Social media perceptions of firearms and U.S. mass shootings: A comparative analysis

matthew d'anna
Department of Criminology
George Mason University
Fairfax, VA 22030
mdanna2@masonlive.gmu.edu

Word count: 5,634

ABSTRACT

The rise in the United States of social media usage and uptick in mass shootings creates an opportunity to use one to study the other. Using the extracted, normalized, and cleaned bio locations of Twitter followers of prominent firearm-related accounts, @NRA and @Everytown, this study seeks to identify a relationship between the locations of pro-firearm supporters and pro-firearm control supporters to the locations of mass shootings. Using kernel density estimation, spatially joined Thiessen polygons, and nearest distance measures, this study finds that the supporters of pro-firearm control live closer to mass shootings events than the supporters of pro-firearm rights. The @NRA network is more densely clustered in non-mass shooting hotspots, while the @Everytown network is more dispersed across mass shooting hotspots. Thus, the geolocated followers of @NRA are located significantly farther away from mass shooting events than the geolocated followers of @Everytown. Future research considerations and limitations of the current study are also discussed.

Keywords

mass shootings, twitter, hotspots, firearms, social network analysis

1. INTRODUCTION

An anecdotal comparison of U.S. mass shootings and social media reveals a startling correlation: since the early 2000s, both have become exceedingly more common. From 1982-2004, there was an annual average of 1.75 mass shootings in the United States. From 2005-2016, that figure has jumped to 3.75 annually [15]. Coincidentally, the two most popular social media platforms in the U.S. – Facebook and Twitter – bookend that statistically breakpoint; Facebook was created in 2004, Twitter in 2006. While this comparison is purely illustrative, the similar rise of mass shootings and social media makes it somewhat logical to use one (social media) to analyze the other (mass shootings). Thus, the motivation for this work is to examine the spatial relationship between public views on firearms and mass shootings. Specifically, the intent is to investigate and identify location-based correlations among individuals supporting pro-firearm organizations and individuals supporting pro-firearm control organizations, with mass shooting events. This study seeks to potentially identify useful patterns of public opinion on a critical national matter. As the number of mass casualty incidents have increased at an alarming rate since the late 1990s, firearm legislation has become a hotly contested issue in local, state, and federal politics. Further, applying spatial analysis of social media data has the potential to lend otherwise unknown insights into the problem. Social media data is often correlated to current events in time, and not space. Given the current U.S. political climate (a Presidential election year), public perceptions of fear (growing concerns of public safety), and substantial increases in social media usage (65% of U.S. adults use some form of social media [11]), this study hopes to provide relevant and timely insights.

2. CURRENT STUDY

2.1 Research Question

This study seeks to determine the extent of the relationship between views on firearm control and the locations of mass shootings. Thus, there is an interest in identifying any correlations between U.S. mass shootings and pro-firearm or pro-firearm control views. Specifically, there are two research questions. First, what is the spatial relationship between the supporters of pro-firearm policies and the supporters of pro-firearm control policies, and U.S. mass shootings? Second, do mass shootings occur closer to clusters of pro-firearm support, or pro-firearm control support?

2.2 Hypothesis

The hypothesis of this study is that areas immediately surrounding mass shooting locations will cluster with pro-firearm supporters. It is hypothesized that these views will experience a distance decay; as the distance from a mass shooting location increases, there will be a decrease in pro-firearm attitudes. Thus, pro-firearm control supporters are expected to cluster farther from mass shootings. The logic supporting this hypothesis is that pro-firearm supporters enable an environment more supportive of increased firearm availability. Such environments have lower legal restrictions to firearm purchase, carry, and concealment. Hypothetically, greater access to firearms makes usage more likely, and increases the probability of a mass shooting event.

3. LITERATURE REVIEW

Given the multi-faceted nature of this research question, a review of the relevant literature on this topic is best grouped into three categories: research on spatial crime analysis, research on spatial network analysis, and research combining both.

3.1 Spatial Crime Analysis

Among spatial crime analysis literature, there are three studies that particularly relate to this topic. Three studies are specific to spatial crime analysis methods, and two studies address the spatial aspects of firearm legislation and prevention. The research on spatial methods serves as a standard for conducting sound analysis. First, Brantingham and Brantingham [2] developed location quotients (LQs) in order to address questions of the relative structure and importance of local economies, while conceptually avoiding some of the stationarity problems of spatial analysis. An area can be compared to its surroundings, and weighted for specific crimes relative to total crime and population. This becomes a more sophisticated mechanism for hot spot identification.

