Centrality and Influence within COVID-19 Anti-vaccine Networks on Twitter

Vlad-Alexandru Costea

Introduction & Research Question

In recent years, COVID-19 anti-vaccine sentiment has emerged as one of the most significant challenges facing public health. An estimated 15 million people have died worldwide from the COVID-19 disease in the 2020-21 period (Adam 2022). This number is increased by vaccine refusal, as in the United States alone, between June 2021 and March 2022, a staggering 234 thousand people are estimated to have died of the COVID-19 disease after refusing to get vaccinated (Amin et al. 2022). As such, misinformation leading to vaccine refusal is a true public health crisis that needs to be tackled immediately (Butcher 2021).

Furthermore, there is evidence that social media platforms such as Twitter contain large networks dedicated to falsehoods relating to the COVID-19 vaccine, which contribute to vaccine hesitancy and refusal. These falsehoods have significantly stifled the public health response to the ongoing COVID-19 pandemic and, by some estimates, are responsible for up to 30% of vaccine refusal (Bruns et al. 2021). It is clear that this misinformation crisis needs to be addressed as part of the wider response to the COVID-19 pandemic.

We can empirically identify networks of misinformation which consist of individuals who like, retweet and reply to each other's posts and spread falsehoods about the COVID-19 vaccine and vaccination campaign. In order to limit the spread of misinformation, software engineers and policy-makers require an accurate understanding of the mechanisms at work within these networks, as well as of their internal structures. This research paper will attempt to further develop the understanding of these networks by looking at their structures and how they relate to social influence.

As shown in previous research, different centrality measures can be used in order to identify highly influential accounts within an online network (Soares et al. 2018; Boulet & Lebraty 2018). Furthermore, discrepancies between the rankings of the accounts with the highest calculated centrality according to different measures indicate that the individuals managing the accounts in question are central within the network in distinct ways (Aggrawal & Anand 2022, 53). As such, we seek to first identify a Centrality measure that best corresponds to social influence within such networks of misinformation. As social influence is not directly measurable, we will empirically test whether the measure proves reliable by comparing it to another previously established measure of influence, i.e. the number of retweets received.

Thus, the following research question will be investigated: within the entirety of the Twitter anti-vaccine network, to what extent does an account's centrality correspond to the account's social influence? Since the entire network cannot be sampled, the study will make inferences from a sample of a subset of the network. As an account's social influence cannot be directly calculated, the proposed measure will be compared to another measure of social influence which has been previously shown to be reliable. As will be elaborated upon, the number of retweets (i.e. the number of times other accounts have reshared the tweet in question for their own followers to see, while still crediting the original poster) an account receives has been shown to be one such measure (Cappelletti 2012). As such, this will be the measurement with which the centrality measures will be related in order to determine their reliability.

Twitter Networks and the Public Arena

The relevance of studying Twitter in order to understand the COVID-19 vaccine misinformation problem stems from the role that social media platforms play within the informational ecosystem. Social media platforms offer a new avenue for social scientists to study social interactions, while still limited by the necessity of making assumptions on how social media and real social behavior interact. There is significant evidence that misinformation about the COVID-19 vaccine has shifted people's real behavior when it comes to getting vaccinated (Bruns et al. 2021, Roozenbeek et al. 2020).

According to Jungherr & Schroeder (2021), Twitter has emerged as a public arena, in which individuals are able to debate their opinions, structure their political focus and interact with political leaders, where people's speech is generally uncensored and participants have significant freedom of expression. Jungherr & Schroeder (2021) use the term "public arena", in order to compare and contrast with Habermas' (1989) concept of a "public sphere". According to Habermas, the public sphere represents a place where private individuals have the ability to discuss public affairs and potentially reach a consensus. In this way, the public sphere is seen as important for keeping authority figures accountable to the general population.

While originally the "public sphere" represented physical locations and traditional media in which such activities take place, the concept can be extended to social media as long as it fulfills a similar role. Furthermore, by switching from "public sphere" to "public arena", Jungherr & Schroeder (2021) emphasize that conflict between individuals with clashing ideologies is inherent within such a space, as people holding different conceptions of the world freely interact with each other. This is in contrast with traditional media outlets where the gatekeepers had control over the information and the types of ideologies promoted, and therefore were able to exert direct censorship upon individuals holding opinions deemed undesirable.

As such, by embodying the spirit of the public arena, Twitter represents an important break with legacy media. Within platforms such as Twitter, attention is not necessarily monopolized by the traditional figures of authority, and instead it can be obtained by ordinary individuals' tweets, as well as advertisers or posts created in order to become viral. In theory, this represents an opportunity for increased democracy, as lower reliance on authority figures creates the opportunity for increased representation for individuals holding ideologies formerly shunned, as well as the opportunity for people to engage with such ideologies (Jungherr & Schroeder 2021).

However, it is up for debate to what extent Twitter is able to properly fulfill the role of a public arena. For instance, while users of Twitter tend to be exposed to more diverse points of view, the ability of these users to hold meaningful debate is reduced (Yardi & Boyd 2010) and the impact of such interactions is highly limited (Moe 2012). This is because these cross-ideological engagements are shorter and oftentimes characterized by disagreement, and as a consequence opinions can rarely be changed through persuasion (Liu 2014).

Furthermore, online bullying and harassment by internet trolls can have the effect of silencing people (Nogrady 2021) and the Twitter algorithm is inherently biased towards right-wing content (Huszár et al. 2022). Overall, it is still unknown whether social media is linked to an increase or decrease in democracy (Lorenz-Spreen 2021), and it is doubtful whether Twitter fulfills the democratic function of a public arena as defined by Jungherr & Schroeder (2021).

