Analyzing the Social Media Footprint of Street Gangs

(Invited Paper)

Sanjaya Wijeratne
Kno.e.sis Center
Wright State University
Dayton OH, USA
sanjaya@knoesis.org

Derek Doran
Kno.e.sis Center
Wright State University
Dayton OH, USA
derek@knoesis.org

Amit Sheth
Kno.e.sis Center
Wright State University
Dayton OH, USA
amit@knoesis.org

Jack L. Dustin
Center for Urban and Public Affairs
Wright State University
Dayton OH, USA
jack.dustin@wright.edu

***Abstract*—Gangs utilize social media as a way to maintain threatening virtual presences, to communicate about their activities, and to intimidate others. Such usage has gained the attention of many justice service agencies that wish to create better crime prevention and judicial services. However, these agencies use analysis methods that are labor intensive and only lead to basic, qualitative data interpretations. This paper presents the architecture of a modern platform to discover the structure, function, and operation of gangs through the lens of social media. Preliminary analysis of social media posts shared in the greater Chicago, IL region demonstrate the platform's capability to understand gang members' social media usage patterns.**

Keywords*—*Gang Activity Understanding, Social Media Analysis, Internet Banging, Platform Design

I. Introduction and Motivation

Street gangs are defined as an affiliated groups of individuals that claim control over physical territory in a community. They express this control by maintaining public threatening presences, and by engaging in violent or criminal activity. They express similar behaviour through online social media as well. Street gangs often express themselves online by sharing provocative, threatening, and intimidating messages publicly in social media [1]. According to a survey 74% of the gang members who participated in it identify themselves as frequent Internet users and had established an online presence to gain respect for their gang [2]. This very high percentage of Internet use is surprising given the illicit activity and negative connotations gangs are affiliated with. If we live in an era of openness, it is surprising that only few segments of the population are more open than 21st-century gang members[1].

The emergence of social media as a tool for gangs to express themselves may be precipitated by the way societies' youngest generations are completely surrounded by technology[2]. Rather than finding public places where like-minded young gang members can congregate [3], they instead exhibit a preference for online public places (e.g., Twitter, Instagram, and YouTube) [4] to express their affiliation, to sell drugs, and to publicize illegal activities. The allure of social media for street gangs is further amplified by the way it quantifies the number of friends, followers, views, and reposts of messages.

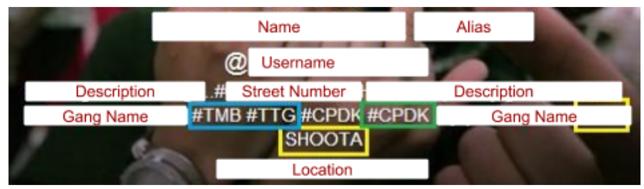

Fig. 1: Twitter profile of a gang member

Such metrics may be used by gangs to approximate their influence, and hence, their perceived power [5].

Law enforcement agencies have recognized the importance of social media analysis to investigate and anticipate gang related crimes. For example, the New York City police department now has over 300 detectives assigned to combat teen violence that is triggered by insults, dares, and threats exchanged on social media[3], and the Toronto police department teaches officers about the use of social media in investigations [6]. Furthermore, a recent survey found 86% of law enforcement officers to use social media at least twice a month in an investigation, and 67% acknowledge its importance as a crime fighting tool [7]. However, an officer's training is generally limited to polices on the use of social media in an investigation, and best practices for post storage and organization [8]. Because no formal methods for the *analysis and interpretation* of social media for crime prevention exists, investigators must create their own ad-hoc methods. For example, a gang member profile shown in Figure 1 could be discovered by manually searching for gang or crime related hashtags (#CPDK, #TTG) and keywords (SHOOTA). An informal process for analyzing social media is not desirable because: (i) the process of tying together information across accounts and posts is difficult to reproduce by others; (ii) the process used is manual and thus unable to gather all accessible data about gang members and their activities; and (iii) it makes crime prevention tasks, where information needs to be collected or updated at regular time intervals, a tedious and inexact process.

