Re-spatializing Gangs: An Exponential Random Graph Model of Twitter Data to Analyze the Geospatial Distribution of Gang Member Connections

Ryan J. Roberts
Gardner-Webb University
Boiling Springs, NC
rroberts3@gardner-webb.edu

ABSTRACT

Gangs are commonly described as loose affiliations of juvenile members whose time in the gang is short. The connections between members are explained as part and parcel of neighborhood effects, which is consistent with the proximity principle. As a result, policies that address gangs are limited to narrowly defined geographic spaces which can have deleterious security consequences that extend beyond local boundaries. This location-based approach to explaining gangs neglects the role of social media in re-spatializing connections between gang members. The present study collects Twitter data to analyze the geospatial distribution of gang member connections using an exponential random graph model (ERGM) of location homophily. An ERG model analyzes network sub-structure to determine the patterns of relationships between vertices. The data collection for this research involves a four-step iterative process. Step one involves detecting initial network seeds using a language algorithm to capture streaming application programming interfaces (APIs), the Twitter search function, and Twitter recommendations. Step two involves manually inspecting and validating gang member profiles. Step three involves searching the Representational State Transfer (REST) API of validated gang members to determine their location and identify gang members from their list of followers using a random snowball sampling. Step four involves building the network and ERG model. The results of this study support the proximity principle but challenge the consensus that gangs are strictly localized.

ABOUT THE AUTHOR

Ryan Roberts is Assistant Professor of Criminal Justice and Homeland Security at Gardner-Webb University. He received his PhD in 2021 from Old Dominion University with concentrations in international security and modeling and simulation. Prior to earning his degree, he worked as a security threat group intelligence officer at Pasquotank Correctional Institution and served on the North Carolina Prison Emergency Response Team.

Re-spatializing Gangs: An Exponential Random Graph Model of Twitter Data to Analyze the Geospatial Distribution of Gang Member Connections

Ryan J. Roberts
Gardner-Webb University
Boiling Springs, NC
rroberts3@gardner-webb.edu

INTRODUCTION

There is a consensus in gang research that gangs are localized (Coughlin & Venkatesh, 2003; Venkatesh, 2000). This location-based perspective on gangs is partially attributed to data limitations in gang research and studies that focus on the cross-section of gangs, social problems, and crime (Pyrooz & Mitchell, 2015). Moreover, gangs are often described as loosely connected, disorganized groups of juveniles whose time in the gang is short (Venkatesh, 2000). The implication of defining gangs as "youth groups" suggests that gang members lack mobility and that their connections to other gang members are limited to narrowly defined geographic spaces. Thus, gangs conceptualized from location-based perspectives explain gang interconnectivity as embedded in the local landscape, an approach that is consistent with the proximity principle. According to the proximity principle, location determines the formation, existence, and maintenance of interpersonal relationships, where connections are more likely to form in environments that foster repetitive socialization (Newcomb, 1960). This often occurs at the local level, where individuals live, work, worship, or attend school.

Research focusing on local conditions has been used to draw inferences about gang formation and participation, which is featured prominently in the neighborhood effects and collective efficacy literature (see Hagedorn, 1988; Miller, 1958; Sánchez-Jankowski, 1991; Short, Jr. & Strodtbeck, 1965; Thrasher, 1927). However, these studies tend to limit our understanding of gangs to a specific time and place (Venkatesh, 2014). Despite the national and transnational security implications of gang proliferation, location-based studies often neglect the interconnectedness of gangs beyond the neighborhood setting. Advances in communication technology have re-spatialized how gang members share information, form connections, and maintain relationships (Pyrooz & Moule, Jr., 2019). In particular, the increased use of social media platforms, such as Facebook and Twitter, enables gang interactions in unbounded geographic spaces which can transpose local security threats.

Spatializing gangs is typically determined by qualitative methods that are influenced by location-based perspectives (Radil, Flint, & Tita, 2010). This study aims to quantitatively analyze the geospatial distribution of gang members in the United States using an exponential random graph model (ERGM) of Twitter data. ERG models analyze the substructures of social networks to determine the patterns of relationships between vertices (Newman, 2015; Robins & Lusher, 2012). The contributions of this study are threefold. First, I examine location homophily by city and state to determine the extent to which location influences gang member connections. If the location-based gang consensus holds, the smaller the geographic space, the more likely we are to observe the interconnectivity between gang members. The second contribution of this study is the discovery of macro-level implications (gang interconnectedness) through the examination of the micro-level processes (gang member interconnectedness). If gang membership is largely homogenous (gang members belong to the same gang), then, by proxy, we can make inferences regarding the (trans)national connectivity of gangs. Finally, this study analyzes the geographic clustering of the population sample and the distribution of gang members across different cities. Gangs in the United States formed in urban areas and spread to other parts of the country (Howell, 2015). If gangs are strictly localized, it would be reasonable to expect the frequency distribution of gang members from the sample population to be concentrated in high-density cities. Although this objective is less related to the ERG model, it is still an important contribution to understanding the geospatial distribution of gangs.

