Chapter 16
Understanding Spatiotemporal Patterns of Multiple Crime Types with a Geovisual Analytics Approach

Diansheng Guo and Jiang Wu

Abstract Comprehensive crime data sets have been collected over time, which contain the location and time of different crime types such as aggravated assault or burglary. To understand the patterns and trends in such data, existing mapping and analysis methods often focus on one selected perspective (e.g., temporal trend or spatial distribution). It is a more challenging task to discover and understand complex crime patterns that involve multiple perspectives such as spatio-temporal trends of different crime types. In this Chapter we used a data mining and visual analytics approach to analyze the crime data of Philadelphia, PA, which has all the crimes reported from January 2007 to June 2011. Specifically, the adopted approach is a space-time and multivariate visualization system (VIS-STAMP) and the analysis examines the spatial and temporal patterns across six crime types, including aggravated assault, robbery, burglary, stolen-vehicles, rape and homicide. The geovisual analytic tool provides the capability to visualize multiple dimensions simultaneously and be able to discover interesting information through a variety of combined perspectives.

Keywords Spatial-temporal analysis • Crime analysis • Visual analytics • Data mining • Multivariate mapping

D. Guo (✉) • J. Wu
Department of Geography, University of South Carolina,
709 Bull Street, Room 127, Columbia, SC, 29208, USA
e-mail: guod@sc.edu

M. Leitner (ed.), *Crime Modeling and Mapping Using Geospatial Technologies*, Geotechnologies and the Environment 8, DOI 10.1007/978-94-007-4997-9_16,

16.1 Introduction

With the increasing academic interests in place-based crime theories since late twentieth century (Anselin et al. 2000), a large body of literature has discussed the relationship between spatial locations and crimes. Crime analyses span across a wide range of topics, such as identifying the crime concentration in a study area (Chainey and Ratcliffe 2005; Craglia et al. 2000; Eck et al. 2005; Murray et al. 2001; Ratcliffe and McCullagh 1998; Wu and Grubesic 2010), discovering the underlying social/physical factors or built environment that may account for spatial patterns of crime activities (Gorman et al. 2001), investigating theoretical roots of how space exerts influences on the crime pattern (Messner and Anselin 2004), establishing effective models used for law enforcement and legitimate prevention programs (Hunt et al. 2008; Ratcliffe 2004), and developing methodologies for spatial and statistical analyses of crime incident data (Anselin et al. 2000; Bernasco and Elffers 2010; Levine 2006).

Crime data may be divided into two major categories based on spatial representation: point data (with point locations of crime incidents) and areal data (crime incident counts aggregated to predefined boundaries). Point crime data can be converted (aggregated) to areal data when needed. To visualize and analyze point patterns, commonly used methods include quadrat count, distance statistics, and kernel density estimation (KDE). Some regard KDE as the most suitable spatial analysis technique for visualizing point data (Chainey et al. 2008; Eck et al. 2005). For visualizing areal crime patterns, choropleth map, scatterplot and/or variogram coupled with classic statistics including Moran's I and Geary's C (Cliff and Ord 1970), distance-based statistics and LISA statistics (Anselin 1995) are frequently applied.

In addition to spatial distribution patterns, the temporal trend of crimes at various temporal scales is also of critical interest to both researchers and law enforcement (Felson and Poulsen 2003; Henry and Bryan 2000; Rengert 1997; Townsley 2008; Townsley et al. 2000; Weisburd et al. 2004; Weisburd et al. 2009). For example, Ratcliffe and McCullagh (1998) propose a framework called 'aoristic crime analysis' to detect spatio-temporal crime patterns where the exact offense time is unknown (Ratcliffe 2000, 2002). To visualize spatio-temporal patterns of crime activities, Brunsdon et al. (2007) compare and evaluate three major techniques, including map animation, the COMAP (Brunsdon 2001) and isosurfaces. In addition, hotspot plot (Townsley 2008), space-time cube (Nakaya and Yano 2010) and CrimeViz (Roth et al. 2010) provide alternative approaches for space-time analysis of crimes.

It has also been recognized that it is important to include additional factors, such as crime types or socioeconomic environment of the crime neighborhood, to understand the context and underlying process that influence the spatial and temporal variation of crime activities (Hagenauer et al. 2011). However, there are not many methods that can simultaneously visualize and analyze the spatial, temporal, and multivariate dimensions related to crime activities. Hangenauer et al. (2011) propose a framework to identify the spatial and temporal characteristics of crime patterns by incorporating the socioeconomic and environmental attributes of the neighborhoods, which consists of three steps. First, a spatial scan statistic is applied to identify significant spatio-temporal crime hotspots. Second, a self-organizing map (SOM) is used to cluster neighborhoods based on their social-economic and environmental attributes. Finally, the hotspots are mapped in the SOM visualization so that one may see the correlation between crime hot spots and contextual factors. The main limitation of this approach is that the visualization (i.e., U-matrix and Component Planes) in the framework cannot display or link to actual spatial locations and therefore it is difficult to perceive spatial distribution patterns or correlations. Additionally, as noted by Hagenauer et al. (2011), different crime types often exhibit different patterns in relation to space, time, and context, and therefore should be separated and compared in analysis, which is not addressed in the framework.

In the research field of geovisualization and visual analytics, it is an active and challenging research topic to develop new approaches for visualizing complex datasets that contain geographic locations, time series, and multiple variables (Andrienko et al. 2010; Guo et al. 2006). To map and visualize patterns across multiple variables and dimensions, dimension reduction techniques are often used. Guo et al. (2005) developed a multivariate mapping approach (named SOMVIS) that integrates self-organizing map (a dimension reduction and clustering method), color encoding, and multidimensional visualization to map multiple variables in a single map. Guo et al. (2006) extended the SOMVIS approach to accommodate the temporal dimension and visualize spatio-temporal and multivariate information simultaneously. This new approach is called VIS-STAMP, an acronym for Visualization for Space Time and Multivariate Patterns. VIS-STAMP first constructs an *overview* of the data, from which the analyst can easily perceive complex patterns across all dimensions and then explore specific patterns through user interactions such as selection and linking different views. In this chapter, we apply VIS-STAMP to the analysis of crime data in Philadelphia (2007–2011) that involves spatial, temporal, and multivariate information. We will briefly introduce the VIS-STAMP approach in Sect. 16.3.

This chapter is organized as follows. Section 16.2 introduces the study area and the crime data. Section 16.3 briefly explains the methodology and how the data were processed and prepared for the analysis. The analysis results are presented in Sect. 16.4, with subsections focusing on different types of patterns being discovered from the data. We provide discussions on the methodology, analysis, and future directions in Sect. 16.5.

