

# Visual Analytics Law Enforcement Toolkit

Abish Malik, Ross Maciejewski, Timothy F. Collins, David S. Ebert  
Purdue University Visualization and Analytics Center  
Purdue University  
West Lafayette, IN 47906  
Email: amalik|rmacieje|ebertd|tfcollin@purdue.edu

**Abstract**—We present VALET, a Visual Analytics Law Enforcement Toolkit for analyzing spatiotemporal law enforcement data. VALET provides users with a suite of analytical tools coupled with an interactive visual interface for data exploration and analysis. This system includes linked views and interactive displays that spatiotemporally model criminal, traffic and civil (CTC) incidents and allows officials to observe patterns and quickly identify regions with higher probabilities of activity. Our toolkit provides analysts with the ability to visualize different types of data sets (census data, daily weather reports, zoning tracts, prominent calendar dates, etc.) that provide an insight into correlations among CTC incidents and spatial demographics. In the spatial domain, we have implemented a kernel density estimation mapping technique that creates a color map of spatially distributed CTC events that allows analysts to quickly find and identify areas with unusually large activity levels. In the temporal domain, reports can be aggregated by day, week, month or year, allowing the analysts to visualize the CTC activities spatially over a period of time. Furthermore, we have incorporated temporal prediction algorithms to forecast future CTC incident levels within a 95% confidence interval. Such predictions aid law enforcement officials in understanding how hotspots may grow in the future in order to judiciously allocate resources and take preventive measures. Our system has been developed using actual law enforcement data and is currently being evaluated and refined by a consortium of law enforcement agencies.

## I. MOTIVATION

The exploration of criminal incident reports for detecting trends, discovering anomalies and evaluating resource usage is an ever expanding issue for law enforcement agencies. It is no longer efficient for a single analyst to pull files, take notes, form hypotheses and request data from different sources. Further, as budgets shrink and departments scale back, the ability of local law enforcement agencies to effectively analyze the data being collected becomes increasingly strained. As such, tools need to be developed that bring varying data sources into a unified framework assisting analysis and exploration in order to speed the analytical process and ease the burden on local agencies. Many of these tool development needs are being addressed by the emergence of a new scientific field, visual analytics. Visual analytics is the science of analytical reasoning assisted by interactive visual interfaces [1].

In order to better facilitate criminal incident analysis, we have extended our previous visual analytics tools [2], [3] for the enhanced exploration of multivariate spatiotemporal data to enable advanced data exploration and analysis of criminal incidence reports. These tools were modified directly for law

enforcement use through collaboration with the VACCINE public safety coalition of four law enforcement agencies. Our system is currently used to explore daily crime report data from the West Lafayette Police Department (Indiana). Our current work focuses on both spatial and temporal modeling of criminal activities as well as the early detection of unusual criminal occurrences. System features include the following:

- Multi-level aggregate views for crime mapping including census tract based choropleth maps and kernel density estimate heat maps
- Linked interactive displays for multi-domain/multivariate exploration and analysis
- Seasonal trend decomposition modeling for temporal trend analysis and prediction
- Multiple time series views for time series exploration and trend analysis including line graphs and calendar views
- Crime clustering based on spatial autocorrelation
- Filter controls that enable database querying and analysis through an intuitive graphical interface

Our work focuses on advanced interactive visualization and analysis methods providing linked environments of geospatial data and time series graphs. Hotspots found in one data display visualization can be selected and immediately analyzed in the corresponding linked view. Further, our system allows analysts to integrate other data streams, looking at community events (such as local football games), weather, and other signals of interest. As such, our system allows users to look for patterns in both the spatial and temporal data domains. Knowledge extracted here can be used to develop hypotheses and future analytic capabilities will provide means for hypothesis testing on data clusters and time series anomalies.

