Spatiotemporal Driven Analysis of Law Enforcement Data

Guizhen Wang* Purdue University

Aubrey Akers† Cornell University

José Florencio de Queiroz Neto ‡ Federal University of Ceara

Chittayong Surakitbanharn§ Purdue University

David S. Ebert¶ Purdue University

ABSTRACT

Identification of crime patterns and trends can help law enforcement agencies to conduct proactive resource allocation and predictive policing practice. To satisfy analytical scenarios across different spatial aggregates (e.g., neighborhoods or census tracts) and temporal localities (e.g., weeks or months), a system, Visual Analytics Law Enforcement Toolkit (VALET), is proposed to interactively explore criminal incident reports in both domains. It is comprised of a suite of analytical tools that can conduct data analysis including spatiotemporal aggregation, correlation, and prediction of incidents. Using VALET, law enforcement agencies can find knowledge, formulate and verify hypotheses, and make decisions in a comprehensive analytical environment.

Index Terms: I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques; I.3.8 [Computer Graphics]: Applications—Visual Analytics

1 INTRODUCTION

Improving community safety and reducing criminal activities is a challenging task for law enforcement agencies across the world. Many factors create complex dynamic crime patterns both spatially and temporally: for example, burglaries may be more likely to happen during the day in residential areas and at night in commercial areas, whereas robberies may occur at night in public places. Data-driven policing has been a growing trend over the past two decades with great potential, but often cumbersome systems. Based on the needs of our law enforcement partners in undertaking this approach to proactive policing, we have developed Visual Analytics Law Enforcement Toolkit (VALET) [6]. The goal of our work is to understand crime occurrence patterns across spatial and temporal dimensions to enable proactive policing with better prepared teams, and more efficient resource allocations to reduce crime and increase community safety.

Law enforcement agencies today are capable of collecting rich, large volumes of data. However, traditional analysis approaches (e.g., R, Excel) and current software (e.g., PREDPOL [4]) often focus on using statistical methods to abstract criminal patterns but lack sufficient interactive methods of linking granularities in both the spatial and temporal domains.

To ease the analysis workflow, we have developed VALET [6], a suite of tools capable of assisting domain users in analyzing criminal, traffic and civil (CTC) incident reports. VALET specializes in dealing with spatiotemporal dimensions and includes functionalities for data aggregation, correlation, and prediction. Conducting data exploration through VALET, domain experts can acquire knowledge, obtain comprehensive situational awareness of CTC incidents and work with hypotheses, both formulating and verifying them.

*e-mail: wang1908@purdue.edu

†e-mail: aca79@cornell.edu

‡e-mail:florencio@lia.ufc.br

§e-mail: csurakit@purdue.edu

¶e-mail: ebertd@purdue.edu

2 RELATED WORK

To aid law enforcement agencies, researchers have tried many approaches to analyze crime patterns across both space and time. PREDPOL [4] focuses on statistical analysis of crime patterns generating predictive real-time maps. CrimeReports [3] proposes basic interactive methods to explore incident reports, shared by law enforcement agencies, citizens, and other institutions. TipSoft [5] focuses on law enforcement data acquisition and management. CommandCentral [1] is a crime mapping tool, allowing end users to compare crime patterns side by side with particular spatial and offense filtering predicate. All of these tools give little attention to interactive spatiotemporal data visualization and analysis. The advantage of VALET over other systems lies in its specialization in receiving and manipulating a large amount of spatiotemporal data, providing sophisticated visual interactive tools, and giving no limits to the domain expert.

3 LAW ENFORCEMENT DOMAIN CHARACTERIZATION

Law enforcement agencies nowadays face a dilemma between a more complicated public safety environment and severe budget restrictions. The traditional approach to fighting crime by using a large number of police officers to conduct random patrols can be impractical. [11] In the absence of a large number of police officers, the agencies have the incentive to deploy their teams to the best possible locations. Compared to random patrols, a directed patrol plan based on crime patterns and trends can help officers conduct a predictive policing practice through allocating their resources based on crime opportunities, and thereby preventing crime from occurring. For example, law enforcement agencies can send a team to patrol neighborhoods during the day with significant historical robbery trends, while at night, the effort can shift to sporting events where disorderly conduct and assaults may be more likely to occur.

