THE GEOARCHIVE: AN INFORMATION FOUNDATION FOR COMMUNITY POLICING

by

Carolyn Rebecca Block
Statistical Analysis Center, Illinois Criminal Justice Information Authority

Abstract: This chapter presents the main concepts of the GeoArchive as an "information foundation for community policing." Based on the Illinois Criminal Justice Information Authority's experience in developing this innovative idea, many agencies are using a GeoArchive to identify problems and develop strategies for crime prevention and intervention at the neighborhood level. The theory of the GeoArchive and practical suggestions and rules of thumb for developing a GeoArchive are illustrated with examples of how the GeoArchive is being used in current crime prevention strategies.

INTRODUCTION

A GeoArchive is a database of community and law enforcement data, organized for use in crime analysis, investigation and community problem solving. A type of geographic information system (GIS), a GeoArchive contains address-level data from both law enforcement and community sources, linked to computer mapping capability and
organized so that the data can be updated, maintained, mapped, analyzed and used by those who are developing and implementing strategies of crime reduction in the community. When combined with a problem-oriented community policing program, a GeoArchive can become an information foundation for community policing.

Crime maps are nothing new. Pin maps have graced walls behind police chiefs' desks since pins were invented. Neither is the "high-tech" version of these pin maps — computer-aided crime mapping — particularly new. What is new is that police districts and community-level organizations now have direct access to and control over computer mapping. This means that those people who have the greatest stake in solving neighborhood problems now have direct access to the information and analysis tools they need to identify and develop effective solutions for specific problems facing their community. The potential effects of this innovation are so fundamental to the nature of local decision making and problem solving that it deserves to be called a "technological revolution."

By itself, however, computer mapping technology will not supply the information needed for problem-oriented community policing. For effective problem identification and problem solving, communities need more than the ability to map data. They must be able to turn spatial data into information. This means that they must compile and organize the vast amount of mapped data generated by day-to-day activity in a neighborhood, relate the data to other information, and then summarize those data quickly and objectively as a basis for making decisions. Increasingly, people eager to meet this challenge have been springing up in a variety of sworn, civilian, technical, practical, academic and community settings across the continent and around the world. These innovative "mapping entrepreneurs" are searching for ways to go beyond making pretty maps towards using spatial data as a foundation for community problem solving. In Chicago, for example, the Illinois Criminal Justice Information Authority is working with the Chicago Police Department and Loyola University Chicago to develop an "Information Foundation for Community Policing," which couples a "GeoArchive" database with spatial analysis and statistical tools (such as the Spatial and Temporal Analysis of Crime [STAC] software package).

Technology by itself is not enough to support problem-oriented community policing, crime analysis, tactical decisions, or the development of investigation strategies and intervention programs. Mapping technology is useful only to the degree that it is coupled with useful data and analysis tools to make sense of those data (Figure 1).
Both community problem solving and law enforcement tactical decisions require a well-organized body of local-level community and law enforcement information (a GeoArchive). In order to develop a strategy for preventing homicide, for example, we must know more than just the patterns and trends of homicide. We also need to know about all the events (lethal and non-lethal, criminal and noncriminal) that may escalate to or presage a homicide. Therefore, a database intended to support such problem solving should link law enforcement and community data at the address level. This we have called a GeoArchive.

In addition to a GeoArchive database, we need tools to help us make sense of all of this information. When it is initially acquired, computer mapping technology is a tremendous boon to crime analysts and local problem solvers; no longer is it necessary to draw maps by hand. All too soon, however, the amount of mapped information becomes too much to handle, many alternative summaries of it are possible, and quick decisions become more and more out of reach. In such situations, an efficient and objective summary of reality provided by statistics and spatial analysis can offer a useful guide to interpretation. Statistics are tools designed to summarize enormous amounts of information and to organize that information to answer specific, practical questions.

But these requirements are difficult to meet. Tools for organizing and analyzing spatial data are still in their infancy, particularly tools that are applicable in practical situations. The GeoArchive and STAC, basic components of the "Early Warning System" project in Chicago Police Area Four and prototypes for the Chicago Alternative Policing Strategy (CAPS) program and the ICAM (Information Collection for Automated Mapping) program, are attempts to develop solutions to these problems. Though we still have much to learn, these pilot projects have taught us a lot about developing and maintaining a GeoArchive, linking it to statistical and spatial analysis tools, and using the two resources together to identify and solve community problems. Using the Chicago experience as an example, this chapter outlines the major things we have learned so far, in the hope that other communities will benefit from our experience.
COMPUTER MAPPING AS A TECHNOLOGICAL REVOLUTION

Only a few years ago, the only mapped information available to most police departments was in the form of cardboard pin maps with colored plastic pins. Computer maps required such expensive equipment and such a high level of expertise that they could be produced only by a central city planning agency outside of the police department or perhaps by a central administrative unit within the police
department. Mapping software and hardware was complex and expensive, and required experts to use. In addition, it was — and still is — tremendously expensive to create the computerized (digitized) street maps that are necessary for mapping. Access to mapping equipment and automated maps was, therefore, beyond the reach of most departments.

Because cities that did have mapping capability usually housed it centrally, often in an agency outside the police department, police access to mapping was indirect, often cumbersome, and usually time consuming. Local or field-level decision makers had to petition a central "data division" to obtain a map meeting their needs, and the response, even if successful, was seldom timely. Furthermore, the digitized maps, mapped data and area boundaries available from a central source are not always the most useful for identifying and solving community safety problems. Maps that are appropriate for central planning might not be appropriate for local law enforcement decisions.

Times have changed. A combination of three recent innovations in computer mapping technology — accessible mapping software, personal computers and work stations that can handle that software and the U.S. Census's TIGER (Topologically Integrated Geographic Encoding and Referencing) [see Guptill, 1988] files — have brought computer mapping capability within the reach of local communities (Stallo, 1995). Though some mapping software is still very expensive, requiring years of training and high-powered hardware, software companies have begun to produce mapping packages that are much cheaper and friendlier, and that need no more than an ordinary personal computer (PC) to run them (Sanford, 1995). Also, the Census Bureau's creation of digitized street maps for every U.S. county was a huge breakthrough for local-level mapping. These "TIGER" street map files, available at low cost from the Census or computer software vendors, eliminate the necessity of digitizing the local street map, an extremely expensive task that had been a formidable obstacle to local mapping.

The advent of accessible, PC-based mapping software and inexpensive automated street maps means that computer mapping capability is now available at the local, district and neighborhood levels. As a result of this technological revolution, the ability to identify and solve problems using spatial information is no longer the exclusive purview of analysts and technical experts in large organizations or in city, state or federal governments. It is also available to people trying to identify and solve problems in their own neighborhood.
Accessible computer mapping generates a need for methods and techniques geared to take advantage of spatial information (Guptill, 1988; Block and Green, 1994). When law enforcement and community agencies begin to shop for computer mapping software, they may encounter a vendor who tries to sell them data. They should realize that they already have a multitude of data, and easy access to even more. A police department does not need to buy back its own offense and arrest data from a vendor. For other kinds of data, agencies generally will obtain much better information from each other if they establish their own data-sharing relationships instead of going through a vendor. Instead of more data, what agencies really need are tools to map, organize and summarize the data they already have. With these tools, maps can go beyond description to become a foundation for community problem solving.

Data Overload

The last 10 or 20 years have seen a quiet revolution in criminal justice. Even though conventional wisdom points to a dearth of high-quality criminal justice data, both the quality and quantity of data have improved tremendously in recent decades, as has their availability to decision makers (Block, 1989). Sufficient information is now available to allow for the measurement of basic indicators with a degree of precision that was not only unknown but even unanticipated a few years ago. Even though there is still a lot of room for improvement in the degree of detail and specificity of data available, the criminal justice system generates a tremendous amount of information. Much of it, however, is unused. There may be so many pieces of information that it is impossible for the human mind to assimilate them, sort them out, and summarize them before the window of opportunity for an effective decision has passed. Thus, rather than the major problem being a dearth of data, often the problem is just the opposite — data overload.

Just when information technology has begun to be widely used in law enforcement to bring data overload under control (Manning, 1992), computer mapping is generating yet another surge of data. This is occurring for two reasons. First, computer mapping adds quantities of new information — mapped data sets, automated street maps and boundary maps — to the data repertoire. Second, accessible computer mapping has changed the nature of that information with an added dimension — space (see Anselin, 1989). Therefore, mapping generates both a quantitative increase in the amount of data and qualitative changes in the character of data.
Having ready access to spatial information for the first time can be compared to suddenly being granted another sense. Someone without the sense of sight, for example, might be quite capable of perceiving the environment with the other four senses, and if granted the ability to see, would need time to learn how to use this fifth sense and integrate it with the other four. Similarly, now that spatial information is generally available, our first inclination might be to maintain, organize and analyze it in the same way as we have always done. The old ways, however, do not utilize the unique character of the information offered by spatial data. For example, an address-based data set of homicides might support many kinds of spatial analysis; we could ask whether the homicide locations are clustered, whether the homicides tend to be located close to other mapped locations such as taverns or gang territories, or whether they tend to occur on the periphery of a city or in the center. Similarly, with a data set of areas defined within boundaries (such as Census tracts, police districts, gang territories, or crime Hot Spot Areas), we might examine the location of high-crime areas relative to transit stops, or the effect of a Hot Spot Area on crime levels in the area surrounding it. But it is difficult to study issues such as these without a database organized to use spatial information, plus statistical tools to summarize that information.