16.2 Study Area and Data

Philadelphia County, PA covers approximately 143 square miles with a population of 1,526,006 according to U.S. Census 2010. Crime incident reports are available from the online service provided by the Philadelphia Police Department (http://citymaps.phila.gov/CrimeMap/StepByStep.aspx). There were in total 169,829 crime incidents for six crime types (aggravated assault, robbery, burglary, stolen vehicle, rape and

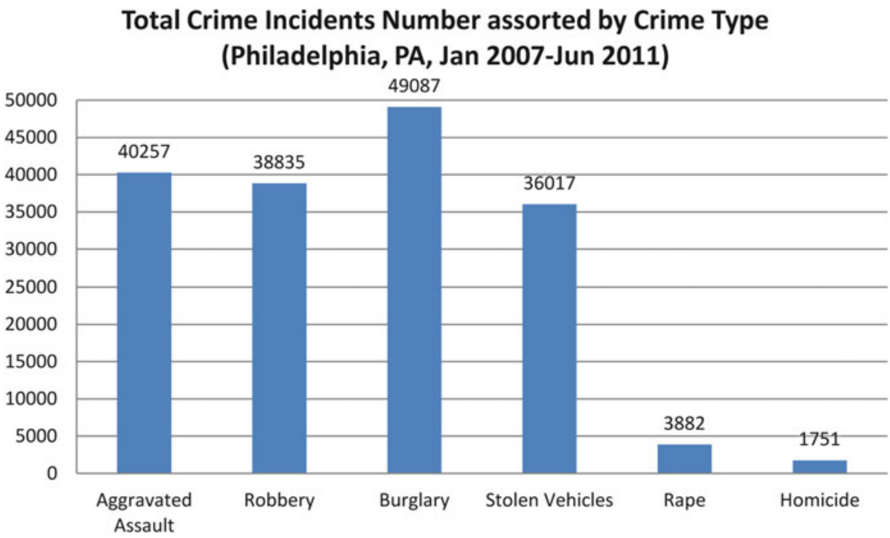

Fig. 16.1 Total crime incidents by crime type (Philadelphia, PA, Jan 2007–June 2011)

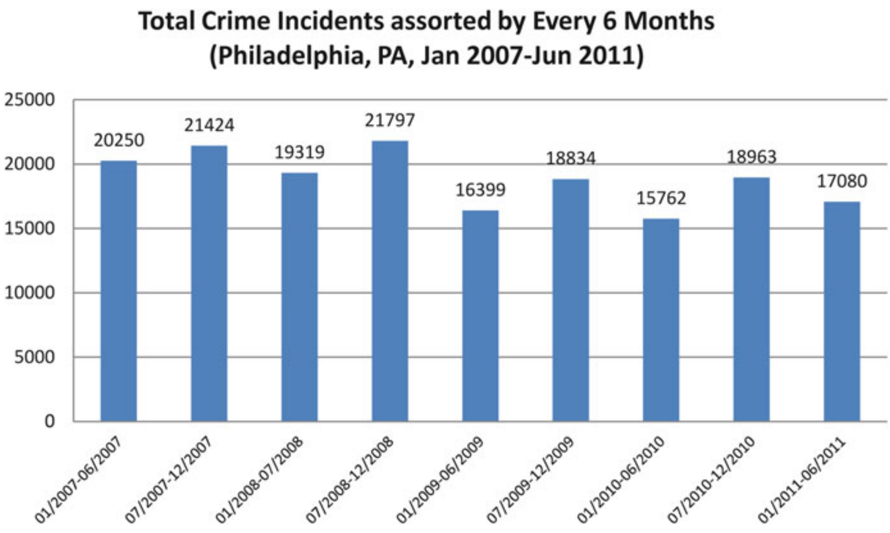

Fig. 16.2 Total crime incidents for every 6 months, since Jan. 2007

homicide) from January 2007 to June 2011. Each crime incident record has a type, date, police dispatch time, and its geographic location (x, y coordinates).

Figure 16.1 shows the total number of crime incidents for each crime type. Aggravated assault, robbery, burglary, and stolen vehicles are the four most frequent types, and rape and homicide have much fewer incidents. Figure 16.2 shows the distribution of crimes (of all six types) across time, aggregated for every six months.