This paper is divided into three sections. In the first section, gang spatialization is explained from a location-based perspective. Absent a unified theoretical framework, various descriptors that underscore the localization of gangs are highlighted. Whereas some gangs fit the "local actor" description, the sophistication and needs of other gangs have evolved. One tool that facilitates gang transformation involves advances in communication technologies. In particular, gangs use social media to promote gang culture and coordinate, recruit, and disparage rival gangs (National Gang Intelligence Center [NGIC], 2015).

In the second part of this paper, the research methodology is discussed. As gang members use social media, it provides a valuable data source for research on gangs. In this study, a network of gang members on Twitter is detected and constructed using a four-step process. The first step is the initial seed discovery, where gang member profiles are identified by capturing streaming application programming interfaces (APIs) with a language-based algorithm, the search function is used, and Twitter recommendations are followed. In stage two, a relevance computation is conducted by manually inspecting each profile to validate the gang members using multiple criteria. The third step involves searching the Representational State Transfer (REST) API to determine the locations of the validated gang member profiles. Additionally, an exponential non-discriminative snowball sampling process is used to randomly draw *followers* from each profile. Out of the randomly selected group, the techniques from stage two are applied to manually validate the gang member profiles. Stages two and three are continued as an iterative process to build a network edgelist in the fourth and final step.

The final section of this paper provides the results of the data collection workflow process as well as the ERG model. The ERG model tests four hypotheses. Three hypotheses use nodal attributes of city, state, and gang affiliation to analyze the impact of homophily on gang member connections. The fourth hypothesis involves an edge attribute to determine the influence of distance (miles) on gang connections. After interpreting the results, the implications of this study are discussed and suggestions for future research are provided. Insofar as the results of this study support the proximity principle, it challenges the location-based gang consensus. Whereas location homophily plays a role in observing shared connections between gang members to an extent, the city level is not as high as one would expect, given the consensus that gangs are local actors. In fact, the state-level and gang affiliation variables appear to better explain gang member connections and, by proxy, demonstrate gang interconnectedness on a larger scale than is represented in location-based studies. Moreover, the results from the data collection process suggest that gang member locations are diffuse, spanning miniscule- to high-density cities both within and outside the United States.

SPATIALIZING GANGS

Gangs are often treated as groups embedded within local geographic spaces (Coughlin & Venkatesh, 2003), where the spatial distribution of gangs is commonly determined through qualitative means (Radil, Flint, & Tita, 2010; Venkatesh, 2000). This strand of gang research assumes that gang interconnectivity is established through neighborhood or community ties, a perspective rooted in the proximity principle. The proximity principle states that interaction at the local level leads to a higher likelihood of forming interpersonal relationships (Festinger, Schachter, & Back, 1950; Newcomb, 1960). Absent a unified theoretical framework, this location-based perspective often applies descriptive language to indicate that gangs are localized. One commonly accepted gang definition uses observable characteristics to qualify gangs as any "durable street-oriented youth group whose involvement in illegal activity is part of its group identity" (Klein & Maxson, 2006; Weerman et al., 2009). Defining gangs as "youth groups" implies a type of impermanence in which member maturation into adulthood leads to gang disintegration (Reiss, Jr., 1988). Moreover, conceptualizing gang members as "juveniles" implies limited mobility, sophistication, and ambition that restrict them to local geographic spaces. Although some gangs fit this description, G. David Curry (2000) and David Pyrooz (2014) deride the term "juvenile gang" as anachronistic. They agree that juvenile membership may have been more prevalent in the past but argue that the gang problem is adult-centric. Survey data from the NGIC (2012) supports their assertion: the results show that 65% of gang members in 2011 were aged 18 years or older. The percentage of adults to youth has been steadily increasing, with approximately three out of every five gang members being adults, an increase of 15% from 1996 when the ratio of adult to youth gang members was 1:1.

