Framing Events of 2020 within Echo Chambers: A Twitter Network Analysis

Submitted as partial fulfilment of a Master's of Science in Geospatial Technologies

Tom Truong[1]

[1]University of Washington Tacoma
School of Urban Studies
1900 Commerce Street
Tacoma, WA 98402-3100
Email: tomtruong03@yahoo.com

Abstract

In the age of digital media, sources of news have diversified both in origin and perspectives. Whether factual or plain misinformation, coverage about the current COVID-19 pandemic, United States' racial unrest, and the presential election have garnered various attitudes. Previous studies about the accuracy of stories have mainly focused on the relationship of where an individual resides, and the news covered in their region. Periodically, emerging alternative online news sources have competed with traditional markets over who is subjectively more truthful and has less perceived bias. This has opened the doors for online communities made up of like-minded individuals, that are now connected beyond proximity.

This study investigates the potential spatial connections between individuals within online echo chambers. Similarly, to how local markets reach their audience within a serviced area, there also could be a possible spatial connectivity within an online community. While online activity is usually concentrated in more populated areas, I theorize that these communities are more likely form from users residing in suburban areas around these large cities. To visualize this practically, a social network analysis through the connection of nodes (users) and edges (online interactions) can tie in this relationship. Throughout this project, we expect to correlate the transition between regional ties to tradition media, to the emerging alternative news sources that are currently formed by less structured spatial relationships.

1. Introduction

The purpose of this capstone is to discover the audience that an online news source attracts, based on where these individuals reside. This will incorporate the relationship of how one chooses their news sources and how that ultimately shapes their views and opinions on current events. With the integration of news within social media, the opinion of the public is more scattered and less tied to geographical regions (Mersey 2009). This brings into question if these audiences tend to come from users within proximity of each other, which is usually the means of older media being serviced within a certain region. Based on how curated news online has become, there has been a demand for more individualistic news for various perspectives and interests (Owen & Wei 2020). This project takes on that fact and investigates to see if there is a correlation of the following of niche ideas and one's geographic location.

Exploring through the dynamics beyond this phenomenon is crucial to the growing issue of misinformation, that has inflated with online news. Misinformation could be more susceptible to certain audiences, due to the variance of news accuracy (Ahmed *et al.* 2019). This capstone aims to make aware of where individuals tend to be more misinformed or indifferent to opposing viewpoints of current events in our country. The following are the research questions that will be discussed upon throughout this project:

- What sort of audiences are more susceptible to misinformation based on the sources of news they consume?
- Does the origin of online news sources indicate the kind of audiences they reach based on geographic proximity?
- How well does alternative news media create echo chambers that shape an audience's views and opinions on current events?

The methods used to answer these questions are utilizing Twitter data from given hashtags to see the mutual online interactions between users and where they reside geographically. The hashtags that are being analyzed are **#Plandemic** (the idea that the current COVID-19 pandemic has been planned out for various political and health related reasons), **#AbolishThePolice** (the call for the abolition or replacement of the current policing system in the United States, amid current racial unrest) and **#TrumpWon** (the idea that the then-president Donald Trump won the 2020 presidential election, due to claims of fraudulent voting). This capstone will seek out how well alternative news online has shaped these echo chamber of ideas, based on if their audiences are concentrated in one region or in specific areas around the country. Ultimately, this project aims to bring awareness to areas that are susceptible to being more ingrained in echo chambers. This also entails a call to action for individuals to become conscious in seeking out more diverse and balanced perspectives in the current market of news media.

2. Literature Review

The targeted audience for traditional media in the form of radio, newspaper, and cable television has been usually their respective serviced region. But with the internet becoming more accessible, news stories are now transforming the public's view of how the press should operate. Standards such as trustworthiness, being instantaneous and speculator news has shifted audiences to online news sources (Elejalde *et al.* 2019). With the emergence of alternative sources online, it has brought more variance to the perspective and accuracy of news stories. These "middle-level gatekeepers" now have the edge of influencing the flow of information to give their audience, "…different preferences for the types of messages they spread…" (Hemsley 2019). With the abundance of information of news spreading on social media, data about the users who partake within these communities are made accessible. Previous studies have analyzed the connections of individuals sharing similar opinions and how it ties in spatially (Bastos *et al.* 2018). This was using Twitter activity stemming from tweeted coverage of an event or issue, that can make bigger rounds nationwide than it would, if it came from traditional media.

The following literature first touches on the historical use of traditional media and how the masses have transitioned to online alternative sources for news. From there, with the emergence of online news sources, a discussion about how it became appealing to media consumers and how it shaped the variance of perspective and accuracy of news stories. Finally, this ties into studies investigating the spatial connections of online communities through the shared connections between users and their respective residences. That brings me into my intention of this capstone, which is to analyze the mutual interactions between Twitter users who share similar settlement of controversial ideas and where they connect by proximity throughout the United States.

2.1 Transitioning from Traditional Media to Online

Starting with the recent past, traditional news markets have attempted to adapt with growing online alternatives, by providing a more "pick and choose" method of news coverage. This is motivated by how within social media, there are news feeds that are usually curated to match an individual's taste (Fitzpatrick 2018). Typically, this is in the form of articles that accompanies a headline and/or image that encapsulates the story to fit the narrative that the source so chooses. The demand for this sensational news has become more prevalent as older news mediums struggle to meet the "thirst to be first" strategy of online news sources (Fitzpatrick 2018). This has resulted in more media outlets becoming victim of "visibility bias" of becoming more indifferent to other perspectives that aren't their prioritized focus (Eberl *et al.* 2015). So, the viewpoints of coverage from traditional media have failed to match the diversity of perspectives offered by online sources. Over time, this has caused the perception of mainstream media to become more biased and less trustworthy to properly inform the public (Ardèvol-Abreu & Gil de Zúñiga 2016).

