Chapter 4
Crime Analysis at Multiple Scales of Aggregation: A Topological Approach

Patricia L. Brantingham, Paul J. Brantingham, Mona Vajihollahi, and Kathryn Wuschke[1]

Abstract Patterns in crime vary quite substantially at different scales of aggregation, in part because data tend to be organized around standardized, artificially defined units of measurement such as the census tract, the city boundary, or larger administrative or political boundaries. The boundaries that separate units of data often obscure the detailed spatial patterns and muddy analysis. These aggregation units have an historic place in crime analysis, but increasing computational power now makes it possible to start with very small units of analysis and to build larger units based on theoretically defined parameters. This chapter argues for a crime analysis that begins with a small spatial unit, in this case individual parcels of land, and builds larger units that reflect natural neighborhoods. Data are limited in these small units at this point in time, but the value of starting with very small units is substantial. An algorithm based on analysis of land unit to unit similarity using fuzzy topology is presented. British Columbia (BC) data are utilized to demonstrate how crime patterns follow the fuzzy edges of certain neighborhoods, diffuse into permeable neighborhoods, and concentrate at selected high activity nodes and along some major streets. Crime patterns that concentrate on major streets, at major shopping centers and along the edges of neighborhoods would be obscured, at best, and perhaps missed altogether if analysis began with larger spatial units such as census tracts or politically defined neighborhood areas.

[1] The authors would like to thank "E"-Division, Royal Canadian Mounted Police, for provision of data and for generous support. The authors would like to thank Jordan Ginther for preparation of the maps shown in Fig. 4.2.

P.L. Brantingham
Institute for Canadian Urban Research Studies, Simon Fraser University, Vancouver, BC, Canada
e-mail: pbranting@sfu.ca

D. Weisburd et al. (eds.), *Putting Crime in its Place*,
DOI 10.1007/978-0-387-09688-9_4,

Introduction

Crime is a complex event occurring in a real spatio-temporal environment. Understanding crime patterns requires both theory and research. In many situations, the requirements of understanding crime patterns lead to the development of new theories or new research methods or techniques. The choice of unit of analysis constitutes a fundamental issue for criminologists interested in spatial patterns in crime. Crime can form very different patterns at different scales of analysis (Brantingham et al. 1976; Lim et al. 2007). Standard spatial aggregations such as census tracts or politically defined neighborhood or city borders often fail to reflect the underlying socio-spatial distributions of people, land uses, or criminal events (Schmid 1960a, b).

Our study builds on criminology's history of interest in spatial patterns in crime (e.g., Quetelet 1842; Shaw and McKay 1942; Brantingham and Brantingham 1984) and focuses on the importance of starting with small units of information and aggregating them in a fashion that permits both the construction of theoretically relevant spatial units of analysis and the maintenance of a capacity for micro analysis of crime. We pay particular attention to crime pattern theory (Brantingham and Brantingham 1993a) to set the stage for presentation of a topological aggregation technique that facilitates understanding where offenders choose targets as they move through their urban surrounds.

Following a brief background review of crime pattern theory our study explores the spatial patterning of residential burglary in a suburban municipality in metropolitan Vancouver, British Columbia. The study analyses address level burglaries over a four year time period in relation to high activity nodes, major travel arteries, and well-defined edges distinguishing topologically constructed, coherent, small neighborhood areas.

Crime Pattern Theory and Scale of Analysis

Crime pattern theory maintains that criminal events occur in persistent, identifiable patterns in time and space. These patterns are temporally structured by routine human social and economic activities and are spatially structured by physical and social nodes, paths, and edges that constrain physical activity. They are shaped more deeply by the cultural, social, economic, and physical backcloth that underlies any place of human habitation (Brantingham and Brantingham 1993a, b, 2008).

Most people spend their days in very routine ways: time is spent at home; in travel to work or school; at work or school; in travel to visit friends or entertainment sites; and in travel back home. This routine may cover a small area or a large area depending on the social context, network of friends, work and home locations, the design of the city, the means of transit, and the reasons for moving around. People learn routes between destinations and tend to follow those routes repeatedly. These routes and their end points or *nodes* form an *activity space* and the basis for an *awareness space*. For most people the activity space stabilizes for long periods of time but changes when there is a life course or lifestyle change.

People who commit crimes mostly engage in non-criminal behavior. Offenders usually base their criminal activities on the time constraints or normal time expenditures of the routines they have primarily developed for their non-criminal activity. (See Wikström and Butterworth 2006; Ratcliffe 2006 for recent research that addresses the role of situational time and time budgeting in offending patterns.) In understanding criminal behavior in an urban environment, it is important to understand general variations both in the legitimate and criminal activities of offenders and in the legitimate activities of people more generally.

In the aggregate, certain locations such as drinking establishments, entertainment areas, large and small shopping areas, major transit stops, and schools are found to be at the center of clusters of crimes. These "hot spots" can attract intending offenders, that is serve as *crime attractors*, or can serve as *crime generators* simply by attracting large volumes of people including some who commit opportunistic crimes (Brantingham and Brantingham 1995; McCord et al. 2007). At the same time, the networks of paths formed by the roadways and transit systems connecting activity points channel and cluster criminal events (Beavon et al. 1994). This study explores general paths and nodes but places special emphasis on neighborhood *edges* to see how they influence patterns in criminal activity.

Crime is a rare event. This poses problems for analysts in several ways: First, because crime is rare analysts are tempted to aggregate information into larger spatial agglomerations in order to achieve sufficient counts for statistical analysis (e.g., Shaw and McKay 1942). Second, in the agglomeration process there is a temptation to use existing spatial units such as census tracts or city planning department "neighborhoods" for ease of statistical comparison between crime and information collected specifically for that spatial unit. Results from such a procedure are often misleading either because they assume a smooth distribution of crime across the entire unit or because they assume that the pre-defined unit correlates spatially with natural social neighborhoods as understood by residents of the area. However, it has long been known that both assumptions are usually wrong (Wilcox 1973; Schmid 1960a, b; Brantingham and Brantingham 1984). Third, analysts are tempted to agglomerate discrete crime categories such as assault and robbery or burglary and theft into larger categories such as "violent crime" or "property crime" even though the spatial and temporal patterns of the specific crime types may be very different. Fourth, spatial and temporal crime patterns can be very different at different spatial and temporal scales: agglomeration obscures these differences (Brantingham et al. 1976; Lim et al. 2007). Moreover, crime patterns vary at the individual location and land use levels. For example, some drinking establishments experience few crimes; some experience a lot of crimes. Similarly, some transit stops experience high crime levels; some do not.

In all aggregate spatial information, there are ongoing issues that relate to modifiable area unit boundaries; to the limitations of statistics in dealing with data with high levels of spatio-temporal autocorrelation; and to the lack of independence that bedevils statistical analysis of data generated by high levels of repeat offending. Contemporary research into spatio-temporal crime patterns has an expanding interest in understanding extremes in the spatio-temporal patterning of crime and in finding ways to move through a continuum of analysis from the micro level individual offender's activities and characteristics through many levels of aggregate patterns without being restricted to choosing either a single micro or meso or macro level.

