For office use only
T1 ________________
T2 ________________
T3 ________________
T4 ________________

Team Control Number

6036

Problem Chosen

B

For office use only
F1 ________________
F2 ________________
F3 ________________
F4 ________________

2010 Mathematical Contest in Modeling (MCM) Summary Sheet
(Attach a copy of this page to each copy of your solution paper.)

Type a summary of your results on this page. Do not include the name of your school, advisor, or team members on this page.

Summary

Here, we offer two new approaches for the local police agency to narrow down the geographical range of serial-crime suspects. These two schemes are based on Criminal Spotlight Model (CSM) and Street Control Model (SCM).

CSM targets at criminals' base (usually their home.) As its name suggests, it marks the neighborhood of criminals' potential base on the map. In contrast with finding the `center of mass' of the locations of attacks, CSM gives a scope of where the criminal would probably live rather than a point. On the other hand, compared with the Rossmo's Criminal Geographical Targeting [22], our method has a more stable theoretic foundation based on the mathematical statistic tools and probability theory.

We classify serial crimes into two categories: serial murder and serial robbery. According to their different patterns, we add two optional techniques ("temporal" factors included) for two kinds crime respectively.

After validating the reliability of CSM through 12 real world serial murder cases (8 of which in USA), we developed a stochastic simulation, which takes advantage of Monte Carlo algorithm framework, to check its competence. The program deals with various levels of police force, and concludes the time length of a successful criminals capture in most of the serial murder cases. This expected time length also assists officers to judge the suitability of our model.

Our second scheme draws attention on another aspect of the problem. This model requires police to abstract an undirected connected planar graph out of a city map and collect the data related to serial crimes, notably the crime rate, `hotspots' distribution, etc. Thereafter, SCM will calculate the probability that the criminal might pass through a given street using a revised Dijkstra's single-source shortest path algorithm. This probability distribution information will shed light on the installation of potential road monitors and other crime obstruction facilities.

Combining these two strategies properly, the police would accumulate more information about a serial crime. Firstly, CSM would help the police to eliminate some possible bases of the criminal. Based on the information given by CSM, SCM will regenerate the results, which would help the local police to narrow down the geographical range of suspects and expedite the collection of evidences after a successful capture.

After all, CSM and SCM have their limitations in different situations. They are evaluated after each part of this paper and complied in the executive summary.

Criminals Under The Spotlight

Team #6036

February 23, 2010

Contents

"We all use math every day; to predict weather, to tell time, to handle money. Math is more than formulas or equations; it's logic, it's rationality, it's using your mind to solve the biggest mysteries we know."

—*Numb3rs*

Part I
Introduction

1 What Is Serial Crime?

Serial criminals refer to the criminals who commit serial or repetitive crimes. Episodic or sequential murders, rapes, arson and robberies are all included in serial crimes.[19]

For the year 1986-90, the average number of murders reported in USA annually, according to the Uniform Crime Reports, was over 20,000 and of these 20,000 about 5,000 are classified as unsolved. It is suspected that many of stranger-related murders are committed by serial killers. Besides, a distinct increase in serial killings has been reported since the 1970s.[7] Compared with other types of crime, serial crime may be rare. However, the great panic of the people who frequent the crime area and the high pressure on law enforcement officers following a series of crime lead the crime event to hold the immediate public spotlight. With so many serial crimes unsettled and the huge public influence, many techniques have been applied in the investigation of serial criminals.

The study of serial crimes has traditionally been the preserve of disciplines such as psychology and sociology and it was not until the late 1970s that the spatial or geographical dimension was introduced to the investigation of serial crimes.[3] In fact, the police have long recognized the intrinsic geographical component of crime by sticking pins to maps displayed on walls, with each pin representing a crime event, but it was from the 1970s that scientific rather than empirical methodology was fully explored.

2 Geographical Profiling

In the first episode of the American television serial *Numb3rs*, the mathematician Charlie assists his brother Don, an FBI agent, on a serial rapist case by calculating a 'hot zone', an area where the rapist is most likely to live. *Geographical Profiling*, the approach Charlie employs, analyzes the locations of a connected series of crimes to predict or provide guidance about the probable area of the offender's residence.

Geographic profilers have developed a range of strategies to narrow the search for the serial criminal's residence. These are described succinctly by Levine and Associates [13], who make the broad distinction between *spatial distribution strategies* and *probability distance strategies*. Spatial distribution strategies predict the offender's residence by calculating the *central point* of the acquired crime sites, which varies among *the center of the circle, centroid, median, geometric mean, harmonic mean, and center of minimum distance* in different strategies. The '*Center of Mass*' strategy was used as early as in the Yorkshire Ripper Series. (Another strategy used in this case concerns about the reasoning that the killer has the tendency to attack later the closer he is to home) [23]. Probability distance strategies assign each point in activity space a small positive real number to depict the likelihood of an offender living at a particular location.[21] Specialized software systems, such as *Rigel* [17] and *CrimeStat* [9], have been developed based on these strategies.

Inspired by the prominent literature book in this area, Rossmo's *Geographical Profiling* [22], we develop a scheme that improves the probability distance strategies. Apart from the currently existing strategies, we also put forward a scheme from a totally different point of view.

Part II
Literature Review

3 Review over Rossmo's Approach

Rossmo's model introduced the concept of *Buffer Zone*,which is centered around the offender's residence. Brantingham and Brantingham [2] postulate that victims are "probably spatially biased toward the offender's home base" because offenders generally commit crimes in areas they are familiar with and concentrate targets within their immediate environment. According to Rossmo, however, within the *Buffer Zone*, the perceived level of risk increases as the crime scene approach to the criminal's home and therefore the target is less desirable; this area represents a balance between the maximization of opportunity and the minimization of risk.

Therefore Rossmo come up with a model named Crime Graphical Targeting (CGT) [22] by introducing his famous formula to describe the probability of a place being the base of the criminal.

If the T crime sites CS_1, CS_2, $\ldots$, CS_T with coordinations

$$(x_1, y_1), (x_2, y_2), \ldots, (x_T, y_T)$$

are given, the probability of a place with the coordination (x, y) being the base of the criminal is

$$P(x,y) = k\sum_{i=1}^{T}\left[\frac{\phi(i)}{ND((x,y),(x_i,y_i))^f} + \frac{(1-\phi(i))\cdot B^{g-f}}{(2B - ND((x,y),(x_i,y_i)))^g}\right]$$

where B stands for the radius of the *Buffer Zone* and

$$\phi(i) = H[ND((x,y),(x_i,y_i)) - B]$$

in which H is the Heaviside step function defined as follows

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The terms k, f and g are both empirically determined exponents. For criminal cases, the empirical values of f and g are chosen as 1.2 and the value of k as 1.0.[5] The value of B varies with time and space, even depending on individual criminals.

