ANALYZING CRIME ON STREET NETWORKS: A COMPARISON OF NETWORK AND EUCLIDEAN VORONOI METHODS

BY

WILLIAM DEBARD MEYER

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Geography
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Professor Sara Mclafferty

Abstract

The analysis of the uneven spatial distribution of crime has been an important area of research investigation and policy analysis for the past several decades. These analyses typically use spatial analytical methods that are based on the assumption of Euclidean (straight-line) distance. However, crime like most social activity is often mediated by the built environment, such as along a street or within a multi-story building. Thus, analyzing spatial patterns of crime with only straight-line Euclidean distance measurement ignores this intervening built landscape and may very possibly introduce error into the ensuing result. The purpose of this research is to compare and contrast the differences in analytical results for spatial analysis techniques that have the capability to use either Euclidean or network distance. Voronoi diagrams which can be implemented utilizing either Euclidean distance or network distance (distance measured along a street) offer a means for performing this comparison. Utilizing Voronoi diagram implementations with Euclidean distance and network distance this thesis will examine the spatial distribution of gun-inflicted homicide locations and the similarity/differences between the results of their application with the aim of informing the spatial analysis of street located homicide.

Table of Contents

Chapter 1 - Introduction

1.1 Introduction

Analyzing the uneven spatial distribution of crime has been an important area of research investigation and policy analysis for the past several decades. These analyses typically use spatial analytical methods that are based on the assumption of Euclidean (straight-line) distance. However, crime like most social activity is often mediated by the built environment, such as along a street or within a multi-story building. Thus, analyzing spatial patterns of crime with only straight-line Euclidean distance measurement ignores this intervening built landscape and may very possibly introduce error into the ensuing result. The purpose of this research is to compare and contrast the differences in analytical results for spatial analysis techniques that have the capability to use either Euclidean or network distance. Euclidean distance is distance measured as a straight line between two points on a plane. Network distance is the distance along a transportation corridor such as a street network. A network is composed of a set of nodes and links, often called edges, which connect nodes (de Smith et al. 2007), and the network distance between two nodes is the shortest sum of distances along links connecting the nodes.

This thesis will examine the spatial distribution of gun-inflicted homicide utilizing Euclidean vs. network distances with respect to the method of Voronoi diagrams which can either be expressed as a polygon diagram or a lattice of polylines (Network voronoi). Voronoi diagrams have been widely used in the natural and social sciences to define optimal regions. These regions define the locations closer to a particular point than to any other point, and they are collectively exhaustive and mutually exclusive except along shared boundaries (Okabe, Boots and Sugihara, 1992). When used with data representing gun-inflicted homicide locations on streets these areas can serve as search zones where evidence pertaining to the crime committed might be sought by investigating officials, or represent the victim hunting zones used by perpetrators of crimes. Depending on the repeat frequency of homicide location one could also use a voronoi diagram, generated from a crime location to represent an area where patrol resources could be focused to deter crime. Until recently Voronoi diagrams could only be used in analyzing Euclidean (continuous) space, but new analytical techniques make it possible to calculate a Voronoi optimized space along a network of streets using distance measurement along the streets. This new technique was developed by Atsuyuki Okabe, Kei-ichi Okunuki, and Shino Shiode and is implemented in an analysis tool called Spatial Analysis on Networks (SANET). It was due to the voronoi diagrams utility in the spatial analysis of crime and its ability to be implemented either with Euclidean distance or network distance that made it an ideal choice for this research.

To accomplish the research, I use SANET to compare Voronoi diagrams based on network and Euclidean distance. The diagrams are calculated for homicide data for St. Louis, MO and are used in evaluating spatial clustering of homicides. I perform a comparison utilizing the Pearson's *r* correlation coefficient to determine the level of similarity/dissimilarity between the two Voronoi diagrams in terms of the length of street network and the area encompassed by each Voronoi polygon. I selected this method of comparison to determine if there was a profound mathematical difference in the results the two methods produce. If there was, the application of each method's results to the task of crime analysis would need to be carefully applied as it likely would indicate that the two methods were not interchangeable and one might have greater value over the other. To support these statistics and provide a clearer picture of differences I provide thematic maps illustrating the magnitude of similar/dissimilarity between the two techniques. The data utilized in the analysis was acquired from the National Archive of Criminal Justice Data. The dataset is entitled "Arrests As Communications to Criminals in St. Louis, 1970, 1972-1982", Kohfeld, and Sprague, (1999). The data were originally used in a series of homicide studies performed at the University of Missouri, St. Louis by Kohfeld and Sprague, the most influential of which was their work examining how unemployment drives urban crime (Kohfeld and Sprague, 1988).

The significance of this research is rooted in a need to understand why the distance measurement commonly used in analytical techniques such as Voronoi is not more accurately represented. This research asks: how much difference it makes in analytical results if the more accurate network distance (distance represented as one's actual path) is used as compared to Euclidean distance? The topic of this research, gun-inflicted homicide, is likely facilitated and constrained by location along streets since streets are the transportation conduits in communities and rarely is it possible for individuals to transit between locations in a straight line (Euclidean). If one accepts this point, then spatial analytical techniques such as Voronoi should utilize distance along streets instead of Euclidean distance. However, very few spatial analysis techniques used today allow for the use of actual street network distance. Instead, these techniques, including Voronoi methods, rely heavily upon the use of Euclidean distance even when applied in urban settings. To date there are only a handful of studies which have attempted to compare the differences and effect of distance measurement between the Euclidean and network methods. Only one of these studies utilizes Voronoi diagrams, but not in respect to gun-inflicted homicide on streets (Groff and McEwen, 2006, Yamada and Thill 2007, Maki and Okabe, 2005 and Okabe et. al 2008).

Chapter Two reviews the literature and the theoretical grounds on which the thesis is based. Attention is directed to the sciences of criminology, sociology, and geography from which the foundation of knowledge is derived. Literature describing the evolution of spatial analysis methods used in analyzing crime represented as an x and y coordinate (point) is presented. The major theories underlying the study of crime as it applies to this thesis research question are reviewed. This thesis does not explore the complete foundation of spatial theories applied to the investigation of crime as the theories extend way beyond the main purpose of this investigation. Interest instead is placed primarily upon several relevant topics including distance to crime studies and the importance of the criminals' neighboring environment on the spatial pattern of crime. Literature is also reviewed on how police departments utilize spatial analytical techniques in assisting in prosecuting and solving crime. Chapter Two also discusses the assumptions and frames the research questions.

Chapter Three describes the data and methodology utilized in the study, including the choice of study area, time period, data and a detailed description of the Voronoi method. Performing the analysis required considerable pre and post processing of data.

The full list of steps and software tools involved is too lengthy to be discussed here and will be given full attention in Chapter Three.

Chapter Four presents the results of the comparison and summarizes the findings. The final chapter concludes the thesis with a discussion of how these findings might inform the current process of homicide analysis where intervening street-networks may pose a factor in the analysis approach.

Chapter 2 - Literature Review

This literature review will concentrate on providing context for the research beginning with a discussion of early works on crime mapping and analysis. The discussion emphasizes research on distance to/from a crime's location and the effect of the ensuing environment on crime. These provide a foundation of historical knowledge for more contemporary spatial analysis techniques used in analyzing crime incident locations when each location is represented by the geographic primitive of an x and y coordinate (point) on a street network. This review will be further supported by a discussion of the predominant geographic and criminological theories underlying the spatial analysis of crime and the predominant analytical techniques used by law enforcement officials in point pattern analysis. Special emphasis will be placed on the use of the Voronoi spatial analysis technique and the importance of distance measurement for the application of this (and other) spatial analysis techniques. Concluding the literature review will be a discussion of the research problem. This section highlights the research topic of how the results of Voronoi analysis differ when utilizing network distance vs. Euclidean distance in the spatial analysis of homicides. This review is not meant to be a comprehensive survey of all the excellent works completed in the fields of geography, criminology, or sociology, but instead it provides a sample of many significant works that create a foundation of knowledge for this thesis.

2.1 Geographic Analysis of Crime as Represented as an X and Y Coordinate (Point): History and Theoretical Foundations

Analysis of spatial phenomena or "why things are where they are" and "why things occur where they do" has always been a keen interest of mine. Place is very important in the human decision making process and one would expect it to be even more important to criminals perpetrating homicides. The study of criminal events has received wide attention in the field of crime geography (Harries 1990, Brantingham and Brantingham 1993, Rengert and Walischick 1989, 1985). From the early days to the present the primary form of geographical analysis of crime has revolved around mapping crime incidents represented as *x* and *y* coordinate locations (points). Some of the earliest attempts to map crime date back to the efforts of Jacques Quetelet and Andre Michel Guerry during the early 1800's in France. These pioneers of crime mapping used crime statistics produced in France in 1827 to construct maps of crime totals for various regions in France over time to try and understand crime behavior (Radzinowicz , 1965). Their efforts were successful as their maps depicted underlying spatial structure in crime incident locations such that incidents were not distributed uniformly across urban areas but tended to cluster in specific places. The clustering of crime forms the theoretical foundation that has become known as "Hot-Spot" or "Micro-Place" theory which will be discussed in more detail later but forms a significant theoretical foundation for analyzing street based crime.

It was not until many years later, beginning with the early work of the social ecology movement in the 1920's and 1930's, that a science began to evolve which would help to explain the underlying causes of the spatial structure of crime, causes that were responsible for creating the complex patterns captured by the Guerry and Quetelet map. The social ecology movement as described by Christopher Dunn (Dunn 1980) saw crime as a complex of social ecological processes occurring in urban areas. Dunn elaborates further by stating that crime is a behavioral phenomenon which consists of a complex set of transactions between an individual and his or her environment. These transactions vary in setting, location, time, objects, participants and activities (Dunn 1980).

Studying this complex interaction of the individual with their environment has required the merging of theories, concepts, and analytical tools from a wide range of disciplines such as “behavior setting” from psychology, “physical setting” “land use,” and “density” from geography, “social disorganization” and “social control” from sociology and “systems theory” from administrative science. These concepts describe detailed epistemologies of activity and process (Dunn 1980). One of the classic works in this area is the theory of social disorganization founded by Emile Durkheim and often synonymous with social analysis of crime. Durkheim saw crime as a consequence of the industrial revolution occurring at the turn of the 19th and 20th centuries. During this time, society underwent a massive change in the means and ways of production from an agrarian based economy to an industrialized economy. This transition created greater wealth within the population in general and caused populations to migrate from rural areas to urban areas where manufacturing employment was rapidly expanding. The redistribution of wealth to a wider base of the population created what Durkheim termed an insatiable appetite for continued wealth and property beyond one’s ability to provide the means for satisfying this appetite. According to Durkheim, this created within the individual a state of “anomie,” an inability to meet one’s needs. Anomie led individuals to pursue criminal paths, which violated the acceptable social organization and norms in order to meet their un-met needs. A society which has a high incidence of anomie is thought to be in a state of social disorganization which often correlates with a high crime rate (Krohn M., 2001). Thus, social disorganization as theorized by Durkheim has been widely applied as an indicator of crime. Many studies have examined the social ecology of crime as exemplified in the work of Curry and Spergel (1988) on new immigrant populations in Chicago. Their argument states: “poverty (the social adaptation to chronic deprivation) and social disorganization (the settlement of new immigrant groups) are most strongly related to the spatial distribution of delinquency and gang homicide in Chicago's communities” (Curry and Spergle, 1988).

One of the most noted works in the social ecology movement was by sociologists Clifford Shaw and Henry McKay who set out to explore the relationship of juvenile delinquency with the physical and social characteristics of neighborhoods (Baldwin 1979). As Baldwin (1979) describes, this was one of the first studies which set out to try and explain underlying spatial phenomena which produced the uneven distribution of crime and delinquency across the urban landscape. The authors focused on the unique physical, social, and economic characteristics of neighborhoods. The principal findings of Shaw and Mckay’s work were as follows: (1) delinquency rates peaked in areas adjacent to central business districts; (2) high rates persisted over time in spite of population change; (3) “social disorganization theory (Emile Durkheim)” explained high delinquency rates; and (4) the spatial correlation of various social pathologies was explained by the social conditions in those communities (Harries 1990). Although the Shaw and McKay work identified potential root causes of the distribution of criminal delinquency, there were many inherent problems and criticisms leveled against the work. Dominant among them was the criticism that their approach was anything but ecological since very little of their process of inquiry and investigation was derived from a biological epistemology (Baldwin 1979, Dunn 1980). Also, their approach failed to provide a direct measure of an individual's interaction with their complex local environment. This problem resulted from Shaw and McKay's methodology which involved analyzing data about individuals aggregated into areas. Using aggregate data to make inferences about individual criminal behaviors led to the problem of *ecological fallacy* (Baldwin 1979, Harries 1990). Ecological fallacy states that one cannot attribute characteristics inherent in a group of people to the individuals which comprise the group (Robinson 1950). Despite these shortcomings Shaw and McKay's work was still highly significant, and although not strictly ecological in a true sense of biological epistemology, its approach of examining the environment as a potential root cause of criminal behavior was innovative and had far reaching effects in the analysis of crime for years to come.

2.2 Contemporary Crime Theories: Routine Activities, Environmental Criminology, Hot Spot, and Micro-Place Theory

Emile Durkheim's theory of social disorganization, as examined in the research of Dunn, Curry and Spergel, Shaw and McKay, forms the basis for examining how and why crime occurs where it does. As evidenced in these works numerous social ecology factors contribute to crime and people's insatiable desires for un-met wealth are ever present, but absent in these theories is the examination of “opportunity.” When does a crime opportunity arise? Recent research argues that opportunity exists when all the elements for crime success are present. The theoretical foundation for understanding crime opportunity is found at the confluence of routine activities theory and environmental criminology theory (Groff and McEwen, 2006, Wiesburd et al. 2004, Anselin et al. 2000, Eck et. al 2005, Tita and Griffiths 2005, Brantingham and Brantingham 1993). Routine activities theory posits that crime occurs at a point in time and place where there is an intersection of motivated offenders and victims with the absence of crime suppressors (capable guardians). The works of Cohen and Felson (1979) and Tita and Griffiths (2005) are good examples of research that demonstrates routine activity theory. Tita and Griffiths explain that routine activities theory is primarily concerned with the daily activities (“routines”) of victims and their inherent victimization risk as they attend to life’s daily rituals. These daily routines include the path taken to and from a work location, which could be a walking route, that presents concealment and attack opportunities at certain times of day and where a place guardian may be absent. As this is a routine part of the victim’s normal daily life, threatening elements can easily go unnoticed.

In contrast, environmental criminology theory focuses on the offender and his/her motivations and opportunities to commit crime via their journey-to-crime trip (Tita and Griffiths 2005, Brantingham and Brantingham 1993). Brantingham and Brantingahm (1993) suggest that the offender is highly motivated to take advantage of environmental conditions that favor crime success. Examples of awareness of environmental conditions include knowing what times of day pedestrian traffic patterns are lowest and which streets/places offer the greatest amount of concealment through a lack of lighting, capable guardians, surveillance or dense vegetation. Tita and Griffiths (2005) and Groff and McEwen (2006) point out that both routine activities theory and environmental criminology involve studying the mobility patterns of participants (victims and perpetrators) and the intersections between these mobility patterns. Because the daily movement patterns of people involved in crimes are constrained by and channeled in transportation networks, streets play a role in shaping the spatial distribution of crime. Most criminals use streets to access a crime location (Tita and Griffiths 2005, Lundrigan and Canter 2001, Santilla et al. 2008, and Snook and Cullen). Snook and Cullen cite work completed by Ressler et. al. (1986) which studied the preferred transportation mode of sexual serial killers and rapists, finding that 85% of those who performed "organized" crimes (a crime which they carefully planned) used a vehicle, compared to 65% for those who committed disorganized crimes (unplanned). In both cases the use of a vehicle was cited as the main form of offender transportation which suggests a street network was used to access victims. The study by Groff and McEwen (2006) noted that certain types of crime occur at significant distance from a victim's home, again suggesting the influence of streets in gaining access to victims. In gang-related homicides and homicides in which firearms are the weapons of choice, victims and offenders travel the farthest of any type of homicide from their home locations making distance measured in terms of street routes a potentially important factor and one which should be considered in applying analytical techniques.

In addition to the importance of distance measured along streets, studies have concluded that streets themselves create a setting for crime as they attract certain types of high intensity crime such as drug markets, property crimes, prostitution, and violent crime (Weisburd. and Green., 1996, (Weisburd et al., 2004, Roncek. and Maier, 1991) . When crimes occur in high intensities at certain locations, this is termed a cluster or hot-spot. The National Institute of Justice dedicated a special report on this topic, *Mapping Crime: Understanding Hotspots* (Eck et al 2005). The report notes that the majority of crime occurs in very few places. Research in Minneapolis and Seattle found that the majority of crime events occurred and reoccurred in just a few locations. Sherman, Gartin and Buerger's (1989) study of police call data for Minneapolis, MN showed that just over half of the calls came from 3.3% of the addresses and intersections in Minneapolis.