Even taking for granted that Twitter plays the role of a public arena, this simple fact can create societal challenges in its own right. The public agenda reflected by the traditional media outlets, while less democratic and prone to filtering content, was made to reflect the diversity of opinions and demographics inherent in society, including groups that are underrepresented by social media group dynamics. Furthermore, in order to maintain societal stability, traditional media censored aggressive, hateful or harmful opinions which have always existed within society, but which social media have made highly visible (Jungherr & Schroeder 2021).

In fact, it is not only misinformation which causes societal issues within these public arenas, but also the increased freedom of expression granted towards extreme and radical groups (Jungherr & Schroeder 2021). Social media can act as a platform in which previously constituted beliefs can make themselves publicly known, while people holding marginal beliefs tend to be more active and aggressive, and in consequence to become highly visible (Gaisbauer et al. 2021). Consequently, the public arena is the locus of a clash of values, not merely of information (Jungherr & Schroeder 2021).

That being said, misinformation and disinformation represent a particularly serious issue when it comes to medical information. It has been shown that individuals who hold inaccurate beliefs about the COVID-19 vaccine are more likely to refuse to get vaccinated or to follow public health guidance, which is a serious threat to both individual and public health (Roozenbeek et al. 2020). Considering that repeated exposure to misinformation from multiple sources can make it appear more believable (Van der Linden 2022), it becomes urgent to understand how networks of misinformation operate and to research the means that can be used to stop the spread of falsehoods about COVID-19 vaccines.

Networks of Misinformation

Misinformation is a prevalent feature of social media, websites and apps. This was the case before the start of the COVID-19 pandemic, and it has continued to be the case to this day. Social media networks tend to contain a small number of accounts with a disproportionate amount of influence, as measured in likes and retweets received. For instance, even before the onset of the COVID-19 pandemic, a small number of Twitter accounts had the most influence within anti-vaccine networks (Featherstone et al. 2020).

These highly influential accounts are likely to be the main drivers of the diffusion of false information on social media. This “Broadcasters” model can be contrasted to the “Virality” conception of misinformation, in which the main drivers of misinformation are one-to-one interpersonal interactions between people that do not hold a high level of influence within the networks. Rather, in accordance with the “Broadcasters” model, informational memes tend to spread through a small number of specific sources, both online and offline, that are able to influence a high number of users (Goel et al. 2015).

However, according to Centola & Macy (2007), social transmissions often operate through the process of complex contagion. Complex contagion is distinct from simple contagion in that it can only occur when a social actor is influenced by others from multiple sources. This is due to the fact that in order for an actor to be influenced by a movement or idea, they must be exposed to it multiple times, which will allow the actor to grasp the idea better and deem it increasingly plausible (Van der Linden 2022). As such, it could be necessary for an idea to be repeated by several of these highly influential sources before it is believed.

The existence of highly influential accounts within misinformation networks on social media has been established (Featherstone et al. 2020), as well as their tendency to act as the main drivers of misinformation (Goel et al. 2015). We expect to find the same dynamic taking place in the specific case of Twitter networks dedicated to COVID-19 anti-vaccine content.

According to Shore et al. (2016), social media networks dedicated to radical opinions are marginal, and the majority of Twitter users do not partake in them. Instead, most social media users tend to post content that is more moderate than the content they receive. Accordingly, fears of increased radicalism and misinformation need to be localized, and the possibility of their general spread must not be erroneously overstated. However, there is a small core of politically engaged users that does indeed lean towards increased radicalism.

As right-wing radicalism and anti-vaccine rhetoric are often interwoven (Ahmed 2021), a core of users who spread highly inaccurate information can be found in the case of COVID-19 misinformation as well. This core represents a small, but substantial part of the overall population which generally distrusts truthful information and trusts misinformation (Roozenbeek et al. 2020). In this view, studying the highly dedicated core Twitter network which spreads misinformation on the COVID-19 vaccine can be a good proxy for studying the misinformation problem as a whole.

Power and Influence

According to Weber (1922), "Power" is defined as the possibility of an actor enacting social influence towards other actors, potentially against their own will. Expanding on this, Bruggeman (2021) adds that such influence must be intentional. In accordance with this definition, Power does not only manifest directly through interpersonal relations, but it can also be enacted through indirect influence, across large network distances, although with diminishing returns as the distance increases. As such, in order for an actor to have a higher level of power, it has to be tied to other actors who are powerful in their own right.

This definition takes into account the exercise of social influence through the sending of messages (such as information, misinformation, propaganda, etc.) meant to enact social influence towards others. This conception is shaped by Lazarsfeld et al.'s (1955) work on political influence, which shows that in order to make political decisions, people tend to rely on opinion leaders they have an interpersonal connection with. As such, the individuals whose messages are widely broadcast can be seen as influential through the "First Dimension of Power" (Lukes 2021), in which one has the power to directly influence another's thinking and ideas.

We first seek to identify the most powerful actors by using a measure based on the network structure. Such a measure must take into account the strength of the ties, the Power of other actors, and both direct and indirect influence (i.e. traveling through two or more edges within a network), while putting less weight on indirect influence due to the waning of influence over longer distances. As will be shown in the "Centrality Measures" section, the network measure of Eigenvector Centrality (Bonacich 1987) can be applied for this purpose (Bruggeman 2021).

In order to gauge whether such a measure is accurate in predicting the most influential accounts, it needs to be compared to a second variable which is independent from the centrality method and which has been previously shown to accurately represent an account's social influence. The variable we will use is the number of retweets the focal account has received on its own tweets from other accounts, either from within the network or from outside the network. This can be deemed a viable measure since an account receiving a high number of retweets on its posts will have its messages spread around further to other accounts that do not necessarily follow the original account. Through the fact that messages are spread to more people, the actor holding the focal account is likely to become powerful and influential, as long as the retweeted messages have an impact on their readers. This is not an ideal measurement of Power due to the limitations inherent in analyzing Twitter data, but it is backed by research such as Cappelletti's (2012).