It should be mentioned that the author chose Manhattan distance, that is, $ND((x_1,y_1),(x_2,y_2)) = |x_1 - x_2| + |y_1 - y_2|$, instead of Euclidean metric because real life distance is always larger than Euclidean distance in the influence of different traffic flows and geography.

4 Flaws in Rossmo's Model

For the author didn't raise enough reasoning for the formula, his conclusion is unconvincing. Here are several points where we find flaws and try to improve.

- CGT method did not take the time factor into account, say, it did not utilize time sequential information of the serial crime. However, as the behavior of serial criminals might change over time, the omission of temporal sequential data may result in the inaccuracy of this model.

- The individual term in the formula is arbitrarily selected, which lacks statistical background.

- Rossmo's formula does reflect the postulate that the possibility: follows some form of distance-decay if the point lies outside the *Buffer Zone*; becomes larger the longer the distance if the point lies inside the *Buffer Zone*. However, there is no solid foundation to assume that the distance is in inverse proportion.

- The formula simply adds rather than products the individual probability together which, on some level, represents overall possibility, but betrays the common knowledge of the probability density function.

The key solution to modify CGT method lays in a detailed analysis of the murder's activity. Moreover, employing the timeline of the serial crime will be beneficial. Inspired by Rossmo's CGT model, we put forward our model.

Part III
Criminal Spotlight Model

5 Overview

5.1 Notation

JTC The journey from the criminal's residence to crime.

T The total number of crime sites as in Rossmo's model.

CS Crime sites for serial crimes including murder, rape, robbery etc.

BD The site where the victim's body is dumped or disposed in murder cases.

BASE The home, or in some cases, the working place of the criminal is the crime base of serial criminals and the crime sites are always centered around the base.

ND($\mathbf{P}_1$, $\mathbf{P}_2$) The network distance between Place 1 and Place 2.

5.2 Hypothesis

1. We assume that all the data obtained by the police are reliable.
2. There exists a BASE, from which the crime set out to commit offenses. The BASE that provides the anchor for the criminal activity may take many forms, including home, workplace or frequently visited sites.
3. Each time the crime commit an offense, he or she would randomly select a direction and the distance from BASE to decide the crime site. The direction uniformly distributes among $[0, 2\pi]$, while the distance follows a fixed distribution. These random variables (the selected directions and distances for each time) are independent to each other.

5.3 Theoretical Analysis

We propose our first model, named the Criminal Spotlight Model (CSM), to analyze previous crime sites and then highlight the probable area of the criminal's residence with the aid of statistical tools.

In the statistical view, JTC is a random variable. And based on empirical data and Rossmo's concept of *Buffer Zone*, the distribution of JTC (to simplify computation, we measure JTC in Euclidean metric without loss of generality) is in the following shape, shown in Figure [1]

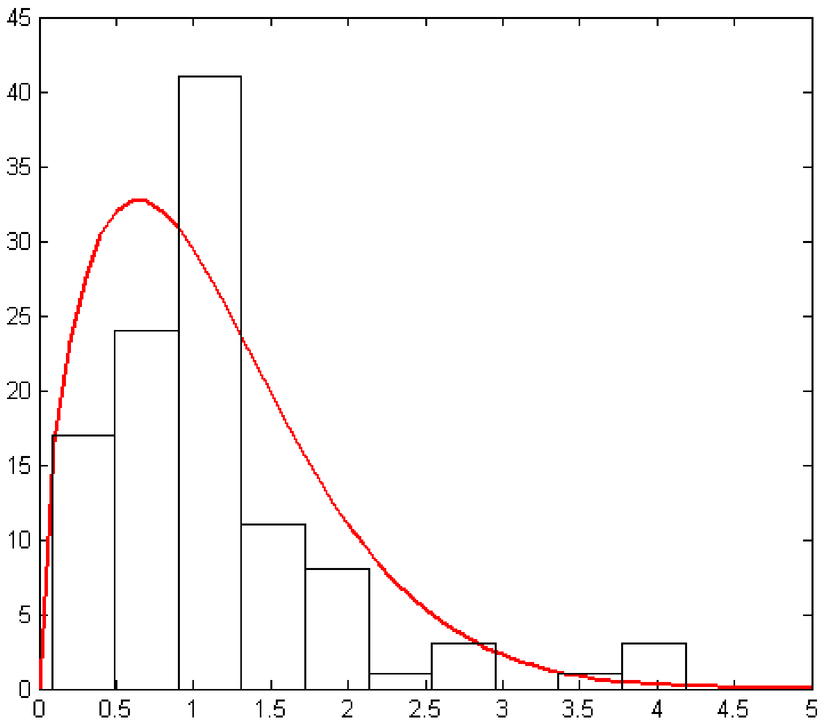

Figure 1: Histogram of data and the MLE of Weibull distribution with parameters $\hat{a} = 1.2518$ and $\hat{b} = 1.5486$.

Among the distributions shaped like this, the Weibull distribution fits our previous data of JTC best, so we can describe JTC as a Weibull random variable. The probability density function (usually abbreviated as pdf) of a Weibull random variable X is

$$WF(x; a, b) = \begin{cases} \frac{b}{a}(\frac{x}{a})^{b-1}e^{-(\frac{x}{a})^b} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $a > 0$ is the scale parameter and $b > 0$ is shape parameter. We can make an estimation of these two parameters based on the JTC data using *Maximum Likelihood Estimation* (MLE).

With the estimated Weibull distribution, we can write out the likelihood function of the *BASE*. Given the *BASE* location, the CSs are independent and identically distributed variables with the joint density function

$$f(CS_1, CS_2, \ldots, CS_n|BASE) = \prod_{i=1}^{n} f(CS_i|BASE)$$

Now we want to look at this function at a different angle: let the observations $CS_1, CS_2, \ldots, CS_n$ be fixed parameters of this function, whereas the *BASE* is allowed to vary freely.

$$\mathcal{L}(BASE|CS_1, CS_2, \ldots, CS_n) = \prod_{i=1}^{n} f(CS_i|BASE)$$

The method of maximum likelihood estimates the *BASE* by finding the location BASE that maximizes $\mathcal{L}(BASE|CS)$

$$\widehat{BASE}_{\text{MLE}} = \underset{BASE}{\arg\max} \mathcal{L}(BASE|CS_1, CS_2, \ldots, CS_n)$$

In the actual operation, we usually plot the likelihood function to aid the investigation. Areas with higher likelihood function value tend to be of higher need for search.