Ideally, analysis would be nested, that is information about individual criminal events – specific location or address, specific time of occurrence, detailed description of crime type – would be the base unit of analysis. This base unit would be aggregated to different, larger units depending on the research question, but the basic information would be maintained for aggregation along a different schema for a different research question. This implies that no single level of aggregation can constitute the "best" unit of analysis for studying the spatial or temporal patterns in crime. Data should be collected at the most detailed level possible and aggregated upward to fit the requisites of theory or the limitations of data unit aggregations of those elements of urban backcloth thought to be important. That is, in looking at different levels of aggregation, researchers must consider aggregation of crime units into different areal units for comparison against the urban backcloth.

It is a simple fact that, until recently, the tedium and expense of manual data collection and analysis as well as the limitations of computer storage and analysis capabilities, meant that aggregation was often necessary to facilitate any type of analysis. Advancements in computational power and the availability of extensive data at detailed spatial and temporal levels now make it possible to start small, at discrete locations in space-time and have theory help direct aggregation into larger units for analysis.

The recent emergence of *computational criminology*, grounded on improvement in the computational power available to researchers, provides, potentially, a way to link theory and research at a micro level with theory and research at the meso levels of analysis. The research made possible by computational criminology is nascent, but rapidly evolving. Several chapters in this book use extensive computing that would not have been possible ten years ago, let alone in the time of Shaw and McKay.

Computational modeling is particularly important in its requirement that the structure of the model and the rules of computation be made explicit. Computational criminology is also an invitation to more criminologists to use artificial intelligence, agent-based modeling, and graph theory in modeling crime and testing related crime theories (See, e.g., Groff 2007; Xue and Brown 2006; Brantingham et al. 2005; Liu et al. 2005; Townsley et al. 2003; Brantingham et al. 2005; Brantingham and Brantingham 2004; Adderley 2004; Brown 1998). This study uses a computationally intensive mathematical technique called *fuzzy topology* to build nested models of paths, nodes, and edges in order to study the discrete distribution of specific crime types across the urban backcloth of a British Columbia municipality.

Methodology

This study attempts to provide an example of a new approach to looking at crime by placing discrete crime locations on a model of the urban backcloth based on common paths and activity nodes, and on identification of the boundaries and cores of neighborhood areas that stand out as different from their surrounds. This chapter should be seen as a compliment to the other chapters in this section: *Oberwittler and Wikström's* study of behavioral contexts and *Groff, Weisburd, and Morris'* exploration of juvenile crime against block level crime trajectories over time.

We explore the patterns of residential burglary for 2004 for a suburban municipality in Metro Vancouver. During 2004, more than 12,000 criminal code offences and more than 350 drug offences were reported to police in this municipality. There were 552 residential burglaries reported to police in 2004 and 2,296 over the four year period from 2001 through 2004.

The rapidly growing suburb used in this analysis is relatively near the core city of Vancouver and is primarily residential. Its population of about 124,000 grew by 24% between 1993 and 2004, and by slightly more than 7% over the period 2001–2004. In 2005, the suburb had almost 37,000 separately tracked parcels of land.

This study suburb is similar to many other Vancouver Metro suburbs with an older section dating back to the first half of the 20th century when this was more on the edge of suburb development. A major growth spurt in the last quarter of the 20th century has seen some increase in higher density housing. Overall, however, this municipality remains primarily a residential suburb with single family dwelling units as the primary type of residence. There is no core business/commercial area.

The analysis presented in this chapter builds on micro level address data for residential burglaries and then analyzes the patterning of these crimes using the ideas of common nodal activity points, routine common pathways, and roads within the suburb and the edges of neighborhood spaces. The analysis is primarily done using an agglomeration algorithm that will be described below.

Data Sources

The analysis was done using four categories of data: officially reported crimes; British Columbia Assessment Authority (BCAA) data that provides individual lot level land use information; detailed street information from GIS Innovations; and Canadian Census data for the 2001 census.

The reported crime data was made available by "E" Division of the Royal Canadian Mounted Police (RCMP). "E" Division provides local municipal policing services to over 180 jurisdictions in British Columbia, including the suburban municipality used in this study. The British Columbia Assessment Authority is a provincial government agency that is responsible for the provision of tax assessment information for every separate parcel of property in the province. More than 200 detailed land use types are tracked. This information is used by municipalities for property taxation purposes. The assessment data set contains detailed land use information that makes it possible to distinguish between and identify many detailed types of residential, commercial, civic, and industrial land uses by address.

The crime and land use data were geocoded using a street network file developed by GIS Innovation and used by many BC government ministries. The information contained in this street network file is very detailed. It even identifies traffic calming speed bumps. For the purposes of this analysis, the street network was used to identify local and arterial roads. More detailed information about foot paths, unpaved logging roads, and other types of pathways is available and will be analyzed in future studies.

The final type of data used in this analysis is Canadian census information. Canada undertakes a census every five years; at the time of this study the latest available census data was for the year 2001. This created a time lag for census data. However, the census provides detailed socio-demographic household and housing information on a five-year cycle; so, data are fresh in comparison with census data used in many criminological studies.

As is common in much criminological research, there was some necessary mixing of time ranges. Crime data was available for 2001–2004; BCAA data was available for 2005; GIS Innovation street files were available for 2006. Census data was from 2001. This creates uncertainty about some of the data and should be kept in mind in interpreting the results. We are engaged in creating archives for these and other data sources; so, over time it should become possible to eliminate such temporal mixing, at least for specific census years.

Micro-Meso-Micro Analysis

The unit of analysis is fundamental in any study and the central issue for this book. As mentioned before, we are entering a period where we can undertake analysis at the most fundamental address level, aggregate to some larger summary unit to look for patterns, and return to a more micro level of analysis to better understand any interesting patterns that are found.

Following geocoding, the burglary data was explored visually at the address level in comparison to the BCAA taxable parcel database.[2] We used census data at the smallest available level. This small unit is called a Dissemination Area (DA) and consists of a cluster of blocks. Eight to ten Dissemination Areas, when combined, form a Census Tract. BCAA data was aggregated to the DA level. The census data was used mathematically to depict an empirical version of the urban backcloth for analysis of crime in relation to major activity nodes, major streets, and the sharpness of the edges of coherent neighborhoods. The technique will be described in

[2] The effective geocoding rate for the burglary data was 94.9%; some 98.3% of the BCAA land use data successfully geocoded.

summary below. Detailed description of the technique is presented in the mathematical appendix.

