

Cartographic Visualization Methods of Geo-Tagged Social Media Data During Civil Unrest

Edwin Roman

Master of Science in Geographic Information Science (MSGISci)
Department of Geography, California State University, Long Beach



Introduction

On April 27, 2015, downtown Baltimore, Maryland (figure 1) spiraled into complete chaos as protesters clashed with Baltimore riot police hours following the funeral of Freddie Gray, a 25 year old African American who died while in police custody earlier in the month. Protesters, many of them high school students, circulated an image on social media calling for a citywide "purge" to start at the Mondawmin Mall (a reference to the 2013 movie "The Purge" in which crime became legal for 12 hours in the US). Similar confrontations between protesters and police spread to other parts of the city that evening. This project assesses how social media data can be spatially analyzed to uncover patterns of sentiment during times of civil unrest by specifically looking at the April 27th riot in Baltimore. It was determined that identifying and visualizing negative sentiment by using geo-tagged social media data can reveal areas in which negative sentiment is intensifying. These "hotspot" areas can signify specific locations where violent behavior or other activity is occurring. Law enforcement agencies and intelligence units can utilize this information to effectively deploy assets during times of civil unrest.

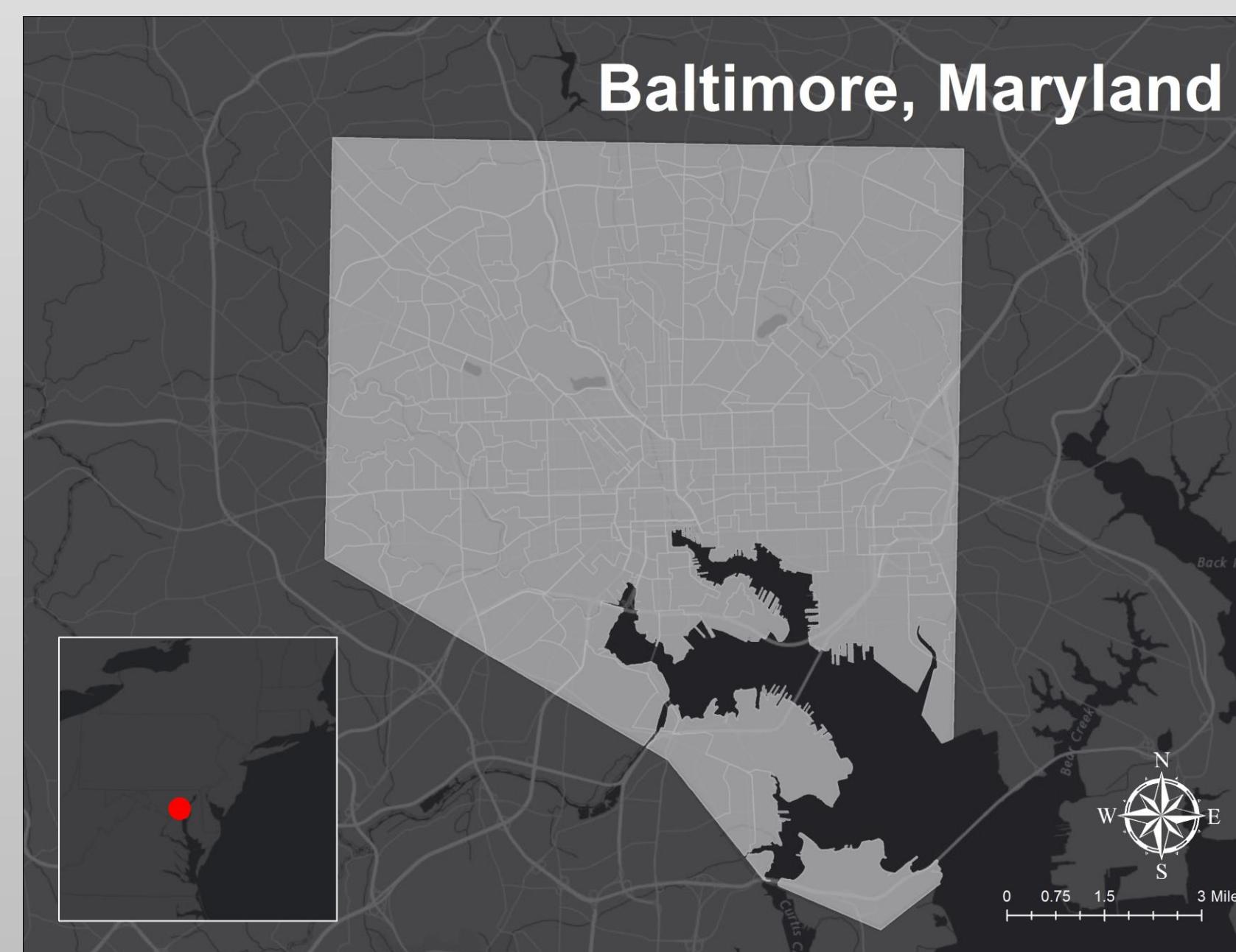


Figure 1. On April 27, 2015 riots erupted in Baltimore, Maryland

Data and Data Sources

Access to social media activity in Baltimore was provided by Geofeedia, a subscription base platform, which monitors and collects geo-tagged social media data. A subscription to such platform can cost agencies between \$75,000 and \$100,000 a year. In addition, a vetting process is required to ensure the ethical use of the platform. Thus, the accessibility to the Geofeedia platform is almost always limited to the public safety sector, corporate security teams, and business intelligence units. Additional data was found on Open Data-City of Baltimore website.