As some accounts could have more tweets retrieved during data collection than others, I will also normalize the number of retweets so as to not disadvantage accounts which had fewer tweets retrieved. This will be done in the following manner: I will first record the total number of retweets that the focal account received from any account on its last 400 Tweets, which is the number of Tweets retrieved by the algorithm by default. However, as due to the Twitter API, occasionally less Tweets can be retrieved, the measure is normalized as retweets_received / tweets * 400, so as to compute the expected number of retweets received had exactly 400 tweets been retrieved. For instance, an account that received 900 retweets but only had 300 tweets retrieved would be expected to have received 900 / 300 * 400 = 1200 retweets, had 400 tweets been retrieved.

Data Collection, Processing and Analysis

The data was collected programmatically using Twitter's REST API (Twitter Developer Platform 2022), specifically the "Elevated" access level that grants access to 2 million tweets per month. All of the data that is analyzed as part of this study was collected on the 6th of April 2022, using the entirety of the 2 million tweets quota for the month. The API was accessed by means of a program written in the Python programming language, using the Tweepy library.

For each account, the data collected consisted of up to 400 tweets, including retweets, quote-tweets and replies. These tweets correspond to roughly the last couple months of the account's content. In the process of accessing every tweet, the following information has been retrieved: the text of the tweet, any referenced tweets (in case of replies, retweets and quote-tweets), and further details about the accounts referenced and public metrics, such as likes (i.e. other users showing appreciation for a post without sharing it on their own account) and retweets received (i.e. other users resharing the post on their own account, while still making the original user visible) on each post.

The method for retrieving accounts that are deemed relevant to the anti-COVID-19 vaccine networks is as follows: firstly, I manually pick an account deemed relevant within the COVID-19 anti-vaccine network. Then, the program finds other accounts which the manually picked account has retweeted at least twice, which are then sorted from most to least retweeted. Each of these accounts will then be analyzed in a similar manner, while also checking if they have tweeted certain anti-vaccine keywords at least five times. For each account, the total number of times its tweets have been retweeted is also collected.

Keywords used to identify COVID-19 anti-vaccine accounts

- ivermectin
- genetherapy
- pfizergate
- vaccinemandate
- medicaltyrany
- vaers
- vaccinegenocide
- phizer
- vaxinjur
- anti-vax
- antivax
- vaccinedeath
- vaxxdeath
- vaxdeath
- vaccineinjur
- vaccinemandate
- sideeffect
- vaxxinjur
- vax-injur
- vaxx-injur
- jabbed
- vaccinationregret
- #pfizer
- bigpharma
- vaxmandateexperimentalvaccine*

*typing error

Figure 1

The keyword search uses keywords which were manually picked as relevant by looking through words and phrases often used by anti-COVID-19 vaccine accounts (Figure 1), and it does not take into account capitalization or spacing differences, in order to find more matches. If the account in question has tweeted at least the minimum number of keywords (specifically, five), it is deemed relevant within the network, and the accounts it links to will in turn be analyzed. The graph searching algorithm used is Breadth-First Search, and the method corresponds to Snowball sampling.

Using this method, 5156 accounts have been retrieved in total. Out of those, 1840 accounts have been deemed relevant. In addition to the 1790 accounts which reached the minimum number of keywords, the 50 most influential accounts that have not met the keyword requirement have also been accepted as relevant. These accounts, while often using language that is not easily detectable as anti-vaccine by the keyword method, are clearly highly relevant within the network as a whole and as such should be included within the analysis.

Centrality Measures

This research will be using the measure of Eigenvector Centrality, as first defined by Bonacich (1987). Eigenvector Centrality is the mathematically defined measure that best corresponds to a Weberian conception of Power, as elaborated in the "Power and Influence" section. It does not take into account merely the number of direct links a node has, as degree centrality does, but also the strength of these ties, as well as the relative Power of other linked nodes.

The Eigenvector Centrality measurement algorithm calculates the relevance of a node based on the summation of the power of adjacent nodes, divided by a constant. This is done recursively for each node, so as to take into account influence over nodes which are not directly tied to the node in question as well. As the measurement is divided by the constant each time, the numerical influence of far away nodes goes down exponentially (Aggrawal & Anand 2022, 57). The formula for Eigenvector Centrality, where C is the centrality value, x is the current node, w is a matrix containing the weights of the edges, v is the set of edges of the node and λ is the constant (representing the largest eigenvalue of the adjacency matrix), is as follows:

$$C_x = \frac{1}{\lambda} \sum_{y \in v} (w_{x,y} * C_y).$$

In order to gauge whether Eigenvector Centrality is a good measure of social influence, it must be correlated with another variable which also indicates, in a different way, social influence within the network. While social influence cannot be directly quantified, by comparing two variables, it can be established to which extent the measurements are reliable. As such, the Eigenvector Centrality measure will be compared to the normalized number of received retweets, which has been established in previous sections as a measure that can model an account's social influence.

Analysis

Using the NetworkX Python library (Hagberg et al. 2008), I first created a graph data structure containing the relevant accounts, in which the nodes represent the accounts and the edges represent one account retweeting the other at least two times. Then, for each account, its Eigenvector Centrality was calculated. Relevant data on the accounts was then saved within an SPSS file, in order to be used for further statistical analysis.

