

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

## **Final Proposal**

### **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 presidential election have garnered various attitudes. Previous studies about the accuracy of stories have mainly focused on the relationship between an individual's political leaning and the news platform they consume (Eberl et al., 2015; Hoffman, 2013). Periodically, emerging alternative news sources online have competed with traditional news markets over who is subjectively more truthful and has less perceived bias. This opened the doors for online communities made up of like-minded individuals, that are now connected beyond the bounds of proximity (Ardèvol-Abreu & Gil de Zúñiga, 2016).

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

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

### Introduction

The purpose of this project is to discover the representation of individuals an online news source attracts, based on where those people reside. This will incorporate the relationship of how one chooses their news sources and how that ultimately shapes their views and opinions of current events. With the integration of news within social media, the opinion of the public is more scattered and less tied to geographical regions. An explanation to this is that certain areas tend to have audiences who see more individualistic news sources outside their local market realm (Mersey, 2009). This brings into question if the audiences of online news sources tend to come from users within a close proximity, which is the traditional means of older media such television or newspaper.

With the growing reliance of news on social media platforms, more consumers tend to explore beyond the coverage of their local markets and look for new takes of journalism that fits more to their individual tastes. Knowing the dynamics beyond this phenomenon is crucial to the growing issue of misinformation that certain audiences could be more susceptible to, due to the variance of news perspective (Ahmed et al., 2019). This project aims to make aware of where individuals tend to be more misinformed or indifferent to opposing viewpoints of current events in our country. Ultimately, this will possibly bring more conscientiousness to seek out fresh perspectives to develop a more informed society. This can be beneficial to organizations or activists that are looking to help their local communities to be a more informed news consumer. This may be in the form of bringing attention to media literacy to groups that may be more enticed by misinformation and echo chambers.

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

### Literature Review

In the past, traditional media in the forms of radio, newspaper, and cable television usually reached either a regional or national audience. But with the internet becoming more accessible, news stories carried in there have widened the public expectations of how the press should be. Standards such as trustworthiness, instant, and spectacular news coverage have 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 for news spreading on social media, data about the users who partake within these communities are made widely available. Studies have analyzed the connections of individuals sharing similar opinions and how it ties in spatially (Bastos et al., 2018). This was through the use of Twitter activity stemming from tweeted coverage of an event or issue that can circle more widely nationwide than it could if it came from traditional media.

### **Section 1 - Transitioning from Traditional Mediums to Online**

Starting from the recent past, traditional news markets have attempted to adapt with growing online news 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 accompany a headline and/or an image that encapsulates the story to fit the narrative that the source so

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

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 being victims 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 mainstream media’s perception to be 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 sources have to the market of distributing news. Also, to how the spatial ties of local markets has now become less geographically specified due to the shift to online communities. This brings into question about research regarding if the origin of the news source matters to what audience they attract. Psychological studies suggest that individuals tend to consume the local news markets if they feel the “neighborhood effect” of being influenced by the people they live around (Althaus et al., 2009; Mersey, 2009). There seems to be a negative correlation in more populated and dense areas, where local television and talk radio news are less consumed with online news being widely available (Althaus, 2009). In the realm of newspapers, a carrier like The New York Times has expanded out of their local market by focusing on more national news topics that an incoming audience would be more keen to. This has led to more negligence of issues at a local level, resulting in more individuals being less informed and focused on local policies and political candidates (George & Waldfogel, 2008). That shows how traditional media in many

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

ways have been indirectly influenced by alternative online media that has become more advanced in adopting a diverse set of content.

## **Section 2 - Emergence of Alternative News Sources Online**

These alternative news sources have bypassed the “gatekeeping” of information originally controlled by traditional outlets. But now, online communities are created to make an accessible and welcoming community to like-minded news consumers. Rather than the motive of maintaining a reputation kept by traditional sources (Ahmed, 2019), these online platforms have more incentive in “ideological inclinations” (Elejade et al., 2019). The mechanisms of liking and following particular 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 its 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 have been detrimental for their market. What has caused individuals to accuse mainstream media to be biased, has now allowed room for echo chambers of information online.

This shows how echo chambers have gathered audiences from various places throughout the country, compared to how local news markets have mainly targeted their regional area. This confronts the challenge to figuring out what regional audiences are more likely to explore out from traditional media to venture into online environments. One article briefly discussed upon

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

this as 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). To be able to visualize how well like-minded online communities correlate to real-world proximity, work done using Twitter data has been a prominent method.

### **Section 3 - Spatial Dimensions of Online Media Communities**

Since the activity on Twitter makes up a fraction of internet users that represents only “about one-third of [the] global population” (Crampton et al., 2013), that means it 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 provided quite an abundance of information, beyond just location coordinates. It also comes with a timestamp, retweet and likes data that can be involved in social network analyses. In regards to proximity, “cultural, economic, political, and social histories” is an aspect that supports the idea that online communities can be geographically tied (Bastos et al., 2018; Crampton et al., 2013). 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; while possibly targeting people in power or people who have more of a following (Graham, 2020).

