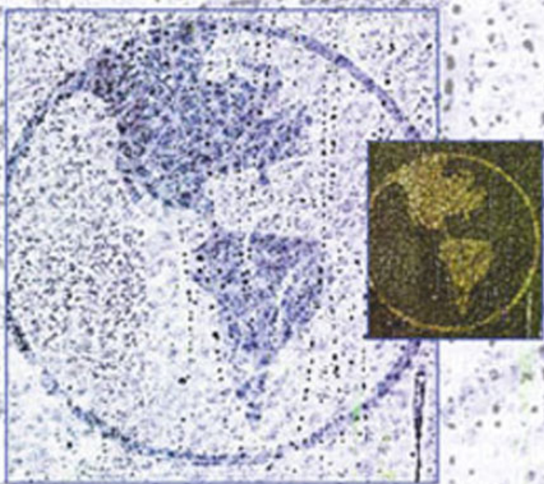
FAHUI WANG

Geographic Information Systems and Crime Analysis



Geographic Information Systems and Crime Analysis

Fahui Wang
Northern Illinois University, USA

IDEA GROUP PUBLISHING
Hershey • London • Melbourne • Singapore

Acquisitions Editor: Mehdi Khosrow-Pour
Senior Managing Editor: Jan Travers
Managing Editor: Amanda Appicello
Development Editor: Michele Rossi
Copy Editor: Dawne Brooks
Typesetter: Marko Primorac
Cover Design: Lisa Tosheff
Printed at: Integrated Book Technology

Published in the United States of America by
Idea Group Publishing (an imprint of Idea Group Inc.)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@idea-group.com
Web site: http://www.idea-group.com

and in the United Kingdom by
Idea Group Publishing (an imprint of Idea Group Inc.)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 3313
Web site: http://www.eurospan.co.uk

Library of Congress Cataloging-in-Publication Data

Geographic information systems and crime analysis / Fahui Wang, editor.

p. cm.

Includes bibliographical references and index.

ISBN 1-59140-453-3 (h/c) -- ISBN 1-59140-454-1 (s/c) -- ISBN 1-59140-455-X (ebook)

1. Crime analysis--Data processing. 2. Geographic information systems. 3. Digital mapping. I. Wang, Fahui, 1967-

HV7936.C88G46 2004

363.25'6'0285--dc22

2004022146

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Geographic Information Systems and Crime Analysis

Section III: Geographic Profiling

Section IV: Crime Monitoring and Tracking

Section V: New Methods and Technologies

Preface

According to Gewin (2004), in early 2004, the U.S. Department of Labor identified geotechnology as one of the most important emerging and evolving fields. For the last decade or so, as Clarke (2003) points out, the Geographic Information Systems (GIS) industry "has seen double-digit annual growth." Among GIS applications related to socioeconomic issues or public policy, crime analysis is one of the most active fields.

GIS and related geotechnologies have turned crime mapping into a powerful decision-making tool for law enforcement agencies. Based on a 1997 survey conducted by the Crime Mapping Research Center (CMRC) of the National Institute of Justice (NIJ, the research arm of the U.S. Department of Justice), computerized *crime mapping* in law enforcement agencies has experienced rapid growth. See Figure 1 according to Weisburd and Lum (2001). *Crime mapping* is often used as a synonym for *GIS-based crime analysis*. However, GIS is beyond mapping, and increasingly so for crime studies and control.
There are three objectives to this book:

- The first is to showcase *a diverse array of GIS applications in crime analysis that are not limited to crime mapping* (though crime mapping remains the primary and fundamental function of GIS). The book covers themes from general issues such as GIS as a communication process and inter-jurisdictional data sharing to specific applications in tracking serial killers and predicting juvenile violence; from routine GIS tasks such as geocoding and buffer analysis, to advanced simulation models; and from neighborhood violence and crime in and around public housing to homicide across southern states in the U.S.
- The second objective is to feature *a broad range of new methods and techniques* including geographic profiling, agent-based modeling, GPS tracking and web GIS.

- The third objective is to *bridge the academics and practitioners* for crime analysis and crime control. Contributors range from university professors, criminologists in research institutes to police chiefs, GIS analysts in police departments and consultants in criminal justice.

I believe that this book is useful to a wide range of audience in both academia and in law enforcement agencies.

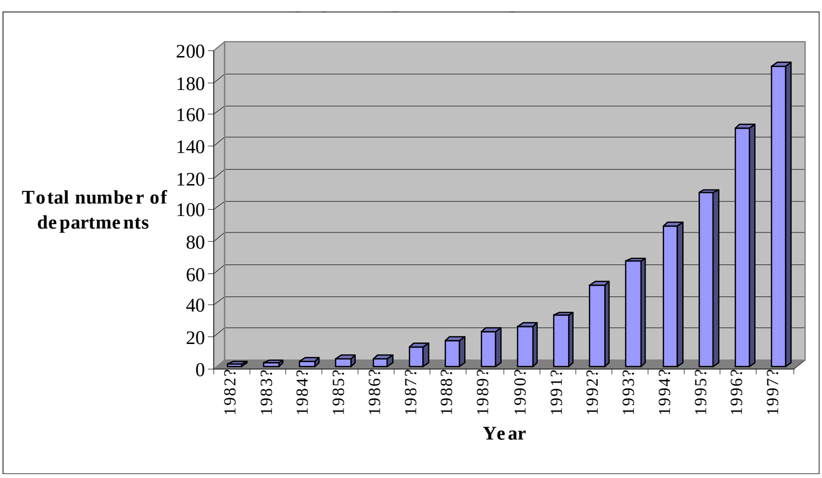
In the following, I will provide a brief overview of all chapters and highlight what I consider the most important contributions. The overview serves as a tour guide in the "hallway", and points the reader to the various "rooms," where one may explore the rich contents that are of most interests to them.

The first section contains two chapters covering a common issue: *data sharing*. GIS plays an important role in enhancing data sharing, improving public access to data, and assisting decision making in local governments. In Chapter I, *GIS as a Communication Process: Experiences from the Milwaukee COMPASS Project*," Jochen Albrecht and James Pingel report their experience from the Milwaukee COMPASS (Community Mapping, Planning and Analysis for Safety Strategies) project. The National Institute of Justice (NIJ) initiated the COMPASS program in 1999, with an explicit emphasis on using GIS for analyzing public safety problems and on collaborative partnerships with actors outside the criminal justice community. The Milwaukee COMPASS project used a web GIS to foster communication among public safety programs as well as between

Figure 1. Growth of computerized crime mapping in law enforcement agencies in the U.S.

1997* CMRC Survey:

Cumulative Distribution of Crime Mapping Adoption



government agencies and local communities, and demonstrated the problem-solving capabilities of GIS. Chapter I details the project's implementation process and various challenges, which may be useful for many agencies that are pursuing similar programs. In Chapter II, *Interjurisdictional Law Enforcement Data Sharing Issues: Benefits of the Use of Geo-Spatial Technologies and Barriers to More Widespread Cooperation*, Mark R. Leipnik and Donald P. Albert cite various examples to illustrate the importance of using GIS to create and disseminate geospatial data across jurisdictions and across departments within a jurisdiction. They also discuss the benefits and barriers for data sharing among law enforcement agencies.

The second section includes three chapters examining several issues related to *data quality*. In Chapter III, *Garbage In, Garbage Out: Geocoding Accuracy and Spatial Analysis of Crime*, Tess McCarthy and Jerry Ratcliffe examine three sources for spatial data inaccuracy (conceptual, positional and attribute) and their impact on crime analysis. Drawn from the experience of geocoding burglary records in an Australian city, they also compile the guidelines for the best practice in geocoding address data. In Chapter IV, *Disaggregating the Journey to Homicide*, Elizabeth Groff and J. Thomas McEwen focus on distances traveled by homicide offenders and by victims. They analyze the differences by homicide motive, by offenders and victims by sex and age, and between Euclidean and street distances. One interesting finding is that both victims and offenders tend to be involved in homicide incidents close to their residences (i.e., a median distance of 0.54 miles for victims and 0.74 miles for offenders). One may find the result useful for justifying a common practice in many studies utilizing aggregate data: using homicide incident locations as approximate locations of offenders. In Chapter V, *Constructing Geographic Areas for Analysis of Homicide in Small Populations: Testing Herding-Culture-of-Honor Proposition*, Fahui Wang and Van M. O'Brien warn the small population problem in analyzing rare events, such as homicide, and propose two simple geographic approaches to mitigate the problem. They apply the techniques to testing the herding-culture-of-honor hypothesis proposed by Nisbett and Reaves, and they find that the herding-culture-of-honor proposition was merely an artifact of unreliable estimate of homicide rates, particularly in areas with small population.

The third section has three chapters on *geographic profiling*, a methodology for analyzing the geographic locations of a linked series of crimes to determine the unknown offender's most probable residence area. In Chapter VI, *Geographic Profiling for Serial Crime Investigation*, D. Kim Rossmo, Ian Laverty and Brad Moore discuss the theoretical foundation for geographic profiling such as the Rigel geographic profiling system: commonly known as the distance-decay law with consideration of a buffer zone around an offender's residence (or workplace). Note that a geographic profile does not pinpoint a single location, but rather provides an optimal search strategy. Its value may be enhanced when it is combined with other information such as a behavioral profile and neighborhood land use and demographic data. In Chapter VII, *Single Incident Geographical Profiling*, Richard Z. Gore, Nicholas J. Tofiluk and Kenneth V. Griffiths explore a data intensive method that applies geographic profiling techniques in a single incident. The method predicts offender residence by computing the relative frequencies of offender residences obtained from arrest record data for all who have committed crimes around the incident location in the past. They evaluate the effectiveness of different geographic filters and examine the boundary effects in each scenario. In Chapter VIII, *Geographic Profiling and Spatial Analysis of Serial Homicides*, Sunghoon Roh and Mark R. Leipnik use the case of serial killer Robert Yates to illustrate how GIS-based spatial analysis techniques, particularly geographic profiling, was used in serial homicide investigations.

The fourth section is on *crime monitoring and tracking* with three chapters. Chapter IX, *Geographic Surveillance of Crime Frequencies in Small Areas*, by Peter A. Rogerson, describes a system for detecting any increase in any area's crime frequency. The system compares cumulative differences between the observed and expected crime frequencies in an area, and identifies areas that experience any significant increases in crime activities. Chapter X, *Application of Tracking Signals to Detect Time Series Pattern Changes in Crime Mapping Systems*, by Wilpen L. Gorr and Shannon A. McKay has a similar objective: detecting areas with significant changes in crime patterns. Gorr and McKay apply a technique used widely in management science, so-called "smoothed-error-term tracking signal," which uses all prior data to estimate trend and seasonality as the counterfactual basis of comparison and quickly detects step jumps and outliers. Both Rogerson's method and the Gorr-McKay method are implemented in GIS. Chapter XI, *Integrating GIS, GPS and MIS on the Web: EMPACT in Florida*, by Gregory A. Frost, describes the EMPACT (Electronic Monitoring Protection and Crime Tracking) project in Florida. The EMPACT automatically correlates data from GPS tracking of offenders (probationers, parolees, and offenders on pre-trial release) and local crime incident data through a web-based interface, and determines whether a tracked offender is at the scene of a crime incident. It demonstrates how geotechnologies, including GIS and GPS, are used on the frontlines for crime prevention and control.

The fifth section showcases five case studies using *new methods and technologies*. Chapter XII, *Simulating Crime Events and Crime Patterns in a RA/CA Model*, by Lin Liu, Xuguang Wang, John Eck, and Jun Liang, use a cellular automaton (CA) model to simulate crime patterns. Based on the routine activity theory, the model considers offenders, targets and crime places as individual agents and simulates crime patterns based on the interaction between these three agents at a specific time. The calibrated simulation model generates crimes similar to actual crimes in both the total number of crimes and their spatial distribution. This demonstrates the promise of using such a model as a virtual laboratory for predicting future crime patterns based on different conditions. Chapter XIII, *Integrating GIS and Maximal Covering Models to Determine Optimal Police Patrol Areas*, by Kevin M. Curtin, Fang Qui, Karen Hayslett-McCall and Timothy M. Bray, applies an optimal covering model (commonly used in location and allocation studies), to finding the most efficient spatial distribution of police patrols. The method takes GIS data layers of incidents and road network, uses linear programming to formulate the optimization problem, and finds heuristic solutions that increase the current level of police service. The method should be useful to decision makers by providing alternative (and better) police patrol covering scenarios. Chapter XIV, *Web GIS for Mapping Community Crime Rates: Approaches and Challenges*, by Tung-Kai Shyy, Robert J. Stimson, John Western, Allan T. Murray, and Lorraine Mazerolle, describes a prototype web GIS for mapping crime rates in Brisbane, Australia. In addition, they discuss the challenges for web GIS applications and offer suggestions to overcome the barriers. Chapter XV, *Identifying "Hot Link" Between Crime and Crime-Related Locations*, by Yongmei Lu, goes one step further than the common hot spot analysis. Her objective is to identify "hot links", i.e., the links between crime incidents and the offenders' residence locations or the links between crime incidents and the victims' residence locations. Like the work by Groff and McEwen (Chapter IV), her work represents the efforts of using GIS in analyzing journey-to-crime (J2C) patterns. Chapter XVI, *Remote Sensing and Spatial Statistics as Tools in Crime Analysis*, by Dongmei Chen, John R. Weeks, and John V. Kaiser Jr., discusses remotely sensed satellite imagery to define land use variables that may be associated with criminal activities. The newly-defined variables, along with traditional socioeconomic variables, are used to explain intraurban variation of crime rates in a regression model that accounts for spatial autocorrelation.

Two case studies on *neighborhood crimes* are collected in the final section. Chapter XVII, *Routine Activities of Youth and Neighborhood Violence: Spatial Modeling of Place, Time and Crime*, Caterina Gouvis Roman, aims to explain neighborhood violence by physical environment, neighborhood characteristics, and routine activities of youth. A new angle of the work is its emphasis on how time influences crime patterns across places and environments. It uses a spatial lag model to account for spatial dependence. Chapter XVIII, *Measuring Crime in and around Public Housing Using GIS*, by Harold R. Holzman, Robert A. Hyatt and Tarl Roger Kudrick, addresses the debate whether high crime rates associated with public housing are attributable to the housing itself or the neighborhood surrounding it. Using extracted crime counts in public housing developments and the surrounding neighborhoods in three cities, the study reveals that the answer depends on the types of crime. The risk of aggravated assault in public housing communities is much higher than in the surrounding neighborhoods, whereas risk of property crimes such as burglary, larceny and car theft is much lower.

References

Clarke, K.C. (2003). *Getting started with geographic information systems*. Upper Saddle River, NJ: Prentice Hall

Gewin, V. (2004). Mapping Opportunities. *Nature*, 427, 376-377.

Weisburd, D. & Lum, C. (2001). Translating research into practice: Reflections on the diffusion of crime mapping innovation. Keynote speech at the *Fifth Annual International Crime Mapping Research Conference*, Dec. 1-4, Dallas, TX.

Acknowledgments

Editing the book turned out to be easier than I had anticipated thanks to the enthusiastic responses to my initial call for chapter proposals. The only way for an editor to ensure the success of any edited volume is through high quality and high quantity submissions. I was happy to receive more than 40 proposals for chapters, which led to 30 draft submissions of full chapters. All submissions were subject to a peer review process and 18 chapters were finally selected after revisions (in some case as many as three to four rounds). Most of the contributors served as reviewers for the others. In addition, I also sought the professional insights and detailed comments from the following outside reviewers: Chris Badurek, Patricia Brantingham, Kevin Bryant, Frank Cullen, Stuart Hamilton, Jaishankar Karuppannan, Peter Schmitz and Yifei Sun, to whom I am grateful. It was my sole decision whether to include a chapter or not while considering reviewers' comments. In many cases, the rejections were a difficult decision, and some were made merely for balancing the book's coverage of themes. I thank all of the authors for their support and diligent participation in this project.

Finally, I would like to thank Mehdi Khosrow-Pour, Acquisitions Editor, for his encouragement and guidance, Jan Travers, Senior Managing Editor, who has always been responsive to my endless questions, and Amanda Appicello, Managing Editor, who worked hard with myself and all of the contributors through the production process.

Section I

GIS and Data Sharing

Chapter I

GIS as a Communication Process: Experiences from the Milwaukee COMPASS Project

Jochen Albrecht, Hunter College CUNY, USA

James Pingel, State of Wisconsin Sentencing Commission, USA

Abstract

We examine the role of GIS in communication and decision-making processes by re-interpreting the experiences of the Milwaukee COMPASS Project (Community Mapping, Planning and Analysis for Safety Strategies) in the light of Enhanced Adoptive Structuration Theory. Using numerous practice-derived examples, we conclude that GIS not only facilitates and strengthens communication, but can be used to defuse political constraints to collaborative decision making.

Introduction

This chapter focuses on the communication process of the Milwaukee COMPASS, a federally funded demonstration project to use GIS to improve data sharing, public access to data, and reliance on data in public decision making. The overall goal of this chapter is an investigation of how well geographic information technologies support communication between city government, citizenry, technical staff, neighborhood organizations, and academic researchers. Probably the single most important aspect of Milwaukee's success story has been the common belief, *a priori*, among the participants, that communication, openness and collaboration are valuable to the policy-making and implementation processes. Without this tenet, they would probably not have been in the position to overcome many of the difficulties inherent to any multi-agency project. The political, technical and financial barriers to implementation are high, and will not be overcome unless the value is readily seen by a majority of participants. The experiences of the Milwaukee COMPASS project illustrate the power of GIS as a tool for improved communication across sectors in a community, and for opening new lines of communication among actors who simply need a common language to begin meaningful dialogs. It is, indeed, the power of GIS as a communication tool that facilitates the shared value among participants and makes community-wide, collaborative problem-solving efforts possible.

This research focuses on the communication process of innovation, diffusion, and adoption of spatial technologies to combat crime and foster healthy neighborhoods. The experiences of the Milwaukee COMPASS project are illustrative of these concepts because GIS played a central role in the project's mission, "to make public safety decision making more collaborative, strategic and data-driven" (City of Milwaukee, 2003).

GIS and Communication Processes

At least four forms of GIS communication processes have been described in the relevant literature:

(a) GIS as a mapping tool, mirroring the cartographic communication process (Foote & Crum, 1995);

(b) GIS as part of a decision support system and facilitating the communication between its various components – usually in a PPGIS (Merrick, 2003) or GIS-in-developing-countries context (Jenssen, 2002);

(c) Emerging as a tool for holding managers accountable for measurable results, beginning in law enforcement (Bratton & Knobler, 1998; Weisburd & Lum, 2001; Stoe, Watkins, Kerr, Rost & Craig, 2003) and spreading to other disciplines (Swope, 2001); and

(d) The low-level technical aspects of (b) and (c), that is, the communication between software objects in a Microsoft COM or Unix-based CORBA environment (Peng & Zhou, 2003).

Of course, we cannot tackle information sharing without thinking about how this information is communicated.

Communicating Values through Information

One of the main issues with establishing a geographic information framework that partners as diverse as the COMPASS participants (see Table 1) are confronted with is the difference in agendas and value systems that these partners bring to the table. Agendas are usually well spelled out – they have been part of the original funding application to the National Institute of Justice (City of Milwaukee, 2001). Values, on the other hand, are usually not part of the communication among technical folks. The broad social values inherent in geospatial databases may be inescapable (Pickles, 1995) and, to the extent that they are taken for granted, not easily documented. However, the values embedded in databases as a function of institutional characteristics can be articulated and documented in metadata and subsequently communicated to the GIS user. This communication process is important since it affects the user's understanding of the limitations of the GIS and facilitates its appropriate use. The primary mechanisms that have evolved to serve this communication process are based on the Federal Geographic Data Committe's (FGDC) descriptors of geospatial data quality.

In general, we must take into account two factors integral to the role of information and technology in communication. First, different tools are used for communicating information. Second, a condition is essential for assuring good communication: mutual understandability among the partners. We have to assure that all partners engage in this mutual understandability. In linguistic or information science terms, the message emitted by the transmitter must be understood by the receiver (Figure 1) – that is, the relation between the signifier and the signified must be the same for all partners. For communication to work, not only the ability of using a code, but also the will (or the obligation) is necessary. The roots of cooperation are found in the very structures of language (Habermas,

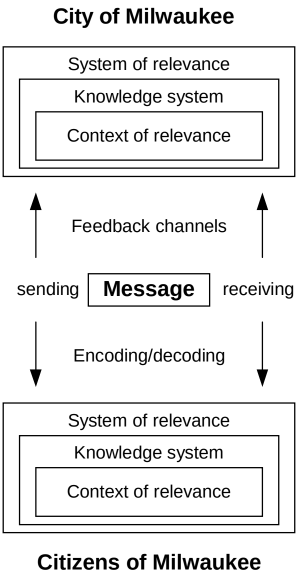
Table 1. Milwaukee COMPASS Partnerships

Data Collection	
Police Department Assessor's Office Department of Neighborhood Services Fire Department Health Department Municipal Court Public Library Citywide Housing Coalition Milwaukee Public Schools Milwaukee County Children's Court Milwaukee County Sheriff's Office District Attorney's Office	Judicial Oversight Demonstration Project Safe & Sound, Inc. Community Partners Department of City Development Department of Corrections UWM - EPIC Northwest Side Community Development Corp. Project UJIMA – Children's Hospital YMCA of Metro Milwaukee Boys & Girls Clubs of Greater Milwaukee Police Athletic League (forthcoming)
Data Entry and Management for Community Groups	
Citywide Housing Coalition • LAND • Sherman Park • Westside Neighbors • ACTS • Harambee • Neighborhood Housing Services • Metcalfe Park • St. Martin DePorres	Community Prosecutors • District Attorney • Milwaukee Alliance • Harambee Ombudsman • Drug Abatement Hotline (pending)
Policy/Research Projects	
Urban League of Milwaukee Milwaukee County Sheriff's Department Weed & Seed / Community Partners City Attorney Department of Neighborhood Services Milwaukee Public Schools – School Safety (NIJ project) Department of City Development – Planning Division Judicial Enforcement Demonstration Initiative Task Force on Family Violence Safe & Sound, Inc. The Mayor's Commission on Crime Boys & Girls Clubs YMCA of Greater Milwaukee County Department on Aging U.S. Attorney, district office Firearm Injury Center	Community Advocates Brighter Futures Initiative Northwest Side Community Development Corp. Sherman Park Residents Association Community Care Organization Third District Community Justice Center Department of Corrections Mercy Memorial Baptist Church Merrill Park Neighborhood Association Metcalfe Park Residents Association Midtown Neighborhood Association Community Block Grant Administration Martin Drive Association Project UJIMA City Clerk – Nuisance Service Calls program

1990). If the project partners do not have an implicit commitment, then disagreement and misunderstandings arise. Hence, for good communication, it is not only important to speak the same language (English, German, Chinese), but also to know the values of the different actors.

Next, we need to explore in general the application of GIS as a tool to enhance the discovery and learning process and for the communication of findings. Nyerges and Jankowski's (1997) Enhanced Adaptive Structuration Theory

Figure 1. Sociointitutional view of the GIS communication process



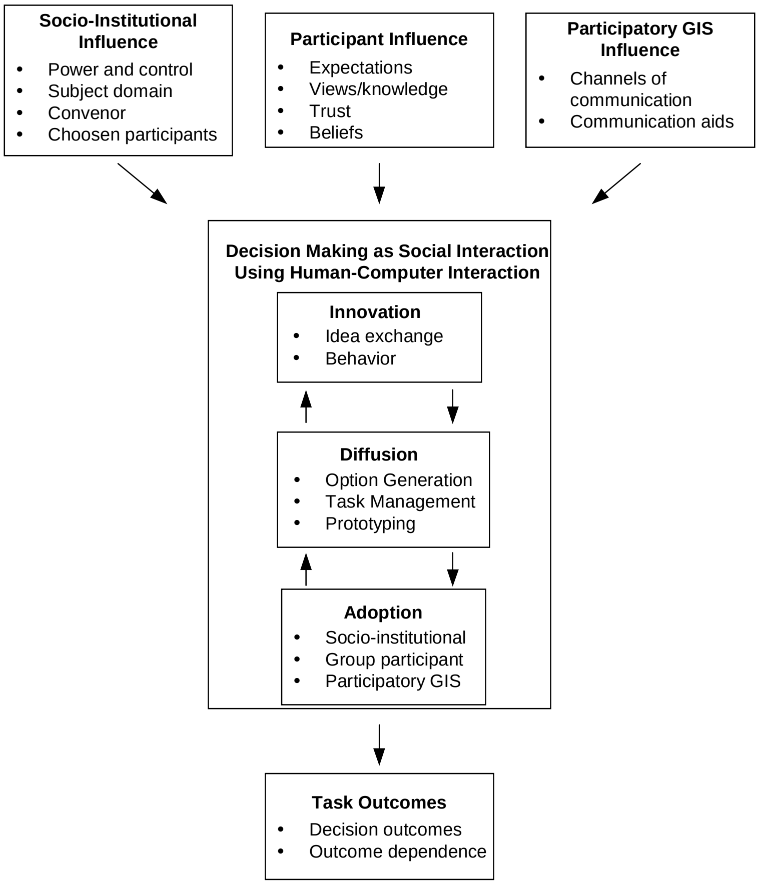
(EAST) is particularly instructive, as well as applicable to the experiences in Milwaukee COMPASS. Derived from Gidden's 1984 Theory of Structuration, it lists 21 aspects of (GIS-based) collaboration in three categories identified as "convening," "social interaction" and "outcome." Our adaptation identifies basically the same phases but puts the emphasis on innovation, diffusion and adoption of communication technologies in general and GIS in particular (Figure 2).

This matches nicely with the works of Ramasubramanian (1995, 1999), whose identification of criteria for the successful adoption and use of spatial technologies in nonprofit organizations prove to be applicable to the communication between (local) government agencies and nonprofits as well.

We argue that in order to foster local ownership of GIS, one must understand the existing networks of communication and cooperation and utilize these to make GIS more relevant to citizens. Good communication is a critical first step to facilitating local ownership of GIS – both real and perceived – which in turn enhances its potential for long-term success at all levels of crime prevention.

Before we can apply these concepts and models to the Milwaukee COMPASS experiences, some history and background are necessary.

Figure 2. Milwaukee COMPASS project interpreted in the light of Nyerges and Jankowski's (1997) Enhanced Adaptive Structuration Theory 2 (EAST2) as a conceptual map for understanding the communication processes that lead to a successful partnership



History of COMPASS

The US Department of Justice, Office of Justice Programs, National Institute of Justice (NIJ) initiated the COMPASS program with a pilot grant award to the Seattle Police Department in 1999 (Pendleton, 2000). In November 2001, after a competitive application process, NIJ awarded Milwaukee, WI, with the second two-year COMPASS grant. A third grant was awarded to the City of Redlands, CA, in spring of 2002.

A brief recounting of the history of the NIJ's efforts to develop and test collaborative, data-driven problem-solving strategies is necessary to fully appreciate the experiences of the Milwaukee COMPASS Project.

The Boston Gun Project

In the late 1990s, the Boston Police Department, the U.S. Attorney for the Eastern District of Massachusetts and others in the community created a partnership with researchers at Harvard University's Kennedy School of Government to address the rampant problem of juvenile gun violence. The researchers performed an intensive analysis of offense and arrest reports, as well as closely guarded intelligence information on gangs, gang members and other actors. Working closely with police officers and other front-line practitioners, they developed a more thorough, shared understanding of gun violence in Boston. The results of this work led the criminal justice community to develop some very targeted, and ultimately very successful, interventions (Kennedy, Braga & Piehl, 2002). Kennedy et al. also argue that the interactive problem-solving *process* was more instructive and more important to replication than the strategies that emerged: "Perhaps the most fundamental lesson here is that the basic approach the project followed – serious, sustained attention to an important problem, with ambitious goals – is worthwhile… One suspects that many difficult problems might appear less so if similarly addressed" (p. 44). In other words, it was a reliance on the data that made for effective, productive communication, which in turn led to progress against a seemingly intractable problem.

SACSI

In March 1998, the NIJ, which serves as the research and development arm of the U.S. Department of Justice, launched the Strategic Approaches to Community Safety Initiative (SACSI) to test a specific framework for combating local crime problems (Solomon, 1997). The stated goals were explicitly developed to replicate the elements of the Boston Gun Project in other communities: a) formation of an interagency working group; b) enhancement of a research and technology infrastructure; and c) use of a defined set of problem-solving process steps. Five cities were chosen to participate in the two-year pilot project, with another five sites selected in 1999. The United States district attorneys' offices served as the coordinating agency for each local initiative, and convened a collaborative group of law enforcement practitioners and criminology researchers in their local communities. Each site selected a general category of public-safety problem on which to focus (for example, gun violence, juvenile violence, sexual assault). Local academics, in an action research role, facilitated a structured, data-driven, problem-solving approach to understanding the selected problem, developing and implementing broad-based, strategic solutions to address the problems defined.

Results varied across the sites (Groff, 2000). But again, it was the *process* that endured: the key innovation of the SACSI process was a problem-solving process that became known as *incident review*. Based on the Boston process, incident review brings many practitioners – and their data – together to understand and solve problems. In terms of Figure 2, the incident review is a prototypical innovation that at the same time allowed the exchange of ideas, focused communication, eased task management, and finally facilitated the adoption solutions by all partners.

And Finally, COMPASS

In partnership with Chief Norm Stamper of the Seattle Police Department, NIJ staff created the COMPASS grant program in 1999. With GIS as a central component, COMPASS could be characterized as a fusion between the problem-solving approach of SACSI and the successful reliance on GIS by the New York Police Department as a tool for both strategic planning and holding managers accountable for results (Bratton & Knobler, 1998; Dussault, 1999; Silverman, 1999).

NIJ's description of COMPASS reveals a widening of the scope of the problem-solving process vis á vis SACSI and the Boston Gun Project:

"In recent years, a shift has occurred in local juvenile and criminal justice policy development toward a more collaborative approach that relies on analyzing public safety problems to develop strategic interventions to address them. This approach needs to be supported by timely, accurate, multi-disciplinary and automated data with a geographic reference. Jurisdictions that have developed such data systems, analytic capacity, and collaborative partnerships have experienced great success in reducing crime and addressing public safety problems." (NIJ, 2000)

The two key differences between COMPASS and the problem-solving efforts that preceded it are an explicit reliance on GIS and an opening of the process to actors outside the criminal justice community. In other words, COMPASS was an attempt to replicate the problem-solving process across many different policy areas and policy issues, within a single local site. And GIS is *the* communication tool that is critical to focusing such a broad, ambitious level of community dialog into productive, problem-solving processes.

It is a poignant illustration of this chapter's key tenets that GIS proves to be the critical element in considering a project as broad and ambitious as COMPASS. The utility of GIS as a problem-solving tool that encourages both collaboration and strategic thinking is one of the most important aspects of the "experiment" that is COMPASS.

Milwaukee, in its application for funding as a COMPASS pilot site, posited that grant resources could be used to apply the community's existing GIS capacity to improving decision making in a number of arenas – basically by providing a platform and a process for more effective communication and collaboration across sectors, and organizational "silos." Milwaukee had the distinct advantage of a strong GIS infrastructure already in place. The city of Milwaukee first implemented GIS in 1976, and has continued to innovate its geographic systems. The University of Wisconsin-Milwaukee's School of Architecture, Regional and Urban Planning, the research partner in the project, offers one of the oldest GIS certificate programs in the nation. Implementation was a matter of using this infrastructure and technological capacity to reach out to policy-makers and problem-solvers.

Implementation

Milwaukee COMPASS Mission and Goals

The Milwaukee COMPASS project team adopted a process-oriented mission: "To make public safety and other decision making in Milwaukee more strategic, collaborative and data-driven." The team adopted four goals in support of that mission:

1. Create a shared data infrastructure.
2. Use the Internet to make data available to the community.
3. Demonstrate the problem-solving potential of GIS.
4. Support ongoing collaborations and evaluations of public-safety programs in the community.

A complete report on implementation of the project is beyond the scope here. It is, however, instructive to describe the COMPASS project in terms of the process of innovation, diffusion and adoption of new technologies. In the sections that follow, we use examples from the COMPASS experiences to illustrate the theoretical concepts discussed above.

Innovations

While the project did not result in the creation of any substantially new technologies, it was essentially an attempt to apply such technologies as GIS, HTML and Java to problems directly relevant to neighborhood public safety in Milwaukee.

- *The Web site: Pushing maps to the public*. COMPASS aggressively exploited the Internet as a means to open new lines of communication with residents, and to improve existing communication between community-based organizations and public agencies. In fact, the first product that the project produced was an Internet Map Server-based (IMS) Web site, pushing crime data and other information to residents (see Figure 3).
- *Java-enabled Web applications: Pulling data from the community*. Eventually, COMPASS staff adopted Java-based tools to enable community groups to both download raw data sets and to enter data directly onto the city's Web site, *www.milwaukee.gov/compass*.

"Community Mapping" is a GIS application using ESRI's ArcIMS technology to give residents the ability to create maps of specific neighborhoods, incorporating data of their choosing. Users include individual residents, community organizers, community-based organizations and even public officials who want a quick, intuitive view of their own data. Figures 4 and 5 illustrate the ability to interactively define different map layers to view, and to zoom-in on specific neighborhoods to reveal more detailed geographic information.

This was in response to the Citywide Housing Coalition, a group of neighborhood organizations that has been working with the city's Department of Neighborhood Services (building code compliance) for several years on prioritizing and addressing dilapidated housing. The Java application streamlines the neighborhood groups' data entry, automatically geocodes the data they collect about problem housing, and tightens the communication between the groups and the city. This helps to both reduce the time frame and increase the accuracy of communication between diverse groups. In terms of the Figures 2 and 6, this is a prime example of an innovative use of technology to improve the timeliness and accuracy of the group processes of idea exchange, task management and behavior.

Figure 3. COMPASS utilizes GIS as a communication tool over the Internet and in focused group settings

Figure 4. Milwaukee is one of a growing number of cities in the U.S. where ordinary citizens have access to both crime and property data.

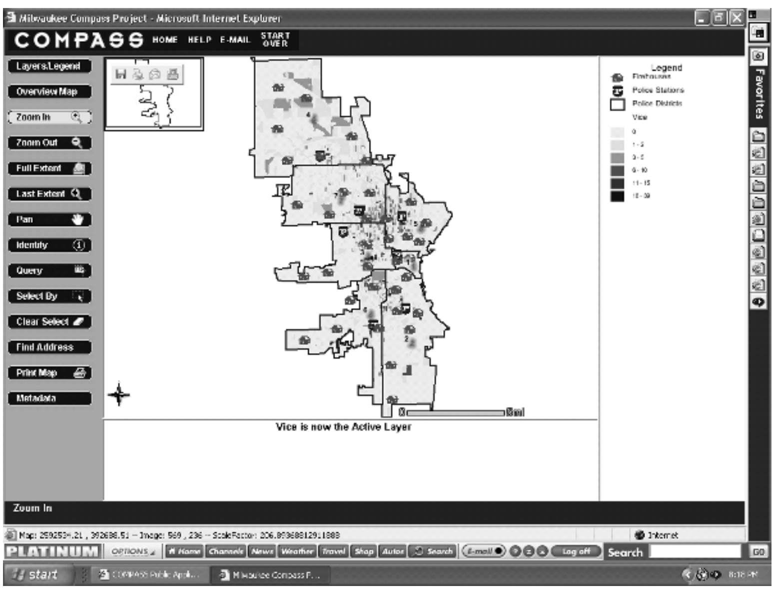
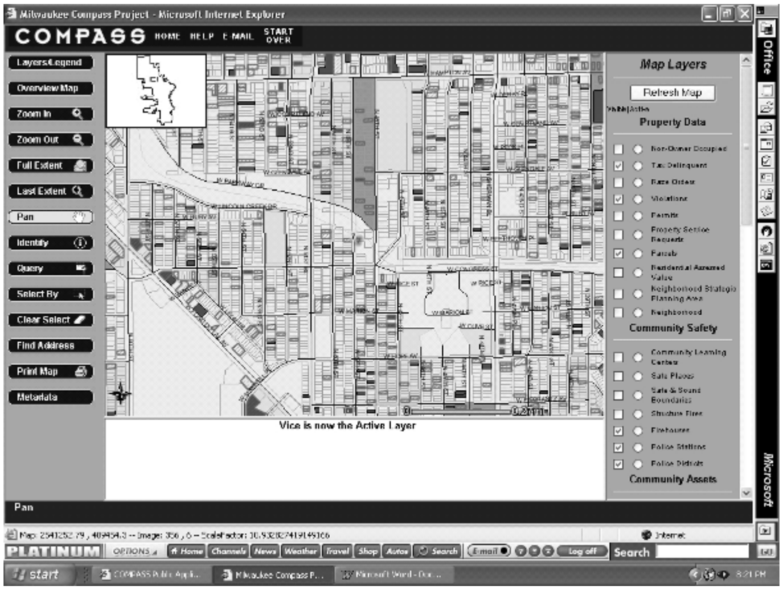


Figure 5. Milwaukee is one of only a handful of cities in the U.S. where ordinary citizens have access to parcel-level crime and property data.



Informal feedback from Internet users told COMPASS staff that people did not necessarily value the mapping capability to discover new things about their neighborhoods. Residents and organizers familiar with a neighborhood and its problems found few surprises in the data that agencies would allow to be published on the public site.

What users did with the maps, however, was to some extent surprising. Many of them used the maps to communicate a problem to their elected officials and other stakeholders not as intimately familiar with a neighborhood's problems. Not only did grant writers use the sight to illustrate need in a particular area, but neighborhood organizations also frequently reported using maps generated on the COMPASS Web site to alert police, building inspectors, and others to patterns that seemed to be emerging. The commissioner of Milwaukee's Department of Neighborhood Services even used COMPASS maps to show aldermen the activities in which his department was engaged in their districts.

"Query & Download" utilizes Java programming to enable users to download raw data sets for their own analysis. Users may select a specific geography, then download selected data sets to an Excel Spreadsheet file. This application demonstrates that geographic analysis is about more than simply map production, since GIS is the technology that enables the user to select a meaningful – and

Figure 6. Repeat cycle of communicative steps in a GIS-based spatial decision support process (important is the wider context, revisiting the original values, and anticipating the political feasibility of policy recommendations)

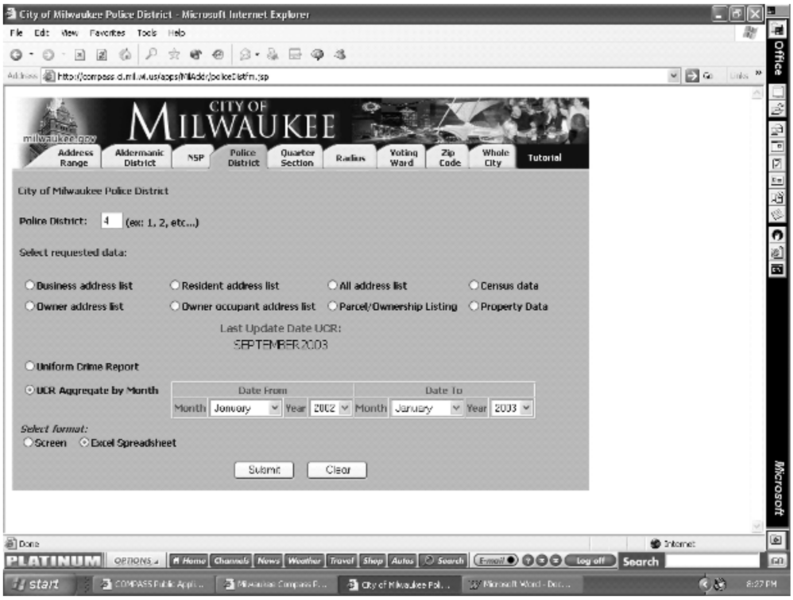
	Project phases		
Reiterated steps	**Innovation**	**Diffusion**	**Adoption**
in a GIS communication process	Values, objectives and criteria	Development of options	The decision
Gather…	And refine value as a basis for objectives	Primary criteria as a basis for option generation	Values, criteria and option list scenarios for an evaluation
Organize…	Objectives as a basis for criteria and constraints	And apply approach(es) for option generation	Approaches to priority and sensitivity analyses
Select…	Criteria to be used in analysis as a basis for generating options	The feasible option list	Recommendation as a prioritized list of options
Review…	Criteria, resources, constraints and standards	Option set(s) in line with resources, constraints and standards	Recommendation(s) in line with original value(s), goal(s) and objectives

The traditional GIS-based decision process

manageable – subset of very large, citywide data sets, and put the data directly to use by downloading to his or her desktop (Figure 7).

- *"CompStat for the Community."* In addition to the Internet, COMPASS staff found it necessary to "take the show on the road," and present its results to its constituent groups (Table 1). COMPASS used specific requests for data or maps as opportunities to demonstrate to Milwaukee's community leaders the power of GIS and of sharing data across organizational boundaries. Perhaps most importantly, GIS – when it is used to integrate diverse sources of data – is a powerful tool for convening diverse constituent groups. Figure 8 is an example of a map that was developed to support a community organizing/public-safety initiative on Milwaukee's south side. The COMPASS director used a laptop and LCD projector to

Figure 7. Query and download option allows citizens to access the raw data for further analysis in a software package of their choice



depict maps of the neighborhood in a meeting that involved residents, faith-based leaders, the local aldermen and the two police captains whose patrol districts included the targeted neighborhood. These people collaboratively interacted with the GIS in real time, reacting to different map layers, and requesting that specific data sources be displayed and the map view be zoomed to a particular section of the neighborhood. The participants' ability to change the map view on the fly, as the discussion around neighborhood safety ebbed and flowed, helped to focus the discussion and led to plans for a collaborative, law enforcement/community response to gang activity and dilapidated housing in a few specific blocks. Again, this is not an innovation in the sense of a new technology, but an application of the technology in a way that was new to the participants and opened up new lines of communication.

Diffusion

In essence, the first and foremost outcome of a successful communication process is the process itself. An answer to the question being sought is only

Figure 8. Example demonstrating the analytical capabilities that COMPASS provides the citizens of Milwaukee (and beyond)

Crime incident trends
January to June, 2002 and 2003

Legend

Safe&Sound_boundaries

citylimits

Density of Shots Fired Incidents

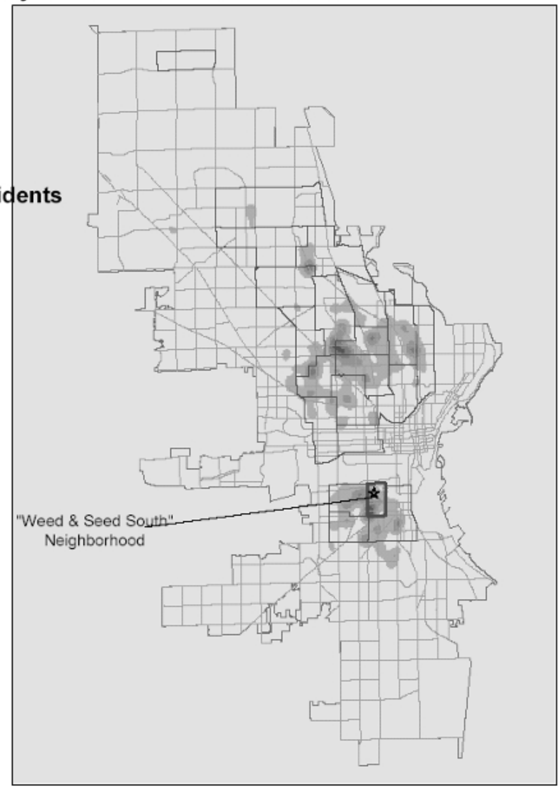
Value

Low Density

High Density

Main Streets

★ Latino Community Center



Analysis
This map shows where shots-fired incidents were most strongly concentrated in 2002. The Latino Community Center is located in the "Weed & Seed South" neighborhood, as defined by the Safe & Sound. It is near, but not directly in one of the City's persistent "hot spots" for violence (depicted as the darkest-red areas).

A comparison of the first half of 2002 to January through June of 2003 suggests that crime is decreasing in this area. And these decreases are steeper than the declines in these crime categories for the entire City, as shown in the table at right

Data Sources: Safe & Sound, Inc. Milwaukee Police Department

Map prepared by: Milwaukee COMPASS Project

	Weed & Seed South Neighborhoo			Citywide		
	Jan-June 2002	Jan-June 2003	Percent Change	Jan-June 2002	Jan-June 2003	Percent Change
Violent Incidents						
Homicide	0	0	-	51	45	-11.8%
Sexual Assault	13	6	-53.8%	471	378	-19.7%
Robbery	29	20	-31.0%	1329	1183	-11.0%
Aggravated Assault	13	8	-38.5%	724	668	-7.7%
"Quality of Life" Incidents						
Burglary	39	23	-41.0%	3092	2567	-17.0%
Retail Theft	0	1	-	542	420	-22.5%
Simple Battery	45	36	-20.0%	2952	2443	-17.2%
Disorderly Conduct	24	17	-29.2%	1458	1217	-16.5%

secondary to that. Diffusion, as a component of communication, is the set of processes employed to ensure that innovations in communication, such as those described above, are applied in practice.

GIS facilitates more complex communication than the two-way dialog modeled in Figure 1. There are often multiple receivers with multiple agendas and environmental constraints. And the communication process is *iterative*. The sender (GIS staff and researchers) often must refine the message (maps, analysis, other portrayals of the data) several times in order to achieve consonance with the receiver (one or more public officials, community leaders or other stakeholders). Thus feedback from the receiver to the sender is a critical part of how the model worked in the case of Milwaukee COMPASS. Several examples from different stages of the project are instructive:

- *Prototyping the Web site.* The diffusion strategy was the technical innovation itself. COMPASS staff very quickly assembled a prototype Web-based, community-mapping interface. This was important to the ability to diffuse the general idea throughout the community – a picture is worth a thousand words, basically. This was viewed as a much more efficient, effective implementation path than developing a comprehensive needs assessment. Community participants have proven to be very effective at fine-tuning the Web tools, as well as specific analytical tools and processes that were developed for specific problem-solving settings. It is doubtful that a structured, thorough (and time-consuming) needs-evaluation process would have yielded this level of buy-in.
- *Requests for specialized maps: Helping agencies attract funds and tell their story.* As the idea became embedded in institutions in the community, the demand rose for COMPASS' GIS staff to provide maps and analysis of an agency's data for that agency's own review and dissemination. Although almost antithetical to the goal of *collaborative* decision making, COMPASS did add considerable value to the policy space, by helping agencies tell the story that is inherent in their own data. An unexpected example was the city of Milwaukee Department of Neighborhood Services. The department provided COMPASS with data on citizen complaints about rats and other "vector" nuisances. COMPASS provided a simple map, plotting all rat complaints citywide. The department then used this map to prioritize neighborhoods for trash and alley clean-up projects. They also displayed the map at a press conference, where they publicly announced their spring cleanup strategies. Thus, GIS also enables a form of political communication: government officials using GIS to "spin" a potentially negative story like rat-infested alleys in their favor. This experience

underlines Ramasubramanian's assertion (1999) that one of the main uses of GIS is its capability to assist organizations in reframing their position.

- *Communication loops.* Another example helps illustrate the iterative nature of the communication process. The county sheriff wanted to embark on an initiative to reduce firearm violence. With no new resources to add to the initiative, the department had to make the most strategic, targeted use of existing resources. They approached the COMPASS project for help. The first step was to plot shootings in the city of Milwaukee for 2002. Since the dots covered large portions of the city, the next step was to develop a smoothed-surface map in an attempt to statistically highlight concentrations of gun violence. The resulting kernel-density map suggested to department officials that gun violence was most heavily concentrated in a few neighborhoods across the city. As the sheriff's department implemented the initiative and gained experience on the street, they inquired about the temporal patterns of gun violence. This led to COMPASS staff analyzing the data by day of week and hour of the day. Further refinements, including mapping specific types of drug arrests and adding tax-delinquent properties, helped to both refine the department's implementation on the street, and further improve the communication process between researcher and practitioner.
- *GIS in support of collaboration.* As noted above, Milwaukee had many collaborative initiatives already in place prior to implementation of COMPASS – working to achieve diverse neighborhood-oriented goals from funding after-school safe places to mobilizing against absentee landlords. The COMPASS team made a concerted effort from the beginning to reach these collaborative groups.

The existence of a partnership across sectors or around a specific policy issue or objective, already made the case for a shared data system: if partners were working together, their data should be integrated as well. GIS allowed for a low-cost means of integration, and protection of confidentiality (Mamalian, LaVigne & Groff, 2001; McEwen & Wartell, 2001).

While it is difficult to get busy public executives to take the time, and to suspend their judgment long enough to step through a structured decision-making process, GIS can be used as an especially effective communication tool to overcome communication barriers. Because it is a graphical representation of a particular phenomenon, a map (or, in the case of interactive GIS presentations, a series of maps) can serve as a shared starting point, a way of grounding the multiple perspectives of diverse stakeholders in a common picture of reality. COMPASS is currently working with a group consisting of several faith leaders, a gang outreach worker from a local youth center, three police captains, a staff member from the U.S. Attorney's Office, and a sheriff's captain to target a very small, problem-ridden neighborhood for a collaborative, coordinated "intervention." The group started this process by viewing maps of a larger area, selecting which problems (specific types of crime, housing violations) and assets they wanted to view, and using the GIS software to "zoom" into several specific neighborhoods, before settling on one that presented both serious crime problems and several opportunities for rebuilding the neighborhood. All of this was accomplished in a two-hour meeting in the back room of a community-based organization a few blocks from the problem area. The ability to interactively produce different map views "on the fly," reflecting and channeling the discussion as it flowed, was a very powerful demonstration of GIS making high-level communication highly productive and effective among very diverse participants.

Adoption

The ultimate test of the value of the innovations, and the effectiveness of the diffusion strategies, is how deeply and widely the system becomes embedded in the community. As a means of tying together many different ideas, we attempt in this final section to identify and summarize a set of critical success factors that enable a GIS to improve community communication.

- *Willing and ready receivers.* Public agencies face a trade-off in committing to a collaborative, data-sharing process. The potential for improved results, and the goodwill of participation, must be placed in the balance against the political and public relations risks of "opening" the agency up to criticism and opening its protected data to interpretation. For some agencies, who may be mired in a traditional isolationist mode or politically embattled, the calculation will tilt in the favor of remaining closed to the process. The key piece of advice, and the lesson learned in the COMPASS project, is to take every opportunity to work with those who are willing to participate. Not only will the process be mutually beneficial, but the communication of successful results will also put political pressure on the non-participants to join the bandwagon.
- *Technical capacity of the sender.* It goes without saying that the GIS professionals and researchers conducting the process must be highly capable, well-equipped technologically, and familiar with their data. It is also important to note that the quality and accuracy of the message emitting

from these professionals improves over time, with experience, and with trial and error at providing maps and analysis that generate real communication.

For example, early on the COMPASS staff attempted to form a partnership with the local planning department. The idea was to incorporate crime analysis into a neighborhood economic development plan. However, the presentation of the crime maps and analysis was not directly on point, and the aggressive time frame of the development plan meant an opportunity was lost. This failure to improve communication, however, led to future successes as the staff improved their ability to meet users' needs in a timely fashion.

- *Timely messages*. Obviously, the data must be timely. "Annual report" data, summarizing last year or the year before, does not lead to productive communication or effective strategy. This alludes to another critical type of communication: negotiation of data-sharing agreements. COMPASS was able to develop protocols for public and nonprofit agencies to provide timely, regular updates of information, sometimes only a week old. This was accomplished only through intensive, iterative communication – both verbal and written – that took into account all of the political, technical and fiscal challenges of such an open-ended data-sharing agreement. This negotiation must be explicit of all risks of information sharing, and therefore must describe the benefits of sharing information (and using GIS) equally explicitly.
- *Time and a process to fine-tune the message.* It bears repeating that the process of communication based on GIS is *iterative***.** As a result, GIS analysts must learn to allow enough time for the receivers to absorb the message and generate meaningful feedback. A structured process helps this endeavor (Figure 3). More importantly, if the process is shaped around a critical public policy issue, the participants are motivated to participate in the process.
- *Metadata as communication.* As noted above, metadata is a key piece of any GIS system or infrastructure. Through COMPASS, we learn that a) metadata can take both formal (written, structured) and informal (spoken, unstructured) forms, and b) metadata itself is an aspect of the communication process. To be sure, formal metadata are provided on the COMPASS Web site and conform to the FGDC standards. And, as in many development projects, a lag in documenting metadata has led to confusion on the part of users, when metadata do not accurately describe such details as the time frame of the data being presented.

  But conversations involving metadata happen on a less formal plane as well. For example, when presenting a map in an interactive meeting, a participant might ask something like, "So, what does this map show?" This question then launches a discussion of data elements, time frame, and even the data-capture process. For example, nearly every projection of a specific crime category on a map is accompanied by a discussion of Uniform Crime Reporting, the process of filing police reports and the suspected patterns of unreported crime. These conversations do not reflect FGDC standards, and unfortunately do not always get captured in writing. However, they are a critical piece of the communication, because they ensure that the participants do in fact have as similar an understanding of the shared picture as possible.

- *Being nimble: Adopting to shifting political, policy and funding priorities*. By now, the dissonance between the original four goals and the diverse array of COMPASS projects is no doubt apparent to the reader. This has occurred because the team employed one overriding imperative: do not pass up an opportunity to demonstrate to the local community the capacity of GIS to improve communication and decision-making processes.

  Thus when an opportunity arose, for example, to map complaints about rats, or to build a data-management application for a community housing survey, the COMPASS project staff was nimble enough to adjust goals, objectives and priorities to accommodate any demand for data-driven decision making that arose (Figure 6). In other words, the philosophical approach to the innovation, diffusion and adoption of GIS as a community decision-making tool was to be opportunistic, as opposed to espousing a structured planning model and process.

  This "open" philosophical approach led to a wide diversity of experiences and opportunities. Some of the attempts to integrate GIS into existing, established problem-solving processes failed. Others took unexpected directions. The result is a wide set of experiences that speak to GIS as a communication tool.

- *Willingness to accept failure*. The general concept of "data-driven problem-solving" is a tougher sell than specific tools, such as an Internet mapping interface. Thus, the COMPASS project team members made a number of early attempts to integrate the idea of GIS-driven decision making, planning or problem solving across a wide variety of settings and actors. Some of these, such as an attempt to integrate crime trend analysis into an economic development plan, did not pan out. Others, such as the partnership with the Citywide Housing Coalition, only came about after several attempts at a more aggressive GIS-focused, problem-solving approach to housing problems. GIS practitioners and researchers alike need to realize that they will not "bat 1,000" in their attempts at using GIS in action research. But they also should realize that success often comes only after several iterations, which give their local community time and exposure to the utility of GIS and the value of sharing data. By understanding the general conceptual theories laid out here, and recognizing the critical success factors for effective GIS-oriented communication, they will, over time, increase their success rate and their value to the local community.

Conclusion

COMPASS Milwaukee provides the necessary communication and data management network to support better access to crime-relevant data and facilitate communication among citizens, the scientific community and policy-makers.

From an organizational perspective, GIS provides a common framework, and beyond that even a common language for the technical staff involved. Whether it is a simple question of data format, more difficult issues of data organization, or a really complex problem such as firearm violence, the information technology background of staff in all organizations involved, combined with their peculiar spatial perspective and local geographic knowledge, formed a valuable supporting structure for the COMPASS project.

Acknowledgments

This project is supported by Grant No. 2001-IJ-CX-K005 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

References

Bratton, W., & Knobler, P. (1998). *Turnaround: How America's top-cop reversed the crime epidemic*. New York: Random House.

City of Milwaukee (2001). *Implementation plan – project COMPASS* (Report). Washington, DC: National Institute of Justice.

City of Milwaukee (2003). Project COMPASS Web site. Retrieved November 3, 2004, from *http://isdweb1.ci.mil.wi.us/compass/*

Dussault, R. (1999). Jack Maple: Betting on intelligence. Interview with Jack Maple in *Government Technology, Vol. 5.* Retrieved June 13, 2004, from *http://www.govtech.net/gtmag/1999/apr/maple/maple.shtm*

Foote, K., & Crum, S. (1995). *Cartographic communication*. Last retrieved July 29, 2003, from *http://www.colorado.edu/geography/gcraft/notes/cartocom/cartocom_f.html*

Groff, E. (2000, December). *Strategic approaches to community safety initiative: Enhancing the analytic capacity of a local problem-solving effort* (Internal report). National Institute of Justice. Last retrieved August 25, 2003, from *http://www.ojp.usdoj.gov/nij/pubs-sum/sacsi.htm*

Habermas, J. (1990). Discourse Ethics: Notes on a Program of Philosophical Justification. In: J. Habernas (Ed.), *Moral consciousness and communicative action* (pp. 43-115). Cambridge, MA: MIT Press.

Huxhold, W., & Levinsohn, A. (1995). *Managing geographic information systems projects*. New York: Oxford University Press.

Innes, J. (1996). Planning theory's emerging paradigm: Communicative action and interactive practice. *Journal of Planning Education and Research, 14*, 183-189.

Jankowski, P., & Nyerges, T. (2001). *GIS for group decision making*. London: Taylor & Francis.

Jenssen, J. (2002). *Down to earth: Geographic information for sustainable development in Africa* (Report). Committee on the Geographic Foundation for Agenda 21, Board on Earth Sciences and Resources. Washington, D.C.: National Research Council of the National Academies.

Kennedy, D., Braga, A., & Piehl, A. (2002). *Reducing gun violence: The Boston gun project's operation ceasefire.* Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Mamalian, C., LaVigne, N., & Groff, E. (2001). *Use of computerized crime mapping by law enforcement in the United States, 1997-1998* (Computer file). Conducted by U.S. Department of Justice, National Institute of Justice, Crime Mapping Research Center. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Masser, I., Campbell, H., & Craglia, M. (Eds.). (1996). *GIS diffusion: The adoption and use of geographical information systems in local government in Europe*. GISDATA 3. London: Taylor & Francis.

McEwen, T., & Wartell, J. (2001). *Privacy in the information age: A guide for sharing crime maps and spatial data* (Report No. NCJ 188730). Washington, DC: National Institute of Justice.

Merrick, M. (2003, July 20-22). *Proceedings of the 2003 URISA public participation GIS conference at Portland State University*. Chicago: URISA.

National Institute of Justice. (2000). *Project COMPASS overview*. Last retrieved August 25, 2003, from *www.usdoj.gov/nij/programs/compass/about.htm*

Nyerges, T., & Jankowski, P. (1997). Enhanced adoptive structuration theory: A theory of GIS-supported collaborative decision making. *Geographical Systems, 4*(3), 225-257.

Pendleton, M. (2000). *COMPASS - Seattle, Washington, process research quarterly report #1* (Draft report). Washington, DC: National Institute of Justice.

Peng, Z., & Tsou, M. (2003). *Internet GIS: Distributed geographic information services for the Internet and wireless services*. New York: Wiley.

Pickles, J. (Ed.). (1995). *Ground truth: The social implications of geographic information systems.* New York: The Guilford Press.

Ramasubramanian, L. (1995). Building communities: GIS and participatory decision making. *Journal of Urban Technology, 3*(1), 67-79.

Ramasubramanian, L. (1999, August 21-24). *Where's the power in empowerment? Case studies of information technology adoption and use by community-based organizations.* Retrieved from Proceedings of the 1999 Urban and Regional Information Systems Association (URISA) conference CD-ROM.

Silverman, E. (1999). *NYPD battles crime: Innovative strategies in policing.* Boston: Northeastern University Press.

Solomon, A. (1997). *Strategic approaches to community safety initiative* (Internal document). U.S. Department of Justice. Washington, DC: National Institute of Justice.

Stoe, D., Watkins, C., Kerr, J., Rost, L., & Craig, T. (2003). *Using geographic information systems to map crime victim services: A guide for state victims of crime act administrators and victim service providers* (Report No. NCJ 191877). Office for Victims of Crime. Washington, DC: National Institute of Justice. Last retrieved August 25, 2003, from *http://www.ojp.usdoj.gov/ovc/publications/infores/geoinfosys2003/*

Swope, C. (2001, April). Restless for results – Baltimore mayor Martin O'Malley is tracking performance on a scale never seen before in local government. He wants change, in a hurry. *Governing*.

Weisburd, D., & Lum, C. (2001, December). *Translating research into practice: Reflections on the diffusion of crime mapping innovation* (Keynote address). Dallas, TX. Fifth Annual International Crime Mapping Research Conference.

Chapter II

Interjurisdictional Law Enforcement Data Sharing Issues: Benefits of the Use of Geo-Spatial Technologies and Barriers to More Widespread Cooperation

Mark R. Leipnik, Sam Houston State University, USA

Donald P. Albert, Sam Houston State University, USA

Abstract

This chapter discusses the use of geographic information systems (GIS) to create and disseminate geospatial data among multiple law enforcement agencies in the same metropolitan area, county, region, state and nation. Cooperation between different agencies of government, such as between a municipal police department and a comprehensive-planning, information technology or public works department, with GIS expertise will be discussed. The benefits derived from sharing human and technical resources, from *using a common set of geospatial data and a common crime records database schema, and from the centralization of activities, such as geocoding, will be emphasized. Issues impeding interjurisdictional use of GIS, such as technical issues of interoperability, confidentiality concerns and cost-sharing problems, are presented. Multiple examples drawn from the United States and several other countries illustrate the universality of interjurisdictional issues and the value of using GIS to facilitate data sharing and cooperation among multiple law enforcement and government agencies.*

Introduction

The majority of GIS users in the field of law enforcement are municipal police departments in urbanized areas. In the past, for clearly delimited metropolitan areas with a suburban fringe and rural hinterland, the need to share data with other law enforcement jurisdictions or cooperate closely with other governmental agencies had generally been rather limited. However, the growth of urban areas and the greater mobility of citizens, including criminals, are complicating that ideal situation. Simultaneously, GIS and related geospatial technologies, in particular digital aerial photography and global positioning systems, are being implemented in an ever-expanding range of agencies that have responsibilities that include multiple jurisdictions or at least impinge on areas of responsibility of multiple law enforcement agencies (Leipnik & Albert, 2003). Even in the case of a police department with well-defined boundaries, there often arise situations where data must be shared between and among various agencies of government and potentially other law enforcement agencies that may have limited jurisdiction within the municipality. Therefore, the need for sharing geospatial data between various law enforcement jurisdictions and among various government agencies has become more pressing (LaVigne & Wartell, 1998, 2000). There are a number of situations where it is beneficial or even essential for law enforcement agencies to share data, including data with a significant spatial component, either across jurisdictional boundaries or with other branches of government (Burka, Mudd, Nulph & Wilson, 1999). The situations where interjurisdictional data sharing related to law enforcement is most desirable are when there are crimes that span jurisdictional boundaries, when there are multiple cities in close proximity, when a regional law enforcement entity such as a metropolitan police or county sheriff wishes to use GIS most effectively, and when a state or national law enforcement agency wishes to develop a GIS (Wilkinson & Ritchie-Matsumoto, 1997). Intergovernmental sharing of GIS data is desirable or necessary in many situations, such as when a police department lacks the internal expertise to develop and maintain a GIS, but another component of local or regional government has that expertise. Also included are situations where some aspect of the application of GIS by law enforcement impinges on the responsibilities of another agency of government. Cooperation can be a two-way street and of mutual benefit to both departments and the city government, as well as its citizens as a whole (Burka et al., 1999; Markovic, 2002).

Interjurisictional Issues

Crimes Spanning Multiple Jurisdictions

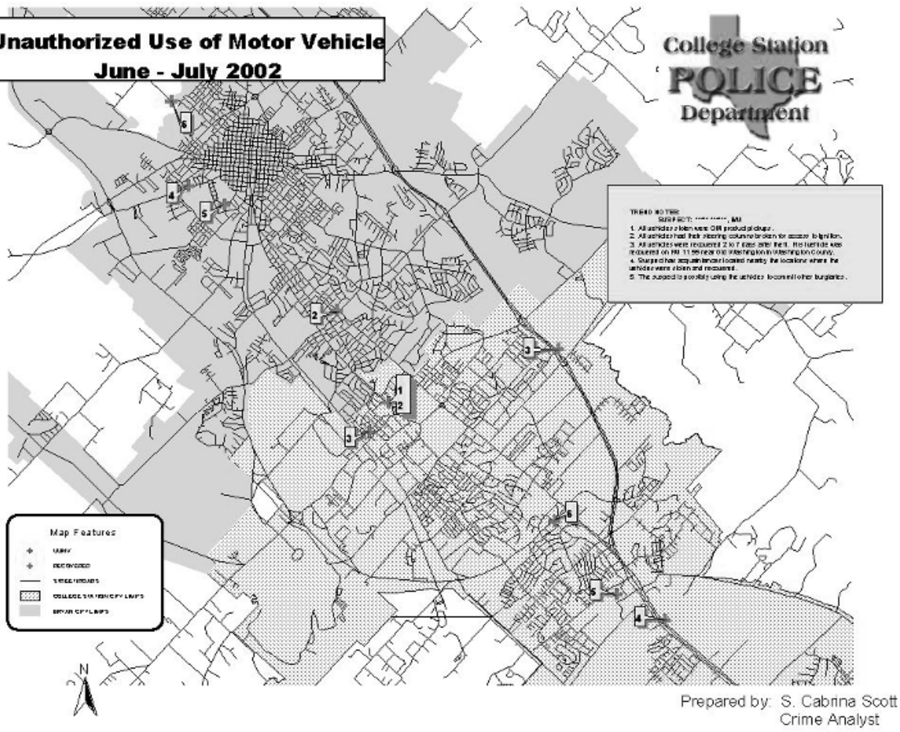
It has been observed that just as criminals are frequently oblivious of legal restrictions, they are likely to be equally contemptuous of jurisdictional boundaries and niceties (Harries, 1999). This is a valid contention with respect to casual or drug-crazed criminals of the lowest order; it is not, however, valid in many instances. For example, organized gangs of criminals engaged in robbery or confidence schemes frequently make a point of moving from one jurisdiction to another. The same can be said of the well-organized rings of prostitutes that circulate call girls around the United States or hereditary quasi-criminal groups that are highly mobile. International money launderers, drug traffickers and terrorists also take advantage of jurisdictional impediments to effect law enforcement in many cases, particularly where there are weak national and regional governments and porous borders.

A situation where a crime spans jurisdictional boundaries is illustrated by Figure 1. In this case, a series of stolen car incidents in College Station, Texas, all appeared to share the same *modus oparandi* and hence were attributed to the same suspect. This suspicion was hardened by the fact that the recovery locations for five of the six stolen cars were in Bryan, Texas (one car was recovered in a rural area of an adjacent county). Since recovery location can yield important clues as to the residence of the car theft suspect and information on businesses engaged in selling stolen parts, knowing the spatial characteristics of the neighboring city is an important asset to investigators.

Complex Boundary Issues

When a single, autonomous city exists with a suburban fringe and a surrounding rural area with the urban area under the jurisdiction of a municipal police department and some portions of the suburban area and all of the rural area being

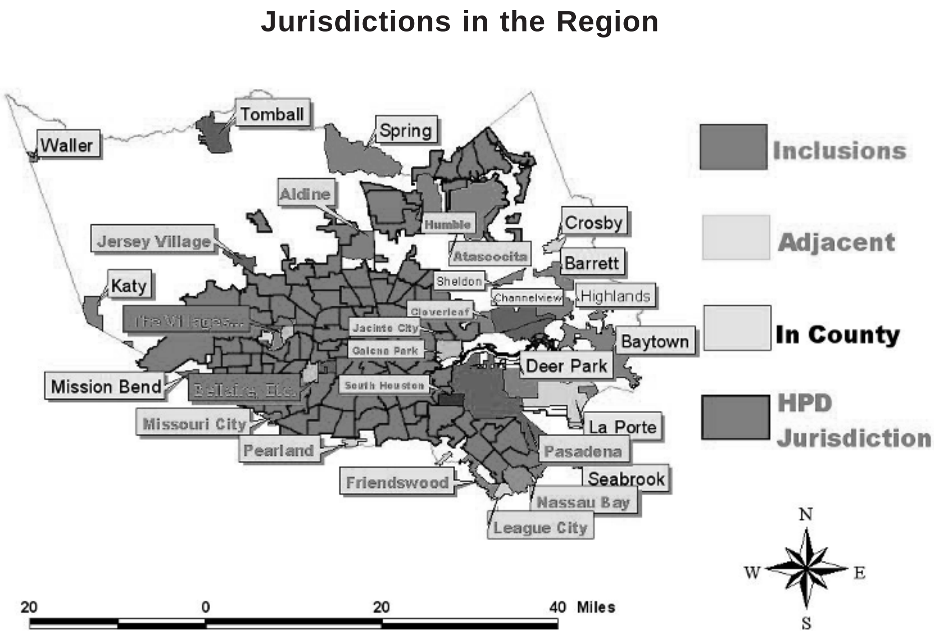
Figure 1. Seen are stolen and recovered car locations spanning Bryan/ College Station, Texas, boundary (courtesy of College Station Police Department)



administered by a county sheriff, interjurisdictional issues related to law enforcement are minimal. However, many urban regions in the United States do not fit this neat pattern. Cases in point include the metropolitan areas of Washington, D.C.; Kansas City, Missouri and Kansas City, Kansas; Cincinnati, Ohio, and the Kentucky border; greater Los Angeles; and Houston, Texas.

Houston, Texas, is an example of another sprawling metropolitan area with significant issues related to sharing geospatial data. Houston is a city that has outgrown its surrounding county. The incorporated area of Houston covers portions of three counties: Harris County, Fort Bend County and Montgomery County. Houston borders Brazoria County and has sprawled to within less than a mile of Galveston County. Houston is adjacent to 28 cities that have their own police forces, and no fewer than seven *enclave* cities that are entirely surrounded by miles of incorporated city of Houston territory. Several of these cities have reputations as "speed traps," but this is a minor issue. Sharing data, cooperation and even knowing where the boundaries actually are located is a

Figure 2. Features are some of the more than 110 jurisdictions in Harris County that can make arrests. State Police and Federal Law enforcement agencies are not depicted (beat boundaries courtesy of Houston Police Department (HPD) and Police Research Institute at Sam Houston State University).



bigger problem. Unfortunately, there is no regional crime mapping solution in Houston and only two of 110 law enforcement jurisdictions that operate in the area are using GIS. The need to better coordinate agencies to respond to homeland security threats is empowering a regional solution that is being developed by the Houston-Galveston Area Council of Governments (COG).

Examples of Interjurisdictional Cooperation

Regional Law Enforcement Sharing

Counties that have a significant degree of urbanization or a large, spatial extent have found that using GIS in a sheriff's office has many benefits. Sometimes the

GIS development effort sweeps up incorporated municipalities that are located in the county, and a truly countywide crime mapping program is initiated. In the case of Pinellas County, Florida, a specialized criminal justice information authority was established within the county government to share and manage geospatial data (Burns, Leipnik, Preston & Evans, 2003).

A regional consortium functions in the inland empire region of southern California. Development of this regional consortium has a great deal to do with the fact that the Environmental Systems Research Institute is headquartered in this area in the city of Redlands, California. This consortium has seven members from six municipal police departments, including San Bernardino, Redlands and the county sheriff's office. Coordination is accomplished through an outside special purpose entity largely funded through federal grants, but drawing from resources and expertise in each city. This effort has been associated with the development of the Regional Crime Analysis GIS in the Washington D.C./Baltimore metropolitan area. Because no previously existing entity is coordinating the effort, jealousies among member agencies seem to have been minimized. This initiative also is looking at crime in the larger context of society and thus is trying to mesh crime-fighting efforts with programs such as youth sports, community redevelopment and economic diversification (Poulsen, 2003). The city of Chicago has a long and proud history of using GIS to map and analyze the occurrence of

Figure 3. Map of all 13 jurisdictions in Pinellas County, Florida, that are sharing data

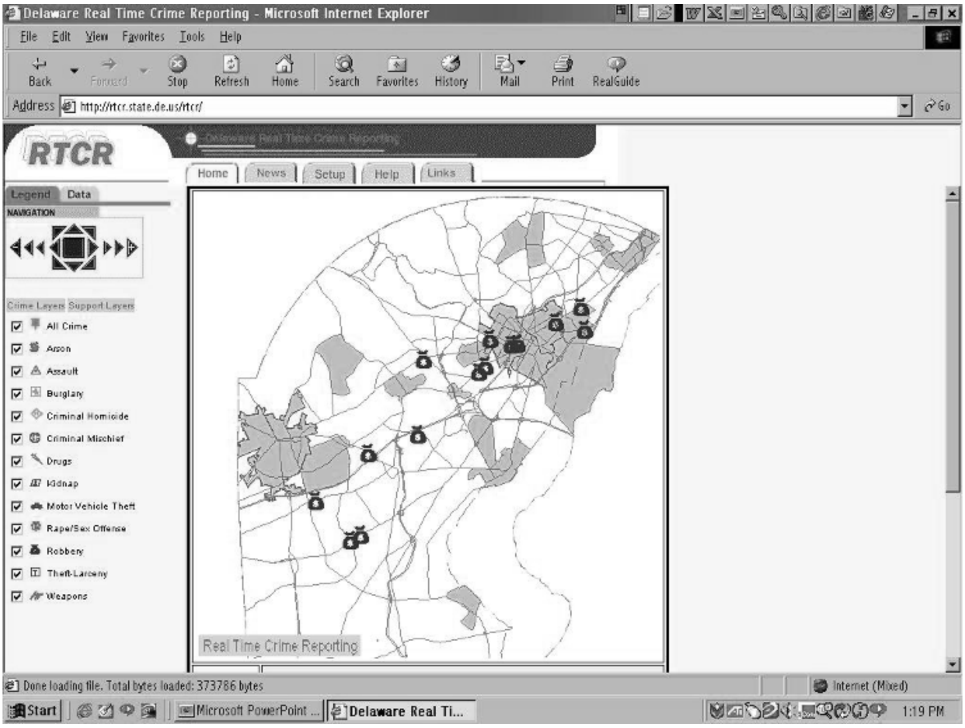
crime. A good deal of this work has been in conjunction with Dr. Richard Block at Loyola University and with the Illinois Criminal Justice Information Authority, which is located in Chicago (Block, 1995). As the incorporated area of the city has grown, it has begun to extend to the boundaries of other jurisdictions, such as Skokie, Cicero, Oak Park and so forth. The need to share data on serial crime, recovery locations for stolen cars and other interjurisdictional issues has led Chicago to take the lead in trying to develop crime-mapping capabilities in Cook County as a whole.

Statewide Data-Sharing Among Law Enforcement Agencies

Just as county governments have lagged behind municipal governments in the use of geospatial technologies in law enforcement, state governments have proven even less likely to use GIS than county government entities in the context of statewide crime mapping or other law enforcement-related purposes. Exceptions include the state of Delaware, and to a lesser extent, Illinois. Illinois should be a candidate for widespread use of GIS, since use of GIS is deeply rooted in its largest city, Chicago, and since the Illinois Criminal Justice Information Authority has been in the forefront of computerizing crime records and has developed tools for spatial analysis of crime patterns. Nevertheless, the early start taken by the Illinois State Police in crime mapping and analysis has subsided due largely to recent budget cuts. Examples of interjurisdictional sharing of data between the Illinois State Police and other agencies included, in past years, support for tracing guns involved in drug- and gang-related shootings and support for the city of Springfield, Illinois, in crime mapping and analysis.

Delaware has the best example of statewide crime mapping and analysis – it uses a Web-based application developed using Intergraph software to provide for real-time crime reporting and context-sensitive mapping of crime (Leipnik, 2003). The reason is Delaware's small size and the fact that except for one county, all law enforcement outside the 39 incorporated cities is provided by the state police. Since the mission of the state police extends beyond traditional duties to include local law enforcement, mapping crime incidents on a statewide basis seemed a logical extension of the state police's activities. Since the system is centralized and Web-based, it is a relatively easy matter for the municipalities and the single county sheriff to submit incident reports in a common format and then be able to view context-sensitive maps of boundaries, street centerlines and selected crime incidents by type and temporal criteria.

Figure 4. Northern Delaware jurisdictional boundaries, highways and robbery locations (the state police have primary jurisdiction in the areas outside the filled polygons)



National-Level Crime Mapping and Analysis

Examples of nationwide crime mapping are limited because of the sheer size of most nations, but include mapping of calls for service and crime-incident locations by the National Police of Iceland. In Iceland, the capital city holds the majority of the population for this sparsely populated country, so extending crime mapping beyond the capital involves a relatively minor increase in effort. Iceland uses a GIS integrated into its computer-aided dispatching system and utilizes customized versions of GIS software from ESRI to map all calls for service to the national police.

In the Republic of China on Taiwan, the National Police Agency in conjunction with the National Fire Prevention Bureau is mapping all structure fires to attempt to detect and prevent arson. Arson is a major problem in the densely populated

Figure 5. Dispatching center of the National Police of Iceland, with a GIS/ emergency coordination analyst at a typical workstation

country of Taiwan. Digital aerial photography is a major component of this effort, which also includes a spatial decision support system with an expert system designed to identify factors related to arson.

The National Assembly of Wales is undertaking an effort termed Project Dragon to establish a crime-end public safety GIS that will allow local constabularies to submit crime incidents, as well as calls for service to fire and other emergencies, to a centralized geocoding function. Then over a secure Internet connection or via wireless Internet from laptop computers and PDAs in the field, crime-incident locations and related attribute data, along with digitized pictures of parolees, can be displayed. Wales is a semi-autonomous area, so this can be viewed as a nationwide approach. Interestingly, despite being half a world apart, arson is a major problem in both Wales and Taiwan (Pan, 2003). To find examples of nationwide crime mapping in the United States, one must look to selected areas of law enforcement responsibility. In the United States, most of the national-level crime-mapping activities are related to counter-drug efforts. GIS supports state and local drug enforcement in every state, a project largely coordinated through the National Guard Bureau and the Office of National Drug Control Policy (Asbell, 2003). GIS is a key component of these efforts, along with remote sensing and aircraft-mounted, forward-looking infrared sensors. The

U.S. Border Patrol is an active user of GIS both to interdict illegal immigrants, as well as interfere with drug smuggling. Since the border patrol is now a part of the Department of Homeland Security, it is likely that nationwide counter-terrorism efforts will begin to incorporate greater use of geospatial technologies.

Examples of Inter-Governmental Cooperation

Municipal Government

Law enforcement jurisdictional boundaries are one barrier that can be crossed with a GIS data set that spans a county, region, state or nation. Another common barrier that can be bridged lies between law enforcement agencies and other components of government. Cooperation can be between an agency, department, office or function of a city government and the municipal police agency covering that city. Thus, a city public works, comprehensive planning department, information resources department or GIS office might cooperate with a police department. This cooperation could take the form of the government component mapping some or all crime incidents for the city police department and providing periodic maps to support the police. This is the case in Ontario, California, where the city GIS planning department performs that function on behalf of the police. In Huntsville, Texas, the city public works and information systems departments, along with nearby Sam Houston State University, supports crime and auto accident mapping for the police.

A good example of the benefits of municipal police and/or other law enforcement agencies collaborating with existing GIS functions within government is offered by the evolution of GIS in Las Vegas, Nevada. ESRI-based GIS solutions have been adopted by numerous public and private entities in southern Nevada. Early on, the benefits of having public works, utilities, the appraisal district and public safety agencies all on the same page in terms of development of base maps, common data base standards and other aspects of GIS development were recognized. Hence, a county GIS Management Office (GISMO) was established to coordinate GIS development for Clark County, Nevada. All municipalities, along with public utilities and many branches of government, participated in planning meetings and in network-based data sharing. Specifically, in the early 1990s a sworn police officer was assigned to work on-site at GISMO to gain technical experience. Soon, this officer returned to work at police headquarters. The Las Vegas Metropolitan Police continue to cooperate closely with GISMO and share data with member organizations. They are also using the CRIMEVIEW community software development by the Omega Group to share crime location data over the Internet to the community as a whole.

County Government

County entities can support city, metropolitan and county law enforcement agencies with GIS. Oakland County, Michigan, presents an exemplarily case of a county government developing a multiple purpose GIS that has a significant law enforcement component. Oakland County is located in the northern fringe of the Detroit metropolitan area. This fast-growing and prosperous county has many of the problems and opportunities associated with suburban sprawl. In particular, it faces the issue of spill over of crime from Detroit into previously relatively crime-free areas. Like many suburban fringe areas, such as Montgomery County, Maryland, it has numerous, small municipal police agencies and many unincorporated areas that are patrolled by the county sheriff. In Oakland County, the county GIS management function is located within a county information technology department. This department has developed up-to-date, countywide base maps generated from purpose-flown digital aerial photography. Street address ranges are also up-to-date, facilitating geocoding of crime incidents. There is a real value to having crime mapping and database maintenance performed by a single, highly competent entity, as in Oakland County.

Appraisal districts can also be an important source of data used in crime mapping. Data used in mapping incidents crossing the Bryan/College Station city boundaries in Texas was provided by the Brazos County Appraisal District. In Pinellas County, Florida, it was possible to analyze buffer zones around registered sex offenders' residences by using parcel mapping data. In some places, such as Orange County, Florida, the appraisal district has developed such an effective GIS capability that their street, parcel and other base-map data, as well as their aerial photography and their expertise are driving use and development of GIS in all parts of county government, including law enforcement.

State Government

Automobile accidents are a major preoccupation of many law enforcement agencies. In some regions, responding to automobile accidents and traffic enforcement is the primary function of both city and county law enforcement agencies. While enforcement of speed limits and other traffic ordinances is primarily the concern of law enforcement agencies, traffic safety is an issue that cuts across the responsibilities of law enforcement agencies and transportation agencies. In many states, the department of transportation is creating and maintaining statewide traffic accident databases. GIS has been used by several state transportation departments to create statewide traffic accident GIS data that is then available to municipal and county law enforcement in planning and enforcement efforts. Examples of states particularly active in using GIS in traffic accident analysis include Iowa, New Mexico and Hawaii. In each of these states, the state's department of transportation has been greatly assisted by transportation safety researchers at the state universities. In the case of New Mexico, the Internet has been used to disseminate the results of GIS-based mapping of automobile accident locations to the general public and to county and city law enforcement.