## II. RELATED WORK

In order to improve public safety and prevent crimes, law enforcement agencies need to analyze the volumes of data from multiple systems, search for trends, and deploy services appropriately. As such, many packages exist for studying spatial relationships between crime and area demographics. Work by Messner and Anselin [4] uses exploratory spatial data analysis to visualize spatial distributions and suggest clusters and hotspots. Specifically they look at spatial autocorrelation and box maps. Other work includes WebCAT by Calhoun et al. [5] which focuses on enhanced data sharing and crime data analysis tools via the web. Their tools include choropleth mapping and capabilities to export records to Excel. Our work



Fig. 1. The Visual Analytics Law Enforcement Toolkit. In this screenshot, the user has adjusted the interactive filter controls (as seen in the map viewing window) to visualize theft. Linked views (above and left) show the line graph and calendar view temporal plots. A legend for the crime incident color mapping is shown in the upper right. The interactive time slider is shown in the lower right. Note the controls for aggregating data by different time scales.

presents similar capabilities to both Messner and Anselin [4] and Calhoun et al. [5]; however, we also include dynamically linked views and advanced hotspot detection tools not found in either of these works.

Other work that pertains to spatial data mapping and analysis include ArcView that aggregates crime incident data by geography (like census tracts, etc.) and displays it in the form of a choropleth map [6]. ArcView provides simple statistical tools like scattergram and histogram plots that help in the visual examination of the data. Some other crime mapping applications include GeoDa [7], developed by Anselin et al., that provides several statistical applications for doing both

exploratory and confirmatory data analysis. Their spatial analysis work includes box plot maps, spatial association, LISA local Moran maps, Moran significance maps, spatial regression residual mapping, etc.

Levine [8] developed the CrimeStat program for the analysis and mapping of crime incident reports. CrimeStat incorporates statistical tools that describe the general properties of the spatial distribution of the crime incidents. These include journey to crime estimation tools that estimate the residence of serial offenders based on the pattern of their crime locations, a space-time analysis tool for analyzing clustering in space and time, spatial autocorrelation, etc. Although the CrimeStat

toolset provides the analysts with spatial analysis tools, it lacks interactive exploration methods that our system provides to the users.

### III. VISUAL ANALYTICS ENVIRONMENT

We have developed a visual analytics system (VALET) for analyzing spatiotemporal law enforcement data that can be used by agencies for detecting anomalies and criminal activity patterns. The criminal, traffic and civil (CTC) data is maintained by the Tippecanoe County Police Department and has been aggregated into different categories including: armed aggravated assault, armed robbery, burglary, homicide, noise, other assaults, rape, attempted rape, residential entry, robbery, theft, unarmed aggravated assault, vandalism and vehicle theft. Currently, we have CTC data starting at the year 2000 with updates ingested monthly.

#### A. System Features

Our current work is based on the system developed by Maciejewski et al. [2], [3], modified to operate for criminal incident report data. Figure 1 shows a snapshot of our system. The main window (Figure 1 - Center) of the system shows the geospatial view that supports the overlay of different maps and CTC incidents along with interactive panning and zooming tools. The top-most window shows the line graph time series view of the CTC data which provides the users with the option of viewing the aggregate of all the selected CTC types or visualizing them separately on one graph. The left-most window is the calendar view of the selected CTC incidents that shows the sum of crime incidents for each day of a calendar year. The calendar view enables the users to visualize special events like football and basketball games on the calendar further allowing them to make a connection between the reported CTC activities and these events. The bottom-most window contains the time slider that is used to temporally scroll through the CTC data while dynamically updating all the other linked windows to reflect the change. This window also contains radio buttons to select the type of temporal aggregation for the CTC incidents. Finally, the right-most window shows a legend of all the CTC incidents with those selected for visualization highlighted.

Another key feature of our system is the interactive demographic and CTC filtering component. Users interactively generate data search queries through the use of check boxes and edit controls to find specific CTC categories. This interaction is shown in the drop-down menu of Figure 1. Other filter options (using the Census Data menu) allow users to plot demographic data on the map, such as median income. Analysts can interactively select various data layers and search for potential correlations. Such interaction furthers hypothesis generation and exploration as users can now quickly filter signals by demographic constraints in order to see if crimes are related to a particular segment of the population. The choices of filters affect both the geo-spatiotemporal viewing area and all unlocked temporal plots.

Other system features include an interactive legend (shown in Figure 1 - Upper Right). By clicking on any of these selected CTC types on the legend, the system fills the CTC circles with a solid color and further dims out the other CTC types on the map. This action spatially highlights the selected CTC types among all the other CTC types selected for visualization.