The term "hot spot" describes a place of high intensity of crime and not a specific spatial entity. The term has been used to describe spatial clustering of crime at a wide range of spatial scales from localized concentrations of individual events to large regions of aggregated events. Sherman, Gartin, and Buerger (1989) discuss the virtues and drawbacks of using data at different scales for hotspot analysis and settle on "police call data" which is recorded at fine scale, an *X*, *Y* coordinate, as the most precise measure for hotspot analysis. However, for some kinds of research, such as the work completed by Block (2000) on activity zones for gangs, a larger aggregated area serves as a more appropriate hotspot definition because gang zones shift over time. The appropriate geographic scale of data for defining hotspots depends greatly on the purpose of the research investigation. For targeting crime prevention policies, highly localized hotspots are appropriate; for understanding broad social and environmental correlates of crime, larger, regional-scale hotspot analysis may be more appropriate.

Several studies have examined the theoretical foundation of crime hotspots (Brantingham and Brantingham 1999, Eck et al. 2000, Weisburd and Green, 1994, Weisburd et al. 2004). The term micro-place refers to locations within larger social systems such as neighborhoods where specific spatial structural components such as buildings, blocks, block faces or streets become sites of high-intensity “hot-spot” crime occurrence (Weisburd et al., 2004). In a study of street crime, hot spots were found to vary in intensity based on the composition of surrounding land uses and facilities at the hotspot location (La Vigne, 1997). These locations were called crime generators and thought to be more prone to street crime because they attracted large numbers of pedestrians which would present potential offenders with suitable targets. A study in Seattle by Weisburd et al. (2004) supported La Vigne’s findings showing that sections of certain street segments where pedestrian traffic was high were locations of high crime incidence. Streets as a location for crime and as an organizing element in the daily lives of offenders and victims should not be overlooked in analyzing crime hotspots. Analysis techniques for street related crime need to make use of algorithms which appropriately measure distance in relation to the to/from distances traveled along street networks by victims and perpetrators.

2.3 Significance of Network Distance in the Spatial Analysis of Crime

The significance of streets as crime generators and routes to crime emphasize the need for research comparing street network and Euclidean measures of crime hotspots. This gets back to the main research question of this thesis: does the measurement of distance, either Euclidean or along a network, significantly affect the spatial analysis of crime patterns and hotspots? Contemporary crime analysis techniques which seek to identify hot-spots mainly employ algorithms based on Euclidean distance. The most widely used of these methods is kernel density analysis (Chainey, Thompson, and Uhlig, 2008, Eck et. al 2005, Anselin et al. 2000, Gatrell et. al., 1996, Wiesburd and Green, 1994, Roncek and Maier, 1991). Kernel density methods enable one to estimate the intensity g(x) of event points in terms of a continuous density surface. This technique describes the mean number of events per unit area measured from a given reference location based on those events falling inside a moving three dimensional window or kernel (Gatrell et al, 1996). The method's purpose is to provide a measure of intensity of the point process that can be used to identify hotspots (areas of high intensity) in the study region.

The method proceeds as follows; a uniform grid of equally-sized grid cells is overlaid upon the areal extent of the study region containing the point process event locations i.e. homicides (Bowers et al, 2004). A circular window centered on each grid cell is moved across the study area. The density of crimes at each grid cell is calculated based on the kernel weighting function represented in its standardized form below. A key parameter is the radius of the window which is represented by a user specified bandwidth (boundary) parameter τ.

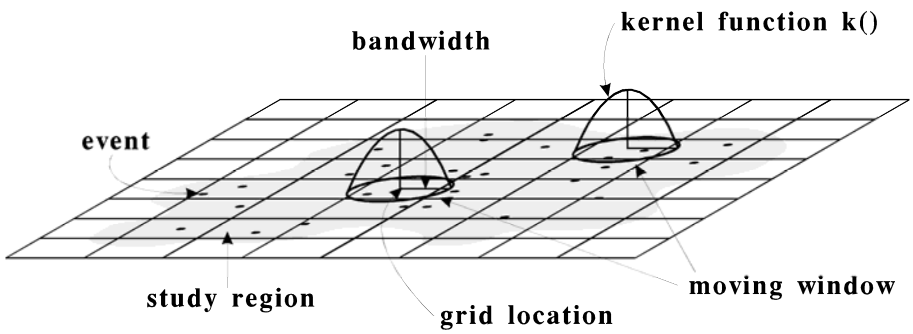

Figure 1. Kernel density estimation

$$\hat{\lambda}_{\tau}(s) = \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{s - s_i}{\tau}\right)$$

(Gatrell et al. 1996)

k – kernel function

s – grid reference point

si – point event being measured for intensity

t – bandwidth

R – study region (not part of formula)

Each point that falls within the bandwidth of the kernel function as it progressively moves from one grid reference point to the next contributes to an estimate of the intensity of the point process (Gatrell et al. 1996). When the estimation is completed for all cells within the matrix, a continuous surface of event point intensities is produced illuminating hotspots as areas of high intensity.

In most applications, kernel densities are estimated based on Euclidean distance; however conceptually it is straightforward to incorporate network distance by having the distance term (s) in the above equation represent network distance. Although this is conceptually straightforward, it is computationally complex and no readily available software exist for kernel density application using network distance.

There are other spatial analysis methods for identifying crime hotspots including the Spatial and Temporal Analysis of Crime (STAC) method which fits a standard deviational ellipse to clusters of crime points. The ellipse provides additional information based on size and orientation that may suggest a directional component to the crime hotspot (Block, 2000, Chainey, Thompson and Uhlig, 2008). Note that the STAC method works with Euclidean distances in computing the standard deviational ellipse. Chloropleth mapping has also been used to identify hotspots. This is accomplished by arbitrarily creating a rating system based on the raw numbers of crimes which occur in areal units such as census tracts, police districts or political wards. The areal units are shaded based on that rating. A similar approach can also be applied to map symbols. Locations with a high number of occurrences have a proportionally sized symbol, or the symbol can be colored -- i.e. green for low, yellow for medium, and red for high (Eck et. al 2005). A street segment (polyline) with a high number of crimes occurring along it might also be treated in this way using a heavier line weight and color to indicate the intensity of crime.

Most hotspot methods rely on Euclidean distance. The only known technique currently available to law enforcement agencies which utilizes both Euclidean and network distance in its calculation and also provides a similar visual representation of spatial clustering is the Voronoi diagram. Given a set of points in a 2-dimensional space, the Voronoi diagram identifies for each point the area closer to that point than to any other (Figure 2).

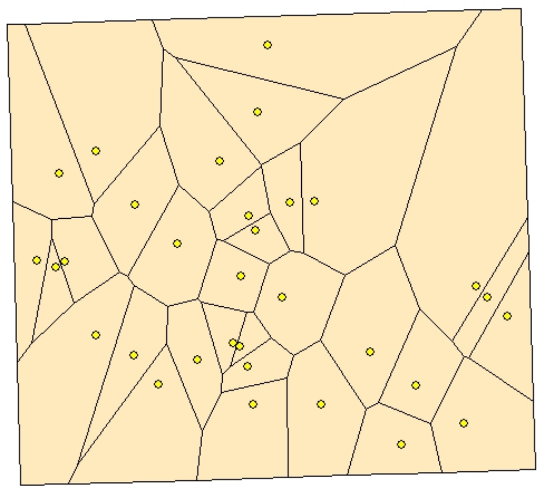

Figure 2. Euclidean Voronoi

Most Voronoi diagrams are based on Euclidean distance, but network distances can be incorporated. Voronoi diagrams were first proven as a measure for hot-spot analysis in the work of Mollenkopf, Goldsmith, McGuire, and McLafferty (2003) in the New York City, Police Department's Crime Mapping, and Analysis Application (CMAA) program.

In that application, each crime event location served as the Voronoi diagram generator point. The area surrounding each event location was distributed into mutually exclusive zones representing the area closest to the generator point. The sizes of the Voronoi diagrams are used in identifying hotspots. If a hotspot exists, the ensuing diagram would have a tightly clustered center of small Voronoi diagrams. Outside the hotspot, the polygons are larger in size because points are farther apart. Thus, the sizes of diagrams serve as an indicator of clustering. Using a Voronoi diagram in this way could provide decision support in allocating patrol resources as the diagrams could easily be overlaid onto a neighborhood map displayed in a geographic information system (GIS) depicting the areal extent of high intensity activity.

Only recently have researchers developed algorithms to construct Voronoi diagrams based on a street network. Atsuyuki Okabe, Kei-Ichi Okuniki, Shino Shiode at Tokyo and the Nagoya Universities in Japan developed a tool for accomplishing this called Spatial Analysis on Networks (SANET). Since its inception this tool has primarily been used for market area studies and for optimal siting of businesses and transportation stations (Okabe and Okunuki 2000). Okabe's technique is different from the traditional method for creating Voronoi diagrams which is based on Euclidean distance. A network based Voronoi diagram treats the crime event, homicide in the case of our data set, as a generator point. The Voronoi diagrams are defined as sub-networks representing the shortest paths to each generator point or homicide event. The Okabe method utilizes a graph solving algorithm for determining the shortest path known as the Dijkstra algorithm. It accomplishes this task by first converting the streets represented by polylines (edges) in a GIS and the intersections represented as nodes (points) in the GIS into the mathematical construct of a graph or network. Once converted the Dijkstra algorithm can be applied to extract network distances utilizing the homicide location as a generator point in the Voronoi calculation. The result is represented as a network of associated streets optimized as the shortest paths to the homicide location.

The result of the network Voronoi method is very different visually from what is produced by a Euclidean Voronoi analysis but one which can be very powerful in terms of analyzing homicide. The network Voronoi diagram not only represents small groupings of street segments that identify hotspot locations but also provides a shortest path analysis to the scene of the crime.

Other research into network spatial data analysis has only recently begun to develop. Researchers in landscape ecology have used network spatial techniques to test for spatial autocorrelation (clustering) in invasive plant species that colonize linear landscape structural features, such as roadsides or canal edges. A study completed by Maheu-Giroux and Blois 2006, on the ecology of Phragmites australis invasion in networks of linear wetlands in southern Quebec' found equal evidence of clustering to be present when using either a network distance based K function or Euclidean distance based K function analysis. The Euclidean K function was also compared to the network K function in an analysis of vehicle traffic accidents in Buffalo NY. The finding indicated that there was a noticeable difference as the Euclidean based K function over-detected the presence of spatial clustering (Yamada and Thill, 2004). These mixed results show that there is still much more work to be done to determine the full value of using network distance in comparison to Euclidean distance. As the K function is quite different from the Voronoi method, only in a study designed specifically for Voronoi can the effect of the two distance measures on the Voronoi method be fully understood.

2.4 Does Network Distance Matter in Voronoi Analysis? The Current Study

The literature review has clearly demonstrated the importance of streets as a place where homicide occurs, and it emphasizes the likelihood that the effect of distance measurement on spatial analysis methods for crime analysis is a non-trivial matter deserving attention. To provide greater insight into how streets and street-distance (network) may affect the analytical results of crime analysis, a comparison was conducted using the Voronoi diagram technique. Specifically, I ask: are the areal extent and lengths of streets contained in a Euclidean Voronoi diagram vs. a network Voronoi diagram statistically similar or dissimilar? The purpose of this research is to extend our understanding of the effects of distance measurement on spatial analysis techniques and to shed light on the potential importance of this for law enforcement agencies who utilize these techniques in analyzing and prosecuting crimes.

Chapter 3 – Data and Research Methodology

The study area for this research is St. Louis, MO, a large Midwestern city that has long been plagued by crime-related violence. Historical crime data that describe point locations of crime events are readily available for St. Louis, making it possible to implement the Voronoi methods to be tested in this research. Like many cities, St. Louis has a grid-like street pattern modified to reflect land use patterns and physical features. With a street network typical of those found in industrial cities in the U.S., St. Louis provides an ideal study area for investigating differences in Voronoi results based on Euclidean and network distances. St. Louis City encompasses the downtown city center of St. Louis with neighborhoods to the North, West and South. At the time of this study, St. Louis was a highly racially segregated city: The population of the southern section was predominantly white and that of the northern section was predominantly black (Figure 3). (Jones, 1976).

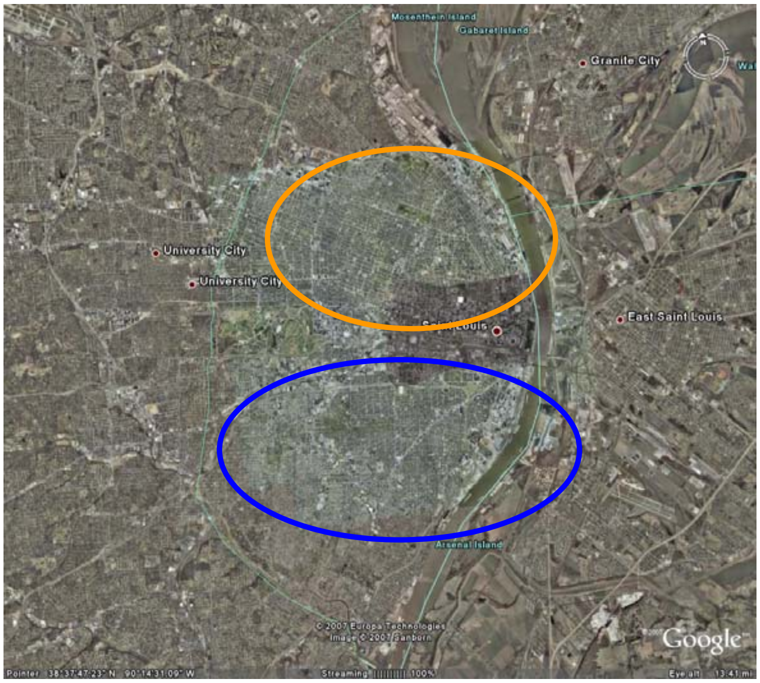

Figure 3. Northern and southern neighborhoods **(Predominantly Black)**, **(Predominantly White)**

3.1 Data Description

The data utilized in the analysis were acquired from the National Archive of Criminal Justice Data. The data set was provided by two political scientists, Carol Kohfeld and John Sprague (Kohfled & Sprague, 1991) who used the data in a series of crime studies performed at the University of Missouri, St. Louis. The data encompassed arrest and crime reports from the St. Louis Police Department for the years 1972-82. Crime incidents were divided into two categories: all Uniform Crime Reporting Program Part I crime reports, including arrests, and Part II felony arrests. The St. Louis Police Department also provided geographical *x* and *y* coordinates corresponding to the longitude and latitude where each crime and arrest took place. With the point coordinate information, the data were well suited for this research study.

The variables collected in this data set included offense code, census tract, police district, police area, city block, date of crime, time crime occurred, value of property taken, and *x* and *y* coordinates of crime. Because of the very large size of the data set, I decided to select a more limited subset of data for analysis. The City of St. Louis is divided into three area patrol stations and 9 police districts. Only two districts were chosen for analysis, police patrol districts 8 and 9 (Figure 4). A subset of data was also chosen based on type of crime. This study only focuses on gun inflicted homicide. As the purpose of the study was a comparison between analytical techniques, it made sense to focus on a limited subset of crime data.   Gun-inflicted homicide was chosen because it was cited in the literature as a crime that is likely to occur away from the offender's home location implying that street travel might be involved in committing the crime or that streets may be important sites for crime occurrence.(Groff and McEwen, 2006). Police patrol districts 8 and 9 were selected because they had a high concentration of homicides for the time periods being analyzed.