6 Case Study

We collect 11 serial murder cases including Richard Chase, Albert DeSalvo, Peter Sutcliffe, etc, with the victim numbers, BD positions, BASE positions. [22]

For a certain case, network distances (ND) between the BASE and each BD are calculated first. Standardization is necessary in order to make the data comparable. Several ways of standardization are available, including divide each ND by their mean, median or maximum. Taking it into consideration that median is more robust than mean and maximum in dealing with abnormal data, we standardize the data by dividing by the median. Secondly, we pool all the data of the 11 cases together, and then use MLE to estimate the parameters of Weibull function, with the result $\hat{a} = 1.2518$ and $\hat{b} = 1.5486$.

Figure [1] shows the data histogram as transparent bars and the estimated Weibull distribution in red curve.

With the estimated Weibull distribution, we can depict the probability density function of BASE over the plane. Let $r = ND(CS_1, BASE) \geq 0$, the distribution of BASE given CS_1 is

$$\frac{WF(r; a, b)}{2\pi r} = \frac{b}{a}\left(\frac{r}{a}\right)^{b-2} e^{-(r/a)^b}$$

because the value of probability density function on r is equally divided within the circle with the radius r.

Finally, we plot the likelihood function of BASE, which is the product of all the WFs given the CSs. From the images, we can find out areas with higher likelihood function value, where law enforcement officers should pay more attention.

7 Further Discussion

7.1 Prior Distribution

Let us go back to the likelihood function of BASE. From the viewpoint of classic mathematical statistics, what we should do is just to maximize the likelihood function to pick out the MLE of BASE. However, from the viewpoint of Bayesian statistics, there should be a prior distribution of BASE and a posterior one updated from the prior by acquired data. $\xi(BASE)$, denoting the prior distribution of BASE, is roughly the crime rate distribution over a city or state, such as Figure [3] and Figure [4].

According to the theorem in Bayesian statistics, $\xi(BASE|CS_1, \ldots, CS_n)$, denoting to the posterior distribution of BASE, can be calculated as:

$$\xi(BASE|CS_1, CS_2, \ldots, CS_n) = \frac{\prod_{i=1}^{n} f(CS_i, BASE)\xi(BASE)}{\int \prod_{i=1}^{n} f(CS_i, BASE)\xi(BASE)dBASE}$$

where

$$f(CS_i, BASE) = WF(ND(CS_i, BASE); a, b)$$

is the joint density function of CS_i and BASE. It is obvious that the denominator of the formula is constant to BASE. Therefore we can revise our likelihood function as

$$\mathcal{L}(BASE|CS_1, CS_2, \ldots, CS_n) = \prod_{i=1}^{n} f(CS_i, BASE)\xi(BASE)$$

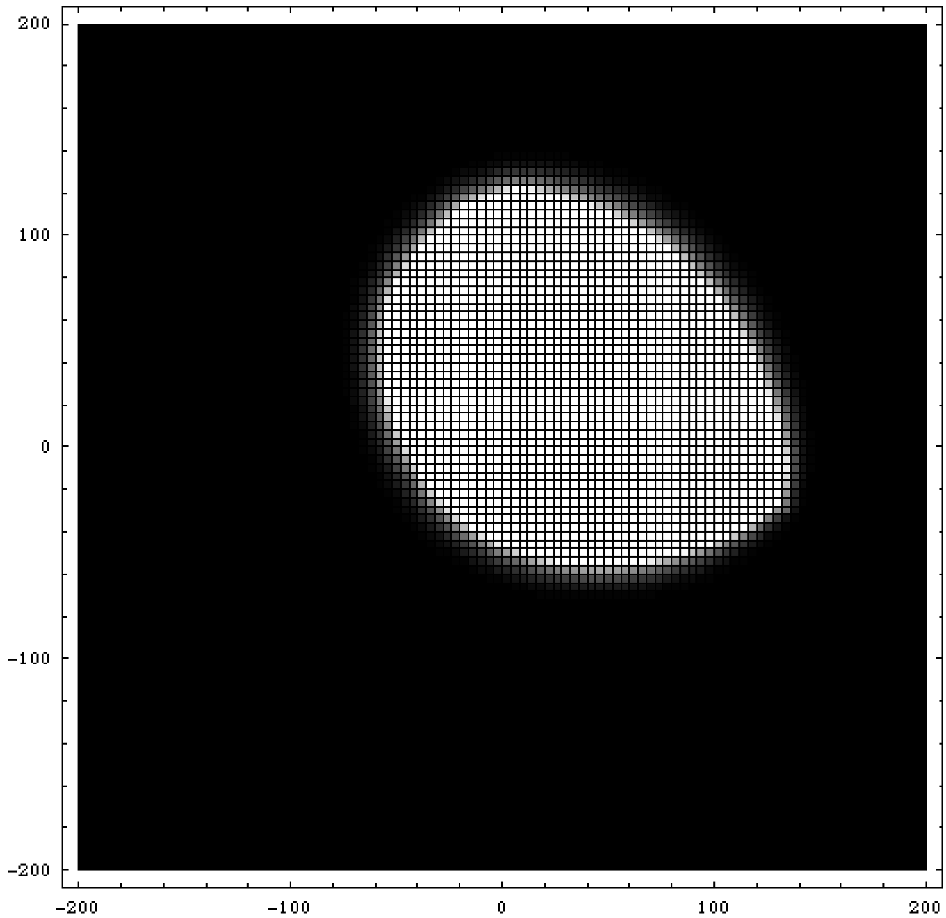

Figure 2: This figure shows the Spotlight in the case of Peter Sutchiffe from England, UK. Areas with higher brightness corresponding to higher likelihood of BASE. The origin represents the ground truth position of BASE. The coordination of CSs are (-254,373.5), (-245,353.5), (-51,-63.5), (-38,68.5), (-24,53.5), (-15,109.5), (13,-10.5), (13,-1.5), (16,6.5), (8,119.5), (31,11.5), (59,-15.5), (100,-17.5), (106,-20.5), (115,-18.5), (115,-14.5) , (125,-20.5), (125,-9.5), (136,-29.5) Given the above data, the likelihood function is calculated according to the Criminal Spotlight Model and the spotlight areas are highlighted.