#### B. Time Series Display and Analysis

Along with the advanced graphical interface, our system contains a variety of visualization features for both spatial and temporal views. For temporal views, three options are provided to the user: the calendar view, the line graph view, and a predictive overlay on the line graph view.

a) **Calendar View:** The calendar view visualization was first developed by van Wijk and Selow [9]. This visualization provides a means of viewing data over time, where each date is shaded based on the overall yearly trends. Here, the max data value is shaded the darkest blue, and the lowest data values are shaded white. Users can interactively control the cycle length of the calendar view. In Figure 1 the user has interactively chosen a cycle length of 14 days. Each row and column of the calendar view also maps to a histogram plot at the bottom and right edges of the calendar. This lets users search for both seasonal and cyclical trends. Furthermore, our system also allows for multi-source data integration, including factors such as weather, school calendars, local sporting events, etc.

b) **Line Graph View:** The line graph visualization allows the user to view the temporal data trends of multiple crime signals at a single time. If a user selects both theft and vandalism, for example, two line graphs would appear on the map, allowing the analyst to compare trends. A user may then interactively choose to combine the input signals together using the ‘Total’ toggle button found on the line graph view. This allows for a variety of ways to view the data. Furthermore, the data is plotted based on the time slider control aggregate (Figure 1-Lower Right Widget). In this case, the data is being plotted by month. Finally, the line graph view also utilizes a predictive analysis metric in order to capture trends and inform analysts of potential future issues.

c) **STL Prediction:** In order to more accurately model the data, we employ a different strategy in which the time series is viewed as the sum of multiple components of variation [10]. Seasonal-trend decomposition based on *loess* (locally weighted regression) [11] is used to separate the time series into its various components. STL components of variation arise from smoothing the data using moving weighted-least-squares polynomial fitting, in particular *loess* [12], with a moving window bandwidth in days. The degree of the polynomial is 0 (locally constant), 1 (locally linear), or 2 (locally quadratic).

Here, it is important to note that in order to appropriately model the time series using STL, the mean and variance of the data needs to be independent. To accomplish this, a power transformation is applied to the data. In time series analysis, the logarithm transformation is widely applied when the mean

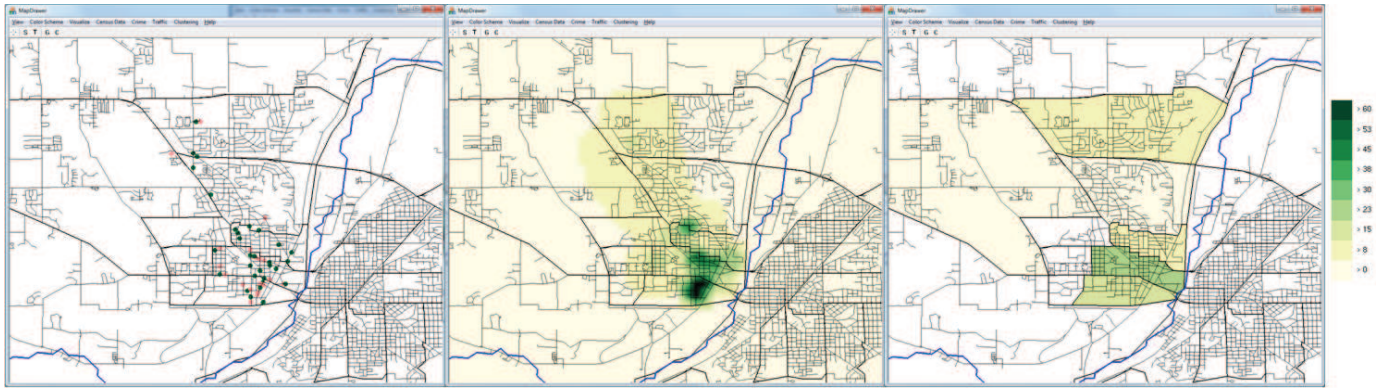


Fig. 2. Screenshots of the available map aggregation levels. (Left) Plotting crimes by point and color. (Middle) Plotting crimes using density estimate heatmaps. (Right) Plotting crimes aggregated over census tract areas.

is proportional to the standard deviation [13]. In cases where the data consists of counts following a Poisson distribution a square root transformation will make the mean independent of the standard deviation.