Second, Rogerson and Sun [13] applied temporal change detection of spatial point patterns using nearest neighbor statistics and cumulative sums to crime data in Buffalo, New York. The result was a new approach for identifying the characteristics of newly emerging clusters of crime. Using arson crimes in 1996, Rogerson and Sun [13] demonstrated how measuring mean nearest neighbors over time can quickly detect new hotspots, as well as deviations from the spatial norm.

Third, Reich and Porter [12] developed a model for spatiotemporal clustering to identify potential suspects for unsolved crimes. Using burglary crimes from 2009-2010 in Baltimore, Maryland, Reich and Porter [12] cluster similar events over space and time. This process results in identifying the most likely crimes to be associated with a solved crime. Reich and Porter [12] is relevant to the current study in the way it handles multivariate data to create hotspots.

Further, research on spatial components of firearm legislature and prevention demonstrates the significance of community perceptions and support. Isenstein [5] examined correlations of state laws to rates of shooting deaths. States with the most restrictions on gun users also have the lowest rates of gun-related deaths. Comparing states with Stand Your Ground (prosecutorial immunity from self-defense with deadly force) versus Duty to Retreat (attempt to flee prior to using deadly force) laws indicated a 1.52 lower homicide rate among Duty to Retreat states. Further, Isenstein [5] found states that conduct extra background checks have homicides rates .97 lower than states that do not.

Second, Koper et al [8] evaluated a community-wide initiative in St. Louis, Missouri to reduce gun violence. The study used statistically rigorous methods to examine the impact on UCR Part I violent crimes during a nine month program in nonequivalent comparison areas based on social characteristics and preprogram violence trends for pre- and post- testing. Koper et al [8] found significant reductions in total violence and gun violence in the targeted areas, noting that “initiatives targeted on high-risk places, behaviors, and actors are effective, particularly when conducted in the context of multiagency problem-solving efforts” (p. 3).

In total, these five studies demonstrate both the power of sophisticated spatial analysis on crime data, and the degree to which location-based analysis is essential to understanding the dynamics of firearm attitudes and activity.

3.2 Spatial Network Analysis

There are three noteworthy studies that address spatial social media visualization. First, Jain [6] discusses the value in still using K-means clusters, more than 50 years after its discovery. Given the unique goal of discovering natural groupings of unstructured objects, clustering is a complicated process. Jain [6] cautions that clustering should be used as an exploratory tool, in order to develop hypotheses. This is an important consideration in the current study: clustering social media users and comparing cluster locations does not provide an answer; rather, it frames a question.

Second, Luo and MacEachren [9] argues for a more coherent interpretation of geographic and network spaces. Often analyzed separately, fusing the two can offer valuable insight into a dataset. Distance, for example, is treated differently by each discipline. However, there is an interaction in geographic and network space that should be understood. Luo and MacEachren [9] offer visual analytics as the solution, which involves data exploration, visual analytics, and predictive analysis. This is essential to the current study; understanding the geographic footprint of a network can identify individuals close in the network *and* physical close. Further, grouping by network enhances geographic cluster detection.

Third, Croitoru et al [3] describes geosocial analysis as a hybrid of spatial and social analysis. The inherent uniqueness of social media data demands a new approach to spatial analysis. Similar to Luo and MacEachren [9], Croitoru et al [3] suggest mapping network connectivity in geographic space, rather than abstract network space. This can allow for deeper analytic dynamics to emerge, enabling richer findings. Croitoru et al [3] serves as a call to creativity, given the complexity and opportunity of geotagged social media data.

In sum, these three studies provide a sample of spatial analytic techniques to apply to social media data. In each, there is a need to apply methods that are flexible and responsive to the underlying data.

3.3 Spatial Analysis of Crime and Social Media

There are two relevant research studies that highlight combined spatial analysis of crime and social media data. First, Gerber [4] used linguistic analysis and statistical topic modeling in a crime prediction model. In 19 of 25 crime types examined, the addition of Twitter data improved predictive performance. By using these methods to identify discussion topics among geotagged tweets, surveillance areas for specific crime types (notably violent person crimes) can be improved.

Second, Signorini et al [14] used Twitter data to track, measure, and forecast the spread of the H1N1 virus. Signorini et al [14] demonstrated that “Twitter chatter” offers unique public insights that other, more traditional data sources, cannot. Specifically, the velocity and volume of real-time data enables faster, more robust insights.