This literature provides insight to how much power alternative news sources have to the market of distributed news. Also, it has been speculated that there is a profit incentive that drive these online sources to try to manipulate public opinion. (Owen & Wei 2010). This could be from third parties that try to skew news stories to fit their agenda, with some actors being the wealthy, corporations, or even political groups. This is certainly prevalent in high stake areas, such as more populated and dense regions where the public opinion is highly influential to the rest of the country. So there seems to be a correlation of less of these dense populations partaking in local television and talk radio news, where online news is more widely available (Althaus 2009). In the realm of newspapers, a carrier like The New York Times have expanded out of their local market by focusing on more national news topic that an incoming audience would be keener to. This had led to more negligence of issues at a local level, resulting in more individuals being less informed and focused on local policies and events (George & Waldfogel 2008). That shows how traditional media in many ways have been directly influenced by alternative online media, that has become more advanced in adopting a diverse set of content.

2.2 Emergence of Alternative News Sources Online

These alternative news sources have bypassed the "gatekeeping" of information, originally controlled by traditional outlets. Now, online communities have been created to make an accessible and welcoming community to like-minded news consumers. Rather than keeping the motive of maintaining a reputation kept by traditional sources (Ahmed 2019), these online platforms have more incentive in "ideological inclinations" (Elejalde *et al.* 2019). The mechanisms of liking and following groups or personalities have allowed users to seek information that amplifies and reinforces their beliefs. This concept has existed before on "closed and symmetrical social media" that are made up of online forums that share controversial and/or questionable news. But now, it has made way into "open and asymmetrical social media", such as Facebook and Twitter (Kim & Ihm 2019). This brings into question if traditional news outlets choosing to adapt attention grabbing tactics has been detrimental to their market. What has caused individuals to accuse mainstream media to be biased, has now allowed room for echo chambers of information online.

This research has shown that news online has made room for echo chambers to develop for like-minded news consumers. Social media has developed into a platform for news stories to be advertised to users, based on interests and views, without that much explicit assistance from the user (Gil de Zúñiga *et al.* 2017). Platforms such as Facebook and Twitter have made it easy for news to find the user, rather than the user having to actively search. This confronts the challenge to figuring out what regional audiences are more likely to explore out from traditional media to venture into online environments. This is questioned by the online news targeted advertising is more scattered geographically, than it would with local news markets targeting their regional area. One article briefly discussed this, as they stated that, "…the geo-localization of users based in their network (based on the assumption that users are more likely to interact with other users that are geographically closer to them) are more accurate at a finer level" (Mersey 2009). Similarly, neighboring users tend to have similar confirmation biases that is leveraged by peer influence and reinforcement (Brugnoli *et al.* 2019), which is parallel to how individuals that tend to consume their local news markets if they feel the "neighborhood effect" of being influenced by the people they live around (Althaus *et al.* 2009; Mersey 2009).

With how information can be spread traditionally by word of mouth, it comes into question if rumors and misinformation can evolve that way through geographic proximity. A sudden event that originated in a concentrated area can develop "merging rumors" or misinformation, that could eventually become a "long-standing rumor" as it spreads throughout the area (Zubiaga *et al.* 2018). To see how well an idea can spread throughout the country from like-minded online communities, analysis on social media interaction such on Twitter has been a prominent method. Usually, it is how well a certain hashtag can spread throughout different locations beyond proximity.

2.3 Spatial Relations of Online News Communities

Since the activity on Twitter makes up a fraction of internet users that represent only "about one-third of [the] global population" (Crampton *et al.* 2013), it means that this doesn't perfectly reflect the real-life public perception of current events. But it does compliment the instantaneous or real time nature of online news, like mentioned before. Using geotagged tweets provide an abundance of information, beyond just location coordinates. It also comes with data including time (Shelton 2017), which would be beneficial to analyze the temporal progression of a hashtag. Regarding topics previously studied using this form of social media analysis, notable subjects related to health (Kim 2015) and controversial conspiracy theories about current events (Zollo 2015) are on the forefront. One case study follows the attitudes of vaccine beliefs within certain demographics across the United States. After the analysis of tweets, it resulted in higher "autism-related anti-vaccine beliefs" in more urbanized and highly populated areas (Tomeny *et al.* 2017). With these online "middle-level gatekeepers" being bulk of the ideological holders on Twitter (Hemsley 2019), they play a big role in dictating the next viral hashtag, which most likely would target people in power or someone who has more of a following (Graham 2020).

Another issue with using Twitter to reflect the real-life perception of current events, is the abundance of unauthentic bot and fringe accounts. This common theme is mentioned in much research regarding social network analysis (Lim *et al.* 2019; Crampton *et al.* 2013; Hemsley 2019; Graham 2020; Tomeny *et al.* 2017). It seems that the main objective of these bot and troll accounts is to amplify polarizing messages and to promote controversial discord (Broniatowski *et al.* 2018). This brings into question of how previous studies have handled bot accounts. One piece of literature referenced how they used a "bot detection protocol" model that identified suspicious bot activity (Bastos *et al.* 2018) and another identified that the bot effect is negligible to their study's results (Hemsley 2019). These readings provided some insight and advice to how bot accounts can be handled in this capstone using Twitter data for analysis. Nevertheless, each article provides evidence that there is some connection between echo chamber communities and their proximity with other members. This is most likely due to each literature showcasing tweeted events that have occurred within a short period of time or within a concentrated area.

With my research, I will investigate more contemporary and long-term events that had more of an effect nationwide. The accompanying literature has provided knowledge of how the geographic proximity of online communities can be derived from local news markets servicing their region. But with alternative sources emerging online, it brings the extra element of echo chambers that are built on individuals looking to curate their news consumption, unaware of possible confirmation biases. That introduces the question of what kinds of audiences are more likely to be susceptible to misinformation, based on what is shared within their respective online communities. Using data provided from the Twitter platform, the analysis will test to see if there's a spatial relationship between users of the same audience. This is to help answer if online atmospheres can amplify political opinions, that could lead to "swirling claims and counterclaims about the facts and fabricated news stories" (Dodge 2018). Whether it's concentrated within one region or mainly coming from certain areas (urban, suburban, rural) scattered across the map, this analysis can speak a lot to who is following into echo chamber sharing information online.