Police Support for GIS in other Areas of Government

Local police departments may create crime maps or other GIS products that can be used by other components of government. For example, the Las Vegas Metropolitan Police mapped automobile accidents and street robberies, and then used database selection and query techniques to determine which accidents and robberies occurred during the hours of darkness. Then, working with the Clark County GISMO and with data provided by Nevada Power on the location and illumination intensity of street lighting, they determined locations where deficient street lighting may have been a contributory factor in crime or accidents. This information fed back into the city and county public works department's efforts and led to installation of additional street lights. Data on automobile accidents in Huntsville, Texas, was used by the planning and public works departments to obtain money and expedited action from the Texas Department of Transportation on accident prone intersections. Information on liquor stores that had multiple alcohol-related crimes in close proximity to them was used by the Charlotte-Mecklenburg Police to successfully convince the North Carolina State Liquor Control Board to revoke liquor licenses. In Lincoln, Nebraska, and Redlands, California, juvenile crime hot spots have been identified and the police have cooperated with community development agencies and neighborhood groups like the Boys and Girls Clubs of America.

Benefits of Cooperation

Developing a Comprehensive, Regional Base-Map

Many law enforcement agencies in the United States rely on street centerline data actually developed by the U.S. Census Bureau. This data set termed TIGER (Topologically Integrated Geographic Encoding and Referencing) has many advantages, notably that it is available at little or no cost and it provides a consistent, nationwide base map (Block, 1995). Unfortunately, TIGER has many problems – most notably with the spatial accuracy of road locations, the inaccuracy of address ranges and names for existing roads, and the frequent absence of newer roads, particularly in the generic and vintage TIGER data that is available at the lowest cost. One benefit from police departments working cooperatively with other jurisdictions and branches of government is that the problem of enhancing and maintaining the GIS base map can be spread across other entities and institutions, many of which have a greater stake in having an accurate, up-to-date base map. Also, these other entities may be able to bring in additional geospatial data sets, such as digital aerial photography, to enhance the GIS. Thus, for example, a county GIS management office serving all aspects of county government may use dedicated aerial surveys to generate street layers where streets are defined by curbs and multiple lanes and where building footprints, hydrography, elevation contours, land use/land cover, zoning and many other features and themes (layers) are mapped. Of these multiple layers of information, law enforcement is likely to benefit most from streets and building footprints. The value of specific buildings being in the GIS is highlighted if buildings such as law enforcement agencies, schools and features in the community like bars, liquor stores, convenience stores and so forth are capable of being identified. With this in mind, digital aerial photography may prove useful for exhibits in criminal cases (Messina & May, 2003).

Saving Money

Cooperation among multiple jurisdictions can help reduce the costs in time, money and resources related to such laborious, ongoing activities as geocoding. There are relatively few examples of specific cost savings that can be tied to interjurisdictional cooperation. Frequently, the costs are apparent, but the savings are harder to quantify. A published example of the costs of development of a comprehensive, interjurisdictional crime-reporting and mapping capability is provided by the Delaware Real Time Crime Reporting System (Leipnik, 2003). This system was developed in 1997 and had specific system-associated costs through fiscal year 2001 of $732,287. Costs could be broken down as follows: software – $245,837; hardware – $96,145; programming – $266,150; training – $74,125; and warranty/service – $65,130. The costs for training and warranty are worth considering. These figures make clear that for a system to continue to function, these costs must be included (Paulsen, 2001). In addition, not apparent in this cost summary were the salaries of users within the organization. While technology may take staff time to utilize, those salary costs are offset by the time savings derived by doing previous tasks more efficiently using the new technology.

Leveraging Technical Expertise

One of the most important components of a GIS is the human user. Human resources have been identified as one of the five essential parts of a GIS (Korte, 1994). Police organizations are generally less likely to have staff or sworn officers with information technology (IT) expertise than many other branches of government. This is particularly true of smaller departments. What is true for IT in general goes double for a specialized technology such as GIS. There is an ongoing shortage of qualified GIS technicians, analysts and managers. Thus, a law enforcement agency is likely to have to really stretch to either develop in-house expertise or recruit a GIS specialist. What tends to happen in many organizations is that someone responsible for crime records management and/or information resources gets GIS added to an already long list of duties. Another common response is to take someone with interests in crime prevention or analysis and train them in GIS. This underscores the need for and value of training. It also shows that particularly for small organizations, lack of trained personnel is likely to be one of the greatest stumbling blocks to using GIS.

SolvingCrimes

The ultimate test of an effective law enforcement GIS is the extent to which it can help solve crimes. In general, the benefits of development of a local or regional GIS for crime mapping and analysis are stated in terms of more efficient deployment of resources or a better understanding of the characteristics of crime occurrence and crime patterns in the community. However, there are instances where use of GIS in a locality or a region has helped to solve a specific crime or crime series. There are also examples where sharing data between various components of government have resulted in apprehension of specific criminals. A case in point is the cooperation between the Knoxville, Tennessee, Police Department and the Tennessee Board of Pardons and Paroles (Hubbs, 2003).

Figure 6. Map of Knoxville with locations of a series of sexual assaults, streets and residences of registered sex offenders, and a home of a convicted rapist and his mug shot highlighted

In Tennessee, information on the residential locations of all parolees, including registered sex offenders, is maintained by the board. In Knoxville, this data has been incorporated into the police department's GIS. When a series of rapes and attempted rapes occurred along a greenway, crime analysts were able to quickly and definitively identify a suspect (later convicted) who lived in very close proximity (see Figure 6).

Fostering Trust

One of the barriers to development of interjurisdictional approaches to GIS in the law enforcement context is a frequent lack of trust and a pervasive lack of communication and mutual understanding among law enforcement agencies and other agencies of government. In order to build a comprehensive, interjurisdictional

GIS, these barriers must be lowered. A common base map must be developed and maintained, common or at least compatible standards for crime records management must be adopted, data dispersal and updating arrangements must be implemented, and personnel may need to be cross trained, reassigned, delegated or even exchanged between organizations. These interactions can have a positive, if unanticipated, benefit. They can help to build mutual trust and create a climate for cooperation in other aspects of law enforcement.

For example, a city redevelopment agency might share address data to help a police department map crime in an area of housing projects. This interaction could help the police to better understand the magnitude of the crime problem there and then work with the housing authority to specifically coordinate security inside the projects with police patrol and deployment strategies. Likewise, being able to see that cars stolen in one jurisdiction are repeatedly being recovered in another jurisdiction could help both jurisdictions to cooperate with each other. It might also facilitate cooperation with the county sheriff who patrols the intervening area between the two cities. This cooperation could, in turn, help to minimize the overall level of car theft, rather than simply attacking the problem in the one jurisdiction where the cars are stolen, which at best would cause displacement of the criminal activity.

Barriers to Cooperation

Technical Issues

The most serious technical issue precluding cooperation between law enforcement agencies in sharing geospatial data relates not to the geographic element of GIS, but to the tabular attribute data that must be linked to features portrayed in the geographic portion of the GIS. More specifically, there is a vast array of crime records management systems and database schema conventions in use by law enforcement agencies. Many of these systems are legacy systems running on mainframe computers and containing proprietary database structures created in obsolete programming languages like COBOL. Other systems are elaborate crime records management systems that involved agencies have sunk major resources into, but which were not envisioned as being used in conjunction with a GIS. Therefore, costly and difficult procedures are required to extract data into a GIS-compatible format. Conversely, many smaller agencies have simple, flat-file database structures. Unfortunately, failure to adopt uniform reporting codes, or a convention for abbreviation of street suffixes or even use of common street names for arteries running through multiple communities, frequently frustrates efforts to share data. In some instances, the crime records management system functions as a virtual black hole, with days required to extract the results of simple queries. This means that in a region with numerous crime records management systems, coming up with a common approach for a regional law enforcement GIS can be a major challenge.

Lack of Money

There are significant costs of maintaining a GIS system, creating or obtaining an accurate and up-to date base map with key cultural features present, and the ongoing burden of geocoding crime incident locations and/or automobile accident locations linked to incident reports stored in a database. What agency or agencies should bear these costs if the mapping extends outside a single jurisdiction? Cost-sharing agreements can be complex, and often times ad-hoc arrangements, such as those of the Illinois State Police supporting the city of Springfield, break down when budgets get tight. No one can argue that GIS is as essential a component of law enforcement as patrol cars, overtime pay for patrol officers or testing of DNA evidence, but many police departments are cutting these types of expenditures. How much easier is it to cut crime-mapping work, especially if it is being done as a courtesy to another jurisdiction? In the event of budget strictures, there is a potential for the perceived “free riders” to receive reduced services or even be entirely cut off from support.

Lack of Trust

Trust among law enforcement agencies is often pretty solid, but there can also be long-standing rivalries. Some police agencies have the reputation of under-reporting crime. A police force that is directed from above to minimize the apparent level of crime is unlikely to freely share data, and hence is very unlikely to cooperate in interjurisdictional crime-mapping efforts. Also, when crime data is released to another party, there is the fear of misuse and litigation. This latter concern is particularly acute in litigious areas. Confidentiality concerns limit all release of data, including sharing among various jurisdictions. A commonly expressed concern among senior law enforcement decision makers is the fear that sharing data between different jurisdictions will bring to light the vast differences in crime rates that exist between nearby areas. Public recognition of these disparities will, it is feared, generate adverse publicity and political pressures. These fears may or may not be well grounded. Placement of maps of crime-incident locations on the Internet in Las Vegas does not appear to have generated adverse publicity; but Las Vegas is far from a typical community. The bottom line is that in the absence of other issues, sharing data on crime incidents whether as summary statistics or as part of a geospatial database is unlikely to cause political issues, but it may exacerbate preexisting tensions, particularly where crime-related data has been deliberately suppressed in a community.

Lack of Knowledge

Finally, there is the barrier of ignorance. Many law enforcement agencies are unaware of the potential benefits that can be derived from development of a GIS that can portray crime both in their community and the surrounding region. Many small law enforcement agencies feel they are fully familiar with their communities and do not need a map to understand their responsibilities. How wrong that contention can be is illustrated by several cases brought to light by recent GIS-based mapping of jurisdictional boundaries by the Montgomery County Emergency Communications District in Texas. In the course of mapping the boundaries of incorporated communities in Montgomery County, Texas, this agency discovered one neighborhood in the town of Cut-N-Shoot (yes, that is its official name!) that was being taxed by the city government for law enforcement services, but which the police chief thought was outside the city limits. Hence, calls to 911 from this area resulted in the dispatch of sheriff's deputies from farther away, but the county did not receive taxes to reimburse them for the resulting expenses. In another case which came to light in 2003 in the same Texas county, an area of approximately 100 homes in Splendora, Texas, was found to have been receiving municipal police and fire protection services, but the residents had never paid a cent in taxes to the city. There is no doubt that more accurate mapping facilitated by GIS of many smaller towns such as these would bring to light similar problems.

Conclusion

GIS can play a vital role in helping to break down jurisdictional barriers. The precedent for cooperation between a wide range of entities sharing a common GIS base map and cooperating synergistically to maximize limited resources has been set. One can only hope that the many barriers to cooperation can be lowered. Multiple benefits can be realized by developing a GIS in a cooperative manner. We can anticipate that interjurisdictional and intergovernmental examples of GIS development will proliferate in the future. This will be to the ultimate benefit of the public, which law enforcement agencies and other branches of government are seeking to serve and protect.

Acknowledgment

Material presented in this chapter has been gathered as a result of research supported by faculty research enhancement grants provided by Sam Houston State University, Huntsville, Texas.

References

Asbell, B. (2003). National Guard Bureau – Counter-drug GIS programs: Supporting counter-drug law enforcement. In M. Leipnik & D. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 211-227). London: Taylor & Francis.

Block, R. (1995). Geocoding of crime incidents using the 1990 TIGER file: The Chicago example. In C. Block, M. Dabdoub & S. Fregly (Eds.), *Crime analysis through computer mapping* (pp. 189-193). Washington, DC: Police Executive Research Forum.

Burka, J., Mudd, A., Nulph, D., & Wilson, R. (1999, MONTH DAY). *Breaking down jurisdictional barriers: A technical approach to regional crime analysis*. Proceedings of 1999 Environmental Systems Research Institute International User Conference.

Burns, T., Leipnik, M., Preston, K., & Evans, T. (2003). The enforcer GIS: Helping Pinellas County, Florida, manage and share geospatial data. In M. Leipnik & D. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 184-196). London: Taylor & Francis.

Harries, K. (1999). *Mapping crime principles and practices*. Washington, DC: U.S. Department of Justice.

Hubbs, R. (2003). Mapping crime and community problems in Knoxville, Tennessee. In M. Leipnik & D. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 127-145). London: Taylor & Francis.

Korte, G. (1994). *The GIS Book*. Santa Fe, NM: Onword Press.

LaVigne, N., & Wartell, J. (Eds.). (1998, 2000). *Crime mapping case studies: Successes in the field, Vol. 1 & Vol. 2*. Washington, DC: Police Executive Research Forum.

Leipnik, M. (2003). The Delaware real-time crime reporting system: A state-wide and enterprise-wide GIS application for law enforcement. In M.

Leipnik & D. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 197-210). London: Taylor & Francis.

Leipnik, M., & Albert, D. (2003). *GIS in law enforcement: Implementation issues and case studies*. London: Taylor & Francis.

Markovic, J. (2002). *Planning and building a cross-jurisdictional crime mapping system for New York state.* Proceedings of the 2002 Crime Mapping Research Conference.

Messina, J., & May, J. (2003). Aerial photography and remote sensing for solving crimes. In M. Leipnik & D. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 232-240). London: Taylor & Francis.

Pan, P. (2003). *Project Dragon: Nationwide crime mapping in Wales*. Proceedings of the 1999 Environmental Systems Research Institute International User Conference.

Paulsen, D. (2001). *Mapping in Podunk: Issues and problems in the implementation and use of GIS in small and rural law enforcement agencies*. Proceedings of the 2001 Crime Mapping Research Conference. Retrieved from *http://www.ojp.usdoj.gov/nij/maps/Conferences/01conf/Papers.html*

Poulsen, E. (2003). Designing a database for law enforcement agencies. In M. Leipnik & D. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 69-80). London: Taylor & Francis.

Wilkinson, R., & Ritchie-Matsumoto, P. (1997). Collaboration and applications: Exchange of information between facilities can improve safety, cut costs. *Corrections Today*, *59*(4), 64, 66-67.

Section II

Data Issues in Crime Studies

Chapter III

Garbage In, Garbage Out: Geocoding Accuracy and Spatial Analysis of Crime

Tess McCarthy, Charles Sturt University, Australia

Jerry Ratcliffe, Temple University, USA

Abstract

Advances in computing technology and analytical techniques have given crime analysts increasingly powerful toolboxes with which to unlock the spatial patterns and processes of crime. However, the utility of such tools is still bounded by the "garbage in, garbage out," maxim, whereby analytical output is only as reliable as the analytical input. Therefore, this chapter reviews some of the sources of spatial data inaccuracy that must be considered when analyzing crime. Given the prevalence of street addresses as a spatial location identifier for crime events, particular attention is given to the accuracy and optimum parameters for geographically referencing address data. Example data drawn from burglary records in the city of Wollongong, Australia, illustrate the significance of the issues and the impact that poor address management can have on the analysis of crime.

The chapter emphasizes the practical, by outlining address correction options and summarizing recent research that identifies optimum settings for geocoding software tools.

Introduction

Modern crime fighting techniques, such as SARA (Eck & Spelman, 1987), problem-oriented policing (Mazerolle & Terrill, 1997) and CompStat (Walsh, 2001), are increasingly dependent on spatial analyses of crime to effectively and efficiently allocate crime reduction resources. The accuracy of these analyses is tied in part to the accuracy of the spatial crime data on which they are based. Not to be confused with precision, accuracy is the degree to which a measurement matches the accepted truth, while precision refers to the level of measurement. The following discussion will focus on three types of spatial data inaccuracy: conceptual, positional and attribute, using the example of burglaries recorded in the city of Wollongong, on the southeast coast of Australia, approximately 40 miles south of Sydney.

Conceptual Accuracy

Conceptual accuracy refers to the legitimacy of conceptual models used to simplify and represent complex, real-world features or events. In spatial analysis, this may relate to the choice of spatial model used to summarize and describe the geography of crime events – the choice of a point, line or area as the basic geographic or topological unit. Points have no spatial size – that is, no length or area – and have come to represent the vast majority of crime events. Lines represent objects that exist in one dimension only, objects with length. There are few examples of crime events being mapped as lines, though LeBeau used road segments to suggest areas of risk for patrolling officers based on the number of incidents occurring along a road segment over a given period of time (LeBeau, 2000). Regions, such as police beats, exist in two dimensions, possessing both length and area, and may be represented by a polygon (in a vector format) or a collection of pixels (in a raster system). Though in reality most crime events exist in three spatial dimensions, finite resources and the scale at which the data will ultimately be visualized and analyzed usually demands that the mapping unit be collapsed and the event represented by a simple point or line. Indeed, of the major crime categories recorded by the New South Wales Police

Service in Wollongong from 1998 to 2002, 93% of all offenses were recorded as a point feature. The majority of the remaining offenses did not possess a readily identifiable location and the police were forced to record position at a generalized regional level, merely noting the police beat or district where the incident occurred. In essence, this lack of geographic reference forces the police to store a crime location (a point in space) as a polygon. These were generally cases involving moving road traffic offenses, or crimes committed on public transport.

Point data used in crime analysis need not be limited to crime events. Points are increasingly being used to represent the individuals of a population at risk. For example, the population at risk for a residential burglary analysis could be represented by a map of individual residential properties mapped as a mosaic of single points. Analyses that generate the population at risk in formats such as this (that is, more precisely than at the census block group level) are more complicated to conduct but are considerably more accurate. This accuracy is gained by reducing reliance on the aggregated and less precise data of census enumeration districts or block groups and in doing so avoiding the problems associated with the Modifiable Areal Unit Problem (MAUP). The MAUP is a potential source of error that can affect spatial studies which utilize aggregated data sources (Unwin, 1996). Geographical data is often aggregated in order to present the results of a study in a more useful context, and spatial objects such as block groups or police beat boundaries are examples of the type of aggregating units used to show results of some spatial phenomena. These zones are often arbitrary in nature and different areal units can be just as meaningful in displaying the same base level data. It is this potential for variation in acceptable areal solution that is the source of the term "modifiable." While the MAUP is becoming more significant in academia, the problem is rarely discussed or considered by crime analysis practitioners. As in many fields of study (including crime), the units of analyses "are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating" (Openshaw, 1984, p. 3).

Points are also regularly used to represent environmental features. The presence of such features may influence local crime, and include street intersections, railway stations, shops, schools, licensed premises, pawnbrokers, fire stations or hotspots of other crime types. The forthcoming study by Rengert, Chakravorty and Ratcliffe of drug arrests in the vicinity of many of these types of locations in Wilmington, Delaware, has increased knowledge of the spatial relationship between illicit drug market activity and rarely studied urban features such as freeway off-ramps and fire stations. In geographic profiling (Rossmo, 2000), points are used to represent the abduction location, body dump site and other important activity or anchor nodes. Point locations can also be used to map suspect home addresses and with geographic profiling, a raster matrix may be developed to represent a probability surface of the likely offender home address or base. The combination of raster surface that depicts a probability of offender residence with an overlay showing individual residences of suspects can be a powerful tool to aid detectives in the investigation of serial crime.

Points are therefore one of the most often used and versatile geographic feature types used in crime mapping. Though some analyses may group and summarize points to a higher order feature type (line or area), they are usually the primary unit of spatial data collection.

Having selected a spatial model, the data under investigation must be geocoded, or assigned standard geographic coordinates. The next section first describes some geocoding techniques for point-based features, and then discusses various sources of positional and attribute-based inaccuracies that can creep into the analytical process through geocoding.

Geocoding Techniques

Geocoding is the "activity of defining the position of geographical objects relative to a standard reference grid" (Burrough, 1986, p. 179). Many techniques may be used to geocode point-based crime data, the most basic of which involves manually recording the x, y coordinates of associated features displayed on an existing map, with the aid of the map's reference grid. For instance, the position of a hit-and-run vehicle accident may be determined from the coordinates of the related street intersection on a road map. Tedious and time consuming, this method is only appropriate for a small number of points, though may be valuable for low-volume crimes of great concern, such as serial rape and homicide. In Wollongong, a city of approximately 115,000 people with a significant burglary problem, it is not feasible to use this method to geocode the large volume of burglary events.

A more efficient technique for obtaining the coordinates of many events using existing maps is digitizing. A digitizer is an electromagnetic device, usually composed of a wire-embedded tablet upon which a paper map is fixed, and a magnetic mouse or "puck" with which users trace and register the map's features. Alternatively, a paper map may be scanned and displayed on a computer monitor rather than fixed to a digitizing tablet, and the points "heads up" digitized in a Geographical Information System (GIS). Control points identified on the map enable the digitizer or GIS to convert the tablet coordinates or scan coordinates defining the traced features, into real world map coordinates. Though digitizing may eliminate the need to manually record the x, y coordinates of points, identification of the features of interest on the map remains a manual and often time-consuming task.

There are automated techniques, known as raster-to-vector conversions, that can identify and record the location of vector features in images. However, their thinning algorithms that fit a line to a collection of pixels in a scanned image are more applicable to the identification of lines or polygon boundaries than individual point locations.

Whether the technique used to geocode a set of points is manual, based on digitizing, or a raster-to-vector conversion, it is dependent on the availability of existing maps and sufficient descriptions of the crime's location by victims, witnesses or police officers to find it on a map. Though the majority of reported crime events occur in urban areas for which maps may be readily obtained, others may occur in unsurveyed or newly developed areas where maps do not exist or are out of date. Out-of-date street directories are a particular problem in urban fringe areas where development and change are most rapid (Ratcliffe, 2001). A particular problem of this type occurs when attempting to map the instances of theft of workmen's tools in new housing developments. The new homes are in streets that have not yet been built, let alone mapped for crime recording purposes.

When feature sets cross a number of different map sheets, the time taken to accurately georegister each sheet may also prove prohibitive if previously georegistered digital maps cannot be obtained. Data recorded on paper maps is also unable to record multiple events in the same location on the same map, while data traced from published maps may (usually outside the U.S.) infringe on the publisher's copyright unless a royalty is paid or permission obtained.

A tool that determines position and is not dependent on the availability of existing maps is the Global Positioning System (GPS). This increasingly precise and economical radio-navigation system uses a constellation of satellites as reference points from which a GPS receiver on the earth's surface can trilaterate its own position. Many police cars are now fitted with GPS receivers (Thompson, 2003), allowing the coordinates of incidents to be quickly obtained once officers have arrived at the scene. This is analogous to the GPS technology used to instantly notify police of the moving position of a stolen car, if the stolen car is fitted with a tracking device and the police car is fitted with equivalent receiving technology.

Approximately one-quarter of the offenses recorded by Wollongong police in this study used GPS or location coordinates stored in a database, to determine incident position. Offenses such as receiving stolen goods, malicious damage, stealing and minor traffic accidents, had a higher than average proportion of events with pre-recorded coordinates, possibly because they were more likely to recur at the same locations. However, if the experience of other countries is similar to that of Australia, the number of crime events that are actually attended by GPS-toting police has been reducing over the years as police concentrate their activities on crime prevention rather than crime recording.

Geocoding within a GIS

Although, as the previous paragraphs have shown, there are a variety of ways to get a crime location on to a map, the majority of offenses in most police jurisdictions are initially identified as street addresses, and subsequently converted into map coordinates using geocoding software within a GIS. Street addresses are currently the most popular identifiers of crime event location as they are often automatically collected, recorded and stored by police computer aided despatch systems and crime recording systems. Using a geocoding program, the downloaded addresses can, in most cases, be converted into geographically referenced locations using indexed street files or parcel boundaries.

Indexed street files such as the TIGER/Line files® (Topologically Integrated Geographic Encoding and Referencing database) developed by the U.S. Census Bureau (USCB, 2002) claim to contain a record for every road segment (segments usually extend between street intersections) in the U.S., with standard fields that describe the street name, street number range ("from" and "to") for both the left and right sides of the street, the street type (sometimes bundled with the street name), and any prefix, suffix or alias. The spatial database contains further information relating to the physical location of the segment, summarised as x and y coordinates for the terminal nodes and the intervening vertices. Based on the information contained in the indexed street table, geocoding software assigns coordinates to a specified property address by first locating the record in the indexed street table with a matching street name, type, prefix, suffix and street number range. The side of the street on which the property is located is then determined from the odd/even status of the street number. The coordinates of the address are then derived from its length along the segment, a proportion of the total length of the segment, with the ratio equal to the difference between the address' street number and the segment's minimum street number, and the difference between the segment's maximum and minimum street numbers. To more accurately portray property position, some geocoding software allows the user to specify an offset distance for the address point from the street centerline, and an inset distance from the terminal segment nodes.

Geocoding or basic database software can also assign coordinates to an address using parcel or block data. Many local governments have created large, cadastral databases that store land information, such as title, value and surveyed location, for the purposes of taxation. In contrast to the old paper based systems, these databases may be easily searched and the centroid coordinates extracted for a parcel record that matches a specified address.

Positional Accuracy

Positional accuracy refers to the closeness of a spatial object's location relative to its true position on the earth's surface. The spatial attributes of parcel data are often derived from very precise site surveys, therefore they are a preferable base data set to indexed street files. Further, parcel data are likely to be updated more frequently than street files, increasing the accuracy of geocoded points. Many indexed street files were created for projects such as the U.S. census whose scale did not require high levels of positional accuracy (Powell & Clifton, 1999). Reduced precision through street centerline simplification can be a particular problem in hilly areas with winding streets. Consequently, more precise, spatially corrected products based on aerial photography or GPS technology have been created by third-party vendors (Powell & Clifton, 1999).

Commercial databases may provide more updated street files than publicly available data, but at a price. The decision to go with a commercial vendor must be weighed against the likely benefits of such an expenditure. If an agency is able to attain a 98% geocoding hit rate with the existing street database, then there will be little or no cost benefit from a new address file. For cash-strapped agencies with lower hit rates, forming a liaison with the local planning department or urban authority may provide a cheaper way to access current street files as well as provide a way to access more regular updates.

Address Matching

In addition to the reduced positional accuracy of street files, the method by which geocoding software uses them to approximate parcel centroids can result in the misallocation of addresses to properties. For instance, to determine the distance between segment parcels, geocoding software must assume the width of parcels along a segment are identical, though this may not be the case. Ratcliffe (2001) compared the locations of geocoded addresses in Sydney (using a 10 meter offset distance hard coded into the geocoding software) with surveyed property boundaries, and found only 10% were located in the correct cadastral parcel. Comparisons between the geocoded point and the parcel centroid revealed an average discrepancy of 31 meters (about 100 feet). Further analysis revealed that geocoded points rarely lay close to a perpendicular line that ran through the parcel centroid and intersected the parcel's street segment. In fact, geocoded points tended to deviate from this line by angles of more than 45 degrees. Therefore the author concluded that a change to the geocoding offset distance would only marginally improve geocoding accuracy. Repetitive testing of the Sydney data determined the optimal settings for geographically referencing addresses, to be a 25 meter offset with a 15 meter inset.

The study of Sydney (Ratcliffe, 2001) assumed that the centroids of the parcel approximated those of the building footprint. In the eastern suburbs of Sydney where individual land parcels are small and dominated by the residential building, visual inspection of aerial photographs found such an assumption to be generally true. However, when land parcels are large or the building is not centrally located within the parcel boundary, the geocoded point or parcel centroid is more likely to lie outside the building footprint (though this issue is only of concern in microanalyses of crime that investigate location at a scale below that of an address).

Though positional accuracy is vital to micro-level geographic analyses, perhaps the greatest limitation to the accuracy and representativeness of geocoded address data lies in the initial documentation of the event's street address. If recorded incorrectly or in a nonstandard format, geocoding software will be unable to match the address to a record in the indexed street table or cadastral database. The various sources of data input error were summarised by Harries (1999) and Ratcliffe (2001), and the following errors may occur in both the address data to be geocoded, and the base street or parcel data upon which the geocoding process depends.

- **Misspelling or typographic errors**. Street names, types, prefixes, suffixes and aliases can be misspelled, while street numbers can be incorrectly entered. More than 6% of geocoded burglaries in Wollongong misspelled the street name, while a further 6% were given the incorrect street type. As an example, the street name "Atchison" was found to have been misspelled as "Achitson," "Aitcheson," "Aitchison," "Atcheson," "Atchinson" and "Atchitson." Of most concern is the possibility that although an address may be incorrectly entered, it may still find a match in the database. Without double-checking the validity of every address entered, erroneous matches may be surreptitiously made. Double-checking a representative sample of addresses is one way to estimate data uncertainty.
- **Abbreviations**. Street prefixes, suffixes and types can be abbreviated. For instance, "Avenue" is commonly abbreviated to "Av" or "Ave," and

"East" is commonly abbreviated to "E." Fortunately, most geocoding programs are built to recognize common abbreviations, or include an alias table the user can edit to enable recognition of less standard abbreviations. For some streets, commonly used local names differ from that recorded in the street directory. For instance, the central shopping strip in Wollongong is known to locals as "The Mall," but is recorded in the indexed street files as "Crown St Ml." In these instances, the alias field in the street file could be edited to include the local name, but this local naming variation is still a common cause of error.

- **Address duplication**. In many cities, roads are named after local figures or landmarks and may be duplicated in other suburbs. As an example, there are eight instances of Station Street in the city of Wollongong. To avoid this problem, most geocoding software allows the user to further define the address by including a suburb, postcode or other type of identifier. Of course, this further identifier may also incur error. Almost 10% of geocoded burglaries in Wollongong were initially mistakenly attributed to an adjacent suburb.
- **Extraneous information**. Nonstandard, additional information in a field can confuse some geocoding programs. However, some of this extraneous information, such as a business name or unit number preceding the address, can be stripped from the address with the use of an "address scrubber."
- **Non-address location**. Though most geocoding programs can interpret street intersections with the use of special characters (for example, "North Street & East Street"), they cannot interpret other locations, such as "cnr North and East Sts" or "50m south of North St" that deviate from the standard street address format.
- **Missing data**. If any of the address data in the street file or the input address is incomplete, the geocoding program will be unable to find a match. Fortunately, many geocoding programs prompt the user with likely alternatives if any part of the input address is missing (or unmatched). Unfortunately, the user is reliant on a complete reference layer.

To remove such errors, "address scrubbers" may be employed to check every value in a street address field. Address scrubbers can perform a variety of tasks, including:

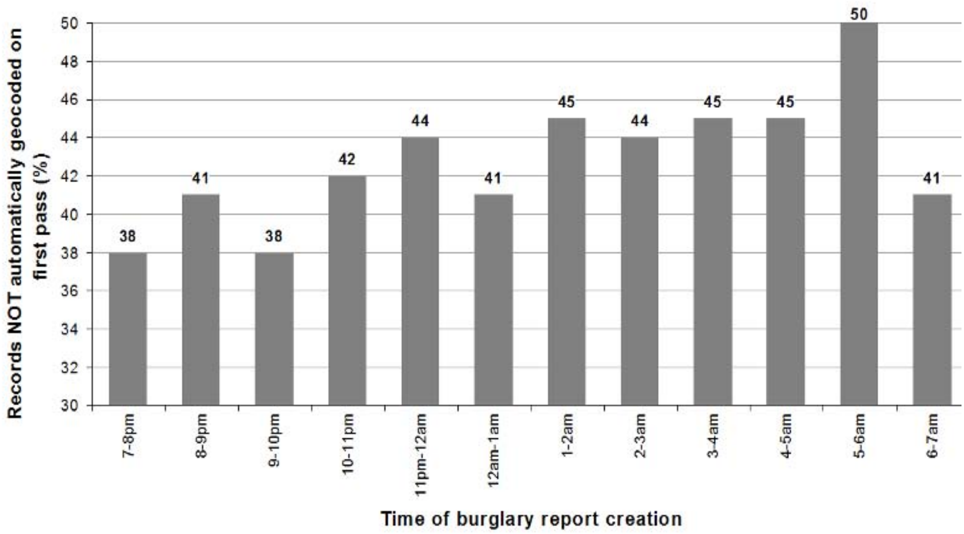
- **Remove all text before the first number**. This solves the common problem of crime events being recorded as "Outside 6 Station Road" or "In the front garden of 6 Station Road." It will also permit the geocoding of "A mile from 6 Station Road," but with the ensuing increase in inaccuracy as the geographical qualifier of the mile is lost.

- **Ignore all text prior to any slashes**. This solves the problem of apartment numbers confusing the numerical routines in the geocoding engine. Addresses such as "12/6 Station Road" that show an apartment number followed by a street address are difficult for a geocoder to interpret. Although the spatial reference of the apartment is lost, the main building is retained and geocoded.
- **Correct common errors**. One way for an address scrubber to improve the geocoding hit rate is to detect errors in spelling and automatically correct them, such as in the "Atchison Street" example above.
- **Summarize problem areas**. In urban areas that are undergoing building and development, the proliferation of new addresses that have to be geocoded can be significant. Address scrubbers can provide a log file facility to summarize the most common problem areas and names encountered, so that a more permanent solution can be found. This facility can be useful to street database managers and analysts who create and use alias tables.

Unfortunately, the errors listed above are rarely randomly distributed in space. Lee and Flewelling (n.d.) investigated the spatial distribution of inaccuracies in Erie County, New York, TIGER files. Points in the urban area were found to be more accurate than those outside the urban area that used TIGER files derived from different source data. Further, points near streams were found to be less accurate than those farther from streams and this was because roads near streams are less able to maintain a straight line due to the changeable nature of stream position.

Just as errors are not evenly distributed in space, there is some evidence that they are not evenly distributed in time. Analysis of Wollongong's recorded burglary addresses (Figure 1) revealed a small but steady and statistically significant increase[1] in the proportion of cases with incorrectly recorded addresses entered into the system, for each hour from 10 p.m. to 6 a.m., reaching a peak between 5 a.m. and 6 a.m., one hour before the end of the night shift. Police in Wollongong run a 12-hour shift roster, from 7 a.m. to 7 p.m. and 7 p.m. to 7 a.m. Address data for burglaries are entered directly onto the department's computerized database by the reporting patrol officer. Although the shift finishes officially at 7 a.m., many officers are released early for two reasons. First, there is very little crime from 6 a.m. to 7 a.m., and secondly, the day shift officers often arrive at the station early and are available for emergency calls. This suggests either a degree of urgency in event reporting from 5 a.m. to 6 a.m. as officers are in a position

Figure 1. Rate of inadequate address referencing for burglaries in Wollongong by time of report creation



to go home soon thereafter, or more simply this is a matter of fatigue due to the long night.

The rate of incorrectly recorded addresses also decreased significantly ($p < 0.001$) from 47% to 35% with the introduction of the Police Assistance Line (PAL). PAL is a telephone number that citizens can call and report crime. It does not replace the 000 emergency number (similar to the US 911 system), but is commonly used as a non-emergency crime reporting system that does not tie up police resources with simple event recording. PAL operators may have a higher level of training with regard to address protocols, or may have a greater degree of oversight during the early months of the system. Like any new feature, PAL will have undergone a fair degree of scrutiny on introduction, and it is certainly possible that more attention was paid to detail.

Attribute Accuracy

Attribute accuracy refers to the truth of information attributed to a spatial object. Given errors in the recording of crime event location, or the reference datasets upon which the geocoding technique is based, only a proportion of the point features may be successfully geocoded. This "missing data" may have a serious impact on subsequent analyses. Using a Monte Carlo simulation of a declining geocoding success rate with a statistical comparison of the complete and reduced sets aggregated to census collection districts, Ratcliffe (2004) determined a minimum geocoding success rate of 85% was required to maintain the statistical integrity of the complete crime dataset. Only 78% of Wollongong burglary records could be geocoded automatically, and the distributions of a large number of burglary *modus operandi* variables were significantly different in the geocoded dataset.

Summary

Geocoding is the first, crucial stage in most crime analysis and yet is fraught with problems of data clarity and accuracy. In this chapter, we found that in one (typical) police district in Australia, 6% of burglary locations had a wrongly spelled address, 6% had the wrong street type (such as "Street" or "Avenue") and 10% had been entered with the wrong suburb or neighborhood. This level of error is substantial, such that when an initial attempt was made to geocode the addresses (without any address scrubbing), an average of 41% could not be geocoded.

Errors and problems at this level place significant time constraints on busy crime analysts. They rarely have the time to scrub and clean address-level data, and yet micro-level studies of the geography of crime are dependent on spatially accurate and representative geocoding of address-level, point-based data. Though street addresses are currently the most popular identifiers of crime event location, they require conversion to geographically referenced coordinates before they can be analysed in a GIS or using spatial statistics. Unfortunately, geocoding programs that use indexed street files to estimate position, do just that: they estimate it. Many indexed street files were created for projects that did not require a high level of positionality. Such errors are compounded by geocoding programs that assume property parcels to be equally distributed along the street segment, and outdated street files caused by rapid urban development. To minimize positional error, addresses should be geocoded against surveyed cadastral data, or when unavailable, up-to-date street files used in conjunction with a 25-meter offset and a 15-meter inset. In addition to positional inaccuracy, incorrectly recorded addresses that confound geocoding software and reduce the hit rate to below 85% limit the representativeness of geocoded data.

Getting the Best Out of Geocoding?

From the analysis of Wollongong burglaries presented, and from other recent research presented in this chapter, the following guidelines for best practice can be formulated for the geocoding of digitally recorded address data:

- Use an address scrubber to clean data prior to geocoding;
- Get the most updated street files possible;
- In urban areas, use a 25-meter offset (about 80 feet) and a 15-meter inset (about 50 feet);
- Aim for a geocoding hit rate of at least 85% to be statistically valid;
- Be aware that as officers become either more tired or closer to the end of their shift, the quality of their data entry may deteriorate;
- Dedicated data-entry personnel (such as those that operate the PAL assistance line mentioned earlier) may increase data-entry quality.

A rigid system of data entry such as that operating in parts of the UK that rejects addresses unknown to the system protects the integrity of the recorded data while allowing addresses to be geocoded at the point of entry. However, this method does require constant maintenance as new addresses have to be cross-checked and added to the master database. When such systems are beyond the current resources of a department, address scrubbers that remove common addressing errors can be built with a minimum of effort, expense and expertise. The appeal of GPS technology should be tempered with the consideration that its proper use adds another module to the already busy training schedule of police officers. It does, however, also hold the promise of accurate geocoding, not only of crime, but of all police activity, and the possibility of comparable accuracy from urban to rural and from country to country.

References

Burrough, P. A. (1986). *Principles of geographical information systems for land resources assessment*. New York: Oxford University Press.

Eck, J. E., & Spelman, W. (1987). *Problem solving: Problem-oriented policing in Newport News.* Washington, DC: Police Executive Research Forum.

Harries, K. (1999). *Mapping crime: Principle and practice*. Washington, DC: National Institute of Justice.

LeBeau, J. L. (2000). *Demonstrating the analytical utility of GIS for police operations: A final report* (Report). National Institute of Justice (Crime Mapping Research Center). Washington, DC: National Institute of Justice.

Lee, B., & Flewelling, D. M. (n.d.). An accuracy assessment of the TIGER files: Spatial pattern of inaccuracy in Erie County, New York TIGER 2000 files. Retrieved January 1, 2004, from *http://www.acsm.net/CensusTiger41.pdf*

Mazerolle, L. G., & Terrill, W. (1997). Problem-oriented policing in public housing: Identifying the distribution of problem places. *Policing: An International Journal of Police Strategies & Management, 20*(2), 235-255.

Openshaw, S. (1984). The modifiable areal unit problem. *Concepts and Techniques in Modern Geography, 38,* 1-41.

Powell, E., & Clifton, M. (1999). *Address-based geocoding: Final report* (Report). Washington, DC: COPS Office.

Ratcliffe, J. H. (2001). On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units. *International Journal of Geographical Information Science, 15*(5), 473-485.

Ratcliffe, J. H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science, 18*(1), 61-72.

Rengert, G., Chakravorty, S., & Ratcliffe, J. H. (forthcoming). *Marketing illegal drugs in America*. Monsey, NY: Criminal Justice Press.

Rossmo, K. (2000). *Geographic profiling*. Boca Raton, FL: CRC Press.

Thompson, K. (2003, July 22). New South Wales Police Force Intelligence Analyst, Wollongong Local Area Command. Personal communication.

United States Census Bureau. (2002). *TIGER®, TIGER/Line® and TIGER-related products*. Washington, DC: U.S. Census Bureau.

Unwin, D. J. (1996). GIS, spatial analysis and spatial statistics. *Progress in Human Geography, 20*(4), 540-551.

Walsh, W. F. (2001). Compstat: An analysis of an emerging police managerial paradigm. *Policing: An International Journal of Police Strategies & Management, 24*(3), 347-362.

Endnotes

[1] The trend for the proportion of incorrectly recorded addresses had an interceptof 38.7 with the number of hours since the start of the shift parameter of 0.6364. $R^2 = 0.46$.

Chapter IV

Disaggregating the Journey to Homicide

Elizabeth Groff, Institute for Law and Justice, USA

J. Thomas McEwen, Institute for Law and Justice, USA

Abstract

This research examines the distance traveled by offenders and victims to their involvement in homicide. Key research topics include (1) the differences in distance traveled by offenders and victims by homicide motive, (2) the differences in distance traveled by offenders and victims by sex and age, and (3) the relationship between street distance and Euclidean distances by type of homicide. Findings indicate that there are clear differences in travel behavior between victims and offenders. In addition, travel distance to event location varies according to the demographic characteristics of the offender and victim. Related to the method of measurement, street distance is always longer than Euclidean distance and there is a strong and consistent linear relationship, making it possible to predict street distance from Euclidean distance. A Pareto-exponential function was determined to be a good model for representing the distances that offenders travel to their crimes. This research will assist police practitioners with respect to

investigations (for example, aid in refining suspect lists) and homicide prevention (for example, by developing richer information about activity spaces of offenders and victims).

Introduction

One aspect of homicide that has rarely been examined is the distance from the residences of the offenders and victims to the location of the offense. The individual travel patterns of victims and offenders intersect at a particular moment in time to set the stage for a homicide to occur. These travel patterns are not random, but instead reflect the individual, purposeful actions by victims and offenders, although their reasons for travel are often different. By achieving a better understanding of these travel patterns, improved strategies for investigation and prevention can be developed.

Geographical theory provides the foundation for research on journeys to crime. For example, Horton and Reynolds (1971) coined the term “action space” to describe the area with which residents are familiar and “activity space” to describe the area in which residents usually conduct their daily lives. Further research (Chapin & Brazil, 1969; Harries, 1999; Orleans, 1973) showed that the sizes and shapes of action spaces vary depending on factors, such as place of residence (suburban versus urban), sex, socioeconomic class and age. For example, the activity spaces of women and children were generally smaller and more compact than those of men and young adults. Both individuals with lower socioeconomic status and those who lived in urban areas had more compact activity spaces than those with higher socioeconomic status and residences in suburban areas.

Two important criminological theories built on this earlier work – environmental criminology (Brantingham & Brantingham, 1981) and routine activity (Cohen & Felson, 1979). As explained by Brantingham and Brantingham (1981), a crime occurs when four things are in concurrence – a law, an offender, a target and a place. The law specifies behaviors that are acceptable to society and prohibits actions that go against those behaviors. Without the existence of a law, a crime does not formally take place. Offenders and targets must also be in concurrence in order for a crime to occur. Targets can, of course, include persons, residences, automobiles and so forth. Criminologists study offenders and targets by asking questions about the motives of offenders, why offenders choose certain targets, how the target can be secured against crime, and related topics. Under environmental criminology, the fourth dimension of crime is place, defined as a discrete location in time and space at which the other three dimensions come together for a criminal event to occur. Environmental criminologists focus on place as the starting point of their studies, asking questions about the physical and social milieu in which crimes occur.

Routine activity theory (Cohen & Felson, 1979) extends the ideas of environmental criminology by examining more closely the activity spaces and daily routines of victims and offenders. Within this framework, routine activity theory identifies the elements of crime as a motivated offender, a suitable victim and the absence of a capable guardian. All three elements must be present for a crime to occur. If the routine activity spaces of potential offenders, victims or guardians change, so does the probability of a crime occurring at a particular place. In other words, it is the intersection between offender and victim activity spaces without a capable guardian that provides fertile ground for a crime to occur.

Within both these theoretical frameworks, home and incident locations provide an approach for determining how the dimension of place interacts with the other dimensions to produce criminal events. Several studies (Bullock, 1955; Capone & Nichols, 1976; Chainey, Austin & Holland, 2001; Rhodes & Conly, 1981; Wiles & Costello, 2000) analyzing distances from home have calculated the Euclidean (i.e., straight-line) distance to the crime location by offenders. For example, Wiles and Costello (2000) analyzed offenses in Sheffield, England, and showed that offenders traveled an average of 1.49 miles for actual bodily harm (ABH) crimes. They also calculated average distances for several property offenses finding a range of 1.83 miles of offender travel for non-domestic burglary to 2.51 miles for shoplifting offenses. Their analysis of the ages of offenders concluded that young offenders do not travel as far as older offenders.

Block, Galary and Brice (2002) analyzed the distances traveled by victims and offenders in Chicago during 1998 using the offenses of sexual assault, robbery and aggravated assault. They calculated the Manhattan, or right-angle, distance from home to incident for these offenses and presented their results as median metric distances. They found, for example, that the median distance for sexual assault victims from home to incident was 825 meters (.51 miles), for noncommercial robbery was 844 meters (.52 miles), and for aggravated assault was 201 meters (.12 miles). For offenders, the median distance from home to incident was 117 meters (.07 miles) for sexual assaults, 1,288 meters (.80 miles) for noncommercial robberies, and 173 meters (.11 miles) for aggravated assaults.

The research reported in this chapter differs from prior research in several ways. Our research is entirely on homicides and includes both victims and offenders. The analysis is not limited to victims and offenders who resided within the city's boundaries, but also includes those outside the city limits. Many times victims and offenders reside in surrounding jurisdictions but travel into the city where the homicide occurs. Our analysis disaggregates homicides into various types based on motive. The disaggregation is important because of the differences in the characteristics of different types of homicides (Wolfgang, 1958; Zahn & Jamieson, 1997). Gang-related homicides differ from domestic violence homicides, which in turn differ from homicides involving robbery. Motives play a crucial role in explaining the distances that victims and offenders travel. We also examine the mathematical relationship between Euclidean and street-network distances from residences to incidents. Euclidean distance is the straight-line distance between two addresses while street-network distance is the shortest street path that one could take from one address to another. Street-network distance usually is longer than Euclidean distance because of the circuitous path that must be taken along streets.

In addition to the theoretical aspects outlined above, this research has practical implications for investigations, problem solving and crime prevention efforts. As part of an investigation, knowledge of offender behavior can be used to construct more accurate suspect lists (Kotake, 2001). Distance traveled to the crime location is an important component of offender behavior. Currently, the majority of crime analysts use Euclidean distance when measuring the distance between events. Recent studies have provided a comparison of the two measures (Chainey et al., 2001; Wiles & Costello, 2000). For example, the analysis by Chainey et al. (2001) concluded that Euclidean distances were about .72 those of street network distances. Their analysis did not attempt to develop functional relationships between the two distributions and did not attempt to fit curves to either distribution.

A quantitative examination of the relationship between Euclidean and street distance will provide valuable information regarding the relative validity of one measure over the other when describing criminal behavior. In addition, problem-solving efforts depend on getting to the root of the problem, a process requiring a good understanding of the dynamics among victim, offender and place (Crime Reduction Toolkits, 2001). Finally, knowledge gained from this research can be used to target crime-prevention efforts and reduce the opportunity for homicide incidents.

Studying Homicide in Washington, D.C.

This research takes advantage of an unusually rich homicide data set from the Metropolitan Police Department of the District of Columbia for the 13 period from 1990 to 2002. One of this chapter's authors supervised the coding of information from the 4,552 homicides that occurred during these years. The source of the information was the master case jackets maintained in the Homicide Division on all homicides. The master case jacket includes the original homicide report, autopsy, investigative narratives and arrest information. The Federal Bureau of Investigation (FBI)'s coding booklet for the national Violence Criminal Apprehension Program (ViCAP) served as the basis for the coding. The police department maintains the database as a caseload management tool for investigating homicides in the city. The database includes all the necessary ViCAP information along with locally needed information, such as detectives assigned to a case and names of offenders. On a periodic basis, the police department transmits records to the FBI for inclusion in the national ViCAP system, which can be accessed by police departments across the country to identify homicide trends and serial incidents.

Of particular interest for this study, the database includes the address where the homicide occurred, the home address of the victim, and for closed cases, home addresses of the offenders. These addresses were geocoded in order to develop distance measures.[1] The addresses were batch geocoded in a two-step process. Initially, all the addresses within the District of Columbia were batch matched to the street centerline file developed by the city's department of transportation. The city's file is more accurate than the TIGER files provided by the U.S. Census Bureau. Addresses outside the District of Columbia were batch matched against the TIGER files. Addresses that could not be matched through the batch process were then interactively geocoded. An attempt was made to locate all addresses that remained unmatched by consulting both the Internet[2] and hard-copy versions of maps. Addresses that were successfully located via these sources were manually matched by dropping a point onto the correct location[3]. While this process was time consuming, it produced a data set with both a high level of accuracy and a high match rate. A total of 4,534 homicide locations (99.6%) were matched through this process. Home addresses were available in the database for 4,151 victims, of which 3,955 (95.3%) were successfully geocoded. There were 3,434 addresses available for offenders of which 3,304 (96.2%) were successfully geocoded.

Euclidean distances and street distances were calculated for this study. As mentioned earlier, Euclidean distance is the straight-line distance between two points while street distance is based on the network of streets that would be traveled from one location to another. Each measure has advantages and disadvantages. The key advantage of the Euclidean distance is that it is easy to calculate using the functionality of a GIS straight out of the box. Consequently, it has become the dominant method of distance measurement among crime analysts. However, Euclidean distance does not take into account the urban transportation network or topography of an area that might lengthen a trip. Street distance offers a more accurate measure of the actual path between two points but is more difficult to determine, especially in an automated manner. If available, it may be more beneficial for police in identifying suspects, canvassing areas, and designing prevention strategies because it provides a more realistic measure of actual distance traveled than does Euclidean. Despite the accuracy gained through analyzing the transportation network, street distance is still only an estimate of the journey to victimization because there is no way of knowing if victim or offender used the shortest path route available. Moreover, victims and offenders are not necessarily traveling from or to their homes in conjunction with the incident. Street distances do, however, serve as good proxy measures for the activity spaces of the participants (Rhodes & Conly, 1981). Routines for calculating these two distances are available from the Web site maintained by ESRI, Inc.[4]

The street networks of Maryland, the District of Columbia and Virginia were used in the study since 40% of the suspects and 17% of the victims lived outside of the District of Columbia. While the street pattern in the District of Columbia is basically a grid network in the downtown area, there are numerous circles, parks and diagonal streets that make it unique (Rhodes & Conly, 1981). In addition, outside the downtown area, the street pattern is a more suburban one. Maryland and Virginia also have suburban street patterns, except in the downtowns of some of the larger cities. Thus, there is significant reason to expect that the standard assumptions regarding the relationship between Euclidean and street distances on a grid network might not hold for homicides in D.C.

An Overview of Findings

Table 1 gives basic information on the homicide victims and offenders that served as the basis for the analysis. Males dominate in both the victim and offender categories, accounting for 3,457 victims (87.4%) and 2,799 offenders (94.4%). With regard to race, African-Americans predominated accounting for 3,679 victims (93.0%) and 2,873 offenders (96.9%). Because of the large percentage of African-Americans, it was not possible to perform any meaningful comparisons with the other races represented in the data. The average age of victims was 28.6 years with a standard deviation of 12.9 years, and the average age of offenders was 23.7 years with a standard deviation of 8.7 years.

One of the advantages of the coding scheme supporting the ViCAP database is that more than one motive can be coded, when applicable, for a homicide[5]. For example, a gang-related homicide over drug turf can be coded as both gang- and drug-related, rather than trying to decide the most important motive. The capability to code more than one motive eliminates the need to make choices about the predominant motive and provides more information on the amount of different types of homicides.

Table 1. Statistics on victims and offenders, Washington, DC, 1990-2002

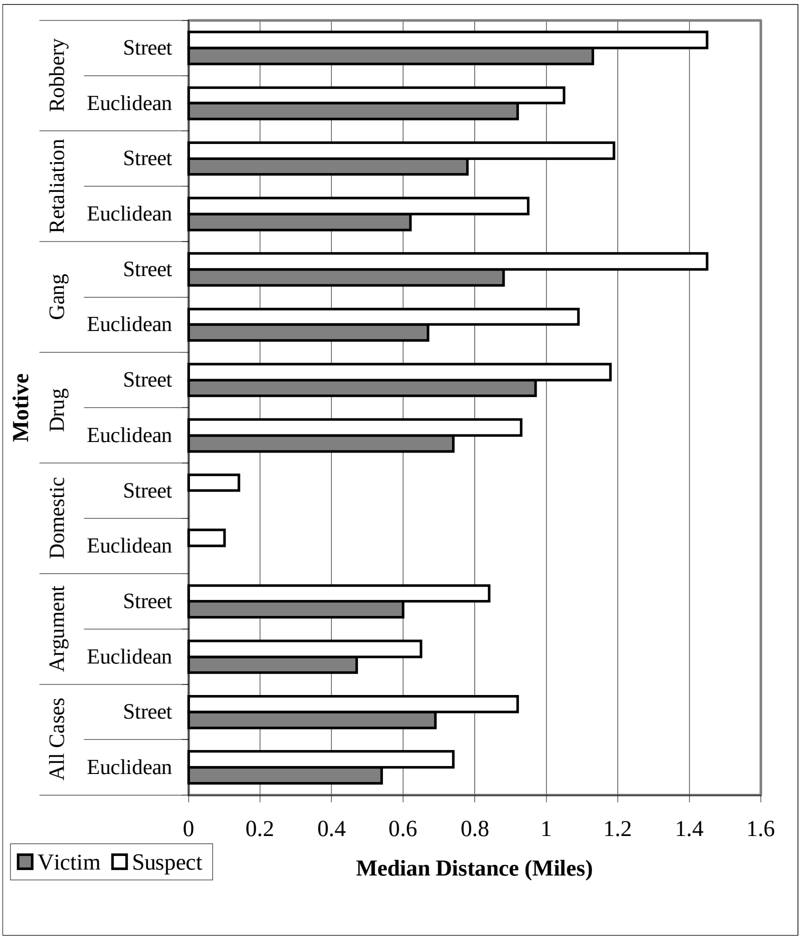
	Victims		Offenders	
Gender	Number	Percent	Number	Percent
Male	3,457	87.4	2,799	94.4
Female	498	12.6	167	5.6
	3,955	100.0	2,966	100.0
Race	Number	Percent	Number	Percent
African-American	3,679	93.0	2,873	96.9
White	138	3.5	26	0.9
Hispanic	107	2.7	61	2.1
Other	31	.8	12	0.1
	3,955	100.0	2,972	100.0
	Average Age (mean)	Standard Deviation	Average Age (mean)	Standard Deviation
Age	28.6	12.9	23.7	8.7

The results on distances show the advantages of disaggregating homicides by motive. Figure 1 gives the median Euclidean and street distances for all homicides in the geocoded database and the distances for the six most frequent subtypes of homicide (arguments, domestic violence, drugs, gangs, retaliations and robberies). Median values were selected for most of our analysis because of the skewness in the distributions. For all homicides, the median Euclidean distance for victims was .54 miles and the median street distance was .69 miles. For offenders, the distances are .74 miles and .92 miles, respectively. As shown in Figure 1, the distances from home for offenders are always longer on average than for victims.

Figure 1 also shows that travel distances vary considerably by type of homicide. Domestic violence homicides obviously have shorter distances because most victims are at home when the homicide occurs. Interestingly, the median distance from home for offenders is about .10 of a mile, reflecting the fact that some offenders had a personal relationship with their victims but were living elsewhere at the time of the incident. For robberies, victims were about a mile from home when they were killed, and the distances for retaliations, drug-related and gang-related homicides were similar at about .67 miles.

Overall, offenders had median distances of .72 Euclidean miles and .94 street miles.[6] The median distances from home to incident for offenders were always greater than for victims. Offenders involved in gang-related homicides averaged 1.09 miles from home and those involved in robbery homicides averaged 1.05

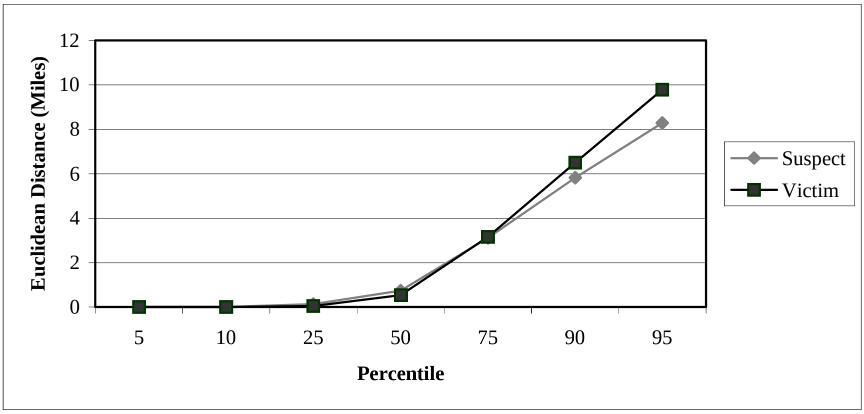
Figure 1. Victim and offender distances to homicide



miles from home. Those involved in either drug-related or retaliation homicides averaged about .94 miles. These distances are in line with the results of previous studies of criminal mobility cited earlier. Domestic crimes and arguments have the shortest journeys to crime, while robberies and gang-related incidents have the longest journeys both for victims and offenders.

As an indicator of the skewness of the data, Figure 2 shows the cumulative distribution functions for Euclidean distances from home to incident for victims and for offenders. For victims, the 25th percentile for victims was .05 miles, the median was .54 miles (as reported above), and the 75th percentile was 3.16 miles. These percentiles for offenders were .13 miles, .74 miles and 3.12 miles,

Figure 2. Cumulative distribution of distances



respectively. In other words, approximately 25% of victims were killed within one block of their own home. One quarter of suspects committed murder within two blocks of their residence. A significant number of homicides are local events. This finding supports earlier studies that found homicide events to be highly concentrated and the journey-to-homicide very short.

As expected from the literature, the preceding results suggest there are differences in the distances under study according to the underlying motives of the homicides. To determine the statistical significance of these findings, a non-parametric test on the differences between medians was conducted. Table 2 shows the results for the distances of victims from home to incident for the selected motives. The distances for homicides with arguments as a motive vary significantly from the other types of homicides. Similarly, domestic violence homicides also differ significantly from the other types. On the other hand, drug-related, gang-related and retaliation homicides have median distances that do not differ significantly from each other based on this non-parametric test.

Table 2. Chi-square values for median tests, victims

Motive	Domestic Violence	Drug-Related	Gang-Related	Retaliation	Robbery
Argument	21.5*	17.8*	6.8*	6.8*	11.6*
Domestic Violence	--	53.7*	57.6*	46.8*	50.4*
Drug-related		--	.3	3.5	0.1
Gang-related			--	0.5	1.1
Retaliation				--	3.0

* p <= .01

Table 3. Chi-square values for median tests, offenders

Motive	Domestic Violence	Drug-Related	Gang-Related	Retaliation	Robbery
Argument	22.0*	8.2*	19.9*	9.9*	15.4*
Domestic Violence	--	38.1*	47.5*	44.2*	42.3*
Drug-related		--	1.0	0.2	4.3*
Gang-related			--	1.6	2.3
Retaliation				--	6.2*

p <= .01

Table 3 gives the results from the test of medians for the distances of offenders according to the underlying motives of the homicides. As with the victims' distances, homicides with motives of arguments and domestic violence have distances that differ significantly from the other types of homicide. The distances of robbery-related homicides differ significantly from drug-related and retaliation homicides. Finally, the median distances for drug- and gang-related homicides are not significantly different from each other or from retaliation homicides.

Euclidean versus Street Distance

In order to better describe the relationship between Euclidean and street distance measures, we used a bivariate linear regression model. Table 4 shows regression results with street distance as the dependent variable and Euclidean distance as the independent variable. The relationship between the two distances is very strong. All R-squared values are .99, indicating a virtually perfect linear relationship between the two distances. What this result means is that a reliable estimate of the street miles can be obtained by applying the regression results. For all victims, the coefficient for the Euclidean miles was found to be 1.142 and the constant as .079. Effectively, this result means that increasing the Euclidean distance by 14.2% and adding .079 miles results in a good estimate for the street miles of a given incident. The remainder of the table can be interpreted in the same manner.

These results suggest that investing in software to measure street distance is unnecessary once the above analysis is complete for a particular city/crime combination. This is good news because the Euclidean or straight-line distance has many advantages for crime analysts – it is easy to calculate using GIS software; it does not require the purchase of additional software extensions; and it is easy to explain to both the community and police. In addition, the Euclidean

Table 4. Regression results for street and Euclidean distances

Victims Motive	Coefficient	Constant	R^2
All Cases	1.142	.079	.996
Arguments	1.125	.113	.997
Domestic Violence	1.130	.046	.995
Drug-related	1.147	.093	.994
Gang-related	1.134	.101	.997
Retaliation	1.179	.006	.993
Robberies	1.148	.073	.994
Offenders			
Motive	Coefficient	Constant	R^2
All Cases	1.174	.008	.997
Arguments	1.175	.015	.997
Domestic Violence	1.133	.078	.995
Drug-related	1.160	.027	.996
Gang-related	1.179	.035	.991
Retaliation	1.161	.043	.999
Robberies	1.122	.130	.997

distance is a mathematical formula that can be included in an analytical routine. As illustrated above, the regression analysis shows a linear relationship between the two distances, which means that if desired, an analyst can obtain a good estimate of the street distance through the regression equation. In combination, the measured Euclidean distance and the calculated street distance provide an envelope that represents the most likely distance traveled and thus the area in which an offender may live. The distance envelope concept needs more testing but if validated could be used to more effectively narrow suspect lists and lead to quicker apprehension of offenders. In sum, initial results indicate that analysts can take advantage of the simpler straight-line distance when measuring distance and not worry about purchasing additional software.

Given the ease of calculation and the linear relationship with street distance, the rest of the results are reported in Euclidean distance only. Street distance can be calculated for any of the Euclidean distances using the formula provided.

Patterns by Sex and Age

Of the 3,955 records with geocoded home addresses for victims, 3,457 (87.4%) were males and 498 (12.6%) were females. The preponderance of male victims typifies homicides across the country (Smith & Zahn, 1999). As seen in Table 5, the Euclidean distances from home to incident location are considerably different by gender with female victims having significantly shorter distances. For all homicides, the median distance for males was .69 miles, compared to .06 for females. Domestic violence cases are an anomaly because the majority of these homicides took place in the victim's residence with a resulting median value of zero distance. The average distances for gang-related homicides are fairly close at .68 miles for males and .54 miles for females. On the other hand, homicides with arguments as a contributing motive had distances of .56 miles for males and .07 miles for females. Many of these homicides for females were also domestic violence cases. Victims where robbery was a motive had distances of .98 miles for males and .12 miles for females.

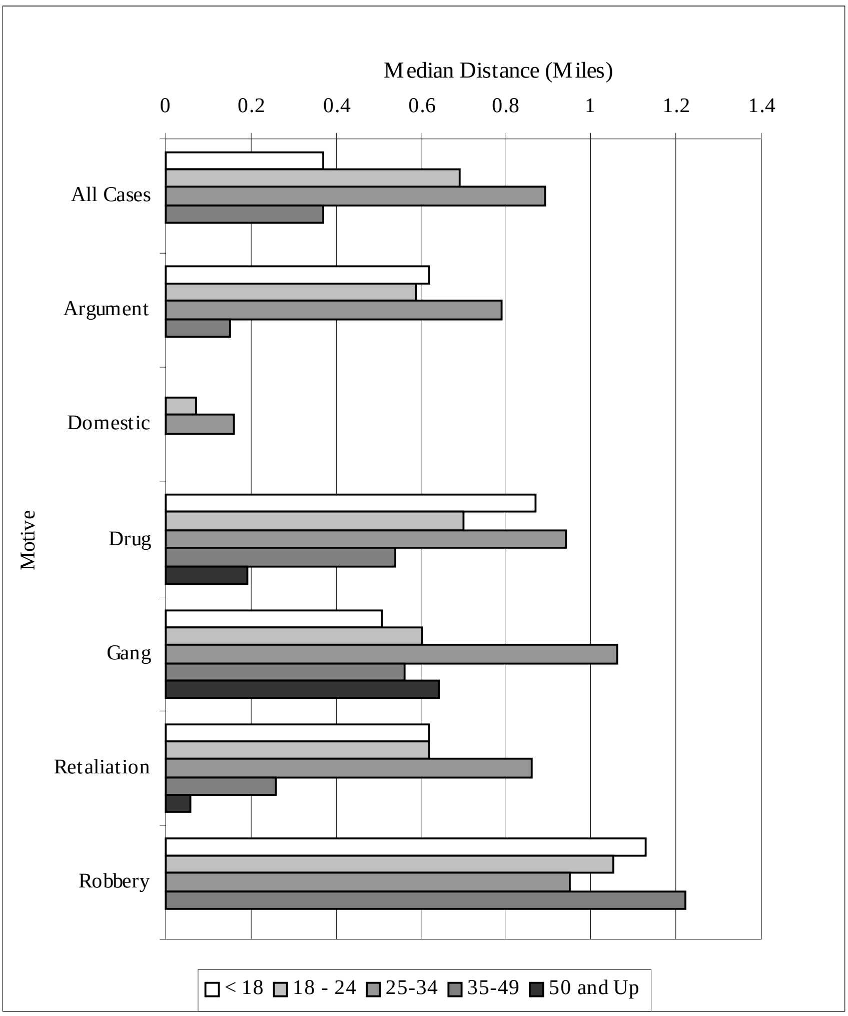
The bottom portion of Table 5 shows the median Euclidean distances for offenders from their homes to the incident location. Overall, male offenders were .82 miles from home and female offenders were only .07 miles.[7] Figure 3 shows the median Euclidean distances for victims by age groupings. The general trend is for increasing distances through the 25- to 34-year-old bracket and then decreasing after 35 years old. For example, the median distance from home to incident for victims under the age of 18 years was .37 miles, compared to .69 miles for 18-24 years old, .89 miles for 25-34 years old, .37 miles for 35-49 years old and .00 for victims 50 years and older. Homicides involving arguments, drugs, gangs and retaliation show similar patterns with variations only on the distances

Table 5. Median distances from home to incident by gender (miles)

Victim's Gender	**Male**	**Female**
All cases	.69	.06
Arguments	.56	.07
Domestic Violence	.00	.00
Drug-related	.83	.16
Gang-related	.68	.54
Retaliation	.64	.23
Robberies	.98	.12
Offender's Gender	**Male**	**Female**
All cases	.82	.07
Arguments	.72	.16
Domestic Violence	.28	.00
Drug-related	.94	.58
Gang-related	1.1	.48
Retaliation	.98	.24
Robberies	1.05	.85

involved. For example, homicides with retaliation as a motive had an average for the age brackets of .61 miles, .62 miles, .89 miles, .37 miles and .06 miles, respectively. Homicides with robbery as a motive are an exception to the above patterns. With these cases, the median values for victims under the age of 18 years was 1.13 miles; 18-24 years old, 1.05 miles; 25-34 years old, .95 miles; 35-49 years old, 1.22 miles; and 50 years or older, .00 miles (i.e., the robbery took place at the victim's residence). Interestingly, homicides involving victims age 50 and older had a median distance of zero for all cases as well as for argument, domestic and robbery motives. Not surprisingly the more constricted activity

Figure 3. Median distances from victim home to incident by age

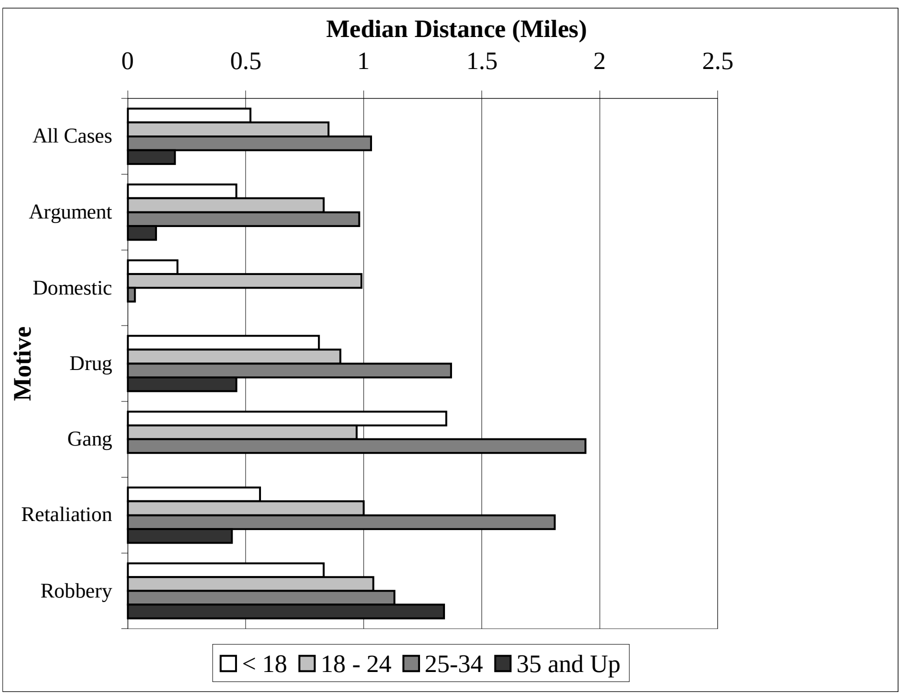


space of older victims means they are frequently victimized at or near their homes while younger individuals spread their risk out over a larger area.

As seen in Figure 4, the median distance from home to incident for offenders under the age of 18 years was .52 miles, increasing to .85 miles for ages 18-24 years, 1.03 miles for ages 25-34 years and then decreasing to .2 miles for 35 years or older. The overall pattern does not hold for gang-related homicides.

From the offender database, we have 2,519 offenders who employed a firearm during the commission of the homicide and 774 offenders with other types of weapons (cutting instruments, fists, ligatures and so forth). Analysis of distances in firearm and other weapons shows significant differences as reflected in Table 6. Overall, offenders were .87 miles from home when firearms were used for the homicides compared to .33 miles with other weapons. With arguments, the distances are .60 miles and .15 miles, respectively. Only gang-related homicides have shorter average distances when firearms are involved. With these cases, the average was .68 miles with firearms and 1.12 miles with other weapons.

Figure 4. Median distances from offender home to incident by age



Note: The gang-related homicides, ages 35 and up category is not shown because it encompassed only three observations.

Table 6. Median distances from home to incident by weapon involved

Motive	Firearm	Other Weapon
All cases	.87	.33
Arguments	.60	.15
Domestic Violence	.07	.00
Drug-related	.74	.40
Gang-related	.68	1.12
Retaliation	.65	.16
Robberies	1.06	.49

Interestingly, the largest distance for firearms was 1.06 miles for robberies. Capone and Nichols (1976) postulated that robberies with a gun had longer travel distances because they were more planned out and often involved a vehicle. It is plausible that armed robbery homicides would follow the same pattern.

Distribution of Distances for Offenders

Figures 5 through 9 show the results of modeling the distances that offenders take to their homicides with motives of arguments, gangs, drugs, retaliation and robbery[8] Developing functional representations of actual distance distributions can be beneficial in modeling offender behavior. We have already established that most homicides occur close to home with the average distances varying by motive. Moreover, it is clear that the occurrences of homicide follow what is generally called *distance decay*, which simply means that fewer homicides take place as distance increases. The analysis of our homicide data reflects distance decay with the shape of the decay function depending on the motive. By modeling the distances for each motive, comparisons can be made with the parameters of the functions to quantify the differences of the decay functions.

By way of background, Capone and Nicholas (1976) studied the distances that robbery offenders went to commit their offenses. They fitted three functions – exponential, Pareto and Pareto-exponential – to the actual distance distribution for 825 offenders. They found that all three functions provided reasonable fits to the actual data, and they favored the mixed Pareto-exponential function because it combined the favorable characteristics of both contributing functions. Our results parallel the work of Capone and Nicholas (1976) with a similar outcome, as explained below.

Analysis of the offenders' distances proceeded in the following manner. First, the distances were aggregated into tenths of a mile with a cutoff point of 10 miles.

The cutoff point was established to eliminate very long distances that would likely skew the analysis. In essence, the 10-mile limit allows for analysis of offenders who generally live in the city or in relatively close proximity to its boundaries.

Frequency distributions of the resulting aggregated distances were then reviewed for each motive, and subsequent analysis conducted using several possible distributions to determine the best fits to the data. Following the lead of Capone and Nichols (1976), we included exponential, Pareto and Pareto-exponential distribution, and then expanded the analysis to include the Beta and Weibull distributions.

The Pareto-exponential function was found to provide the best fit over the complete set of motives. In functional form, it is defined as follows:

$$y = a\, D^{-b}\, e^{-cD}$$

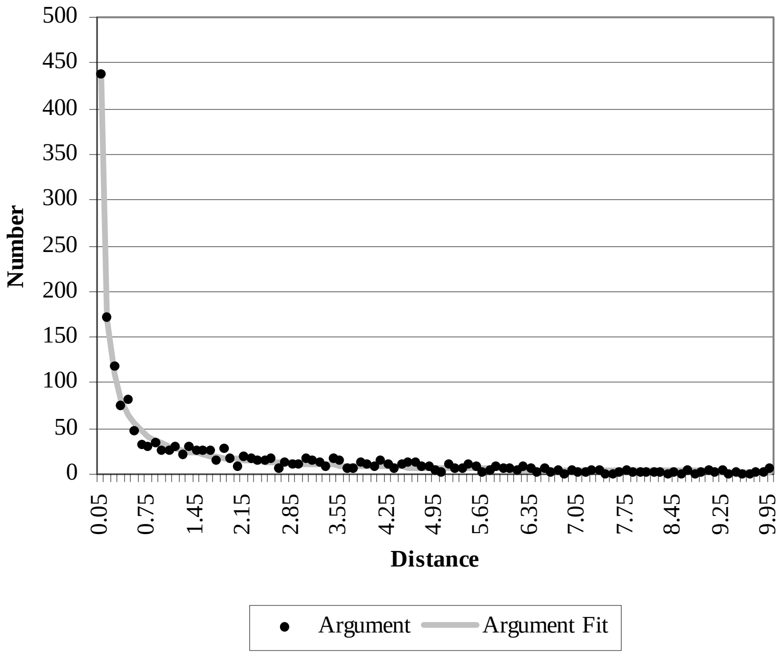
where y is the number of homicides that occur for a given distance D, and *a*, *b* and *c* are the estimated parameters for the function. The Pareto-exponential curve usually lies between the Pareto and exponential curves that are fit to the same set of data. It therefore moderates the effects of the two distributions.

Table 7 gives the results for the estimated parameters of the Pareto-exponential functions fitted to distance distributions for each motive. These estimated parameters were used to develop the curves in Figure 5 showing the frequencies of the actual distances and the fitted curve for argument homicides. The Pareto-exponential function gives a good fit for the distributions with each of the five motives (Figures 5-9).

Table 7. Parameters for pareto-exponential function figure

Motive	***A***	***b***	***c***	***R²***
Arguments	33.75	.86	.05	.992
	(1.22)	(.013)	(.050)	
Drug-related	27.91	.65	.22	.928
	(2.72)	(.036)	(.069)	
Gang-related	16.94	.47	.33	.931
	(1.37)	(.032)	(.058)	
Retaliation	23.99	.58	.27	.953
	(1.75)	(.028)	(.052)	
Robberies	19.61	.61	.19	.943
	(1.48)	(.029)	(.048)	

Figure 5. Distance fit for argument motive



Implications of Findings

The journey to crime of both the victim and offender is important for the study of homicide in the context of criminological theory. Both environmental criminology and routine activities approach the study of crime from the viewpoint of the relationship between victim, offender and place. For the purposes of these theories, information on the distances between the places contributes to the understanding of crime from these viewpoints.

From the research presented in this chapter, several important distinctions that contribute to the theoretical literature can be drawn. First, as with other crimes, the distribution of distances for victims and offenders is skewed. That is, both victims and offenders tend to be involved in homicide incidents when they are relatively close to their residences. As reported, victims had median distances of .54 miles and offenders .74 miles from their homes. For victims, the 25th percentile was .05 miles and the 75th percentile was 3.16 miles. Thus, as stated earlier, homicides are local problems. On average, both victims and offenders tend to be close to home when the homicide occurs. This quality of the spatial

Figure 6. Distance fit for drug motive

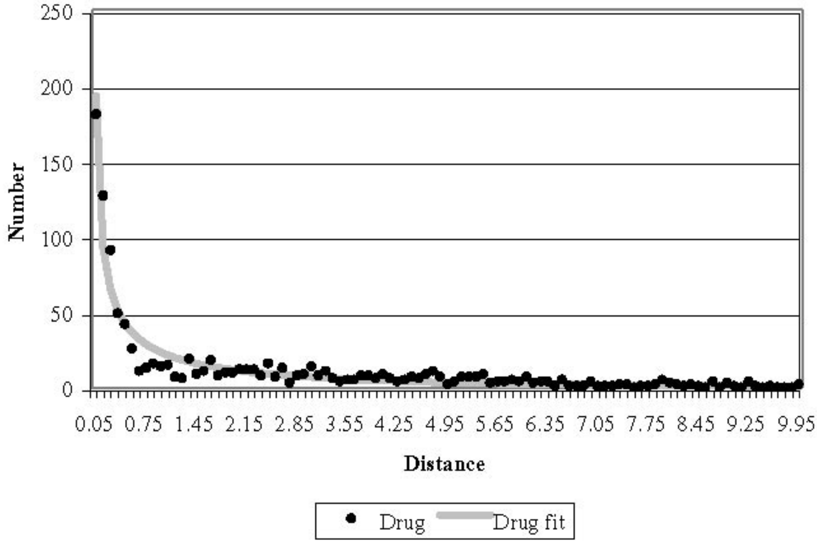
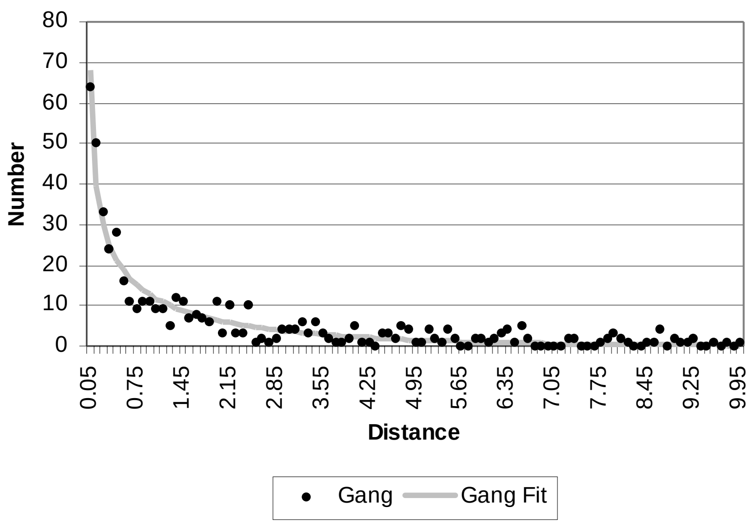


Figure 7. Distance fit for gang motive



distribution of crime stems from the daily activities of both victims and offenders noted in the introduction. People, in general, tend to frequent familiar areas. Both victims and offenders use their knowledge in their decisions about where to work, shop, recreate and in the case of offenders commit crimes. These decisions are in turn, shaped by the urban structure (Rhodes & Conly, 1981).

Figure 8. Distance fit for retaliation motive

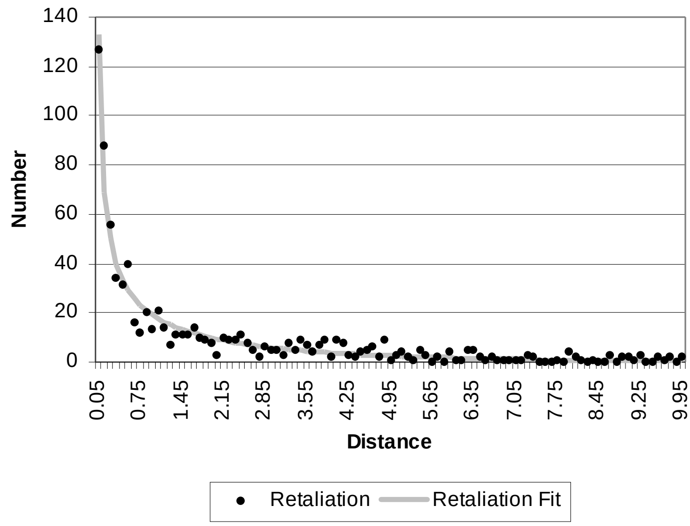
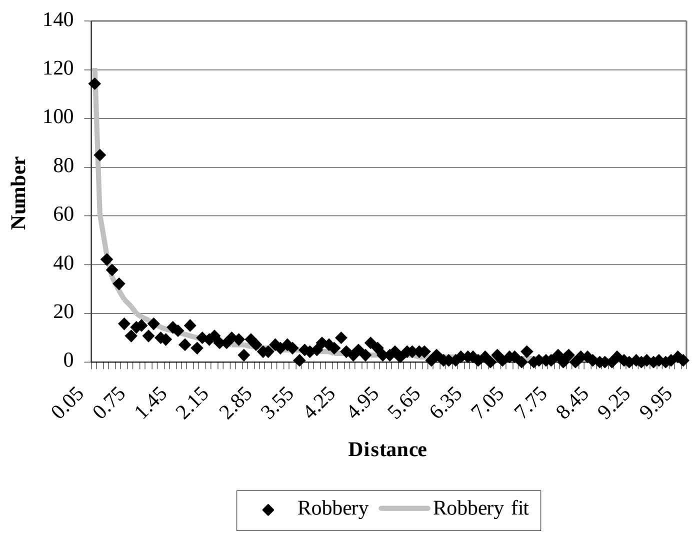


Figure 9. Distance fit for robbery motive



This finding has implications for both crime prevention and homicide investigations. Prevention efforts can be targeted to the areas surrounding hotspots of homicide activity since nearby residents are at highest risk of being victims. Investigators can use travel distances to narrow suspect lists to the most likely candidates. While these findings are informative, future research on modeling distances should be based on fitting these skewed distributions to appropriate mathematical functions, such as the Pareto-exponential function described in this article. Development of the functional representations of the distances provides a way of describing these distances beyond just providing averages and variances. The functions give a theoretical representation of the distances that can be beneficial in modeling the routine activities of offenders. Moreover, comparison of the functions for different motives provides additional insight into the crime of homicide.