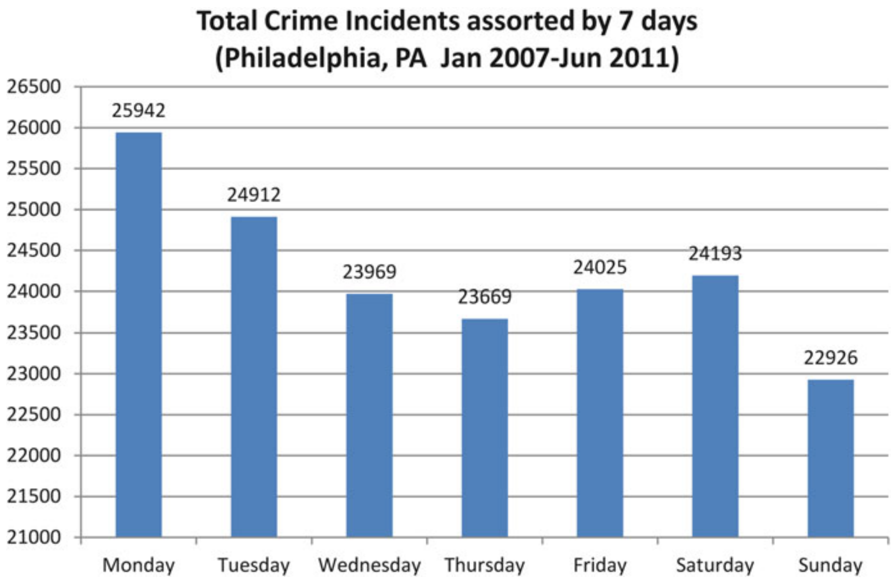

Fig. 16.3 Total crime incidents number assorted by 7 days per week

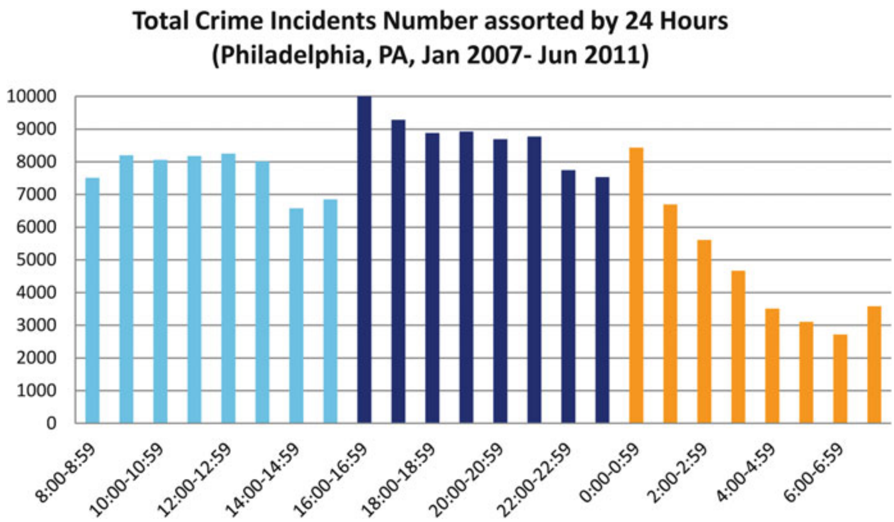

Fig. 16.4 Total crime incidents number assorted by 24 h per day

The annual crime totals dropped a little bit after year 2008. For each year, there appeared to be more crimes in the second half (Jul–Dec) than in the first half (Jan–Jun). In addition to these general trends over years, we also explored the trends over 7 days of a week (Fig. 16.3) and 24 h in a day (Fig. 16.4). Figure 16.3 shows an interesting trend over the 7 ways in a week, with Monday and Tuesday having the

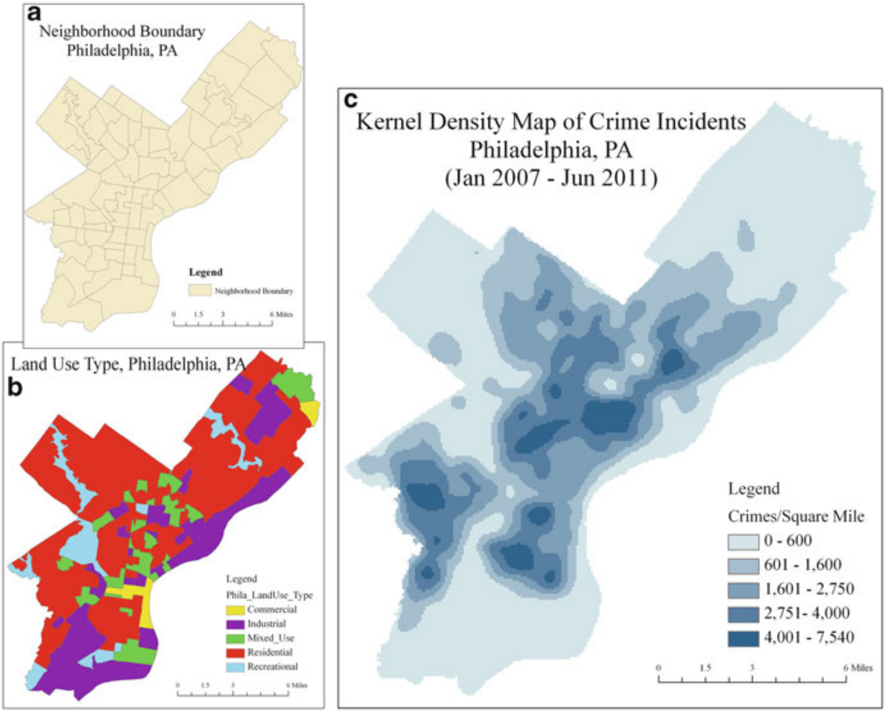

Fig. 16.5 (**a**) NIS neighborhood boundaries (Neighborhood Information System, http://cml.upenn.edu/nis/), Philadelphia, Pennsylvania. (**b**) Land use types aggregated from census tracts. (**c**) Crime density, including 169,829 crime incidents from Jan. 2007 to June 2011

most crimes while much less on Sunday. Figure 16.4 shows the aggregated crime counts for each hour, regardless of day and type. Based on three abrupt surges in crimes as shown in Fig. 16.4, in our analysis we divided a day into three 8-h time periods: Early Morning (0:00–7:59), Day Time (8:00–15:59) and Evening/Night (16:00–23:59).

The overall spatial distribution of the 169,829 crimes is presented by a kernel density map (Fig. 16.5c). To help understand the context of the crime locations, five major land use types (commercial, industrial, mixed use, residential, recreational) are mapped for the study area (Fig. 16.5b), which are extracted and generalized based on the Philadelphia Zoning Code. The land use type data is aggregated with census tracts and parcel data. If the area of a specific land use type exceeds 50% of the total properties within a tract, this particular land use type is assigned to the entire tract. If none of four types covers more than 50% of a tract, the tract is deemed ‘Mixed Use’. The land use map (Fig. 16.5b) shows that the residential area covers most of the western and northern Philadelphia. Recreational areas mainly include public waterfront parks and sport facilities in the county. The water front belt in the northeast and the lower south are occupied by industrial area.

The research question is that, although Figs. 16.1, 16.2, 16.3, 16.4, and 16.5 can help understand crime patterns or trends from a specific perspective (e.g., crime type, time, or space), it is difficult to examine the data across multiple perspectives, such as the patterns of different crime types and their distribution and trends over space and time. For example, Fig. 16.5c may show the spatial concentration of crimes, but it cannot reveal the characteristics of crimes (such as composition of crime types) or temporal trends at different places. In this research, we use the VIS-STAMP method to gain insights on crime patterns that involve multiple perspectives. We use 69 neighborhoods (see Fig. 16.5a for a map of neighborhoods and detailed definition at http://cml.upenn.edu/nbase/nbAbout.asp) as the spatial units, along time and crime types, to aggregate crimes for further analysis (see Sect. 16.3 for details on different data aggregations). We use neighborhoods as the base units for two reasons. First, each neighborhood has a meaningful community definition that directly supports policy analysis and planning efforts in the city. Second, neighborhoods are sufficiently large to examine its internal crime characteristics across several dimensions such as time or crime type.

16.3 Data Preprocessing and Methodology Overview

16.3.1 Data Aggregation and Preprocessing

The input data is a set of 169,829 geocoded crime incidents, which are aggregated into a *data cube* and transformed in different ways depending on the analysis task. Figure 16.6 shows an illustrative view of the data cube. The three dimensions in the data cube include: *spatial dimension* (69 neighborhoods), *temporal dimension* (which can be of three different temporal scales), and *multivariate dimension* (e.g., crime types). The three temporal scales are semi-annual (i.e., half-year periods as shown in Fig. 16.2), day of the week (regardless of month and year, as shown in