3. Description of Application/Intervention

This capstone's technical portion consists of two parts: social network analysis and time-lapse heat map analysis. As mentioned in the beginning, the three hashtags being analyzed are **#Plandemic**, **#AbolishThePolice**, and **#TrumpWon**. The network analysis takes about a sample size around two thousand tweets per hashtag to be graphed. The way it is laid out is that each user's tweet is represented by a node. The way these nodes are connected are by edges that connect by mutual retweets, replies, user mentions, and media shared between the users. The graph's layout is revolved around a central node that holds the most interconnectivity, based on the highest number of mutual retweets, replies, and user mentions shared with other users of this hashtag.

With the map analysis, the time-lapsed heat maps analyzes the progression of a hashtag as it spreads throughout the country over time. Since the tweets now must be filtered down by being geolocated as a recognizable location within the contiguous United States, the sample size has been condensed around one thousand tweets per hashtag. These geotagged tweets have been aggregated by general location and then by day, within a short period of days. The heat map will then show how concentrated the tweets are on a given day, based on where those tweets were located. Overall, these two visualizations support in viewing how well a viral hashtag reaches an audience outside the origin region of an event. Also, it helps support in seeking the well connected these users who have similar ideologies are, by the mutuality of interactions on Twitter.

4. Methods

The technical component of this capstone is composed of multiple data visualizations to show how echo chambers are spatially represented in the United States. This process is planned out to be implemented in three phases. First is to collect data from Twitter, then compose that data for network analysis, and finally implement it into a spatial representation. Extracting geotagging data from the tweets will tie into answering is such echo chambers are made up of individuals residing in certain areas. My hypothesis is that the reinforcement of these controversial ideas will mainly comes from suburban areas around major metropolitans in the United States. This experiment will also support the idea that alternative news mediums have some sort of proximity connection, like traditional regional news markets.

To start off, tweet data need to be scrapped through Twitter. Through a Python working environment, the library *snscrape* is used to compile the needed information from each tweet, most importantly the location. For location coordinates, their only available for tweets sent from GPS enabled devices that allow Twitter that accessed information, according to Twitter documentation (Vivek 2021). That's why only about 1-2% of tweets are geotagged (Tomlins 2020), the closest best method is extracting the location given on a user's profile. The scraping library provides an abundance of attributes that can be extracted from tweets, but what is mainly needed is timestamps, information of retweets, replies, user mentions, and types of media linked from tweets. Scrapping will allow to compile all the tweets under one of the three hashtags and within a time period during peak activity of the respective event.

As for the hashtags mentioned before, they represent major events of 2020 that created an isolated audience to believe in such ideas or theories. This is inspired from what phrases have snowballed to be a defining perspective of these controversial events. Temporal configurations will be made to filter down the tweets during the duration of the respective events of each hashtag. The starting point of the timeline will start from the date of the real-world event, to when the hot topic starts fading out. As mentioned before, location of said tweets will be geolocated by the user's stated location from their profile and as of 2012, "approximately 60% of all tweets can ultimately be associated with a physical location with some degree of confidence" (Crampton *et al.* 2013).

Once the tweets are collected, they will be used to generate visualizations for a non-spatial network analysis. Using the library NetworkX, users sharing the same hashtag can be graphed as nodes and they will be interconnected by edges based on their common Twitter activity. This is from sharing mutual retweets, replies, user mentions, and media shared from tweets. This analysis will support answering if there's any correlation between the sharing of one idea from a current event or situation, would lead individuals to become part of the same echo chamber. It will prove whether alternative news sources have the power to influence echo chambers and do so spatially. The distance between these connected nodes will dictate how well common media activity is throughout this echo chamber.

To prepare the dataset for the cartographic visualization, it needs to be filtered down within a bounded area of the contiguous United States (created by an array set of coordinates), so the later map analysis could avoid any geographic outliers. This is done by geocoding each tweet from location name into coordinate pairs. This is made possible from a Geolocator library, that "can identify a real-world location and provide some extra details such as latitude and longitude" (Lungu 2019). From there, the pair is made into a point and the filtering process indicates if that point is contained in the United States polygon's array of coordinates. This smaller sample size of geotagged tweets is now ready for map analysis.

The last step is to use the tweets for the cartographic visualization. For the time-lapse functionality, the tweets will be grouped by the date it was sent within a certain time period. For the heat map to operate, each tweet will be aggregated within its general region to visualize the density of tweets in a particular area. There will be a snapshot of the map for each day the sample set of tweets are within, each visualizing the concentration of tweets within an area based on the coloring. From there, each frame will come together to make an animated .gif that visualizes the temporal progression of a certain hashtag. As will be discussed later, a hashtag could stem from a particular area or region and this visualization shows how well an idea could spread throughout the country.

This three-step process will help answer what possible audiences are more likely to be part of an echo chamber of homogenous ideas, based on where they reside from. The idea of echo chambers being spatially connected will be supported if the users share common activity of retweeting, replying, mutual user mentioning and media sharing; signifying that they are like-minded in a way. The social network analysis will first begin testing if there is a particular interconnectivity of Twitter activity between users. If this checks out, then visualizing it on a map could indicate if there is a spatial pattern for an online echo chamber community. Also, it indicates how well a hashtag can be distributed throughout the country by staying maintained in some areas and spreading to new ones.