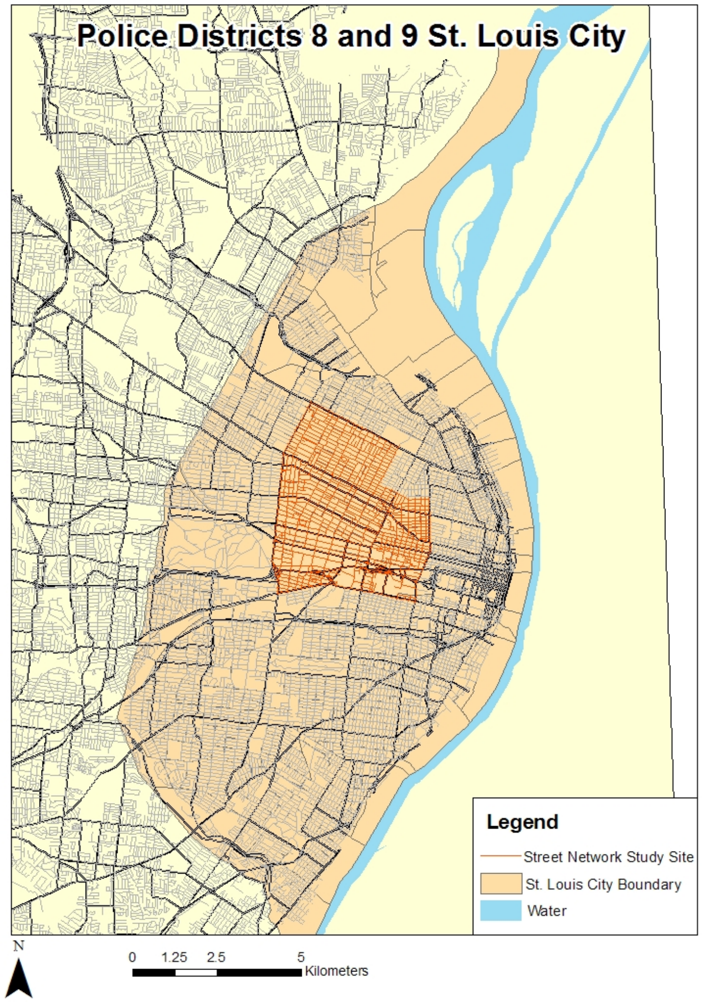

Figure 4. Research study site

Street network data was collected from the Environmental Systems Research Institute (ESRI) Street Map data set. The Street Map data set is based on a U.S. Census Topologically Integrated Geographic Encoding and Reference (TIGER 2000) road data set with a scale of 1:20,000. Additional geospatial data used for thematic map display was also collected from ESRI datasets provided with the company's ArcView version 9.2

GIS software. This data included state boundary delineation, surface hydrology, and St. Louis City boundary information. It provided a base map for displaying the street network and homicide location data which were used in the analysis. All geospatial data including the *x*, *y* coordinate data for homicide location were projected into Universal Transverse Mercator (UTM) Zone 15 North American Datum 1927 (NAD27) which is a cylindrical projection where the cylinder is longitudinal along a meridian rather than being associated with the equator (ESRI, 2006). This results in a conformal projection which minimizes the distortion in area and distance within the study area. The UTM projection is widely used in the United States by state and regional planning agencies because of its properties which minimize these distortions. NAD27 was used because positional accuracy was not a concern in the research. What was most important was that all the data had the same projection. As the data was extensive and much of the original data came already in the NAD27 projection no attempt was made to re-project the data in the more recent NAD83 projection.

Three years of homicide data were selected for the purposes of comparison: 1970, 1972 and 1982. The overall goals were to compare several different homicide patterns and to make the year to year comparison as uniform as possible. These three years were selected because they had the least variation in the total number of homicides among them. Thus, differences between the years are related primarily to differences in homicide locations, not in the overall numbers of homicides. The year 1970 had the greatest number of homicides with 36, while 1972 and 1982 had 32 and 33 respectively.

3.2 Research Methodology

The research methodology seeks to provide a comparison of the length of street segments and area of polygons encapsulated by Euclidean Voronoi diagrams to that of network generated Voronoi diagrams for homicide locations in each year. Both types of Voronoi diagrams had to be generated for each year and the results compared. Euclidean Voronoi diagrams can be generated using ArcINFO GIS, however it is only possible to generate a network Voronoi diagram through the use of the specialty software product, SANET, designed to work within ArcGIS.

The standard Euclidean Voronoi technique in ArcINFO proceeds by assigning all locations in the Euclidian plane to the closest event within the homicide point-set. After locations are assigned to each point, the result is a set of regions (Voronoi polygons), centered on each point, that are collectively exhaustive and mutually exclusive except along shared boundaries (Okabe, Boots, and Sugihara, 1992). Each region identifies the locations closer to its generating point (in this case, homicide location) than to any other generating point.

In comparison, the network Voronoi algorithm based in the SANET software, proceeds as follows: the set of homicide locations located on the street network each serve as a generating point. The network Voronoi region around each generating point is constructed by creating shortest paths outward along the street network. The paths stop when the network distance from the generating point equals that to some other point – i.e. the point reached, *pi,* is equidistant from two or more generating points. In effect, the network Voronoi regions are constructed by “growing” outward from generating points along the street network   As in the Euclidean Voronoi, the network segments contained within these shortest path calculations create sets of streets (regions) which are collectively exhaustive and mutually exclusive except at the boundaries between networks (Okabe et al., 2008).

After the network and Euclidean Voronoi diagrams were created for each year, it was necessary to compare the two diagrams to assess the degree of similarity/difference. Two indicators were used for comparison: The areal size of each Voronoi region (polygon) and the total length of street network within each polygon. Both indicators measure the sizes of the Voronoi polygons, and as discussed in Chapter 2, size is an important indication of the spatial organization of crimes. Calculating the area size of each Voronoi region is straightforward for the Euclidean Voronoi, but more complex in the network case. On the other hand, calculating the length of encapsulated streets in each Voronoi region is straightforward in the network case, but more complex for the Euclidean Voronoi. In some cases, complex GIS operations were required. Sections 3.3 and 3.4 explain the procedure for calculating the two indicators with each type of Voronoi diagram, beginning with the Euclidean Voronoi.

3.2.1 Procedure to generate Euclidean Voronoi diagram and calculate the length of encapsulated streets and area of Voronoi regions

Calculating the area and street lengths contained in the Euclidian Voronoi regions requires the utilization of the ESRI ArcGIS Desktop and ArcINFO workstation GIS. To generate the Euclidean Voronoi diagram for each of the three years of data, I used the ArcINFO command "Thiessen_arc <in_cover> <out_cover>" The ensuing diagram was then fitted to police districts 8 and 9 utilizing the ArcGIS "CLIP" command. The result was further trimmed to fit within the full extent of the Euclidean Voronoi diagram which in all three years analyzed was smaller than the full boundary of police districts 8 and 9. This small discrepancy was due to boundary effects. This became the new area of analysis for comparing the Euclidean Voronoi and network Voronoi. The new area encompassed all generating points and all Euclidean polygons, so it was appropriate for the analyses. Unfortunately, there was no capability with in the ArcINFO "Thiessen arc" command to specify a large enough bounding box to encompass all of police districts 8 and 9 so this was a limiting factor. Once this operation was completed, the area contained within each Euclidean Voronoi region was determined. ArcGIS calculates this as a standard part of the attribute table for this type of polygon output.

The length of streets contained within each Voronoi region was more difficult to determine and required further processing. First, the street network had to be overlaid on the Voronoi diagram, and street segments falling within each zone were selected for measurement. ArcGIS Desktop does not have an easy way to accomplish this. The standard "CLIP" tool does not provide output for each individual zone. Instead, a command in ArcINFO workstation called "SPLIT" was used to accomplish this task. The resulting output (Figure 5) contained polyline coverages of only those streets which fall within each individual zone. Since the "SPLIT" command can only produce a coverage output for each individual Voronoi region this was a very computationally intensive task as it needed to be repeated for each homicide location in each year. As a result 1970 required 36 coverages, 1972 required 32 coverages and 1982 required 33 coverages. Once these polyline coverages were created, street length was acquired from the coverage attribute table. The end result was a "street length" measurement for each Euclidean Voronoi region.

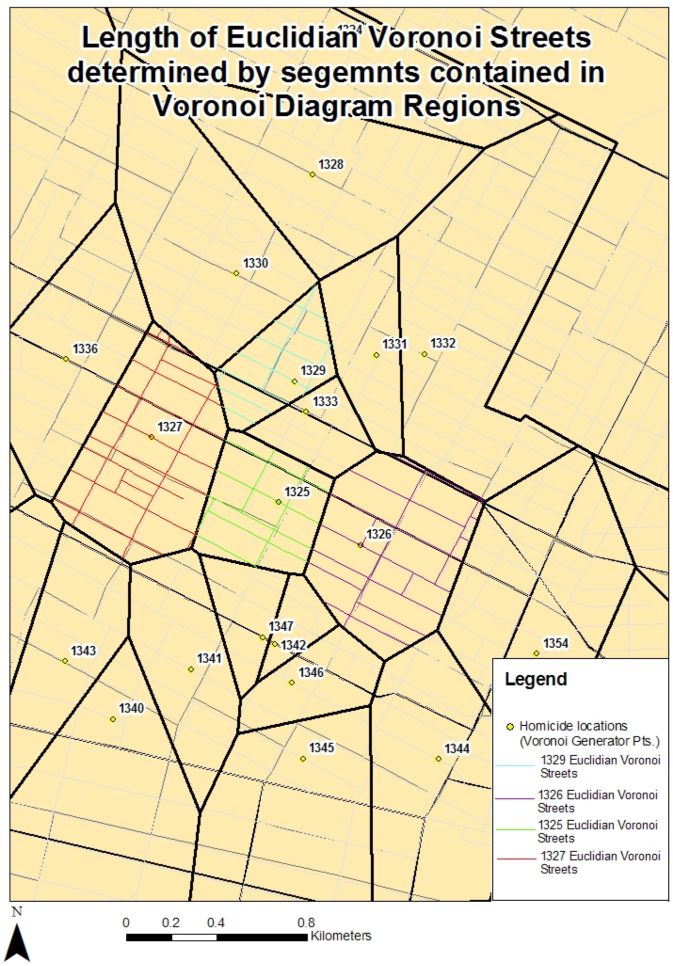

Figure 5. Street lengths contained within each zone of the Euclidian Voronoi diagram created using the ArcINFO workstation "SPLIT" command.

3.2.2 Procedure to generate Network Voronoi diagram and calculate the length of encapsulated streets and area of network Voronoi regions

The network Voronoi diagram consists of webs of street segments emanating out from each generating point (i.e. homicide location) that represent the locations closest to that generating point in terms of network distance. Both the total street length and area size had to be computed for each Voronoi region. To accomplish these tasks required the utilization of the ESRI ArcGIS Desktop and two custom extensions to the ArcGIS environment: SANET, which was used to generate the network Voronoi diagram, and a software extension called HAWTH'S Tools. HAWTH'S Analysis Tools version 3.26 extends the functionality of ArcGIS with spatial statistical functions commonly used in ecology. This software was used in determining the area of each Voronoi region. A function known as the minimum convex polygon (MCP) generation tool was used to convert each web of streets into a polygon whose size could be calculated. A minimum convex polygon is a technique for fitting a bounding box around a point distribution when calculating area or intensity of the encapsulated point process. The polygon is considered the tightest boundary around a set of points. MCP was utilized to calculate the area over which the network Voronoi street segments extended – in effect, the area bounding the ends of the street segments. In producing the network Voronoi diagram utilizing SANET, one of the interim GIS layer outputs is a file of intersections and end-points for each generating point. Utilizing this GIS layer file as an input to the MCP function in HAWTH'S tools, a minimum convex polygon was created which encapsulated all the street segments in the network Voronoi diagram. Once the polygons were generated for the network Voronoi segments the areas of these polygons were obtained from the newly created MCP GIS layer attribute table. This like the “SPLIT” function was very computationally intensive as it also required producing a single MCP for each of the individual Voronoi regions for each year. An example of the application of the MCP is depicted in Figure 6. Note how the boundaries of the MCPs connect the endpoints of the street segments associated with a particular homicide location in the network Voronoi diagram.

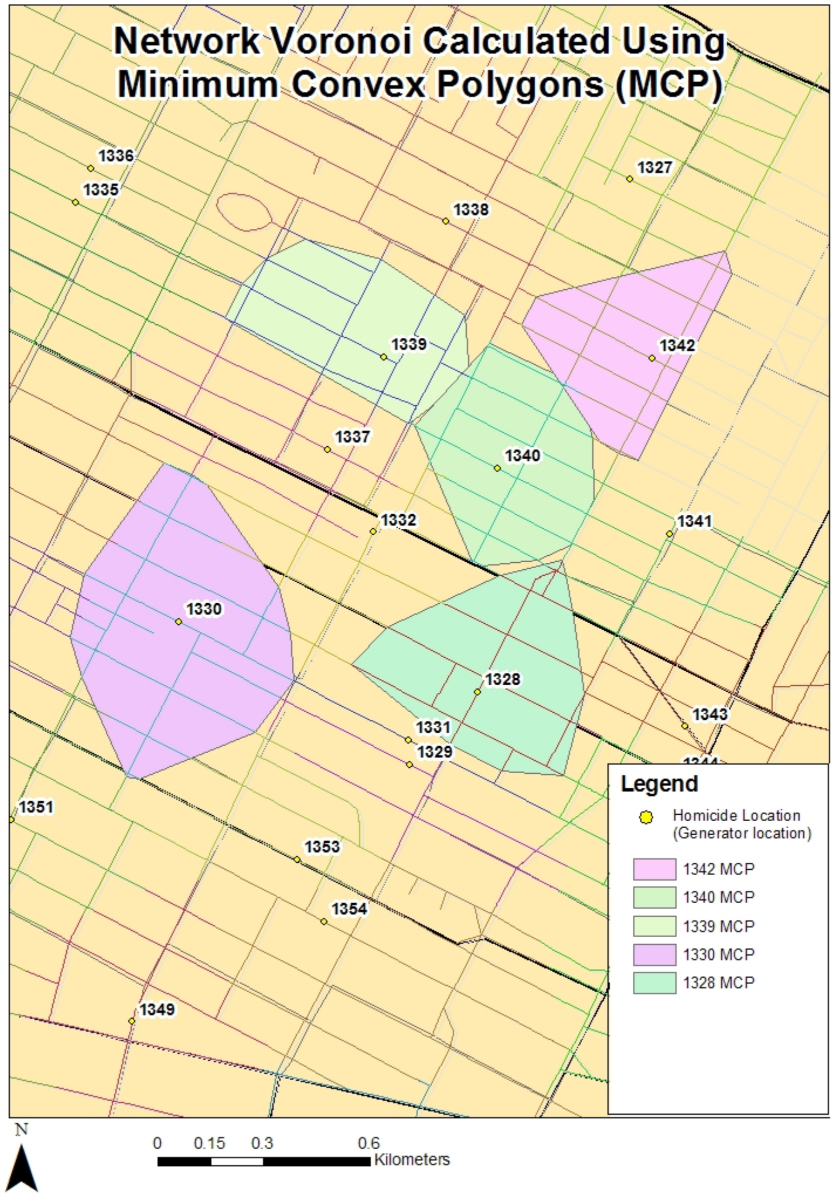

Figure 6. Minimum convex polygons representing areal extent of network Voronoi regions.

Extracting the network Voronoi street lengths to compare with the Euclidian Voronoi street lengths was simpler than extracting area. The polyline output produced by the SANET extension in ArcGIS provides length as a standard attribute. An example of the Network Voronoi polyline output produced by SANET is illustrated in Figure 7.

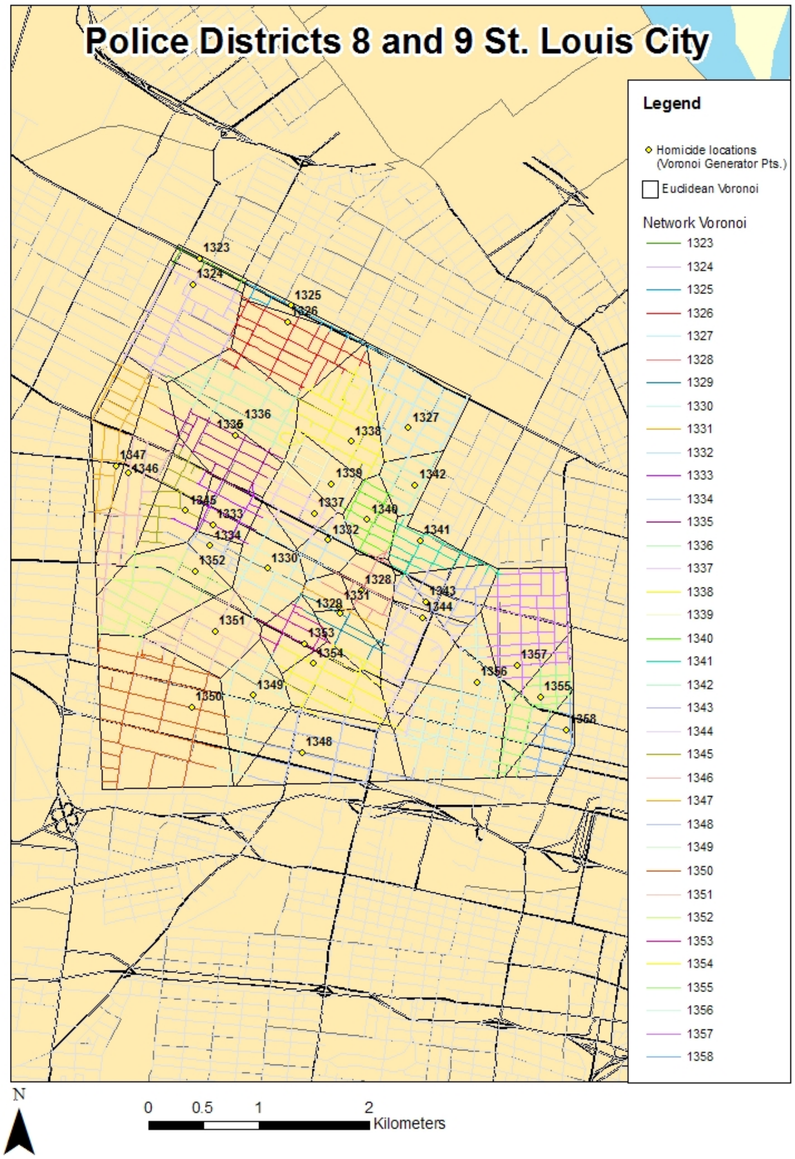

Figure 7. Network Voronoi result from SANET overlaid on corresponding Euclidean Voronoi result

Figure 7 shows the network Voronoi diagram overlaid on the Euclidean Voronoi diagram and thus provides a visual comparison of the two methods. Under close examination one can see that the network Voronoi street segments generated for each homicide location create street groupings which extend beyond, or fall inside the boundary of the corresponding Euclidean Voronoi. Some of the network Voronoi regions are irregularly shaped. It is precisely these differences that this research aims to evaluate.