In addition to the data overload precipitated by the advent of accessible computer mapping, problem-oriented community policing can precipitate its own inundation of data. Though many people have some of the information necessary to identify and solve a community problem, no single individual is likely to have all of it. A tactical officer, a patrol officer, a narcotics officer, a long-time resident and a community worker are all likely to have differing sets of information about patterns of street gang violence, and officers working the night shift may be aware of very different aspects of the neighborhood's problems than officers working the day shift (Block and Green, 1994). Some of this knowledge is spatial; individuals have "cognitive maps" that may differ, even for the same area (Rengert, 1995b; Rengert and Greene, 1994; Mattson and Rengert, 1995). In addition, community information might be forever lost when an especially knowledgeable person moves, retires or is promoted out of the area. In principle, then, a complete problem analysis could require the compilation and evaluation of the body of knowledge representing the experience of all aspects of a community, past and present — an overwhelming task.

Like computer mapping, community policing can produce an enormous increase in the quantity of data as well as a qualitative change in the nature of data. Information necessary for problem-
oriented community policing is often different from information usually collected in law enforcement (Sparrow, 1994:124-126). To identify the problems facing a neighborhood and to describe those problems with enough detail to support effective intervention programs, we must have some way to organize and sift through the vast amount of information available about an area from a multitude of sources, each event anchored by location and time, and to make this information easily and readily available to local problem solvers.

No matter how sophisticated it may be, computer mapping technology is not enough by itself to control and manage data overload. In addition to technology, we need tools that can manage data that is organized in different spatial units and that changes over time, and that can link spatial and other kinds of information, such as individual, incident, location and situation characteristics. We also need spatial analysis and statistical tools that can summarize a vast amount of spatial information for quick and objective decision making. Database tools such as the GeoArchive and statistical tools such as STAC can control and manage data overload, so that law enforcement and community information can become a foundation for the tactical, crime analysis and policy decisions of problem-oriented community policing (see Figure 1, above).

In his analysis of evolving interaction between two Chicago street gangs (the Black Gangster Disciple Nation and the Vicelords) from 1987 to 1992, David Curry (1995) used a GeoArchive and STAC to deal with data overload. In the original map (Figure 2), showing all street gang-related offenses attributed to the two gangs in an area on Chicago's West Side, with locations of four schools, Humboldt Park, and the grid of local streets, there is so much data that it is difficult to perceive any pattern. However, when Curry organized the data to examine specific hypotheses, he began to see a pattern that tells a story. First, he separated violent gang offenses attributed to the two gangs, and separated crimes occurring in the earlier years from those occurring in the later years. Second, using consistently defined boundaries and search parameters, Curry used STAC to identify the densest concentrations (Hot Spot Areas) of gang-related crimes attributed to each gang in the two periods.

Figures 3a and 3b indicate that two changes took place between 1987-1988 and 1991-1992: first, a sharp decline in violent gang activity, and, second, a shift in gang territory. The earlier years saw four dense concentrations of violent Black Gangster Disciple Nation (BGDN) activity, centered in two locations to the northeast and southeast of Humboldt Park, and around and between four area
schools. Violent Vicelord activity was much less densely compressed, and the most concentrated areas centered on the schools (where they overlapped with BGDN Hot Spot Areas), not the park. By 1991-1992, the level of violent gang-related offenses in the area had fallen sharply, and the BGDN offenses were so widely scattered that STAC did not find a Hot Spot Area. In contrast, Vicelord offenses continued to be concentrated around the schools.

Figure 2: Gang Crimes in Study Area around Four Schools and Park

To move from a data overload situation to a problem analysis, Curry (1995) used a GeoArchive of community data (school and park locations) and law enforcement data (gang offenses), as well as STAC Hot Spot Area analysis. Equally vital to his analysis, however, was the hypothesis that spatial patterns and concentrations of gang-
related offenses attributed to the two gangs had changed over time. Such a theoretical perspective, like mapped data, is necessary for effective spatial analysis. However, finding the best theory for a given application can be as problematic as managing and summarizing the data.

**Figure 3a: Density Ellipses for Violent Gang Crimes in Study Area for Two Gangs in 1987-88**

**Theory Overload**

For law enforcement and community information to make sense as a basis for local-level decisions, we need more than "just a pretty map." As John Eck argues elsewhere in this volume and others have argued previously (Roncèk and Maier, 1991; Maltz et al., 1991), the
successful analysis of spatial patterns of crime requires that mapping technology be guided by theory that can link place to crime, can unravel the spatial characteristics of different types of crime, and can provide explanations and suggest prevention strategies for the high vulnerability of some neighborhoods or demographic groups. Nevertheless, it is much easier to assert that computer mapping technology must be linked to an information-organizing framework built on theory than to actually do it. Which theory do we choose? What general theoretical framework(s) or guiding construct(s) apply best to the specific, local problem at hand?

**Figure 3b: Density Ellipses for Violent Gang Crimes in Study Area for Two Gangs in 1991-92**

![Density Ellipses for Violent Gang Crimes in Study Area for Two Gangs in 1991-92](image)
For almost every community safety problem, there are numerous alternative theories that might be used to guide problem identification and intervention strategy development, theories that may suggest different and even completely contradictory interventions. This is equally true, perhaps even more true, when the people involved in problem solving are not aware that they have a theoretical framework. Each individual experiences somewhat different aspects of the same community, and has different ways of understanding community events. Such a perceptual framework is an implicit theory. These diverse but sometimes competing ways of thinking about a community and its problems can be a rich source for innovative problem analysis and solution development (Fisher, 1994). On the other hand, a plethora of theories, whether or not they are formally stated as such, can lead to stagnation or produce only sound and fury. This "theory overload," if not harnessed to the process of community problem solving (Goldstein, 1990), can be a formidable obstacle.

Problem-oriented community policing (Sparrow, 1994; Moore; Goldstein, 1977, 1990) provides a potential solution to theory overload. In community policing, the police department and the community collaborate to set and achieve community safety priorities (Sparrow, 1994). In problem-oriented (or problem solving) policing, the emphasis has evolved from the traditional focus on handling individual incidents as they arise to identifying, analyzing and solving the general problem leading to similar incidents (Goldstein, 1990). Though the concepts of community and problem-oriented policing overlap (Moore and Trojanowicz, 1988), the existence of one does not necessarily imply the existence of the other. Community policing programs vary widely in the degree to which the police department and citizens collaborate to identify and analyze problems and develop solutions for them. If not combined with problem solving, community policing alone is likely to yield only limited and ephemeral benefits (Sparrow, 1994). By the same token, police problem solving may not always involve the community (Ward et al., 1995; Goldstein, 1993; Sparrow, 1994), thus losing access to information and resources available only from people intimately familiar with a specific neighborhood. However, when they do occur together as problem-oriented community policing, the resulting police and community collaboration draws on the knowledge and resources of those who know a community best — people working together to solve its problems. This can be a powerful mechanism for solving theory overload.
Implementation of this solution is not simple, however. Community-centered problem-solving is an art, not a science. It involves many steps, summarized by Spelman and Eck’s (1987) acronym SARA (Scanning, Analysis, Response, Assessment). Each of these elements presents a different organizational and logistical challenge. Of the steps in problem solving, perhaps the biggest challenge is analysis, which requires not only compiling and organizing information (grouping incidents as problems), but also relating this information to alternative theories (analyzing relevant interests in the problem, critiquing the current response, and searching for innovative responses). One of the basic tasks of analysis is to collate and compile disparate theories about a problem and its solutions into a common working definition that can be a basis for a collaborative solution — a strategy for intervention or prevention.

As Sparrow (1994:46) puts it, "...to find broadly acceptable solutions to problems (police) need, first, to find broadly acceptable definitions of the problems." But Sadd and Grinc (1996:14) point out that, "One of the principles guiding community policing is recognition that the police must be guided by the values of the community. Identifying these values may not be easy." Pulling the community’s theoretical perspectives together to identify and define the problem is a foundation for developing solutions. But theory overload makes this difficult, because there may be as many theories about a problem as there are community members, sometimes more.

