2021

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

Ryan J. Roberts
Old Dominion University, rrobe031@odu.edu

Original Publication Citation

Roberts, R. J. (2021). Re-spatializing gangs: An exponential random graph model of Twitter data to analyze the geospatial distribution of gang member connections. *International Journal of Cyber Criminology, 15*(2), 18-43. https://cybercrimejournal.com/menuscript/index.php/cybercrimejournal/article/view/17

Copyright © 2020 International Journal of Cyber Criminology – ISSN: 0974–2891
July – December 2021. Vol. 15(2): 18–43. DOI: 10.5281/zenodo.4766541
Publisher & Editor-in-Chief – K. Jaishankar / Open Access (Authors / Readers No Pay Journal).

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

Ryan J. Roberts[1]
Old Dominion University, United States of America

Abstract
Gang studies often use location-based approaches to explain gang members' interconnectedness. Although this perspective remains consistent with the proximity principle that the smaller the geographic space, the greater the likelihood of observing connections between individuals, location-based studies limit our understanding of gang member connections to narrowly defined geographic spaces at specific points in time. The advent of social media has re-spatialized gang member interconnectedness to unbounded geographic spaces, where the preservation of online activity can extend indefinitely. Despite having an online presence, most research examining the digital footprint of gangs tends to be descriptive. This study collects Twitter data to analyze the geospatial distribution of gang member connections using an exponential random graph model (ERGM) of location homophily. An ERGM analyzes network substructures to determine the patterns of relationships between vertices. In this case, the extent to which homophily by city, state, and gang affiliation determine gang member connections. The results of this study support the proximity principle but challenge the assertion that gangs are strictly localized.

Keywords: gangs, location, Twitter, exponential random graph model (ERGM)

1. INTRODUCTION

There is a consensus in gang research that gangs are localized (Coughlin & Venkatesh, 2003; S.A. 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 (David C. Pyrooz & Mitchell, 2015). Moreover, gangs are often described as loosely connected, disorganized groups of juveniles whose time in the gang is short (S.A. 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

[1] Old Dominion University, United States of America. Email: rrobe031@odu.edu

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 and Macon (1988); Jankowski (1991); Miller (1958); Short and Strodtbeck (1965); Thrasher (1927)). However, studies that aggrandize local conditions limit our understanding of gangs to a specific time and place (S. A. Venkatesh, 2014). Moreover, location-based studies can neglect the interconnectedness of gangs beyond neighborhood settings. Advances in communication technology have re-spatialized how gang members share information, form connections, and maintain relationships (David C 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.

Spatializing gangs is typically determined by qualitative methods that are influenced by location-based perspectives (Radil et al., 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 API 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 REST API to determine the locations of the validated gang member profiles. An exponential non-discriminative snowball sampling process is used by randomly drawing *followers* from the initial seeds. 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 two separate sets of results. The first part includes the data collection results. These include descriptive statistics on gang member Twitter profiles, as well as the gangs and locations discovered from the workflow process. The other set of results includes calculations from the ERG model that aim to test the four hypotheses. Three hypotheses use the 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 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 the gang member location is diffuse. The sample population used for this research shows gangs concentrated in small- and mid-sized cities rather than in highly populated cities.

2. 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 et al., 2010; S.A. 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 et al., 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, the Eurogang 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" (M.W. Klein & Maxson, 2006; Medina-Ariza 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 C 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, 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 takes 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.

The contribution of location-based studies to our understanding of gangs cannot be overlooked. This strand of research features prominently in the neighborhood effects and collective efficacy literature (Papachristos & Kirk, 2006), where gang formation and participation are derived from negative local stimuli. Theories such as social disorganization, concentrated disadvantage, and social inequality use neighborhood effects to explain *how* the failure of social institutions at the local level leads to deviance and other high-risk activities (Sampson et al., 2002). Collective efficacy, in contrast, "refers to the process of activating or converting social ties to achieve any number of collective goals, such as public order or the control of crime" (Papachristos & Kirk, 2006), and explains behavioral outcomes as an adaptive response to deficiencies in local conditions (Sampson et al., 1997). In short, where the government has failed to provide public goods such as security or economic opportunity, individuals facing shared abject conditions at the local level take collective action to improve their situation.

Within this strand of gang research, several motivational factors have a higher intrinsic value within local geographic spaces. For example, gangs claim territory to provide members with a safe area to congregate and conduct illicit business activities. The geographic concentration of gangs results in turf wars (Campbell, 1984; Vargas, 2016), where competition over local resources drives rivalries (Brantingham et al., 2012). Within these gang-controlled territories, Tita et al. (2005) further compartmentalize the geography of gangs into what they refer to as "gang set spaces." Rather than the total area claimed by a gang, they argue that gang set spaces are smaller subsections within a territory reserved for gang activity. In addition to territorial motivation and material benefits, psychological factors at the local level help explain gang participation. For some individuals, gangs satisfy status-seeking behavior and help people meet their peer group needs (Cohen, 1955; Shaw & McKay, 1942; Thrasher, 1927). In some cases, gangs provide a source of friendship, mutual trust, and identity (Malcolm W Klein, 1995); in other cases, they provide a path for individuals to gain power (Knox, 1994) or respect (Anderson, 2000).

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. Like 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.

Research that examines the online behavior of gangs tends to be descriptive (Moule Jr et al., 2014). Leverso and Hsiao (2020) use a digital trace web to analyze Hispanic gangs in Chicago. One of their findings shows that the "digital street" extends to proximate *and* distant geographic spaces. On the "digital street," some gangs use social media to collect intelligence data and denigrate rival gangs. The extension of gang rivalries into the cybersphere is often referred to as "cyber banging." Desmond Upton Patton et al. (2013) refer to this as "The phenomenon of gang affiliates using social media sites to trade insults or make violent threats that lead to homicide or victimization." According to them, the three features of "cyber banging" include the following: "(1) promote gang affiliation and/or communicate interest in gang activity; (2) gain notoriety by reporting participation in a violent act or communicating an impending threat; (3) share information about rival gangs or network with gang members across the country" (p. A55). Examining whether online hostilities translate into offline violence, Stuart (2020) finds that gangs "cyber bang" by attacking their rival's reputation through "cross-referencing," "calling bluffs," and "catching lacking." When an online conflict escalated to offline violence, gang intelligence data were extracted from social media to target rival gang members. Whittaker et al. (2020)

discuss two gang types and their disparate social media usage. One is the "traditionalist" gang, which operates with discretion and largely avoids social media. The other is the "digitalist" gang, which uses social media as a form of branding. In other words, digitalists use social media to promote their gangs, coordinate activities, recruit new members, gain reputation, and expand territory. According to the authors, age and longevity influence whether gangs function as traditionalists or digitalists. Younger and newer gangs tend to garner attention by expanding their digital footprint.