In analyzing crime patterns and aggregating data, different naturally occurring patterns must be considered (e.g., seasonal, monthly, weekly, day of week, time of day), as well as events and holiday calendars. These temporal moments can reveal CTC concentrations, e.g., moon phases, holidays, weekdays, weekends, and weather changes. For example, the number of assault cases increases during the football season [2].

Similarly, it is important to aggregate and analyze the data at appropriate and significant spatial resolutions that are related to the driving factors of crimes (e.g., man-made and natural boundaries, neighborhoods). These different spatial aggregates usually have different geographic attributes, and as such, behave differently regarding crime patterns and trends. For example, criminal patterns at university campuses are different from those in commercial districts.

Inspired by the analytical needs of our law enforcement agencies in the Midwest, VALET enables end users to conduct many different analytical tasks, including the following three types of data analysis:

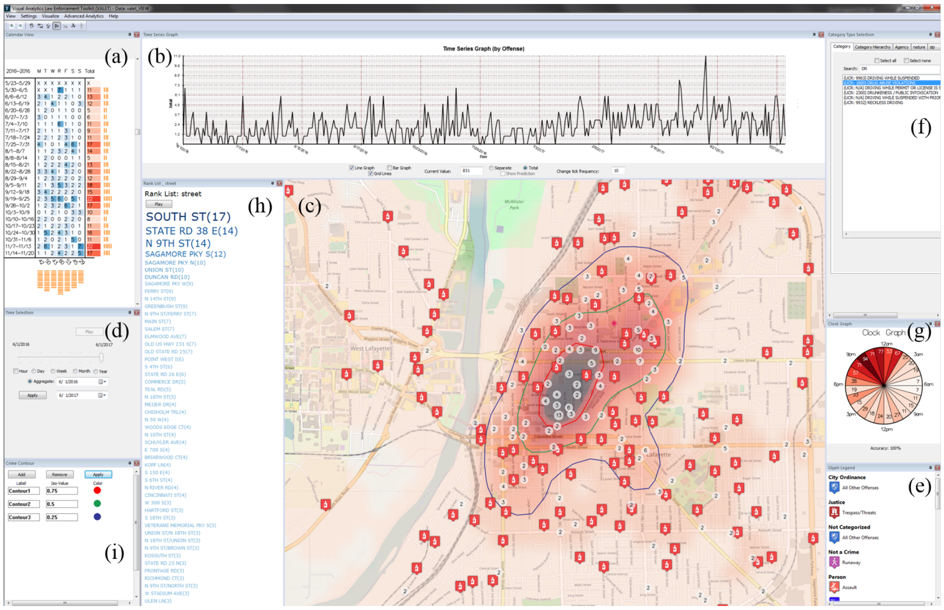

Figure 1: The Visual Analytics Law Enforcement Toolkit with (a) the calendar view to visualize incident counts in the weekly layout, (b) the time series view to show daily incident counts, (c) the map view to show spatial incidents, (d) the time slider to aggregate data in multiple temporal granularities, (e) the glyph view to show offense types, (f) the menu selection widget to filter incidents by categorical attributes, (g) the clock view to visualize data by hours, (h) the rank list to visualize attributes in the descending order of frequencies and (i) the setting to outline areas with high data densities. This picture shows the drug abuse cases from Jun 1st, 2016 to Jun 1st, 2017 in the cities of Lafayette and West Lafayette, IN.

- **Spatiotemporal aggregation analysis**, which can reveal crime concentrations and patterns at different resolutions. For example, a neighborhood focuses on security analysis in small spatial units, like streets and houses, while police officers allocate their resources in spatial units of patrol beats and districts, such as downtown areas and different neighborhoods. These different scales need to be accommodated and correlated.

- **Temporal correlation analysis**, since temporal correlations exist among different offense types. Occurrence of one offense may be a precursor of other offenses. For example, drug abuse is related to theft and burglaries.

- **Crime prediction** that can forecast future crime incidents based on historical records. For example, there is a near-repeat effect in burglaries [9], that is, a neighborhood with a burglary is likely to have another one.