3.4 Application

The current research relates to aspects of these previous studies in multiple ways. First, it expands on concepts introduced by Signorini et al [14], by examining the power of spatial correlation of social media data to current events and social problems. Second, an extension of Rogerson and Sun [13] can be applied to mass shooting events. In particular, by calculated nearest neighbor statistics in conjunction hotspot density analysis, clusters of pro-firearm and pro-firearm control individuals can be statistically compared to mass shooting locations. Third, similar to Gerber [4], augmenting mass shooting events with social media data may help determine predictive characteristics for future occurrences. While topical content is beyond the scope of this study, location-based analytics can help identify neighborhood characteristics similar across the U.S., and may identify other at risk for mass shooting activity.

4. METHODOLOGY

This study follows a four-step process for identifying, collecting, cleaning, and analyzing data.

4.1 Mass shooting data

First, a dataset for mass shooting events must be acquired. “Mass shootings,” for the purpose of this study, are defined as a firearm incident involving four or more casualties previously unknown to the offender. Once obtained, data will be geocoded and mapped as points at the address level. Given the distinct location information of each event, this should be possible. In the occasion of a multi-location incident, an event will be mapped at the centroid of all locations involved. Further, if a mass shooting has vague spatial reference information, it will be mapped at the locale (city/county) center; given the public notoriety of each event, this is highly unlikely.

4.2 Social media "seeds"

Second, the appropriate social media platform must be identified. Given the plethora of options, the best choice should enable communication among users without mandating reciprocation. Thus, the number of followers for a user can serve as a proxy measure of message broadcasting reach. In this model, it is important to note that data interpretation implies that a follower is a supporter of an account. While this is not always the case, this study assumes that the recipient of a message is, at a minimum, an interested listener. Additionally, the appropriate social media platform should have an application program interface (API) accessible to common computer coding languages, in order to acquire samples of data.

After choosing a platform, key profiles – "seeds" – should be identified. These seed accounts should represent each view on firearm control; one pro-firearm, and one pro-firearm control. While each seed does not need to be the authority on the issue, it should at minimum be a major stakeholder on the subject.

Part of this process of choosing the appropriate seeds is identifying key similarities and differences. For similarities, meta-data should be extracted from each account, in order to conduct basic statistical analyses. These seed accounts should have basic usage commonalities. As for differences, a sample of the social media content of each seed account should be extracted and analyzed. Thus, the two accounts should be communicating different views on the same topic at similar rates.

Finally, once seeds have been identified and statistically and contextually compared, a sample of followers of each account should be extracted. Applying the principles of social network analysis and graph theory [1], the amount of common associates should be used to determine distinctly different communities.

4.3 Follower locations

The sample sets of followers of each seed previously extracted should be used to derive location information. Specifically, the location listed by the user in their bio/home page. This data is provided manually by the user. This is different than the location information found in the metadata of a social media post, which is derived from global positioning system (GPS) coordinates from the electronic device. By using the bio location, this study is subject to more manual data normalization and geocoding; however, it is hypothesized that this location information is more readily available than geotagged metadata, which is rarely enabled by a user [10], [16]. Once extracted, this data is cleaned and geocoded to the locale (city/county) level. For this study, only U.S. locations are derived and analyzed; otherwise international distances to U.S. mass shootings would likely skew the findings. The result is two point files, one set for each seed's followers.

4.4 Spatial analysis

Once compiled, robust spatial analysis is conducted across the three datasets. This includes multiple clustering techniques, distance calculations, and polygon generation. Each method compares the locations of social media users to locations of mass shootings. The combination of tests should reveal the strongest spatial relationships.

5. DATA AND PROCESSING

The data used and the processing applied in this study is best understood in the same four steps as the methodology.

5.1 Mass shootings

U.S. mass shooting data was obtained from the news and political website *Mother Jones* [15]. Researchers for the website maintain a database of mass shootings events since 1982. The website uses the same definition as this study for mass shootings (firearm event with at least four previously unknown victims). Each event has information on location, date, suspect, casualties, and other details; for this study only location was relevant. The location fields included locale, state, latitude, longitude. Fewer than ten percent of all records needed to be cleaned and manually geocoded. When necessary, this process was conducted with Google Maps. This study captured events until February 2016, analyzing 80 total U.S. mass shootings. *Mother Jones* makes the data publicly available through a variety of formats, including a comma separated file. Data manipulation and analysis was conducted in Microsoft Excel and ESRI ArcGIS.