Commercial areas reported in the BCAA data set are used to identify likely activity nodes for the study city. Schools, parks, and recreation areas were not used in this initial beta testing of the new algorithm. In this study city, road arteries were consistent with commercial areas. Census information about age of housing and type of housing was used in this exploratory analysis. Older housing built before 1946 and middle aged housing built between 1946 and1960 and small apartment buildings were identified as reasonably good visual identifiers of residences that would be more likely to experience burglary (See, e.g., Brantingham and Brantingham 1977; Waller and Okihiro 1978; Bennett and Wright 1984; Clarke and Hope 1984; Cromwell et al. 1991; Rengert and Wasilchick 2000).

Urban Backcloth

What surrounds us in an urban environment includes centers of activity, roads and pathways, well known landmarks, and parks as well as neighborhoods with different socio-economic and demographic character. We move around in the urban environment from one activity node to another sometimes with fixed location goals (such as a specific restaurant) and sometimes with general area goals (the entertainment district). This movement takes people through well defined areas with crisp, clear borders, and through less clearly identifiable areas. Crimes occur within this backcloth, and can even shape the backcloth. Individuals have personal nodes, paths, and edges that shape their activities within the backcloth. In the aggregate some nodes, paths, and edges stand out.

Common aggregate activity nodes that are studied in criminology are shopping areas, entertainment districts (including pubs or bars), and schools. Aggregate awareness spaces are likely to be located the areas around these types of nodes. Of course, activity nodes vary by individual, but areas of activity concentration reflect the activity nodes for many people. In a similar way, major roads and mass transit shape and reflect paths used by many people. Individual movement patterns vary, but groups of individuals shape aggregate patterns of movement (WAAG Society 2007).

What is of particular importance is what shapes the edges of the awareness spaces around major paths and nodes. A new algorithm described below is being developed to model the shape of the aggregate awareness spaces around major activity centers, major roads, and homogenous neighborhoods. The algorithm helps identify sharp or gradual breaks or barriers between aggregate awareness spaces. The algorithm is designed to reflect the dynamics of an urban environment and move towards softer definitions of distance.

We are currently in a period of innovation in spatial analysis. Part of the originality of research is in the development of new measures to articulate ideas. (Bittner 2001; Bafna 2003; Elffers 2003; and Weisburd et al. 2004 for examples of the creation of interesting measures to address differing theoretical approaches in spatial analysis.)

Common GIS software makes it relatively easy for researchers to create fixed distance buffers around points, streets, or shapes (polygons) and look at crime concentrations within these buffered areas. This chapter tries to go beyond the fixed distance buffers, utilizing a new algorithm, TOPO©, for aggregation of fine grained spatial units into larger, coherent aggregates that reflect probable awareness spaces – cognitive buffers of flexible size and power. The algorithm has value in understanding the impact of perceptual edges on crime patterns. In highly varied areas in particular, sharp perceptual edges may block awareness and activity at the boundary, keeping people and events from crossing into spatially adjacent but perceptually different areas. Flexible buffer size may be of particular importance where space, place, and context are important and where there are high levels of dissimilarity from spatial unit to spatial unit.

TOPO provides a tool for research construction of cognitive buffer areas around activity locations. Cognitive buffers need not be of equal length in all directions but can reflect the actual structure and use of areas around an activity node or common path. Rengert and Wasilchick (2000) show a cardinality in the direction of offending patterns consistent with a non-circular awareness space (Pyle 1974; Costanzo et al. 1986). Ratcliffe (2006) defines time constrained buffers that take a teardrop shape.

TOPO is a fuzzy topology algorithm. The details of the algorithm are set out in appendix A. As a brief summary, topology is an area of mathematics that has a focus on continuity and discontinuity. Continuity within TOPO means that one unit of analysis is similar to an adjacent unit to some degree. This area of mathematics has links to interpretation of images in medicine, satellite images, and in linguistics, among others.

Fuzzy topology measures the amount of difference from block to block. When changes between adjacent blocks are gradual, someone travelling past them may have trouble noticing the change until it has cumulated into a very large difference. It is possible to move from one neighborhood to the next; recognize when you are in the core area of a neighborhood; but not be able to say exactly when you move from one neighborhood to another. Gradual change produces a *fuzziness* in neighborhood membership in that a block may partially belong to two or more neighborhoods.

Fuzzy topology also tracks crisp or sharp changes between neighborhoods.[3] When there is a crisp change from one block to the next, people tend to notice the difference. The border between an industrial district and an adjacent residential area is sharp and noticeable; everybody experiences the sharp change when they turn off a commercial block with stores and move into a quieter residential area.

In TOPO, an adjacency table is used to compare a series of adjacent units and measure the rate of change in the variables of interest as observation moves from one adjacent unit to another. This measure of rate of change from one unit to the next adjacent unit makes it possible to identify core areas (groups of contiguous spatial units) with higher homogeneity and to identify spatial transitions between core areas by gradual change or by sharp, abrupt difference, that is, where there

[3] These are the sorts of neighborhood edges described by Kevin Lynch (1960) in his classic book *The Image of the City*.

is much less similarity from block to block on the variable of interest. The places on the fuzzy edges or border areas may partially belong to two or more disparate but homogenous neighborhoods. The places on crisp edges represent a sharp break between adjacent neighborhoods.

To help measure the crispness or softness of the borders between adjacent units, the algorithm is written so that it allows for consideration of neighborhood sets across a range of levels of difference in the variable of interest. In this study, a lattice was used that agglomerated sets of similar blocks and identified their fuzzy and crisp boundary edges by allowing 30% variation in value of the variable of interest, 20% variation, and 10% variation to be set as the measure of similarity. Using this lattice (30%, 20%, and 10%), the sharpest edges would be dissimilar at 30% variation (which means that the adjacent units would be dissimilar at 20% and 10% as well). A softer edge would be similar in adjacent units at 30% difference and 20% difference and only show dissimilarity at 10% variation.

The softness or crispness of neighborhood edges can vary greatly depending on the number of variables considered and the number of levels of dissimilarity considered. In this study four primary variables with three dissimilarity levels are used on the lattice. As a result, the number of edge components assigned to a particular unit could range from 0 to 16. The lowest value would be *zero* where adjacent units were similar for all variables of interest and at the 30%, 20%, and 10% difference levels of variation. The highest measure in the sharp edges or border areas would be 16 where there was a difference for adjacent units for 10, 20, 30% variation levels. This provides an index of the extent to which any given unit belongs to a single unique neighborhood.

This approach is different from traditional clustering algorithms where a set of units is divided into well-defined subsets. A unit either belongs to a subset or it does not. TOPO assigns a *level* of membership to basic sets. It is developed to handle, in spatial analysis, something equivalent to how it is possible to say someone is "tall" or "short" in some situations, but that there are many people who are neither "tall" nor "short" but are "taller" or "shorter". As you might infer from the example, fuzzy topology and fuzzy logic are used heavily in such fields as computational linguistics.

The TOPO algorithm is used in this study to identify the clear centers or interiors of neighborhoods within a suburb of Vancouver and to identify both the gradual and sharp changes between neighborhoods, that is, the fuzzy edges. The edges define many of the structures of the urban backcloth.