The set of findings related to variation in distances by the demographic characteristics of sex and age has additional implications for understanding homicide events. As expected, females have shorter travel distances than males regardless of role in the homicide. Among offenders there is generally a positive correlation between age and average distance traveled until age 34 with domestic violence and gang homicides as the exceptions. This finding is in line with the expected distances based on the variation of the size of activity space with age (Chapin & Brazil, 1969; Harries, 1999). Younger people tend to have larger activity spaces that gradually shrink as they get older. Aggregate victim travel distances followed the same pattern but there was variation by motive. Victims of argument, drug and robbery homicides are the exceptions to the overall pattern with victims under 18 having longer travel distances than 18-24 year old individuals. In the case of drug and robbery homicides, the longer travel distances for individuals under the age of 18 may be due to the extensive public transportation system in the Washington, D.C. area. The metro makes it easy for individuals to travel relatively long distances from their homes without a vehicle.

A final comment is in order on the relationship between the age of the victim and the distance from home when they were killed. From a statistical view, the correlation between the two is small. However, closer examination between the two shows a curvilinear relationship. That is, the distances for victims tend to increase until ages 25 to 34 and then decrease.

Travel distances of both victims and offenders are significantly different based on motive. Not surprisingly, travel distances for homicides related to domestic violence and arguments are significantly different from all other homicide motives and tend to be shorter. In addition, the offender travel distances for robbery and retaliation related homicides are significantly different from all other motives considered. These differences offer important clues for investigators and can inform crime prevention efforts.

The research also demonstrates the strong and consistent linear relationship that exists between Euclidean distance and street network distance even in an area without a uniform, grid street network. Both measurement techniques have their advantages and disadvantages. Selection should depend on the particular problem at hand and the resources available to address it. Crime analysts may want to depend on the more easily measured Euclidean distance, while investigators may want to use the formulas provided to estimate the distance actually traveled by offender and victim. Our research on homicides shows there is such a close relationship between the two that knowing one virtually ensures good information about the other assuming the acceptance of regression results. In other words, crime analysts in Washington, D.C., can measure Euclidean distance using GIS and then use the formula provided to calculate probable street distance traveled. We propose that crime analysts consider using both measures to create an envelope representing the most probable distance traveled. However, this suggestion requires additional empirical tests to validate its utility.

We also note that the research was based on homicides that occurred within the boundaries of Washington, D.C. The extent to which these homicides and their associated distance distributions are representative of other cities cannot, of course, be determined. Replications of this research – especially on the distances that victims travel to homicides – is encouraged for other researchers. As it stands, we are unable to generalize our findings to a larger universe.

One of the most intriguing findings was that offenders with firearms tend to be further from their residences when they committed their crimes than offenders with other weapons. Following Capone and Nichols (1976), we note that these individuals may be committing crimes that involve specific targets, significant planning and a vehicle. We also conjecture that offenders feel more confident to venture further from their homes when they are carrying a firearm. More research could be done in this area by interviewing offenders to determine their reasons for carrying firearms and the role that firearms played in their daily routines.

In sum, this initial foray into disaggregating travel to homicide has yielded some interesting findings that contribute to achieving a better understanding of travel behavior for both theoretical criminology and police practice. Additional research is needed to improve generalizability of these findings and to enhance our knowledge of our how victims and offenders end up at the same place. Following Bullock (1955), a clear next step is to measure the proximity of residence of offender and victim. An analysis of these distances will provide important information that will aid in the investigation of neighborhood effects on homicide and ultimately contribute to the formation of more effect strategies for preventing homicide.

Acknowledgments

The authors would like to thank the Metropolitan Police Department of the District of Columbia for providing access to homicide reports, and James

Cardona of ESRI, Washington, D.C., for programming assistance. The analysis, results and conclusions reported in this chapter are the sole responsibility of the authors.

References

Block, R., Galary, A., & Brice, D. (2002). *Place space and crime revisited: Targets and offenders converge in violent index offenses in Chicago* (Unpublished paper).

Brantingham, P., & Brantingham, P. (1981). Introduction: The dimensions of crime. In P. Brantingham & P. Brantingham (Eds.), *Environmental criminology* (pp. 7-26). Prospect Heights, IL: Waveland Press.

Bullock, H. A. (1955). Urban homicide in theory and fact. *Journal of Criminal Law, Criminology and Police Science, 45*, 565-575.

Capone, D. L., & Nichols, W. W. (1976). Urban structure and criminal mobility. *American Behavioral Scientist, 20*, 199-213.

Chainey, S., Austin, B., & Holland, F. (2001). *Insights into the analysis and patterns of offender journeys to crime* (unpublished manuscript).

Chapin, F. S. J., & Brazil, R. K. (1969). Human activity systems in the metropolitan United States. *Environment and Behavior, 1*(2), 107-130.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review, 44*, 588-608.

Crime Reduction Toolkits. (2001). *Patterns of offending*. Retrieved February 12, 2002, from *http://www.crimereduction.gov.uk/toolkits/p0315.htm*

Harries, K. (1999). *Mapping crime: Principle and practice*. Washington, DC: National Institute of Justice.

Horton, F. E., & Reynolds, D. R. (1971). Action space differentials in cities. In H. McConnell & D. Yaseen (Eds.), *Perspectives in geography: Models of spatial interaction* (pp. 83-102). Dekalb, IL: Northern Illinois University Press.

Kotake, J. (2001). Psychological and criminal profiling. In *Forensic Science* magazine online, The Bronx High School of Science. Retrieved Aprilo 4, 2003, from *http://www.bxscience.edu/publications/forensicbio/articles/pyschologicalprofiles/r-pscy04.htm*

Office of Quality Assurance. (2001). *Murder analysis: A study of homicides in the District of Columbia*. Washington, DC: Metropolitan Police Department, District of Columbia.

Orleans, P. (1973). Differential cognition of urban residents: Effects of social scale on mapping. In R.M. Downs & D. Stea (Eds.), *Image & environment: Cognitive mapping and spatial behavior* (pp. 115-130). Chicago: Aldine Publishing Company.

Rhodes, W. M., & Conly, C. (1981). Crime and mobility: An empirical study. In P. Brantingham & P. Brantingham (Eds.), *Environmental Criminology* (pp. 167-188). Prospect Heights, IL: Waveland Press.

Wiles, P., & Costello, A. (2000). *The road to nowhere: The evidence for traveling criminals* (Research study 207). London: Home Office.

Wolfgang, M. E. (1958). *Patterns in criminal homicide*. Montclair, NJ: Patterson Smith.

Zahn, M. A., & Jamieson, K. M. (1997). Changing patterns of homicide and social policy. *Homicide Studies*, *1*, 190-196.

Zahn, M. A., & McCall, P. L. (1999). Homicide in the 20th-Century United States: Trends and Patterns. In D. M. Smith & M. A. Zahn (Eds.), *Studying and preventing homicide: Issues and challenges* (pp. 10-30). Thousand Oaks, CA: Sage Publications.

Endnotes

[1] All geocoding was done in ArcGIS 8.3©.

[2] The Mapquest site (*http://www.mapquest.com*) was used to identify the location of addresses during the interactive geocoding process.

[3] The PickAddress script available at /arcgis/arcexe83/ArcObjects Developer Kit/Samples/Geocoding/Pick was accessed for this purpose.

[4] See *http://support.esri.com*. The script used in this study for calcualting Euclidean distances is called pt2pt_distance.ave. The new script developed by James Cardona to measure street distance is under ‘Calculate Network and Euclidean Distance’.

[5] Definitions for the motive categories are as follows (Office of Quality Assurance 2001): **Argument:** A disagreement between two or more parties with intent to provoke a breach of the peace by annoying, disturbing, interfering, or offending others. **Domestic:** The deliberate and premeditated killing of another family member, to include: any person with whom the offender is related by blood, legal custody, marriage, having a child in common, or with whom the offender shares or has shared a mutual residence; or any person with whom the offender maintains or maintained a romantic relationship not

necessarily including a sexual relationship. **Drug-Related:** Any criminal act that directly or indirectly involves substances recognized as a controlled substance.**Gang-Related:** A group of two or more individuals involved in any type of criminal activity, typically recognized as gang membership by their neighborhood. **Retaliation:** A violent act committed against another person as the result of retaliation for a perceived wrong done to the offender. **Robbery:** The taking of anything of value fro another person or their immediate possession by force or violence, whether against resistance or by sudden or stealthy seizure or snatching, or by putting the person in fear.

6 As with victims, the distribution of distances is skewed, with the 25th percentile for Euclidean distances at .13 miles and the 75th percentile at 3.12 miles.

7 Statistics for gang-related homicides are not shown for offenders because of the small number of female offenders.

8 Domestic violence homicides were not included in this analysis because the majority of them occur within the residence of the offender and victim, or in very close proximity.[1]

Chapter V

Constructing Geographic Areas for Analysis of Homicide in Small Populations: Testing Herding-Culture-of-Honor Proposition

Fahui Wang, Northern Illinois University, USA

Van O'Brien, The Sidwell Company, USA

Abstract

The rate estimates for rare events like homicide in small populations are very susceptible to data errors, and thus compromise the validity of inferences. This chapter discusses two geographic information systems (GIS)-based methods for constructing geographic areas with sufficiently large base populations to permit reliable estimates of homicide rates to be obtained. One is the spatial order method, and the other is the ISD method (after the Information & Statistics Division of the Health Service in Scotland, where it was devised). Both methods construct new analysis areas *based on spatial proximity of basic units. For demonstration, the methods are applied to testing the herding-culture-of-honor hypothesis proposed by Nisbett and Reaves, and the result shows that the herding-culture-of-honor proposition is merely an artifact of unreliable estimate of homicide rates. The methodology, in general, is applicable to analysis of any rates with small base populations.*

Introduction

The study of homicide rates across geographic units and for demographically specific groups often entails analysis of aggregate homicide rates in small populations (Loftin and McDowall, 2000)[1]. This presents three serious analytical problems:

(1) *Sensitivity of Rate Estimates to Data Errors*. Research shows that significant *reporting errors* exist in major U.S. crime data systems including those for homicides. In addition, researchers find it necessary to eliminate some homicide cases because information about certain demographic variables may be missing. This can lead to the *error of missing data*. Homicide rates from small populations can be very sensitive to these data errors.

(2) *Fairness of Statistical Sampling*. Crime rates, including those for homicide, are usually based on geopolitical entities (for example, county) or predefined arbitrary census units (for example, census tract), which are markedly heterogeneous in population size. An area that has a very large population may thus be equated with one that has a very small population, causing "over-sampling" of sparsely populated areas.

(3) *Violating Assumptions in Ordinary Least-Squares (OLS) Regressions*. Researchers often use OLS regressions in analysis of homicide rates. The existence of larger errors of prediction for homicide rates based on small populations than for large populations violates the assumption of *homogeneity of error variance* in OLS regressions. As population decreases, one expects to see more and more cases of zero homicide. The increasingly skewed error distribution towards the lower bound of zero violates the assumption of *normal error distribution* in OLS regressions.

Several strategies have been attempted to mitigate or resolve the problem of small population-base rates by criminologists. The first is to use homicide counts instead of per capita rates. However, most studies are interested in the offense or victimization *rate* relative to population size. The second is to simply delete outliers or unreliable estimates in areas with small population in order to reduce their influence. Obviously exclusion of selected observations undermines the validity of analysis. The third is to use larger units of analysis (for example, states, metropolitan areas or large cities) or aggregate over more years. But this strategy prevents one from analyzing the variation within the large unit or within the time period. The fourth is to use more advanced methods such as Poisson-based regressions to better capture the non-normal error distribution pattern. This approach only applies to regression analysis.

Researchers in health-related fields have long sought solutions to a similar problem: many diseases like cancer are rare events, and rate estimates in small populations are unstable. They have developed several methods of constructing geographic areas, two of which will be tested here for homicide analysis. By constructing new analysis areas, homicide rates should be less sensitive to data errors because these areas have larger base populations; they should also have more comparable population size and thus generate a fairer statistical sampling; and conceivably, heteroscedasticity and skewedness of error distribution from regressions should be mitigated or eliminated. For demonstration, the methods are used to reexamine the culture-of-honor hypothesis – one of many influential explanations for the pattern of violence in the southern U.S. This technique can be applied to any research that uses data susceptible to the problem of small population-base rates. Results from this research also shed light on the "modifiable areal unit problem" (MAUP), well known to geographers and spatial analysts (Openshaw, 1984; Fotheringham & Wong, 1991).

Literature Review

This section briefly reviews literature on the following areas: (1) problems associated with homicide analysis in small populations, (2) approaches employed by criminologists to address the problems, (3) the spatial analytical methods of constructing geographic areas developed in health-related fields, and (4) the herding-culture-of-honor hypothesis. It is not intended to be an exhaustive review, and only exemplary references are cited.

Problems Associated with Homicide Analysis in Small Populations

Criminologists are confronted with the task of measuring homicide in areas with small population sizes or among small populations in various contexts. Several problems are associated with the task.

The first problem is that measures of homicide rates are sensitive to data errors in small populations. The two main sources of homicide data in the U.S. are the Uniform Crime Reporting (UCR) and the National Vital Statistics System (NVSS). Using line-of-duty homicides by law enforcement officers as an example, Dobrin and Wiersema (2000) found that the UCR system tends to report more homicides than the NVSS, but in some states the pattern is reversed. This confirms that reporting errors exist in both systems. In addition, a sizeable and growing number of unsolved homicides have to be excluded from studies analyzing offender characteristics (Fox, 2000). For example, in their studies of homicide rates across 318 counties in 14 southern U.S. states, Chu, Rivera and Loftin (2000) had to delete 173 homicide offenders from their analysis because the race or gender of the offender was unknown. Yet, in the case of Kennedy County, Texas, if only one of these offenders were a white, non-Hispanic male, the rate would change from zero to 312 homicides per 100,000 residents.

The second problem concerns a fair representation of statistical samples. Homicide rates are usually based on geopolitical entities (for example, county) or arbitrary census units (for example, census tract), which are markedly heterogeneous. Brantingham and Brantingham (1984) questioned the usage of census tracts as the unit of analysis because of the lack of uniform population sizes. Typical urban census tracts have population ranging from fewer than 100 to several thousands. Even among the 264 counties that are all rural in the study by Osgood and Chambers (2000), population varies from 560 to 98,000. As a result, statistical analysis may be over-influenced by areas with small population. For instance, a region may consist of five areas: one area has 200,000 residents and 50 homicides (that is, a homicide rate of 25 per 100,000 residents), and each of the other four areas has 50,000 residents and zero homicide. Either the mean (five per 100,000) or median (zero) of homicide rates in the five areas significantly underestimates the region's homicide rate (12.5 per 100,000). Ideally, analysis areas should be of approximately equal population size (Black, Sharp & Urquhart, 1996).

The third problem, related to regression analysis, was pointed out by Osgood (2000) who noted that aggregate crime rates from small populations cause two problems for ordinary least-squares (OLS) regressions. One violates the assumption of homogeneity of error variance because the precision of the estimated crime rate depends on population size (that is, larger errors of prediction for crime rates based on small populations than for rates based on large populations). The other violates the assumption of normal error distribution because the distribution becomes increasingly skewed as crime rates approach the lower bound of zero (that is, more crime rates of zero are observed as populations decrease).

Various Approaches by Criminologists

Criminologists have used a number of approaches to address the problems associated with homicide analysis in small populations.

The first has been to use homicide counts instead of per capita rates. In the study of homicide and neighborhood transition in Chicago, Morenoff and Sampson (1997) argued that using per capita rates would create many outliers among census tracts with very small population sizes, and therefore they chose to use homicide counts. This strategy was perhaps less problematic for their study since their primary interest was in locational features. In other words, it might be the frequency of homicides (not necessarily the rate) in an area that created fear and thus led to neighborhood transition. Most studies are interested in the offense or victimization rate relative to population size, for which per capita rates instead of counts should be used.

The second approach has been to simply delete samples of small populations. Harrell and Gouvis (1994) used various crime rates at the census tract level to study crime patterns and community decay in Washington, D.C., and Cleveland, Ohio. In processing the Cleveland data, they encountered 10 tracts (out of 203 tracts) with fewer than 100 residents, and simply deleted these tracts because of their concerns about unreliable crime rates in these tracts. Morenoff and Sampson (1997) also deleted tracts with less than 100 residents from their Chicago study. In doing so, however, they risked overlooking data that could have been valuable or even critical to their studies.

The third approach has been to aggregate over more years or increase the geographic level of aggregation. In a study of homicide patterns around the St. Louis region, Messner, Anselin, Baller, Hawkins, Deane and Tolnay (1999) aggregated homicides for 17 years in order to achieve a greater degree of stability in the county-level homicide rates. Such an approach assumes that homicide and its causal factors have been stable over time. This assumption is questionable and the approach prevents analysis of variations over time. In the 21 studies of homicide rates across social space surveyed by Land et al. (1990), most used large units of analysis such as states, Standard Metropolitan Statistical Areas (SMSAs) and large cities. There may be more variation of homicide rates *within* unit than *between* them. The resolution of analysis needs to be sharpened (Harries, 1997).

The fourth approach uses Poisson-based regressions to remedy the two problems associated with OLS regressions. Land et al. (1996) compared various Poisson-based regressions for crime analysis. However, it was not until recently that criminologists specifically linked the models to the problem of small-population-base rates (Osgood, 2000; Osgood & Chambers, 2000). Regressions based on a Poisson distribution recognize only a small range of counts having a meaningful probability of occurrence when the mean crime count is low, which is more likely to be so for a small population, and thus may lead to a skewed distribution of errors (Osgood, 2000). The basic Poisson regression model has traditionally been used to analyze counts of events, and needs to be modified for analysis of rates of events. Errors in the modified model are inversely related to population sizes. This important property acknowledges the greater precision of rates based on larger populations, and thus addresses the problem of heterogeneity of error variance (Osgood, 2000).

For clarification and comparison, all approaches are summarized in Table 1. Approaches 1-3 appear to have critical weaknesses, and the fourth approach seems to offer the most appropriate solution. However, not all questions can be answered by regressions, and many studies depend on reliable estimates of crime rates at a fine geographic resolution.

Table 1. Approaches to the problem of small population-base rates

	Approach	**Examples**	**Comments**
1	Use homicide counts instead of per capita rates	Morenoff and Sampson (1997)	Not applicable for most studies that are interested in the offense or victimization rate relative to population size
2	Delete samples of small populations	Harrell and Gouvis (1994); Morenoff and Sampson (1997)	Deleted observations may contain valuable information
3	Aggregate over more years or to a high geographic level	Messner et al. (1999); most studies surveyed by Land et al. (1990)	Impossible to analyze variations within the time period or within the large areal unit
4	Poisson-based regressions	Osgood (2000); Osgood & Chambers (2000)	Effective remedy for OLS regressions; not applicable to non-regression studies
5	Construct geographic areas with large enough populations	Haining et al. (1994); Black et al. (1996); Sampson et al. (1997)	Generate reliable rates for statistical reports, mapping, regression analysis and others

Methods of Constructing Geographic Areas from Small Areas

In their study of homicides in Chicago, Sampson, Raudenbush and Earls (1997) combined 865 census tracts to create 343 "neighborhood clusters" (NC). Each NC was composed of geographically contiguous and socially similar census tracts. The construction of NCs was based on major geographic boundaries (for example, railroads, freeways and parks), knowledge of local neighborhoods and cluster analysis of census tracts, and thus was a time-consuming, manual process. Although the purpose of defining NCs was stated to approximate local neighborhoods, it implied that NCs were large enough (averaging around 8,000 people) to generate reliable homicide rates.

Researchers in health-related fields face a similar problem when they attempt to measure the rates of diseases such as cancer and AIDS in areas of small population size. Geography-trained researchers have developed several spatial analytical methods by constructing geographic areas that have larger populations. The problem of constructing geographic areas has much common ground with the long tradition of regional classification (regionalization) in geography (Cliff, Haggett, Ord, Bassett & Davis, 1975). Some methods, such as the *spatial-order method* used by Lam and Liu (1996) and the *ISD method* (after the Information & Statistics Division of the Health Service in Scotland where it was devised; see Black et al., 1996) emphasize spatial proximity. Lam and Liu (1996) faced with the challenge that some rural counties had insufficient HIV-cases to sample from, and used the spatial-order method to generate a national rural sampling frame for HIV/AIDS research. They used the method to create clusters of rural counties so that each cluster had approximately 50 new AIDS cases. The spatial-order method uses space-filling curves to determine the nearness or spatial order of areal units. Given a capacity constraint, the areal units are grouped consecutively according to their spatial order values. Black et al. (1996) attempted to reduce the complex spatial distributions of disease at the small-area level to a form more amenable to analysis. They used the ISD method to group the census enumeration districts (ED) into larger analysis units of approximately equal population size, and analyzed the spatial distribution of disease rates across the redefined analysis areas. The ISD method constructs an analysis area by beginning with a tract and adding tracts nearest to the starting point to form until the desired threshold within the area is reached. Neither of these methods, however, considers within-area homogeneity of attribute.

Another group of methods place the first priority on attribute similarity within areas. A good example is the work by Haining, Wises and Blake (1994), who attempted to consolidate 1,159 enumeration districts (ED) in the Sheffield Health Authority Metropolitan District in the U.K. to a manageable number of regions for health service delivery (referred to hereafter as the "Sheffield method"). At the first stage, they mapped the EDs using four classes based on the Townsend deprivation index scores (Townsend, Phillimore & Beattie, 1988), and dissolved boundaries between adjacent EDs in the same class to form the initial regions. At the second stage, they used several subjective rules and utilized some local knowledge to further reduce the number of new regions to 48. The Sheffield method started by complying with *within-area homogeneity*, and then made adjustment by achieving *spatial compactness*. The method attempted to balance these two competing goals.

This chapter illustrates the spatial order method and the ISD method, which emphasize spatial proximity in constructing new areas. Both are simple and easy to implement in a GIS environment. In the application of testing the Nisbett-Reaves herding-culture-of-honor proposition, counties are classified into three categories. The task is to merge counties within the same category (that is, homogenous), and thus focuses on spatial proximity only. Incorporating both spatial proximity and attribute similarity remains the major challenge for constructing geographic areas and will be studied elsewhere (Wang, 2004). The spatial order or the ISD method can be built into other more complex regionalization models.

The Herding-Culture-of-Honor Theory

Nisbett's culture-of-honor theory (Reaves, 1992; Nisbett, 1993; Nisbett and Cohen, 1996) is one of many influential explanations for the pattern of violence in the southern U.S. The theory can be divided into two major propositions. The *culture-of-honor-violence proposition* attributes the high levels of assaultive violence in the south to the so-called "culture of honor" that supports assaultive behavior in defense of one's reputation, family and other sacred values. According to the *herding-culture-of-honor proposition*, areas where topography or precipitation limited agriculture to marginal farming and herding had a chronic threat of livestock theft, and should have higher homicide rates than those areas where farming was the major form of agriculture. The pattern is most likely to persist in white, non-Hispanic populations that have been most exposed to the culture. This chapter is particularly interested in testing the *herding-culture-of-honor* proposition.

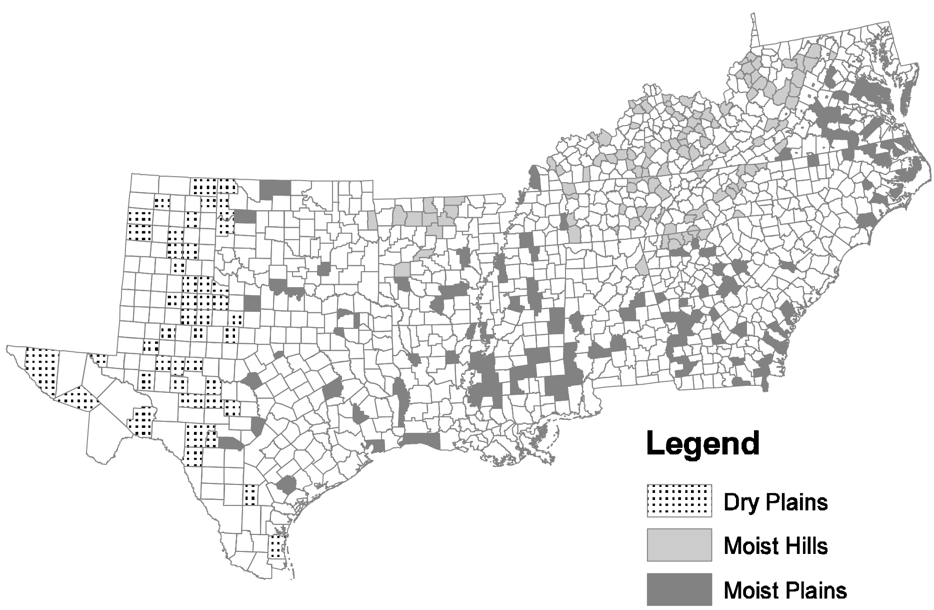
Based upon the intersection of *topology* (hills versus plains) and *precipitation* (moist versus dry if area received more or less than 23-inch annual rainfall), Reaves (1992) divided counties into three environmental categories: dry plains, moist plains and moist hills (dry hills are a null set). The "moist-plain" region is used for raising crops and supports farming; the "moist-hill" and "dry-plain" regions are used for raising livestock and support herding. Homicide rates were then analyzed to see whether the rates were statistically different among these three categories. Most recently, Chu et al. (2000) further revised the classification of counties and tested whether the homicide rates in herding counties were statistically higher than those in farming counties in the south, particularly among non-Hispanic, white offenders. This chapter uses the revised classification by Chu et al. (2000), courtesy of Colin Loftin.

Methods of Constructing Geographic Areas

Data and Study Area

Chu et al. (2000) used two data sets to measure homicide rates: the Uniform Crime Reports Supplementary Homicide Reports (SHR) (referred to hereafter as the "SHR Data") and the National Vital Statistics System (NVSS). The

Figure 1. Environmental classification of counties



former is available through the National Archive of Criminal Justice Data Web site at *www.icpsr.umich.edu/NACJD/home.html*; and the latter is from NCHS (1993). This research uses homicide rates by offenders based on the SHR data only, as the *victims*-based homicide measure by the NVSS (the only available measure from the vital statistics) is not relevant to the proposition.

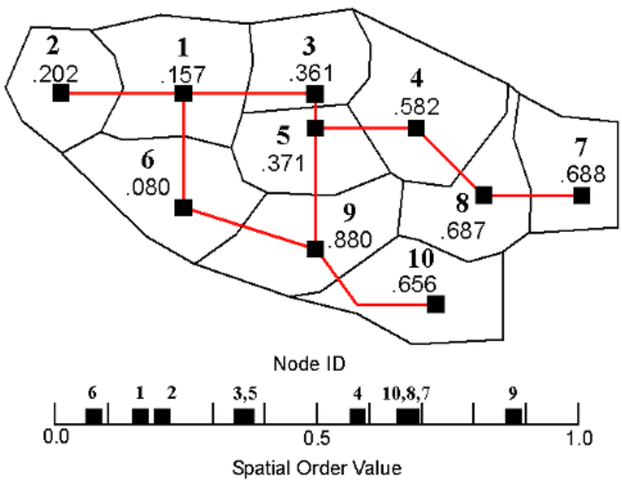
The SHR Data set was compiled by Fox (2000) based on the Federal Bureau of Investigation's (FBI) Uniform Crime Reports (UCR). The data provide incident-level information on *criminal homicides* including location, circumstances and method of offense, as well as demographic characteristics of victims and perpetrators and the relationship between the two. Criminal homicides include murders and non-negligent manslaughters, but exclude negligent manslaughters and justifiable homicides. For the years 1976-1998, the file includes 439,954 of the estimated 481,500 murder victims, and 486,359 of the estimated 532,463 offenders, covering about 92% of homicides in the U.S. The geographic location for each incident includes the county code, which enables us to aggregate the data at the county level. Like Chu et al. (2000), this chapter's study uses only the 1976-83 homicide rates in 318 rural counties in 14 southern states (Alabama, Arkansas, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia). Figure 1 shows the classification of counties (dry plains, moist plains and moist hills). Unlike other such studies, data from historical context ("where dependence on herding is more complete and contemporaneous," (Chu et al, 2000, p. 983)) are preferable for testing the herding-culture-of-honor hypothesis.

Before discussing procedures, we need to set *threshold population* – that is, the minimum population size in the new analysis area for reliable rate estimates. Black et al. (1996) suggested the use of one standard deviation below the mean as a beginning point. In this chapter's study, we have used (1) the mean and (2) the mean minus one standard deviation in order to examine the impact of threshold population.

The Spatial Order Method

The spatial order method uses space-filling curves (discovered by the mathematician Peano) to determine the nearness or spatial order of tracts. Space-filling curves traverse space in a continuous and recursive manner to visit all tracts, and assign a spatial order (from 0 to 1) to each tract based on its relative positions in a two-dimensional space. In general, tracts that are close together have similar spatial-order values, and tracts that are far apart have dissimilar spatial-order values (ESRI, 1998). See Figure 2 for an example based on the ArcInfo online manual by ESRI (1998) with modifications. The method provides a first-cut

Figure 2. Example of assigning spatial-order values to tracts



measure of closeness. The SPATIALORDER procedure in ArcInfo (using the ArcPlot module) is based on one of the algorithms developed by Bartholdi and Platzman (1988). Once the spatial-order value of each tract is determined, the COLLOCATE command in ArcInfo will assign nearby tracts one group number. The threshold population is fed into the COLLOCATE command as the capacity of each group formed by tracts. Finally, tracts are dissolved based on the group numbers. The whole process is automated in ArcInfo Macro Language (AML).

In this particular application, merged counties must be from the same environmental category. In practice, the study area is divided into three regions: one of dry plains, one of moist plains and one of moist hills. The process is repeated on each region. A similar strategy is used in the ISD method. Figure 3 shows the 283 new analysis areas (in bold lines) constructed from the original 318 counties using the mean as the threshold population. Based on the mean minus one standard deviation as the threshold population, 311 areas are constructed. See Table 2 for details. Note that a smaller threshold population leads to less aggregation of counties, and most aggregation occurs in dry-plain counties, which have less population.

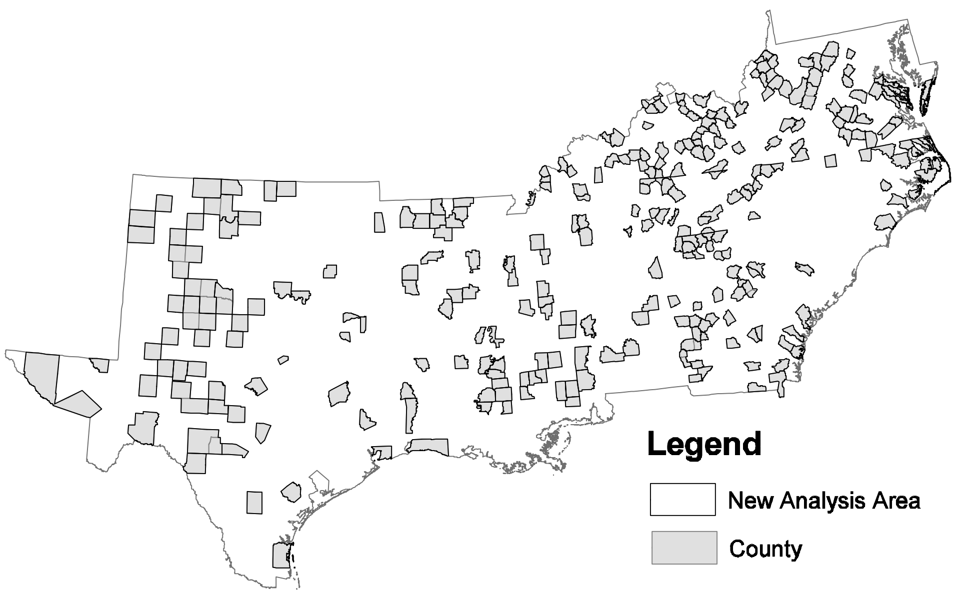
The ISD Method

The ISD method is based on a similar rationale as the spatial order method. The procedure is summarized in Figure 4, which is based on Black et al. (1996) with

Table 2. Analysis areas constructed by the spatial order and the ISD methods based on different threshold population

	Original Counties	Spatial Order Method		ISD Method	
		Mean	Mean-SD	Mean	Mean-SD
Dry plain	47	21	40	12	24
Moist hill	118	115	118	87	115
Moist plain	153	147	153	102	143
Total	318	283	311	201	282

Figure 3. Areas constructed by the spatial order method (mean as threshold)



modifications. A starting tract (for example, the southernmost tract) is selected first, and its nearest and contiguous tract is added. If the total population is equal to or more than the threshold population, the two tracts form an analysis area. Otherwise, the next nearest tract (contiguous to either of the previous selected tracts) is added. The process continues until the total population of selected tracts reaches the threshold value and a new analysis area is formed. The whole procedure is repeated until all tracts are allocated to new analysis areas. The process utilizes two matrixes (one of distances between any two tracts and another of adjacency relationship between tracts) generated by ArcInfo. The method is implemented in a C program.

Based on the mean as the threshold population, 201 areas are constructed from the original 318 counties. Based on the mean minus one standard deviation as the threshold population, 282 areas are constructed. See Table 2.

Result and Discussion

To test the Nisbett-Reaves hypothesis (or the herding-culture-of-honor proposition), we can construct a simple regression model with two dummy variables to implement the traditional ANOVA (analysis of variance). The three environmental categories can be coded by two dummy variables: x_1=0 and x_2=0 for moist plains, x_1=1 and x_2=0 for moist hills, and x_1=0 and x_2=1 for dry plains. Using y to denote white male homicide rates, the model is written as:

$$Y = b_0 + b_1 x_1 + b_2 x_2 \ .$$

A positive and statistically significant x_1 implies that homicide rates in moist-hill areas are higher than moist-plain areas. A positive and statistically significant x_2 implies that homicide rates in dry-plain areas are higher than moist-plain areas. The model's F value indicates whether the homicide rates are significantly different overall among three types of areas.

Table 3 shows the regression results. The regression on data of the original counties shows that moist-hill counties indeed have higher white male homicide

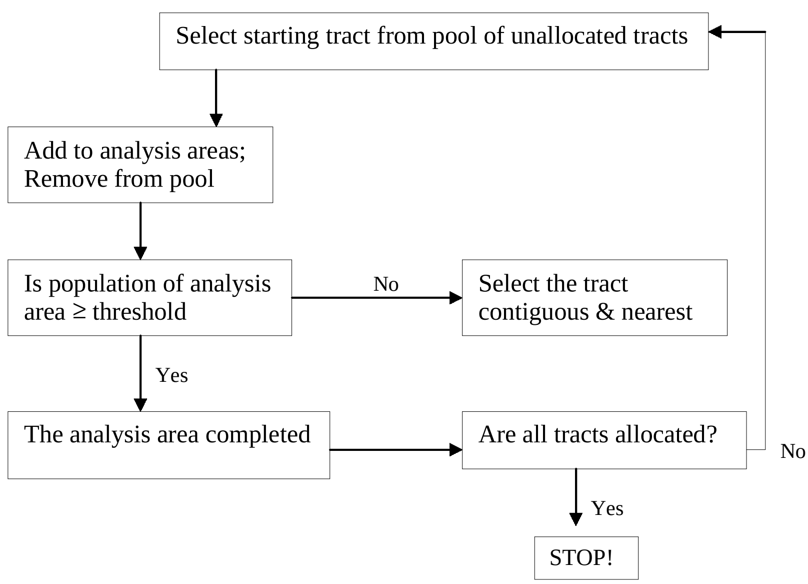
Table 3. Regression results of testing the Nisbett-Reaves hypothesis

	n	b_0	b_1	b_2	F	R^2
Original	318	2.06 (1.02)	5.63 (2.35)*	2.61 (1.13)	3.18*	0.020
ISD (mean)	201	8.08 (1.59)	2.73 (0.50)	-0.79 (-0.15)	0.91	0.009
ISD (mean-SD)	282	4.04 (1.35)	3.85 (1.17)	0.96 (0.30)	1.50	0.011
Spatial Order (mean)	283	4.62 (1.44)	3.28 (0.94)	0.24 (0.07)	1.49	0.011
Spatial Order (mean-SD)	311	2.42 (1.09)	5.27 (2.05)*	2.25 (0.90)	2.67	0.017

*Note: * indicates statistically significant at 0.05* rates than moist-plain counties, and the overall F is also statistically significant. However, none of the F values in the four regressions based on new analysis areas is statistically significant. Only in the case of spatial order method using the mean minus one standard deviation as the threshold population, areas of moist hills have higher homicide rates than areas of moist plains. This is close to the regression result using the original county unit as the aggregation is minimal (that is, from 318 counties to 311 areas). The results lead us to draw a conclusion similar to Chu et al. (2000) that the herding-culture-of-honor hypothesis was more likely an "artifact of unreliable estimates" (p. 982). Most likely, it is caused by several, less-populated counties (mainly in dry plains or moist hills), in which homicide rates are very high with a small base population.

Crime analysts should be alerted to possible outliers of crime rates in small population areas. For example, area-based cluster analysis may reveal false hot spots with high crime rates. Conventional regression analysis may yield spurious results if the issue of small population problem is not addressed adequately. The main objective of this chapter is to demonstrate the two simple methods of constructing geographic areas for rate estimates in small populations. Both methods are sensitive to the choice of threshold population. In this case study of testing the herding-culture-of-honor hypothesis, new analysis areas are constructed from counties of the same environmental category. The within-area homogeneity of attributes is maintained by applying the spatial adjacency and proximity criterion to each region composed of counties from the same environ-

Figure 4. Procedure of the ISD method



mental category. In other cases, balancing the two competing goals of within-area homogeneity and spatial compactness is more complex and remains the main challenge for regionalization.

Acknowledgments

An earlier version of this chapter was presented at the 2002 American Society of Criminology Annual Meeting in Chicago. Thanks are due to Colin Loftin, who provided the environmental classification codes for the 318 counties in 14 U.S. southern states.

References

Bartholdi, J. J. III, & Platzman, L. K. (1988). Heuristics based on space-filling curves for combinatorial problems in Euclidean space. *Management Science, 34*, 291-305.

Black, R. J., Sharp, L., & Urquhart, J. D. (1996). Analyzing the spatial distribution of disease using a method of constructing geographical areas of approximately equal population size. In P. E. Alexander & P. Boyle (Eds.), *Methods for investigating localized clustering of disease* (IARC Scientific Publications No. 135, pp. 28-39), Lyon, France: International Agency for Research on Cancer.

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

Chu, R., Rivera, C., & Loftin, C. (2000). Herding and homicide: An examination of the Nisbett-Reaves hypothesis. *Social Forces, 78*, 971-987.

Cliff, A., Haggett, P., Ord, J., Bassett, K., & Davis, R. (1975). *Elements of spatial structure*. Cambridge, MA: Cambridge University Press.

Dobrin, A., & Wiersema, B. (2000, November 15-18). *Measuring line-of-duty homicides by law enforcement officers in the United States, 1976-1996*. Paper presented at the 52nd Annual Meeting of the American Society of Criminology, San Francisco, CA.

Environmental Systems Research Institute (ESRI). (1998). Arc/Info version 7.2.1 online help menu. Redlands, CA: ESRI.

Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning, A*(23), 1025-1044.

Fox, J. A. (2000). Uniform crime reports (United States): Supplementary homicide reports, 1976-1998 (Computer file). ICPSR version produced by Northeastern University, College of Criminal Justice, Boston. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research.

Haining, R., Wises, S., & Blake, M. (1994). Constructing regions for small area analysis: Material deprivation and colorectal cancer. *Journal of Public Health Medicine, 16*, 429-438.

Harrell, A., & Gouvis, C. (1994). Predicting neighborhood risk of crime: Report to the National Institute of Justice. Washington, DC: The Urban Institute.

Harries, K. D. (1997). Social stress and trauma: Synthesis and spatial analysis. *Social Science and Medicine, 45*, 1251-1264.

Lam, N. S. N., & Liu, K. (1996). Use of space-filling curves in generating a national rural sampling frame for HIV-AIDS research. *Professional Geographer, 48*, 321-332.

Land, K. C., McCall, P. L., & Cohen, L. E. (1990). Structural covariates of homicide rates: Are there any invariances across time and social space? *American Journal of Sociology, 95*, 922-963.

Land, K. C., McCall, P. L., & Nagin, D. S. (1996). A comparison of Poisson, negative binomial and semiparametric mixed-Poisson regression models: With empirical applications to criminal careers data. *Sociological Methods and Research, 24*, 387-442.

Loftin, C., McDowall, D. (2000, November 15-18). *Problems associated with the analysis of homicide in small populations* (Paper). Presented at the 52nd Annual Meeting of the American Society of Criminology, San Francisco, CA.

Messner, S. F., Anselin, L., Baller, R. D., Hawkins, D. F., Deane, G., & Tolnay, S. E. (1999). The spatial patterning of county homicide rates: An application of exploratory spatial data analysis. *Journal of Quantitative Criminology, 15*, 423-450.

Morenoff, J. D., & Sampson, R. J. (1997). Violent crime and the spatial dynamics of neighborhood transition: Chicago, 1970-1990. *Social Forces, 76*, 31-64.

National Center for Health Statistics (NCHS), U.S. Department of Health and Human Services. (1993). Mortality Detail Files, 1968-1991 (Computer file). ICPSR edition produced by NCHS, Hyattsville, MD. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research.

Nisbett, R. E. (1993). Violence and U.S. regional culture. *American Psychologist, 48*, 44-149.

Nisbett, R. E., & Cohen, D. (1996). *Culture of honor: The psychology of violence in the south*. Boulder, Colorodo: Westview Press.

Openshaw, S. (1984). *Concepts and techniques in modern geography, number 38. The modifiable areal unit problem*. Norwich, UK: Geo Books.

Osgood, D. W. (2000). Poisson-based regression analysis of aggregate crime rates. *Journal of Quantitative Criminology, 16*, 21-43.

Osgood, D. W., & Chambers, J. M. (2000). Social disorganization outside the metropolis: An analysis of rural youth violence. *Criminology, 38*, 81-115.

Reaves, A. L. (1992). *The cultural ecology of rural white homicide in the southern United States* (Ph.D. dissertation). University of Michigan.

Sampson, R. J., Raudenbush, S., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277*, 918-924.

Townsend, P., Phillimore, P., & Beattie, A. (1988). *Health and deprivation: Inequality and the north*. London: Croom Helm.

Wang, F. (2005). Job access and homicide patterns in Chicago: An analysis at multiple geographic levels based on scale-space theory. *Journal of Quantitative Criminology, 21* (forthcoming).

Section III

Geographic Profiling

Chapter VI

Geographic Profiling for Serial Crime Investigation

D. Kim Rossmo, Texas State University, USA

Ian Laverty, ECRI, Canada

Brad Moore, Ontario Provincial Police, Canada

Abstract

This chapter describes the technique and application of geographic profiling, a methodology for analyzing the geographic locations of a linked series of crimes to determine the unknown offender's most probable residence area. The process focuses on the hunting behavior of the offender within the context of the crime sites and their spatial relationships. Rather than pinpointing a single location, it provides an optimal search strategy by making inferences from the locations and geometry of the connected crime sites. Geographic profiling can therefore be thought of as a spatially based information management tool for serial crime investigation. Tools based on geographic information systems (GIS), such as the Rigel geographic profiling system, allow the rapid computation and visualization

of the geographic profile as a three-dimensional probability surface, which can then be combined with other geographically based information to narrow the offender search parameters for the criminal investigator.

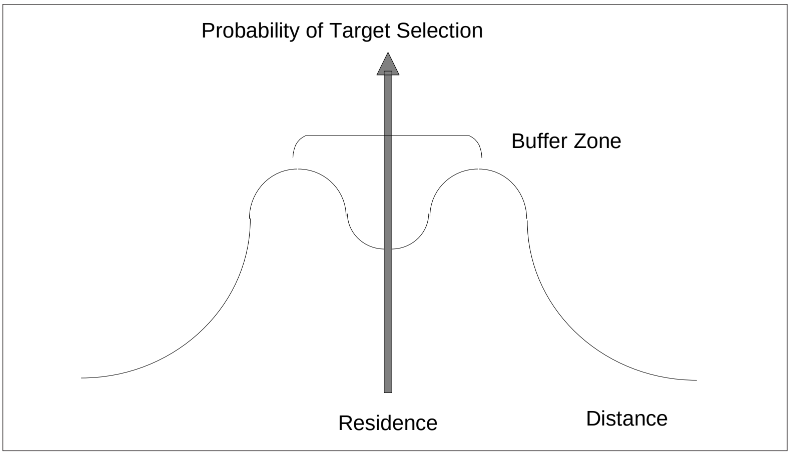
Introduction

Perhaps the most empirically certain aspect of any crime is the location where it occurred. It is therefore natural to consider whether crime locations can be used to help identify unknown criminals. This question becomes even more relevant when a linked series of crimes is identified. Is there a significance to the geographic pattern of the crimes? Everyone is familiar with the concept of putting pins in a map to mark crime locations, and then studying the pattern formed to see if some something obvious jumps out at the observer. At the very least, we understand that when crimes cluster around a particular area, there is something about that area which draws the criminal. But the question becomes a more complex one when we try to define a specific empirical process to follow. What do we mean by "cluster"? How should multiple clusters be interpreted? Should anomalous outliers be ignored? Does background geography or street layout distort the pattern? Should the timing or order of the crimes be considered? Should some crimes be weighted more heavily than others? And what should we be looking to find in this pattern?

Geographic profiling tries to answer these questions by starting from the basics – what influences a criminal's journey to a crime site? We understand certain things intuitively, and these tend to be confirmed by studies. Nobody wants to go further than necessary to accomplish his or her goal; this is known as the *least effort principle* (Zipf, 1950). If the opportunity and the desire to commit a crime exist, the criminal is more likely to take the first, or closest, opportunity. On the other hand, there exists a well-documented aversion to committing crimes too close to home, within a mental "*buffer zone*" (Brantingham and Brantingham, 1981). In practical terms, this can be seen as the desire for anonymity, or diversion of attention away from one's home location. This conflicts with the desire to travel no further than necessary, resulting in a lower probability of crime site selection close to the criminal's home, and then a more typical *distance-decay function* further away (see Figure 1).

Based on these principles of environmental criminology, *geographic profiling* evaluates possible journey-to-crime scenarios for every point on the map, and produces a probability surface indicating the likelihood that any given point is the offender's base. This anchor point is most commonly the home or work site of the offender.

Figure 1. Crime distance-decay function with buffer zone



Geographic profiling is intended to be used as part of an investigative strategy, in combination with other information and other investigative techniques. Rather than pinpointing a single location, it provides an optimal search strategy. This information may be useful in prioritizing leads or directing the geographic focus of an investigation. As it is based on statistical probabilities, not known certainties, it cannot stand on its own as evidence. It is an intermediate step in the investigation of a crime series and the collection of evidence that will solve the case. Not all types of offenders or crimes can be geographically profiled, but in appropriate cases the process produces practical results.

Geography of the Crime Theory

Criminals do not choose their targets at random, and an understanding of the underlying processes involved can assist in decoding their offense patterns. Geographic profiling is based on crime pattern, routine activity, and rational choice theories from *environmental criminology*, a field of study interested in the interactions between criminals and the physical environment that surrounds them. Crime is viewed as the product of potential offenders and their setting or place – the “where and when” of the criminal act. Any understanding of the target patterns and hunting behavior of criminal predators must be aware of the dimensions of offender, victim, crime and environment.

According to *routine activity theory*, for a predatory crime to occur the paths of the offender and victim must intersect in time and space, within an environment suitable for criminal activity. A crime requires: (1) a motivated offender; (2) a suitable target; and (3) the absence of capable guardians (Felson, 1998). Rhythms are important for understanding the ebb and flow of people through an environment. A given location may range from crowded to deserted, depending upon the time, day of week, month or weather. Geography is not independent of time.

Crime and criminal behavior can be viewed as the outcome of rational choices based on the efforts, rewards, and costs involved (Cornish & Clarke, 1986). Criminal offenders are rational and make choices and decisions that benefit themselves. This framework places importance on situational variables, such as scene and victim characteristics, and their choice properties. Pathological crimes involve non-pathological behavior; violent criminals, sex offenders and even psychotic individuals with unfathomable motives exhibit a substantial degree of rationality. This is particularly true for serial offenders who, by definition, are intelligent and successful enough to avoid capture for a time.

Experience changes an individual's information processing and a criminal may improve his or her decision making over time. Learning is an integral part of *rational choice theory*, which sees behavior as interactional and adaptive. But rational does not equal intelligent or sophisticated. Most crime is quick, easy and unskilled. It is typically spontaneous or, at best, only casually planned; it is rarely well thought out. The choices of offenders are often based on decisions that exhibit bounded rationality, limited by constraints of time, effort and information.

Brantingham and Brantingham (1981, 1984, 1993) provide a model of offense site selection called *crime pattern theory*, applying the principles of environmental criminology to understanding the geometry of crime. Their model combines elements of rational choice, routine activity theory, and environmental factors to explain the distribution of crimes. It suggests that criminal acts are most likely to occur in those areas where the offender's awareness space intersects with an environment containing suitable targets at an acceptable level of risk.

Offenders do not choose their crime sites randomly. While any given victim may be selected by chance, the process of such random selection is spatially structured whether the offender realizes it or not. Target choice is affected by the interactions of offenders with their physical and social environments. Potential victims are not considered in isolation from their surrounding environment, and the entire "target situation" must be seen as acceptable before a crime will occur.

A person's *awareness space* forms part of his or her mental map and is constructed primarily from the spatial experiences and routine activities of the individual. An awareness space is derived from an activity space, the latter being comprised of various activity sites (for example, residence, workplace, social activity locations and so forth) and the connecting network of travel and commuting routes. This is similar to the concept of a comfort zone. Well-known locations (landmarks, tourist sites, important buildings) may also become part of a person's awareness space without actually being a component of the activity space.

Offenders search outward from these areas, usually following some form of distance-decay function. In predatory crimes, however, there is typically a buffer zone around the offender's residence. Within this area targets are viewed as less desirable because of the risk in operating too close to home. Offenders seek to maximize opportunity and minimize risk by optimally balancing the desire to operate within their comfort zone with the need for anonymity.

Several other factors must be taken into consideration in a geographic profile. Urban areas usually conform to some form of grid or Manhattan street layout, with dissimilar traffic flows and travel times associated with different streets. Movement as the crow flies is the exception rather than the rule. Consequently, a geographic profile must consider street layout, physical environment, date, time weather, mode of transportation and any physical or mental barriers.

There are often different locations connected to a crime, each with a different geographic meaning. While all are important in geographic profiling, these site types can have varying relevance in different cases. Investigators may also not know the location of every site.

The target or victim backcloth is important for an understanding of the spatial distribution of crime sites. The locations and availability of targets determine where offenses occur. Victim selections that are nonrandom or based on specific and rare traits will require more searching than those that are random, nonspecific and common. If victim selection is specific, as in a series of prostitute murders, then the encounter locations will be restricted, influenced more by the target backcloth (that is, where red light districts are located) then by the offender's activity space. In this type of situation, the body disposal sites become more significant. Victimology thus plays an important role in the development of a geographic profile.

Internal and external influences shape a criminal's hunting process. Serial offenders gain knowledge with each new crime and often learn from their experiences. Media disclosures and certain investigative strategies, such as patrol saturation tactics, may create spatial displacement, which might hinder or delay apprehension. Such influences, while often unavoidable, can change or alter geographic crime patterns.

Geographic Profiling

A *geographic profile* (or *geoprofile*) is focused around the concept of a base or *anchor point* – the single most important place in a person's spatial life. For the vast majority of people this is their residence, therefore the terms anchor point and residence are sometimes used interchangeably. But it must be kept in mind that in certain cases the offender's anchor point is their work site or immediate past residence. Some street criminals do not have a permanent residence, but rather employ a bar, drug corner, or similar location as the base for their activities. Others may be transient, either living on the street or mobile to such a degree that they lack any real base. Investigators should always keep these possibilities in mind.

Crime Sites

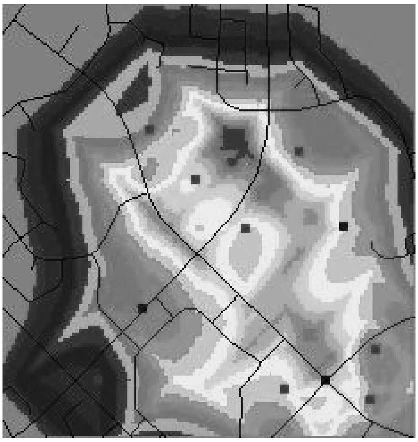
While a street robbery or arson may involve only one location, other crimes may involve multiple related locations. A murder, for example, involves encounter, attack, murder and body disposal sites, which may be in the same or different locations. A rape has encounter, attack, rape and victim release sites. An automobile theft has vehicle theft and drop sites. The various locations connected to a crime have different meanings to the offender, and may be used differently in geographic profiling. Breaking an offense down into its constituent locations and examining the possible groupings is called *crime site parsing*.

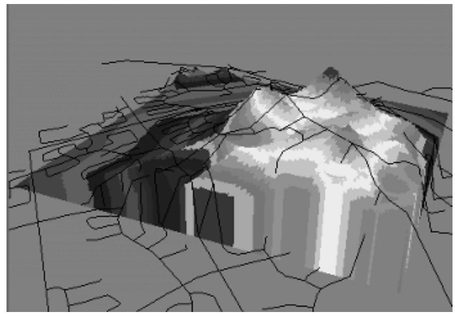
To construct a geographic profile the coordinates of the offense sites must be entered as data points in the analysis software. The site addresses are usually available from police records systems, but they may not be *geocoded* to provide usable geographic coordinates. Geographic information systems (GIS) software is normally used to plot the sites on a map or geocode them automatically from an address database. The *Rigel* geographic profiling software employed by police geographic profilers can use common GIS systems such as ArcView, MapInfo or MapPoint for this purpose.

CGT Analysis and Visualization

Key to the analysis is an algorithm called *CGT* (*Criminal Geographic Targeting*) that is based on the spatial nature of the criminal hunt process. It divides the hunting area (the area enclosing all of the crime sites) into a fine grid, and then calculates the probability that each individual grid point is the offender's anchor point.

Figure 2. A Geoprofile in 2D and 3D





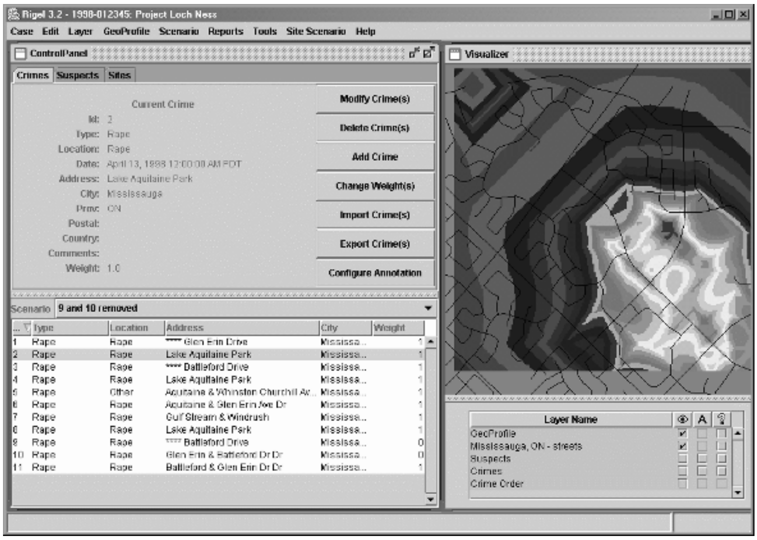
This produces a three-dimensional probability graph, termed a *jeopardy surface*, showing scores (heights) that represent the relative probabilities that given areas contain the offender's anchor point (see Figure 2).

By turning the jeopardy surface into a two-dimensional color probability map, and superimposing it over a street map of the area of the crimes, specific locations can be assessed in terms of likelihood of offender anchor point. The resulting probability map is termed a geoprofile. The computationally intensive calculation of the jeopardy surface and the production of the geoprofile are normally handled by the *Rigel* computer software (see Figure 3).

The distributions of these scores can be plotted on a graph to show the number of pixels (that is, area size) with a given score, which assist in optimizing cutoff points for area prioritization. The geographic profile may indicate a specific high probability offender anchor point region, although it is just as commonly used to prioritize other data. When focussing on a high probability region, it is necessary to also consider neighborhood characteristics, such as socioeconomic status and land use, when assessing the likelihood of offender residence in a given area.

The performance of a geographic profile is measured by what is termed the *hit-score percentage*. This is defined as the ratio of the area searched in the prioritized order given by the geoprofile before the offender is located, to the total area covered by the crimes. A recent review of all now-solved operational geographic profiles from four different police agencies (covering cases in several different countries), combined with the findings from the initial research project at Simon Fraser University, resulted in a mean hit-score percentage of

Figure 3. Calculating a geoprofile with Rigel



approximately 4.7%, and a median of 3.0%, with a standard deviation of 4.4% (N = 1,426 offenses, and 1,726 crime locations) (Rossmo, 2001).

Assumptions

The validity and reliability of the process is dependent upon the accuracy of certain theoretical and methodological assumptions, the violation of which distort the statistical process and affect the accuracy of the results. In all cases, the following assumptions must be considered:

- the *linkage analysis* is accurate (that is, the crimes were committed by the same offender, or team of offenders working together);
- the series is relatively complete (that is, there are not a significant number of offenses that have not been connected);
- the offender is a local hunter and not a poacher (that is, the offender is not commuting into the area to commit these crimes); and
- the offender has not moved residence during the time period of the crimes.

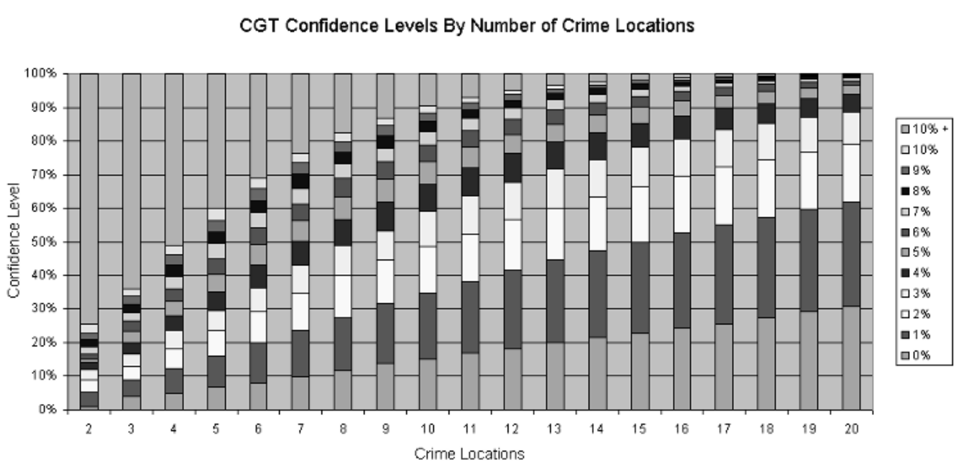
The more locations in the analysis, the more accurate and reliable the process. Statistically the peak area of the geographic profile is considered to be the total area required to give a 70% likelihood of including the offender residence. For cases with a large number of crime sites, this peak area may be as small as a few percent of the total "hunting area" covered by the crimes. For example, with as few as 10 crime sites, a peak area of 5% can be identified with a 70% likelihood (see Figure 4).

Geographic Profiling Techniques

Research into alternative techniques of geographic profiling has tended to focus on three areas in particular:

1. How can distance best be computed? Should it be based on *Manhattan distance* (right-angle distances representative of a city street grid) or straight-line distance (as the crow flies)? Should exact distance by road be computed from a street map if possible, and should it be modified for mobility factors (for example, driving versus walking, heavy traffic versus light)?
2. How can parameters for the buffer zone and distance-decay function best be determined for a specific case? Clearly fixed values do not work well when different crime series involve varying geographic ranges.

Figure 4. Confidence levels by number of crime site locations



3. How can the correct subset of crime sites be selected for geographic profiling when some of the crime sites may be linked in error or may violate underlying assumptions?

An important practical question must also be answered – how much difference does each of these factors make to the actual results of the geographic profile?

Calculation of Distance

It would seem that certain distance measures would be more accurate than others in specific areas. The difficulty is to know which is most appropriate for the specific case, especially when necessary information is not available. For example, the suspect's method of travel may be unknown, or local street maps may be unreliable because they are out of date, fail to note one-way streets, or do not include construction.

The theory of modeling human behavior discusses the desirability of three qualities: (1) realism; (2) generality; and (3) precision (Levins, 1966). Unfortunately, improvement in one quality usually results in deterioration in another. Greater precision, for example, often means a loss of generality. In the real world of policing, much is often unknown during a criminal investigation. Consequently, the approach in geographic profiling has been to emphasize generality, realism and robustness, even at the expense of precision. The only way to tell how much the geographic profile may be affected by the chosen method is to actually try the experiment on a significant number of cases. This testing has shown that improvements in precision are marginal after a certain level. It makes little sense to achieve a 1% performance improvement at the expense of a 15% loss in applicability.

Research has shown that Manhattan distance gives the most accurate result in the greatest number of cases, while not being significantly worse than other methods across the entire spectrum of cases – a finding true for both North America and Great Britain. This basic distance measure is surprisingly forgiving of such major factors as large excluded areas within the hunting area (for example, lakes, industrial parks and so forth), barriers such as rivers and railroad tracks, and bottlenecks such as bridges. There are specific geographic cases where simple methods break down, however, such as along a seacoast. As long as these specific exceptional cases can be identified and recognized through training and experience, the most reliable and practical method involves the use of Manhattan distance.

Determining Function Parameters

Various methods have been proposed for the process of determining parameters for the buffer zone and distance-decay function, generally relying on calibration against a large dataset of similar cases (for example, the method proposed in *CrimeStat II*) (Levine, 2002). The obvious difficulty is the specific case under investigation may differ significantly from the available datasets (for example, the crimes may cover a much larger or smaller area than average for that type of offense). Also, as a geographic profile is based on an individual crime-distance probability distribution, using group crime-distance probability distributions based on samples of offenders creates an ecological fallacy. This can significantly reduce accuracy.

It would be best if these parameters could be adjusted automatically to fit the pattern of crimes seen in each specific case. The only way to devise such an algorithm, and to test the relative effect on the geographic profile of various parameters, is to examine a large number of actual cases. The commercial geographic profiling system *Rigel* incorporates one such algorithm (details are proprietary). In general, the noncommercial systems leave it to the user to determine these parameters for each individual case, an impractical approach for a police operational tool.

Scenario Selection

In real criminal investigations of serial crimes involving unknown suspects and uncertain events, it is possible for some crimes to be incorrectly linked. Also, the key underlying assumption of independent journey from anchor point to crime site may be violated for certain offenses. While it is possible to prune the data to eliminate "suspicious" sites, how can this be done in an unbiased way?

Professional police geographic profilers spend a large proportion of their training on studying this problem, and learning how to go about developing a valid "*scenario*" (that is, the optimal subset of crime sites to be profiled). Some factors, such as determining the validity of crime linkage, depend on experience and are outside the scope of the geographic profiling system. It is possible for the geographic profiling computation to include a relative probability of linkage, but it has been found in practice that it is difficult to guess at this (25%?; 75%?). Professional profilers generally prefer to leave out any data points that they are not reasonably certain are linked to the series.

Practical rules have been developed to filter out most nonindependent crime sites, based on proximity in space and time. These rules are intelligently employed by the police profilers. The commercial *Rigel* geographic profiling system incorporates these rules in an automatic *Expert System*, which makes recommendations to the profiler.

Anomalous outliers are a problem for geographic profiling, as they tend to have a disproportionate effect on the profile. A small number can be tolerated, but it is best to identify and eliminate such outliers, as practical experience over a large number of cases has shown that they are most often in error or represent genuinely anomalous behaviour (for example, an unusual trip outside of normal activity space by the suspect). Rules for what constitutes an outlier can be developed by studying a large number of cases. The police geographic profilers learn and employ a standard set of criteria in their work. The *Rigel* Expert System used by police profilers incorporates these rules, and recommends the elimination of specific outliers as well as nonindependent sites.

Using the Geographic Profile

Investigative Strategies

Once a geographic profile has been constructed, a variety of criminal *investigative strategies* may be more effectively and efficiently employed.

The geographic profile, in conjunction with a psychological or behavioral profile, can help prioritize individuals for follow-up investigation. The problem in most serial violent crime investigations is one of too many suspects, rather than too few. Profiling can help manage hundreds or thousands of suspects, leads and tips. In addition to focusing an investigation, a geographic profile may also assist in the trial stage by analyzing the spatial relationship of a crime site pattern and its congruence with an accused offender's activity space.

Areas most likely associated with the offender can be used to help direct patrol efforts. This strategy is particularly effective if the offender is operating during specific time periods. Prioritized areas may also be employed for neighborhood canvassing efforts, area searches, information sign posting and community cooperation and media campaigns. Police departments have used this approach to focus directed community mailings to obtain suspect information from targeted neighborhood areas.

Additional investigative leads can be obtained from information contained in various computerized police record systems (for example, automated jail booking systems, records management systems, computer-aided dispatch systems and so forth). Offender profile details help focus the search. Police agencies with computerized records containing offender addresses, physical descriptions and modus operandi details of local criminals can use profiling information, including residence area, as the basis for search criteria.

Data banks are often geographically based and valuable information may be obtained from parole and probation offices, mental health outpatient clinics, social services offices and private businesses based in the peak geoprofile area. Some police and correction agencies keep information on known violent criminals, and most U.S. states have mandatory sex offender registry programs. Linkage analysis systems, such as *ViCLAS* (*Violent Crime Linkage Analysis System*) in Canada, or *ViCAP* (*Violent Criminal Apprehension Program*) in the United States, contain detailed data on sexually motivated violent crimes and criminals. It has been estimated that approximately 85% of records – and almost all police records – contain address information.

A geographic profile can prioritize zip or postal codes. If suspect description or vehicle information exists, prioritized zip or postal codes can be used to conduct searches of registered vehicle or driver's licence files contained in state or provincial department of motor vehicle records. Even with just a few search parameters (for example, vehicle type and color), this usually produces a reasonably small set of records with the appropriate responses. Depending on the number of specifiers, and the size of the ranges, thousands of cases may be narrowed down to a few dozen vehicles or drivers. This strategy can produce significant results by focusing on areas of a size manageable for a major crime investigation.

Task forces investigating specific series of crimes often collect and collate information in a major case management system. Geographic profiling can prioritize street addresses, zip or postal codes, and telephone numbers (NNXs). The details of the specific computer database software, including information fields, search time, number of records and correlational abilities, determine the most appropriate form that the geographic profile should take to maximize usefulness to the police investigation.

In cases where the identity, but not the whereabouts, of a criminal fugitive is known, geographic profiling may be able to assist in determining probable hiding places. Sightings, purchases, credit or bank card transactions, telephone calls, cellular telephone switch sites, crimes and other locational information can be used as input for the profile. This process is also applicable to extortion and kidnapping investigations.

In certain missing person cases that are suspected homicides, geographic profiling can help determine probable body dump site areas.

While geographic profiling is primarily an investigative tool, it also has a role in the courtroom. In addition to analyzing the geographic patterns of unsolved crimes for investigative insights, the spatial relationship between the locations of a crime series and an accused offender's activity sites can be assessed in terms of the probability of their congruence. When combined with other forensic identification findings (for example, a mitochondrial DNA match), such information increases evidential strength and likelihood of guilt. Geographic profiles have also been used as supporting grounds for search warrant affidavits.

Case Example

In 1998 there were a series of 11 linked sexual assaults, including one rape, in Mississauga, a suburb of Toronto, Ontario, Canada. The linkage was made through offender description, behavior and speech pattern, and the timing and location of the attacks. The Peel Regional Police, the responsible agency, instigated *Project Loch Ness* to investigate this case. Investigators developed a list of 312 suspects. A DNA scene sample was obtained from the last crime, the rape, but testing all the suspects was too expensive and time-consuming to be feasible with the technology available at the time. The detectives therefore prioritized the suspects based on several factors, including an uncorroborated alibi, similarity to the composite sketch, the results of initial interviews, a behavioral profile and the geographic profile (the same profile shown in Figures 2 and 3).

This process was able to eliminate 144 of the 312 suspects, and the remaining 168 were assigned priority levels. In the initial batch of four forensic samples sent by police to the laboratory for testing, one was identified as top priority. The lab reported a match after eight days. Lee Marvin Payne was arrested in September 1998. He confessed and then pled guilty in February 2000. Payne was sixth in the prioritized suspect list from the geographic profile alone (a hit-score percentage of 1.9% measured by suspects and 2.2% measured by area). The key evidence used in convicting him was the DNA match and the confession; profiling and suspect prioritization, however, resulted in a faster investigation, substantial cost savings and possibly fewer victims.

This case illustrates some of the most important points about geographic profiling, in particular, that it is most effective when employed in conjunction with other techniques, and that it can be used to prioritize investigative information and direct search strategies.

Conclusion

Geographic profiling is a sophisticated, spatially-based information management tool with particular value for serial crime investigation. A geographic profile can determine the most probable area of offender residence, providing a framework for prioritizing suspects and tips, and serving as a basis for developing new investigative strategies. The value of a geographic profile may be enhanced when it is combined with other independent sources of information, such as a behavioral profile or neighborhood information on land use and demographics. By employing a geographic profile to focus on the most likely area of offender residence, an investigation can better manage and assess its information, reducing costs and helping to apprehend criminals in a timely manner.

References

Brantingham, P. J., & Brantingham, P. L. (Eds.). (1981). *Environmental criminology*. Beverly Hills, CA: Sage.

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

Brantingham, P. L., & Brantingham, P. J. (1993). Environment, routine and situation: Toward a pattern theory of crime. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice* (pp. 259-294). New Brunswick, NJ: Transaction.

Cornish, D. B., & Clarke, R. V. (Eds.). (1986). *The reasoning criminal: Rational choice perspectives on offending*. New York: Springer-Verlag.

Felson, M. (1998). *Crime and everyday life* (2nd ed.). Thousand Oaks, CA: Pine Forge Press.

Levine, N. (2002). *CrimeStat: A spatial statistics program for the analysis of crime incident locations (v. 2.0)*. Washington, DC: National Institute of Justice.

Levins, R. (1966). The strategy of model-building in population biology. *American Scientist*, *54*, 421-431.

MacKay, R. E. (1999, December). Geographic profiling: A new tool for law enforcement. *The Police Chief*, 51-59.

Rossmo, D. K. (1995). Place, space, and police investigations: Hunting serial violent criminals. In J. E. Eck & D. A. Weisburd (Eds.), *Crime and place: Crime prevention studies, Vol. 4* (pp. 217-235). Monsey, NY: Criminal Justice Press.

Rossmo, D. K. (1997). Geographic profiling. In J. L. Jackson & D. A. Bekerian (Eds.), *Offender profiling: Theory, research and practice* (pp. 159-175). Chichester, UK: John Wiley & Sons.

Rossmo, D. K. (2000). *Geographic profiling*. Boca Raton, FL: CRC Press.

Rossmo, D. K. (2001, November). *Evaluation of geographic profiling search strategies* (Paper). Presented at the meeting of the American Society of Criminology, Atlanta, GA.

Zipf, G. (1950). *The principle of least effort*. Reading, MA: Addison Wesley.

Chapter VII

Single Incident Geographical Profiling

Richard Z. Gore, University of Massachusetts-Lowell, USA

Nikolas J. Tofiluk, West Midlands Police, UK

Kenneth V. Griffiths, West Midlands Police, UK

Abstract

This chapter describes the results obtained by using simulation software to determine the ability to rank the suspects of a single incident based on the geographic information derived from arrest records. The current software uses three different geographic filters. These geographic filters were based on the standard distance decay curve (DDC), an incident based distance decay curve (IBDDC) and the incident-based offender residence probability surface (IBORPS). These filters were rated on their ability to order suspect lists. Boundary effects due to the crude core-periphery population gradient characteristic of cities and sub-areal heterogeneities were found to be associated with the standard DDC analysis. The results indicated a definite utility value in these filters, which tends to support crime theories based on the premise that criminal activity patterns are systematically influenced by an offender's geographic setting. These results also have strategic significance in the formulation of policies that prescribe the assembly and

processing of suspect lists and serve as the basis for geographic profiling based on a single incident.

Introduction

Geographic profiling has been defined by Rossmo (2000) as “an information management strategy for serial violent crime investigation that analyzes crime site information to determine the most probable area of offender residence” (p. 259). The computer software portion of Rossmo’s procedure, Rigel, recalculates a generalized distance-decay curve, adjusting it to fit the dimensions and pattern distribution of the crimes believed to be part of a series. The advantages of this method are that it only needs the locations associated with a series of crimes as input and the adjusted DDC is now tailored to the behavior of that particular individual. The disadvantages of the Rigel calculation is that unless the series contains a sufficient number of crime-related locations the results will be poor. Although the necessary number of crimes for good results varies with local conditions (Rossmo, 2000), it should not be expected to perform well on only one or two cases.

This chapter describes a method that attempts to extend quantitative geographic profiling into this case gap by describing a data intensive method that can be used with a single incident. This single-incident method makes the prediction of offender residence based on establishing the locational relative frequencies of offender residences obtained from arrest record data for all who have committed crimes around the incident location in the past. This method draws its justification heavily from the principals of environmental criminology with its emphasis on environmental setting and routine activity theory. In the application described here, it is assumed that a list of suspects has been generated which now has to be ordered using the empirically established probabilities linking offender residence location to incident location.

Conceptually, this method could be expanded for use in typical serial applications. However, the fact that the single-incident method employed here showed effectiveness does not imply that the compounded application with serial incidents would produce improved results. Only further research can answer this question. Should it show utility in a serial incident application, there would still be a question as to whether or not it offered any advantage to the widely used theory based technique employed by Rigel. Even if it turned out that these methods, based on data from aggregated arrest records, was superior only when the number of incidents was small, it would extend the conditions under which geographic profiling was useful.

Attempts to model the journey-to-crime (J2C) process, a necessary precursor for the development of any geographic profiling method, have involved empirical, theoretical and combinational formulations of the problem. In the past, limited data sets of adequate size and/or the lack of availability of powerful and user-friendly GIS software prevented extensive spatial analysis of this problem. Empirical formulations have employed the construction of frequency curves from arrest data. These plots record the number of crime trip distances falling into a series of fixed-distance intervals. Because these curves generally show a consistent falling frequency with increasing travel distance they are called distance-decay curves (DDCs). The interpretation and merits of using this aggregated data are still debated (Van Koppan & DeKeijser, 1997; Rengert, Piquero & Jones, 1999; Levine, 2000). Attempts to understand the underlying significance of these curves has seen the wide use of theoretical models based on gravity analogs. Smith (1976) used migration rules and intervening opportunity theory within a gravity model. He attempted to improve the correlation between the observed offender flows in Rochester, New York, and model predictions by using different opportunity surrogates in place of the standard distance measure. Rengert (1981) developed an opportunity model that starts with deterministic expressions based on gravity and potential theory analogs, which are postulated to describe the journey to crime process. He then used these theoretical results and compared them with empirical observations on burglaries in Philadelphia in order to evaluate the findings of previous investigators and to comment on the complexities of the process. The value of theoretical formulations as pointed out by Rengert (1981) are that they can be adjusted to predict future crime patterns based on expected or postulated changes in the underling controlling factors. Levine (2000) gives an excellent summary of the use of theoretical gravity models, as well as fitting mathematical expressions to empirical data. This paper will show that some applications of DDCs have often been based on hidden assumptions probably not generally true. These conclusions obviously have implications for attempts to model the J2C process using gravity models and other variables.

A somewhat different empirical approach has been advanced (Gore and Pattavina, 2001) that uses a clustering metric based on offender residences. This method first defines a target neighborhood around the incident or hot spot under investigation. Then a search of the arrest records is made for either all or just similar incidents occurring in that target neighborhood over some period of time. Next the offender residences associated with those incidents are determined. Lastly, these residences are used to create a contoured point intensity surface whose values are converted to empirical probabilities. As an empirical method, it can be argued that incident-based offender residence probability surfaces (IBORPSs) should have an advantage over the traditional DDCs in that the associated probabilities are specific to the area around the incident location and not the aggregated sum over a large area that could have a substructural or domain (Turner & Weiss, 1963) character. A domain is a statistically homogeneous subarea with regard to some metric defined within a larger spatial framework.

This chapter will introduce the use of an incident-based distance-decay curve (IBDDC) as a potentially improved distance based J2C metric. This method uses the same logic as that described above for the creation of IBORPSs, but instead of calculating the density of the offender residences, the distances to their respective incident locations are calculated and the results combined exactly as described above for the creation of the standard DDC. The IBDDC potentially can be used with the IBORPS probability measures of offender residence density to obtain a more complete numerical description of the information implicit in the J2C geometry.

Description of the Database and the Computer Software

The database was supplied by the West Midlands Police and contains all the incidents reported in the city of Wolverhampton, UK, during 2001. Wolverhampton is an old, industrial city with a population of a 250,000 people. The city is divided into two operational command units (Figure 3) or policing districts at the northwest end of the West Midlands force area, which also includes Birmingham and Coventry.

Wolverhampton is part of the Birmingham Conurbation where extensive road and public transportation facilities are present, which potentially offers a motivated offender access to multiple target localities. This chapter study is part of a long-term Force-University cooperative effort to create new applications for the West Midlands Police's Force Linked Intelligence System (FLINTS) software.

From the database, all incidents for which an offender had been detected were extracted. This list was further condensed by removing all incidents where the journey-to-crime distance was zero. It was felt that, as we are primarily interested in the combined effects of local geography and the friction-of-distance on the offender's behavior, disputes at the offender's residence would not be influenced by these factors. This resulted in a list of 7414 incidents. This list of crime-criminal pairings was restricted to crimes committed in Wolverhampton, but the offenders can come from anywhere. The boundary effect problem referred to later in this paper is not that which is typically due to an increasing number of incidents and suspects unpaired as one approaches the boundary of a police enforcement area but probably relates mainly to a decreasing population density when moving away from the center of a typical city and speculatively to factors related to the transportation network and the behavioral legacy imposed by experiences, such as schools attended, which are often controlled by jurisdictional boundaries.

Software was developed to test the relative ability of three geographic filters to improve the position of a suspect on a randomly ordered suspect list. These filters were based on the standard DDC, the IBDDC and the IBORPS.

Procedure Followed by the Computer Program for this Chapter Study

1. An initial suspect is chosen from the database. It is their associated crime that is assumed to be the crime under investigation.
2. Nineteen other suspects are randomly chosen from the database.
3. These 20 suspects becomes the suspect list. This list is randomly shuffled 2,000 times. This process produces a random list.

   The actual offender by chance could be at the very top of list, the bottom or anywhere in between.
4. This random list is passed through each of the geographic filters, which assign a relative probability measure to each suspect based on the location of their residence relative to the incident location. The geographic filters are regenerated before each simulation leaving the actual offender's residence out of the respective curves or density surface. This procedure results in generation of four separate lists: the original random list and a list from each of the three filters.
5. Steps 2 through 4 are repeated 24 more times using the same offender but different "other suspects." A single simulation in this procedure generates 100 lists. This includes the 25 original random lists, all containing the offender, and an additional 75 lists, 25 for each filter. Effectively, we will be comparing the offender to 475 other randomly chosen suspects. The position of the offender on each group of 25 lists of the same type is averaged. If the randomizing process is true, we know that the offender should average out at 10.5 on the random lists. If an offender's position on filtered lists is below 10.5 by a significant amount, we can conclude that the filter is implicitly reflecting the existence of geographic information that has predictive value.

6. Steps 1 through 5 were repeated 100 additional times producing 101 total simulations. Each simulation is based on a different incident having a different offender. All crime types in the database can be chosen.

This process has used no other filters, such as only choosing other suspects who have committed offenses of the same type, to reduce the potential list of suspects. Everyone in the database who had committed a crime in Wolverhampton was considered a potential suspect to be added to the list with the actual offender. This rational is employed to test the proposition as to whether or not there is any systemic relationship existing, based on the employed geographic metrics alone, to preferentially link incident location and offender residence. This link is measured by the average improvement in list position the actual suspect might experience after filtering over the expected random placement of 10.5. If this preferential link exists, it would be supportive of crime theories that maintain that geographic considerations influence criminal activity.

All references above to J2C have avoided the question of the starting point or anchor point. The anchor points for crime are often the offender's home, but can be many and varied. The analysis employed in this paper has assumed that the home address contains information about where an offender commits his crime. The degree to which this is true will determine how much better than random will be the average result of offender list placement. For example, if the anchor point for all offenders was always work, then no geographic association would likely exist between home address and incident location and the application of geographic filters based on home address should yield results no better than random. As the results discussed below show that list placement is significantly improved by use of the geographic filters, several possibilities exist. Among these are that the anchor point of most offenders is their home, a location close to their home or a location not near their home but systematically linked to this home neighborhood.

Even assuming the anchor point is strongly linked to the home for most offenders, data quality issues exist. The offender may give a false address, an old address or a relative's address, but may actually be living elsewhere when the crime was committed.