3.2.3 Comparison of Euclidean and network Voronoi in terms of street length and area

Two well-known statistical measures were used in comparing the area and street length associated with the two analytical techniques, the product moment correlation (Pearson's *r* statistic) and root mean square error (RMSE). The analysis was performed using Microsoft Excel. These two statistical techniques were selected because they provide a measure of correlation (Pearson’s *r*) and similarity or dissimilarity (RMSE) which represents the correspondence between street length and area measurements between the network and Euclidean Voronoi methods. Based on these measures, conclusions can be drawn about the degree of difference between findings for the analysis of homicide on street networks.

Pearson’s *r*, developed by Karl Pearson and often referred to as the product-moment correlation coefficient, measures the degree of linear association between two variables (Blalock, H. M., 1979). This statistic is expressed as a value between -1 and +1. A value close to +1 suggests the variables are strongly, positively and linearly correlated, indicating that a change in one variable is associated with a constant increase in the other variable. If the result is closer to -1 then the variables are negatively correlated meaning that an increase in one variable is associated with a constant decrease in the other variable. It is important to note that Pearson's *r* provides no measure of causality in regards to why the variables affect each other in a positive or negative manner.

In this research, I expect to find a strong, positive correlation between the network and Euclidean Voronoi measurements. The two sets of measurements should be very similar, and therefore we should see Pearson's *r* values close to +1.0 for both area and street length.

The second statistical measure is the root mean square error (RMSE) which is the square root of the sum of the mean squared error. If E_i is the measurement of street length (or area) for polygon *i* in the Euclidean Voronoi, and N_i is the corresponding measurement for polygon *i* in the network Voronoi, the RMSE is (Cromley & McLafferty, 2002):

$$\sqrt{\sum \frac{(E_i - N_i)^2}{n}}$$

The result is expressed in the units of the measure being compared, for example, kilometers for street length. The magnitude of the RMSE determines relative similarity or dissimilarity between the methods. The larger the RMSE, the larger the difference in street length or area measurements between the network and Euclidean methods.

Through a combination of the results from these two statistical tests and plots graphing the differences between the areas and lengths, the two spatial analytical techniques are compared.

In addition to these two statistical measures, I created maps for street length and area for each year showing the spatial distribution of error. The difference between measurements was depicted on bar graphs centered on the corresponding homicide location. Maps provide a visual display of the places where network and Euclidean measurements are similar (bar graphs are approximately equal) and places where the two differ (bar graphs are not equal). This information is useful in identifying characteristics of local street networks that result in significant differences in measurement.

Chapter 4 – Discussion of Results

This chapter describes the findings of a comparison between the two Voronoi diagram techniques -- one which is calculated using Euclidean distance (straight-line) and the other which is calculated using shortest-path distance measured along a street (network distance). The data set analyzed with these two techniques consists of *x* and *y* coordinates of gun-inflicted homicides locations in St. Louis, MO for the years 1970, 1972 and 1982.

The first section describes the results of Pearson's *r* and root mean square error (RMSE) statistics which were used in measuring the difference between the two techniques with respect to length of street segments and area. The sections which follow seek to provide a visual comparison using graphs and maps to highlight specific instances of similarity/ dissimilarity while also attempting to suggest possible underlying cause needing further investigation.

4.1 Analysis of Length and Area using Pearson's r and RMSE

Overall, statistical measures of similarity/difference indicate that there is little difference between the area and length results produced by the two methods. The Pearson's *r* coefficients calculated for the three years show that the methods are highly correlated as all values are +.90 or better for both area and length (Tables 1 and 2) These high, positive values indicate a strong linear association between network and Euclidean measurements of street length and area for the Voronoi diagrams. The measures of association vary slightly from year to year (Table 1 and 2), but no clear pattern is represented in the statistical measures. The goodness of fit is the strongest in 1970, and somewhat weaker in 1972 and 1982. This weaker fit in 1972 and 1982 results from some unusual abnormalities which will be discussed in the following sections.

Area Sq. KM		
Year	**Pearson's *r***	**RMSE**
1970	0.96	0.49
1972	0.90	0.97
1982	0.95	0.83

Table 1. Area comparison using Pearson's r and Root Mean Square Error (RMSE)

Length KM		
Year	**Pearson's *r***	**RMSE**
1970	0.98	0.31
1972	0.90	0.78
1982	0.95	0.64

Table 2. Length comparison using Pearson's r and Root Mean Square Error (RMSE)

The calculated root mean square error (RMSE) values show the difference between the two methods in the actual units of measurement for length and area, kilometers and square kilometers respectively (Cromley and McLafferty, 2002). The RMSE values in Tables 1 and 2 indicate that the two measures are not very different for both length and area, with all results differing by less than one unit of respective measurement. According to both measures, the fit is weakest in 1972. The RMSE values for 1970 indicate the best fit especially in the comparison of length measurements while area measurements have a slightly poorer fit. This difference in fit could be due to several factors. These include limitations in the technique utilizing the minimum convex polygon (MCP) to produce area measurements or the effects of underlying street geometry and intervening landscape structures.

Figure 8 shows an example of how the limitation of the MCP method for producing area measurements may have contributed to measurement error affecting the research results. When the underlying network Voronoi had an unusual shape, the MCP over-estimated the area encompassed by the street network, leading to overlap between polygons.

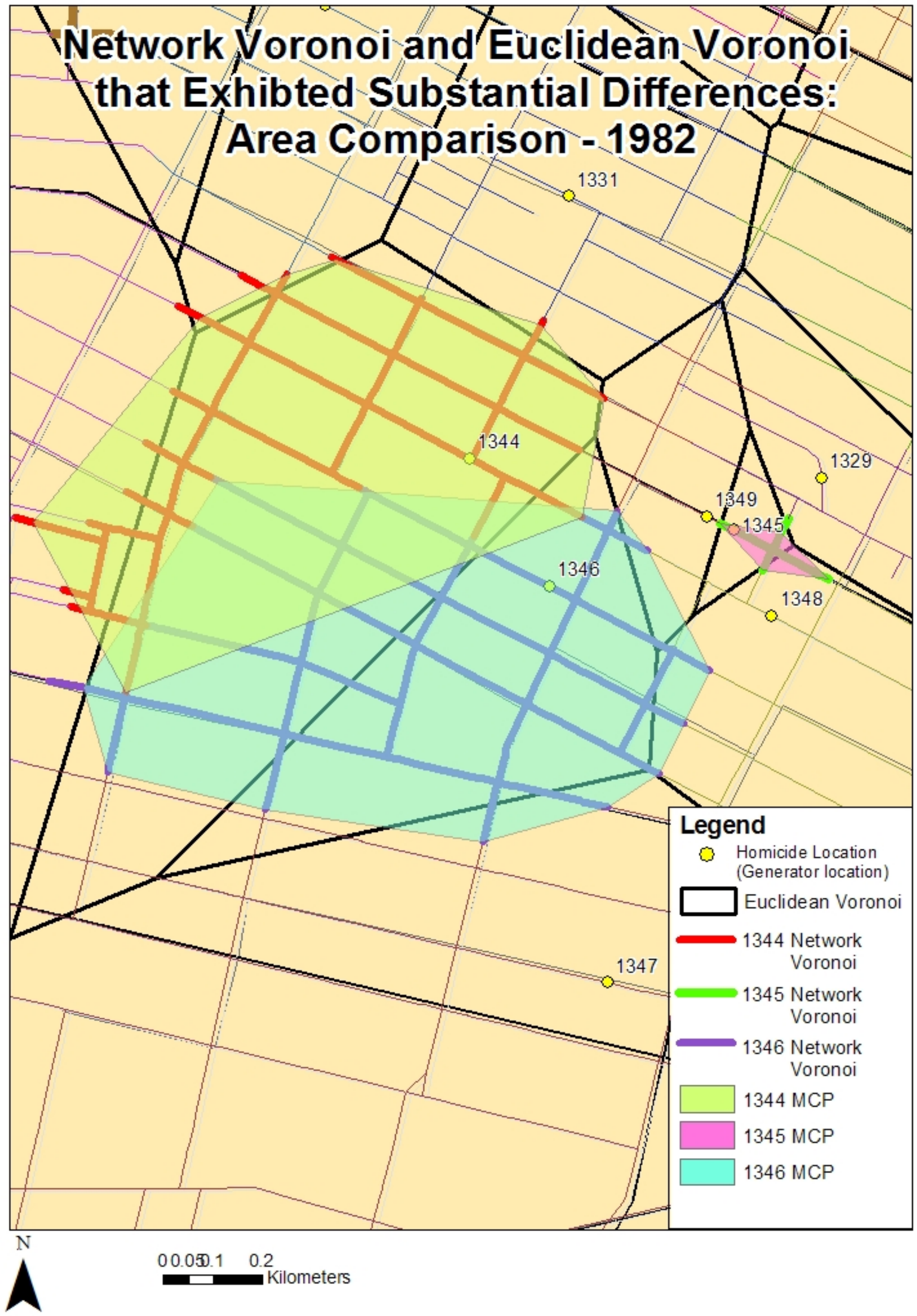

Figure 8. Minimum convex polygon (MCP) over-estimation of area

Figure 8 shows an example of this in the 1982 homicide data. Specifically, for homicide number 1344, the MCP produces a convex shape rather than concave shape, and the convex shape overlaps the adjacent polygon. This is an outcome of the MCP procedure in Hawth's Tools. The SANET software was not designed to produce area measurements for network Voronoi, so the MCP procedure had to be used. The only alternative to achieve greater accuracy would have been to manually digitize a boundary around each network Voronoi which was not a practical solution. Fortunately, these overlapping polygons were rare and therefore are not likely to have had a major impact on the results.

Police districts 8 and 9 where the study took place held a wide variety of street geometries based on the intervening landscape. These geometries have important effects on the observed differences between the Euclidean and network Voronoi. They affect the measurement of shortest path by network Voronoi which in turn can lead to local differences in results between the two methods. Examples of how street geometry may have affected the measurement of length are presented later in section 4.2.2.

The Pearson's r and RMSE findings are not surprising as the two methods are essentially designed to do the same thing -- an optimal distance/region analysis. What these statistics do not readily reveal is the effect of scale on producing differences in the methods. In research done by Maki and Okabe (2005) comparing street length measurements generated by the two methods, the scale of the analysis played a significant role. They found network Voronoi street lengths were substantially different from Euclidean Voronoi streets lengths at short distances (500 meters and under)

(Figure 9.). In contrast, at longer distances the two methods produced very similar street length measurements.

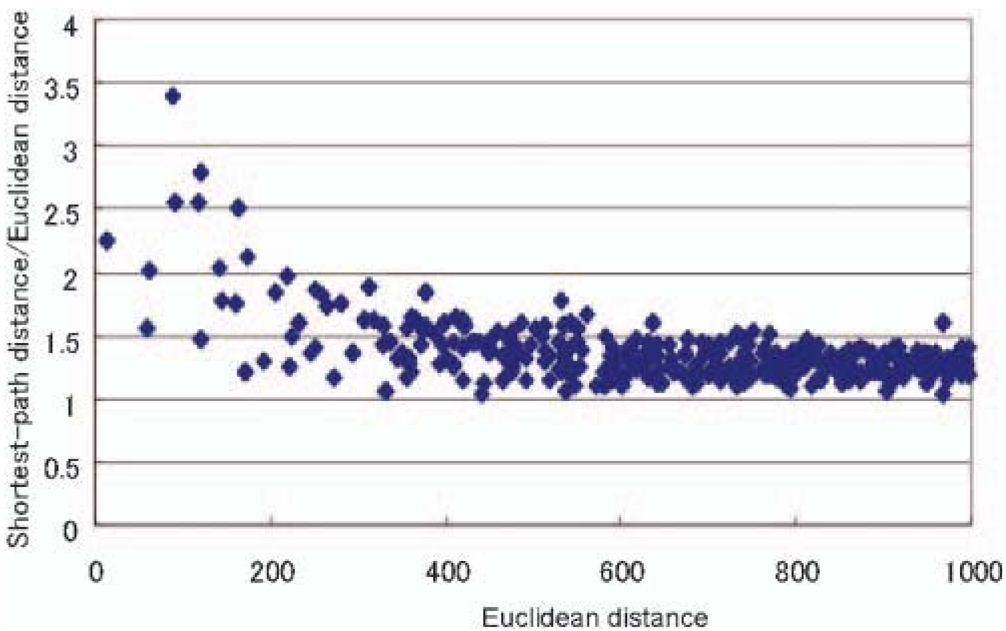

Figure 9. Ratio of the shortest-path distance to Euclidean distance on the street network in Kokuryo, a suburb of Tokyo (from Maki and Okabe 2005).

Of the three years of data analyzed in this research only one network Voronoi resulted in a distance below 500 meters (Figure 10) and based on visual examination the result was substantially different. The remainder of network Voronoi diagrams generated were 500 meters and above. Thus, the close fit observed in this study supports Make and Okabe's (2005) finding, indicating that longer distances do not vary much between the network and Euclidean Voronoi methods.

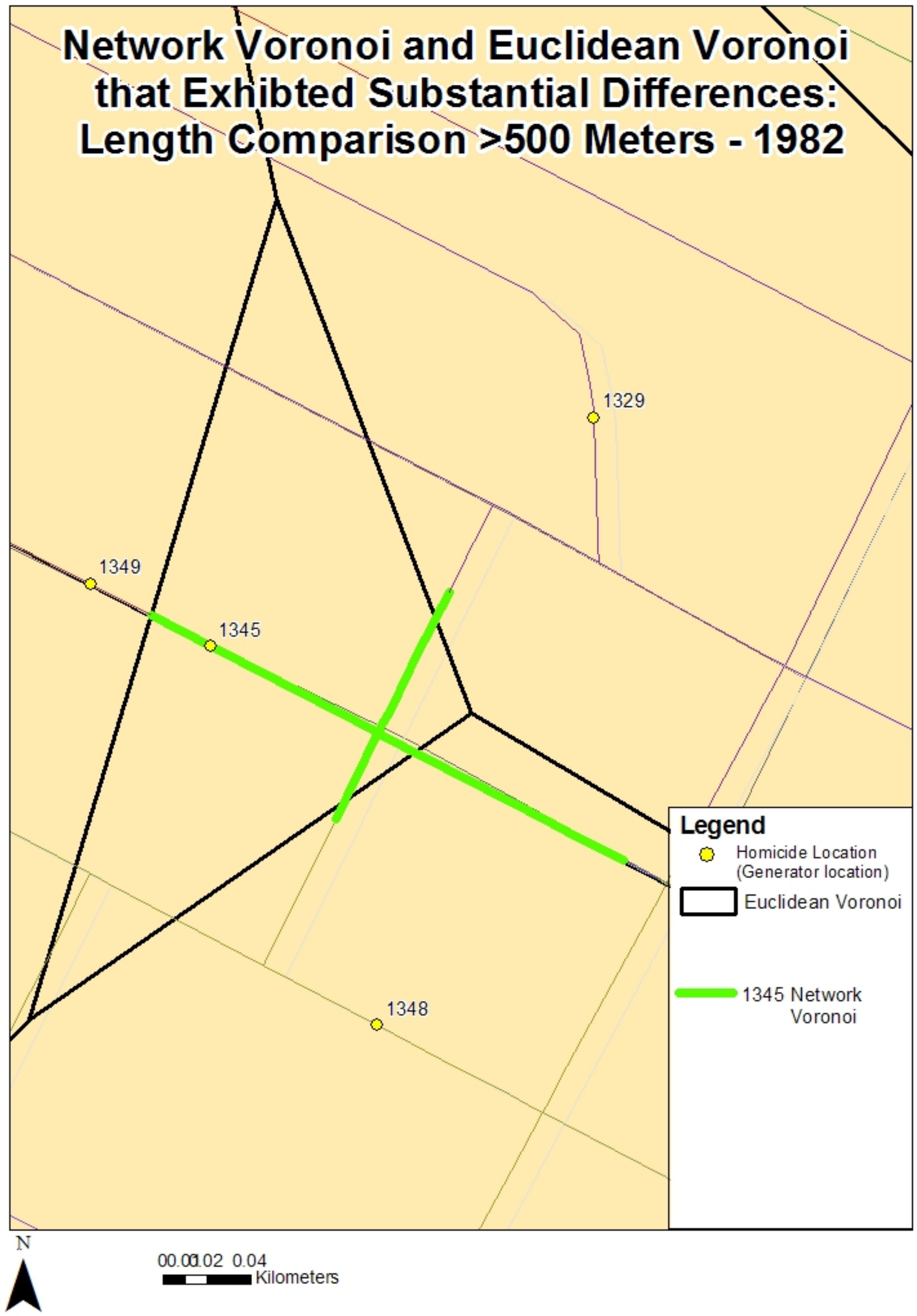

Figure 10. Difference comparison between Euclidean Voronoi and Network Voronoi when Network Voronoi is less than 500 meters in overall length

4.2 Differences in Fit Among Locations

Overall the network and Euclidean methods produce very similar results, but in each of the years analyzed there were a few data points which exhibited substantially different area and length measurements. Utilizing opposing bar graphs which organized the data from smallest area/length to largest showed the overall trend in difference, if any, among the data points. Maps depicting the opposing bar graphs helped to identify which locations resulted in wide differences between the two techniques. In the following paragraphs is a presentation of the common differences found among these locations.

