Guidebook for Measuring Crime in Public Housing with Geographic Information Systems

Guidebook for Measuring Crime in Public Housing with Geographic Information Systems

Prepared for:

U.S. Department of Housing and Urban Development
Office of Policy Development and Research

Prepared by:

Robert A. Hyatt, Ph.D.
Research Triangle Institute

Harold R. Holzman, Ph.D.
U.S. Department of Housing and Urban Development
Office of Policy Development and Research

August 1999

TABLE OF CONTENTS

CHAPTER 1 - Introduction

1.1 Purpose of this Guidebook

This guidebook is about the computerized mapping of crime locations. The guidebook's fundamental objective is to provide public housing authorities (PHAs), local police departments (PDs), and criminal justice researchers with a straightforward, relatively inexpensive method for measuring crime in **public housing developments** and other multifamily housing complexes (see Glossary for definitions of bold faced words). The approach presented here was the focus of recent research sponsored by the U.S. Department of Housing and Urban Development's (HUD) Office of Policy Development and Research (PD&R). That research project, which was conducted by Research Triangle Institute (RTI), is part of PD&R's ongoing criminal justice research agenda, aimed at creating "tools" to facilitate the evaluation of crime prevention programs at the individual housing development level. The widespread commercial availability of easy-to-use geographic information systems (GIS) and the subsequent embracing of GIS by PDs as an important crime analysis tool were crucial in laying the groundwork for the approach presented here.

At first glance, some criminal justice researchers familiar with GIS might be inclined to dismiss the task of measuring crime in public housing using GIS as so technologically straightforward as to warrant the characterization of "simple." Yet, to date, this particular application of GIS technology is exceedingly rare. Although the computer programming *per se* is indeed straightforward, *the fundamental implementation of the data collection effort required to support GIS analysis will likely be more complex.* The following analogy seems appropriate here. To those individuals unfamiliar with the ways of tigers, the task of bandaging a tiger's paw might appear technologically uncomplicated. But, once engaged in that task, positioning the tiger might prove quite challenging. Similarly, a familiarity with the public housing universe and the historically uneasy state of relations that sometimes prevails between PHAs and their local PDs suggests that partnerships between these groups must be developed carefully.

With rare exceptions, where GIS is currently being used to map crime in a local jurisdiction, it is the PD rather than the PHA that is the focus of this activity. Therefore, the information in this guidebook is structured on the existence of that condition. Thus, there are two basic requirements for successfully using GIS to measure crime in public housing. First, the local PD must already be routinely recording the eight major Federal Bureau of Investigation (FBI) Part I Crimes (i.e., homicide, rape, robbery, burglary, larceny, assault, motor vehicle theft, and arson) in a GIS database. In other words, they must be routinely **geocoding** locations of reported Part I Crimes. Second and equally important, the local PD and the PHA in its jurisdiction must formally agree to a well-defined partnership for the systematic sharing of both crime and geographic data.

The production of some of the GIS "deliverables" to be derived from a PD/PHA data-sharing partnership will require a fair level of sophistication with respect to the application of this type of software. Geographic information systems are a relatively new adjunct to the business of crime analysis and therefore experienced crime mappers are in short supply. Therefore, it is also important to note here that, depending on the level of GIS expertise at the participating PD, it may be necessary for the initial stages of the partnership to include the services of an experienced GIS consultant (see Section 4.3). Ideally, this consultant should also have strong skills in social science research and statistics. Although the PD/PHA partnership will be using data that are already routinely collected by the PD, those data will need to be manipulated to produce a new set of tables, graphs, and rates. If either partner employs social science researchers with experience in the use of applied statistical analysis software (e.g., SAS or SPSS), then assistance from outside consultants may safely be limited to GIS. Once the process for creating these new "deliverables" has been standardized, the need for any outside assistance is likely to involve only occasional troubleshooting.

At present, few PD/PHA partnerships for the exchange of geocoded crime data exist. However, the above-mentioned HUD-sponsored research project (the HUD Study), conducted by RTI in 1997-98, created such partnerships in three cities and proved that teaming PDs and PHAs for measuring crime in public housing is practical. This guidebook draws heavily on the experience gained in that project with respect to establishing and testing the administrative and technical arrangements necessary for the implementation of HUD's GIS measurement approach. It is hoped that this guidebook will assist PDs and PHAs as they embark on data-sharing partnerships.

One of the major law enforcement innovations of the last decade has been *community policing*. This law enforcement approach focuses on customizing police responses for small areas such as neighborhoods. Community policing also attempts to join community groups and police departments together in cooperative, problem solving initiatives. This HUD-funded project that developed a collaborative relationship between PHAs and PDs in three cities is very much in the spirit of community policing. In a very practical sense, the development of the capability to map the local public housing community's crime problem begins with a series of conversations that might reasonably be expected to serve as the basis for the future development of the cooperative crime control initiatives that are often a key ingredient in community policing.

1.2 Why Use GIS to Measure Crime in Public Housing?

The incidence of crime in public housing has yet to be routinely and systematically measured. In the vast majority of jurisdictions with public housing, official police statistics on crimes specific to those areas are just not available. The majority of law enforcement organizations base their statistics on relatively large geographic areas often called precincts or districts. Police rarely publish official crime statistics for small parcels of land such as public housing developments. Few of the biggest public housing developments (i.e., 1000+ units) come close to even qualifying for consideration as distinct crime reporting zones. In any event, these 1000+ unit developments are exceedingly rare, appearing in less than one-half of one percent of the Nation's 3300+ public housing authorities. Furthermore, these very large developments constitute less than one-half of one percent of the roughly 14,700 public housing developments in the United States. While a handful of criminologists have attempted to gauge the levels of some crimes in public housing, even the most rigorous attempts have been unable to generate comparisons between ostensibly high-crime public housing developments and their adjacent neighborhoods or even adjacent blocks. That being the case, it is possible that many inner-city public housing developments—by comparison—may prove to be calm islands in the midst of neighborhoods beset by crime and disorder.

GIS has the ability to generate crime statistics in arbitrarily defined geographic areas, including small areas such as individual public housing developments and/or neighboring areas within an arbitrary distance from a development. In the recently completed HUD study, RTI performed such tasks in the context of an investigation of GIS-based methodologies for measuring crime in and around individual public housing developments. In the course of that research, GIS was used to aggregate data from all of a city's public housing developments. In effect, a single imaginary public housing police "precinct" was created in each of the HUD study's three cities.

1.3 What This Guidebook Does

This guidebook addresses the following questions:

1. How can the incidence of crime in public housing be measured easily and inexpensively using GIS?

2. What are the fundamental requirements for a partnership between a PHA and a local PD that measures crime in public housing using GIS?

3. What are some useful ways in which public housing crime data can be packaged for both the PHA and the PD?

4. What are the specific roles of PHAs and PDs in a crime-data-sharing partnership? Will it be necessary for the PHA/PD partnership to obtain the

services of an “outside” GIS consultant to assist in designing and programming the database and statistical/graphic output unique to the endeavor?

5. What portions of the GIS data creation and sharing process are likely to prove most troublesome? What are the most important “lessons learned” from the HUD study with respect to avoiding these trouble spots?

In attempting to answer these questions, some fairly technical “nuts and bolts” of GIS need to be presented. While both PHA/PD managers and their respective technical staffs are encouraged to read this guidebook in its entirety, detailed knowledge of the mechanics of computer mapping is not necessary to design, negotiate, and oversee the PHA/PD partnership for the collection and sharing of public housing crime data. Therefore, the chapter devoted primarily to technical information has been placed near the end of this guidebook, and a glossary of technical terms has been provided.