5.2 Social media "seeds"

Twitter was chosen as the ideal social media platform. As identified by Kietzmann et al [7], the nature of social media conversations, specifically the rate, direction, and flow of content dictate the push-pull dynamics of engagement. Twitter's method of following and followers, which enables users to follow others without reciprocation, enables the ideal messaging for this study. Additionally, access to the free API for users enabled code to extract relevant data.

After choosing Twitter, seed users were identified. The account associated with the National Rifle Association (@NRA) was selected as the pro-firearm seed, and the account associated with Everytown for Gun Safety (@Everytown) was selected as the pro-firearm control seed. Both users have significant presence and reputation, as described by Kietzmann et al [7], as highly symbolic, polarizing, representations of the U.S. firearm debate.

Next, the data generated to derive significant similarities and differences between @NRA and @Everytown came from two sources. First, python code was written to extract the user profile metadata for each account. This included counts for the number of followers, following, favorites, lists, and tweets. The code ("md_user_data.py," provided as an addendum to this paper) output a comma separated file, statistically compared in Microsoft Excel. Second, python code was written to extract the 3,200 most recent tweets from each account. The code ("md_user_tweets.py," provided as an addendum to this paper) output a comma separated file for each account, where each row represented a tweet. This was used to create a visualization.

Further, social network analysis was conducted on each accounts' followers. Sample sets of followers were generated with python code. The code ("md_list_followers.py," provided as an addendum to this paper) output a list of accounts. These lists were manipulated in Microsoft Excel into relational tables, and analyzed in Gephi.

Samples were used in lieu of entire follower lists. Using the free API service to extract data limits the ability to conduct large queries over time. Even with adequate time lags, the connection can be broken eventually. However, given the findings of Morstatter et al [10] and Valkanas et al [16], the veracity of samples have shown to be nearly identical to full datasets. Given the amount of time and cost to obtain the full follower lists, samples were preferred.

5.3 Follower locations

As previously mentioned, the location data associated with sample followers of @NRA and @Everytown was derived from the bio of each user. Python code extracted the raw location data into a comma separated file. The location field was extensively cleaned, normalized, and geocoded at the locale (city/county) level. Locations more precise than locale were adjusted, and locations less precise (e.g. state, country) were excluded. Geocoding was done by joining multiple tables of open source U.S. city and county latitude and longitude coordinates. This process involved constant manual data interventions; however, this was by design, since a similar process has not to be employed. After geocoding, mapping and analysis was conducted in ESRI ArcGIS.

5.4 Spatial analysis

Three different spatial analysis methods were applied to the three datasets (mass shootings, @Everytown followers, and @NRA followers). First, kernel densities were created in ESRI ArcGIS. The resultant hotspots for each dataset were compared, specifically to identify overlap and gaps. Overlap describes common hotspot areas; gaps are mass shooting hotspots that do not have a corresponding social media user hotspot.

Second, mass shooting points were transformed into Thiessen polygons in ESRI ArcGIS. Thiessen polygons are naturally self-organizing areas created by perpendicular bisectors of nearest neighbors. Points for each set of social media followers were spatially joined to their corresponding Thiessen polygon (the area they fell into). The raw difference in each polygon was determined for each area by subtracting the sum of @Everytown followers from the sum of @NRA followers. The result, a national map of color-coded irregular polygons, identified the size of "neighborhoods" surrounding a mass shooting, and as well as the dominant social media support.

Third, nearest distance was calculated in ESRI ArcGIS. Specifically, the distance from each social media user to its closest mass shooting was determined. These distances were grouped by @NRA and @Everytown followers, and a two-tailed t-Test for significant differences in the mean distances was determined ($p=0.05$). The null hypothesis was there was no significant difference between the mean distances of @NRA and @Everytown followers to a mass shooting.

6. RESULTS

To begin, a basic statistical comparison of @NRA and @Everytown highlights key similarities (Figure 1). For similarities, neither account is relatively new to Twitter; both were established at least five years ago.

	NRA	Everytown
Screen Name	NRA	Everytown
Date Created	2/25/2009	2/11/2011
Followers	342,549	66,157
Followers Collected	17,999 (5%)	6,000 (9%)
Potential Followers Located	7,919 (44%)	3,515 (59%)
U.S. Geocoded Followers	3,916 (49%)	2,354 (67%)
Following	493	8,458
Tweets	9,342	7,759
Favorites	121	29,721
Lists	3,130	975
Private Account	No	No
Location	Fairfax, VA	n/a

Figure 1. Statistical comparison of seed profiles.