5. Discussion

This section is going over the results of the analysis from the visualizations. Starting with the network analysis, this will go over how the graphs ended up being structured. Again, the graphs visualize the mutual interactions between users of a given hashtag. The three hashtags are **#Plandemic**, **#AbolishThePolice**, and **#TrumpWon**. The graph is built around what is called the center node. Table 1 shows the interconnectivity within the hub of this center node with total amount of nodes and edges for this largest component of the graph:

Table 1. Largest Component of each Graph

Hashtag	Sample Size	Total Nodes	Total Edges
#Plandemic	1,859 tweets	773	779
#AbolishThePolice	1,897	684	683
#TrumpWon	1,719	463	530

This center hub (pictured down below), represents a pivotal tweet that gathered a lot of attention between users who used the same hashtag. This attention is built on common Twitter interaction, which indicated before, was mutual retweets, replies, user mentions, and media shared. The chart indicates that this central hub makes up at least a quarter of the Twitter interactions within the sample sizes for each tweet. In Figures 1-3 down below, the network graphs show the center hub of interconnectivity for each of the three hashtags:

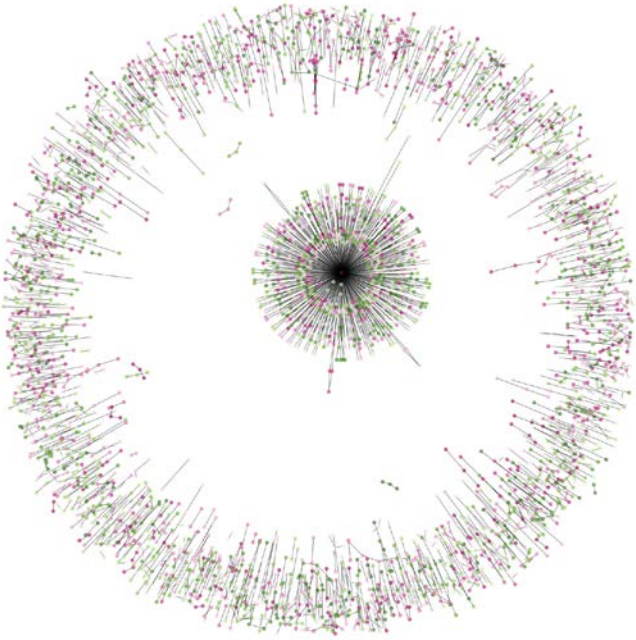

Figure 1. Network Graph of #Plandemic *(n=1859)*

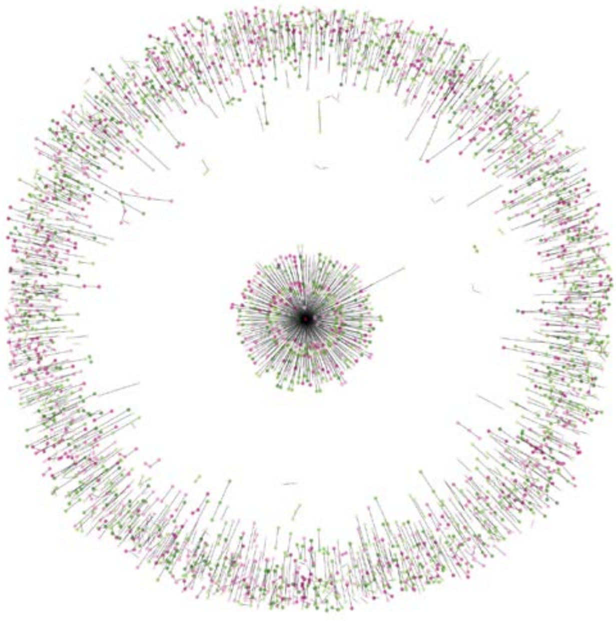

Figure 2. Network Graph of #AbolishThePolice *(n=1897)*

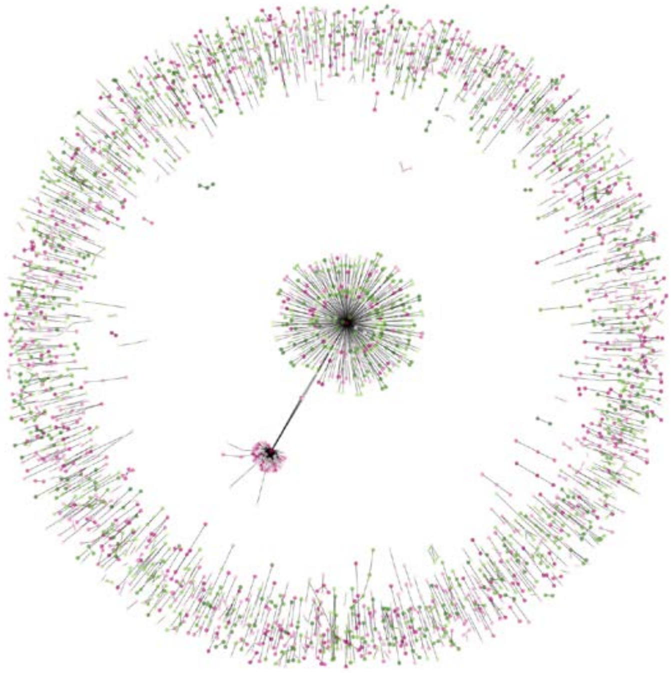

Figure 3. Network Graph of #TrumpWon *(n=1719)*

The central node is known to be the “Network Influencer”. The influence they have within this communication sphere is measured by three centrality measures: Degree Centrality, Closeness Centrality, and Betweenness Centrality. **Degree Centrality** is measured to be “the number of connections a particular node has in the network” (Kapoor 2018). The **Closeness Centrality** is known to be a measure of how important a node is if it’s connected to other important nodes. Finally, the **Betweenness Centrality** indicates “the frequency at which a point occurs on the geodesic (shortest paths) that connected pair of points” (Kapoor 2018). So, this means the node with the highest betweenness centrality essentially has the most influence in information flow within the network, which means the biggest influence on other users. Table 2 goes over the measurement outputted for each measure for each of the three hashtags discussed:

Table 1. Centrality Measures of Center Node

Hashtag	Degree	Closeness	Betweenness
#Plandemic	0.09	0.12	0.04
#AbolishThePolice	0.05	0.07	0.01
#TrumpWon	0.05	0.07	0.02

Based on these numbers, the Network Influencer for **#Plandemic** can be viewed as the node which has the most pull of influence towards the rest of the network. Even looking at the graph visualization itself, the central hub has quite a higher density of nodes and edges compared to the other two hashtags. Even though it is slightly skewed north-eastward, there is still a fair number of connections on that general of the outer ring. With the other two hashtags’ central hub, they came close in terms of centrality measures, but compares less in terms of influence in information flow for their respective central node. This brings into question, what was the tweet and the user that is represented as the central node for each of the three hashtags? Figures 4-6 will show off which tweets garnered the most combined Twitter activity within the networks:

> We've had Cancer for many years and have never developed a vaccine and all of a sudden we're gonna have one for #COVID19 ?
>
> "We'll know our disinformation program is complete when everything the American public believes is false."
> CIA Director William Casey
>
> #Plandemic
>
> - @GoodShepherd316 ***(51.6k followers)***
>
> May 27th, 2020. Land of Lincoln, USA.

Figure 4. Network Influencer Tweet of #Plandemic

> Y'all. Y'ALL. They cancelled the police contract with public schools!! Ed folks wondering how to make an impact - THIS IS HOW. Abolition aint gonna happen in one fell swoop - cops out of schools keeps our kids safe AND cuts the police revenue #defundthepolice #AbolishThePolice https://t.co/MGNTa3mv62
>
> - @savannahshange ***(7.5k followers)***
>
> June 3rd, 2020. Zamunda.

Figure 5. Network Influencer Tweet of #AbolishThePolice

> There's going to be a lot of people who believe in God once this is all said and done. #TrumpWon
>
> - @Annakhait ***(295.2k followers)***
>
> November 11th, 2020. United States.

Figure 6. Network Influencer Tweet of #TrumpWon

As seen from these tweets, these "Network Influencers" have a rather good following being in the thousands of followers. We see a wide range of followers for each hashtag, determining that a following of any size can bring in a remarkable amount of Twitter activity within their niche communities. This brings back the idea of "middle-level gatekeepers" from before, that can influence the flow of information with ideas that aren't mainstream that is covered on traditional media.

Moving on, is to show how these ideas spread out on a geographic level. The following Figures 7-9 will visualize a time-lapsed heat map of the distribution of tweets for each hashtag within a certain time period. The time period of **#Plandemic** was set for May 15^{th}, 2020 (shortly after the documentary of the same name was released) to June 1^{st} (during the first weekend of racial unrest in 2020). Slightly overlapping with **#AbolishThePolice**, the dates were set for May 25^{th}, 2020 (the murder date of George Floyd) to June 5^{th} (a week into the civil unrest). Finally, for **#TrumpWon** the days start from November 4^{th}, 2020 (the day after the election night) to November 14^{th} (the day of the 'Million MAGA March' event). The figures below demonstrate the growth of the hashtag throughout time:

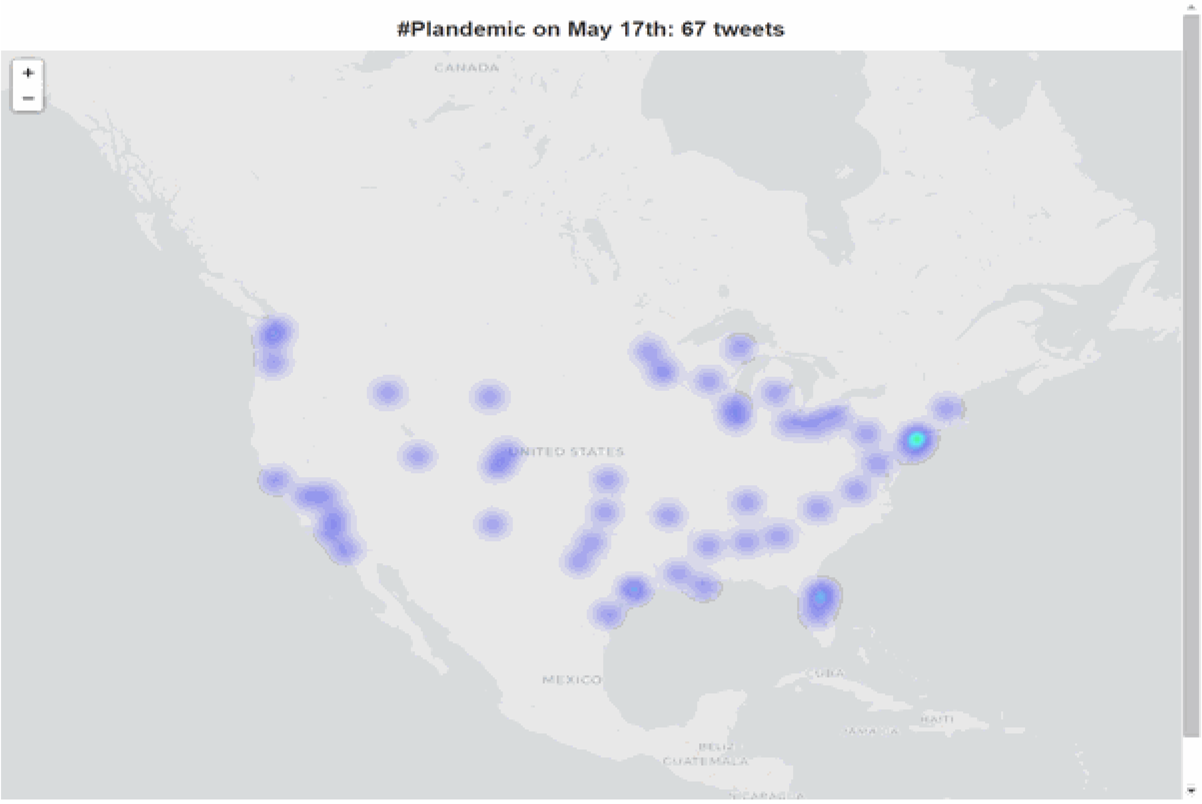

Figure 7. Screenshot of the use of #Plandemic throughout United States

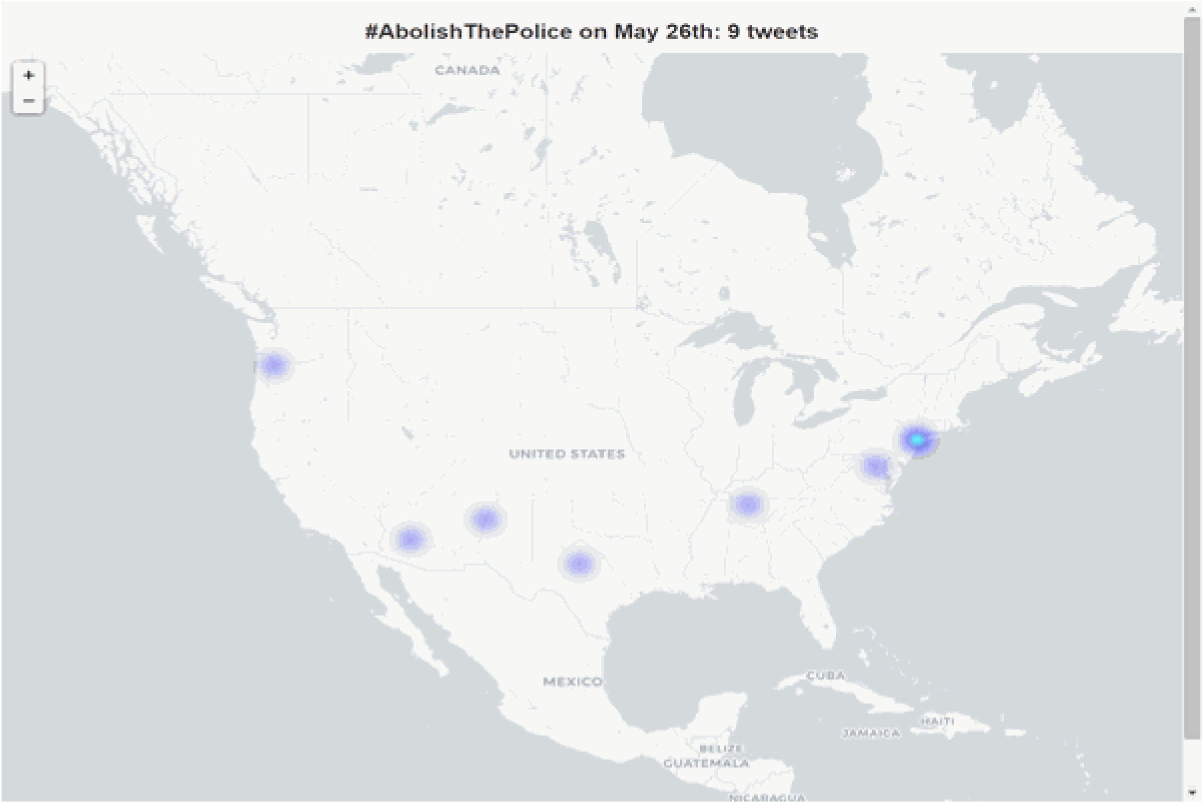

Figure 8. Screenshot of the use of #AbolishThePolice throughout United States

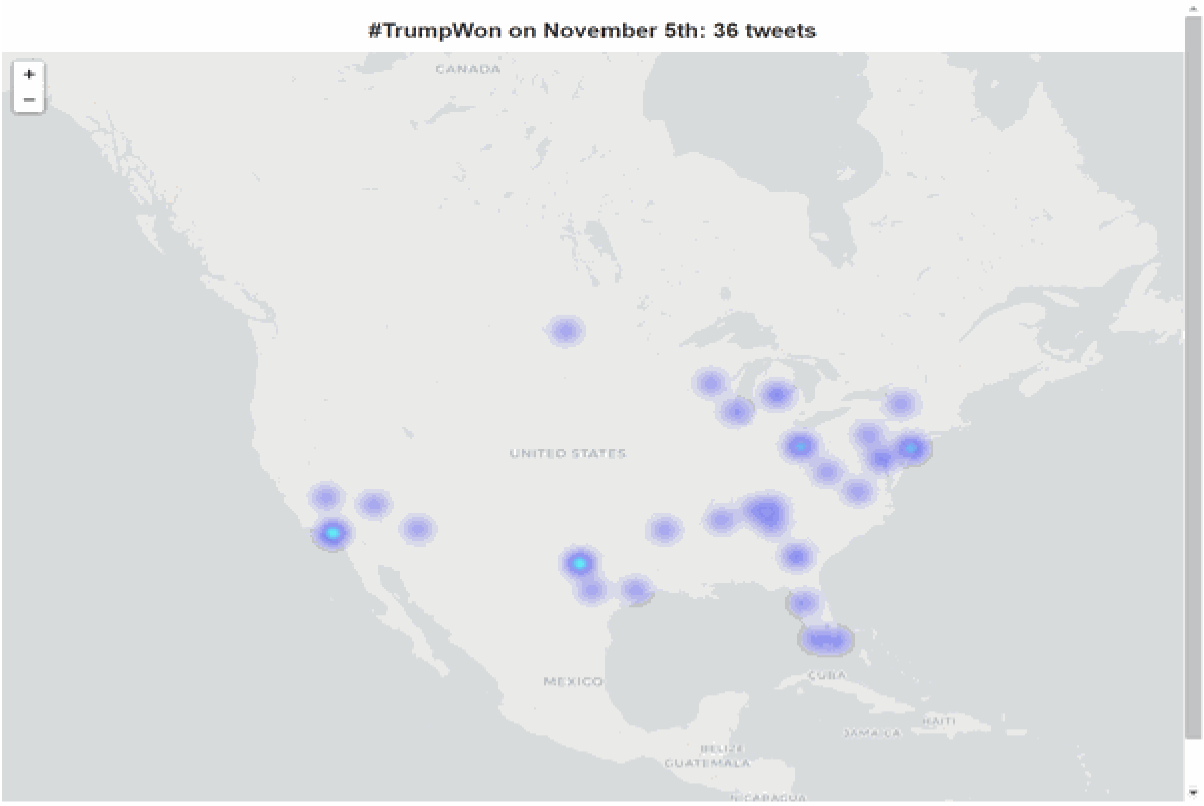

Figure 9. Screenshot of the use of #TrumpWon throughout United States

These still screenshots represent the initial days of the time period for these hashtags. In terms of how well each hashtag is distributed throughout the country within time, it general is concentrated is particular areas and may spread to regions around its proximity. From the hashtag **#Plandemic**, initial activity has come from more urban environments, possibly from states with stricter public guidelines due to the pandemic. That activity sometimes reaches more remote areas but is mainly concentrated in urban centers across the country like parts of California, Texas, New York, and Florida. For **#AbolishThePolice**, initial results were dense around states near Minnesota (the location of George Floyd's murder) and continued to have a presence throughout time. Also more urban, and liberal areas around the country saw prominent use of the hashtag which reflects well off the real-world unrest happening in these areas. The areas being the Pacific Northwest (Washington and Oregon), Southern California, and New York. Finally, with the hashtag **#TrumpWon**, it is more scattered than just the urban centers, but notably areas that were contested races for the presidential election. These states consist of the Rust Belt region (Wisconsin, Michigan, Pennsylvania) and states who were considered swing states (Arizona, Texas, Florida).

Despite seeing high concentration overall in major metropolitans in the United States (California, Texas, New York), these three hashtags provided some variance of tweet distribution in respective areas. We see that places that seen initial activity either stay prominent or spreads to a nearby location. But overall, we see a chain reaction trend of tweets being spread within proximity and it makes logical sense in respect to the real-world events that create these hashtags.

6. Conclusion

This capstone is consisted of research into the traditional of media power to online communities and how they influence the flow of news information. This is by answering the questions of how well alternative media creates echo chambers to shape their audience's point of view, to if the origin of online news indicates the type of audiences they reach based on proximity. This research sets out to answer those questions by two types of analyses: social network and heat map analysis. Using the network analysis, it will help support if a viral hashtag has the influence to garner many users within its network. On top of that, if those users have some sort of connectivity based on mutual Twitter interactions of retweets, replies, user mentions, and media shared with that given hashtag. For the map analysis, it was more to see if a specific hashtag can distribute out to audiences outside the realm of where it originated. This would be in the form of concentration within proximity of the origin or in significant areas around the country.