CHAPTER 2 - Critical Issues for the PHA/PD Partnership

2.1 Why Use the FBI Part I Crimes to Measure Crime in Public Housing?

The research that resulted in the production of this guidebook was restricted to generating statistics for Part I Crimes reported to the police. This same group of reported crimes is a major feature of the FBI's annual publication *Crime in the United States*. This FBI report is an integral part of the statistical program known as the Uniform Crime Reports (UCR).

Besides being generally recognized as including the most serious types of violent and property crime, the FBI's Part I Crimes were chosen for the following reason: The overwhelming majority of U.S. police agencies participate in the UCR crime data collection program and, therefore, use the same definitions for each of the Part I Crimes. This uniformity of definition also extends to the crime reports that participating PDs use in generating their own crime statistics. Therefore, Part I Crime data represent the standardized depiction of crime, and their use allows for valid crime comparisons across the highly diverse universe of law enforcement agencies in the United States.

Also important, PDs and criminologists typically manipulate UCR data in more or less standardized ways in order to calculate rates of reported crimes. For example, statistical reports often present data as "crimes per 1000 population." Hence, meaningful comparisons of rates of particular offenses between cities or counties are possible. In summary, Part I Crime categories and the associated rates are the "industry standard" for the measurement of crime in the U.S.; comparisons of Part I Crime rates across jurisdictions are both reliable and valid. Therefore, their use also allows for meaningful comparisons of crime levels across PHAs and between types of housing in the public housing universe (e.g., specific types of development such as high rise versus low rise), between PHAs, and between PHAs and other geographic entities such as neighborhoods.

The major limitation of using the FBI's UCR is that crimes not reported to the police are not included, and it is well known that many crimes, even serious ones, are not reported to the police. National victimization surveys show that fewer than half of all violent crimes are reported to the police. However, the extensive victimization survey research, conducted since 1970, suggests that the reported proportions of various Part I crimes have stayed fairly constant. Hence, knowledge of the "dark figure" of crime allows one to generate realistic estimates of the actual incidence of serious offenses.

2.2 Mapping Drug Offenses

Unquestionably, drug abuse violations are a major focus of most PHA crime prevention programs. In the UCR taxonomy of offenses, drug violations fall into the group of offenses termed Part II crimes. While the UCR program reports information on the number of Part I offenses known to the police, the number of Part I cases cleared by arrest or other means, and the number of persons arrested for Part I offenses, only arrest data are reported for Part II offenses. Besides drug violations, there are some 17 other groupings of offenses that are classed as Part II offenses, including weapons offenses, simple assault, prostitution, vandalism, disorderly conduct, drunkenness, and vagrancy. The reader should understand that these offenses are as easily mapped as any of the Part I crimes discussed in the preceding section of this guide. Indeed, there is nothing inherent in GIS that would prevent a PHA/PD partnership from mapping any particular type of crime or any other event of interest, such as calls for emergency medical service or incidents of gun violence. Furthermore, a PHA may wish to map maintenance requests so that the association between "broken windows" and drug violations could be tracked.

While the occurrence of Part I offenses is often linked to drug abuse violations under the rubric of "drug-related" offenses or drug "hotspots," it is important to note a crucial difference between Part I offenses and drug crimes. Rates of drug crime (e.g., the incidence of possession and/or trafficking in illegal drugs) are largely a result of police enforcement activity since the complainants in drug cases are usually the police. The level and location of drug enforcement activity is discretionary. Law enforcement agencies determine where they will try to suppress illicit drug traffic. The geographic concentrations of such police activities then become a major factor in determining concentrations of known drug crimes. In contrast, a substantial proportion of Part I Crimes become known to the police because citizens report them. Here, citizens are crucial to initiating police intervention. For example, when a serious violent crime is reported to the police, there is relatively little discretion involved in the PD's initial involvement. Although the police themselves sometimes do come upon Part I Crimes and then, in effect, report them to their PDs as part of their job, the general public is the source of most Part I Crime reporting. Hence, maps displaying known drug crimes and drug arrests more closely reflect the selective allocation of law enforcement resources than the actual geographic dispersion of illicit traffic.

2.3 PD and PHA Responsibilities

As noted earlier, a PD should already be using GIS to track crime in order to successfully partner with a PHA in the approach to measurement of crime in public housing discussed in this guide. Successful GIS operations require significant investments in hardware, software, and training. Therefore, unless the police have already made these investments for other purposes, using GIS to generate crime data specific to public housing will likely be beyond the ability of their crime analysis personnel. When viewed as discrete tasks, the steps required to move the information provided on police reports into GIS, analyze the data, and generate meaningful output are complex and labor intensive. An experienced GIS staff would have automated much of this process so that the data processing would become relatively routine. However, the training and experience required for accomplishing this automation cannot be obtained "overnight."

The GIS approach presented here is practical and relatively inexpensive because the data collection and analysis apparatus already exists at the PD. Thus, in formally contracting with the PD for crime statistics, the PHA is essentially paying the labor costs associated with the generation of a particular set of statistics from data that are already being collected and mapped routinely. In large urban jurisdictions, there may be such a high volume of crime in some parts of town that the different map symbols for various types of crime cover one another and even coalesce, thereby obscuring the view of the distribution of specific crimes such as robbery or burglary. Here, the GIS analyst may be asked to produce maps for each Part I Crime so that its geographic distribution is visible and therefore more amenable to study (see Figure 1).

The primary role of the PHA is to communicate their statistical needs to the PD and to provide accurate up-to-date "maps" of their public housing developments plus the number of residents in each of these developments—the latter is used to calculate crime rates. Chapter 4 of this guide deals with computer mapping technology. However, no such technology is required of the PHA, only of the PD. In fact, the PHA "maps" shared with the PD need only be reasonably high quality street maps of the local jurisdiction with the PHA properties outlined in a highly visible but non-opaque soft-tip pen. The PD GIS analyst can then geocode the boundaries of each PHA property, creating what GIS analysts call a **polygon** for each respective property. These polygons appear as outlines of the PHA properties on computer-generated maps and also serve a cookie-cutter function with respect to the PD's crime maps, allowing the extraction of information about reported crime in each development (see Figure 2).

2.4 Opportunities for More Precise Mapping of Public Housing Developments

The creation of PHA/PD crime mapping partnerships may present an opportunity to generate detailed maps of PHA properties, involving roadways, walkways, and footprints of buildings with entrance/exits marked. In many cases, the PHA may only have original blueprints. Some GIS packages provide for the insertion of such features and may even include an "addressing" capability so that incidents on the property but outside buildings can be geocoded independently of a nearby structure. Just the creation of such maps—independent of GIS benefits—will represent a "first" for many jurisdictions and is likely to result in immediate