This data also includes the normalized number of total retweets received on Tweets by the focal account in question, which is defined as $retweets/tweets * 400$. As previously explained, this calculates the expected number of retweets, had 400 tweets been retrieved, in case a smaller number of tweets have been retrieved. This is done in order to not disadvantage accounts which had fewer tweets retrieved by the algorithm. This second variable will be used in order to gauge the reliability of the Eigenvector Centrality measure.

In calculating the Eigenvector Centrality of different nodes, the weight assigned to edges corresponds to the base 2 logarithm of how many times one node has retweeted the other. This is an empirical choice, as issues arose when not taking the logarithm of the number. Namely, two accounts which retweeted each other extensively were both measured as having very high centralities, which does not satisfy common sense.

Ethics of Data Collection

I do not believe there are significant ethical concerns when it comes to the methods used in this research. All information analyzed is publicly available on Twitter, and no private information was used. While the research was done on messages having been posted by persons who did not consent beforehand to being part of this research, this is a limitation inherent to a research design which seeks to analyze large amounts of information scraped from the internet.

Furthermore, all participants are completely anonymized and no identifiable information, such as usernames, display names, profile pictures or tweets, is used in whole or in part in this paper. The data regarding this research is stored privately and is not made accessible to anyone except the researcher. In conclusion, due to the strong confidentiality measures used as well as due to the fact that research entails evaluating graph structures without taking into consideration content (beyond specific keywords), I don't believe this research runs the risk of harming any participants.

Results

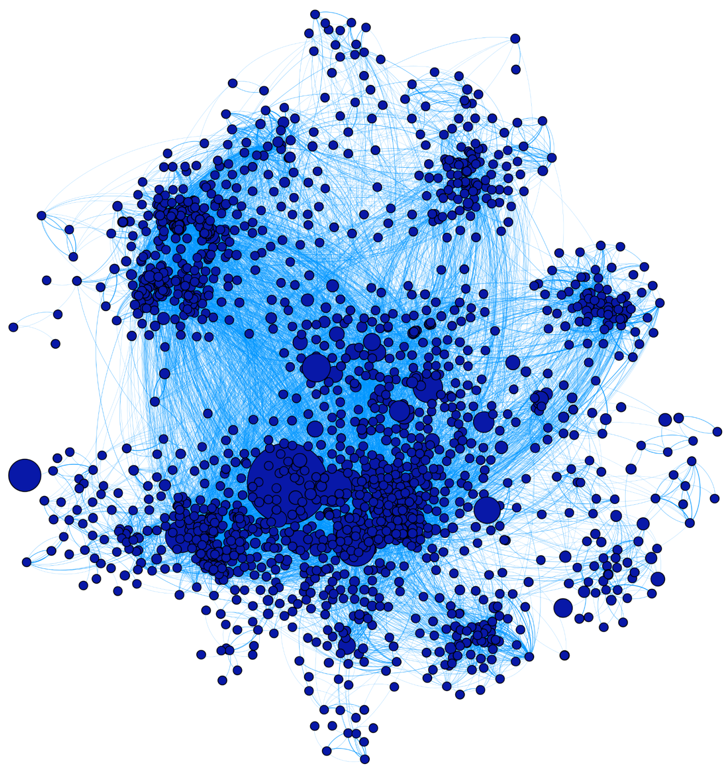

Figure 2: Visualization of the network

I used the graph visualization software Gephi in order to illustrate the network, as seen in Figure 2. Within the illustration, larger nodes represent more influential accounts and more opaque edges represent stronger ties, corresponding to more retweets from one account to another. Most highly influential nodes are visually identifiable as central, which is in agreement with what one would expect if there is a relation between centrality and influence. In order to see whether this tendency is supported by data, we will now turn to statistically analyzing the network.

Descriptives

Within this statistical analysis, two variables were selected as relevant to the research question: *eigen_centrality* (representing the computed Eigenvector Centrality) and *retweets_received_normalized* (the number of retweets the account in question received, normalized as previously established). They are taken as alternative measures of social influence within the network. Both variables are ratio variables, as there is a clear definition of a zero (i.e. zero centrality and zero retweets). *Eigen_centrality* has a minimum of 0.00 and a maximum of 0.34, for an average value of 0.0059 with a standard deviation of 0.02256. *Retweets_received_normalized* has a minimum of 0 and a maximum of 22839234, for an average value of 15848 with a standard deviation of 83727.

I decided to also compute the natural logarithm of the two defined variables, which will be used in a second statistical analysis. This is done as both variables do not vary simply linearly, but exponentially, in other words, through different orders of magnitude. A linear regression runs the risk of not taking the scale of the variation into account, and as such it could lead to misleading statistical results as the real statistical trends might not fit best within a simple linear model. In the following sections, I will illustrate the results of both a simple linear model (model 1) and a logarithmic model (model 2).

As 50 of the *retweets_received_normalized* values are zero, their natural logarithm is minus infinity, which cannot be used as part of a regression. In order to adjust for this, their values will be set to -1, so that these data points can still be used as part of the regression. The value of -1 is lower than the logarithm of any other value of *retweets_received_normalized* that is bigger than 0, and as such it can be used to map a value for which *retweets_received_normalized = 0* that is smaller than any normally occurring value.

After calculating the natural logarithm of the variables, I store the values within two new variables: *eigen_centrality_log* and *retweets_received_log*. These newly defined variables have the following descriptions: *eigen_centrality_log* has a minimum of -14.57 and a maximum of -1.07, for an average value of -8.3589 with a standard deviation of 2.92424, meanwhile, *retweets_received_log* has a minimum of -1 and a maximum of 14.64, for an average value of 6.2504 with a standard deviation of 2.98577.