Crime is relatively rare. It is expected that sharp edges form cognitive barriers and locations where crime is likely to cluster (remembering that relationships are not isomorphic). A sharp border does not necessarily produce crime, just a locale that is frequently compatible with crime because the activity of both neighborhood insiders and outsiders is channeled and held there. Insiders are reluctant to cross the sharp border into a very different neighborhood; outsiders are reluctant to leave the sharp border zone. A soft border – a highly fuzzy area – and less crime would be expected in the broader fuzzy areas. In a topological sense this could be visualized by considering a buffer along a road with equal risk within the buffer. A crisp edge compacts the buffer, moving the crimes into a smaller area. A large, soft, fuzzy area expands the buffer and spreads the crime over a larger area.

Results

Overview of Residential Burglary

Figure 4.1 provides a kernel density map for residential burglary in the suburb of Vancouver covered in this chapter. Residential burglary is varied across the city with a concentration along its south western jurisdictional limit. As presented in this chapter, data from adjacent municipalities were not used, but some preliminary testing indicated that the adjacent municipal areas represented crisp, sharp breaks with the area covered in the analysis particularly along the western edge. One municipal boundary is on a sharp transition to industrial storage; the other municipal boundary runs along a major road with a large regional shopping complex with mass transit and bus exchange stations across the street in the other municipality; shops associated with the mall complex are found on the study community side of this road.

Figure 4.2 provides additional general display information for the study municipality. The upper panel shows the general distribution of residential developments. The lower panel shows the distribution of multi-family residences and higher density dwelling types. As is typical of many North American suburban cities, there is no high density core area. The study city is dominated by single family residential land uses. It should be noted that as part of a planning scheme to develop a city

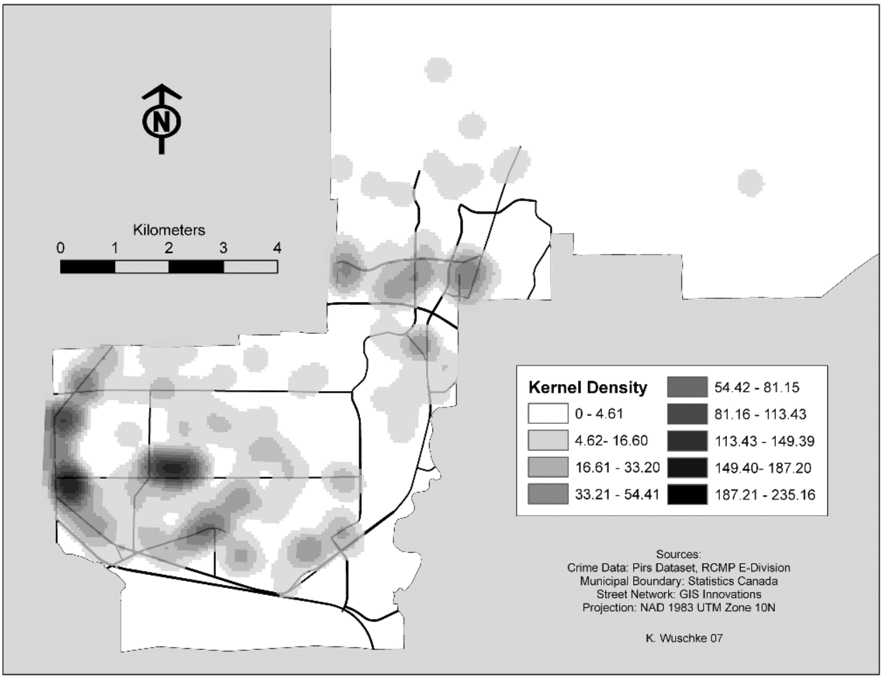

Fig. 4.1 Repeat residential burglary: B.C. Municipality, 2001–2004

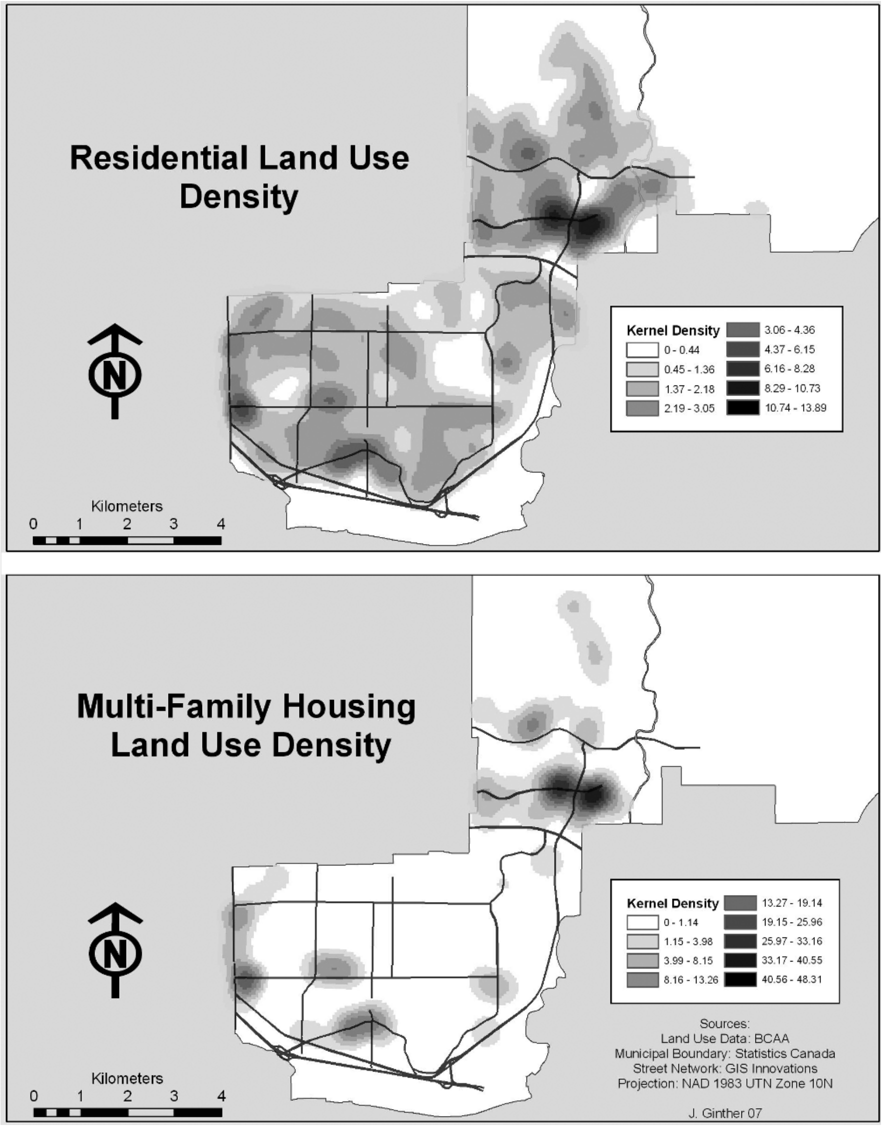

Fig. 4.2 Study area land use, 2005

center, the municipal government has encouraged construction of new multi-story condominium developments in an area near a shopping mall where it has also built a large park and recreation center, a series of new city service buildings including the city hall, the police station, a library, and allocated land to a new community college. There is a clustering of some residential burglaries in this area, but, as will be shown, the volume is small in relation to the number of residential units.