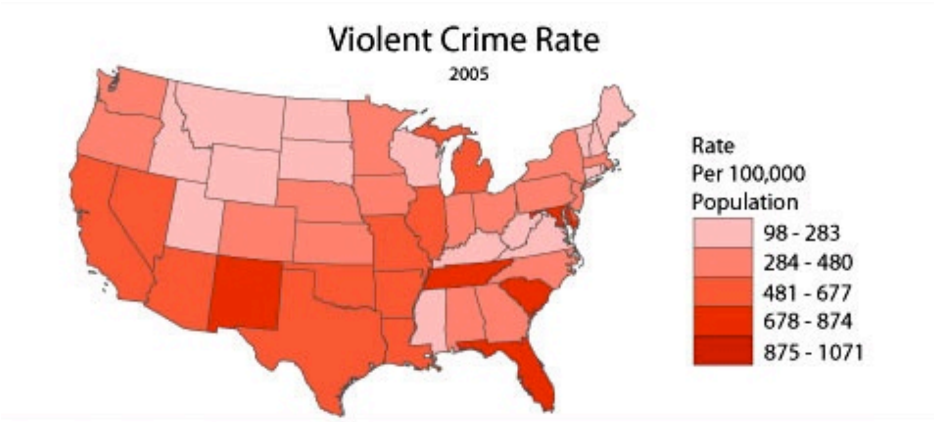

Figure 3: Violent Crime Rate of US in2005. Source: `http://geoggeol.mansfield.edu/media/images/crime-rate-map.jpg` [2010]

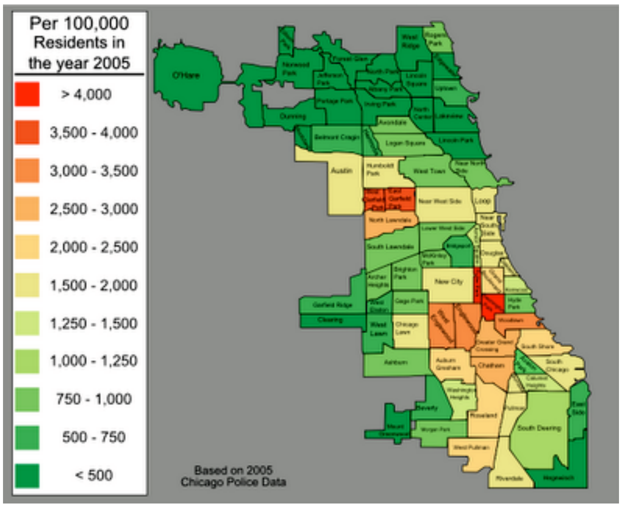

Figure 4: Crime Rate of Chicago in 2005. Source: `http://iguide.travel/Chicago/Safety` [2010]

In the actual operation, police could make use of the crime rate distribution to conduct a better estimation of the likelihood function of the BASE.

7.2 Temporal Influence

7.2.1 Serial Murders

In a recent paper [7] concentrating on the spatial behavior of US serial murder, a remarkably consistent closeness to home in the distances the criminals traveled to the body disposal site is revealed. The study is based on the crime data of 54 male US serial killers. Maurice Godwin and David Canter examined the mean distances traveled to BD of the 54 killers for each time they commit a crime. For example, the mean distance traveled to abduct the first victim is 0.5 miles while the mean distance of serial killers traveled to abduct the tenth victim decreases dramatically to 0.2 miles. Detailed results are showed in the following table.

Offense	1	2	3	4	5	6	7	8	9	10
JTC	23.4	16.9	20.7	14.5	14.4	12.7	14.6	10.6	9.23	5.28

Table 1: The mean distances serial killers traveled from their home base to body disposal site for each offense.

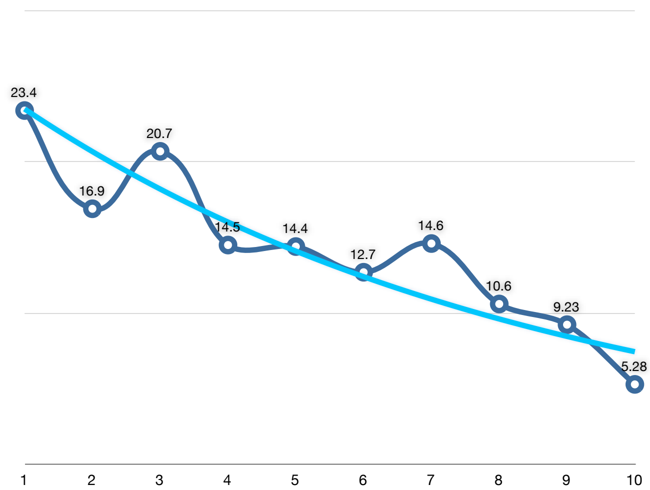

Figure 5: Relationship between mean of JTC and offense times.

As the regression equation shows in the graph, we have reason to assume there is a constant proportion between two successive crime sites based on time series data. That is to say, the expected distance from BASE to BD for the ith offense ($MJTC_i$) can be determined in the following formula.

$$MJTC_i = c \cdot d^i$$

c and d are empirically determined and also influenced by the crime area. In our regression equation, the values of c and d are respectively 26.70 and $e^{-0.12}(0.8869)$. We name this formula as the *Time-decay Formula*.

Many skewed distributions are best described by their mean [**?**], that explains why we pay attention to the mean distance to BD. Besides, the mean of the Weibull distributed variable JTC is calculated in the following formula:

$$\int_0^\infty \frac{b}{a}(\frac{x}{a})^{b-1}e^{-(\frac{x}{a})^b}dx = a\Gamma(\frac{1}{b}+1)$$

Combining this formula and the *Time-decay Formula*, we have the equation

$$a\Gamma(\frac{1}{b}+1) = c \cdot d^i$$

which means that the relationship between the two parameters of Weibull distribution is clear given the empirical values of c and d. Thus we can rewrite the probability density function of JTC for the ith offense as

$$WF(x; \frac{c \cdot d^i}{\Gamma(\frac{1}{b}+1)}, b)$$

.

Hence, using the Maximum Likelihood Estimation approach similar to method in section [5.3] we can calculate the likelihood function of the criminal's base, that is

$$P(x,y) = \prod_{i=1}^{T} WF(ND((x,y),(x_i,y_i)); \frac{c \cdot d^i}{\Gamma(\frac{1}{b}+1)}, b)$$

The coordinations of BASE and CSs are collected from [7]. Using the method mentioned above, we calculate the likelihood function of BASE and plot it. The following figures [6] and [7] show the spotlight area.