A study conducted by Way and 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. Transnational connections are discovered using the workflow process for this study and are discussed in the data collection results section.

3. METHODOLOGY

This study aims to quantitatively test the impact of location on gang member connections. To achieve this, Twitter data were mined to examine the geospatial distribution of gangs using an ERGM 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 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 into "gang set spaces" proposed by Tita et al. (2005).

For this study, data were collected using Twitter. "Twitter is a real-time global information network that lets users create and share ideas and information instantly. People and organizations send messages through our website and mobile site, client applications (e.g., Twitter for Android; Twitter for iOS), SMS, or any variety of third-party applications" (Twitter Help Center, n.d.). I use R-Studio, an integrated programming environment for R, to capture the Twitter streaming API and generate this study's results. "R is a language and environment for statistical computing and graphics" (The R Foundation, n.d.).

3.1 Workflow Process

The methods for conducting a social media analysis are well established. They typically involve stages of discovery, relevance computation, inspection, and, if applicable, network modeling (see Décary-Hétu and Morselli (2011); Desmond U Patton et al. (2015); Way and Muggah (2016); Wijeratne et al. (2015)). The study's workflow includes the following four-step process.

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. Décary-Hétu and 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 detection strategy used 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 Way and 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, Wijeratne et al. (2015) 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. Table 1 provides a sample of the words and phrases used to capture the Twitter streaming API of gang members.

Table 1.Language Sample for Four of the Largest Gangs in the United States

Gang	Language	Meaning
Folk Nation	All is one	We're all together and OK
	GD	Gangster Disciples
	74	Gangster Disciples
	Vicky Lous	Insult to Vice Lords/People Nation
People Nation	(G)DK	(Gangster) Disciple Killer
	5 in the sky, 6 must die	Revenge against Folk
	ALKN	Almighty Latin King's Nation (a member of the People Nation)
Crips	Slob, Sloob	Disrespect to Bloods
	Adidas	All Day I Destroy a Slob (Blood)
	B/K	Blood killer
	What it C Like	Crip greeting
Bloods	Crab	Disrespectful name for Crip
	Damu	Swahili for Blood
	Snoovers	Insult to Hoover Street Crips
	Krab	Insult to Crips

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 the following Twitter recommendations.

2. Relevance Computation – The second stage 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: 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). Table 2 shows the breakdown of the gang member validation criteria. As the use of text data to detect gang members is central to this study, it is not surprising that the largest factor across all validated gang member profiles is language. Of the total gang members, 80.30% of the gang member profiles include language as one of the validation criteria. Among all validated gang members, 32.37% met at least two criteria, and 33.88% met three criteria.

Table 2. Gang Member Validation Criteria

Validation Criteria	**Total Validation Criteria**	**Validation Criteria as % of Total Gang Members**	**Validation Criteria Met**	**Total Validation Criteria Met by Gang Members**	**Validation Criteria Met as % of Total Gang Members**
Self-Identification	293	40.36%	Two	235	32.37%
Language	583	80.30%	Three	246	33.88%
Hand Signs	237	32.64%	Four	125	17.22%
Tattoo	14	1.93%	Five	84	11.57%
Media	375	51.65%	Six	30	4.13%
Colors	186	25.62%	Seven	6	0.83%
Associates	301	41.46%			
Hashtag	158	21.76%			
Emoji	176	24.24%			
News	25	3.44%			

A further breakdown of those gang members that only met the two-criteria threshold shows that 10.64% were validated because they self-identified *and* used the gang language. Another 78.3% of gang members who met at least two validation criteria included either self-identification *or* language. Those that self-identify and include some other criteria represent 18.72% of the sample population, and 59.57% include language and some other criteria. For all pairs of criteria, substantive evidence was used to validate a gang member. For example, no gang members were validated using only a hashtag or an emoji. If supporting evidence to validate a gang member could not be found, then the profiles were excluded from the dataset. Table 3 provides a breakdown of the validation criteria for gang members meeting the two criteria as a subset of the total sample population.

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.

Table 3. Validation Criteria for Gang Members Meeting Two Criteria as a Subset of the Total Sample Population

Validation Criteria Combinations	Total Validation Criteria Combinations	Combination of Validation Criteria as % of Total Gang Members Meeting Two Criteria
Self + Language	25	10.64%
Self + Other Criteria	44	18.72%
Language + Other Criteria	140	59.57%
Hand Sign + Media	2	0.85%
Hand Sign + Colors	1	0.43%
Hand Sign + Associates	2	0.85%
Hand Sign + Hashtag	2	0.85%
Hand Sign + Emoji	1	0.43%
Tattoo + Emoji	1	0.43%
Media + Colors	3	1.28%
Media + Associates	5	2.13%
Media + Hashtag	2	0.85%
Media + Emoji	2	0.85%
Media + News	1	0.43%
Colors + Emoji	1	0.43%
Associates + Hashtag	1	0.43%
Associates + Emoji	1	0.43%
Associates + News	1	0.43%

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 the Twitter accounts inspected. One of the weaknesses of relying on the geodesic code is highlighted in Wijeratne et al. (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 et al. (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 were collected between June 1 and June 30, 2019. Network data are analyzed using an ERGM. Like 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 gang member connections are localized. The ERGMs are explained in more detail after discussing the data collection results. Figure 1 illustrates the workflow process. The same process presented here can be used to identify gang members on other social media platforms, such as Facebook and Instagram.