To address the above limitations, this paper presents a platform for the automatic analysis of social media messages published by gangs over a geographic region. Using Twitter as a source of data, the framework captures tweets posted by gangs and uses an automated analysis to discover gang structures, functions, and operations. Preliminary results, using

[1] http://www.wired.com/2013/09/gangs-of-social-media/all/
[2] http://www.hhs.gov/ash/oah/news/e-updates/eupdate-nov-2013.html
[3] http://bit.ly/1m80pZ9

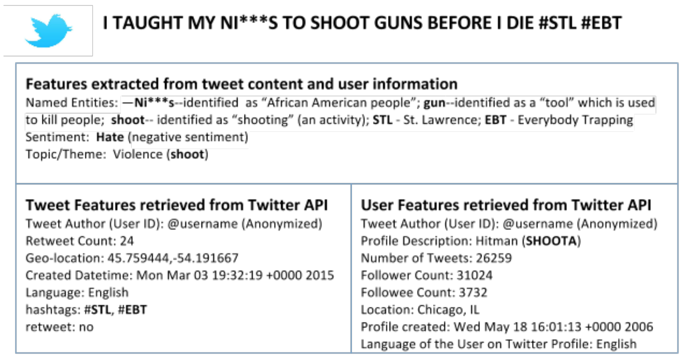

Fig. 2: Metadata and content analysis of a gang tweet

multiple datasets of social media posts published by gang members, reveal the types of thematic, sentiment, and network knowledge the platform can automatically extract.

II. Related Research

Gang violence is a well studied topic in Social Science research, dating back to 1927 [9]. Historical reviews portray American gangs emerging along racial and ethnic lines and developing into organizations designed for illegal business [10]. Even though gang violence is a well studied topic, "cyber" or "Internet banging" [1] has not been fully explored [11]. The presence of gang violence depicted online and realtime violence caused by online gang activities deserves our attention.

Research has shown that street gangs use social media mainly to post videos depicting their illegal behaviours, watch videos, threaten rival gangs and their members, display firearms and money from drug sales [12], [13], [14]. For example, a recent article reported an interview with a known gang member from Chicago who explained that social media is a way to publicize their weapons, so as to protect themselves against rival gangs[4]: *"Someone says something to me on Facebook, I dont even write a word. The only thing I do is post my 30-popper, my big banger"*. As another example, studies confirm that about 45% of gang members now participate in at least one form of online offending such as selling drugs online, threatening or harassing individuals, posting violent videos online or attacking someone on the street for something they said online [15].

Researchers have undertaken few efforts to build cyber systems for monitoring street gang activities using social media. Early work by Decary *et al.* used a small numbers of keywords to identify gang related posts in order to measure the level and type of information they share online [16], [12]. Those studies also reviled the significant increase in Internet use by street gang members in recent years. Patton *et al.* [14] explored distinct ways with which gang members use Twitter, including to grieve, argue with rivals, disclose weapons, and display illicit substances and alcohol. A major drawback of these efforts are the authors' decision to identify social media posts by searching for those that satisfy a small list of keywords related to violence, drugs, and weapons. Such filtering may significantly add bias to the topics and modes of use discovered. In order to address this limitation, our work instead considers geo-tagged tweets collected from the Twitter Streaming API[5] across "hot-spot" neighborhoods where gang activities concentrate[6].

Instead of focusing on a small set of keywords and exploratory analysis of how gangs generally use social media, the platform proposed in this paper focuses on understanding and analyzing the posts of specific gangs, operating in specific neighborhoods, in ways that may be useful to design a specific judicial services program. The platform discovers keywords and phrases by analyzing the frequency of terms made in geo-tagged tweets within specific locations. It also considers the terms used in Tweets made from gang member accounts to discover the topics they discuss, and to build a location-specific term list for gang tweet searching. Finally, the platform is able to carry out complex analysis on the content of tweets, including sentiment and emotion analysis. Analysis of metadata about a social media account is also used to build networks of relationships that visualize the structure and social operation of gangs.