Thus, we come full circle: as a potential solution to theory overload, the analysis step in problem-oriented community policing requires that the problem solvers identify and address the various theoretical frameworks in the community. They must, in other words, resolve theory overload in order to resolve theory overload. Herman Goldstein and his colleagues have compiled and tested numerous techniques and social mechanisms aimed at overcoming this dilemma (for an overview, see Goldstein, 1993.) As with many tasks, the effort is aided by the proper tools — database management and statistical tools for compiling, storing, summarizing, analyzing and communicating information.

As a tool for overcoming theory overload, a GeoArchive works in several ways. First, it can serve as a community memory bank (Maltz et al., 1991) — a device for storing, linking and sharing enormous amounts of community information from diverse sources. Further, because maps can be so compelling, particularly maps of someone’s own neighborhood, an accessible GeoArchive can increase the amount of interaction among neighborhood players, the clarity of
their communication with one another, and the resulting degree of consensus on problem-solving strategies. In addition, a GeoArchive can actually motivate community problem solvers to share information with each other. Thus, a GeoArchive can provide a springboard for problem solvers to identify and evaluate those alternative theoretical frameworks that should be an integral part of problem analysis and the search for solutions.

Building an Information Foundation for Community Problem Solving

No matter how innovative or revolutionary, no technology or theory by itself is a panacea for solving community problems. The most effective problem analysis and problem solving will not emerge from technology or theory alone, but only when they support each other. We need conceptual resources to utilize technological resources effectively, and we need technological resources to make it easier to compile, evaluate, communicate and utilize conceptual resources. Together, computer mapping technology linked to an information-organizing framework that encompasses both law enforcement and community information can become an information foundation for community problem solving.

Data overload and theory overload present obstacles to linking theory and technology. However, two kinds of tools designed to accommodate the unique aspects of spatial data can integrate and make sense of the enormous amounts of information generated by daily interaction in a neighborhood (see Figure 1). Database management and statistical tools, used in combination, can compile, summarize and communicate spatial and other information. With these tools, one of which is the GeoArchive, we can link technology to theory and form a foundation upon which practical applications can be built.

WHAT IS A GEOARCHIVE?

Spatial data overload calls for tools that can do more than just manage large databases. To turn "spatial data into spatial information," we must be able to assimilate, sort, link and summarize data over several dimensions: individual characteristics, spatial relationships, and trends over time. However, the development of data management and statistical tools for geographic analysis has not kept pace with the technological revolution in computer mapping. Data-
base and statistical tools, particularly spatial analysis tools applicable to practical law enforcement situations, are still in their infancy. In response to this situation, the Illinois Criminal Justice Information Authority, working with the Chicago Police Department, Loyola University Chicago and mapping entrepreneurs around the world, has been compiling a portfolio of database and statistical tools to manage, organize and summarize spatial data as a basis for practical community safety decisions. The GeoArchive is one of these tools.

A GeoArchive is a particular kind of GIS. Like all GISs, a GeoArchive is especially organized for spatial data, and contains a digitized map and data geocoded to be located on that map. It can be seen as a large set of map transparencies that can be overlaid on each other. But a GeoArchive has several characteristics that distinguish it from other GIS databases (see Figure 4). A GeoArchive links address-based local-level data from a variety of law enforcement and community sources, and is organized so that it can be updated, maintained, mapped, analyzed and used by those who are developing and implementing strategies of crime reduction in the local community.

Address-Based, Neighborhood-Level Information

Geographic point (address-based) data and area data are key components of GIS databases. In point data, the spatial unit is a dot on the map, representing a single location such as an offense, an offender's residence or a tavern or abandoned building. In area data, the spatial unit is a two-dimensional area surrounded by an enclosed boundary, such as a zip code, a police district or beat, a Census tract or a gang territory. In contrast to other GISs, a GeoArchive must contain both area and point data sets, with a database of information behind each. In fact, the most important geographic information in a GeoArchive is point (address or pin) data.

In some GIS databases, the area is the smallest spatial unit of analysis. Any point in such a GIS is either only a map data location with no information behind it, or actually represents an area (i.e., a centroid). Both point data and area data have information behind them in a GeoArchive; they are much more than locations on a map. In the Early Warning System for Street Gang Violence project, for example, each crime incident has about 50 variables associated with it (such as offense type, weapon, number of offenders); each Census tract has numerous demographic variables associated with it (such as total population, percent under age 15); each street gang territory
Figure 4: The GeoArchive and Community Problem-Solving
boundary is linked to information about the gang; and so on (Jacob and Block, 1995; Green and Whitaker, 1994; Block and Green, 1994; Bobrowski, 1988).

In a GIS in which area is the smallest spatial unit of analysis, the unit scale is often too large and the boundaries of the units too arbitrarily drawn for the data to be useful for analyzing community problems. Local problems require local address-level information, not address information summarized across areas defined by artificial boundaries. We do not send a squad car to a Census tract to answer a call, but to an address. A GIS that depends too completely on Census or district-level data may be useful for some purposes, such as writing annual reports or long-term forecasting of resource utilization at the district or city level. It will not, however, be a useful tool for daily decisions and community problem solving.

Moreover, address-level information is lost with an area-level GIS. Points can be aggregated to areas, but areas cannot be disaggregated to points. With an address-level database of criminal incidents linked to an area-level database of police districts, it would be easy to count, for example, the number of criminal incidents in every police district. In contrast, with a strictly area-level database of total incidents for each police district, there is no way to determine the exact location of each incident. If forced to aggregate address-level to area-level data, we will lose valuable information.

For example, the Community Areas in Chicago vary widely in their population-based rates of street gang-motivated homicide (Figure 5). The number of homicides occurring between 1987 and 1994 ranged from zero to 63, and the rate per 100,000 population ranged from zero to 21. However, aggregating the data by Community Area obscures the actual pattern of homicides, and may even be misleading (Brantingham and Brantingham, 1981). An address-level map of the same street gang homicides (Figure 6) shows that the overall homicide rate for a Community Area may be a misleading representation of concentrations of homicides (shown as triangles on the map). This happens when arbitrary area boundaries divide a homicide cluster (for example, where Areas 22, 23, and 24 meet), or when a cluster is confined to a small part of a larger area (for example, Areas 8 or 46). Vice Lord activity, for example, is not determined by Community Area boundaries.

Do the densest concentrations of street gang non-lethal violence and drug activity produce a higher risk of street gang homicide? Actually, the risk of street gang-motivated homicide is higher in Hot Spot Areas of street gang-motivated non-lethal violence (narrow-line
ellipses in Figure 6), calculated independently of the homicide locations, than in the densest concentrations of street gang drug offenses (heavy-line Hot Spot Areas). For example, the small Hot Spot Area of non-lethal violence in Community Area 8 had 45.7 homicides per square mile, while the large drug offense Hot Spot Area covering most of Areas 23, 25, 26 and 27 contained only 9.1 per square mile. The greatest risk of street gang homicide, however, occurs when non-lethal violent and drug offense Hot Spot Areas intersect. For example, along the border between Areas 23 and 24, where the large drug Hot Spot Area intersects with a violent Hot Spot Area, there were 18.8 homicides per square mile. The portion of Community Area 61 where a violent and drug Hot Spot Area overlap experienced 37.5 homicides per square mile. Thus, neighborhoods unfortunate enough to have both Hot Spot Areas of gang turf violence and of street gang-related drug offenses tend to have a very high risk of lethal gang violence. This kind of analysis would be impossible without address-level data.

But this is not to say that area-level data should be excluded from a GeoArchive. For many reasons, both are necessary. The most obvious reason is that some information that may be vital to informed decisions is defined or available only at the area level. Population data and street gang territory data, for example, are defined only within the boundaries of areas. By linking Census data to the non-lethal street gang data summarized by the Hot Spot Areas in Figure 6, for example, we can measure the per capita rate of non-lethal violent or drug offenses within each Hot Spot Area. In addition, many agency or community decisions are focused on a specifically defined area, such as a beat or a ward. If the GeoArchive is to support analysis in response to specific local-level questions, it must be organized so that information can be aggregated and presented according to the area units that are appropriate to the particular problem.

In addition, the point and area data sets of a GeoArchive should be related to each other, so that analysis can benefit from both. It is often necessary to relate data from one agency to another, such as police data to court data, taking into account that the two agencies "think" according to different district systems. With geocoded addresses of incidents plus the boundaries of police, court, or other areas, this is easy to do. Police incidents, for example, may be aggregated to court areas, or court events (addresses of probationers, for example) aggregated to police areas. The homicides shown in Figure 6
Figure 5: Gang-Motivated Homicides, Mean Annual Rate Chicago Community Areas, 1987-1994

Source: Chicago Homicide Dataset, Illinois Criminal Justice Information Authority
could be aggregated across Community Areas, police districts, or any set of areas having available mapped boundaries. Therefore, a GeoArchive should be organized so that it is easy to aggregate points across a variety of areas (police districts, probation districts, school
districts, Census block groups, Boys and Girls Club neighborhoods, health center catchment areas, and so on). In this way, spatial analysis can focus specifically on the question at hand and on the intended audience for the results, and the results can be easily used and readily interpreted by those making the decisions.