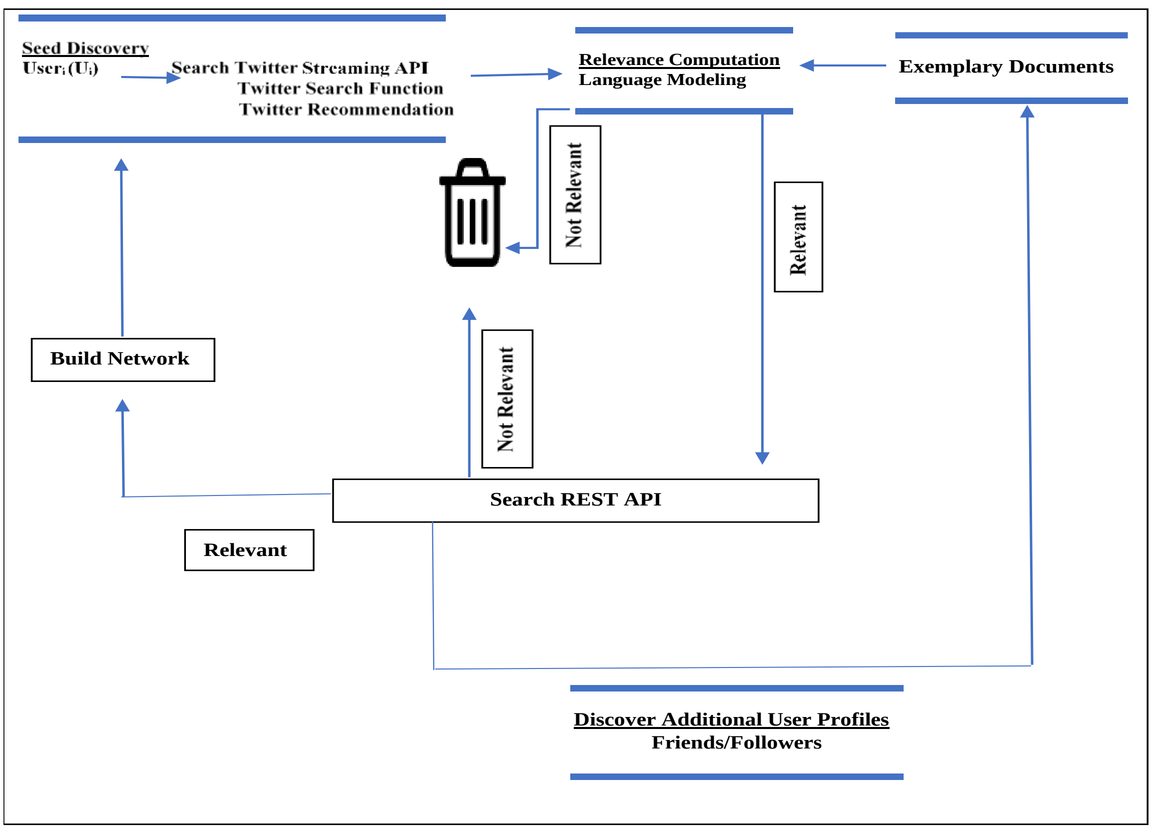


Figure 1. Workflow Process to Collect Twitter Data

3.2 Limitations

The workflow process has three primary limitations. The first was regarding identifying the level of gang involvement. Manual inspection of each profile helps identify false-positive gang member profiles but does not account for the level of gang involvement. According to the Santa Cruz County Gang Task (2018), gang members can be one of three levels. At the lowest level are the *Wannabes*. A *Wannabe* has no formal ties to a gang but expresses an interest in gang culture and often fits the profile of gang members. The second level of gang involvement is an *Associate* characterized by having a personal relationship with a gang member, adopting gang colors and symbols, and considering joining a gang. *Gang Members* are at the highest level of gang involvement. These individuals have gone through the initiation of becoming gang members, pledged their commitment to the gang, frequently engaged in illicit activities, and fully adopted the gang's language, symbols, and rituals. This study does not measure the magnitude of gang involvement but seeks to detect those who identify as gang members. All three levels give the appearance of gang membership by explicitly promoting, disseminating, and supporting gang culture, which is consistent with other forms of radicalization (Crone & Harrow, 2011; Moghaddam, 2005; Silber et al., 2007).

Second, the workflow process does not consider the magnitude of the “connection” between gang members. Twitter users can follow public accounts without knowing or interacting with the primary account holder. This study assumes that the connection between gang members, at a rudimentary level, occurs through the implicit communication of digital media, as *followers* are exposed to gang content by following a gang-affiliated Twitter profile. Moreover, some profile connections are more thoroughly represented in the sample population. Whereas profiles with fewer than 200 *followers* had all their connections reviewed, profiles with more than 200 *followers* did not. Future iterations of this research should focus on the level of communication between gang members and use a more comprehensive survey of *followers* to draw.

Finally, this study isolates gang affiliation and location as separate attributes. In reality, gang members belong to both a gang *and* location. Due to data limitations, the interaction effect of these variables is not considered. It would be beneficial for future research to analyze the extent to which both variables collectively impact the geospatial distribution of gang member connections.

4. DATA COLLECTION RESULTS

The workflow process used resulted in the discovery of 1,636 connections between 726 gang and cartel members in 135 cities (18 international), 35 U.S. states, and 13 countries (including the United States). See the appendix for the distribution of gang members by gang affiliation, city, state, and country. Cartels are included in the sample population for two reasons. First, cartels feature prominently in the structure of the gang network (DEA, 2018; NGIC, 2011, 2013, 2015). Second, these connections were formed as part of the discovery process. Except for the Red Command, the discovery of cartels in the sample population is consistent with the findings of Way and Muggah (2016). Connections between gang members and cartels further challenge the location-based gang consensus by highlighting their geospatial diversity. The average activity of Twitter users in this dataset includes 4.22 years and 12,220 tweets, with an average of 38,492 followers. Compared to the median, the years of activity are close to the mean at four years, but the number of tweets and followers are 2,250 and 355, respectively. This suggests that some Twitter accounts in the sample population are more influential than others. Whereas the median provides a better descriptive indicator for this study, the mean provides a snapshot of gang content exposure to Twitter followers. Table 4 provides information on the Twitter profile data discovered during the workflow process.