In addition to age, James Howell (2012) further describes gangs as loosely affiliated, disorganized groups that lack definitive leadership. One observation he makes about local gangs is that they often adopt the names of nationally recognized gangs to deter confrontation with other local gangs. This creates the illusion of being "connected" and "dangerous" (Felson, 2006). The Drug Enforcement Agency [DEA] (2018) conceptualizes neighborhood-based gangs (NBGs) similar to Howell but makes an important distinction between NBGs and national-level gangs. They explain, "NBGs operate mainly in the specific jurisdictions where they live. Many take on the names of national-level gangs and attempt to emulate them, but they rarely display the same level of sophistication or structure as national-level gangs" (p. 107). In contrast, "National-level gangs are often highly structured; maintain a strict hierarchy, a constitution, and definitive set of rules; and share common tattoos and symbols. They have a presence in many jurisdictions around the country. Many of these national-level gangs work in conjunction with their counterparts in other locations to benefit the whole gang" (p. 108). Although both gang types exist simultaneously, gang research tends to frame gangs as neighborhood-based.

Despite improving our understanding of gangs, location-based gang research tends to neglect gang interconnectedness beyond the mutually constitutive social conditions at the local level. Gangs transform along different trajectories across space and time (Howell, 2015). For example, the commercialization of cocaine and other narcotics in the 1970s and 1980s fundamentally transformed gangs into market-oriented groups motivated by profits rather than territory (Coughlin & Venkatesh, 2003). More recently, social media sites, such as Facebook and Twitter, have re-spatialized how individuals interact, allowing users to form and maintain relationships in unbounded geographic spaces. Cyberspace has transformed the "local gang," once isolated by geography, into a national and transnational web of interconnected communities. A 2015 survey on gang member social media participation conducted by the NGIC shows that nearly 100% of agencies report street gang members having a Facebook account. The same survey shows that a little over 60% of gang members have Instagram and Twitter accounts. Another NGIC survey included in the same 2015 report reveals that gang member social media usage continues during incarceration. Similar to street gang members, Facebook is the most preferred social media platform for prison gang members. Nearly 100% of the agencies reported that their inmates have an active Facebook account. Additionally, 50% of prison gang members use Twitter, while another 45% use Instagram.

A study conducted by Julian Way and Robert Muggah (2016) demonstrates the application of social media as a data collection tool to study the interconnectivity of gangs. They find that gangs and cartels coordinate criminal activities through social media platforms. Although their initial research focuses on the U.S.–Mexico border, they detect a transnational network of connections that extends throughout the United States and Central and South America. Some of the connections they identify include the Skyline Pirus, Los Ántrax, Gente Nueva, and the Black Disciples.

METHODOLOGY

This study aims to quantify the relationship between location and gang member connections. To achieve this, I mine Twitter data to examine the geospatial distribution of gang member connections using an exponential random graph model to test location homophily. The following four models and hypotheses are considered:

Node Attribute Models
Model 1: Location by City
H_0 – City attributes do not impact gang member connections.
H_1 – Gang members in the same city are more likely to form connections.
Model 2: Location by State
H_0 – State attributes do not impact gang member connections.
H_1 – Gang members in the same state are more likely to form connections.
Model 3: Gang Affiliation
H_0 – Gang affiliation does not impact gang member connections.
H_1 – Gang members with the same gang affiliation are more likely to form connections.

Edge Attribute Model
Model 4: Location by Distance (Miles)
H_0 – Distance between gang members does not impact their connection.
H_1 – The smaller the distance between gang members, the more likely they are to form a connection.

Model 1 tests the interconnectedness of gang members by focusing on city homophily. The null hypothesis posits that there is no relationship between city location and observing gang member connections, whereas the alternative hypothesis proposes a positive correlation between city location and gang member connections. According to the location-based gang perspective and proximity principle, we should be able to reject the null hypothesis as gang members are considered local actors. Widening its geographic scope, Model 2 tests the interconnectedness of gang members by focusing on state homophily. The null hypothesis posits no relationship between state location and observing gang member connections. However, it can be inferred from the location-based perspective that if gang members from the same city are connected, gang members from the same state will be connected. The alternative hypothesis for Model 2 proposes a positive correlation between gang member connections from the same state. Model 3 tests the interconnectedness of gang members from the same gang. The null hypothesis posits that gang affiliation does not impact observing gang member connections, whereas the alternative hypothesis proposes a positive correlation between gang affiliation and gang member connections. Although determining the magnitude of these connections is beyond the scope of this study, observing national connections among gang members of the same gang would further challenge the location-based gang perspective by showing that these connections are decentralized.