7.2.2 Robbery and Burglar

Different from the situation mentioned above, robbers and thieves have their own modus operandi. In [2], P.L. Brantingham and P.J. Brantingham illustrate the following picture to describe the pattern of serial crimes.

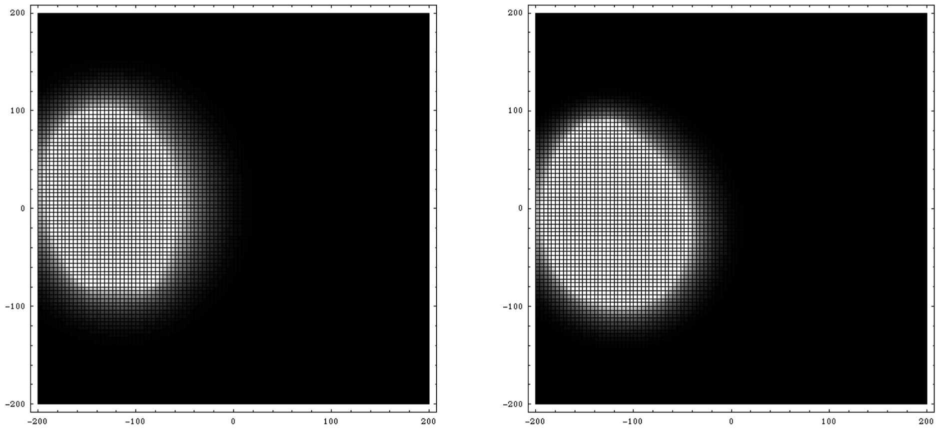

Figure 6: The left spotlight is plotted according to the basic Criminal Spotlight Model, while the right one is plotted with the introduction of time factor. Similar with, the origin represents the ground truth position of BASE. The coordination of CSs are (-122,84),(-160,59),(-149,79),(-150,-67),(111,260),(-217,-254),(253,-150),(-321,114),(-138,-13),(-90,-73).

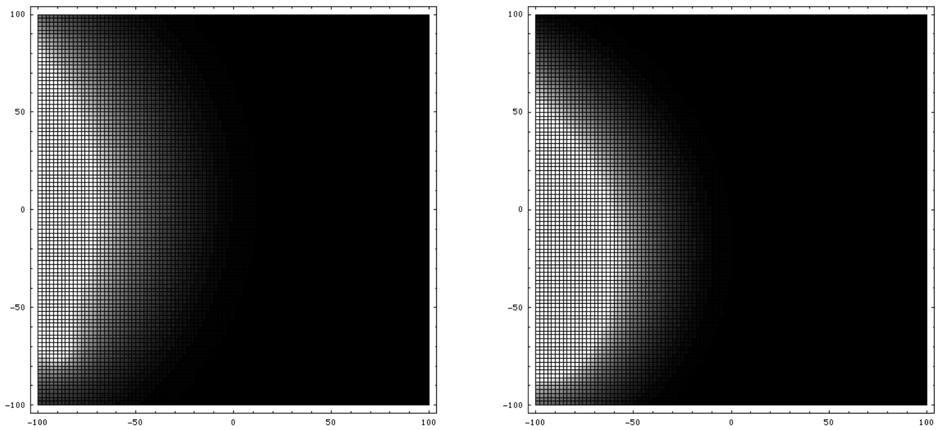

Figure 7: Zooming in the above images, we can find that although they looks alike, the right one is brighter near the origin, which indicates the improved one preforms better than the basic one.

Their pattern theory implies that robbers and thieves tend to commit crimes away from their earlier crime sites, which leads to another scheme for the police to capture the criminals. Besides searching the base of the criminal, the law enforcement officers can relocate the distribution of the police force. Furthermore, less police force would be put around the earlier crime sites and more would be arranged near other hot spots.

8 Simulation

In this part, we make a computer simulation to test the criminal spotlight model and to estimate the relationship between the number of victims and police force used. We first make a simulation of the criminal's behavior and then apply the Criminal Spotlight method to find out the criminal. Finally we use Monte Carlo to estimate the time needed to catch the criminal.

8.1 Criminal's Behavior

In order to simulate crime behavior, the offender's activity area is laid out on a 11×11-grid map, in the center of which the criminal's BASE locates with the coordination $(6, 6)$. There is a correspondence between the grids and the actual scene: unit Manhattan distance between grids represents half the median of JTC, which is determined by local crime statistics.

Here the previous data (Figure [1]) is used to indicate the relationship between the value of JTC and the possibility for the crime incident to happen in a certain area. Then we fill each grid with the frequency in the histogram, according to the Manhattan distance between the grid and the central grid.

In our simulator, each time the criminal randomly selects the victim in a grid according to the probability distribution on grids as shown above.

8.2 Algorithm

According to the Criminal Spotlight method, given the existing crime sites we can calculate the probability that the criminal's BASE locates in a certain point. Suppose police search for the criminal on the basis of the probability on each grid. Therefore, the probability that the criminal is caught equals the probability that the BASE locates in grid $(6, 6)$. P_i denotes the probability that the criminal is caught after his or her ith offense. Therefore $1 - (1 - P_1) \cdot (1 - P_2) \cdots (1 - P_i)$ indicates the possibility that the criminal has been captured before the $(i + 1)$th body dump. While the probability is not large enough, the criminal will make another body dump, that is increase i by 1.

0.00	0.00	0.00	0.00	0.05	0.39	0.05	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.05	0.39	0.67	0.39	0.05	0.00	0.00	0.00
0.00	0.00	0.05	0.39	0.67	3.35	0.67	0.39	0.05	0.00	0.00
0.00	0.05	0.39	0.67	3.35	2.94	3.35	0.67	0.39	0.05	0.00
0.05	0.39	0.67	3.35	2.94	4.17	2.94	3.35	0.67	0.39	0.05
0.39	0.67	3.35	2.94	4.17	0.00	4.17	2.94	3.35	0.67	0.39
0.05	0.39	0.67	3.35	2.94	4.17	2.94	3.35	0.67	0.39	0.05
0.00	0.05	0.39	0.67	3.35	2.94	3.35	0.67	0.39	0.05	0.00
0.00	0.00	0.05	0.39	0.67	3.35	0.67	0.39	0.05	0.00	0.00
0.00	0.00	0.00	0.05	0.39	0.67	0.39	0.05	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.05	0.39	0.05	0.00	0.00	0.00	0.00

Figure 8: 11×11-grid map indicating the activity area. Number in the grid indicates the probability that an offense happens here.