III. Platform Requirements and Architecture

Social media profiles, friendship or follower/followee relationships between accounts, and the content of posts all contain information relevant to understanding gang operations, locations, functions, and participants. Based on our studies, we identify the following requirements for a social media powered platform to discover the structure, function, and operation of gangs:

1. Monitor negative community effects of gang activities. This monitoring requires finding entities relevant to gang activities in posts as well as references to communities. Such information is essential to measure the impact of gangs, and to design community support programs countering this effect.

2. Discover opinion leaders who influence the thoughts and actions of other gang members. This includes tracking the creation and diffusion of information across a network of related social media accounts. Discovering opinion leaders, their network and the reach of their influence can help to understand the social structure of gangs, which may be essential during designing crime prevention programs.

3. Evaluate the sentiment of posts targeting communities, locations, and groups (including rival gangs). Such sentiment analysis may be useful to identify rivalries, and may be used to anticipate crimes.

4. Monitor community and gang responses to community support programs. This monitoring may be useful in the evaluation of interventions and other programs executed by a town or city.

To achieve these requirements, a platform must be able to analyze the spatio-temporal-thematic (where, when, what), people-content-networking (who and how), and emotion-sentiment (perceptions and impact) dimensions of social media posts. The necessary capabilities for analysis along each of these dimensions are elaborated below.

Spatio-Temporal-Thematic analysis: Monitoring the effects of gang activities and intervention programs (e.g., gun violence, drug dealing, shooting) based on the content of gang

[4] http://www.wired.com/2013/09/gangs-of-social-media/all/

[5] https://dev.twitter.com/streaming/overview

[6] http://bit.ly/19DxIVq

social media posts require the platform to process them using machine learning, natural language processing, and semantic technologies [17] that infer associations between words and concepts. It also requires the integration of knowledge bases capable of translating slang terms, since gangs typically use their own shorthand to describe rivals, crimes, and violent activities. For example, Figure 2 shows a tweet posted by a known gang member who operates in the Chicago area. Interpreting the tweet requires a knowledge base such as Urban Dictionary[7] or the Internet Slang acronym dictionary[8] to map terms to concepts (STL: St. Lawrence, EBT: Everybody Tripping). Further analysis of content may reveal the gang the user belongs to, their role, and skill set. For example, the profile of the Twitter user shown in Figure 1 indicates his affiliation (gang names), the location where he operates (street number), his skill set (e.g. TTG: Trained to Go), his role in the gang (Shoota: Hitman/Drug dealer), and groups he associates with (CPDK: Chicago Police Department Killer)[9].

People-Content-Network analysis: Discovering opinion leaders and the social structure of gangs can be inferred through network analysis of friendship or follower/followee relationships across member profiles. The platform should therefore be capable of querying a social media service's API or scraping its Web pages to infer this network of relationships. Social network analysis methods may then be applied to discover social connections such as groups that discuss similar concepts [18], leaders or influencers [19], and patterns of coordination [20].

Emotion-Sentiment analysis: Evaluating the emotion (a person's feelings when a post is published) and sentiment (the degree to which the post exhibits positive or negative emotion) of posts targeting groups of others, the law, and places may help to anticipate crimes and other events. Such analysis requires the platform to possess tailor made sentiment extraction techniques adapted to gang-related keywords and expressions. These adaptations may automatically classify social media posts into those targeting different groups with expressions of positive, negative or neutral attitude [21], [22].

