

Signals Revealing Street Gang Members on Twitter

Lakshika Balasuriya, Sanjaya Wijeratne, Derek Doran, and Amit Sheth

Kno.e.sis Center, Wright State University, Dayton, Ohio, USA
{lakshika, sanjaya, derek, amit}@knoesis.org

Abstract. We study the problem of automatically finding gang member profiles on Twitter. We outline a process to curate one of the largest sets of verifiable gang member profiles that has ever been studied. A review of these profiles establishes differences in the language, images, YouTube links, and emoji features gang members use compared to the rest of the Twitter population. We generate word embeddings that translate these features into a real vector format amenable for machine learning classification and use them to train a series of supervised classifiers. Our classifiers achieve promising F_1 scores with low false positive rates.

Keywords: Street Gangs on Twitter, Gang Activity Understanding

Gang affiliates have joined the masses who use social media to share thoughts and actions publicly. Interestingly, they use this public medium to express their illegal actions, to intimidate others, and to share outrageous images and statements [2]. Their tweets may thus be useful to law enforcement agencies for uncovering clues about recent crimes or to anticipate ones that may occur. Finding these posts, however, requires a method to discover gang member Twitter profiles. We developed a method based on word embeddings to identify street gang member profiles on Twitter [3]. We curated what may be the largest dataset of gang member profiles on Twitter that consists of 400 authentic gang member profiles and 2,865 non-gang member profiles [1]. We then analyzed the text in their tweets and profile descriptions, their emoji use, their profile images, and music interests (see Figure 1). We found that 5.72% of all words posted by gang members were curse words, which is nearly five times more than the average curse word usage on Twitter among the general population. We found that the terms *rip* and *free* appear in approximately 12% of all gang member Twitter profile descriptions, which suggests that gang members use their profile descriptions as a space to grieve for their fallen or incarcerated gang members. We also found that 51.25% of the gang members in our dataset had a tweet that linked to a YouTube video and 76.58% of them were related to gangster rap music. Emoji analysis revealed that the fuel pump emoji was the most frequently used emoji by the gang members, where it was often used in the context of selling or consuming marijuana. The pistol emoji was the second most frequent emoji, which was often used with the police cop emoji in an ‘emoji chain’ to reflect anger at police officers. We found that 32.25% of gang members had chained together the police and the pistol emoji, compared to just 1.14% of non-gang

members. Gang members were often seen holding or pointing weapons in groups displaying gangster culture, or showing off graffiti, hand signs, tattoos, and bulk cash in their profile pictures.



Fig. 1. Gang Member and Non-gang Member Twitter Data Analysis.

To automatically discover gang member profiles, we converted all non-textual features such as emoji and images into textual formats, pre-processed the text, and then fed all unigrams collected from Twitter profiles to Word2Vec tool and trained a Skip-gram model [3]. Each word was then represented using a vector of size 300 learned by the Skip-gram model. The word vectors were combined to get a final vector of size 300 which we used to represent a Twitter profile. The final vector was then used to train a series of machine learning classifiers. When evaluated using 10-fold cross validation, the Logistic Regression classifier performed the best with an F_1 score of 0.7835 when identifying gang members [3]. We plan to extend this work to predict social media wars that could end up in real world fights by analysing hateful tweets posted by gang members.

Acknowledgement. This work was supported by Grant 2014-PS-PSN-00006.

References

1. Balasuriya, L., Wijeratne, S., Doran, D., Sheth, A.: Finding street gang members on twitter. In: IEEE/ACM ASONAM 2016. vol. 8, pp. 685–692. San Francisco, CA, USA (08 2016)
2. Wijeratne, S., Doran, D., Sheth, A., Dustin, J.L.: Analyzing the social media footprint of street gangs. In: IEEE ISI 2015. pp. 91–96 (May 2015)
3. Wijeratne, S., Balasuriya, L., Doran, D., Sheth, A.: Word embeddings to enhance twitter gang member profile identification. In: IJCAI Workshop SML 2016. pp. 18–24. CEUR-WS, New York City, NY (07 2016)