Model 1

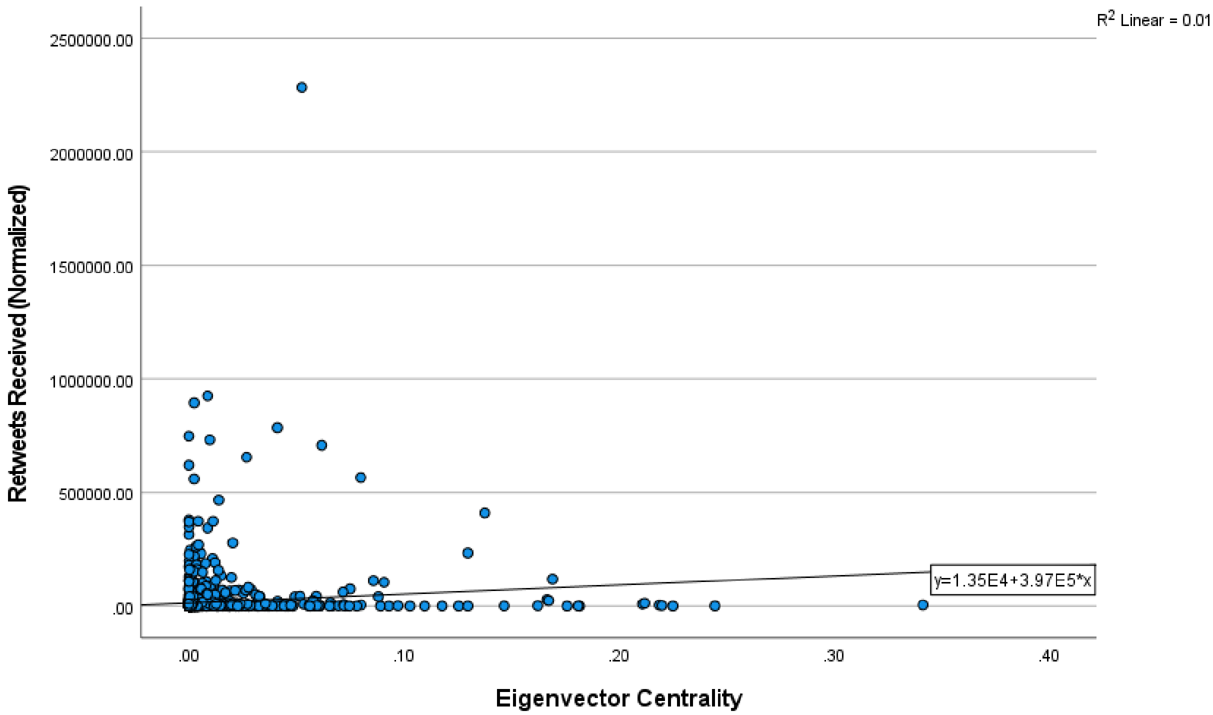

Figure 3: Chart for the Linear Model

For the linear model, I will be using two variables: *eigen_centrality* as the independent variable and *retweets_received_normalized* as the dependent variable. These variables will be used as part of a linear regression analysis using the software SPSS. As such, the model will find the linear mathematical formula that best predicts *retweets_received_normalized*, given *eigen_centrality*.

Firstly, The ANOVA test is significant with a p-value of less than 0.001, meaning that it is likely that the two variables differ significantly, a requirement for the linear regression to be meaningful. The R^2 value, which indicates to what extent the variation in the dependent variable (i.e. *retweets_received_normalized*) can be explained by the independent variable (i.e. *eigen_centrality*), is 0.011. In other words, only 1.1% of the variation in influence can be explained through Eigenvector Centrality alone.

By running the regression analysis (Figure 3), we find that the best fitting line is *retweets_received_normalized* = 13505.656 + 396818.534 * *eigen_centrality*. In other words, for each unit of increase within *eigen_centrality*, we would expect an increase of 396818.534 in *retweets_received_normalized*. This linear regression is also highly statistically significant, with a p-value of less than 0.001. As such, we find evidence against the null hypothesis and we conclude that it is likely that there is a statistically significant correlation between *retweets_received_normalized* and *eigen_centrality*.

The corresponding chart (Figure 3) does not graphically appear to show a linear correlation, despite one being statistically found. This is because many of the values tend to be near 0 for both variables, and as such the chart is heavily concentrated in one area and the dots tend to overlap with each other. Furthermore, since many data points are close to 0 on Eigenvector Centrality, there is a lot more variation within the data range close to 0 on *eigen_centrality*, erroneously making it seem like an inverse correlation would provide significantly better results. In order to get a clearer picture of what is occurring, we will turn to the logarithmic model.