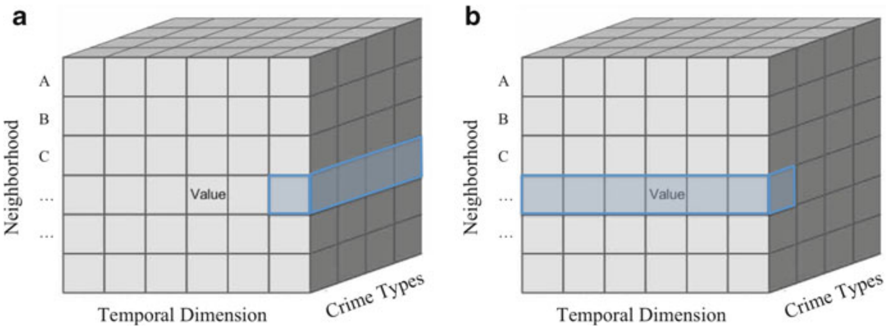

Fig. 16.6 The data cube, which is a space-time-attribute aggregation of crime data: (**a**) A composition of crime types for the same place and time is highlighted. (**b**) A time series for a place and crime type is highlighted

Fig. 16.3), and hour of the day (including three 8-h periods, regardless of day, month, and year, as shown in Fig. 16.4). Other temporal scales can be easily accommodated as well. For the multivariate dimension, additional crime-related variables (such as offender's age, offender's modus operandi, victim's information, etc.) can be used, which unfortunately are not available in the Philadelphia data. Each *cell* in the data cube is a unique combination of a spatial unit (e.g., neighborhood A), a time period (e.g., Monday), and a crime type (e.g., aggravated assault). The value for a cell is the total number of crimes in the cell, e.g., the total number of aggravated assaults occurred in neighborhood A on Monday.

Once the data cube is constructed as described above, it should be normalized or standardized, which depends on the analysis tasks. In this chapter, we primarily focus on two different analysis tasks:

1. <u>Task 1: How crime characteristics vary across space and change over time</u>. Figure 16.6a highlights a sequence of crime counts (i.e., a composition of different crime types) for a specific neighborhood at a specific time. We want to understand how such crime compositions change from place to place and from time to time. For this type of analysis, we normalize the data cube by dividing the crime count of each cell in the sequence by the total crimes of the sequence. In other words, after the normalization each cell value becomes a percentage value representing the proportion of each crime type in the sequence (for a specific neighborhood and at a specific time). There two reasons for choosing this standardization procedure instead of dividing the crime composition by area size or population. First, crime activities are not necessarily related to population density or area of predefined boundaries. Second, we are more interested in the crime composition (or characteristics) of a place and time.
2. <u>Task 2: How temporal trends differ in different places and for different crime types.</u> This is to look at the data cube from a different perspective by focusing on the time series, one for each neighborhood/crime type combination. Figure 16.6b highlights a time series for a specific neighborhood and a specific crime type. Here we want to understand how the temporal trend varies across space and crime types. For example, one type of crime may have been increasing while another type(s) of crime(s) may have been declining for some part of the city. For this purpose, we divide the crime count in each cell in a time series by the total crime count of the time series. In other words, each cell value now becomes a percentage value representing the proportion of crimes in each time period for a specific neighborhood and crime type.

Given a normalized data cube, the VIS-STAMP system will view it differently depending on the analysis task. For the first task described above, VIS-STAMP treats the cube as a set of multivariate vectors arranged in a space-time matrix (see Fig. 16.6a), where a multivariate vector is a sequence of percentage values representing proportions of crime types for a place and time. For the second analysis task, VIS-STAMP treats the cube as a set of time series arranged in a space-crime matrix (see Fig. 16.6b), where each time series is a set of percentage values representing the temporal trend of crimes for a place and crime type. VIS-STAMP performs clustering with the set of vectors (which are either multivariate vectors or time series) and visualizes them across two other dimensions.

16.3.2 Multivariate Mapping and Space-Time-Attribute Visualization

The VIS-STAMP approach extends the Self-Organizing Map (SOM) (Kohonen 2001) to extract clusters from the set of multivariate vectors, project the clusters onto a two-dimensional space, and use a 2D color scheme (Brewer 1994) to color each cluster so that similar clusters have similar colors. Clusters are visualized in map(s) and other visual representations, such a re-orderable space-time matrix. Each cluster and the data items in the cluster are of the same color (assigned by the SOM) in all visual components. In other words, similar colors in a map represent clusters of similar multivariate vectors. A parallel coordinate plot (PCP) is used as the 'legend' to show the multivariate meaning that each color represents. Details on the methodology can be found in (Guo et al. 2005, 2006). Below we use an example analysis to help explain the methodology.

Figure. 16.7 presents a multivariate map of the crime data described in the previous section, *without considering temporal variations*. In other words, this is a special case of the cube in Fig. 16.6a, where there is only one time period (entire 54 months, January 2007– June 2011). Therefore, each neighborhood has a multivariate vector, representing the composition of different crime types in the neighborhood. A weight can be assigned for each variable. The analysis involves five variables, including four major crime types (i.e., aggravated assault, robbery, burglary, and stolen vehicles) and the Median Household Income for each neighborhood. The goal is to see the crime compositions in different neighborhoods and their possible relation with income level. We included the Median Household Income in the PCP visualization but did not use it in the SOM clustering so that clusters are constructed independent of income levels. Due the limitation of the software, which does not allow a zero weight, we assigned an extremely small weight to Median Household Income to effectively exclude it in the clustering step. All other variables, i.e., crime types, are assigned the same weight (see Fig. 16.7a).

SOM groups the 69 neighborhoods into 38 clusters based on their multivariate vectors (i.e., crime compositions). SOM arranges the clusters with a 2D U-matrix so that similar clusters are close to each other (Fig. 16.7b) without considering their locations. Based on a 2-D color scheme, each cluster is assigned a color so that similar clusters have similar colors (Guo et al. 2005). The parallel coordinate plot (PCP, see Fig. 16.7c) shows the mean vector of each cluster. Each axis in the PCP represents a variable and uses a nested-means scaling (Guo et al. 2005), which puts the mean value of that variable at the midpoint of its axis. The PCP also provides several other linear scaling options. Each cluster is a string in the PCP, with the same color as it has in the SOM. The width (thickness) of the string represents the number of neighborhoods in the cluster, i.e., larger clusters (having more neighborhoods) is represented with a wider string.

Fig. 16.7 A multivariate map of four major crime types. (**a**) Data processing and configuration. (**b**) Self-Organizing Map, where the each circle is a cluster. (**c**) PCP. (**d**) Multivariate map

From the PCP in Fig. 16.7c, for example, we can tell that the reddish clusters represent neighborhoods with a relatively high proportion of crimes being aggravated assault (about 30%) and a low median household income (less than 20 k/year). The map (Fig. 16.7d) shows each neighborhood in the same color as that of the cluster that the neighborhood belongs to. With the PCP and the map, we understand not only the crime compositions of the reddish clusters but also where they are in the geographic space (i.e., concentrating in West Philadelphia, such as Strawberry Mansion, Mill Creek and Haddington). The greenish clusters represent neighborhoods with primarily burglary crime threat (which accounts for nearly 50% of the crimes in those neighborhoods) and such neighborhoods mainly locate in the northwestern and northern parts of Philadelphia, which have a higher median household income (more than 40 k/year). Similarly, one can interpret the meaning of other clusters, such as the purple and blue clusters, in both the geographic and multivariate spaces.