The architecture of a platform exhibiting the above capabilities is presented in Figure 3. Its components are divided into four stages: (1) Data collection and filtering; (2) Data processing; (3) Data access tools for exploration and visualization; and (4) Data analysis and interpretation. The architecture focuses on the analysis of the Twitter social media platform due to its popularity and widespread use [16]. The first stage continuously collects tweets from the accounts of gang members, across geographic spaces, and based on keywords learned by analyzing previously captured tweets. It will capture high quality topic-descriptors (e.g. hashtags such as #BDK, #GDK) used by gang members to disseminate information about specific activities. A slang term dictionary based on online dictionaries will incorporate the knowledge of experts familiar with slang terms used by gang members into the filtering step. The second data processing stage adapts entity recognition with the slang dictionary to disambiguate the context with which entities (e.g. a location) are referred to within. It will also perform sentiment analysis, user location estimation, and social network analysis using machine learning, natural language processing, and network analysis tools. The third data access stage will store the analysis results from the previous stage into a database. Visualization and information retrieval front-ends help an analyst study the results in the database and offer comparisons over time and geography. The separation of the data processing and access tool stages enable developers to add their own custom search and visualization tools; thus, data summaries and queries can be customized for specific agencies and tasks.

IV. Preliminary Findings

This section offers a preliminary analysis of social media posts shared by gang members across the Chicago, IL region. The analysis was generated using a prototype version of the platform described above. We first discuss data collection for the analysis, and show initial results from thematic, sentiment, and network analysis. When reporting our findings, we have removed personally identifiable information such as gang member names and slightly altered the tweet text in a way that it does not change the original meaning so that the poster cannot be identified.

A. Data collection

Since development of an automated method for collecting gang related tweets is still a work in progress, we consider different strategies to obtain tweets from users associated with gangs. The first approach uses Followerwonk[10] with pre-identified keywords that members from Gang A[11] mostly used to identify themselves on Twitter. Gang A resides in South Side of Chicago where their members are associated with the Gangster Disciples gang. We selected Gang A because they are well known for their gang related activities on Twitter. The second approach found tweets that contained either #BDK, Gangster Disciples, #GDK, Black Disciples or one of the 28 keywords used in [16]. Combined, these two methods lead to a dataset of over 105,447 *gang-related* tweets collected during a 10 day time period in March 2015. We additionally collected all tweets sent within a geographic boundary of 10 neighborhoods in South Side of Chicago known for gang related activities[12], namely South Landale, North Landale, West Elsdon, Gage Park, West Lawn, Chicago Lawn, New City, Humboldt Part, Logan Square and Belmont Cragin. In addition, 383,656 *location-related* tweets were collected across these neighborhoods during the 10 day time period.

Future developments for collecting data will strongly consider the content of known gang member profiles to build databases of keywords and phrases related to gang activity posts. This is because, based on manual analysis of the content within the datasets collected, knowing that a tweet is from a gang member can provide additional contextual clues to understand the message conveyed by tweet content. For an example, consideration of the fact that a gang member published a tweet with the number '7414' makes it more likely to refer to GDN[13]

[7] http://www.urbandictionary.com/

[8] http://www.internetslang.com/

[9] Gang names and street number are not shown as to protect the identity.

[10] https://followerwonk.com/bio

[11] Gang name has been anonymised due to privacy concerns.