Both accounts have a similar amount of tweets; @NRA averages 3.6 tweets per day, and @Everytown averages 4.1. Both accounts are "verified," which is a significant status symbol in Twitter; this indicates that Twitter vouches for the veracity of the account. Further, while the amount of followers may seem significantly different, in actuality it is not. Both accounts are in relatively rare ranks, compared to the average Twitter user. Thus, there is enough similar about these accounts to reasonably deem them appropriate to compare. This leads to further analysis, in three dimensions: content, network, and space.

6.1 Content

The content of each user was derived from the 3,200 most recent tweets on their timeline. Figure 2 highlights the key words from @Everytown. The shape of the word cloud interestingly organized into a heart; given the pro-firearm control, liberal stance of the organization, this seems fitting. Shape aside, the content is therapeutic and relatively caring in nature. There is a theme of safety and concern.

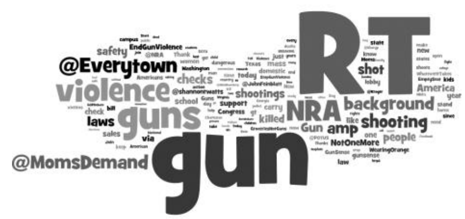

Figure 2. Word cloud of recent @Everytown tweets.

Conversely, as Figure 3 demonstrates, the content from @NRA is distinctly different. The word choices appear more blunt and direct. One of the largest terms is "2A," which is a popular NRA hashtag for the Second Amendment – a reference to the portion of the Constitution calling for the "right to bear arms." Further, the shape is telling: this word cloud appears to resemble a gun.

Figure 3. Word cloud of recent @NRA tweets.

Therefore, based on this broad content analysis, it is reasonable to assume the content delivered by these accounts offers different perspectives on the same issues.

6.2 Network

As far as their network connections are concerned, Figure 4 illustrates the common connections of both follower groups. In this network, the OutDegree for each node is either 0, 1, or 2; where 0 represents the seeds, 1 represents following only one seed, and 2 indicates following both seeds. The sample data consisted of 17,999 @NRA followers (5% of their overall network), and 6,000 @Everytown followers (9% of their overall network). Of these more than 23,000 total nodes, only 91 had an OutDegree of 2, meaning they are common to both accounts.

Those 91 users are further interesting in that their statistics differ greatly from the average @NRA and @Everytown follower. Followers of @NRA usually follow approximately 813 other users, and are followed by 1,206. On average, followers of @Everytown follow 1,997 users and are followed by 5,643 others. However, the 91 common followers of @NRA and @Everytown follow 17,808 users on average, and are followed by 60,278 users. Part of this discrepancy is a single user; a parody account for National Basketball Association (NBA) player and worldwide icon LeBron James follows both accounts, and has 3.6 million followers.

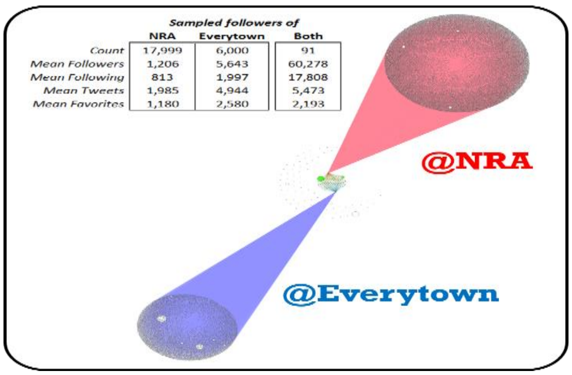

Figure 4. Social network comparison of @Everytown and @NRA.

Regardless, if these sample sets were extrapolated to the entire follower population for @NRA and @Everytown, we should expect between 1000-1,700 common followers. This range still represents only a small portion of the network (only 2.6% of Everytown followers, for example).

6.3 Space

Spatially analysis accounted for the most robust techniques to compare these social media users to mass shooting events. Figure 5 displays a kernel density estimation of all 80 U.S. mass shootings from 1982-2016. There are approximately six hotspots, or clusters of elevated activity: the northeast (Connecticut, New York, New Jersey); the southeast (Florida, Georgia, South Carolina); the Midwest (Illinois, Indiana); Texas; southern California; and the Pacific Northwest (Oregon, Washington).