Fuzzy and Sharp Borders

The paths, nodes, and edges in this municipality are explored using the TOPO© algorithm. For Dissemination Areas (DAs), the smallest census area unit available, burglaries are analyzed for clustering in areas along major arteries, near shopping areas (common activity nodes) and for their border/edge areas.

The results of this exploratory testing are very interesting. Most Dissemination Areas have some measure of dissimilarity with contiguous areas, that is, there are many fuzzy boundary areas in the city. The fuzzy borders or edges for commercial areas, older housing, and small apartments were calculated allowing a 30% change from adjacent unit to adjacent unit. These borders or edges show a difference in the average number of burglaries (see Table 4.1). There is enough diversity in this suburb that there are few areas that are interiors for all variables and many more that are borders or edges for all three variables. Even with this diversity, the range of values runs from 1.67 for the interiors or areas surrounded by similar areas to 13.81 where the fuzzy borders cumulate to produce strong edges.

The variation, however, is large within each fuzzy boundary category and with some outliers or extreme values. Figure 4.3 presents the box-plot for the fuzzy boundary/edge values. The results are very interesting given that this exploratory analysis was planned primarily to test the algorithm. It is expected that results should be even more interesting using a broader range of variables with smaller units of analysis for the aggregation into homogeneous and less clear, fuzzy areas.

For descriptive purposes, Table 4.2 shows the impact of both individual and combined boundaries for the three variables. Zero represents an interior; the value 1 is used for a single boundary. As can be seen from the figure, the highest average number of burglaries (about 14) is for Dissemination Areas that are fuzzy borders/edges of commercial, older housing, and small apartment areas. Areas that are interiors for all variables or a boundary for only one of the variables have much lower numbers of burglaries than the high boundary areas.

The small apartment variable and the older housing variable were used with the boundary counts as a factor in a General Linear Model (GLM). As is usually the case with crime data, the assumptions of GLM could not be met. The variances are unequal (Levene's Test); cell sizes vary; and clear outliers exist. The largest variance is, as expected, with the boundary category with the largest n. This and the small

Table 4.1 Mean and standard deviation of number of burglaries (2001–2004) by number of fuzzy boundaries

Number of fuzzy boundaries	Mean	Standard deviation	N
0	1.67	2.887	3
1	4.3	3.323	23
2	9.29	6.979	62
3	13.81	9.594	93
Total	10.85	8.788	181

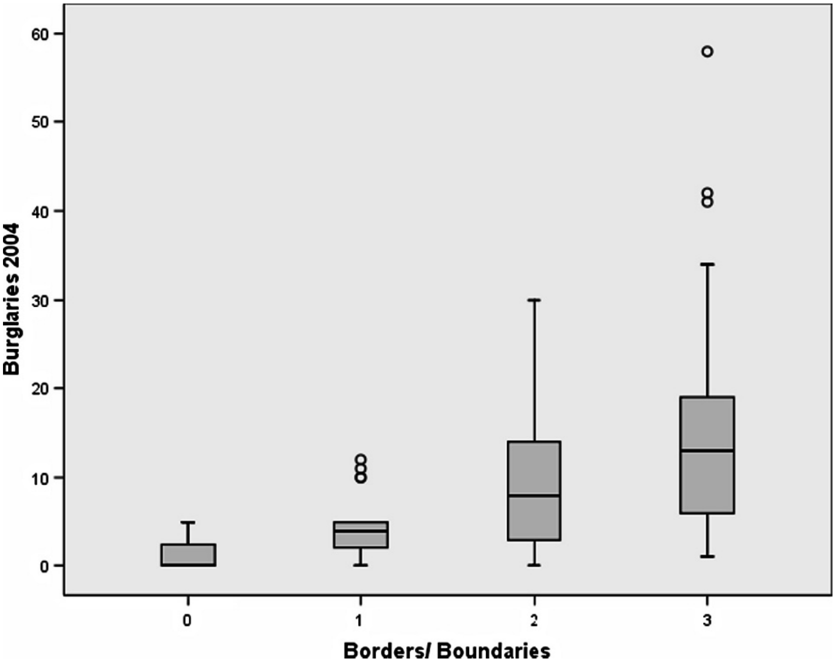

Fig. 4.3 Box-plot for the fuzzy boundary/edge values

Table 4.2 Mean and standard deviation of number of burglaries (2001–2004) by boundaries/borders

Construction 1946–1960 boundary	Commercial boundary	Small apartment boundary	Mean	Std. deviation	N
0	0	0	1.67	2.89	3
		1	4.07	2.95	14
	1	0	4.50	0.71	2
		1	7.86	6.97	29
1	0	0	4.71	4.61	7
		1	10.68	7.04	31
	1	0	8.50	2.12	2
		1	13.81	9.59	93

n for interior areas are likely to under-identify relationships. With the exclusion of the three outliers (worth special study separately in a crime attractor analysis), the boundary impact is reasonable. The boundary/interior difference was significant for the number of borders and for the covariates. The η_p^2 values, however, are small for all the variables.

An additional GLM was run with natural log transformations, exclusion of outliers, and a collapsing of the boundary edge variable into two categories (zero and one boundary; and two or three variables). The GLM model, consistent with the variance equality assumption, continues to show significance for the transformed variables. The η_p^2 values remain small.

The analysis was repeated for different lattice values (20% and 10%, as well as 30%). The results were similar.

Crime is rare, but clustered. Research needs a focus on extreme values. Table 4.3 shows the difference in the extreme values for the interior (0/1) and boundary (1/2) divisions used in the GLM just described. As can be seen by the information in the table, there is a large difference in the high and low values. The five high values for the interiors range from 5 to 12. The five high values for the borders range from 33 to 58. The low values for both categories are zeros. There are low crime borders or edges just like there are low crime interiors, but the high crime edges are of a magnitude greater than high crime interiors. Similarly, the weighted averages for the 95th percentile for interiors start at 11.65 for interiors and 26 for borders.

In general, this initial testing provides support for expanding work using fuzzy set theory and topology in computational criminology. The importance of moving to block level analysis is strongly supported. Dissemination Areas are aggregations of blocks. By using street blocks, it will be possible to create a model that reflects micro level perceptual changes as well as one that moves closer to block level concentration of crime and to begin to add individual cognitive awareness spaces.