For a given time series, we decompose our data into a day-of-the-week component, a yearly-seasonal component that models seasonal fluctuations, and an inter-annual component which models long term effects:

$$\sqrt{Y_t} = T_t + S_t + D_t + r_t \quad (1)$$

where for the  $t$ -th day,  $Y_t$  is the original series,  $T_t$  is the inter-annual component,  $S_t$  is the yearly-seasonal component,  $D_t$  is the day-of-the-week effect, and  $r_t$  is the remainder.

The procedure begins by extracting the day-of-the-week component,  $D_t$ . First, a low-middle frequency component is fitted using locally linear fitting. Then  $D_t$  is the result of means for each day-of-the-week of the  $\sqrt{Y_t}$  minus the low-middle-frequency component. Next, the current  $D_t$  is subtracted from the  $\sqrt{Y_t}$  and the low-middle-frequency component is re-computed. This iterative process is continued until convergence. After removing the day-of-the-week component from the data, we use loess smoothing to extract the inter-annual component,  $T_t$ . Finally, we apply loess smoothing to the data with the day-of-week and inter-annual components removed, thereby obtaining the yearly-seasonal component,  $S_t$ , using local quadratic smoothing. After removing the day-of-week, inter-annual, and yearly-seasonal components from the time series, the remainder is found to be adequately modeled as independent identically distributed Gaussian white noise, indicating that all predictable sources of variation have been captured in the model. Details of the methodology and means of appropriate parameter choices can be found in [10] and is not the main focus of this work. However, the extension of this method to criminal incidence reports is novel.

For prediction using the STL method, we rely on some statistical properties of loess, namely that the fitted values  $\hat{Y} = (\hat{Y}_1, \dots, \hat{Y}_n)$  are a linear transformation of the observed data,  $Y = (Y_1, \dots, Y_n)$ . Each step of the STL decomposition involves a linear filter of the data. In other words, an output

time series  $x = \{x_1, \dots, x_n\}$  is produced by an input time series  $w = w_1, \dots, w_n$  through a linear combination

$$x_i = \sum_{j=1}^n h_{ij} w_j. \quad (2)$$

If we let  $H$  be a matrix whose  $(i, j)$ -th element is  $h_{ij}$ , then we have

$$x = Hw. \quad (3)$$

Further details of utilizing STL for prediction can be found in an application to syndromic surveillance work by Maciejewski et al. [3], and we extend this method to use with CTC data.

In Figure 1 (Top), we see the predicted rates plotted as the red line, with the upper and lower confidence bounds plotted in purple. We then compare this prediction to the actual data in Figure 1 (Top). Here we can see that our data falls within the confidence interval provided and the rises and falls of the data are relatively well captured.

### C. Geospatial Displays

Along with the various temporal viewing and analysis algorithms, VALET also provides analysts with various spatial mapping and analysis components. In the main map window (seen in Figure 1), the user is allowed to plot incidents as points (Figure 2 - Left), density estimated heatmaps ((Figure 2 - Center) or choropleth maps using census tract boundaries (Figure 2 - Right).

For the density estimated heatmaps, we employ a modified variable kernel method [14] which scales the parameter of the estimation by allowing the kernel width to vary based upon the distance from  $X_i$  to the  $k$ -th nearest neighbor in the set comprising  $N - 1$  points.

$$\hat{f}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\max(h, d_{i,k})} K\left(\frac{\mathbf{x} - X_i}{\max(h, d_{i,k})}\right) \quad (4)$$

Here, the window width of the kernel placed on the point  $X_i$  is proportional to  $d_{i,k}$  (where  $d_{i,k}$  is the distance from the  $i$ -th sample to the  $k$ -th nearest neighbor) so that data points in regions where the data is sparse will have flatter kernels, and  $h$  is the minimum allowed kernel width.