Table 4. Twitter Profile Descriptions

Average Twitter Profiles Following	901
Median Twitter Profiles Following	453.5
Average Twitter Followers	38,492
Median Twitter Followers	355
Average Year Joined	2013
Median Year Joined	2013
Average Years of Activity	4.22
Median Years of Activity	4
Average Tweets	12,220
Median Tweets	2,250
Average Likes	2,221
Median Likes	267
Gang Members	726
Connections	1636
Gang Total	42
Established Gangs	38
"New" Gangs	5
Cartels	6
Location	
City	135
US	117
Average Population	329,969
Median Population	111,398
International	18
Average Population	1,904,832
Median Population	539,624
State	48
US	35
International	13
Country	13

In the sample population, 27.76% homophily ties (gang members in the same city shared a connection) were detected compared to 72.24% heterophily ties (gang members in different cities shared a connection). The edgelist to calculate the distribution frequency of city ties is used in Model 1 to determine the significance of location by city on gang member connections. Homophily ties detected for gang members in the same state were 35.58%, compared to 64.42% heterophily ties. The edgelist to calculate the distribution frequency of state ties is used in Model 2 to determine the significance of location by state on gang member connections.

Finally, the frequency distribution of gang member connections of the same set was 64.3%, compared to 35.7% heterophily ties to different sets. However, when considering the frequency distribution of gangs from the same primary gang, the homophily and heterophily ties change significantly to 82.04% and 17.96%, respectively. Although some gangs claim the same primary gang affiliation, there is a higher degree of rivalry compared to the gang set. Both Rollin' 60s Neighborhood Crips and Eight Tray Gangster Crips, for example, claim Crip affiliation. However, a dispute in 1979 turned each set into rivals. As the division widened, other Crip sets joined either the Neighborhood Crips (Rollin' Os) or the Gangster Crips (United Gangs, 2020). Therefore, in this study, rather than primary gangs, gang sets were used as nodal attributes. The high percentage of gang members connecting to other members of the same gang indicates that gang homophily may be a strong predictor of shared connections between gang members.

Moreover, increasing connections between members of the same set and members of the same primary gang reinforces the importance of understanding gang relationships at the macro level, an under-researched area of gang studies. Gang sets appear fragmented in the overall network structure but appear to share more connections when gangs connect with members of the same alliance. For example, at the city-level connections of the Gangster Disciples, the study finds that they have 86 heterogeneous ties. As members of the Folk Nation, several of these ties include members within their gang alliance. When consolidating the Gangster Disciples and other sets into their primary gang, the Folk Nation, these ties represent 221 homogenous connections in the sample population. Table 5 shows the frequency distribution of homophily and heterophily ties among gang members by city, state, and gang affiliation.

Table 5. Frequency Distribution of Gang Member Connections (Location & Gang Affiliation)

	Frequency of Homophily Ties	**Frequency of Heterophily Ties**
City	27.76%	72.24%
State	35.58%	64.42%
Gang Set	64.30%	35.70%
Gang Primary	82.04%	17.96%

The distribution of gangs is not limited to highly populated urban areas. Gang members were detected evenly between mid-density (population of 100,000–999,999) and small-density (population of 1,000–99,000) cities at 44.44%, with a low percentage of gang members discovered in high-density (population of 1–3 million) and minuscule-density (population < 1,000) cities. Table 6 shows the frequency distribution of gang members by city size measured by population density.

Table 6. Frequency Distribution of Gang Members Across City Size (Measured by Population Density)

	City Population
High-Density (1–3 million)	5.98%
Mid-Density (100,000–999,999)	44.44%
Small-Density (1,000–99,999)	44.44%
Minuscule-Density (< 1,000)	1.71%

5. EXPONENTIAL RANDOM GRAPH MODEL (ERGM)

ERGMs analyze the substructures of social networks to determine the patterns of relationships between vertices. Robins and 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 Hunter et al. (2008). ERGMs function in a manner quite like 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 Toivonen 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 et al. (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 et al. (2009) for a detailed explanation of ERGM degeneracy).

5.1 ERGM 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. In other words, individuals concentrated in a geographical space are more likely to develop interpersonal relationships. When considering 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 and state that gang members from the same city are likely to form connections. Model 1 includes 634 edges between 335 vertices. By taking the log-odds of the coefficient, we can predict that the probability of a connection forming between gang members from the same city is 59.12% in this model.[2]

When considering 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 ($p < 0.0045$) with a probability of 57.25% that a connection between gang members will form. 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. Moreover, the results suggest that a more diffuse population across a broader geographic space reduces the likelihood of interactions. Defining location on a larger scale appears to contribute to a lower probability of connections forming between gang members when comparing city attributes (59.12%) and state attributes (57.25%). 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 ($p < 0.0002$) and accounts for a 56.78% probability that connections between gang members form based on similarities in gang affiliation. 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. This is an important distinction when considering the probability of connection formation. If gangs were consolidated into their primary gang's affiliation, then it is likely that the probability of connection formation would be greater than 56.78%.

Unlike the three nodal attribute models, Model 4 uses an edge attribute to test the 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 gang member connection, 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

[2] The plogis function in R-Studio generates a log-odds likelihood ranging from 0–1.

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 geographical space. Table 7 provides the ERGM results for the individual effects of attribute homophily (city, state, gang affiliation, and distance [miles]) on gang member connections.

Table 7. ERGM Results: Individual Effects Model 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.	-4.4848	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

6. DISCUSSION

Gang members commit crimes at a higher rate than do non-gang criminal offenders. "Effective use of SNA 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). This is especially more acute in a globalized world where criminal connections have become transnational (Brewster et al., 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 connection formation. Gangs exist within these public spaces, making it unsurprising that city and state attributes help explain gang member connections to some extent. However, location 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 and 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.

In addition to challenging the location-based consensus on gangs, this study suggests further areas of research. For example, some gangs, such as the Grape Street Crips, appear to be more geographically concentrated than other gangs, such as the Gangster Disciples and Five Percenters. Distinctions between gang typologies could help explain the geospatial distribution of gang member connections. Moreover, some gangs are easier to detect on social media than others, allowing for gang-specific studies that examine how micro-level behavioral processes influence macro-level outcomes within a specific subset of gangs. Finally, this study can be used to identify other potential research areas at the local level. The discovery of "new" gangs and their whereabouts provides an opportunity to analyze gang formation and behavior in a contemporary context. Similarly, the sample population includes several cities that are not typically associated with gang activities. The results of the data collection process in this study can expand on work that compares emerging gang cities to established gang cities (Decker et al., 1998). Working with local law enforcement in these cities can help improve our understanding of gangs outside studies that privilege highly populated cities, such as Los Angeles, Chicago, and New York.