Table 4.3 Extreme values for borders and interiors

				Case Number	Value
Burglaries	Interior	Highest	1	72	12
			2	75	11
			3	83	10
			4	87	10
			5	124	5[a]
		Lowest	1	181	0
			2	178	0
			3	177	0
			4	174	1
			5	171	1
	Border	Highest	1	1	58
			2	2	42
			3	3	41
			4	4	34
			5	5	33
		Lowest	1	180	0
			2	179	0
			3	176	0
			4	175	0
			5	173	1[b]

[a] Only a partial list of cases with the value 5 are shown in the table of upper extremes.
[b] Only a partial list of cases with the value 1 are shown in the table of lower extremes.

Table 4.4 Number of land uses and burglaries at selected high and low border areas

Dissemination area	Borders	Land uses	Crimes	Residential units	Crime rate
High Border 1	14	5	58	270	21.48
High Border 2	13	8	42	350	12.00
High Border 3	14	9	41	370	11.08
Low Border 1	4	2	3	125	2.40
Low Border 2	4	3	3	195	1.54
Low Border 3	5	1	4	115	3.48

Return to the Micro Level of Analysis

The difference between the crisp border areas and the more homogeneous interiors is rather strong. This exploratory analysis used a limited number of variables. To bring the study full circle, we identified Dissemination Areas with the highest and lowest counts and explored how the residential burglaries varied from building to building, that is, in essence we took the similarity idea in TOPO and applied it to lot or individual parcel level data in some selected locations.

As Table 4.4 shows, the three highest residential burglary Dissemination Areas had border/edge counts of 14, 13, and 14. The three lowest Dissemination Areas had border/edge counts of 4, 4, and 5.

Conclusions

This research strongly supports the move in criminology towards using micro units of analysis, aggregating them when necessary, but maintaining the detailed units of analysis for additional research purposes. One level of natural aggregation is to small locales such as neighborhoods and to major roads and activity centers. We explored using fuzzy topology to develop small locales or neighborhoods and identify the edges or borders of these neighborhoods. We found that even using an aggregate unit of a Dissemination Area (the smallest census area with available data) there were fairly strong differences in the amount of burglaries in homogeneous interiors and in their fuzzy edges or borders. Burglaries were higher for crisp borders than for more gradually changing edges.

It was particularly interesting to explore the residential burglary patterns within a selection of high border/ high crime Dissemination Areas (DAs) and the low border/ low crime DAs. While the TOPO algorithm only used three variables, the actual land uses in the high border/high crime DAs were highly varied. There were many micro edges within these areas. The land uses in the lower edge/lower crime DAs were less varied, more homogeneous. Our conclusion is that the basic spatial unit of analysis in crime pattern research should be the individual address or parcel of land. Larger aggregates should be constructed from this spatial level in a way that makes it possible to look back within the aggregates. Future research in crime patterns should begin by using a fine, small unit of analysis, and aggregating up to street blocks, neighborhoods, major activity spaces, or other larger units when the theoretical orientation calls for it.

Mathematical Appendix: Fuzzy Topology Algorithm

This appendix contains a brief summary of the fuzzy topology algorithm developed for this study. Topology is an area of mathematics that has a focus on sets, continuity, and discontinuity. In fuzzy topology, sets can have partial membership in a set instead of the traditional set theory where an element is either a member of a set or not a member. The degree of membership fits well into the use of words like “near” or “similar”. The fuzzy set membership fits well into urban concepts like “neighbourhood” where there may be agreement about the core of a neighbourhood, but lack of agreement about its borders or edges. Similarly, a person’s awareness space can have well-defined centres but fuzzy edges.

Background

A *topology* on a space T is the collection of subsets, X_i, such that:

- The $\emptyset \in T$ and $T \in T$
- If $X_1, \ldots, X_n \in T$ then $\cup X_i \in T$
- If $X_1, \ldots, X_n \in T$ then $\cap X_i \in T$

An important concept which must be considered is that of a *basis* or *base*. A basis of a topology T is a sub-collection B of T with the property that every open set X of T is the union of the basis sets. Formally, if $X_i \in T$ then $X_i = \cup B_j$ where B_j is a basis set. Another important concept is the difference between finite point-set topology and infinite topology. Topologies in criminology are built on a finite number of sets when the unit of analysis is something like crimes or addresses or where the unit of analysis is a city block.

Probably one of the most important aspects of topology useful to criminologists is that the individual sets can be of different sizes. Unions of adjacent sets, that are sets themselves, create a natural way to aggregate information. With property lots, for example, these can be aggregated to block faces, to blocks, to nearest intersection, to groups of blocks joined by specified criteria, to neighbourhoods, to urban areas, and to larger units. While the modifiable area unit problem in spatial analysis of government data sources cannot be eliminated completely, aggregating and disaggregating elements of topology provide the potential for forming natural neighbourhoods or creating a cognitive unit like an awareness space.

Edges (*boundary*), *interiors*, and *neighbourhoods* are concepts in topology that are very helpful in crime analysis and spatial analysis. These concepts will be defined mathematically. These are concepts that distinguish between the relative position of points in sets and sets that have been joined into a union of sets

($S_i = \cup X_j$ where $X_j \in T$). Fundamentally, in topology, you can have a point x that is in a set. That set may be surrounded by other sets. When a set is surrounded by other similar sets in the topology it is called an open set. The open sets around point x form the neighbourhood of x. Generally for sets a neighbourhood has the following properties: if T is a topological space and S is a subset of T, then X is a neighbourhood of S if X is open and X is contained in another subset Y that is contained in T. Symbolically, $S \subseteq X \subseteq Y \subseteq T$. The *interior* of a set ($int(S)$) is the union of open sub-sets within the set and points within these open sub-sets. The *boundary* or *edge* of a set ($bd(S)$) is the union of the closed sub-sets, that is, sets that are not open sets.

Topological Aggregation Algorithm

The focus of this algorithm is to develop aggregate awareness spaces shaped by a distinction between neighbourhoods or districts with sharp or fuzzy boundaries defined by gradual change. This topology algorithm is different from the common statistical methods of clustering where there are classifications of areas using some k-mean values, high or low values, density/connectivity measures or grid-based methods as the basis of alternative mathematical rules for similarity that use a set common value. TOPO© uses rules for determining similarity that are based on the differences in adjacent block units only. What this means is that there can be a series of adjacent blocks where a highly visible attribute like the age of the buildings can vary in small amounts from one block to another so that the block at the beginning of the series is very different from the block at the end, but where there is little difference between the any two adjacent blocks. Strong differences are more easily recognized than small changes.

Fuzzy set theory was developed by Zadeh (1965). Fuzzy topology is a growing area of applied research. Readers who have an interest in fuzzy logic and fuzzy topology should review the work of Li and Li (2004), Winter (1998), Haq and Zimring (2003), Yeung et al. (2005), and Liu and Shi (2006) to see the uses of fuzzy sets and fuzzy topology. The article by Liu and Shi (2006) is of particular interest. It describes a fuzzy topology algorithm with some similarity to the algorithm presented in this chapter.