[12] http://bit.ly/19DxIVq

[13] http://www.urbandictionary.com/define.php?term=7.4.14

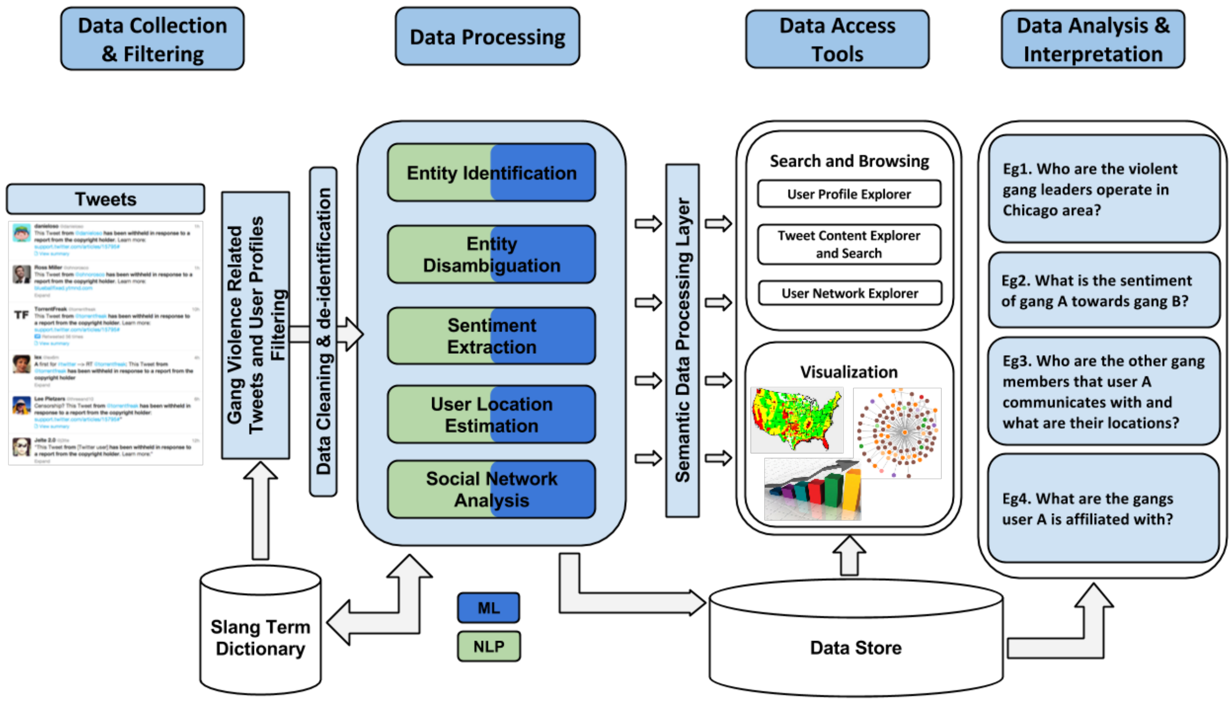

Fig. 3: Architecture of the proposed platform

(7th, 4th and 14th letters of the alphabet form GDN, which refers to the Gangster Disciple Nation) than a house address or digits of a phone number. To automatically identify gang member profiles, this study briefly explored the features of gang members' Twitter profiles that may be useful to train a classifier. We collected the profile descriptions from 91 Twitter users in our gang-related datasets, who were members of Gang A, and compared them with the profile descriptions obtained from the authors of tweets in the location-related dataset. Figure 4 compares word clouds based on the frequency with which unigrams from each type of profile were used. When constructing these word clouds, we removed stopwords and performed word stemming using the typically used Porter's Stemmer[14]. We also removed the seed words used to collect gang member profiles of Gang A from the words list and anonimized any personally identifiable names. Comparison of the word clouds shown in Figure 4 clearly reveal how gang members consistently use words associated with their gangs (Fallen gang members, Gang names, Curse words - Fuck, Shit, Fto etc., Gang related slang - Nolackin, CPDK etc.) to identify themselves. On the other hand, profiles from the location-based dataset use general terms relating to where they are from (e.g. - Chicago), their interests (e.g. - Music, Sports), roles in their families (e.g. - Mom, Father, Husband) or their occupations (e.g. - Student, Writer, Director, Artist etc.). Such lexical features may therefore be of great importance to automatically find gang member profiles. When collecting tweets based on keywords, we found out that some keywords bring noisy (unwanted) tweets due to the fact that those words had multiple meanings. Future enhancements to keywords based collection of tweets will focus on implementing automated methods to filter out noisy tweets [23].