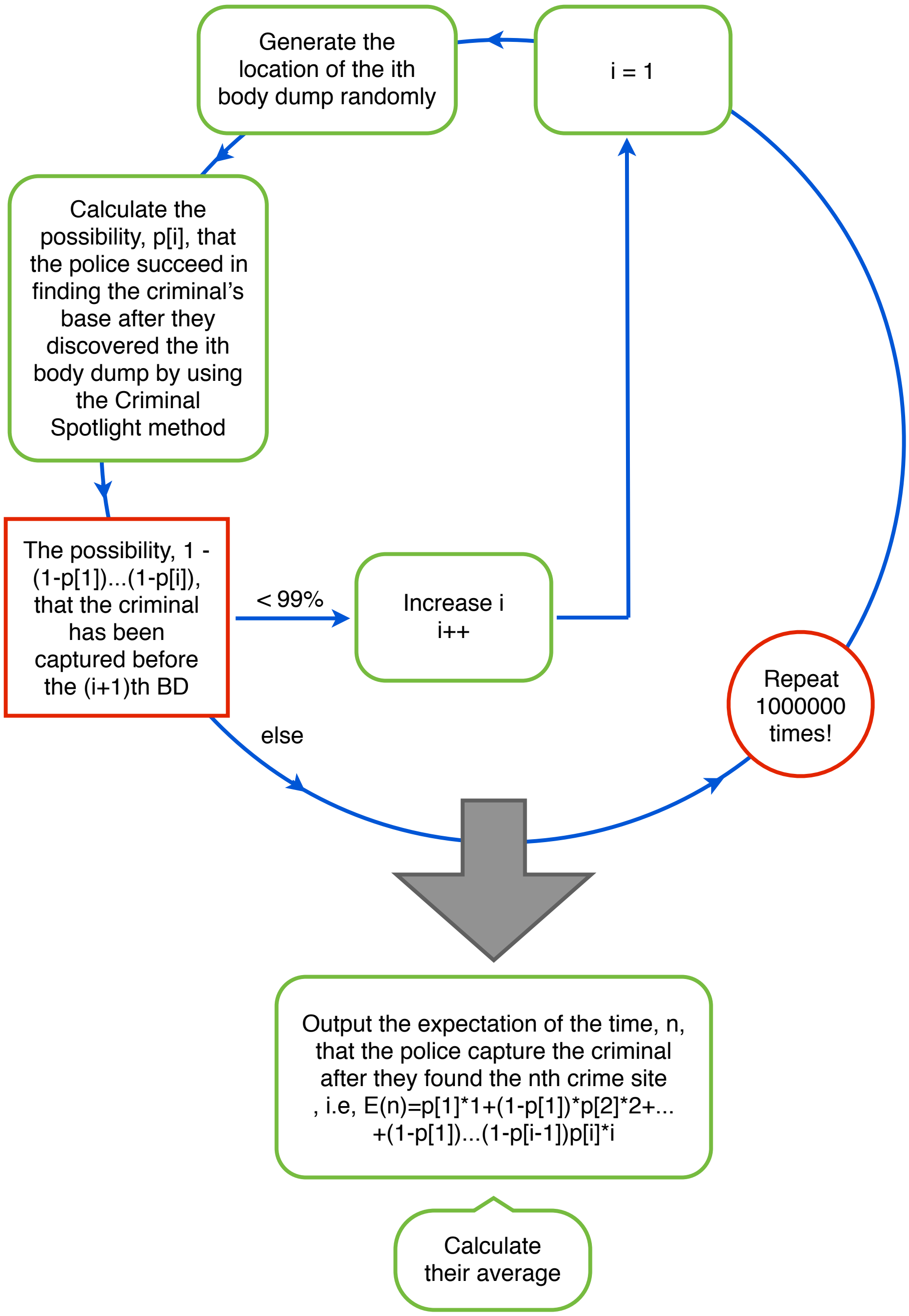

Figure 9: Algorithm Flow Chart of the simulation of Criminal Spotlight Model.

Otherwise, the criminal is caught before the $(i+1)$th body dump. In this situation, the expectation of n, the time to capture the criminal is

$$E(n) = P_1 \cdot 1 + (1-P_1) \cdot P_2 \cdot 2 + \cdots + (1-P_1) \cdots (1-P_{i-1}) \cdot P_i \cdot i$$

Repeat this procedure for 1000000 times, we can estimate the time needed to capture the criminal with considerable accuracy.

Next, we are introducing the concept Police Force Level. Previously, the probability that the criminal is caught equals the probability that the BASE locates in grid $(6,6)$, which is called Police Force Level 1. If the officer decide to put more police in the case, there will be more grids to be checked. If the Police Force Level is m, the probability that the criminal is caught equals the probability that the BASE locates in grids whose Manhattan distance to $(6,6)$ are $m-1$, which means that the police can check a larger square of grids centered in $(6,6)$ once a time.(Indicated in different color levels in Figure [8]).

According to the method mentioned above, we obtain the relationship between Police Force Level and the time needed to catch the criminal. With Police Force level 1, the expectation of time needed to needed to catch the criminal is 7.3; while with Police Force level 2, the expectation of time is 3.7.

From the above results, there is an estimation of the Police Force level needed to solve the case within certain days or certain number of victims.

8.3 Results

In order to make a more precise simulation, we enlarge the activity area into 21×21 grids, that is, unit Manhattan distance between grids represents 0.25 median of JTC.

Similarly we obtain the relationship between Police Force levels and time needed to catch the criminal, shown in the following histograms [11] [12] [13] [14].

There is an empirical formula between Police Force Level and mean of time need. Let x be the Police Force Level and y the mean of time need. Then

$$y = 21.97 \times 0.576^{\frac{x}{4}}$$

The expected time provides a criterion for the police to decide whether to change their method or not. Given a certain Police Force Level, if the time need to catch the criminal goes far beyond the mean, there possibly be problem with our assumption. Therefore it might be better for them to alter their method.

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00
0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00
0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.86	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00
0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.86	2.45	0.86	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02
0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.86	2.45	0.00	2.45	0.86	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12
0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.86	2.45	0.86	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02
0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.86	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00
0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.14	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00
0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.61	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	1.37	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.53	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.28	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.09	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.08	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Figure 10: A 2121-grid map indicating the activity area of the criminal. The number in grid is the probability that a crime would take place there. Column 11, row 11 is the criminals base.