In general, a GeoArchive should be able to handle community safety problems requiring different kinds of information and different levels of aggregation. Analysis of an immediate threat to neighborhood safety often requires information that is timely and specific to a particular situation, while analysis of long-term trends and patterns requires data covering a long time span across a variety of areas. Similarly, solving crime (investigation) and preventing crime may involve different kinds of analysis and require different information. For example, detailed information about individual offenses, offenders and victims is very important to the investigation of a pattern of serial offenses, while crime pattern and crime analysis decisions may call for more information about areas such as drug markets or gang territories.

**Community and Law Enforcement Data**

The events and environmental situations that surround violence or property crime are not limited to those recorded in official law enforcement statistics. By the same token, community data alone do not provide enough information for community problem solving. To identify community problems, neither criminal justice nor community data are enough by themselves; both are needed (see Figure 4). In addition, the most effective solutions will draw on the resources of both (Sadd and Grinc, 1996:11-12). Though in principle most community policing projects recognize the necessity of community collaboration, there are many difficulties in actually achieving it (for a discussion of some of these, see Sadd and Grinc, 1996). The process of obtaining information from citizens and community agencies and utilizing community resources to solve problems is often *ad hoc* and poorly documented. The *mechanism* of sharing information is seldom addressed. A GeoArchive provides such a mechanism. This section is a quick outline of some of the major law enforcement and community data sets most valuable for a GeoArchive.

Law enforcement information includes a vast amount of point data (location of offenses and arrests, addresses of victims and offenders, citizen calls for service, police response information) and area data (districts, beats, wards). Point data sets, such as crimes known to the police, accumulate at a great rate. By developing and instituting an
efficient system for geocoding, however, mapped data can be available to local decision makers within no more than 24 hours after the event.\textsuperscript{20} To law enforcement data can be added point and area data from other criminal justice agencies, such as corrections or probation, and from organizations affiliated with the criminal justice system, such as drug abuse intervention agencies or organizations working with ex-offenders.

Data sets of streets, bodies of water and major landmarks (called map data sets) are fundamental to mapping. The accuracy of digitized street map data affects not only the display of streets on a map but the geocoding of address data. Erroneous maps are, therefore, a serious threat to accurate decisions. In the U.S., the availability of Census TIGER street maps is one of the main reasons that computer mapping is now accessible at the community level. Like all data sets, however, the TIGER files do contain some errors. Streets may be missing, either because new streets have been added to the area or because the original map was erroneous; street names may be missing, misspelled or inaccurate; and locations important to police work (e.g., under a viaduct, along a park road, at the lakefront) may not be recognized by the map.\textsuperscript{21} This can be rectified, however, by editing the street map file. Therefore, for accurate geocoding and accurate maps, it is vital that the users or managers of a GeoArchive have access to and are capable of modifying the base street map file.

Community point and area data sets that are particularly useful in building a GeoArchive (see Block and Green, 1994, for detail) include: (1) land use data sets containing information on each parcel of land in the city (for example, vacant or not, abandoned or not, residential or commercial, state of repair, specific function such as tavern or convenience store, and so on); (2) public transit data sets (train or bus stops and routes); (3) schools (grammar, high schools, private); (4) community organizations (block clubs, churches, social service agencies); (5) parks and other open areas (with park roads, field houses, lagoons); (6) emergency locations (hospitals, fire houses, police stations); (7) public housing (by type, showing roads and play lots); (8) places holding liquor licenses (by type of establishment and license); (9) public health data (area-level mortality and morbidity rates by specific cause, point data on health problems such as infant mortality or fatal firearm accidents); and, of course, (10) Census data.

In sum, the first and second criteria for a GeoArchive maintain that it should contain a wide variety of data sets, organized to be mapped at various scales and levels of detail and as point, area or line data, depending on the application. As John Eck points out else-
where in this volume, theory should govern the choice of details to place on a map. For example, transit routes may be mapped as point files (the stops) or as line map files (the routes). Land use data sets may be point files, area files, or both; for example, a point file of convenience stores extracted from a land use file of parcels. Schools, public housing, or city parks may be mapped as point or area files, depending on the scale of the map, and routes to and from school can be shown with a line map file. Area boundaries showing each building and the major entrances and exits of a school campus or public housing complex are more informative for community decisions, but it is easier to create a point file than to create boundaries; updating is also easier with a point file. Census data should be available for a wide choice of scales (block, block group, tract, Community Area). Though tracts cover such a large area that it may be cumbersome to relate them to neighborhood problems, detailed information may be available only at the tract level, not at the smaller block level (Green and Whitaker, 1994).

Easy Accessibility to Local Decision Makers on a Timely Basis

The third criterion for a GeoArchive is that it is a local resource for crime analysis and decision making. In contrast to many GIS databases, a GeoArchive should be developed and controlled at the local level (from the bottom up rather than from the top down). Control means that local decision makers can change, update and manipulate GeoArchive data. It is this local, neighborhood control of information that makes the GeoArchive an information foundation for community problem solving. This does not mean that every local GeoArchive should be built and maintained independently of all others, solely by local efforts and resources; that would be nonsensical. There is no reason for each local decision maker to be required to learn the techniques of geocoding or the intricacies of editing street map files. Similarly, there is no reason why citywide mapped data sets, such as abandoned buildings or liquor licenses, should not be shared across all city neighborhoods. On the contrary, local GeoArchives should be supported by data sets and skills provided by centrally located experts. However, these technocrats should serve the needs of the local-level decision makers, not vice versa.

Local-level data accessibility raises two issues. First, it could be argued that central maintenance is necessary to check for and control errors, and to eliminate the discrepancies that would inevitably
occur if every locality were to make independent changes in a data set. If data were corrected and updated at the local level without central coordination, maps in neighboring areas would become increasingly incompatible with one another. On the other hand, those most likely to discover data errors are those people who use the data, in this case, the individuals who work the streets and know them well and the decision makers who analyze the data and map the patterns, not the central data coordinators. One solution to this dilemma is a data clearinghouse, in which local-level users regularly send data corrections and enhancements to a central "holding file," where a holding file manager checks for accuracy and consistency and then makes the enhanced data generally available. With such a clearinghouse system, local decision makers have free rein to create and define data sets that best fit their purposes, but only data approved by the clearinghouse is shared citywide.

A second consideration raised by locally accessible data is data security. This can be a problem not only for law enforcement data, but for community data as well. For example, in the Early Warning System for Street Gang Violence project, an agency's volunteer staff was willing to provide street gang territory information to the GeoArchive, but was concerned that gang members would learn they had done so. It is also important to recognize that officially verified data may not require the same level of security as unverified data. Officially verified data have gone through a review process, are usually standardized in format, have standard codes and an identification number of some kind, and are often considered public information. Whereas unverified investigation data such as lists of suspects, contacts or citizen tips might require high security, officially verified data usually does not.

In the GeoArchive created by the Loyola Community Safety Project for the Rogers Park/Edgewater community (Block et al., 1993), consideration for security was balanced with a high degree of access to the data. The Rogers Park/Edgewater GeoArchive was queried regularly by community groups, beat committees, aldermen and state representatives. Because of problems of confidentiality and concerns about the value of this information for real estate speculators, all requests for information were approved by the project director or technical coordinator. However, this did not inhibit the community from using the GeoArchive or requesting specific additions and expansions. For example, when Rogers Park community groups and the police district became concerned about the effect that taverns and liquor stores were having on crime in the neighborhood, the Commu-
The GeoArchive — 51

Community Safety Project added information on crime in liquor outlets to the GeoArchive and analyzed the relationship between concentrations of liquor stores or taverns and concentrations of crime in these places (Block and Block, 1995). Acting in part on the results of this analysis, the community organizations and the police district launched coordinated projects aimed at reducing levels of violence in the specific places and areas identified as generating the most serious problems.

Though some skeptics might argue that neighborhood-level decision makers such as patrol officers or social agency workers are not capable of using mapped information to identify and solve problems, the ICAM system in Chicago provides empirical evidence to the contrary (Rich, 1996). Developed to support the ambitious CAPS approach to problem-oriented community policing, under which beat officers city-wide focus on problem solving (Rodriguez, 1993; Rich, 1996), the primary goal of ICAM is to bring mapped data to all beat officers in the city quickly — within 24 hours at most. Early, though anecdotal, results indicate that ICAM's accessibility encourages beat officers to analyze and solve problems, and that officers regularly take ICAM maps to community beat meetings (Rich, 1996).