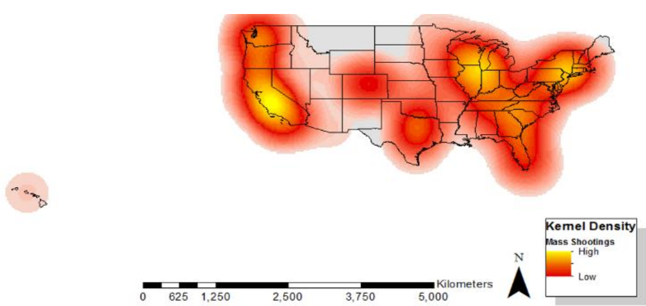

Figure 5. Kernel density of U.S. mass shootings, 1982-2016.

Given the relatively low number of events across a large geographic space, the search radius is based on nearest neighbor calculations. Thus, it is a large number, making the clusters potentially artificially larger and subject to event-to-event variation. Regardless, six key areas (with a seventh in Colorado not quite hot enough to be deemed a hotspot) are identified.

Figure 6 illustrates an identical process for @Everytown followers. This map consists of 2,354 social media users clustering into three distinct hotspots, and multiple lesser significant areas. The three clusters of highest interest are: the northeast (Connecticut, New York, New Jersey); the Midwest (Illinois, Indiana, Wisconsin); and southern California. These areas appear to align closely with typically Democratic political strongholds.

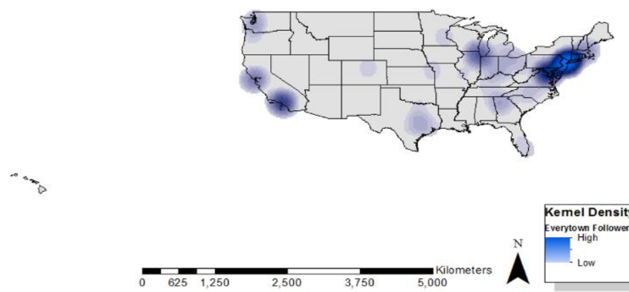

Figure 6. Kernel density of @Everytown followers.

Figure 7, a kernel density estimation of 3,916 @NRA followers, demonstrates a more dispersed network. There are at least six high priority hotspots: the northeast (Connecticut, New York, New Jersey); central Florida; northern Georgia; the Midwest (Illinois, Indiana, Wisconsin); Texas; and southern California. The size of these clusters is consistent with @Everytown followers; this is both a function of the amount of data and the precision.

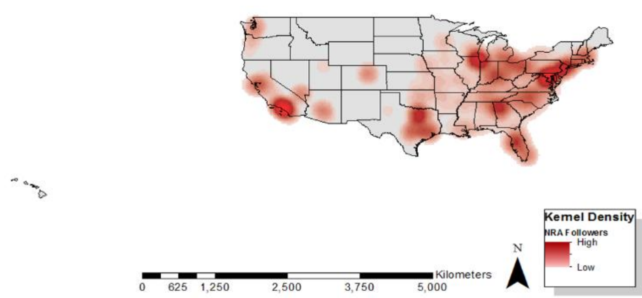

Figure 7. Kernel density of @NRA followers.

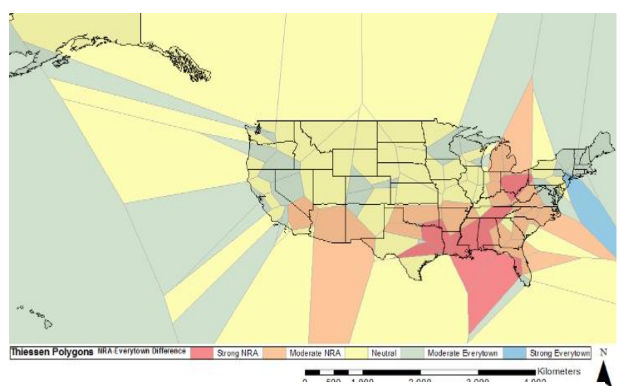

Figure 8. Mass shooting Thiessen polygons, colored by @NRA and @Everytown majorities.

Further, Figure 8 is a portrayal of Thiessen polygons of mass shootings. These areas are irregular by design, naturally organized relative to the surrounding points. Additionally, these polygons are color-coded by the number of each group of followers. Thus, @NRA and @Everytown "win" a polygon by having more followers in that area. Based on this analysis, @NRA significantly wins approximately six areas: five in the southeast, ranging from Tennessee to Florida to Texas; and one in the Midwest, basically consuming Ohio. There are other, less substantial pro-NRA areas in the southwest, mid-Atlantic, and Oklahoma/Arkansas area as well. As for @Everytown, there is only one significant win: the northeast. There are many other less significant pro-Everytown areas, including much of the West Coast; portions of Colorado; a sliver of the Pacific Northwest; Minnesota/Wisconsin; as well as everything north of Pennsylvania.