Table 1. List of data and data sources used in the project

Dataset	Source	Type
Baltimore City Boundaries	Open Data-City of Baltimore	Polygon
Baltimore Crime Data	Open Data-City of Baltimore	Point
Geo-Tagged Social Media Data	Geofeedia	Excel

Methodology

The collection of geo-tagged social media data specific to the Baltimore riots came from Geofeedia a location based social media monitoring company based in Chicago, IL. Geofeedia's platform allows users to capture in real time social media activity of a specific location. On the evening of April 27, 2015 and going into the early morning of the April 28, 2015, Geofeedia's platform captured geo-tagged social media data from Twitter, Instagram, Flickr, Picasa, and YouTube related to the riots in Baltimore. A limitation of the Geofeedia platform that may impact the results of this study is that Geofeedia only captures data from users who have location sharing turned on, and have their profile open to the public. Therefore, Geofeedia admits it only collects approximately 2% of all social media activity occurring at a specific location.

Roughly 797 geo-tagged social media posts were captured between the hours of 6pm EDT (April 27, 2015) and 3am EDT (April 28, 2015). The Excel file included the following information, which assisted in categorizing, analyzing, and mapping the social media activity:

- Date and Time Group
- Source i.e. Twitter, Instagram, Picasa, Flickr, and YouTube
- X Y Coordinates
- Author Name and Author Profile URL
- Comment Description

The comment description was vital in analyzing the sentiment at a particular time and place. Sentiment analysis refers to the evaluation of the attitude, opinion or feeling about a person, group or event. This type of analysis is beneficial as it provided some insight into the general attitudes being expressed location or event. In events like the Baltimore riots, analyzing social media sentiment can provide public safety officials an awareness of the mindset of individuals participating in social unrest. Data were processed using Geofeedia and ArcGIS to perform hotspot analyses and represent contemporaneous locations of looting and fires (figure 2)