In the United Kingdom in general and within the West Midlands Police in particular, data quality issues have been a high priority. In addition to police enforcement and prevention programs, various crime prevention and revenue distribution programs of other government departments rely on accurate address data for offenders.

When a person is taken into custody, his/her address details are generally checked by a visit to the given address. If the person lives in the district in which they are detained, often a police officer will call at their home address and confirm that they are resident there. They often check bedrooms for clothing and documents such as letters. If there is no one at home, the offender is often taken home to ensure that they go in the door using their key. Sometimes authority is given to search the address and, if available, their key will be used. These methods are nearly always undertaken by the arresting officer. If the person lives outside the district but within the West Midlands force area, a radio "command and control" message is sent to the local police station and the local officers undertake the address check. Again, this is done physically by an officer and sometimes that can take an hour or two. Checks for addresses outside the West Midlands force area are done in a similar way and again a personal visit is requested. The officer verifying the address is always recorded for accountability. Of course, some people will move after the charge procedure and try to avoid going to court, or they have several places of residence in an attempt to avoid detection. The FLINTS database labels offenders with these histories and includes all the addresses they have given in the past. Offenders so labeled tend not to be believed the second time around. If there is any doubt, the custody officer at the police station responsible for the detained person can impose conditions on his release, including a surety, residence curfew and so forth. These restrictions are limited and if they are not suitable, the offender is kept in police custody and a court can impose further conditions or remand them into prison custody. Each address within the West Midlands is assigned an Easting and Northing from a database derived from the Ordnance Survey and is accurate to 1 meter.

Analysis of Results

Generalized Results

The average for each simulation list type was calculated over the 101 simulations (Table 1). The average offender position on the random lists was exactly as expected – 10.50 with a standard error of 1.29. Each average is effectively derived from the results on 2,525 lists, 25 lists for each particular filter type for each simulation and 101 simulations. The average list placement for all three filters (5.29-6.28) was more than three standard errors away from the expected random position. These results would be expected less than one percent of the time, if the geographic filters were doing no better than random. The similarity of results from the three filters, each derived from somewhat different geographic assumptions, also suggests that the results reflect a systematic association of incident location and offender residence. It should be emphasized that it is not the list placement improvement of the offender that is the important result of this simulation. That result was totally expected as DDCs and list placements are just two different ways of expressing the same frequency distribution. What the simulation has revealed is a quantitative estimate of the amount of benefit associated with filter use in ordering suspect lists.

Another measure of relative effectiveness between the filters would be a tally of the number of lists on which the offender showed improvement over the expected random position. The Standard DDC produced improvement on 76 of the 101 simulations. The IBDDC placed 81 of the offenders in improved positions. The IBORPS produced 77 improvements. This result suggests that some additional benefit might derive from devising an empirically defined composite filter combining the attributes of the IBDDC and the IBORPS.

Implications for Procedures Controlling Suspect-List Generation and Use

If a list of usual suspects is routinely assembled (the list can be any length) for every crime and the actual offender is on it, than the application of one of these geographic filters should, on average, significantly improve the actual offenders position on the filtered list. If we assume some average time to investigate each suspect on a list, these results from Wolverhampton suggest that, if the offender is always one of the suspects on the list and the investigators work the list in sequential order, a savings in time ranging from 41% (DDC Filter), 43% (IBORPS) to 48.5% (IBDDC) would be achieved over using the random list. In the real world it is unlikely that the offender would be on every suspect list. It is well known that for any random crime the chances that the offender is a without a previous record is low. However, if the offender was, on average, on only one third of the lists, this would bring them to the attention of the investigator 13.7 % to 16.2% sooner. It is important to recognize that this result is of strategic significance and not tactical significance for a single, specific crime. For example, randomly the actual offender may be at the very top of a specific list and the application of a geographic filter could move them down. The time savings is an average result that should be expected over 50 or more cases.

These results suggest additional qualifications regarding the modification of the list and the assignment of list suspects to investigators. If attribute filters are now to be used based on a suspect's modus operandi (MO) and description, these must be binary filters resulting only in removal of a suspect from the list based on their failure to meet filter criteria. No sorting filters that reorder the list in any way can be employed. This prohibition applies only to attempts to sequentially

Table 1. Average list position using different filters

	Filter Type			
	Random List	IBORPS Filter	DDC Filter	IBDDC Filter
Offender List Placement	10.50	5.99	6.19	5.41

compound sorting procedures and does not preclude the use of the filter probability values in a more sophisticated model where a score would be compiled from weights calculated for each relevant factor, both geographic and descriptive. When dividing this list, a card dealing process should be employed. For example, if a list of 80 suspects is to be divided among four investigators, investigator one gets the first suspect on the list, investigator two gets the second and so forth until the whole 80 suspects have been dealt out. This process gives each investigator a list that benefits equally from the geographic filtering.

Research is currently underway on how to use these simulation results based on past events as a basis for establishing criteria for shortening the suspect list, as well as improving suspect list placement. This is particularly important when applying these methods to volume crime, as was the initial goal of this work. If a list of all burglars with a particular MO yields 58 names, this is clearly too many suspects to warrant the investigative costs. If it is found that for a given neighborhood, the actual burglar of a given crime was in the top 40% of the list 99% of the time, then the lowest 35 names could be dropped with only a 1% change that the actual offender was eliminated from the list.

Boundary and Domain Consequences in the Use of Distance-Decay Curves

This study has revealed complexities associated with DDC usage not mentioned in any of the published work known. Some debate (Van Koppan & DeKeijser, 1997; Rengert et al., 1999; Levine, 2000) has arisen about how much of the distance-decay phenomena based on aggregated data is reflective of individual behavior or how much is an artifact of the method of aggregation. Without elaborating, this paper appears to support Rengert et al.'s conclusion:

In other words, while the aggregate-level analysis clearly indicated distance decay, this pattern may not be characteristic of every single burglar (for example, burglar 6). Yet, the aggregate distance-decay function appears to mimic the majority of the burglars who individually exhibit distance decay in their offending. (1999, p. 434)

We would expand this statement, substituting the word "offender" for "burglar." The distance-decay pattern is evident in curves produced from aggregating the journey-to-crime distance for all crime types over different levels of areal aggregation. The debate as to how much of the apparent distance decay is reflective of individual behavior and how much is due to the aggregation process is not a critical issue in the application of suspect-list filtering for a single crime. No assumption about individual behavior is made or implied. This application interprets the DDC exactly as intended in any usage that assumes that aggregated trials are a measure of empirical probability.

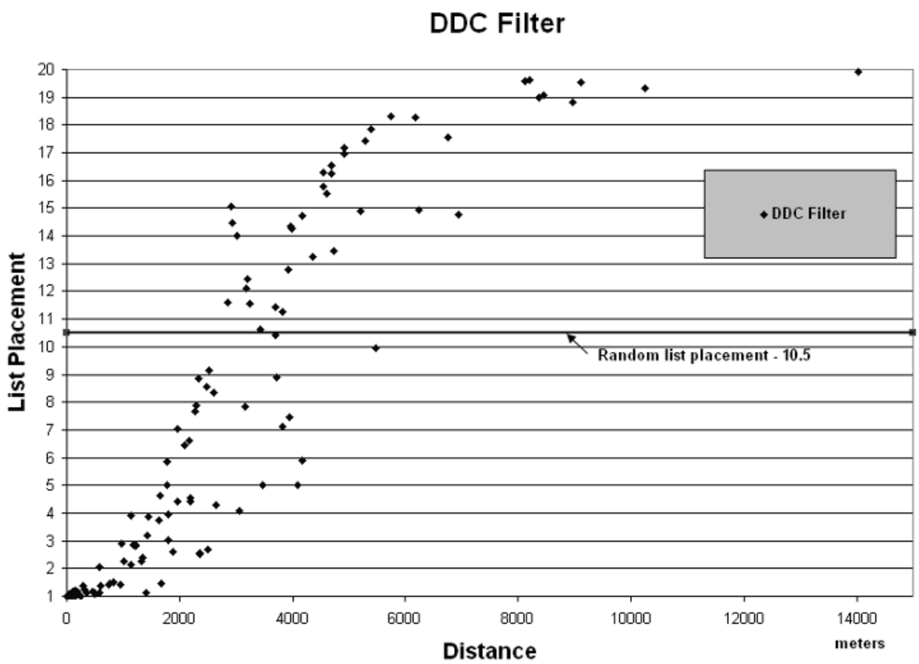
The usage of DDCs in serial criminal investigation is tacitly assuming that the probabilities associated with this aggregated frequency distribution assembled from individual crime trips are generally representative of what should be expected if multiple trips by a single individual or corporate individual (gang) could be observed. The major practical issue here is not whether distance decay characterizes individual behavior, but whether it is justified to use the empirical probabilities that derive from aggregating individual crime trips over a large jurisdiction as measures of individual behavior in sub-areal analysis.

Boundary Effects on DDC Effectiveness

A scatter plot (Figure 1) of the 101 simulations for the DDC filter reveals the expected characteristics. The filter assigns higher probabilities to persons living closer to the crime with the result that these offenders will place higher (closer to 1) on the list. Offenders who lived far away (over 8,000 meters) place nearer the bottom (closer to 20) of the list. The black horizontal line at 10.5 is the position that would be expected from a random process. The DDC filter placed 76 of the 101 offender simulations below this random value. This result in itself, of course, is not surprising as it is simply an alternate means of expressing the probabilities of individual crime trips in Wolverhampton and is the justification of using the DDC filter in the first place. However, this plot also reveals an inherent weakness of the DDC filter. For those, although fewer, offenders who place over 10.5, the DDC filter has actually degraded their position relative to that expected on a random list.

As expected, the plot (Figure 1) shows a large, tight cluster of offenders in the first 500 meters where the average list placement is below 1.5. Somewhere between 600 and 1000 meters, the offender list placement begins to spread out for a given distance. For offenders who traveled about 3000 meters to the

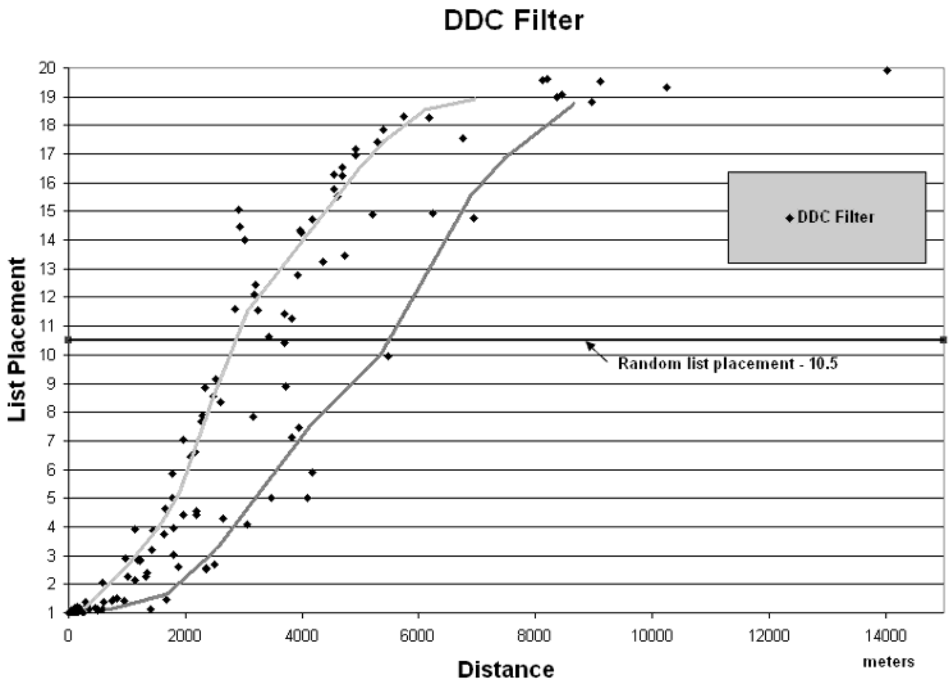
Figure 1. X-Y scatter plot of the list placement versus journey-to-crime distance for the 101 incidents used in the simulation



incident location, the average list placement varied from about 4 to 15. Similar large variations are seen with decreasing magnitude as the distances increase and the list placement range is restricted by the upper limit of 20. If this large range is reflective of randomly differing levels of association between a specific incident location and offender residence at distances greater than 3000 meters, then the method would be quite noisy and of little practical value beyond this distance.

Analysis of the x-y scatter plot (Figure 2) was made by mapping the incident locations near the bottom of the range (below the lower curve) and those near the top of the range (above the upper, lighter curve). This map (Figure 3) reveals that the observed list placement range for a given distance is not mainly noise, but is a boundary effect inherent in applying the distance-decay logic to urban incident locations. Those locations (stars in Figure 3) that plotted closer to the top of the list (list position 1) for a given distance (below the lower, darker curve in Figure 2) are all on the periphery of Wolverhampton. Those locations (circles in Figure 3) that plotted closer to the bottom of the list (above the upper curve in Figure 2) are locations near the center of the city. Note that one circle occurs

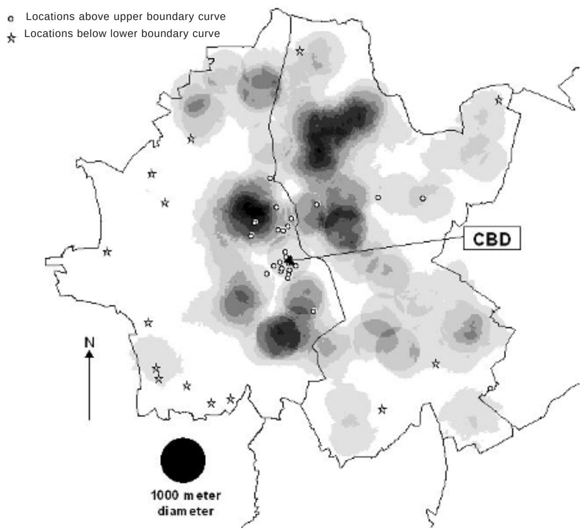
Figure 2. Shown is the partitioning of scatter plot data into three regions (the top of list range region lies in the area below the lower, darker curve; the bottom of list range (nearer 20) region lies above the upper, lighter curve)



near the periphery (on the southeast side), which has an associated J2C distance of 8,112 meters. This point was incorrectly associated, but is useful as it is indicative of where the two extreme sets of data begin to remerge against the lower list limit of 20.

Once this core-periphery pattern of relative effectiveness of the DDC filter was recognized, its explanation is fairly simple. When DDC logic is employed starting at the incident location, the radius of the medial circle determines a rough estimate of filter effectiveness. In this instance, the medial circle is defined as that circle around the incident location containing 50% of potential suspect residences. For an incident location on the periphery, the number of suspects will generally be smaller portion of the suspect pool for any given travel distance and the distance range larger over the whole suspect pool as compared to a location in the center of the city. This will give the peripheral incident a larger medial distance than an incident location at the city center. The fact that neighborhoods

Figure 3. Map showing the two extreme data sets partitioned within the distance-decay curve scatter plot in Figure 2 (core set - circles; periphery set - stars; CBD - central business district)



with high suspect densities often asymmetrically occur closer to the city center may accentuate DDC filter effectiveness differences between the core and specific peripheral locations. It is important to reemphasize that this is a boundary effect inherent in the distance-decay nature of offender residences around any fixed area and is not due to lack of data beyond the boundaries of that fixed area.

A polynomial curve was fitted to the core and periphery data sets. The core curve crossed the random list placement line at about 3,000 meters. The periphery curve crossed at about 5,500 meters. Relative to the core locations in Wolverhampton, the DDC filter can be used at peripheral locations to distances 83% greater with some expectation of benefit. The form of the DDC curve, the shape of the aggregating spatial unit containing all of the incidents and the nature of offender residence distribution will determine the degree of core-periphery effectiveness difference when using the standard DDC filter.

Domain Structure Effects

Experience from many disciplines has found that making the assumption of a homogenous distribution for virtually any characteristic over a large area is very risky.

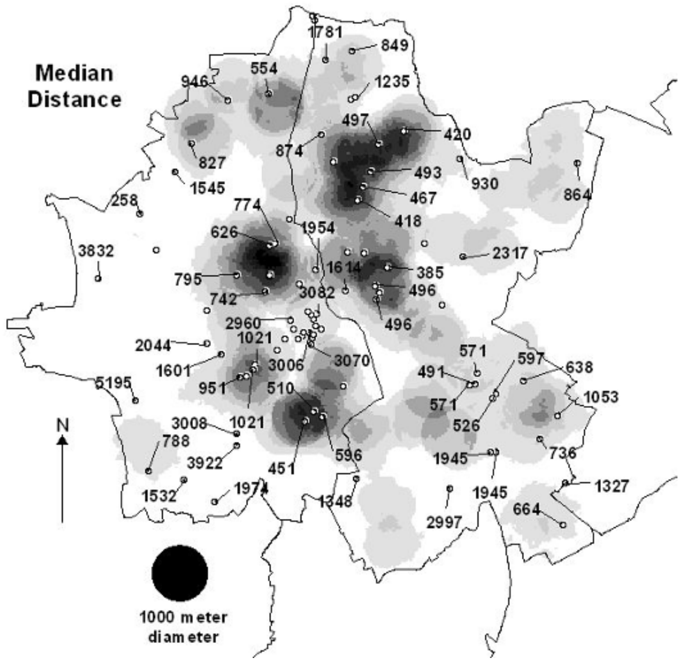
Often applications using the standard DDC assume that this curve or an appropriately warped version of an idealized DDC curve can be used in connection with a specific incident location to assign probabilities to each of the different potential J2C distances. In the case of serial offender applications (an in-depth review of the literature and history of DDC usage is contained in Rossmo, 2000, particularly, pp. 197-201), this procedure is repeated at each of the crime sites. The continuous band of rings around each crime site can be visualized as the ripples around a stone's impact on quiet water. In the crime application, the circular bands will be of different heights reflecting the probabilities derived from the DDC. Every location in the area can be assigned a probability for each crime based on the probability value of the ring that crosses it. The sum of these values at any given location for all the intersecting rings (one ring for each crime location) determines its locational score for that series of crimes. The locations with the highest scores are presumed to be the most likely places from which the offender begins the journey to crime. The Rigel application of these concepts appears to assume an idealized DDC for all serial offenders. The only difference between offenders is the actual distance values for the various percentile totals along this curve. A distance histogram can be computed from every possible location for the crime series. That location giving the best fit to the idealized curve stretched to fit the series is the more likely location for the offender's residence. Rigel's success for any particular serial offender would seem to be determined by how closely the shape of their actual DDC approaches Rigel's idealized curve.

An assumption often made in other applications is that the DDC generated from information aggregated over the whole jurisdiction is applicable to subareas of that jurisdiction. This implies that the offender population is homogeneously distributed with respect to their J2C distance preferences. This uniform application of the standard DDC is rather curious and seemingly atheoretical. It is often postulated that the geographic patterns of individual criminal behavior reflects both the increasing time cost with distance and local geographic factors considered by criminal theories, such as routine activity theory (1979) and environmental criminology (Brantingham & Brantingham, 1981), which place emphasis on the unequal distribution of criminal opportunity. The assumption that DDCs can be locally applied uniformly over the whole jurisdiction from which they were derived would seem to be saying in effect that all conducive or opportune crime locations are so ubiquitous within a criminal's mental map that locational choices are made for reasons not influenced by what effectively has become a geography of uniform opportunity. Geographic control crime theories would seem to predict the expectation of significant variations in local J2C characteristics reflective of the heterogeneous geographic patchwork of functionally controlled land use and uneven distribution of offender residence clusters. In order to test for uniformity in local J2C distance behavior, a local measure of actual J2C distance preferences, the incident-based distance-decay curve (IBDDC, derivation described in the Introduction), is introduced.

IBDDCs were generated for all 101 incidents in the study. A thorough statistical study comparing each of these curves to the standard distance-decay curve to determine the existence and degree of significant difference has not been undertaken. A preliminary rough measure of difference has been employed which compares the median values of the curves. A large difference in medial values is evidence for concluding the curves are different; no difference is not necessarily evidence that they are equivalent. The median J2C distance for the standard DDC was 1,654 meters. The median values from the IBDDCs exhibited a large range, varying from 258 to 5,195 meters. In addition, the spatial distribution of these median values (Figure 4) appears not to be random. Distinct medial J2C distance domains appear to exist within Wolverhampton. A domain is a relatively homogeneous subarea with regard to the variable or variables under examination. The shaded surface in Figure 4 is the density map of all year 2001 offender's residences. Note that those neighborhoods with high offender densities tend to be associated with lower and more consistent medial J2C distances then seen for the city center or other residential areas. These neighborhoods appear to define distinct J2C distance domains. These high offender-density areas that appear to define a specific domain show variations less than 200 meters. Collectively, these high offender residential neighborhoods have average J2C distance values that range from about 300 to 1000 meters. The lower J2C distances for those neighborhoods corroborates the findings of Rengert and Wasilchick (2000) where they noted that burglars tended not to travel into neighborhoods that were perceived to contain gang populations. This study suggests that their observation may be extended to an aggregation of all offense types. The central business district (labeled CBD on Figure 3) is also a well-defined domain having significantly higher values (around 3,000 meters) than the residential neighborhoods with high offender populations. There is also some suggestion from the spatial distribution (Figure 4) that higher value J2C medial distance domains exist in the peripheral areas, but that the domain diversity may be large. A great many more simulations will have to be undertaken in the peripheral areas to delineate their domain structure.

The lack of homogeneity in IBDDC J2C medial distances suggests that some applications using the standard DDC subjects that analysis to a variant of the modifiable areal unit problem (MAUP). In this instance, the analysis assumes that an attribute distribution seen within the larger population contained in a large area of enumeration also applies to the smaller populations contained in areal subdivisions of the larger unit.

Figure 4. Median distances from incident-based distance-decay curves from selected incidents (the underlying density surface is based on all offender residences for incidents committed in 2001)



Effectiveness of Distance-Decay Curves in a Domain-Structured Environment

The existence of J2C distance domains complicates the problem of estimating the effectiveness of a particular filter. The use of a medial circle (contains half of all offender anchors inside and outside) based on the total pool of possible suspect anchor points has proved insufficient alone to estimate how effective a DDC-type filter will be in a local application.

The apparent effectiveness of the standard DDC filter (Table 1), despite the potential problems arising from boundary, suspect residence distributional and domain effects, does not necessarily imply that these factors are of marginal importance. List ordering is a restricted use of the distance-decay concept. As long as we are modeling a steady decline in probability with increasing J2C distance, any steadily declining curve could have been used to order the list. The Wolverhampton DDC showed no lower probability buffer in the interval nearest zero. The first interval (0-250 meters) was the interval containing the mode. Even a straight line with a negative slope would have obtained much the same result. Position on a list is an ordinal metric scale where only relative position is important. The original DDC probability values are ratio scale values, which are effectively converted to ordinal scale values by the process of list ordering. Any DDC application that depends on the actual magnitudes of the DDC empirical probabilities may be materially impacted by the effects enumerated above.

The implications of domain structure can be seen in a local analysis of the important central business district of Wolverhampton. Twenty-five of the simulation incidents fell within this district. This agreed with, depending on boundary definitions, the determination that 21% to 24% of the 7,414 incidents occurred within the CBD. Many of the simulation failures (based on all 101 simulations) to improve a suspect's list placement over a random expectation occurred in the CBD. In this very important district, the standard DDC improved the suspect's list placement on only 11 of the 25 lists. The other two filters performed about the same – the IBDDC produced improvement on 11 lists and the IBORPS on 10 lists. This result indicates that the use of all three filters, on average, somewhat degraded the position of the suspect over what would be expected (improved on 12 to 13 lists) from random lists. Their use in this district would be no better than random and actually slightly detrimental to the investigative process. Regarding the Wolverhampton CBD, we must conclude that the geographic restraints inherent in these three filters have no predictive value that can associate incident location with the offender's residence. This local result clearly contrasts with the overall positive value of filter use over the 101 simulations spread through out the city and indicates the complexities to be expected by the existence of J2C distance domains.

Conclusion

This chapter study has established the following:

1. On average, the three geographic filters analyzed can offer significant improvement in offender placement on suspect lists or in the ordering of mug shots for viewing.
2. A boundary effect exists where distance-decay methods tend to be more effective near the periphery of a typical urban jurisdiction than in the center.
3. Sub-areal domains exist with regard to the medial J2C distances associated with locally defined (IBDDC) distance-decay characteristics. This finding

appears to indicate that significant variations in journey to crime characteristics exist between populations residing in different subareas of the larger jurisdiction. These variations would make use of the standard DDC problematic in sub-areal analysis.

4. Filter effectiveness differences seen between local applications of a particular type of geographic filter indirectly reflect the degree to which the geographic premises inherent in that filter models local criminal behavior.

This chapter study has suggested the following:

1. Additional improvement in offender placement on a suspect list may be possible through a composite metric based on an empirically defined equation combining IBDDC and IBORPS metrics.
2. An analysis of the nature of the distribution and boundary characteristics of J2C distance domains may allow the delineation of neighborhoods based on a crime mobility classification scale. These crime mobility neighborhoods would likely be more reflective of the crime dynamic than the arbitrary census track enumeration units now widely used in analysis.

References

Brantingham, P. L., & Brantingham, P. J. (1981). Notes on the geometry of crime. In P. J. Brantingham & P. L. Brantingham (Eds.), *Environmental criminology (*pp. 27-54). Prospect Heights, IL: Waveland Press.

Brantingham, P. L., & Brantingham, P. J. (1993). Environment, routine and situation: Toward a pattern theory of crime. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice: Advances in criminological theory, Vol. 5.* New Brunswick, NJ: Transaction Publishers.

Brantingham, P. L., & Brantingham, P. J. (1995). Location quotients and crime hot spotsin the city. In C. R. Block, M. Dabdoub & S. Fregly (Eds.), *Crime analysis through computer mapping.* Washington, D.C.: Police Executive Research Form.

Gore, R. Z., & Pattavina, A. (2001, December). *Linking offender residence probability surfaces to a specific incident location: An application for tracking temporal shifts in journey to crime relationships and prioritizing suspect lists and mug shot order* (Paper). Presented at the 5th Annual International Crime Mapping Research Conference, Dallas, TX.

Levine, N. (2000). Chapter 8. In *CrimeStat 1.1 user guide*. Washington, D.C.: National Institute of Justice.

Rengert, G. F. (1981). Burglary in Philadelphia: A critique of an opportunity structure model. In P. L. Brantingham & P. J. Brantingham (Eds.), *Environmental criminology* (pp. 189-201). Prospect Heights, IL: Waveland Press.

Rengert, G. F., Piquero, A. R., & Jones, P. R. (1999). Distance decay re-examined. *Criminology*, *37*(2), 427-445.

Rengert, G. F., & Wasilchick, J. (2000). *Suburban burglary: A tale of two suburbs* (2nd ed.). Charles C. Thomas Publishers.

Rossmo, D. K. (2000). *Geographic profiling*. Boca Raton, FL: CRC Press.

Smith, T.S. (1976). Inverse distance variations for the flow of crime in urban areas. *Social Forces*, *35*(4), 802-815.

Turner, F. J., & Weiss, L. E. (1963). *Structural analysis of metamorphic tectonites*. McGraw-Hill.

Van Koppen, P. J., & De Keijser, J. W. (1997). Desisting distance decay: On the aggregation of individual crime trips. *Criminology*, *35*(2), 505-513.

Chapter VIII

Geographic Profiling and Spatial Analysis of Serial Homicides

Sunghoon Roh, Sam Houston State University, USA

Mark R. Leipnik, Sam Houston State University, USA

Abstract

The characteristics of the crime of serial homicide are *examined, the concept of geographic profiling is presented, and the application of geographic information system (GIS)-based spatial analysis techniques to the investigation of serial homicide investigation are discussed in this chapter. The case of serial killer Robert Yates, in which geospatial technologies played a prominent part, is detailed.*

Introduction

Of all the many crimes committed in our diverse and all-too-often-violent society, none elicits more attention from both the public and the police than homicide. This is particularly true of the nefarious activities of multiple, mass and serial murderers. Serial killings titillate the public, fixate the media, and complicate the task of law enforcement, not the least because they may involve crimes committed in multiple jurisdictions. The recent case of the Washington, D.C., area snipers perfectly illustrates these phenomena, including the challenges posed to law enforcement by offenders that commit crimes in multiple jurisdictions. A large literature has developed around the theoretical under-pinnings of serial homicide investigation, in particular the topic of profiling (Turvey, 1999). More recently, a concept termed geographic profiling has been propounded (Rossmo, 2000). GIS have been helpfully applied to the analysis of a wide variety of crimes (Leipnik & Albert, 2003), but analysis of serial homicides covering a wide geographic area highlight several important advantages of the technology. These include the use of GIS to perform spatial and geo-statistical analysis, to portray multiple features in co-registered layers, to show the spatial proximity of features in the community such as areas of potential offender residence, access and/or concealment to crime incident locations, and not least, to display spatial data for areas extending beyond the jurisdictional boundaries of a single city, county, state or even a single nation. In addition to GIS per se, related technologies such as global positioning systems (GPS) and, to a lesser extent, digital aerial photography are figuring more frequently of late in serial homicide investigations (Leipnik & Albert, 2003).

Background: Natural History of Serial Killers

Homicide takes a number of diverse forms, ranging from reckless and drunken driving to domestic and alcohol-fueled disputes, to murder for gain, and finally, to not only premeditated crimes, but to multiple murders where murder is the primary objective of the killer's life. These latter crimes are universally regarded as among the worst of offenses. Even among premeditated killings of multiple individuals there is a range of behaviors exhibited. Thus one can differentiate among mass murders, sequential murders of known victims and serial killings of strangers. Mass murders tend to be short duration outbursts of often long-suppressed rage. Usually, solving mass murders is straightforward for police with the by-then suicidal killer frequently being the last victim of the spree. Multiple murders of victims well known to the killer include the classic crime of poisoning relatives and the increasing number of killings by health professionals. The Dr. Harold Shipman case in Hyde, England, may be the most egregious example. Even ascertaining the number of victims in such cases is problematic.

For example, Dr. Shipman was convicted of 15 killings, but suspected of between 215 and 260 killings (Hill, 2004).

Then there is the crime of serial killing of strangers. The most prolific murderers (except for war criminals and health professionals) tend to be in this last category and include such infamous cases as Jack the Ripper, Ted Bundy and Gary Ridgeway (Keppel, 1995). Even within this relatively rare subset of killers, one can differentiate organized and disorganized types. Organized serial killers are methodical, carefully planning their crimes, selecting vulnerable victims and relishing their activities. The Robert Yates case, which is detailed below, fits this mold, as does Clifford Olson's activities. Clifford Olson killed 11 victims over the course of less than nine months, making him Canada's second most prolific serial killer. Olson had a plan to attract his young female victims: he had a drugged drink ready and a body disposal location selected in advance of having a victim. Spatial analysis of his activities indicates proximity of his activities to his home (Mulgrew, 1990). Both Olson and Yates preyed on vulnerable women, including runaways and what has been a victim since at least the 19th century: the sex worker.

Conversely, disorganized serial killers (and most serial rapists) tend to attack impetuously. Something about the victim induces an irresistible urge to rape, maim or kill. Thus, in the most infamous Korean serial killing case (which is at present unsolved), 10 women all wearing red clothes were attacked in the open in the rural Hwa-Sung region from 1986 to 1991. In general, disorganized serial killers are easier to catch unless the spree stops, because they also have disorganized mental processes and lack the careful selection of victims seen in organized serial killers.

The Terry Driver case is an example of the behavior of a mixed category sexual killer. Terry Driver attacked two teenage girls, murdered one, and then engaged in bizarre activities such as calling the police, writing letters, and defacing his victim's tombstone. GIS-based geographic profiling using the Rigel software program from Environmental Criminology Research Inc. in Vancouver, British Columbia, and analysis of digital aerial photography was employed in this case. It is worth noting in passing that these categories are somewhat artificial, as demonstrated by the change in *modus operandi* (MO) of Ted Bundy from a tentative killer to an extremely careful, organized serial killer to a maniac who went on a murderous rampage in a sorority house, a behavior more likely to be associated with a disorganized killer or even a mass murderer.

The behavior of serial killers can also be examined in terms of its evolution over the course of time and its patterns in space. Many serial killers begin with minor crimes committed close to home, like the torture of animals and sexual exploitation of playmates (Edgar, 2002). Some serial killers, like Jeffery Dahmer, begin by killing at a young age. Dahmer committed at his first murder at home and all of his later victims were lured into his apartment (Baumann, 1991). Even though the initial killing is rarely in the home, it is frequently in very close proximity to the home or workplace. These are areas with which the killer is familiar – a "home range" – and as they enter the tantalizing underworld of murder, many serial killers seek the comfort of familiar surroundings. Thus Ted Bundy first killed near his Seattle, Washington, home and Andrei Chikatilo committed the first of his 53 murders within a few yards of his home in Rostov, Ukraine. As serial killers become more comfortable in successfully getting away with their crimes, they tend to expand their range both spatially and in terms of the selection of victims. Ted Bundy initially only killed young women, but eventually he began to murder children. He expanded his scope of operations from Seattle to western Washington state, the to the whole Pacific Northwest region and eventually his field of operations expanded to the point where, while he studied law in Salt Lake City, Utah, he took "hunting expeditions" to Idaho and western Colorado. After next fleeing on to the greener pastures of Florida, he was finally captured and confessed to 28 murders over a five-year span (although he is suspected in as many as 100 killings). Andrei Chikatilo expanded beyond young women (first to girls and then to young boys). His first killing was in his neighborhood and he was even a suspect in the crime; soon he progressed to killings in proximity to the rail system in Rostov (linked to his employment with a rail equipment supplier rather than to tied to his home, but still familiar ground), and then to a larger area of the Ukraine over a 14-year span. In a somewhat similar case from a differing culture and economic system, Xihai Yang employed a bicycle to travel across four eastern Chinese provinces between 1999 and 2003, where this itinerant construction worker used a series of job site tools to kill 67 victims prior to his recent arrest. Another example of a spatial progression can be seen in the crimes of serial sniper John Allen Mohammed and his accomplice, Lee Malvo. His first killing was linked to his ex-wife and was committed in the Tacoma, Washington, home of his ex-wife's friend. Then there was a killing in Baton Rouge, Louisiana, where he had relatives and was familiar with the area. The pair (led by Mohammed) then progressed to murders in Birmingham, Alabama, before their activities metastasized into an unprecedented series of sniper attacks in the Washington, D.C., area. Even within this later series of 11 shootings, one can see a spatial progression from a limited area (Prince Williams County, Maryland) with which Mohammed was somewhat familiar, to a roving hunt over an ever-expanding swath of two states and the District of Columbia. This case also illustrates two important characteristics of serial murder progressions: one is that the crimes can become more brazen as the killer or killers gain confidence, and secondly, that the crimes may become simultaneously harder to solve (Horwitz, 2003).

The reason that serial killings may be harder to solve as the series progresses is that serial killers are likely to strike farther affield and evolve strategies to thwart detection, such as dumping bodies in bodies of water, moving from one area to another, or choosing a broader range of victims. Some particularly cunning killers like Unabomber Ted Kaczynski can even deliberately leave false clues as to their base of operations, relocate to an isolated area, and interrupt their activities for a long hiatus until the trail grows cold (Waits, 1999). Examples of this peripatetic behavior include Christopher Wilder, the Australian millionaire that traveled across the United States leaving a string of killings in his wake (Gibney, 1990); and Kenneth McDuff, the Texas killer who struck in numerous separate counties, but never more than once in any county and also prepared hidden graves before he selected victims (Lavergne, 1999).

Of course, one can argue that those killers who succeed in carrying out extended killing series are the ones who adopt these measures and are not exemplars of the evolution of a perfected technique. At any rate, an important principle for investigators is that it is extremely valuable to focus on the early crimes in a murder sequence since these may yield more valuable clues about the "home range" and proclivities of the killer, than would the freshest case in a given series, which often garners disproportionate attention. By the same token, it would appear that careful investigation of any unsolved killing may help preclude the escalation of the crimes into serial murder. If Tacoma Police had carefully investigated the seemingly random killing of Keenya Cook (John Allen Mohammed's first victim), the Washington, D.C., sniper attacks might never have occurred. If Soviet Ministry of Interior apparatchiks had not so quickly arrested a neighbor of the first victim of Andre Chikatilo and then physically coerced a confession from him, the deaths of 52 more victims of the "Red Ripper," along with the executions of two innocent men found wrongly guilty for Chikatilo's crimes, could have been averted (Lourie, 1994).

Geographic Profiling and Investigation of Serial Homicides

Keppel and Weis (1994), for example, found that a murder case is more likely to be solved when investigators analyze detailed spatial information. The spatial information associated with crime has been regarded as one of most crucial elements in criminal investigation. Investigation of serial homicides is particularly likely to be thorough and often have spatial aspects that may benefit from detailed analysis. Such spatial information consists of the place where the victim was last seen, the location where the offender and the victim initially may have encountered each other, the crime scene (if known), and the place where the body was recovered. Other locations may also be included. In the Terry Driver case in Abbotsford, British Columbia, the location of the phone booths from which taunting calls to police were made and the location of the defaced tombstone were used in the geographic analysis. Law enforcement agencies may use the information to discover the criminal's identity, the probable offender's residence or the potential place of employment (Canter, 1994; Rossmo, 2000).

The rationale for the relationship between spatial characteristics and crime investigation of serial homicides is grounded upon the assumption that criminals do not choose crime locations at random. Criminals are usually constrained by many factors when they make decisions regarding where they seek victims, commit crimes, and dispose of the bodies. Even different transportation opportunities caused by criminals' economic conditions may affect their decision making regarding crime locations (Rengert, Piquero & Jones, 1999). *Geographic profiling* is a term coined to describe how various factors concerning crime locations are associated with criminal investigation. Rossmo (2000) defines geographic profiling as "a strategic information management system designed to support serial violent crime investigations" (p. 211). More broadly it is part of "a framework to understanding how an offender searches for victims and the social and the physical environment, the way that the offender understands this environment as well as the offender's motives" (Santtila, Zappala, Laukkanen & Picozzi, 2003, p. 43).

One basis of geographic profiling is the assumption that criminals do not choose crime locations at random and the likelihood of crime commission can be partially determined based on a serial killer's home location. According to the "circle hypothesis," a serial killer's home is located within a circle formed by the offender's two farthest offenses (Canter and Larkin, 1993; Canter and Gregory, 1994). This indicates that crime locations are most likely to be found within the area with which serial killers are familiar, and therefore, in which they may feel more comfortable. This "home range" or "activity space" can cover a large or small area depending on factors such as the criminal's mobility and cunning. In some cases, such as that of the Knoxville, Tennessee, "Greenbelt" serial rapist, each of the three attacks were within easy walking distance of the offender's home. In other cases, such as that of Robert Yates, where a Corvette was used for collection and transport of victims, greater mobility was exhibited (Leipnik & Albert, 2003). The probability of crime commission is also related to the distance between an offender's home and the crime location. Researchers found that the greater the distance, the less likely an offender will commit a crime there. This finding is termed the "distance-decay function," indicating that offenders have a tendency to select their victims in the regions near their residence (Turner, 1969; Van Koppen and De Keijser, 1997). Within the geographic area closest to an offender's home, however, there is an area where crime commission may be less likely to occur, termed a "buffer zone" (Brantingham and Brantingham, 1981; Turner, 1969). An offender is less likely to a commit crime within the area very near to their home for fear of being easily detected. Given the principle that an offender tends to commit crime near to his (or her, although except for Aileen Wournos, very few serial killers are women) residence (*the distance-decay function*) and the tendency to maintain a certain minimum distance away from the home (*the buffer zone*), it may be possible to narrow down the area (in theory, a sort of donut) where the offender's home is most likely to be located.

The "home-base" principle, however, may not be applied to every serial homicide case. Obsessed organized serial killers, who often prepare a carefully designed crime plan, commit crimes further away from home rather than simply choose the victim from the same neighborhood in which they reside. Some serial killers separate murder locations intentionally in order to hamper the investigation. The separation of locations inhibits an effective investigation not only because it will take more time for police agencies to discover each body from disparate locations, but also because multiple involved jurisdictions may have problems with communication and cooperation in performing an interjurisdictional investigation (Keppel & Weis, 1994).

Homes and Homes (1994) distinguish two types of serial murderers based on their spatial mobility and stability. "Geographically transient serial murderers" commit murders in geographically different areas. They travel from one area to another for the primary purpose of killing people (Ted Bundy or Kenneth McDuff would be examples). In contrast, "geographically stable serial murderers" stick to familiar areas near their homes (or in their homes, as in the cases of Jefferey Dahmer and John Wayne Gacy). In accordance with the home-base principle, the selection of victims and commission of murders are accomplished close to home.

Godwin and Canter (1997) added a notable elaboration to the somewhat simplistic home-base principle. Differentiating the location of an initial contact between the murderer and the victim and the body disposal site, they found that offenders tend to place the victims' bodies further away from their homes than the initial contact location. This behavior was observed in the Robert Yates case. Offenders may encounter victims in the area where they are normally involved with routine activities or they may hunt for victims in an area where potential victims congregate. Thus many serial killers like Robert Yates or Gary Ridgeway seek out an area such as a "red light" district where their victims of choice (sex workers) are readily available. After the offenders abduct and kill the victims, however, they tend to dispose of the bodies farther away from the areas that they frequent. This may be partially due to the fear of being observed during the time-consuming process of disposing of a body in a public place. Thus rural and unfrequented wastelands and places like forested tracts, public parks, brownfield sites, logging roads and riparian zones are favored.

Robert Yates Case: Spatial Analysis of Serial Murder

Organized serial killers are among the most frightening of all criminals. They are frequently cunning predators, striking when and where they choose, and often employing elaborate precautions to preclude detection. Ted Bundy, Clifford Olson and Gary Ridgeway exemplify these sociopathic monsters. Robert Yates fits this mold, and GIS-based spatial analysis helped to provide important clues to locations where he might be found, and also, where he may have sought additional victims and discarded his victims. As in many murder series, the earliest two crimes in the series of brutal murders committed by Yates (which were committed in Walla Walla, Washington, in the 1970s) could have provided investigators with valuable clues, had they been recognized as part of a series, as they were committed before Yates had expanded his "horizons" and sophistication. Unfortunately, these early crimes were not recognized as belonging to the series until after the apprehension of Yates. Rather, it was 10 murders with the same MO committed in Spokane, Washington, beginning in 1996 that aroused the keen attention of law enforcement authorities.

Spokane is the major metropolitan area of eastern Washington State. Spokane County covers an area of 1,800 square miles, 10% of which is an urban center along the banks of the Spokane River. The region boasts a population of approximately 500,000 and includes wheat farms and rugged wilderness.

Beginning in 1996, bodies of murdered female sex workers began to be discovered in outlying areas around Spokane. It quickly became evident to local and regional law enforcement that a serial killer was active. In addition to the discovered victims, several women missing from the region were suspected of being victims of the serial killer. Ultimately, the bodies of 10 victims were recovered and linked by investigators to the same series. Another victim was discovered buried in the killer's garden outside the bedroom window after he was apprehended.

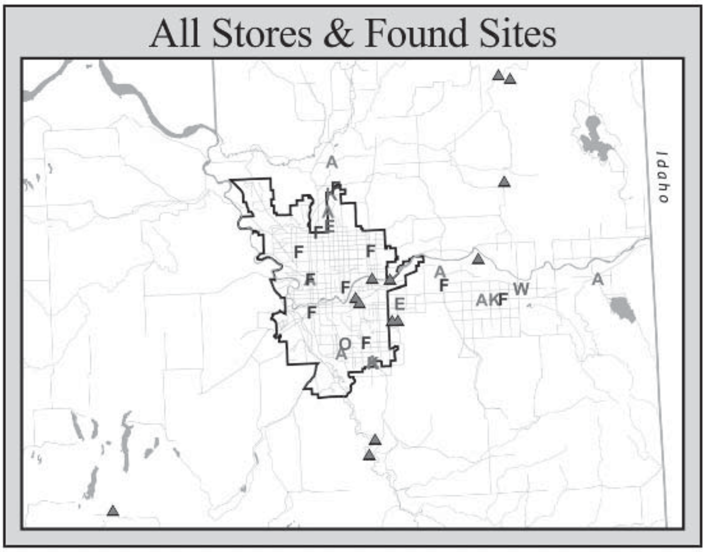
The most important clue for investigators was the ballistics evidence, which indicated that each victim had been shot in the head with the same .38 caliber pistol. Another vital clue was that the head and occasionally the hands of the victims were wrapped in plastic supermarket shopping bags. This second clue showed that the serial killer was highly organized, since the bags were being employed to prevent DNA evidence, such as blood, from contaminating the killer's vehicle, and also to minimize the possibility of fiber evidence being left on the bodies. Adoption of such precautions showed that a sophisticated and methodical killer was at work. Another factor apparent from the outset was that cooperation between municipal and regional law enforcement would be necessary. While most of the victims had presumably plied their trade in the small but active Sprague Street "Red Light" district in the urban core of Spokane, many of the bodies were recovered from areas of Spokane County outside the city limits in the jurisdiction of the Spokane County Sheriff's Office.

A joint city-county serial homicide task force was established to investigate the killings and possible related disappearances. Given the wide spatial distribution of sites at which the bodies were being found and the clear implication that a private vehicle was being employed both to pick up the victims and then to convey them to the disposal sites, the serial homicide task force decided to use GIS to map these disparate locations and then employ spatial analysis to try to narrow the focus of their investigations. The team utilized the Spokane County GIS Management Office's Arc/Info-based GIS, which was then running on UNIX workstations, along with a number of related Environmental Systems Research Institute (ESRI) products, such as ArcView. The coverages utilized by analysts included road networks with address ranges, city boundaries and topographic and hydrographic data layers.

Evidence from the shopping bags provided an inadvertent, though significant, spatial clue to investigators. Each shopping bag was imprinted with the name of the store, which had sold the merchandise it had once contained. It was hypothesized by investigators that while the locations that bodies were disposed of might provide some insights into the "home range" of the killer, these locations tended to be ones where the bodies could be discarded or buried unobserved. On the other hand, the shopping bags provided clues to the locations where the killer (or possibly the killer's spouse) regularly shopped. Eighteen shopping bags from five companies were recovered. Shopping bags from Albertsons (A), Eagle (E), Safeway (F), ShopKo (K), Super-One (O) and Wal-Mar t(W) stores were recovered by investigators. Several of these companies had multiple outlets in the Spokane area, but interestingly, Super-One had only one store in the area and Eagle had only two outlets. The GIS analysts geocoded the locations where the bodies were recovered and the locations of the stores from which the bags may have originated, using the street centerline layer.

The team of investigators from the serial homicide task force and spatial analysis experts from the Spokane County GIS Management Office were faced with the issue of how to use the spatial clues inherent in the locations of stores and body disposal locations to help narrow the focus of the investigation geographically. Study of the behavior of serial criminals indicates that they have an activity space

Figure 1. Map showing body disposal locations and store locations (letters A, E, F, K, O & W are abbreviations of store names listed on previous page)

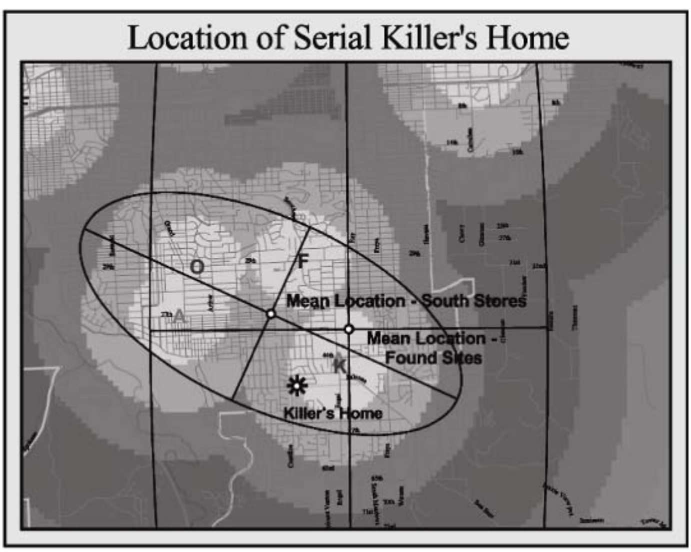


in which they feel a higher comfort level in committing crimes. This activity space often includes their home, their workplace, areas they travel through and areas they frequent in the course of shopping and recreational activities. Also, analysis of the behavior of numerous serial killers indicates that crimes early in the series are likely to provide the best clues, both because they are likely to be more amateurish and because they are likely to be committed within a more geographically limited activity space, in other words, closer to home. These tendencies are reflected in behavior patterns that can be analyzed spatially with GIS, based on principles of geographic profiling, to yield an understanding of spatial and temporal factors that influence aberrant behavior. On May 12, 1998, Dr. Kim Rossmo made a presentation highlighting use of spatial analysis to perform geographic profiling in cases of serial homicides to the Washington Association of Sheriffs and Police Chiefs in Spokane. This presentation appears to have prompted the Federal Bureau of Investigation Special Agent in Charge to recommend the use of spatial analysis in the case of the Spokane-area murders.

In order to use spatial analysis to gain insights into the behavior of the Spokane-area serial killer, the Spokane serial homicide task force opted to employ the existing GIS capabilities of the County GIS Management Office, but supplement them with specialized modeling software. They utilized ESRI's Arc/Info GIS along with an ArcView extension called the animal movements model, which was developed by the United States Geological Survey's Alaska Field Office. This model uses spatial analysis to estimate the home range of large animals, such as caribou or grizzly bears. It is, indeed, ironic that a geospatial model originally intended to study animals can be applied to the study of predatory and animalistic humans. The animal movements model extension was utilized to generate probability ellipses for the likely "home range" of the serial killer based on body disposal locations and on the location of all stores from which bags found at murder scenes might have been obtained.

Inferences drawn from the spatial analysis of the distribution of body disposal locations and store locations yielded probability ellipses that, while having differing orientation of their axis, were both centered in the same neighborhood of Spokane. This was the upper middle class South Hill neighborhood. This analysis helped focus interest on this area. Other evidence obtained from interviews of acquaintances of the victims led investigators to link the likely suspect to a white Chevrolet Corvette. Investigators eventually tracked down a white Corvette previously belonging to Robert Yates. Mr. Yates was a retired 18-year U.S. Army veteran, a middle-class metal industry crane operator and Army Reserve helicopter pilot/instructor. Yates resided with his wife and children in a well-maintained home in the comfortable South Hill neighborhood, as predicted by the spatial analysis. It is noteworthy that the area within the larger

Figure 2. Zoomed map of probability ellipses for selected group of stores with killer's home location highlighted



probability ellipse generated from the body disposal locations covers approximately eight square miles, while the intersecting probability ellipse for the selected store locations covers only four square miles. The area where both ellipses intersect occupies less than one square mile. It is in this small area where Robert Yates' residence is located. This square mile area represents less than one-tenth of 1% of the total area within Spokane County and is home to less than 1% of the region's population.

DNA evidence recovered from the Corvette linked him to one murder victim, and the body of another victim was subsequently discovered buried on his property. Faced with this overwhelming evidence of complicity in at least two killings, he confessed to 13 murders: two earlier crimes in Walla Walla, Washington, one in Skagit County, Washington, and 10 murders in Spokane. In 1998, he was sentenced to 408 years in prison.

If this had been the end of the story, the spatial analysis and GIS-based mapping that was performed would be an affirmation of the value of this technology in serial murder investigations. However, the case has an added wrinkle that again called on the expertise of the GIS analysts at the Spokane County GIS Management Office to help in its elucidation. It turns out that after his arrest, Robert Yates was found to have possessed a Magellan GPS unit, and stored in its memory were 72 waypoints scattered around Washington state. Given the methodical nature of this killer, and his clear obsession with hunting female victims, investigators theorized that the waypoints might provide spatial and temporal clues to Yates' criminal activities. The mapping of the waypoints was a laborious undertaking, since Yates' Magellan was a "recreational-grade" unit. A statewide base map containing features of possible relevance such as roads, hydrography and political boundaries was generated and used to portray the coordinates of each waypoint in context. Waypoints can help determine where and when a suspect was using the GPS unit, and this in turn can be linked in time and space to actual or potential crimes. Investigators used the extracted coordinates and related maps in ground searches.

Self-incrimination by GPS data, although highly unusual, is not unprecedented. Canada's most-wanted fugitive, a financier named Albert Walker, was linked to a homicide in Great Britain by Scotland Yard investigators who extracted coordinates from a marine GPS unit on a yacht owned by the suspect. The coordinates extracted from this unit plotted out the course of a cruise made by the suspect's yacht. The boat had spent time at a point in the English Channel where, subsequently, a fisherman had snagged the weighted body of a murdered man. A Rolex watch on the body eventually led sleuths, with the help of Interpol, to identify the victim as the man whose identity Albert Walker had assumed to evade arrest by authorities seeking him. In another Spokane County case, Brad Jackson was tracked to the location along an isolated logging road where he had reburied the body of his suffocated daughter (Leipnik, Bottelli, Von Essen, Schmidt, Anderson & Cooper, 2001). A Silent Position Monitor 2000 wireless internet GPS-based tracking device had been attached magnetically to the undercarriage of his pick-up truck and his movements in the weeks following her disappearance were monitored and mapped using GIS. Such GPS-based tracking and mapping of a homicide suspect's movements was also a feature of the much-publicized Laci Peterson murder case in Modesto, California.

Of particular note, in the map of waypoints retrieved from Yates' GPS unit were several waypoints left by visits to Pierce County, Washington, where Fort Lewis and Tacoma are located. These waypoints were an important verification of the presence of Yates in the vicinity of Tacoma during the time period when two sex workers were murdered there. This evidence, along with other information subsequently developed, was used to help convict Robert Yates of two additional murders in a trial during the spring of 2001. In these cases, the defendant received a death sentence. The imposition of these sentences will put a definitive end to the case of Robert Yates. But the employment of GIS for spatial analysis of his activities and the utilization of GPS data in the case is likely to presage more widespread adoption of geospatial technologies in bringing serial murders to book, or rather, to map.

Acknowledgments

The co-authors would like to acknowledge first and foremost the assistance of Kim Rossmo, currently on the faculty of Criminal Justice at Texas State University in San Marcos, and of Ian Von Essen with the Spokane County GIS Management Office. Also, thanks to John Bottelli, Ariane Schmidt, Laurie Anderson and Tony Cooper, all currently or formerly associated with government agencies in the Spokane, Washington, area that were involved in the spatial analysis discussed in this chapter.

References

Barer, B. (2002). *Body count*. New York: Pinnacle.

Baumann, E. (1991). *Step into my parlor: The chilling story of serial killer Jeffrey Dahmer*. Chicago: Bonus Books.

Brantingham, P. J., & Brantingham, P. L. (1981). *Environmental criminology*. Prospect Heights, IL: Waveland Press.

Canter, D. (1994). *Criminal shadows*. London: Harper-Collins.

Canter, D., & Gregory, A. (1994). Identifying the residential location of rapists. *Journal of the Forensic Science Society, 34*, 169-75.

Canter, D. V., & Larkin, P. (1993). The environmental range of serial rapists. *Journal of Environmental Psychology, 13*, 63-69.

Cohen, L., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review, 44*, 588-608.

Cornish, D., & Clarke, R. (1986). *The reasoning criminal: Rational choice perspectives on offending*. New York: Springer-Verlag.

Edgar, S. (2002). *The killers among us: An examination of serial murder and its investigation*. Upper Saddle River, NJ: Prentice Hall.

Gibney, B. (1990). *The beauty queen killer*. New York: Pinnacle.

Godwin, M., & Canter, D. (1997). Encounter and death: The spatial behavior of U.S. serial killers. *Policing: An International Journal of Police Strategy and Management*, *20*(1), 24-38.

Hill, M. (2004). *Addicted to murder: The true story of Dr. Harold Shipman*. London: Virgin Press.

Holmes, R., & Holmes, T. (1994). *Murder in America.* Thousand Oaks, CA: Sage.

Horwitz, S. (2003). *Sniper: Inside the hunt for the killers who terrorized the nation.* New York: Random House.

Keppel, R. (1995). *The riverman: Ted Bundy and I hunt the Green River killer*. New York: Pocket Books.

Keppel, R., & Weis, J. (1994). Time and distance as solvability factors in murder cases. *Journal of Forensic Sciences*, *39*(2), 386-401.

Lavergne, G. (1999). *Bad boy from Rosebud: the murderous life of Kenneth Allen McDuff*. Denton, TX: University of North Texas Press.

Leipnik, M., & Albert, D. (2003). *GIS in law enforcement.* London: Taylor and Francis.

Leipnik, M., Bottelli, J., Von Essen, I., Schmidt, A., Anderson, L., & Cooper, T. (2001). Coordinates of a killer. *Geospatial Solutions Magazine*.

Lourie, R. (1994). *Hunting the devil.* New York: Harper-Collins.

Mulgrew, I. (1990). *Final payoff: The true cost of convicting Clifford Robert Olson*. Toronto: Seal Books.

Rengert, G., Piquero, A., & Jones, P. (1999). Distance decay reexamined. *Criminology*, *37*(2), 427-445.

Rossmo, K. (1999). *Geographic profiling.* Boca Raton, FL: CRC Press.

Santtila, P., Zappala, A., Laukkanen, M., & Picozzi, M. (2002). Testing the utility of a geographical profiling approach in three rape series of a single offender: A case study. *Forensic Science International*, *131*, 42-52.

Turner, S. (1969). Delinquency and distance. In M. E. Wolfgang & T . Sellin (Eds.), *Delinquency: Selected studies* (pp. 11-26). New York: John C. Wiley.

Turvey, B. (1999). *Criminal profiling*. London: Academic Press.

Van Koppen, P., & De Keijser, J. (1997). Desisting distance decay: On the aggregation of individual crime trips. *Criminology, 35*(3), 505-515.

Waits, C., & Shors, D. (1999) *Unabomber: The secret life of Ted Kaczynski*. Helena, MT: Farcountry.

Section IV

Crime Monitoring and Tracking

Chapter IX

Geographic Surveillance of Crime Frequencies in Small Areas

Peter A. Rogerson, University at Buffalo, USA

Abstract

In this chapter, I describe a system for monitoring crime frequencies for a set of small areas. The objective is to detect as quickly as possible any increase in any area's crime frequency, relative to a specified expected frequency. The system uses a cumulative sum approach, cumulating differences between the observed and expected frequencies of crime in each area. The approach is illustrated using 1996 burglary data from Buffalo, which is available by census tract. Computer code associated with the geosurveillance program is provided in the appendix.

Introduction

An important aspect of crime analysis is the daily, weekly and monthly monitoring of crime reports. Quick detection of significant increases in the frequency of crime and/or changes in the geographic pattern of crime can lead to efficient reallocations of enforcement effort. There may also be interest in detecting any *decreases* in crime frequencies that may, for example, result from crackdown efforts. Rogerson and Sun (2001) provide an example of such monitoring in the context of crime analysis.

In this chapter, I describe a simple system for monitoring crime frequencies in a set of small geographic areas. The purpose of the system is to signal to the crime analyst any significant changes in the frequency of crime that may occur in any of the small areas. Crime frequencies will exhibit natural fluctuations over time, and it is important to be able to distinguish significant changes in the underlying rate of crime from these natural fluctuations. The statistical detection of a significant increase in burglaries could imply, for example, that criminals had recently targeted the area; quick reaction to the statistical increase could lead to added patrols and, ultimately, to either arrests or declines in criminal activity.

The geosurveillance system is based upon methods developed in industry for the quick detection of faults in manufacturing processes (see, for example, Montgomery, 1996; Wetherill & Brown, 1991; Ryan, 1989). It makes use of the *cumulative sum* of the excess of observed frequencies over some expected, baseline frequency. If the cumulative sum of these excesses reaches some critical value, an “alarm” is sounded, and the analyst then investigates the nature of the change. In the criminological context, the alarm would imply that the frequency of criminal activity had increased by a statistically significant amount. The increase presumably would be worth investigating – it might ultimately be traced to some known cause (for example, there may have been a long spell of hot, dry weather; weather conditions are known to be correlated with criminal activity). But the cause also might be due to one or more criminals who have begun to operate in the area. Quick detection could conceivably lead to quicker arrests.

Section 2 describes the monitoring system and section 3 provides an illustration using burglary data from Buffalo, New York. The final section contains a discussion of potential extensions of the method, as well as a summary of its limitations. Computer code is included in the appendix.

Geographic Surveillance of Crime Frequencies in Small Areas

Suppose that we wish to monitor crime frequencies for a set of small geographic areas. A specific goal is the quick detection of increases in the frequency of a particular crime type. We divide the following discussion into situations that involve either small expected frequencies (less than about two or three), and high expected frequencies.

Small Frequencies

The method that follows is closely related to the work of Lucas (1985), who has described how to monitor variables that represent small frequencies; the variables have a Poisson distribution, and these variables are used with the cumulative sum methods employed in quality control.

Let $x_{i,t}$ be the observed number of crimes (of a given type) committed in area i during time period t. Furthermore, let λ_i be the number of crimes that are *expected* in area i (for simplicity we will initially assume that these expectations are not dependent on time). Intuitively, we wish to monitor the degree to which the observed values are exceeding these expected values as information is collected over time. If excesses cumulate sufficiently (and in particular, beyond some threshold value h_i), we will infer that a significant increase in the crime rate has occurred in that area. There are two types of errors that can be made. First, if the threshold is set too low, then "false alarms" will occur, where a significant change is indicated by the system, but where, in fact, no change has actually occurred. If, however, the threshold is set too high, then although the number of false alarms will be minimized, true change in the crime rate may not be detected when it does occur. The choice of a threshold will be discussed later in this section; it represents a balance between these two types of errors.

To implement the cumulative sum system, we calculate:

$$S_{i,t} = \max(0, S_{i,t-1} + x_{i,t} - k_i) \qquad (1)$$

where $S_{i,t}$ is the cumulative sum at time t and where we begin with an initial cumulative sum of $S_{i,0} = 0$ in each subregion. Note that we are cumulating the number of crimes that is in excess of k_i. An intuitive value of k_i would be λ_i, since the difference $x_{i,t} - \lambda_i$ is the difference between what is observed and what is expected. But in practice, an hypothesized alternative frequency is specified; $\lambda_i^A > \lambda_i$ is the increased level of crime that we wish to detect as quickly as possible. To minimize the time it takes to detect a shift in the mean level of crime from λ_i to λ_i^A, an approximate value of k_i that achieves this objective may be determined by taking it to be the average of λ_i and λ_i^A; that is:

$$k_i \approx \frac{\lambda_i + \lambda_i^A}{2} \tag{2}$$

In practice, we often wish to minimize the time it takes to detect a one standard deviation change in the mean, and so this implies that we should use $\lambda_i^A = \lambda_i + \sqrt{\lambda_i}$ when looking for increases in crime frequencies, since the standard deviation of a Poisson variable is equal to the square root of its mean. When this choice is used in Equation (2), we have $k_i \approx \lambda_i + 0.5\sqrt{\lambda_i}$. A somewhat more accurate value of k_i (in the sense that it will lead to slightly quicker detection of changes when they occur) may be determined from the formula:

$$k_i = \frac{\lambda_i^A - \lambda_i}{\ln \lambda_i^A - \ln \lambda_i} \tag{3}$$

An additional, practical adjustment for implementation can be made by rounding the value of k to the nearest integer; when expected frequencies are small, k is often rounded to a fraction.

Finally, we need to specify a threshold for the cumulative sum, h_i, in each region. To do this, one first needs to decide on the *average run length* (ARL_0), which is the average number of observations until a change is signaled, when in fact no change has occurred (that is, ARL_0 is the average number of observations until a false alarm). Then, to find the corresponding value of h_i, one can either (a) use tables such as Table 1 or the table given by Lucas (1985), or (b) a computer program such as the one in the appendix, which is based on the program of White and Keats (1996).

Table 1. Average run lengths for values of λ, h, and k in the Poisson cusum

λ	k \ h	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0.25	0	12	16	21	25	28	32	37	41	45	49	53	57	61	65	69	73	77	81	85	89	93
0.50	1	174	633	2249	7934	27905	98068															
0.75	1	37	76	147	274	496	884	1559	2733	4770	8306	14437	25070	43507	75471							
1.00	1	15	23	34	47	62	78	97	117	140	165	192	220	251	284	318	355	394	434	477	522	568
1.25	2	55	139	345	842	2038	4912	11817	28407	68262												
1.50	2	26	51	97	177	319	565	994	1739	3032	5275	9164	15909	27603	47877	83025						
1.75	2	14	24	38	58	84	120	167	230	313	421	563	749	992	1308	1720	2256	2953	3860	5039	6572	8563
2.00	3	37	85	188	412	894	1927	4144	8896	19087	40939	87790										
2.25	3	22	43	80	145	259	458	804	1404	2446	4253	7387	12821	22242	38575							
2.50	3	14	24	39	63	96	144	214	315	459	665	960	1380	1981	2837	4058	5799	8280	11818	16860	24045	34286
2.75	4	31	66	139	288	591	1208	2461	5004	10165	20640	41901	85053									
3.00	4	20	39	71	129	230	405	710	1240	2158	3752	6514	11304	19609	34007	58968						
3.25	4	14	24	40	65	103	160	246	376	570	860	1295	1946	2918	4373	6547	9798	14657	21920	32776	49002	73252
3.50	4	10	16	25	37	53	74	103	140	190	254	339	449	593	781	1025	1343	1756	2293	2991	3889	5078
3.75	5	20	36	67	120	213	375	657	1146	1994	3464	6014	10435	18100	31388	54425	94362					
4.00	5	25	41	67	108	172	270	422	655	1016	1571	2425	3740	5763	8877	13668	21040	32383	49835	76687		
4.25	5	11	17	27	41	60	88	125	178	250	350	488	678	939	1297	1789	2466	3395	4670	6422	8826	12127
4.50	5	8	13	19	27	37	49	66	86	112	144	184	233	295	370	464	580	723	900	1117	1385	1715
4.75	6	15	25	42	69	112	181	289	459	725	1144	1802	2834	4455	6997	10986	17244	27063	42468	66637		
5.00	6	11	18	29	44	67	99	146	213	309	446	642	921	1320	1888	2699	3855	5502	7850	11196	15965	22761
5.25	6	9	14	20	30	42	59	80	109	147	196	261	345	454	597	783	1025	1340	1748	2280	2970	3867
5.50	7	15	25	43	71	116	189	305	490	785	1255	2002	3192	5085	8097	12890	20516	32651	51957	82674		
5.75	7	12	9	30	47	72	110	165	245	364	539	794	1169	1718	2522	3699	5424	7950	11648	17064	24995	36607
6.00	7	10	15	22	33	47	67	95	132	183	253	347	474	646	879	1194	1620	2196	2974	4026	5447	7367
6.25	7	8	12	17	23	32	44	58	77	100	130	168	215	274	348	441	558	704	886	1115	1401	1758
6.50	8	12	20	32	50	77	119	182	276	417	628	943	1415	2121	3175	4752	7109	10632	15898	23769		
6.75	8	10	16	24	35	52	76	109	156	221	312	439	616	862	1206	1685	2352	3280	4572	6372	8877	12363
7.00	8	8	12	18	26	36	50	69	93	125	166	220	291	383	503	659	862	1126	1469	1915	2495	3247
7.25	9	13	21	33	52	82	128	198	305	467	714	1088	1658	2523	3837	5833	8866	13471	20467	31092	47229	71739
7.50	9	11	16	25	38	57	84	123	179	258	372	535	767	1098	1570	2243	3202	4570	6518	9295	13253	18893
7.75	9	9	13	19	28	40	57	79	110	151	206	280	379	512	690	930	1251	1681	2256	3028	4061	5445
8.00	9	8	11	15	22	30	40	54	71	93	120	156	201	257	329	419	534	678	860	1090	1380	1745
8.25	10	11	17	27	41	61	92	136	202	296	434	635	926	1349	1964	2857	4155	6040	8778	12574	18529	26916
8.50	10	9	14	21	31	44	64	90	127	178	248	344	476	658	907	1250	1721	2367	3255	4474	6147	8443
8.75	10	8	12	17	24	33	45	62	83	111	148	196	259	340	446	584	763	997	1300	1694	2205	2870
9.00	10	7	10	14	19	25	33	44	57	73	93	118	149	187	235	293	366	455	565	700	867	1072
9.25	11	10	15	22	33	48	70	101	145	206	292	412	581	818	1151	1616	2269	3184	4465	6260	8776	12299
9.50	11	9	12	18	26	36	51	70	96	131	178	241	324	436	585	783	1047	1399	1868	2492	3324	4432
9.75	11	7	10	15	20	28	38	50	66	87	113	147	190	244	314	402	513	655	835	1062	1351	1718
10.00	12	11	16	24	35	52	77	112	162	234	337	484	694	992	1418	2025	2890	4124	5881	8386	11956	17043

Table 1. Average run lengths for values of λ, h, and k in the Poisson cusum (continued)

λ	k \ h	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
10.25	12	9	13	19	28	40	56	79	110	153	211	290	397	544	744	1015	1385	1888	2572	3503	4768	6490
10.50	12	8	11	16	22	31	42	57	77	102	136	180	237	311	408	533	697	910	1186	1545	2011	2617
10.75	12	7	10	13	18	24	32	42	55	71	91	116	147	186	235	295	370	464	580	725	904	1128
11.00	13	10	14	21	30	43	62	88	124	175	245	342	477	664	923	1281	1777	2465	3417	4736	6561	9088
11.25	13	8	12	17	24	34	47	64	87	119	160	216	289	387	517	690	919	1223	1626	2162	2872	3815
11.50	13	7	10	14	20	27	36	48	63	83	108	141	182	235	302	387	497	635	812	1037	1324	1688
11.75	13	6	9	12	16	21	28	36	47	60	75	95	119	148	184	228	282	347	428	526	647	794
12.00	14	9	13	18	26	36	51	71	99	136	187	255	348	473	642	871	1180	1598	2163	2926	3957	5350
12.25	14	8	11	15	21	29	40	54	72	96	127	168	221	290	380	497	650	848	1105	1439	1873	2437
12.50	14	7	9	13	17	23	31	41	53	69	89	114	145	185	234	295	372	469	590	741	930	1166
12.75	14	6	8	11	15	19	25	32	41	51	64	80	99	121	149	183	223	272	330	401	487	589
13.00	15	8	12	16	23	32	44	60	81	110	148	198	265	354	471	626	832	1103	1463	1940	2570	3403
13.25	15	7	10	14	19	26	34	46	61	80	104	136	176	227	292	376	483	619	794	1016	1300	1662
13.50	15	6	9	12	16	21	28	36	46	59	75	95	120	150	188	234	291	361	448	555	687	849
13.75	16	9	12	17	24	34	48	66	91	125	170	231	313	424	573	774	1043	1406	1893	2548	3430	4614
14.00	16	8	11	15	20	28	38	51	68	91	121	159	210	275	360	471	615	802	1044	1360	1770	2303
14.25	16	7	9	13	17	23	31	40	52	68	88	112	144	183	233	295	374	472	596	752	948	1193
14.50	16	6	8	11	15	19	25	32	41	52	65	81	101	125	155	191	235	288	353	432	528	645
14.75	17	8	11	16	22	30	41	57	77	103	139	185	247	329	437	580	769	1018	1347	1781	2354	3111
15.00	17	7	10	14	19	25	33	45	59	77	101	132	171	221	285	367	472	606	778	998	1279	1639
15.25	17	7	9	12	16	21	27	36	46	59	75	95	120	152	190	239	298	372	464	578	719	894
15.50	17	6	8	10	14	18	23	29	36	46	57	71	87	107	131	160	194	236	287	347	420	507
15.75	18	8	11	15	20	27	37	49	66	87	116	153	201	263	345	450	588	766	998	1300	1691	2200
16.00	18	7	9	13	17	23	30	39	52	67	86	111	142	182	232	295	374	474	601	761	961	1215
16.25	18	6	8	11	15	19	25	32	41	52	66	82	103	128	160	198	245	302	372	459	564	693
16.50	18	6	7	10	13	16	21	26	33	41	51	62	76	93	113	136	164	198	238	286	343	410
16.75	19	7	10	14	18	25	33	43	57	75	99	129	167	216	279	360	463	596	766	984	1262	1619
17.00	19	7	9	12	16	21	27	35	46	59	75	96	121	153	193	242	304	382	478	597	746	932
17.25	19	6	8	10	14	18	23	29	37	46	58	72	90	111	136	167	205	251	306	373	455	553
17.50	20	8	11	14	20	26	36	48	64	85	112	148	194	254	332	434	567	739	962	1253	1630	2120
17.75	20	7	9	13	17	22	30	39	51	66	85	110	141	181	231	294	375	476	605	767	972	1232
18.00	20	6	8	11	15	19	25	32	41	52	66	83	105	131	164	204	253	314	389	481	595	735
18.25	20	6	8	10	13	16	21	27	33	42	52	64	79	97	118	144	175	212	257	310	375	452
18.50	21	10		13	18	24	32	43	56	74	97	126	164	212	274	354	456	587	756	972	1249	1604
18.75	21	7	9	12	16	21	27	35	46	59	75	96	122	154	195	246	310	389	489	614	770	964
19.00	21	6	8	11	14	18	23	29	37	47	59	74	92	114	141	174	214	264	324	397	486	595
19.25	21	6	7	9	12	15	20	25	31	38	47	58	70	86	104	126	152	182	219	263	315	377
19.50	22	7	10	13	17	22	29	39	50	65	85	109	141	180	230	294	375	478	608	773	982	1247
19.75	22	7	9	11	15	19	25	32	41	53	67	84	106	133	167	209	261	325	404	502	623	773
20.00	22	6	8	10	13	17	21	27	34	43	53	66	82	100	123	151	184	225	274	333	404	491

High Frequencies

As λ_i becomes large, the observed frequency becomes approximately normally distributed. In this case, we again use the cumulative sum (Equation 1), with the modifications that (a) the observed frequency X_i is transformed to a variable $\tilde{X}_i$ that is normally distributed with mean zero and variance one, and (b) k_i is set equal to one half of the size of the effect one wishes to detect (in terms of standard deviation units). For example, in the common case where one wishes to minimize the time to detect a change of one standard deviation in the mean frequency, k_i is set equal to 0.5 (since the standard deviation of the transformed

Table 2. In-control ARLs (false alarm rates) for various values of h for transformed frequencies

h	ARL_0
2.5	68.9
2.6	76.9
2.7	85.8
2.8	95.6
2.9	106.5
3.0	118.6
3.1	131.9
3.2	146.7
3.3	163.1
3.4	181.2
3.5	201.2
3.6	223.4
3.7	247.9
3.8	275.0
3.9	304.9
4.0	338.1
4.1	374.7
4.2	415.3
4.3	460.1
4.4	509.6
4.5	564.4
4.6	625.0
4.7	691.9
4.8	766.0
4.9	847.8
5.0	938.2
5.1	1038.2
5.2	1148.7
5.3	1270.9
5.4	1405.9

variable is equal to one). The suggested normalizing transformation is (Rossi, Lampugnani & Marchi, 1999):

$$\tilde{x}_i = \frac{x_i - 3\lambda_i + 2\sqrt{\lambda_i x_i}}{2\sqrt{\lambda_i}} \tag{3}$$

This may be used to transform frequencies to approximate normality for use in cumulative sum methods as long as λ_i is greater than about two or three.

When monitoring normally distributed variables, the approximate relationship between the average run length and the signaling parameter h_i is (Siegmund,1985):

$$ARL_0 \approx 2(e^{h+1.166} - h - 2.166). \tag{4}$$

To find h directly from a specified value for ARL_0, one may solve, approximately, Equation (4) for h (Rogerson, 2003).

$$h \approx \frac{\mathrm{ARL}_0 + 4}{\mathrm{ARL}_0 + 2} \ln\left(\frac{\mathrm{ARL}_0}{2} + 1\right) - 1.166 \tag{5}$$

Table 3. Monthly frequency of burglary in Buffalo, New York (1996)

Month	Frequency
January	494
February	445
March	425
April	452
May	551
June	536
July	514
August	546
September	585
October	517
November	487
December	559

To summarize, when expected frequencies are more than two or three, the observed frequency should be transformed using Equation (3), a value of $k_i = 0.5$ should be used (when minimizing the time to find a change of one standard deviation in the mean frequency), and h_i should be chosen using Equation 5 to give the desired ARL_0. Table 2 portrays the average run length for given choices of h.

Application to Burglary Data from Buffalo, New York

Data were available on the number of burglaries by census tract and by month for Buffalo, New York, in 1996. There were 87 census tracts with burglaries, and there were a total of 6,111 burglaries. The breakdown of burglaries by month (Table 3) shows that there is not any strong monthly variation, though the frequencies are slightly lower in the winter months.

The objective is to determine any tract-specific increase in burglaries as quickly as possible. For each tract, the mean monthly expected number of burglaries was taken to be the observed annual total, divided by 12. In monitoring situations where more data are available, it would be more natural to use historical data to determine the average for each tract (λ_i). For each census tract, the value of k_i was found by first taking $\lambda_i^A = \lambda_i + \sqrt{\lambda_i}$, and then using Equation (2), rounding to the nearest integer.

Next, it was decided to allow for the possibility of one false alarm per year. If only one census tract was being monitored monthly, this would imply $ARL_0 = 12$. But since 87 tracts are being monitored simultaneously, we use $ARL_0 = 87 \times 12 = 1{,}044$. This choice implies that, on average, one false alarm per year will occur somewhere among the set of 87 tracts. The computer program found in the appendix was used to construct Table 1, and to find the h_i values that yielded an ARL_0 close to the desired value of 1,044. For the few tracts with expected monthly frequencies greater than 20, we transformed the frequencies using Equation (3), used $k_i = 0.5$, and used Table 2 to find a signaling parameter of $h_i = 5.0$. (For regions with expected frequencies between about 2 or 3 and 20, making the transformation to normality may be regarded as optional; either the Poisson or normal approach will give similar results. When frequencies are less than 2 or 3, the normalizing transformation is not accurate, and when frequencies are greater than about 20, there is no need for the Poisson approach since normality is approximately satisfied).

Table 4. Parameters for census tracts in Buffalo, New York

Tract	λ	*k*	*h*	Tract	λ	*k*	*h*	Tract	λ	*k*	*h*
1	2.5	3	13	30	6.8	8	16	59	4.5	5	21
2	4.2	5	15	31	21.9	0.5	5	60	2.2	3	9
3	1.5	2	9	32	17.8	20	--	61	1.1	2	5
4	2.3	3	9	33	8.8	10	20	62	4.3	5	15
5	3.5	4	17	34	2.3	3	9	63	6.7	8	15
6	2.2	3	9	35	7	8	19	64	4.4	5	21
7	1.4	2	9	36	1.8	2	15	65	4	5	11
8	3	4	10	37	6.7	8	16	66	7.9	9	21
9	1.8	2	15	38	2.6	3	13	67	8.4	10	16
10	7.3	9	13	39	6.4	8	13	68	11.3	13	19
11	3.7	5	10	40	6.2	7	21	69	10.5	12	19
12	8.4	10	16	41	8.8	10	20	70	0.5	1	4
13	0.7	1	8	42	9.6	11	19	71	0.4	1	4
14	1.7	2	15	43	9.9	11	23	72	8.1	9	21
15	0.7	1	8	44	3.5	4	17	73	2.7	4	18
16	3.8	5	10	45	6.3	7	21	74	0.8	1	11
17	5.8	7	14	46	8.7	10	20	75	4.8	6	12
18	30.8	0.5	5	47	9.3	11	16	76	4.3	5	15
19	6.3	7	21	48	7.7	9	16	77	5.4	7	12
20	2	3	7	49	8.5	10	16	78	2.5	3	13
21	2.3	3	9	50	9.1	10	23	79	8.5	10	16
22	2.5	3	13	51	4.3	5	15	80	4.8	6	12
23	1.8	2	15	52	1.3	2	6	81	13.8	16	18
24	1.9	3	7	53	3.8	5	10	82	26.8	0.5	5
25	2.2	3	9	54	10.9	12	25	83	10.3	12	17
26	8.8	10	20	55	2.5	3	13	84	13	15	19
27	1.6	2	9	56	4	5	12	85	7.5	9	15
28	5.2	6	18	57	2.3	3	10	86	6.1	7	17
29	3	4	10	58	1.8	2	15	87	0.4	1	4

Results

Table 4 displays the parameters for each of the 87 census tracts. An alarm signaling a change in the mean number of monthly burglaries was sounded in only one tract (tract no. 47). Data for this tract is shown in Table 5. This tract has a mean of 9.3 burglaries per month. For this tract, $k = 11$. The cumulative sum represents a tally of the number of monthly burglaries that is in excess of 11. In August, the cumulative sum is positive for the first time, at 15-11 = 4. In September, the cumulative sum increases to 4 + (22-11) = 15. In October, the cumulative sum reaches the critical value of 16, and an alarm is sounded, indicating that a significant increase over the base rate of 9.3 burglaries per month has occurred. At this time, further investigation may be warranted to

Table 5. Tract No. 47

$\lambda = 9.3; k = 11; h = 16$

Month	Number of Burglaries (X_t)	(X_t -k)	Cumulative Sum (S_t)
January	10	-1	0
February	8	-3	0
March	5	-6	0
April	4	-7	0
May	9	-2	0
June	9	-2	0
July	3	-9	0
August	15	4	4
September	22	11	15
October	12	1	16*
November	7	-4	12
December	7	-4	8

Note: The asterisk denotes the first time period where the cumulative sum has reached or exceeded the critical threshold.

The cumulative sum accumulates the number of monthly burglaries that is greater than k = *11.*

Thus in August, 4 = 15-11. In September, the cumulative sum is equal to the August cumulative sum, 4, plus the September excess, 11 (= 22 - 11), which totals 15 (= 4 + 11).

attempt to ascertain the cause of the increase in this tract. The alarm is not sustained over time; the cumulative sum declines below the critical value of h=16 in November and December.

One important limitation of this illustration is the relatively long time delay between the month of most pronounced increases (September, where there were 11 more burglaries than expected), and the time the alarm is sounded (presumably November 1, after data for October have been tabulated). Thus, an individual burglar who is active during September would not even be noted until the beginning of November, and it would by then be difficult to act on this information. One potential remedy would be to use daily or weekly data if it were available. With daily data, the number of expected burglaries in particular census tracts would be extremely low, and the alternative approaches suggested by Lucas (1989) for data consisting of extremely low counts might be used.

