Twitter users. Identifying tweet contagion from the follower's subgraph of users lead the study to detect different communities. These communities are detected based on their modularity values which assisted to differentiate the hoax communities from neutral and reliable communities. Key findings of this study are listed as following-

- *Users with higher closeness centrality are less likely to deviate from original information.*
- *Users with higher average neighbor degree and harmonic centrality (more connected and closer to other nodes) are more likely to deviate from original information.*
- *Users with higher degree and eigenvector centrality (connected with influential nodes) are more likely to spread information quickly.*
- *Users with higher average neighbor degree and indegree centrality are less likely to spread information quickly.*
- *Users having similar sentiments in their opinion tend to cluster in same community, where relatively negative communities are responsible for spreading misinformation.*

The findings and proposed methodology of this study may help policy makers, public health, transportation managers, and emergency organizations to better understand the propagation of misinformation through social media as well as to develop effective strategy to deal with this. Future studies may perform advanced machine learning techniques to learn the suspicious information spreading behavior and predict misinformation ahead of time which will help policy makers and agencies to develop effective strategy to track, identify and predict misinformation. An artificial intelligent (AI) based tool might be developed following the proposed framework to combat with misinformation in social media. The limitation of the study is that the data used for the analyses is relatively small as well as the streaming API used for the data collection could not fetch all the tweets generated at that time.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. IIS-2027360. However, the authors are solely responsible for the findings presented in this study. The analysis and results section are based on the limited Twitter dataset and authors' opinion and cannot be expanded to other datasets without detailed implementation of proposed methods. Other agencies and entities should explore these findings based on their application/objectives before using these findings for any decision-making purpose. Any opinions, findings and conclusions or

recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

AUTHOR CONTRIBUTIONS

The authors confirm the contributions to the paper as follows: study conception and design: A. M. Sadri; data collection: M. A. Ahmed, A. M. Sadri; analysis and interpretation of results: M. A. Ahmed, A. M. Sadri; draft manuscript preparation: M. A. Ahmed, A. M. Sadri. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Singh, L., et al., *A first look at COVID-19 information and misinformation sharing on Twitter.* arXiv preprint arXiv:2003.13907, 2020.
2. Baum, M., et al., *The State of the Nation: A 50-State Covid-19 Survey Report# 14: Misinformation and Vaccine Acceptance.* Homeland Security Digital Library, 2020.
3. Sadri, A.M., et al., *Exploring network properties of social media interactions and activities during Hurricane Sandy.* Transportation Research Interdisciplinary Perspectives, 2020. **6**: p. 100143.
4. Sadri, A.M., et al., *Crisis communication patterns in social media during Hurricane Sandy.* Transportation research record, 2018. **2672**(1): p. 125-137.
5. Sadri, A.M., S. Hasan, and S.V. Ukkusuri, *Joint inference of user community and interest patterns in social interaction networks.* . Social Network Analysis and Mining, 9(1), 11., 2019.
6. Ukkusuri, S.V., et al., *Use of Social Media Data to Explore Crisis Informatics-Study of 2013 Oklahoma Tornado.* Transportation Research Record, 2014.
7. Jain, S., V. Sharma, and R. Kaushal. *Towards automated real-time detection of misinformation on Twitter*. in *2016 International conference on advances in computing, communications and informatics (ICACCI)*. 2016. IEEE.
8. Pazzanese, C. *Battling the 'pandemic of misinformation'*. 2020; Available from: https://news.harvard.edu/gazette/story/2020/05/social-media-used-to-spread-create-covid-19-falsehoods/.
9. Sadri, A.M., et al., *Understanding information spreading in social media during Hurricane Sandy: user activity and network properties.* arXiv preprint arXiv:1706.03019., 2017a.
10. Brown, S. *MIT Sloan research about social media, misinformation, and elections*. 2020; Available from: https://mitsloan.mit.edu/ideas-made-to-matter/mit-sloan-research-about-social-media-misinformation-and-elections.
11. Vosoughi, S., D. Roy, and S. Aral, *The spread of true and false news online.* Science, 2018. **359**(6380): p. 1146-1151.
12. Pennycook, G. and D.G. Rand, *Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning.* Cognition, 2019. **188**: p. 39-50.
13. Eastwood, B., *When the "lying demagogue" is the authentic candidate*. 2018, MIT Management Sloan School, April.

14. Fussell, S. *This AI Helps Police Monitor Social Media. Does It Go Too Far?* 2021; Available from: https://www.wired.com/story/ai-helps-police-monitor-social-media-go-too-far/.
15. Bursztyn, L., et al., *Misinformation during a pandemic.* University of Chicago, Becker Friedman Institute for Economics Working Paper, 2020(2020-44).
16. Shah, D. and T. Zaman, *Rumors in a network: Who's the culprit?* IEEE Transactions on information theory, 2011. **57**(8): p. 5163-5181.
17. Kouzy, R., et al., *Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter.* Cureus, 2020. **12**(3).
18. Buntain, C. and J. Golbeck. *Automatically identifying fake news in popular Twitter threads*. in *2017 IEEE International Conference on Smart Cloud (SmartCloud)*. 2017. IEEE.
19. Hjouji, Z., et al., *The impact of bots on opinions in social networks.* arXiv preprint arXiv:1810.12398, 2018.
20. Mesnards, N.G.d., Z.e. Hjouji, and T. Zaman, *Detecting Bots and Assessing Their Impact in Social Networks.* arXiv preprint arXiv:1810.12398, 2018.
21. Hutto, C. and E. Gilbert, *Vader: A parsimonious rule-based model for sentiment analysis of social media text.* Proceedings of the International AAAI Conference on Web and Social Media, 2014. **8**(1).
22. Ahmed, M.A., A.M. Sadri, and M.H. Amini, *Data-driven Inferences of Agency-level Risk and Response Communication on COVID-19 through Social Media based Interactions.* arXiv preprint arXiv:2008.03866, 2020.
23. Shang, R., et al., *Community detection based on modularity and an improved genetic algorithm.* Physica A: Statistical Mechanics and its Applications, 2013. **392**(5): p. 1215-1231.