In fuzzy set theory there is a membership function that is used to assess a level of membership. That is, there is a membership function μ such that $\mu \to [0, 1]$ where the values 0 and 1 are non-membership and complete membership. The 0/1 is like traditional Cartesian *true/false* logic. The values between 0 and 1 measure the degree of membership.

A fuzzy set on the types of sets previously described uses the following notation: $\tilde{A} = \{(x, \mu_A(x))|x \in X\}$. For all elements of a set, X, the fuzzy set is the elements and their associated membership based on the fuzzy membership function, μ_A. The intersection of two fuzzy sets is denoted as $(A \cap B)(x) = \min [A(x), B(x)]$. The union of two fuzzy sets is denoted as: $(A \cup B)(x) = \max [A(x), B(x)]$.

The associated membership functional value that is used to set the limit for similarity can be changed. Every time the inter-unit variation is changed, new basis sets are constructed. Many sets are created as the percentage variation is allowed to range up or down. For example, let B_i be a basis set and b_j be a block. Let $f(b_j)$ be a functional value associated with block b_j, such as average cost of housing, average rent, or percent apartment houses. Then a basis set is:

$$B_j = \left\{ b_i \,\middle|\, \left\| f(b_j) - f(b_i) \right\| \leq \max\{af(b_j), af(b_i)\} \right\} \text{Where}$$

$$b_i \in B \text{and } b_i \cap b_j \neq \phi; \text{and } 0 < a < 1; b_i \neq b_j;\ i = 1, \ldots, n;\ j = 1, \ldots, m$$

The contours of the *neighbourhood* change and develop as the permitted inter-block variation is changed and new basis sets are formed. A set constructed from a fixed level of inter-block variation contains sets constructed from lower levels of inter-block variation. If $a_{i-1} < a_i < a_{i+1}$ are real numbers between zero and one and $B_j(a_j)$ is a set formed by allowing a fixed a_i inter-block variation then:

$$\ldots \subseteq B_j(a_{i-1}) \subseteq B_k(a_i) \subseteq B_l(a_{i+1}) \ \ldots$$

It is worth special consideration to note that as the associated membership function value decreases additional boundary blocks will be created. When a block is a boundary block for one level of variation it will be a boundary block for a set created by a smaller inter-block variation. Let $\bar{b}_j$ be a boundary block in a basis set constructed by allowing an a_i inter-block variation, then another chain is formed as the inter-set variation changes:

$$\ldots \subseteq \left\{ \bar{b}_j(a_{i+1}) \right\} \subseteq \left\{ \bar{b}_j(a_i) \right\} \subseteq \left\{ \bar{b}_j(a_{i-1}) \right\} \subseteq \ldots$$

It should be noted that boundary blocks for sets constructed from lower levels of variation will not always be boundary blocks for sets constructed from higher levels of variation. Boundaries may change as the value of a changes; this is a gradual change. The boundaries will be the same when there is a sharp, crisp, difference between neighbourhoods.[4]

TOPO© uses multiple characteristics in a *product topology,* that is, we consider different characteristics of urban areas co-jointly. If we let $\{X_\alpha\} \quad \alpha \in J$ be a finitely indexed family of topological spaces. $\prod_\alpha X_\alpha$, the Cartesian product of the X_α's, is the product space. The basis for this product space is the collection of all sets of the form $\prod_\alpha B_\alpha$ where B is an open set in X_α and J is a finite index set.

Once again, as with the simple topologies, the interesting properties emerge as the permitted variation is allowed to increase and decrease in the basis sets for the component topologies. For example, in a given residential area, the component topologies may have the same basis sets; that is, within an area B_α the basis set for characteristic α (topology α) could contain the same blocks as B_β, a basis set for topology T_β for many levels of contiguous variation. This area would have a very

[4] See Brantingham and Brantingham (1978) for a more detailed exploration of the boundary effect.

high level of internal homogeneity and clear-cut boundaries with adjacent areas. An area that has the same basis sets for all the component topologies at many levels of variation has a high level of perceptual distinctiveness.

Many complex nests and chains are created when the boundary blocks are different for each component topology (each characteristic or attribute) and when the boundary blocks change as the permitted variation increases or decreases. The range of types of transitions between the centres or interiors of neighbourhoods is part of how this type of neighbourhood model may approach cognitive images. Sharp or crisp boundaries are relatively rare except when there is a physical feature such as a lake or a highway or an abrupt land use change such as a move from a shopping area to a residential area.

For the purposes of analysis, the ideas just described are written in the following functional form: Consider f a functional value for the unit of analysis. Let the membership function μ have two values: $\mu = 0$ when $b_i \cap b_j = \phi$; and has the following value when the basis sets intersect.

$$\mu = 1 - \left(\frac{\| f(b_j) - f(b_i) \|}{\max(f(b_j), f(b_i))} \right)$$

The fuzzy sets have values for every adjacency in the matrix described earlier. For example, when $\text{block}_1$is adjacent to block_2, block_3, and block_4, then fuzzy values would be numbers such as {.8, .7, .3} when the functional value between $\text{block}_1$and block_2 is an 80% similarity, the similarity for $\text{block}_1$and block_3 is 70%; and the similarity between $\text{block}_1$and block_4 is 30%. In a fuzzy sense, blocks 1, 2, and 3 are highly similar; block_4 is dissimilar.

The topology just described creates an urban backcloth that identifies sharp changes from one spatial area to the next. It also identifies gradual changes. The model identifies neighbourhood interiors for single and multiple variables. It creates a fabric of changes, a fabric that may start to identify the location of potential cues and cue clusters that identify locations as unique. The nodes (other than home locations) are likely to be in boundary areas; paths may fall along boundaries, particularly when in strip commercial development or when the main roads run along the edge of water or open spaces. The paths, when large enough, create a boundary themselves. It is expected that sharp boundaries influence the permeability of urban areas. The sharper the boundary, the more likely it is that people passing through an area stay in the boundary area. It is also expected that the boundary areas are areas where everyone feels like an outsider. There is lack of similarity. Each turn or block can seem different. With most people feeling like outsiders there is little likelihood that there are many natural guardians.

References

Adderley, R. (2004). The use of data mining techniques in operational crime fighting. *Intelligence and Security Informatics*. Second Symposium on Intelligence and Security Informatics, Tucson, AZ, USA, June 10–11, 2004. Proceedings.

Bafna, S. (2003). SPACE SYNTAX A brief introduction to its logic and analytical techniques. *Environment and Behavior*, 35(1), 17–29.

Beavon, D. J. K., Brantingham, P. L., & Brantingham, P. J. (1994). The influence of street networks on the patterning of property offenses. *Crime Prevention Studies*, 2, 115–148.

Bennett, T., & Wright, R. (1984). *Burglars on burglary: Prevention and the offender*. Brookfield, Vermont: Gower Publishing Company.

Bittner, T. (2001). The qualitative structure of built environments. *Fundamenta Informaticae* 46, 97–128.