4.2.1 Differences among fit for location between area comparison of the two methods

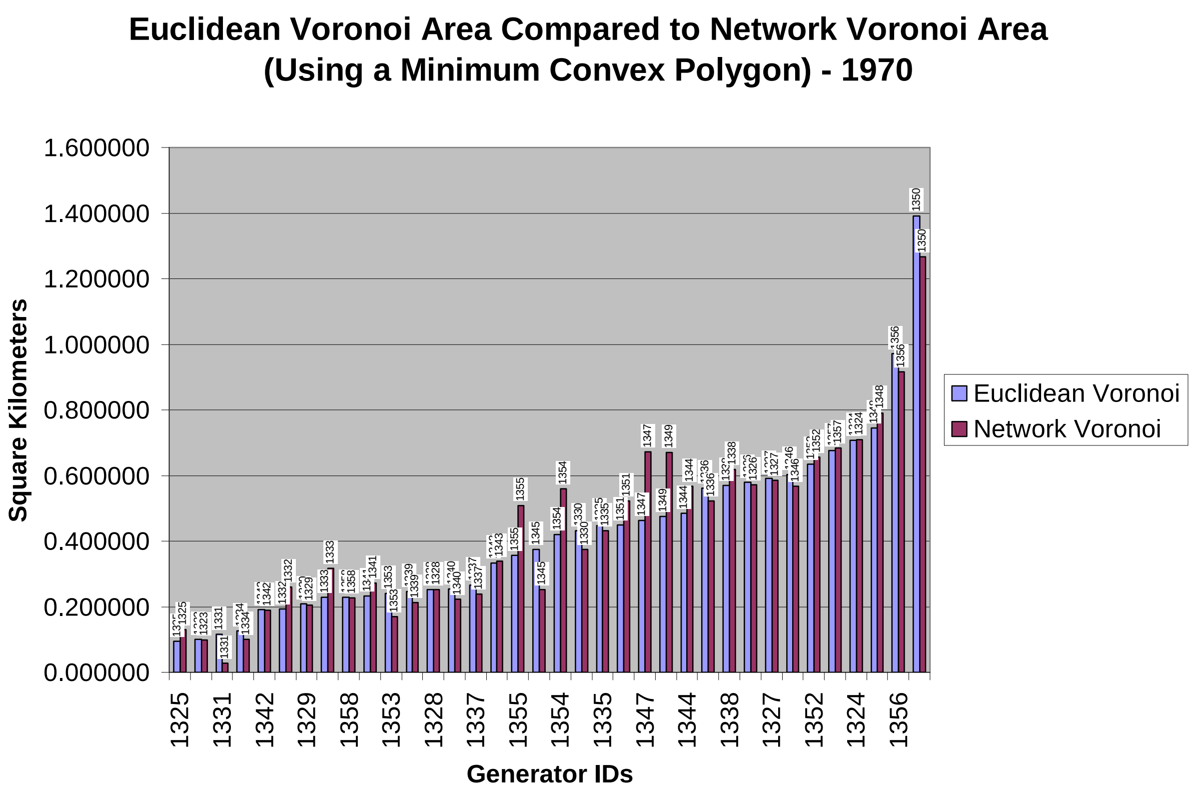

Figure 11. Euclidean Voronoi area compared to network Voronoi area using Minimum Convex Polygon – 1970

Figure 11 shows that for the 1970 data there was little trend in the difference in area measurements based on the size of the Voronoi polygon. The disparity in measurements is proportionally large for some large polygons, but also for some medium and small-sized polygons. Specific examples can be viewed in the map in Figure 12. Many of the large differences are related to the problem mentioned earlier that was caused by encapsulating an irregularly shaped network Voronoi with an MCP. The red circle on Figure 12, identifies an example of this problem, and a zoomed-in view is provided in Figure 13..

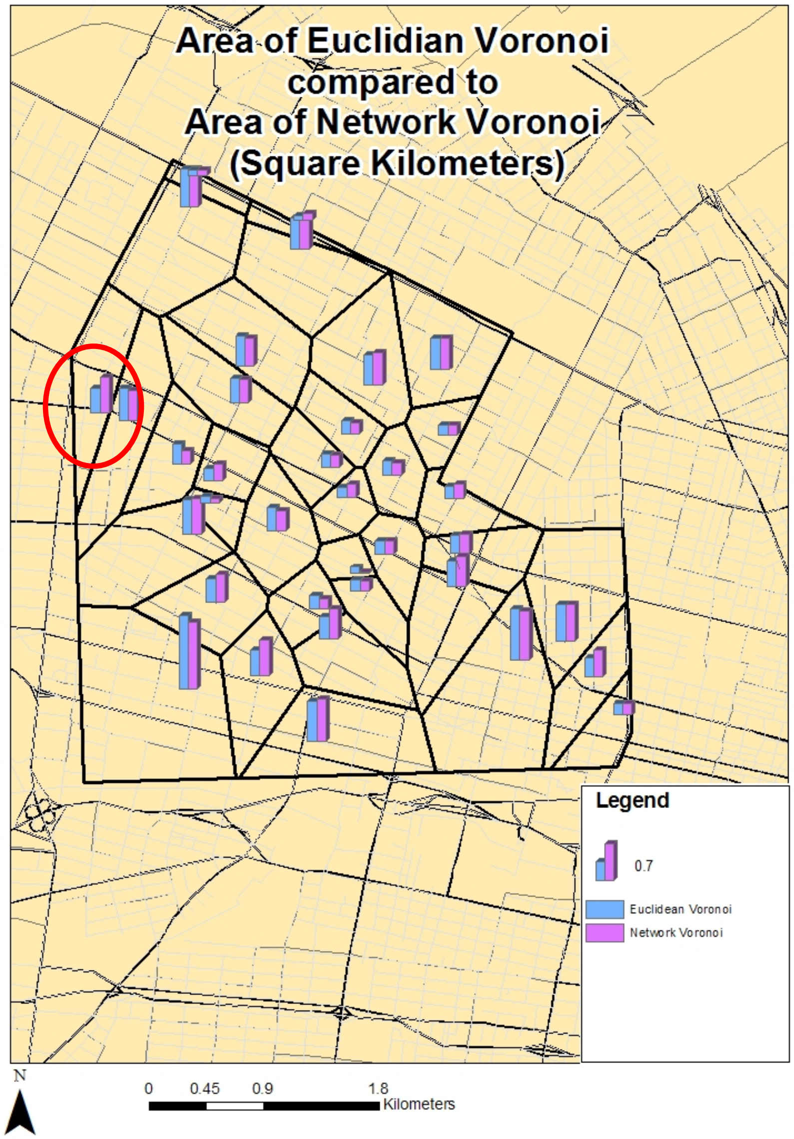

Figure 12. Map of Euclidean Voronoi area compared to network Voronoi area with bar graph symbology centered on homicide locations – 1970

Homicide location 1347 indicates how the encapsulating MCP over-encloses the foot-print of the network Voronoi (Figure 13.).

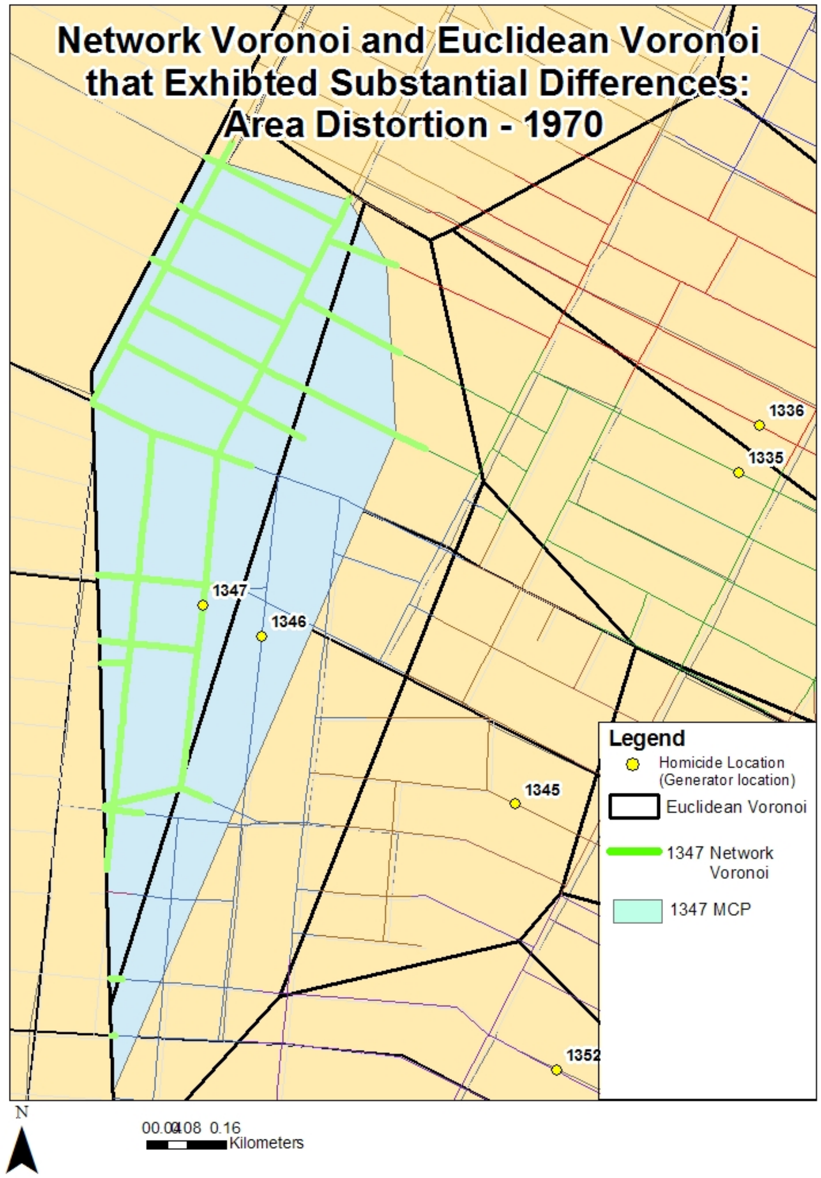

Figure 13. Area measurement distortion, 1970.

The over-enclosure of area by the MCP is likely what caused the network Voronoi area measurement for this homicide to be over-stated in the graph. Of additional interest is the basic incompatibility of the network Voronoi compared to that of the Euclidean Voronoi. In Figure 13 the network Voronoi extends well beyond the foot print of the Euclidean Voronoi, following the street network to the point where a different homicide is closer.

Not only does this contribute to forming an MCP larger than the area actually occupied by the network Voronoi but it reveals a difference in the network Voronoi which was not captured in the statistical measurements performed in this thesis or in the Maki and Okabe (2005) work: The layout of streets produced by the network Voronoi calculation is very different from the layout of streets encapsulated by a Euclidean Voronoi boundary even though their corresponding length and are measurements may be similar. In the following example from the analysis of 1972 data, the full impact of how these topological differences affect the results of each method and the potential ramifications for police work is discussed.

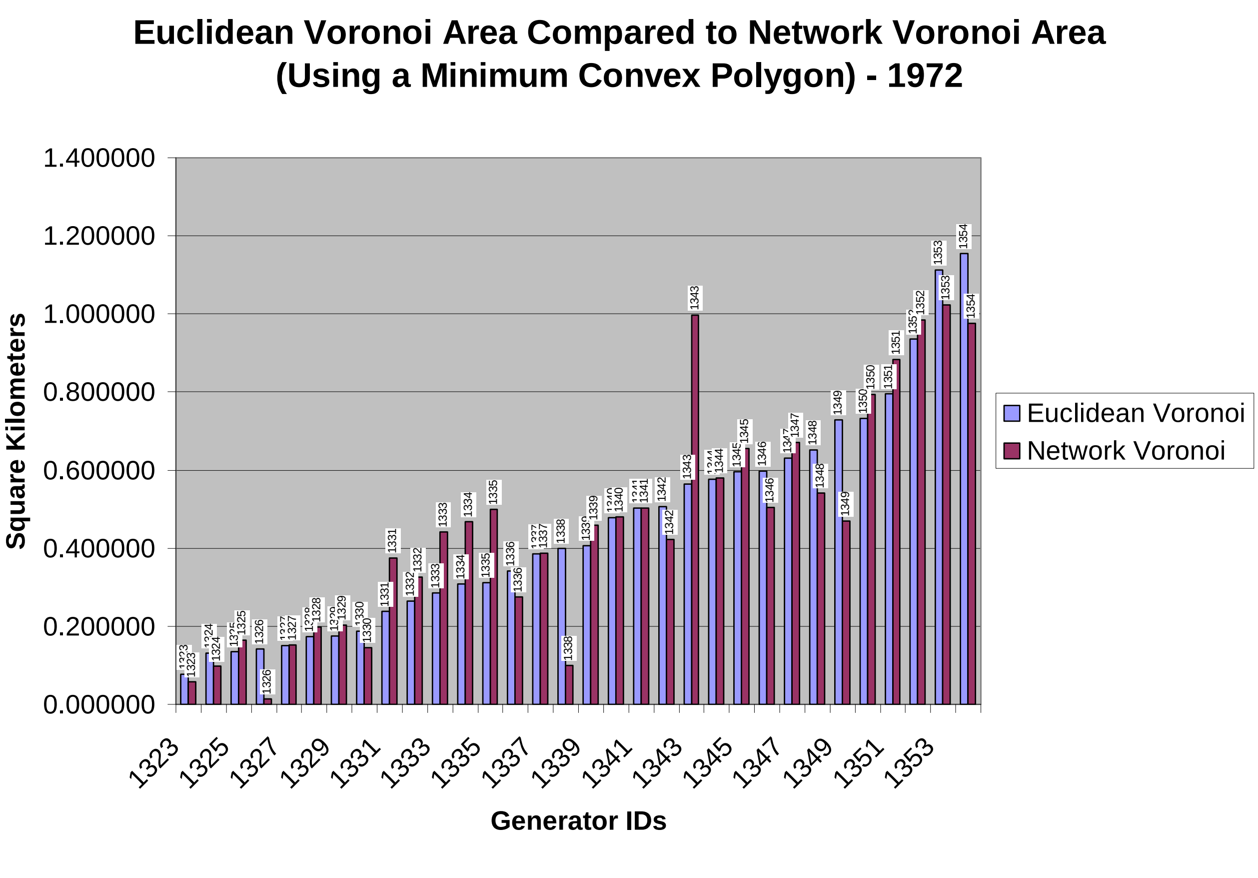

Figure 14. Euclidean Voronoi Street Area compared to Network Voronoi Area using Minimum Convex Polygon – 1972

In Figures 14 and 15 comparing the area differences of the 1972 data, homicide location number 1343 especially stands out. The difference in area measurements is extremely large for this homicide.

