Social Network Analysis for Coronavirus (COVID-19) in the United States

Article *in* Social Science Quarterly · May 2020

DOI: 10.1111/ssqu.12808

CITATIONS
23

READS
634

1 author:

Seungil Yum
University of Michigan
25 PUBLICATIONS **49** CITATIONS

SEE PROFILE

Social Network Analysis for Coronavirus (COVID-19) in the US

Introduction

Coronavirus (COVID-19) is an infectious disease occurred in December 2019 in Wuhan, China. Coronavirus rapidly spreads to the world, resulting in the ongoing 2019-20 coronavirus pandemic. Coronavirus has been the most serious issue in the world between 2019 and 2020 and is considered as one of the worst viruses in human history. This is because coronavirus has the high infectious fatality rate, reproduction number, environmental viability, and diversity. For example, according to the World Health Organization (WHO) officials, the mortality rate for COVID-19 is 3.4% globally as of March 3, 2020. As of April 17, 2020, more than 2.2 million cases have been reported across 185 countries and territories, resulting in more than 150,000 deaths (see e.g., World Health Organization, 2020).

Among the countries, the US experiences the worst situation in the world. Coronavirus spread to the US in January 2020 and occurred in all fifty US states, the District of Columbia, and all inhabited US territories except American Samoa. The US ranks first both in confirmed patients and deaths among all countries as of April 17, 2020. The US has confirmed patients of about 3.5 times and deaths of about 1.6 times higher than the second ranked country.

Therefore, coping with coronavirus is the most important issue in the US. This study explores how people access information on coronavirus via social networks. This study specifically explores how public key players, such as the US president, the White House, WHO and its regional offices, health organizations, Centers for Disease Control (CDC), and news channels, play an important role in communications based on Social Network Analysis (SNA) for Twitter, which is one of the most popular Social Network Systems (SNS).

This study aims to highlight some important issues as follows: first, this study identifies public key players and their roles for social networks. Specifically, this study measures the magnitude of in-degree centrality of public key players and categorizes the characteristics of the Twitter users. Second, this study highlights how in-degree centrality of public key players plays a different role in communication networks, compared to the number of followers. This study particularly shows the relationship between in-degree centrality and the number of followers based on all Twitter users in this study. Third, this study employs the sentiment analysis of people to understand the important issues and needs of citizens against coronavirus in the US. This study specifically utilizes word clouds analysis to visualize the interests and topics of individuals for coronavirus. To the best of my knowledge, this is the first article exploring online social networks for coronavirus based on a multitude of social network analyses for Twitter.

Social Network Analysis for coronavirus

Social Network Analysis is the process of exploring social structures via the networks and graph theory (see e.g., Otte & Rousseau, 2002). There has been a considerable growth of interest in SNA since it could visualize networks among individual actors, people, or other things. During the last decade, many books and articles deal with aspects of social network theory, application, and method (see e.g., Carrington et al., 2005). Many scholars have tried to employ SNA for their research goals to understand the social networks among entities (see e.g., De Nooy et al., 2018; Freeman, 2004; Marsden & Lin, 1982).

This study utilizes Twitter data to analyze social network among public key players for coronavirus. Twitter is known as one of the most useful resources for big data analyses (see e.g., Broniatowski et al., 2014). This study monitors Twitter data stream between April 16 and April 22, 2020 based on the keyword “US coronavirus” and chooses the best data

set for the analysis (April 19 and April 20) based on some important criteria (e.g., the number of Twitter users, communications, and suitable contents). This study employs social network analyses based on 2,864 Twitter users and 2,775 communications.

This study employs NodeXL for exploring how public key players play a pivotal role in social networks for coronavirus. NodeXL is a visualization software program, which supports social networks and content analysis. NodeXL has been widely used as a social network tool in a variety of articles (see e.g., Bonsignore et al., 2009).

This study explores in-degree centrality to analyze the relationship between public key players and social networks. It highlights how public key players gain attention to their tweets from communication networks. This study chooses public key players among all Twitter users based on the magnitude of their in-degree centrality.