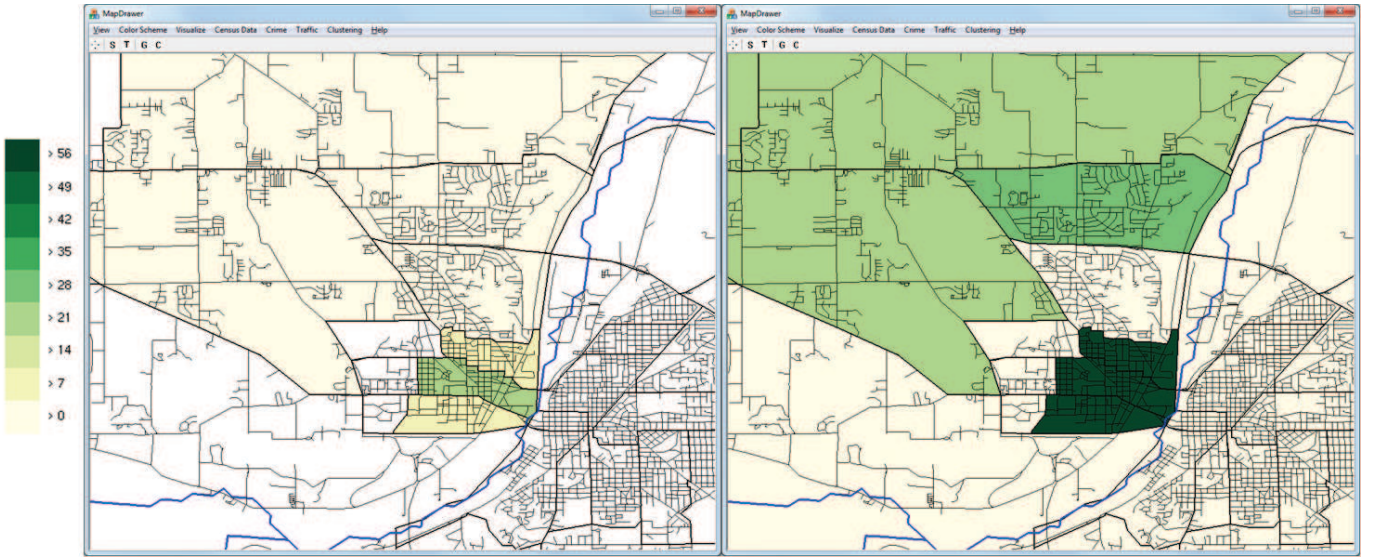


Fig. 3. Applying the AMOEBA algorithm to crime data. On the left, census tracts are colored based on the number of noise and vandalism complaints within the area. On the right, census tracts are grouped together based on their spatial autocorrelation in terms of the number of noise and vandalism complaints and then colored based on the aggregate.

We utilize the Epanechnikov kernel [14], Equation 5:

$$K(\mathbf{u}) = \frac{3}{4}(1 - \mathbf{u}^2)\mathbf{1}_{(\|\mathbf{u}\| \leq 1)} \quad (5)$$

where the function  $\mathbf{1}_{(\|\mathbf{u}\| \leq 1)}$  evaluates to 1 if the inequality is true and zero for all other cases.

#### D. AMOEBA Clustering

Along with the various geospatial display capabilities, our system also provides clustering based on spatial statistics. In order to group data based on spatially similar population statistics, we utilize the AMOEBA algorithm for creating spatial weights matrix developed by Aldstadt and Getis [15]. AMOEBA (A Multidirectional Optimum Ecotope-Based Algorithm) procedure is designed to identify hot and cold spots in mapped data by assessing the spatial association of a particular mapped unit to its surrounding units. It is able to aid in the demarcation of clusters of related spatial units, and we utilize this fact to group counties based on population statistics.

AMOEBA maps clusters of high and low values by creating a spatial weights matrix based on the Getis-Ord  $G_i^*$  statistic. For a given location  $i$ ,  $G_i^*$  is defined as

$$G_i^* = \frac{\sum_{j=1}^N w_{ij}x_j - \bar{x} \sum_{j=1}^N w_{ij}}{S \sqrt{\frac{[\sum_{j=1}^N N_{j=1} w_{ij}^2 - (\sum_{j=1}^N w_{ij})^2]}{N-1}}} \quad (6)$$

Here,  $N$  is the number of spatial units,  $x_j$  is the value of interest within the areal unit at location  $j$ ,  $\bar{x}$  is the mean of all values, and

$$S = \sqrt{\frac{\sum_{j=1}^N x_j^2}{N} - (\bar{x})^2} \quad (7)$$

$w_{ij}$  is used as an indicator function that is one if  $j$  is a neighbor of  $i$  and zero otherwise.

The AMOEBA algorithm develops a cluster from a selected seed location by evaluating  $G_i^*$  at all locations surrounding this seed location, and if the addition of a neighbor to the cluster increases the  $G_i^*$  value, then the neighbor is added. Details of this algorithm and the use of it in other visualization applications can be found in [15], [16], [17].

