

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

Seungil Yum , *University of Florida*

Objectives. This study explores how public key players play an important role in social networks for coronavirus (COVID-19). **Methods.** This study employs social network analyses based on 2,864 Twitter users and 2,775 communications of Twitter. **Results.** This study finds that President Trump plays the most important role in social networks among the top 20 key players for both in-degree centrality and content in tweets. Second, Donald Trump and Barak Obama show the opposite result for the in-degree centrality and follower analysis. The result shows that the topic-based networks and the person-based networks play a different role in social networks. This study demonstrates that the presidents, the World Health Organization (WHO) and its regional offices, the Centers for Disease Control, and news channels play a crucial role in the news of COVID-19 for people. Key players, such as Donald Trump, Barack Obama, and BBC, are located in the central networks. In contrast, U.S. news channels and WHO and its regional offices have independent channels. **Conclusions.** Governments should understand the characteristics of public key players to provide information for COVID-19 in a timely manner.

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

Among the countries, the United States experiences the worst situation in the world. COVID-19 spread to the United States in January 2020 and occurred in all 50 U.S. states, the District of Columbia, and all inhabited U.S. territories except American Samoa. The United States ranks first both in confirmed patients and in deaths among all countries as of April 17, 2020. The United States has confirmed patients of about 3.5 times and deaths of about 1.6 times higher than the second-ranked country.

Therefore, coping with COVID-19 is the most important issue in the United States. This study explores how people access information on COVID-19 via social networks. This study specifically explores how public key players, such as the U.S. 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

Direct correspondence to Seungil Yum, Design, Construction, and Planning, University of Florida, Gainesville, FL 32611 (yumseungil@ufl.edu).

SOCIAL SCIENCE QUARTERLY, Volume 101, Number 4, July 2020

© 2020 by the Southwestern Social Science Association

DOI: 10.1111/ssqu.12808

social network analysis (SNA) for Twitter, which is one of the most popular social network systems.

This study aims to highlight some important issues as follows: first, this study identifies public key players and their roles in social networks. Specifically, this study measures the magnitude of in-degree centrality of public key players and categorizes the characteristics of 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 sentiment analysis of people to understand the important issues and needs of citizens against COVID-19 in the United States. This study specifically utilizes word clouds analysis to visualize the interests and topics of individuals for COVID-19. To the best of my knowledge, this is the first article exploring online social networks for COVID-19 based on a multitude of SNAs for Twitter.

Social Network Analysis for Coronavirus

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

This study utilizes Twitter data to analyze social networks among public key players for COVID-19. Twitter is known as one of the most useful resources for big data analyses (see, e.g., Broniatowski, Paul, and Dredze, 2014). This study monitors the Twitter data stream between April 16 and April 22, 2020, based on the keyword “U.S. 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 SNAs based on 2,864 Twitter users and 2,775 communications.

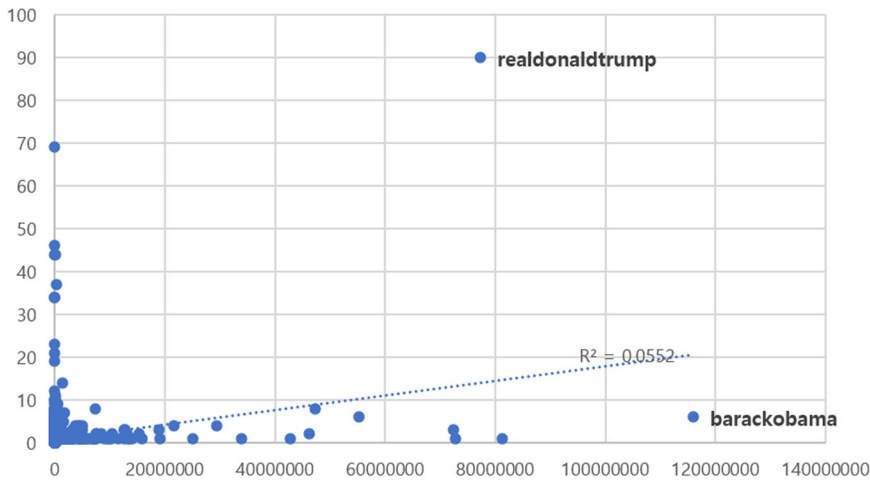
This study employs NodeXL for exploring how public key players play a pivotal role in social networks for COVID-19. NodeXL is a visualization software program that supports social networks and content analysis. NodeXL has been widely used as a social network tool in a variety of articles (see, e.g., 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.