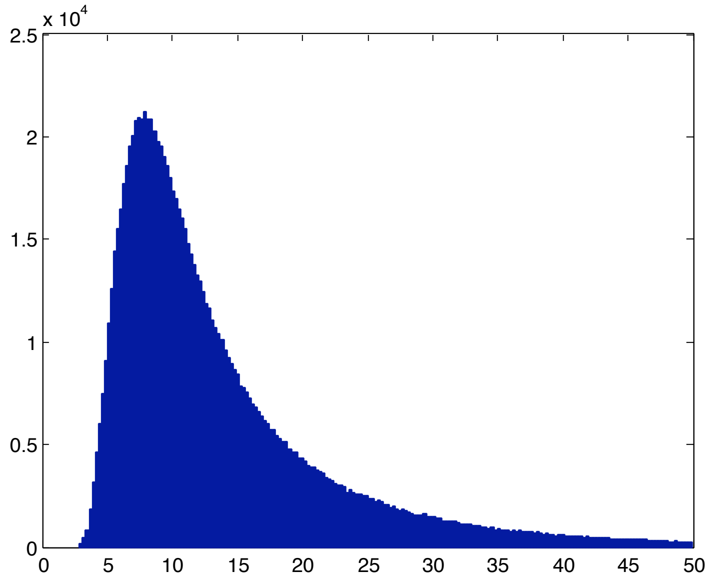

Figure 11: Police Force Level 1. Mean: 13.5344, Standard Deviation: 8.2457, Median:10.9543

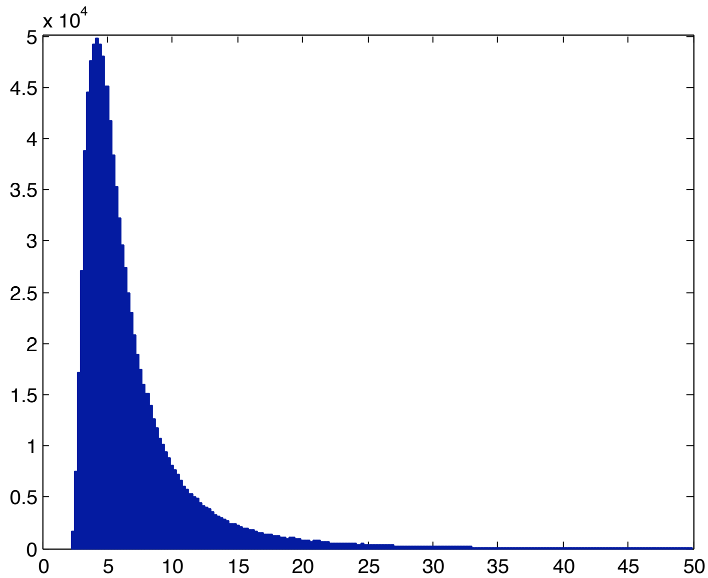

Figure 12: Police Force Level 2. Mean: 6.9581, Standard Deviation: 4.7654, Median:5.5084

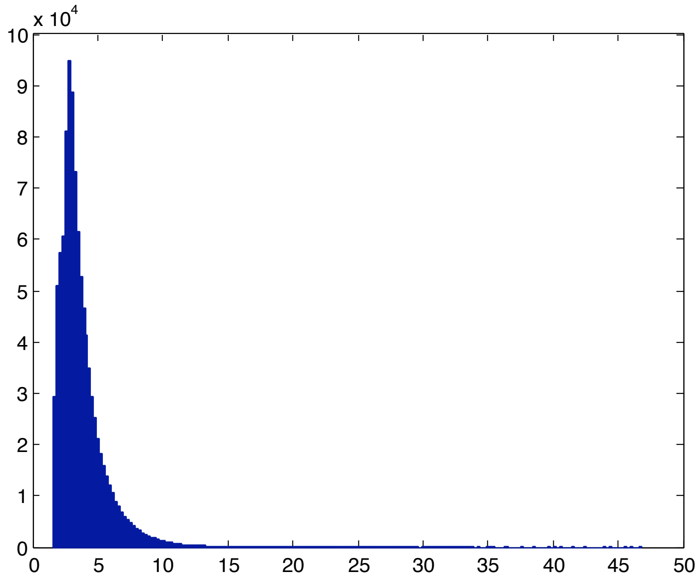

Figure 13: Police Force Level 3. Mean: 3.7337, Standard Deviation: 1.8520, Median: 3.2362

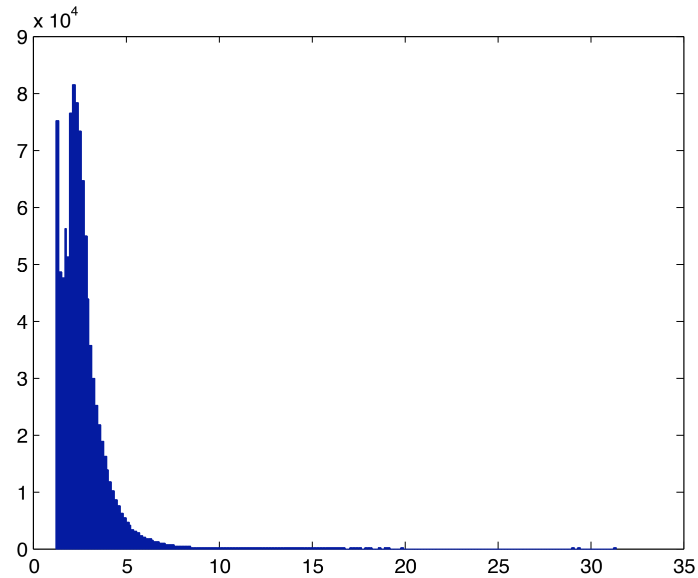

Figure 14: Police Force Level 4. Mean: 2.6397, Standard Deviation: 1.1186, Median: 2.4230

9 Model Analysis

9.1 Reliability

Apart from theoretical analysis of the model, we tested our method on several cases in 6. Within the 11 cases, there are 7 cases perfectly encompassed in the spotlight area(Richard Chase, Albert DeSalvo, Angelo Buono and Kenneth Bianchi, Peter Sutcliffe, Richard Ramirez, Aileen Wuornos, Ian Brady and Myra Hindley), which indicates that the model is of acceptable reliability in dealing with real cases.

