Document Title: Identification, Development and Implementation of Innovative Crime Mapping Techniques and Spatial Analysis

Author(s): John Mollenkopf ; Victor Goldsmith ; Philip McGuire ; Sara McLafferty

Document No.: 198475

Date Received: January 2003

Award Number: 97-LB-VX-K013

198475

FINAL REPORT
JULY 1, 1997 TO JUNE 30, 1999

PROJECT 97-LB-VX-K013

IDENTIFICATION, DEVELOPMENT AND IMPLEMENTATION OF INNOVATIVE CRIME MAPPING TECHNIQUES AND SPATIAL ANALYSIS

TO:
DR. NANCY LA VIGNE
NATIONAL INSTITUTE OF JUSTICE
US DEPARTMENT OF JUSTICE
WASHINGTON, DC

FROM:

PROFESSOR SARA MCLAFFERTY
DEPT. OF GEOGRAPHY
HUNTER COLLEGE, CUNY

PROF. JOHN MOLLENKOPF
CENTER FOR URBAN RESEARCH
GRADUATE SCHOOL AND UNIVERSITY CENTER, CUNY

PROF. VICTOR GOLDSMITH
CENTER FOR APPLIED STUDIES OF THE ENVIRONMENT, CUNY
DEPARTMENT OF GEOGRAPHY, HUNTER COLLEGE, CUNY

ASST. COMMISSIONER, PHILIP MCGUIRE
OFFICE OF MANAGEMENT, ANALYSIS AND PLANNING
NEW YORK CITY POLICE DEPARTMENT

March 10, 2000

FINAL REPORT

The objectives of Grant 97-LB-VX-K013 were to explore, analyze and develop innovative tools for crime mapping and analysis, and to incorporate selected tools in a Crime Mapping and Analysis Application (CMAA) that can be used by the New York City Police Department at both the COMPSTAT and precinct levels. The grant was a collaborative effort involving participation and interaction between NYPD's Office of Management, Analysis and Planning (OMAP) and researchers at the City University of New York. Throughout the project we emphasized the importance of creating an application that is understandable for non-specialists, easy-to-use and fits in NYPD's existing crime mapping and analysis efforts. The CMAA is finished and is currently being tested and refined at the NYPD with the assistance of members of the CUNY team. We view the CMAA as a work in progress that will be enhanced and updated as it is used at NYPD. The NYPD is currently negotiating with GIS vendors to develop an enhanced crime mapping system, and we expect that this new system will incorporate several of the tools that were developed in our CMAA. Thus, in the long run, we anticipate that the CMAA will evolve into a system that is fully integrated with a commercial GIS and that is an essential component of NYPD's GIS/mapping efforts.

The project was divided into two phases. During the first year of the project we developed and tested a series of methods for spatial analysis and mapping of crime patterns. That work was described in progress report for phase I of the grant, which was submitted to NIJ in 1998. At the end of the first phase of the project, we chose four tools for implementation in the CMAA. The tools are: block aggregation, kernel smoothing, Voronoi diagrams/coverage curves, and animation. The four tools are discussed in detail in Appendix A of this report. In addition to developing the CMAA, our team presented the results of our collaborative work to the crime mapping community in publications (see Appendix B -- List of Publications) and presentations at national and international crime mapping conferences (Appendix C – list of presentations).

How were the four tools selected? First and foremost they were chosen to meet NYPD's needs for clear, understandable and easy-to-use mapping tools. NYPD's input, especially guidance from Assistant Commissioner Phil McGuire of OMAP, was crucial in this process. Many theoretically sophisticated methods were rejected because they were 'too difficult to explain and understand' or because the 'insights they provided would not be very useful in NYPD decision-making.' The result was four fairly simple tools that seemed to be especially promising. In February 1999, we demonstrated the tools to Commissioner Safir and other high-level NYPD officials. The response was very positive (Appendix D – Letter from Safir).

The second phase (Year 2) of the project involved creating the Crime Mapping and Analysis Application. The CMAA incorporates the four mapping tools along with basic querying and mapping capabilities, and it is designed for use in the COMPSTAT process and for precinct-level mapping and analysis. Several general considerations influenced the design and development of the application:

- The CMAA had to be clear and easy-to-use for both crime analysts and non-specialists
- The CMAA had to rely on the NYPD's current GIS/mapping system, MapInfo, so that it can be used in precincts and system-wide.
- The application should be well-integrated with the NYPD's existing Crime Analysis Package (CAPS) and involve similar kinds of querying and selection capabilities.

Each of these parameters greatly influenced the design and features of the CMAA. The programming and initial testing of the CMAA were done at CUNY, Hunter College. The application is now being tested and enhanced in the Office of Management, Analysis and Planning (OMAP) at the New York City Police Department. The CMAA is described below.

Overview of NYPD Crime Mapping and Analysis Application

Introduction

The purpose of the NYPD Crime Mapping and Analysis Application is to give members of the NYPD, specifically crime analysts and the COMPSTAT unit, the ability to analyze crime patterns and generate several types of maps in a user-friendly environment. The application allows users to query the NYPD's COMPSTAT data and perform four different types of mapping and spatial analysis. Specifically, the application: 1) performs 'block aggregation' to generate choropleth maps of incidents aggregated to several types of geographic units (census blocks, police sectors, police precincts, etc.), 2) creates smoothed density maps of crime incidence using kernel estimation, 3) prepares Voronoi diagrams (Theissen polygons) of individual crimes and performs a coverage analysis based on the areas of Voronoi regions, and 4) creates map animations to show changes in crime patterns over time. This section describes the overall Crime Mapping and Analysis Application (CMAA), noting the challenges faced in creating the application. Afterwards, each of the three 4 tools is discussed in depth

Although the NYPD Crime Mapping and Analysis Application incorporates several complex spatial analysis methods, the goal was to create tools for rigorous display of distribution and change in crime patterns within an easy-to use system aimed at non-specialists in NYPD.

Wherever possible, technical details were hidden from the user, parameters were 'hard-coded', and options were presented as clearly as possible. The application utilizes MapInfo, currently the NYPD's main mapping tool, used in both the COMPSTAT process and in precinct-level analysis and decision-making.

The application was developed in Microsoft Visual Basic 6.0 and relies heavily on MapInfo's MapX (v.3.5.2) Software Developer's Kit (SDK). The decision to use MapX was a significant one. The NYPD felt MapX was the optimum solution primarily because its employees were already familiar with MapInfo and would find it relatively easy to grasp the new environment. The CMAA also utilizes several other important applications such as MapInfo Professional, MapBasic, Vertical Mapper SDK, and Microsoft Excel. A major challenge in developing the application was to link all these systems and join them in an easy to use, seamless application (Figure 1).

In its simplest form, the application consists of four Visual Basic forms or windows: the main parent window, the selection/query window, the map window, and the tool window. There are a number of other minor windows that perform a variety of smaller tasks, such as confirming a query, selecting layers to operate on, changing the appearance of layers, etc. Each of these four windows will be discussed here in detail.

The Main Parent Window

Most users who work in the Windows environment will be familiar with this window. This is the main window where all other smaller windows reside. The primary function of this window, aside from acting as a parent to all the 'child' windows, is to give the user some basic functionality common to most Windows programs (see Figure 2). For example from the File pull-down menu, users can exit the application, print maps, export maps and save layers. From the Select menu pull-down, users can open the query builder window, or run existing queries that have been saved from an earlier session. The final menu pull-down, the Window pull-down allows users to open or bring to the front any of the windows that are present.

The Selection/Query Window

The primary function of this window is to allow the user to create a subset of the COMPSTAT data by performing a query based on SQL (Structured Query Language). From this window the user can query the data based on patrol borough, precinct, jurisdiction, type of crime, date of incident (or report), and time of incident, as well as combinations of these attributes (Figure 2). The window also allows the user to save a query. This is particularly useful when the same query will be used a number of times. The specifications for this window came from extensive discussions with the NYPD. They envisioned this phase of the analysis to be similar to an older application that offered a flexible environment for the creation of queries. This older application, which was created in-house at the NYPD, was known as the Crime Analysis Package (CAPS). Its primary functions were the querying and reporting of COMPSTAT data, but it did allow for the creation of simple pin maps. The new CMAA gives the user the same power and flexibility in querying, but with an enhanced set of mapping and analysis tools.

The Map Window

This window is the heart of the application. It is here where the maps are displayed and more importantly, where the users can specify which type of map they would like to generate (see Figure 3). The mapping 'engine' utilized is MapX SDK v.3.5.2 by MapInfo. In general, MapX SDK allows programmers to deploy mapping applications utilizing technology from the popular Geographic Information System (GIS), MapInfo Professional. MapX does not have all the functionality of MapInfo Pro. Instead, it incorporates the most common attributes, such as creating dot maps and thematic maps, for example.

The limits imposed by MapX's functionality are important for two reasons. First, since functions and commands are limited, users do not have to spend a lot of time familiarizing themselves with the software. Therefore, the learning curve is not nearly as steep as it is with a full-fledged GIS, so resources do not need to be squandered on training. On the other hand, since MapX lacks functionality, it is not nearly as powerful as a full-blown GIS. Because of this, many of the tasks and tools that one would normally expect from a GIS are not present. For example, in the version of MapX used by the application, users can not directly query native MapInfo tables (.tab files). Instead, the programmer must use a different query environment such as Open Database Connectivity (ODBC), or in this case, MapBasic scripts to perform the query in MapInfo.

The four types of analysis, or mapping that can be performed in the application are block aggregation, density estimation, voronoi diagrams, and animation (see Appendix A for details of each). The block aggregation method is probably the simplest to comprehend. In this case, a map is generated in which geographic units are shaded based on the number of incidents within them. The name, block aggregation, is a misnomer of sorts because any geographic unit can be used, not just blocks, but census tracts, police sectors, police precincts, or any other polygon layer in the application (see Figure 4). This type of analysis is a useful tool for data mining. It enables the user quickly to determine which areas have a high incidence of crime and allows them to 'zero in' on those areas and perform further analysis. It is also useful for creating tables that show counts of crime by area and changes in crime incidence over time.

The block aggregation portion of the program depends completely on MapX. This is not to say that it was an easy task, quite the contrary. Since this type of analysis – overlaying polygons on points and counting the numbers of points per polygon -- is fundamental to any GIS, one would expect it to be quite simple. However, this functionality, point in polygon analysis, is not built in to MapX. Therefore, using Visual Basic and MapX, a workaround was created in which each polygon was selected and the number of points within it was calculated. Once this information was known, an array was created which contained each polygon ID and its associated count. This array was then 'bound' to the polygon layer so that a thematic map could be prepared based on the count.

The second type of analysis, kernel density estimation, is slightly more complicated, theoretically and practically. Density estimation computes the number of crimes per unit area from the dot map of crime events. The result is a smooth surface of density values (crime per unit area) which show how the density varies over space. This is useful, because unlike block aggregation, the analysis is not limited to some arbitrary geographic boundary, and it is much easier to discern patterns than on a complex point map (see Figure 5).

From a practical standpoint however, this method was rather difficult to implement. The main reason is that neither MapInfo nor MapX supports this type of analysis. Neither product has the functionality for creating grids or continuous data surfaces, and such grids are an essential output of kernel estimation. Therefore another product was required, namely Vertical Mapper SDK. Vertical Mapper SDK is similar to MapX, except that it is the programming environment for the Vertical Mapper software from Northwood Geosciences, Inc. At the time this project was undertaken, kernel density estimation had not yet been implemented into either Vertical Mapper or Vertical Mapper SDK. We discussed this situation with people at Vertical Mapper, and they expressed interest in incorporating kernel estimation as a tool in VM, using VM SDK. The team at CUNY worked closely with developers at Northwood Geosciences to create the algorithm to perform kernel density estimation. This involved a lot of additional time in learning a new programming environment as well as validating, testing and debugging the algorithm they provided. Since the VM developers were unfamiliar with kernel smoothing, there were many errors, and we spent a great deal of time trying to determine if the algorithm was working correctly. Based on our collaborative efforts kernel density estimation is available as an option in new releases of Vertical Mapper.

As currently implemented, the CMAA performs kernel density estimation by calling the new procedure in Vertical Mapper. The algorithm in VM uses a quartic kernel density function as the default option. Default bandwidth and grid cell size values are provided in a dialog box in the application, but knowledgeable users can vary both the bandwidth and grid cell size within the dialog to create the most effective density surface.

The third method, Voronoi analysis, was also rather complicated. Voronoi diagrams, or Theissen polygons, are regions drawn around points such that each region is the smallest area that can be drawn around its corresponding point. The region around a point includes all areas that are closer to that point than to any other point (see Figure 6). As with the two previous methods, neither MapInfo nor MapX supports Voroni analysis; however, it is available in Vertical Mapper SDK so an additional link to Vertical Mapper was created.

There were several challenges to using Vertical Mapper for the Voronoi analysis. Vertical Mapper requires that there be no duplicate points. That is, two or more crime points cannot occur at the same location. This is significant when dealing with the incidence of crime, because it is well known crimes often occur at the same address or location, particularly when one is analyzing crime data over longer time periods (months and/or years). This is especially true in densely populated cities like New York where it is common to find as many as ten crimes at the same address, in part because of the large number of multi-story apartment buildings that include many residences at a single street address. Deleting duplicate points was not an option because that would result in a loss of data, and an underestimation of incidence. To overcome this problem, a script was written in MapBasic for dispersing duplicate points. The script takes a layer of points, checks for duplicates, randomly disperses any duplicates around the initial point and then writes out a new point layer. A point is dispersed by randomly varying the last digit of its geographical coordinates, which moves the point a small distance, no more than 60 feet for the current New York City base map. The point dispersal application calls MapInfo, runs the script for dispersing points, returns control to the application, then calls Vertical Mapper to create the Voronoi polygons. Transferring control was quite tricky, but was finally resolved using OLE (Object Linking and Embedding). OLE gives programmers the ability to place common applications within another application. For example, in Microsoft PowerPoint, the user can imbed a Microsoft Excel chart in a presentation, even though Excel is not running.