The analysis of the magnitude of in-degree centrality shows that 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 two and top 15 key players, meaning that people's interests are heavily concentrated on the behavior of President Trump for the U.S. corona problem. The next important finding is that many organizations in Africa play a crucial role in communication networks. For instance, the Nigeria CDC as well as the Africa CDC rank the second and fifth, respectively. In addition, other organizations, such as eHealth Africa, the WHO in Nigeria, the

FIGURE 1

The Relationship Between In-Degree Centrality and Followers



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 20 key players. The results show that organizations and news channels of Nigeria and Africa play an important role in the information hubs for COVID-19, 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.

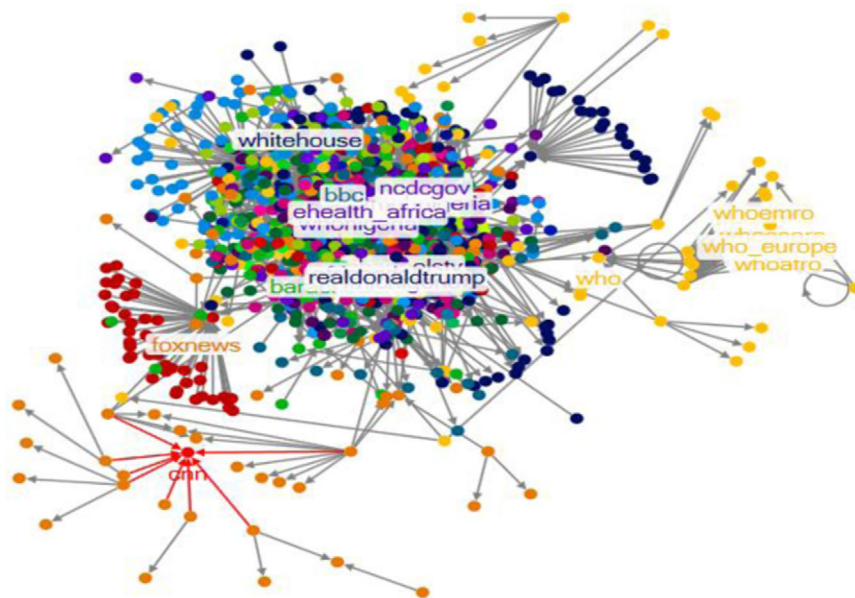
The third remarkable finding is that not only news channels in the United States, such as CNN and Fox News, but also a news channel in the United Kingdom, the BBC, play a crucial role in the information nodes. For example, CNN ranks third, BBC places fifth, and Fox News takes the 20th rank. BBC ranks even higher than Fox News, even though BBC is a British public service broadcaster.

The fourth notable finding is that WHO and its regional offices, such as the WHO Regional Office for Eastern Mediterranean, Europe, South-East Asia, and Africa, play a crucial role in the information of COVID-19. For example, WHO ranks third, and eight of the top 20 key players are organizations of WHO. The results show that people rely heavily on WHO for sharing information on COVID-19.

The fifth interesting finding is that Barack Obama, who is the former U.S. president, ranks 11th, showing that he still exerts a powerful impact on public influence, even though he finished the presidential job a long time ago. Also, the White House places 18th, as well as Barack Obama and Donald Trump, meaning that people are highly dependent on the governmental key players to deal with COVID-19.