Figure 1

Figure 1
Sample Map of Crime Locations

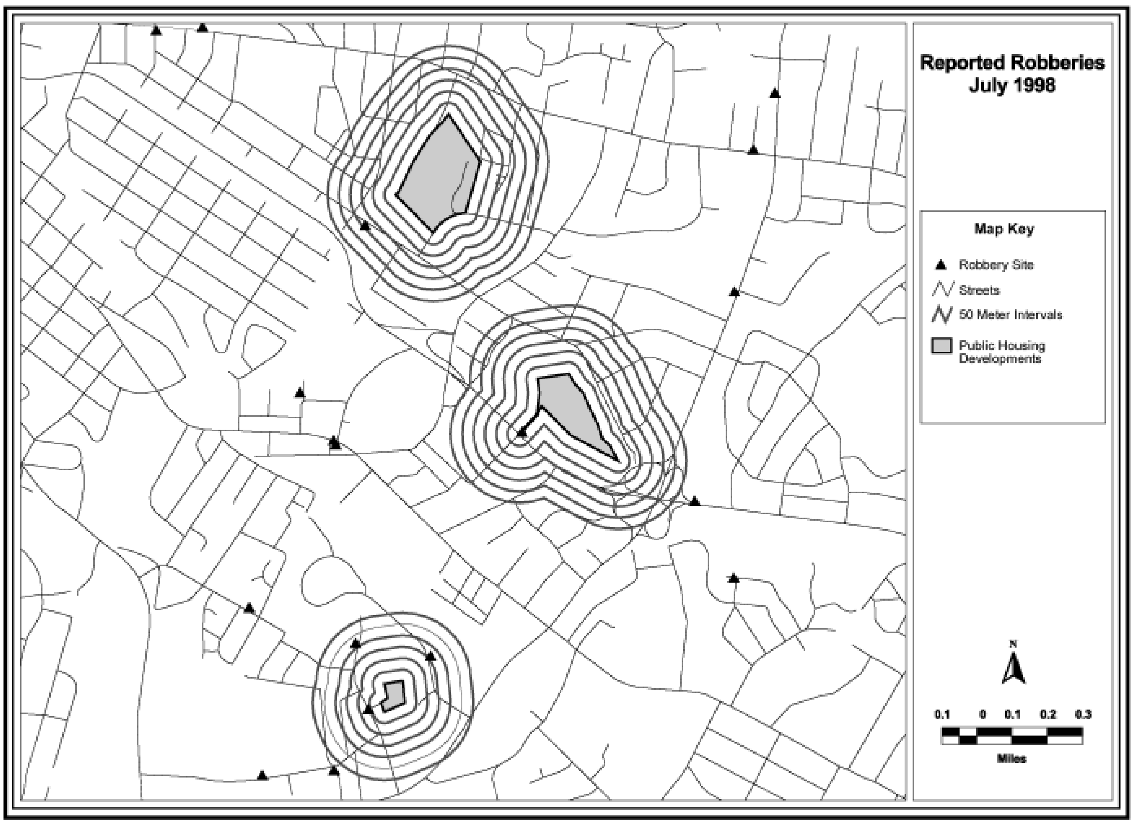
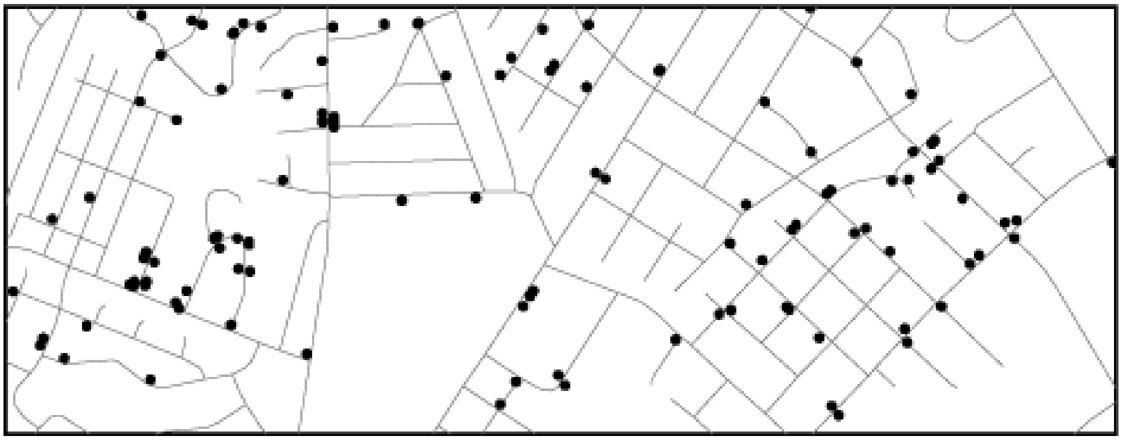


Figure 2

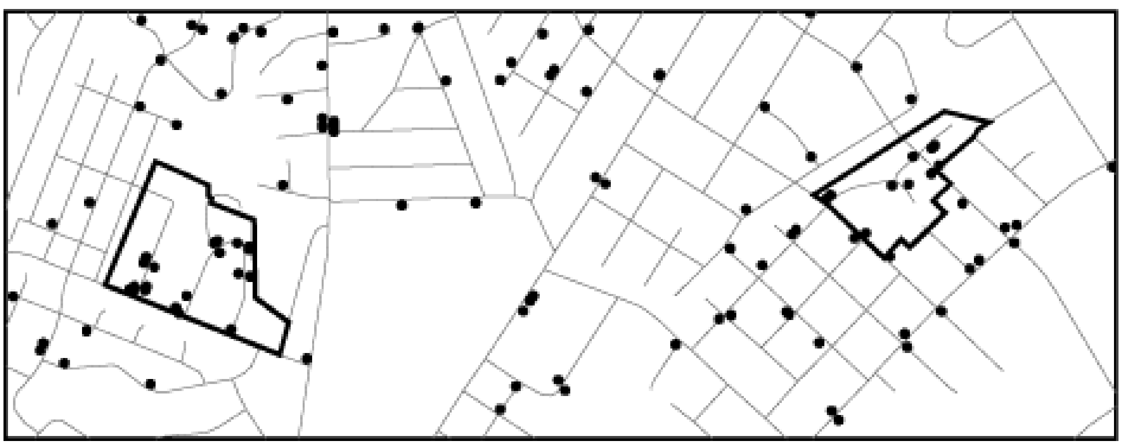
Figure 2.

Sample of "Cookie-Cutter" Process in Crime Mapping

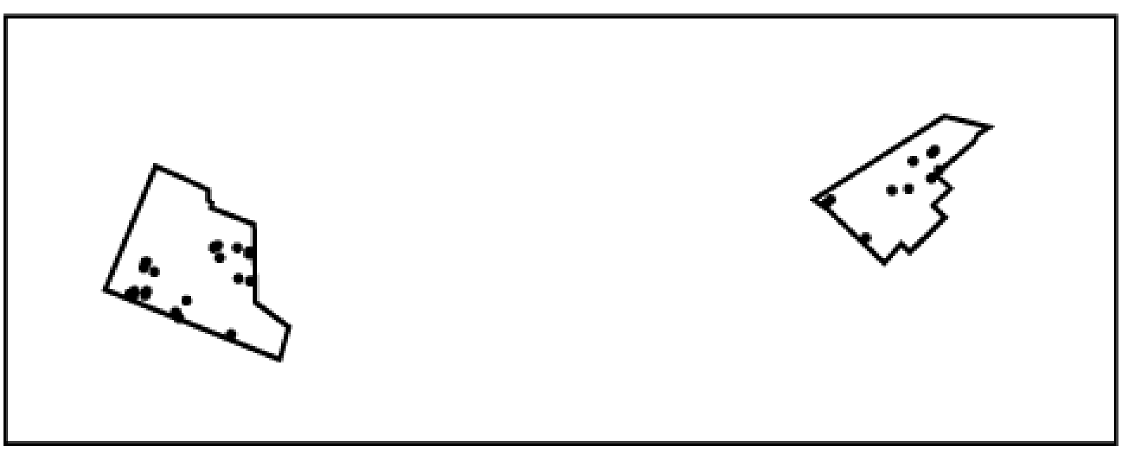
Frame 1. Part I Crimes, July 1998



Frame 2. Crime Points Overlaid by Public Housing Development Polygons



Frame 3. Crimes Found Within Public Housing Development Using Point-in-Polygon Overlay



11

use by the local PDs, facilitating routine patrol activities as well as responses to "911" calls-for-service. More detailed mapping of PHA properties also may enhance residents' access to emergency medical and fire protection services if the new maps are made available to these service providers.