Brantingham, P. J., & Brantingham, P. L. (1977). Housing patterns and burglary in a medium-sized American city. In: J. Scott & S. Dinitz (Eds.), *Criminal justice planning* (pp. 63–74). New York: Praeger.

Brantingham, P. J., & Brantingham, P. L. (1984). *Patterns in crime*. New York: Macmillan.

Brantingham, P. J., Dyreson, D. A., & Brantingham, P. L. (1976). Crime seen through a cone of resolution. *American Behavioral Scientist*, 20, 261–273.

Brantingham, P. L., & Brantingham, P. J. (1978). A topological technique for regionalization. *Environment and Behavior*, 10, 335–353.

Brantingham, P. L., & Brantingham, P. J. (1993a). Environment, routine and situation: Toward a pattern theory of crime. *Advances in Criminological Theory*, 5, 259–294.

Brantingham, P. L., & Brantingham, P. J. (1993b). Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology*, 13, 3–28.

Brantingham, P. L., & Brantingham, P. J. (1995). Criminality of place: Crime generators and crime attractors. *European Journal on Criminal Policy and Research*, 3, 5–26.

Brantingham, P. L., & Brantingham, P. J. (2004). Computer simulation as a tool for environmental criminologists. *Security Journal*, 17(1), 21–30.

Brantingham, P. L., & Brantingham, P. J. (2008). The rules of crime pattern theory. In: R. Wortley, L. Mazerolle, & S. Rombouts (Eds.), *Environmental criminology and crime analysis*. Devon, UK: Willan Publishing.

Brantingham, P. L., Brantingham, P. J., & Glässer, U. (2005). Computer simulation as a research tool in criminology and criminal justice. *Criminal Justice Matters*, 58, 19–20.

Brantingham, P. L., Glässer, U., Kinney, B., Singh, K., & Vajihollahi, M. (2005). A computational model for simulating spatial aspects of crime in urban environments. *Proceedings of the IEEE international conference on systems, man and cybernetics* (pp. 3667–3674). Hawaii, October 2005.

Brown, D. (1998). The Regional Crime Analysis Program (ReCAP): a framework for mining data to catch criminals. *Proceedings of the 1998 IEEE international conference on systems, man and cybernetics*, 3, 2848–2853.

Clarke, R. V. G., & Hope, T. (1984). *Coping with burglary: Research perspectives on policy*. Boston: Kluwer-Nijhoff.

Costanzo, C. M., Halperin, W. C., & Gale, N. (1986). Criminal mobility and the directional component in journeys to crime. In R. M. Figlio, S. Hakim, & G. F. Rengert (Eds.) *Metropolitan Crime Patterns*. (pp. 73–95). Monsey, New York: Criminal Justice Press.

Cromwell, P. F., Olson, J. N., & Avary, D. W. (1991). *Breaking and entering: An ethnographic analysis of burglary*. Newbury Park, CA: Sage Publications.

Elffers, H. (2003). Analysing neighbourhood influence in criminology. *Statistica Neerlandica*, 57(3), 347–367.

Groff, E. R. (2007). Simulation for theory testing and experiments: An example using routine activity theory and street robbery. *Journal of Quantitative Criminology*, 23(2), 75–103.

Haq, S., & Zimring, G. (2003). Just down the road a piece: The development of topological knowledge of building layouts. *Environment and Behavior*, 35(1), 132–160.

Li, Y., & Li, S. (2004). A fuzzy sets theoretic approach to approximate spatial reasoning. *IEEE Transactions on Fuzzy Systems*, 12(6), 745–754.

Lim, M., Metzler, R., & Bar-Yam, Y. (2007). Global pattern formation and ethnic/cultural violence. *Science*, 317, 1540–1544.

Liu, K., & Shi, W. (2006). Computing the fuzzy topological relations of spatial objects based on induced fuzzy topology. *International Journal of Geographical Information Science*, 20(8), 857–883.

Liu, L., Wang, X., Eck, J., & Liang, J. (2005). Simulating crime events and crime patterns in a RA/CA model. In: F. Wang (Ed.), *Geographic information systems and crime analysis* (pp. 197–213). Reading, PA: Idea Publishing.

Lynch, K. (1960). *The image of the city*. Cambridge, Massachusetts: MIT Press.

McCord, E. S., Ratcliffe, J. H., Garcia, R. M., & Taylor, R. B. (2007). Nonresidential crime attractors and generators elevate perceived neighborhood crime and incivilities. *Journal of Research in Crime and Delinquency*, 44(3), 295–320.

Pyle, G. F. (1974). *The spatial dynamics of crime*. Chicago: Department of Geography, University of Chicago.

Quetelet, L. A. J. (1842). *A treatise on man and the development of his faculties*. Edinburgh: W & R Chambers.

Ratcliffe, J. H. (2006). A temporal constraint theory to explain opportunity-based spatial offending patterns. *Journal of Research in Crime and Delinquency*, 43(3), 261–291.

Rengert, G. F., & Wasilchick, J. (2000). *Suburban burglary: A tale of two suburbs*. Springfield, IL: C.C. Thomas.

Schmid, C. (1960a). Urban crime areas – Part I. *American Sociological Review*, 25, 527–543.

Schmid, C. (1960b). Urban crime areas – Part II. *American Sociological Review*, 25, 655–678.

Shaw, C., & McKay, H. D. (1942). *Delinquency and Urban Areas*. Chicago: University of Chicago Press.

Townsley, M., Homel, R., & Chaseling, J. (2003). Infectious burglaries: A test of the near repeat hypothesis. *British Journal of Criminology*, 43, 615–633.

WAAG Society. (2007). *Amsterdam real time project*. http://realtime.waag.org. Accessed November 27, 2007.

Waller, I., & Okihiro, N. (1978). *Burglary: The victim and the public*. Toronto: University of Toronto Press.

Weisburd, D., Bushway, S., Lum, C., & Yang, S.-U. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology*, 42(2), 283–321.

Wikström, P. O., & Butterworth, D. A. (2006). *Adolescent crime: Individual differences and lifestyles*. Portland: Willan Publishing.

Wilcox, S. (1973). *The geography of robbery*. [The Prevention and Control of Robbery, Vol. 3]. Davis: The Center of Administration of Justice, University of California at Davis.

Winter, S. (1998). Location-based similarity measures of regions. In: D. Fritsch, M. Englich, & M. Sester (Eds.), *ISPRS Commission IV Symposium "GIS Between Visions and Applications"* (Vol. 32(4), pp. 669–676). International Archives of Photogrammetry and Remote Sensing, Stuttgart, Germany.

Xue, Y., & Brown, D. E. (2006). Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decision Support Systems*, 41(3), 560–573.

Yeung, D., Chen, C., Tsang, E., & Lee, J. (2005). On the generalization of fuzzy rough sets. *IEEE Transactions on Fuzzy Systems*, 13(3), 343–361.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.