With extremely low counts, the number of time periods needed to detect a change will increase. The viability of implementing the cusum approach will ultimately depend upon the combination of the region size, the width of the time interval used to collect observations, and the detection delay that is acceptable for effective response. For a region of given size, the time window used for data collection can not be too large, for reasons seen in this example. If the time window is very small, the extremely low counts that result may imply longer detection times. Similar comments may be made about spatial aggregation. If the region size is too large, changes occurring on smaller spatial scales may be missed. But if the regions are too small, low expected frequencies will again imply longer detection times.

Summary and Discussion

Summary

To summarize the steps in carrying out the geosurveillance program:

1. Begin by determining the average expected number of crimes expected in each subarea (λ_i). For example, this might be equal to the areawide rate of crimes per person, multiplied by the number of persons in subregion *i*. More interesting would be to take λ_i as the crime frequency that prevailed during some recent historical period.
2. If λ_i is sufficiently large, then transform the observed frequencies for that region to $\tilde{X}_i = (X_i - 3\lambda_i + 2\sqrt{\lambda_i X_i})/(2\sqrt{\lambda_i})$, and use $k_i = 0.5$. Otherwise, find k_i either from Equation 2, or the more specific and approximate $k_i \approx \lambda_i + 0.5\sqrt{\lambda_i}$. Round k_i to the nearest integer.
3. Decide on the false alarm rate (that is, the value of ARL_0) and find the corresponding value of h_i from either (a) Equation 5, when the normalizing transformation has been used, or (b) a table (such as Table 1, or the one given by Lucas) or a computer program for threshold values associated with Poisson cumulative sums.
4. Use the parameters found in steps 1 through 3 to construct a cumulative sum for each region (using Equation 1). Declare an alarm when any of the cumulative sums exceeds its signaling value of h_i.

Discussion

In this chapter, I have described a system and computer program designed to detect quickly any changes in crime frequencies relative to what is expected for a set of regions.

There are a number of potential extensions to the system. First, it may be of interest to look for changes on larger spatial scales. This would entail monitoring the crime frequency in some small neighborhood defined around each region. For example, the observed and expected crime frequencies could be compared for the area defined by the region and its adjacent regions (instead of simply using the region itself). Alternatively, when defining the cusum for a region, the observed and expected frequencies for all regions could make weighted contributions, with larger weights applied to those regions that were close to the primary region, and small weights (possibly equal to zero in some cases) applied to those that were farther away. Thus we could monitor the cumulative sum for the quantity $x_1 + \sum_{j \neq 1} w_j x_j$ instead of x_1. Essentially this describes the scenario where one is monitoring "local" statistics (see Anselin [1995] and Getis and Ord [1992] for a discussion of local statistics).

It will also often be of interest to take into account changes that are expected over time in the crime frequencies. Seasonal or daily fluctuations are a good example; there is no reason to believe that λ_i necessarily remains constant over time. If the expectations are sufficiently high, taking such fluctuations into account is straightforward – one simply carries out the transformation described above in Equation (3), using a time-specific value of the expectation. One then compares the transformed observed frequency with the (time-varying) expected frequency.

For Poisson cumulative sums, the required adjustments to the procedure are more detailed when the expectations vary over time. Suppose the expected value associated with the Poisson expectation varies with time ($\lambda_{i,t}$; $t = 1,2,...$). We now use values of the parameters k and h that are time-specific. The observed values, X_t, may then be used in the cumulative sum as follows:

$$S_t = \max(0, S_{t-1} + c_t(X_t - k_t)) \qquad (6)$$

where the parameters c_t and k_t change from one period to the next, and their values are now discussed. First h is chosen based upon the mean of the time-varying Poisson parameter, an associated value of k, and the desired ARL. Once h is chosen, next choose k_t based upon λ_t and λ_t^A:

$$k_t = \frac{\lambda_t^A - \lambda_t}{\ln \lambda_t^A - \ln \lambda_t} \tag{7}$$

Then c_t is chosen as the ratio h to h_t, the value of the signaling parameter that would have been chosen in a usual Poisson scheme with desired ARL, k_t, and constant values of λ_t and λ_t^A (designated h_t). Thus $c_t = h/h_t$. The quantity c_t is therefore chosen so that observed counts, X_t, will make the proper relative contribution toward the signaling parameter h that is used in the actual cusum. If for example $h > h_t$, then the contribution $X_t - k_t$ is scaled up by the factor h/h_t (see Rogerson and Yamada 2003 for additional details and an example).

It is also important to note that the illustration here has been made using retrospective data in a prospective manner. That is, we have used past data (from 1996) to create a scenario "as if" we were carrying out prospective monitoring in 1996. In more typical applications it is expected that cumulative sums will be kept up-to-date by adding new, current observations as they become available.

In practice, it may be important to account for the changing spatial distribution of changing population. This will be particularly true for study areas that are growing rapidly, or for study areas that are monitored for long periods of time. Any changes in population distribution should be reflected in the specification of expected frequencies. More generally, more sophisticated models of expectations based upon covariates such as age, weather and other variables could be developed.

Acknowledgments

This chapter is based upon work supported by Grant No.98-IJ-CX-K008 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. It was also supported by the National Science Foundation under Award No. BCS-9905900. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice or the National Science Foundation. I am grateful to the referee for several helpful suggestions.

References

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis, 27*, 93-115.

Getis, A., & Ord, J. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*, 189-206.

Lucas, J. M. (1985). Counted data cusums. *Technometrics, 27*, 129-144.

Lucas, J. M. (1989). Control schemes for low count levels. *Journal of Quality Technology, 21*, 199-201.

Montgomery, D. C. (1996). *Introduction to statistical quality control*. New York: Wiley.

Rogerson, P. (2003). Formulas for the design of CUSUM quality control charts (Unpublished article, submitted for publication).

Rogerson, P., & Sun, Y. (2001). Spatial monitoring of geographic patterns: An application to crime analysis. *Computers, Environment, and Urban Systems*, *25*(6), 539-56.

Rogerson, P. & Yamada, I. 2004. Approaches to Syndromic Surveillance When Data Consist of Small Regional Counts. Morbidity and Mortality Weekly Report. vol. 53 (supplement), September 24, 2004, pp. 79-85.

Rossi, G., Lampugnani, L., & Marchi, M. (1999). An approximate cusum procedure for surveillance of health events. *Statistics in Medicine, 18*, 2111-2122.

Ryan, T. P. (1989). *Statistical methods for quality improvement*. New York: Wiley.

Siegmund, D. O. (1985). *Sequential analysis: Test and confidence intervals.* New York: Springer-Verlag.

Wetherill, G. W., & Brown, D. W. (1991). *Statistical process control: Theory and practice.* New York: Chapman and Hall.

White, C. H., & Keats, J. B. (1996). ARLs and higher order run length moments for Poisson CUSUM. *Journal of Quality Technology, 28*, 363-369.

Appendix

GAUSS program code

```
new;

/*  this program should be run first; it finds the values of h for user input values of k and lamda (lam) .  It does not need to be run if values of h are simply found from Table 1.   */

mm=1; do while mm<33;
h=1+mm;
/* SET k and lambda HERE */  k=1; lam=.7;

p=zeros(51,1);  tot=zeros(51,1);   tot2=0;
 i=1; do while i<52;
j=i-1;  p[i,1]=exp(-lam)*lam^j/j!;  tot2=tot2+p[i,1];  tot[i,1]=tot2;
i=i+1;  endo;

h1=h+1;
t=zeros(h1,h1);
v1=seqa(1,1,int(k)+1);  t[1,1]=sumc(submat(p,v1,0));
t[h1,h1]=1;

m=2; do while m<h1;
if k-m+1<0; t[m,1]=0; goto aa; endif;
v4=maxc(1|int(k-m+1)+1);
v2=seqa(1,1,v4);
t[m,1]=sumc(submat(p,v2,0));
aa:   m=m+1; endo;

m=1; do while m<h1;
j=2; do while j<h1;
```

```
if k+j-m<0; t[m,j]=0; goto bb; endif;
v6=maxc(1|int(k+j-m+1));
t[m,j]=p[v6,1];
bb: j=j+1; endo; m=m+1; endo;

m=1; do while m<h1;  t[m,h1]=1-sumc(t[m,.]’);  m=m+1; endo;

v9=seqa(1,1,h);
t1=submat(t,v9,v9);
u=inv(eye(h)-t1);
u1=u*ones(h,1);
“h    ARL”; h~u1[1,1];
mm=mm+1; endo;
new;

/* this program finds the cumulative sum once parameters have been chosen
and input below */
/* LOAD DATA AND CHOOSE PARAMETERS HERE */

/* r is the number of regions; t is the number of time periods */
r=87; t=12;
/* load observed burglary data; data should consist of r rows and t columns
*/
load x[r,t]=e:burgmonth4.txt;
/* input the threshold values (h); one for each region */
let h[87,1]=13 14 9 9 17 8 8 10 19 13 9 16 7 12 7 10 14 5 21 7 9 13 19 7
8 20
10 17 10 16 5 23 20 10 19 19 15 15 12 19 20 19 23 17 21 18 16 16 16 14
15 7
10 25 13 12 10 19 20 8 5 15 15 18 12 19 16 19 19 4 4 22 18 11 12 16 11
13 16 12 18 5
17 19 15 17 4;
```

```
/* now the program begins */
v1=seqa(1,1,t); rr2=submat(x,0,v1);

/* here we compute expected values as the monthly averages */
lam2=sumc(rr2')./t;

/* next we compute the k parameters for each region */
k=round(sqrt(lam2)./(ln(lam2+sqrt(lam2))-ln(lam2)));

q=1; do while q<r+1;

if lam2[q,1]<21; goto bb; endif;
k[q,1]=0.5;
x[q,.]=sqrt(x[q,.])+sqrt(x[q,.]+ones(1,t))-sqrt(4*lam2[q,1]+1).*ones(1,t);

bb: s=0; t3=0;

i=1; do while i<t+1;
s=s+x[q,i]-k[q,1];
s=maxc(s|t3);
if s>=h[q,1]; "alarm:"; "region: " q; "time:" i; goto aa; endif;
i=i+1; endo;
aa: q=q+1; endo;
```

Chapter X

Application of Tracking Signals to Detect Time Series Pattern Changes in Crime Mapping Systems

Wilpen L. Gorr, Carnegie Mellon University, USA

Shannon A. McKay, Massachusetts Institute of Technology, USA

Abstract

Tracking signals are widely used in industry to monitor inventory and sales demand. These signals automatically and quickly detect departures in product demand, such as step jumps and outliers, from "business-as-usual". This chapter explores the application of tracking signals for use in crime mapping to automatically identify areas that are experiencing changes in crime patterns and thus may need police intervention. Detecting such changes through visual examination of time series plots, while effective, creates too large a workload for crime analysts, easily on the order of

1,000 time series per month for medium-sized cities. We demonstrate the so-called smoothed-error-term tracking signal and carry out an exploratory validation on 10 grid cells for Pittsburgh, Pennsylvania. Underlying the tracking signal is an extrapolative forecast that serves as the counterfactual basis of comparison. The approach to validation is based on the assumption that we wish tracking signal behavior to match decisions made by crime analysts on identifying crime pattern changes. We present tracking signals in the context of crime early warning systems that provide wide-area scanning for crime pattern changes and detailed drill-down maps for crime analysis. Based on preliminary results, the tracking signal is a promising tool for crime analysts.

Introduction

Police generally know the current crime patterns in their jurisdictions and accordingly allocate manpower to precincts and shifts, target patrols to hot spots, and take other tactical actions. What is less well known to police is how crime patterns are changing, so that police can reallocate manpower in response to changes. We learned this lesson in the early 1990s when we built a crime mapping system for the Pittsburgh, Pennsylvania, Bureau of Police under a drug market analysis program (DMAP) grant funded by the National Institute of Justice. Many times, our DMAP crime mapping system detected enforcement-induced displacement of street-level drug dealing before narcotics detectives were able to do so. Follow-up surveillance of new drug dealing locations detected by our system always proved the maps to be right.

From this experience, we learned the value of building crime early warning system (CEWS) maps. These maps display crime changes to provide a jurisdiction- wide scan for areas needing changes in tactical deployment of police. Used on an interactive basis in a geographic information system, the maps provide drill down to areas of high change to provide detailed, diagnostic information. We provide example maps in this chapter, but before proceeding to them, it is important to distinguish two types of change: experienced and forecasted change.

Experienced change is the sort mentioned above, which has the objective of quickly detecting any sort of crime innovation (departures from business-as-usual crime patterns), such as crime displacement in response to enforcement. Underlying analytic problems are 1) to provide counterfactual forecasts (business as usual) as the basis of comparison for the most recent historical crime data and 2) to sort out true pattern changes from random variations. More is on these issues below. Detecting experienced change is a major activity in Comstat meetings (see Henry & Bratton, 2002).[1]

The second type of change - not the subject of this chapter - is forecasted change, which provides some capacity for crime prevention. Recently there has been success on developing crime forecasting as an applied research field (for example, see Gorr and Harries (2003), which introduces a special section on crime forecasting in the *International Journal of Forecasting*). Extrapolation of crime seasonality and time trend one month ahead have proven to be accurate enough for use in tactical deployment given adequately high crime rates in areas investigated (Gorr, Olligschlaeger & Thomson, 2003).

Our purpose in this chapter is to introduce and examine tracking signals as a potential tool for crime analysts for automatically detecting crime innovations. We undertake an exploratory empirical validation of one of the best tracking signals. We have not seen any validation studies in the literature using real data such as used here. All have relied on simulated data with known pattern changes for validation. Instead, we use judges (ourselves) to visually identify pattern changes. The next section provides examples of CEWS maps to provide the context for tracking signals (and crime forecasting) as tools for use by crime analysts in mapping crimes. The third section makes the case that an automated approach is needed to detect experienced crime pattern changes. The fourth section of this chapter briefly reviews tracking signals. Following that is a section on our research design for validation, followed by a section on results, and then a conclusion.

Crime Early Warning System Maps

Next, we provide examples of CEWS maps. Such maps appear similar in format whether using experienced or forecasted change. Thus, while we do not have good example maps for experienced change at this time, the ones provided for forecasted change next are representative of change maps in general.

Figure 1 is a CEWS map for Pittsburgh, Pennsylvania, displaying one-month-ahead crime forecasts where the areas are uniform grid cells 4,000 feet on a side (Gorr et al., 2003). The plotted values are forecasted changes in part 1 property crimes in December made at the end of November in a particular year. Increasingly dark, solid-fill shading shows areas of increasingly larger forecasted increases and increasingly dark cross-hatching shows areas of increasingly larger forecasted decreases. While there are 103 grid cells, only nine have forecasts of sizable increases and of those only two have large increases (grid cells 61 and 77). Thus crime analysts would likely start working with the two worst cases, and then proceed to the other seven.

CEWS includes drill down to individual crime points of the most recent month – either for the crime type of the grid cells (part one property crimes) or corresponding leading indicator crimes (such as criminal mischief and disorderly conduct). Figure 2 is a drill down (zoom in) to grid cell 77 showing crime points for two part one property crime types, burglary and larceny, in November. Clearly, there are hot spot clusters for both crime types. Based on an assumption of persistence for the hot spots (for example, Block ,1995; Harries, 1999; Liu & Brown, 2003), and a study of corresponding crime reports and modus operandi data (for example, place of entry, time of day and so forth), crime analysts can suggest places and times to patrol hot spot areas within grid cells.

Need for Automated Detection

A problem with attempting to identify crime time series pattern changes for current conditions is that the analyst must examine time series plots of about five years length each month. Analysts have to account for regular noise versus departures from established time trend patterns, such as a sudden discrete change (step up or down) or a turning point (for example, change from a

Figure1. Early warning system with forecasted change in serious property crimes for December made at the end of November in Pittsburgh, Pennsylvania

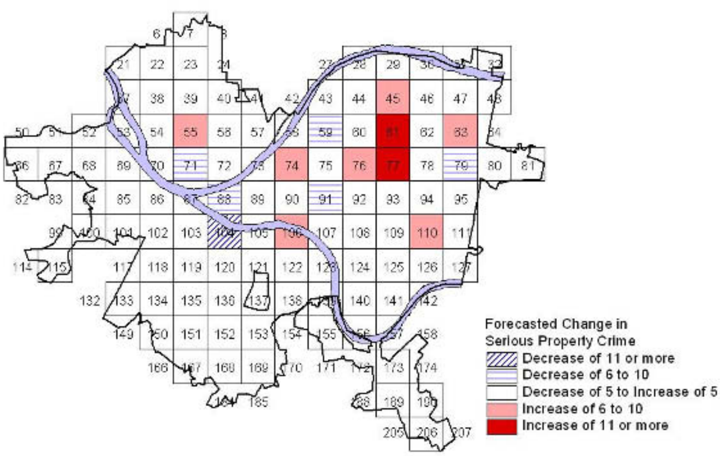


Figure 2. Zoom-in to grid cell 77 to view November crime points

decreasing time trend to an increasing trend). This work can be done by visual examination, but generates an unacceptably large workload because analysts must work with relatively small geographic areas, such as grid cells or census tracts. For example, in Pittsburgh, there are approximately 100 grid cell areas to examine and at least 10 crimes of interest, yielding roughly 1,000 crime series plots to generate and examine each month. Clearly it is infeasible to implement pattern change detection with visual examination. This is where tracking signals come into play. They automatically flag exceptional time series.

Time series tracking signals are widely used by businesses for sales forecasting and inventory control to generate exception reports of time series that have likely deviated from their historical time trends. Next is a brief review of tracking signals.

Tracking Signals

An approach to evaluating a phenomenon at a point in time is to make a counterfactual forecast for the point, which predicts the point under business-as-usual conditions. Then a tracking signal can be established, based on the actual crime data point and in reference to the corresponding counterfactual, so that if the tracking signal exceeds a selected control limit, an exception report is tripped for a potential time series pattern change. We use extrapolative time series forecasts to make counterfactual forecasts; namely, the most accurate extrapolative forecast method as determined by Gorr et al. (2003) for one-month-ahead crime forecasts. This is Holt exponential smoothing with smoothing parameters optimized and using time series data deseasonalized with multiplicative seasonal factors estimated from jurisdiction-wide data and by classical decomposition (Bowerman & O'Connell, 1993). Thus, business-as-usual is defined to be a time series pattern following a smoothed linear time trend (straight line fitted to the time trend, placing most weight on the most recent data points) and monthly seasonal factors such as 1.25 (25% higher seasonal effect) or 0.80 (20% lower seasonal effect). The counterfactual forecast extends the estimated time trend ahead to the point being analyzed and applies the corresponding seasonal multiplier, using all prior data to estimate trend and seasonality.

Tracking signals generally are ratios in which the numerator is a sum or weighted sum of counterfactual forecast errors that has an expected value of zero when time series patterns (time trend and seasonality) are stable. When there is a pattern change, such as a step jump or turning point, the numerator moves away from zero. The denominator's purpose is to normalize by the long-term average variability of forecast errors. Of the common tracking signals, the smoothed error tracking signal due to Trigg (1964) is a good choice for practitioners (McClain, 1988). The equations are as follows:

$$E_t = \alpha_1 e_t + (1- \alpha_1)E_{t-1} \tag{1}$$

$$MAD_t = \alpha_2|e_t| + (1- \alpha_2)MAD_{t-1} \tag{2}$$

$$T_t = |E_t/MAD_t| \tag{3}$$

where

MAD = mean absolute deviation of forecast errors

E_t = smoothed forecast error

T = tracking signal

t = month being evaluated

e_t = counterfactual forecast error

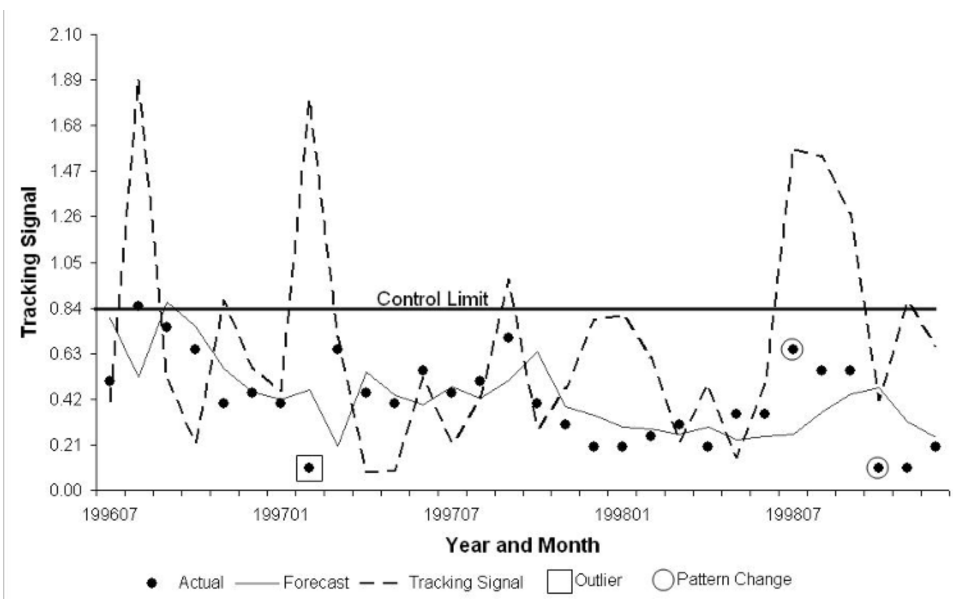
α_1 = smoothing factor for numerator

α_2 = smoothing factor for denominator

We implement this signal with smoothing parameter values as suggested by McClain (1988): α_1=0.40 for the smoothed sum of errors for the numerator (in order to quickly detect pattern changes) and α_2=0.05 for the denominator of smoothed mean absolute deviations of forecast errors. The initial value for E_0 is assumed to be 0, so that there is a burn-in period during which the tracking signal adapts to the actual pattern and forgets the initial value. In addition to computing the tracking signal, the analyst must also choose critical values, which, if exceeded, trip an exception report. We make the critical value an experimental treatment, trying a range of critical values in an attempt to tune tracking signal behavior to match crime analysts' judgment on crime pattern changes.

These equations are easily implemented in a spreadsheet package for experimentation, but normally would be programmed to work automatically within a

Figure 3. Sample tracking signal for 911 drug calls in grid cell 120 with marked pattern changes and outlier



CEWS. Figure 3 is an example of equations 1-3 applied to monthly time series data for 911 drug calls in grid cell 120 of Figure 1. Each tracking signal value has five years of historical data behind it in order to estimate corresponding counterfactual forecast models.

Marked for comparison purposes are two pattern changes and an outlier. The actual and forecasted crime levels have been rescaled to match the vertical scale of the tracking signal. When the tracking signal crosses above the control limit line, it issues (trips) an exception report, warranting analysis of this time series. As J. McClain (1988) states, "A perfect tracking signal would detect an out-of-control forecast (that is, a time series pattern change) immediately, and would never give a false alarm." Of course, this is not possible, so in Figure 3 the reader can see false positives (the first and third trips), but also actual positives detected immediately (the second and fourth trips), and a delay in detecting an actual positive (the last data point which appears as if it would be detected if one more data point were available).

Research Design

This section addresses the question of whether tracking signals really perform well for detecting changes in crime series patterns. We must account for false positives and determine if tracking signals reduce workloads adequately. We have not seen any attempts in the literature to validate tracking signals with actual data, as in Figure 3. All validations appear to have used simulated data with known pattern changes and outliers. It is very desirable, however, to use actual data in order to assess value in a given context; namely, will tracking signals adequately reduce workload and not miss actual positives? Thus, we assumed that the purpose of tracking signals is to match behavior of trained, human judges (crime analysts), and simply automate their decisions on pattern changes and outliers.

We did not have the resources to embark on a full-scale validation; hence, we decided to carry out an exploratory study to determine the feasibility of our approach and provide preliminary results. We chose 10 crime time series from the Pittsburgh grid system of Figure 1. They consist of a variety of crime types with five time series having pattern changes and the other five not having any. It is important to include time series with no pattern changes to assess false positive rates.

Both authors independently marked-up each of the time series for pattern changes and outliers, as in Figure 3, under the guideline that we would only mark those that are large and obvious. We then compared results and reconciled differences. One of us had merely admitted some smaller pattern changes in interpreting “large and obvious.” The result was 18 instances of pattern changes or outliers in five of the time series used in our analysis.

Our treatment of the smoothed signal tracking signal was to use it with a variety of control limits, searching for the control limit that best matches detection of our judged, true pattern changes. After some trial and error, we decided to use values of 0.84, 1.05, 1.26 and 1.47. This range starts at a low value (0.84) that detects most of the actual positives, but at the cost of tripping many false positives (false alarms). At the other extreme (1.47), there are fewer detections of actual positives, but also many fewer false positives.

Results

We applied equations 1-3 on the 10 time series over the 36-month period in which counterfactual forecasts were made. In reporting results, we decided to exclude the first six months of tracking signals for burn-in so that the tracking signal could forget arbitrary initial values and start tracking correctly. Hence there were 10 time series times 30 months each for a total of 300 signal values estimated. Also, this translates to 300 time series plots that a crime analyst would have had to examine to accomplish the same task.

We define an exception report “epoch” to be the total number of time periods that the tracking signal is above its control limit, including the first month that it trips. We assume that the crime analysis protocol is that the crime analyst must investigate each time series plot and corresponding crime maps for each month of epochs. Hence the count of all epoch months is a measure of the workload that the crime analyst would have to do when using tracking signals. The comparison without a tracking signal is 300 or 10 per month.

Table 1 is the result of our research. For a control limit of 0.84, the tracking signal detects 17 (94%) of the 18 actual positives, which appears to be quite good. It also does so with no lag or one period lag. The cost is, that of the average total of four time series per month to be examined (instead of 10), 2.9 are false positives. At the other extreme, with a control limit of 1.47, only 11 (61%) of the actual positives are detected, but the total workload per month is down to 1.6 time series, 1 of which is a false positive. The number of false positives falls quickly between the first two control limits in Table 1 and then flattens out.

We believe that these results are promising because they show a 60% work reduction for the most stringent case and up to an 84% workload reduction for the least stringent case.

Table 1. Final results on validation research

Control Limit	True Positives Detected	Average Workload (Time Series/Month)	Average False Positives (Time Series/Month)
0.84	17 (94%)	4.0	2.9
1.05	13 (72%)	2.8	1.9
1.26	12 (67%)	2.1	1.4
1.47	11 (61%)	1.6	1.0

Conclusion

This chapter has discussed crime early warning systems and introduced an application of tracking signals for detecting experienced time series pattern changes in crime maps. The basis of the tracking signal is information obtained from counter-factual forecasts for each point examined. These are forecasts providing business-as-usual estimates for a point in time, as if no pattern changes existed. The tracking signal automates detection of pattern changes by matching the decisions of crime analysts as to what data points constitute the start of a new time series pattern. We varied the control limit of the tracking signal, making it more or less sensitive to information in the time series data in attempting to tune the tracking signal to match crime analysts' decisions.

In future work, it will be necessary to take a large sample of time series, have crime analysts mark them up for pattern change points and outliers, and rerun the research study. Additional tracking signals may be tried, as well as varying the tracking signal numerator's smoothing factor (which we did not do) for further tuning and attempting to improve performance.

Acknowledgments

This project was supported by Grant No. 2001-IJ-CX-0018 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of

Justice. This project was also supported by a research fellowship from the Heinz Endowments awarded to Shannon McKay. We are grateful for the assistance of Commander Kathleen McNeely of the Pittsburgh Police Bureau and John Shultie of Pittsburgh City Information Systems for providing crime data. We wish to thank two anonymous referees and the editor for valuable suggestions.

References

Block, C. (1995). STAC hot spot areas: A statistical tool for law enforcement decisions. In C. Block, M. Dabdoub & S. Fregly (Eds.), *Crime analysis through computer mapping*. Washington, DC: Police Executive Research Forum.

Bowerman, B. L., & O'Connell, R. T. (1993). *Forecasting and time series: An applied approach*. Belmont, CA: Duxbury Press.

Gorr, W. L., & Harries, R. (2003). Introduction to crime forecasting. *The International Journal of Forecasting, Special Section on Crime Forecasting, 16*, 551-555.

Gorr, W. L., Olligschlaeger, A. M., & Thompson, Y. (2003). Short-term time series forecasting of crime. *The International Journal of Forecasting, Special Section on Crime Forecasting, 16*, 579-594.

Harries, K. (1999). *Mapping crime: Principle and practice* (Research report).Washington, DC: National Institute of Justice.

Henry, V. E., & Bratton, W. J. (2002). *The Compstat paradigm: Management accountability in policing, business and the public sector*. Flushing, NY: Looseleaf Law Publications.

Liu, H., & Brown, D. E. (2003). Criminal incident prediction using a point-pattern-based density model. *The International Journal of Forecasting, Special Section on Crime Forecasting, 16*, 603-622.

McClain, J. O. (1988). Dominant tracking signals. *The International Journal of Forecasting, 4*, 563-572.

Trigg, D. W. (1964). Monitoring a forecasting system. *Operational Research Quarterly, 15*, 271-274.

Endnotes

[1] Note that this chapter pursues experienced change relative to geographic areas, such as grid cells or census tracts. Another important pattern to establish, as an innovation, may cut across several geographic areas and is the identification of a serial criminal. In this case, analysis surrounds the matching of physical descriptions and modus operandi of perpetrators.

Chapter XI

Integrating GIS, GPS and MIS on the Web: EMPACT in Florida

Gregory A. Frost, Tallahassee Police Department, USA

Abstract

Computer applications for conducting complex spatial analysis of crime data are widely used by law enforcement agencies. By combining sophisticated geographic information systems with global positioning satellite tracking devices, a new tool is emerging that will remove the criminal anonymity of probationers, parolees and offenders on pretrial release. Every year, an ever-increasing number of offenders are set free on either probation or parole within our nation's communities. As the number of offenders on our streets grows, the need for the criminal justice system to hold these offenders accountable and exert some level of control also increases. Florida's Electronic Monitoring Protection and Crime Tracking (EMPACT) project is breaking new ground in an effort to use technology as an effective way to remove the anonymity of crime. Through the automated correlation of GPS tracking data and local crime incident data, participating criminal justice agencies are able to determine if a tracked offender was at the scene of a crime when it occurred. In addition, because EMPACT uses a Web-based interface, participating agencies also have

access to each other's data. This creates a crime-mapping environment where crime analysts and investigators have the opportunity to evaluate, at the click of a button, multi-jurisdictional crime patterns and offender track data.

Introduction

The use of sophisticated geographic information systems (GIS) as a part of the crime analysis process has significantly impacted law enforcement's understanding of offender behavior and its relationship to crime. This increased understanding has, in turn, translated into more effective enforcement operations and crime prevention strategies. Today's complex analysis of spatial factors such as crime location, offender residence, transportation routes, police patrol zones and residential and commercial areas enables problem solving at a level that was unheard of just a dozen years ago. And more advanced systems are being continually developed. Keith Harries accurately forecasted how crime mapping would evolve:

The hallmark of the first decade or so of the modern era of crime mapping was the use of geographic information systems (GIS). Perhaps the next decade will see the integration of previously separate technologies such as global positioning systems (GPS)…and a wide range of local databases with relevance to policing – and the World Wide Web. (1999, p. 151)

The Florida Electronic Monitoring Protection and Crime Tracking (EMPACT) project is making Harries' forecast a reality. By combining GIS technologies with GPS data from tracked probationers and parolees, and crime data extracted from local law enforcement agencies' records management systems, a new tool has been created that enables the criminal justice system to effectively remove anonymity and reduce the likelihood of criminal behavior.

Criminologists have acknowledged for decades that the best predictor of future criminal activity is an offender's past criminal behavior (MacKenzie, 1997). Law enforcement officers also know that only a small percentage of the community members they police are responsible for the majority of criminal offenses – and it is these offenders who continually cycle through the criminal justice system. As a society, we cannot lockup all criminal offenders and throw away the key. Not only is the "throw away the key" proposition economically unfeasible, it is not justified under our nation's legal philosophies. One of the results, however, is an increasing reliance on probation and parole as way to exert at least some level of control over criminal offenders without the associated costs of incarceration.

Data released by the U.S. Department of Justice, Bureau of Justice Statistics clearly reflect the criminal justice system's widespread use of community supervision, both probation and parole, as an alternative sanction. Between 1995 and 2002, the number of offenders on probation or parole increased from 3,757,282 to 4,748,306 – almost an additional million offenders nationwide (Glaze, 2003). The Bureau of Justice Statistics also documented, through a series of reports, the level of criminal recidivism for offenders sentenced to community supervision. Nationwide, 43% of adult felons sentenced in state courts to supervised probation in the community were arrested within 36 months for a new felony offense, and almost 50% of those arrested were arrested more than once (Langan & Cunniff, 1992). In another study, Cohen (1995) reviewed the records of over 300,000 state inmates who were in prison during 1991 for committing a new crime while serving a sentence of parole or probation for a previous conviction. The group of studied offenders committed at least the following serious crimes – 13,200 murders, 12,900 rapes, 39,500 robberies, 19,200 assaults, 39,600 burglaries and 7,900 vehicle thefts.

History of EMPACT

At the center of the EMPACT system are GPS devices worn by tracked offenders. The emergence of small, low-cost GPS recording equipment provided the enabling technology to create devices capable of tracking the whereabouts of criminal offenders. One of the first portable GPS recording apparatuses was developed in the late 1970s and was a 25-pound backpack (Rand Corporation, 1995). Several manufacturers subsequently refined the emerging portable technology primarily for military and commercial survey applications.

The first GPS tracking device developed specifically for tracking offenders under community supervision was operationally deployed in 1997 by the Florida Department of Corrections. The two-piece system, which is still in use today, consists of an ankle bracelet and a four-pound box carried by the offender. System integrity is monitored by features that ensure alarms are recorded if the device or the ankle bracelet is tampered with, or the two pieces are more than approximately 150 feet apart. Because of its cellular telephone connectivity, this tracking device is considered an "active" system. The near real-time notification capability for alarms is also important for reporting "exclusion" and "inclusion" zone violations. Discussion of zone use to control tracked-offender behavior and as an investigative tool is included in the "EMPACT Functionality" section of this chapter.

As with most new technology, there is an evolutionary cycle during which the equipment continues to be refined in size, functionality and cost to meet market needs. In 2002, a more cost effective "passive" GPS tracking device was developed. This device still requires an ankle bracelet to electronically tether the tracking device to the offender. Functionality of the passive system is almost identical to the active system except for cellular connectivity. All track data and alert notifications for the passive system are downloaded to the monitoring center when the device is placed in its charging stand at the offender's home. Over the last 12 months several vendors have developed advanced GPS based tracking devices – including a one-piece "active" device that is worn on the offender's ankle. These new devices provide the criminal justice system with alternatives that best fit jurisdictional needs based upon the type of offenders being tracked.

Using the Department of Defense's constellation of 24 orbiting global positioning satellites, all of the current tracking devices worn by offenders collect latitude and longitude points with 10- to 20-meter accuracy. The accuracy and reliability of GPS tracking technology has gained acceptance within Florida's criminal justice system. Evidence of this acceptance is reflected in the 56.6% increase in use of GPS tracking by Florida circuit judges between fiscal year 1999/2000 and fiscal year 2001/2002 (Florida Department of Corrections, 2002), as well as implementation of EMPACT in five Florida counties for county level offenders.

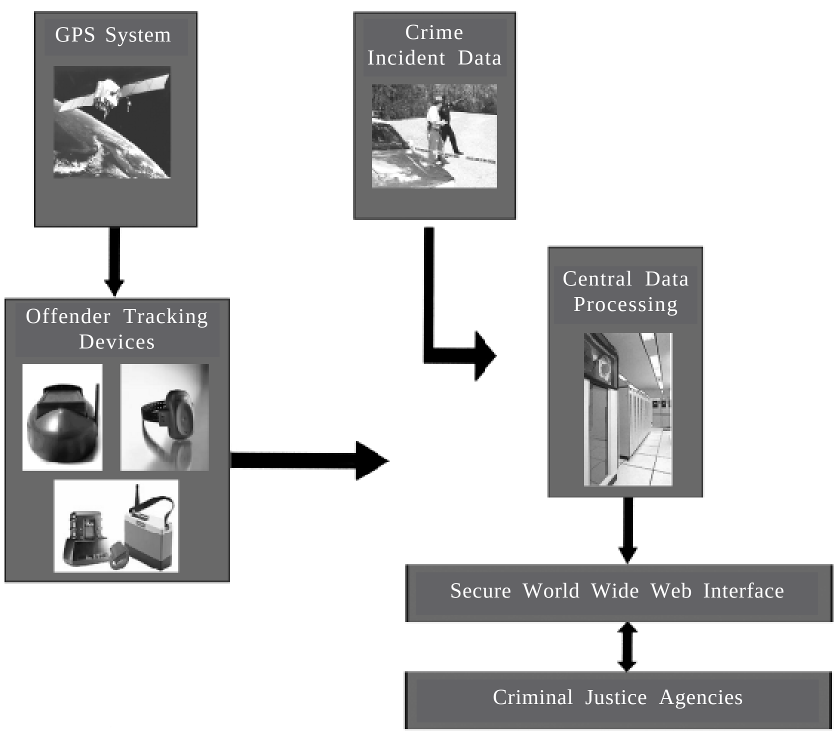
EMPACT Functionality

Offender track data and crime incident data are uploaded every night into the EMPACT system through automated extract, transform and load (ETL) programs. These programs allow crime incident data to be directly extracted from the local law enforcement agency's records management system, thereby eliminating redundant data entry. Automated analysis of the combined data reveals if a tracked offender was within a specified radius of a reported crime. If an incident "hit" is found, the reporting law enforcement agency is notified by automated e-mail message.

Through a secure user interface on the World Wide Web, EMPACT uses five distinct databases for its primary correlation feature – crime incident data, GPS track data, zone data, incident "hits" and zone "hits." The system was designed with extensive input from crime analysts, criminal investigators, intelligence officers and probation/parole officers. Because of their input, the interface is user-friendly for operational staff and allows queries to be answered in either text format or through detailed maps. Information technologies, such as EMPACT, offer limited value unless they can be used to clearly answer critical questions. As described in the following pages, the various ways EMPACT can be queried provides answers that criminal justice agencies need to control tracked offenders' behavior, to assist in identifying criminal suspects, and to hold tracked offenders more accountable.

Automated records management systems enable law enforcement agencies to compile a large number of crime incident records and effectively generate maps showing "hot spots" and the geographic migration of crime patterns over time.

Figure 1. EMPACT system overview



This data, however, is usually limited to a single jurisdiction. Interagency conflicts and disparate records management systems have limited the number of successful interjurisdictional automated crime analysis projects. In our post-9/11 terrorist environment, intelligence sharing on both a regional and national basis has made significant progress through interagency information networking. Even though GIS technology has taken crime analysis into a new era of automation, the sharing of incident level crime data has not made the same progress as law enforcement intelligence systems. The broad sharing of crime incident data is one of the valuable features of the EMPACT system. Data provided by participating law enforcement agencies is uploaded to the central server and is available to all other criminal justice users – whether it is a neighboring jurisdiction or one across the state.

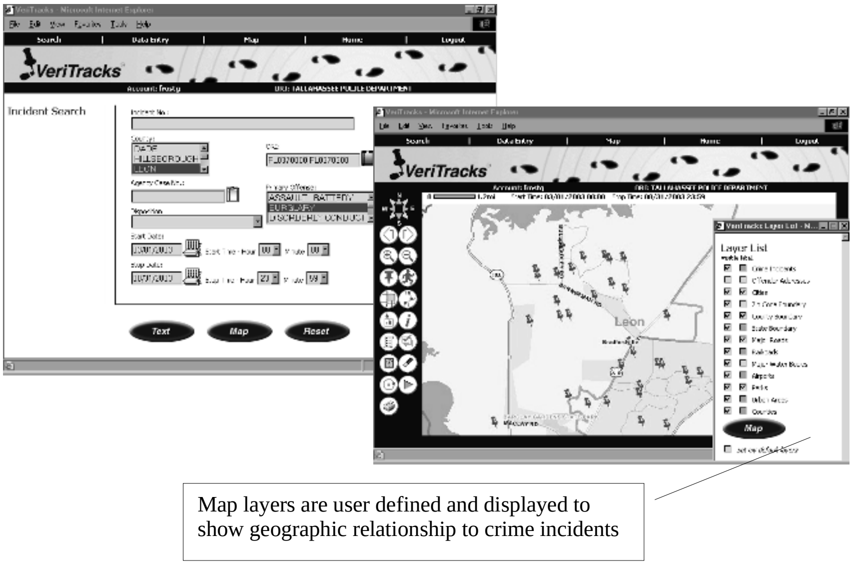
Common access creates the need for standardization of data. The ETL programs ensure that data from disparate records management systems are transformed to meet agreed upon system standards. The incident level data captured from each agency includes:

- Case Number
- Primary Offense
- Incident Location
- Start Date/Time
- Stop Date/Time
- Disposition
- Suspect Race
- Suspect Sex
- Suspect Age

Users query incident data through the system's incident query screen with query parameters set using a series of drop-down menus. Multiple parameters can be used based upon the question being asked. The capability of selecting multiple agencies allows the analyst to retrieve data showing any crime type, as well as determine interjurisdictional geographic patterns.

The following graphics show the query screen and resulting map of burglary offenses clustered around the jurisdiction boundary of the city of Tallahassee and unincorporated Leon County. The shaded area represents the city limit line. Finding this cluster allows for investigative follow-up and coordination between the two involved law enforcement agencies.

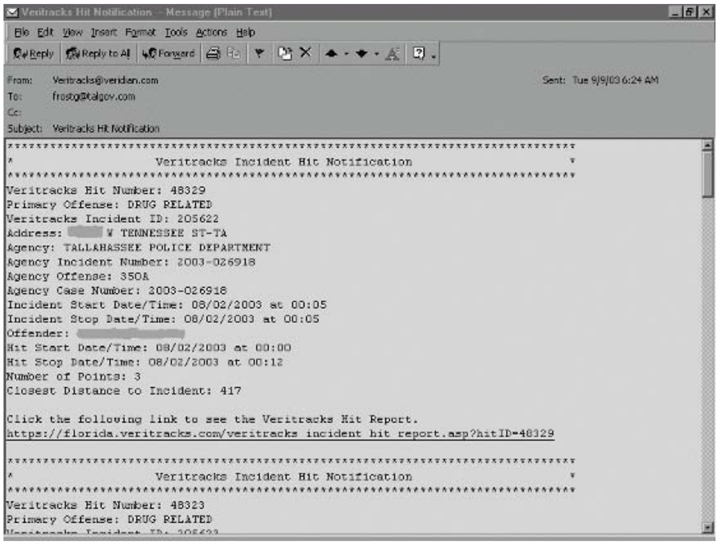
Figure 2. Multi-agency crime incident data



Knowing if a tracked offender was at the scene of a reported crime is one of the primary components of EMPACT. This is a benefit for law enforcement and community supervision agencies because it provides the leverage of telling tracked offenders that if they commit a crime and it is reported to a law enforcement agency, the agency will know that they were at the scene by the next day. While the impact of this on offender behavior has not been formally studied, common sense would indicate that only a small percentage of offenders will continue to commit crimes while being tracked.

The various GPS tracking devices currently on the market collect latitude and longitude points at least every minute as offenders move around their community. These track points, along with any device tampering alerts, are uploaded to the EMPACT server through the device providers' monitoring centers. Each night the uploaded crime data and offender track data are processed through a "hit engine" to determine if a tracked offender was within 1,000 feet of a reported crime, and if the offender was there within 30 minutes of the time it occurred. When an incident hit is detected, the information is e-mailed to the law enforcement agency reporting the crime and the hit's detailed data is recorded in the server for querying through the Web interface. Analysts or investigators

Figure 3. Incident hit e-mail notification



reviewing the e-mailed hit reports are able to quickly determine if additional information is needed based upon the basic information contained in the incident hit e-mail. A hyperlink is included in the e-mail to immediately link to the EMPACT Web interface so users can begin reviewing more detailed information.

Incident "hits" are queried through the Web interface using multiple user-defined parameters. Using the various fields shown in Figure 4, single or multiple offenders can be queried across various crime types and multiple jurisdictions.

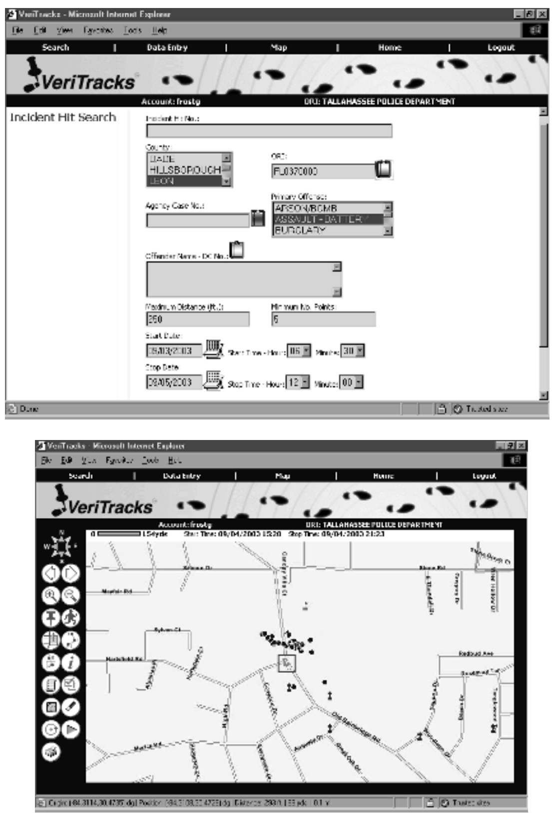
Correlation of offender tracks with crime incidents allows analysts and investigators to not only identify possible suspects, it also allows them to exclude specific tracked offenders from the suspect list by determining the offender was not near the crime scene when it occurred. Using EMPACT to quickly obtain exculpatory information saves overworked law enforcement agencies a significant amount of investigative work.

EMPACT contributes to the modification of behavior for tracked offenders through the use of "exclusion" and "inclusion" zones. Rhodes and Conly (1981) describe environmental factors used by criminals as part of their target identification process. Within any specific geographic area these factors build over time and provide criminal offenders with a sense of confidence that their criminal activity will not be detected. Using zones to restrict access to areas where offenders may have established patterns of criminality and victimization significantly enhances crime prevention by forcing tracked offenders into areas where they are less likely to commit new crimes.

Creating exclusion zones and informing the tracked offender of the zones' locations puts the offender on notice they must avoid the designated area in the case of an exclusion zone, or for inclusion zones, that they must stay within the zone. Any zone violation results in an alert being recorded in the EMPACT system and is included as part of the daily hit notices. If the offender is assigned an active tracking device and violates a zone restriction, an immediate alert is sent to the offender's community supervision officer for follow-up.

Neighborhoods with high incidents of crime, such as drug sales and prostitution, are areas where tracked offenders should not be allowed unless they have an

Figure 4. Correlation of crime incident data and GPS track points



approved, legitimate reason for being there. Creating exclusion and inclusion zones helps provide generalized crime control and offender surveillance in high crime neighborhoods. One of the ways the Tallahassee Police Department has successfully used exclusion zones is in an inner-city neighborhood with significant levels of drug dealing. Analysts and community policing officers established a zone that had positive results in reducing the ability of tracked offenders to easily obtain illegal drugs. Several days after activating the zone a hit was received showing a tracked offender traveling through the area when a more natural path of travel would have taken them through a different neighborhood. The hit provided information showing suspicious movement consistent with a drug purchase. Follow-up coordination with the offender's community supervision officer resulted in the offender not going back into the area.

Inclusion zones are used when a condition of community supervision includes curfews or monitoring of the offender's presence at work, school or social service facility. In these cases, the zone will have specified times when the tracked offender must be present inside the zone or alerts will be sent. Inclusion zones provide enhanced accountability for the offender without significantly impacting the workload of the supervising officer. Once the zone is created, the officer knows the tracked offender is meeting scheduled responsibilities and no monitoring is needed unless a violation alert is received.

Zones are created by users through the Web interface with the user able to define the location and size of the zone, the offender or offenders the zone applies to, and the days of the week and times of day the zone is to be active. Once

Figure 5. Exclusion zone violation

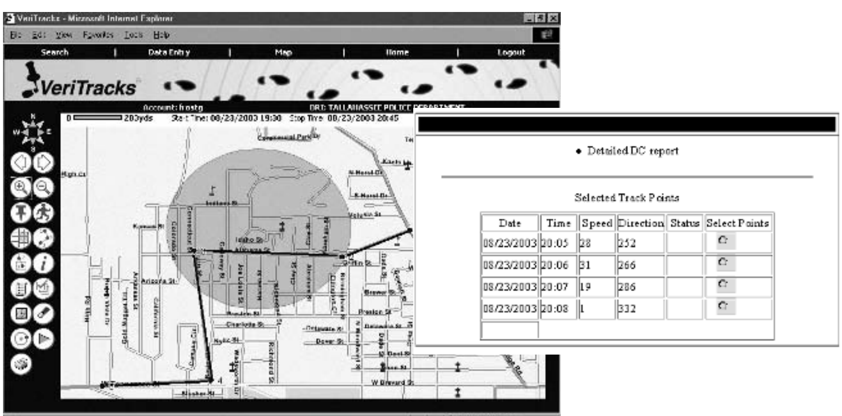
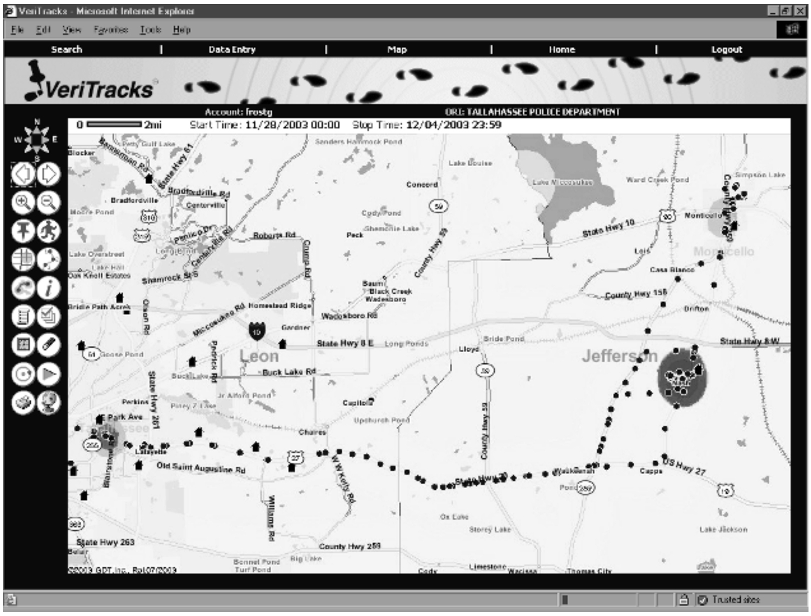


Figure 6. Offender home range map



created, any violation of the zone results in an e-mailed hit notice to the agency that created the zone and a zone hit record is automatically entered into the CrimeTrax system. Whether queried through the Web interface or directly through the hyperlink in an e-mail notice, the resulting map shows the offender's track points entering, traveling through, and exiting the zone.

By using the "identify" function of the mapping application, specific track point text information is available reflecting date and time of the points, speed, direction of travel and status of the GPS device including loss of GPS signal or tamper alert.

In addition to programmed zones, ad hoc geographic areas are queried to determine if a tracked offender was in a specific area. Users designate the date and time parameters for the map and then click and drag a box defining the geographic area to be questioned. The results reflect if any tracked offender(s) where present at the location and time selected.

The ability to quickly determine the locations where tracked offenders are spending time is important for probation and parole officers. EMPACT has a "home range" function that creates a density map for each offender using all the tracking points for that offender contained in the database. The home range is automatically recalculated for all offenders on a weekly basis to ensure accuracy overtime. Figure 6 shows the home range map – with three clearly identified "hot spots" – for a tracked offender along with specific track points for the time period queried.

Conclusion

Combining powerful technologies has led to the creation of a new criminal justice tool – one that will hopefully benefit our communities by reducing crime and by building stronger partnerships between corrections and law enforcement agencies. EMPACT is just beginning to emerge as a viable program within the United States and its full potential is still ahead. The ability to efficiently correlate GPS track data with multi-agency crime data and zone data has not been possible until recently. The immediate challenge is for criminal justice professionals, private technology firms and academic researchers to work together to determine the best way to maximize EMPACT's contribution to improving public safety.

References

Cohen, R. L. (1995, August). *Probation and parole violators in state prison, 1991: Survey of state prison inmates* (Report No. NCJ-149076). Washington, D.C.: Bureau of Justice Statistics.

Florida Department of Corrections. (2002). *Global positioning and radio frequency, 2002* (Table 2E). Retrieved from *http://www.dc.state.fl.us/pub/gpsrf/2002/tab2e.html*

Glaze, L. E. (2003, August). *Probation and parole in the United States, 2002* (Report No. NCJ-201135). Washington, DC: Bureau of Justice Statistics.

Harries, K. (1999). *Mapping crime: Principle and practice.* Washington, DC: National Department of Justice.

Hart, T. C, & Rennison, C. (2003, March). *Reporting crime to the police, 1992-2000* (Report No. NCJ-195710). Washington, DC: Bureau of Justice Statistics.

Langan, P.A., & Cunniff, M. A. (1992, February). *Recidivism of felons on probation, 1986-89* (Report No. NCJ-134177). Washington, DC: Bureau of Justice Statistics.

MacKenzie, D. L. (1997). *Preventing crime: What works, what doesn't, what's promising* (Ch. 9). Washington, DC: National Institute of Justice.

Rand Corporation. (1995). *GPS history, chronology, and budgets (Appendix B)*. Retrieved from *http://www.rand.org/publications/MR/MR614/MR614.appb.pdf*

Rhodes, W. M., & Conly, C. (1981). The criminal commute: A theoretical perspective. In P.L. Brantingham & P.J. Brantingham (Eds.), *Environmental Criminology* (Chapter 8). Beverly Hills, CA: Sage Press.

Ringel, C. (1997, November). *Criminal victimization 1996: Changes 1995-96 with trends 1993-96* (Report No. NCJ-165812). Washington, DC: Bureau of Justice Statistics.

Section V

New Methods and Technologies

Chapter XII

Simulating Crime Events and Crime Patterns in a RA/CA Model

Lin Liu, University of Cincinnati, USA

Xuguang Wang, University of Cincinnati, USA

John Eck, University of Cincinnati, USA

Jun Liang, University of North Carolina at Chapel Hill, USA

Abstract

This chapter presents an innovative approach for simulating crime events and crime patterns. The theoretical basis of the crime simulation model is routine activities (RA) theory. Offenders, targets and crime places, the three basic elements of routine activities, are modeled as individual agents. The properties and behaviors of these agents change in space and time. The interactions of these three types of agents are modeled in a cellular automaton (CA). Tension, measuring the psychological impact of crime events to human beings, is the state variable of the CA. The model, after being calibrated by using a real crime data set in Cincinnati, is able to

generate crime patterns similar to real patterns. Results from experimental runs of the model conform to known criminology theories. This type of RA/CA simulation model has the potential of being used to test new criminology theories and hypotheses.

Introduction

Many crime analyses attempt to model the relationships among various factors contributing to crime and measures of crime aggregated to an area units of analysis (Swartz, 2000). These types of analyses avoid modeling the processes of how individual participants' decisions give rise to crime events and how these events coalesce to form crime patterns. Although they are useful for describing aggregate crime patterns, they cannot reveal the underlying processes that generate crime patterns.

This chapter demonstrates an ongoing collaborative project started in 2000 on crime simulation between the Department of Geography and the Division of Criminal Justice at the University of Cincinnati. The goal of this project is to explore the possibility of simulating individual crime events and generating plausible crime patterns. It applies criminology theories and reasonable assumptions to explicitly model crime processes. These processes generate individual crime events, and the accumulation of these events then forms crime patterns. This project also aims to explore the potential of using this type of simulation models as a virtual laboratory for testing new crime theories and hypotheses and predicting future crime patterns based on different scenarios of policing and law enforcement strategies.

The RA/CA crime simulation model presented in this chapter is based on the integration of routine activities (RA) theory (Cohen & Felson, 1979), and cellular automaton (CA) simulation in GIS (Wu, 1999). Street robbery is used as an example to illustrate the characteristics of the model.

Modeling Crime with the Routine Activities Theory

The routine activities theory (RA), introduced by Cohen and Felson (1979) and expanded by Felson (1995), is a micro-level crime explanation theory that seeks to explain the occurrence of crime events as the confluence of several conditions. The first condition is that there must be a motivated offender. The second condition is that there must be a desirable target. Third, the target and offender must present at the same place and at the same moment (Felson, 1995; Eck & Weisburd, 1995).

Based on RA, a crime event is resulted from the interaction of an offender, target and a place where offender and target meet (Figure 1). In addition to the obvious role of offender in a crime event, the role of target and place is also important in explaining the crime. Targets may contribute to crime by creating opportunities for criminals. A place with good accessibility may facilitate crime because it provides convenient escaping routes for offender. The management of places can also be related to crime (Eck, 1994). Weak management facilitates crime because it lowers the level of control by place managers and guardians on crime.

Following the routine activities theory, Eck (1995) developed a formula to evaluate the instantaneous propensity for a crime event:

$$L(S_{tijk}) = \frac{\delta_{tik} T_{tik} \mu_{tijk} O_{tijk} \alpha_{ti} P_{ti}}{(1+\gamma_{tik} G_{tik})(1+\beta_{tijk} H_{tijk})(1+\varepsilon_{ti} M_{ti})} \qquad (1)$$

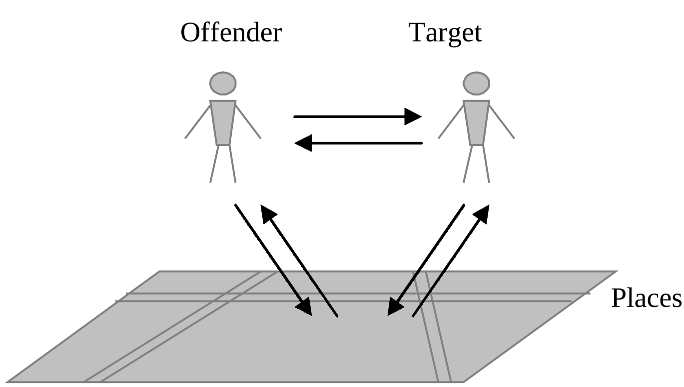
T: Target

δ: Desirability

G: Guardians

γ: Capability

Figure 1. Three basic elements of the routine activities theory (RA)



O: Offender

μ: Motivation

H: Handler

β: Intimacy

P: Place

α: Accessibility

M: Managers

ε: Effectiveness

$L\ (S_{tijk})$ is the crime likelihood for situation S_{tijk}. t is time, i is place, j is offender, k is offense type. The variables T, G, O, H, P and M represent respectively the presence of a target, a guardian, an offender, an intimate handler, a place and a place manager. If any of these six elements presents in situation S_{tijk}, the element takes the value of 1; otherwise it takes the value of 0. For a crime to occur, target (T), offender (O) and place (P) must be present. The presence of guardian (G), handler (H) and place manager (M) decreases the likelihood of crime occurrence. δ, γ, μ, β, α and ε represent target desirability, guardian capability, offender motivation, handler intimacy, place accessibility, and management efficiency. In street robbery, targets are people walking on streets; offenders are street robbers; guardians are target themselves, their companions and others who protect them; place managers are people who manage the place, such as store owners; intimate handlers are those who have direct personal influence over an offender, such as teachers, parents, friends and so forth.

In addition to explaining crime events and crime patterns, RA also has the potential of helping understand the behavior of offender, target and place manager. However, despite its significant theoretical contribution, RA has been rarely applied in crime models. Eck (1995) has documented the difficulties of applying RA into models. First, the theory is inherently nonlinear. The interaction between offender, target and place is complex. Second, data about the status of offender (motivation), target (desirability and guardian capability) and place (management efficiency) is usually unavailable. These difficulties are overcome with a simulation approach in this study.

Simulation with Cellular Automata in GIS

GIS-based simulation approach is suitable for studying space-time processes that involve a high level of complexity (Wu, 1999). The integration of GIS and cellular automata (CA) has drawn great attention for the past two decades (Takeyama & Couclelis, 1997). CA models have been used in ecological modeling (Maxwell & Costanza, 1997; Balzter, Braun & Kohler, 1998; Rajar, Cetina & Sirca, 1997), urban growth simulation (Batty & Xie, 1999; Clarke & Hoppen, 1997; Clarke & Gaydos, 1998), forest fire spread simulation (Clarke, Riggan & Brass, 1995), and transportation simulation (Wahle, Neubert, Esser & Schreckenberg, 2001; Blue & Adler, 2001) and so forth.

A CA has four basic elements that form a tuple (X, S, N, f). X is a set of cells that constitute the space of a study area. S is a nonempty finite set of automaton states. Each cell at a given time can be in only one state. N defines a neighborhood for a given cell. Typically, a cell has eight neighboring cells. The transition function f consists of a set of rules that determine the state of a cell at a future time based on the state of the cell at the current time and the state of its neighborhood cells at the current time. Wu (1999) defines a two-dimensional CA model as:

$$S_{rc}^{t+1} = f(S_{rc}^{t}, N_{rc}^{t}) \quad (2)$$

where $S_{rc}{}^{t+1}$ is the state of the cell at time t+1 and at cell (r, c), $S_{rc}{}^{t}$ is the state of the cell (r, c) at time t, $N_{rc}{}^{t}$ is the neighborhood space of cell (r, c), and f ( ) is the transition function.

The work of Liang, Liu and Eck (2001) represents the first application of CA to crime simulation, specifically to the simulation of commercial robberies. Their CA model can be used for examining crime theories and hypotheses on commercial robberies, but it cannot be calibrated with real crime data due to the limitation in its conceptual design. Therefore, its potential in real world crime modeling and analysis is limited. The model introduced in this chapter aims to overcome this limitation.

The RA/CA Crime Simulation Model

Based on the routine activities theory, crime events are directly related to the interaction of offender, target and place in a complex human-environment system. Whether a crime occurs depends on the motivation and capability of the offender, the desirability and protection of the target, the characteristics of the place, among other factors.

This study extends Eck's static formulation and applies it to a dynamic system to simulate individual crime events, which form crime patterns over time. Offender, target and place are modeled as artificial agents, each of which has its own properties and behaviors. The interaction of these agents mimics the interaction between offender, target and their environment.

The following sections use street robbery as an example to describe a prototype crime simulation model. Robbery is defined as "theft and attempted theft, directly from a person or commercial establishment, of property or cash by force or threat of force, with or without a weapon" (Ward & Ward, 1975, p. XX). Street robbery is a type of robbery whose victims are pedestrians.

I. Tension Generated by Crime

A crime event such as an incidence of street robbery, whether successful or failed, causes anxiety, fear, depression and hostility (Hollway & Jefferson, 2000; Norris, 1997). Following the work of Liang, Liu and Eck (2001), this study uses "tension" as a surrogate concept to represent the overall psychological reaction to a crime event. When a crime occurs, it adds tension to the crime location and to the target/victim.

II. The Agents in the Model

This project models place, offender and target as different types of agents. Each type has its own properties and behaviors.

(a) Place Agent

A place agent has the following properties: the accessibility of the place, place tension, and management effectiveness. It also has two behaviors to update tension and management effectiveness.

When a crime occurs in a place, the tension of the place increases. The neighboring areas of the place also experience increases in their tension. Tension gradually decreases in time. This behavior of a place agent is modeled in a cellular automaton.

Whenever tension increases or decreases in a place, the place manager may increase or decrease security to adjust the protection of the place. This is made possible by increasing or decreasing the place management effectiveness.

The accessibility of a place depends on its connectivity to streets and the capacity of the streets. A place at an intersection of two or more streets is more accessible than a place in the middle of a street, which is in turn more accessible than a place at the dead end of a street. The capacity of a street is related to the speed limit and the number of lanes. Broader streets with higher speed limits are more accessible than small alleys.

(b) Offender Agent

An offender agent has two properties: location and motivation. It also has two behaviors to change these two properties.

The location of an offender at a given day and time is related to the offender's routine activities. Empirical evidence suggests that many offenders do not travel a great distance to commit robbery (Wright & Decker, 1997). Most street robberies are committed in neighborhoods near, but not too close to, an offender's residence. It is reasonable to assume that the probability of an offender going to a place is inversely related to the distance from offender's home to the place (Brantingham & Brantingham, 1993; Block & Block, 2000), when the distance exceeds a small threshold value (approximately 0.5 miles). Based on this probability, a random process (Hagerstrand, 1967) assigns the location of an offender.

Motivation is updated according to the experience of an offender. An offender can increase its motivation after being encouraged by a successful robbery or decrease its motivation after being discouraged by a failed robbery. Typically, a novice offender tends to change its motivation at a faster rate than an experienced offender.

(c) Target Agent

A target agent has four properties: location, tension, desirability, and guardian capability. It also has four corresponding behaviors that update the value of these properties.

Ideally, a target's routine activities should be the determining factor in identifying the location of a target given the day and time. When these routine activities are unknown, a random process determines the placement of targets on streets. More targets are placed at the places that are more accessible.

Only the targets that have been attacked by street robbers have target tension. Target tension, like place tension, also decays in time.

We assume that most targets do what is necessary to decrease their desirability and increase their guardian capability after being robbed, to avoid future attacks. Those who cannot take proper protection may become repeat victims.

III. A Crime CA Model for Tension Propagation in Space and Time

This study applies a cellular automaton (CA) to model the propagation of tension in space and time. Like any other CA, this crime CA model also has four elements: cellular space, state variable, neighborhood template and transition function.

In street robbery, the cellular space consists of street locations. Non-street locations are not represented. Street segments are mapped into interconnected cells. Each cell location is associated with a place agent. Because a suitable spatial unit in the context of routine activities and street robbery should be not larger than the size of individual addresses and street corners (Eck, 1995), the cell size of this type of crime CA should be less than 100 feet.

Place tension, a property of the place agent, is the state variable in this crime CA model. A crime event at a place increases the tension of the place. Place tension decreases in time.

The neighborhood template of the crime CA model is a Moore neighborhood. Each cell has eight neighboring cells. The tension of these eight neighbors influences the future tension of the central cell.

The transition function consists of a set of transition rules that determine the place tension of a cell at time (t+1) based on two inputs: one is the tension of itself and its neighborhood cells at previous time (t); and the other is whether a crime is committed at the cell at the current time (t+1).

Following the work of Liang, Liu and Eck (2001), this CA model assumes that place tension decays in time and space. Let $S_{P0}{}^{t}$ denote the place tension value of cell (i, j) at time t, and $S_{P0}{}^{t+1}$ denote the place tension value of cell (i, j) at time t+1. Temporal decay of place tension is represented as:

$$S_{P0}^{t+1} = S_{P0}^{t}(1 - K_{Ptd}) \tag{3}$$

where K_{Ptd} is a temporal decay coefficient. A larger coefficient results in faster temporal decay. The value for K_{Ptd} is larger than 0 but smaller than 1.

The spatial diffusion of place tension is determined by the difference of tension between the cell (*i*, *j*) and its eight neighboring cells at time *t*:

$$Sigma = \sum_{k=1}^{8} (S_{Pk}{}^{t} - S_{P0}{}^{t}), where (S_{Pk}{}^{t} > S_{P0}{}^{t}) \tag{4}$$

where $S_{Pk}{}^{t}$ denotes the place tension of neighborhood cells of cell (*i*, *j*) at time *t* and *Sigma* the sum of the difference between $S_{Pk}{}^{t}$ and $S_{P0}{}^{t}$. Because people tend to pay attention to bad news or high tensions, this model only considers those $S_{Pk}{}^{t}$ values that are larger than $S_{P0}{}^{t}$. The role of *Sigma* in spatial diffusion of tension is formulated as:

$$S_{P0}^{t+1} = S_{P0}^{t+1} + K_{sd} \times Sigma \tag{5}$$

where K_{sd} is a spatial diffusion coefficient, representing the portion of *Sigma* propagated from the neighboring cells to the cell (*i*, *j*). The range of K_{sd} is (0, 1). The $S_{P0}{}^{t+1}$ on the right side of Equation (5) is the same as the $S_{P0}{}^{t+1}$ on the left side of Equation (3).

Finally, if a robbery occurs at cell (*i*, *j*) at time *t*+1, a tension M_P caused by this robbery is added to the exiting tension.

$$S_{P0}^{t+1} = S_{P0}^{t+1} + M_P \tag{6}$$

The $S_{P0}{}^{t+1}$ on the right side of Equation (6) is same as the $S_{P0}{}^{t+1}$ on the left side of Equation (5). The $S_{P0}{}^{t+1}$ on the left side of Equation (6) is the final tension value of cell (*i*, *j*) at time *t*+1.

To explain a crime event, the time interval in the model should be not longer than the time required to commit a crime (Eck, 1995). In reality, a street robbery usually takes only a few seconds. Ideally, the time interval of the CA model should be set to a few seconds. If the time interval is set to five seconds, 6,307,200 iterations (365 days * 24 hours/day * 60 minutes/hour * 60 seconds/minute / five seconds/iteration = 6,307,200 iterations) are needed for the CA to simulate crime for an entire year. It would take several days or weeks on a personal computer to complete the simulation. This seems impractical. Based on the crime database provided by the Cincinnati Police Department, the city has about 450 street robberies per year on average. It is reasonable to consider street robberies as rare events in the space-time continuum. For the practicality of simulating such rare events, the time interval of this CA model is set to no more than 15 minutes. This time interval is still considerably shorter than what is observed in traditional criminological studies.

IV. Likelihood of Crime

In Equation (1), if a target (*T*), an offender (*O*), or a place (*P*) is not present, the crime likelihood is 0. In addition, a crime is not likely to occur if a handler (*H*) is present (Eck, 1995). For the purpose of street robbery simulation, this study only focuses on situations where the crime likelihood is not 0, which means that *T*, *O*, and *P* must be set to 1 and *H* to 0. Because Guardians (*G*) can be targets themselves in street robbery, G is 1. The number "1" in Equation (1) does not represent anything meaningful other than guarantees the denominator not become 0. Here it is changed to "0.1" to make the effect of management efficiency ε and guardian capability [3] more significant. With these simplifications, Equation (1) becomes:

$$L = \frac{\delta\mu\alpha}{(0.1+\varepsilon)(0.1+\gamma)} \tag{7}$$

The crime likelihood is now positively related to target desirability δ, offender motivation μ and place accessibility α, and inversely related to management efficiency ε and guardian capability γ. If δ or μ or α is 0, *L* is 0; if δ, μ and α are all set to 1, ε and γ are set to 0, *L* becomes 100. The range of *L* is thus [0, 100].

It is reasonable to assume that a crime does not occur when *L* is equal or less than a randomly generated threshold value b_0. If *L* is larger than b_0, a robbery occurs, but the outcome of the robbery can be successful or unsuccessful (Wright & Decker, 1997). To simulate these two scenarios, a smaller random value b_1 is generated and added to the threshold value b_0. If *L* is larger than $b_0 + b_1$, a crime is successful. Conversely, a street robbery fails if *L* is in-between b_0 and $b_0 + b_1$. These two random values control the random errors of the model.

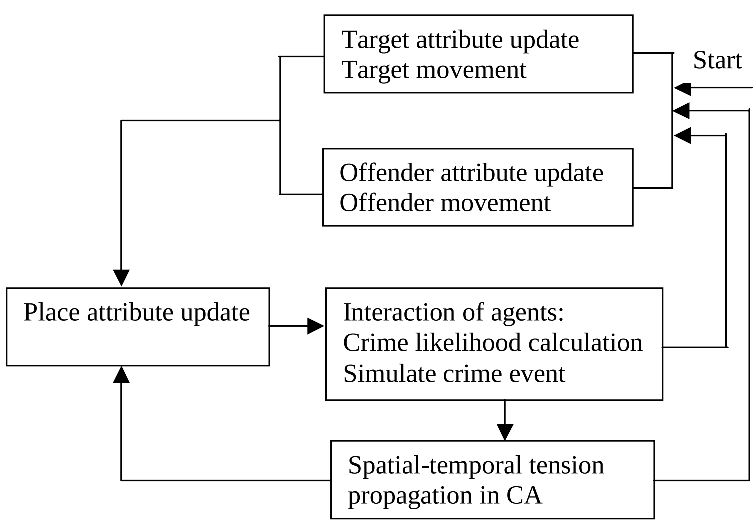
When a street robbery occurs at a location, all three agents react to the crime. For the place agent, its tension increases, and the increased tension influences its management practice. The target agent also has increased tension, which further impacts its desirability and guardian capability. For the offender agent, a successful robbery increases its motivation, while a failed attempt decreases its motivation.

The above has described the interaction between offender agent, target agent and place agent at a specific time. In reality, such interaction changes in time and space. To mimic the changing interaction, agents with updated properties are applied to the next iteration of the crime CA model, which propagates tension in space and time. The change of tension in the CA model again updates the properties of the agents. These updated properties again serve as the inputs to the next iteration of the CA model. This continuously changing interaction mimics the dynamic system in reality.

V. Model Summary

Figure 2 illustrates the interconnections between various components of this RA/CA crime simulation model and how the model works in a series of iterations. The model starts with setting the attributes of the agents. For each interaction of offender, target and place, the model calculates crime likelihood and determines whether a crime event occurs. The outcome of each interaction is then used to update the properties of place, offender and target. The CA simulates the spatial

Figure 2. RA/CA crime simulation model



propagation of tension and temporal decay of tension, which further updates the attributes of the agents. When the iteration is completed, another iteration follows and this process is repeated until the specified number of iterations is reached.

Model Calibration

The above RA/CA crime simulation model has been implemented in Visual C++ and applied to a small neighborhood area near downtown Cincinnati. This is mainly a residential area, with a number of convenient stores, bars, clubs and so forth. Mapping the calls for service data related to street robbery from 1997 to 1998 reveals that street robberies are concentrated near bars, clubs, abandoned houses, worn-out buildings, large public housing projects and street intersections.

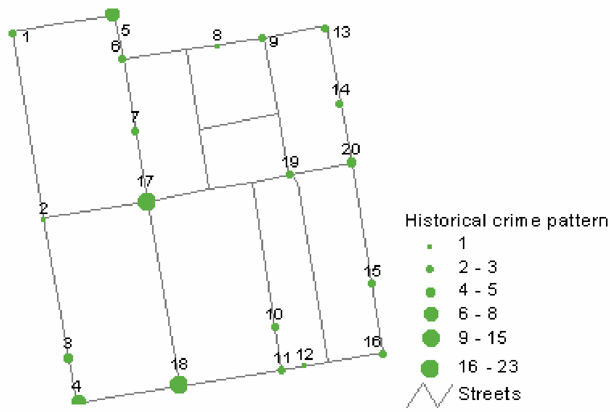
Calibrating the parameters of a simulation model to fit a real data set involves a complex process. A vision test (Clarke & Hoppen, 1997) is used to obtain an initial estimate of the parameters. Then fine-tuning ensures that the total number of simulated crime is close to the actual number and the spatial distribution of simulated crime similar to that of real crime.

In addition to model calibration, the stability/consistency of simulation results is also important in establishing the validity of the model. If such calibration fails to produce patterns that reasonably mimic reality, then the model is seriously flawed. Because the model contains random processes, multiple runs of a valid model will produce small variations in the results.

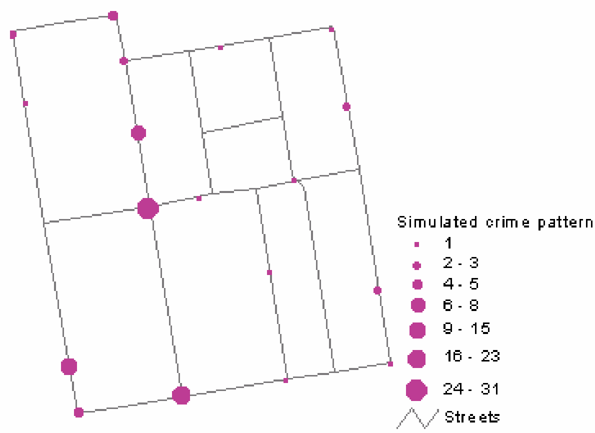
The number of crimes in the area, generated by the model in 500 different runs of the model for a two-year period, has a mean of 98 and a standard deviation of 10. The mean is very close to the actual number of 94 robberies in this area. The relatively small deviation suggests that the model generates consistent results.

Figure 3 shows the spatial pattern of actual street robbery in the area (3a) and that of simulated street robbery (3b). Visual comparison shows that the two patterns are similar. The calibrated simulation model is capable of generating crimes similar to actual crimes in both the total number of crimes and their spatial distribution. These results suggest that the model is valid within a reasonable degree of confidence, even though the confidence cannot be tested statistically.

Figure 3. Simulated crime spatial pattern compared with actual crime pattern



(a) actual spatial pattern of street robbery from 1997 to 1998



(b) simulated crime spatial pattern

Results

Using the calibrated parameters, experimental runs of the model yield the following plausible results:

- **Repeat location:** it is evident on Figure 3b that the simulated street robberies are concentrated in a few locations, which are repeat locations. This is consistent with research on crime places (Eck & Weisburd, 1995).

- **Repeat victimization:** more desirable targets are victimized more frequently than less desirable targets. In one of the experiments, the target population is divided into four groups, each group having 25 people and a different desirability coefficient. The targets in the first group, the most desirable group, are victims of 42% of all simulated street robberies. This is consistent with research on victimization (Farrell, 1995).
- **Repeat offending:** a small number of offenders are responsible for a disproportionately large amount of crime. The result of an experimental run of this model indicates that 35% of the offenders committed 68% of street robberies. This is consistent with the research on offending (Spelman, 1994).

Situational crime prevention claims opportunity for crime can be reduced by increasing risks and difficulties and reducing rewards of crime (Clarke, 1992). The experimental results of the simulation model support the following:

- Increasing management efficiency of all place agents by 25% (for example, install surveillance devices or police hotspot patrolling), as a means of increasing crime risks, reduces the total number of crime by 42%. Increasing management efficiency of place agents at site #5, #17 and #18 (the three sites with the highest crime counts) by 25% cuts the total crime by 31%. This provides insight on where limit resources should be allocated to achieve the most impact on crime reduction.
- Increasing targets' guardian capability by 8% (for example, walking with a friend instead of walking alone), as a means of increasing crime difficulties, reduces the total number of crime by 17%.
- Decreasing targets' desirability by 10%, as a means of reducing the rewards of crime, reduces the total number of crime by 10%.

These results are qualitatively consistent with the crime prevention evaluation literature (Eck, 2002; Sherman & Eck, 2002). Overall, the results of this RA/CA model are highly plausible. This suggests that research using RA/CA models has potential as a tool for improving our understanding of the development and control of crime patterns. In the long run, models such as this could be used for "bench testing" policies prior to field experimentation and implementation.

Conclusion

The RA/CA crime simulation model presented in this chapter is based on the integration of routine activities theory (RA) and cellular automaton (CA). It has been implemented by using the Visual C++ computer language. Results from experimental runs of the model are plausible and conform to known criminology theories. When such model is extensively tested, it has the potential of being used as a virtual laboratory for testing new crime theories and hypotheses.

The limitation of such model lies in its parameter calibration. At the present time, the calibration of any CA based simulation model with a large number of parameters is not an exact science. It very much relies on the experience and expertise of the user. The (global) best calibration may never be attainable due to the computation complexity. However, an (local) optimal calibration may be sufficient, just like any heuristic numeric method that often finds a local optimal solution, yet scholars use these heuristic methods in scientific research on a daily basis.

References

Balzter, H., Braun, P. W., & Kohler, W. (1998). Cellular automata models for vegetation dynamics. *Ecological Modeling, 107*, 113-125.

Batty, M., Xie, Y. & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems, 23*, 205-233.

Block, R., & Block, C. R. (2000). The Bronx and Chicago: Street robbery in the environs of rapid transit stations. In V. Goldsmith, P. G. McGuire, J. H. Mollenkopf & T. Ross (Eds.), *Analyzing crime patterns: Frontiers of practice* (pp. 137-152). Thousand Oaks, CA: Sage.

Blue, V. J., & Adler, J. L. (2001). Cellular automata microsimulation for modeling bi-directional pedestrian walkways. *Transportation Research Part B: Methodological, 35*, 293-312.

Brantingham, P. L., &. Brantingham, P. J. (1993). Environment, routine, and situation: Toward a pattern theory of crime. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice: Advances in criminological theory, vol. 5.* (pp. 259-294). New Brunswick, NJ: Transaction.

Clarke, K. C., & Gaydos, L. J. (1998). Loose-coupling a cellular automata model and GIS: Long-term urban growth prediction for San Francisco and

Washington/Baltimore. *International Journal of Geographical Information Science, 12*, 699-714.

Clarke, K. C., & Hoppen, S. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay Area. *Environment and Planning B: Planning and Design, 24*, 247-261.

Clarke, K. C., Riggan, P., & Brass, J. A. (1995). A cellular automaton model of wildfire propagation and extinction. *Photogrammetric Engineering and Remote Sensing, 60*, 1355-1367.

Clarke, R. V. (1992). Part I: Introduction. In R. V. Clarke (Ed.), *Situational crime prevention: Successful case studies* (pp. 3-36). New York: Harrow and Heston.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activities approach. *American Sociology Review*, *44*, 588-608.

Eck, J. E. (1994). *Drug markets and drug places: A case-control study of the spatial structure of illicit drug dealing* (Ph.D. dissertation). College Park, MD: University of Maryland, Department of Criminology and Criminal Justice.

Eck, J. E. (1995). Examining routine activities theory: A review of two books. *Justice Quarterly, 12*, 783-797.

Eck, J. E. (2002). Preventing crime at places. In L. W. Sherman, D. Farrington, B. Welsh & D. L. MacKenzie (Eds.), *Evidence-based crime prevention* (pp. 241-294). New York: Routledge.