CHAPTER 3 - Shaping Your Data Request: What Kind of Information is Useful for Measuring Crime in Public Housing?

3.1 The PHA as a Crime Reporting District

The HUD study that led to the preparation of this guidebook had as one of its primary objectives the use of GIS to generate crime statistics for PHAs that would be comparable to those typically available for PD precincts or districts. The GIS cookie-cutter approach mentioned above would allow PD crime analysts to aggregate the reported crime data for all of the local PHA's developments, creating a PHA "precinct." Then, current counts of residents of the constituent developments could be summed and used to produce crime rates for Part I Crimes. This procedure entails drawing from the same data stream that the PD (1) sends to the UCR and (2) uses to produce its own crime statistics for precincts/districts and for the municipality as a whole. Thus, crime in public housing would be measured and could be compared with precinct and city crime rates. The first measures of crime in public housing generated in the HUD study then were crime counts and crime rates (rate = crime count/population) for the PHA as a whole, using development-based data.

Noteworthy here is the fact that the PHA "precincts" created in the HUD study did *not* contain "scattered site" properties. As the term implies, these dwellings tend literally to be spread all over the local jurisdiction's map and are virtually indistinguishable from neighboring homes. Often, PHAs acquire buildings directly from private owners or from local governments that have taken possession due to nonpayment of taxes. Also, from time to time, some PHAs have been able to acquire small multifamily properties through new construction. Since it tends to blend so well with its surroundings, scattered-site housing is largely invisible from both the perspective of criminological theory and research and is almost never mentioned in discussions of crime in the public housing universe.

Exhibits 1 and 2 contain actual monthly crime data collected in the HUD study for the PHA (development-based) "precinct" as a whole. The reader will note that homicide, rape, and arson instances were excluded from these two exhibits. This was done because these offenses are statistically rare occurrences. Hence, a few offenses would cause a major fluctuation in rates but would not yield much information of practical value. Data on these offenses are included in Exhibit 3.

Part I crime rates also were graphed across the 6-month test period so that the fluctuations in offenses for the PHA "precinct" could be compared with the jurisdiction as a whole (see Exhibit 4). Fluctuations in crime counts for specific developments also could be used to evaluate crime prevention programs. Such analyses will be discussed later in this chapter.

Exhibits 1 and 2

Exhibit 1.

Number of Reported Crimes by Offense for PHA Developments and the City for April 1998

Offense	All Developments	All Developments, Plus Surrounding Block Faces	Entire City
Robbery	36	61	663
Aggravated Assault	46	90	610
Burglary	26	59	953
Larceny	85	168	2598
Vehicle Theft	14	32	658

Exhibit 2.

Crime Rates Per 1,000 Residents by Offense for PHA Developments and the City for April 1998

Offense	All Developments	All Developments, Plus Surrounding Block Faces	Entire City
Robbery	1.415	1.046	0.901
Aggravated Assault	1.808	1.544	0.829
Burglary	1.022	1.012	1.295
Larceny	3.341	2.881	3.530
Vehicle Theft	0.550	0.549	0.894

14

Exhibit 3

15

Exhibit 3.

Number of Reported Crimes by Offense for PHA Developments and Surrounding Geographic Rings for February 1998

Offense	PHA Developments	50m from PHA Development	51-100m PHA Development	101-150m PHA Development	151-200m PHA Development	201-250m PHA Development	251-300m PHA Development	Entire City
Homicide	0	0	0	0	0	0	0	1
Forcible Rape	1	1	1	0	0	0	0	24
Robbery	1	4	1	0	4	3	4	169
Aggravated Assault	13	5	2	8	3	7	4	328
Burglary	12	8	8	15	9	13	14	674
Larceny	7	36	19	28	42	15	34	1942
Vehicle Theft	5	4	3	5	6	6	7	281
Arson	0	0	0	0	0	1	1	19

Exhibit 4

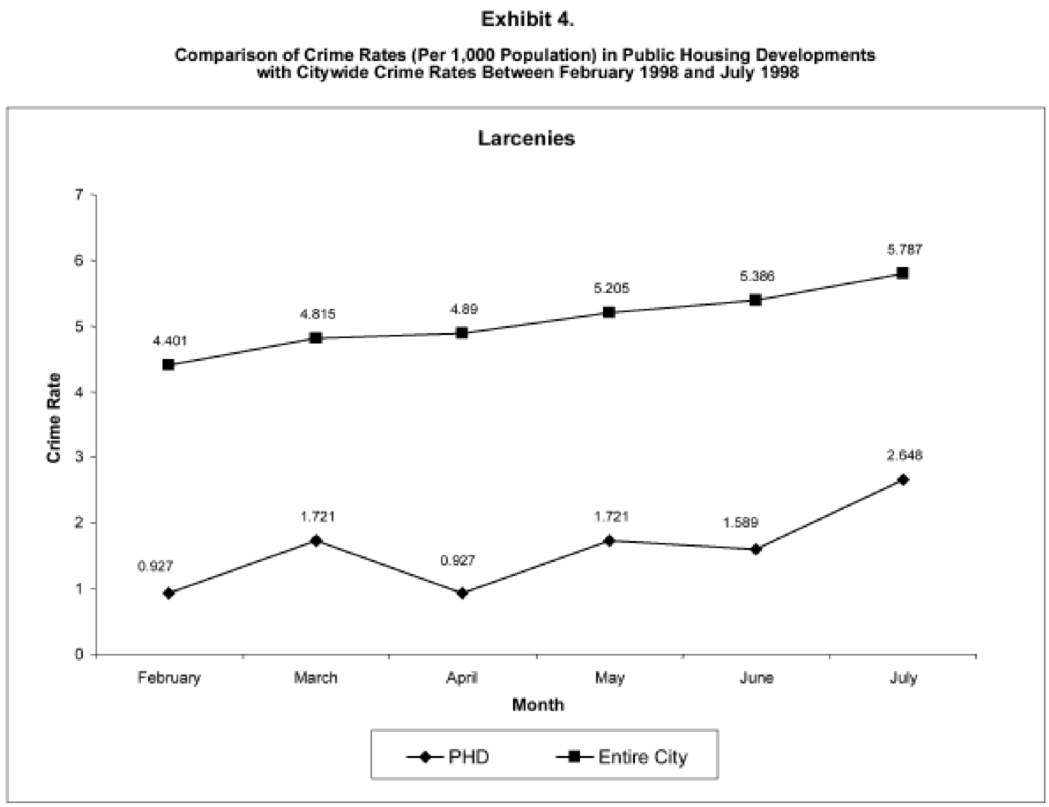


Exhibit 4.
Comparison of Crime Rates (Per 1,000 Population) in Public Housing Developments with Citywide Crime Rates Between February 1998 and July 1998
Larcenies
Crime Rate
Month
February
March
April
May
June
July
4.401
4.815
4.89
5.205
5.386
5.787
0.927
1.721
0.927
1.721
1.589
2.648
PHD
Entire City

3.2 An Alternative to a Solely Development-Based Approach

An analysis based only on crimes reported within a PHA's properties might exclude some important portions of the typical public housing community's living space, namely the streets and shops immediately adjacent to the developments. Interviews with police officers who patrol public housing communities supported the notion that it would be useful to create an expanded version of the development-based approach since these adjacent areas sometimes proved more troublesome from a crime prevention point of view than the PHA properties themselves (see Exhibit 1). Therefore, it was decided to extend each development's GIS polygon out to the immediately adjacent **block face,** which includes the sidewalks and roadway as well as the houses and shops on the opposite side of the street.