Model 2

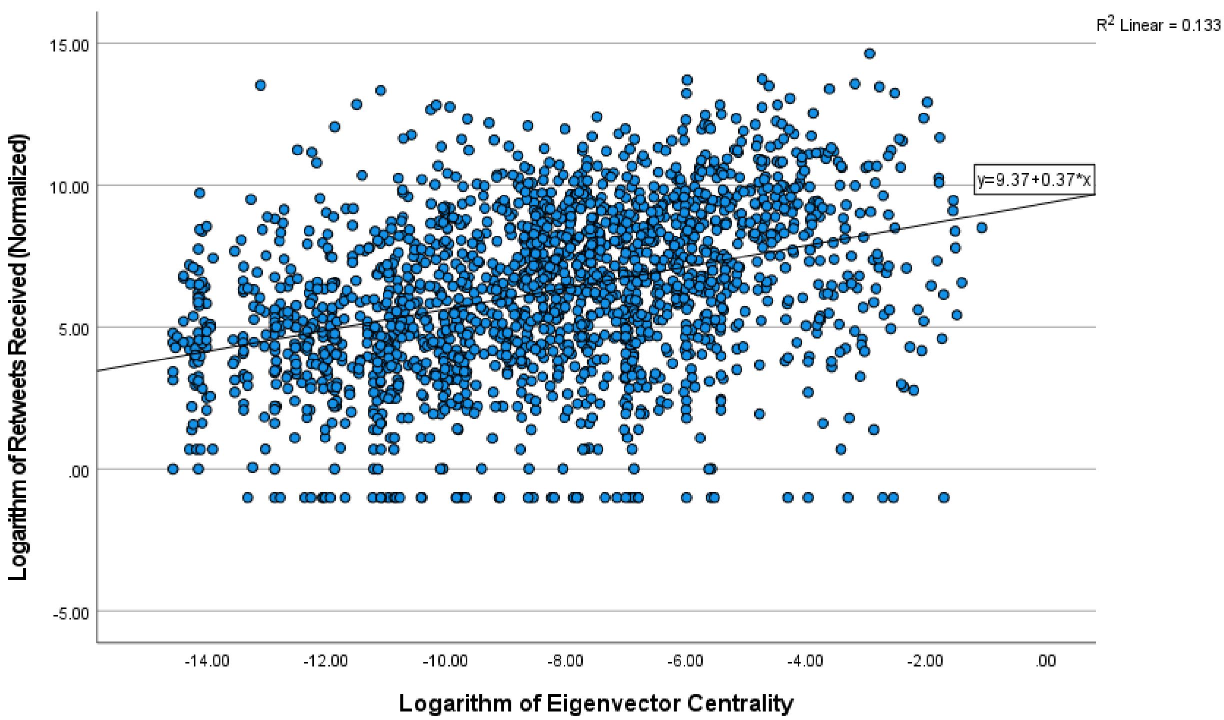

Figure 4: Chart for the Logarithmic Model

For the second model, which is a logarithmic model, I will be using *eigen_centrality_log* as the independent variable and *retweets_received_log* as the dependent variable. The model is still a simple linear regression model, however, as both variables are logarithmic in nature, it will find the linear formula that best predicts the logarithms of *eigen_centrality* (stored within *eigen_centrality_log*), given the logarithm of *retweets_received_normalized* (stored within *retweets_received_log*). After running the ANOVA test, it results in statistical significance ($p<0.001$), therefore, the two variables significantly differ and the linear regression model can proceed.

The linear formula found by the regression analysis is *retweets_received_log* = 9.367 + 0.373 * *eigen_centrality_log*. In other words, for each increase by 1 in *eigen_centrality_log*, we would expect *retweets_received_log* to increase by 0.373. This represents the influence of the logarithm of *eigen_centrality* on the logarithm of *retweets_received_normalized*. The correlation is statistically significant, with a p-value of less than 0.001.

The resulting chart and linear formula can be seen within Figure 4. The R^2 value is 0.133, meaning that 13.3% of the variation within *retweets_received_log* can be explained through the impact of *eigen_centrality_log*, which is significantly stronger than the 1.1% correlation found in Model 1. Given that the linear regression is statistically significant and that the ANOVA test succeeds, we find evidence that there is a correlation between *eigen_centrality_log* and *retweets_received_log*.

Interpretation of Results

The first model provides statistically significant ($p<0.001$) evidence for a correlation between Eigenvector Centrality and normalized number of retweets received. However, only 1.1% of the variation in retweets received can be explained by Eigenvector Centrality alone. This indicates that, while there is a correlation between the two variables, this is not the ideal manner to model the social influence of Twitter actors within the network under examination.

The second model, which involves doing a linear regression on the logarithm of both values, also proved to be statistically significant ($p<0.001$). Furthermore, 13.3% of the variation can be explained using this method, strongly improving on the 1.1% of variation explained through the first model. This indicates that Eigenvector Centrality proves to be a better predictor of the normalized number of retweets received once the logarithm of both values is taken.

As both Eigenvector Centrality and the normalized number of retweets received are taken as measures of social influence, this shows that the two measures are reliable with regards to each other, as they correlate to each other. While social influence as such is an intangible concept and cannot be directly measured as it is an intangible concept, the reliability of the two variables indicates that it is likely that they both correlate with the aforementioned influence. This would find evidence for Lazarsfeld et al.'s (1955) vision of how social influence spreads, as well as for Bruggeman's (2021) framework on Power, built upon Weber's (1922) original conception.

Discussion

I believe the methods used for data collection and processing, while far from perfect, do well in retrieving a representation of the network that fits our purposes. Empirically, labeling as relevant the accounts which have tweeted a keyword at least 5 times appears to run the risk of excluding other relevant accounts, rather than including irrelevant ones, a choice which I deem preferable. While it is possible that some false positive accounts could be accidentally included within the network, I could not find any evidence suggesting this is the case.

Furthermore, counting as an edge when one account retweeted another at least two times empirically is also effective. An account retweeting another only once can potentially correspond to a sporadic retweet irrelevant to the topic at hand, such as a widespread viral tweet. However, this is significantly less likely to be the case if one account retweeted the other at least two times, and it can be assumed there is a real tie between the two.

That being said, there are clear limitations to the methods used here. Firstly, as Snowball Sampling was used, the network found is not necessarily representative of the anti-COVID-19 vaccine Twitter network as a whole. As I have started the breadth-first network search from one account, all accounts found within this graph will correspond to accounts close-by in the network to it and highly distant accounts can not be found. However, due to the Six Degrees of Separation principle, in which even by searching with a small number of total edges one can find a large number of nodes (Watts and Strogatz 1998), this is unlikely to be a major issue, as the Snowball Sample is able to quickly identify nodes which are far away from the starting node.