An important component of the Voronoi procedure is the ability to perform coverage analysis. As discussed later, coverage analysis uses the results from Voronoi analysis to determine the trade-off between the number of crimes and the amount of area covered by the corresponding Voronoi regions. Coverage analysis allows the user interactively to select the percentage of crime to be analyzed, and then it identifies which Voronoi polygons cover this percentage of crimes and the area encompassed (see Figure 7). This was a complicated method from a programming standpoint because the functionality of sorting and cumulating Voronoi polygon areas does not exist anywhere in MapInfo or Vertical Mapper.

To accomplish this, Microsoft Excel was used. Excel was chosen because of the ease with which many of the tasks could be performed, such as sorting and calculating new values based on values in other cells. A macro was written in Visual Basic for Applications (VBA) which carried out all the necessary tasks. This script takes the resulting layer from the Voronoi analysis and performs all the required Excel functions. Then based on the user input, the specific Voronoi polygons are identified. Once they are identified, MapInfo is called. There the polygons are selected based on percentage of crimes and a new layer is created that contains only those polygons. From here, the layer is returned to the application (and MapX) so that the selected polygons can be shaded and displayed on a map. The user chooses the desired percentage of crimes to be covered by moving a sliding bar along the coverage curve to the desired percentage. Then, polygons are selected and the map is drawn.

The final tool to be added to the CMAA was animation. This tool results in the creation of an animated sequence of density maps. The resulting animation can be viewed from the application itself, or by using a standard Windows media player. The final file format of the animation is a Windows video file (.avi). To create the animation, the user is prompted to perform a selection similar to the one described above, but with fewer options for querying. Only certain variables can be queried, such as precinct, crime type and date. The user is also asked to define a 'time window' that specifies the duration of time to be represented in each frame of the animation, i.e., five days, seven days, etc. They are also prompted for an incrementor that specifies how much each animation frame will be advanced by. This in essence determines if the animation will show a moving average or not. For example, if the time window is seven days and the incrementor is one day then there will be six overlapping days in the animation.

Once the user specifies all the necessary parameters, the application begins querying the data, creating the maps, and exporting the maps as images. Once all the images are created, a shareware product known as VFD (Video for DOS) assembles the frames (images) into one seamless video file. The process can be CPU intensive as well as memory intensive.

The animations themselves show density maps, rather than just point maps. This was done because visually, it is extremely difficult to discern a pattern in point maps over time, especially in a high-crime density area like New York City. Changes in the size, shape and location of hotspots are easily seen on the smooth density maps. The application also allows the user to incorporate or draw a cosmetic layer that is displayed on top of all images in the animation. This can be used with any spatial data layer -- bus stops, housing projects, subway lines, etc. to provide a geographical reference base during the animation. One can also do a density or Voronoi analysis of crime data for a long period of time and then draw a hotspot based on the historical analysis. Using that historical hot spot as the cosmetic layer for the animation, makes it possible to view the changes in crime density patterns over time in relation to historic areas of high crime intensity.

In addition to the four tools described above, there are a number of other features within the Map Window. From the map window, the user can call dialogs to change the appearance of layers, set layers to visible or invisible, or change the drawing order of layers. The user can also exit the application, or reset the map to its initial state. Also on the map window are a status bar, which relays messages to the user and a progress bar, which shows the amount of time left for a process.

The Tool Window

The last window to be discussed is the tool window, which is familiar to MapInfo users. It contains the basic tools and functionality required for navigating in a map window. It includes tools for zooming in, zooming out and panning (see Figure 8). It also contains a tool for labeling features and adding text as well as standard tools for selecting features.

Requirements for Running the Crime Mapping and Analysis Application

The CMAA was written using Microsoft Visual Basic 6 on the Windows NT v4.0 (Service Pack 4) operating system. It is recommended that the application be used on a machine configured similarly. Although it has not been tested, the application should run on any 32-bit operating system such as Windows 98. The application utilizes MapInfo and therefore requires MapInfo and related software in order to operate. The CMAA is dependent on the following: MapInfo v5.0, Vertical Mapper 2.1, MapInfo MapX v4.0 and Microsoft Excel 97. It is recommended that all these programs be installed before attempting to install the application. Although not tested, it may be possible for the application to run without MapX; however, again, this is not recommended since it has not been tested. The application can be installed by running the setup.exe file. This will guide the user through the installation process. The installation program will create a number of directories that contain the application, the associated data, and the source code. All necessary registry entries will be made by the setup.exe program. One new directory will need to be created by the user, if it does not already exist, namely c:\temp

Conclusion

The Crime Mapping and Analysis Application represents a substantial advance over the New York City Police Department's previous capacity to create computerized point maps of crime locations. It offers a series of technically sophisticated, yet easily used, tools to understand when and where crimes are clustered in space, how those clusters change over time, and how a given level of police resources might be allocated to cover the largest amount of crimes.

As the NYPD implements CMAA within its crime analysis procedures, the project team will monitor how different types of users react to the different functionalities of the package. This will enable us to identify and correct any problem areas as well as to determine whether additional training might help the NYPD to realize the full potential of the application. It will also be important to determine, at some future point, the extent to which this technological innovation has helped the NYPD continue its remarkable record of reducing the level of crime reported in New York City.