To add the time dimension to the analysis, each neighborhood will have a unique multivariate vector for each time period (for the first analysis task) or a unique time series for each crime type (for the second analysis task). VIS-STAMP will then group similar crime compositions or time series into clusters. Essentially, the set of multivariate vectors or time series are reduced to a set of clusters and encoded in colors, which are visualized in other components such as maps (with one map for each time period or for each crime type) and a re-orderable matrix, where the vertical dimension is ordered by spatial units and the horizontal dimension is either the set of time periods or the set of crime types. We will explain this in detail in the next section with a variety of analysis results.

16.4 Analysis Results

16.4.1 Spatio-Temporal Patterns of Crime Compositions (Task 1)

We first examine how crime compositions change across space (neighborhoods) and time (every six months). Figure 16.8 shows the result with VIS-STAMP, which includes a re-orderable matrix (and in this case it can be called a space-time matrix), a map matrix (each map represents a crime composition map for a time period), and a PCP. In the re-orderable matrix, the rows represent the 69 neighborhoods and the columns represent nine 6-month periods from January 2007 to June 2011. The rows are ordered so that similar neighborhoods in terms of crime compositions over time are next to each other. Columns are in the naturally temporal order. Each column in the space-time matrix corresponds to a map in the map matrix. In other words, the re-orderable matrix and the map matrix show the same data from two different perspectives, with the former focusing on revealing temporal patterns while the latter focusing on spatial and spatio-temporal patterns. Essentially, with colors representing multivariate information (i.e., crime compositions in this case), the re-orderable matrix or the map matrix is an *overview* of the data cube. To ensure meaningful analysis, if a cell in the data cube has less than 20 crime incidents, the cell (i.e., the multivariate vector) will be excluded from the analysis and its corresponding elements in the map matrix or re-orderable matrix will be colored in gray.

In Fig. 16.8, the analysis again uses the four major crime types, same as in Fig. 16.7. The difference is that, in Fig. 16.7, there are 69 multivariate vectors (i.e., crime compositions) while here there are $69 \times 9 = 621$ vectors. From the PCP, one can understand the meaning of each cluster (and thus the meaning of each color). For example, reddish clusters represent a composition with a high proportion of aggravated assault, average on robbery and low on stolen vehicle. In the

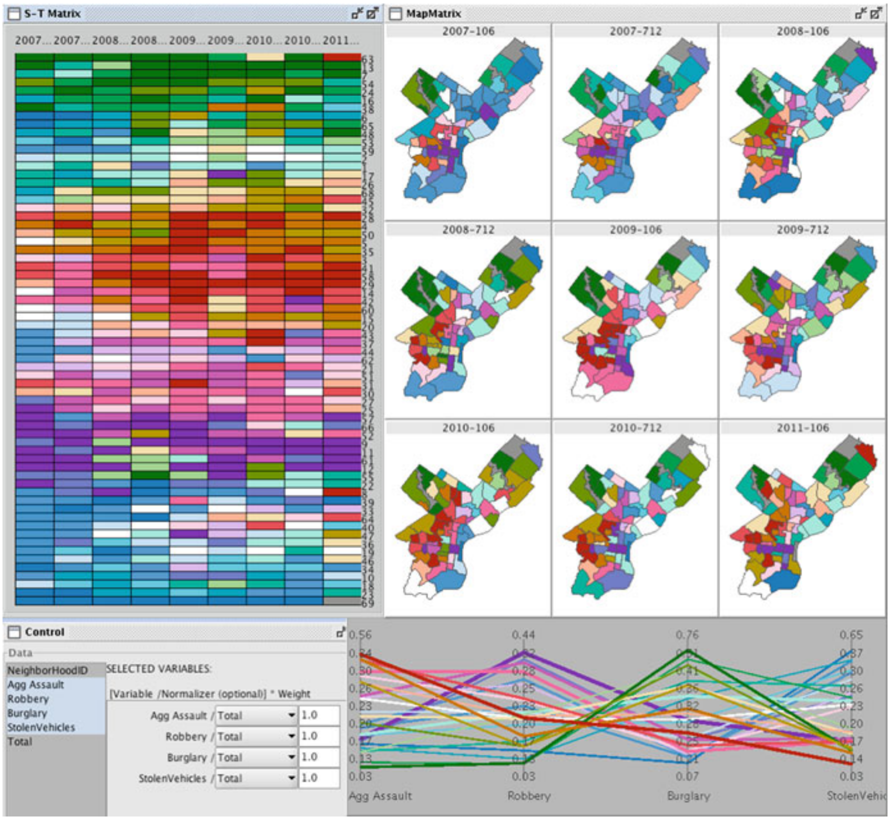

Fig. 16.8 Multivariate crime patterns across space and time (every 6 months from Jan 2007 till June 2011). The view includes a re-orderable matrix (*top-left*), map matrix (*top-right*), self-organizing map (not shown, see Fig. 16.7b from an example), and a parallel coordinate plot (*bottom-right*)

re-orderable matrix, it is clear that there are more reddish cells after 2008. In the map matrix, it is also evident that reddish neighborhoods are expanding. On the other hand, bluish (from light blue to navy blue) clusters represent high proportion of stolen-vehicle crimes while relatively less of other crimes. From both the re-orderable matrix and the map matrix, it is evident that the number of bluish neighborhoods had decreased over time.

To take a closer look at these two opposite trends, one can select the two groups of clusters in the PCP (see Fig. 16.9). In other words, both the reddish clusters and the bluish clusters are highlighted. Now it is very easy to perceive these two types of patterns described above, involving space, time, and crime types. From Jan 2007 to June 2011, the neighborhoods with high percentage of stolen vehicle crimes are shrinking dramatically, especially since Jan 2009. On the opposite, the threat of aggravated assault surged since January 2009. Comparing Figs. 16.9 and 16.5b, we may explore the possible relation between crime types and land uses. For example,