After analyzing the magnitude of in-degree centrality, President Trump plays the most important role in the communication nodes among the key players. The magnitude of in-degree centrality of President Trump is equal to the sum between the top second and top fifteenth, meaning that people's interests are heavily concentrated in the behavior of President Trump for the US corona problem. The next important finding is that many organizations in Africa play a crucial role in communication networks. For instance, the Nigeria Centre for Disease Control as well as Africa Centers for Disease Control rank second and fifth, respectively. In addition, other organizations, such as the eHealth Africa, the World Health Organization in Nigeria, the Federal Ministry of Information and Culture, Nigeria, and the Federal Ministry of Health in Nigeria, and a news channel (Channels Television in Nigeria) place within the top 20th. The results show that organizations and new channels of Nigeria and Africa play an important role in the information hubs for the coronavirus, and people in Nigeria and Africa have a high interest in the virus in the study

period. This is because the corona situation in Nigeria and Africa has become more serious since April 2020.

The third remarkable finding is that not only new channels in the US, such as CNN and Fox News, but also a news channel in the UK, such as BBC, play a crucial role in the information nodes. For example, CNN ranks third, BBC places fifth, and Fox news takes twentieth. BBC even ranks higher than Fox news, while the BBC is a British public service broadcaster.

The fourth noticeable finding is that WHO and its regional offices, such as WHO Regional Office for Eastern Mediterranean, Europe, South-East Asia, and Africa, play a crucial role in the information of coronavirus. For example, WHO ranks third, and eight out of the top 20 key players are the organizations of WHO. The results show that people rely heavily on WHO for sharing information on the coronavirus.

The fifth interesting finding is that Barack Obama, who is the former US president, ranks eleventh, showing that he still exerts a powerful impact on public influence, whereas he finished the presidential job a long time ago. Also, the White House places eighteenth, as well as Barack Obama and Donald Trump, meaning that people are highly dependent on the governmental key players to deal with coronavirus.

This study highlights the differences between in-degree centrality and followers since They have a similar concept. They both measure the number of incoming links incident to the node, whereas the former one is the number of the topic-based networks, and the latter one is the number of the person-based networks in this study. Figure 1 shows that in-degree centrality and the number of followers are lowly correlated for coronavirus. For example, Donald Trump and Barack Obama show the most different characteristics. When we see the trend line, they are located in the opposite position. To be specific, Donald

Trump shows the highest in-degree centrality (90), whereas he has quite a lower number of followers (77.3M) than Barack Obama (116.1M). In contrast, Barack Obama has the highest followers, while he shows quite low in-degree centrality (6) than Donald Trump. This is because Donald Trump has the highest power for solving the coronavirus problem, whereas Barack Obama became a normal citizen who does not have administrative power for the problem. The results show that in-degree centrality and the number of followers (the topic-based networks and the person-based networks) play a different role in social networks.

[Figure 1 about here]

Next, this study utilizes cluster analysis by employing the Clauset–Newman–Moore cluster algorithm. Cluster analysis is a methodology for the task of assigning a set of objects into groups so that objects in the same cluster are more similar to each other than those in other clusters. The Clauset-Newman-Moore cluster algorithm is one of the most useful cluster methodologies for big data analysis (see e.g., Vieira et al., 2014). Figure 2 shows the Clauset–Newman–Moore cluster algorithm for social networks of COVID-19. Donald Trump places the center of the whole social network, and other key players are located in the central networks. In contrast, the US news channels and WHO and its regional offices have some independent channels

[Figure 2 about here]

Lastly, this study visualizes the word frequency of tweets for coronavirus in social networks by employing word clouds. The word clouds draw a collection of words in different sizes, which is commonly used to depict keyword metadata in tweets. The bigger and bolder keywords mean that the more often it is mentioned within the tweets, and the more

important it is. The study excludes the keyword “coronavirus” for word clouds since it is already used for collecting tweets.

The highest keyword is “realdonaldtrump.” People strongly focus on the US president’s action and policy to cope with coronavirus (see Figure 3). The second-highest keyword is “people.” Twitter users are highly interested in how many people are infected and dead by the virus in their country or the world. The third-highest keyword is “protect.” People share the method how to protect themselves against coronavirus. The fourth-highest keyword is “China.” This is because coronavirus is also known as the “China virus” and is originated from Wuhan in China. The fifth-highest keyword is “world” since coronavirus is a pandemic virus across the world.

Also, people are highly interested in and “work” because many of them cannot work because of the virus. In fact, according to the Department of Labor, 6.6 million US workers filed for their first week of unemployment benefits in the week ending March 28, which is more than 3,000% the pre-pandemic levels and one of the most devastating periods in history for the American job market.