4 OVERVIEW OF VALET

As previously mentioned, based on law enforcement agency needs, we developed the VALET system (seen in Fig. 1) with linked views and interactive displays. In the temporal domain, incidents can be aggregated by hour, day, week, month, or year, which enables analysis and trending over appropriate temporal scales and granularities. In the spatial domain, to allow crime analysts to conduct analysis at appropriate spatial scales, VALET provides two ways to specify spatial aggregates: (1) allowing free-from hand-drawing sketches on the map to specify regions, and (2) providing APIs to integrate boundary shapefiles, (e.g., census tracts, neighborhood watch boundaries, patrol beats). Then, based on the appropriate spatiotemporal predicates, VALET enables analysis and prediction for specific officer and analyst tasks.

5 INTERACTIVE SPATIOTEMPORAL INCIDENT ANALYSIS

Our system focuses on both spatial and temporal modeling of criminal activities. In this section, we introduce three methods to address analytical tasks proposed in Section 3, including (1) spatiotemporal data aggregation, (2) temporal correlation analysis, and (3) crime prediction.

5.1 Spatiotemporal data aggregation

VALET assists our law enforcement partners in observing incident aggregations across geospace and time by providing multiple linked views. For spatial aggregation, information in the format of categories (e.g., zip code and streets) can be visualized in rank lists, prioritizing high-frequency values in descending order, while contiguous information can be visualized on the map view, using heatmaps and contours to enhance regions with high data densities. For temporal aggregating, data densities across multiple granularities can be found in the three views, including a calendar view for daily incident counts (a matrix layout with rows as weeks and columns as days-of-week), a clock view for hourly distribution, and the time series graphs for temporal daily trends. We demonstrate these features for drug abuse cases in the Tippecanoe County from Jun 1st, 2016 to Jun 1st, 2017 in Fig. 1 . In terms of spatial measurements, the map view shows that spatial hotspots occur in the downtown area, and contours sketch regions with counts in the upper 25% (red), upper 50% (green) and lower 25% (blue) quantiles. The rank list shows the streets with the highest drug abuse frequencies. Our temporal views show that drug abuse cases are concentrated at night (clock view), while the calendar view shows that Thursday has the highest number of cases, and the time series view shows a cyclical pattern in daily occurrences.

5.2 Temporal correlation analysis

VALET uses a visual analysis approach to explore the temporal correlation among crime categories [7]. To find out whether one offense may be a precursor to a series of other offenses, we measure the correlation coefficient between the time series with different lead and lag periods, using the Pearsons product-moment correlation. As described in [7], you see correlations between drug abuse violations and burglaries (both residential and commercial) of Tippecanoe

County, IN in 2011. Fig. 2 shows a negative correlation at a lead value of -1 days between drug abuse violations and burglaries, indicating that an increase in drug abuse violation arrests or reports lead to a decrease in burglaries after one day. Another significant correlation is that drug abuse violations lead burglary crimes by 11 and 26 days (seen by the coefficient at a lag value of -11 and -26 days in Fig. 2). The crime analyst in the local police department hypothesized that these two discoveries may come from the fact that drug offenders need fast money to buy drugs, and thus, burglaries can be reduced when these offenders get arrested with drug abuse violations in jail, and increase again once they get bailed out of jail. Although the causal relation between drug abuses and burglaries cannot be completely confirmed by this case, crime analysts indeed gained new knowledge to support their subsequent research.

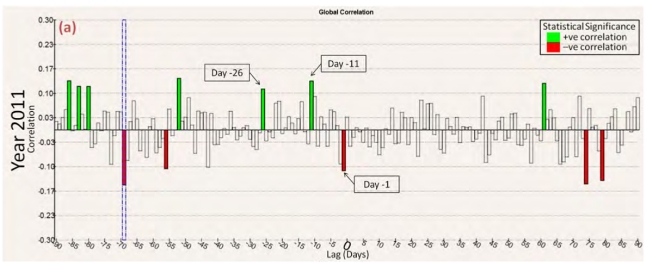

Figure 2: Correlating drug abuse violations and burglary crimes in Tippecanoe County, IN, 2011. The x axis encodes the lead/lag value in the unit of days, while the y axis shows the coefficient value.