Clearly, criminal activity and disorder in the block face can reasonably be seen to affect the residents of the adjacent development. Of course, the block face is not PHA property and is therefore not under the PHA's direct control. Similarly, some crimes such as residential burglary directly involve individual rental units on public housing property and stand apart from a social world that might encompass the block face. Domestic violence would be another such example of a matter largely confined to particular households. Furthermore, the calculation of the development plus block face crime rate is much more complicated than the calculation of the development-based crime rate and will be discussed later. Despite the increased complexity of calculation and the fact that specific types of crimes are most often associated with private living space, the collection of block face data seemed to offer potentially valuable information. In this connection, HUD plans to sponsor research on the very important question of the role that public housing residents play as victims and offenders with respect to crime in the developments themselves and in adjacent areas.

3.3 Multifamily Developments as Distinct Units of Analysis

Development-level crime data are potentially valuable in alerting PHA managers to the existence and nature of emerging crime problems. Furthermore, development-level data allow for establishing baseline information prior to the implementation of a crime prevention program. While first and foremost development-based data are central in yielding "before," "during," and "after" measures for program evaluation, they can also serve the role of "controls." Specifically, the data on developments that are *not* involved in a particular implementation, plus data on the crime levels in the jurisdiction as a whole, offer opportunities to observe naturally-occurring trends outside a program's "treatment" areas. For example, a drop in street robberies may be experienced by *all* the developments and the city as a whole as a direct result of unusually cold weather rather than the implementation of better lighting at several test sites. Conversely, upswings in robberies in untreated “control” developments would help validate declines in crime on better-lighted PHA properties.

Also, development-level data collection can be designed to encompass areas beyond the property. Looking at reported crime in concentric rings around a development has the potential for highlighting nearby “hot spots” that may be associated with current problems or may point to situations that could have a negative impact in the future. The HUD study collected monthly Part I Crime data for six concentric 50-meter rings around each development property (see Figure 1 and Exhibit 5). Thus, the total observation area around each property was 300 meters. In Exhibit 5, the reader will note that block face data do *not* appear since the width of block face rings can vary considerably across developments while the application of increments of 50 meters allow for greater precision in comparing and analyzing crime patterns. In other words, the block face buffer subsumes the roadway, green space, and any idiosyncratic features of the territory between the PHA property boundary and the shops and other buildings immediately “across the street.” Fifty-meter rings offer a more standardized perspective. Furthermore, their true concentric nature allows for additivity across the rings. Exhibit 3 offers summary data for Part I Crimes across the range of 50-meter rings. The reader will note that all eight crimes are being tracked in the summary analysis.

Experience gained in the HUD study suggests a particular approach to counting crimes within the concentric 50-meter ring. It is advisable to query a GIS database as to the number of offenses (e.g., robberies) at a precise distance from a PHA development boundary. For example, the first query might be from the boundary to precisely 50 meters, the next from the boundary to 100 meters, etc. Then, simply by subtracting the counts for successive rings, the 50+ to 100-meter ring is derived by subtracting the 50-meter count from the 100-meter count, and so on.

This approach is advisable because it was found that simply having the GIS count crime incidents in a particular buffer may cause errors in ascribing “buffer membership” when the offense location falls near the buffer boundary. This situation occurs because the buffer “ring” boundaries are not actually smooth curves but a series of short, straight line segments sketched next to one another to approximate the curved buffer boundary. Occasionally, a “crime address” will fall on the “wrong” side of one of these short line segments (see Figure 3). Crime counts at precise distances avoid this problem.

In this connection, however, it may still prove useful to generate a series of concentric 50-meter buffers using the GIS software package. While—for the reason stated above—the within-buffer counts may differ slightly from those

Exhibit 5

Exhibit 5.

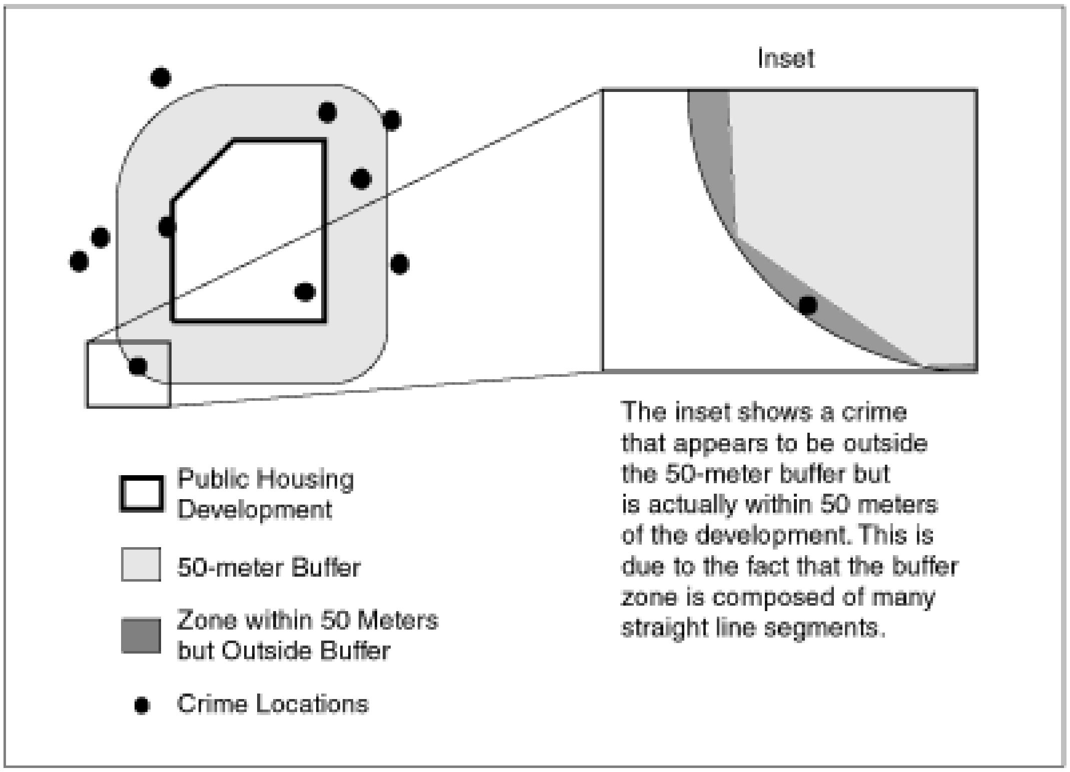
Number of Reported Aggravated Assaults by Individual Development for PHA Developments and Surrounding Geographic Rings for March 1998

Name of PHA Development	Development Population	PHA Developments	50m from PHA Development	51-100m from PHA Development	101-150m from PHA Development	151-200m from PHA Development	201-250m from PHA Development	251-300m from PHA Development
Smith Homes	36	0	0	1	1	0	0	1
John M. McDonald	75	0	0	2	0	0	0	1
Mansfield Apartments	187	0	0	0	0	0	0	0
The Lakes	255	0	0	0	2	0	2	0
Durland House	216	0	0	0	0	0	0	0
Hutchinson Homes	1105	4	0	0	1	0	0	1
Mignot House	51	0	2	2	0	3	2	1
Taylor Hill Homes	177	0	0	0	0	1	1	1
Taylor Hill Homes Extension	192	0	0	0	0	0	0	0
Albertson Homes	1503	2	1	0	1	0	0	0
Stateland	1270	2	0	0	0	0	1	1
Hill Top Apartments	148	0	1	0	0	1	0	0
David K. Johnson	733	2	0	0	0	0	0	0
Hamilton Homes	667	2	0	0	1	0	2	1
Swift Place	102	0	0	0	0	3	7	1
Hedges House	124	0	4	0	1	2	0	0
Randall Julian Mur	63	1	1	0	0	1	0	0

Figure 3

Figure 3.