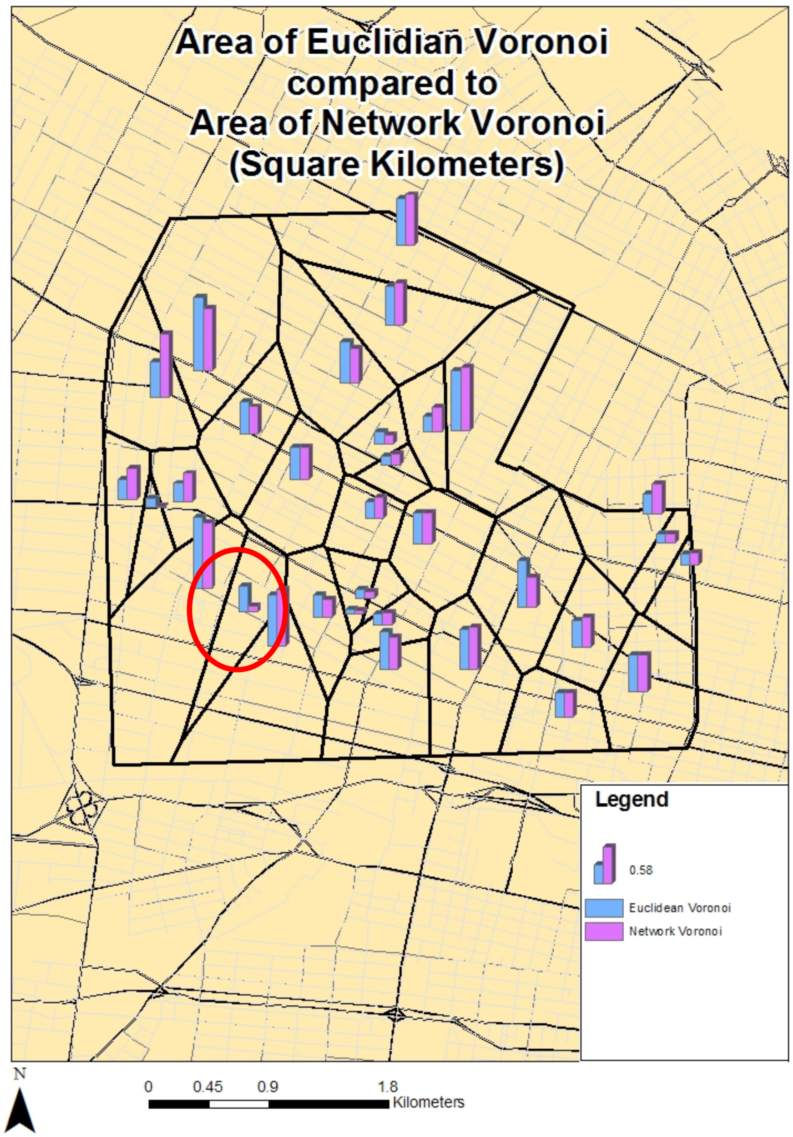

Figure 15. Euclidean Voronoi Street Area compared to Network Voronoi Area using Map with Bar Graph symbology centered on Homicide locations – 1972

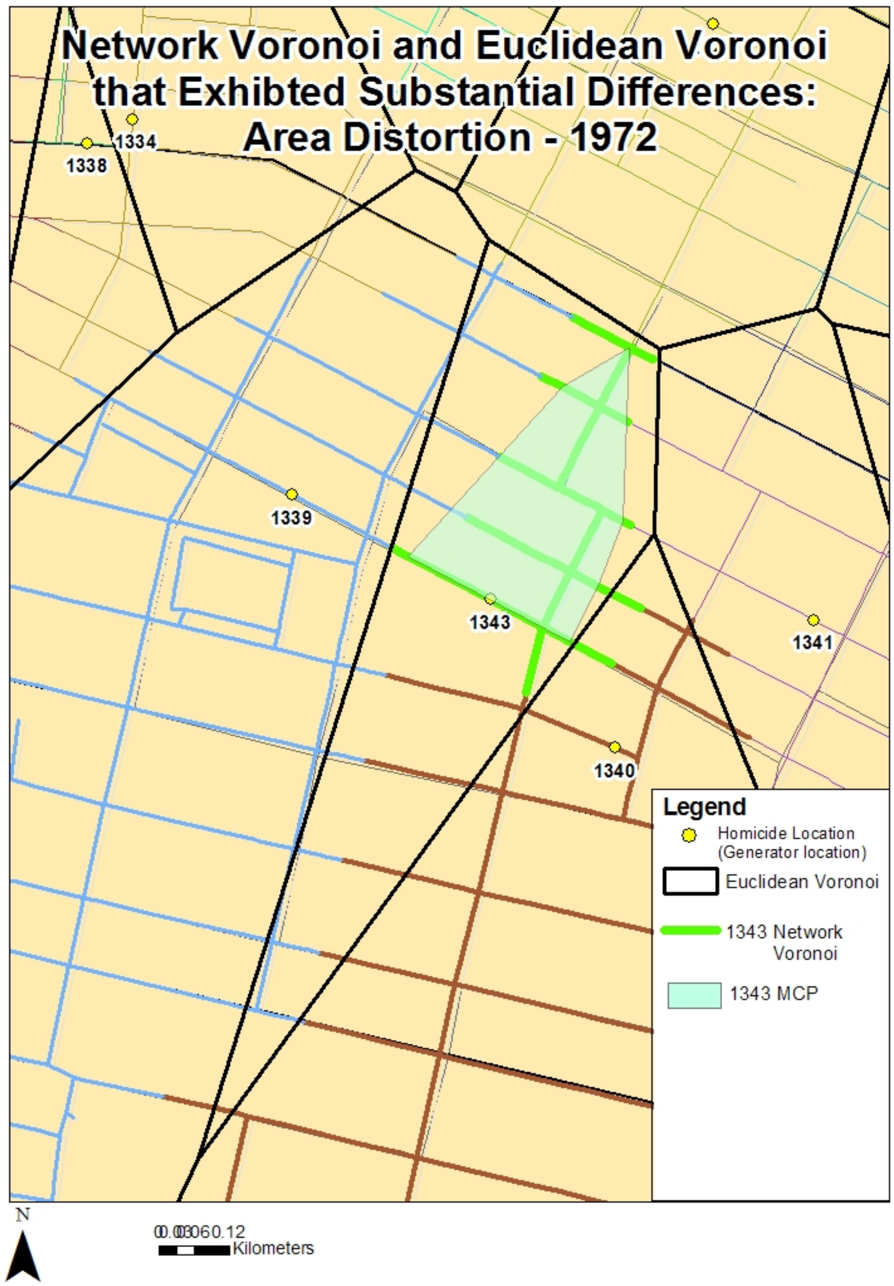

Figure 16. Area distortion and mis-match with Euclidean Voronoi

Figure 16 shows the network Voronoi for homicide location 1343 and the corresponding Euclidean Voronoi. The large discrepancy between the Voronoi regions for location 1343 indicates that even at scales above 500 meters in length the two methods are not as interchangeable as one might conclude from the statistical analysis with Pearson's *r* and RMSE. Note the four homicides indicated on Figure 16 and how the layouts of their network Voronoi diagrams differ from those of their corresponding Euclidean Voronoi diagrams. Although the Euclidean region for location 1343 extends far to the south-west, the corresponding network Voronoi region is truncated. The large block on which the homicide is located creates a barrier to movement in a south-easterly direction along the street network. Large discrepancies are also evident for several of the other homicide locations shown in Figure 16.

The research of Groff and McEwen (2006) based on 3293 offenders who committed homicide in Washington D.C. during 2001, found that the median distance from an offenders' home was 1030 meters for acquaintance-based homicide and 1866 meters for stranger-on-stranger homicide. Utilizing this information, along with the network shortest path measurements like those presented in Figure 16, could potentially enable police investigating the homicides to approximate where perpetrators live. Assuming that the perpetrator didn't deliberately obfuscate their journey to crime, as might occur in the case of an organized offender such as a serial murder (Ressler et al. 1986), using network distance measurements in tracking perpetrators makes sense. Using the Euclidean Voronoi for this same analysis would have yielded entirely different and very possibly incorrect results, leading the police to focus their investigation in the wrong location.

The data for 1982 present similar examples of the issues mentioned thus far which may have contributed to the fit differences between locations (Figure 17).

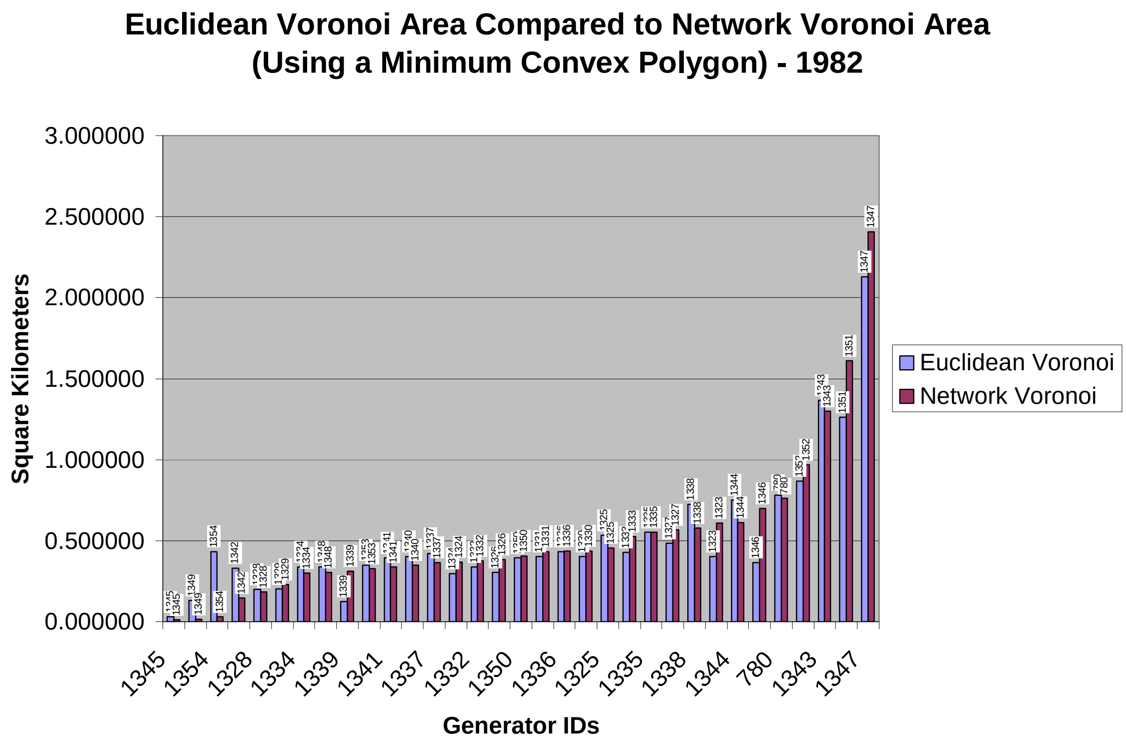

Figure 17. Euclidean Voronoi street srea compared to network Voronoi area using Minimum Convex Polygon – 1982

For example, MCP area measurements for locations 1347 and 1346 (Figure 17) differ greatly due to overlap and topological differences. For location 1346, the network Voronoi region extends far beyond the boundary limit of the Euclidean Voronoi, reflecting the layout of the street network (Figure 18).

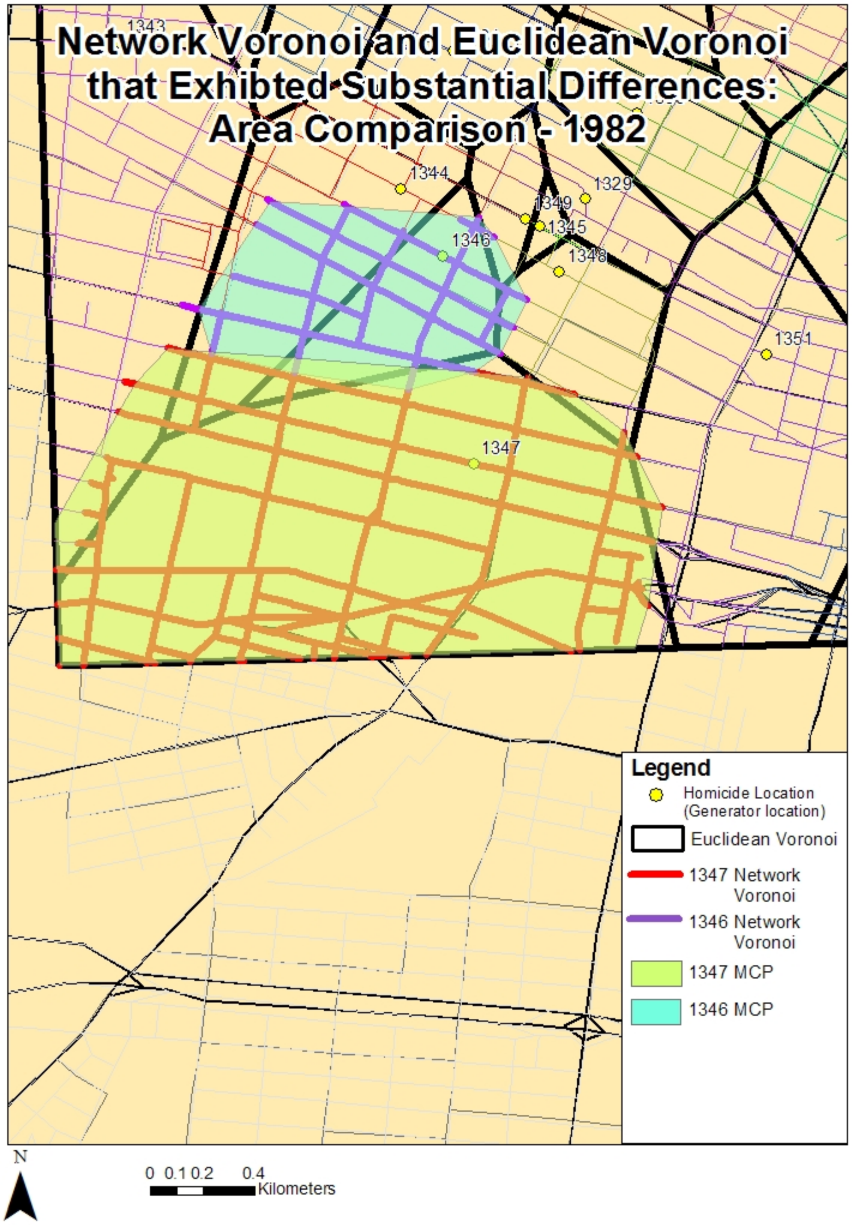

Figure 18. Network Voronoi Homicide location 1346 and 1347 produce MCP distortion

4.2.2 *Differences among fit for location between length comparison of the two methods*

Examination of the graphs and maps comparing length show similar kinds of incongruence as were discussed above with area. Although most Euclidean Voronoi and network Voronoi regions aligned fairly well, there were always a few which did not. Figures 19 through 24 provide more examples of the large disparity in street length measurements for each year.