Keywords the “save,” “more,” “now,” “kill” also rank within top tenth because coronavirus drastically kills people, and people are highly interested in the number of dead people as of now. In addition, keywords “party” place within the top twentieth since people cannot enjoy partying and drinking alcohol with their friends caused by the effects of curfews and quarantine.

[Figure 3 about here]

Conclusions

This study explores the key players for online social networks of coronavirus in the US by employing SNA. This study suggests some important findings as follows: First, while Barack Obama plays a key role in the networks, his centrality is relatively low, compared with his followers. In contrast, Donald Trump exerts the strongest impact on it, whereas his followers are quite lower than Barack Obama. Second, this study finds that President Trump plays the most important role in social networks among the top 20 key players. The magnitude of in-degree centrality is equivalent to the sum between the top second and top fifteenth. Third, this study highlights that many organizations in other countries, such as Nigeria and the UK, play a pivotal role in communication networks.

Fourth, news channels in the other countries as well as in the US are key players for the corona issue in the US. Fifth, Barack Obama still exerts a powerful impact on communication networks, while he completed the presidential job a long time ago. Sixth, this study demonstrates that key players, such as Donald Trump, Barack Obama, Africa organizations, and BBC, are located in the central networks. In contrast, the US news channels and WHO and its regional offices have some independent channels. Seventh, the highest keyword in the social networks is “realdonaldtrump,” followed by “people,” “protect,” and “China,” “world.”

References

Bonsignore, E. M., Dunne, C., Rotman, D., Smith, M., Capone, T., Hansen, D. L., & Shneiderman, B. (2009). First steps to NetViz Nirvana: evaluating social network analysis with NodeXL. In 2009 International conference on computational science and engineering, 4, 332-339.

Broniatowski, D. A., Paul, M. J., & Dredze, M. (2014). Twitter: big data opportunities. Science, 345(6193), 148-148.

Carrington, P. J., Scott, J., & Wasserman, S. (2005). Models and methods in social network analysis. Cambridge: Cambridge university press.

De Nooy, W., Mrvar, A., & Batagelj, V. (2018). Exploratory social network analysis with Pajek. Cambridge: Cambridge University Press.

Freeman, L. (2004). The development of social network analysis. Vancouver: Empirical Press.

Marsden, P. V., & Lin, N. (1982). Social structure and network analysis. Beverly Hills: Sage.

Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. Journal of information Science, 28(6), 441-453.

Vieira, V. D. F., Xavier, C. R., Ebecken, N. F. F., & Evsukoff, A. G. (2014). Performance evaluation of modularity based community detection algorithms in large scale networks. Mathematical Problems in Engineering, 2014, 1-16.

World Health Organization. (2020). COVID-19 STRATEGY UPDATE. Geneva: World Health Organization.