5.3 Crime prediction

For crime prediction, VALET utilizes natural scale spatial templates generated with domain expertise [8]. The template is a group of spatial hotspot regions generated from historical data utilizing the density based measurements, as well as the end users' knowledge and analytical requirements. In a case study we conducted with one police department [8], we forecasted CTC incidents and helped police officers to plan patrol routes in Tippecanoe County, IN. One example of the prediction results is shown in Fig. 3, using a heatmap to visualize crime opportunities in the city downtown regions. The selected spatial template in Fig. 3 is based on census blocks, which are the spatial aggregates that the crime analyst wants to analyze. Therefore, the crime opportunities are shown by block. The analyst can use these results to formulate patrolling plans.

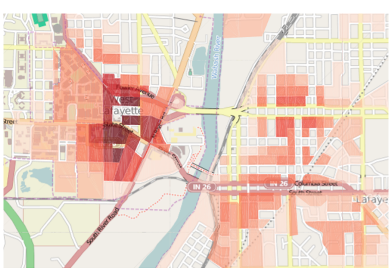

Figure 3: Crime prediction results for 3/11/2014 using 10-year incident data from 3/11/2004 through 3/10/2014 in Tippecanoe County, IN, visualized by heatmaps and refined by the census block boundaries.

6 Collaboration with Law Enforcement Agencies

VALET was developed in strong collaboration with our Purdue VACCINE public safety Consortium law enforcement partner agencies, including three local police departments around Purdue University and several police departments in other cities of Indiana and the Midwest. Domain users from law enforcement agencies have used VALET to strengthen their effectiveness (e.g., supporting new dynamic police beat concepts in Lafayette, IN, officers in Evansville, IN using VALET for daily crime analysis). They usually generate some hypothesis from their domain knowledge and use VALET's rich interactive tools to look at incident details to validate or explore their hypothesis. Our system has also been used in the Lafayette Police Department to assist discussions and increase community-based engagement at their monthly community meetings.

From the start, our system and its features have been driven by these police departments. Several components in our system are from these domain users' suggestions. For example, the chief of the Lafayette Police Department proposed the idea of using the rank list to order streets (seen in Fig. 1(h)) for patrolling. Having seen the benefits of VALET in their desktop operations, police officers requested a mobile version of VALET, iVALET [10], that runs on an iPad/iPhone. Different from VALET, which is often used in a controlled office environment for a full and thorough analysis, iVALET is more geared towards real-time situation awareness, proving in-field officers with on-the-go risk assessment and easier interactions during patrolling. Particularly, iVALET has a risk profile design to show historical incidents on the map concerning current location and time of officers, and proposes a customized heatmap to encode the temporal distances of those events from the current time. Constrained by the limited display area in mobile devices, iVALET condenses the layout of multiple views to optimize space utilization. Considering the distraction for police officers to drive a car while simultaneously looking at the system, we are currently developing a voice command component to allow officers to communicate vocally with the system.

The data analysis across the spatial aggregates and temporal localities are inspired by domain users as well. Here, we summarize two analytical scenarios that demonstrate the impact of spatial granularities on data analysis. The first case is for the monthly community meetings with the Lafayette Police Department, where the spatial granularity of their analytical task focuses on a specific neighborhood. A neighborhood is a small region that can only be shown at a very detailed zoom level, including houses, street corners, shops, etc. Using a heatmap to show the incident distribution in a neighborhood is insufficient due to the difficulty to choose appropriate spatial binning size. Fig. 4 (a) shows the spatial incident distribution by the heatmap across the entire Tippecanoe County in Indiana, with the blue rectangle drawing the neighborhood that their analysis focuses on. Fig. 4 (b) shows the expected zoom level appropriate for their analysis, clearly showing the details, like street corners. From Fig. 4 (c), we can find that the spatial granularity of the heatmap is not useful in their analysis since the hotspot covers the entire neighborhood and users cannot tell the incident distribution across each small location. In order to allow crime exploration in a small location, we developed the function to allow end users to draw a polygon on the map to query data, such as drawing a rectangle (shown in Fig. 4 (b)) or manually drawing a polygon (a green region shown in Fig. 1 (c)). Fig. 4 (d) shows the spatial analysis process used in a recent community meeting to discuss their community security issues. After drawing a rectangle (in green) on the map to select data about their neighborhood, they can use glyphs and bubbles to visualize incident locations, where a glyph indicates a single incident report, and a bubble is an aggregate of crimes occurring in the same location with the number shown in the bubble center. User's attention is attracted to bubbles with high numbers and some hypothesis are generated to explain the high frequency of incidents, e.g., a shop may have more 911 calls than a house. The second case uses the crime prediction feature, a popular function in police departments, since officers want to use this function to plan their patrolling strategies. The idea of using natural scale templates in our prediction algorithm in Section 5.3 is from the feedback of crime analysts in our local police departments. Officers want to plan their daily patrolling routes in the unit of a small region (e.g., a residential area). To satisfy the requirements, we enabled end users to specify a spatial aggregate to show the distribution of crime opportunities.