References

Anderson, E. (2000). *Code of the street: Decency, violence, and the moral life of the inner city*. WW Norton & Company.

Brantingham, P. J., Tita, G. E., Short, M. B., & Reid, S. E. (2012). The ecology of gang territorial boundaries. *Criminology, 50*(3), 851-885. https://doi.org/10.1111/j.1745-9125.2012.00281.x

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. *International Conference on Conceptual Structures* (pp. 104-117). Springer. https://doi.org/10.1007/978-3-319-08389-6_10

Campbell, A. (1984). The Girls in the Gang. In: Oxford, U.K.: Basil Blackwell.

Cohen, A. K. (1955). *Delinquent boys; The culture of the gang*. New York, NY: Free Press. https://psycnet.apa.org/record/1955-07594-000

Coughlin, B. C., & Venkatesh, S. A. (2003). The urban street gang after 1970. *Annual Review of Sociology, 29*(1), 41-64. http://dx.doi.org/10.1146/annurev.soc.29.101602.130751

Crone, M., & Harrow, M. (2011). Homegrown terrorism in the West. *Terrorism and Political Violence, 23*(4), 521-536. https://doi.org/10.1080/09546553.2011.571556

Curry, G. D. (2000). Self-reported gang involvement and officially recorded delinquency. *Criminology, 38*(4), 1253-1274. https://doi.org/10.1111/j.1745-9125.2000.tb01422.x

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. https://doi.org/10.1002/9781118726822.ch2

DEA. (2018). *National Drug Threat Assessment*. Washington, DC: U.S. Department of Justice.

Décary-Hétu, D., & Morselli, C. (2011). Gang Presence in Social Network Sites. *International journal of cyber criminology, 5*(2), 876-890. http://cybercrimejournal.com/davidcarlo2011julyijcc.pdf

Decker, S. H., Bynum, T., & Weisel, D. (1998). A tale of two cities: Gangs as organized crime groups. *Justice Quarterly, 15*(3), 395-425. https://doi.org/10.1080/07418829800093821

Felson, M. (2006). *Crime and Nature*. Thousand Oaks, CA: Sage Publications. https://books.google.com/books?id=Fz92AwAAQBAJ

Festinger, L., Schachter, S., & Back, K. (1950). *Social pressures in informal groups; a study of human factors in housing*. New York, NY: Harper. https://psycnet.apa.org/record/1951-02994-000

Hagedorn, J. M., & Macon, P. (1988). *People and Folks. Gangs, Crime and the Underclass in a Rustbelt City*. Chicago, IL: Lake View. https://www.ojp.gov/ncjrs/virtual-library/abstracts/people-and-folks-gangs-crime-and-underclass-rustbelt-city

Howell, J. C. (2012). *Gangs in America's Communities*. Thousand Oaks, CA: Sage Publications. https://books.google.com/books?id=5wJEo3YSSHQC

Howell, J. C. (2015). *The History of Street Gangs in the United States: Their Origins and Transformations*. Lanham, MD: Lexington Books.

https://books.google.com/books?id=mc5_CgAAQBAJ
Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of statistical software, 24*(3), 1-29. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2743438/
Jankowski, M. S. (1991). *Islands in the Street: Gangs and American Urban Society*. Los Angeles, CA: University of California Press. http://www.jstor.org/stable/10.1525/j.ctt1ppdz1
Klein, M. W. (1995). *The American Street Gang. Its Nature, Prevalence, and Control*. New York, NY: Oxford University Press.
Klein, M. W., & Maxson, C. L. (2006). *Street Gang Patterns and Policies*. Oxford, UK: Oxford University Press. https://books.google.com/books?id=DXNLojfCtJoC
Knox, G. W. (1994). *An introduction to gangs*. Bristol, UK: Wyndham Hall Press. https://www.ojp.gov/ncjrs/virtual-library/abstracts/introduction-gangs-5th-edition
Leverso, J., & Hsiao, Y. (2020). Gangbangin on the [Face] Book: Understanding Online Interactions of Chicago Latina/o Gangs. *The Journal of Research in Crime and Delinquency, 58*(3), 239-268. https://doi.org/10.1177/0022427820952124
Medina-Ariza, J., Weerman, F. M., Maxson, C. L., Esbensen, F. A., Aldridge, J. A., Medina, J., & Van Gemert, F. (2009). *Eurogang program manual: Background, development and use of the Eurogang instruments in multi-site, multi-method comparative research*. Eurogang Network. https://www.research.manchester.ac.uk/portal/files/45767040/FULL_TEXT.PDF
Miller, W. B. (1958). Lower Class Culture as a Generating Milieu of Gang Delinquency. *Journal of Social Issues, 14*(3), 5-19. https://doi.org/10.1111/j.1540-4560.1958.tb01413.x
Moghaddam, F. M. (2005). The staircase to terrorism: a psychological exploration. *The American psychologist, 60*(2), 161-169. https://doi.org/10.1037/0003-066x.60.2.161
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), 1548. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2481518/pdf/nihms-54862.pdf
Moule Jr, R. K., Pyrooz, D. C., & Decker, S. H. (2014). Internet adoption and online behaviour among American street gangs: Integrating gangs and organizational theory. *British Journal of Criminology, 54*(6), 1186-1206. https://doi.org/10.1093/bjc/azu050
Newcomb, T. M. (1960). Varieties of interpersonal attraction. In D. C. A. Zander (Ed.), *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.
NGIC. (2011). *The 2011 national gang threat assessment*. Washington, DC: Federal Bureau of Investigations.
NGIC. (2012). National youth gang survey analysis.