Furthermore, the decisions made as part of the operationalization are based on common-sense assumptions of how social interactions on Twitter work. For instance, taking one account retweeting another numerous times as a sign of the latter influencing the former makes intuitive sense, however there can be alternative explanations, such as someone using the retweet functionality for different purposes. Furthermore, while there is prior backing for the idea that a high number of retweets received corresponds to high social influence (Cappelletti 2012), this is only the case if the individuals who have retweeted the tweet actually paid attention to the content of the tweet. It is also possible that a retweet intends to denounce opinions expressed in the original tweet, however, quote-tweeting is more commonly used for this function.

There is also the issue of the possibility that large parts of the networks could be bot accounts, that is, accounts that do not belong to an individual, but rather to automated programs. It has been shown that bots contribute to the amplification of misinformation, which is achieved by tweeting and retweeting misinformation (Shao et al. 2018). As the current research does not attempt to identify such bot accounts, their impact on the network cannot be assessed.

Finally, some coding errors have occurred that, while I don't believe run the risk of heavily influencing the results, must be taken into account. Namely, two of the keywords were accidentally combined due to a typing error and three accounts were accidentally retrieved twice and as such the second listing was manually deleted. Furthermore, as an oversight, within the received retweet count for each account, retweets on reply-tweets were not counted.

It is also possible that the statistical analysis could be flawed. Firstly, it could be that the results of the linear regressions are merely coincidental and do not represent the real situation for the entirety of Twitter. Furthermore, it is possible that the statistical analysis design is flawed itself and the findings are merely a result of the ways in which the data was operationalized and collected. However, I have no evidence for this being the case.

Conclusion

This thesis sought to analyze the relation between two different measures of Social Influence in anti-COVID vaccine networks of misinformation, namely Eigenvector Centrality and retweets received, in order to determine whether the former can be found to be a reliable measurement of social influence when correlated to the latter. Understanding which types of accounts tend to have the most influence within misinformation networks is an important part of creating measures that can dampen the misinformation crisis on social media, which has led to widespread vaccine hesitancy and refusal (Amin et al. 2022).

For these purposes, we used the Twitter REST API and the Python programming language, thus retrieving a network of 1840 relevant anti-COVID vaccine accounts to be analyzed. The retrieval algorithm used snowball sampling, starting from one account and expanding outward to identify an entire network. The accounts were identified as being part of the network either because they have tweeted anti-vaccine keywords at least 5 times, or otherwise due to their central position within the network.

Based on previous theoretical work (Weber 1922; Bruggeman 2021; Lazarsfeld et al. 1955), we established the relevance of Eigenvector Centrality in measuring an account's social influence within a network. We sought to analyze whether this measure empirically does well in identifying influential accounts. Since a direct way of determining how influential an account is does not exist, this research instead compares Eigenvector Centrality with another variable which this paper establishes based on previous theoretical work (Cappelletti 2012) as a measure of social influence, namely, the normalized number of retweets received.

We found a statistically significant ($p<0.001$) correlation between Eigenvector Centrality and retweets received within a linear regression model. However, merely 1.1% of the variation can be explained within this model. After taking the natural logarithm of both values in order to account for the exponentiality of both values, the amount of variation that can be explained rises significantly to 13.3%. This suggests that there is a significant correlation between the two measures used to model Social Influence within the network, and that the correlation between the two is stronger when modeled through a logarithmic approach.

As Eigenvector Centrality proves to be a reliable indicator of social influence, when compared to a second potential measure of such influence, we find some evidence for Lazarsfeld et al.'s (1955) theory of political influence within the networks of misinformation studied, according to which opinion leaders become influential through the spreading of their messages to people they hold a connection with. That Eigenvector Centrality is a reliable measure of influence also matches a Weberian (1922) framework of Power, as elaborated on by Bruggeman (2021). In this view, Power spreads through both direct (one-to-one) and indirect (through one or more intermediaries) connections, and as such actors that have a higher centrality tend to be more powerful.

This research is limited by the ways in which factors such as anti-vaccine sentiment and social ties were operationalized. As the study was done on Twitter data, it is impossible to know to what extent the interactions found can be taken as a proxy for real social interactions between individuals. The data collected and analyzed consist of social media interactions, and the ways in which social media data can be translated into sociological data on real persons remain open to debate. Furthermore, the role that bot accounts play within this network cannot be assessed using the current methods.

Further research is needed in identifying the other factors which impact an account's social influence beyond the centrality measure considered here, as well as whether other centrality measures can be shown to be more closely linked with people's social influence. Further research is also necessary to confirm whether the statistical correlations identified here can be replicated, or whether they represent a statistical fluke.