A concomitant goal of CAPS, still in the beginning stages of realization, is to increase the problem solving collaboration between community members and the police, in part by increasing community access to ICAM and other information (Rodriguez, 1993; Baladad, 1996). In some districts, citizens access ICAM information at kiosks and officers regularly sit down with community members around ICAM maps to study a problem. A recent review in the newsletter of a coalition of Chicago neighborhood organizations (Baladad, 1996:11) found that "significant gains have been made in community access to information" under ICAM and CAPS, but noted that access "is still marred by inconsistency" from district to district and called for more citizen use and requests for ICAM information.

A typical example of a community concern is depicted in Figure 7, which was created at the request of a coalition of community organizations concerned about organizing a task force to confront safety problems in some high-risk areas on the Near North Side. Working with the two local police districts, these organizations used the Hot Spot Areas analysis of gang-related offenses to identify and focus on specific problem areas in their neighborhood.

There may be some disagreement whether the increased availability of information through ICAM is driving increased citizen interest in problem solving, whether — as Baladad argues — the increasing
Figure 7: Uptown Hot Spot Areas for Total Gang Crime: 1990-1994

Data Source: Chicago Police Department
Illinois Criminal Justice Information Authority
awareness of the value of mapped information on the part of community groups and citizens is driving greater access to ICAM, or whether the two processes drive each other. Whatever the reason, it seems clear that ICAM is becoming a catalyst for the decentralized use of information to solve problems. According to Manning (1992), centralized control of information has been an obstacle preventing law enforcement from benefiting from innovative technology. He argues that one reason why information technology has not had much effect on policing so far is that its basic assumption "is a belief in the centrality of information" (p.390).

However, overcoming information centrality is not easy (Hasson and Ley, 1994). On the one hand, to avoid data incompatibility across local-level GeoArchives and to make large and widely used data sets (such as street files, geocoded incidents, land use or other community data) easily available at the district or neighborhood level, it is better to maintain many large, important data sets centrally. On the other hand, the best source for data validity-checking and correction is at the user level, which is often the local level. Building a system to meet these conflicting needs requires creative and innovative management. It may be best to develop this system incrementally, with the collaboration of a Working Group of local GeoArchive users across the city who experiment to find the best system to insure that local GeoArchives would really support local decision making. Instead of being predetermined, the needs of GeoArchive users would become apparent as the project progresses. The evaluators of CAPS have found that such flexibility and willingness to experiment and learn from experience has played an important role in its development (Chicago Community Policing Evaluation Consortium, 1994).

THE GEOARCHIVE AND COMMUNITY POLICING: PRACTICAL SUGGESTIONS

In creating and working with the GeoArchive in Chicago, we have discovered a number of helpful hints, suggestions and rules of thumb about combining law enforcement and community data. This paper is too short to treat all of these in detail, but several that may be useful to others are discussed below.
Spatial Linking Can Generate Information that Previously Did Not Exist

A common dilemma in data set management is that two data sets cannot be matched, because they do not contain the same identifier. This obstacle may be overcome by matching the coordinates of geocoded addresses. For example, the Authority needed data on convenience store robbery for a cross-state comparative analysis (Amandus et al., 1996), but the location variable in Chicago Police Department incident data does not include a "convenience store" code. The solution was provided by Richard Block, who matched six months of Chicago Police Department address-based incident data compiled for the Chicago CAPS evaluation (approximately 300,000 incidents) to the addresses of convenience stores, creating a data set of offenses occurring at convenience stores. This data set was then used to identify the robberies and their characteristics. A similar technique was also used to identify offenses occurring at taverns or liquor stores, in independent studies by Block and Block (1995) and by Florence (1995).

In addition, a GeoArchive can be used to add new fields, such as incidents occurring at or near an establishment, to the data set behind a point or area file. For example, the Early Warning System project (Jacob and Block, 1995) added a "number of offenders" variable to the incident file, by counting the number of records in the offender-based file with the same incident identification number, and writing the total as a new field in incident file.

Coordinate matching is thus a very powerful technique. With a GeoArchive, you can create new data in two ways: by linking spatial data sets to create a third data set, and by using one data set to create new fields behind another data set. However, as we discovered when applying coordinate matching to criminal incidents and liquor license addresses (Block and Block, 1995), this valuable tool should not be used blindly. In the first matching attempt, only 60% of the addresses of incidents that the police had recorded as occurring at a tavern or liquor store matched the liquor license addresses in a file from the state Department of Revenue. After investigation, we discovered that very few of these mismatches were due to inaccurate information in either data set. Instead, most of the mismatches were caused by a combination of definition differences and map accuracy problems. When these problems and incompatibilities were corrected, the coordinate matching "hit rate" approached 95%.
A GeoArchive Can Stimulate Cooperation among Agencies and between Agencies and the Community

Information sharing is a necessary foundation for community-based crime reduction programs, and, therefore, necessary for a GeoArchive. By definition, a GeoArchive consists of a compilation of spatial databases originally collected by many different agencies and organizations for many different purposes. This is one of its most valuable attributes. In the mid-1980s, for example, a pilot mapping project in selected Chicago police districts (Maltz et al., 1987) found that information from neighborhood and community groups added to the richness of the spatial database, and allowed officers to identify high-activity areas more accurately.

Though conventional wisdom has it that agencies, and even departments of the same agency, resist sharing data with one other, we have found that maps provide a great incentive for data sharing and communication. For example, Sparrow (1994:121) notes that GISs "constitute an appealing technology around which to form cross-functional teams." This does not happen automatically, however. In our experience, two things will help open and maintain lines of communication. First, the agency requesting data should do everything possible to reduce the provider's cost of data transfer. For example, accept data in the format that is easiest for the provider, avoid requests for any special "selects" or other analysis, and do your own geocoding (but avoid hard-copy data if at all possible). Second, the requesting agency should treat the project as a team effort. This means giving the agency who provided the data copies of the enhanced, geocoded data set and the printed analysis, offering to conduct analysis for them on request, and citing their generosity in publications.

Under these circumstances, most community, city and state agencies or organizations will be more than happy to contribute data to a GeoArchive, especially when they know that it is being developed as an "information foundation" for community problem solving and to reduce levels of violence. In turn, a GeoArchive can increase collaboration and cooperation among community agencies, by providing a storehouse of community information (a community memory bank) that combines and relates information across agencies, and by making it possible to produce maps that can present this information in a visual, readily understandable form. As law enforcement and community groups begin to use GeoArchive maps for decision making, their stake in the quality and availability of that data will increase. As
a result, they will share data more readily and become more interested in maintaining data quality. Thus, though cooperative relationships between the police and local communities are not easy, they are possible (Kennedy, 1993; Kennedy et al., 1996) and a GeoArchive can facilitate their development.

Chicago's Early Warning System for Street Gang Violence project is based on the assumption that information compiled by community and neighborhood organizations, as well as by law enforcement, can be used to develop an "early warning system" of neighborhoods in crisis (see Jacob and Block, 1995; Spergel and Curry, 1990). Since much street gang violence is spatially anchored and occurs as the culmination of escalating incidents of revenge and retaliation, continuing escalation would then be prevented by crisis intervention and dispute mediation, using both internal community influences and external police support. Such a program, which requires the strong support of neighborhood agencies, churches, community groups and the police department, showed success in pilot projects in Chicago's Humboldt Park and in Philadelphia (Spergel, 1984; Spergel et al., 1984; Spergel, 1986). As the Violence Reduction Project, the concept is currently being tested in a Chicago neighborhood plagued by extremely high rates of street gang-related violence (Jacob and Block, 1995), and is being replicated in five cities around the country.

One result of Early Warning System analysis useful for the Violence Reduction Project was to relate the home addresses of the most active street gang offenders to the location of the most serious street gang offenses (Jacob and Block, 1995). In the Little Village area on Chicago's West Side (Figures 8 and 9), the predominant street gang-related activity is turf battles between the Latin Kings and the Two-Six street gangs. The densest concentrations of residential addresses of those Latin Kings and Two-Sixers who were identified by police investigation as offenders in serious gang-related violence (homicide, firearm-aggravated battery or firearm-aggravated assault) are defined by the Hot Spot Areas in Figure 8, with the clusters of Latin King residences (narrow-line ellipses) lying generally to the east and the densest clusters of Two-Sixer residences (wide-line ellipses) generally lying to the west. When this map was shown to Violence Reduction Project street workers, they identified most clusters as locations of a hub of a specific faction of the Two-Six or Latin King street gang. Thus, for turf gangs, the core of gang territory can be identified by finding the densest clusters of addresses where the most active gang members live.
However, for turf street gangs, the hub of gang territory does not necessarily define the center of street gang-related violent offenses. The densest concentrations of serious violence committed by the Latin Kings on Two-Six victims or by Two-Sixers on Latin King victims do not coincide with the residence clusters of the offenders, but rather with the residence clusters of the victims (Figure 9). Latin King violent attacks on Two-Sixers (dashed-line ellipses) tend to coincide with Two-Six hub turf, and Two-Six violent attacks on Latin Kings (wide-line ellipses) tend to coincide with Latin King hub turf. (The Hot Spot Area of Latin King attacks on Two-Sixers around Harrison High School are an exception to this pattern.) This suggests a "marauder" pattern, in which members of rival gangs travel to the hub of their enemy's territory in search of potential victims. The location of the Two-Sixer and Latin King hub turf is known not only to project street workers, but to other Little Village residents, and certainly to the rival gang members. The two rival gangs search out their enemies where they believe they are most likely to be found, often in or near the center of their turf.