https://www.nationalgangcenter.gov/Survey-Analysis/Measuring-the-Extent-of-Gang-Problems
NGIC. (2013). *The 2013 national gang report*. Washington, DC: Federal Bureau of Investigations.
NGIC. (2015). *The 2015 national gang report*. Washington, DC: Federal Bureau of Investigations.
Papachristos, A. V., & Kirk, D. S. (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. https://www.google.com/books/edition/Studying_Youth_Gangs/vC-VVeql37MC?hl=en&gbpv=0
Patton, D. U., Eschmann, R. D., & Butler, D. A. (2013). Internet banging: New trends in social media, gang violence, masculinity and hip hop. *Computers in Human Behavior, 29*(5), A54-A59. https://doi.org/10.1016/j.chb.2012.12.035
Patton, D. U., Patel, S., Sung Hong, J., Ranney, M., Crandall, M., & Dungy, L. (2015). Gang violence, crime, and substance use on twitter: A snapshot of gang communications in detroit. *Society for Social Work and Research 19th Annual Conference: The Social and Behavioral Importance of Increased Longevity*. New Orleans, LA: SSWR.
Piquette, J., Smith, C., & Papachristos, A. (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. https://doi.org/10.1007/s10940-013-9206-1
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. https://doi.org/10.1002/9781118726822.ch3
Pyrooz, D. C., & Moule Jr, R. K. (2019). Gangs and social media. In *Oxford research encyclopedia of criminology and criminal justice*. Oxford University Press. 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 American Geographers, 100*(2), 307-326. https://doi.org/10.1080/00045600903550428
Reiss Jr, A. J. (1988). Co-offending and criminal careers. *Crime and justice, 10*, 117-170. https://doi.org/10.1086/449145
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. https://doi.org/10.1214/08-EJS350
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.
Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual review of sociology, 28*(1), 443-478. https://doi.org/10.1146/annurev.soc.28.110601.141114

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *science, 277*(5328), 918-924. https://doi.org/10.1126/science.277.5328.918

Santa Cruz County Gang Task. (2018). Levels of gang involvement. http://www.basta.santacruz.k12.ca.us/levels_of_gang_involvement.html

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. Chicago, IL, US: University of Chicago Press. https://psycnet.apa.org/record/1943-00271-000

Short, J. F., & Strodtbeck, F. L. (1965). *Group process and gang delinquency*. University of Chicago Press Chicago. https://www.ojp.gov/ncjrs/virtual-library/abstracts/group-process-and-gang-delinquency

Silber, M. D., Bhatt, A., & Analysts, S. I. (2007). *Radicalization in the West: The homegrown threat*. Police Department New York. https://www.brennancenter.org/sites/default/files/legacy/Justice/20070816.NYPD.Radicalization.in.the.West.pdf

Stuart, F. (2020). Code of the tweet: Urban gang violence in the social media age. *Social problems, 67*(2), 191-207. https://doi.org/10.1093/socpro/spz010

The R Foundation. (n.d.). What is R? https://www.r-project.org/about.html

Thrasher, F. (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*(2), 272-299. https://doi.org/10.1525/sp.2005.52.2.272

Toivonen, R., Kovanen, L., Kivelä, M., Onnela, J.-P., Saramäki, J., & Kaski, K. (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social networks, 31*(4), 240-254. https://doi.org/10.1016/j.socnet.2009.06.004

Twitter Help Center. (n.d.). Guidelines for law enforcement: What is Twitter? https://help.twitter.com/en/rules-and-policies/twitter-law-enforcement-support#1

United Gangs. (2020). Deuces and Trays. https://unitedgangs.com/deuces2x-and-trays3x-gangster-crips/

Vargas, R. (2016). *Wounded City: Violent Turf Wars in a Chicago Barrio*. Oxford University Press. https://books.google.com.pk/books?id=k2HmCwAAQBAJ

Venkatesh, S. A. (2000). *American Project: The Rise and Fall of a Modern Ghetto*. Harvard University Press. https://books.google.com.pk/books?id=rk7Dd0k_SG8C

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.

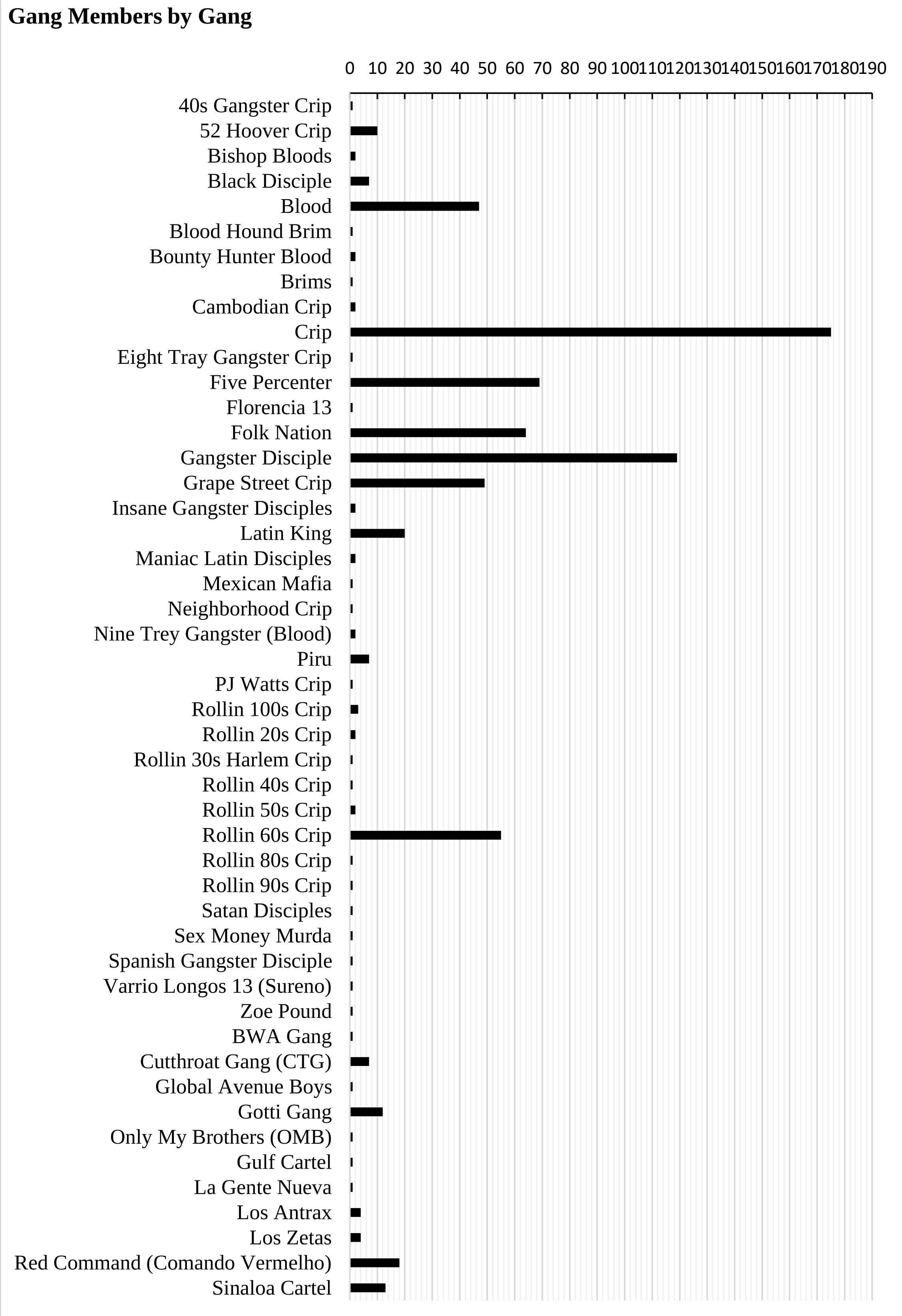
Way, J., & Muggah, R. (2016). Charting out the digital ecosystem of gangs in the US and Mexico. *Small Wars Journal, 11*(12). https://smallwarsjournal.com/jrnl/art/charting-out-the-digital-ecosystem-of-gangs-in-the-us-and-mexico