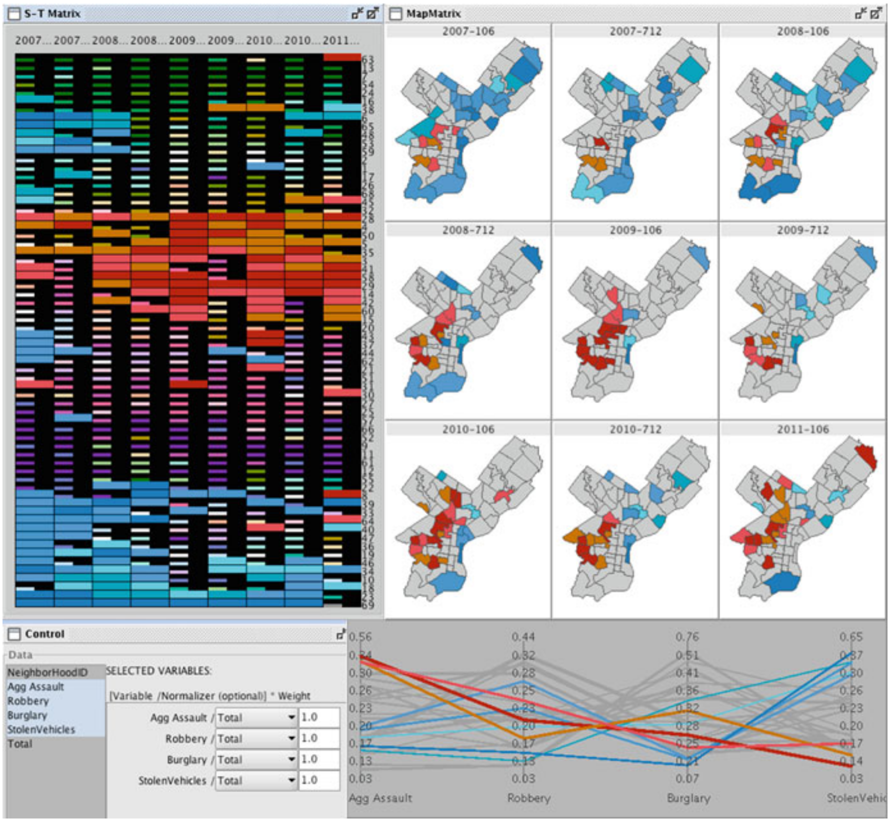

Fig. 16.9 Aggravated assault and stolen vehicles exhibit opposite spatio-temporal trends. Bluish clusters on average have at least 30% of their crimes being stolen vehicle, while reddish clusters are those with at least 30% being aggravated assault. While bluish clusters have been diminishing over time and space, reddish clusters are expanding and growing

neighborhoods with high percentage of stolen vehicles are primarily of residential and industrial land use types. Through interactive exploration, one may also examine many other patterns that are present in the overview (Fig. 16.8).

By changing the temporal scale to seven days of the week (regardless of month and year), we can discover different patterns from the data. Figure 16.10 shows the result of the same data except that it uses seven weekdays on the temporal dimension. The clusters and colors for this analysis are very similar to those in the previous analysis, e.g., reddish clusters are dominated by aggravated assault and bluish clusters mainly represent stolen vehicles activities. It is interesting to see that reddish clusters grow/expand steadily from Monday to Sunday, with weekend days threatened most by such violent crimes. This specific pattern becomes even more evident if we select the reddish clusters with 25% or more crimes being aggravated assault (Fig. 16.11). Both the re-orderable matrix and the map matrix show that

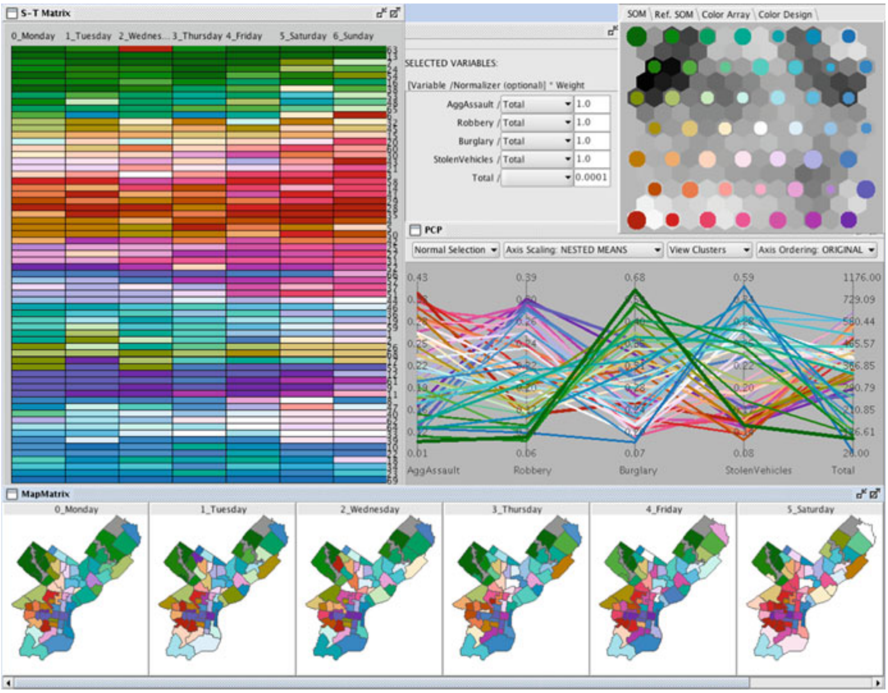

Fig. 16.10 Crime composition patterns across space and 7 days a week

spatial concentration and escalating temporal trend of the crime. Note that, in Fig. 16.11, the PCP shows each individual vector instead of the clusters, which is an option that the user can choose.

One may also focus on the greenish clusters, which represent a high percentage of burglary crimes, which exhibit a different trend, i.e., more incidents during the week days (Monday-Friday) than weekends. As for robbery crimes (in purple), it shows a persistent spatial concentration at the downtown area of Philadelphia but no noticeable temporal variation. One may also use 24 h or several time periods of a day to analyze daily patterns of crimes. Due to limited space, we do not include these analyses here.

16.4.2 Spatial-Crime Differences of Temporal Trends (Task 2)

VIS-STAMP can also treat time series as “multivariate vectors” and support the analysis of temporal trends and their differences in space and for different crime types. As introduced in Sect. 16.3.2, this is a different view of the same data cube as used in the previous analyses, with a slightly different normalization procedure. Each time series is normalized so that each value represents the percentage of total crimes (for a specific spatial unit and a crime type) in each time period. For example,