[14] http://tartarus.org/martin/PorterStemmer/java.txt

(a) Gang-related tweets (b) Location-related tweets

Fig. 4: Comparison of terms in user profiles

B. Spatio-Temporal-Thematic Analysis

We asked our prototype platform to identify tweets containing GPS coordinate information in its metadata or with location information included in the poster's profile from the gang-related tweet dataset. It was able to identify geographic location information within 3.62% of these tweets. Of those identified tweets, the platform highlighted interesting examples where people threaten their rival gang members when looking at tweets from Chicago area. For an example, "lemme hear you say Gang X and we finna murder you" is a tweet from a gang member from Chicago who threatens members from Gang X. "On location_name We Drillin Fuck Da Opps" is another tweet originated from a Chicago south side neighborhood that threatens the opponents in general (Fuck Da Opps). "@user1 @user2 check out this 7414 track, url_to_file" is another example discovered, but requires the use of a knowledge base to decode the slang and understand that the tweet talks about

Fig. 5: Follower network of 91 Gang A members

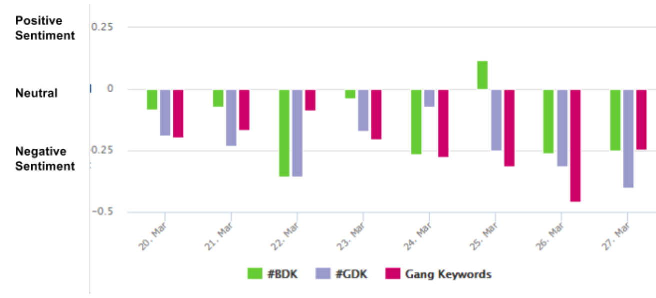

Fig. 6: Sentiment of gang member tweets

Gangster Disciples Nation (7-4-14). These examples shows the additional knowledge that one could glean out of tweets by associating the space (location) and theme (gang violence) of posts with entities and slang terms in tweet text. Had we not identified that the above tweets were originated from Chicago or some of the entities/slang terms in those tweets are gangs from Chicago (using entity identification and disambiguation components powered by DBpediaSpotlight[15] and gang related slang terms extracted from Urban Dictionary[16]), we would not be able to understand the messages they conveyed. This shows how location information found in tweets (spatial) and gang related (theme) knowledge available in online slang term dictionaries can be used towards understanding gang members' chatter in social media at a given time (temporal).

C. *People-Content-Network Analysis*

The friend and follower relationships of Twitter accounts can also be studied by our prototype. For this preliminary analysis, we considered the friend and follower connections of the 91 gang members whose posts are contained in the gang-related dataset. A network of these relationships was constructed by assigning a directed edge from one node A to another B if user A is a follower of B. Figure 5 visualizes the structure of this network as drawn by the network analysis tool Gephi[17]. In total, this network consists 1,322 nodes, 19,539 edges and has an average degree of 14.77. This tells us a Twitter user in our dataset is at least connected to 15 other Twitter users. A small number reflecting the exclusivity of gang follower activity on Twitter may suggest that follower/followee represent very strong offline bonds, and be potential witnesses or informants when a user is under investigation. Furthermore, we applied a modularity based community detection algorithm in Gephi which identified 72 different clusters with a modularity of 0.859. This tells us that Twitter users in the 72 communities are well connected among each other. Further we noticed that each cluster has an average of 19 users in them. This observation further strengthens our earlier hypothesis based on small number of followers per user. Since networks with high modularity have dense connections, and the average number of users in a cluster is 19, we may find these strong online bonds in the real world too. We noticed that most groups formed consisted of at least few gang members. When we studied the nodes in communities manually, we identified some nodes belonging to Twitter users who could have some possible affiliation with gangs based on what they tweet, but we could not verify their affiliations as they have not explicitly defined those in their user profiles. This will be an interesting research area we would like to explore in future.