In Models 1–3, the nodal attributes of city, state, and gang affiliation, respectively, are considered to test homophily; however, in Model 4, I test location homophily using an edge attribute that considers the distance (miles) between gang members. The null hypothesis posits that there is no correlation between the distance in miles and gang member connections. On the other hand, the alternative hypothesis proposes a positive correlation between distance in miles and gang member connections. In addition to the location-based perspective, Model 4 accounts for the compartmentalization of gangs "gang set spaces" proposed by Tita, Cohen, and Engberg (2005). For this study, data was collected using Twitter, a social media website, and R-Studio, an integrated programming environment for R, was used to capture the Twitter streaming API and generate the results of the ERG model.

Workflow Process

The process for conducting a social media network analysis are well established. They typically involve stages of discovery, relevance computation, inspection, and network modeling (see Balasuriya, Wijeratne, Doran, & Sheth, 2016; Décary-Hétu & Morselli, 2011; Patton, 2015; Way & Muggah, 2016). The workflow process for this study includes the following four-steps:

1. Seed Discovery – In the initial seed discovery stage, gang member profiles were identified using three strategies. One detection method used is typing gang names in the Twitter search function. David Décary-Hétu and Carlo Morselli (2011) apply a similar approach when mining gang data on Twitter and Facebook to comparatively analyze the gang groups and pages of each platform. Another gang detection strategy borrows from the authors' recommendations. An automated algorithm is used to capture the Twitter streaming API coded in R-Studio from a bounding box targeting the continental United States. When attempting to analyze human trafficking on the southern border, the use of language was effective for Julian Way and Robert Muggah (2016) in the initial seed discovery process. Gangs use language as a method to establish and reinforce a distinct identity. At times, gang members use a unique set of words and phrases to greet friends, denigrate enemies, or reference people, places, and events. Although not predicated on text data, Balasuriya, Wijeratne, Doran, and Sheth's (2016) study utilizes hashtags like #BDK (Black Disciple Killer) and #GDK (Gangster Disciple Killer) in the discovery stage of their workflow process. Unlike these other studies, however, this study uses language configurations that target a broader spectrum of gangs. The list of words and phrases this study uses to capture tweets are both general and gang-specific to the Bloods, Crips, People Nation, Folk Nation, Five Percenters, Black Guerilla Family, Hispanic gangs, White gangs, Jamaican gangs, Outlaw Motorcycle Gangs, and Asian gangs. Finally, Twitter uses an algorithm to recommend user profiles based on one's Twitter activity. The final detection method used in the discovery process involves following Twitter recommendations.

2. Relevance Computation – The second step involves relevance computation based on the initial seed discovery from the first stage, referenced against exemplary documents. This stage is conducted manually to validate the gang members' Twitter accounts. G. David Curry (2015) emphasizes self-identification as important in the validation process. When inspecting the profiles, self-identification is sought out in addition to other indicators. Gang member profiles with two or more of the following criteria are included in the data set: self-identification, language, hand signs, tattoos, media illustrating gang culture/symbols, gang colors, associates, hashtags, emojis, or external news sources (primarily used for gang-affiliated celebrities).

A part of the identification process also involves determining the gang to which a Twitter user belongs. For instance, the six-pointed star is a symbol used by Jewish practitioners and members of the Folk Nation. Manual inspection of Twitter profiles allows for ascertaining the context of these symbols. Emojis are another symbol that can have multiple applications. The handicap or grape emojis can have one meaning for non-gang members but are also used by the Crips and Grape Street Crips, respectively. Therefore, the inclusion of false-positive profiles is mitigated by focusing on at least two validation criteria.

3. Search REST API – After validating the profiles in the second stage, I search the Twitter REST API to determine the location of gang members and discover other gang member accounts. The location was manually identified for all Twitter accounts inspected as opposed to using the geodesic code. One of the weaknesses of relying on the geodesic code is highlighted in Sanjaya Wijeratne, Derek Doran, Amit Sheth, and Jack Dustin (2015), whose study results only produced a location in 3.62% of the detected profiles. In cases where multiple locations were discovered, I code them as primary or secondary. Additionally, other gang member accounts are extracted through retweets, user mentions, and a list of followers. The data selection process uses an exponential non-discriminative snowball sample, where referrals are randomly drawn from the initial seeds and their *followers*. I consider the list of *followers* as opposed to the list a user is *following* because this signals the intent to subscribe or receive notifications from a specific Twitter user. As the *followed* can choose to block a *follower*, allowing an account to follow is an implicit acceptance of that connection. Finally, after discovering additional profiles from the Twitter REST API, I validate these accounts using the same criteria as in stage two of this workflow process. I continue this as an iterative process for up to 200 followers, or until the discovery of *follower* profiles is exhausted. Additionally, all non-relevant profiles are discarded, and relevant profiles are added to the dataset.