This pattern is not seen for the entrepreneurial gangs with territories to the north of Little Village. For these gangs, the densest concentrations of street gang-related activity (serious drug offenses such as manufacture and delivery) tend to coincide with the residential clusters of gang member residences. In contrast to gangs in which violent turf battles are the predominant activity, entrepreneurial gangs tend to commit gang-related offenses close to where they live.

**Diverse Data Sets Present Not Only Benefits but Also Great Technical Difficulties**

The somewhat Panglosian scenario described above, in which agencies cooperate to compile data sets from many sources into a single, related GeoArchive, carries with it some technical obstacles. Some of the greatest difficulties in GeoArchive management and interpretation stem from linking, combining and merging data across agencies and between local and central sources.

First, even the most basic data elements of a GeoArchive may be defined differently from agency to agency. Street address information, for example, which is fundamental to a geocoded address-based data set, may be used differently in two agencies (Block and Block, 1995). In a data set maintained for tax revenue purposes a tavern address might be the mailing address for accounting, while in a data set maintained for police purposes a tavern address could be the actual
Figure 8: Latin King and Two-Six Offender Residency Hot Spot Areas, 1987-1994
Figure 9: Latin King and Two-Six Violent Hot Spot Areas, 1987-1994
location of the bar; these two street addresses might be several doors apart or around the corner from each other. Though both addresses are accurate, they are not the same. As a result, the accuracy, precision and even definition of address may differ from one agency to another.

Second, for most data sources, information is not static, but is continually being corrected and updated at the source. This raises two issues: the degree to which the GeoArchive data should be synchronized with the source data, and the degree to which the database system should emphasize current versus past data. One trade-off is having GeoArchive data that is consistent with "official" data from the source, versus the dangers of data overload. Every time new source data are received by an archived data set, the new data must be integrated with earlier data so that archived information will not be written over and destroyed (Miller, 1995). Another trade-off is the availability of current versus past information. Timeliness can be vital, not only for investigation but also for targeting areas at high risk of an escalating crisis. On the other hand, past information is often vital to the identification and analysis of a problem.

Chicago's ICAM system approaches this problem by emphasizing current information. Incident data are available to ICAM users no more than 24 hours after the incident occurs (usually much sooner). This emphasis on speedy access is a large factor in the success of ICAM, and a major reason for its widespread use by district-level officers. To achieve this speed, ICAM captures incident information as it travels from the district to the central data division of the police department, before the central office has added any additional information, enhancements or corrections (Rich, 1995). While this creates some differences between the central and local data sets, complete synchronization in this case is less important than current availability of data. For similar reasons, ICAM data sets currently contain only three months of data. This not only helps to avoid data overload, but also the lack of a long time series of archived data in ICAM makes it less necessary to worry about synchronizing ICAM data with centrally archived data sets.

**Triangulation: Combining Data from Diverse Sources to Improve the Accuracy of Measurement**

Outside of textbook examples, the perfect measure seldom exists. In practical applications, a combination of two or three indicators may each capture a different aspect of the variable to be measured.
This is one of the main benefits of linking law enforcement and community data in a GeoArchive. For example, there is no address-level data on firearm availability in Chicago, though this variable has always been a high priority for the Early Warning System. In a current project, we are hoping to use data sets from three sources — firearm confiscations (Chicago Police Department), multiple buys (Bureau of Alcohol, Tobacco and Firearms) and firearm injuries (Cook County Trauma Registry) — to create a neighborhood-level indicator.

**Multiple Data Sets, Users and Uses Create Opportunities But Pose Data Management Problems**

Multiple, related databases offer great potential benefit for GeoArchive users, but only if they are organized and managed so that the data sets are efficiently related to each other but still timely and accessible to users. Multiple GeoArchive databases present several challenges: linking diverse databases to each other and linking the same database to updated versions, assimilating information that accumulates at a great rate, synchronizing multiple and changing data sources within the GeoArchive, and balancing some users' needs for quickly accessible information with other users' needs for exhaustive detail (see Block and Green, 1994, for more detail).

The GeoArchive must be able to generate spatial analysis and maps or sets of maps targeted to meet the varied needs of many users: community agency workers and public officials, detectives and tactical officers, neighborhood patrol officers, crime analysts and others. Various problems may require information at a different level of aggregation, at a different degree of timeliness, or in more or less detail. The information needs of tactical and investigatory support officers are not the same as those of crime analysis; the needs of short-term, crisis decisions and long-term planning vary; and the information necessary to develop a crime prevention strategy may not be the same as that necessary to apprehend an offender.

A GeoArchive with multiple data sets from numerous sources, each with different fields and definitions covering different time spans and requiring periodic updating according to different cycles, creates considerable opportunity for confusion and error. Many of the data sets in a GeoArchive file are huge and detailed. Chicago, for example, accumulates 50,000 to 65,000 incident records a month with many fields of information in each record. To attempt to provide instant and immediate access to such a large, complex database might increase data overload to the point where the GeoArchive becomes too difficult
and too slow to be used in practical situations. On the other hand, for many decisions — such as solving an investigation puzzle — minute detail is vital. Similarly, when law enforcement data are being generated every minute, it may be vital for beat officers to have the most current information possible. Yet, solving some neighborhood problems may require comparative information for past months or years. Thus, different users and different decisions require different levels of detail and emphasize either current or long-term data.

The development of street gang territory maps (Block and Block, 1993; Block and Green, 1994; Jacob and Block, 1995) provides an example of multiple data and multiple users, and of combining law enforcement and community data. In the Early Warning System project, the Gang Investigation Section of the police department provided initial information about street gang territory locations (Block and Block, 1993). However, local detectives saw greater detail, with more specific factions, and community leaders saw somewhat different boundaries. In general, the profusion of cognitive maps in a neighborhood produces a variety of perceptions of where street gang territory boundaries lie.

Some of these differences are due to perspective or experience, some due to a focus on the general versus the specific, and others due to actual changes in the territories over time. (One purpose of turf battles is, after all, to change turf boundaries.) However, none of these perspectives is wrong; the problem is how to combine all of these truths into a territory map that is "best" for most decisions. For some decisions it is important to know the specific, current detail of each faction's boundaries, while for others this degree of detail would present a data-overload obstacle to quick decisions. Perhaps the best, but not the easiest, solution is to respond to both needs — to create multiple, related gang territory map files containing multiple levels of detail, update them on a regular basis while maintaining earlier files in an archive, and devise and maintain a simple, current, summary map for rapid tactical decisions.²⁹

Methods of Combining Both Point and Area Data in the Same Analysis

The first tenet of a GeoArchive is that it contains both point- and area-level information. To exploit all the information in these point and area data sets, we need to do more than simply describe them on overlapping maps. We need to be able to relate point to area data, and vice versa, in a multivariate analysis. One way to do this is to
turn area data into points (for example, population potential; see Choldin and Roncek, 1976; Felson, 1986). Another approach is to turn points into areas.

STAC Hot Spot Areas turn point data into areas that reflect the actual scatter of points over the map (the point pattern). These areas are not arbitrarily determined by boundaries of political units, Census tracts or street patterns, but represent those areas where the points being analyzed (events, places, activities, etc.) are most densely clustered. For mapping, each of these dense clusters is bounded by the best-fitting standard deviational ellipse.30

Unlike procedures that are ultimately based on area-level data, such as isolines connecting area centroids (for example, Curtis, 1974) or population potential measures, STAC Hot Spot Areas do not lose the detailed information of point pattern scatter. Unlike topographical isoline maps (see LeBeau, 1995), STAC Hot Spot Area ellipses have defined boundaries that are easy to visualize in relation to other point, area or boundary information on the map. Unlike cognitive maps or expert opinion, STAC Hot Spot Area analysis is an objective, quick, database-driven method for finding dense areas that the expert in question may not know about.31

By itself, a point data set cannot define a particularly dense area. Though a particular address may be a high-crime place, it is not a Hot Spot Area. If applied to an address, the term "area" takes on a qualitatively different meaning. In addition, a high-crime address could reflect some unique characteristic of the particular location, and irrelevant variables (such as the presence of a pay phone from which calls for service are made) could easily obscure the measurement.32 As Block and Block (1995) found in the analysis of liquor-license crime, a single address with more crimes than any other address may or may not be located within a high-density crime area. Some high-crime liquor outlets are located in dense areas of liquor-outlet crime, and others are not. By the same token, some low-crime places are located in high-crime areas and others are not. Both the characteristics of the place and the area are important in determining the most effective strategy for intervention in a particular case.