The Problem with Using Buffers to Count Crimes



20

generated using a command to count offenses at a precise distance, the visual display of crime patterns may be helpful in analyzing "hot spots" for tactical or strategic purposes.

3.4 Tailoring GIS Data to Meet Your Needs

Because of fiscal and time constraints, the HUD study experimented with a limited range of approaches to generating and tabulating crime counts and rates. Potential users of the GIS approach should select and/or create the measures and associated comparisons that best suit their respective information needs and research budgets. For example, it might be useful to include time of day, day of week, and day of month along with the location and type of crime. Such information might bolster efforts to improve security arrangements. Furthermore, in portions of the community where the PHA has an interest, absent the presence of developments (e.g., elderly high-rises, scattered sites or PHA-administered but privately-owned Section 8 rental housing stock), geocoding emergency medical and/or police calls-for-service might prove useful.

CHAPTER 4 - Technical Issues

4.1 Capabilities of GIS

Simply put, GIS was chosen for the task of measuring crime in public housing because this technology can manipulate information in unique ways. However, besides its more exotic abilities, GIS can do some simple but important jobs as well. Like many other programs that generate statistics and pictorial representations of information (e.g., pie charts), most off-the-shelf GIS programs can perform arithmetic functions and even generate tables, graphs, and other types of charts. Hence, in the midst of this guidebook's discussion of **layers** and other GIS-specific terms, it will be useful to remember that crime counts as well as maps can be produced automatically by most GIS software packages. Lastly, before plunging into the more technical aspect of computer mapping, the lay reader might keep in mind that much of the following discussion can be understood as (1) drawing special maps and (2) "cutting and pasting" those parcels that represent PHA properties.

4.2 Base Map Data Layers

Within a GIS, each type of data is organized into specific layers. One could think about these layers as sheets of clear plastic (all being imprinted with the same basic map) upon which special information has been added with colored grease pencils. The base map layers include jurisdictional boundaries; streets; rivers, lakes and streams; and any other layers that are useful for orientation. Layers can also be devoted to population data. The PD/PHA partnership seeks to create a "crime in public housing" layer as shown in Frame 3 of Figure 2. Frame 3 involves two layers. One layer contains the polygons that represent public housing developments and the other layer contains the locations of Part I Crimes. They have been combined in cookie-cutter fashion so that the crime pattern can be seen and analyzed. The layers mentioned below all need to be included in the crime measurement data collection process. There is, however, an important distinction to be made between the "clear plastic overlay" layers and other GIS layers. The points, lines, or polygons representing map features (e.g., crimes, streets, public housing developments) on a clear plastic overlay generally contain little, if any, **attribute** information. That is, there may be a name or a number associated with each map feature, but there is unlikely to be much detailed information about the features on a plastic overlay. In a GIS, however, each feature is linked to a database record that can contain many different pieces of information about each particular feature like the location of a Part I Crime. For example, a point representing a crime location may be linked to a database record that contains all the information in the original police report, such as crime type, street address, victim's name, time, and date. Therefore, a

GIS layer consists of both the map representation of the features and the database, with each feature being linked to a specific database record.

4.2.1 Streets

The **street centerline layer** contains **line segments** representing the locations of streets. Each street segment is attributed with its name, address range, street type, and perhaps other information. In addition to being used for map orientation, this layer is required to convert the addresses listed on police reports into points at the locations of each crime.

4.2.2 Jurisdictional Layer

This layer contains the jurisdictional boundaries including those of the county, city, precincts, and/or other jurisdictional limits for which comparison statistics are to be generated.

4.2.3 Public Housing Development Boundaries

The GIS layer containing the boundaries of public housing developments can be viewed as one data element that is unique to the task at hand. Because it is seldom a standard feature of even the most detailed police maps, care must be taken that this information is accurate and complete with respect to all existing developments — including those under construction or undergoing sweeping redesign. Both the PHA development boundaries layer and the jurisdiction layer will be used in "cookie-cutter" fashion to determine what crimes occur within the respective entity (see Figure 2). The public housing layer is also used to generate new layers containing the boundaries of areas that are within a specified distance from a public housing development. For example, the GIS can measure outward from a public housing development to create the outline of a larger area that lies within 100 meters of that development. The new boundary can then be used to select only those crimes that occur within 100 meters of the development.

As noted above, the local PD will likely not have computer-mapped public housing developments prior to entering into a partnership with the PHA. Because most public housing authorities do not use GIS, and the GIS departments in many cities have not created a “PHA” layer, it may be necessary to create this layer from scratch. Most commonly, the source documents available for layer creation will be hard-copy maps or architectural drawings. Often, these are not referenced to any real-world coordinate system. Usually, however, most of the boundaries are streets, making it possible to quickly extract the relevant streets from the street centerline layer to create the boundaries of the public housing developments. An incidental advantage of this approach to layer creation is that the mapped locations of the development boundaries will be exactly coincident with the mapped locations of the streets that define the boundaries. Furthermore, a small **offset** used during the geocoding process will help ensure that the sites of crime incidents are clearly depicted as being either on or off PHA property rather than being recorded as directly on a street centerline boundary.

Generally, those boundaries of developments that are not coincident with streets need not be precisely located because the address matching process will only generate points near streets. Obviously, the more precise the estimate of the boundary location, the more likely that crimes reported within the boundary of a development will be mapped within the boundary.

The polygons within the GIS layer containing public housing development boundaries should be attributed with the names of the developments and their actual or estimated populations. Assigning arbitrary sequential integers to each development may facilitate automating data processing within the GIS but is not required.

4.2.4 Other Layers

Additional base layers such as lakes, streams, and railroads are not required, but most users find it much easier to use a map that contains layers such as these. Several such layers may already exist within a local government's GIS.

4.3 Converting Address Data into Map Locations

Police crime reports include the physical address of the crime location. Important tasks here are to locate these address data on a map in order to display the map locations and to determine their locations relative to public housing—for instance, whether a crime fell within the boundaries of a public housing development or within a specified distance of a public housing development boundary. Fortunately, the GIS provides the tools necessary to convert the physical address to a map coordinate location.

The process of converting address data to coordinate data is called geocoding or address matching. It relies on the GIS's ability to compare the elements of each address to the attributes associated with each line segment until the elements match (or in the case of the street number, until the number falls within the range of addresses associated with a particular line segment). The exact location is determined by interpolating the available range of addresses to determine where the specific street number lies along the line segment.

The output of the geocoding process is a point layer showing the location of each address in the same coordinate system as the street centerline layer. In other words, the output can be used to create a map showing the crime locations and the streets. The map will show each crime as a point on or near the actual street address of the crime. In addition, each point is linked to any other attribute data contained in the original database—that is, all the other information extracted from the written crime reports.