Figure 3 illustrates the application of AMOEBA clustering in our system. In Figure 3 (Left) we have the choropleth map of vandalism counts in West Lafayette, Indiana. In Figure 3 (Right) we show the results of an AMOEBA clustering. Groups are colored based on their  $G_i^*$  values, and census tracts that connect to other census tracts of the same color are considered to be a cluster.

#### IV. EXPLORING CTC WITH VALET

By using a combination of geospatial and temporal visualization and analytics tools, our system provides crime analysts with tools for real-time hypothesis generation and exploration. To better illustrate the hypothesis testing phase, we walk through a typical analysis scenario using feedback from a state police detective. During this interview, we discussed how he searches for problem areas, creates an initial hypothesis, and what steps are taken in an attempt to confirm or deny this hypothesis.

Traditionally, the first steps taken to identify problematic regions are to explore the spatial distribution of CTC incident reports aggregated over time followed by a temporal analysis to identify the time periods with peaked activity. These steps allow the analysts to better understand the nature of the incidents and their correlations with the underlying spatial and temporal elements. The initial spatial exploration of the incident reports over time usually reveals regions with higher concentrations of CTC incidents. These regions, referred to as hotspots, allow the analysts to narrow down their attention

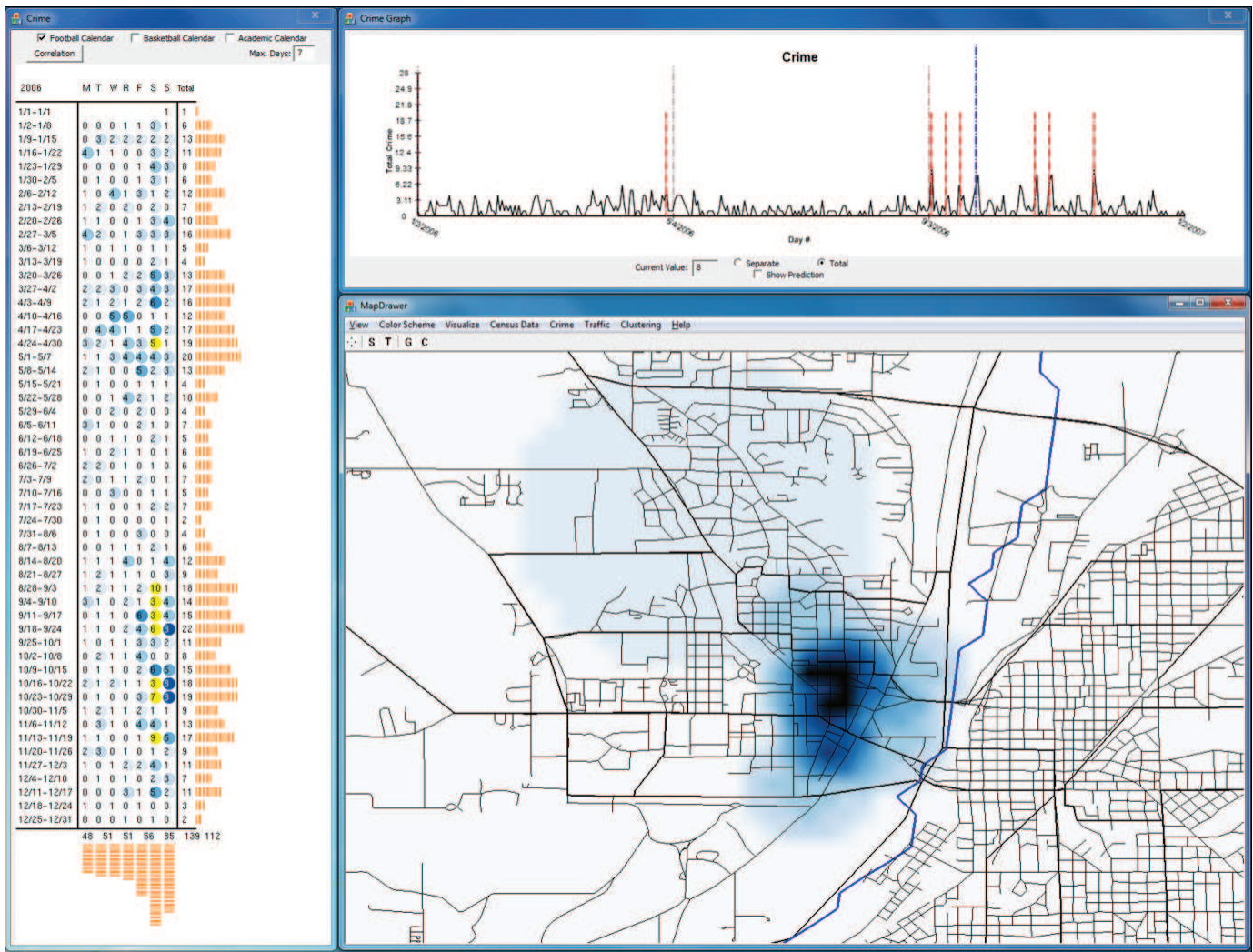


Fig. 4. Analyzing hotspots in spatial and temporal views. In the main map window, the analyst is exploring the large spatial hotspot formed by noise and vandalism complaints in West Lafayette, Indiana. The analyst hypothesizes that this hotspot may be correlated to weekend gatherings, particularly home college football games. The analyst explores this hypothesis through the use of the linked calendar view and line graph plots

and guide them through their process of hypothesis generation. Once these hotspots are identified, the analysts then explore whether these regions show a correlation with the underlying spatial elements and also whether certain temporal periods, like large community events or weekends, tend to trigger such incidents. Some CTC incidents generally tend to spike around certain geographic locations and at certain times of the year, and prior knowledge of such information can help the officials to be better prepared and equipped to deal with such situations. This process of finding hotspots followed by observing temporal trends in data generally leads the analyst to formulate a hypothesis which can then be tested over time.

We demonstrate this process of hypothesis generation and testing by analyzing an aggregate of noise and vandalism reports for West Lafayette, Indiana using our system. First, in order to narrow down the region of analysis, the analyst chooses an appropriate time aggregation level and observes the spatial hotspots form over time. A snapshot of the system in