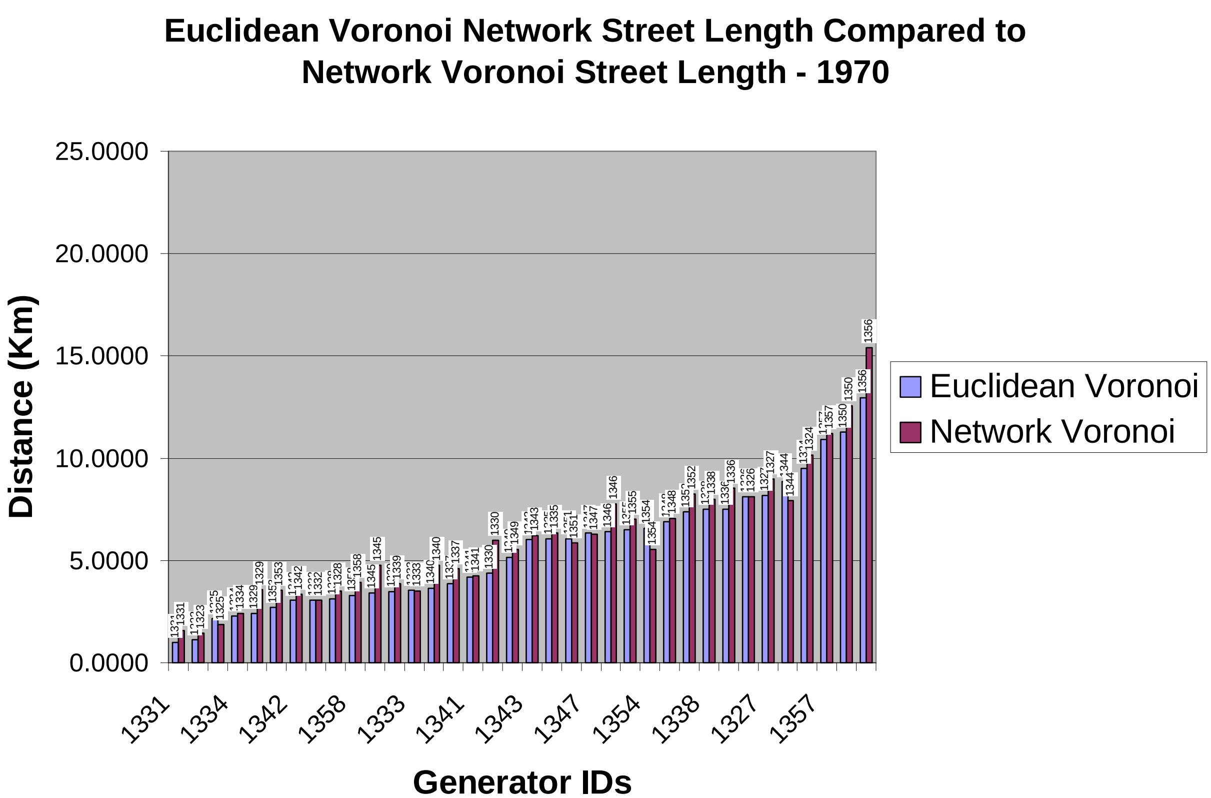

Figure 19. Euclidean Voronoi Street Length compared to Network Voronoi Street Length – 1970

For the 1970 data, location 1356 shows the greatest in street length measurements (Figure 20.). The network Voronoi region for homicide location 1356 does not cover the same streets as enclosed in the Euclidean Voronoi region. Characteristics of the local landscape such as larger blocks or diagonal streets as evident in Figure 20 may play a role in the large difference between the two street length measurements.

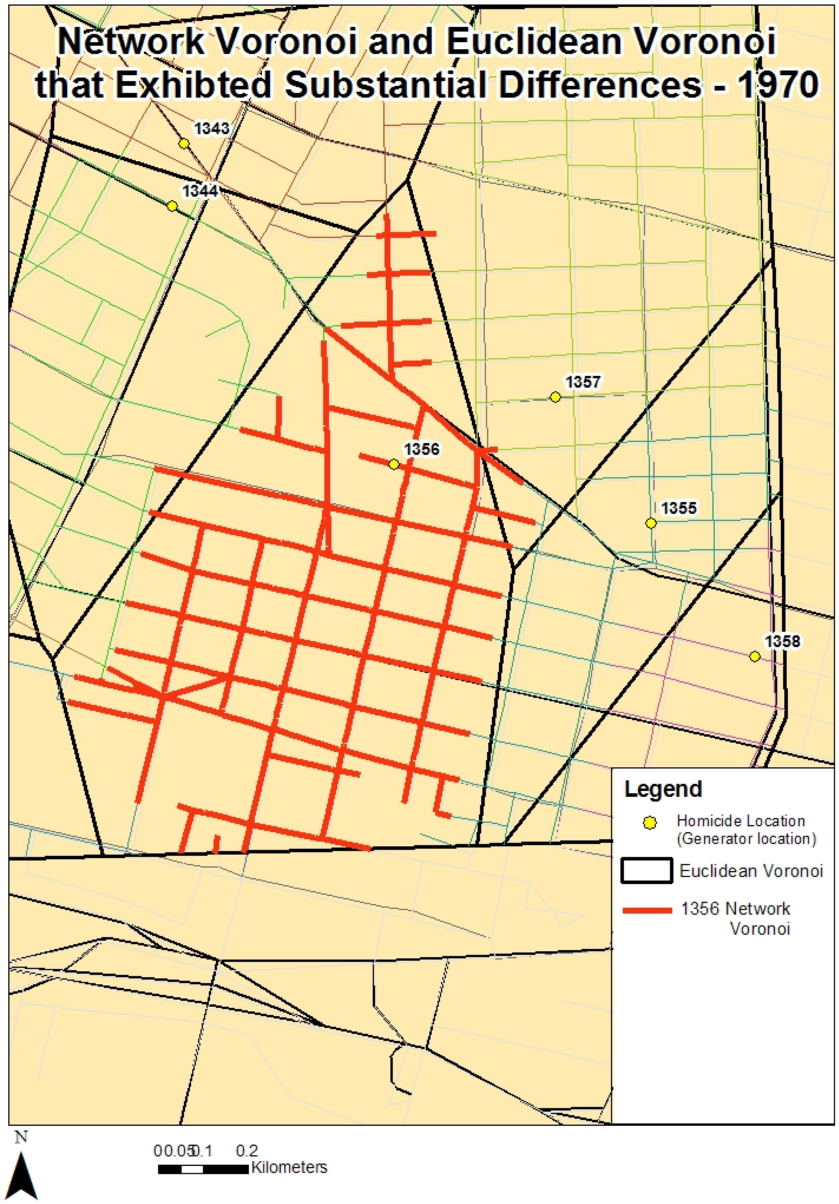

Figure 20. Euclidean Voronoi Street Length compared to Network Voronoi Street Length – 1970

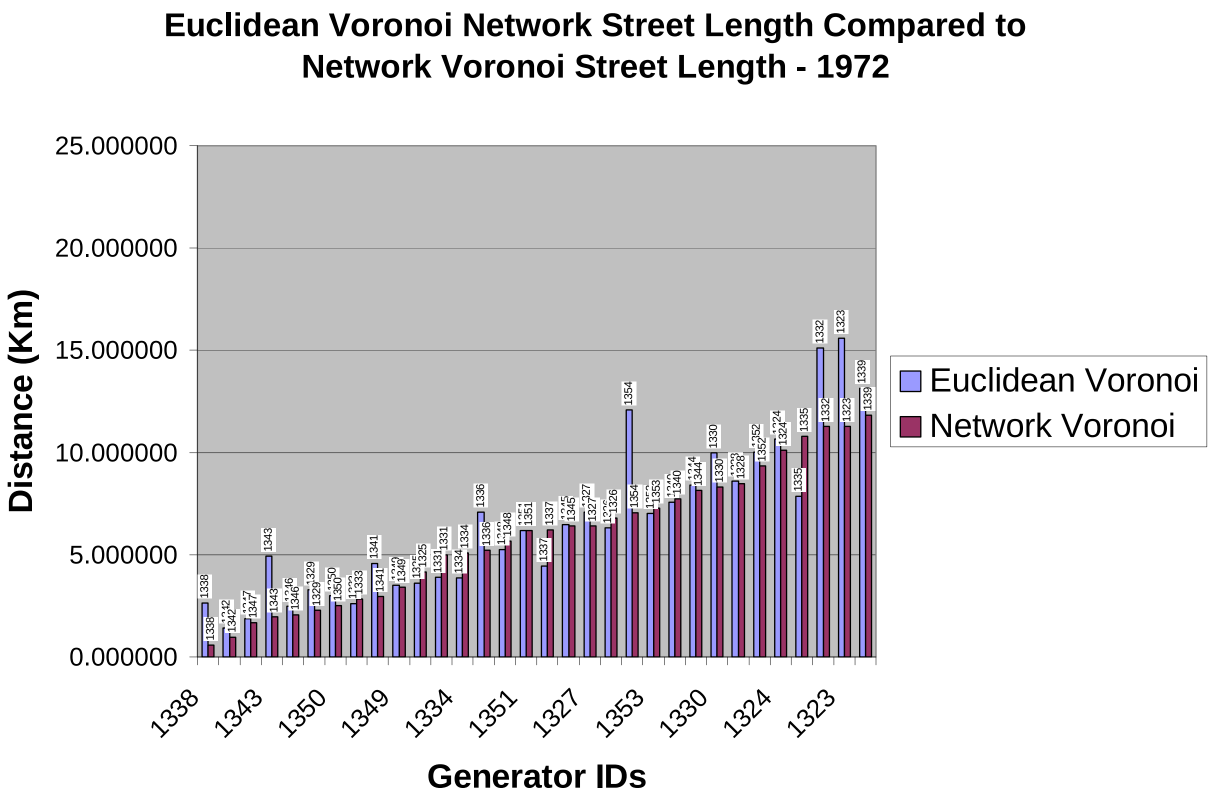

Figure 21. Euclidean Voronoi Street Length compared to Network Voronoi Street Length – 1972

For the 1972 data, six network Voronoi failed to correspond well in terms of length comparison with their Euclidean Voronoi counter-parts. Homicide location 1354 exemplifies the disparity and its relationship to the underlying street grid (Figure 22). This homicide is located on a cul-de-sac with few connecting streets. This results in a truncated and irregularly-shaped network Voronoi region that contains fewer street segments than the corresponding Euclidean Voronoi region.

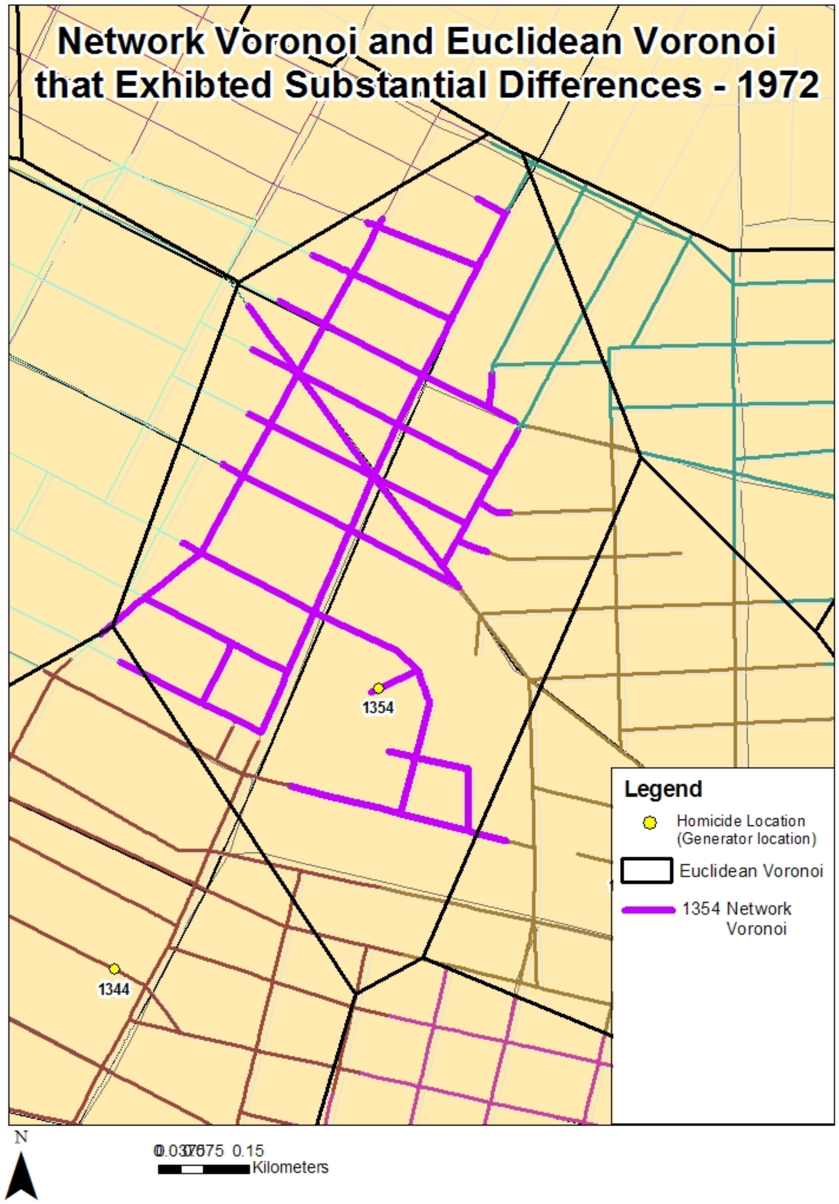

Figure 22. Euclidean Voronoi Street Length compared to Network Voronoi Street Length – 1972

Again the presence of curved streets, diagonal streets and large distances without any streets suggests the possibility that this unusual street geometry may play a role and deserves further investigation in future research.

Euclidean Voronoi Network Street Length Compared to Network Voronoi Street Length - 1982

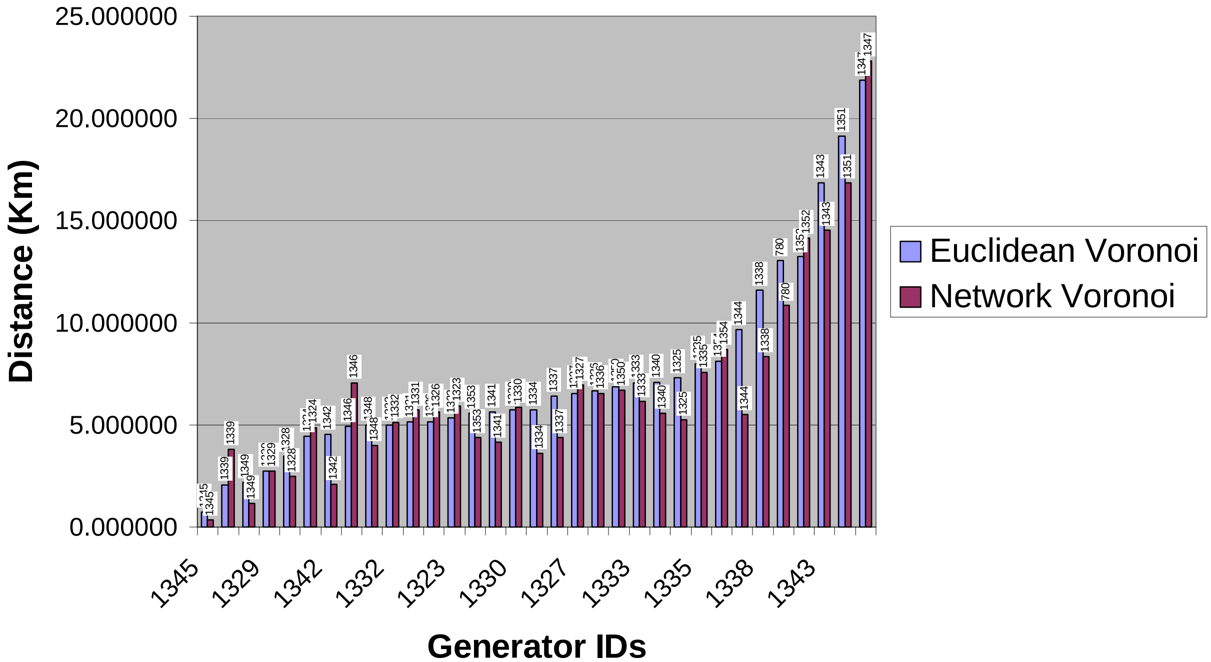

Figure 23. Euclidean Voronoi Street Length compared to Network Voronoi Street Length – 1982

Homicide locations 1344, 1345 and 1346 standout in the 1982 data (Figure 24) confirming how some network Voronoi can be completely different from their Euclidean Voronoi counter-parts. Figure 24 also shows how the presence of a large block creates a barrier to movement which cuts off the network Voronoi region. Situated on a large block, homicide location 1344's network Voronoi region is truncated to the south and west (Figure 24).