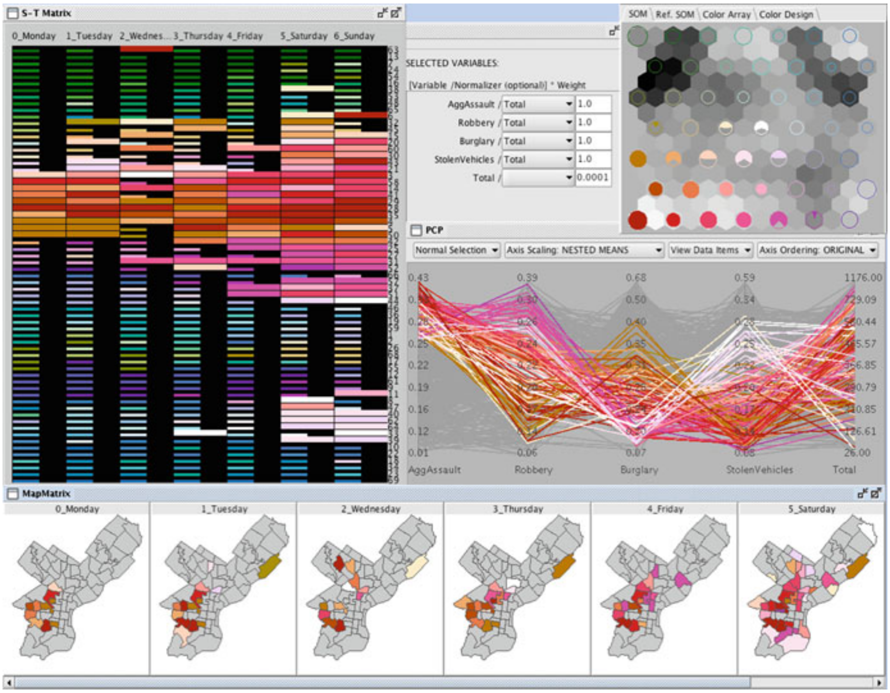

Fig. 16.11 Space-time trend of aggravated assaults (>25 %) over seven weekdays

neighborhood A had 100 robbery crimes, out of which 10 crimes occurred for time period T1, then the value for T1 is 10% for this time series for neighborhood A and robbery crime.

VIS-STAMP extract clusters with all time series and assign a color to each cluster. As such, similar colors now represent similar temporal trends. Figure 16.12 shows the analysis result with Philadelphia data, with four major crime types and nine 6-month periods. The PCP shows the clusters of time series, with each axis representing one time period. For example, a dark green cluster represents a declining trend, with more crimes in earlier times than in later time periods. The reddish clusters, on the other hand, represent a surging trend with more crimes in recent time periods. Each map in the map matrix shows the overall spatial distribution of temporal trends for a specific crime type. For example, it is obvious that the stolen vehicle crime has dropped significantly over time, as evident in its map that is dominated by greenish colors. Conversely, aggravated assault and burglary crimes have been rising lately, as shown in the “reddish” maps for both crime types. If we take a closer look at the two maps, we can also notice the spatial differences in temporal trends, with some neighborhoods getting better while others getting worse. For example, burglary surged in recent months for some neighborhoods in the central and the northeast portion of Philadelphia, while the northwest in purples and blues has improved lately (relative to their worse time around 2008 and 2009). Blueish and purplish clusters represent trends with a peak in the middle (in 2008

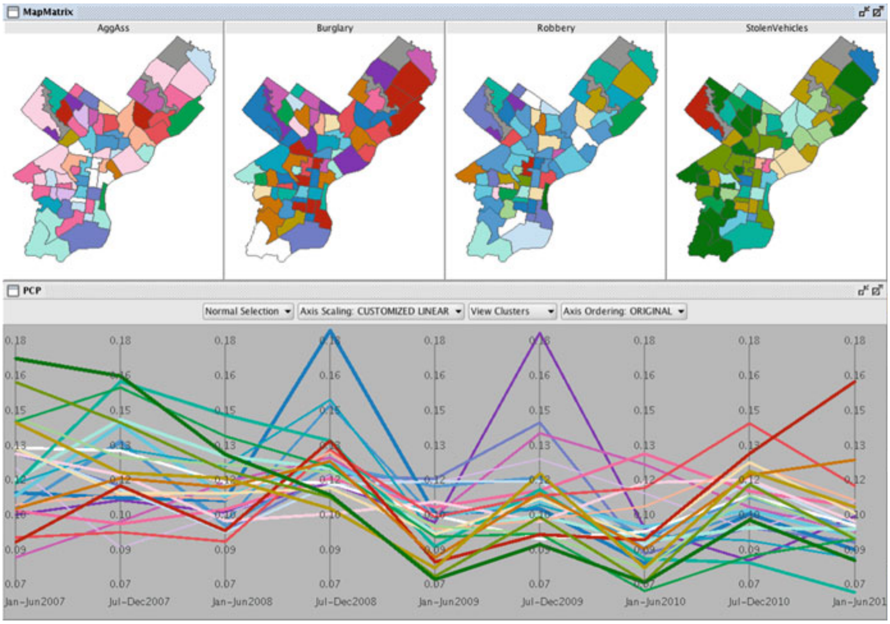

Fig. 16.12 Temporal trends vary across space and crime types. Different from the PCP in other figures, the PCP here uses a linear scaling with fix minimum and maximum values on each axis so that the height of a curve can be compared

and 2009) but relatively low crimes recently. Therefore, the robbery map indicates that robbery crimes have dropped as well, relative to the peak 2 or 3 years ago.

There is also an obvious temporal oscillation pattern shown in Fig. 16.12, where the second half (Jul–Dec)of a year often has more crimes than in the first half (Jan–Jun). Similar to the diagram presented in Fig. 16.12, we can also change the temporal scale to weekdays or hours of the day to analyze difference temproal patterns and their variation over space and crime types. One can also examine the patterns interactively through selection and linking. Due to space limitation, we will not present more and different analysis results.

16.5 Discussions and Conclusion

This Chapter presents an exploratory approach to discover and understand complex crime patterns that involve multiple perspectives such as spatio-temporal trends across different crime types. The VIS-STAMP approach is adopted to analyze the crime data in Philadelphia reported from Jan. 2007 to June 2011. The analyses focus on four major crime types: aggravated assault, robbery, burglary, and stolen-vehicles. All crimes are aggregated into a data cube with space (69 neighborhoods), time, and crime types as the three dimensions. VIS-STAMP can effectively construct an overview of the major patterns in the data cube, allowing the analysis and understanding of complex patterns across all dimensions and supporting interactive exploration of specific patterns through highlight selection and multiple-view linking. A variety of interesting patterns have been found in the Philadelphia crime data, including spatio-temporal variations of different crimes and the shifting temporal trends across space and crime types. Compared to conventional methods such as density mapping or temporal analysis alone, the VIS-STAMP approach provides an alternative way to discover more complex patterns across multiple perspectives.

The crime data is originally a point data set, which is converted and aggregated to an areal data set based on neighborhoods and other dimensions. This may be a limitation since it reduces the data resolution by using a predefined set of boundaries. An alternative solution, which we do not include in the analysis, may be to create a raster density surface for each crime type and for each time period, and then treat each raster pixel as a "spatial unit" in subsequent analysis. However, this approach has its own limitations, such as the excessive spatial autocorrelation introduced by the kernel density estimation and the uncertainty in the "interpolated" data especially when there are no sufficient data points for certain locations and time periods. Although crimes may be influenced by different factors at different scales, neighborhood is a reasonable choice in examining crime patterns in Philadelphia for two reasons, as we explained earlier. First, neighborhoods are meaningful communities that are directly related to policy analysis and planning efforts in the city. Second, neighborhoods are sufficiently large to examine its internal crime pattern across several dimensions such as time or crime type. Choosing a suitable spatial scale is important in using the VIS_STAMP approach in order to avoid statistically unstable crime measures (such as percentages), which requires that each areal unit contains a sufficient number of crime incidents.