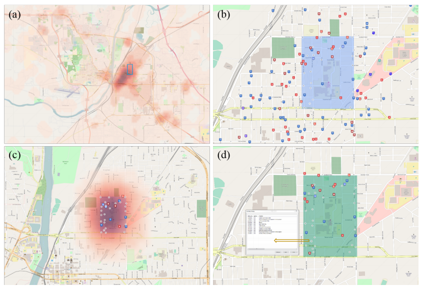

Figure 4: Spatial analysis at the community meeting in Lafayette, Indiana. (a) shows the spatial incident distribution of the entire dataset visualized with a heatmap, with a blue rectangle to highlight the community region. (b) shows a close-up of the community meeting neighborhood highlighted in a blue rectangle (the blue region in (a)). (c) shows the visualization using both the heatmap and glyphs in the neighborhood. (d) shows the actually used spatial analysis process.

In addition to law enforcement agencies, we also help experts in other domains to conduct criminal analyses. Purdue Veterinary Medicine researchers currently utilize VALET to analyze the correlation between animal abuse and domestic violence. Animal abuse reporting is not mandatory in Indiana, so VALET is being used to provide research to inform Indiana politicians, which might encourage them to begin the push for mandatory reporting of animal abuse by veterinarians, as well as cross-reporting neglect and abuse cases between child and animal protection agencies. VALET provides the ability to explore spatiotemporal aggregations regarding cases of child abuse, domestic violence, and animal abuse, as well as locations of vet clinics and human societies. Fig. 5 demonstrates the adopted primary visualizations, from which some analytical achievements have been made by researchers to summarize the correlation.

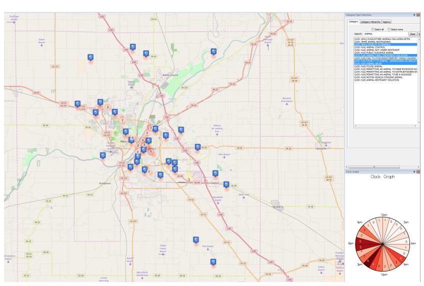

Figure 5: Spatial data distribution of cases of animal abuse, classified under "cruelty to animals", "animal care/treatment violation", "cruelty to animals - abandoning or neglect", or "animal abuse" between Jaunary 1st, 2013 and June 14th, 2017. The clock view represents one form of a temporal data visualization that can be analyzed with VALET.

Overall, VALET received positive feedback from our collaborators who felt that the interactive spatiotemporal data analysis methods in VALET can help them more conveniently conduct analysis, compared to other standard methods (e.g., using Microsoft Excel). One police officer in Evansville, IN, used VALET to successfully identify a habitual theft, and regarded VALET as "probably the most useful tool he had". The temporal correlation analysis method was motivated by the experience of a former police officer who wanted to compare activities and patterns of several types of crime, since criminals are often engaged in multiple criminal activities. After exploring our developed method, the officer thought the proposed visualization could intuitively present correlations between different crime activities, and quickly test their hypotheses, compared to their standard ways of looking at statistical values. Furthermore, a Crime Prevention Specialist, aware of the clarity and visual appeal of our solution, recommended VALET as a software for public demonstration, and actively involved our system in Lafayette, IN neighborhood meetings. Fascinated by the interactive analysis capabilities in VALET, the meeting attendees showed strong interests in using our system, inquired about public availability, and actively suggested new features. Our law enforcement collaborators believe that VALET can be used for daily briefings, community outreach, crime analytics, as well as shift scheduling and officer allocation.