Eck, J. E., Gersh, J. S., & Taylor, C. (2000). Finding crime hotspots through repeat address mapping. In V. Goldsmith, P. G. McGuire, J. H. Mollenkopf & T. Ross (Eds.), *Analyzing crime patterns: Frontiers of practice* (pp. 49-62). Thousand Oaks, CA: Sage.

Eck, J. E., & Weisburd, D. (1995). Crime places in crime theory. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 1-33). Monsey, NY: Criminal Justice Press.

Felson, M. (1995). Those who discourage crime. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 53-66). Monsey, NY: Criminal Justice Press.

Farrell, G. (1995). Preventing repeat victimization. In M. Tonry & D. P. Farrington (Eds.), *Preventing repeat victimization* (pp. 469-534). Chicago: University of Chicago Press.

Hagerstrand, T. (1967). *Innovation diffusion as a spatial process*. Chicago: University of Chicago Press.

Hollway, W., & Jefferson, T. (2000). The role of anxiety in fear of crime. In T. Hope & R. Sparks (Eds.), *Crime risk and insecurity* (pp. 31-49). New York: Routledge.

Liang, J., Liu, L., & Eck, J. (2001). Simulating crimes and crime patterns using cellular automata and GIS. *UCGIS Student Award Papers 2001*. Retrieved from *http://www.ucgis.org/f2oppor.html*

Maxwell, T., & Costanza, R. (1997). A language for modular spatio-temporal simulation. *Ecological Modeling, 103*, 105-113.

Morrill, R., Gaile, G. L., & Thrall, G. I. (1988). *Spatial diffusion*. Newbury Park, CA: Sage.

Norris, F. H. (1997). The psychological consequence of crime. In R.C. Davis, A. J. Lurigio & W. G. Skogan (Eds.), *Victims of crime* (pp. 146-166). Thousands Oaks, CA: Sage.

Rajar, R., Cetina, M., & Sirca, A. (1997). Hydrodynamic and water quality modeling: Case studies. *Ecological Modeling, 101*, 209-228.

Sherman, L. W., & Eck, J. E. (2002). Policing for crime prevention. In L. W. Sherman, D. Farrington, B. Welsh & D. L. MacKenzie (Eds.), *Evidence-based crime prevention* (pp. 295-329). New York: Routledge.

Spelman, W. (1994). *Criminal incapacitation*. New York: Plenum.

Swartz, C. (2000). The spatial analysis of crime: What social scientists have learned. In V. Goldsmith, P. G. McGuire, J. H. Mollenkopf & T. Ross (Eds.), *Analyzing crime patterns: Frontiers of practice* (pp. 33-46). Thousand Oaks, CA: Sage.

Takeyama, M., & Couclelis, H. (1997). Map dynamics: Integrating cellular automata and GIS through geo-algebra. *International Journal of Geographical Information Science, 11*(1), 73-91.

Wahle, J., Neubert, L., Esser, J., & Schreckenberg, M. (2001). A cellular automaton traffic flow model for online simulation of traffic. *Parallel Computing, 27*, 719-735.

Ward, R. H., Ward, T. J., & Feeley, J. (1975). Chapter 1. Introduction. In *Police Robbery Control Manual*. Washington, D.C.: U.S. Department of Justice.

Wright, R. T., & Decker, S. H. (1997). *Armed robbers in action: Stickups and street culture*. CITY, STATE: Northeastern University Press.

Wu, F. (1999). GIS-based simulation as an exploratory analysis for space-time processes. *Journal of Geographical Systems, 1*, 199-218.

Chapter XIII

Integrating GIS and Maximal Covering Models to Determine Optimal Police Patrol Areas

Kevin M. Curtin, University of Texas at Dallas, USA

Fang Qiu, University of Texas at Dallas, USA

Karen Hayslett-McCall, University of Texas at Dallas, USA

Timothy M. Bray, University of Texas at Dallas, USA

Abstract

This chapter presents a new method for determining the most efficient spatial distribution of police patrols in a metropolitan region, termed the police patrol area covering (PPAC) model. This method employs inputs from geographic information systems (GIS) data layers, analyzes that data through an optimal covering model formulation, and provides alternative *optimal solutions for presentation to decision makers. The goal of this research is to increase the level of police service by finding more efficient spatial allocations of the available law enforcement resources. Extensions to the model that incorporate variations in the priority of calls for service based on the type of crime being committed, and the need for an equitable distribution of workload among police officers are discussed. Examples of the inputs from – and outputs to – GIS are provided through a pilot study of the city of Dallas, Texas.*

Introduction

Virtually all metropolitan police departments in the U.S. create a geographic division of their area for the purposes of administration and patrol. The way in which this spatial division is made influences the provision of police services. An optimal spatial division could equitably distribute limited police resources throughout the city, reduce response times, save money through efficient deployments, and create a fair division of risk among police officers.

The primary goal of this research is to provide suggestions that will lead to an increase in the level of police service by finding more efficient spatial allocations of law enforcement resources. We take care to note here that efficiency can be a broadly defined term – with multiple metrics – and we restrict ourselves to the spatial efficiency of the patrol-area boundaries. The resulting alternatives take the form of optimal solutions to maximal covering problems, which delineate police patrol areas. These problems must be solved in the context of a major metropolitan area, with a large population and a concomitantly large number of potential police patrol areas. It is proposed that covering models hold the greatest promise for determining optimal solutions to the problem of delineating police patrol areas. Given the difficulties inherent in optimally solving large instances of such problems, both GIS and integer programming software must be integrated in order to allow these problems to be re-examined when the initial conditions change. Given the many people and organizations who must approve the changes suggested by the optimal solutions that are found, the results must be presented as a series of alternatives from which a best arrangement can be selected that satisfies the greatest number of people and that can be efficiently implemented.

The relevant literature that forms the basis for this research falls into three major categories: the discipline of location science and its integration with GIS, the determination of police patrol areas, and the formulation of covering models for service provision to geographic areas. In addition to these three categories, there are pertinent bodies of literature surrounding the methods for solving problems optimally, which will be discussed in more detail in the methodology section below.

Location Science and Optimization

The field of operations research (OR) – also termed management science – is defined as the generation and application of advanced analytical techniques in order to solve complex problems and make organizational decisions (Curtin, 2004). Often these problems involve the allocation of scarce resources in such a way as to maximally achieve a goal (such as profit or level of service) or minimize a negative consequence of the operation of an organization (such as cost or environmental degradation). Although this discipline has its origins in the application of problem-solving techniques in a military context, a wide range of industrial, transportation, social and ecological applications have been developed over the past several decades (Hillier & Lieberman, 1995).

A substantial subdiscipline within OR is the field of location science, where the geographic location of facilities or activities within the system or organization is a primary determinant of the optimal solution (Hale, 2004). Within location science, a sub-set of problems and methods are concerned with both the location of facilities *and* the allocation of demand to those facilities, and research in this area is termed location-allocation theory. Unfortunately, large instances of many problems in location science are practically impossible to solve optimally. While it is conceivable that every possible competing solution could be compared through an evaluation of the objective function (known as enumeration of all possible solutions), this method quickly becomes impractical. The number of alternatives is a function of the number of facilities to be located and the number of potential facility locations. In general the number of alternatives will be:

$$\left(\frac{n}{p}\right) = \frac{n!}{p!(n-p)!}$$

where n = the number of potential facility locations, and p = the number of facilities that one wants to locate among those potential sites. The number of alternatives can thus grow rapidly as the values of n and p increase. So rapidly, in fact, that as the size of a problem of this type increases, it becomes impossible to enumerate all answers in a reasonable amount of time. If a decision problem is intrinsically harder to solve than those that can be solved by a non-deterministic

Turing machine in polynomial time is classified as NP-Hard (NIST, 2001). Optimization versions of these decision problems are termed NP-Complete.

In order to solve the complex problems posed within location science, these problems must be formulated in such a way that they can be efficiently analyzed, and this generally requires the formulation of a mathematical model. These models commonly take the form of an objective function, which defines the goal of the organization (or one of many goals), and a set of constraints representing the conditions within which the system must operate. Once the general version of a problem is formulated, individual instances of that problem can be solved optimally in order to suggest specific distributions of the organization's resources. The optimal solution will be the one that best satisfies the objective that had been optimized. In the context of the research presented here, one potential objective is to provide the highest level of service by delineating the optimal arrangement of police patrol areas. The success of that solution can be measured by the number of incidents to which the police can respond within an acceptable service time. We caution that this approach does not suggest that all administrative issues be simplified through a broad acceptance of response-time reduction. Instead, this approach, like others, should be used to determine appropriate alternative deployments for particular types of calls, and used in the context of overall agency objectives. The constraints on this system include the number of police patrols that are available, the level of workload that can be assigned to each patrol, and numerous other economic and legal restrictions on police activity.

The application areas to which location science has been applied are far too numerous to list here, but it is clear that the efficiencies that can be gained by mathematically modeling a system and solving that model optimally are significant and can in some cases be extraordinary. Recent findings show that major corporations have realized savings in the hundreds of millions of dollars attributable to the implementation of optimization techniques (INFORMS, 1999). In the context of the research presented here it is reasonable to expect that these methods could result in savings in terms of money spent on officer salaries, decreased response times and increased revenues from citations based on the more efficient deployment of officers.

Location Science and GIS for Law Enforcement

Although operations research has proven its worth in applications to a wide range of problems, the issue at hand is whether or not it can benefit an application in the field of law enforcement, and particularly in the determination of police patrol areas. The division of an area by a police force is fundamentally a geographic problem. Commonly, a city or metropolitan area is divided into police command areas, variously termed precincts, districts or divisions. These command areas are further divided into patrol areas, sometimes called beats or sectors (Larson, 1978). Several sources confirm that prior to 1972 police patrol areas were determined "by hand," where a person was responsible for drawing police patrol areas on a map based on their knowledge of the total area to be patrolled by the police force and the available police resources (Mitchell, 1972; Taylor & Huxley, 1989). No other description of formal analysis entered into the procedure for determining police patrol areas, nor was there a quantitative method for evaluating how the hand drawn boundaries compared to an optimal arrangement or for comparing alternative deployment schemes. As late as 1986, a study of departments found that no integer optimization models were in use. This pervasive and persistent lack of formal procedures for police patrol area development has been seen to complicate higher-level policy decision making due to the lack of objective quantitative measures of efficiency (Taylor & Huxley, 1989).

The first application of mathematical modeling to the determination of police patrols known to the authors was a formulation and application for Anaheim, California (Mitchell, 1972). This application used a version of the p-Median problem to minimize the demand weighted distance between sectors of the city. The demand weights were a function of the expected number of incident calls in each sector. Several distance metrics were presented and suggestions were made for the differentiation of incidents based on type, for multiple unit response to incidents, and for constraints on work load and maximum response distance. In that application no proven optimal solutions were found, but heuristic solutions based on the Maranzana heuristic (Maranzana, 1964) for dividing a geographic region were determined. Even so, the application of heuristic methods to solve this mathematical model resulted in a 13% to 24% reduction in average response distance when compared to the hand drawn district boundaries.

At about the same time a series of publications of research supported by the Rand Corporation presented the development of a hypercube queuing model for the deployment of police assigned to different police patrol areas (Larson, 1975). A mathematical model of the hypercube could be used to find the optimal deployment pattern *for a pre-determined set of police patrol areas*. An approximation algorithm was also developed to solve these problems more quickly, albeit without a guarantee of optimality. Once a beat plan has been established (by hand) the hypercube queuing model is used to distribute police resources to calls in an efficient manner. Several performance measures of the efficiency of the patrol area plan can be generated (Chaiken & Dormont, 1978b). These measures can be used to compare different beat plans. However, the hypercube queuing model does not determine the optimal arrangement of police patrol areas, rather it allocates police resources within an existing arrangement.

Therefore the performance of a particular patrol area arrangement cannot be compared to the optimal spatial arrangement.

Variations of the hypercube queuing model have been applied in St. Louis County, Missouri (Kwak & Leavitt, 1984), and New Britain, Connecticut (Sacks, 2000). In these applications, it is suggested that the police department ought to try redrawing different arrangements of districts in order to compare the solution values with the performance measures generated by the hypercube queuing model. Given the combinatorial complexity of the districting problem it is highly unlikely that the optimal solution will be determined in this way, and there is no way of knowing if the optimal solution to the problem of drawing district boundaries has been reached through trial and error.

Where operations research techniques have been employed in law enforcement contexts it has been noted that the results of these analyses are sometimes the only quantitative information provided to decision makers (Aly, Litwhiler & Heggy, 1982). However, virtually all of the applications of operations research have focused on scheduling (Chaiken & Dormont, 1978a; Taylor & Huxley, 1989), and it appears that additional models are necessary to determine the optimal arrangement of police patrol areas which can then be used to allocate police resources with an efficient schedule. This research employs a set of models of the type known as covering models to address the current deficiency in the literature and in practice.

In contrast to operations research techniques, geographic information systems have become widely accepted among police departments as a valuable tool for a wide range of applications. Perhaps most importantly, GIS has been used to assist in the determination of clusters of crime activity, termed hot-spots (Craglia, Haining & Wiles, 2000; Harries, 1999). However, even well-informed and experienced GIS users labor under the misconception that state-of-the-industry GIS packages can optimally solve combinatorial optimization problems with the push of the button. While they are extraordinarily useful tools, a close examination of the supporting documentation for GIS software packages demonstrates that they are NOT designed to solve such problems optimally (ESRI, 2004). In fact, only a very limited number of problems can be solved, and even this limited set of problems can only be solved heuristically, generally using versions of interchange heuristics such as the Tietz and Bart (Teitz & Bart, 1968) or GRIA (Zanakis, Evans & Vazacopoulos, 1989) heuristics. The difficulties in solving combinatorially complex location science problems within GIS have been well documented (Church, 2002), and the heuristic solution procedures have been described as providing "a minefield of local optima" (Church & Sorenson, 1994), which can lead to substantially suboptimal solutions. Even if the heuristic solution procedures embedded in out-of-the-box GIS packages provide a good solution, there is no way of knowing whether or not the optimal solution has been determined, or how close the solution is to optimal. Although there are superior heuristic procedures such as those that employ simulated annealing (D'Amico, Wang, Batta & Rump, 2002), these heuristics require sophisticated users to test and apply parameters. For this reason they are not built into off-the-shelf GIS software packages. Due to this limitation, this research demonstrates that GIS can be integrated with integer programming solution software to find optimal solutions to the problem of delimiting police patrol areas. This research presents an optimal covering model as a reasonable and flexible choice for designing police patrol areas when integrated with GIS.

Covering Models

The maximal covering location problem (MCLP) was first formulated in 1974 (Church & ReVelle, 1974). The MCLP seeks to find the solution to the problem of locating facilities (such as police cars on their beats) in such a way as to maximize the number of incidents that can be served within a given service distance (or response time). Because the MCLP has been shown to be NP-Complete, robust solution procedures must be developed to allow the optimal solution to be found. A number of solution procedures and reformulations of the problem to facilitate solution have appeared in the literature. Additionally, the MCLP has been related theoretically to other prominent location models including the p-Median model and the location set covering model (Church & ReVelle, 1976). These links between models allow solution procedures developed for one of the problems to be applied to the others. Heuristic solution procedures such as the TABU search heuristic (AdensoDiaz & Rodriguez, 1997) and Langrangean relaxation heuristics (Galvao, Espejo & Boffey, 2000; Galvao & ReVelle, 1996) have also been developed for the MCLP.

Additional areas of research that can provide insights into potential problems when applying covering models to the determination of police patrol areas include an analysis of data aggregation errors (Current & Schilling, 1990) and the inclusion of capacities on workloads (Pirkul & Schilling, 1991). A variant of the MCLP has been developed to ensure that not only is coverage maximized but that travel times or distances to service demand outside the maximal covering distance is minimized (Church, Current & Storbeck, 1991). The formulation of the maximal conditional covering problem suggests that "backup" coverage can also be modeled for police patrol areas (ReVelle, Schweitzer & Snyder, 1996).

Covering models have been applied to the location of emergency warning sirens (Current & Okelly, 1992), the location of ambulance bases in a rural region (AdensoDiaz & Rodriguez, 1997), integrated fire and ambulance siting (ReVelle & Snyder, 1995), the location of retail facilities (Berman & Krass, 2002), and ecological reserve selection (Church, Stoms & Davis, 1996). A review of applications of the MCLP that do not involve geographic location (Chung, 1986) found that the model was proven useful for data abstraction and statistical classification. To date no application of the MCLP to the determination of patrol areas has appeared in the literature.

Police Patrol Area Covering (PPAC) Model

Maximal covering models can be applied to the problem of generating optimal police patrol areas with the following formulation:

$$\text{Maximize } Z = \sum_{i \in I} a_i y_i$$

Subject To :

$$\sum_{j \in N_i} x_j \geq y_i \quad \text{for all } i \in I \qquad (1)$$

$$\sum_{j \in J} x_j = P \qquad (2)$$

$$x_j = (0,1) \quad \text{for all } j \in J \qquad (3)$$

$$y_i = (0,1) \quad \text{for all } i \in I \qquad (4)$$

Where:

I = the set of known incident locations;

J = the set of potential locations for police patrols;

S = the acceptable service distance (surrogate for desired response time);

d_{ij} = the shortest distance from incident location i to police patrol location j;

x_j = 1 if a police patrol is located at potential site j, and 0 otherwise;

y_i = 1 if an aggregated crime location at i is covered by at least one l located police patrol area, and 0 otherwise;

$N_i = \{j \text{ in } J \mid d_{ij} \leq S\}$;

a_i = weight of crime incidents at incident location i;

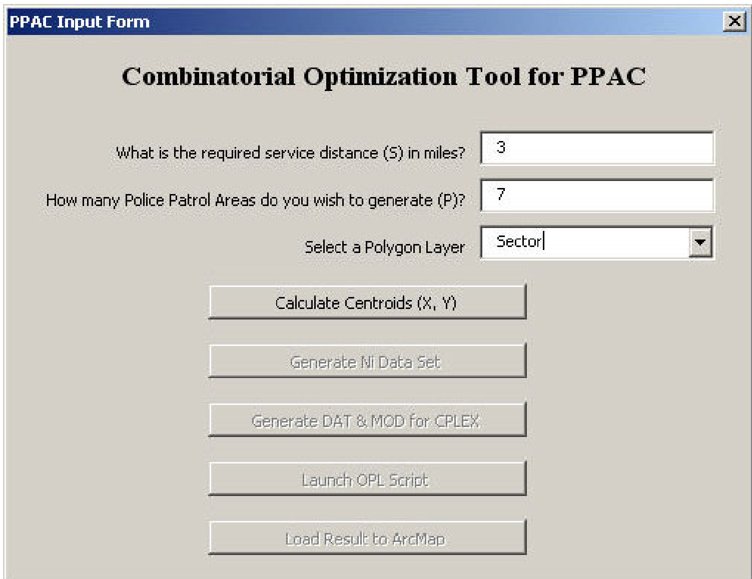
P = the number of police patrol areas to be located.

In this formulation N_i is the set of facility sites eligible to provide "cover" to incident location i. In the context of patrol area development, N_i is the set of crime incident locations that can be served within the acceptable response time, S. S can vary for different types of incidents, which may require faster response times. Keep in mind that although d_{ij} and S do not appear directly in the formulation, they are included in constraints (1) through the inclusion of the sets N_i. The objective is to maximize the number of incidents served or "covered" within the acceptable response time. Any subset of crime incidents may be used to populate the set I. As an example, if there are seasonal trends in crime incidents, it may be appropriate when defining patrol areas for a given week (or month) to consider just those incidents that occurred during the same week (or month) of the previous year. Constraints of type (1) allow y_i to equal 1 only when one or more patrol cars are established at sites in the set N_i. The number of patrol areas to designate (P) is limited to the number of available patrol cars by constraint (2). Constraints (3) and (4) require that only integer values are included in the solution. That is, police patrols cannot be split between patrol areas.

The PPAC model assumes a priori that an acceptable level of service (measured as a response *distance*) has been agreed upon as representing an acceptable level of citizen safety. This assumption is reasonable based on recent findings that police response time can be a significant determinant in the evaluation of police performance (Priest & Carter, 1999). While research has shown that response times have little bearing on the volume of crime in a jurisdiction (Sherman, Gottfredson, MacKenzie, Eck, Reuter & Bushway, 2004), we caution the reader to regard an important point. While crime reduction is clearly of paramount importance to policing, their efforts to reach this goal do not occur in a vacuum. As with all government agencies, police departments are subject to a number of resource constraints and political realities that make the efficient delivery of services important. Even with little or no reduction in crime rates, an increase in operational efficiency may lead, in turn, to improved effectiveness. Given the limit on police resources, the implementation of PPAC also requires that the number of police patrols is known in advance. This is, in fact, one of the models' strengths given that the amount of financial resources to be allocated to police protection may change quickly and often.

There are three primary objectives in police deployment: citizen safety, cost of operations, and workload balance (Taylor & Huxley, 1989). Consider that citizen safety is generally accepted to be greater when the number of police on duty is increased and therefore response times are decreased, and consider also that the cost of operations are directly related to the number of officers on duty. These objectives are clearly opposed to one another, and therefore solutions must be determined that reconcile these goals based on an acceptable level of service

Figure 1. GIS data collection tool



provision. The PPAC objective of maximal coverage is one attempt to resolve this internal competition.

Solving PPAC by Integrating GIS and Optimization

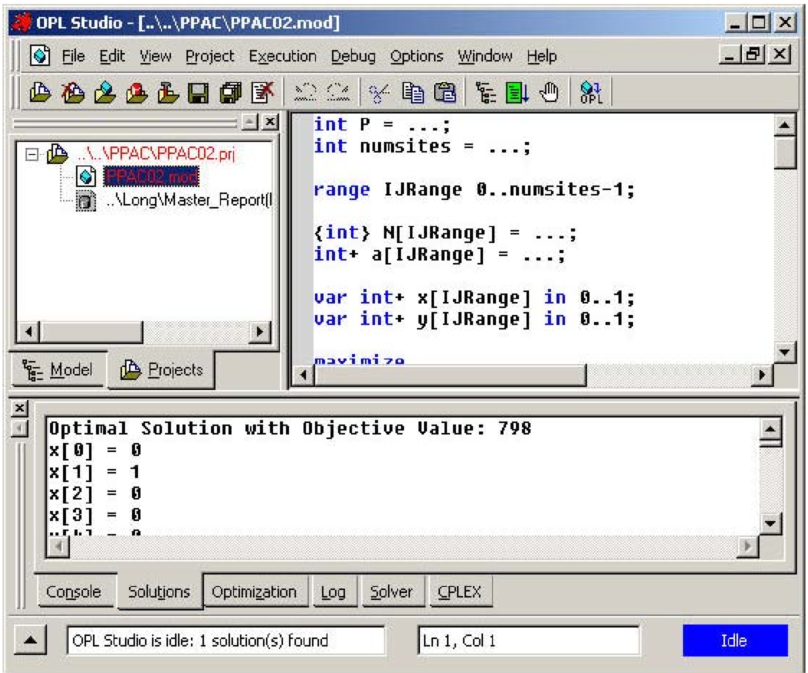
While GIS alone cannot provide optimal solutions to combinatorially complex location problems, it does provide the ideal platform for the assembly of geographic data layers, the collection of model instance parameters from users, and the output of cartographic representations of optimal solutions to decision makers. By integrating GIS with software designed to determine optimal solutions for complex systems, alternatives can be generated in order to inform police policy and practice.

Consider how such integration takes place in the context of solving the PPAC model. For this example and the pilot study in the following section, the ArcGIS Software (specifically, the ArcMap module) and the ILOG Optimization Programming Language (OPL) Studio and CPLEX optimization software for integer programming applications was employed. In order to begin the process of solving PPAC, several model parameters must be collected from the user. ArcMap provides a Visual Basic for Application (VBA) platform where interactive menus can be built to query the user for this data. In the case of PPAC the user must input values for the acceptable service distance (*S*), the # of patrol areas to generate (*P*), and the polygon layer from which the optimal spatial arrangement will be built (Figure 1). If the centroids of the polygons are to be used as potential patrol area centers, these can be generated within ArcMap as well. Once the service distance is provided, the sets N_i denoting the possible coverage relationships can be generated within the GIS. This involves the calculation of distances between potential patrol area locations and crime incidents and a pairwise comparison of locations to determine possible coverage. If the crime incident data is associated with the polygons, the a_i values can be computed with standard point-in-polygon GIS query tools. If addresses are available for the crime incidents they can be address-geocoded with the assistance of appropriate GIS tools. All of these functions are well within the capabilities of industry standard GIS software packages. At this point in the process, however, the GIS must surrender this information to the optimization software in order for optimal solutions to be determined. The VBA programming platform can be used to export the model parameter information to a data file that can be read by OPL Studio.

Optimal solutions can be obtained for many location problems by combining the use of a version of the simplex solution method (Dantzig, 1957) on linear programming relaxations of the problem, with a complementary branch and bound technique for dividing the original problem into more solvable subproblems (Hillier & Lieberman, 1995). The ILOG CPLEX software package employs these methods in combination with procedures for advanced preprocessing of the problems to be solved. Techniques are embedded within the software to reformulate problems to encourage integer solutions and reduce the amount of time and the level of computer resources that are needed to determine optimal solutions. The OPL Studio software allows the general version of the PPAC model to be formulated (Figure 2). It can be launched with commands given from within the GIS. The output parameter file for the particular problem instance can be read and the resulting problem instance submitted to CPLEX for solution with no further input from the user. When an optimal solution is found, the locational decisions (the values of the decision variables) can be exported back to the GIS for display.

Of utmost importance when implementing these models is that they must be accessible to police and associated staff as the primary users. In the past it has been suggested that the mathematical formulations common to operations research are beyond the understanding of all but a few experts in the field (Aly et al., 1982). In order to ensure that this is not the case, the models must be

Figure 2. Optimization software for the processing of GIS-generated model parameters



presented in a clear, and easy to understand format, with cartographic output that can be generated for those who must make the final decision on which solution to implement. In the case of the PPAC model the solution is a set of decision variable values that represent the locations of patrol area centers. The GIS provides the tools for selecting the areas that can be serviced by the chosen patrol area locations, and for dissolving the polygon layer to generate an intuitive cartographic representation of the alternative spatial arrangement. If the boundaries of the patrol areas must conform to other land features (such as census area polygons or traffic analysis zones) then GIS contains functions for the overlay of such layers and the selection of polygons for each police patrol area.

Pilot Study for the City of Dallas, Texas

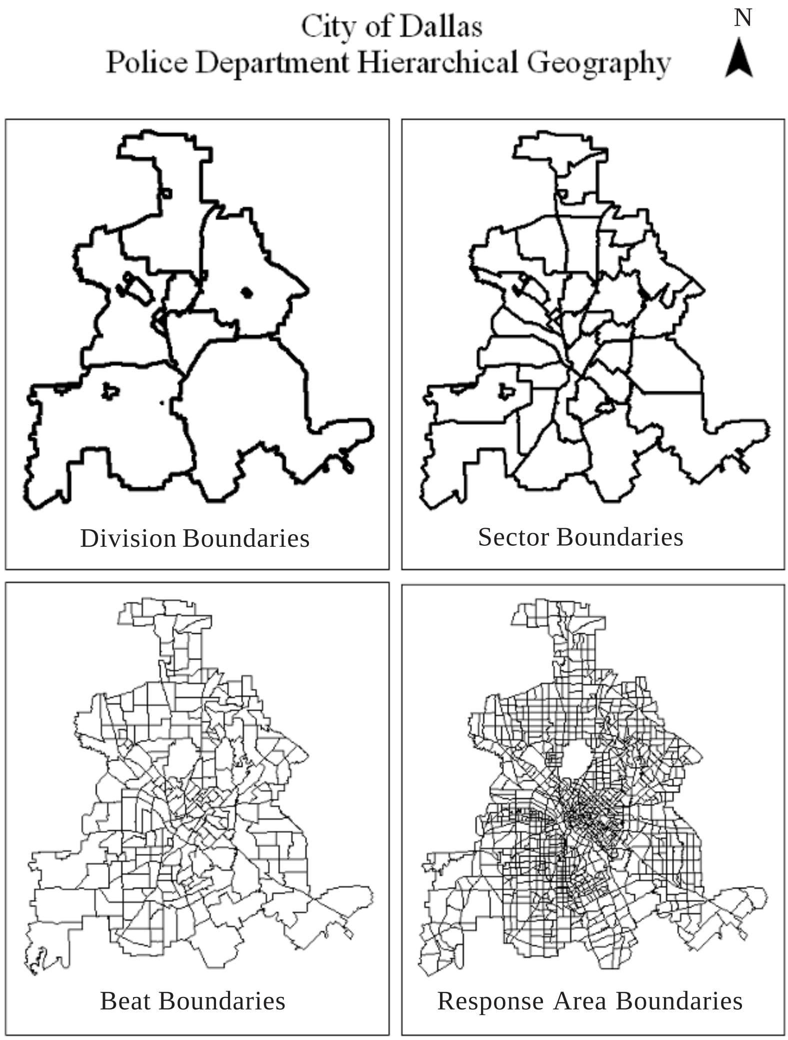
A pilot study of the applicability of the PPAC Model has been conducted as a proof-of-concept using the geographic boundaries employed by the city of

Dallas, Texas, Police Department, and a subset of the incidents to which that department has been required to respond. In the case of Dallas, Texas, there are seven divisions, 33 sectors, 233 patrol areas (beats) and 1,176 response areas (Dallas Police Department, 2002a) as shown in Figure 3. A sample of 798 calls for service from January 1, 2002, through January 7, 2002, where Dallas Police Department officers responded to 911 hang-up phone calls are used for this test of the PPAC model (Dallas Police Department, 2002b). For any level of the police geographic hierarchy, a spatial division can be created based on any of the lower levels in the hierarchy. That is, beats can be built from response areas, sectors can be built from either response areas or beats, and divisions can be built from any of the three other geographic layers. Each of these permutations was solved optimally in order to generate alternative spatial divisions for Dallas, but for purposes of graphical clarity, and given the space constraints in this format, we present only the instance where divisions are built from the sector level geography.

For the purpose of determining the optimal arrangement of divisions based on sectors, the incident locations were aggregated to the centroid of the sector within which they occurred. This approach exposes the results to aggregation errors, and it would be preferable to have address-geocoded locations of the incidents. Unfortunately such locational specificity was unavailable for this pilot study. The values of i and j can therefore vary from 1 to 33 (the number of sectors). In the case of i, the sector centroids represent the aggregated incident locations. In the case of j, they represent the potential locations for the police patrol areas. Since police patrols are presumed to be traversing their area of patrol responsibility until assigned to a call (Bodily, 1978) there is no way to know in advance the precise location of the police patrol, and therefore the sector centroids represent as reasonable an assumptive location as any others. The value of a_i becomes the sum of the 911 hang-up incidents that occurred in each sector. For this problem instance, one division (and its associated single sector) and another sector were removed from the dataset. These areas represent suburban lakes where no incidents occurred. Therefore, an optimal spatial arrangement of six divisions should be created from a set of 31 sectors used as building blocks. In the PPAC model P is equal to six.

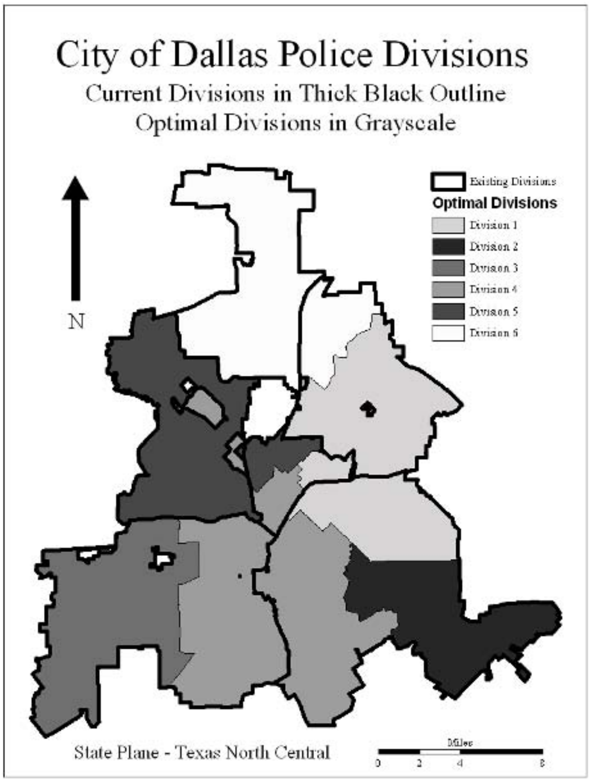
This problem was solved with increasing service distances (S) in increments of one-tenth of a mile until a spatial arrangement was found that covered all of the incidents. The smallest service distance that covers all incidents is 5.1 miles. This service distance was measured using the Euclidean distance between the sector centroids representing the patrol area locations and the aggregated incident locations. Clearly, network distance would be a preferable metric but, once again, address-geocoded locations were unavailable. Once the optimal solution to the PPAC model was found that could cover all incidents with the smallest possible response distance, the optimal spatial arrangement was compared to the

Figure 3. City of Dallas police geography



current one (Figure 4). In terms of the total distance traveled from police patrol location to incident locations, there was a 3.66% reduction using the PPAC optimal solution. In terms of average distance traveled, there was a 3.54% reduction with the PPAC solution. The worst case distance (meaning the longest single distance from a police patrol location to an incident location that it covers) increased by 0.59% with the PPAC solution. Perhaps the most important comparisons are made with the total and average weighted distances, where the

Figure 4. Current and optimal police divisions



weights are the number of incidents at each location. Both of these distance measures showed and improvement of 4.85% with the PPAC solution when compared to the current spatial arrangement.

Although it is satisfying that the PPAC model generated optimal arrangements that improved on the current spatial arrangement, for this particular problem instance (generating division boundaries from sectors) the results could well be interpreted as a verification of the appropriateness of the current arrangement. It would likely not be appropriate to redesign long-standing division boundaries – and incur the concomitant expense – for only a 4.85% reduction in weighted response distances. This highlights the importance of understanding that the optimal solutions to location problems represent alternatives to be presented to decision makers for evaluation, and perhaps for implementation.

Model Refinement

One of the most significant benefits of finding optimal solutions to this problem is the ability to evaluate the models in terms of the improvement that they provide when compared to the existing arrangement of police patrol areas, or when compared to alternative arrangements. However, the optimal solution is only a reflection of the ability of the model to replicate a real-world system. The process of testing the model involves refinements of the model to more accurately reflect the system that the model attempts to represent. The example we present above is designed simply to show the potential power of this approach. The true power lies is developing solutions steeped within the operational and administrative priorities of the agency.

There are several components of the model that can be immediately refined. First, the model formulated above considers each crime location to be of equal importance. Since the test case in Dallas considered only a single type of call (911 hang-up responses) this assumption is reasonable. However, the delineation of patrol areas must consider all of the types of crime incidents that occur. Therefore the measure of demand (a_i) should be a variable that represents the importance of a timely response (a covering) of that incident. Secondly, the model formulation above does not include any constraint on the number of calls that will be located within any particular police patrol area. In other words, there is no guarantee of equity among the police patrol areas. In order to address this deficiency the model can be refined with the addition of a set of constraints of the following type:

$$\sum_{i \in N_j} a_i x_j \leq M \quad \text{for all } j \in J$$

where:

$N_j = \{i \text{ in } I \mid d_{ij} \leq S\}$;
M = the maximum incident load that each patrol area can serve;

The set N_j is defined as all of the crime incident sites (i) that can be served from a potential patrol area centroid (j). There is one constraint for each potential patrol area centroid (j). If a patrol area is centered at j the value of x_j will be 1. If this is the case the constraints will require that the sum of the crime incident values (a_i) for all of the sites i that are covered by j must be less than or equal to the maximum crime incident load that can be handled by any single patrol area.

It is presumed that there are a large number of other potential constraints, and they will fall into several general categories – physical resource constraints, economic constraints, scheduling constraints and legal constraints. Physical resource constraints include the number of patrol cars available, the number of police officers available to operate those cars and the availability of support staff and safety equipment that the police officers need to complete their duties. Economic constraints are presumed to be a function of the budgets available for police officers and support staff salaries, for the purchase of equipment and for the ongoing operation of police activities. Scheduling constraints will likely be imposed by the contracts of the police officers and their union, and by the desire to equitably distribute the workload among the officers who are assigned to the designated police patrol areas. Finally, it is presumed that there are numerous legal constraints on the level of police service that must be provided by the police and that control the areas that they must patrol.

Conclusions, Discussion, and Future Research

Since GIS software is not designed to solve computationally complex problems such as the police patrol area problem, a GIS cannot be used in isolation to find optimal solutions. By integrating GIS with software that is capable of determining optimal solutions to such problems, a system can be built that can take advantage of the strengths of both to create understandable output for decision making once the optimal solution has been found. In this context GIS provides an interface for querying users regarding model instance parameters, and the functionality for both pre-processing the geographic data to be submitted to the solution software, and post-processing the results of the solution procedure.

Complementing the GIS is the process of developing a mathematical model to represent – as best as is possible – the system within which police patrol areas must be developed. This chapter suggests the PPAC model as an appropriate representation. PPAC seeks to maximize the coverage of crime incidents within an acceptable service distance as defined by the user, and will do so by creating a user-defined number of police patrol areas.

Such an integration of GIS and optimization software has been developed and presented in this chapter. In order to test the ability of covering models to solve the problem of determining optimal police patrol areas, a pilot study was conducted for the city of Dallas, Texas. One of the problem instances was presented here with the results showing an improvement in service distance over the current spatial arrangement.

However, there are many possible improvements to the model and its implementation. First, an ideal interface with GIS would consist of a custom application that could input the updated crime incident locations as they are made available by the police department, as well as variables such as the number of available police officers or special events requiring additional police support. The system could then recompute an optimal solution and transmit the results to patrol areas or to decision makers as image files or geographic data files. Industry standard GIS programs contain all the necessary functionality for the dissemination of the results in this way.

Secondly, although each of the problem instances using the Dallas dataset were able to generate optimal solutions (even with 1,167 response areas used to generate 233 beats) the model should be more rigorously tested with larger samples, particularly with larger sets of crime incidents. Moreover, if address-geocoded locations of instances become available they should be used, which would also allow the use of network distances.

The presentation of the PPAC model is not intended to suggest that beat-based patrolling is superior to any other approach for police organization. PPAC is not presented – and should not be interpreted – as an alternative to hot-spot analysis or any other means of determining the concentration of crime locations. Such techniques have proven their usefulness in suggesting intelligent distributions of tactical police resources, including education and crime prevention initiatives (Sherman, Gartin & Buerger, 1989). PPAC is simply a model that can be used to present alternative spatial partitionings of existing administrative or patrol boundaries to decision makers. PPAC generates optimal solutions to the problem of covering crime locations given an acceptable service distance specified by the user. Given that such boundaries exist in virtually all metropolitan police departments, alternatives allow for improved efficiencies in the deployment of limited police resources. Even recent comparative analyses of crime patterns suggest that current research could benefit from the development of alternative geographic frameworks for analysis (Craglia et al., 2000).

The solution of these models has several implications for policy and practice. By using an optimally determined set of patrol areas, police departments can better serve the population and use their resources more efficiently. If these patrol areas can be changed rapidly through repeated solution of the PPAC model under changing conditions, the police department can more readily respond to these changes. Moreover, given the many people and organizations who will need to approve the changes suggested by the optimal solutions that are found, the results must be presented as a series of alternatives from which a best arrangement can be selected that will satisfy the greatest number of people and that can be efficiently implemented.

References

AdensoDiaz, B., & Rodriguez, F. (1997). A simple search heuristic for the MCLP: Application to the location of ambulance bases in a rural region. *Omega-International Journal of Management Science, 25*(2), 181-187.

Aly, A., Litwhiler, D., & Heggy, T. (1982). Operations research: A useful police management tool. *Journal of Police Science and Administration, 3*(3), 279-283.

Berman, O., & Krass, D. (2002). The generalized maximal covering location problem. *Computers & Operations Research, 29*(6), 563-581.

Bodily, S. E. (1978). Police sector design incorporating preferences of interest groups for equality and efficiency. *Management Science*, *24*(12), 1301-1313.

Chaiken, J. M., & Dormont, P. (1978a). Patrol car allocation model: Background. *Management Science, 24*(12), 1280-1290.

Chaiken, J. M., & Dormont, P. (1978b). Patrol car allocation model: Capabilities and algorithms. *Management Science, 24*(12), 1291-1300.

Chung, C. H. (1986). Recent applications of the maximal covering location planning (Mclp) model. *Journal of the Operational Research Society, 37*(8), 735-746.

Church, R. L. (2002). Geographical information systems and location science. *Computers & Operations Research, 29*(6), 541-562.

Church, R. L., & ReVelle, C. S. (1974). The maximal covering location problem. *Papers of the Regional Science Association, 32*, 101-118.

Church, R. L., & ReVelle, C. S. (1976). Theoretical and computational links between p-median, location set-covering, and maximal covering location problem. *Geographical Analysis, 8*(4), 406-415.

Church, R. L., & Sorenson, P. (1994). *Integrating normative location models into GIS: Problems and prospects with p-median model* (Report No. 94-5). Santa Barbara, CA: National Center for Geographic Information and Analysis.

Church, R. L., Current, J., & Storbeck, J. (1991). A bicriterion maximal covering location formulation which considers the satisfaction of uncovered demand. *Decision Sciences, 22*(1), 38-52.

Church, R. L., Stoms, D. M., & Davis, F. W. (1996). Reserve selection as a maximal covering location problem. *Biological Conservation, 76*(2), 105-112.

Craglia, M., Haining, R., & Wiles, P. (2000). A comparative evaluation of approaches to urban crime pattern analysis. *Urban Studies, 37*(4), 711-729.

Current, J. R., & Okelly, M. (1992). Locating emergency warning sirens. *Decision Sciences, 23*(1), 221-234.

Current, J. R., & Schilling, D. A. (1990). Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis, 22*(2), 116-126.

Curtin, K. M. (2004). Social measurement in operations research (In press). In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement*. San Diego: Elsevier Academic Press.

D'Amico, S. J., Wang, S., Batta, R., & Rump, C. M. (2002). A simulated annealing approach to police district design. *Computers & Operations Research, 29*, 667-684.

Dallas Police Department (Cartographer). (2002a). *Cartographic Boundary Files*, Dallas, TX.

Dallas Police Department. (2002b). *Crime Incident Database*, Dallas, TX.

Dantzig, G. (1957). Discrete-variable extremum problems. *Operations Research, 5*, 266-277.

ESRI. (2004). ArcGIS help: LOCATEALLOCATE (Version ARC/INFO help). Redlands, CA: Environmental Systems Research Institute.

Galvao, R. D., & ReVelle, C. (1996). A Lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research, 88*(1), 114-123.

Galvao, R. D., Espejo, L. G. A., & Boffey, B. (2000). A comparison of Lagrangean and surrogate relaxations for the maximal covering location problem. *European Journal of Operational Research, 124*, 377-389.

Hale, T. (2004). *Location Science References*. Retrieved November 4, 2004, from *http://www.ent.ohiou.edu/~thale/thlocation.html*

Harries, K. (1999). *Mapping crime: Principle and practice* (Report No. NCJ178919). Washington, DC: National Institute of Justice.

Hillier, F. S., & Lieberman, G. J. (1995). *Introduction to operations research*. New York: McGraw-Hill.

INFORMS. (1999). *AT&T honored in operations research award*. Cincinnati, OH: The Institute for Operations Research and Management Sciences.

Kwak, N. K., & Leavitt, M. B. (1984). Police patrol beat design: Allocation of effort and evaluation of expected performance. *Decision Sciences, 15*(3), 421-433.

Larson, R. C. (1975). Approximating performance of urban emergency service systems. *Operations Research, 23*(5), 845-868.

Larson, R. C. (1978). Police deployment - introduction. *Management Science, 24*(12), 1278-1279.

Maranzana, F. (1964). On location of supply points to minimize transport costs. *Operational Research Quarterly, 15*, 261-270.

Mitchell, P. (1972). Optimal selection of police patrol beats. *The Journal of Criminal Law, Criminology, and Police Science, 63*(4), 577-584.

NIST. (2001). *Dictionary of algorithms, data structures, and problems*, 2001, from *http://www.nist.gov/dads/*

Pirkul, H., & Schilling, D. A. (1991). The maximal covering location problem with capacities on total workload. *Management Science, 37*(2), 233-248.

Priest, T. B., & Carter, D. B. (1999). Evaluations of police performance in an African American sample. *Journal of Criminal Justice, 27*(5), 457-465.

ReVelle, C., & Snyder, S. (1995). Integrated fire and ambulance siting: A deterministic model. *Socio-Economic Planning Sciences, 29*(4), 261-271.

ReVelle, C., Schweitzer, J., & Snyder, S. (1996). The maximal conditional covering problem. *Infor, 34*(2), 77-91.

Sacks, S. R. (2000). Optimal spatial deployment of police patrol cars. *Social Science Computer Review, 18*(1), 40-55.

Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of crime: Routine activities and the criminology of place. *Criminology*, *17*, 69-100.

Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (2004). *Preventing crime: What works, what doesn't, what's promising*. Department of Criminology and Criminal Justice: University of Maryland. Retrieved November 4, 2004, from *http://www.ncjrs.org/works*

Taylor, P. E., & Huxley, S.J. (1989). A break from tradition for the San Francisco Police: Patrol officer scheduling using an optimization-based decision support system. *Interfaces, 19*(1), 4-24.

Teitz, M. B., & Bart, P. (1968). Heuristic methods for estimating the generalized vertex median of a weighted graph. *Journal of the Operational Research Society of America, 16*(5), 955-961.

Zanakis, S. H., Evans, J. R., & Vazacopoulos, A. A. (1989). Heuristic methods and applications: A categorized survey. *European Journal of Operational Research, 43*(1), 88-110.

Chapter XIV

Web GIS for Mapping Community Crime Rates: Approaches and Challenges

Tung-Kai Shyy, University of Queensland, Australia

Robert J. Stimson, University of Queensland, Australia

John Western, University of Queensland, Australia

Alan T. Murray, Ohio State University, USA

Lorraine Mazerolle, Griffith University, Australia

Abstract

This chapter describes a prototype Web geographic information system (GIS) and spatial model application for mapping person crime rates in Brisbane, Australia. Our application, which integrates GIS functionality, a clustering model, client/server technology and the Internet, can generate useful documents such as maps and tables to examine and present crime patterns in space and time. Our chapter also demonstrates the usefulness and appeal of the Web GIS application as an information dissemination and spatial data analysis tool for promoting public awareness of social

conditions. This chapter argues that Web-based data access is a better approach to delivering large volumes of crime data and geographical information to the public. We expect that police, community workers and citizens could utilize the application and associated maps to facilitate and enhance crime prevention activities. We note, however, that further development of Web-based GIS applications need to answer a number of pertinent questions regarding system maintenance, data integrity and neighborhood crime prevention.

Introduction

Over the last decade, Australian crime rates have risen noticeably. With the exception of homicide, Australia's crime rates are among the highest in the industrialized world (Graycar & Grabosky, 2002; Morgan, 2003). Among the more serious property offences, the crime rate for unlawful entry with intent increased 13% between 1995 and 2000 (Australian Bureau of Statistics, 2002). During the same period, motor vehicle theft increased by 9% and the rate of "other theft" increased by 38%. In the year 2000, one in 28 persons would have been a victim of "other" theft (Australian Bureau of Statistics, 2002)[1]. Numerically, crimes against property outweigh crimes against the person by about 10 to one and while homicide rates have not changed markedly over the last seven years, non-aggravated assaults and sexual assaults have shown some increase (Morgan, 2003).

Commensurate with the rise in crime rates in Australia, crime mapping and spatial analysis, which is the process of turning raw data into useful information (Longley, Goodchild, Maguire & Rhind, 2001), for examining urban crime have grown in importance across Australia in recent years (Morgan & Fernandez, 2000; Murray, McGuffog, Western & Mullins, 2001). This trend mirrors the growing interest British and American policing has shown in crime mapping over the last decade (Rich, 1995; McEwen & Taxman, 1995; Anselin, Cohen, Cook, Gorr & Tita, 2000). Indeed, crime mapping has become an integral tool in the development and advancement of problem oriented policing in particular and crime prevention in general (Maltz, Gordon & Friedman, 1991; Weisburd & Green, 1995; McEwen & Taxman, 1995). With ready access to inexpensive and easy to use spatial analysis and crime mapping programs, police agencies throughout the world can easily produce computer generated crime maps that can help the police in partnership with the public to identify and respond to ongoing crime problems (Green, 1994; McEwen & Taxman, 1995; Maltz et al., 1991; Rich, 1995).

The development and increased use of GIS in policing has intensified the demand for public access to digital spatial information. Currently there are a variety of World Wide Web GIS applications ranging from city guides, digital libraries, economic development, ecotourism, location services, traffic information, white pages directory, geographical analysis machine, local planning and safe city[2] (Doyle, Dodge & Smith, 1998; Kirkby & Pollitt, 1998; Openshaw, Turton, Macgill & Davy, 1999; Peng, 1999; Shyy, Stimson, Davis, Murray, Baum & Barker, 2001; Peng & Tsou, 2003; Shyy, Stimson & Murray, 2003). Peng and Tsou (2003) uses the term Web-based GIS to refer to the use of the Web as a primary means to exchange data, perform GIS analysis, and present results. These applications provide interesting possibilities for the use of both the Web and GIS in terms of accessing different kinds of geographical information for a wider audience regardless of their physical locations. The general public may complain about the lack of crime information, including the dearth of crime maps, which is specifically related to their local area or region. Often the information they seek exists, such as the *Atlas of Crime in Australia 2000* (Australian Institute of Criminology, 2000) and annual police reports, but the information is often not accessible in a form that potential users can easily assess or comprehend. We suggest, therefore, that there are better ways (for example the Internet and the Web) of delivering large volumes of crime data and geographical information for public consumption.

Our chapter presents a prototype Web GIS application that allows citizens to access spatial crime data and use GIS technology without the demand on users to learn a commercial GIS package. We begin our chapter with a discussion of the development of a prototype Web GIS application for depicting or mapping person crime rates interactively in Brisbane[3], Australia. We describe how measures of attribute similarity and spatial proximity are combined using a clustering model to support the identification of spatial crime patterns. Application results highlight the flexibility and potential usefulness of the developed prototype Web GIS as an online spatial data analysis tool for mapping urban crime rates. We conclude with a discussion of how this application might be best utilized by the general public and describe some of the practical implications of providing this type of Web GIS application to the public.

Web GIS Development

Our Web GIS application uses an ACER Altos Server[4] running ESRI's MapObjects Internet Map Server (IMS) and the Microsoft Internet Information Server (IIS). MapObjects, which contains an ActiveX map control and a set of ActiveX objects such as data access and map display, are used in a Visual Basic programming environment running on a Microsoft Windows 2000 platform. The mapping functionalities developed within the application include zoom, pan, classify, select and identify polygon, and label features. The maps and analyses provided by the application highlight magnitude and patterns of person crime rates in Brisbane between 1996 and 1997. MapObjects IMS, which provides software components for linking MapObjects to the Web server (ESRI 1998), is used to deploy applications on the Internet. ESRI's Shapefile format, defining the geometry and attributes of geographically referenced features, is used to store spatial information.

A client interface using a Java applet was implemented for our Web GIS application because Java applet has the advantages of being versatile, platform neutral and secure (Peng, 1999; Peng and Tsou, 2003). This interface facilitates human-computer interaction in the analysis of crime rates in Brisbane. Java, an object-oriented, Internet programming language, is used for creating the Java applet. The Java applet is embedded in a Web page and downloaded over the Internet to the client's computer. Source code is first compiled into Java bytecode (.class extension). With this the Web browser is then able to execute the bytecode using a Java interpreter (Java Virtual Machine or JVM) (Decker & Hirshfield, 1998; McFarlane, Chiarelli, De Carli, Li, Wilcox, Wilton, Wootton & Updegrave, 1999).

One of the major features of the developed Web GIS application is the ability to generate simple and complex choropleth map displays of crime rates. Thematic classification using equal interval, quantile and bicriterion median clustering problem (BMCP) approaches for choropleth display is provided. The first two are standard GIS display options (Robinson, Morrison, Muehrcke, Kimerling & Guptill, 1995), whereas the third is based on spatial modeling (Murray & Shyy, 2000). Equal interval classifies attributes into equally divided ranges. Quantile classifies approximately the same number of features in each identified class. The BMCP is a spatial optimization approach that uses attribute similarity and spatial proximity for class grouping. Each display option facilitates pattern identification. However, the BMCP may be considered a more spatially based approach. As a result, identified patterns are likely to have greater meaning in a spatial context. The objective of BMCP is to minimize total within group difference of attribute similarity and spatial proximity. This is similar to minimizing within group variance. Thus, both measures of relative performance for the classification approaches are reported in the developed Web GIS application.

Supporting pattern display functionality is the ability to assess the degree of spatial autocorrelation, a prerequisite for spatial patterning/grouping. A global measure of spatial autocorrelation is the Moran coefficient (MC), which indicates the degree of grouping of spatial units with like attribute values

Figure 1. Map pattern examples for the Moran Coefficient index (modified from Griffith, 1987)

2	3	4
3	4	5
4	5	6

(a) Similar Values Clustering

5	2	5
3	6	3
4	4	4

(b) Dissimilar Values Clustering

4	5	4
5	4	3
2	3	6

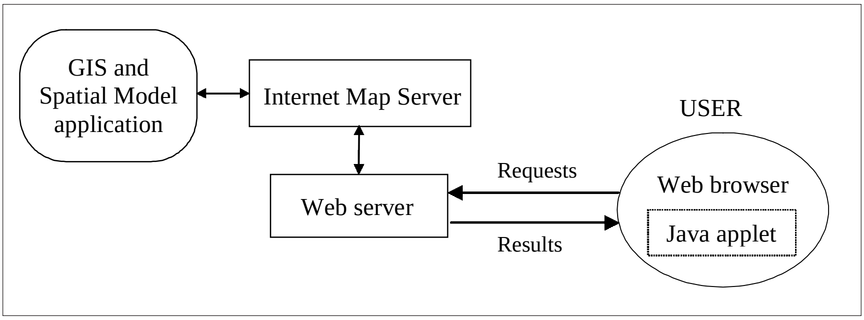
(c) No Pattern

(Griffith, 1987). Similar values tend to cluster together on a map when MC value closes to 1 (see Figure 1a). Dissimilar values tend to cluster together when MC value closes to -1 (see Figure 1b). And, there is no pattern on a map when MC value closes to 0 (see Figure 1c). The developed system provides capabilities for computing the MC. For 1996 person crime rates in Brisbane, a moderate degree of spatial autocorrelation was found (spatial neighbors are suburbs sharing at least one common boundary), thus warranting further analysis of spatial grouping using choropleth approaches.

The Web GIS Applications

The Web GIS application was developed for informing the general public of observed crime rates in Brisbane. Technical details of the crime data and detailed spatial analysis of crime in this region may be found in Murray et al. (2001). The relationship between the Internet map server and Web browsers is displayed in Figure 2. The general public may access the developed Web GIS application at *http://www.uq.edu.au/cr-surf/brisbane_crime.htm* using a Web browser (such as Internet Explorer or Netscape Navigator). A Java applet serving as the client interface is loaded by the browser from the server, and subsequently executed on the client's computer (see Figure 2). Users may then select desired functionality from the interface and send their requests for information through the Internet to the Web server. The server interprets and distributes requests to the specified GIS application. The application performs its tasks and sends results back to the Web server. The Web server then returns results such as maps and tables to the client's browser. This is an example of a server-side application, which depends on a powerful GIS Server to perform GIS data processing and analysis. Anyone who knows how to use a Web browser will

Figure 2. Server-side GIS and spatial model application and client-side Java applet interface



be able to access our Web GIS application interactively through user-friendly interface from the Internet with minimal learning overhead. The client-side applications conduct GIS analysis on the client's computer. Users will need to install plug-ins or help programs and download spatial data from the server in order to interact with the spatial features on the map. This may not be a user-friendly approach to the public. A detailed review of the server-side and client-side approaches to Web GIS development can be found in Peng and Tsou (2003).

Once a user makes a uniform resource locator (URL) request to the Web GIS server, a map of Brisbane with suburb boundaries is displayed. Eight options of functionality remain constant throughout the user's visit to the server. *Zoom in* and z*oom out* center the map at the point where a user clicks with the mouse. *Pan* allows the user to move around the map by dragging the display in any direction using the mouse. *Label on* displays the name of each of the subareas (suburbs) in the mapped region. *Full extent* displays the entire region being analyzed. Finally, *restart* takes the user back to the opening page.

The remaining options for this application are *identify* and *classification*. When the user selects the *identify* option, they are asked to select a suburb of interest on the map. Once the user selects a suburb, crime rate information is displayed for the selected suburb. *Classification* is the major developed modeling functionality, enabling choropleth display modification using the equal interval, quantile or BMCP methods. When a user selects the classification option, they are then presented with three drop down menus from which they select a desired classification method. The first enables the equal interval or quantile approaches, followed by selection of the desired number of classes from the second drop down menu. The third drop down menu queries for the attribute to display. With this information the system then presents a classification map of Brisbane crime rates. If the BMCP option is chosen from the first drop down menu, a fourth and fifth drop down menu will be presented. These menus enable the user to choose/

vary relative weights (attribute similarity and distance) for determining grouping similarity. The user is then prompted by a “go” button before model optimization is carried out.

The BMCP classifies features by simultaneously maximizing attribute homogeneity and spatial proximity. For classification using BMCP, we focus on a single attribute and one spatial scale for illustrative purposes. Specifically, person crime per 1000 residents is demonstrated for 1996 in Brisbane, Australia, at the suburb level of analysis. For spatial proximity, the Euclidean distance metric in kilometers was utilized. There are 178 suburbs in Brisbane. The analysis was carried out on an ACER Altos Server. The time required to solve the BMCP and to display classification result was about 15 seconds for the application reported in this chapter. The Moran coefficient for person crime was 0.132. This is indicative of positive spatial autocorrelation and we would expect to see a certain degree of grouping of suburbs with like person crime rates. This suggests that the use of choropleth display and spatial clustering would be revealing.

Other currently available online options for choropleth display are equal interval and quantile. However, these approaches do not account for spatial proximity. For this comparison, weights of attribute (Wa = 1) and distance (Wd = 0) for the BMCP have been utilized. Thus, initially we are only concerned with patterns typically produced by traditional choropleth display. As an example, for seven classes the BMCP total within group difference is 332 (Figure 3), which is less than the differences found by the equal interval and quantile approaches of 824 and 1210 (Figures 4 and 5). Each class shown in Figure 3 is represented as a unique color and/or pattern. The more traditional way of creating choropleth display classes is to minimize total within group variance, rather than difference. As an example, for seven classes the BMCP total within group variance is 4169, which is less than the variances found by the equal interval and quantile approaches of 9412 and 210558.

The visual implications of the classification produced using the BMCP is also of great interest. Figure 6 illustrates the seven classes using weights of Wa = 1 and Wd = 0.4. The BMCP total within group difference is 390, which is less than the differences found by the equal interval and quantile approaches. The BMCP total within group variance is 4444, which is also less than the variances found by the equal interval and quantile approaches. Interactive exploration of the classes shown in Figure 6 reveals the attribute similarity between some suburbs is significantly influenced by spatial proximity considerations. Another interactive examination is presented in Figure 7 using weights of Wa = 1 and Wd = 1. The BMCP total within group difference is 627, which is again less than the differences found by the equal interval and quantile approaches. The increased importance of space in the classes is very apparent in Figure 7 (for example, grouping of north, east, west and northwest suburbs) as compared to Figure 6

Figure 3. Seven group classification of person crime using BMCP

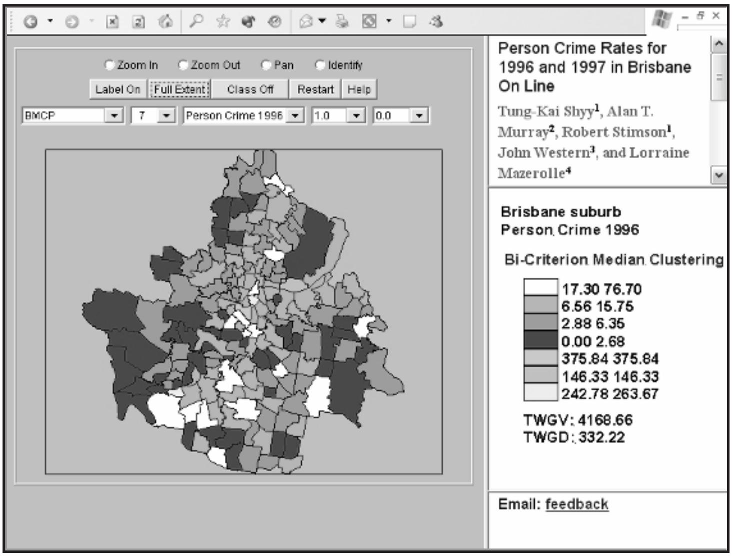


Figure 4. Seven group classification of person crime using Equal Interval

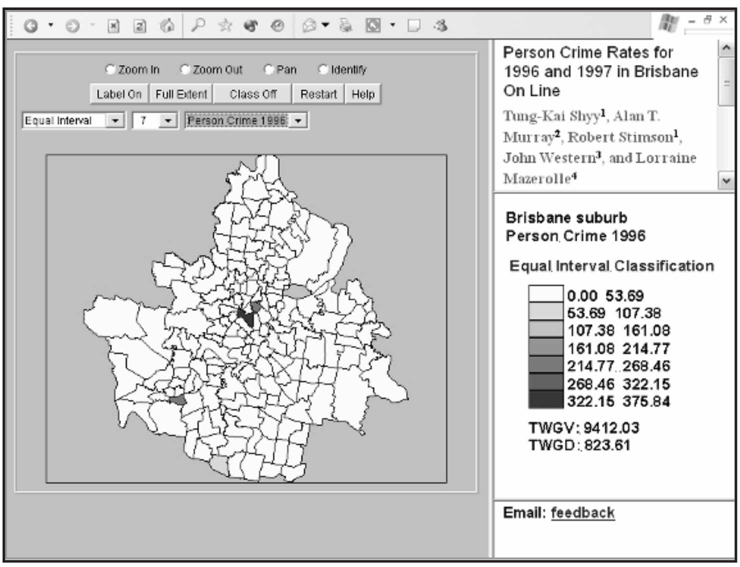
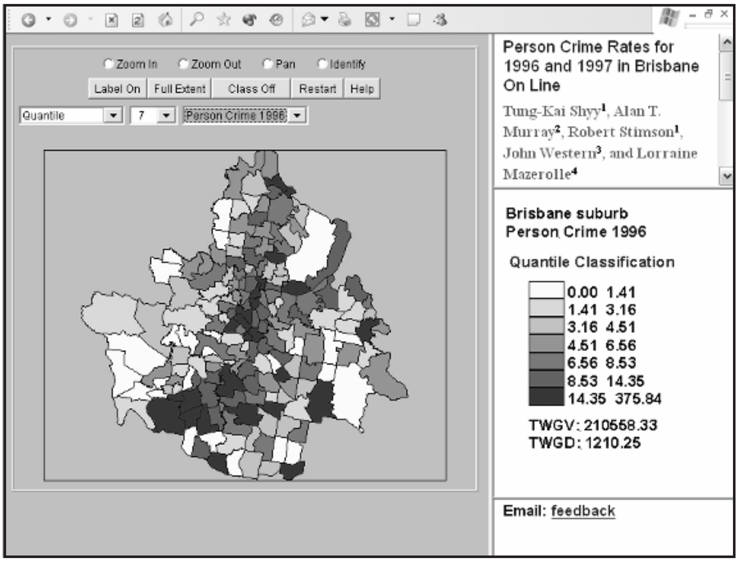


Figure 5. Seven group classification of person crime using Quantile

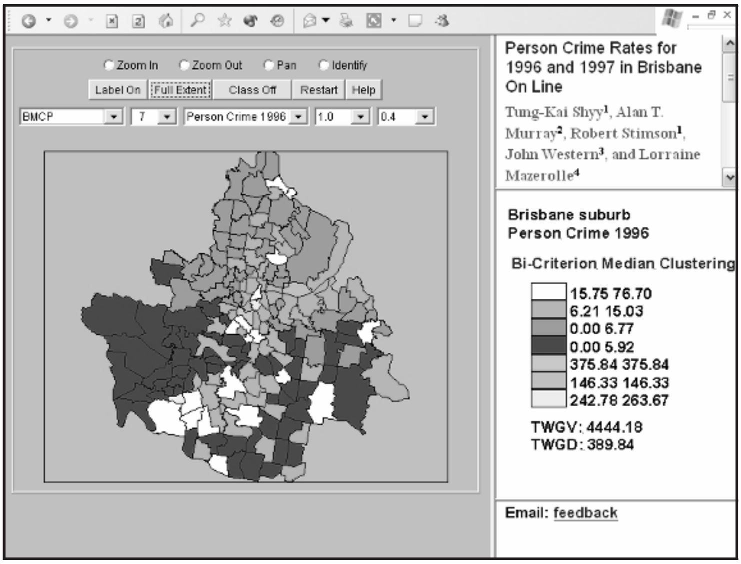


(for example, grouping of north and west suburbs). This may be a more revealing classification for identifying spatial grouping of suburbs with similar person crime rates. There are some overlaps in attribute values between classes in Figures 6 and 7 because spatial relationships are included in these classifications. Equal interval and quantile approaches, which are standard GIS options, are unable to display spatial grouping of suburbs well because they do not account for spatial proximity. A detailed discussion of the impacts of different Wa and Wd in BMCP approach can be found in Murray and Shyy (2000).

Challanges of Web GIS Applications

One of our prime motivations for developing the Web GIS prototype is to demonstrate how GIS technology can be used to provide crime data and spatial analytic tools to the general public. Our basic argument is that crime data (principally calls for police service, offense and arrest data) should be provided to the public in a way that is both easy to access and easy to interpret. We suggest that GIS technology, available via the Web, can fill this void and enable the public to access police and other government data and utilize GIS technology to analyze the spatial distribution of crime. In this section, we identify several reasons why

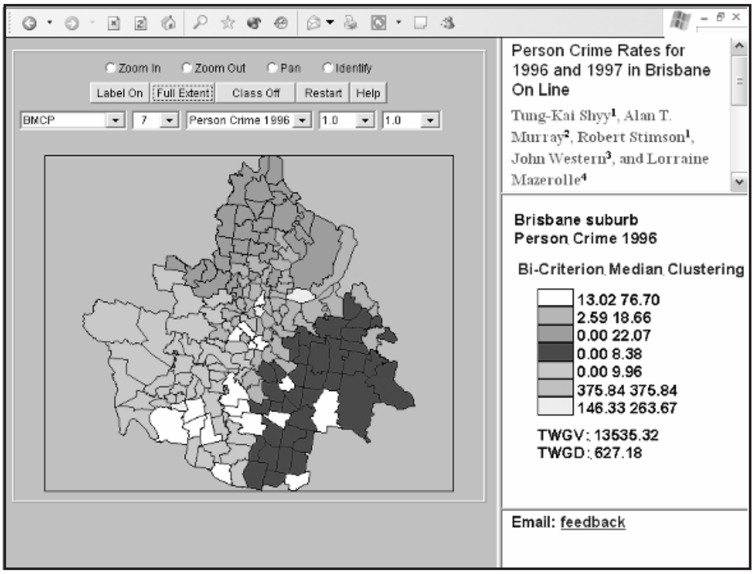
Figure 6. Alternative seven group classification of person crime using BMCP



the police and others (for example, census bureaus and criminal justice agencies) have been slow to provide crime data to the public in a GIS environment and why, conversely, the public might be slow to start using the type of Web GIS prototype we describe in our chapter. We suggest ways that the technology might best be implemented to increase the utility of Web GIS applications.

There are many reasons why the police may be reticent in cooperating with efforts to build a Web GIS application and provide spatial crime data available in the public domain. First, the lack of accuracy and difficulties in interpretation of police data gives the police reason to pause before offering up their data for public consumption. Inaccuracies in the precise location of where an offense occurred do not create problems when data are reported at the regional or state level. But as the use of the data moves down to what geographers refer to as a “cone of resolution” (Brantingham & Brantingham, 1975; Harries, 1980) from national, state and regional levels to suburb, neighborhood and place levels (Eck & Weisburd, 1995), the police become more concerned about inaccuracies in their data. If large numbers of cases cannot be geocoded, then biases arise in the data. When the police attending the scene of a crime fail to enter the correct location, this too can lead to inaccuracies in the spatial distribution of crime. These data problems are highlighted in the use of GIS in mapping crime and create challenges for those wanting to deliver Web GIS applications available in the public domain. Obviously, any delivery of Web GIS would need to acknowl-

Figure 7. Another seven-group classification of person crime using BMCP



edge the potential limitations of the data such as over/under reporting, location accuracy and so forth. Perhaps a warning message that prompts acceptance of these limitations would at least highlight the point.

Second, police and other government agency crime data have been typically unavailable for public consumption in a form that might identify individuals or disadvantage groups of people. The police have, for years, published annual crime statistics for the state, region or even district. Since GIS requires location information for geocoding purposes (for example, address of where a crime occurred), police agencies have been understandably unwilling to provide police data in the public domain at the address level unit of analysis. In our prototype Web GIS application, we display the data at a level of aggregation (that is, the neighborhood) such that individuals or individual households cannot be identified through the maps. Indeed, the zoom feature of the application only allows end-users to examine crime rates at the suburb unit of analysis. We point out, however, that regardless of the zoom feature on a Web GIS application, the police will need to be convinced that individuals or households cannot be identified through the maps.

Third, police ownership of crime data raises several complex questions: If a Web-based crime mapping facility is developed for public consumption, who will maintain a Web GIS application? If it is not the police, then who? University-based researchers? GIS companies? How might this compromise the integrity of the data? How do the police ensure that they do not lose ultimate control over the data? And what about confidentiality issues when address-level data are provided as part of the Web GIS contract? What type of maintenance issues should be considered? Cost? Size of the data archive? Updating the system with new data? Purging "old" data from the system to avoid data overload and reductions in the Web GIS performance? How often should data be transferred and subsequently purged?

Finally, one of the reasons why the police have been reticent in the past at supplying detailed spatial information about crime distributions is that they worry about the possibility that spatial crime maps could potentially aggravate, rather than reduce, crime problems in some communities. Insurance companies could increase their premiums in high risk communities and real estate companies could use the data to artificially inflate prices in some neighborhoods and deflate prices in other neighborhoods. Perhaps the best example of this type of abuse is what Skogan (1986) refers to as "demagoguery." Skogan (1986) warns that, "…cagey real estate agents can reap enormous profits trading on fear. Stirring concern about crime and racial change, they frighten white residents into selling their homes at reduced prices; then the homes are re-sold at inflated prices to Blacks and Hispanics desperate for better and safer housing, a practice often known as 'blockbusting'" (p. 207). Skogan (1986) goes on to describe how the practice of demagoguery can be a triggering event that leads to neighborhood decline.

There are also several reasons why the public might be slow to utilize a Web GIS application for crime mapping purposes. First, research suggests a certain level of apathy in citizens mobilizing to reduce and control crime problems (Davis & Lurigio, 1996). Indeed, recent research suggests that when a neighborhood has a high level of what Sampson, Raudenbush and Earls (1997) refer to as "collective efficacy," then there are generally lower than expected levels of crime, regardless of the social capital and social structural conditions of the community (p. 918) (see also Sampson, Morenoff & Earls, 1999). Collective efficacy is described as a *process* for mobilizing social capital to tackle specific neighborhood problems. It is a mechanism that facilitates social control without requiring strong ties or associations (Sampson et al., 1997, 1999). As distinct from other ecological constructs such as informal social control, community capacity and social capital, collective efficacy is a *task-specific* construct that exists relative to particular, perhaps episodic, neighborhood problems. It highlights shared expectations and mutual engagement by residents in their efforts to impose local social control for specific crime or social problems (Morenoff, Sampson & Raudenbush, 2001; Sampson et al., 1999). Research exploring the spatial distribution of collective efficacy using data from Chicago has found that collective efficacy is the most "…proximate social mechanism for understanding between-neighborhood variation in crime rates" (Morenoff et al., 2001, p. 521).

We suggest that those communities that score high on the scale of collective efficacy are likely to be the types of communities that are more interested in accessing Web GIS applications that allow for a spatial analysis of crime problems. However, we suggest that arguably the communities that would most benefit from an analysis of spatial crime patterns might be those communities that are least likely to access the information. That is, we know that economic disparity is reflected in overall Web use and IT literacy (National Telecommunications and Information Administration, 2001). We would expect this "digital divide" (Servon, 2002, p. 1, 24) to limit the use of a Web GIS system in some communities. As such, we suggest that any implementation of a Web GIS application must be accompanied by training of the police particularly in these lower socioeconomic areas, to encourage them to use the facility for problem-oriented policing activities and stimulate collective involvement in crime prevention activities.

Second, we suggest that as people become aware of the Web GIS facility[5] and as the application integrates more spatially relevant data (for example, crime distributions, proximity to mass transit, proximity to shopping centers), then we would hope the community members and citizens would increase their interest in GIS. That is, Web GIS applications will be of more interest to the public as the systems include more spatially relevant data, more sophisticated (yet easy to use) analysis tools and with better statistical and visual qualities than traditional approaches. We suggest, however, that regular input into a Web GIS system by crime analyst (via meetings and Web announcements) will be crucial in bridging the gap between crime mapping use and interpretation. Evolution of these technological and communication systems will help laypersons to understand gradually both the spatial and temporal patterns of crime occurrence.

Conclusion

In our chapter, we have described a prototype Web GIS application that we believe is useful for promoting a greater public awareness of social conditions. Equal interval, quantile and BMCP approaches were employed to support the identification of spatial crime patterns in Brisbane. Incorporating both spatial proximity and attribute similarity is significant for evaluating relationships in spatial information. The Moran coefficient, which indicates the degree of grouping of suburbs with like attribute values, is an important reference measure for BMCP approach. By placing GIS application online, from the operator's computer one can gain an idea of which suburbs or urban regions have been changing over time.

The power of our prototype is in its simplicity to use: Anyone who knows how to use a Web browser will be able to access the Web GIS application interactively from the Internet with minimal time spent learning how to use the application. Moreover, we point out that the GIS database can be updated as new crime and social data become available. Adding more data sets, such as access to shopping centers, access to bus stops and other mass transit centers, will also facilitate better spatial decision making. Notwithstanding the enhancements that are possible to our prototype, we believe that our chapter demonstrates the usefulness and appeal of the developed prototype Web GIS and spatial model application as an information dissemination and spatial data analysis tool for crime mapping.

Nonetheless, the further development of Web-based GIS systems needs to explore and answer a number of pertinent questions: who will maintain such a system? Who will have responsibility for the integrity of the data? How do you protect individuals? How do you protect neighborhoods from further decline? How do you further integrate spatially relevant data into the system? In addition, how do you enable and motivate citizens, particularly those in low-efficacious neighborhoods, to access the information and ultimately use the information to solve and prevent neighborhood crime problems?

Acknowledgment

The authors thank Queensland Police Service for helpful support and provision of unofficial Brisbane crime data.

References

Anselin, L., Cohen, J., Cook, D., Gorr, W. & Tita, G. (2000). Spatial analyses of crime. In Crime and justice 2000, vol. 4, *Measurement and analysis of crime and justice* (pp. 1-24). Washington, DC: U.S. Department of Justice.

Australian Bureau of Statistics (ABS). (2002). ABS crime and justice data 2002. In A. Graycar & P. Grabosky (Eds.), *The Cambridge Handbook of Australian Criminology*. Cambridge: Cambridge University Press.

Australian Bureau of Statistics (ABS). (2003). *2001 census of population and housing*. Canberra: ACT.

Australian Institute of Criminology (AIC). (2000). *Atlas of Crime in Australia 2000.* Canberra: ACT.

Brantingham, P. J., & Brantingham, P. L. (1975). The spatial patterning of burglary. *Howard Journal of Criminal Justice, 14*, 11–23.

Davis, R. C., & Lurigio, A. J. (1996). *Fighting back: Neighborhood antidrug strategies.* Thousand Oaks, CA: Sage Publications.

Decker, R., & Hirshfield, S. (1998). *An introduction to programming using java.* Boston: PWS Publishing.

Doyle, S., Dodge, M., & Smith, A. (1998). The potential of Web-based mapping and virtual reality technologies for modeling urban environments, *Computers, Environment and Urban Systems*, *22*(2), 137–155.

Eck, J. E., & Weisburd, D. (1995). Crime places in crime theory. In J.E. Eck & D. Weisburg (Eds), *Crime prevention studies, vol. 4: Crime and place* (pp. 1-33). Monsey, New York: Criminal Justice Press.

Environmental Systems Research Institute (ESRI). (1998). *Using MapObjects on the Internet.* Redlands, CA: ESRI.

Graycar, A., & Grabosky, P. (2002). Introduction. In A. Graycar & P. Grabosky (Eds.), *The Cambridge handbook of Australian criminology* (pp. 1-6). Cambridge, MA: Cambridge University Press.

Green, L. (1994). Drug nuisance abatement, offender movement patterns, and implications for spatial displacement analysis. In C. R. Block & M. Dabdoub (Eds.), *Crime analysis through computer mapping* (pp. 21-32). Chicago: Illinois Criminal Justice Authority.

Griffith, D. A. (1987). *Spatial autocorrelation: A primer*. Washington, D.C.: Association of American Geographers.

Harries, K. D. (1980). *Crime and the environment*. Springfield, Illinois: Charles C. Thomas.

Kirkby, S. D., & Pollitt, S. E. P. (1998). Distributing spatial information to geographically disparate users: A case study of ecotourism and environmental management, *Australian Geographical Studies*, *36*(3), 262–272.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2001). *Geographic information systems and science*. Chichester,West Sussex: John Wiley & Sons.

Maltz, M., Gordon, A., & Friedman, W. (1991). *Mapping crime in its community setting: Event geography analysis*. New York: Springer-Verlag.

McEwen, J. T., & Taxman, F. S. (1995). Applications of computer mapping to police operations. In J. E. Eck & D. Weisburd (Eds.), *Crime and place:*

Crime prevention studies, vol. 4: Crime and place (pp. 259-284). Monsey, NY: Criminal Justice Press.

McFarlane, N., Chiarelli, A., De Carli, J., Li, S., Wilcox, M., Wilton, P., Wootton, C., & Updegrave, S. (1999). *Professional javascript*. Birmingham, UK: Wrox Press.

Morenoff, J., Sampson, R. J., & Raudenbush, S. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology, 39*, 517-560.

Morgan, F. (2003). The extent and location of crime. In A. Goldsmith, M. Israel & K. Daly (Eds.), *Crime and Justice* (pp. 11-38). Pyrmont, NSW: Law Book Company.

Morgan, F., & Fernandez, J. (2000). Mapping crime in western Australia. *GIS User, 43*, 24-26.

Murray, A. T., & Shyy, T. K. (2000). Integrating attribute and space characteristics in choropleth display and spatial data mining. *International Journal of Geographical Information Science*, *14*(7), 649-667.

Murray, A. T., McGuffog, I., Western, J. S., & Mullins, P. (2001). Exploratory spatial data analysis techniques for examining urban crime. *British Journal of Criminology, 41*, 309-329.

National Telecommunications and Information Administration. (2001). Falling through the Net: Defining the digital divide. In B. M. Compaine (Ed.), *The digital divide: Facing a crisis or creating a myth?* (pp. 17-46). Cambridge, MA: The MIT Press.

Openshaw, S., Turton, I., Macgill, J., & Davy, J. (1999). Putting the geographical analysis machine on the Internet. In B. Gittings (Ed.), *Innovations in GIS 6* (pp. 121-131). London: Taylor & Francis.

Peng, Z. R. (1999). An assessment framework for the development of Internet GIS. *Environment and Planning B: Planning and Design, 26*, 117–132.

Peng, Z. R., & Tsou, M. H. (2003). *Internet GIS: Distributed geographic information services for the Internet and wireless networks*. Hoboken, NJ: John Wiley & Sons.

Rich, T. (1995, July). The use of computerized mapping in crime control and prevention programs. *Research in action*. Washington, D.C.: National Institute of Justice.

Robinson, A.H., Morrison, J. L., Muehrcke, P. C., Kimerling, A. J., & Guptill, S.C. (1995). *Elements of cartography*. New York: John Wiley & Sons.

Sampson, R., Morenoff, J., & Earls, F. (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review, 64*, 633-660.

Sampson, R., Raudenbush, S., & Earls, F. (1997). Neighbourhoods and violent crime: A multilevel study of collective efficacy. *Science*, *277*, 918-924.

Servon, L. J. (2002). *Bridging the digital divide: Technology, community, and public policy*. Malden, MA: Blackwell Publishing.

Shyy, T. K., Stimson, R. J., & Murray, A. T. (2003). An Internet GIS and spatial model to benchmark local government socio-economic performance. *Australasian Journal of Regional Studies, 9*(1), 31-47.

Shyy, T. K., Stimson, R. J., Davis, R., Murray, A. T., Baum, S., & Barker, R. (2001). An Internet GIS for improving access to regional performance data: A decision support system to enhance local development and planning. *Australian Planner, 38*(3/4), 158-166.

Skogan, W. (1986). Fear of crime and neighborhood change, in A. J. Reiss & M. Tonry (Eds.), *Communities and crime* (pp. 203-230). Chicago: University of Chicago Press.

Weisburd, D., & Green, L. (1995). Policing drug hotspots: Findings from the Jersey City DMA experiment. *Justice Quarterly*, *12*(4), 711-735.

Endnotes

1 It is worth noting that these figures refer to offences recorded by police. The actual rates for property crime are likely much higher since crime victim surveys reveal that significant numbers of the victims of property crimes fail to report incidents to the police (Morgan, 2003).

2 http://64.218.68.50/stlouis/newslmpd/viewer.htm

3 Unofficial Brisbane crime data were provided by the Queensland Police Service. The analysis of the suburb crime data conducted by the research group should be treated as estimates only. The Queensland Police Service and the research group disclaim all responsibility and all liability (including without limitation, liability in negligence) for all expenses, losses, damages and costs you might incur as a result of the information being inaccurate or incomplete in any way, and for any reason.