As an exploratory approach, VIS-STAMP currently lacks rigorous statistical testing procedures to evaluate the significance level of discovered patterns such as clusters and trends. Future work may integrate statistical testing through Monte Carlo simulation to assess patterns, in addition to visual exploration. Findings discovered through VIS-STAMP and confirmed with subsequent testing may help in crime modeling and prediction and related policy making for crime control. Given the complexity, unknown factors, and time-varying characteristics of crimes, data driven and exploratory approaches are indispensible for understanding crime data and providing timely information for response.

Acknowledgments This work was supported in part by the National Science Foundation under Grant No. 0748813. We greatly appreciate the constructive comments and suggestions from anonymous reviewers.

References

Andrienko G, Andrienko N, Bremm S, Schreck T, Von Landesberger T, Bak P, Keim D (2010) Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns. Comput Graph Forum 29:913–922. doi: 10.1111/j.1467-8659.2009.01664.x

Anselin L (1995) Local indicators of spatial association—LISA. Geogr Anal 27:93–115

Anselin L, Cohen J, Cook D, Gorr W, Tita G (2000) Spatial analyses of crime. Crim Justice 4:213–262

Bernasco W, Elffers H (2010) Statistical analysis of spatial crime data. In: Piquero AR, Weisburd D (eds) Handbook of quantitative criminology. Springer, New York, pp 699–724

Brewer CA (1994) Color use guidelines for mapping and visualization. Visual Mod Cartogr 2:123–148

Brunsdon C (2001) The comap: exploring spatial pattern via conditional distributions. Comput Environ Urban Syst 25:53–68

Brunsdon C, Corcoran J, Higgs G (2007) Visualising space and time in crime patterns: a comparison of methods. Comput Environ Urban Syst 31:52–75

Chainey S, Ratcliffe J (2005) GIS and crime mapping. Wiley, Chichester

Chainey S, Tompson L, Uhlig S (2008) The utility of hotspot mapping for predicting spatial patterns of crime. Secur J 21:4–28

Cliff AD, Ord K (1970) Spatial autocorrelation: a review of existing and new measures with applications. Econ Geogr 46:269–292

Craglia M, Haining R, Wiles P (2000) A comparative evaluation of approaches to urban crime pattern analysis. Urban Stud 37:711–729

Eck JE, Chainey S, Cameron JG, Leitner M, Wilson RE (2005) Mapping crime: understanding hot spots. In: NIJ special report. https://www.ncjrs.gov/pdffiles1/nij/209393.pdf

Felson M, Poulsen E (2003) Simple indicators of crime by time of day. Int J Forecast 19:595–601

Gorman DM, Speer PW, Gruenewald PJ, Labouvie EW (2001) Spatial dynamics of alcohol availability, neighborhood structure and violent crime. J Stud Alcohol 62:628

Guo D, Gahegan M, MacEachren AM, Zhou B (2005) Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach. Cartogr Geogr Inf Sci 32:113–132

Guo D, Chen J, MacEachren AM, Liao K (2006) A visualization system for space-time and multivariate patterns (VIS-STAMP). IEEE Trans Vis Comput Graph 12:1461–1474

Hagenauer J, Helbich M, Leitner M (2011) Visualization of crime trajectories with self-organizing maps: a case study on evaluating the impact of hurricanes on spatio-temporal crime hotspots. In: Proceedings of the 25th conference of the International Cartographic Association, Paris

Henry LM, Bryan BA (2000) Visualising the spatio-temporal distribution of motor vehicle theft in Adelaide, South Australia. In: National Centre for Social Applications of GIS (GISCA). http://pandora.nla.gov.au/nphwb/20010320130000/http://www.aic.gov.au/conferences/mapping/henry.pdf

Hunt ED, Sumner M, Scholten TJ, Frabutt JM (2008) Using GIS to identify drug markets and reduce drug-related violence. In: Thomas YF, Richardson D, Cheung I (eds) Geography and drug addiction. Springer, Dordrecht

Kohonen T (2001) Self-organizing maps. Springer, Berlin/New York

Levine N (2006) The CrimeStat program: characteristics, use and audience. Geogr Anal 38:41–56

Messner SF, Anselin L (2004) Spatial analyses of homicide with areal data. In: Janelle DG, Goodchild MF (eds) Spatially integrated social science. Oxford University Press, Oxford

Murray AT, McGuffog I, Western JS, Mullins P (2001) Exploratory spatial data analysis techniques for examining urban crime. Br J Criminol 41:309–329

Nakaya T, Yano K (2010) Visualising crime clusters in a space time cube: an exploratory data analysis approach using space time kernel density estimation and scan statistics. Trans GIS 14:223–239

Ratcliffe JH (2000) Aoristic analysis: the spatial interpretation of unspecific temporal events. Int J Geogr Inf Sci 14:669–679

Ratcliffe JH (2002) Aoristic signatures and the spatio-temporal analysis of high volume crime patterns. J Quant Criminol 18:23–43

Ratcliffe JH (2004) Crime mapping and the training needs of law enforcement. Eur J Crim Policy Res 10:65–83

Ratcliffe JH, McCullagh MJ (1998) Aoristic crime analysis. Int J Geogr Inf Sci 12:751–764

Rengert GF (1997) Auto theft in central Philadelphia. In: Homel R (ed) Policing for prevention: reducing crime, public intoxication and injury. Criminal Justice Press, Monsey, p 7

Roth R, Ross K, Finch B, Luo W, MacEachren A (2010) A user-centered approach for designing and developing spatiotemporal crime analysis tools. In: Proceedings of GIScience, Zurich, Switzerland

Townsley M (2008) Visualising space time patterns in crime: the hotspot plot. Crime Pattern Anal 1:61–74

Townsley M, Homel R, Chaseling J (2000) Repeat burglary victimisation: spatial and temporal patterns. Aust N Z J Criminol 33:37–63

Weisburd D, Bushway S, Lum C, Yang SM (2004) Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. Criminology 42:283–322

Weisburd D, Morris N, Groff E (2009) Hot spots of juvenile crime: a longitudinal study of arrest incidents at street segments in Seattle, Washington. J Quant Criminol 25:443–467

Wu X, Grubesic T (2010) Identifying irregularly shaped crime hot-spots using a multiobjective evolutionary algorithm. J Geogr Syst 12:409–433