4. Build Network – The relevant profiles discovered from the workflow process are used to build a network using an edgelist, where the vertices or nodes represent Twitter users, and an edge indicates a tie between vertices (see Piquette, Smith, & Papachristos [2014] for a discussion on the benefits of social network analysis [SNA] to gang studies). The network used is an undirected graph that assumes reciprocity between gang members. To conceal the identity of Twitter users, I designate each node with a numerical value. The data for this study was collected between June 1 and June 30, 2019. Network data are analyzed using an ERGM. Similar to regression analysis, ERGMs examine the influence of an independent variable on a dependent variable. However, while statistical regression assumes independence between nodes, ERGMs account for their interrelatedness. It is the dependence between nodes that forms the structural foundation of a network and the point of interest for an ERG model. The ERGM used in this study tests the location homophily of gang member connections, or the extent to which we are likely to observe connections between gang members from similar locations.

Table 1 shows the results of the workflow process which led to the discovery of 1,636 connections between 726 gang members in 135 cities, 48 states, and 13 countries, including the United States.

Table 1. Workflow Process Results

Gang Members	726
Connections	1636
Gangs	48
Cities	135
States	48
Countries	13

Exponential Random Graph Model (ERGM)

ERGMs analyze the substructures of social networks to determine the patterns of relationships between vertices. Garry Robins and Dean Lusher (2012) provide the following definition of ERGMs:

> Exponential random graph models (ERGMs) are statistical models for network structure, permitting inferences about how network ties are patterned. Put another way, ERGMs are tie-based models for understanding how and why social network ties arise. This focus aligns ERGMs with a principal goal of much empirical social network research, which is to understand a given "observed" network structure (i.e., a network on which a researcher has collected data), and so to obtain insight into the underlying processes that create and sustain the network-based social system (p. 9).

A more formal explanation of ERGMs can be found in David Hunter, Mark Handcock, Carter Butts, Steven Goodreau, and Martina Morris (2009). ERGMs function in a manner quite similar to linear regression models with one distinct feature: they account for path dependencies in the network structures. This can be accomplished by measuring the impact of nodal attributes. For further explanation and a comparison between nodal attribute models and evolutionary models, see Riitta Toiven et al. (2009). In addition to node attributes, edge attributes (also referred to as relational attribute effects) can be used to determine the probability distribution of a graph (see Morris, Handcock, & Hunter [2008] for a more detailed explanation).

For this study, an ERGM is used with an undirected network graph to test the location homophily of shared gang member connections. By using the ERG model, this study aims to understand the extent to which location impacts gang member connections. Although there is a degeneracy problem in ERGMs, this relates to the issues of transitivity in social networks. Transitivity analyzes the likelihood that a friend of a friend is your friend. For this reason, triadic closures or network clustering are not relevant to this study but should be considered in future research. ERGs that model homophily, however, do not suffer from the same limitation (see Rinaldo, Feinberg, & Zhou [2009] for a detailed explanation of ERGM degeneracy).

RESULTS

For each calculation, there is a null model that shows the probability of a connection forming between gang members without considering the attributes. For example, the edgelist used in the city attribute model shows a 1.12% probability of a connection being formed between two nodes. This means that, in the absence of any identifiable criteria, there is a low probability of observing a connection between two individuals in the network. The edgelists used in the state and gang affiliation nodal attribute models and the edge attribute model also show a low probability of observing connections between nodes when only edges are considered.

We can observe the relevance of the attributes by comparing them to the null models. This study's results support the proximity principle to some degree. Individuals concentrated in geographic spaces are more likely to develop interpersonal relationships. When considering the individual effect of nodal attributes, location has an impact on the formation of gang connections. In the first model, city attributes are statistically significant at the 95% confidence interval ($p < 0.0139$). We can reject the null hypothesis that city location does not impact gang member connections. Model 1 includes 634 connections (edges) between 335 gang members (vertices).