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

A recurring issue with quantifying the influence of Twitter personalities in real world beliefs, is the abundance of unauthentic fringe accounts and also manipulation by using bot accounts. This has been a common theme in almost all the research regarding network analysis (Crampton et al., 2013; Hemsley, 2019; Graham 2020; Tomeny et al., 2017). But regardless of that, some case studies have provided great insight into how bot accounts were handled (Bastos et al, 2018) or how this effect is negligible to the study's results (Hemsley, 2019). Each article provides strong evidence that there is some connection of echo chamber communities and their proximity with other members. This is mostly likely to do with the literature showcasing tweeted events that occurred within a short period of time or within a concentrated area.

With my research, I will look into 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 my question of what audiences are most likely to be susceptible to misinformation based on what is shared within the online community. This brings Twitter into this, to test out if there is a spatial relationship between the users of this audience. 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 chambers sharing information online.

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

### Methods

The technical component of my project is composed of multiple data visualizations to show how echo chambers are spatially represented in the United States. This process is planned to be implemented in three phases. First is to collect data from Twitter, then compose that data for network analysis, and to finally execute it into a spatial representation. Extracting geotagged data from the tweets will tie into answering if such echo chambers are made up of individuals residing in certain areas. My hypothesis is that the reinforcement of these fringe ideas will usually come from suburban areas around major metropolitans. This experiment will also support the idea that alternative news media have a proximity connection, similar to traditional regional news markets.

To start off, data needs to be collected through Twitter. Through a Python working environment, libraries will be utilized to access the streaming functionalities of the Twitter API and to geocode such tweets for the spatial visualization. Tweepy will be the library used to work with the API and Geopy for the spatial elements of the tweets. There is an abundance of information that can be extracted from tweets, but what is mainly needed is timestamps, location, and information about likes and/or retweets. Using the API's filtered stream will help a collection of real-time tweets during the peak activity of an event. This is where rules can be set to match tweet attributes that I'm looking into, such as hashtags.

At least two viral hashtags will be used to filter and collect the needed tweets, utilized as the keyword. The choices of hashtags will whatever stemmed from last year's major events that have created an isolated audience or audiences to believe such ideas or theories. This will be



Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

inspired from what phrases has snowballed to be a defining perspective from the controversial events. From there, temporal configurations will be made to filter down the tweets during the duration of the respective events of that hashtag, usually spanning months. The starting point of the timeline will start from the date of the real world event or the anticipation leading up to it, to when the hot topic starts fading out. The location of said tweets will be located by the user's location stated in their profile. While the location of a Twitter user can essentially be set to anything and is not verifiable, "...approximately 60% of all tweets can ultimately be associated with a physical location with some degree of confidence" (Crampton et al., 2013). Also, elements such as detecting bot or spam accounts will need to be considered, which may affect the results especially with more controversial topics or hashtags (Bastos et al., 2018).

Tweets would be further filtered down within a bounded area of the contiguous United States (possibly from an existing shapefile with array of coordinates), so the latter map analysis would avoid any geographic outliers. Depending on how much geotagged tweets can be collected within a certain amount of time, a timer for running the script can be implemented. This can also prevent unresponsiveness from the collector script, possible from a processing overload or an excessive amount of storage built up from many tweets. The goal is to collect around the ballpark of five to seven thousand geotagged tweets per hashtag.

Once the tweets are collected, they will be used for visualizations with a non-spatial social network analysis being next. 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 can be in the lines of retweeting or liking the same tweet or mutually

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

following someone. This analysis will support answering if there's any correlation between the sharing of one peculiar idea from a source would lead individuals to become part of the same echo chamber. It will prove whether or not an alternative source has the power to influence an echo chamber and also do so spatially. The distance between these connection nodes will be tightly woven together if that common media activity is shown.

The last step is to derive that network connection into a cartographic visualization. Still within the Python environment, the libraries of ArcGIS API, GeoPandas, or Arcpy can help towards mapping the tweets. Having the boundary set within the lower 48 contiguous states like mentioned before, will allow the nodes to maintain their interconnectedness. With the tweets applied with geotagged data onto a map, a scattered relation of points can be spatially tied together by looking at the possible patterns of what areas these points are coming from. One way to do this is to choose one particular viral tweet as the origin and have its retweets or like string out from that point. This would be useful if a hashtag has developed from a particular place based on a real life event.

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 liking, retweeting or have mutually following others; 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 that is confirmed for certain hashtags, then visualizing it on a map could indicate that there is a common spatial pattern of a like-minded online community.