D. *Sentiment-Emotion Analysis*

Finally, we explored the sentiment and emotions of tweets as quantified by our platform, which are based on algorithms [21], [22]. The sentiment analysis algorithm implemented in our platform tries to identify sentiment of a tweet related to a target entity (target dependent sentiment) where target entities being the entities found in the tweet [21]. Our platform can identify seven emotions: joy, sadness, anger, love, fear, thankful and surprise as discussed in [22]. As shown in Figure 6, the sentiment analysis results by our platform reports most tweets to exhibit negative sentiment, including ones such as "He took @user1 #Gang_1K He killed @user2 #Gang_2K he murder @user3 #CPDK #RIP". The general negative sentiment assigned to gang tweets may be related to the excessive use of curse words, which our sentiment analysis algorithm maps to negativity. Furthermore, gang members have a penchant for using social media to share anger or sad emotions in the threats they express [14]. Even though sentiment and emotion analysis gave us negative sentiment and anger emotion for most tweets, combining this data with known entities in tweets (e.g. - gangs) can lead to interesting findings. For an example, once the above tweet's sentiment (negative), emotion (anger) and entities are identified (Gang_1 and Gang_2 as gangs, CPD as Chicago Police Department), we can understand that the user who posted it hates the identified entities. This tells us that this Twitter user belongs to one of the rival gangs of Gang_1 and Gang_2 (The letter "K" after Gang_1 and Gang_2 stands for killer, so the terms read as

[15] http://spotlight.dbpedia.org/
[16] http://www.urbandictionary.com/)
[17] http://gephi.github.io/

“Gang_1 Killer” and “Gang_2 Killer”). On the other hand, sentiment analysis algorithms implemented in our platform work well with tweets in general; however, the excessive use of curse words in gang members’ tweets may limit our platform’s ability to analyze sentiment in tweets. We plan to customize the algorithms for sentiment analysis in the future to better identify sentiment of gang members’ tweets by adjusting the weights given for curse words when deciding a tweet’s sentiment.

V. Conclusion and Future Work

This paper introduced a platform for the automatic analysis of social media posts to understand the structure, function, and operation of street gangs. We identified (i) monitoring negative community effects of gang activities, (ii) discovering opinion leaders who influence the thoughts and actions of other gang members, (iii) identifying the sentiment and emotion of posts targeting communities, locations, and groups (including rival gangs) and (iv) monitoring community and gang responses to community support programs to be the basic requirements of such platform and discussed how spatio-temporal-thematic, people-content-network, and sentiment-emotion analyses may be used to meet those requirements. Analysis of tweets made in regions that were known to contain significant gang activity demonstrate the capabilities of the platform as a tool for gang activity identification and monitoring. Specifically, the analysis identified the hateful messages exchanged among gang members threatening their rival gangs, clues to identify a user’s gang affiliations (if there are any) based on what he tweets, and follower network analysis revealed insights about who are the possible gang members associated with a known gang member. Our study also reviled lexical features one could extract from gang members’ tweets or Twitter profile descriptions that can be of great importance to build systems to automatically identify social media profiles of gang members.

The implementation of the platform presented in this paper is still at an early stage of development. Future work will improve the data collection system, featuring classifiers that automatically identify gang member profiles, algorithms to filter unwanted tweets and will expand our slang dictionaries using non-traditional knowledge bases such as HipWiki[18]. The enhancements to the Sentiment-Emotion analysis of our framework, as discussed above, will also be explored. Evaluation of the effectiveness of our platform will be based on empirical methods, where we will develop methods to evaluate the gang member and network identification, in future work.

VI. Acknowledgements

We are grateful to Gary Alan Smith for helping us with data pre-processing. We also like to thank the anonymous reviewers for their useful comments.

References

[1] D. U. Patton, R. D. Eschmann, and D. A. Butler, “Internet banging: New trends in social media, gang violence, masculinity and hip hop,” *Computers in Human Behavior*, vol. 29, no. 5, pp. A54–A59, 2013.

[2] J. E. King, C. E. Walpole, and K. Lamon, “Surf and turf wars online-growing implications of internet gang violence,” *Journal of Adolescent Health*, vol. 41, no. 6, pp. S66–S68, 2007.