When considering the individual effect of state location, the statistical significance of connections forming between gang members is higher. Model 2, which includes 771 edges connecting 385 vertices, measures state attributes and is statistically significant at the 99% confidence interval ($p < 0.0045$). In model 2, we can reject the null hypothesis that state location has no impact on gang member connections forming. Although a national model is not included in this study, it can be inferred that connections based on country are highly statistically significant at the 99.99% confidence interval ($p < 0.001$), especially considering that of the 726 gang members detected, 672 are from the United States. The third model that tested individual effects is gang affiliation homophily with 1,538 edges connecting 717 vertices. Gang affiliation is highly statistically significant at the 99.99% confidence interval ($p < 0.0002$). For Model 3, it is important to note that the results are based on gang sets rather than their primary affiliation. The Rollin' 60s Neighborhood Crips, for example, are treated as separate entities from the Crips. Therefore, it can be inferred that primary gang affiliation is also highly statistically significant.

Unlike the three nodal attribute models, Model 4 uses an edge attribute to test the individual effect of distance between vertices (measured in miles). The miles between the gang members tested in Model 4 do not significantly impact the formation of a connection. Although the distance in miles is not a good predictor of gang member connections, we can still make inferences about the location-based perspective. If gangs are localized, we would expect to see higher clustering in terms of distance. The miles between nodes might be too scattered to make a statistical determination of the impact of distance and the formation of gang member connections; however, this is not necessarily a reflection of proximity. Gang members that are 2, 3, 5, or 10 miles apart can be considered geographically proximate. However, the dataset for Model 4 (the same dataset used in Model 1) shows that the distance between the nodes is decentralized rather than clustered. The average distance between vertices is 963.24 miles, with a range of 0–12,863 miles. Though we may not be able to reject the null hypothesis for Model 4, the distance between nodes challenges the idea that gangs are localized. Rather than clustering, the mileage between gang members suggests that they occupy a more diffuse geographic space. Table 2 provides the ERGM results for the individual effects of attribute homophily (city, state, gang affiliation, and distance [miles]) on gang member connections.

Table 2. Individual Effects of Attribute Homophily

	Model 1		Model 2		Model 3		Model 4	
	Null 1	City Nodal Attribute Model	Null 2	State Nodal Attribute Model	Null 3	Gang Nodal Attribute Model	Null 4	City Edge Attribute Model
Vertices	335	335	385	385	717	717	335	335
Edges	634	634	771	771	1538	1538	634	634
Estimate Std.	-4.485	0.3691	-4.574	0.2921	-5.13	0.4978	-4.4848	22.5093
Error	0.0403	0.15	0.0366	0.1027	0.0259	0.0672	0.0403	210.3468
p-Value	<1e-04***	0.0139*	<1e-04***	0.0044**	<1e-04***	0.0002***	<1e-04***	0.915
Probability	0.0112	0.5912	0.0102	0.5725	0.0058	0.5678	0.0112	1
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1							

DISCUSSION

Gang members commit crimes at a higher rate than do non-gang criminal offenders. "Effective use of SNA (social network analysis) techniques to mine criminal network data can have important implications for crime investigations. The knowledge gained may aid law enforcement agencies fighting crime proactively" (Xu & Chen, 2005, p. 106). This is especially more acute in a globalized world where criminal connections have become transnational (Brewster, Polovina, Rankin, & Andrews, 2014). In addition to SNA as a resource for learning about the interpersonal relationships of gang connections, open-source data and text analytics facilitate sociometric analysis to mitigate criminal threats. One method of understanding the gang threat is to study the interconnectedness of gangs in the social media era. This study's findings are consistent with the proximity principle. In other words, location homophily plays a role in the formation of gang member connections. It is reasonable to expect that people living close together are more likely to have interpersonal relationships. Social interaction at school, work, and place of worship, or in shared residential spaces increases the likelihood of localized connections forming. Gangs exist within these public spaces, making it unsurprising that city and state attributes help explain gang member connections to some extent. However, city homophily is not as strong a predictor of gang member interconnectivity as one would expect to observe, given the location-based consensus in gang studies. Depending on the unit of analysis or how location is defined (e.g., public housing complex, street, city, county, state), this study shows that the wider the geographic space, the greater the likelihood of observing a shared connection between gang members. Hence, gang member connections appear to be less localized than the extant literature suggests. Definitions that describe gangs as loosely organized groups of juveniles seeking to protect territory discount their national and transnational connections. Instead, advances in communication technology like social media platforms have enabled gang members to re-spatialize how they form and maintain friendships in unbounded geographic spaces.