This study highlights the differences between in-degree centrality and followers since they have a similar concept. Both of them measure the number of incoming links incident to the node, whereas the former one is the number of topic-based networks, and the latter one is the number of person-based networks in this study. Figure 1 shows that in-degree centrality and the number of followers are lowly correlated with COVID-19. 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

FIGURE 2
The Social Network



followers (77.3M) than Barack Obama (116.1M). In contrast, Barack Obama has the highest number of followers, while he shows quite low in-degree centrality (6) compared to Donald Trump. This is because Donald Trump has the highest power for solving the COVID-19 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.

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 appears at the center of the whole social network, and other key players are located in the central networks. In contrast, U.S. news channels and WHO and its regional offices have some independent channels.

Lastly, this study visualizes the word frequency of tweets for COVID-19 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 the keyword, 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 U.S. president’s action and policy to cope with COVID-19 (see Figure 3). The second highest keyword is “people.” Twitter users are highly interested in how many people are infected by and dead

FIGURE 3



from the virus in their country or in the world. The third highest keyword is “protect.” People share methods of how to protect themselves against COVID-19. The fourth highest keyword is “China.” This is because COVID-19 is also known as the “China virus” and originated in Wuhan in China. The fifth highest keyword is “world” since COVID-19 is a pandemic virus across the world.

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

The keywords “save,” “more,” “now,” and “kill” also rank within the top 10 because COVID-19 drastically kills people, and people are highly interested in the number of dead people as of now. In addition, the keyword “party” places within the top 20 since people cannot enjoy partying and drinking alcohol with their friends, which is caused by the effects of curfews and quarantine.

Conclusions

This study explores the key players for online social networks of COVID-19 in the United States by employing SNA. This study suggests some important findings: first, while Barack Obama plays a key role in the networks, his centrality is relatively low compared to his followers. In contrast, Donald Trump exerts the strongest impact on networks,

although his followers are quite fewer in number than those of 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 two and top 15 players. Third, this study highlights that many organizations in other countries, such as Nigeria and the United Kingdom, play a pivotal role in communication networks.

Fourth, news channels in other countries as well as in the United States are key players for the corona issue. Fifth, Barack Obama still exerts a powerful impact on communication networks, even though he completed the presidential job a long time ago. Sixth, this study demonstrates that key players, such as Donald Trump, Barack Obama, African organizations, and the BBC, are located in the central networks. In contrast, U.S. 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,” “China,” and “world.”

REFERENCES

- Bonsignore, E. M., C. Dunne, D. Rotman, M. Smith, T. Capone, D. L. Hansen, and B. Shneiderman. 2009. “First Steps to NetViz Nirvana: Evaluating Social Network Analysis with NodeXL.” Pp. 332–39, Vol. 4, in 2009 International Conference on Computational Science and Engineering, Vancouver, BC.
- Broniatowski, D. A., M. J. Paul, and M. Dredze. 2014. “Twitter: Big Data Opportunities.” *Science* 345(6193):148.
- Carrington, P. J., J. Scott, and S. Wasserman. 2005. *Models and Methods in Social Network Analysis*. Cambridge: Cambridge University Press.
- De Nooy, W., A. Mrvar, and V. Batagelj. 2018. *Exploratory Social Network Analysis with Pajek*. Cambridge: Cambridge University Press.
- Freeman, L. 2004. *The Development of Social Network Analysis*. Vancouver, BC: Empirical Press.
- Marsden, P. V., and N. Lin. 1982. *Social Structure and Network Analysis*. Beverly Hills, CA: Sage.
- Otte, E., and R. Rousseau. 2002. “Social Network Analysis: A Powerful Strategy, Also for the Information Sciences.” *Journal of Information Science* 28(6):441–53.
- Vieira, V. D. F., C. R. Xavier, N. F. F. Ebecken, and A. G. Evsukoff. 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.