Geocoding normally requires some amount of data preprocessing to achieve acceptable results. The address given on the police report commonly does not match an address available within the street centerline layer. Reasons may include that the street name does not exist, the name was spelled incorrectly on the report, the street number does not fall within a valid address range, the street type is incorrect, or the street's locational designation (e.g., North Main Street) is omitted or does not exist. Also, police reports often give the location as a name (e.g., county jail) or a street intersection rather than an address. Finally, the report may contain a valid address that is missing from the street centerline layer (e.g., an address within a new subdivision that has not yet been included in the street centerline layer or the street centerline attributes may be incomplete).

Fortunately, the PD is already geocoding crime data for its own purposes and therefore has likely had to wrestle with address problems. In some cases, unmatched addresses are manually researched until a match is obtained (for example, the officer submitting the original report is contacted to ascertain a valid address). Also, the coordinate system being used to geocode crime data must be the same as the coordinate system chosen for the analysis. If it is not, it will be necessary to **project** crime locations from the original coordinate system to the one currently being used for analysis. In most instances, the GIS personnel assigned to this task will be able to detect the problem and perform these conversions. *The fact that geocoding problems are virtually inevitable represents a strong argument for retaining the services of an experienced GIS consultant very early in the life of the PHA/PD partnership. This consultant can then assume the role of an ongoing data processing link between the PD and the PHA; this role was played by RTI in the HUD study. Of course, a tripartite data delivery operation (PD ? consultant ? PHA) will likely be more expensive than relying only on the PD's GIS staff because the PHA must then pay both the PD (data collector) and the consultant (data analyst).*

In those cases where crime locations have not been geocoded for other reasons, the initial match rate of the addresses on police reports may be very discouraging. Usually this is as much a result of a poor street centerline layer as it is of poor address reporting procedures. Both problems need to be resolved to achieve reasonable match rates.

4.4 More on Ensuring Accuracy of Street Centerlines

The importance of having an adequate street centerline map cannot be underestimated. Although it is possible to create a street centerline layer from the U.S. Census Bureau's **TIGER** data, the result is sometimes not accurate enough for geocoding crime incidents. From their inception prior to the widespread use of commercial GIS, TIGER files were never meant to have an extremely high degree of spatial accuracy. The reader should understand that these files can be reasonably accurate but were really meant to convey a *relational accuracy for data management associated with the decennial census.* Hence, TIGER coordinates that define the streets sometimes may prove troublesome, making it difficult to *accurately* map those addresses that do match (i.e. matching an address to a point along an inaccurately mapped street results in a location that is inaccurate).

Alternatives to using the original TIGER data layers include purchasing commercially "enhanced" TIGER data (layers), improving the spatial and attribute accuracy of the TIGER data (either in-house or through contracting), generating and attributing a street centerline from other existing sources (e.g., engineering maps or **orthophotomaps**), and adjusting a TIGER layer to other sources. Given the widely recognized value of precision here, the costs associated with the creation of an accurate street centerline layer sometimes may be shared by other potential users within local government. Attempts to use an inadequate street centerline layer are inevitably plagued by inaccurate results.

4.5 Determining How Many Crimes Were Reported In and Near Public Housing Developments

Counting the number of crimes that were reported within a specific area is accomplished within the GIS through point-in-polygon overlay (see Frame 3 in Figure 2). That is, the crime points are overlaid by the polygons representing the area(s) of interest (e.g., one or more public housing developments) to create an output layer containing only those points that fall within the polygons. The procedures necessary to accomplish this vary from one GIS to another, but the results are similar: the output layer excludes points that lie beyond the boundaries of the input polygons.

The total number of crimes of each type that were reported within the polygons used during the overlay process can be determined by selecting the crime type. The details of the process are specific to the GIS being used, but any good GIS will have that capability.

If the data analysis includes 50-meter concentric rings beyond the development property line, situations may arise where the rings emanating from two neighboring developments intersect. When this occurred in the HUD study, the decision was made to assign crimes to the closest development. This strategy avoids double counting of crimes in the calculation of rates for the PHA as a whole.

4.6 Obtaining Data for the Population Layer

To calculate crime rates, it is necessary to estimate the total population of the area of interest in addition to the total number of crimes of each type reported within that same area. Public housing authorities have fairly precise population data for individual developments. But population estimates beyond the boundaries of the developments need to be derived from another source.

City and county population data are routinely available from the U.S. Census Bureau. This information may come directly from a recent decennial census or from the Census Bureau's biennial estimate population for cities (even years) and counties (odd years). Current population estimates for smaller geographic units, such as census tracts, may not be easy to find. Occasionally, a city or county may have relatively current detailed population data available on geographic areas such as census tracts and police reporting districts. Private data vendors who provide information to the survey research industry also sometimes have constructed demographic databases for varying types of geographic subdivisions.

Note: If there is no plan to calculate crime rates other than those for local jurisdictions and PHA properties themselves, much of the remainder of this chapter may not apply. The reader may be assured, however, that other highly informative and potentially useful material lies ahead.

Obtaining population data for the **block face buffers** around PHA properties will likely be more expensive than city or county data. The U.S. Census Bureau calculates population for individual blocks (termed "census blocks") and for geographic units termed "block groups," which, in general, contain roughly 10 residential blocks. Population information for these blocks and block groups is updated by the Census Bureau once every 10 years following a decennial census. Therefore, unless a decennial census has recently been conducted, census block and block group population information may be somewhat imprecise. If current census block and block group data are not available, one could purchase such information from a private vendor who has fashioned more up-to-date populations estimates using a variety of social indicators such as birth and death rates, employment rates, school enrollments, and the like.

4.7 Estimating Population Counts for Calculating "Public Housing Precinct" Crime Rates

Because the database associated with the public housing developments layer includes the population of each development, these data can be used to calculate crime rates within the developments taken together (the development-based “precinct”) or within any particular development. Thus, calculating crime rates for the development-based public housing *precinct* is straightforward. However, blockface population estimates must be derived with the aid of the U.S. Census Bureau population layer or some other more recent data source.

In calculating the population of an area that lies entirely outside of public housing developments such as the blockface, one makes the tacit assumption that the residents within each census block or block group (depending on the specificity of the census data in one’s possession) are evenly distributed within that geographic area. In reality, highways, parks, and schools account for sizable chunks of many neighborhoods and the land actually devoted to homes may be tucked into a relatively small space that is much more densely populated than the overall density of the area suggests. In any event, it is obvious that not all of the space in block faces or in the other individual concentric rings is devoted to people’s homes. Hence, this population estimation procedure is affected by land use patterns in a blockface and the general pattern of land use surrounding public housing developments. Therefore, the accuracy of this approach can vary by neighborhood and thus influence the calculation of aggregate crime rates for a “development plus block face” public housing precinct.

4.8 Calculation Procedures

The assumption of uniform population density within block groups allows calculation of the population density of each block group. The process used to generate the population estimates for areas that lie partially or entirely outside of public housing developments is similar to counting crimes but uses polygon-on-polygon (*vs*. point-in-polygon) overlay. In this type of overlay, the GIS is used to create an output layer containing the polygons only in those areas that contain polygons in both of the input layers (i.e., the census block or block group polygons and the block face or **buffer** polygons).

This new layer will generally consist of several contiguous polygons within each block face or buffer. These smaller polygons are the block or block group polygons, or portions of these polygons, that were coincident with the block face polygons or buffers. The population density of each of these smaller polygons is used along with its area to estimate its population. The sum of these population estimates provides the population estimate of the whole block face or buffer.

For purposes of this example, one could assume that the areas of interest are the areas within 100 meters of any public housing development in a particular jurisdiction, *excluding* the developments themselves. The first step is to create a polygon layer that encloses all those areas lying within 100 meters of a public housing development boundary. This is a "buffer" operation and is part of the functionality available with a complete GIS.