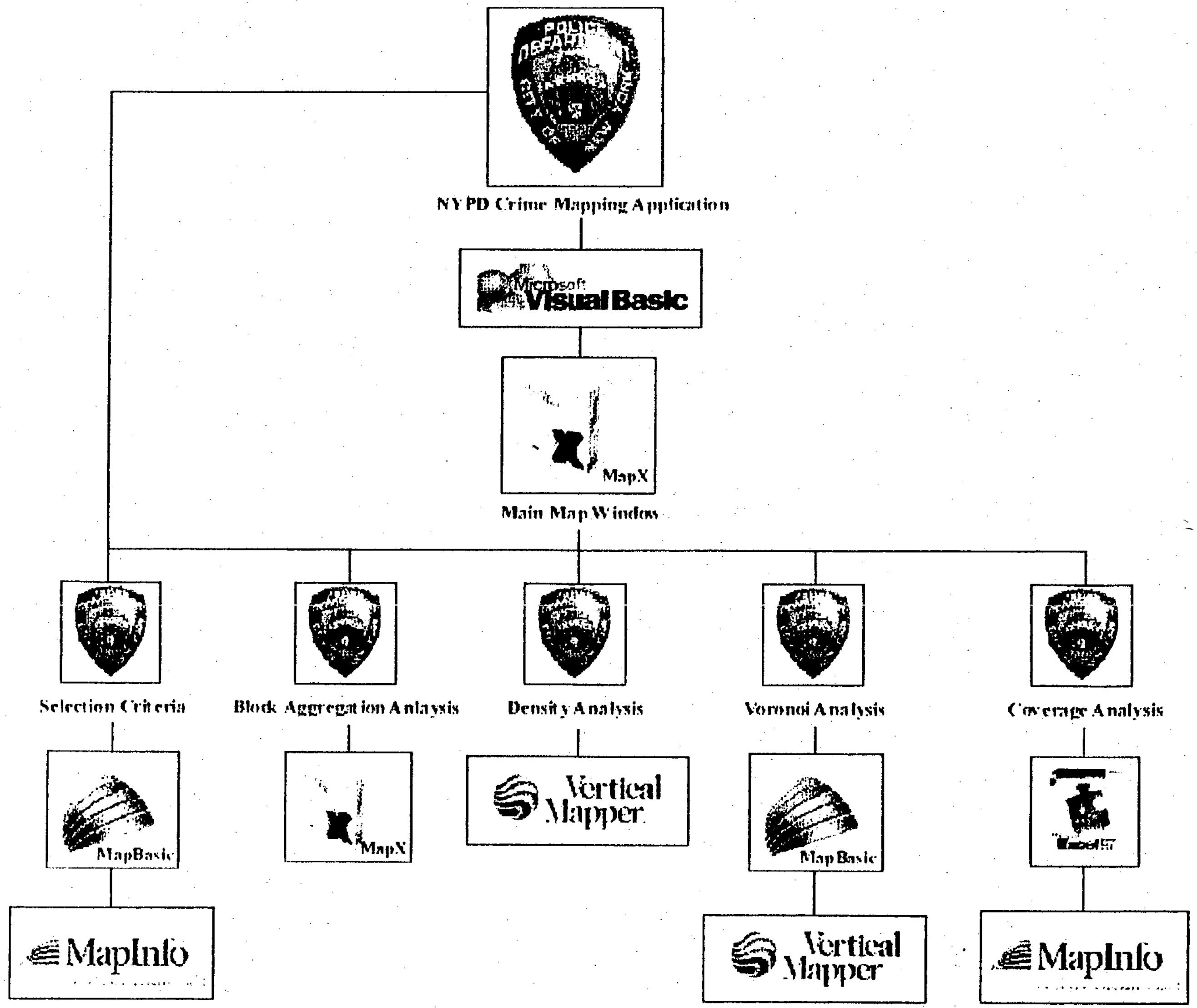

Figure 1: Organizational schematic of NYPD Crime Mapping Application

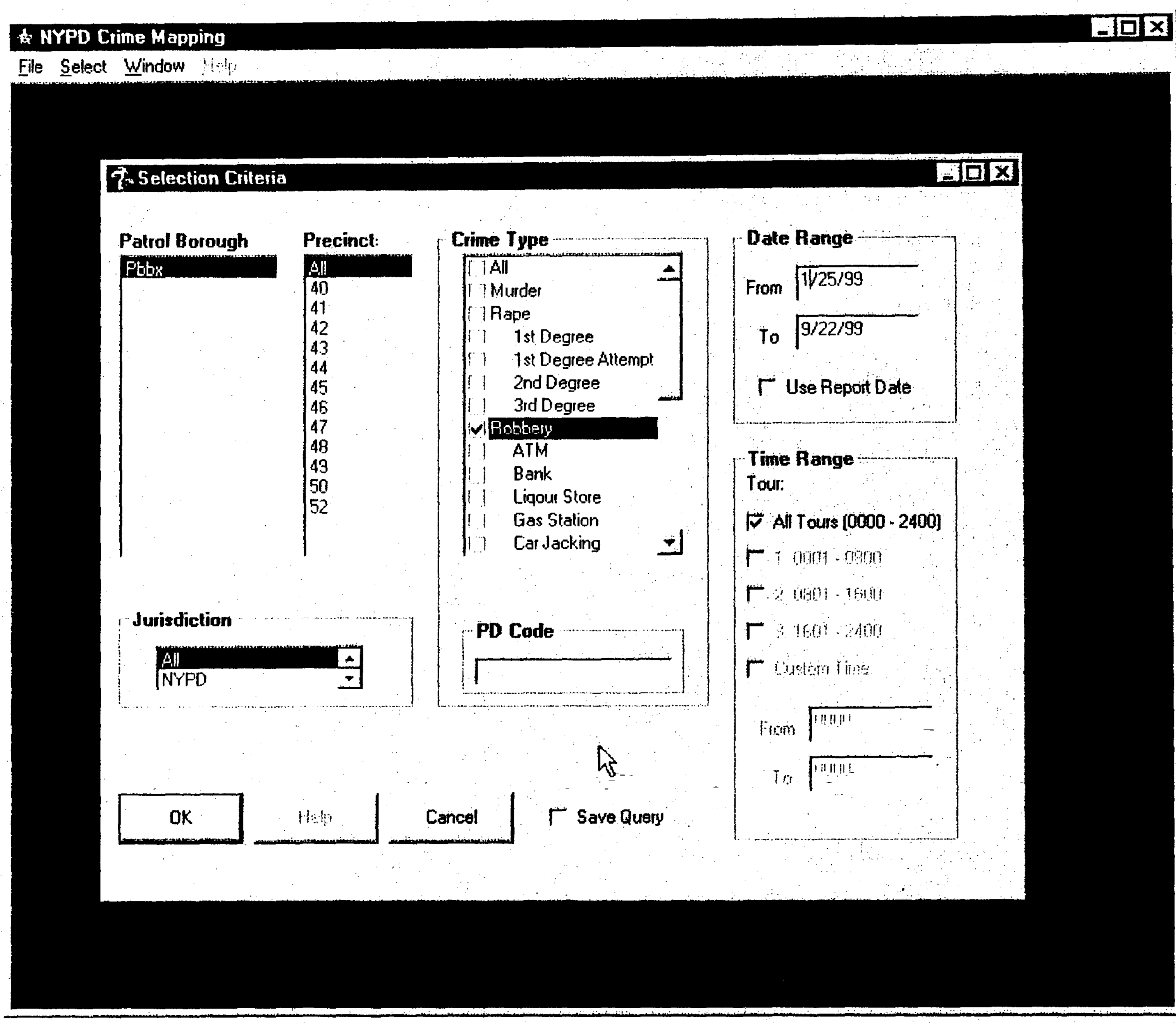

Figure 2: Main/Parent window and query builder window

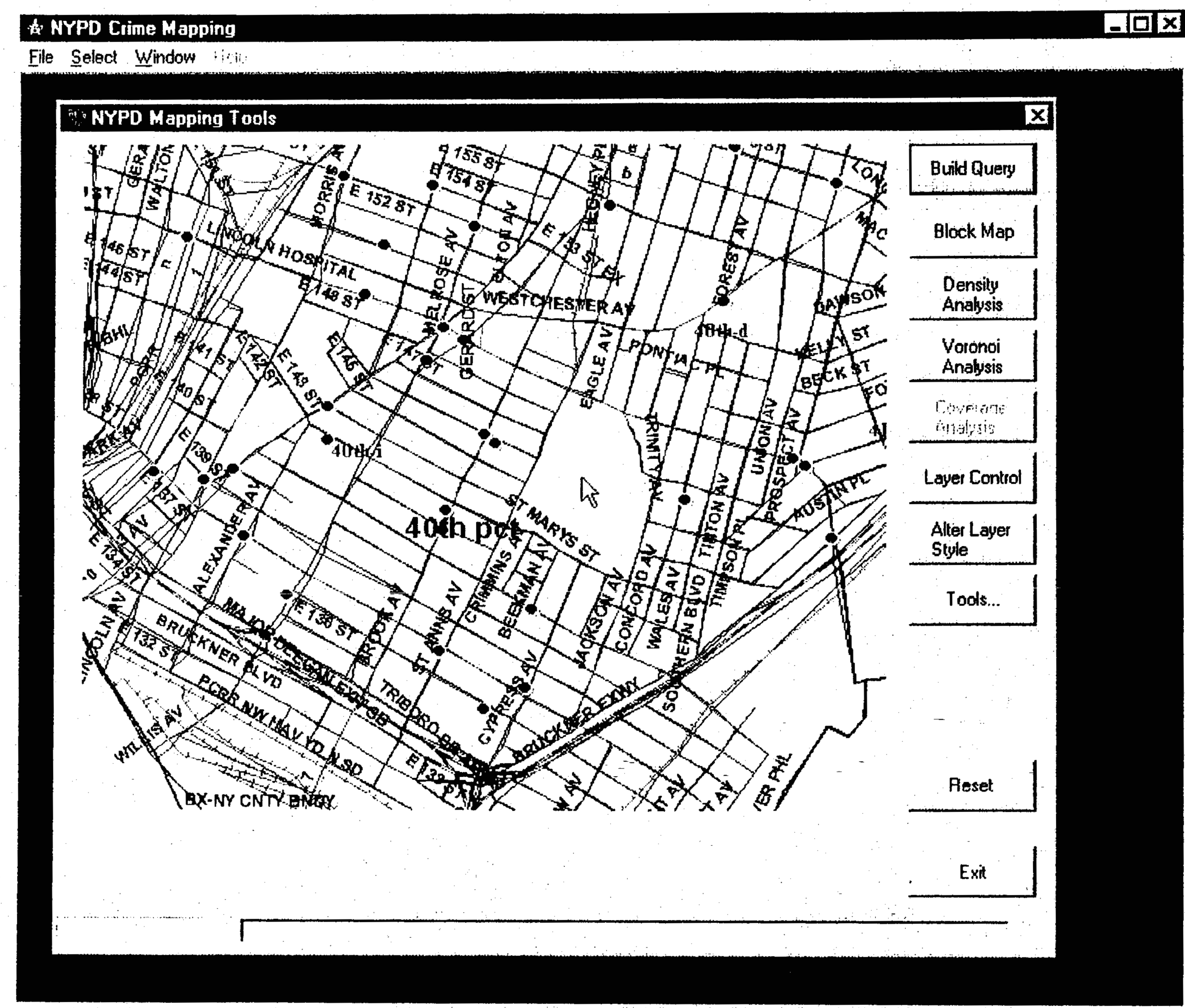

Figure 3: Main map window showing results from query

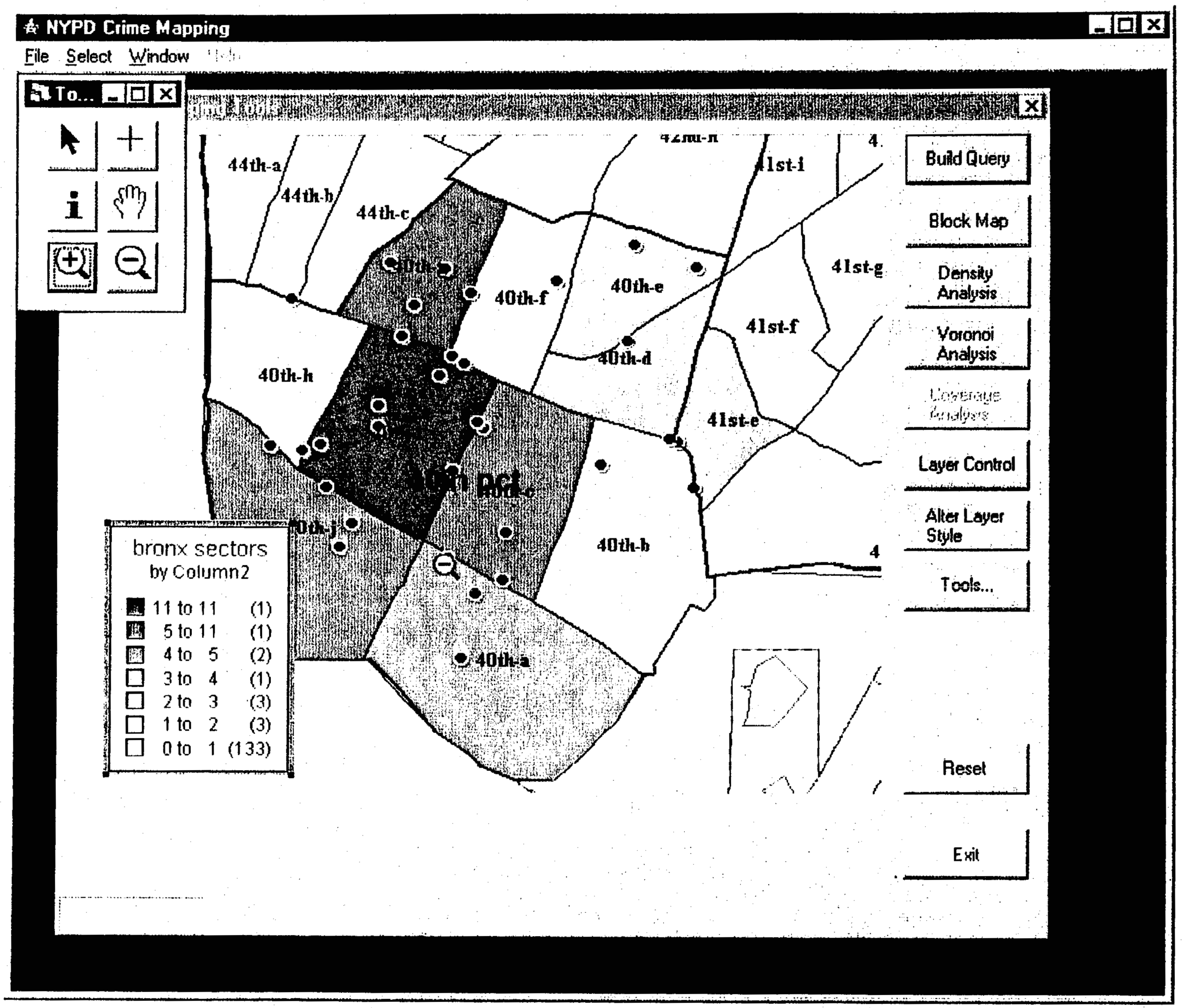

Figure 4: Block aggregation of points to police sectors

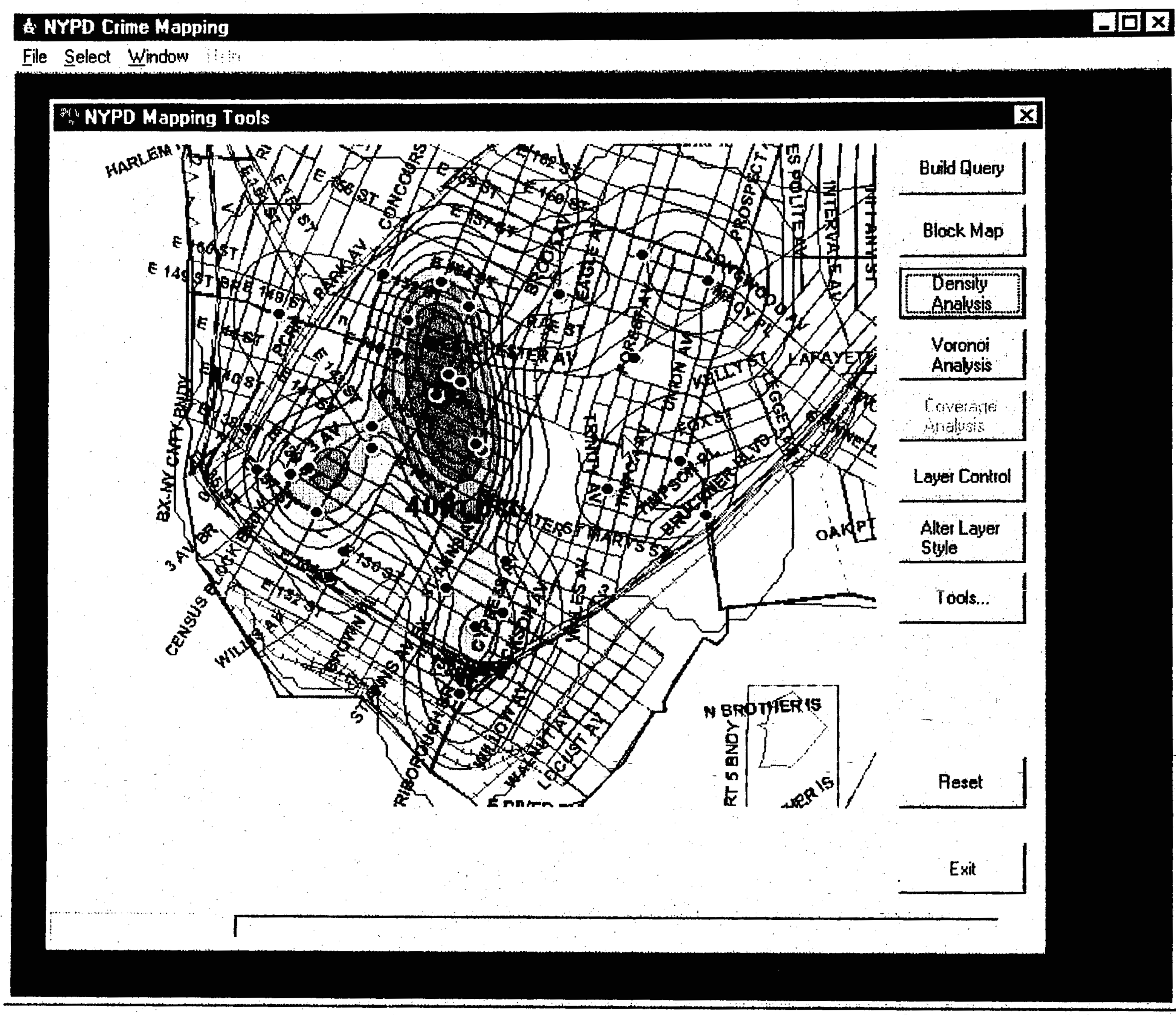

Figure 5: Density analysis of points, including street names

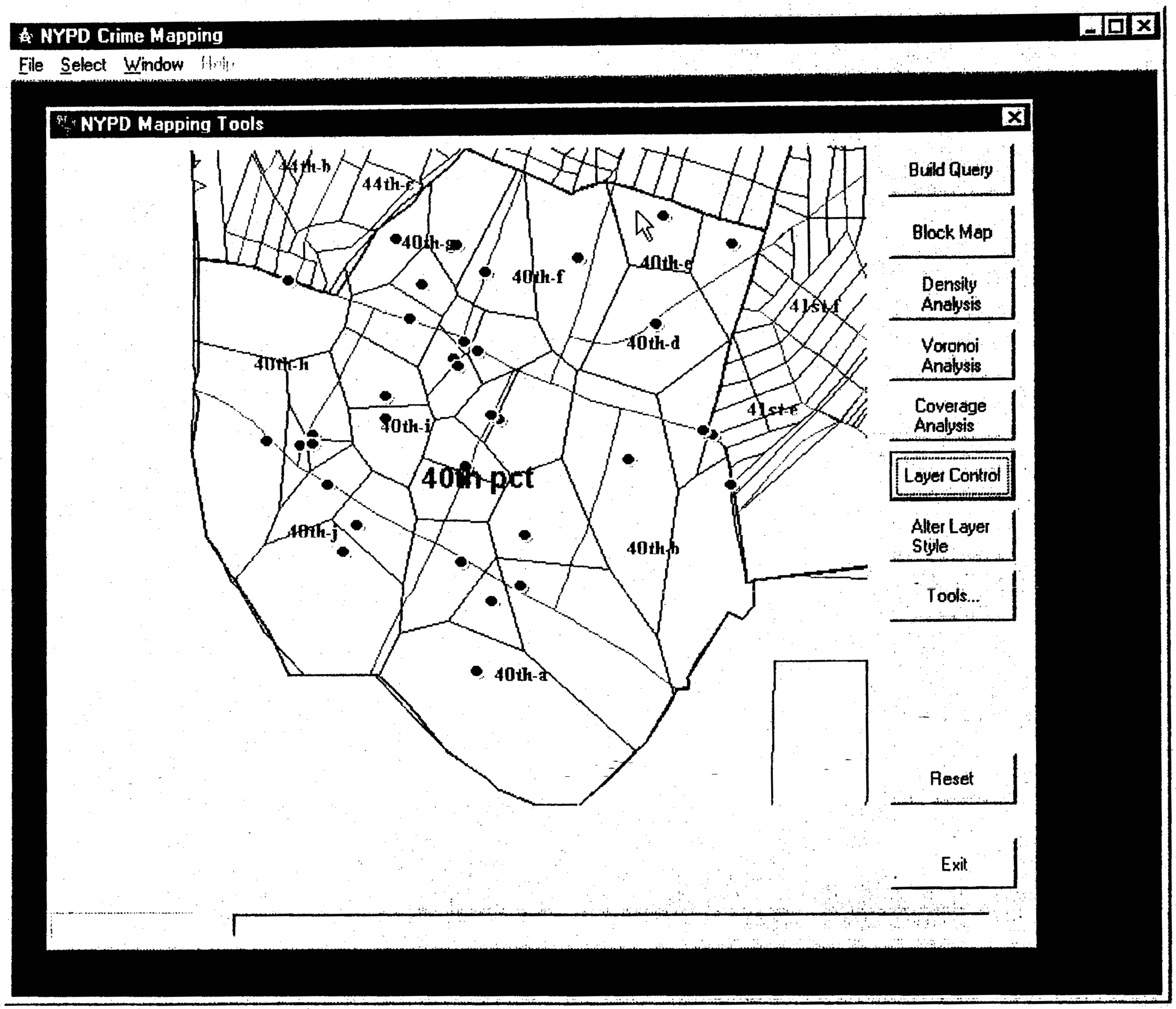

Figure 6: Voronoi diagrams of points

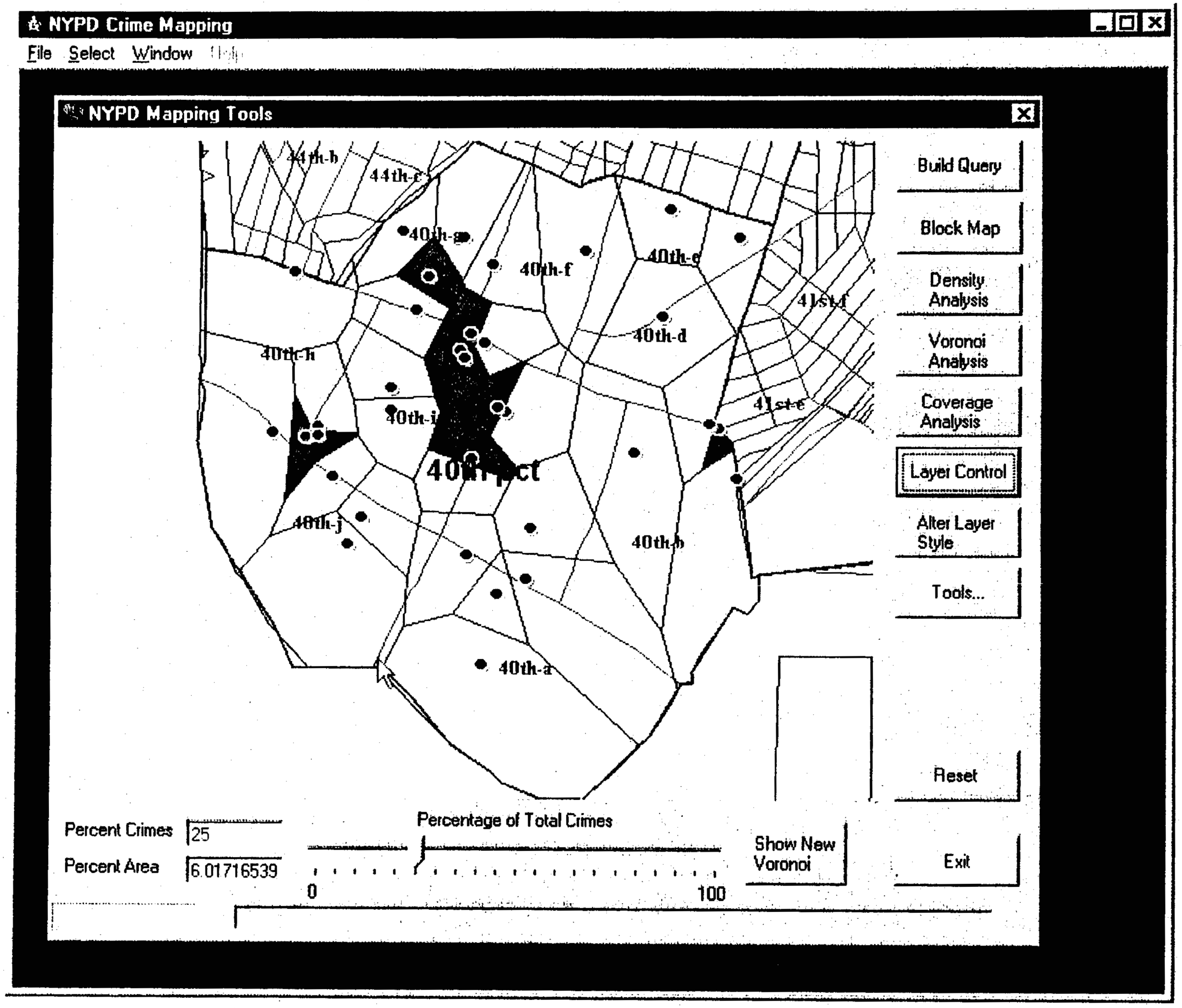

Figure 7: Coverage analysis based on results of voronoi analysis

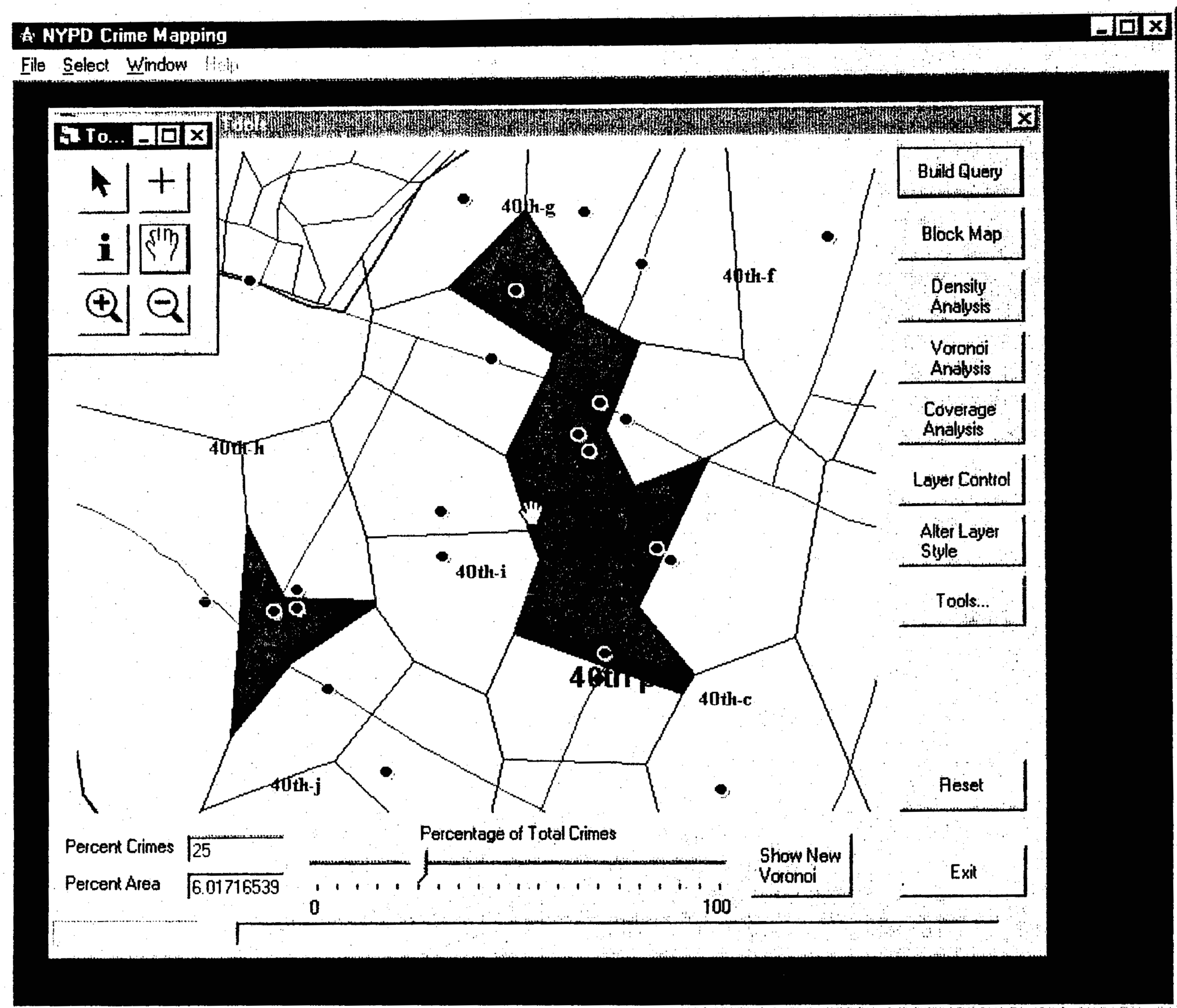

Figure 8: Example of zooming and panning using the tool window

Appendix A

Detailed Descriptions of Tools in CMAA

Block Aggregation

The CUNY/NYPD team developed a simple yet useful technique we call block aggregation. After geocoding crime incident locations, the number of incidents per block is aggregated to Census blocks.[1] To aggregate the data, we perform a spatial join between the crime points and the polygon Census blocks . We then create color-coded thematic maps that delineate hot spots using deciles.[2] The Map Basic code developed for this project performs these operations automatically.

Figure A.1 is an example of a block aggregation map that shows burglary data for 1998 in Brooklyn, New York. The map shows that the area east of Brooklyn's largest park, Prospect Park, is a center for burglary activity. Figure A.2 shows a magnified portion of that map that shows significant variation in burglary activity among blocks in that neighborhood. In other work, we used a burglary rate by incorporating the number of housing units (Kamber, Mollenkopf, Ross and Swartz, 1999). It turns out that this area of Brooklyn is thickly populated, and there is a large pool of potential targets as well as perpetrators here.

---

[1] Census blocks boundaries generally, though not always, conform to the street grid. Census block maps are available from many standard vendors as part of larger coverage packages that usually cost under $1,000 for a city. While it is theoretically possible to extract a Census block layer from a TIGER file, in practice we have found this to be a difficult and time consuming task. Another difficulty that may be encountered is that Census blocks have identifiers that include string characters, such as block 201A. It may be necessary to modify the identifier for some applications.

[2] A decile is similar to a median, except that it has ten divisions rather than two. When using a median, the data is divided into the those cases falling into the highest fifty percent, and those falling into the lowest fifty percent. When using deciles, the data is divided by ten percent increments, so that a hot spots can be defined as those blocks that fall into the top ten percent most crime ridden blocks. Though some analysts might prefer to use equal area, equal interval or some other break point, our experience suggests that marking blocks that fall into the top three deciles as hot spots is a useful operational definition.

The block aggregation technique has several advantages. First, it is easy calculate. Other than the acquisition of a Census block map, there is no additional cost to employing the method. In addition, the output is easy to explain. Whether the audience is executive staff, beat cops, or community groups, block aggregation maps are easily understood, and do not require technical expertise to interpret.

The technique is precise yet flexible. While many smoothing methods and other cluster identification techniques incorporate non-hot spot areas into their hotspots, the block aggregation technique precisely identifies which blocks have lots of crime and which do not. It is easy to identify situations where a crime-free block sits next to a crime-ridden one. The flexibility allows for the display of many different types of hotspot maps -- equal area and equal interval for example. Like other techniques, block aggregation can be used to create graphs showing the percentage of blocks that would need to be patrolled to cover a given percentage of crime.

Finally, Census block maps can easily be linked with other data sources, including the decennial census, the number of bars and/or liquor outlets, and zoning data. We found, for example, that once adjusted for the number of housing units, the concentration of residential burglaries all but disappears. Similar analyses could determine the importance of subway stops, public housing or schools in predicting crime rates (see Block and Block 1999; Kamber, Ross and Mollenkopf 1999; and Roncek, 1999).

The block aggregation technique also contains several weaknesses. First, the technique does not handle small amounts of data especially well. In a month time period for example, most blocks experience relatively few crimes. Increasing the cut-off point by one additional crime may drastically reduce the number of blocks identified as hot spots, and vice versa. This often results in maps with either too few or too many hot spots to be operationally useful.