4 Intel XEON, 36 GB hard disk, 1024 MB RAM

5 More than 78% of Brisbane suburbs have above national average household Internet access (Australia Bureau of Statistics, 2003). People can easily find our Web GIS application at home, at work or at library by entering the words “Web GIS for Mapping Community Crime Rates” and select “search page from Australia” using Google search engine.

Chapter XV

Identifying “Hot Link” Between Crime and Crime-Related Locations

Yongmei Lu, Texas State University-San Marcos, USA

Abstract

Crime is typically a multi-location event where multiple locations are associated through one crime incident. Understanding the patterns of the spatial association between crime locations and their corresponding crime-related locations (for example, the residence location of an offender or that of a victim) can enhance our capability to explain and predict crime patterns. GIS technologies coupled with spatial statistics have been widely used to model areas of high crime (that is, crime hot spot). But very limited effort has been spent to investigate the spatial association between crime locations from a crime hot spot and the corresponding crime-related locations. This chapter introduces the concept of “hot link” to describe the spatial autocorrelation of the one-to-one relationship between a crime location and a crime-related location. It develops an approach for the hot link analysis problem. Related techniques are applied to the hot link analysis between auto theft and recovery locations in the city of Buffalo, New York. Identifying the patterns of hot links from auto theft locations to

recovery locations is important for understanding auto thieves' travel behavior. Related findings can contribute to the law enforcement's effort to fight against auto theft. The hot link analysis method can be easily extended to analyzing spatial patterns of other types of crime and crime-related locations.

Introduction

It is until the last decade of the 20th century did the systematic applications of *Geographic Information Systems* (GIS) and related spatial analysis methods to crime pattern analysis become widely accepted (for example, Harries, 1999; Getis, Drummy, Gartin, Goor, Harris, Rogerson, 2000). One important topic in crime spatial pattern analysis and also one that has benefited from GIS and spatial statistics largely is crime "hot spot" analysis (Sherman, 1995). Investigating the spatial autocorrelation of crime incidents, hot spot analysis has been an important approach for the explanation and prediction of crime spatial patterns. Instead of focusing on crime location solely, another group of crime spatial analysis tries to explain the spatial patterns of crime through investigating criminals' mobility. Both journey-to-crime research (for example, Capone & Woodrow, 1976; Rossmo, 2000) and journey-after-crime analysis (for example, Lu, 2003) belong to this group. Moreover, there are studies looking into the spatial patterns of crime triangle defined by three anchor locations – offence location, criminal's residence, and victim's residence (for example, Rand, 1986). By linking offense locations with other crime-related locations, these analyses attempt to explain crime concentration as related to the spatial patterns of other anchor locations.

However, there is one important aspect that has been missing so far – the stability and intensity of the spatial association between different crime anchor locations. Are there consistent patterns showing that criminals who commit at certain crime hot spot are more likely to come from a hot spot of offender's residence locations? Do criminals who commit property crimes at certain hot spot tend to travel to clustered fence locations? Searching answers to these questions is a process of examining the spatial autocorrelation of the spatial association between crime and crime-related locations. This study refers to the spatial autocorrelation of the linkage between a crime hot spot and a crime-related location hot spot as "hot link," emphasizing that there are certain factors purporting the consistent links between these two hot spots. Identifying crime hot links can advance the understanding of crime by steering resources towards explaining hot links. Moreover, police manpower can be deployed more efficiently to break down the hot links so as to interrupt the underlying process purporting the patterns of crime and criminal activity.

Spatial Pattern Analyses of Crime

There is a rich literature on the spatial pattern analysis of crime and criminal's mobility. Related research has been devoted to revealing the patterns of crime and crime related locations from the following aspects:

(1) **Crime hot spots analysis – Where do criminals commit crime?**

This group of analyses focuses on identifying the high crime areas in a statistical manner. Early studies of crime hot spots lent support mainly from traditional descriptive statistics and concluded that the presence of certain facilities can increase the risk of vehicle theft: large, non-residential business areas (Eck & Spelman, 1988), street blocks with bars and taverns (Roncek & Maier, 1991), big parking lots (Clarke, 1983), among others. With the support of GIS and spatial statistics, the new generation of hot spot analysis follows *Exploratory Spatial Data Analysis* (ESDA) and investigates the spatial concentration of crime based solely on the information of geographical location (Ratcliffe & McCullagh, 1999; Craglia, Hanning & Wiles, 2000). These studies leave the explanation of crime hot spots to later investigation which links crime patterns with the patterns of ecological characteristics and land use. But crime hot spot analyses can not explain how and why areas of high offenses are not necessarily areas of high offenders (Sherman, 1995). They fail to link these two types of crime-related locations and hence can not provide a comprehensive picture of how and why crime hot spots are formed at certain places.

(2) **Analysis of crime related journeys – How far and to which direction does a criminal tend to travel before and after crime?**

Aiming at explaining how criminal's mobility impact the spatial patterns of crime, this group of analyses looks into the inter-location relationship between paired crime and crime-related locations. Many studies have investigated the travel distances (for example, Capone & Woodrow, 1976; Rossmo, 2000; Lu, 2003) and directions (for example, Costanzo, 1986; Lu, 2003) at a global scale. These studies generally concluded that criminals' trips are short and are directed towards certain types of land use. But they failed to reveal the spatial variations of the crime-related trips across the study area. They did not tell if the trips with their origins clustered together tend to go to similar distances in similar directions. This group of approaches are not able to disclose if a crime hot spot contains offenses committed by criminals who live close to each other (that is, related offenders' residences forming hot spots).

(3) **Crime triangle analysis – What is the spatial relationship between an offense location, an offender's residence location, and a victim's residence location?**

The major advance of this group of studies is that multiple crime-related locations are linked together when describing crime patterns. Rand (1986) examined three types of crime-related locations (the location of an offense, the residence of an offender, and the residence of a victim) for every offense to reveal if they are in the same census tract. It is actually a simple extension of the techniques for journey-to-crime analysis. Leitner and Binselam (1998) investigated the spatial association between the anchor locations for homicide triangle: the homicide locations, the victims' residences, and the offenders' residences. The spatial relationship between every two types of locations was evaluated using a global measurement of spatial association (Sorensen, 1974). But the one-to-one relationship between a crime location and the related offense's residence, as well as that between a crime location and the victim's residence, were overlooked. Similar to journey-to-crime approaches, this group of studies fails to provide information about the local statistics of criminal's mobility patterns. They do not examine if the offenders tend to travel from clustered residence locations to clustered offense locations.

Based on the above discussion, it can be concluded that (1) understanding criminal's travel behavior is important for explaining both crime hot spot and the underlying procedure generating the hot spot, and (2) the majority of current research of criminal's mobility stays at global description and is lacking an effective way to describe the spatial variation of criminals' trips at a local scale. Existing studies can not reveal if there is a statistically significant amount of crime-related trips that start from clustered locations and end at clustered locations. Therefore, there is an evident need to expend current crime analyses and to investigate the spatial patterns of the spatial association between crime and crime-related locations. Lu and Thill (2003) are among the first to develop techniques to investigate the correspondence of spatial clusters. Applied to crime and crime-related location analysis, this technique enables them to examine if vehicles that were stolen at clustered locations were recovered at clustered locations as well. However, this technique does not reveal where the recoveries are clustered. In another word, their approach does not show how vehicle theft hot spots and vehicle recovery hot spots are related on a spatial scale. Nevertheless, as discussed later in this chapter, a modified algorithm based on this technique can reveal the patterns of explicit hot link – a strictly defined crime hot link.

The need to investigate the spatial autocorrelation of the linkage between certain offense locations and their corresponding crime-related locations can also be justified by the progress in both empirical and theoretical research. Studies already confirmed that hot spots are formed for not only offense locations (for example, Sherman, 1995; Craglia et al., 2000) but also other crime-related locations including the residences of offenders and those of victims (Leitner & Binselam, 1998). It has also been found that the locations for dismantling stolen-vehicles tend to cluster (LaVigne, Fleury & Szakas, 2000). On the theoretical side, theories of crime and place, including routine activity theory (Cohen & Felson, 1979), rational choice theory (Cornish, 1993) and crime pattern theory (Brantingham & Brantingham, 1993), emphasize the impact of environment on a criminal's location decision. Despite their difference in assessing how socio-economic and behavioral factors may impact crime decision, these theories advocate that a crime location is closely related to a criminal's perception and knowledge of the space around and the locations near by. Compared with criminals who live far apart, it is reasonable to expect that criminals living close to each other tend to have similar activity space and awareness space (Brantingham & Brantingham, 1993). Hence, the offenders living close to each other tend to be impacted by similar environment factors and share similar spatial knowledge; they tend to make similar trips for crime. When linking the offenders' residences with the corresponding offense locations, one might find spatial autocorrelation among these crime-related trips.

The term, *hot link*, is used in this chapter to refer to the spatial association between a group of crime locations that are from one crime hot spot and a corresponding group of crime-related locations that are from one hot spot of crime-related locations. Different from a hot spot that contains locations of a single type, a crime hot link represents the relationship between two types of locations - certain crime locations and their corresponding crime-related locations. Put another way, the presence of a crime hot link is equivalent to the existence of spatial autocorrelation of the vectors linking one group of crime locations with the corresponding crime-related locations. However, it is important to notice the linkage and difference between *hot link* and *high cluster correspondence* defined by Lu and Thill (2003). The later was defined to represent the spatial autocorrelation of correspondence between paired locations forming two clusters (Lu & Thill, 2003); it is equivalent to a "narrowly defined" hot link. If a hot link connects two location clusters where every location in any cluster is related to one and only one location in the other cluster, high cluster correspondence exists, and there is an *explicit hot link* between these two clusters. Otherwise, if only a significant number of locations from one cluster are related to some locations in another cluster, there is no cluster correspondence, but a *relaxed hot link* can be defined.

A hot link between a group of crime locations and a group of crime-related locations might suggest the existence of common factors and processes that impact criminals' mobility. Policies can then be designed and manpower be directed to break down the hot link through mediating the factors and processes that sustain the related trips. The rest of this chapter demonstrates a pioneer study of crime hot link analysis. The method should be applicable to other crime hot link analyses.

Methodology

Crime events have been conceptualized as points on map ever since pin-maps were used to represent the distribution of crime. But the pin-map technology is restricted to simple patterns and smaller distributions (Sadahiro, 1997). Point pattern analysis methods (Diggle 1983; Boots & Getis, 1988; Fotheringham, Brunsdon & Charlton, 2000) have been used as major technical support for crime spatial analysis. However, the traditional point pattern analysis techniques are weak for crime hot link study. For example, for journey-to-crime research, the hot link analysis would reveal the existence of multiple trips that start from clustered criminal's residences and end at clustered offense locations. Traditional univariate point pattern analyses investigate single point set and can not handle this type of studies. Multivariate point pattern analysis (Diggle, 1983) treats a multiple point set as a simple mixture of points from different point sets without considering the pairing relationship between a point in a first set and one in a second set. Applying a multivariate point pattern analysis to examining the spatial relationship between a set of offenders' residences and a set of offense locations, one can only conclude on how the crime locations are spatially mixed with the residence locations at a global level. For example, Leitner and Binselam (1998) were only able to assess if a homicide location is closer to an offender's residence location than to any other homicide locations. This technique considers away the association of a homicide location with an offender's residence location through a homicide offense.

A hot link analysis of crime assesses the spatial autocorrelation of the spatial association between corresponding crime and crime-related locations. By associating multiple locations through a single crime event, crime hot link analysis deals with a typical multi-location problem (Lu & Thill, 2003). Lu and Thill (2003) refer to the spatial relationship between two point sets with every element in set one associated with an element in set two as "spatial correspondence." When both point sets are clustered, a cluster correspondence exists. The method proposed by Lu and Thill (2003) for local analysis of cluster correspondence between two point sets (referred to as *Lu and Thill algorithm* in this chapter) consists of the following four general steps: (1) to evaluate the clustering of any set of type I points, denoted as P_i, (2) to identify the corresponding type II points to form a paired set of points for P_i, denoted as Q_i, (3) to evaluate the clustering of points in Q_i, and (4) to link the clustering evaluation for P_i and Q_i to assess the type of cluster correspondence between them. If P_i and Q_i are both clustered at high significant level, a high degree cluster correspondence can be claimed. Among the various types of cluster correspondence, only the high degree cluster correspondence matters for crime hot link analysis. A high degree of cluster correspondence between crime locations and the corresponding crime-related locations would be equivalent to an *explicit hot link* between them. However, for crime hot link analysis, there is no need to evaluate the clustering of crime-related locations unless the related crime locations are clustered – a crime hot link exists only when both sets of locations are clustered. Put another way, there is surely no hot link if the crime location set does not show clustering in the first place. Hence, the Lu and Thill algorithm for cluster correspondence can be modified in the following way for crime hot link analysis:

Algorithm 1.

(1) to identify crime hot spots for the known crime locations,

(2) for a group of crime locations forming a crime hot spot, C_i, to identify the crime-related locations, offenders' residence locations as an example, so as to establish the corresponding group of residences, R_i,

(3) to assess the clustering of R_i,

(4) to define a crime hot link, $\{C_i, R_i\}$, if residences in R_i cluster, and

(5) to repeat step (2) through step (4) until all crime hot spots are evaluated.

This algorithm is strict and a hot link identified using this algorithm is an *explicit hot link* between the corresponding hot spots of different types. $\{C_i, R_i\}$ is a hot link when and only when both sets cluster and all elements in both location sets, C_i and R_i, are exclusively paired to each other. However, for the purpose of crime control, as long as a meaningful number of locations from a crime hot spot have their associated residence locations located in a hot spot of offender's residence, the linkage between these two hot spots warrants intensive investigation from law enforcement agencies. Hence, a hot link between C_i and R_i can be defined as long as a significant number of locations in C_i are paired with locations in R_i.

In another word, two hot spots linked by a hot link do not necessarily have all locations from one hot spot paired with locations from the second hot spot. If they do, the two paired hot spots are associated through an explicit hot link; otherwise, the two hot spots are associated through a *relaxed hot link*. The following algorithm can be used to identify relaxed hot link:

Algorithm 2.

(1) to identify all crime hot spots, HC_m, from a set of crime locations,

(2) to identify all hot spots of offender's residence locations, HR_n,

(3) for every HC_i, $i \in \{1…m\}$,to check the number of its location elements associated with locations in HR_j for all $j \in \{1…n\}$, and

(4) to define a hot link, $\{HC_i, HR_j\}$, if a significant number of crime locations in HC_i are paired with offenders' residence locations in HR_j.

It must be noticed that when implementing the above algorithms, several technical challenges exist. First of all, the techniques used to identify hot spots impact the final results of hot link analysis. Both algorithms start with identifying crime hot spots; algorithm 2 is based on the identification of both crime and residence hot spots. Although it is not a concern of this chapter to compare different techniques for hot spot analysis, it is well recognized that different techniques might reveal different hot spot patterns, which will define sets of crime locations for further hot link analysis. Secondly, when it comes to assessing the significance of concentration of point locations, as conducted in step (3) of algorithm 1 and step (4) of algorithm 2, different statistical or empirical methods might be applied. It is of high-priority for further research to evaluate the appropriateness of different methods for different data context. The technique presented in the following case study is but one implementation of the above algorithms.

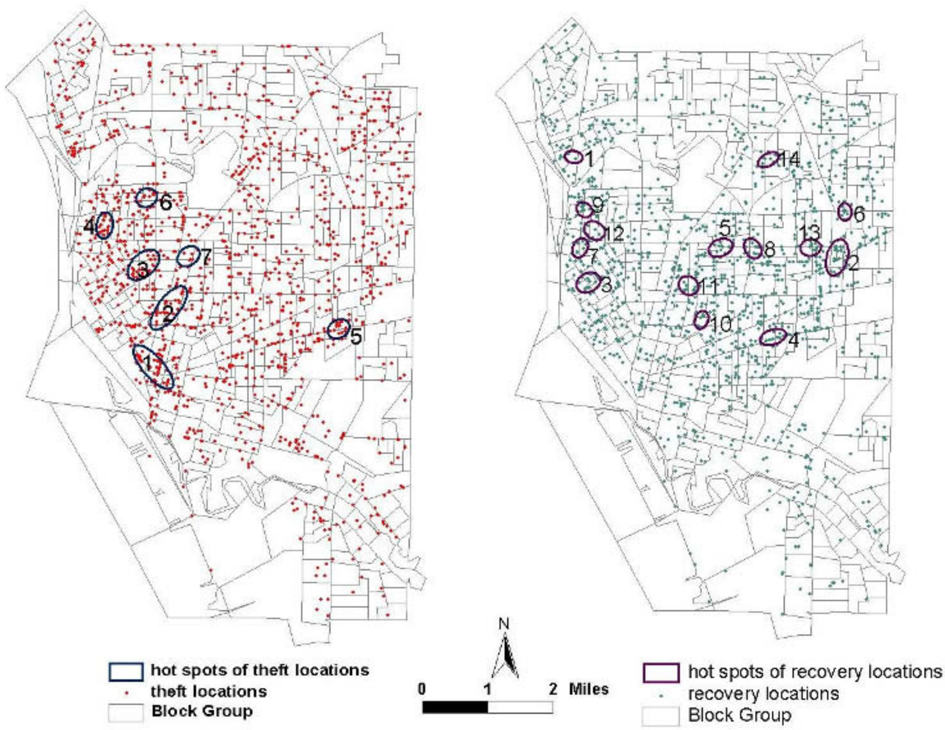
Crime Hot Link Analysis for Auto Thefts

Auto theft is the number one property crime in the United States. While many studies examined the spatial patterns of auto theft (for example, Copes, 1999; Potchak, McCloin & Zgoba, 2002; Rice & Smith, 2002), only a few have paid attention to the spatial patterns of stolen-vehicle recoveries (LaVigne et al., 2000; Lu, 2003). To the author's knowledge, Lu and Thill (2003) conducted the only analysis linking the spatial autocorrelation of auto theft locations with that of the corresponding recovery locations. The rest of this chapter reports two approaches for the spatial patterns of crime hot links between auto theft locations and the corresponding recovery locations. The first one is an investigation of explicit hot links – the hot link patterns that can be identified following algorithm 1 proposed in the previous section; the second one is a relaxed hot link analysis implementing algorithm 2.

More than 3271 auto thefts were reported in the city limit of Buffalo, New York, in 1998, resulting in a rate of 23.9 auto thefts per 1,000 households. In the same year, auto theft rate in urban area of the United States was 17.8 nationwide and 13.2 in the northeast states (Bureau of Justice Statistics, 2000). According to police records, there were 2284 recoveries of stolen vehicles in the same year in Buffalo. Through a series of GIS and database operation, a total of 1,467 auto theft-recovery location pairs were selected and geocoded to a TIGER/line street map for further analysis [1].

Although hot spot analysis is an important part for crime hot link study, it is not the focus of this study to evaluate the many different techniques for crime hot spot analysis. This study employs STAC (Block, 1994), a software program developed by the Illinois Criminal Justice Information Authority, to identify hot spots of both auto theft and recovery locations. Based on a scan-type algorithm, STAC repeatedly lays a circle on a grid and counts the number of points within the circle. A cluster is defined if more than expected number of points fall into a circle; circles of overlapping clusters are combined to form large clusters until there is no overlapping circles. Figures 1 displays the distributions of auto theft and recovery locations and the hot spots of these two types of locations. For explicit hot link analysis, auto theft hot spots displayed in Figure 1 are further investigated one by one to see if the corresponding recovery locations are statistically closer to each other than expected. The key step in this process is to assess the clustering of a group of recovery locations corresponding to the theft locations forming a hot spot. Adapting Lu and Thill algorithm for evaluating the closeness among paired points (Lu & Thill, 2003), it is assumed that the observed 1,467 recovery locations provide a good approximation of the possible recovery locations in the city; Monte Carlo simulations are then conducted to generate random distributions of 1,467 possible recovery locations; the confidence level of clustering among the observed recovery locations is derived by comparing the average distance among the observed recoveries with 99 distances each of which is an average distance among the recovery locations of a simulation. Table 1 reports the evaluation of the closeness of the seven groups of recovery locations, each of which is corresponding to one of the seven auto theft hot spots. It can be seen that there are hot link from auto theft hot spots 1, 2, 3 and 7 to the respective groups of corresponding recovery locations at 5% significance level. There exists spatial autocorrelation among the linkage from

Figure 1. Distributions of auto theft and recovery locations in the city of Buffalo



the auto theft locations forming one of the four theft hot spots to their corresponding recovery locations. Vehicles that are stolen from one of these hot spots are more likely to be recovered at or close to the area defined by the observed recovery locations. However, it should be pointed out that, any methodological change to identify theft hot spots and to evaluate the closeness among the corresponding recovery locations may result in different auto theft-recovery hot links.

However, the analysis of explicit crime hot link is restricted due to its requirement for an explicit correspondence between all locations in a theft hot spot and all locations in the corresponding recovery hot spot. Figure 2 shows the recovery locations for auto thefts forming theft hot spot number 4. Despite the low significant level of clustering among these recoveries (see Table 1), it is visually

Table 1. Evaluating the closeness of recovery locations that are corresponding to the auto theft locations making up the seven auto theft hot spots

hot spot of auto theft locations	number of auto thefts in hot spot	average distance among the corresponding recovery locations (feet)	Confidence level of clustering among recovery locations
1	36	10200.16	100
2	36	12491.99	96
3	31	12320.44	98
4	21	12869.82	73
5	16	14287.43	58
6	12	11074.86	94
7	10	11196.22	96

Figure 2. Distribution of recovery locations corresponding to auto theft locations forming hot spot 4

evident that a majority of these recoveries cluster in the central-west part of the city. As mentioned before, it is common for both law enforcement and crime research to investigate the spatial patterns of relaxed hot links. More specifically, if a number of recovery locations cluster together and their corresponding theft locations belong to a same theft hot spot, although these thefts do not make the full set of thefts in the theft hot spot, a relaxed auto theft-recovery hot link can be claimed. This hot link justifies a concentrated investigation from police and crime analysis on the relationship between these two areas associated by the auto theft hot link. A relaxed auto theft hot link analysis tries to identify the spatial association of a set of theft locations belonging to a theft hot spot with the corresponding recovery locations that cluster together. Hence, the result of a relaxed hot link analysis is closely related to both the way in which a theft hot spot is divided into subsets and the method that is used to assess the clustering of the corresponding recovery locations. Aiming at promoting the research of crime hot link than evaluating the merit of different techniques in handling related analysis, this study adopts a straightforward way to define the subsets of a theft hot spot in order to examine hot links.

As shown in Figure 1, STAC is used to identify hot spots of both auto theft locations and recovery locations. To identify a relaxed hot link, this study examines the intensity of the spatial association between each theft hot spot and each recovery hot spot indicated by the number of location pairs linking them. A hot link is defined between a theft hot spot and a recovery hot spot when the number of theft-recovery location pairs is significantly larger than the expected number. The "expected number of theft-recovery location pairs" is defined using a relative criterion – the total number of thefts within a theft hot spot divided by the number of recovery hot spots in the study area. By rule of thumb, a relaxed hot link is claimed when the number of paired locations from a theft hot spot to a recovery hot spot is equal to or greater than twice the expected number of links. Accordingly, a hot link might be claimed with a small absolute locations pairs (for example, between theft hot spot 7 and recovery hot spot 1 and 2, considering that there are only 10 thefts forming theft hot spot 7). The related results are reported in Table 2. Each bold number in the table indicated one presence of hot link from a theft hot spot (indicated by row) to a recovery hot spot (indicated by column). The value of the number represents the number of theft-recovery locations pairs linking a theft hot spot and a recovery hot spot. For example, the results shows that auto theft hot spot 4 has a hot link with recovery hot spot 12, which is in the central west part of the city (see Figure 1). This hot link is important because not only does it reveal the significant spatial association between these two hot spots, it also calls for further investigation to explain the consistent number of criminal's trips from auto theft hot spot 4 to recovery hot spot 12 (see Figure 1 for their respective locations).

It can be seen that the analysis of explicit hot link would grab only those very strict hot links between a theft hot spot and a recovery hot spot. While the definition is strict, the analysis result might be volatile since the way a theft hot spot is defined determines the subsets of auto theft and recovery locations that will be further evaluated for hot link. Applying different point analysis techniques, one might find more or less different patterns of crime hot spots. While an explicit hot link is denied by one method, a different definition of theft hot spot consisting of different offenses around the same theft location might lead to conclusion of a hot link with the corresponding recovery locations. The analysis of relaxed hot link relaxes the strict definition and allows for the establishment of a hot link when a significant number of locations from a theft hot spot are associated with certain locations from a recovery hot spot. By comparing the results reported in Tables 1 and 2 and by further linking them with the patterns revealed in Figures 1 and 2, one can see the difference between explicit auto-theft hot link and relaxed hot link between auto theft and recovery location pairs in the city of Buffalo. It is important to emphasize that the two approaches complement each other and are important to serve different levels of crime research and law enforcement practice.

Conclusion and Discussion

Starting from discussing the current research of spatial patterns of crime and criminal's mobility, this chapter advocates the expansion of crime spatial analysis from one that investigates the spatial autocorrelation of crime locations and the spatial relationship between crime and crime-related locations at a global level to one that examines the spatial autocorrelation of the spatial association between crime and crime-related locations at a local level. Crime hot link is introduced as a new concept to refer to the consistent and stable spatial association between a group of clustered crime locations and a corresponding group of clustered crime-related locations. Two types of crime hot link are further discussed, namely explicit hot link and relaxed hot link. The former requires that the hot link associates a crime hot spot with a hot spot of crime-related locations with every location in the crime hot spot corresponding to one location in the other hot spot and vice verse. The relaxed crime hot link recognizes a hot link between these two hot spots as long as a significant number of crime events link a crime hot spot with a hot spot of crime-related locations. It is "relaxed" since the requirement for the exclusive one-to-one relationship between the two hot spots, as defined by explicit hot link, is removed.

Table 2. Number of auto theft and recovery locations pairs that have the theft location in a theft hot spot and the corresponding recovery location in a recovery location hot spot

theft hot spot	hot spots of recovery locations (see map 1 for their spatial location information)														outside of any hot spot	Twice the Expected locations pairs	Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14			
1	1	0	**5**	1	0	0	3	1	1	3	1	3	0	0	17	5	36
2	0	4	0	0	0	0	1	1	1	1	1	0	1	0	26	5	36
3	0	0	2	1	2	0	0	0	1	1	2	1	0	0	21	4	31
4	0	0	2	0	1	0	2	0	1	0	0	**4**	0	0	11	3	21
5	2	0	0	**3**	0	2	0	1	0	0	0	0	0	0	8	2	16
6	2	0	1	1	1	0	1	1	2	1	0	0	0	0	2	2	12
7	**2**	**2**	1	1	0	0	0	0	0	0	0	0	0	0	4	1	10

The general algorithms for both types of crime hot link analysis are proposed in the chapter. One implementation of the algorithms on auto theft-recovery location pairs is reported. As pointed out early in the chapter, the application of univariate point pattern analysis techniques may impact the results of crime hot link analysis. First of all, since hot link analysis starts from hot spot identification, different crime hot spot analysis may provide different definitions of crime events set for hot link investigation. Secondly, for an explicit hot link analysis, the single point set analysis method to be used to evaluate the clustering among the crime-related locations is critical for evaluating the presence of hot link. Thirdly, it is central to a relaxed hot link analysis to determine if the number of location pairs between a crime hot spot and a crime-related location hot spot is more than "expected." The "expected number" of links between the two corresponding hot spots may be defined statistically or based on empirical evidence. The analysis of auto theft and recovery location pairs illustrated in this chapter aims at promoting hot link analysis. The evaluation of different techniques for identifying patterns of crime hot link is beyond the scope of this chapter. Therefore, the implementation of related algorithms in the chapter is by no means the only way or the best way for such a study. A comparison of different implementation of the hot link analysis algorithms warrants further investigation.

References

Block, C. R. (1995). STAC hot spot areas: A statistical tool for law enforcement decisions. In C. R. Block, M. Dabdoub & S. Fregly (Eds.), *Crime analysis* *through computer mapping* (pp. 15-32). Washington, DC: Police Executive Research Forum.

Boots, B., & Getis, A. (1988). *Point pattern analysis*. Newberry Park, CA: Sage Publications.

Brantingham, P. L., & Brantingham, P. J. (1993). Environment, routine, and situation: Towards a pattern theory of crime. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice: Advances in criminology theory vol. 5* (pp. 259-294). New Brunswick, NJ: Transaction Publications.

Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. (2000). *National crime victimization survey: Criminal victimization in United States, 1998 statistical tables*. Washington, DC: U. S. Department of Justice.

Capone, D., & Woodrow, W. (1976). Urban structure and criminal mobility. *American Behavioral Science, 20*(2), 199-213.

Clarke, R. V. (1983). Situational crime prevention: Its theoretical basis and practical scope. In M. Tonry & N. Norris (Eds.), *Crime and justice: A review of research vol. 4* (pp. 225-256). Chicago: University of Chicago Press.

Copes, H. (1999). Routine activities and motor vehicle theft: A crime specific approach. *Journal of Crime and Justice, 22*(2), 125-146.

Cornish, D. B. (1993). Theories of action in criminology: Learning theory and rational choice approaches. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice* (pp. 352-382). New Brunswick, NJ: Transaction Publishers.

Costanzo, C. M., Halperin, W. C., & Gale, N. (1986). Criminal mobility and the directional component in journey to crime. In R. M. Figlio, S. Hakim & G. F. Rengert (Eds.), *Metropolitan crime patterns* (pp. 73-93). Monsey, NY: Criminal Justice Press.

Craglia, M., Hanning, R., & Wiles, P. (2000). A comparative evaluation of approaches to urban crime pattern analysis. *Urban Studies, 37*(4), 711-729.

Diggle, P. J. (1983). *Statistical analysis of spatial point patterns*. London: Academic Press.

Eck, J., & Spelman, W. (1988). *Problem solving: Problem oriented policing in Newport News*. Washington, D.C.: Police Executive Research Forum.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2000). *Quantitative geography: Perspectives on spatial data analysis.* Thousand Oaks, CA: Sage Publications.

Getis, A., Drummy, P., Gartin, J., Goor, W., Harris, K., & Rogerson, P. (2000). Geographic information science and crime analysis. *URISA Journal, 12*(2), 7-14.

Harries, K. (1999). *Mapping crime: Principle and practice*. Washington, D.C.: National Institute of Justice, Crime Mapping Research Center.

LaVigne, N. G., Fleury, J. K., & Szakas, J. (2000). Auto theft and detecting chop shop locations. In L. S. Turnbull, E. H. Hendrix & B. D. Dent (Eds.), *Atlas of crime: Mapping the criminal landscape* (pp. 60-67). Phoenix, AZ: Oryx Press.

Leitner, M., & Binselam, S. A. (1998, November 10-12). *The application of spatial association techniques between point sets to crime analysis systems*. Proceedings of GIS/LIS 1998, Fort Worth, TX.

Lu, Y. (2003). Getting away with the stolen vehicle: An investigation of journey-after-crime. *The Professional Geographer*, *55*(4), 422-433.

Lu, Y., & Thill, J. (2003). Assessing the cluster correspondence between paired point locations. *Geographical Analysis*, *35*(4), 290-309.

Potchak M., McCloin, J., & Zgoba, K. (2002). A spatial analysis of criminal effort: Auto theft in Newark, New Jersey. *Criminal Justice Policy Review*, *13*(3), 257-285.

Rand, A. (1986). Mobility triangle. In R. Figlio, S. Hankim & G. Rengert (Eds.), *Metropolitan Crime Patterns* (pp. 117-136). Monsey, NY: Criminal Justice Press.

Ratcliffe, J. H., & McCullagh, M. J. (1999). Hotbeds for crime and the search for spatial accuracy. *Journal of Geographical Systems, 1*, 385-398.

Rice, K., & Smith, W. (2002). Socioecological models of automotive theft: Integrating routine activity and social disorganization approaches. *Journal of Research in Crime and Delinquency*, *39*, 304-336.

Roncek, D., & Maier, P. A. (1991). Bars, blocks, and crime revisited: Linking the theory of routine activities to the empirical of hot spots. *Criminology*, *29*, 725-753.

Rossmo, K. (2002). *Geographic profiling*. New York: CRC Press.

Sadahiro, Y. (1997). Cluster perception in distribution of point objects. *Cartographica*, *34*(1), 49-61.

Sherman, L. W. (1995). Hot spots of crime and criminal careers of places. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 35-52). Monsey, NY: Criminal Justice Press and Washington, D.C.: Police Executive Research Forum.

Sorensen, A. D. (1974). A method of measuring the spatial association between point patterns. *The Professional Geographer, 26*(2), 172-176.

Endnotes

1 Cases were selected according to the following criteria: (1) only those auto thefts with both stealing and recovery happened in 1998 were included for the analysis; (2) only those cases with both theft and recovery locations within the city limit of the city of Buffalo were selected; (3) only those cases with recovery locations different from the reported auto theft locations were selected, since cases having same address for both theft and recovery locations are more likely to be false reported cases according to the police.

Chapter XVI

Remote Sensing and Spatial Statistics as Tools in Crime Analysis

Dongmei Chen, Queen's University, Canada

John R. Weeks, San Diego State University, USA

John V. Kaiser, Jr., San Diego State University, USA

Abstract

This chapter explores the feasibility and utility of using aerial photography or remotely sensed satellite imagery to identify geographic or "place" features that may be associated with criminal activity. It assesses whether or not variables derived from satellite images can provide surrogate relationships between land use and crime. A review of the remote sensing literature suggests two basic approaches to the use of remotely sensed images in law enforcement: (1) tactical; and (2) analytical. The tactical approach uses the imagery as a background to the maps and other spatial information that an officer on the beat might have as he or she is investigating a crime or emergency situation. The analytical approach uses the remotely sensed images to create new variables that may serve as proxies for the risk of crime in particular locations. In this study we employ the analytical approach to the use of remotely sensed images, classifying

images according to the presence or absence of vegetation within a pixel, as well as the classification of specific urban attributes, such as parking lots. We also employ spatial statistics to quantify the relationship between features of the images and crime events on the ground, and these analyses may be particularly useful as input to policy decisions about policing within the community.

Introduction

The concept of place is essential to crime pattern theory because the characteristics of place influence the likelihood of a crime. Most crimes are not random events, nor are they randomly distributed in terms of where they occur (Rossmo, 1995). Some areas are more prone to criminal activity than are others (Coomb, Wong, Charlton & Atkins, 1994; Roncek & Maier, 1991). This spatial variability is a result of the spatially non-random distribution of people who will be motivated to perpetrate a crime, and the spatially non-random distribution of factors (the opportunities) that increase the odds that a person or property will be victimized (Hakim & Rengert, 1981). Motivation tends to be person-specific, whereas opportunity tends to relate more specifically to the characteristics of place (Eck & Weisburd, 1995). These place-specific characteristics may be institutional (such as the amount of police activity oriented toward preventing crime or arresting criminals) and/or they may be more environmental (such as the presence of a large parking lot full of automobiles).

Crime literature has abundant references relating crime patterns to specific geographic features. For example, opportunities for some crimes, such as burglary and robberies, may be particularly enhanced by the existence of commercial areas and parking lots (Canter, 1997; Hill, 2003). Brantingham and Brantingham (1994) reported crack houses induce crimes (dealing in illegal drugs) that have a multiplier effect in the neighborhoods in which they are located, raising the burglary and theft volumes in their vicinity as customers raise the money to buy the drugs. The recognition of the concept of place in crime theory allows a new dimension to implementing crime prevention. Mapping crime locations and associating crime activities to mapped urban features offers the potential to enhance an understanding of the non-random nature of crime locations and to improve crime prevention measures.

Crime event maps provide only a portion of the context of place. Context is provided more meaningfully through the use of remotely sensed images (aerial photographs and satellite images), which are then combined with the crime event maps in a geographic information system (GIS). Hirschfield and others (1995), in their study of GIS analyses of spatially-referenced crimes, report that the use of maps as backdrops to plots of spatially referenced crime events led to a dramatic increase in the number of visual clues presented to investigators and thus facilitated the interpretation of the pattern and location of crime incidents. As Hirschfield and Bowers (1997) and Olligsclaeger (2003) observed, crime pattern maps alone do not allow an investigator to analyze in depth the relationships between levels of crime and the social and physical environment. In addition to maps, information is also needed about the types of social and physical environment, which characterize areas of high crime. A potential solution to these crime-pattern map limitations is to use the powerful tools of remote sensing (RS) and aerial photography.

An extensive set of RS tools and techniques have been developed and are widely used by numerous disciplines to capture the characteristics of the physical environment (Jensen & Cowen, 1999). These same tools offer the potential to enhance the understanding of the relationships of crime to the physical environment and improve the understanding of place and geographic perspective in crime analysis. Aerial photo and remote sensing imagery can be a good source of data to provide the information on physical environment for law enforcement agencies.

Our interest in this research is to identify aspects of the natural and built environment that may be conducive to crime or may impose barriers to crime and thus will influence the opportunities for crime, and thereby will provide an independent determinant of the local crime rate. More specifically, we are interested in discovering whether remotely sensed images can provide information about this relative risk of crime that might not otherwise be available to crime analysts.

The objectives in this study are two-fold: (1) to use remotely sensed images to create new variables that may help to identify geographic locations that have a lower or higher relationship to crime; and then (2) to quantify the spatial mix of propensity and opportunity by bringing both sets of variables into a geographic information system for spatial statistical analysis. We do this with data for Carlsbad, California – a suburban community in San Diego County.

Background

Our spatial approach to crime analysis generally follows the human ecological paradigm of behavior (Poston & Frisbie, 1998) – that where you live influences your life chances and social contacts, which in turn influence a wide range of behavioral patterns, including criminal activity. The crime rate is determined by the combined effects of motivation, opportunity, and the distance between the geographic areas in which criminals reside and those in which opportunities for crime exist. For certain classes of crime, the local crime rate will be a function of the geographic concentration of people who fit the descriptive profile of those persons with a greater propensity to commit crime, and the geographic concentration of opportunities for crime.

A review of the remote sensing literature suggests two basic approaches to the use of remotely sensed images in law enforcement: (1) tactical; and (2) analytical. The tactical approach uses the imagery as a background to the maps and other spatial information that an officer on the beat might have as he or she is investigating a crime or emergency situation. Imagery could be used in tactical operations such as deploying law enforcement resources at the scene of an ongoing crime event (for example, bank robbery or hostage event). A tactical approach example using aerial photography and remote sensing in crime analysis was reported by Messina and May (2003) in a case of carjacking in Overland Park, Kansas. With the aid of aerial photo, the prosecution was able to help recreate the scene and provide the parties involved with a better visualization of the area. Tactical applications might also include crime event assessment, visibility analysis, situational awareness, ingress/egress control, and related uses.

The analytical approach uses the remotely sensed images to create new variables that may serve as proxies for the risk of crime in particular locations. In these applications the images are used to add data to the analysis of crime in an attempt to better understand the spatial distribution of crime within a community. In this study we employ the analytical approach to the use of remotely sensed images, classifying images according to the presence or absence of vegetation within a pixel, as well as the classification of specific urban attributes, such as parking lots.

Opportunities for some crimes, such as burglary and car theft, may be particularly enhanced by the existence of commercial areas and parking lots, and we propose to measure these characteristics with parcel maps and with satellite images from which we can determine parking lots and places with substantially reduced vegetation. Especially in southern California, vegetation will be denser in less densely settled areas, and will be less dense in commercial and multiple-family dwelling areas. Since the latter two categories of land use represent higher-than-average opportunities for several crimes, the vegetation index from a satellite image should provide a good surrogate measure of such risks.

Study Area

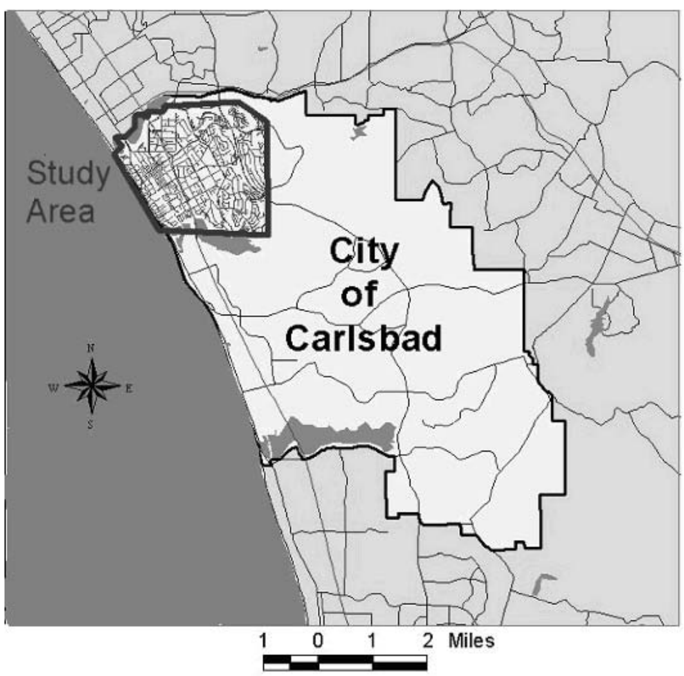
A subset of the city of Carlsbad in the county of San Diego was selected for study. This subset, bounded on the west by the Pacific Ocean, on the north and south by the Buena Vista and Aqua Hedionda lagoons and on the east by the El Camino Real highway, provided the desired diversity of land use categories. The study area includes 7,369 parcels. Figure 1 shows the spatial extent of the city of Carlsbad with an outline defining the area selected for image analysis.

Data and Methods

Dependent Variable – Spatial Pattern of Crime Events

Crime events recorded for the period from June 1995 through December1998 were used in this study. This period was selected to assure that the data

Figure 1. Map showing the location of the city of Carlsbad and the study area located in the northwest portion of the city



represented recent trends in both criminal activity and land use patterns within the study area. Several of the FBI Part I crimes, especially robbery, aggravated assault, and the four property crimes of burglary, larceny-theft, motor vehicle theft and arson, were the focus of this analysis. Analysis of these crimes has the advantages that (a) they are serious enough so that under-reporting should not be a major methodological issue (Coombs et al., 1994); and (b) community concern about them means that local police departments and elected officials have an interest in new analyses that might offer insights into lowering the risk of such crimes occurring. The database includes both arrests and reported crimes, but our focus in this analysis is on the location of reported crimes – where did the police go when notified of a crime? As shown in Table 1, there are data on 10,256 crime events within the categories listed above during the period under investigation for the geographic area within the study site. These crime events represent the dependent variables in our analysis.

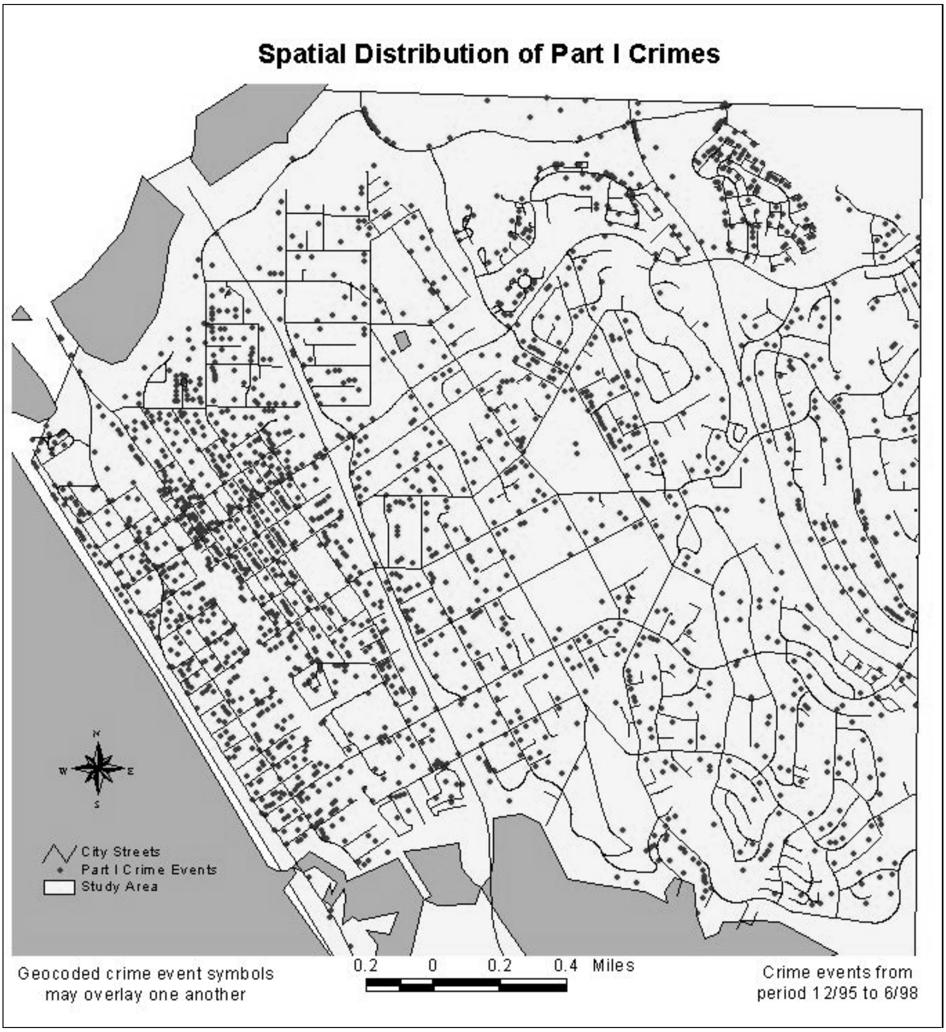
Figure 2 shows the spatial distribution of part I crime events geocoded by street address within the study area for the period of December 1995 through June1998. The map shows that the geocoded crime events are not uniformly distributed throughout the study area. It is important to note that geocoding generally assigns street addresses proportionally along street centerlines based upon street address numeric values and can result in lateral displacement when house addresses are not uniformly spaced along a street. Another potential limitation of the address geocoding process is that addresses are located along a street centerline and then offset to the left or right a uniform distance to place the crime location within a parcel. The off-set distance used in the crime database of this study was 45 feet.

There are more crime events in the urbanized western portion of the map and in multifamily residential areas in the northeastern corner of the map. The south

Table 1. Crime categories and counts in Carlsbad, California, for the period from December 1995 through June 1998

Part I Crime Category	Count	Percent
Arson	30	0.3%
Assault	1,717	16.7%
Burglary	1,861	18.1%
Larceny (Theft)	5,574	54.3%
Robbery	285	2.8%
Vehicle Theft	789	7.7%
	10,256	100.0%

Figure 2. Distribution of part I crime events by geocoded street address



central and southeastern portions of the study area are predominantly single-family, residential areas and show a both a lower density of events and more uniform distribution. There are also small clusters of crime events at the intersections of major thoroughfares. The spatial patterning of these crime events represents the dependent variable of interest in our analysis.

Independent Variables: Propensity and Opportunity for Crime

The data for the independent variables used in this study come from three different sources: (1) census data at the block group level and parcel map data used as a layer within the GIS; and (2) the classification of a satellite image. The sections below describe each source of data.

Census Data

From the 1990 census this source we derived variables that are related to the propensity to crime, often used to profile an area for crime risk. These demographic and housing characteristics include the percent Black, percent Hispanic, percent of the population aged 15-24, percent unemployed, percent with only a high school education or less, the percent that had been living in the same house five years prior to the census, the percent of the population that was not proficient in English, the percent of the population that was at or below the poverty level, the percent of homes that were occupied by renters, the percent of homes that were small (less than three bedrooms), and the percent of households that were one person. Data were aggregated at the block-group level, which was the smallest geographic unit of analysis available for these variables.

Several variables from the census were also used to measure opportunities for crime. These included data on the percent of households that are multiple-family dwellings, percent of housing units that are vacant, and the percent of housing units that were apartments.

Remotely Sensed Images

The ability to relate map features to some dependent variable (in this case, crime) is sometimes called spatial proximity analysis (O'Sullivan & Unwin, 2003). We have approached the application of information from remotely sensed images to crime analysis from two slightly different directions: (1) a generalized scheme of pixel classification designed to reduce the image to understandable patterns or spectral signatures that can then be tested for their association with hot/cold spots for specific types of crime; and (2) the "heads up" digitizing of specific environmental features that may relate to crime, such as the existence of alley-ways, parking lots, open spaces, proximity to freeway ramps, vacant lots and house setbacks, which may either raise or lower the opportunity for crime at a particular location.

A merged 10m SPOT PAN and 20 3-band SPOT XS image of the study area taken in 1995 was generated by RGB to IHS transformation of a three-band SPOT XS image and using SPOT PAN instead of intensity band in order to take advantage of the fine spatial resolution of PAN and the high radiometric resolution of XS. The merged SPOT PAN and XS image was classified for vegetation/non-vegetation in ERDAS Imagine software using an unsupervised classification scheme. First an ISODATA (Iterative Self-Organizing Data Analysis Technique) clustering method was performed on the entire study area to create ten different spectrally homogeneous classes (Jensen, 1996). These ten classes were then reduced to two categories (vegetation and non-vegetation) by visually examining the ten spectral classes in the screen. In this application, vegetation mainly indicates the areas covered by trees, brushes and grasses while the non-vegetation includes various buildings, roads, cleared lands and water. Several spectral classes generated by the classification might have mixed pixels or boundary pixels that can not be identified as being either vegetation or non-vegetation, thus producing an error term. The aid of 1m resolution scanned color aerial photos was used for these classes to identify the majority of classes in these pixels. For each class, the pixels belonging to it were displayed and geolinked to those in the 1m scanned color image. If the majority of pixels in the scanned image could be seen visually to be vegetation, then that class was assigned to vegetation.

A 300-meter grid was then draped over the classified imagery to create a set of data for spatial data analysis. The 300-meter size was chosen as a size large enough to ensure that there was a statistically adequate number of crime events within each cell, yet small enough to produce a sufficient number of cells (n = 285) for the statistical analysis. The classified image thus produced a variable representing the proportion of pixels within each grid cell that were classified as non-vegetated. As can be seen in Table 2, the average proportion of non-vegetation was 0.52, with a median of 0.50, a standard deviation of 0.18 and very little skewness to the distribution. These data are shown graphically in Figure 3.

From the aerial image we were also able to digitize the plots of land that were used for open, off-street parking. From the resulting polygons we calculated the percentage of the total area of each grid cell that was devoted to parking. These variables are also summarized in Table 2.

Spatial Data Analysis

The same 300 meter grid was then intersected with the crime data (the set of dependent variables – see Table 2), and all other independent variables.

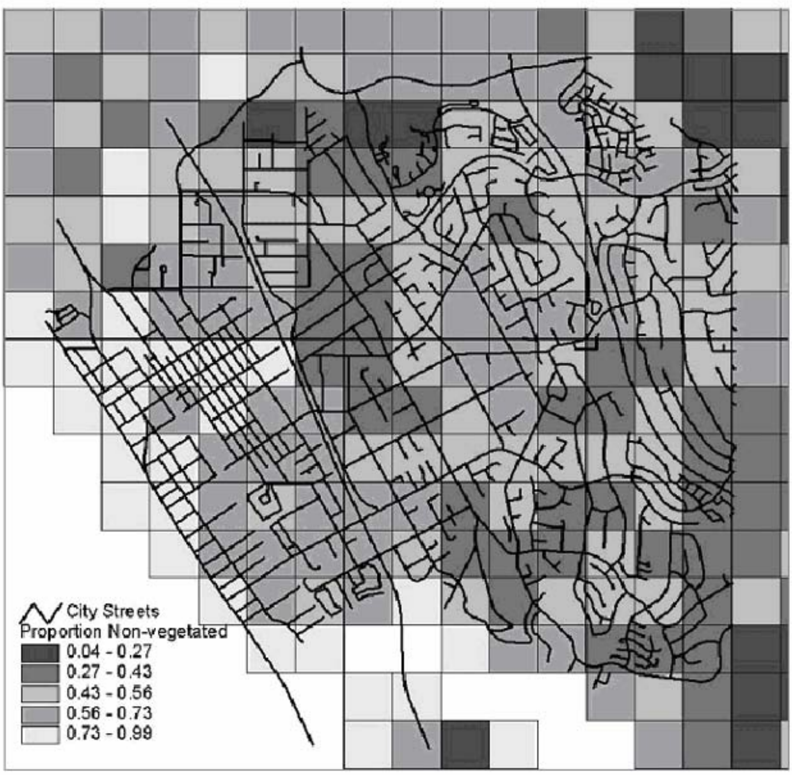
Table 2. Summary measures used in data analysis, based on 300m grid cell for study site

Type of variable	Variable name	Short name	Mean	Median	Standard Deviation
Dependent	All Part I crimes	ALLCRIM	41.58	11.5	95.43
	Robbery	ROBBREY	1.13	0.00	3.75
	Assault	ASSAULT	4.92	1.00	9.94
	Burglary	BURGLRY	6.55	3.00	9.57
	Larceny	LARCENY	24.42	6.00	69.30
	Auto Theft	AUTOTHFT	4.32	1.00	12.77
	Arson	ARSON	0.11	0.00	0.42
Independent - Propensity	Percent at or below poverty level	PCTPOVTY	6.94	4.61	5.89
	Percent unemployed	PCTUNEMP	5.40	4.26	8.57
	Percent high school or less	PCTHS	31.55	28.10	10.37
	Percent in same house 5 years ago	PCTNOMIG	32.91	34.23	15.43
	Percent not proficient in English	PCTNOENG	17.86	16.58	13.91
	Percent aged 15-24	PCT1524	13.76	13.55	4.77
	Percent Hispanic	PCTHISP	14.98	10.02	13.28
	Percent Black	PCTBLCK	1.62	1.29	1.10
	Percent housing units that are small (less than 3 bedroom)	PCTSMHS	17.12	15.45	13.84
	Percent housing units that are renter-occupied	PCTRENT	36.39	36.73	19.36
	Percent households that are one-person	PCT1PHH	21.97	19.05	9.54
Independent - Opportunity	Proportion of are classified as non-vegetation	NONVEG	0.52	0.50	0.18
	Percent of area used for parking	PCTPARKG	3.04	0.00	8.17
	Percent of area used for commercial	PCTCOMM	6.35	0.00	19.45
	Percent of area used for multiple- family dwellings	PCTMULTI	9.82	0.00	22.88
	Percent of housing units that are apartments	PCTAPTB	29.76	31.05	21.07
	Percent of housing units that are vacant	PCTVACNT	7.53	6.53	7.20

Organizing data in this way permits modeling with later regression techniques to test hypotheses about the environmental influence on the density of crime events. Given the potential inaccuracies in the geocoding of crime events, such an aggregation perhaps reflects a reasonable approximation of the environmental influence on criminal activity. Using ArcView we laid 300 meter grids over the study area, and then added the number of crimes in each grid. The dependent variable is thus the number of crime events occurring in each grid cell, and the independent variables are the proportions or rates of each variable occurring in each grid cell. These variables are summarized in Table 2. The census variables for each grid were aggregated from the block-group level using the weighted linear method. The weights were decided by the area percentages of blocks falling in each grid. This simple interpolation did not consider the spatial relationship between blocks and may over- or under-estimate values in some grids. However, these over- or under-estimated values should be small. The creation of the grid means that the dependent variable (number of crimes in each grid) is equivalent to a rate, since each number is implicitly normalized by an area of equivalent size.

The analysis proceeds in three steps: (1) an analysis of the spatial clustering of crimes using Gi* statistics (Getis & Ord, 1992) – are there some places within the study site where crime is much more likely to occur than others? (2) a traditional regression analysis in which spatial location is not taken into account—how much of the variability in crime can be explained by the predictor variables that attempt to measure the propensity and opportunity for crime; and

Figure 3. Percent non-vegetation by grid cell in study area



(3) a spatially filtered regression analysis in which we quantify the importance of spatial clustering as a predictor of the incidence of crime.

The data set includes a fairly large number of independent variables grouped under the broad headings of propensity and opportunity. Since each of the variables listed under these headings has a high likelihood of being correlated with the others, we decided to reduce the number of variables by means of a principal components factor analysis. Using a varimax rotation with 25 iterations, the number of variables under "propensity" was reduced to three components, using combinations of the variables shown in Table 2. The combination of these three components explained 80% of the total variation in the eleven independent

Table 3 shows the rotated factor loadings for each variable within each of the three statistically significant components. The factor coefficient scores for each variable were then used as weights to produce a weighted sum score for each component, thus creating a set of new, combined variables, which were then used in the remainder of the analysis. Thus, the initial eleven variables measuring propensity to crime have been reduced to three variables.

Using the same technique of factor analysis, we reduced five of the six variables that measured "opportunity" into two components, as shown in Table 4. The five variables reduced to two components, which together explained 66% of the variation in the constituent variables. The first component consisted of the variables "percent of area devoted to parking" and "percent of area in commercial property." We called this component "commercial." The second component included the variables "percent of housing units that are apartments," "percent of housing units that are vacant," and "percent of area in multiple family dwellings." We called this component "residential." These results are consistent with our more qualitative analysis of the data described above, in which commercial parcels and residential parcels seemed clearly to differ in the risk of crime events.

Several new spatial statistics permit increased quantitative sophistication of Crime Pattern Analysis (CPA) – a variation on Point Pattern Analysis (PPA) – which has long been an informal staple of efficient community policing (Openshaw et al., 1993). The Illinois Criminal Justice Information Authority has developed a program called STAC (Spatial and Temporal Analysis of Crime) to help detect clusters or "hot spots" of crime (Illinois Criminal Justice Information Authority, 1998).

Using raster grid data, we then introduce the spatial component at the local level. The local spatial statistic utilized is the $G^*_i(d)$ statistic (Getis & Ord, 1992; Ord & Getis, 1995), which measures the clustering of similar values around a given point at a specified distance from that point, relative to the point pattern in the entire geographic surface. There are two uses to which we shall put the $G^*_i(d)$ statistic. First, it has the ability to locate "hot spots" where low or high values are

Table 3. Rotated component matrix reducing the number of independent "propensity" variables

Original Variable:	Component 1 Immigrant Poor	Component 2 Poor, Young, Hispanic	Component 3 Mobile and Black
PCT1PHH	.933	.022	-.057
PCTSMHS	.884	.355	.163
PCTRENT	.757	.401	.410
PCTHS	.710	.496	-.133
PCTNOENG	.703	.517	.185
PCTPOVTY	.505	.767	.113
PCTUNEMP	.097	.736	-.106
PCTHISP	.549	.697	.071
PCT1524	.404	.612	.471
PCTBLCK	-.132	.197	.893
PCTNOMIG	-.258	.361	-.719

Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization; Rotation converged in seven iterations

clustered. These spatial clusters may then be investigated further (either qualitatively or quantitatively) to discover the sources of the clustering. The second use of the $G^*_i(d)$ statistic is as a spatial filter to extract the spatially autocorrelated portion of each of the variables in the regression variable, and then to reintroduce the spatial variable into the regression equation as a separate factor (Getis, 1995).

Results

Crime Event Distribution

Of the 7,369 parcels within the study area, only 1589 (21.6%) experienced a crime incident. Only 764 (10.3%) experienced more than one crime event. The five parcels with the highest crime events are located in the top central portion (major commercial land use) of the study area and account for over 3,100 crime events representing nearly one third (31.2%) of all crimes in the study. Of the 12 parcels having histories of greater than 50 crime events, a review of land use and aerial photographs revealed that nine parcels (75%) were associated with a shopping center land use and associated businesses. Of the remaining three parcels, two were associated with schools and one with a multi-family residence.

Table 4. Rotated component matrix reducing the number of independent "opportunity" variables

	Component	
	1	2
Original variable:	**Commercial**	**Residential**
PCTPARKG	.916	.007
PCTCOMM	.909	.007
PCTAPTB	.252	.798
PCTVACNT	-.120	.775
PCTMULTI	.387	.425

Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization; Rotation converged in three iterations

Examination of the aerial photography associated with these twelve high-crime incident parcels revealed several common geographic features that may be associated with increased crime activity. It is suspected that no single attribute is a dominant cause of increased crime activity by itself, but in combination with other elements, the net result is to make an individual parcel attractive as a location for perpetrating a criminal act. Eight of the nine shopping center parcels are associated with rapid freeway access. The single parcel without immediate freeway access possessed the lowest repeat crime incidence of the nine parcels. It is believed that this attribute contributes to concern about rapid ingress/egress by potential perpetrators. All shopping center parcels were adjacent to a major thoroughfare and had commercial parking lots serving multi-businesses.

It is believed that these attributes contribute to concerns about access and an ability to observe potential business targets or pedestrian victims. Parking lots appear to allow perpetrator presence and observation without raising suspicion. With multiple businesses serviced by a single parking lot, the presence of strangers (others business establishment customers) is common place and would not raise concern. All shopping center parcels with high crime rates have at least on parcel border consisting of a "dead" zone in which visibility and pedestrian access or both are limited. Those community shopping center parcels with the more than one parcel border bounded by a "dead" zone have higher crime rates than community shopping centers with a single "dead" zone borders.

Two of the 12 parcels having repeat-crime histories greater than 50 events are schools. It is suspected that the single greatest attribute associated with criminal activity on school parcels is not the physical environment of the school, but rather the concentration of individuals present in one place on a recurring basis. Review of the database's crime event records show that the majority of incidents occur on weekdays when schools are in session. Peak event times during school days correspond to the midday lunch hour and the end of the school day. The ability to recognize the location of schools from aerial or remotely sensed images may suffice as a surrogate indicator of increased crime at the school parcel and in immediately surrounding parcels within a temporally short walk time of the school.

One of the 12 parcels having a crime event history greater than 50 events was a multi-residence parcel. This parcel shared similar characteristics in terms of legitimized public space (proximity to a communal parking lot and a public park), barriers limiting visibility and restricted pedestrian thoroughfare (“dead” zone) and proximity to a major thoroughfare as higher crime rate parcels. However, this single land use observation precludes drawing any conclusions, although the similarities are intriguing.

Statistical Analysis Results

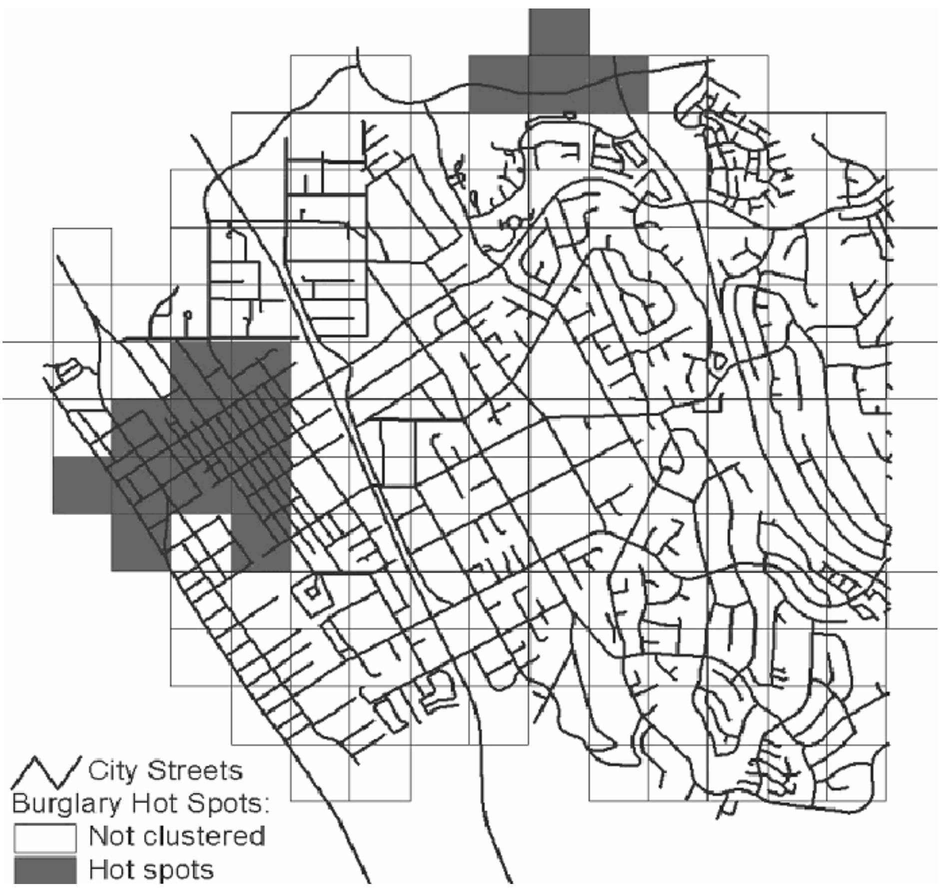
Geographic information systems have been increasingly applied to crime statistics to map “hot spots” – places where crimes tend to cluster spatially (Johnson, 2000). The analysis showed that burglaries and assaults were most closely related to the proportion of land in each grid cell that was classified as non-vegetated, and a qualitative sense of clustering can be gleaned from examining that figure. However, the statistical analysis of clustering, using the G* statistic as discussed above, produced the results in Figure 4, which plots the grid cells in which the number of burglaries is clustered to a statistically significant degree (at the .05 level of statistical significance). These cells represent the burglary “hot spots” in Carlsbad.

Regress analysis was used to test if independent variables can explain this spatial patterning of crime in Carlsbad. The independent variables are reduced to three indices of propensity after the principal component factor analysis to commit crime (based on demographic descriptors – “immigrant poor,” “poor, young, and Hispanic,” and “mobile and black”) and three indices of the opportunity to commit crime (based on census data, parcel data, and imagery data—“commercial,” “residential,” and percent non-vegetated). These predictor variables were regressed on each type of crime. Table 5 summarizes the results in terms of the adjusted R^2 and the statistically significant predictor variables.

The data in Table 5 show that the opportunity factor of being commercial was the single most important predictor of crime in the Carlsbad. This variable emerged as the most important statistically significant predictor of each type of crime under analysis in this study. Robbery was also influenced by the factor of poor-young-Hispanic, and the opportunity factor of residential areas that were disproportionately multiple family dwellings and vacant dwellings. Assault was influenced by the factor of poor-young-Hispanic, and also by the opportunity factor of the index of non-vegetation, meaning that areas that were less vegetated were more likely to be associated with assaults. Larceny and auto theft were influenced only by the commercial factor.

The results for burglary showed the highest level of prediction by our assembled set of variables, and the detailed results of the regression model for burglary are shown in Table 6. As is true with each type of crime, the identification of a grid cell as being commercial was the single most important predictor of where burglaries were occurring, but the socio-demographic profile associated with the immigrant-poor was also a significant predictor, as was the percent of the grid cell that was classified as being non-vegetated. The image classification appears

Figure 4. Burglary hot spots in the study area using G statistic*



to be a surrogate for the areas that are more densely built-out and that are proximate to the commercial areas. On its own, the index of non-vegetation is able to explain 22% of the variation in crime in Carlsbad. In combination with the other variables, it explained nearly half (48%) of the spatial variability in crime in Carlsbad.

The regression model for the prediction of burglaries, as shown in Table 5, was tested for autocorrelation in the residuals. If a statistically significant level of autocorrelation were present, it would indicate that a correction would be necessary in the regression results and thus the spatial filtering process described above would have been appropriate. The possibility of autocorrelation was heightened, of course, by the fact that the census data were collected at the block group level, but we applied those data to a smaller areal unit of the 300 meter grid cell. Thus, contiguous grid cells might have identical values for census-derived variables. However, the residuals were not spatially autocorrelated, indicating that the spatial component of the relationship is included within the data themselves. Therefore, no additional adjustment was necessary. As a result, we can proceed directly to an assessment of the usefulness of the regression results. Could we predict the burglary hot spots in Carlsbad based solely on the combination of census data and the data derived from the remotely sensed images?

The answer is shown graphically in Figure 5, where it can be seen that the hot spots predicted by the regression model, calculated using the G* statistic, are centered on the same hot spots generated by the crime data alone. The predicted values are not quite so tightly focused geographically as are the actual data, but nonetheless the data suggest that the combination of census and remotely sensed data may, on their own, be powerful predictors of where burglaries are occurring in a community, largely because they help to identify those places in which the apparent opportunities for burglaries exist.

It should be noted that our model does not include a feedback mechanism and so it is not able directly to capture the dynamic nature of crime as outlined by Freeman and his associates (1996), in which it is hypothesized that high crime

Table 5. Predicting crimes in Carlsbad

Type of Crime	**R**	**R^2**	**Statistically significant predictor variables (in order of size of standardized beta coefficient)**
Burglary	0.700	0.486	Commercial, Immigrant Poor, Non-vegetation index
Robbery	0.610	0.356	Commercial, Poor-Young-Hispanic, Residential
Larceny	0.599	0.346	Commercial
Assault	0.562	0.303	Commercial, Non-vegetation index, Poor-Young-Hispanic
Auto Theft	0.484	0.225	Commercial
ALL PART I	0.642	0.400	Commercial

Table 6. Which factors best predict where burglaries occur in Carlsbad?

Variable	Beta	T	Significance
Commercial *(opportunity=pct commercial+pct parking)*	0.486	7.850	0.000
Immigrant Poor *(propensity=pct no english+pct high school or less+pct renters+pct small house+pct 1 person hh)*	0.250	2.334	0.021
Percent non-vegetated	0.168	2.388	0.018
R = .704; Adjusted R^2 = .483			
Dependent variable = Number of burglaries in grid cell			

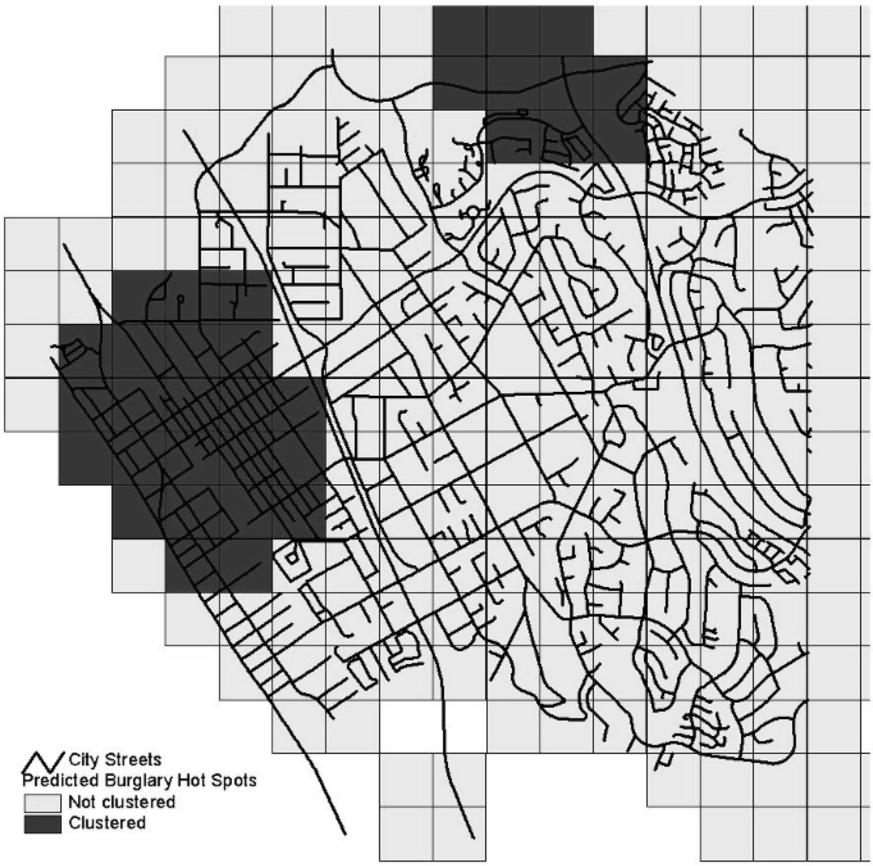
areas beget additional crime because the probability of arrest declines as the number of criminals increases, thereby increasing the motivation to engage in criminal behavior. However, to the extent that those neighborhoods in which crime does exist are identified by the variables we have included in the model, then the feedback should be incorporated indirectly into the standardized beta coefficients of the neighborhood variables.

Summary and Conclusion

This chapter illustrates that crime locations are not spatially random and that place characteristics influence the decision to commit a crime. The study also suggests that certain land use activities are more attractive to criminal activity than are others, including commercial areas and residential areas dominated by apartments. Additional location factors such as proximity to freeway on/off ramps, location adjacent to a major thoroughfare, the presence of parking lots that serve multiple businesses and the presence of visibility and pedestrian-thoroughfare barriers also contribute to making some shopping center locations more attractive to the commission of crime than parcels lacking these attributes.

The study also suggests that many of the location factors which make certain land parcels attractive to committing a crime can be observed and mapped from aerial photographs and high-resolution remotely sensed imagery. Aerial photography and remotely sensed imagery having spatial resolutions sufficient to detect and identify features with spatial dimensions less than one meter are necessary to identify the surrogate crime variables identified in the study. Color imagery was superior to panchromatic imagery in the detection and identification process. These approaches to the use of remotely sensed images will tend to be community-specific and require detailed interpretation of the images.

Figure 5. Burglary hot spots in Carlsbad predicted by the combination of the factors "commercial," "immigrant-poor," and "percent non-vegetated"



It is demonstrated that the classification of a satellite image into an index of non-vegetation facilitated the prediction of where burglaries occurred in the study site, and increased the prediction of geographic hot spots of burglary. The prediction of these hot spots did not reveal any startling new information to the local police department when this information was presented to them, but it did quantify their otherwise *ad hoc* impressions of where criminal activity was concentrated within their community. This process of quantification can be important as a policy tool to direct attention to the exact policing needs in the community.

Acknowledgments

This research was supported in part by a grant from the NASA Earth Science Enterprise Applications Division to the Affiliated Research Center in the Department of Geography at San Diego State University (Douglas Stow, principal investigator). The authors wish to thank Tim Dolan, Milan Mueller and especially Arthur Getis for their valuable assistance. The authors also would like to thank the anonymous reviewer and Fahui Wang for their insightful comments and suggestions. The support from the Academic Research Committee, Queen's University, to the first author is also gratefully acknowledged.

References

Brantingham, P. L., & Brantingham, P. J. (1994). Location quotients and crime hot spots in the city. In C. R. Block & M. Dabdoub (Eds.), *Proceedings of workshop on crime analysis through computer mapping* (pp. 129-149). Chicago: Illinois Criminal Justice Information Authority and Loyola University.

Canter, P. R. (1997). Geographic information systems and crime analysis in Baltimore County, Maryland. In R. V. Clarke (Ed.), *Crime prevention studies* (pp. 157-190). Monsey, NY: Willow Tree Press.

Coombs, M., Wong, C., Charlton, M., & Atkins, D. (1994). Crime risk in urban and rural neighbourhoods: An experimental analysis of insurance data. *Environment and Planning B: Planning and Design, 21*, 489-504.

Eck, J. E., & Weisburd, D. (1995). Crime places in crime theory. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 1-33). Monsey, NY: Criminal Justice Press.

Freeman, S., Grogger, F., & Sonstelie, J. (1996). The spatial concentration of crime. *Journal of Urban Economics, 40*, 216-231.

Getis, A. (1995). Spatial filtering in a regression framework: Examples using data on urban crime, regional inequality, and government expenditures. In L. Anselin & R. Florax (Eds.), *New directions in spatial econometrics* (pp. 172-188). Berlin: Springer-Verlag.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*, 189-206.

Hakim, S., & Rengert, G. F. (1981). Introduction. In S. Hakim & G. F. Rengert (Eds.), *Crime spillover* (pp. XX-XX). Beverly Hills, CA: Sage Publishers.

Hill, B. (2003). Operationlizing GIS to investigate serial robberies in Phoenix, Arizona. In M. R. Leipnik & D. P. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 146-158). New York:Taylor & Francis.

Hirschfield, A., &. Bowers, K. J. (1997). The development of a social, demographic and land use profiler for areas of high crime. *British Journal of Criminology, 37*(1), 108-121.

Hirschfield, A., Brown, P., & Todd, P. (1995). GIS and the analysis of spatially referenced crime data: Experience in Merseyside, U.K. *International Journal of Geographical Information Systems, 9*(2), 191-210.

Illinois Criminal Justice Information Authority. (1998). *STAC users manual*. Chicago: Illinois Criminal Justice Information Authority.

Jensen, J. R. (1996). *Introductory digital image processing: A remote sensing perspective, second ed*. Upper Saddle River, NJ: Prentice-Hall.

Jensen, J. R., & Cowen, D. C. (1999). Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing, 65*(5), 611-624.

Johnson, C. P. (2000, January). Crime mapping and analysis using GIS. In *Geomatics 2000: Conference on geomatics in electronic governance*, Pune, India.

Messina, J., & May, J. (2003). Aerial photography and remote sensing for solving crimes. In M. R. Leipnik & D. P. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 232-240). New York: Taylor & Francis.

O'Sullivan, D., & Unwin, D. J. (2003). *Geographic information analysis*. Hoboken, NJ: John Wiley & Sons.

Olligschlaeger, A. (2003). Future directions in crime mapping. In M. R. Leipnik & D. P. Albert (Eds.), *GIS in law enforcement: Implementation issues and case studies* (pp. 103-109). New York: Taylor & Francis.

Openshaw, S., Cross, A., & Waugh, D. (1993, March 13-15). *Round up the usual suspects*. GIS Europe conference proceedings, London.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis, 27*, 286-306.

Poston, D. L., & Frisbie, W. P. (1998). Human ecology, sociology, and demography. In M. Micklin & D. L. Poston (Eds.), *Continuities in sociological human ecology* (pp. 27-50). New York: Plenum Press.

Roncek, D., & Maier, P. (1991). Bars, blocks and crimes revisited: Linking the theory of routine activities to the empiricism of hot spots. *Criminology, 29*(4), 725-751.

Rossmo, D. K. (1995). Place, space, and police investigations: Hunting serial violent criminals. In M. R. Leipnik & D. P. Albert (Eds.), *Crime and place, crime prevention studies, vol. 4* (pp. 217-235). Monsey, NY: Criminal Justice Press.

Section VI

Crime and Community Characteristics

Chapter XVII

Routine Activities of Youth and Neighborhood Violence: Spatial Modeling of Place, Time, and Crime

Caterina Gouvis Roman, The Urban Institute, USA

Abstract

This chapter discusses how a geographic information system (GIS) and spatial analysis are used to model the relationship between the daily routine activities of youth and rates of violence, and provides an example of how these techniques can be applied to analytical studies examining violence in places. Most research informing hotspots and related crime prevention efforts focuses on the contribution of places and the physical environment to crime levels. Yet very little is known about how time influences patterns of crime across places and environments. This chapter discusses how time of day, week and year can be incorporated into spatial analysis of crime patterns to further inform crime prevention. A model of opportunity factors is developed to predict the spatial and temporal relationship among violence, schools, youth hangouts, retail properties *and neighborhood disorganization across census blocks. Instrumental variables regression is used to estimate spatial lag models of violence.*

Introduction

Over the past few decades, spatial analysis has gained prominence within criminological research (Anselin, Cohen, Cook, Gorr & Tita, 2000; Taylor, 1998). Researchers often use spatial analysis to examine environmental and contextual factors that influence the distribution of crime and violence across neighborhoods. An increasing number of studies are examining how areas smaller than neighborhoods—places such as blocks or street corners—either inhibit crime or create opportunity for it (Block & Block, 2000; Brantingham & Brantingham, 1982; 1995; LaGrange, 1999; Roncek, 2000; Smith, Frazee & Davison, 2000).

Many of these place-based studies are examining whether places and particular facilities such as schools, bars, liquor stores, transit stations and public housing complexes are generating crime in their surroundings. These studies often examine how place features interact with larger neighborhood characteristics, such as residential mobility and economic deprivation, to determine the structure of opportunity for crime. This small-area research is a needed step to inform the neighborhood-place-crime nexus. Studying this nexus narrows the long list of characteristics associated with high-crime areas and provides guidance to a broad range of program initiatives and public safety strategies.

However, the body of empirical literature informing crime opportunity in places rarely takes into account how opportunity structures vary by time of day. Examining how crime clusters in both space and time can lead to more effective crime prevention and suppression policies. Furthermore, many of the crime and place studies examine bars, liquor stores and other retail establishments in isolation from each other and other potential crime-generating institutions. In addition, relatively little attention has been given to institutions that are present in all types of communities where people congregate, such as schools, recreation centers and malls. This is a critical oversight because these institutions are not zoned or restricted to certain neighborhoods, as bars or liquor stores are. Examining risk of crime in places across varying times of the day will allow for the assessment of a greater variety of interactions between places and their social context, and, in turn, allow for more insight into possible causal mechanisms that create the opportunity for crime. Effective crime prevention requires identification of times and places where people are likely to gather as potential criminal offenders or victims.

This chapter discusses how time of day, week and year can be incorporated into a geographic information system (GIS)-based analysis of aggregate crime patterns in small areas to further inform crime prevention. The research discussed here employs geographically referenced data, including incident-based crime data to model aggregate violent crime patterns based on the routine activities of youth. Routine activity patterns generate changes in the flow of potential victims and offenders that can facilitate or inhibit the opportunity for violence. A model of opportunity factors is developed to predict the spatial and temporal relationship among violence, schools, youth hangouts, retail properties and neighborhood disorganization across census blocks. GIS offers the ability to examine new relationships among variables—bringing together many types of data (for example, crime, transportation, census, housing and land use) to enhance crime analysis. Knowledge of specific contexts of violence can inform problem solving to increase public safety. For instance, do certain types of areas and neighborhoods need solutions for youth-oriented crime? Should public safety solutions be time specific, place specific, or both? How can communities and government agencies maximize prevention and enforcement efforts?

Background

For the research discussed here, two theories provide the basis for examining criminal opportunity posed by places. First, social disorganization theory posits that crime and violence occur in neighborhoods characterized by economic deprivation, racial and ethnic heterogeneity, and residential instability (Park, 1926; Shaw & McKay, 1942). Current social disorganization theorists study these structural characteristics of neighborhoods with an added focus on informal social control processes influenced by networks of relationships (Bursik, 1999; Bursik & Grasmick, 1993; Sampson, Raudenbush & Earls, 1997). Historically, the unit of analysis has been an ecological area larger than the neighborhood – the city, county or census tract.