Bibliography

- Adam, David, and others. 2022. “15 Million People Have Died in the Pandemic, WHO Says.” *Nature* 605 (7909): 206.
- Aggrawal, Niyati, and Adarsh Anand. 2022. *Social Networks: Modelling and Analysis*. CRC Press.
- Ahmed, Imran. 2021. “Dismantling the Anti-Vaxx Industry.” *Nature Medicine* 27 (3): 366. https://doi.org/10.1038/s41591-021-01260-6.
- Amin, Krutika, Jared Ortaliza, Cynthia Cox, Joshua Michaud, and Jennifer Kates. 2022. “COVID-19 Mortality Preventable by Vaccines.” Peterson-KFF Health System Tracker. 2022. https://www.healthsystemtracker.org/brief/covid19-and-other-leading-causes-of-death-in-the-us/.
- Bonacich, Phillip. 1987. “Power and Centrality: A Family of Measures.” *American Journal of Sociology* 92 (5): 1170–82. https://doi.org/10.1086/228631.
- Boulet, Romain, and Jean-Fabrice Lebraty. 2018. “A New Influence Measure Based on Graph Centralities and Social Network Behavior Applied to Twitter Data.” *Communications of the Association for Information Systems* 43 (1): 21.
- Bruggeman, Jeroen. 2021. “Power.” Network Analysis.
- Bruns, Richard, Divya Hosangadi, Marc Trotochaud, and K Sell. 2021. “COVID-19 Vaccine Misinformation and Disinformation Costs and Estimated $50 to $300 Million Each Day.” *Baltimore, MD: The Johns Hopkins Center for Health Security. Https://Www. Centerforhealthsecurity. Org/Our-Work/Publications/Covid-19-Vaccine-Misinformation-and-Disinformation-Costs-an-Estimated-50-to-300-Million-Each-Da*.
- Butcher, Paul. 2021. “COVID-19 as a Turning Point in the Fight against Disinformation.” *Nature Electronics* 4 (1): 7–9.
- Cappelletti, Rafael, and Nishanth Sastry. 2012. “Iarank: Ranking Users on Twitter in near Real-Time, Based on Their Information Amplification Potential.” In *2012 International Conference on Social Informatics*, 70–77.

- Centola, Damon, and Michael Macy. 2007. “Complex Contagions and the Weakness of Long Ties.” *American Journal of Sociology* 113 (3): 702–34.
- Featherstone, Jieyu D, George A Barnett, Jeanette B Ruiz, Yurong Zhuang, and Benjamin J Millam. 2020. “Exploring Childhood Anti-Vaccine and pro-Vaccine Communities on Twitter--a Perspective from Influential Users.” *Online Social Networks and Media* 20: 100105.
- Goel, Sharad, Ashton Anderson, Jake Hofman, and Duncan J Watts. 2016. “The Structural Virality of Online Diffusion.” *Management Science* 62 (1): 180–96.
- Habermas, Jurgen. 1991. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. MIT press.
- Hagberg, Aric, Pieter Swart, and Daniel S Chult. 2008. “Exploring Network Structure, Dynamics, and Function Using NetworkX.”
- Huszár, Ferenc, Sofia Ira Ktena, Conor O’Brien, Luca Belli, Andrew Schlaikjer, and Moritz Hardt. 2022. “Algorithmic Amplification of Politics on Twitter.” *Proceedings of the National Academy of Sciences* 119 (1).
- Jungherr, Andreas, and Ralph Schroeder. 2021. “Disinformation and the Structural Transformations of the Public Arena: Addressing the Actual Challenges to Democracy.” *Social Media+ Society* 7 (1): 2056305121988928.
- Lazarsfeld, Paul F, Bernard Berelson, and Hazel Gaudet. 1968. *The People’s Choice*. Columbia University Press.
- Liu, Zhe, and Ingmar Weber. 2014. “Is Twitter a Public Sphere for Online Conflicts? A Cross-Ideological and Cross-Hierarchical Look.” In *International Conference on Social Informatics*, 336–47.
- Lorenz-Spreen, Philipp, Lisa Oswald, Stephan Lewandowsky, and Ralph Hertwig. 2021. “Digital Media and Democracy: A Systematic Review of Causal and Correlational Evidence Worldwide.”
- Lukes, Steven. 2021. *Power: A Radical View*. Bloomsbury Publishing.
- Moe, Hallvard. 2012. “Who Participates and How? Twitter as an Arena for Public Debate about the Data Retention Directive in Norway.” *International Journal of Communication* 6: 23.

- NetworkX. 2022. “Eigenvector_centrality.” NetworkX. 2022. https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.eigenvector_centrality.html.
- Nogrady, Bianca, and others. 2021. “‘I Hope You Die’: How the COVID Pandemic Unleashed Attacks on Scientists.” *Nature* 598 (7880): 250–53.
- Roozenbeek, Jon, Claudia R Schneider, Sarah Dryhurst, John Kerr, Alexandra L J Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden. 2020. “Susceptibility to Misinformation about COVID-19 around the World.” *Royal Society Open Science* 7 (10): 201199.
- Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. “The Spread of Low-Credibility Content by Social Bots.” *Nature Communications* 9 (1): 1–9.
- Shore, Jesse, Jiye Baek, and Chrysanthos Dellarocas. 2016. “Network Structure and Patterns of Information Diversity on Twitter.” *ArXiv Preprint ArXiv:1607.06795*.
- Soares, Felipe Bonow, Raquel Recuero, and Gabriela Zago. 2018. “Influencers in Polarized Political Networks on Twitter.” In *Proceedings of the 9th International Conference on Social Media and Society*, 168–77.
- Tremayne, Mark. 2014. “Anatomy of Protest in the Digital Era: A Network Analysis of Twitter and Occupy Wall Street.” *Social Movement Studies* 13 (1): 110–26.
- Twitter Developer Platform. 2022. “Twitter API Documentation.” Twitter. 2022. https://developer.twitter.com/en/docs/twitter-api.
- Linden, Sander Van der. 2022. “Misinformation: Susceptibility, Spread, and Interventions to Immunize the Public.” *Nature Medicine*, 1–8.
- Watts, Duncan J, and Steven H Strogatz. 1998. “Collective Dynamics of ‘Small-World’ Networks.” *Nature* 393 (6684): 440–42.
- Weber, Max. 1922. *Wirtschaft Und Gesellschaft*. Tübingen: Mohr.
- Yardi, Sarita, and Danah Boyd. 2010. “Dynamic Debates: An Analysis of Group Polarization over Time on Twitter.” *Bulletin of Science, Technology & Society* 30 (5): 316–27.