The block aggregation method is particularly appropriate for crimes like burglary that typically occur in residential buildings located within the block perimeter. Outside crimes -- crimes that take place on or along street segments -- are arbitrarily assigned to one block or another depending on the address or intersection information used in locating the crime, and this in turn affects the block-level counts.

The sizes of the blocks may skew the data and affect the appearance of the block-level map. Some blocks incorporate malls or parks, while others refer only to small underpass areas. Large blocks, not surprisingly, often have a lot of crime, but may not be especially hot compared to similar areas that are chopped up into 3 or 4 blocks. Large blocks also dominate the map visually and exaggerate the amount of crime, whether large or small, that exists in those large geographical areas.

In addition, the block aggregation technique can miss some hot spots. For example, imagine that the cut-off point for a map is 10 burglaries for a block to be included as a hotspot. One block, or even one block face may have 9 burglaries, and there may be 9 burglaries in the block next to it as well. Though clearly such an area has a crime problem, it does not show up on the hot spot map.

We included the block aggregation technique into our software because we believe the advantages easily outweighed the weaknesses. It is a useful tool for studying historic crime patterns, and provides an easy "first cut" at the data before additional analysis using more sophisticated techniques.

Kernel Smoothing

The second method incorporated in the crime analysis system is kernel estimation, or kernel smoothing. Kernel smoothing is a statistical method for determining the density of crimes or other point events at different locations (Bailey and Gatrell 1995). The method is used to generate a continuous crime density surface from crime point data. The analyst begins with a dot map of crime events. Kernel smoothing results in a continuous 'weather map' that shows geographic variation in the density or intensity of crime. Peaks on the map represent areas of high crime (crime hot spots) and valleys represent areas of low crime. Increasingly crime analysts are employing kernel smoothing to visualize and analyze crime patterns (Williamson et al. 1999; Dalton and Brown 1998). The method is widely available to users of ArcView, because it is an option in the Spatial Analyst extension; however, until very recently, the method was not available to users of MapInfo. This changed recently with the introduction of kernel smoothing as an option in Vertical Mapper (see earlier discussion) which was stimulated by this collaborative project, and also by the very recent introduction of the 'hot spot detective' (Ratcliffe 1999).

In kernel estimation we begin by laying a fine grid across the study area. A circular window with a constant radius or *bandwidth* is moved across the study area, centered at each grid point (Figure A.3). The density of events is computed within this circular window. Events within the window are weighted according to their distance from the center of the window, the point at which density is being estimated (Bailey and Gatrell 1995). Events located near the center have a greater weight than those distant from the center. In this way, kernel estimation reflects the underlying geographic locations of events within each window.

Let s refer to a grid point location. The density of crime events at grid point s, $\lambda(s)$, is estimated as:

$$\lambda(s) = \sum_{d_i<\tau} 1/\tau^2 \; k(d_i/\tau)$$

Where:
$\lambda(s)$ = estimated density at grid point s
d_i = distance from point i to grid point s
τ = bandwidth

The bandwidth, τ, defines the radius of a circle, centered on s, within which density is being estimated. The kernel function, k(), describes mathematically the weight assigned to points (crime events) within the circle in calculating density. Weight varies inversely with distance so that the weight assigned to any crime event decreases as its distance (d_i) from the grid point increases. Thus, nearby events are given more weight in the density calculation than those farther away. Common mathematical forms for the kernel function include the gaussian and quartic functions. Research indicates little difference in density estimates for these two functional forms (Diggle 1983).

This means that density depends on both the number of crimes within a circle and their spatial configuration. Even if two grid points have the same numbers of crimes within their respective circles, if the spatial patterns of crimes differ, density will differ. A grid point with crimes clustered nearby will be assigned a much higher density than one with crimes more widely dispersed. In this respect, kernel smoothing differs from more traditional "binning" methods that simply count the number of events within regularly spaced circles, hexagons, or squares.