This 100-meter buffer layer is one of the inputs to the overlay process. The other input is the population layer. Because the population layer covers the entire jurisdiction, the output layer will contain polygons over the same area(s) as the polygons in the 100-meter buffer layer. However, any population polygon bisected by the boundary of a 100-meter buffer will be represented by a smaller polygon in the output layer than it was in the input population layer. In other words, the parts of the population layer that lie more than 100 meters from any development will be missing from the output layer.

Next, because the housing authority’s estimate of the population of each development is probably better than any estimate obtainable from the Census data, the public housing development boundary layer is used to "erase" those polygons and parts of polygons that lie within a development boundary. This is done whether the desired population estimate is for the entire area within 100 meters of a development (including the developments themselves) or only those areas that are within 100 meters of, but outside, the developments. In the first case, the known population of the developments is added back into the estimate of population in the surrounding area to produce a better estimate of population than would be derived by using the population layer data only.

The GIS is then used to calculate the area of each output polygon. With this information, it is possible to estimate the populations of the output polygons by multiplying their areas by their population densities and summing the results. The accuracy of the estimate depends on the validity of the implicit assumption that the population density is uniform across the input polygons. Although this will almost always be an invalid assumption, the input polygons are generally small (meaning that there will usually be at least several represented in the output coverage) and the non-uniformity of their population densities is non-systematic. In other words, the errors will probably cancel each other, resulting in reasonable estimates for the aggregate total population. Alternative approaches that might result in more accurate estimates of block face or buffer zone populations are suggested at the end of this section.

The aggregate total population is the sum of the population estimates for the contiguous polygons surrounding the developments in the output coverage. Any GIS will provide a method of summing the population estimate for each small polygon to obtain totals for the areas extending from the development boundaries to the 100-meter buffer. The population of all areas outside developments (but within 100-meters of the development boundaries) can be obtained by summing the population of the isolated buffer zones, and the development populations can be added in those areas.

4.9 Alternative Approaches to Calculating Block Face Populations

The general approach for estimating block face populations, presented earlier, can be modified to achieve greater accuracy. However, the modified approach involves estimating the population of each block face or ring individually and therefore is more time-consuming and may be more expensive as well. Local PD and PHA staff will know the land use patterns around each development. It might be possible for staff to attempt rough counts of housing units and then apply any other relevant information such as average number of residents per unit. In areas where large proportions of the block face are unoccupied, such as park lands or highways that abut the development, the estimation process would be less complicated. This approach is sometimes referred to as the “housing unit” method of population estimation.

In any estimation procedure, the availability and currency of population data will determine accuracy. In some jurisdictions, reasonably current block-level data might be available from the local government and thus could become the population layer for the GIS approach detailed above. As a general rule, PD and PHA staff would do well to confer with other local government departments if population estimation procedures are planned.

CHAPTER 5 - Technical Assistance Resources

The National Institute of Justice (NIJ) is one of the U.S. Department of Justice's most important research components. Several years ago, NIJ established the Crime Mapping Research Center (CMRC). This NIJ component sponsors professional conferences and training sessions on crime mapping as well as funding cutting edge research on the application of GIS technology to the needs of the criminal justice community. Information on CMRC activities may be obtained from NIJ's Web site (www.ojp.usdoj.gov/nij). Furthermore, materials from recent CMRC conferences may help PHAs and PDs identify potential sources of consulting help such as GIS researchers at local universities or PDs in nearby communities that are also engaged in the business of crime mapping.

A good many books on GIS are now available. Below is a small sample of books that may prove useful in furthering GIS users' familiarity with crime mapping and associated analytical issues.

LaVigne, Nancy and Julie Wartell, Eds. (1998), *Crime Mapping Case Studies: Successes in the Field*, Police Executive Research Forum, Washington, D.C.

Weisburd, David and Tom McEwen, Eds. (1997), *Crime Mapping and Crime Prevention,* Criminal Justice Press, New York.

Eck, John E. and David Weisburd, Eds. (1995), *Crime and Place,* Crime Justice Press and the Police Executive Research Forum, New York and Washington, D.C.

Block, Carolyn R., Margaret Dabdoub and Suzanne Fregly, Eds. (1995), *Crime Analysis Through Computer Mapping,* Police Executive Research Forum, Washington, D.C.

GLOSSARY

Attribute: A data value associated with a particular feature in a GIS layer—for example, the name associated with a particular street. Each feature in a GIS layer (each point, line, or polygon) can have its own, unique set of attributes.

Block Face: The properties facing the streets defining the boundaries of a public housing development. With respect to crime counts, the sidewalk and street between the PHA property and the actual block face properties are also included (see Block Face Buffer).

Block Face Buffer: A polygon surrounding and enclosing a public housing development that is designed to enclose the block face and any additional properties located within one-half block of a public housing development.

Buffer: A polygon whose boundary is defined by measuring outward a specified distance from some other polygon boundary. The distance, with respect to a public housing development, may be variable (e.g., one-half block) or it may be a fixed distance in meters.

Geocoding: Sometimes called address matching, geocoding is the process of assigning map coordinate locations to addresses in a database. A GIS is capable of doing this by comparing the elements of each address in the database (e.g., the number, prefix direction, street name, suffix direction, and type) to the attributes of each line segment in a street centerline layer to find a match. The starting and ending address of the matched line segment is used to locate the precise coordinate position. The output of the geocoding process is a point layer attributed with the crime data from the input database.

Layer: A GIS database containing point, line, or area (polygon) features representing a particular class or type of real-world entities such as crime locations, streets, or police precincts. A layer contains both the visual representation of each feature and a link from the feature to its database attributes.

Line segment: An unbranching line. A line segment can be used to represent the location and shape of a street extending from one intersection to the next, but representing the streets extending outward from those intersections would require different line segments. In a GIS, each line segment can have its own set of attributes (e.g., name, beginning address number, or ending address number).

Offset: The ability of software to place a crime location within a property boundary rather than directly on a street centerline. This process reduces the confusion that arises when a crime location is geocoded as lying directly in the middle of a street. Offsetting clearly shows that geocoded crime locations are near (but not on) streets and that a crime may be located “across the street” from another crime.

Orthophotomap: A geometrically corrected photographic image. Orthophotomaps reduce the scale, direction, and shape distortions inherent in aerial photographs to those associated with maps.

Polygon: A closed area (or the boundary of a closed area).

Project: Here, used as a verb. To change the underlying coordinate system describing map locations so that features originally mapped with one coordinate system can be displayed correctly in their relative locations using another coordinate system. The map itself never changes (e.g., downtown Dodge City still looks the same) but the coordinate system that pinpoints locations has changed. Since the GIS software does all the work, the user is not involved in the actual projection process with all its complex mathematical operations.

Public Housing Development: A PHA-owned housing development. Specifically excluded here are (1) Section 8 privately-owned residences and developments and (2) PHA-owned scattered-site housing.

Street centerline layer: A GIS database containing lines representing streets (along with the attributes of each line segment, e.g., name, address range, etc.). The street centerline is used to represent the streets (instead of a pair of lines representing the edges of the pavement) because the GIS needs only a single line to represent the street and link to each street’s attribute information. Additionally, the GIS user who wishes to make a map showing the streets can select from a wide variety of symbols, including single and double lines, to categorize streets based on their attributes.

TIGER: Refers to Topologically Integrated Geographic Encoding and Referencing, the GIS products produced by the U.S. Census Bureau. TIGER data include, among other things, street centerline layers and population layers.