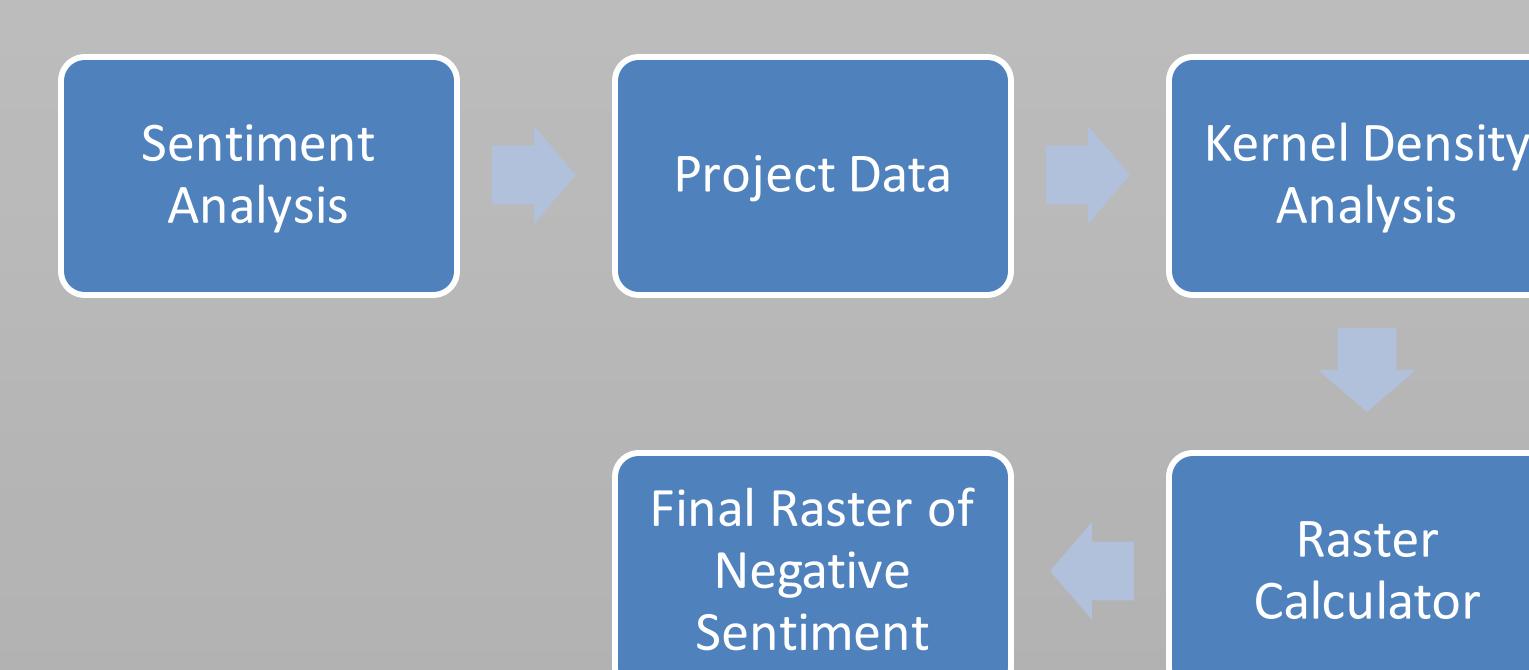


Figure 2. Overall workflow from raw data to final results

Timeline

Data collection occurred on April 27, 2015 the first night of the Baltimore riots through 3am EDT, April 28, 2015. Using the Geofeedia platform a polygon was created around downtown Baltimore to capture all social media activity. Social media activity relating to the Baltimore riots was exported and saved the following day. However, crime data was not available until two weeks after the riots.

Results

The progression of negative sentiment on the evening of the April 27th and going into the morning hours of the 28th demonstrates numerous "hotspots" that shifted throughout the nine hour span (figure 3). A correlation between areas of high negative sentiment and reported incident of looting and fires is also evident (figure 4). Despite the unknown time frame in which these incidents occurred, one can estimate the time in which these incidents occurred based on the sudden uptick in negative sentiment in the immediate area. Overall, the violence in Baltimore the first night was largely concentrated on the edges of the business and industrial districts.