Finally, the distance of each geocoded social media user to their nearest mass shooting event was calculated and grouped by @NRA and @Everytown followers. A two-tailed t-Test for statistical significance at p=0.05 indicated a statistical significant difference and rejection of the null hypothesis. The followers of @NRA are located significantly farther away from mass shooting events than the followers of @Everytown.

t-Test at p =0.05

	NRA follower distance	*Everytown follower distance*
Mean	1.220060542	0.799353278
Variance	3.646730892	3.085862334
Observations	3916	2354
Hypothesized Mean Difference	0	
df	5282	
t Stat	8.884819642	
P(T<=t) one-tail	4.31866E-19	
t Critical one-tail	1.645142161	
P(T<=t) two-tail	8.63731E-19	
t Critical two-tail	1.960413209	

Figure 9. Kernel density of U.S. mass shootings, 1982-2016.

7. FINDINGS

The findings of this study are rather straightforward. The original hypothesis offered in the research question is rejected; in fact, based on average distance to nearest mass shooting, the supporters of pro-firearm control live closer to mass shootings events than the supporters of pro-firearm rights. However, there is some nuance in that finding.

First, given the parameters of the research question, methodology, and data collected, it is reasonable to ascertain that the social media platform used was appropriate. Other services, with either reciprocal communication, newer implementation, or problematic APIs would not suffice. For measures such as this, Twitter will likely remain the most optimal choice. Second, the use of @NRA and @Everytown as seed accounts appears sound. A brief statistical comparison indicates the accounts are similar in reputation [7] and connectivity. However, the accounts are not *too* similar. Text analysis of recent tweets indicates their differences in word choice and messaging. Social network analysis demonstrates their nearly independent networks of followers. Even those users that are common to both networks appear to be anomalies. In short: @NRA and @Everytown represent two ends of firearm spectrum.

Moreover, while not as cleanly separated, these networks operate in unique spaces. The @Everytown network is characterized by its highly clustered distribution; it only strongly maps to three distinct parts of the country. Conversely, the @NRA network maps to twice as many areas – which include all three of the @Everytown locations! While this degree of overlap is likely a "big city" function – those three areas include Los Angeles, New York City, and Chicago – it remains noteworthy when compared to mass shootings. Neither network is a perfect overlap of mass shooting clusters. Mass shootings have six significant clusters. The three clusters of the @Everytown network overlap three of these (northeast, Midwest, southern California); the @NRA network can account for five of the six (the same three, as well as the southeast, and Texas). One mass shooting cluster is relatively un-accounted for by either network: the Pacific Northwest. Approximately 10% of the mass shooting events (8 incidents) occur in this area, and while both networks have a presence there, it is faint.

This notion of a faint presence is an important distinction. As evidenced by Figure 8, the @NRA network has more strong presence areas; a six to one advantage of @Everytown. These six areas are heavily clustered in the southeast, spilling into Texas and the Midwest. It is essentially one really strong area. From there, the @NRA network only has 15 other areas in the country with weaker (but still pro-NRA) support. Conversely, the @Everytown network has 28 of these faint-but-still-majority areas. Those areas are scattered across the country align with five of the six mass shooting hotspots. The only area not accounted for is Florida/Georgia/South Carolina; there is a small portion of Florida that is pro-Everytown, but not enough to count. Thus, the @NRA network has its strongest support spread over several areas; the @Everytown network has its weaker support spread over many areas. Not only does this seemingly create a larger geographic social network, it also fills in the gaps of mass shooting clusters.

This becomes even clearer with the mean distance t-Test. Members of the @Everytown network are almost twice as close to a mass shooting location as members of the @NRA network. When comprehensively examining the results of the three sets of analyses, it becomes clear how significantly different each network is clustered.

8. CONCLUSIONS

In conclusion, the original hypothesis of this study is rejected: using extracted, normalized, and cleaned locations of sample Twitter followers for @NRA and @Everytown, this study finds that the supporters of pro-firearm control live closer to mass shootings events than the supporters of pro-firearm rights. The @NRA network is more densely clustered in non-mass shooting hotspots, while the @Everytown network is more dispersed across mass shooting hotspots. Thus, the geolocated followers of @NRA are located significantly farther away from mass shooting events than the geolocated followers of @Everytown. Given these findings, there are weaknesses in this study and opportunities for future research to address. Any and all weaknesses in this study revolve around data integrity. First, the method of data sampling used in this study was the least intrusive and least expensive method; it was not the most robust and complete. Using a full set of follower data may reveal more common followers than expected, and additional locations that nullify these results. Second, the decision to use bio locations rather than geotagged metadata potentially sacrificed quality for quantity. While the average Twitter sample garners 1-3% in geotagged metadata [10], [16], this study had significantly higher returns. However, the returns are based on users – who lack any incentive to be accurate or truthful – self-reporting. Third, and related, in the process of cleaning, normalizing, and geocoding. This study compare address level data (mass shootings) to locale level data (social media users). Such a comparison artificially adjusts distances; some shorter, some farther.