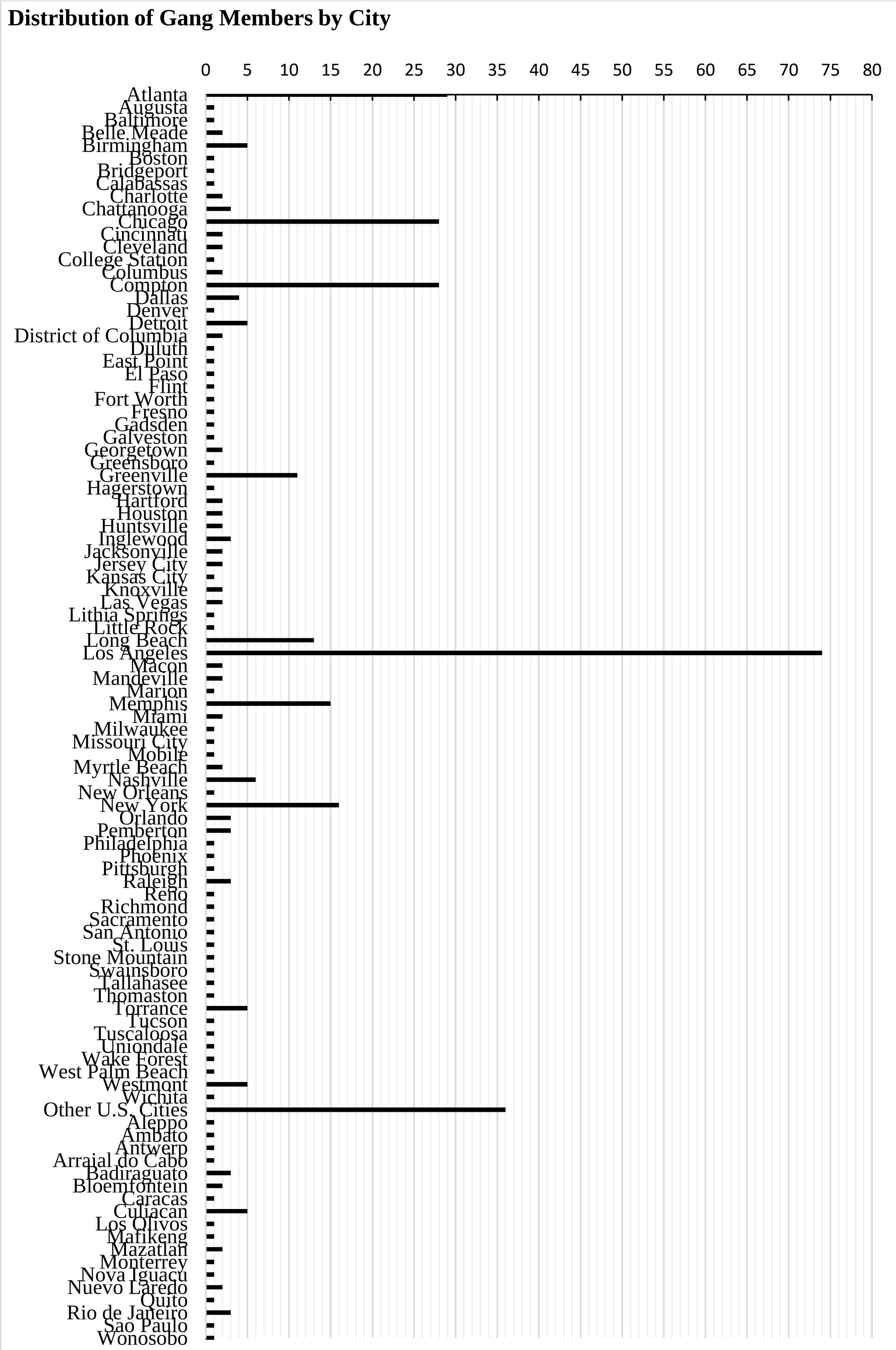
Whittaker, A., Densley, J., & Moser, K. S. (2020). No two gangs are alike: The digital divide in street gangs' differential adaptations to social media. *Computers in human behavior, 110*, 106403. https://doi.org/10.1016/j.chb.2020.106403

Wijeratne, S., Doran, D., Sheth, A., & Dustin, J. L. (2015). Analyzing the social media footprint of street gangs. *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)* (pp. 91-96). IEEE. https://doi.org/10.1109/ISI.2015.7165945