this process can be seen in Figure 4. The dark blue color on the map indicates a hotspot for these crime types on that particular day. The analyst observes consistent hotspots form in this region over time, which shows an unusual behavior in this region for these incidents. However, the geographic knowledge of this region tells him that this region has numerous bars and clubs which may explain why this region is a hotspot for these crime types. The analyst can now hypothesize that this region could be a hotspot for several other types of crimes as well.

The analyst now focuses his attention on the temporal view of the dataset (shown by the calendar view (Figure 4- Left) and time series plot (Figure 4 - Top)) to observe temporal patterns in the selected dataset. He first notices an unusual surge in the numbers of these crime types from 1/9/2006 to 5/7/2006 and 8/14/2006 to 12/17/2006 from the weekly histogram plots on the calendar view. However, he observes that this behavior follows closely to the spring and fall semesters of the university academic year, which indicates

a higher number of people in town during this period. The analyst can further observe patterns within the semesters that closely follow the academic year calendar (for example, notice a drop in crime rates for the weeks of 3/13/2006 to 3/19/2006 (spring break) and 11/20/2006 to 11/26/2006 (thanksgiving break)). Moreover, the histogram plots by days (seen at the bottom of the calendar view) reveal that noise and vandalism complaints increase over the weekends (a fact known to the law enforcement community from experience). This yearly trend can also be observed from the time series plot (Figure 4 - Top).

In order to test whether community events (e.g. football and basketball home games) may have an influence over noise and vandalism incidents, the analyst now chooses to turn on the football home games using the check box on the calendar window. This action highlights the football events on both the calendar (indicated by the yellow colored days) and time series plot (indicated by the red dashed-lines). The analyst notices a spike in the number of crime types on the game days on the time series plot indicating a potential positive correlation between them. We also provide the analysts with the ability to click on any of the days on the calendar which brings up a dialog box showing more information about the events on that day (for example who won the football game). This allows them to further adjust their formulated hypothesis based on the outcomes of the events.

## V. CONCLUSION

Our current work demonstrates the benefits of visual analytics applied to advanced crime mapping. By linking a variety of data sources and models, we are able to enhance the hypothesis generation and exploration abilities of law enforcement officials. Our initial results show the benefits of linking traditional time-series views with spatiotemporal views for enhanced exploration and data analysis. Other future work includes advanced modeling of geo-spatiotemporal data for enhanced data exploration and hotspot detection. Furthermore, we plan to include a suite of aberration detection algorithms and their corresponding control charts for enhanced alert detection in the temporal domain. We also plan on employing spatiotemporal clustering algorithms for syndromic event detection as well as correlative analysis views within the temporal domain.