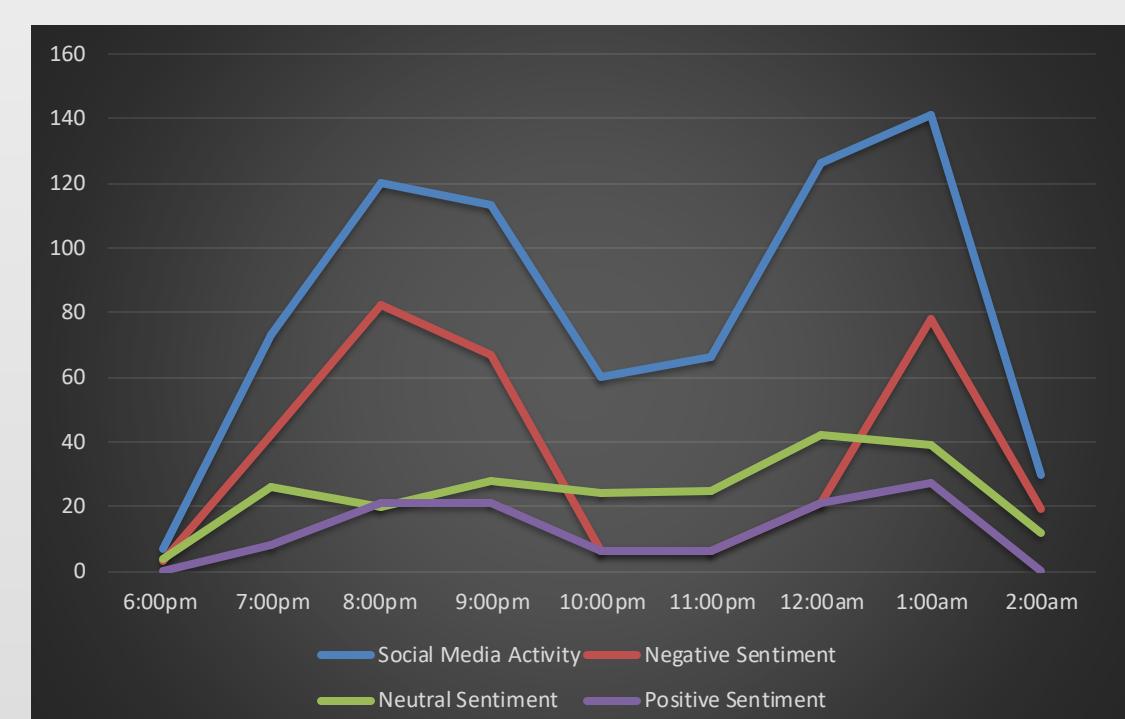


Figure 3. Overall social media activity and progression of social media sentiment.