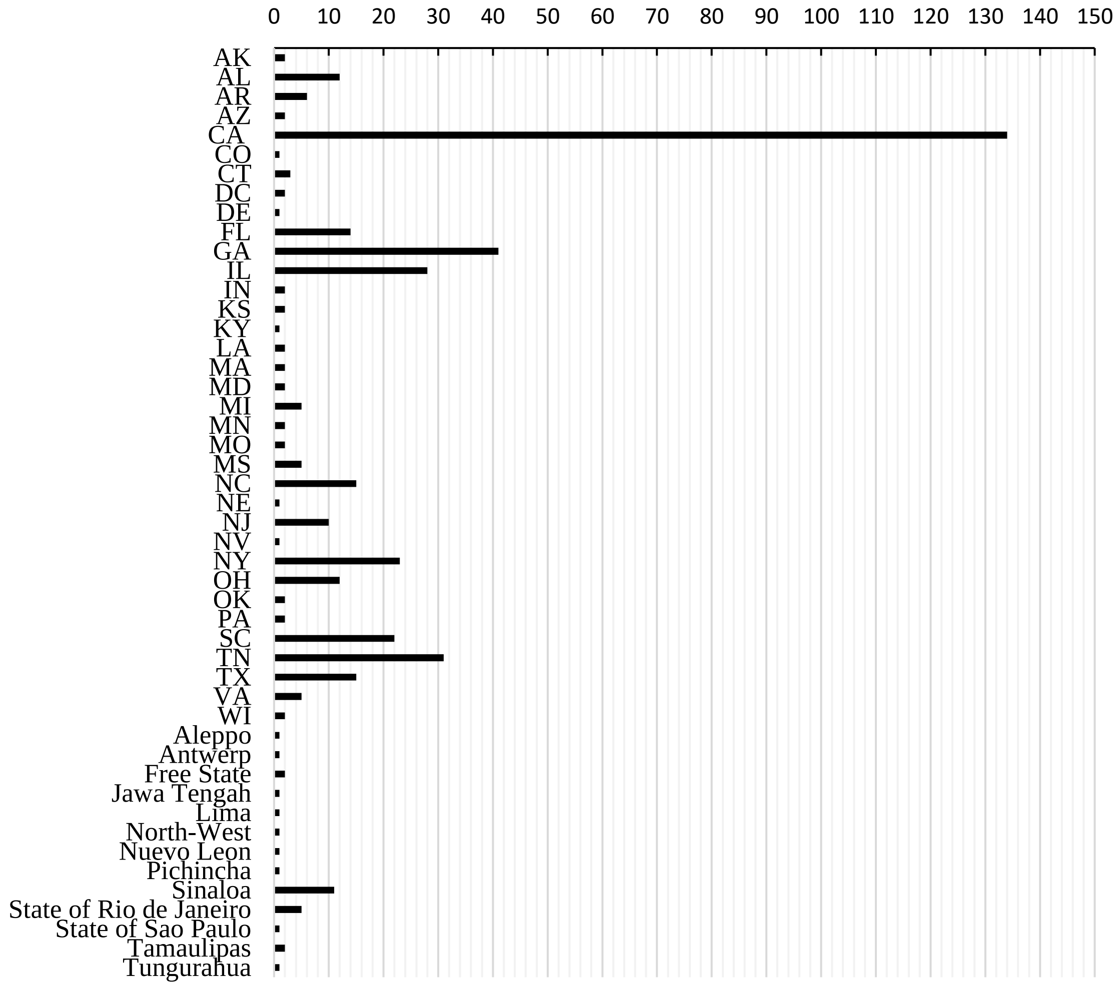
Xu, J., & Chen, H. (2005). Criminal network analysis and visualization. *Communications of the ACM, 48*(6), 100-107. https://doi.org/10.1145/1064830.1064834

Appendix


Gang Members by Gang
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190
40s Gangster Crip
52 Hoover Crip
Bishop Bloods
Black Disciple
Blood
Blood Hound Brim
Bounty Hunter Blood
Brims
Cambodian Crip
Crip
Eight Tray Gangster Crip
Five Percenter
Florencia 13
Folk Nation
Gangster Disciple
Grape Street Crip
Insane Gangster Disciples
Latin King
Maniac Latin Disciples
Mexican Mafia
Neighborhood Crip
Nine Trey Gangster (Blood)
Piru
PJ Watts Crip
Rollin 100s Crip
Rollin 20s Crip
Rollin 30s Harlem Crip
Rollin 40s Crip
Rollin 50s Crip
Rollin 60s Crip
Rollin 80s Crip
Rollin 90s Crip
Satan Disciples
Sex Money Murda
Spanish Gangster Disciple
Varrio Longos 13 (Sureno)
Zoe Pound
BWA Gang
Cutthroat Gang (CTG)
Global Avenue Boys
Gotti Gang
Only My Brothers (OMB)
Gulf Cartel
La Gente Nueva
Los Antrax
Los Zetas
Red Command (Comando Vermelho)
Sinaloa Cartel


Distribution of Gang Members by City
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
Atlanta
Augusta
Baltimore
Belle Meade
Birmingham
Boston
Bridgeport
Calabassas
Charlotte
Chattanooga
Chicago
Cincinnati
Cleveland
College Station
Columbus
Compton
Dallas
Denver
Detroit
District of Columbia
Duluth
East Point
El Paso
Flint
Fort Worth
Fresno
Gadsden
Galveston
Georgetown
Greensboro
Greenville
Hagerstown
Hartford
Houston
Huntsville
Inglewood
Jacksonville
Jersey City
Kansas City
Knoxville
Las Vegas
Lithia Springs
Little Rock
Long Beach
Los Angeles
Macon
Mandeville
Marion
Memphis
Miami
Milwaukee
Missouri City
Mobile
Myrtle Beach
Nashville
New Orleans
New York
Orlando
Pemberton
Philadelphia
Phoenix
Pittsburgh
Raleigh
Reno
Richmond
Sacramento
San Antonio
St. Louis
Stone Mountain
Swainsboro
Tallahasee
Thomaston
Torrance
Tucson
Tuscaloosa
Uniondale
Wake Forest
West Palm Beach
Westmont
Wichita
Other U.S. Cities
Aleppo
Ambato
Antwerp
Arraial do Cabo
Badiraguato
Bloemfontein
Caracas
Culiacan
Los Olivos
Mafikeng
Mazatlan
Monterrey
Nova Iguacu
Nuevo Laredo
Quito
Rio de Janeiro
Sao Paulo
Wonosobo

Distribution of Gang Members by State



Distribution of Gang Members by Country (Excluding the U.S.)