After computing kernel estimates of crime density for each regularly spaced grid point, one can generate a smooth map of density values. The results may be displayed as a standard contour map, a 3-dimensional surface, or as a continuously shaded raster map with gray- or color-tones representing density levels (see LeBeau, 1995).

Figure A.4 presents a point map that shows the locations of robberies in the command of Brooklyn South, New York for a four-week period. During this period, 650 robbery incidents occurred in the command area. Because of the large number of incidents, it is difficult to interpret the point map. Although several concentrations of robberies are evident, differences in the numbers and intensity of events are unclear. In the densest areas, dots conceal other dots giving a false impression of the true crime density. Although this problem can be addressed by using graduated circles – circles whose size is proportional to the number of incidents -- when density is very high, even graduated circles can hide each other. By comparison, Figure A.5 shows a continuous map of crime density for the same data set that was created using kernel estimation. Distinct hot spots appear as shaded peaks on the map. Geographical variation in crime density is clearly visible on the smooth density map. An advantage of kernel smoothing over traditional methods for identifying crime hot spots, like STAC (Spatial and Temporal Analysis of Crime), is that the hot spots can be irregularly shaped and need not follow regular geometric shapes like circles or ellipses (Block 1995).

An important issue in kernel smoothing is the choice of bandwidth. Large bandwidths smooth the data, revealing broad changes in crime intensity; whereas small bandwidths produce less smoothing, resulting in a map that is 'spiky' in appearance. Most existing methods for defining bandwidth are either ad-hoc or are based on the overall density of points in the study area. In this project, we devised a new method for defining bandwidth based on the average kth nearest neighbor distance among the points (Williamson et al. 1999). Doug Williamson wrote an ArcView script to implement the method. We have received requests for the script from crime analysts all across North America, and the response has been very positive.

Our recent publications discuss additional uses of kernel smoothing in crime analysis beyond the simple visualization of crime patterns (Williamson et al. 1999). By 'cutting off' the kernel density map at a particular value to pull out the 'peaks' of crime intensity, we can identify hot spot areas that can be studied in their own right. Characteristics of crimes in the hotspots, including numbers and types of crimes, times of occurrence, victims, weapons, etc. can be analyzed and used in planning appropriate interventions (McLafferty, Williamson and McGuire, 1999). To analyze changes in density over time, density maps can be compared on a grid cell by grid cell basis. As discussed later in this document, we have prepared animated sequences of the density maps to show spatial and temporal trends in crime intensity, as have analysts in other cities (Tempe Police Department web site). Thus, the maps generated by kernel smoothing have value for crime analysis beyond visualization, and in the long run the maps may contribute to improved decision-making at the precinct level and in NYPD's COMPSTAT process.

Voronoi Diagrams and Coverage Curves

The third tool incorporated in the crime analysis system are Voroni diagrams and the related concept of coverage analysis. One technique of spatial analysis that has seen wide application across many disciplines is the Voronoi diagram. This geometric construct splits a space, typically a mapped area, into a number of polygons. Each polygon is constructed around a generating point, here the location of a crime incident, in such a way that every other point within the polygon is closer to the enclosed generating point than to any other generating point (crime incident). The area within the polygon surrounding each generating point has been described as the point's natural neighborhood or Theissen polygon (see Figure A.6). As the intensity of crime in a local area rises the Voronoi polygons become smaller and more closely packed. The Voronoi process can in a sense be interpreted as a clustering method.

Once a geographic area has been identified as the scene of specific criminal activity, that area will be targeted for special attention. In those situations in which the attention involves special patrol resources such as foot, bicycle, scooter, or motor patrol these deployments can be thought of as "covering" a specified area. The extent of the area of coverage is typically based upon an assessment of the characteristics of the criminal activity, the likelihood of continued activity in the area and the expected success of police patrol coverage options in disrupting continued criminal activity.

We propose that a specific interpretation of the Voronoi diagram associated with a crime incident map may be quite helpful to police analysts and managers who are confronted with the problem of deciding on the extent of the area to be assigned some special patrol resources. Specifically we present a method for constructing what we have termed a "coverage curve." A coverage curve relates the percentage of the total incidents mapped to the percentage of the total map area covered by the sum of the Voronoi polygon areas.

The relationship between the percentage of incidents and the percentage of area is based upon two factors:

-Each Voronoi polygon is associated with one and only one generator point (crime incident).

-The Voronoi polygons associated with clustered crime incidents are smaller than those associated with widely spaced crime incidents.

To create a coverage curve for a specific set of mapped crime incidents we first create the Voronoi diagram for the set of incidents (see Figure A.7). Then we rank the generated polygons from the smallest to largest based upon each polygon's area. The ranked list can then be summed moving from the smallest area to largest area row by row giving an accumulation for both incidents and for polygon areas. The cumulative sum in the last row containing the largest polygon and its associated incident will total to the sum of all incidents represented in the map and the total area of the map respectively. The accumulation of each sum can be changed into a percentage by dividing the cumulative row values by the respective totals (see Table A.1.). When the row by row percentages, of cumulative count and matching cumulative area, are plotted against one another in scatter plot format we have constructed what we have termed a coverage curve (see Figure A.8).

Note that the plot is actually discrete and defined only at those integer points between 1 and N, (the total number of crime incidents recorded for the period in question). The plot resembles a smooth curve when the number of incidents is so large that the resolution of the computer charting program used to plot them is unable to depict each individual point. However the use for which we have constructed the curve is most appropriately conveyed by the term "coverage curve."

Once the coverage curve has been constructed we have a direct relationship at each plot point between the percentage of the total incidents displayed in the specific map and the percentage of the total area covered by the Voronoi polygons generated by those points. We propose that by examining the coverage curve, and particularly a specific graphic representation of the curve, crime analysts and police managers will be better able to select areas within their jurisdiction requiring special patrol attention.

In the coverage curve presented in Figure A.9 we see that approximately 50% of the precinct's robbery is contained within approximately 17% of the precinct's area. The area sum being the area of the smallest Voronoi polygons accounting for 50% of the total incidents. Similarly 75% of the precinct's robbery is contained within approximately 36% of the precinct's area. Therefore any approximate percentage of the precinct's crime can be associated with a coverage area based upon the natural neighborhoods surrounding the most densely packed group of incidents that can be assembled to give that percentage of total incidents (see Figures A.9 and A.10).

We can therefore envision a crime analysis tool that allows an analyst to move up and down the coverage curve while simultaneously viewing the appropriately shaded group of Voronoi polygons overlaid on a map of the jurisdiction. In this manner the analyst can gain useful insight into the clustering of incidents within the jurisdiction and be assured that the coverage area at each level has been generated to include the most closely packed grouping of incidents that can be assembled from the set of incidents under examination.

The concept of coverage curve raises a number of interesting questions:

1. Could the concept of the discrete coverage curve be used to create an automated review process that was able to sort through hundreds of crime incident maps and identify those with the most concentrated clusters of incidents? The shape of the coverage curve, the greater its' deviation from the 45 degree line in the plot of the two percentages, indicates greater clustering of incidents.

2. Does the coverage curve represent a fundamental characteristic of the underlying geography or possibly the specific type of crime incident under analysis?

3. Can the concept of the coverage curve be used to help forecast future incidence levels? If a coverage curve analysis is performed for a specific crime in a specific area for a lengthy

historical period how frequently will future incidents occur within and outside the identified historical coverage area.

4. Can coverage curve analysis help to determine if certain fixed sites are surrounded by specific types of incidents beyond what would be expected by pure chance?

5. Can the changes in the characteristics of the coverage curve or coverage area be related to changes in police interventions (enforcement and prevention programs) in the area?

These questions clearly await future investigation. Using a common data set, we have compared coverage curves to similar types of plots generated for smooth density maps, and we have found a remarkable similarity between the two types of curves. The structure of such curves relates to the underlying geography of the study area and the spatial distribution of crime incidents within the area. Our future research will also explore the range of applications for the Voronoi tool. As crime analysts begin to utilize the tool embedded in the CMAA, it will be important to monitor their responses to the tool and make note of the types of policy questions to which the tool is applied.

Animation

Animated maps are: "maps characterized by continuous or dynamic change (Slocum, 1998, 222)." The maps are displayed dynamically, in sequence, forming a constantly changing image or animation. The field of animation has advanced rapidly in recent years, stimulated by developments in computer hardware and graphics. These advances are fueling changes in map making as cartographers gain access to one of the first effective tools for representing continuous change through space and time.

Analysts are just beginning to explore the use of animation in crime mapping. The city of Tempe Police Department, for example, put an animated map sequence on the web to depict changing patterns of crime density over time. Our research team has prepared animated sequences of NYPD crime data that generated much interest. One member of our team (Doug Williamson) administered a web-based questionnaire to examine the effectiveness of animated maps for displaying changing crime patterns over time. An animated sequence of crime data for Brooklyn South is available from the project team on the web.

Creating animated map sequences involves decisions about the timing of maps in the sequence and their rate of change. Duration refers to the length of time each map is in view (Slocum 1998). A short duration means that each image disappears quickly producing a smooth, but constantly changing animation, while a longer duration gives the viewer more time to study each map, but the animation appears choppy. Rate of change describes the smoothness of the animated map sequence. It is computed as the amount of change between maps divided by the duration. If the positions or attributes of features on the map change substantially during the animation, the animation has a high rate of change. One can reduce rate of change by increasing the duration of each map and thus smoothing the transition from map to map. Similarly, reducing the amount of change between maps gives a lower rate of change and a smoother animation. One way to accomplish this is by using overlapping time intervals for the maps rather than discrete intervals. For example, the first map might show crime incidence for weeks 1-4, the second for weeks 2-5, the third for weeks 3-6 and so on. The one-week overlap means that some of the data on a map is shown on the next map so that the rate of change is gradual.

The animation component of the CMAA is designed to provide flexibility in defining duration and rate of change for a particular animated sequence. The crime analyst first selects the time interval for maps in the sequence and then defines the degree of overlap among the maps. This way one can control the smoothness or roughness of change in the animation and the time periods to be displayed. The animation works with crime density maps generated by kernel smoothing, because they are visually appealing and easy to comprehend in the short time interval of the animation. We experimented with animating dot maps of crime, but found them too complex to understand when viewed in sequence.

Animation is an effective way of displaying changes in crime through time and space; however, it is primarily a visual tool. Animated maps clearly show regular patterns, but if events move or vary in intensity unpredictably over time, the animation will be difficult to comprehend. Viewers often have trouble analyzing information on animated maps. The images move by quickly and are difficult to compare. Therefore we expect that the animation tool in the CMAA will primarily be used for visualization and display, for 'gee-whiz' presentations, and to give the viewer a sense of general space-time trends.

References:

Bailey, T. and A. Gatrell 1995. *Interactive Spatial Data Analysis.* New York: Wiley.

Block, C. R. 1995. "STAC hot-spot areas: A statistical tool for law enforcement decisions." In C. Block, M. Daboub and S. Fregly, eds., p. 15-32, *Crime Analysis Through Computer Mapping,* Washington, D.C.: Police Executive Research Forum.

Block, Carolyn and Richard Block. 1999. "The Bronx and Chicago: Street Robbery in the Environs of Rapid Transit Stations." In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice.* Thousands Oaks, CA: Sage Publications.

Brown, D. and J. Dalton 1998. *Regional Crime Analysis Program (RECAP).* Presented at Crime Mapping Specialists Meeting, National Institute of Justice, Washington, D.C., February.

Diggle, Peter 1983. *Statistical Analysis of Spatial Point Patterns.* London: Academic Press.

Kamber, Tom, John Mollenkopf and Timothy Ross. 1999. "Crime, Space and Place." In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications.

LeBeau, J. 1995. "The Temporal Ecology of Calls for Service", in *Crime Analysis Through Computer Mapping*, Police Executive Research Forum, 111-128.

McLafferty S, Williamson D and P McGuire. 1999. "Identifying Crime Hot Spots Using Kernel Estimation," In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications.

A.Okabe, B.Boots and K Sugihara 1992. ***Spatial Tessellations: Concepts and Applications of VoronoiDiagrams*** John Wiley & Sons.

Ratcliffe, J. (1999) 'Hot Spot Detective,' http://www.bigfoot.com/~hotspot.detective.

Roncek, Dennis. 1999. "School and Crime." In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications.

Slocum, T. 1998. *Thematic Cartography and Visualization*, Upper Saddle River, NJ: Prentice Hall.

Williamson, D, McLafferty, S, McGuire, P, Goldsmith, V, and J Mollenkopf 1999. "A Better Method to Smooth Crime Incidence Data" *ArcUser*, January.