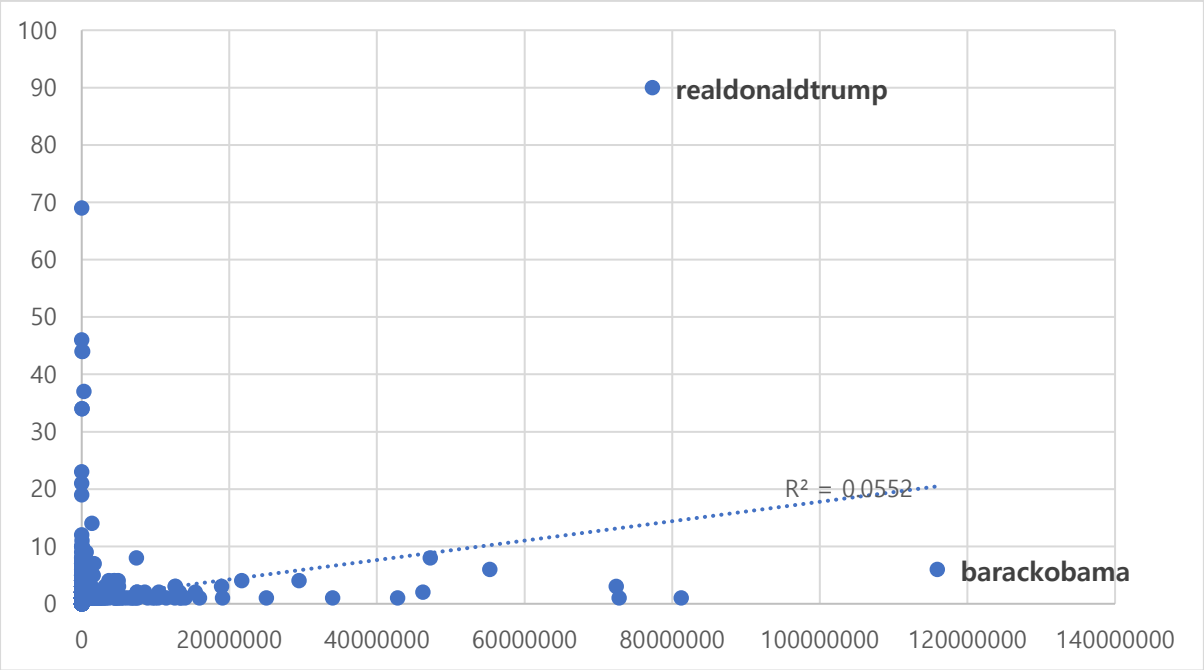

Figure 1. The relationship between in-degree centrality and followers

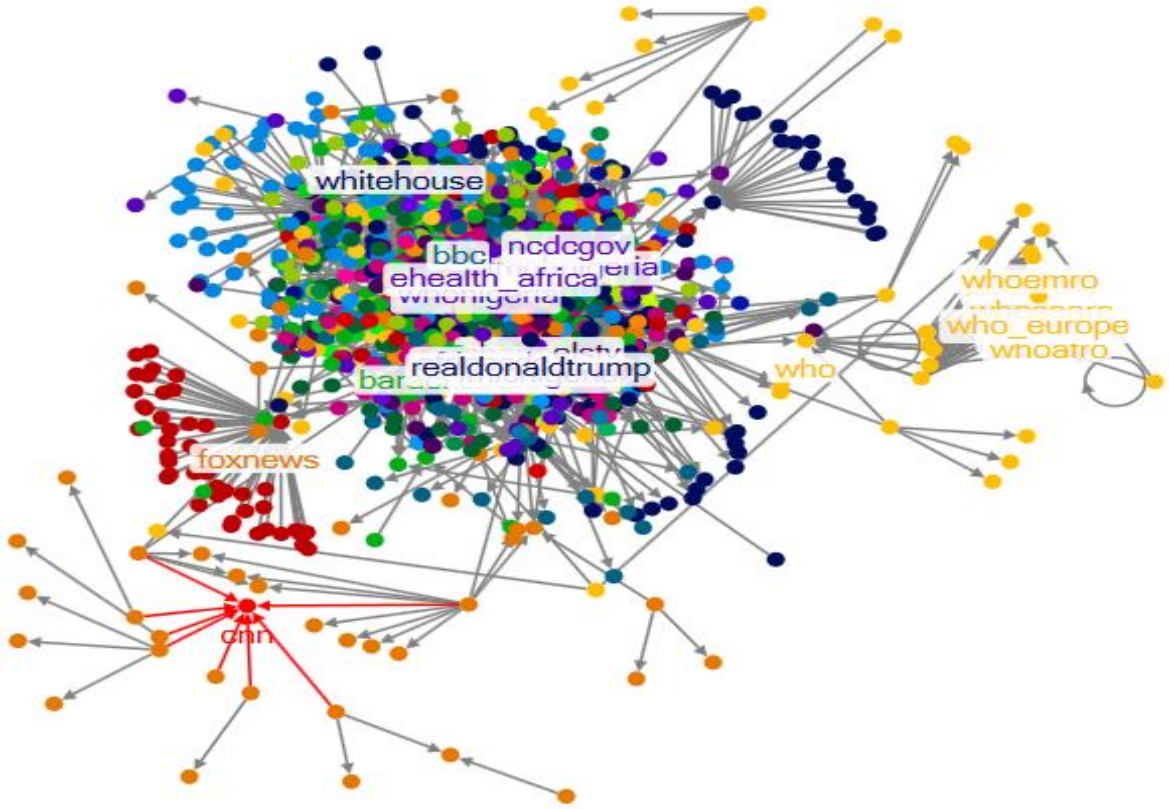

Figure 2. The social network

Figure 3. world clouds