From looking at the results, I view that the social network analysis provided ample results to give a sufficient answer. Just from the discovery of the central node of these so-called "Network Influencers", it is clear to see that these middle-level gatekeepers have a prevalent influence within their base. Within their respective network, we saw there was a significant amount of interaction within the central hub of the network graph that came out to be a lot of connectivity between the users. This really shows the impact of how alternative mediums for news like on Twitter, can be influential on the flow of information online. Also, with how just one tweet can skew the opinions and views of an event to like-minded users.

In regard to the map analysis, I believe the time-lapsed heat map was a good starting point to visualize the spread of a tweet geographically. Especially with utilizing the timestamps of the scrapped tweets, it was used well to animate the progression of a hashtag throughout time. But in terms of answering if this solution answers how well the origin of a news events reaches an audience based on proximity, the visualization can slightly answer that question. I think that the use of a heat map wouldn't garnered the best results compared to a network analysis but on a cartographic scale. First pinpointing a tweet origin and seeing the spread of it through mutual Twitter interactions through edges and nodes, would have been more sufficient and provide a clearer answer. While the animated map was intriguing to view to see how tweet count varied day by day in certain regions, it would have been clearer if the concentration on the map was cumulative over time to see the overall density of tweets for the time period. Especially with the small sample size of geotagged tweets, that doesn't do the map analysis justice to answer the question with best confidence.

In terms of what is next for this project, it is mainly to continue what was ideally planned for this capstone. Much of the ideals had to be limited due to time constraints and the overall time consumption throughout this process to get the visualizations properly. First enhancement would be taking a bigger sample size of tweets and possibly adding more hashtags, some that could be opposing ideas from the hashtags presented in this project. Also, the network graphs could be woven together more properly to properly see the distance between nodes by the actual connectivity between them. But I think a lot more work can go into the map portion of this analysis. As stated before, having a network analysis implemented on a geographic level would have answered one of the questions clearer. While the time-lapse heat map was a good starting point, having it build over time to show cumulative growth would also be better analysis.

References

Ahmed S; Cho J; Jaidka K, 2019, Framing social conflicts in news coverage and social media: A multicountry comparative study. SAGE Publications, *The International Communication Gazette*, London, England, 81(4): 346-371.

Althaus S; Cizmar A; Gimpel J, 2009, Media Supply, Audience Demand, and the Geography of News Consumption in the United States. Taylor & Francis Group, *Political Communication*, Philadelphia, USA, 26(3): 249-277.

Ardèvol-Abreu A; Gil de Zúñiga H, 2016, Effects of Editorial Media Bias Perception and Media Trust on the Use of Traditional, Citizen, and Social Media News. SAGE Publications, *Journalism & Mass Communication Quarterly*, Los Angeles, USA, 94(3): 703-724.