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

### Summary

This project is made up of research into the transition of media power to online communities and how they influence the flow of news information. Then, I connect how previously traditional news markets are tied to a region based on their serviced area and how that can also be applied to these online communities. Through mentioning how these groups can be susceptible to perceived biases, it can lead to echo chambers that are built on misinformed or misleading news stories. To be able to visualize this phenomenon spatially, I first did a network analysis in gathered tweets that soon later will be applied to a cartographic representation. This is to show that a scattered internet environment can have individuals who are closely connected by location through their shared beliefs and positions of hot topics. This is to bring awareness to possible groups that are more likely to fall victim to misinformation within echo chambers and how local organizations can get involved with media literacy. But overall to the general public, this project aims to bring more consciousness to discovering different perspectives to controversial ideas and events by seeking out new sources to challenge one's ideas.

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

Timeline

Task	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Sign up for Twitter Dev Account	X							
Research into Twitter libraries		X						
Obtain API Keys and Tokens		X						
Choose hashtags and time frames			X					
Test initial data collector script				X				
Run collector to gather final tweets					X			
Read into NetworkX library docs					X	X		
Write script to do network analysis						X		
Read in into Python geo libraries							X	
Apply network analysis onto maps							X	
Adjust and finalize visual maps								X

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

## References

- Ahmed, Saifuddin; Cho, Jaeho; Jaidka, Kokil. 2019. "Framing social conflicts in news coverage and social media: A multicountry comparative study." *The International Communication Gazette*, 81 (4): 346-371. London, England: SAGE Publications.
- Althaus, Scott L.; Cizmar, Anne M.; Gimpel, James G. 2009. "Media Supply, Audience Demand, and the Geography of News Consumption in the United States." *Political Communication*, 26 (3): 249-277. Informa UK Limited.
- Ardèvol-Abreu, Alberto; Gil de Zúñiga, Homero. 2016. "Effects of Editorial Media Bias Perception and Media Trust on the Use of Traditional, Citizen, and Social Media News." *Journalism & Mass Communication Quarterly*, 94 (3): 703-724. SAGE Publications.
- Bastos, Marco; Mercea, Dan; Baronchelli, Andrea. 2018. "The geographic embedding of online echo chambers: Evidence from the Brexit campaign." *PLoS ONE*, 13 (11): e0206841. United States: Public Library of Science (PLOS).
- Crampton, Jeremy W.; Graham, Mark; Poorthuis, Ate; Shelton, Taylor; Stephens, Monica; Wilson, Matthew W.; Zook, Matthew. 2013. "Beyond the geotag, situating 'big data' and leveraging the potential of the geoweb." *Cartography and Geographic Information Science*, 40 (2): 130-139. Taylor & Francis.
- Eberl, Jakob-Mortiz; Boomgaarden, Hajo G.; Wagner, Markus. 2015. "One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences." *Communication Research*, 44 (8): 1125-1148. SAGE Publications.

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

- Elejalde, Erick; Ferres, Leo; Schifanella, Rossano. 2019. "Understanding news outlets' audience-targeting patterns." *EPJ Data Science*, 8 (1): 1-20. Heidelberg: Springer Science and Business Media LLC.
- Fitzpatrick, Neil. 2018. "Media Manipulation 2.0: The Impact of Social Media on News, Competition, and Accuracy." *Athens Journal of Mass Media and Communications*, 4 (1): 45-62. National Library of Greece.
- George, Lisa M.; Waldfogel, Joel. 2008. "National Media and Local Political Participation: The Case of the New York Times." In *Information and Public Choice: From Media Markets to Policy Making*, edited by Roumeen Islam, 40-54. World Bank.
- Graham, Timothy; Bruns, Axel; Angus, Daniel; Hurcombe, Edward; Hames, Sam. 2020. "#IStandWithDan versus #DictatorDan: the polarised dynamics of Twitter discussions about Victoria's COVID-19 restrictions." *Media International Australia*, 0 (0): 1-22. SAGE Publications.
- Hemsley, Jeff. 2019. "Followers Retweet! The Influence of Middle-Level Gatekeepers on the Spread of Political Information on Twitter." *Policy and Internet*, 11 (3): 280-304. Wiley.
- Hoffman, Lindsay H. 2013. "When the World Outside Gets Inside Your Head: The Effects of Media Context on Perceptions of Public Opinion." *Communication Research*, 40 (4): 463-485. SAGE Publications.

Tom Truong

3/15/21

Instructor Jim Thatcher

T GIS 507

- Kim, Eun-mee; Ihm, Jennifer. 2019. "More Than Virality: Online Sharing of Controversial News With Activated Audience." *Journalism & Mass Communication Quarterly*, 97 (1): 118-140. SAGE Publications.
- Mersey, Rachel D. 2009. "ONLINE NEWS USERS' SENSE OF COMMUNITY." *Journalism Practice*, 3 (3): 347-360. Taylor & Francis Group.
- Tomeny, Theodore S.; Vargo, Christopher J.; El-Toukhy, Sherine. 2017. "Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15." *Social Science & Medicine*, 191: 168-175. Elsevier BV.