The study findings challenge the location-based perspective that asserts gang localization in two important respects. First, the frequency distribution of the sample population suggests that gang affiliation is a strong indicator of gang member connectivity. Approximately 60 percent of gang members from the same set share a connection. These connections increase to 82 percent when gang members are consolidated into the primary gang with which that set is aligned. The increase of shared connections between gang members from "gang set" to "primary gang" supports the value of understanding the (trans)national relationship between gangs. There is a high degree of homogenous ties between gang members of the same gang or the alliance with which their gang belongs. The ERGM results support gang homophily as a strong indicator of shared gang member connections.

Second, the concentration of gang members in the sample population reveals that gang members are primarily located in mid- to small-density cities. If gang members were localized, we would expect to see more gang members concentrated in large-density cities because gangs originated in large urban centers (Howell, 2015). In the sample population for this study, there are nearly just as many gang members in high-density cities as there are in minuscule-density cities. Similarly, the locations represented in this study are geospatially diverse. Gang member connections are domestically and internationally more diffuse than is currently represented in location-based gang studies. By proxy, the interconnectedness of gangs at the macro level is dispersed over a larger geographic space. The consequence of this transposes localized security threats to the (trans)national consciousness by facilitating recruitment opportunities, disseminating gang culture, and enabling the coordination of criminal gang activity across city, state, and (trans)national borders. On the whole, these activities contribute to the threat of gangs as described by Max Manwaring (2005) whereby they challenge law and order, weaken institutions, and impact the structural integrity of the state.

REFERENCES

Balasuriya, L., Wijerante, S., Doran, D., & Sheth, A. (2016). Finding street gang members on Twitter. *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 685-692). San Francisco, CA: IEEE Press.

Brewster, B., Polovina, S, Rankin, G., & Andrews, S. (2014). Knowledge management and human trafficking: Using conceptual knowledge representation, text analytics and open-source data to combat organized crime. In N Hernandez, R. Jaschke, & M. Croitoru (Eds.), *Graph-based representation and reasoning: Lecture notes in computer reasoning* (pp. 104-117). New York, NY: Springer Publishing.

Coughlin, B. C. & Venkatesh, S. A. (2003). The urban street gang after 1970. *Annual Review of Sociology*, *29*, 41-64.

Curry, G. D. (2000). Self-reported gang involvement and officially recorded delinquency. *Criminology, 38*, 1253-1274.

Curry, G. D. (2015). The logic of defining gangs revisited. In S. H. Decker & D. C. Pyrooz (Eds.), *The handbook of gangs* (pp. 7-27). West Sussex, UK: Wiley-Blackwell.

Décary-Hétu, D., & Morselli, C. (2011). Gang presence in social network sites. *International Journal of Cyber Criminology*, *5*(2), 876-890.

Decker, S. H., Bynum, T., & Weisel, D. (1998). A tale of two cities: Gangs as organized crime groups. *Justice Quarterly*. *15*(3), 395-425.

Drug Enforcement Administration. (2018). National drug threat assessment. Washington, DC: U.S. Department of Justice.

Felson, M. (2006). *Crime and nature*. Thousand Oaks, CA: Sage.

Festinger, L., Schachter, S., & Bach, K. (1950). *Social pressures in informal groups*. New York, NY: Harper.

Hagedorn, J. M. (1988). *People and Folks: Gangs, crime and the underclass in a rustbelt city*. Chicago, IL: Lake View.

Howell, J. C. (2012). *Gangs in America's communities*. Thousand Oaks, CA: Sage.

Howell, J. C. (2015). *The history of street gangs in the United States: Their origins and transformation*. Lanham, MD: Lexington Books.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2009). ERGM: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, *24*(3), 1-29.

Jankowski, M. S-, (1991). *Islands in the street: Gangs and American urban society*. Los Angeles, CA: University of California Press.

Klein, M. W., & Maxson, C. L. (2006). Street gang patterns and policies. Oxford, UK: Oxford University Press.

Miller, W. B. (1958). Lower class culture as a generating milieu of gang delinquency. *Journal of Social Issues*. *14*(3), 5-19.

Manwaring, M. G. (2005). Street gangs: The new urban insurgency. Army War College Strategic Studies Institute Carlisle Barracks PA.

Morris, M., Handcock, M. S., & Hunter, D. R. (2008). Specification of exponential-family random graph models: Terms and computational aspects. *Journal of Statistical Software, 24*(4), 1-24.

National Gang Intelligence Center. (2012). National youth gang survey analysis. Retrieved from https://www.nationalgangcenter.gov/Survey-Analysis/Measuring-the-Extent-of-Gang-Problems.