Bastos M; Mercea D; Baronchelli A, 2018, The geographic embedding of online echo chambers: Evidence from the Brexit campaign. Public Library of Science, *Public Library of Science ONE*, San Francisco, USA, 13(11): 1-16.

Broniatowski D; Jamison A; Qi S; AlKulaib L; Chen T; Benton A; Quinn S; Dredze M, 2018, Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. American Public Health Association, *American Journal of Public Health*, Washington D.C., USA, 108(10): 1378-1384.

Brugnoli E; Cinelli M; Quattrociocchi W; Scala A, 2019, Recursive patterns in online echo chambers. Nature Publishing Group, *Scientific Reports*, England, UK, 9(1): 1-18.

Bruns A; Woodford D, 2014, Identifying the Events That Connect Social Media Users: Charting Follower Accession on Twitter. SAGE Publications, *SAGE Research Methods Cases*, London, England, 1-13.

Crampton J; Graham M; Poorthuis A; Shelton T; Stephens M; Wilson M; Zook M, 2013, Beyond the geotag, situating 'big data' and leveraging the potential of the geoweb. Taylor & Francis Group, *Cartography and Geographic Information Science*, Philadelphia, USA, 40(2): 130-139.

Dodge M, 2018, Mapping II: News media mapping, new mediated geovisualities, mapping and verticality. SAGE Publications, *Progress in Human Geography*, London, England, 42(6): 949-958.

Eberl J; Boomgaarden H; Wagner M, 2015, One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences. SAGE Publications, *Communication Research*, Los Angeles, USA, 44(8): 1125-1148.

Elejalde E; Ferres L; Schifanella R, 2019, Understanding news outlets' audience-targeting patterns. Springer, *EPJ Data Science*, Heidelberg, Germany, 8(1): 1-20.

Fitzpatrick N, 2018, Media Manipulation 2.0: The Impact of Social Media on News, Competition, and Accuracy. Athens Institute for Education and Research, *Athens Journal of Mass Media and Communications*, Athens, Greece, 4(1): 45-62.

George L; Waldfogel J, 2008, National Media and Local Political Participation: The Case of the New York Times. In: World Bank Group, *Information and Public Choice: From Media Markets to Policy Making*, Washington D.C., USA, 40-54.

Gil de Zúñiga H; Weeks B; Ardèvol-Abreu A, 2017, Effects of the News-Finds-Me Perception in Communication: Social Media Use Implications for News Seeking and Learning About Politics. Wiley-Blackwell, *Journal of Computer-Mediated Communication*, Oxford, England, 22(3): 105-123.

Graham T; Bruns A; Angus D; Hurcombe E; Hames S, 2021, #IStandWithDan versus #DictatorDan: the polarised dynamics of Twitter discussions about Victoria's COVID-19 restrictions. SAGE Publications, *Media International Australia*, London, England, 179(1): 127-148.

Hemsley J, 2019, Followers Retweet! The Influence of Middle-Level Gatekeepers on the Spread of Political Information on Twitter. Wiley, *Policy and Internet*, Berkeley, USA, 11(3): 280-304.

Hoffman L, 2013, When the World Outside Gets Inside Your Head: The Effects of Media Context on Perceptions of Public Opinion. SAGE Publications, *Communication Research*, Los Angeles, USA, 40(4): 463-485.

Kapoor A, 2018, *Social Network Analysis in Python*. DataCamp, Delhi, India.

Kim E; Ihm J, 2020, More Than Virality: Online Sharing of Controversial News With Activated Audience. SAGE Publications, *Journalism & Mass Communication Quarterly*, Los Angeles, USA, 97(1): 118-140.

Kim S, 2015, Attracting Views and Going Viral: How Message Features and News-Sharing Channels Affect Health News Diffusion. Wiley, *Journal of Communication*, Hoboken, USA, 65(3): 512-534.

Lim J; Liu Z; Zhou L, 2019, Detection of Fraudulent Tweets: An Empirical Investigation Using Network Analysis and Deep Learning Technique. Institute of Electrical and Electronics Engineers, *International Conference on Intelligence and Security Informatics*, Shenzhen, China, 1-3.

Lungu C, 2019, *Exploring Twitter data using Python: Analyzing the data*. Medium, Edinburgh, Scotland.

Mersey R, 2009, ONLINE NEWS USERS' SENSE OF COMMUNITY: Is geography dead? Taylor & Francis Group, *Journalism Practice*, Philadelphia, USA, 3(3): 347-360.

Owen A; Wei A, 2020, Inequality and Bias in the Demand for and Supply of News. Wiley-Blackwell, *Social Science Quarterly*, Oxford, England, 101(1): 91-106.

Shelton T, 2017, Spatialities of data: mapping social media 'beyond the geotag'. Springer, *GeoJournal*, Dordrecht, Netherlands, 82(4): 721-734.

Tomeny T; Vargo C; El-Toukhy S, 2017, Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. Elsevier, *Social Science & Medicine*, Oxford, England, 191: 168-175.

Tomlins S, 2020, *Scraping Tweets by Location in Python using snscrape*. Medium, Bristol, England.

Vivek S, 2021, *Twitter Location Analysis*. Towards Data Science, Atlanta, USA.

Zollo F; Novak P; Del Vicario M; Bessi A; Mozetic I; Scala A; Caldarelli G; Quattrociocchi W, 2015, Emotional Dynamics in the Age of Misinformation. Public Library of Science, *Public Library of Science ONE*, San Francisco, USA, 10(9): 1-22.

Zubiaga A; Aker A; Bontcheva K; Liakata M; Procter R, 2018, Detection and Resolution of Rumours in Social Media: A Survey. Association for Computing Machinery, *Association of Computing Machinery Computing Surveys*, New York, USA, 51(2): 1-36.