Hot Spot Area ellipses are area summaries of point patterns.33 To complete the cycle, Hot Spot Areas can be related, in turn, to another set of point data. We can explore, for example, the location of street gang-related homicides relative to Hot Spot Areas of street gang non-lethal violent offenses versus drug offenses (see Figure 6). In addition, two sets of Hot Spot Areas can be related to each other, as in the
comparison of clusters of serious street gang offenders (see Figure 8) to clusters of street gang-related offenses (see Figure 9).

Hot Spot Areas can be related to areas defined within arbitrary boundaries, such as Census blocks, Community Areas (see Figure 5 and 6), and more complex aggregates of area-level data (for example, see Hirschfield et al., 1995). In addition, it is simple to calculate a variety of density rates for any Hot Spot Area, not only density per square mile (as in Figure 6), but also rates based on the occurrence of any other appropriate information that can be summed within the Hot Spot Area boundary. In the Early Warning System project, for example, we have found it useful to calculate the density of drug-related offenses in a Hot Spot Area, using the number of abandoned buildings in the area as the denominator of the rate. Finally, as the next section will show, population-based densities can be calculated using Census block data. In general, the opportunities for combining area and point data in the same analysis are numerous and the potential benefits are great.

A Good Descriptive Map May Be Enough for Communicating Information, but for Effective Decision Making Spatial Analysis Tools Are Also Needed

In Manning's (1992:380) review of technology and the police, he complains that "very few uses of information technology are analytic, strategic, or tactical." "Tertiary" information that goes beyond administrative records to support complex analysis and problem solving "is rarely found in policing and, when available, is rarely used" (1992:380). This situation is not confined to law enforcement, but is also seen in other government agencies (see Sparrow, 1994, for a review of several) and among users of spatial information in general (Bailey, 1994). Though not intended to be a complete introduction to spatial analysis, this section outlines a few of the things we have learned.

The analysis of spatial patterns, like the analysis of other kinds of patterns, can be roughly categorized into descriptive and exploratory techniques versus analytical or hypothesis-testing techniques. For example, cartography — the art of creating maps that present spatial information clearly and concisely — tends to be more descriptive, while spatial statistics describing and evaluating point patterns (Boots and Getis, 1988) tend to be more analytical. However, the line of demarcation is vague. As with other kinds of analysis, a good description can, and indeed must, be the foundation of a causal analysis. In Mapping It Out, one of the best practical reviews of techniques
for "expository cartography," Monmonier (1993) presents ways to map such analytical concepts as change over time, variation in intensity, flows and processes, and causal models. A "pretty map," therefore, may be the basis for a causal analysis as well as a means of communicating the results of the analysis.

An information foundation for community policing (see Figure 1, above) requires both a GeoArchive and spatial statistical tools, and the statistical toolbox should include not only STAC Hot Spot Area analysis but others as well. The focus of the 1993 Workshop on Crime Analysis through Computer Mapping was "going beyond pretty maps," and the presentations and discussions were organized around techniques to answer three types of spatial analysis questions: analysis of differences across areas, point pattern analysis, and analysis of sequential travel or attack patterns. In addition to tools that approach spatial analysis questions from each of these three perspectives, as Monmonier points out, we also need tools that can present and analyze spatial data over time (see Langran, 1993).

Analysis of differences across areas, often depicted as "thematic" maps with each area shaded according to the intensity of some indicator (for example, see Figure 5), is often the most accessible kind of spatial analysis. Area maps have long been available in automated mapping systems, and they are relatively easy to do by hand. However, statistical methods for the analysis of relative crime density across arbitrary areal units suffer from strong aggregation biases and serious problems in interpretation (Brantingham and Brantingham, 1984). Also, they cannot deal with a reality in which dense areas cross boundary lines or occur along a boundary line. STAC offers a resolution to this problem by turning points into areas. Because it does not rely on predetermined boundaries, STAC fulfills Openshaw's (1994:87) call for spatial analysis techniques that "impose as few as possible additional, artificial, and arbitrary selections on the data."

A common problem in GeoArchive analysis is the calculation of population-based rates for areas where the boundary does not coincide with a standard Census area. Richard Block has developed a method that will produce useful rate estimates (Block, 1995b) when the area in question is sufficiently large relative to the size of a Census block. Given a GeoArchive containing boundaries for the larger area(s) plus Census block data, he calculates the sum of the populations of all Census blocks in which the centroid (the geographical center) is located within the area boundary. This is possible even without mapping the boundary of each Census block, because the centroid coordinates of each block are included in the Census data.
set. Estimates of the population within a police district, a street gang
territory, a Hot Spot Area, or any enclosed boundary can be calculated
with this method, thus making it possible to compare the relative
density per resident of events or incidents across many variously
defined areas.

The identification and description of dense clusters on the map is
just one of many practical spatial analysis questions, each requiring
a different statistical tool. There are two types of point pattern analy-
sis (Boots and Getis, 1988): those that describe arrangements of
points in space (such as the STAC Hot Spot Area ellipse), and those
that indicate degree of dispersion versus clustering (such as Nearest
Neighbor Analysis). For overviews of spatial statistical tools for point
pattern analysis, see Boots and Getis (1988), Block (1995), Canter
statistical tools for tracing travel patterns from one point to the next,
see Rossmo (1995), and, for early applications, see Pyle et al. (1974)
or Brantingham and Brantingham (1981). In general, however, spa-
tial analysis techniques are still underdeveloped, especially tech-
niques that are accessible to practical decision makers. As Openshaw
(1994:84) argues, we need techniques "that are able to hunt out what
might be considered to be localised pattern or 'database anomalies' in
geographically referenced data but without being told either 'Swhere' to
look or 'Vwhat' to look for, or 'Vhen' to look."

It is also important to remember that the spatial analysis toolbox
needs more than one tool; a single kind of analysis may not be
enough. Monmonier (1993) strongly urges map makers to create and
present more than one kind of map, and argues that it would be un-
ethical not to do so.

Because the statistical map is a rhetorical device as well as an
analytic tool, ethics require that a single map not impose a de-
ceptively erroneous or carelessly incomplete cartographic view
of the data (Monmonier, 1993:185).

Richard Block's current study of patterns of street robbery relative
to the locations of rapid transit stops utilizes not only several kinds of
maps but also several kinds of spatial analysis. Patterns of street
robberies in two Chicago police districts from 1993 to 1994 are
shown in Figure 10, which includes locations of parks, large ceme-
teries and major institutions as well as transit stations. Two kinds
of spatial analysis are depicted here. First, the number of robberies
occurring at a single address is shown by a circle scaled to represent
the number. (The legend shows the relative size of the circles from
only one incident to a maximum of seven incidents.) Second, Figure 10 shows the Hot Spot Area ellipses resulting from a STAC search for the densest clusters of robberies throughout the entire map.

Figure 10: Northeast Side (Districts 20 and 24) Street Robberies 1993-1994: Number at a Location and Hot Spot Areas
The graduated circles in Figure 10 are a cartographic device developed to handle a problem that is very common in maps of law enforcement data. Criminal incidents do not tend to be randomly scattered across the map. Certain places, like certain groups of people, are relatively vulnerable to crime. In a small-scale map where the maximum number of occurrences at each address are also small, it is possible to use symbols or even actual numbers to depict the number of occurrences at an address. In maps such as Figure 10, however, with counts from zero to seven, symbols or numbers would be difficult to see and interpret. In such cases, we have found that symbols of graduated size are clear and unambiguous (circles or triangles work best).

The graduated circles convey two pieces of information: both the risk of ever having a robbery at an address and the risk of multiple robberies increase toward the eastern border of the map. The heaviest concentration runs along the lakefront, with intermittent dense areas around transit stations. A secondary line of high-risk addresses runs parallel to the first, along a major north/south street (Clark Street). The analysis of STAC Hot Spot Areas takes this a step further. There are ten transit stations on the map, and every one is located within a Hot Spot Area of street robbery offenses. Even though the STAC search encompassed the entire area of both districts, it found a series of small Hot Spot Area ellipses, all but one of which contains a transit station, all located along the rapid transit line.

What mechanisms would drive these patterns? To answer this question, the Community Safety Project measured the Manhattan distance between the location of each street robbery in 1993 and 1994 and the nearest transit station. Because Chicago streets are laid out on a grid with an eighth of a mile (about 600 feet) between each block, the distance in feet between a transit stop (usually located at a corner) and a street robbery is a rough indicator of the number of city blocks a victim would walk away from the station before being robbed. The resulting graph (Figure 11) is an example of analysis based on spatial information, but not presented as a map.