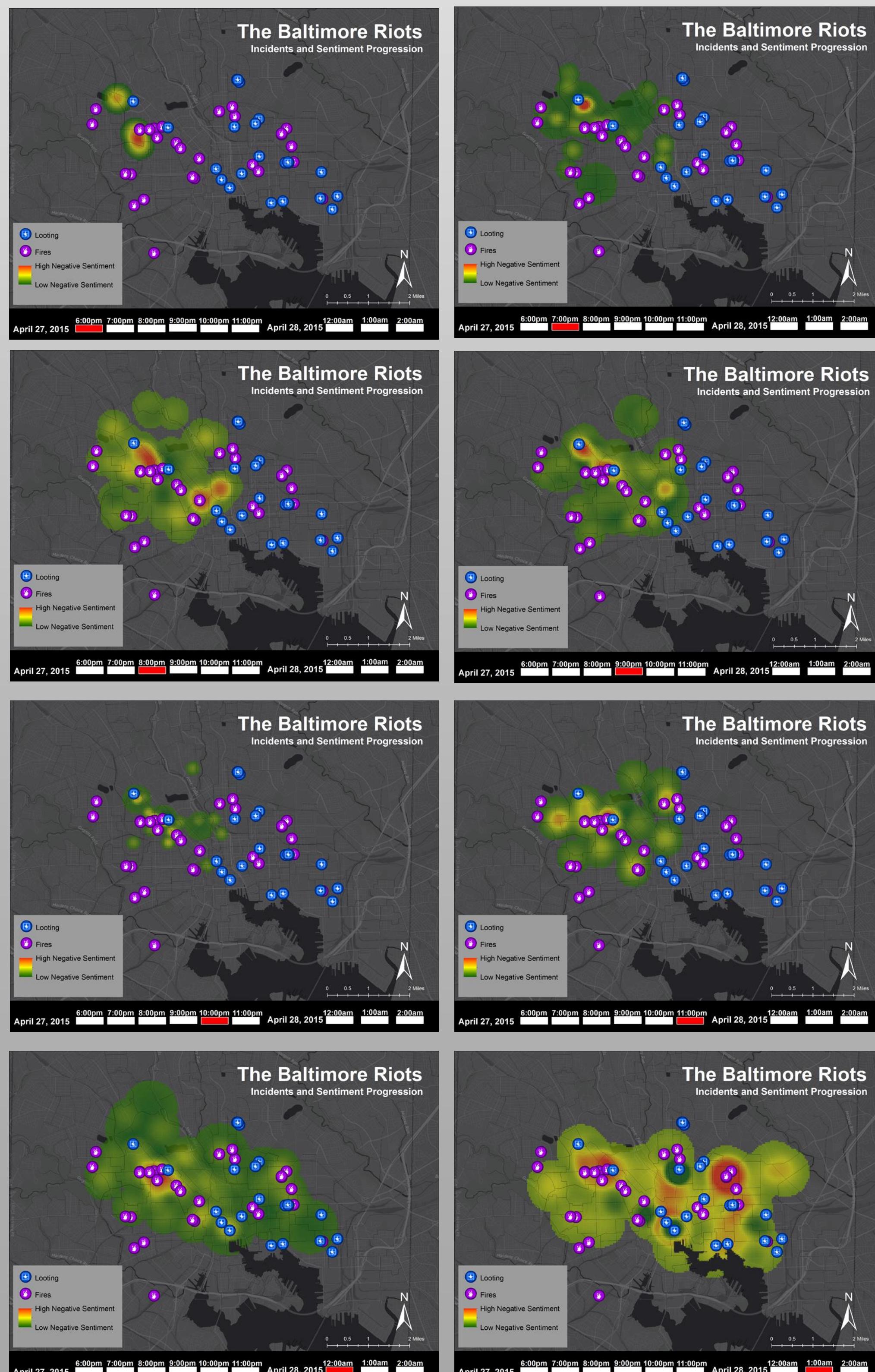


Figure 4. The progression of sentiment through the evening of April 27, 2015 and reported incidents of looting and fires.

Discussion

The geospatial combination of geo-tagged social media data, sentiment, with locations of confirmed incidents of fire and looting resulted in a visual correlation between negative sentiment and violent behavior. The most critical portion of this project was identifying and categorizing negative, positive, and neutral posts. Instead of analyzing specific words or hashtags contained in the post, an analysis of the overall contexts of the post to include the hash tag was done. In order for a post to be categorized as negative it had to meet a set criteria I customized based on factors influencing civil unrest. This type of sentiment analysis is very simple, yet did prove to be fairly accurate in finding correlations between locations of violent behavior and negative sentiment. The criteria were specific to the first day of the riots in Baltimore and may not be an effective means to measure sentiment the next day, as outside elements may impact or influences violent behavior like the implementation of a curfew or the deployment of the National Guard. Due to data limitations I was unable to analyze social media activity on the days following April 27, 2015.

This project demonstrates the effectiveness of using geo-tagged social media data and sentiment analysis to possibly forecast areas that may experience violent behavior during times of civil unrest. Thus, allowing security measures to be implemented or deployment of riot police to areas, which have a high concentration of negative sentiment

Conclusion

Social media has proven to be an effective communication platform for organizing gatherings and invoking civil unrests. It has also proven to be extraordinary real time intelligence tool to understand the complexity and the dynamics of a crowd. The ability to capture and analyze social media content can assist law enforcement in taking a more proactive strategy instead of a reactive one during times of civil unrest.

Going forward it is worth attempting to stream in real time geo tagged social media data into a customized online web map platform. This platform allows analyst to provide a weighted or numerical value on selected words or phrases that the analyst defines as negative, positive or neutral. Thus, based on the value or frequency of a specific term or activity a heat map can be created from the data being streamed and categorized. Such platform would provide the analyst the flexibility to change or add words or phrase that may impact sentiment results. This ability to capture the constant stream of geo-tagged social media data can uncover where violent behavior is likely to occur next.

Submitted in partial fulfillment of the requirements of the Masters of Science in Geographic Information Science(MSGISci), August 15, 2015.

For additional information please contact: Edwin Roman
Eroman03@hotmail.com