## ACKNOWLEDGMENTS

The authors would like to thank the VACCINE Public Safety Coalition which includes the Tippecanoe County Police, West Lafayette Police, Lafayette Police and the Purdue Police for providing data and valuable user feedback. We would also like to thank Bill Cleveland and Ryan Hafen for their STL development work. This work has been supported by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001.

## REFERENCES

[1] J. J. Thomas and K. A. Cook, Eds., *Illuminating the Path: The R&D Agenda for Visual Analytics*. IEEE Press, 2005.

- [2] R. Maciejewski, S. Rudolph, R. Hafen, A. Abusalah, M. Yakout, M. Ouzzani, W. S. Cleveland, S. J. Grannis, M. Wade, and D. S. Ebert, "A visual analytics approach to understanding spatiotemporal hotspots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, pp. 205–220, Mar. - Apr. 2010.
- [3] R. Maciejewski, R. Hafen, S. Rudolph, S. G. Larew, M. A. Mitchell, W. S. Cleveland, and D. S. Ebert, "Forecasting hotspots - a predictive analytics approach," *IEEE Transactions on Visualization and Computer Graphics*, To appear.
- [4] S. F. Messner and L. Anselin, "Spatial analyses of homicide with areal data," in *Spatially Integrated Social Science CD-ROM*. Oxford: Oxford University Press, 2002, pp. 127–144.
- [5] C. C. Calhoun, C. E. Stobbart, D. M. Thomas, J. A. Villarrubia, D. E. Brown, and J. H. Conklin, "Improving crime data sharing and analysis tools for a web-based crime analysis toolkit: Webcat 2.2," in *Proceedings of the 2008 IEEE Systems and Information Engineering Design Symposium*, 2008.
- [6] A. R. Gonzales, R. B. Schofield, S. V. Hart, J. E. Eck, S. Chainey, J. G. Cameron, M. Leitner, R. E. Wilson, and S. V. Hart, "Mapping crime: Understanding hot spots."
- [7] L. Anselin, "Geoda 0.9 user's guide," 2003.
- [8] N. Levine, "Crime mapping and the crimestat program," *Geographical Analysis*, vol. 38, 2006.
- [9] J. J. Van Wijk and E. R. Van Selow, "Cluster and calendar based visualization of time series data," in *INFOVIS '99: Proceedings of the 1999 IEEE Symposium on Information Visualization*. Washington, DC, USA: IEEE Computer Society, 1999, p. 4.
- [10] R. P. Hafen, D. E. Anderson, W. S. Cleveland, R. Maciejewski, D. S. Ebert, A. Abusalah, M. Yakout, M. Ouzzani, and S. Grannis, "Syndromic Surveillance: STL for Modeling, Visualizing, and Monitoring Disease Counts," *BMC Medical Informatics and Decision Making*, vol. 9, 2009.
- [11] R. B. Cleveland, W. S. Cleveland, J. McRae, and I. Terpenning, "Stl: A seasonal-trend decomposition procedure based on loess," *Journal of Official Statistics*, vol. 6, pp. 3–73, 1990.
- [12] W. S. Cleveland and S. J. Devlin, "Locally-weighted regression: An approach to regression analysis by local fitting," *Journal of the American Statistical Association*, vol. 83, pp. 596–610, 1988.
- [13] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting (2nd edition)*. Springer, 2003.
- [14] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986.
- [15] J. Aldstadt and A. Getis, "Using amoeba to create a spatial weights matrix and identify spatial clusters," *Geographical Analysis*, vol. 38, 2006.
- [16] M. Jankowska, J. Aldstadt, A. Getis, J. Weeks, and G. Fraley, "An amoeba procedure for visualizing clusters," in *Proceedings of GIScience 2008*, 2008.
- [17] R. Maciejewski, T. Drake, S. Rudolph, A. Malik, and D. S. Ebert, "Data aggregation and analysis for cancer statistics - a visual analytics approach," in *HICSS 2010*, 2010, pp. 1–5.