National Gang Intelligence Center. (2013). *The 2013 national gang report*. Washington, DC: Federal Bureau of Investigations.
National Gang Intelligence Center. (2015). *The 2015 national gang report*. Washington, DC: Federal Bureau of Investigations.
Newcomb, T. M. (1960). Varieties of interpersonal attraction. In D. Cartwright & A. Zander (Eds.), *Group dynamics: Research and theory* (2nd ed., pp. 104-119). Evanston, IL: Row, Peterson.
Newman, M. E. J. (2015). *Networks: An introduction*. Oxford, UK: Oxford University Press.
Papachristos, A. V. & Kirk, D. (2006). Neighborhood effects on street gang behavior. In J. F. Short & L. A. Hughes (Eds.), *Studying youth gangs* (pp. 63-84). Lanham, MD: Alta Mira Press.
Patton, D. U. (2015). Gang violence, crime, and substance use on Twitter: A snapshot of gang communication in Detroit. *Proceedings of the Society for Social Work and Research 19th Annual Conference: The Social and Behavioral Importance of Increased Longevity*. New Orleans, LA: SSWR.
Piquette, J. C., Smith, C. M., & Papachristos, A. V. (2014). Social network analysis of urban street gangs. In G. Bruinsma & D. Weisburd (Eds.) *Encyclopedia of criminology and criminal justice* (pp. 4981-4991). New York, NY: Springer.
Pyrooz, D. C. (2014). From your first cigarette to your last dyin' day: The patterning of gang membership in the life-course. *Journal of Quantitative Criminology*, *30*(2), 349-372.
Pyrooz, D. C., & Mitchell, M. M. (2015). Little gang research, big gang research. In S. H. Decker & D. C. Pyrooz (Eds.), *The handbook of gangs* (pp. 28-58). Hoboken, NJ: Wiley-Blackwell.
Pyrooz, D. C., & Moule, R. K. (2019). Gangs and social media. In *Oxford research encyclopedia of criminology and criminal justice*. Oxford University Press. Retrieved from https://doi.org/10.1093/acrefore/ 9780190264079. 013.439.
Radil, S. M., Flint, C., & Tita, G. E. (2010). Spatializing social networks: Using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. *Annals of the Association of Geographers*, *100*(2), 307-326.
Reiss, A. J., Jr. (1988). Co-offending and criminal careers. *Crime and Justice*, *10*, 117–170.
Rinaldo, A., Fienberg, S. E., & Zhou, Y. (2009). On the geometry of discrete exponential families with application to exponential random graph models. *Electronic Journal of Statistics*, *3*, 446-484.
Robins, G. & Lusher, D. (2012). What are exponential graph models? In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential random graph models for social networks* (pp. 9-15). New York, NY: Cambridge University Press.
Short, J. F. Jr., & Strodtbeck, F. L. (1965). *Group process and gang delinquency*. Chicago, IL: University of Chicago Press.
Thrasher, F. M. (1927). *The gang: A study of 1,313 gangs in Chicago*. Chicago, IL: University of Chicago Press.
Tita, G. E., Cohen, J., & Engberg, J. (2005). An ecological study of the location of gang "set space." *Social Problems*, *52*, 272–299.
Venkatesh, S. A. (2000). *American project: The rise and fall of a modern ghetto*. Cambridge, MA: Harvard University Press.
Venkatesh, S. A. (2014). The inevitable gang. In J. M. Hazen & D. Rodgers (Eds.) *Global gangs: Street violence across the world* (pp. 281-287). Minneapolis, MN: University of Minnesota Press.
Way, J., & Muggah, R. (2016). Charting out the digital ecosystem of gangs in the U.S. and Mexico. Retrieved from https://smallwarsjournal.com/jrnl/art/charting-out-the-digital-ecosystem-of-gangs-in-the-us-and-mexico.
Weerman, F. M., Maxson, C. L., Esbensen, F-A., Aldridge, J., Medina, J. & van Gemert, F. (2009). Eurogang program manual: Background, development, and use of the Eurogang instruments in multisite, multi-method comparative research. Retrieved from: https://www.umsl.edu/ccj/Eurogang/EurogangManual.pdf.
Wijeratne, S., Doran, D., Sheth, A., & Dustin, J. (2015). Analyzing the social media footprint of street gangs. *Proceedings of the IEEE International Conference on Intelligence and Security Informatics (ISI)*. Baltimore, MD: IEEE Publishing.
Xu, J., & Chen, H. (2005). Criminal network analysis and visualization. *Communications of the ACM*, 48(6), 100-107.