ID	cum percent count	cum percent area
700	0.000978	0.000022
542	0.001957	0.000056
442	0.002935	0.000094
177	0.003914	0.000134
33	0.005871	0.000174
432	0.006849	0.000215
30	0.007828	0.000257
363	0.008806	0.000302
576	0.009785	0.000351
350	0.010763	0.000403
449	0.011742	0.000456
606	0.012720	0.000511
129	0.013699	0.000568
677	0.014677	0.000628
414	0.015656	0.000688
422	0.016634	0.000749
626	0.017613	0.000811
338	0.018591	0.000875
497	0.019569	0.000940
600	0.024462	0.001009
320	0.025440	0.001079
640	0.026419	0.001151
7	0.027397	0.001223
425	0.028376	0.001295
319	0.029354	0.001368

Table A.1 Sample table for creation of coverage curve

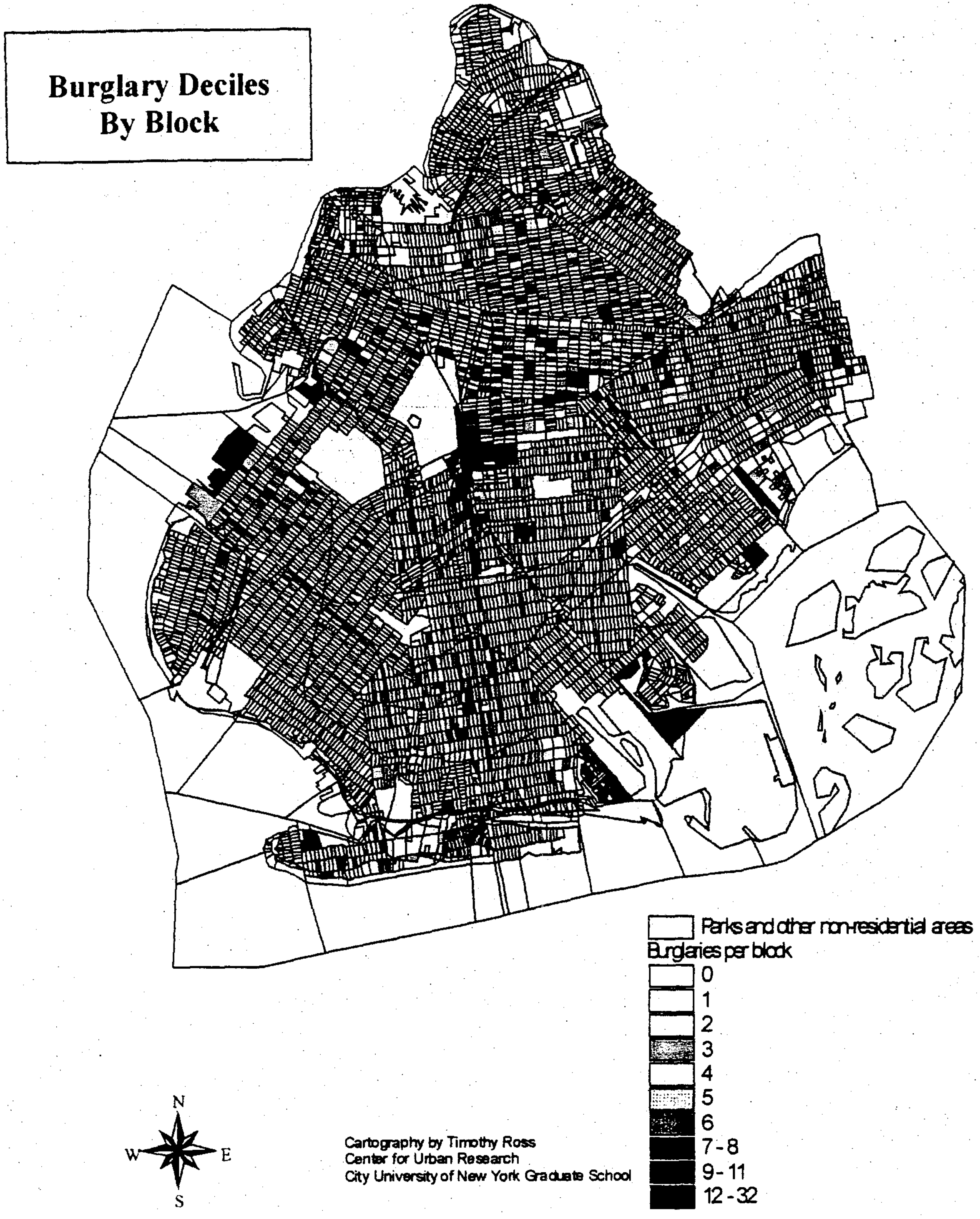
Burglary Deciles
By Block
Parks and other non-residential areas
Burglaries per block
0
1
2
3
4
5
6
7-8
9-11
12-32
N
W
E
S
Cartography by Timothy Ross
Center for Urban Research
City University of New York Graduate School