Taken together, these three kinds of spatial analysis indicate that the most dangerous addresses (based on the analysis-level count analysis) and the most dangerous areas (Hot Spot Area analysis) for street robbery in this neighborhood are located close to a rapid transit station. However, the stations themselves are relatively safe (distance analysis). People are less at risk of street robbery when they are at or still very close to a station, but their risk increases rapidly as they move a block (about 600 feet) to two blocks (1,200 feet) away.
Figure 11: Northeast Side (Districts 20 and 24) Distance From Elevated Station by Number of Incidents for 1993-1994

(There is also a secondary peak at about 3,600 feet, which reflects the secondary line of high-risk places along Clark Street.) Block (1995c) concludes that surveillance is a key issue here (see Felson, 1987). The highest risk seems to occur not at the station itself, where there are likely to be other passengers and the ticket-taker, but after the victim has walked a block, where presumably the crowd, if any, has thinned out. The fact that the peak occurs after the first block may reflect the decreasing surveillance after victims turn the corner.
FINAL WORDS

Recent research (Berry et al., 1993) suggests that, contrary to conventional wisdom, strong neighborhood participation in local problem solving does not produce more conflict or delay in policy making. Decentralization is built on the assumption that the most efficient way to define local problems and to develop solutions may be to provide information to local decision makers. This is because the boundaries and concerns of a community do not necessarily coincide with the arbitrary boundaries defined by Census tracts or police districts (Suttles, 1972), and the problems and resources of a local community do not necessarily correspond with citywide definitions and priorities.

But decentralization of problem solving cannot occur without decentralization of access to information. If problem-oriented community policing is to be built upon community and police collaboration in problem solving at the local level, it must also be built upon access to and control of information at the local level. In other words, successful community policing must be built upon a foundation of information that is address-based and focused on the neighborhood level, that contains both community and law enforcement data, and that is organized so as to be accessible for local problem solving. This chapter has described such an "information foundation for community policing" and has detailed practical suggestions for developing such a GeoArchive.

NOTES

1. This paper is based in part on the experience of the Early Warning System for Street Gang Violence project of the Illinois Criminal Justice Information Authority and the Chicago Police Department (CPD) Area Four Detective Division, which was partially supported by the Bureau of Justice Statistics (BJS). Many Authority analysts, including Robert Whitaker, Lynn Higgins, Anthony Mata, Graham Taylor and Michael Maly, helped with the Early Warning System project in its early stages, but the most crucial persons were the project manager, Lynn A. Green, and CPD Area Four personnel Sgt. Ronald F. Rewers, Det. Richard Respondi and James Elliot. In addition, the project would not have been possible with-
out the advice and support of Paul White of BJS. Richard Block, who pioneered automated crime mapping in Chicago (see Block, 1977), was instrumental in developing the GeoArchive and continues to create innovative applications for community problem solving. Currently, Daniel Higgins, Teresa Hirsch and Ayad Jacob support STAC, the Early Warning System and the GeoArchive at the Authority, and are responsible for many of the ideas and some of the analysis described here.

2. A database (or data table) is a set of information, including an identifier, that is organized in a file structure with fields (variables) displayed in columns and types of information such as offense, date, address or offender's street gang displayed in rows.

3. For a discussion of community policing, problem-oriented policing, and their combination, see the upcoming "Theory Overload" section.

4. Many of these mapping entrepreneurs participated in the Workshop on Crime Analysis through Computer Mapping held in August 1993, and contributed to the proceedings of that seminar (Block et al, 1995.) Sponsored by the Illinois Criminal Justice Information Authority and Loyola University Chicago, with support from the Innovations in State and Local Government program of the Ford Foundation and the JFK School of Government, Harvard University, the theme of the workshop was "more than just a pretty map" (Rengert, 1995a).

5. STAC is a toolbox of spatial analysis statistics designed to support practical law enforcement decisions. STAC is a stand-alone spatial analysis package, not a mapping package. It was developed by the Illinois Criminal Justice Information Authority with the collaboration of STAC users around the world, and is available from the Authority at no cost to law enforcement agencies (Higgins et al., 1995).

6. Spatial analysis is "a general ability to manipulate spatial data into different forms and extract additional meaning as a result" (Bailey, 1994:15). Statistical spatial analysis uses statistics to this end, and consists of spatial summary statistics and spatial analysis statistics.


8. In a digitized map, features such as streets, rivers or political boundaries have x-coordinates and y-coordinates that place them on the map. A file of such features may be called a street file or a map data file.
9. Goldstein (1990) defines and explains the steps of problem solving in more detail, including the following: grouping incidents as problems, labeling the problems, analyzing the relevant interests in the problem, critiquing the current response, searching for innovative responses, establishing accountability, implementing solutions, and measuring their impact.

10. "Individual characteristics" here include characteristics of individual persons, places, and situations.

11. For discussions of some deficiencies of and problems with current GIS and spatial statistical packages, see Levine (1996), Levine et al. (1995) and Block (1995).

12. GIS stands for Geographic Information System. Geographical information systems, which by definition are capable of storing and manipulating point, line and area data (Guptill, 1988:3), are database management systems specifically developed to handle spatial data in an efficient relational database system. Database management systems are "normally designed to handle numeric and textual information and are not capable of manipulating spatial data" (Guptill, 1988:2). See the essays in Fotheringham and Rogerson (1994) for more detail.

13. Geocode means (1) to assign x- and y-coordinates (e.g., longitude, latitude) to an address; (2) to assign an event, incident or map feature to an area; or (3) the x- and y-coordinates corresponding to a given address. The elements of a geocoded dataset are related to x- and y-coordinates so that they may be placed on a map that has the same coordinate system.

14. In "thematic maps" of area data, each area is categorized on a scale, represented on the map by different colors or shading or by an icon (symbol) placed within the area. In thematic point maps, categories of events or places are represented by an icon located at the x- and y-coordinates of that event or place.

15. Even when area information is linked to a central point (centroid) or an intersection within the area, the spatial unit of analysis is still area, not point. "Map data" is a drawing of features such as streets, bodies of water or landmarks, with no information behind the drawing other than the location on the map.

16. Figure 5 was produced by Antigone Christakos; Figure 6, by Ayad Paul Jacob.

17. The Hot Spot ellipses in Figure 6 do not represent clusters of homicides, but rather clusters of non-lethal violence or drug offenses. To locate the Community Areas mentioned here, see Figure 5.

19. Area dataset boundaries usually do not change as frequently, though the information behind the areas often does.

20. Since 1992, Richard Block has geocoded the approximately 50,000 criminal incidents occurring each month in Chicago, with a "hit rate" (incidents placed on the map) of 97% completely automated and close to 99% after hands-on geocoding. This is possible because of a street file that has been corrected and enhanced with places appropriate to police work, and a program that corrects misspellings and non-standard street name references (R.L. Block, 1995a). The Chicago Police Department's ICAM system achieves an automatic 96% hit rate (Rich, 1996). Again, the high rate is due to improved street files and an address-correction program.

21. See Block (1995a) for a discussion of errors in the TIGER files and how to correct them.

22. In contrast, see the discussion in Guptill (1988) of conducting a User Requirement Analysis to determine user needs, an example of control from the top down rather than from the bottom up.

23. In addition to ICAM and the Loyola Community Safety Project, the new Hartford, CT Comprehensive Communities Program, a collaboration of Hartford, the U.S. National Institute of Justice and Abt Associates, plans to provide mapping capability and training to the 17 neighborhood "problem-solving committees" in the city, in the hope that the organizations will use mapping to increase the effectiveness of their community problem-solving.

24. This map was created by Daniel F. Higgins.

25. For more details of the problem and how to solve it, see Block and Block (1995).

26. Figures 8 and 9 were created by Ayad Paul Jacob.

27. This system is not only quick, but because data are entered only once (traveling to both ICAM and the central repository) it is also accurate and relatively inexpensive.

28. In ICAM II, the capabilities and level of detail in ICAM are being expanded (Rich, 1996).

29. For an innovative approach to this problem, see Kennedy and Braga (this volume).
30. A standard deviational ellipse rotates two axes around the cluster of points, until the variance of the $X_i$'s are maximized along one axis and the variance of the $Y_i$'s are maximized along the other axis (see Stephenson, 1980; LeBeau, 1987; Ebdon, 1985; Levine et al., 1995).

31. Though cognitive maps and expert opinion can provide other valuable information (see Maltz et al., 1991; Weisburd et al., 1993; Buerger et al., 1995), they are not objective, automated analysis tools.

32. For a review of other problems encountered in using frequency across specific addresses to define dense areas, see Buerger et al. (1995).

33. STAC can handle much larger datasets (up to 16,000) than most other point pattern analysis programs.

34. Figure 10 was created by Richard L. Block.

35. For detailed advice and examples on mapping count data, see Monmonier (1993.)

36. Manhattan distance measures the distance "around the block"; Euclidean distance, "as the crow flies."

37. Sean Davis calculated the distance data represented in Figure 11.

REFERENCES


Carolyn Rebecca Block


Criminal Justice Information Authority and Loyola University Chicago.)