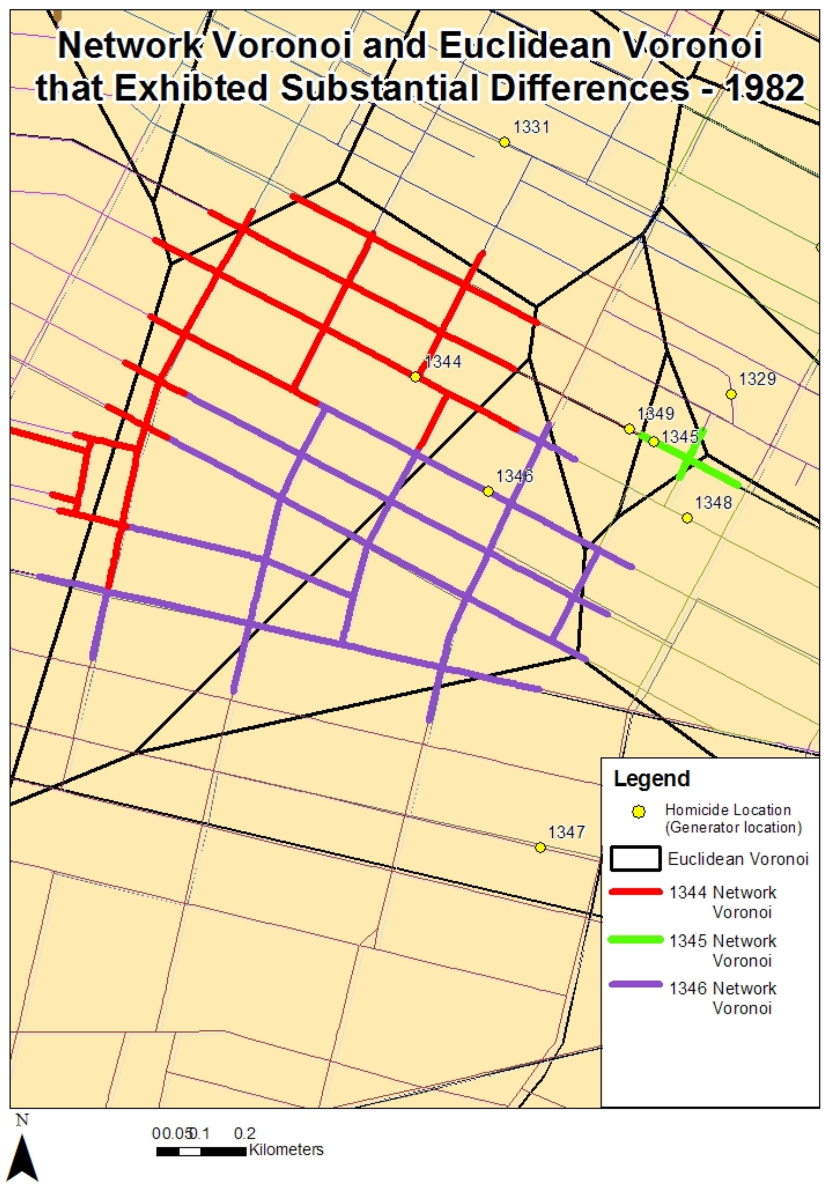

Figure 24. Euclidean Voronoi Street Length compared to Network Voronoi Street Length – 1982

In summary, at a local scale, the two methods yield very different results, identifying Voronoi regions that differ greatly in shape, size and extent. These differences appear to be related to the geometry of the street network and to the locations of homicides within that network. To explore the precise causes of these differences, research would need to be conducted in places with different street geometries, including the grid like structures present in many urban settings as well as more unusual, irregular geometries. This was beyond the scope of this thesis, but it would provide a good basis for future research analyzing the effect of urban street geometry on network Voronoi calculations.

Despite the overall similarity between the two methods at lengths greater than 500 meters, there were sizable local differences in the geographic extent and shape of the regions constructed by the two Voronoi methods. Can a law enforcement agency really afford to utilize techniques which can take them in the wrong direction, costing valuable resources, while also leaving a lethal offender at large? If law enforcement were investigating homicide location 1344 and 1346 in Figure 24, and applied the Euclidean voronoi technique they likely would have focused their search in the wrong place. Had they used the network technique based on the predominant theories of how environment and behavior influences crime choice and location their investigative efforts could have been more effective. The results of this research provide crime analysis professionals with new information useful to the analysis of crime mitigated by street networks and a better understanding of the effect of network distance on voronoi diagram techniques.

Chapter 5 – Conclusion

This research has found a strong similarity between the characteristics of Euclidean and network Voronoi diagrams based on homicide data for 3 years for a St. Louis, MO neighborhood. Pearson's r and RMSE indicate a close correspondence between area and street length measurements in all years. Although the overall similarity was high, the degree of correspondence varied across the study area, revealing substantial local variation in fit. In areas where the street network is poorly connected, with cul-de-sacs, dead-end streets and large, irregularly-shaped blocks, the fit between network and Euclidean measures tends to be poor. These landscapes create barriers to movement along streets, resulting in a larger disparity between straight-line and network distances. Thus, although the network and Euclidean Voronoi results were very similar globally, they were sometimes quite dissimilar at the local scale of particular homicide locations.

These results demonstrate that either method when applied in a dense urban setting using a unit of measure of one kilometer or greater will produce length and area measurements of comparable magnitude in regards to the street networks where the crimes have occurred. Based on the work by Maki and Okabe (2005) we see that scale plays a role and although only one of the locations analyzed produced a result below 500 meters it did corroborate their work. More importantly their work found that at distances above 500 meters the Euclidean and network results tended to be similar – a finding also borne out by this research. Despite the lack of significant difference in the results of the two techniques in terms of statistical analysis, one cannot overlook the unique parameters such as street segment configuration which make one method a better choice over the other in certain contexts. Specifically, when it is likely that a perpetrator could not have traveled through the ensuing space in a straight line (Euclidean). In these instances as demonstrated with homicide locations 1344 and 1346 in Figure 24, the use of the Euclidean voronoi technique would have erroneously lead investigating officers to search a space not likely to have been traveled by the crime's perpetrator.

Although, the Euclidean Voronoi technique does an adequate job of approximating the distances and areas of the network Voronoi technique, it is inherently clumsy when applied in the urban environment. Mollenkopf, Goldsmith, McGuire, and McLafferty (2003) in the New York City, Police Department's Crime Mapping, and Analysis Application (CMAA) had to make considerable changes to the software to account for the possibility of crime occurring in a multi-story building. The Euclidean Voronoi technique as it currently stands was not designed to accommodate the three dimensional space presented by the structure of the urban environment. Similarly, this research shows that the Euclidean method often does a poor job of representing regions around homicides based on travel along street networks. Euclidean Voronoi certainly has its uses, but those uses primarily occur when the problem and data are not highly constrained by the intervening built environment.

Atsuyuki Okabe (2008) in a recent study comparing and contrasting the application of both Euclidean and network Voronoi techniques very strongly believes that the Euclidean technique is entirely inappropriate in urbanized areas citing specifically the substantial differences demonstrated by the Maki and Okabe (2005) work. He adds that the degree of accuracy needed in today's competitive business environments, where knowing precisely the extent of marketing areas is critically important, requires methods which yield improved positional accuracy, as the findings of Maki and Okabe (2005) demonstrate.

This conclusion by Okabe is important in respect to the analysis of crime locations for law enforcement purposes. According to a study by Groff and McEwen (2006), which compiled all research relating to the spatial characteristics of crime and place, the majority of all crime occurs in a very short distance from either a victim's or perpetrator's home location. Specifically, on average, victims were only .4 miles from their homes when a crime was committed, and offender's journey to crime was only .45 miles from their home location. Knowing the proximity of crimes to the locations of perpetrator's and victim's home locations would provide useful information in building a profile of the criminal. Network distance makes sense in measuring distance, because victim's and perpetrator's movements are channeled by these networks. Using the network Voronoi technique to identify a crime's potential context is very appropriate, because travel in urban spaces is largely based on networks. The widespread availability of accurate distance data makes crime network techniques far superior to Euclidean methods for analyzing crime. Network techniques have become an essential tool in geographic criminal profiling.

Distance to crime is not the only useful application of the network Voronoi technique in urban settings. In examining the diffusion of homicides related to gang rivalries, Cohen and Tita (1999) revealed two distinct clusters of youth-gang homicide and youth non-gang homicides which provided evidence of diffusion into new locations non-contiguous to existing home territories. Given the knowledge that gun-inflicted homicide occurs at a greater distance from the perpetrator's home than do other types of crime, use of a network Voronoi technique could have helped in tracking homicide diffusion in terms of shortest-path to crime location.

This research has provided useful results despite the existence of unavoidable introduced error by aspects of the research methodology. As revealed in earlier chapters, the use of the minimum convex polygon (MCP) to provide a measure of areal extent for streets occasionally resulted in flaws that produced a slight over-measurement of area. Had it been possible to manually digitize a boundary around each street network the results would have been more accurate, but still not entirely perfect. Manually digitizing the boundary still would have required that assumptions be made about the exact placement of the boundary. Despite its short-comings, for efficient and consistent processing of boundary placement, the minimum convex polygon approach was the best choice.

The study demonstrates the effect of the intervening landscape and street geometry on the differing street lengths and area measurements identified at some locations. Although the specific relationships between configuration of streets and length/area measurements were not addressed by this research, these relationships should be considered in future studies. This research shows that the unique nature of the intervening landscape plays a role in shortest-path optimization techniques like network Voronoi and this impact cannot be ignored in analysis.

Given that travel is not conducted on a featureless plane, network techniques are more appropriate as they allow for the intervening landscape to be considered. Network techniques should replace Euclidean techniques and become the gold standard for crime analysis in urban landscapes. If everyone was applying network techniques we would be making much faster progress and more accurate predictions in crime analysis. However, network tools have not been made widely available. The reasons for this are complex. We have become a slave to the commercial software development cycle which fails to keep pace with the development and dissemination of the latest analytical techniques being produce in academia. Society, cannot afford to depend on commercial software any longer as its main delivery platform for new research. We need a better software delivery model so techniques and tools such as network voronoi can be widely distributed when first developed, not 10 or 20 years later. If new tools and techniques were made more widely available, the pace of knowledge generation would accelerate across all disciplines of science. The algorithms for network analysis have been available for quite sometime. They have been discussed in the literature for more than 20 years and computers are more than capable of handling the task as demonstrated in this thesis. Unfortunately, software like SANET quickly become obsolete when it is tied to commercial software. So, what do we do now?

The answer should be to invest more heavily in Open Source software initiatives. Having a commonly accepted Open Source platform for software delivery of research will place techniques like network voronoi in the hands of practitioners sooner. This research has shown how important it is to use network methods in analyzing urban crime, but implementing these techniques is still a challenge.

References

Anselin L., Cohen J., Cook D., Gorr W., Tita G. (2000). Spatial Analyses of Crime, Measurement and Analysis of Crime and Justice, Criminal Justice, 4: 213-262

Baldwin J., (1979). Ecological and Areal Studies in Great Britan and the United States, Crime and Justice, 1: 29-66

Beyer H. L. (2002-2006). Hawth's Analysis Tools, version 3.26

Blalock, H.M., Social Statistics, McGraw-Hill Inc., New York 1979

Block R. (2000). Gang Activity and Overall Levels of Crime: A New Mapping Tool for Defining Areas of Gang Activity Using Police Records, Methodological Note, Journal of Quantitative Criminology, 16(3): 369-383

Bowers, K. J., Johnson, S.D., Pease, K., (2004). Prospective Hot-Spotting
British Journal of Criminology, 44, 641-658

Brantingham, P., and Brantingham, P. (1993). Nodes Paths and Edges: Considerations on the Complexity of Crime and the Physical Environment.
Journal of Environmental Psychology, 13, 3-28

Chainey S., Tompson L., Uhlig S.(2008). The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime, Security Journal, 2008, 21, 4 -28

Cohen L.E., Felson M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach, American Sociological Review, 44 (August): 588-608

Cromley E. K., McLafferty S. L. (2002). GIS and Public Health, The Guilford Press, New York

Cohen J., Tita G. (1999). Diffusion in Homicide: Exploring a General Method for Detecting Spatial Diffusion Processes, Journal of Quantitative Criminology, 15(4): 451-493

Curry G.D., Spergel I.A. (1988). Gang Homicide, Delinquency, and Community, Criminology, 26(3): 381-405

de Smith M. J. , Goodchild M., and Longley P. (2008). Geospatial Analysis: A Comprehensive Guide to Principals, Techniques and Software Tools. Leicester: Winchelsea Press.

Dunn C.S. (1980). Crime Area Research, Chapter 1, Crime: A Spatial Perspective edited by Georges-Abeyie D.E. and Harries K. D., Columbia University Press, New York

Eck J. E., Chainey S., Cameron J. G., Leitner M., Wilson R. E. (2005). Mapping Crime: Understanding Hotspots, National Institute of Justice – Special Report, August

Environmental Systems Research Institute, ESRI, (2006). 380 New York Street, Redlands, CA, User's Manual ArcView GIS 9.2

Gatrell A. C., T. C. Bailey, P. J. Diggle and B. S. Rowlingson. (1996). Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology, Transactions of the Institute of British Geographers, New Series, 21(1): 256-274

Groff, E. and McEwen, T. (2006). Exploring the Spatial Configuration of Places Related to Homicide Events – Final Report. Alexandria: Institute of Law and Justice

Hannon L. (2002).Criminal Opportunity Theory and The Relationship Between Poverty and Property Crime, Sociological Spectrum, 22: 363-381

Harries K. D. (1990). Serious Violence: Patterns of Homicide and Assault in America, Charles C Thomas, Springfield Illinois

Jones, T. (1976). The Press as Metropolitan Monitor, The Public Opinion Quarterly, 40(2): 239-244

Kohfeld, C. W., and Sprague, J. (1999). Arrests as Communications to Criminals in St. Louis, 1970, 1972-1982. [Computer file]. ICPSR09998-v1. Los Altos, CA: Sociometrics Corporation [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], doi:10.3886/ICPSR09998

Kohfeld, C.W., and Sprague, J. (1988). Urban Unemployment Drives Urban Crime. Urban Affairs Quarterly, 24(2), 215-241

Krohn M. D. (1978). A Durkheimian Analysis of International Crime Rates, Social Forces, 57(2): 654-670

La Vigne N. (1997). "How Hot Is That Spot?: The Utility and Application of Place-based Theories of Crime." Paper presented at the International Situational Crime Prevention Conference, June, Oslo, Norway

Lundrigan S, and Canter D. (2001). Research Report: Spatial Patterns of Serial Murder: An Analysis of Disposal Site Location Choice, Behavioral Sciences and the Law, 19: 595-610

Maki, N. and Okabe, A. (2005). A Spatio-Temporal Analysis of Aged Members of a Fitness Club in a Suburb. Proceedings of the Geographical Information Systems Association, 14: 29–34

Maheu-Giroux M., de Blois S. (2006). Landscape ecology of Phragmites australis Invasion in Networks of Linear Wetlands, Landscape Ecology, DOI 10.1007/s10980-006-9024-z: 1-17

Mollenkopf J., Goldsmith V., McGuire P., McLafferty S., (2003). Identification, Development and Implementation of Innovative Crime Mapping Techniques and Spatial Analysis: 1-32

Okabe, A., Boots, B. and Sugihara K. (1992). Spatial Tessellations Concepts and Applications of Voronoi Diagrams, New York: John Wiley and Sons

Okabe A. , Okunuki K.I. (2001). A Computational Method for Estimating the Demand of Retails Stores on a Street Network and its Implementation in GIS, Transactions in GIS, 5(3): 209-220

Okabe, A., Satoh, T., Furuta, T., Suzuki, A. and Okano, K. (2008). 'Generalized Network Voronoi Diagrams: Concepts, Computational Methods, and Applications', International Journal of Geographical Information Science: 1-30

Radzinowicz L. (1965). Ideology and Crime: The Deterministic Position, Columbia Law Review, 65:1047-1060

Rengert G, Wasilchick J (1985). Suburban Burglary: A time and a Place for Everything. Charles Thomas, Springfield

Rengert G,Wasilchick J (1989). Space, Time and Crime: Ethnographic Insights into Residential Burglary - Final Report to The National Institute of Justice, Washington, DC

Robinson W. S. (1950). Ecological Correlations and the Behavior of Individuals, American Sociological Review, 15: 351-357

Roncek, D.W., and Maier, P.A. (1991). Bars, Blocks and Crimes Revisited: Linking the Theory of Routine Activity to the Empiricism of “Hot Spots”, Criminology, 29 (4): 725-753

Ressler R. K., Burgess A. W, Douglas J. E, Hartman C. R, and D”Agostino R. B. (1986). Serial killers and their Victims: Identifiying Patterns Through Crime Scene Analysis, Journal of Interpersonal Violence, 1(3): 288-308

Santilla P, Laukkanen M, Zappala A, and Bosco D. (2008). Distance Travelled and Offence Characteristics in Homicide, Rape, and Robbery Against Business, Legal and Criminological Psychology, 13: 345-356

Sherman L.W., Gartin P. R., and Buerger M. E. (1989). Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place, Criminology, 27(1): 27-55

Snook B, Cullen R.M., Markros A. and Harbort S. (2005). Serial Murderers Spatial Decisions: Factors that Influence Crime Location Choice, Journal of Investigative Psychology and Offender Profiling, 2:147-164

Weisburd, D. and Green, L. (1995). Policing Drug Hot Spots: The Jersey City Drug Market Analysis Experiment, Justice Quarterly, 12(4): 711-735

Weisburd, D., Bushway, S., Lum, C. and Yang, S.M. (2004). Trajectories of Crime at Places: A Longitudnal Study of Street Segments in the City of Seattle, Criminology, 42(2): 283-321

Yamada, I., and Thill, J.C. (2007). Local Indicators of Networked-Constrained Clusters in Spatial Point Patterns. Geographical Analysis, 39: 268-292