Figure A.1

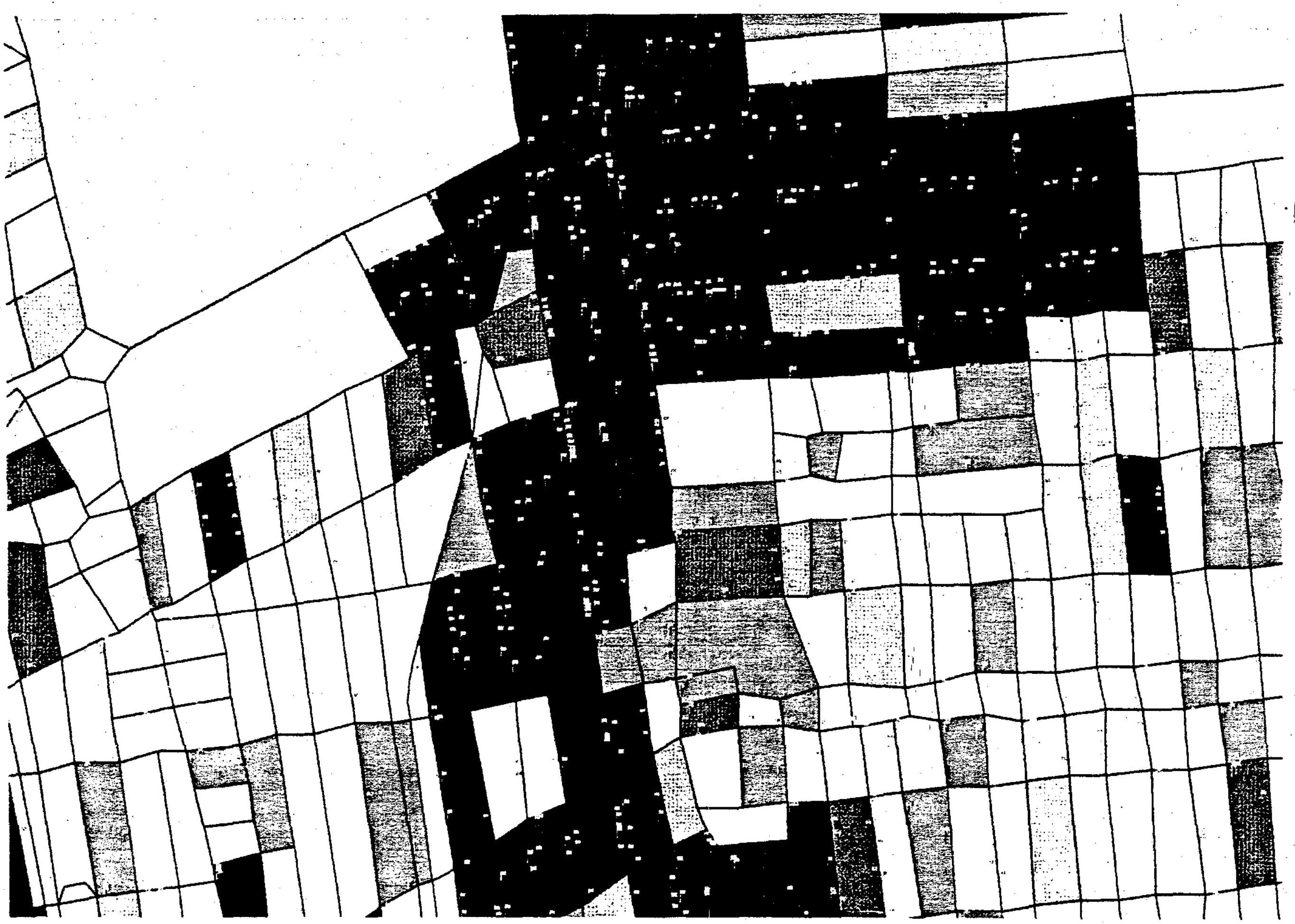

Figure A.2

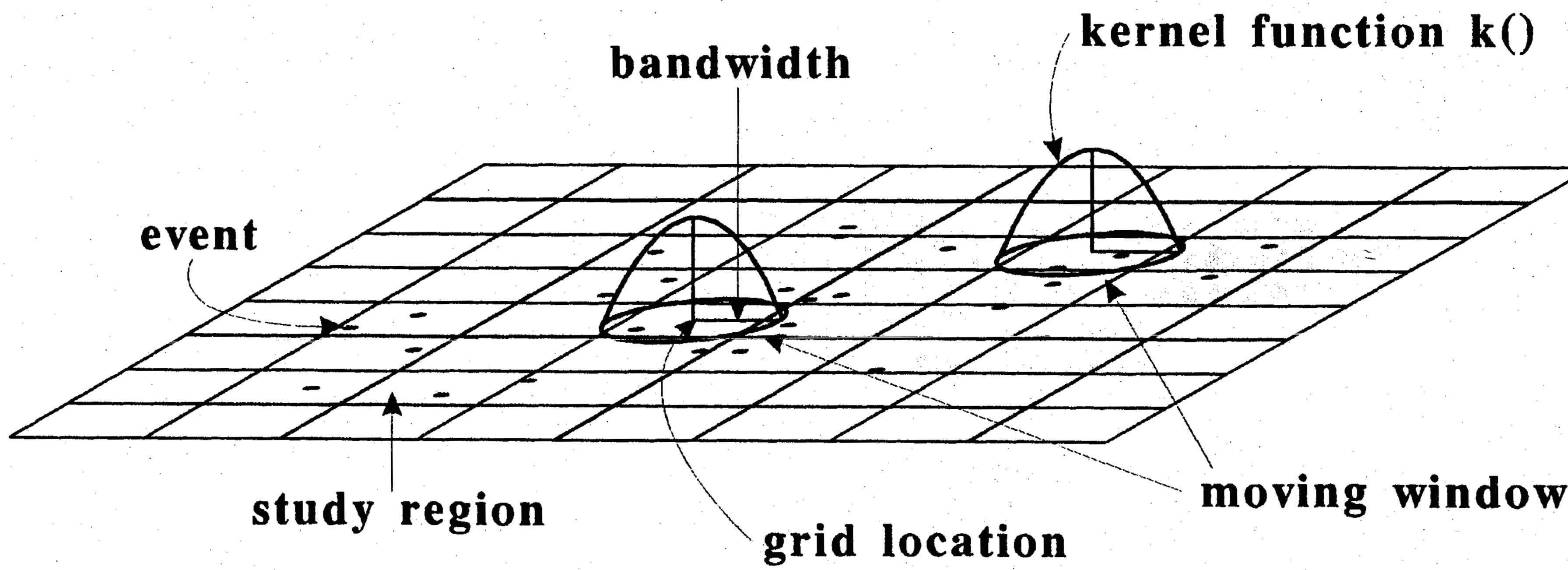

Figure A.3: Kernel smoothing process

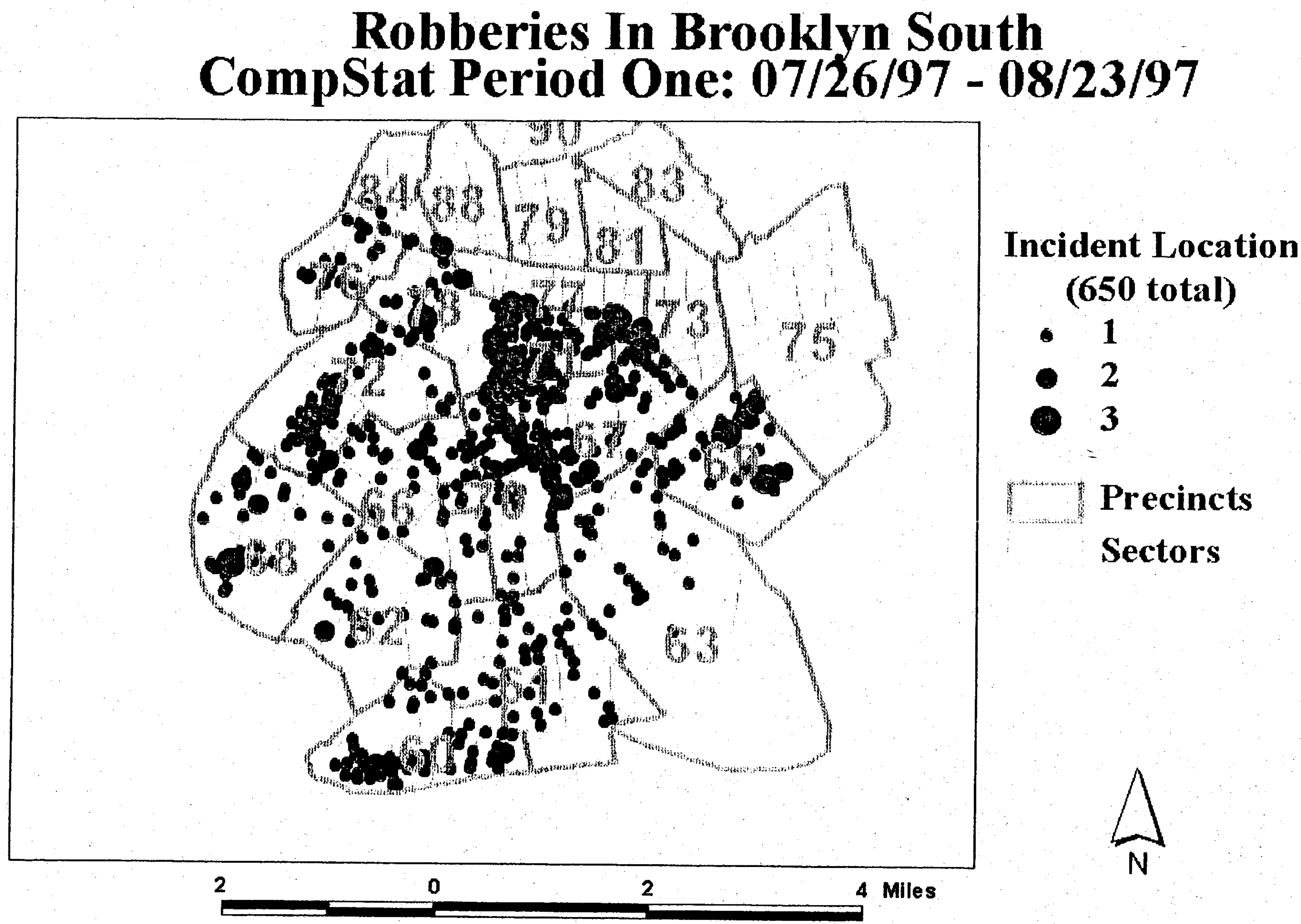

Figure A.4: Graduated symbol map of robberies in Brooklyn South

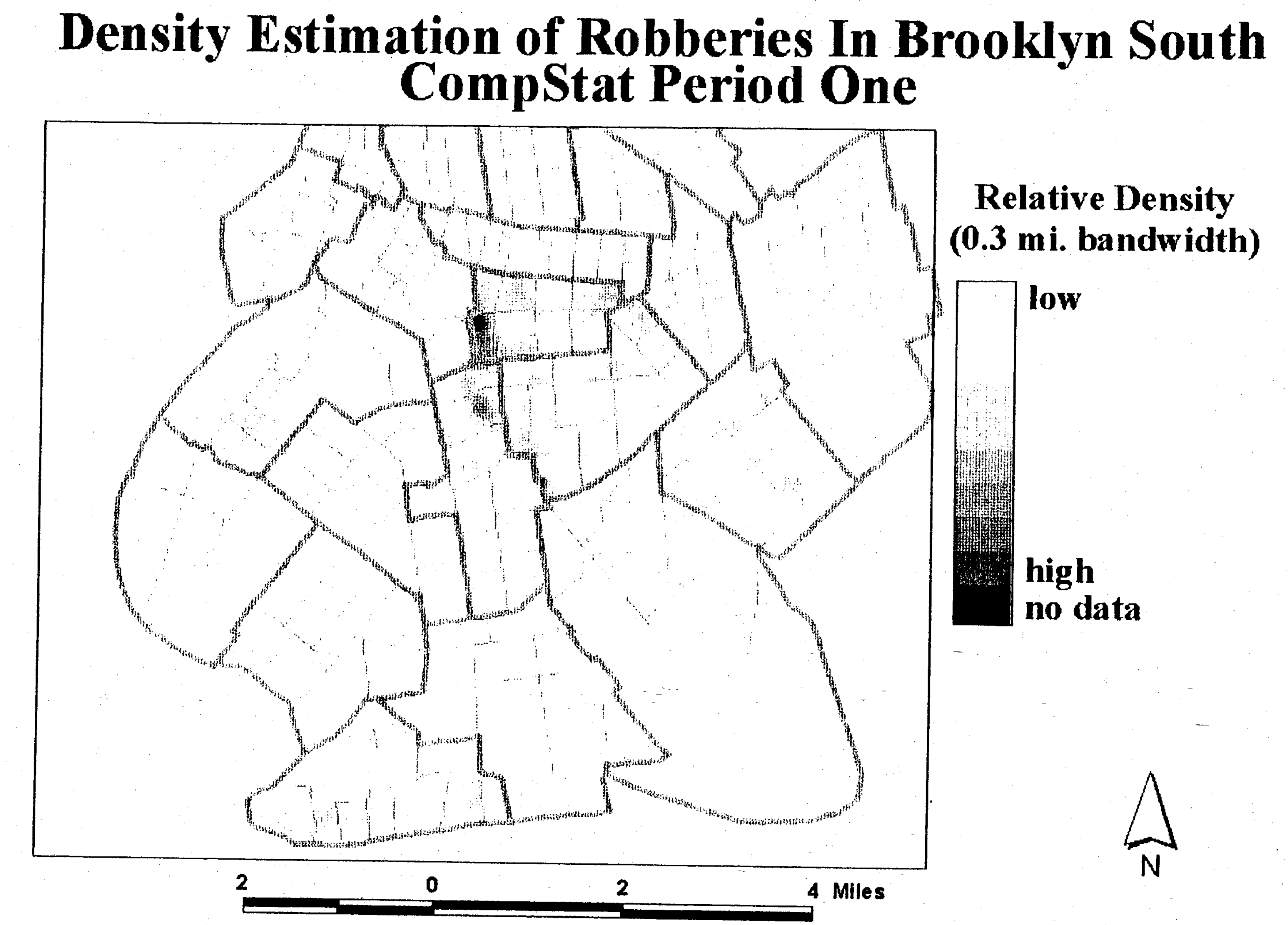

Figure A.5: Robbery density in Brooklyn South from kernel estimation

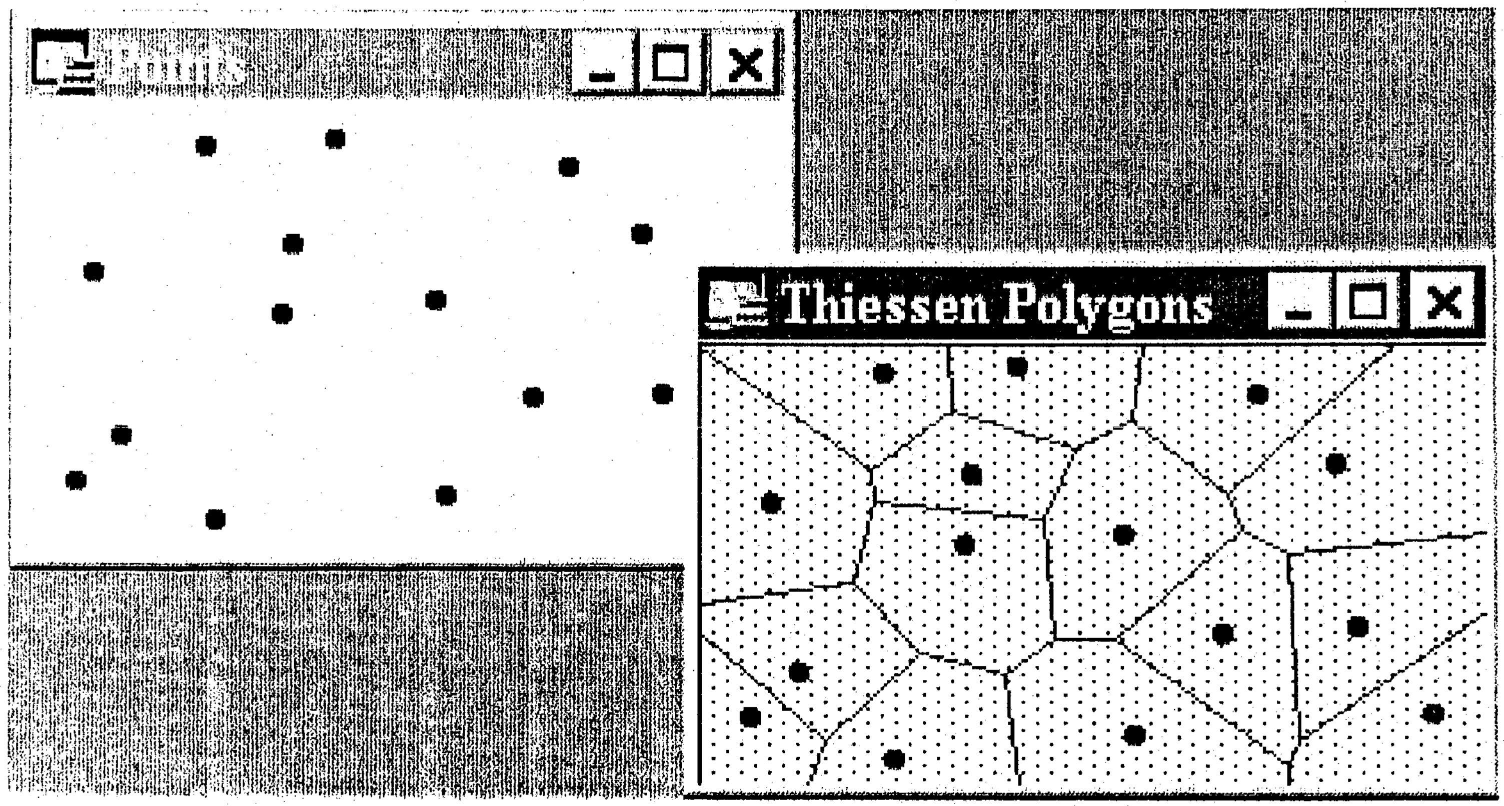

Figure A.6: Construction of Theissen polygons

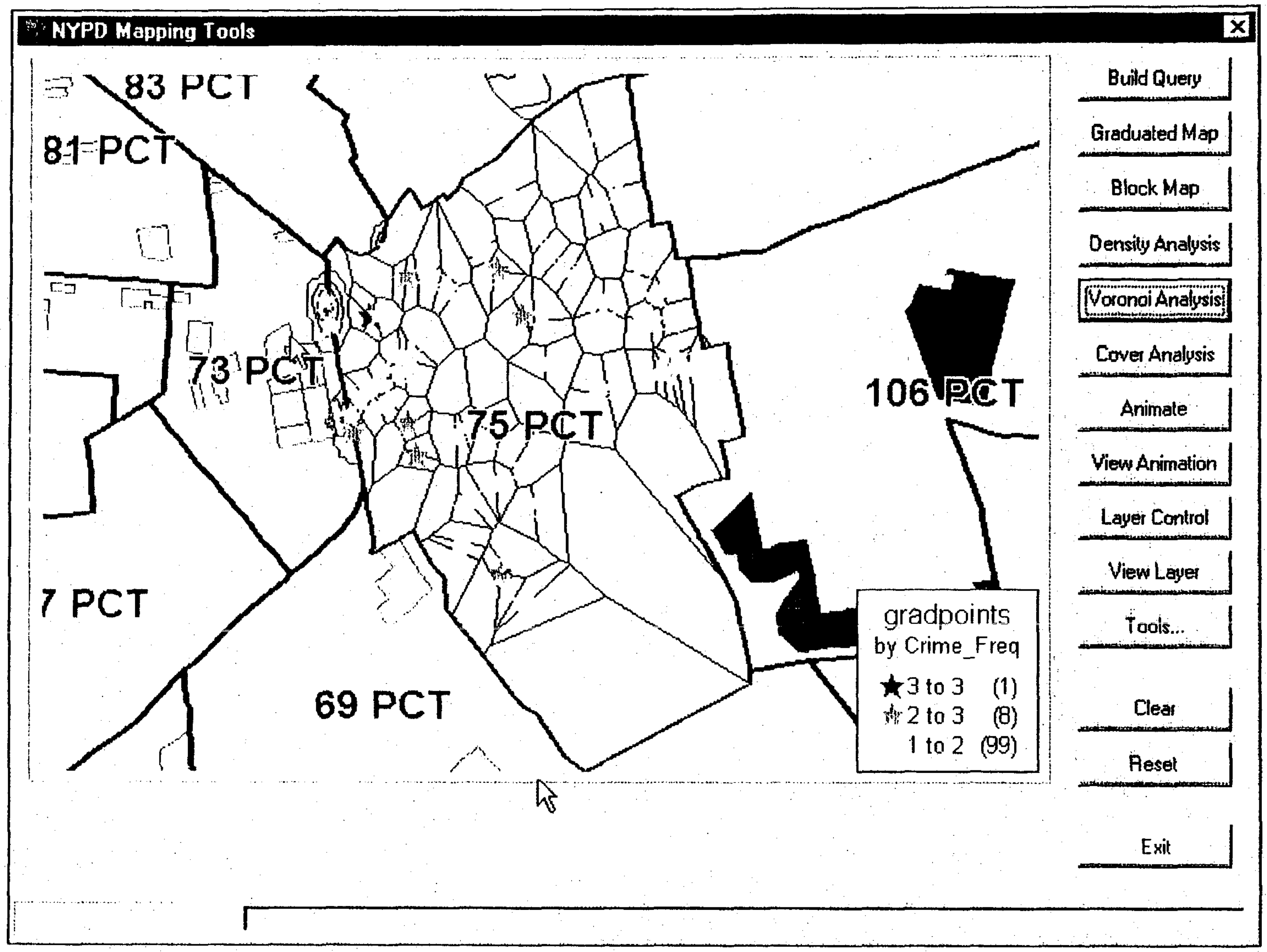

Figure A.7: Voronoi map of crime locations

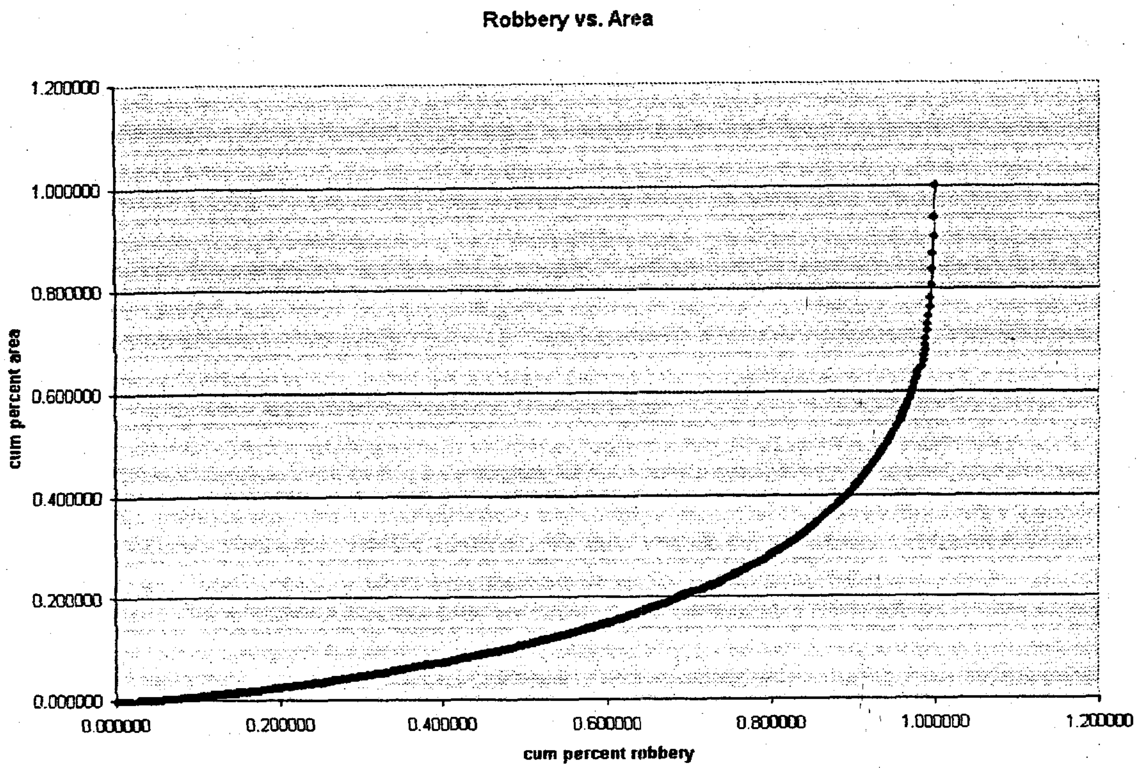

Figure A.8 Coverage Curve

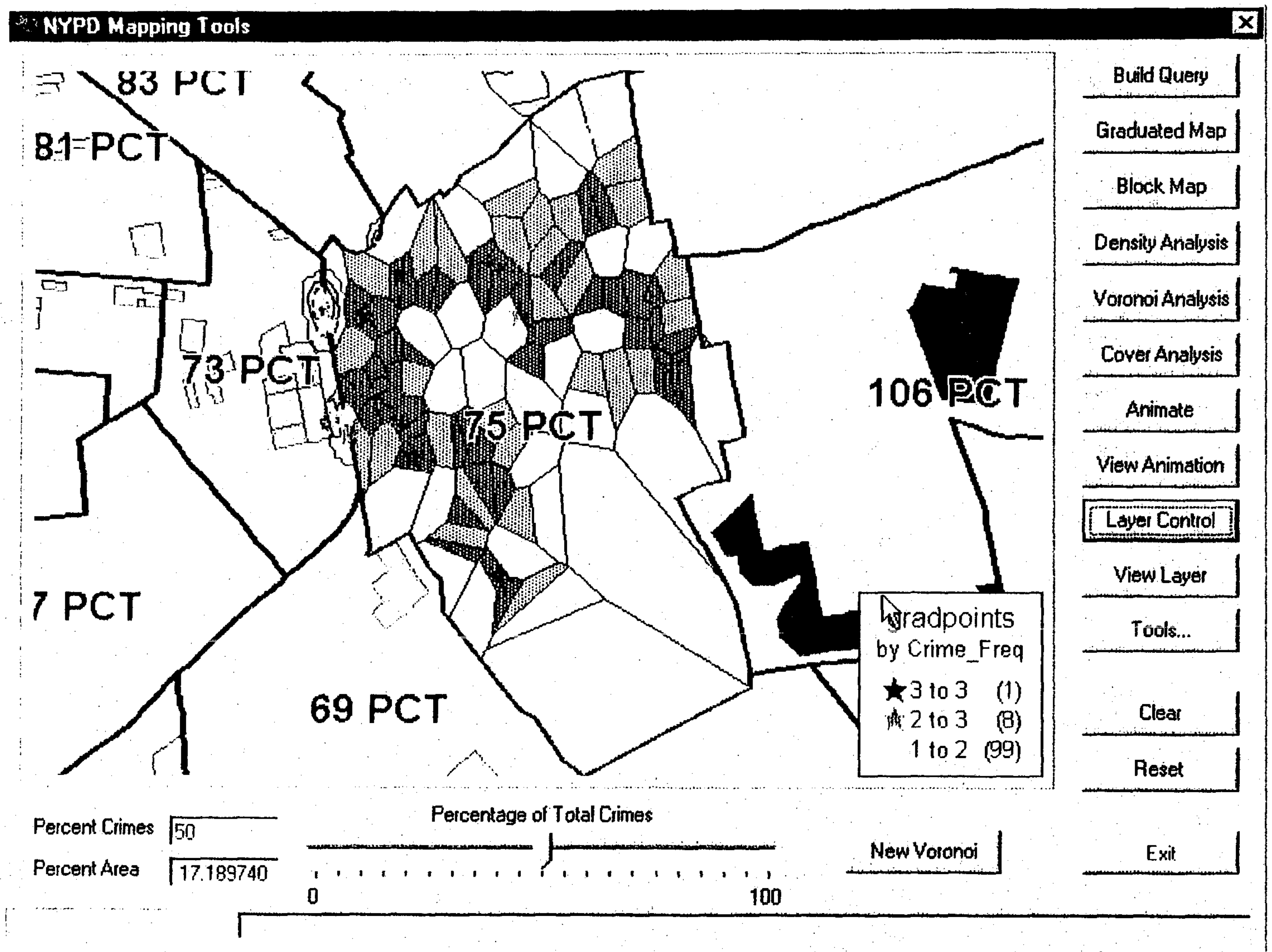

Figure A.9: Smallest Voronoi polygons that contain 50 percent of crimes

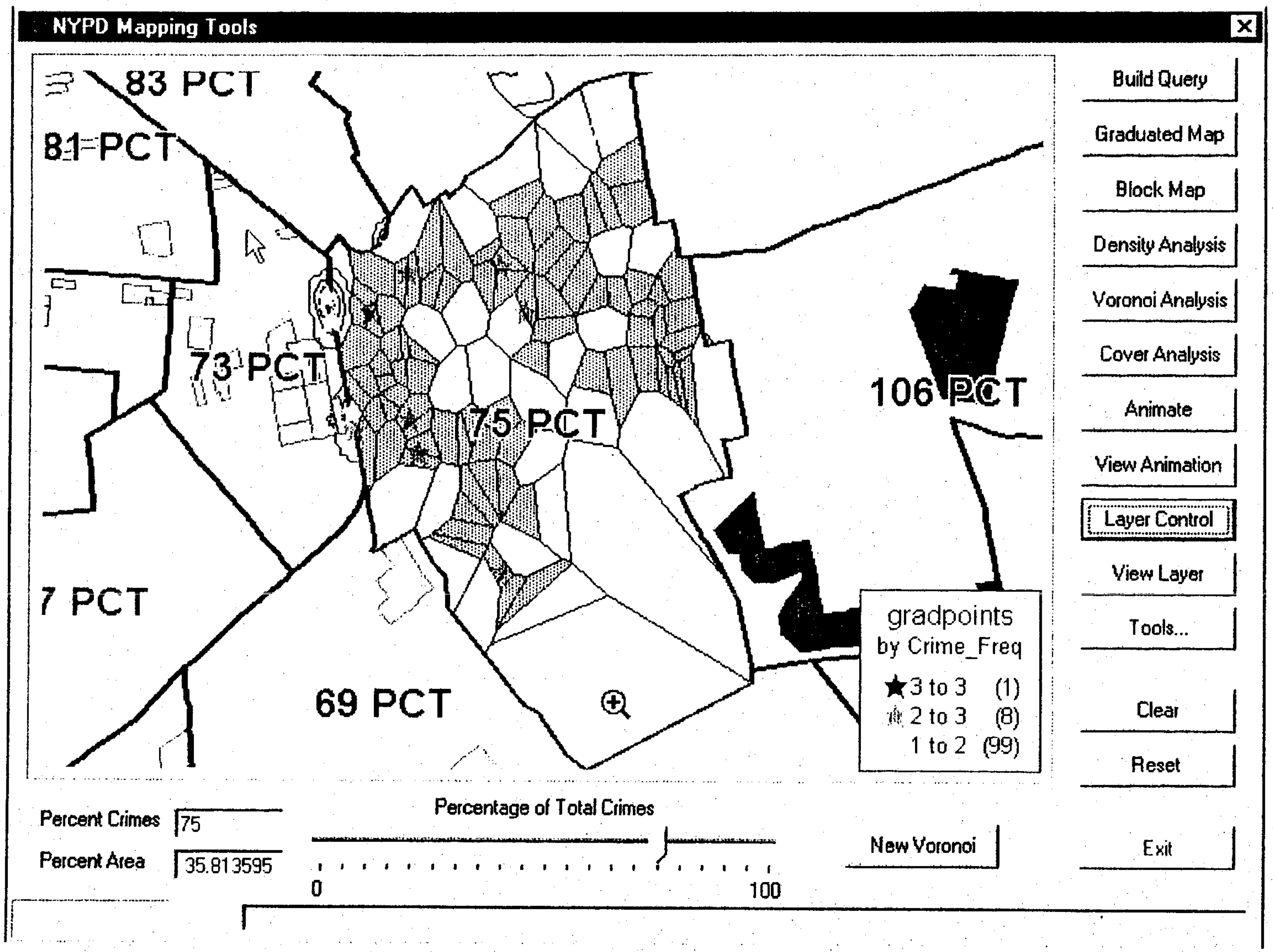

Figure A.10: Smallest Voronoi polygons that contain 75 percent of crime locations

Appendix B
List of Presentations

Research Team, briefing on crime mapping tools for Commissioner Safir, NYPD, February, 1999.

V. Goldsmith, "The Future Role of Crime Mapping in Reducing Crime in New York City," 7th International Seminar on Environmental Criminology, Barcelona, Spain, 1998.

V. Goldsmith, symposium at University of Nanterre, Paris, France, 1999.

V. Goldsmith, symposium at the Prefecture de Police of Paris, 1999.

V. Goldsmith, meeting with Central Directorate of Criminal police, Paris, 1999.

V. Goldsmith, "Analyzing Crime Patterns: Frontiers of Practice," Fatih University, Istanbul, Turkey, 1999.

V. Goldsmith, "New Methods in GIS Spatial Pattern Analysis With Applications From New York City," Eurasian Institute of Earth Sciences, Istanbul, 1999.

P. McGuire and V. Goldsmith, 'The NYPD-CUNY Partnership," Crime Mapping Research Conterence, Denver, 1997.

P. McGuire, presentation at 'Designing the Digital Government for the 21st Century-Digital Government a Multi disciplinary Workshop,' NSF sponsored meeting at the State University of New York Center for Technology in Government, 1998.

P. McGuire "Tool Development for the CUNY-NYPD Partnership," Cluster Conference on Spatial Analysis Tools in Crime Mapping, Washington, D.C., February 1998.

P. McGuire, presentation in Workshop on COMPSTAT Technology, Data Gathering and Staff Preparation, 1999.

P. McGuire "Mapping Tools for Management Accountability, " Crime Mapping Research Conference, Orlando, 1999.

S. McLafferty "Advanced Spatial Clustering Methods," Crime Mapping Research Conference, Denver, 1997.

S. McLafferty "Spatial Smoothing and Hot Spot Delineation," Cluster Conference on Spatial Analysis Tools in Crime Mapping, Washington, D.C., February 1998.

S. McLafferty "Spatial Clustering Methods: Issues and Applications," Crime Mapping Research Conference, Washington, D.C., 1998.

J. Mollenkopf, "NYPD and CUNY Collaborative Mapping Research," Crime Mapping Research Conference, Washington, DC, 1998.

T. Ross, "Crime and Public Housing," American Political Science Association, 1999.

T. Ross, invited presentation, Department of Sociology, Hunter College, 1999.

T. Ross, invited presentation, Department of Sociology, Queens College, 1999.

T. Ross, presentation in Panel on Trends in Crime Prediction and Prevention, MapWorld, 1999.

D. Williamson, "Crime Mapping and the Internet," Crime Mapping Research Conference, Denver, 1997.

D. Williamson, "Smoothing Crime Incident Data," ESRI Users Conference, San Diego, 1998.

D. Williamson, "A Multimethod Exploration of Crime Hotspots," Academy of Criminal Justice Sciences, Albuquerque, 1998.

D. Williamson and T. McEwen, "NYPD and CUNY Collaborative Mapping Research," Institute of Law and Justice, Washington, DC, 1998

D. Williamson and T. Ross, "Visualizing Crime Hotspots," New York City Data Connections Conference, New York City, 1999

