Title: INVESTIGATING CRIME-TO-TWITTER RELATIONSHIPS IN URBAN ENVIRONMENTS - FACILITATING A VIRTUAL NEIGHBORHOOD WATCH

Abstract:

Social networks offer vast potential for marketing agencies, as members freely provide private information, for instance on their current situation, opinions, tastes, and feelings. The use of social networks to feed into crime platforms has been acknowledged to build a kind of a virtual neighborhood watch. Current attempts that tried to automatically connect news from social networks with crime platforms have concentrated on documentation of past events, but neglected the opportunity to use Twitter data as a decision support system to detect future crimes. In this work, we attempt to unleash the wisdom of crowds materialized in tweets from Twitter. This requires to look at Tweets that have been sent within a vicinity of each other. Based on the aggregated Tweets traffic we correlate them with crime types. Apparently, crimes such as disturbing the peace or homicide exhibit different Tweet patterns before the crime has been committed. We show that these tweet patterns can strengthen the explanation of criminal activity in urban areas. On top of that, we go beyond pure explanatory approaches and use predictive analytics to provide evidence that Twitter data can improve the prediction of crimes. Keywords: Decision Support, Predictive Analytics, Social Media, Big Data.

Authors: Johannes Bendler, Tobias Brandt, SebastianWagner, Dirk Neumann

Link: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.666.4085&rep=rep1&type=pdf

Summary:

The ability to crowdsource information pertinent to crimes is very useful to authorities and there are even crime watch crowdsourcing platforms in existence today such as postacrime.com, spotcrime.com and CrimeReports. The data from these sites are useful to both authorities and data analysts that can use this information to find crime patterns. The rough location for where the post was sent can be determined by the social network provider or by geotags from the users phone. This is useful because if there are a large amount of tweets in a certain area, an event of some importance is most likely occurring in that area. This paper tries to find out whether the relationship between tweets and crime exists and is usable, whether anything can be inferred by the location/concentration of tweets, and whether twitter data can be used to not only explain criminal activities but predict it. This paper tries to predict and explain crimes in urban areas through tweet volume. It is inferred again that social activity is highly tied to the location of the user when the tweet is sent and if many people are tweeting in the same location around the same time then something of some importance is most likely happening. Tweets and crime data are linked by time and location in this study.

For this study, tweest and crime data were collected at Market Street in San Fransisco from August 15, 2013 to November 14, 2013 and this data was split into hourly blocks. There were a lot of variables though, because people are prone to tweet more on certain days of the week and not as much on others, and the same goes for criminal activity. Each tweet is

represented by a 3-tuple with latitude, longitude and time block. Crime data is represented as a 4-tuple with latitude, longitude, time block and type of crime. It was found that crime would decrease when there were more tweets and a Poisson regression was used to verify a correlation between tweets and criminal activity. Only 1% of the tweets collected had geotags so the study was limited in that way but in future research, the authors wish to utilize points of interest as well.