[3] P. E. Owens, “Natural landscapes, gathering places, and prospect refuges: Characteristics of outdoor places valued by teens,” *Children’s Environments Quarterly*, pp. 17–24, 1988.

[4] M. Ito, S. Baumer, M. Bittanti, R. Cody, B. Herr-Stephenson, H. A. Horst, P. G. Lange, D. Mahendran, K. Z. Martínez, C. Pascoe *et al.*, “Hanging out, messing around, and geeking out,” *Digital media*, 2010.

[5] R. Wolff, J. McDevitt, and J. Stark, “Using social media to prevent gang violence,” Northeastern University Insitute on Race and Justice, Tech. Rep., 2011.

[6] P. E. R. F. (PERF), “Social media and tactical considerations for law enforcement,” United States Office of Community Oriented Policing Services (COPS) and United States Department of Justice, Tech. Rep., 2013.

[7] R. M. Divison, “Social media use in law enforcement: Crime prevention and investigative activities continue to drive usage,” LexisNexis, Tech. Rep., 2014.

[8] J. Brunty, L. Miller, and K. Helenek, *Social media investigation for law enforcement*. Routledge, 2014.

[9] F. M. Thrasher, *The gang: A study of 1,313 gangs in Chicago*. University of Chicago Press, 1963.

[10] R. L. Bonn, *Criminology*, ser. McGraw-Hill Series in Criminology and Criminal Justice. McGraw-Hill Publishing, 1984.

[11] S. J. R. M. P. S. Patton, D.U. Hong, “Social media as a vector for youth violence: A review of literature,” *Computers in Human Behavior*, 2014.

[12] C. Morselli and D. Décary-Hétu, “Crime facilitation purposes of social networking sites: A review and analysis of the cyberbangingphenomenon,” *Small Wars & Insurgencies*, vol. 24, no. 1, pp. 152–170, 2013.

[13] S. Decker and D. Pyrooz, “Leaving the gang: Logging off and moving on. council on foreign relations,” 2011.

[14] D. U. Patton, “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, jan 2015, https://sswr.confex.com/sswr/2015/webprogram/Paper23817.html.

[15] D. C. Pyrooz, S. H. Decker, and R. K. Moule Jr, “Criminal and routine activities in online settings: Gangs, offenders, and the internet,” *Justice Quarterly*, no. ahead-of-print, pp. 1–29, 2013.

[16] D. Décary-Hétu and C. Morselli, “Gang presence in social network sites,” *International Journal of Cyber Criminology*, vol. 5, no. 1, 2011.

[17] D. Gruhl, M. Nagarajan, J. Pieper, C. Robson, and A. Sheth, *Context and domain knowledge enhanced entity spotting in informal text*. Springer, 2009.

[18] H. Purohit, Y. Ruan, A. Joshi, S. Parthasarathy, and A. Sheth, “Understanding user-community engagement by multi-faceted features: A case study on twitter,” in *WWW 2011 Workshop on Social Media Engagement (SoME)*, 2011.

[19] H. Purohit, J. Ajmera, S. Joshi, A. Verma, and A. P. Sheth, “Finding influential authors in brand-page communities.” in *ICWSM*, 2012.

[20] H. Purohit, A. Hampton, V. L. Shalin, A. P. Sheth, and J. M. Flach, “Framework for the analysis of coordination in crisis response,” 2012.

[21] L. Chen, W. Wang, M. Nagarajan, S. Wang, and A. P. Sheth, “Extracting diverse sentiment expressions with target-dependent polarity from twitter.” in *ICWSM*, 2012.

[22] W. Wang, L. Chen, K. Thirunarayan, and A. P. Sheth, “Harnessing twitter” big data” for automatic emotion identification,” in *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. IEEE, 2012, pp. 587–592.

[23] S. Wijeratne and B. R. Heravi, “A keyword sense disambiguation based approach for noise filtering in twitter,” 2014.

[18] http://www.hipwiki.com/