Valuable experiences and lessons have been accumulated from our long-term collaboration with law enforcement agencies and other collaborators. Firstly, we learned that different criminal offense types pertain to diverse spatial and temporal variants, e.g., days-of-week, moon phases, streets, and neighborhoods. Therefore, VALET has provided multiple options to conveniently allow officers to define spatial and temporal specifications, e.g., the calendar view in Fig. 1 not only sets one row to denote a week with the bottom to show day-of-week patterns, but also allows users to change the number of days a row can have based on other criminal occurrence periods. Secondly, we learned that each analytical scenario has specific data confidentiality requirements in law enforcement agencies. To satisfy this necessity, VALET was designed with flexible data importing and display options, e.g., merely loading non-sensitive data fields for the public demonstration, and loading sensitive data fields in internal analysis tasks. Lastly, the iterative design and refinement of the system with police officers helped us learn and incorporate officer knowledge into VALET. Our team made many efforts to build a productive working relationship with our collaborators, including hosting monthly public safety luncheons to introduce our latest achievements and plans, as well as attending monthly internal meetings of law enforcement agencies to understand their working styles. Those efforts effectively helped us exchange information, find their essential problems, and develop systems in accordance with their mental models. Nonetheless, domain experts indicated several areas of improvement. One area that VALET needs to advance is in comparison among multiple visualizations. For example, in the animal abuse project, researchers identified the necessity to compare spatial hotspots between animal abuse and domestic violence side by side. However, the available way in VALET does not fit this scenario, since users need to save visualizations as images, and manually compare pictures. Another point for improvement is a better way to search smaller spatial units in a finer zoom level. Currently, end users have to zoom in/out and pan the map view in Fig. 1 to find the desired location, (e.g., a street corner).

7 Conclusion and Future Work

Spatiotemporal driven data analysis can help law enforcement agencies identify crime patterns across the spatial aggregates and temporal localities. Law-enforcement partnership led to the development of VALET, a visual analytics system, to provide these capabilities to law enforcement agencies. VALET has a rich set of interactive analysis tools to aid domain users in conducting their data analysis at the appropriate spatiotemporal resolutions, including data aggregation, temporal correlation, and crime prediction. VALET has been assisting law enforcement agencies by transforming their decision making processes from reactive to proactive. VALET will continue to develop new data analysis components to enhance domain users' analytical performance. Two ongoing research directions include helping police officers plan their daily patrolling routes, and allowing police officers to communicate with VALET using speech.

REFERENCES

[1] Commandcentral. http://classic.crimereports.com/lea/ccscreencast. Accessed: 2017-06-09.

[2] Crime and football: Domestic violence rises 10 percent after nfl upsets. http://ucsdnews.ucsd.edu/archive/newsrel/soc/03-24CrimeandFootball.asp. Accessed: 2017-06-09.

[3] Crimereports. https://www.crimereports.com. Accessed: 2017-06-09.

[4] Predpol. http://www.predpol.com/. Accessed: 2017-06-09.

[5] Tipsoft. http://www.tipsoft.com/index.aspx. Accessed: 2017-06-09.

[6] A. Malik, R. Maciejewski, T. F. Collins, and D. S. Ebert. Visual analytics law enforcement toolkit. In *2010 IEEE International Conference on Technologies for Homeland Security (HST)*, pp. 222–228, Nov 2010. doi: 10.1109/THS.2010.5655057

[7] A. Malik, R. Maciejewski, N. Elmqvist, Y. Jang, D. S. Ebert, and W. Huang. A correlative analysis process in a visual analytics environment. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 33–42, Oct 2012. doi: 10.1109/VAST.2012.6400491

[8] A. Malik, R. Maciejewski, S. Towers, S. McCullough, and D. S. Ebert. Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1863–1872, Dec 2014. doi: 10.1109/TVCG.2014.2346926

[9] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108, 2011.

[10] A. M. M. Razip, A. Malik, S. Afzal, M. Potrawski, R. Maciejewski, Y. Jang, N. Elmqvist, and D. S. Ebert. A mobile visual analytics approach for law enforcement situation awareness. In *2014 IEEE Pacific Visualization Symposium*, pp. 169–176, March 2014. doi: 10.1109/PacificVis.2014.54

[11] R. B. Santos. *Crime Analysis With Crime Mapping*. SAGE Publications, 2012.