Second, the routine activities perspective is based on the premise that opportunities for crime arise when three elements converge: a motivated offender, a suitable target and a lack of capable guardians. The routine activities approach states that the conduct of daily activities or "routine" activities delivers the opportunities for crime to occur. Routine activities theory provides the framework to understand how facilities or block features can be attractors or generators of crime and is also useful in understanding why violence might vary over different times of the day (Felson, 1987, 1994). Site features or facilities that draw people to places like schools and bars can increase the likely convergence of motivated offenders and suitable targets. With regard to time, there are periods of the day or week when motivated offenders can commit crimes out of view of potential guardians. With a decrease in capable guardianship, motivated offenders have more opportunity to victimize potential targets. Hence, certain routine activities, by definition, evade informal control. With regard to youth, this is the case for youth walking to school, congregating at bus stops, or hanging out with peers, out of the range of parental supervision. In sum, patterns of routine behavior are linked to times of day and the routine activity patterns generate changes in the flow of people through places. As the number of potential person targets increases (holding constant the presence of motivated offenders and absence of capable guardians), the opportunity for violence increases.

A few studies analyzing patterns of youth violence have provided some evidence of how opportunity structures can vary by the time of day (Garofalo, Siegel & Laub, 1987; McManus, 2001; Snyder, 1999; Wiebe & Meeker, 1998). Examining South Carolina incident-based data on youth, McManus (2001) found that for violent gun crimes in school or on school grounds, the crimes were most common on weekdays when school was in session. Research also has confirmed that patterns of crime are different depending on the school activity in which the youth is participating. An examination of youth victim narratives from a 1982-1983 sample from the National Crime Victimization Survey (NCVS) found that students were more often victims of violent crime while traveling to and from school or waiting for a school bus (Garofalo et al., 1987) than while in the classroom during the day. Other studies examining nighttime activities and crime have shown a relationship between time of day and patterns of assault and residential burglaries (Massey, Krohn & Bonati, 1989; Roundtree & Land, 1996; Sampson & Wooldredge, 1987).

Temporal Model of Violence in Places

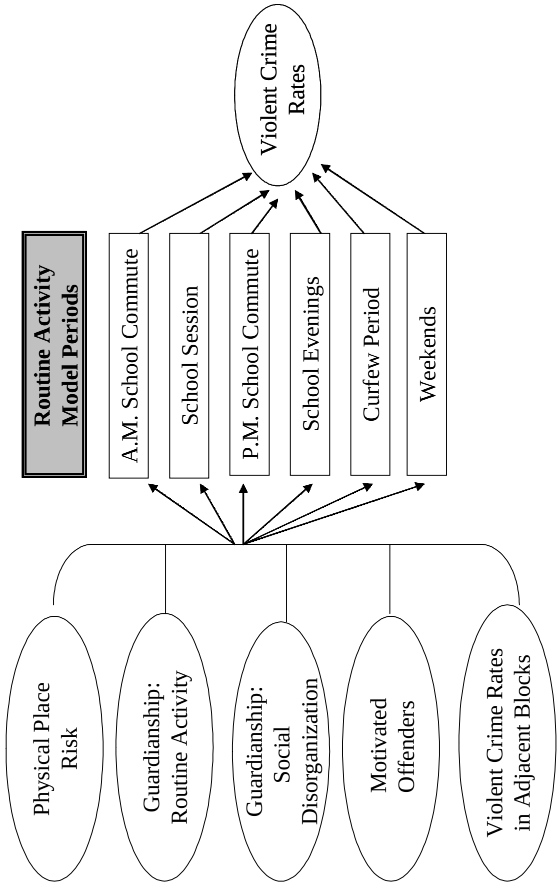
Research suggests that domain-specific models (for example, work, leisure, school) will facilitate drawing the causal link between opportunity measures and victimization (Garofalo et al., 1987; Gottfredson, 1984; Lynch, 1987). The analysis of criminal violence presented here focuses specifically on the routine activities of youth attending school. Following the daily activities or paths of teenagers can provide better understanding of when and where a large portion of violent crime occurs. In 1999, youth between the ages of 12 and 15 and between 16 and 19 had an annual rate of violent crime of 67 and 71 per 1,000 persons, respectively, much higher than any other age group (Rennison & Rand, 2003). Because offenders generally commit their offenses near places where they spend most of their time (Brantingham & Brantingham, 1991; Cohen & Felson, 1979), it can be reasoned that youth offenders commit a portion of their offenses near schools or on pathways to and from school. Similarly, large schools where youth are not closely monitored may provide adult potential offenders situations that facilitate violence. Restricting the scope of inquiry to one specific domain of life activity – in this case, attending school – can increase the explanatory power of routine activity models and at the same time provide useful information that might lead to specific solutions to crime problems (Lynch, 1987). The models tested here are designed to capture other relevant features of the environment that create the opportunity for violence. Furthermore, incorporating time of day into the models will assist with understanding potentially different activity patterns.

To examine neighborhood violent crime, a model of opportunity factors is developed that is divided into variable clusters grounded in routine activities theory and social disorganization theory. The variable clusters represent (1) the risk associated with the physical space or setting, (2) the potential for human surveillance or guardianship, (3) the larger neighborhood structural characteristics that represent the potential for informal social control or guardianship, (4) the potential for motivated offenders to be present, and (5) the influence of violence in adjacent neighborhoods. The model is useful to examine whether time of day adds an important dimension to understanding geographic locations of violence. Given that there are different flows of targets and offenders throughout the day following the routine activities of youth attending school, a key research question is “how does the physical place risk change across time periods?” One can hypothesize that blocks with schools would witness higher rates of violence during the school session and commuting periods, and that youth hangouts would have the highest rates of violence in the after-school commute period. Commercial establishments would likely have high rates of violence later into the evening or during the weekends. Similarly, school bus stops, though not physical *structures* or *facilities,* would provide the opportunity for crime, given the large number of youth congregating without capable guardians.

The conceptual model is presented in Figure 1. The ellipses represent the variable clusters that are hypothesized to influence violence in places. The variable cluster for guardianship can be disaggregated depending on the theory that informs variable selection – either routine activities theory or social disorganization. Within the model, time of day mediates the relationship between the independent variables and violent crime. Following the routine activities of youth, the time periods are as follows:

- Morning school commute (Mon., Tues., Wed., Thr., and Fri. 6 a.m. to 9:59 a.m.);

Figure 1. Temporal model of violent crime in places



- School day (Mon., Tues., Wed., Thr., and Fri. 10:00 a.m. to 1:59 p.m.);
- Afternoon school commute (Mon., Tues., Wed., Thr., and Fri. 2 p.m. to 5:59 p.m.);
- School evening (Sun., Mon., Tues., Wed., and Thr. 6:00 p.m. to 9:59 p.m.);
- School night curfew (Sun., Mon., Tues., Wed., and Thr. 10:00 p.m. to 11:59 p.m.); and Mon., Tues., Wed., Thr., and Fri. 12:00 a.m. to 5:59 a.m.); and
- Weekends (Friday 6:00 p.m. through Sunday, 5:59 p.m.).

Study Area and Data

The study area comprises all census blocks in Prince George's County, Maryland, (7,334 blocks). Prince George's County is a large (488 square miles), high-crime county. The county surrounds the District of Columbia along both the District's northeast and southeast quadrants. Home to the University of Maryland, half of the county's 767,413 people are black (51%), 43% are white, 4% are Asian and 4% are Hispanic. In 2000, there were 166,860 youth between the ages of 5 and 19. The average household income is roughly $45,000, and 16% of households are female-headed households (Gaquin & Littman, 1999; Maryland Department of Planning 2000). Figure 2 shows the distribution of public middle and high schools across the county.

Using blocks as the unit of analysis provides the level of detail on places needed to capture variation in the independent variables hypothesized to be related to crime, and at the same time is a unit of analysis for which neighborhood characteristic data are available from the U.S. Census Bureau.

The rate of reported violent crimes per 1,000 population is the dependent variable. Violent crime includes homicide, rape, robbery, aggravated assault and simple assault. Crimes are incident-based data for all person offenses recorded by the Prince George's County Police Department from 1997 through 2000. For stability purposes, to obtain school year 1998–99 crime counts, the violent crime data are aggregated using a three-year period (school years 1997–98, 1998–99, 1999–2000).

To create a population base, this research uses the average number of block residents (109), based on 2000 U.S. Census data, added to the number of individuals who reside in a block. The estimated population was then multiplied by three to be used as a denominator for the three-year aggregate of crime. This method takes into account the large number of instances when crime occurs on blocks where there is foot traffic but where no individuals reside, and greatly reduces the skewness of the dependent variable.[1]

To account for varying victimization risk by time of day, the number of victimizations is divided into periods of the day corresponding to school schedules and curfews. To reduce the skewness of the data caused by the large number of blocks with no crimes, the variables were transformed using their natural logs.[2] After logging, interpretation of regression coefficients can be expressed as the effect of a unit change in an independent variable on the proportional change in the dependent variable (Cohen & Cohen, 1983).

Independent variables are grouped into five variable clusters or constructs: (1) physical place risk, (2) guardianship-routine activities, (3) guardianship-social

Figure 2. Distribution of high schools and middle schools across research site

disorganization, (4) motivated offenders, and (5) adjacent area violent crime rates. *Physical place risk* includes three measures: a dummy measure to reflect whether blocks have schools; a dummy measure to reflect whether blocks have places where youth are likely to populate or "hang out" (including malls, recreation centers, movie theaters and video arcades); and an index that counts all liquor-license establishments (including restaurants), gas stations and mini-markets. The data were collected using public records from the county, PhoneDisc 2000 listings and a listing of all liquor licenses as of December 2000 provided by Prince George's County Board of License Commissioners.

Within the *guardianship-routine activity* variable cluster are measures that represent the number of residents who would likely be capable guardians and a proxy for the number of youth passing through areas. The variables include the percentage of one-person households, housing density, the percentage of renter households that are overcrowded and the number of youth using school bus stops. All data other than bus stop data were obtained from the 2000 U.S. Census. The number of youth using school bus stops was calculated from a database provided by the Prince George's County Public School System that included bus stop address data and number of youth.

Variables within the *guardianship-social disorganization* cluster are racial/ethnic heterogeneity, percent owner-occupied housing, percent black residents, percentage of households headed by females, median value of owner-occupied housing (as a measure of economic well being) and distance to Washington, D.C. With the exception of distance to Washington, D.C., all social disorganization variables were calculated using the 2000 U.S. Census. Because Prince George's County does not have a large city, the variable for distance to Washington, D.C., represents distance to an urban core. The variable is a measure of the distance in miles from every block centroid to the nearest border of Washington, D.C., which was created using the nearest feature extension for ArcView. Following the empirical literature, factor analysis was conducted on the social disorganization variables to reduce the number of variables and overcome potential problems with multicollinearity. Principal components analysis (Dunteman, 1989) with oblique rotation revealed two factors consisting of three variables each. The first factor, named disorganization, consists of the variables percent black, percent female-headed households and distance to Washington, D.C. The second factor, named gentrification, consists of the variables percent owner-occupied housing, racial heterogeneity and median home value.

The fourth variable cluster measures the presence of *motivated offenders,* one of the three requisites for crime under routine activities theory. This variable is operationalized as the number of all arrests of youth ages 17 and under aggregated for calendar years 1997 through 2000. It is reasonable to assume that youth arrested (the majority of arrests are for minor crimes, such as trespassing) are not incarcerated for any long periods and hence, the variable captures potential to offend again, rather than capturing a deterrent element.

The last construct represents the spatial lag of violent crime rates, modeled as an independent variable (Morenoff & Sampson, 1997; Morenoff, Sampson & Raudenbush, 2001; Smith, Frazee & Davison, 2000). According to routine activities theory, motivated offenders will commit crimes along the paths that coincide with their routine activities. Crime should occur among frequently used blocks or streets. These frequently used blocks or streets will be adjacent to or near each other, because offenders diffuse from where they live to where their daily activities take them. Essentially, then, the amount of crime in one area can be expected to affect the amount of crime in adjacent or nearby areas through diffusion-type processes. For a given observation (i), a spatial lag $\sum_j, w_{ij} y_j$ is the weighted average of the crime rates in neighboring locations (note that the spatial lag does not contain y_i). The weights matrix used defines neighboring locations (j) as third-order contiguity of those block neighbors who share a common node (as opposed to a common border). Regressions run using first-order and second-order contiguity did not adequately account for (that is, significantly reduce) autocorrelation. Third-order contiguity signifies that, for every block, neighbors are all the surrounding (first-order) blocks, plus those blocks surrounding the first-order neighbors, plus those blocks surrounding the second-order neighbors. W_{ij} equals 1 if i and j are contiguous. The spatial autoregressive coefficient represents the effect of a unit change, for a given neighborhood i, in the average crime rate of the third-order neighbors on the crime rate of i.

In addition to the variables listed above, the study controls for the size of each block in square miles and for prior victimization as measured by violent crime counts for the calendar years 1992 through 1995.

Analytic Approach

Models are estimated by means of instrumental variables (IV) methods. Ordinary least squares regression (OLS) is not appropriate because the OLS estimator will be biased as well as inconsistent for the parameters of the spatial model. The multidirectional nature of the spatial dependence limits the type of statistical procedures that will lead to consistent estimates. Although IV estimation for models with spatial dependence has not often been applied in criminological studies, it has been shown to be appropriate and functional (Anselin, 1980, 1988, 1984).

IV estimation is based on the principle that a set of instruments, Q, are strongly correlated with the original variables, Z, but asymptotically uncorrelated with the error term. After identification of instruments, the instruments are used to construct a proxy for the endogenous variables, which consists of their predicted values in the regression on the instruments and the exogenous variables. The proxy variable can then be used in least squares regression. There exists little formal guidance in the selection of instruments for the spatially lagged variable. Anselin (1980, 2000) suggests that the use of the spatial lags of the exogenous variables will provide satisfactory results. Following Anselin's suggestion, the lags of all exogenous variables were used as instruments in regression equations. All models are run using SpaceStat software Version 1.91 (Anselin, 1992).

For this study, eight models are tested. The first reference model (Model I) regresses violent crime rates for the full calendar year (January through December 1999) against the variables within the five variable clusters. The second reference model (Model II) is similar to the first with the exception that only violent crimes committed during the school year (September through mid-June) are used as the dependent variable. Succeeding models each represent violent crime only within subsets of time periods within the school year. The subsets are school morning commute (Model III), afternoon school commute (Model IV), school session (Model V), curfew (Model VI), weekday evenings (Model VII) and weekends (Model VIII).

All 7,334 census blocks were included in the analysis. Blocks with no housing units (that is, no residential population; n = 1,136) were kept in the models because elimination of these blocks would exclude an important type of place from the analysis. Theoretically and practically, it remains vital to assess the influence of factors in areas that have foot traffic but where people do not reside. Values for variables such as percent owner-occupied housing and percent female-headed households in unpopulated blocks are zero. A dummy variable is included to capture the effects of the zero-value variables.

Results

The results discussed here mainly focus on the difference across models for the place risk variables. The place risk variables represent the influence of physical site features on violence. Table 1 shows the results for Model I (violence rates aggregated for the entire calendar year regressed on independent variables) and Model II (violence rates aggregated by the school year). Table 2 provides the results of the six time-period models regressing violent crime rates on the independent variables. In Table 1 (the reference models), the regression coefficients and significance levels for each model are similar to each other for

Table 1. Instrumental variables regression models of violent crime rates, full year and school year models, 1999

Variables	**I. Full Year**	**II. School Year**
Place Risk		
Schools	0.672	0.603
Youth Hangouts	0.613*	0.542*
Retail Places	0.600***	0.595***
Guardianship- Routine Activity		
One-Person Households	-0.004*	-0.003
Renter Crowding	0.028**	0.030**
Bus Stop Count	0.005*	0.006**
Housing Density†	-0.013	-0.011
Guardianship-Soc. Disorg.		
Disorganization	0.661***	0.635***
Gentrification	-0.331***	-0.327***
Offender Presence		
Youth Arrests	0.020***	0.022***
Control Variables		
Block Size	1.235***	1.184***
Prior Victimization	0.002***	0.019***
Zero Population	-1.578***	-1.500***
Spatial Lag	0.526***	0.515***
Constant	-0.714***	-0.912***
R^2	0.28	0.28

N=7334; * p<0.05; **p < 0.01; ***p<0.001. †Coefficient has been multiplied by 1,000.

all of the variables, suggesting that the factors influencing violence during the school year are generally equivalent to those for the entire calendar year. In addition, neither Model I nor Model II has a significant coefficient for the presence of schools variable.

Table 2, however, reveals that, as hypothesized, schools are a significant predictor of violence during the periods of the day that coincide with youth attending school (morning school commute, after-school commute and school day). The presence of a school increases a block's risk of violence by more than 200% during the school day and by 120% during the after-school period. During the evening or curfew period, when youth would not necessarily be near school property, the variable has no significant effect on violence. In addition, the influence of the other two place risk variables – youth hangouts and retail liquor establishments – changes depending on the time of day. The youth hangout variable is a significant predictor of violence only in those time periods when youth would have easy opportunity to congregate at these establishments. As shown in Model III, hangouts are not a significant factor predicting violence

Table 2. Instrumental variables regression models of violent crime rates, time period models, 1999

Variables	III. AM Commute	IV. School Session	V. PM Commute	VI. School Night	VII. Curfew	VIII. Weekend
Place Risk						
Schools	1.446***	2.170***	1.212***	0.168	-0.250	-0.638*
Youth Hangouts	0.162	0.507**	0.634**	0.415*	0.388*	0.243
Retail Places	0.089***	0.261***	0.361***	0.395***	0.408***	0.521***
Guardianship- Routine Activity						
One-Person Households	0.001	-0.001	-0.001	-0.002	0.001	-0.000
Renter Crowding	0.005	-0.001	0.006	0.007	0.009	0.018*
Bus Stop Count	0.006*	0.004**	0.011**	---	---	---
Housing Density†	0.003	-0.005	-0.012*	0.005	0.006	0.000
Guardianship- Social Disorganization						
Disorganization	0.192***	0.188***	0.382***	0.354***	0.382***	0.445***
Gentrification	-0.124**	-0.154**	-0.280***	-0.268***	-0.320***	-0.441***
Offender Presence						
Youth Arrests	0.006**	0.009***	0.020***	0.022***	0.012***	0.019***
Control Variables						
Block Size	0.278***	0.407***	0.475***	0.642***	0.744***	0.836***
Prior Victimization	0.025***	0.028***	0.025***	0.026***	0.030***	0.027***
Zero Population	-0.211*	-0.333**	-0.632***	-0.703***	-0.669***	-1.312***
Spatial Lag	0.188***	0.179***	0.293***	0.291***	0.303***	0.309***
Constant	-3.500***	-3.448***	-2.664***	-2.665***	-2.706***	-2.243***
R^2	0.19	0.24	0.24	0.24	0.26	0.25

N=7334; * $p<0.05$; **$p < 0.01$; ***$p<0.001$. †Coefficient has been multiplied by 1,000.

during the morning commute, but hangouts have a significant influence on violence during the school session (Model IV), after school (Model V) and then less so as the night progresses. Youth hangouts do not have a criminogenic effect during the weekend. This fact could be due to the possibility that youth activity is more spread out during the weekends, as youth participate in any number of activities, including activities that take them out of the county.

The influence of retail establishments on violent crime crimes also varies by time of day. Retail establishments have a much smaller effect on violent crime during the morning commute period than they do in other periods when youth are likely to patronize these places. Retail establishments exhibit the largest effect on violence within the model for the weekend period (Model VIII). Table 2 also confirms that school bus stops are vulnerable places at particular times of the day. As the number of students at bus stops increases, so do the rates of violence, and violence at bus stops is more likely during the after-school period (Model V), when adult supervision is likely low. Similarly, the presence of motivated offenders – a strong predictor of violent crime across all time periods – has the largest influence at times when guardianship is low (after school and evening).

The results also reveal, as the study hypotheses suggest, that variables that are not heavily influenced by the routines of attending school show little variation across time periods. Renter crowding is not a significant predictor of violence across the different time-period models, with the exception of the weekend model (Model VIII). One-person households do not influence violence in any of the time-period models, but are slightly significant when violence rates are aggregated across the calendar year (Model I).

Discussion and Conclusion

The ebb and flow of youth going about their daily routines coincide with levels of violence. The study results provide strong support for routine activities theory, as well as support for theoretical models that explicitly account for time of day. Individuals are vulnerable to violence during times when the flow of youth is highly concentrated. At certain times of day there will be places with high concentrations of youth and limited adult supervision. Youth hangouts, schools and busy retail establishments all influence levels of violence, but their impact on violence is mediated by the time of day. For instance, places with youth hangouts generate 40% more crime in the after-school period than during the weekend. In addition, the findings suggest that weekend routine activities may bring about very different types of opportunity for violent crime.

More important, the lack of significant coefficients for some of the place risk variables in the two reference models reveals that failing to accurately model the phenomenon – violent crime – using time of day can mask significant factors that may be attracting, generating, or facilitating violence. Spatial analyses incorporating time of day as a mediating factor can reveal problem areas more accurately. The research findings suggest that adequate supervision of youth during vulnerable *times* and in vulnerable *places* is critical to public safety. Understanding how time intersects with other variables has implications for policing and community problem solving. Supervision at particular times, whether it is increased police surveillance, parental oversight, or other adults acting as capable guardians (for example, recreation center staff, security at malls), can be important in limiting opportunities for violence.

Researchers and police crime analysts currently using advanced techniques for hotspot mapping should consider incorporating aggregate data models of small areas into hotspot analyses. Even without the use of advanced methods, time of day should factor into the analytical techniques used by police departments and researchers examining crime in small places. Spatial analyses allow one to model how multiple features of the physical and social environment converge to influence violence. By incorporating time of day, analyses can shed insight into particular patterns that might otherwise go unnoticed. The increasing ability to access new and improved analytical tools and layers of spatial data will provide unlimited opportunity for crime analysts to make important contributions to policy and practice.

Acknowledgments

The research for this chapter was supported in part by a number of research grants: Grant #2000-JR-VX-0003 from the Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice; Grant #2000 IJ-CX-0012 from the National Institute of Justice, U.S. Department of Justice; and Grant #11-284001-54103-34714 from the National Science Foundation. Points of view or opinions in this document are those of the author, and do not necessarily represent the official position of the U.S. Department of Justice, the National Science Foundation or the Urban Institute. The author would also like to thank the Prince George's County Police Department for sharing their data and their time.

References

Anselin, L. (1980). *Estimation methods for spatial autoregressive structures* (Regional science dissertation and monograph series 8). Cornell University, NY: Field of Regional Science.

Anselin, L. (1984). Specification tests and model selection for aggregate spatial interaction: An empirical comparison. *Journal of Regional Science*, *24*,1-15.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic.

Anselin, L. (1992). *SpaceStat tutorial: A workbook for using SpaceStat in the analysis of spatial data*. Urbana-Champaign, IL: University of Illinois.

Anselin, L., Cohen, J., Cook, D., Gorr, W., & Tita, G. (2000). Spatial analysis of crime. In D. Duffee (Ed.), *Measurement and analysis of crime and justice: Criminal justice 2000, vol. 4* (pp. 213-262). Washington, DC: US Dept. of Justice, National Institute of Justice.

Block, R., & Block, C. R. (2000). The Bronx and Chicago: Street robbery in the environs of rapid transit stations. In V. Goldsmith, P. G. McGuire, J. H. Mollenkopf & T. A. Ross (Eds.), *Analyzing crime patterns: Frontiers of practice*. Thousand Oaks, CA: Sage.

Brantingham, P. J., & Brantingham, P. L. (1982). Mobility, notoriety and crime: A study in crime patterns of urban nodal points. *Journal of Environmental Systems, 11*, 89-99.

Brantingham, P. J., & Brantingham, P. L. (1991). Introduction. In P. J. Brantingham & P. L. Brantingham (Eds.), *Environmental criminology, second ed*. Prospect Heights, IL: Waveland Press.

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

Bursik, R. J. (1999). The informal control of crime through neighborhood networks. *Sociological Focus, 32*, 85-97.

Bursik, R. J., & Grasmick, H. G. (1993). *Neighborhoods and crime: The dimensions of effective community control*. New York: Lexington Books.

Cohen, J., & Cohen, P. (1983). *Applied regression/correlation analysis for behavioral science.* Hilldale, NJ: Lawrence Erlbaum.

Cohen, L., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review, 44*, 588-608.

Dunteman, G. H. (1989). *Principal components analysis* (Sage University Paper Series). Quantitative applications in social sciences. Newbury Park, CA: Sage.

Felson, M. (1987). Routine activities and crime prevention in the developing metropolis. *Criminology, 25*(4), 911-931.

Felson, M. (1994). *Crime and everyday life: Insight and implications for society*. Thousand Oaks, CA: Pine Forge Press.

Gaquin, D.A., & Littman, M. S. (Eds.). (1999). *1999 county and city extra: Annual metro, city and county data book.* Washington, D.C.: Berman Press.

Garofalo, J., Siegel, L., & Laub, J. (1987). School-related victimization among adolescents: An analysis of National Crime Survey (NCS) narratives. *Journal of Quantitative Criminology, 3*(4), 321-338.

Gottfredson, M. R. (1984). *Victims of crime: The dimensions of risk* (Home Office Research Study No. 81). London: Her Majesty's Stationery Office.

LaGrange, T. C. (1999). The impact of neighborhoods, schools and malls on the spatial distribution of property damage. *Journal of Research in Crime and Delinquency, 36*, 393-422.

Lynch, J. P. (1987). Routine activity and victimization at work. *Journal of Quantitative Criminology, 3*, 283-300.

Maryland Department of Planning. (2000). *Demographic and socio-economic outlook for Prince George's County*. Maryland Department of Planning, Planning Data Services. Retrieved April 21, 2001, from *http://www.mdp.state.md.us/MSDC/County/prin02.pdf*

Massey, J. L., Krohn, M. D., & Bonati, L. M. (1989). Property crime and the routine activities of individuals. *Journal of Research in Crime and Delinquency, 26*, 378-400.

McManus, R. (2001). *Juvenile and young adult firearm use in South Carolina*. South Carolina Department of Public Safety. Retrieved September 5, 2001, *http://www.scdps.org/ojp/Final%20-%209697.pdf*

Morenoff, J. D., & Sampson, R .J. (1997). Violent crime and the spatial dynamics of neighborhood transition: Chicago, 1970-1990. *Social Forces, 76*, 31-64.

Morenoff, J. D., Sampson, R. J., & Raudenbush, S.W. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology, 39*, 517-60.

Park, R. (1926). The urban community as a spatial and a moral order. In E. Burgess (Ed.), *The urban community* (pp. 3-18). Chicago: University of Chicago Press.

Rennison, C. M., & Rand, M. R. (2003). *Criminal victimization, 2002* (Report No. NCJ 199994). Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Roncek, D.W. (2000). Schools and crime. In V. Goldsmith, P.G. McGuire, J. H. Mollenkopf & T.A. Ross (Eds.), *Analyzing crime patterns: Frontiers of practice.* Thousand Oaks, CA: Sage Publications.

Roundtree, P., & Land, K. (1996). Burglary victimization, perceptions of crime risk, and routine activities: A multilevel analysis across Seattle neighborhoods and census tracts. *Journal of Research in Crime and Delinquency, 33*, 147-180.

Sampson, R. J., & Wooldrege, J. (1987). Linking the micro- and macro-level dimensions of lifestyle-routine activity and opportunity models of predatory victimization. *Journal of Quantitative Criminology, 3*, 371-93.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multi-level study of collective efficacy. *Science. 277*, 918-24.

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency in urban areas*. Chicago: University of Chicago Press.

Smith, W. R., Frazee, S. G., & Davison, E. (2000). Furthering the integration of routine activity and social disorganization theories: Small units of analysis and the study of street robbery as a diffusion process. *Criminology, 38,* 489-521.

Snyder, H. N. (1999). *Juvenile offenders and victims: 1999 national report.* Washington, D.C.: National Center for Juvenile Justice.

Taylor, R. B. (1998). Crime in small places: What we know, what we can do about it. In *Crime and place: Plenary papers of the 1997 Conference on Criminal Justice Research and Evaluation,* Washington, D.C.

Wiebe, D.J., & Meeker, J.W. (1998). *Hourly trends of gang crime incidents, 1995-1998.* Irvine, CA: University of California, Irvine.

Endnotes

[1] Most routine activities studies have operationalized the rate of violence as the number of offenses divided by number of residents or physical structure, such as street lengths or housing units. Attempts to create population estimates were made using assumptions about the flow of people at different times of day, but insufficient data exist for creating estimates at a low level of aggregation (that is, the block level).

[2] The traditional R^2 measures of fit are not appropriate when using an instrumental variables approach. The R^2 reported in the tables is the ratio of the variance of the predicted values over the variance of the observed values for the dependent variables. See Anselin (1992) for more information.

Chapter XVIII

Measuring Crime in and around Public Housing Using GIS

Harold R. Holzman,
U.S. Department of Housing and Urban Development, USA

Robert A. Hyatt, Austin Peay State University, USA

Tarl Roger Kudrick, Independent Contractor, USA

Abstract

For many people, the phrase "public housing" conjures up images of serious violent crime. However, the neighborhood surrounding public housing may be a greater factor in crime than the housing itself. Because most police departments do not routinely keep statistics on small parcels of land like public housing developments or neighborhoods, measuring the incidence of crime in public housing has proved difficult. Consequently, there is little hard evidence with respect to whether public housing is more or less crime-ridden than the neighborhoods that surround it. This chapter explores the application of geographic information systems (GIS) technology in measuring reported crime levels in and around public housing developments. GIS technology was used to extract crime counts from police *data bases of reported incidents for (1) public housing developments and (2) the surrounding neighborhoods. Rates of reported Part I crimes in public housing developments are compared with those in the surrounding neighborhoods and in the respective municipal jurisdictions. Odds ratios are used to compare the risk of victimization in public housing with that in the respective neighborhood and municipal catchment zones. The GIS-based analysis of reported crime in and around public housing communities reveals that risk of falling victim to aggravated assault in public housing communities is much higher than in the surrounding neighborhoods or in the parent jurisdictions as whole. Conversely, risk of property crimes such as burglary, larceny and car theft appears to be much lower. These crime patterns are discussed in the context of routine activity theory.*

Introduction

Today, over two million people reside in public housing in the United States. Public housing in America has long been recognized as the housing of last resort for the poorest of our fellow citizens (President's Commission on Housing, 1982). Much has been written in both scholarly journals and the popular press about the problem of crime in public housing, but relatively little substantive information is available with respect to crime rates in public housing developments. The absence of such data has thwarted efforts to rigorously evaluate crime prevention initiatives in public housing developments. This chapter describes the methodology and findings of a three-city study of the application of geographic information systems (GIS) technology to measure reported crime levels in and around public housing developments, using counts of reported crimes extracted from official police databases.

Public Housing: Uncharted Islands in Stormy Seas

Since the late 1970s, a body of research on crime in public housing has begun to take shape, funded in large part by the federal government through the U.S. Department of Housing and Urban Development (HUD) and the National Institute of Justice (NIJ). Collectively, these studies present a picture of deeply troubled communities, beset by disorder and violent crime (Newman, 1972;

Newman & Franck, 1980; Roncek, 1981; Farley, 1982; Weisel, 1990; Keyes, 1992; Dunworth & Saiger, 1994; Skogan & Annan, 1994; Popkin, Olson, Lurigio, Gwiasda & Carter, 1995; Kamber, Mollenkopf, Ross & Swartz, 1999; Fagan & Davies, 2000; Holzman, Hyatt & Dempster, 2001). Most of this criminological research has been done on public housing developments in America's largest, older, core cities which have tended to struggle with crime and disorder for many decades, that is, long before the advent of public housing in the 1930's (Holzman, 1996; Holzman & Piper, 1998).

Furthermore, despite the criminological research that documents crime as a problem in public housing, the paucity of systematic efforts to generate crime rates for public housing has left open the question of whether public housing or the disordered, densely populated neighborhoods that surround many public housing developments are responsible for the crime-ridden image of public housing enclaves. The need for such information assumes added importance when one realizes that some 60% of the nation's supply of public housing units is located in our largest and oldest urban centers (Holzman, Kudrick & Voytek, 1996).

GIS as a Research Methodology Uniquely Suited to Measuring Crime in Public Housing

The task of measuring crime in public housing has proved a challenge. Traditionally, local police departments have tended to collect crime statistics for relatively large portions of their respective jurisdictions such as precincts or districts. Little crime data were collected for smaller areas like neighborhoods. In nearly all municipal and county jurisdictions, except in the some dozen localities where the PHA had its own police (for example, New York City, Chicago, Los Angeles), public housing has tended to be virtually ignored in the record keeping associated with official crime statistics.

In the mid-1990s, HUD's Office of Policy Development and Research (PD&R) revived its involvement in studies of crime in public housing at the suggestion of HUD's Office of Public and Indian Housing (PIH), provider of subsidies to the majority of PHAs in the U.S. In the 1970s and early 1980s, PD&R had formally supported research on this topic, but had done little here for almost a decade (Brill and Associates, 1975, 1976, 1977a, 1977b, 1977c; U.S. Dept. of HUD, 1985). HUD's renewed involvement in public housing crime studies was marked by a

PD&R/PIH partnership that produced a 1994 national survey of public housing residents that focused on residents' problems with crime and their attitudes toward crime prevention. This survey yielded findings that highlighted both longstanding problems and potentially innovative solutions and served as an impetus for further research (Holzman et al., 1996). Central to PD&R's new "crime in public housing" research agenda was the question of how best to measure crime in public housing. Public housing crime measurement was recognized as a crucial component of HUD's effort to evaluate its Public Housing Drug Elimination Program (PHDEP) which, since the late 1980s, had annually dispensed hundreds of millions of dollars in crime prevention grants to PHAs.

The long term objective of PD&R's crime research agenda was to assemble a methodological tool kit for crime measurement in public housing. The first effort was to test the feasibility of using victimization survey methodology in public housing (Holzman & Piper, 1998). This effort was undertaken with technical assistance from the U.S. Department of Justice's Bureau of Justice Statistics. While the research suggested that the approach was feasible and resulted in the publication of a manual for conducting victimization surveys in public housing (Piper et al., 1997), this methodology was deemed too expensive for use by cash-strapped PHAs.

Next, PD&R staff sought to examine the feasibility of using GIS to measure crime in public housing. Most of the data presented here is derived from a pioneering 1998 HUD study that probed GIS as a technique for measuring crime in and around public housing (Hyatt & Holzman, 1999). A GIS-based approach appeared unusually well-suited to the task at hand since it allowed for counting reported crimes in small, arbitrarily defined parcels like public housing developments. Furthermore, at the time, it was noted that NIJ-sponsored support for crime mapping was reaching an ever increasing number of police departments (PDs) in metropolitan areas. An NIJ study indicated that roughly one of every eight PDs was using GIS mapping technology (Mamalian & LaVigne, 1999). It was reasoned that PHAs in jurisdictions where the local PDs were involved in GIS-based crime mapping might be able to partner with the local PDs to map crime. Such a partnership would make GIS technology available for PHAs' crime control programs.

Study Design and Research Methods: Applying GIS to the Measure of Crime in and around Public Housing

Site Selection

As stated above, the basic objective of the 1998 study was to examine the feasibility of using GIS to measure crime in public housing. The actual research was carried out by the Center for Geosciences at Research Triangle International, Inc. (RTI) under contract to HUD. A purposive sampling design was used to select three research sites. Five basic site selection criteria were used: (1) candidate jurisdictions, that is, a county or a city, were required to have a police department (PD) that was known to be an experienced user of GIS who routinely geocoded reported crime; (2) the local PD had to serve a jurisdiction with a substantial volume of serious crime so that sufficient data would be available for crime-specific analysis of serious offenses, that is, Part I crimes as defined in the Federal Bureau of Investigation's (FBI) Uniform Crime Reporting Program (UCR); (3) the local PHA's properties had to include multifamily apartment complexes so that there would be clearly defined public housing "developments" for crime mapping; (4) the PHA had to be interested in obtaining crime maps and agree to provide the local PD with maps/site plans of its developments so that these areas could be identified and marked for analysis on police maps; and lastly (5) when independently queried about the nature of their working relationship, both the PHA and PD both had to characterize that relationship as a good one. The last criterion was deemed necessary because, if the public housing crime mapping enterprise proved (1) feasible for both partners and (2) was judged by PHA to be of practical benefit with respect to crime prevention, it was important that the two agencies be open to continuing the data sharing arrangement.

Obtaining Crime Data from the Police

As part of the research plan, the three (3) PDs in the 1998 study were offered reimbursement for the costs involved in data sharing and providing the associated documentation and support. Each PD contracted with RTI to provide these services under a Memorandum of Understanding (MOU) that, in essence, made these agencies subcontractors to RTI in the research process. The MOU described the responsibilities of both parties and worked well. The PDs and the PHAs in the study were offered anonymity in connection with their participation, which they accepted.

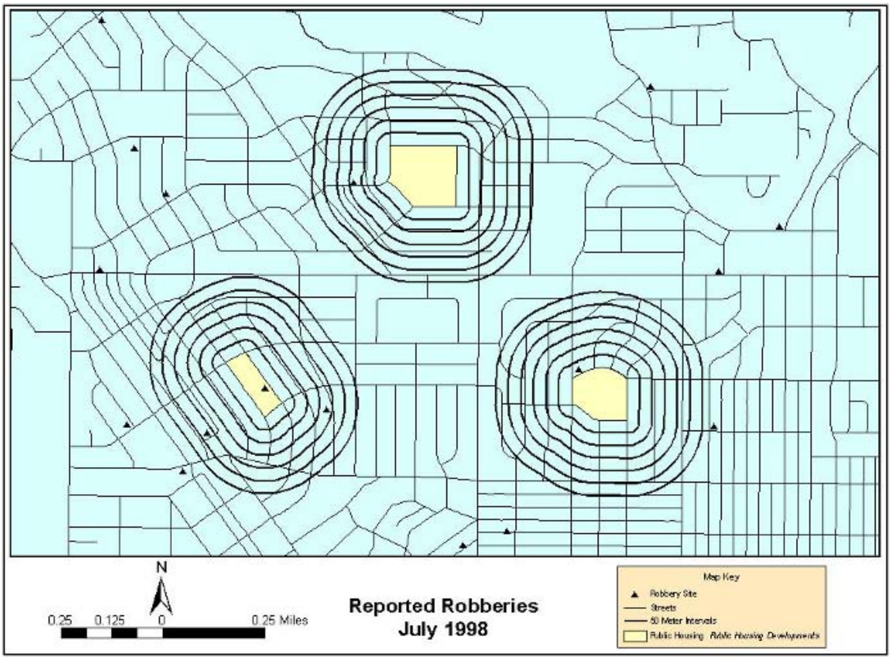
Key Elements of Public Housing Crime Mapping Process

All of the actual crime mapping was done by RTI's GIS specialists with crime data provided by the PDs. The PDs agreed to provide crime data for six contiguous months so offense patterns in each PHA could be tracked over time. At the outset, RTI staff created a map layer for each of the three jurisdictions, depicting its respective public housing developments. These newly created PHA map layers were shared with the participating PDs. Only one of the participating jurisdictions was able to provide RTI with an accurate, up-to-date *street center file,* that is, a computerized map displaying all streets and thoroughfares. These computerized street maps were purchased from private vendors for the other two jurisdictions. Only one of the participating PDs provided geocoded crime data files (that is, files with incident addresses linked to map locations) that were deemed sufficiently accurate for use in the study. RTI staff address-matched and geocoded the crime data for the other two jurisdictions.

After consultation with PHA and PD managers at each of the three sites, it was decided to map crime in all of the respective jurisdictions' public housing developments and also in a 300-meter buffer surrounding each development. The 1998 GIS study was a pioneering effort – such research had not been done previously – so the 300-meter size of the buffer was arbitrary. Crimes were counted in concentric 50-meter buffers within the 300-meter buffer in an effort to detect the presence of any naturally occurring patterns (see Figure 1). No such patterns were observed, so crime counts were collected for both the PHA developments and for the entire 300-meter buffers. In a few instances, the 300-meter buffers around different developments intersected. In such cases, the developments were treated as one property and the 300-meter buffers were redrawn to encompass the whole property.

Research was focused on public the housing developments, and not the three PHAs' individual single family homes and multi-family homes. These free-standing buildings (as opposed to apartment complexes) are known as "scattered sites." This decision was made so findings could be more easily compared to the existing body of research on crime in public housing. Beginning with Oscar Newman's (1972) seminal work on *defensible space,* such research has focused almost exclusively on large, innercity apartment complexes (for example, HUD, 1985; Fagan et al., 1998; Popkin et al., 1999). The domination of these multifamily housing complexes in PHA housing stock in older, major urban centers, where crime is an ongoing problem, highlights the measurement issues that are central to the present study and begs the previously-stated question, "Are levels of crime in public housing developments higher than those in the neighborhoods that surround them?"

Figure 1. Sample map of crime locations



Part I Crimes

For several reasons, Part I Crimes were used as the offense data for analysis. First, these offenses are generally recognized as including the most serious types of violent and property crimes. Second, the overwhelming majority of U.S. police agencies participate in the FBI's uniform crime report data collection program, and use the same definitions for each of the Part I crimes when generating crime statistics. Part I offense data represent a standardized depiction of crime and allow for valid comparisons across the highly diverse universe of law enforcement agencies in the United States.

The "Cities" Studied

Population data for public housing were provided by the PHAs in each city. Population data for the parent jurisdictions were obtained from the Census Bureau's biennial published estimates. Buffer populations were derived using

1990 Census data which was the best data available at the time (1998). With respect to the use of 1990 Census data, PHA and PD personnel concurred that the urban neighborhoods under study were exceptionally stable, having changed little in the preceding decade.

The three respective cities will be referred to as “City A,” “City B” and “City C,” and will be described only in the most general terms. Each of the PHAs in the study had more than 3,000 public housing units and thus was well within the largest 15% of the public housing universe. The populations of the parent jurisdictions each numbered in the hundreds of thousands. The total populations of the public housing developments in each respective city ranged from approximately 7,000 persons (Cities B and C) to over 25,000 (City A). All developments (that is, family, elderly and mixed) in each of the three cities were included in the study. African-American residents outnumbered all other racial/ethnic groups in all three PHAs, accounting for some 70% to 90% of the tenants in each set of PHA developments. The racially segregated nature of public housing developments as well as of innercity neighborhoods has been well documented (Georing, Kamely & Richardson, 1994; Massey & Denton, 1993; Massey & Kanaiaupuni, 1993).

Measuring Crime in Public Housing as a Single Police Precinct

Many offenses, including serious violent crimes, are relatively rare events that often need to be aggregated across geographic areas such as police precincts to have substantive meaning. A single city block in a residential neighborhood or a single public housing development may be the site of only one robbery or aggravated assault in a year. Thus, in designing the present study, the data from all developments in each respective city were aggregated into “public housing precincts.” Crime rates in these precincts were then compared with those of the parent jurisdiction as a whole, and also with those from an aggregation of the surrounding neighborhoods.

The average population per development was highest in City C at approximately 650 persons. The average number of residents in each of City A’s developments was slightly over 500, and City B’s was the smallest (240 persons). Across the three sites, the populations of individual developments ranged from less than 40 residents to nearly 3,000; both the upper and the lower bounds were found in City A.

Table 1. Number of reported Part I crimes by offense for PHA developments, 300-meter buffer zones and entire cities

	City A			City B			City C		
	PHAs	300M	City	PHAs	300M	City	PHAs	300M	City
Homicide	9	30	149	1	5	26	0	0	3
Forcible Rape	10	48	222	8	27	171	5	13	49
Robbery	123	739	3,839	28	128	1,066	26	136	648
Aggravated Assault	228	778	3,770	145	269	2,554	65	146	766
Burglary	161	946	6,408	96	348	4,633	19	191	1,339
Larceny	342	3,210	7,823	102	1,236	13,591	29	154	1,085
Vehicle Theft	70	375	3,750	30	215	1,994	42	280	2,081

Findings

Overview

Crime data were examined with respect to three types of geographic entities: (1) public housing developments, (2) 300-meter buffer zones from the outer boundary of public housing developments, and (3) the parent jurisdiction as whole (See Table 1). In each city, the crime counts and population counts for all public housing developments were aggregated to produce crime statistics for the previously described "public housing precinct." Data for buffer neighborhoods were also aggregated, and crime rates were calculated for each parent jurisdiction. The Part I Crimes data were furnished to the research team by the PD in each city every month. Each monthly "download" included the data from the preceding months so the research at hand would reflect the most up-to-date information in police files (some crimes are re-classified as investigations continue). The 1998 study's databases were the same ones used by the PDs for both their own official statistics and for reports submitted to the Uniform Crime Reports.

Table 2. Crime rate odds ratios for PHA developments versus cities versus 300-meter buffer zones

	PHAs vs. Cities			PHAs vs. 300M			300M vs Cities		
	City A	City B	City C	City A	City B	City C	City A	City B	City C
Homicide	*	*	*	*	*	*	*	*	*
Forcible Rape	*	*	*	*	*	*	*	*	*
Aggravated Assault	1.58	3.28	3.80	1.24	1.79	1.82	1.29	1.83	2.09
Robbery	0.80	1.51	1.80	0.69	0.7	0.78	1.15	2.08	2.30
Burglary	0.67	1.19	0.64	0.71	0.92	0.41	0.91	1.29	1.56
Larceny	0.50	0.43	1.20	0.45	0..27	0.77	1.19	1.57	1.52
Vehicle Theft	0.49	0.87	0.91	0.79	0.41	0.61	0.62	2.10	1.47

** Note: Too few cases for meaningful data analysis*

Measuring and Comparing Risk: Crime Rates and Odds Ratios

Comparisons between the crime rates in the "public housing precincts" and the other two geographic entities are the central focus of analysis. Rather than use the actual rates per thousand population for the six-month period under study, odds ratios are used. As a tool for comparing for crime rates, odds ratios indicate proportional differences in risk level in one area versus another. Like baseball scores (for example, "8 to 3"), crime rates (for example, 4.33 robberies per thousand population) have limited meaning without a frame of reference.

The odds ratio of public housing crime to crime in the parent jurisdiction were calculated using crime rates in public housing as the numerator over the denominator of crime rates in the parent jurisdiction. If the robbery rate in the numerator were twice as large as that of the denominator, then the odds ratio would equal two and the risk of robbery in public housing would be twice that of the parent jurisdiction. Table 2 contains odds ratio comparisons for Part I crimes in each respective city. For example, the risk of having one's car stolen from a public housing parking lot in City A is half that of parking elsewhere in the parent jurisdiction. The reader should note that the differences between crime rates were compared to one another through the use of odds ratios.

Using two-sample difference of proportions tests, the rates for specific offenses in each geographic areas (for example, aggravated assaults rates in public housing developments versus the parent jurisdiction) were tested for statistically significant differences. The crime rates themselves were calculated by placing the number of a specific type of crime in a given zone such a public housing developments (numerator) over the population in that same geographic zone (denominator), that is, public housing developments. For example, in City C, the number of reported aggravated assaults in public housing was 65. If the number of persons living in City C's public housing developments were 5000, then the crime rate for aggravated assault in the PHA would be the proportion 65/5000 or .013. Hypothetically (remember that the research sites were given anonymity), if the population of City C were 205,000 and one used the actual number of aggravated assaults for City C in Table 1, that is 766, then the rate would be 766/205,000 or .004. Before these two rates (.013 versus .004) were compared in an odds ratio as .013/.004 (odds ratio = 3.25), the two rates would be tested in a two-sample difference of proportions tests to see if there was a statistically significant difference between them. This procedure was followed for all the comparisons presented in Table 2. All crime rate comparisons were subjected to this statistical testing and all such tests were statistically significant at the .05 level, with most significant at the .01 level. Given the relatively large number of crimes involved (see Table 1) and the concomitantly large populations noted earlier, attainment of statistical significance was not surprising. More important from a substantive perspective, the reader will note that many of the odds ratios themselves reflect genuinely sizable differences in the crime rates being compared.

Crime in Public Housing Developments Versus Crime in the Parent Jurisdiction

In general, the risk of falling victim to a violent crime is higher in public housing developments than for their parent jurisdiction as whole (see Table 2). However, since so few homicides and forcible rapes were reported in public housing, it was decided to exclude these offenses from public housing crime rate portion of the analysis, lest too much be inferred from the small "cell sizes." For example, City C's public housing developments experienced no homicides during the six months under study, and City B's public housing had only one. If reported crime data for homicide and forcible rape had been collected for a longer period of time, larger samples might have been available for analysis and inclusion in the study.

Aggravated assault rates were higher in all three cities' public housing developments than in their parent jurisdictions. Two out of three cities reported rates that were over three times as high, yielding odds ratios of 3.3 (City B) and 3.8 (City C).

The pattern for reported robberies in public housing developments in comparison with their parent jurisdiction differed somewhat from that of aggravated assault. The public housing development rates predominated in two of three cities, displaying ratios of 1.5 and 1.8. In City A, robberies in public housing were less prevalent (odds ratio = 0.8).

Clearly, the image of public housing developments as violent, dangerous places finds support in this study. Aggravated assaults were markedly more prevalent. Although not as high, the greater risk of robbery is also troubling since public housing developments are almost exclusively residential in character. The absence in public housing developments of such "commercial" targets as liquor stores, bars and convenience stores does not seem to deter robbers. Instead, they prey on people.

In contrast to crimes of violence, the risk of Part I property crimes (that is, burglary, larceny and vehicle theft) tended to be lower in public housing developments compared to that in the parent jurisdiction. Auto theft was lower across all three cities, while burglary and larceny were lower in two out of the three research sites. In City B, the odds of falling victim to burglary were about 20% higher in public housing than in the parent jurisdiction, and in City C, the risk of falling victim to larceny were some 20% higher in public housing. However, these slightly elevated risks pale before the substantially higher risk for violent victimization reported above.

In all three cities, aggravated assault rates were higher in public housing developments than in the 300 meter buffers (see Table 2, columns 4-6). However, the higher risk was not nearly of the magnitude evidenced in the earlier comparison of the developments versus the parent jurisdictions. In all three cities, risk of robbery in public housing developments was lower than in the surrounding 300 meter buffer areas. Again, it is important to note here that the 300 meter buffers contained commercial targets, while land use in the developments was almost exclusively devoted to housing stock. Thus, from the perspective of robbers, the neighborhoods surrounding public housing developments could be termed "target rich" environments compared to residential enclaves at their center. The risk of robbery was also higher in the surrounding neighborhoods than in the parent jurisdictions. This last finding could be interpreted as evidence that the neighborhoods surrounding public housing developments themselves have serious crime problems.

With respect to property offenses, the risk across public housing developments versus the surrounding buffers was consistently lower in contrast to the largely higher risk of aggravated assault. The data suggest that there is a lot more stealing (as well as robbery) occurring in the zones surrounding the developments. Comparatively speaking, the risk of being "ripped off" is higher in the buffers, while there is a greater chance of falling victim to interpersonal violence in public housing. To some degree, the presence of such targets as shops and businesses in the buffers, but not in the developments, probably accounts for some of the higher risk of property crime . For example, retail stores in the buffer neighborhoods are targets for shoplifters and burglars while customers' cars are vulnerable to theft and "smash and grab" larcenies.

Columns 7-9 of Table 2 show the risk of criminal victimization in the buffer areas compared to the respective parent jurisdictions. Clearly, one is confronted with a more dangerous landscape in the buffer areas. The odds of falling victim to a violent crime are substantially higher here than in the city as whole. One fares little better with property crimes. Only in the City A does one face a substantially better chance of escaping vehicle theft in the buffers compared with the risk in the jurisdiction as whole. Otherwise, the buffer areas present a rather bleak picture for all types of crimes. These spaces might reasonably be termed "bad neighborhoods."

Lessons Learned when Working with Police Crime Data

The 1998 study offers some valuable lessons for criminologists who might wish to conduct GIS-based research with PDs who crime map. Perhaps the most crucial lessons involve one's "mind set" in approaching such a collaboration. The police collect, process and map crime data for their own purposes, not researchers'. PD data files, geocoded or not, may not be as complete or as "clean" as one might wish. Therefore, researchers should request as much documentation as is available. This information (for example, coding protocols) will greatly assist the interpretation and editing of crime data files. But, even when one is in receipt of reasonably well prepared "code books," it is almost inevitable that one will need to confer with the PD's crime analysts for technical support.

When seeking such technical assistance, be patient with police managers and the crime analysts in their employ. Build extra time in the research schedule to wait for technical support and data "downloads," since fulfilling research requirements generally not among the PD's top priorities. Similarly, one needs to recognize that GIS analysts in PDs tend to perform a small, but labor intensive, group of tasks. Such analysts are not likely to have the time, breath of knowledge, or wide-ranging experience to serve as GIS experts.

Interpreting the Findings through the Prism of Routine Activity Theory

The present study's findings indicate that, while public housing developments are unusually violent places, the risk of falling victim to property crime is relatively low. Routine activity theory offers some insights into such seemingly contradictory crime patterns. In his book *Crime & Everyday Life*, Marcus Felson (1998) states the three basic elements of routine activity theory's paradigm for the commission of a predatory crime: a likely offender, a suitable target; and the absence of a capable guardian against the offense (p. 53). In discussing this fundamental tripartite paradigm, Eck (1995) notes that the number of targets and the presence of guardians can explain crime levels (p. 784).

Having found low rates of property crime in public housing developments, Dunworth and Saiger (1994) hypothesize that such offenses are underreported. Also, since public housing households are among the poorest Americans with median incomes of some $7,547 per annum (HUD, 1998), one could argue that attractive targets are few, (that is, there is not much to steal). However, in the case of public housing, routine activity theory offers a highly plausible, but largely overlooked, hypothesis for the low rates of property crime: "capable guardians" are numerous in and around public housing.

Socioeconomic data suggests that the majority of public housing's adult residents do not hold jobs and therefore are available to serve as guardians. In 1997, just a year prior to this study, only a third (36%) of income of public housing households was derived from salaries and wages, with nearly half (48%) of all income coming from public assistance, and the remainder from social security and pensions (HUD, 1998). Roughly three-quarters of all public housing households are female-headed (Holzman, 1996; Holzman & Piper, 1998), and about half (46%) of the 1997 households contained children. In that same year, the elderly (that is, 62 years old or above) accounted for 30% of all public housing households. Overall then, a relatively substantial proportion of adult public housing residents are likely to be homebound either because of unemployment, infirmity, retirement or because they are caretakers of children. Hence, there is no shortage of potential guardians in public housing developments to discourage property crime.

The same condition hypothesized to account for the low rate of public housing property crime – householders being home during the day – may also contribute to the high rate of aggravated assault in public housing developments. A substantial proportion of the "capability guardians" of property in this study (that is, unemployed, African-American female heads of households) share characteristics of groups that are at relatively high risk for aggravated assault victimization. In its 1998 report from the National Crime Survey, *Violence by Intimates*, the Bureau of Justice Statistics (BJS) reports that young adult African-American females living in low income, urban households have much higher rates of violent victimization by husbands and boyfriends than women in the general population. Furthermore, three out of four incidences of violence by intimates against women occur in or near the victim's homes (BJS, 1998: 11).

The “capable guardian” who, by her very presence, protects her home and the surrounding areas against thieves, is therefore vulnerable to physical attack by persons known to her. Here, the availability of “stay-at-home” guardians translates into the availability of targets. By admitting a friend, relative, or acquaintance into the privacy of her dwelling, she places herself at risk for criminal violence. The protective surveillance that prevented property crime no longer exists (Gabor, 1990; Felson, 1998; Renzetti, 2000; Holzman et al., 2001).

Conclusion

The present study was originally designed to explore the applicability of GIS for measuring levels of crime in and around public housing. The authors believe that GIS proved to be a useful tool for gauging crime levels in these rather special communities and in the neighborhoods that surround them. Perhaps the most unique aspect of the findings involves the comparisons made between risk levels in the developments and the adjacent areas. No prior study has sought the PHA-wide collection and analysis of a broad range of Part I crimes data for both public housing developments and the 300-meter buffers surrounding them. The odds ratios for aggravated assault in the developments versus those for the buffers showed public housing to be more physically dangerous than the adjacent neighborhoods. Similarly, these neighborhoods themselves were found to be more dangerous than their parent jurisdiction as whole. Clearly, the public housing developments in this study can not be characterized as islands of calm in otherwise rough neighborhoods.

Without the use of GIS as an analytical tool, it would have been quite difficult to measure crime in and around public housing. The applicability of GIS for gauging crime levels in non-traditional statistical catchment zones (such as public housing) renders the technique ideal for evaluating crime local control initiatives as well as for the ongoing observation of crime patterns.

Acknowledgments

The material in this chapter represents the views of the authors and not necessarily those of Research Triangle Institute International nor the U.S Department of Housing and Urban Development.

References

Brill and Associates (1975). *Victimization, fear of crime, and altered behavior: A profile of the crime problem in four housing projects in Boston*. Washington, DC: U.S. Department of Housing and Urban Development.

Brill and Associates (1976). *Victimization, fear of crime, and altered behavior: A profile of the crime problem in William Nickerson Gardens, Los Angeles, California*. Washington, DC: U.S. Department of Housing and Urban Development.

Brill and Associates (1977a). *Victimization, fear of crime, and altered behavior: A profile of the crime problem in Capper Dwellings, Washington, DC* Washington, DC: U.S. Department of Housing and Urban Development.

Brill and Associates (1977b). *Victimization, fear of crime, and altered behavior: A profile of the crime problem in Murphy Homes, Baltimore, Maryland*. Washington, DC: U.S. Department of Housing and Urban Development.

Brill and Associates (1977c). *Victimization, fear of crime, and altered behavior: A profile of the crime problem in Scott/Carver Homes, Dade County, Florida*. Washington, DC: U.S. Department of Housing and Urban Development.

Bureau of Justice Statistics (2000). *Criminal victimization in the United States, 1998*. Washington, DC: Bureau of Justice Statistics.

Craven, D. (1997). *Sex differences in violent victimization, 1994*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.

Dunworth, T., & Saiger, A. (1994). *Drugs and crime in public housing: A three-city analysis*. Washington, DC: National Institute of Justice.

Eck, J. E. (1995). Examining routine activity theory: A review of two books. *Justice Quarterly, 12*, 783-796.

Fagan, J., & Davies, G. (2000).Crime in public housing: Two-way diffusion effects. In V.Goldsmith, P. McGuire, J. Mollenkopf & T. Ross (Eds.), *Analyzing crime patterns: Frontiers of practice* (pp. 121-135). Thousand Oaks, CA: Sage.

Fagan, J., Dumanovsky, T., Thompson, J. P., & Davies, G. (1998) Crime in public housing: Conceptual and research issues. *National Institute of Justice Journal, 36*, 1-8.

Farley, J. E. (1982). Has public housing gotten a bum rap? The incidence of crime in St. Louis public housing developments. *Environment and Behavior, 14*(4), 443-477.

Felson, M. (1998). *Crime & everyday life.* Thousand Oaks, CA: Pine Forge Press.

Gabor, T. (1990).Crime displacement and situational prevention: Toward the development of some principles. *Canadian Journal of Criminology, 32*, 41-74.

Goering, J., Kamely, A., & Richardson, T. (1994). *The location and racial composition of public housing in the United States.* Washington, DC: U.S. Department of Housing and Urban Development.

Gwiasda, V., Taluc, N., & Popkin, S. J. (1997). Data collection in dangerous neighborhoods: Lessons from a survey of public housing residents in Chicago. *The Evaluation Review*, *21*(1), 813-8.

Holzman, H. R. (1996). Criminological research on public housing: Toward a better understanding of people, places and spaces. *Crime & Delinquency*, *42*(3), 361-378.

Holzman, H. R., & Piper, L. (1998). Measuring crime in public housing: Methodological issues and research strategies. *Journal of Quantitative Criminology*, *14*(4), 361-378.

Holzman, H. R., Hyatt, R. A., & Dempster, J. (2001). Patterns of aggravated assault in public housing: Mapping the nexus of offense, place, gender and race. *Violence Against Women: An International and Interdisciplinary Journal, 7*(6), 662-684.

Holzman, H. R., Kudrick, T. R., & Voytek, K. P. (1996). Revisiting the relationship between crime and architectural design: An analysis of data from HUD's 1994 survey of public housing residents. *Cityscape: A Journal of Policy Development and Research*, *2*(1), 107-125.

Hyatt, R. A., & Holzman, H. R. (1999). *Guidebook for measuring crime in public housing with geographic information systems.* Washington, DC: U.S. Department of Housing and Urban Development.

Kamber, T., Mollenkopf, J., Ross, T., & Swartz, C. (1999). *Crime and assisted housing: A spatial analysis.* New York: Center for Urban Research, CUNY.

Keyes, L. C. (1992). *Strategies and saints: Fighting drugs in subsidized housing*. Washington, DC: The Urban Institute.

Mamalian, C. A., & LaVigne., N. C. (1999). *The use of computerized crime mapping by law enforcement.* Washington, DC: National Institute of Justice.

Massey, D. S., & Denton, N. (1993). *American apartheid: Segregation and the making of the underclass.* Boston: Harvard University Press.

Massey, D. S., & Kanaiaupuni, S. M. (1993). Public housing and the concentration of poverty. *Social Science Quarterly, 74*, 109-22.

Newman, O. (1972). *Defensible space: Crime prevention through urban design*. New York: Macmillan.

Newman, O., & Franck, K. A. (1980) *Factors influencing crime and instability in urban housing developments.* Washington, DC: National Institute of Justice.

Piper, L., Lucas, R., Shirey, J., & Rohe, W. (1997*). How to conduct victimizations survey: A workbook.* Washington, DC: U.S. Department of Housing and Urban Development.

Popkin, S. J., Gwiasda, V. E., Rosenbaum, D. P., Amendolia, J. M., Johnson, W. M., & Olson, L. A. (1999). Combating crime in public housing: A qualitative and quantitative longitudinal analysis of the Chicago Housing Authority. *Justice Quarterly, 16*(3), 520-557.

Popkin, S. J., Olson, L. M., Lurigio, A. J., Gwiasda, V. E., & Carter, R. G. (1995). Sweeping out drugs and crime: Residents' views of the Chicago Housing Authority's public housing drug elimination program. *Crime and Delinquency, 41*(1), 73-99.

President's Commission on Housing. (1982). The report of the president's commission on housing, Washington, DC: White House.

Renzetti, C. M., & Maier, S. L. (1999). *Private crime in public housing: Fear of crime and violent victimization among women public housing residents* (Paper). Presented at the annual meetings of the Academy of Criminal Justice Sciences, March, Orlando, Florida.

Roncek, D. W., Bell, R., & Francik, J. M. A. (1981). Housing projects and crime. *Social Problems, 29*, 151-166.

Skogan, W. G., & Annan, S. O. (1994). Drugs in public housing: Toward an effective police response. In D. MacKenzie & C. Uchida (Eds.), *Drugs and crime: Evaluating public police initiatives* (pp. 129-148). Thousand Oaks, CA: Sage.

U.S. Department of Housing and Urban Development. (1985). *Final report of the evaluation of the urban initiatives anti-crime demonstration*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

U.S. Department of Housing and Urban Development. (1998). *Characteristics of households in public and assisted housing* (Fact sheet). Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Weisel, D. L. (1990). *Tackling drug problems in public housing: A guide for police*. Washington, DC: Police Executive Research Forum.

About the Authors

Fahui Wang is an associate professor in the Department of Geography at Northern Illinois University (USA). His research interests include applications of quantitative methods and GIS in crime, health care and cancer analyses. His research has been funded by the National Institute of Justice, U.S. Department of Health and Human Services and U.S. Department of Housing and Urban Development. He is the author of more than 30 refereed articles and the book, *Quantitative Methods and Applications in GIS*, forthcoming in 2005 by the CRC Press.

* * *

Donald P. Albert is an assistant professor in the Department of Geography and Geology at Sam Houston State University in Huntsville, Texas (USA). He received his PhD from the University of North Carolina at Chapel Hill in 1996. Dr. Albert is co-editor of *Spatial Analysis, GIS, and Remote Sensing Applications in the Health Sciences* (2000) and *GIS in Law Enforcement: Implementation Issues and Case Studies* (2003), both published by Taylor and Francis. He has contributed numerous articles, chapters and reviews on applications of GIS to health services research, geographic education and geography of crime.

Jochen Albrecht is an assistant professor in the Department of Geography at the University of Maryland in College Park (USA). His research focuses on the formalization of spatial process models and the development of measures of "change." His two main application areas are landscape ecology and crime analysis. The latter is exemplified by five years of research on Los Angeles hate crime.

Timothy M. Bray is an assistant professor of crime and justice studies at the University of Texas at Dallas (USA). His research interests span both the criminology and criminal justice arenas. His recent work has focused on the effects of socioeconomic disadvantage, residential instability and racial isolation on violent crime in urban America. He has also worked with large criminal history data sets to study recidivism patterns among sex offenders and non-sex offenders. His methodological interests include the application of multilevel modeling techniques and procedures to measure and control temporal and spatial autocorrelation in data. He has employed spatial analysis and HLM techniques to explore juvenile certification decisions, and is currently exploring the separation of race and poverty effects on violent crime using multilevel models.

Dongmei Chen received her PhD in geography from the joint doctoral program of San Diego State University and the University of California, Santa Barbara (2001). She is currently an assistant professor in the Department of Geography at Queen's University, Kingston, Canada. Her research area is in the application of remote sensing and GIS for environmental monitoring and management. Her recent research has been published in *Geocarto International*, *Photogrammatric Engineering and Remote Sensing*, *International Journal of Remote Sensing*, and *Progress in Physical Geography*.

Kevin M. Curtin is an assistant professor in the geographic information science program at the University of Texas at Dallas (USA). He has been involved in the research of geographic information science for more than 10 years. He specializes in transportation GIS and location science. He has developed innovative data models for GIS and has participated in the development of new GIS functionality. He has formulated a family of facilities location problems that optimize maximal distance for multiple types of facilities, and has produced optimal and heuristic solution procedures for these problems. He has taught extensively at both the graduate and undergraduate university levels in GIS, planning, transportation and location science.

John Eck is a professor in the division of criminal justice at University of Cincinnati (USA). His research interests include policing, crime prevention and environmental criminology.

Gregory A. Frost has worked in the law enforcement field for more than 20 years in operational and managerial positions. His professional experience includes serving as a commissioned security police officer in the United States Air Force, research and planning positions with the Florida Department of Law

Enforcement, and he currently serves as the executive services director for the Tallahassee Police Department in Tallahassee, Florida. Mr. Frost holds a BS in psychology with a minor in criminal justice from Colorado State University and is a graduate of the Florida Criminal Justice Executive Institute's Senior Leadership Program. He is co-inventor of the CrimeTrax concept.

Richard Z. Gore spent most of his career as a geo-scientist with a strong interest in spatial problems and GIS. He has a secondary interest in geography, and established the GIS laboratory at the University of Massachusetts-Lowell (USA) that serves the GIS needs of several departments, including criminal justice. He received his BS from the City College of New York in geology, an MS from Boston College in geophysics, and a PhD from Boston University in geology. He is currently working on a cooperative project with the West Midlands Police, UK, exploring the possibilities for expanding the spatial analysis portion of their FLINTS (Force Linked Intelligence System) police intelligence software.

Wilpen L. Gorr is a professor of public policy and management information systems at the H. John Heinz III School of Public Policy and Management at Carnegie Mellon University (USA) and is past editor of the *International Journal of Forecasting*. His research on forecasting, geographic information systems and crime analysis has been supported by a series of grants from the U.S. National Institute of Justice and the Sloan Foundation. His publications have appeared in *International Journal of Forecasting*, *Management Science*, *Geographical Analysis*, *Journal of Geographical Systems*, *Criminology*, *Journal of Criminal Justice* and other journals.

Kenneth V. Griffiths is a serving detective officer in the West Midlands Police, UK. His 25 years with the force have mainly been as an operational officer engaged on the investigation of major incidents, including rape, murder and bomb scene management. He has been a consultant/trainer on the FLINTS project since its inception in 1998 and has influenced the application methods and design. The aims of the project are to link the major policing systems of the UK and provide a national facility with search, analytical and mapping capabilities.

Elizabeth Groff is an analyst at the Institute for Law and Justice in Alexandria, Virginia (USA). She has 10 years of experience working with geographic information systems (GIS). She has spent the last seven years of her career applying spatial analysis techniques to criminal justice issues. She initiated the use of GIS in the Charlotte-Mecklenburg Police Department during the mid-

1990s and gained national level experience while working at the National Institute of Justice's Crime Mapping Research Center, where she promoted the use of analytic mapping in criminal justice agencies. Her current research interests include: offender and victim travel distances, multilevel modeling of individual and community-level processes to predict macro-level crime patterns, and the adoption of technology by police departments. Elizabeth is a PhD candidate in geography at the University of Maryland, College Park (USA). She has a BS and an MA in geography from the University of North Carolina at Charlotte.

Karen Hayslett-McCall is an assistant professor of crime and justice studies, sociology and political economy in the School of Social Sciences at the University of Texas at Dallas (USA). She has research interests in the neighborhood effects on crime, policing, criminological theory, the refinement of contextual definitions through the application of neighborhood infrastructure data and the integration of multilevel models and spatial statistics with GIS. She has presented research at the annual meetings of the American Society of Criminology, Academy of Criminal Justice Sciences and the National Institute of Justice's Crime Mapping Research Conference. She earned a PhD in crime, law and justice from the Pennsylvania State University, where she also served as a GIS research analyst in the Population Research Institute.

Harold R. Holzman received a PhD in criminal justice and criminology from the University of Maryland in 1979. He is presently employed in the Office of Policy Development and Research, U.S. Department of Housing and Urban Development. His professional involvement in criminal justice research began in 1971 when he joined the statistics division of the U.S. Department of Justice's (USDOJ) Law Enforcement Assistance Administration. Over the years, he worked in the U.S. Census Bureau's Crime Analysis Division and in USDOJ's National Institute of Justice, and also spent nearly two years as associate director of a research center at the Rutgers School of Criminal Justice. He has published articles in a number of scholarly journals, including *Crime and Delinquency*, *Journal of Criminal Law and Criminology*, *Journal of Quantitative Criminology* and *Violence Against Women*.

Robert A. Hyatt received a PhD in geography from the University of Georgia. He has more than 25 years of experience in geographic analysis including geographic information systems, mapping, remote sensing and crime pattern analysis.

John V. Kaiser Jr. is a San Diego State University (SDSU) Foundation (USA) researcher and coordinator of the NASA-sponsored Affiliated Research Center (ARC) program. He holds a BS in geology and master's degrees in both geography and business from SDSU. He coordinates the day-to-day remote sensing and GIS research activities on ARC projects. He also conducts training programs for SDSU faculty and staff to learn GIS and desktop mapping using ArcGIS, ArcView and ArcView extensions. His recent research has been published in *Photogrammetric Engineering and Remote Sensing*, *Geocarto International*, *Geospatial Solutions* and NASA Affiliate Research Center Technical publications.

Tarl Roger Kudrick has a PhD in personality psychology from Howard University and more than 12 years of experience in data analysis. He began an independent consulting practice in 2001, and joined the U.S. Office of Personnel Management as a research psychologist in 2004.

Ian Laverty is president and co-founder of Environmental Criminology Research, Inc. (ECRI) (Canada). ECRI developed the *Rigel* geographic profiling system currently used by all the professional geographic profilers in the world.

Mark R. Leipnik is an associate professor in the Department of Geography and Geology at Sam Houston State University (SHSU) in Huntsville, Texas (USA). He received his PhD in geography from the University of California, Santa Barbara (1995) and has been a faculty member and GIS lab director at SHSU since 1996. He is co-editor of *GIS in Law Enforcement: Implementation Issues and Case Studies*, published by Taylor and Francis of London (2003), and has authored more than 50 journal articles, chapters, magazine articles and papers related to GIS applications in law enforcement, homeland security, marketing and environmental planning and management. He has served as a consultant and instructor for the Police Foundation, the Law Enforcement Management Institute of Texas, The Police Research Institute, and federal, state and local law enforcement agencies, as well as the U.S. Department of the Interior and the U.S. Army Corps of Engineers.

Jun Liang is a lecturer and GIS technician in the Department of Geography at the University of North Carolina at Chapel Hill (USA). His current research interests include spatial modeling, crime simulation, GPS and remote sensing.

Lin Liu is an associate professor of geography at the University of Cincinnati (USA). His research interests include crime analysis, GIS and spatial data analysis. He has received research grants from National Science Foundation, Ohio EPA and other funding agencies. He served as the president of the International Association of Chinese Professionals in Geographic Information Science (CPGIS).

Yongmei Lu is an assistant professor in the Department of Geography at Texas State University – San Marcos (USA) and the Texas Center for Geographic Information Science. She holds a PhD from University at Buffalo (2001) in geographical information science and urban and regional studies. Her major research interests are the integration of GIS with spatial analysis and modeling on socio-economic and environmental applications, theoretical issues of geographic information science and spatial data modeling.

Lorraine Mazerolle is an associate professor in the Department of Criminology and Criminal Justice at Griffith University (Australia). She received her PhD from Rutgers University, New Jersey (1993). Before coming to Griffith University in 2000, she was an associate professor at the University of Cincinnati. During her 10-year tenure in the U.S., she received numerous federal grants on topics such as problem-oriented policing, police technologies (for example, crime mapping, gunshot detection systems, 3-1-1 call systems), civil remedies, street-level drug enforcement and policing public housing sites. She is the author of *Policing Places with Drug Problems* (Sage Publications) and a co-editor (with Jan Roehl) of *Civil Remedies and Crime Prevention* (Criminal Justice Press). She has written many scholarly articles on policing, drug law enforcement, displacement of crime and crime prevention. In 2003, she was admitted as a member to the Academy of Experimental Criminologists.

Tess McCarthy is a PhD candidate in the School of Policing Studies at the New South Wales Police College, Charles Sturt University, Australia. With a background in environmental science and geography, her current research is focused on the modus operandi (MO) of serial burglary offenders, and the utility of MO-based models for case linkage and offender detection. Having spent an extraordinary amount of time scrubbing burglary addresses and correcting street-based geo-datasets, she is well aware of the trials and tribulations of geocoding crime data.

J. Thomas McEwen is a principal at the Institute for Law and Justice in Alexandria, Virginia (USA). He has more than 30 years experience in criminal justice and is a nationally recognized expert in quantitative analysis applied to law enforcement. He provides expert consultation in evaluation techniques, crime mapping, staffing and resource allocation, management information systems, community policing and computer crime. He has directed numerous national evaluations and assessments for the National Institute of Justice, U.S. Department of Justice, on such topics as police acquisition of technology, differential police response, community policing, police use of less than lethal weapons, gang crime prosecution and criminal justice modeling. He has a master's degree in statistics and a PhD in mathematics.

Shannon A. McKay is a PhD candidate in the Department of Urban Studies and Planning at the Massachusetts Institute of Technology (MIT) (USA). Her research interests include land use policy, growth management issues, sprawl simulation models, public finance and fostering city and suburban cooperation. She has an M.S. in public policy and management from the H. John Heinz III School of Public Policy and Management at Carnegie Mellon University in Pittsburgh, Pennsylvania. Prior to attending Carnegie Mellon, she worked for three years at the Urban Institute in Washington, DC, on education policy issues.

Brad Moore is a detective sergeant of the Ontario Provincial Police, Canada. He is a working geographic profiler and has lectured and presented extensively on the topic. He was among the first group of geographic profilers trained by the VPD Geographic Profiling Section in 1998. He frequently consults to other police forces around the world on serious criminal cases.

Alan T. Murray is an associate professor in the Department of Geography at Ohio State University (USA). He obtained a BS in mathematical sciences, an MA in statistics and applied probability and a PhD in geography, all from the University of California, Santa Barbara. His research and teaching interests are in geographic information science, urban/regional planning and development, and transportation. His recent research has focused on issues of sustainability and urban patterns of crime. He has published on a range of technical and application oriented topics in journals such as *British Journal of Criminology*, *International Journal of Geographic Information Sciences, Geographical Analysis, Journal of Geographical Systems, Journal of Urban Planning and Development*, *European Journal of Operational Research, Transportation Research* and *Urban Studies,* among others.

Van M. O'Brien is currently a GIS project manager for The Sidwell Company in St. Charles, Illinois (USA). He is responsible for coordinating and implementing geographic information systems for local government departments. He also provides software training for clients to utilize and maintain the system. Prior to his employment at Sidwell, he held an internship at the DeKalb County Economic Development Corporation. He is a candidate for a master of science degree at Northern Illinois University.

James Pingel has been deputy director for the state of Wisconsin's sentencing commission since January 2004. Prior to that, he spent nearly three years directing Milwaukee, Wisconsin's National Institute of Justice pilot project, COMPASS (Community Mapping, Planning and Analysis for Safety Strategies). He has also served as finance director at the Milwaukee Police Department. He has a master's degree in public administration from the LaFollette Institute of Public Policy at the University of Wisconsin.

Fang Qiu is an assistant professor in the geographic information sciences program at the University of Texas at Dallas (USA). He earned his PhD in geography from the University of South Carolina in 2000. His current interests include geographic information science, remote sensing, digital image processing, spatial analysis and modeling, spatial database management, neural network, fuzzy logic, data mining, GIS application software development, Web-based mapping and distributed GIS.

Jerry Ratcliffe is an associate professor of criminal justice at Temple University in Philadelphia (USA). Prior to that, he was a senior research analyst with the Australian Institute of Criminology and a lecturer in policing (intelligence) with Charles Sturt University. In the dim and distant past, he spent a decade as a police officer in London, before it dawned on him that academics did not have to do night shifts. His research interests include environmental criminology, crime mapping and police intelligence practice.

Peter A. Rogerson is a professor of geography at the University at Buffalo, State University of New York (USA). He is also a member of the National Center for Geographic Information and Analysis. His interests are in the areas of spatial analysis and spatial statistics. His recent research has focused on the development of methods for the detection of geographic clustering in spatial data, and on the development of surveillance methods for detecting change in geographic patterns over time. He is also the author of *Statistical Methods for Geography*, from Sage Publications.

Sunghoon Roh is a detective inspector with the Police of Pusan, Republic of Korea, and a PhD candidate in the College of Criminal Justice at Sam Houston State University in Huntsville, Texas (USA). His research interests include crime prevention, community-oriented policing, serial homicide and the application of geospatial technologies to law enforcement.

Caterina Gouvis Roman is a senior research associate in the Justice Policy Center at The Urban Institute (USA). In addition to spatial analysis of crime and violence, her research interests include community crime prevention, criminal justice theory, at-risk youth programming, offender reentry and program evaluation. Her book, *Schools, Neighborhoods, and Violence: A Study in the Sociology of Place*, will be published by Lexington Books in 2004. She received her PhD in sociology and justice, law and society from American University.

D. Kim Rossmo is the leading expert on geographic profiling. He first developed the technique and coined the term "geographic profiling" during his doctoral thesis work at Simon Fraser University in the early 1990s. Subsequently, as a detective inspector of the Vancouver Police Department (VPD), he founded the VPD Geographic Profiling Section in 1995, which worked on cases and trained other geographic profilers for several police forces. He has recently completed a term as director of research for the Police Foundation in Washington, DC, and is currently a professor of criminal justice at Texas State University (USA). He is the author of a textbook and many academic papers on geographic profiling.

Tung-Kai Shyy is a research associate in the Centre for Research into Sustainable Urban and Regional Futures, School of Geography, Planning and Architecture at the University of Queensland, Australia. He obtained a PhD in geography from the University of Oklahoma. His research interests include GIS-based spatial modeling and visualization, Internet GIS applications and remote sensing. He has published articles in *International Journal of Geographical Information Science*, *International Journal of Remote Sensing*, *Australian Planner*, *Australasian Journal of Regional Studies* and a book chapter in *Planning Support Systems in Practice*.

Robert J. Stimson is a professor of geographical sciences and planning and director of the Centre for Research into Sustainable Urban and Regional Futures at the University of Queensland. He has extensive teaching, research and consulting experience in a wide variety of urban-related fields. He has conducted research and teaching in Australia, the U.S. and a number of Asian countries. He is recognized as one of Australia's leading urban researchers, and is author/joint author of more than 30 books or monographs, and more than 200 journal articles, technical papers and conference papers. His most recent book (with

Kevin O'Connor and Maurice Daly) is *Australia's Changing Economic Geography: A Society Dividing,* published by Oxford University Press.

Nikolas J. Tofiluk is an assistant chief constable with responsibility for the intelligence operations and systems of the West Midlands Police (UK), and leader of the FLINTS project since 1999. His police career includes experience over a wide range of operational duties throughout his service with West Midlands for 23 years. He obtained a bachelor's degree in education/economics from Birmingham University and a master's degree in applied criminology from Cambridge University, and also attended the UK police national strategic command course in 1999.

Xuguang Wang is a PhD candidate in the department of geography at the University of Cincinnati (USA). His research interests include GIS applications, spatial statistics and crime analysis.

John R. Weeks is a professor of geography and director of the International Population Center at San Diego State University (USA). He holds a PhD in demography from the University of California, Berkeley. He has written extensively on the interrelationships between remote-sensing, GIS, and social and demographic issues, including crime. He is currently the principal investigator on a three-year NSF grant. His recent research has been published in *Environment and Planning A*, *Geocarto International*, *Population-Espace-Societes*, *International Journal of Geographical Information Science*, *Journal of Immigrant Health*, *Professional Geographer*, *Photogrammatric Engineering and Remote Sensing* and *the Annals of the AAG*. His textbook, *Population: An Introduction to Concepts and Issues,* is now in its ninth edition.

John Western obtained his PhD from Columbia University in the early 1960s. He has more than 30 years of experience in basic and applied social research, with an emphasis on the professions and work, social stratification and social inequality and the criminal justice system. He is the author of *Social Inequality in Australian Society* and a number of other monographs, as well as more than 75 refereed journal articles. He is a joint editor of *Understanding Youth Crime: An Australian Study* (2003).

Index

D

E

M

N

O

P