Appendix C
List of Publications

Goldsmith, V., McGuire,P.,Mollenkopf, J. and Ross, T., Editors. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications, 1999.

Kamber, T., Mollenkopf, J. and Ross, T. "Crime, Space and Place." In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications, 1999.

McGuire, P. The New York Police Department COMPSTAT Process: Mapping for Analysis, Evaluation and Accountability," In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications, 1999.

McLafferty S, Williamson D and P McGuire.. "Identifying Crime Hot Spots Using Kernel Estimation" In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications, 1999.

Williamson, D, McLafferty, S, McGuire, P, Goldsmith, V, and J Mollenkopf "A Better Method to Smooth Crime Incidence Data" *ArcUser*, January, 1999.

Williamson, D., Ross, T., McLafferty, S., and V. Goldsmith. "Evaluating Statistical Software for Analyzing Crime Patterns and Trends," In Victor Goldsmith, Phil McGuire, John Mollenkopf and Timothy Ross, eds. *Analyzing Crime Mapping: Frontiers of Practice*. Thousands Oaks, CA: Sage Publications, 1999.

Williamson, D., McLafferty, S., McGuire, P., Ross, T., Mollenkopf, J., Goldsmith, V. and S. Quinn. "Tools in the Spatial Analysis of Crime," In A. Hirschfield and K. Bowers, eds. *Mapping and Analysig Crime Data: Lessons from Research and Practice*, forthcoming

Williamson, D. "Robbery" in, B. Dent, Turnball, L. and Hallisey, E. eds., *The Atlas of Crime*, Oryx Press, forthcoming

Appendix D

Letter from Commissioner Howard Safir, NYPD

#6

Visit congratulations DAC

THE POLICE COMMISSIONER
CITY OF NEW YORK

March 23, 1999

Dr. David A. Caputo, President
Hunter College
The City University of New York
695 Park Avenue
New York, New York 10021-5085

Dear President Caputo:

As you know, the New York City Police Department has adopted a management system referred to as COMPSTAT, that helps managers throughout the Department identify emerging public order problems, develop intervention strategies, rapidly deploy resources and then assess impact. Automated crime mapping has been a key tool in this process.

I believe that mapping analysis has the potential to make more than just a "points on a map" contribution to police decision making. To that end, and with support from the National Institute of Justice, Hunter College Professors Victor Goldsmith and Sara McLafferty along with Doug Williamson, a graduate student in the Geography Department, have been collaborating with NYPD staff in a search for spatial analysis tools that can make a real contribution to the NYPD's COMPSTAT process.

I am writing to you to bring their work to your attention. Law enforcement's use of spatial analysis tools, specifically tailored to police decision-making, can help to quickly identify crime hot spots, and more complex relationships between crime and environmental characteristics. The tools can also help us to identify changes in spatial crime patterns and more effectively judge the impact of our crime control and enforcement efforts.

I have seen some of the results of the collaboration and I believe the spatial analysis techniques, identified and customized for police use by the research team, will be enthusiastically accepted within the law enforcement community.

I hope the National Institute of Justice will also recognize the contribution made by this collaboration and continue to support this important work. Regardless of future funding outcomes, the research team is to be commended for their work to date and the contribution it will make to public safety decision-making.

Sincerely,

Howard Safir

Howard Safir
POLICE COMMISSIONER

bf