However, having outlined those weaknesses, this study is also one of the first to attempt to identify nationwide patterns of social networks relative to crime. In doing so, the findings reveal, at minimum, opportunities for future research. Such future research should include replications of this methodology on larger (possibly full) samples, different social media accounts and followers, as well as geotagged tweets instead of bio locations. The consistency of those findings may help validate this study. Further, and of most interest, would be the addition of temporal analysis. Specifically, do the areas surrounding mass shootings change in firearm support *after* an incident? That question of change detection could lead to substantial insight on how a neighborhood's opinion on firearm control is effected by a mass shooting.

At a maximum, the findings of this study reveal distinct patterns of social networks over space. Those patterns may contribute to the conditions leading to future events; likewise, those patterns may be indicative of reactions to what has already happened.

9. REFERENCES

[1] Barnes, J.A., and Harary, F. 1983. Graph theory in network analysis. *Social Networks*, 5, 235-244.

[2] Brantingham, P.L., and Brantingham, P.J. 1998. Mapping crime for analytic purposes: Location quotients, counts, and rates. *Center for Problem-Oriented Policing*. Retrieved from http://www.popcenter.org/library/crimeprevention/volume_08/09-brantingham.pdf.

[3] Croitoru, A., Crooks, A., Radzikowski, J., and Stefanidis, A. 2016. Geovisualization of social media. *The International Encyclopedia of Geography: People, the Earth, Environment, and Technology*, 1-31.

[4] Gerber, M.S. 2014. Predicting crime using Twitter and kernel density estimation. *Decision Support Systems*, 61, 15-125.

[5] Isenstein, L. 2015. The states with the most gun laws see the fewest gun-related deaths. *National Journal*. Retrieved from http://www.nationaljournal.com/s/53345/states-with-most-gun-laws-see-fewest-gun-related-deaths.

[6] Jain, A.K. 2010. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31, 651-666.

[7] Kietzmann, J.H., Hermkens, K., McCarthy, I.P., and Silvestre, B.S. 2011. Social media? Get Serious! Understanding the functional building blocks of social media. *Science Direct*, 54, 241-251.

[8] Koper, C.S., Woods, D.J., and Isom, D. 2015. Evaluating a police-led community initiative to reduce gun violence in St. Louis. *Police Quarterly*, DOI: 10.1177/1098611115608506.

[9] Luo, W., and MacEachren, A.M. 2014. Geo-social visual analytics. *Journal of Spatial Information Science*, 8, DOI: 10.5311/JOSIS.2014.8.139.

[10] Morstatter, F., Pfeffer, J., Liu, H., and Carley, K.M. 2013. Is the sample good enough? Comparing data from Twitter's streaming API wit Twitter's firehose. *Association for the Advancement of Artificial Intelligence*. arXiv preprint arXiv: 1306.5204.

[11] Perrin, A. 2015. Social media usage: 2005-2015. *Pew Research Center*. Retrieved from http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/.

[12] Reich, B.J., and Porter, M.D. 2015. Partially supervised spatiotemporal clustering for burglary crime series identification. *Journal of Royal Statistical Society*, 178(2), 465-480.

[13] Rogerson, P., and Sun, Y. 2001. Spatial monitoring of geographic patterns: An application to crime analysis. *Computers, Environment and Urban Systems, 25*(6), 539-556.

[14] Signorini, A., Segre, A.M., and Polgreen, P.M. 2011. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS ONE*, 6 (5): e19467. DOI 10.1371/journal.pone. 0019467.

[15] US mass shootings, 1982-2016: Data from Mother Jones' investigation. 2016. *Mother Jones*. Retrieved from http://www.motherjones.com/politics/2012/12/mass-shootings-mother-jones-full-data.

[16] Valkanas, G., Katakis, I., and Gunopulos, D. 2014. Mining Twitter data with resource constraints. *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*-Volume 01.