9.2 Strengths

- The model is solidly based on standard mathematical statistical framework. Within the model, we utilize the quintessence of maximum likelihood estimation, Parzen window and even Bayesian Statistics, which make it both rational and efficacious.
- Compared with previous strategies such as centroid and minimum distance that gives an estimation of the BASE point, our method highlights an area of suspicion, which is definitely more precise than previous ones.
- This model makes better use of potential criminal information, such as time and local crime rate distribution, than other strategies, which leads to better results.
- Weibull distribution, the probability density function we adopt, is much more flexible than Rossmo's formula, which makes our result more suitable to the real situation.

9.3 Weaknesses

- Although Weibull distribution is of great flexibility, there might be other probability density function with better property for our model.
- For simplicity, it is assumed that the crimes are independent from each other, which might be doubtful because although some criminals commit their offenses randomly, there are still several criminals organized to killing certain victims, such as prostitute or the drunk.

- Our model ignores the influence of the action of police to the criminal. Suppose the criminal is wise enough or familiar with counterreconnaissance, he or she would change his behavior to evade the capture.
- Terrain is neglected for simplicity, which may lead to a unavailable spotlight area.
- Our method fails to distinguish various types of criminals, including murderers, rapists and burglars.

9.4 Warnings

In spite of the efficaciousness and reliability of the method, there are still several warnings when applying it to realistic questions.

- **Requirement for a minimum number of offenses.**

  Our method makes an estimation based on the information about offenses already happened. Therefore it is dependent on the number of existing offenses. The more information about offenses, the better accuracy of the estimation.

- **Update local demographic information as soon as possible.**

  The method is dependent on empirical data to a large extent, including mean of local JTC, local crime rate distribution, etc. In order to obtain a precise result, a set of up-to-date information is necessary.

- **Mind the "commuter".**

  It is assumed that there exists a BASE as an anchor for the criminal activity. However, it is possibly the case that there are more than one BASE for the criminal, so-called "commuter". To handle this situation, appropriate enlargement of spotlight area is needed.

Part IV
Street Control Model

10 Overview

10.1 Motivation

As stated above, [21] the mainstream Geographic Profilers assume that home is the criminal's base and the object of the police's investigation is to find out the criminal's living area (or in some cases, the possible location of the next crime). However, we have watched a range of crime films which reminds us that the police take some street control measures, such as the set up of roadblocks and the use of networks of closed-circuit television (CCTV) cameras, to reduce crime in addition to search for the suspect's home or the next probable crime site. These street control measures not only provide support in the navigation of a certain case, but also benefit a lot in the preventing and control of crime. This reminder throws light upon the thought that apart from focusing on the criminal's residence, the surveillance on the streets the criminal is likely to go across also works. Furthermore, many strategies across the study area simply ignore the restriction of street network and assume that the point locations can be located anywhere. Therefore we come up with our second model, the Street Control Model, based on the Graph Theory.

10.2 Terminology

In this model we will employ the notations in Graph Theory to denote the street network system.

Node An activity center in the considered city or town, where the population density is apparently higher than other areas, such as a residential area or a commercial district.

Edge A street that adjoins two activity centers in the city or town.

Path A sequence of streets the criminal may take from his or her home to another activity center.

Graph Combining all the activity centers and streets, the city or town considered is simplified to an undirected graph, which reflects the street network framework of the city or town.

PDBase[u] The probability density of the vertex u to be the BASE.

PDCS[u] The probability density of the vertex u to be a crime site.

10.3 Assumption

1. We assume that all the data obtained are reliable.
2. Similar to the assumption of our CS Model, we assume that there exists a BASE, in most circumstances the BASE refers to the criminal's home, from which the criminal sets out to commit offenses.
3. Each time the criminal commits an offense, he or she would randomly select a node, that is, an activity center, to search for his or her victim. The decision-making process involves the opportunity as well as the risk to do violence. Therefore we assume that this process is based on the crime rates of the neighborhood of different nodes.
4. According to Rational Choice Theory, the criminal has the tendency to select the shortest path from the home base to the crime site.

10.4 Theoretical Development

According to the assumptions, we can depict a city as an undirected connected graph, in which the vertices stand for the activity centers of the city and the edges the streets. The criminal will choose a vertex as his or her BASE randomly. The probability density of vertex u to be the BASE is $PDBase[u]$. Also, the criminal will choose a vertex as his or her crime site. The probability density will be denoted as $PDCS[u]$. We assign to each edge e a weight, $w[e]$, which is network distance (ND) between two ends of the edge. Till now, we have abstract a weighted undirected connected graph $G(V, E; w)$ from the map of a city. Here, V denotes all vertices in the graph and E denotes all edges.

According to the probability theory, the probability $p[e]$ that the criminal would pass through the edge e is

$$\sum_{u,v \in V} \delta(e; u, v) \times PDBase[u] \times PDCS[v]$$

where the function $\delta(e; u, v)$ equals 1 if the edge e lies in the optimal path from u to v, and equals 0 otherwise. Denote the individual term in the formula as $\psi(e; u, v)$.

The time complexity of calculating $p[e]$ of a given edge e would be $O(|V|^3)$ [24] if we use the simplest implementation of the Dijkstra's single-source shortest-paths algorithm to get the value of $\psi(e; u, v)$. It would cost us $O(|V|^3 \times |E|)$ to compute all values of $p[e]$. However, we can simply revise the Dijkstra's algorithm to lower the time complexity dramatically.

11 Implementation

11.1 Revised Dijkstra's Algorithm

In the chapter 24 of [6], Dijkstra's algorithm is sketched by the following pseudocode

```
function Dijkstra(Graph, source) :
  for each vertex v in Graph:
      dist[v] := infinity
      previous[v] := undefined
  dist[source] := 0
  Q := the set of all nodes in Graph
  while Q is not empty:
      u := vertex in Q with smallest dist[]
      if dist[u] = infinity: break
      remove u from Q
      for each neighbor v of u:
          alt := dist[u] + dist_between(u, v)
          if alt < dist[v] :
              dist[v] := alt
              previous[v] := u
  return dist[]
```

Through the process above, one can get the optimal path from the source to each vertex of the graph. Inspired by the algorithm above, we revised it to meet the need of computing the individual term $\psi(e; u, v)$. Thus we get the probability density $p[e]$ by adding the individual terms.

Figure [15] shows the flow chart of the Revised Dijkstra's Algorithm(RDA), in which the text in red presents how we refresh the probability density of each edge along the optimal path.

We run the RDA for each node whose PDCS is positive

In order to demonstrate the algorithm clearer, we investigate the hotspot region in Fyshwick, Canberra, Australia. The original data is excerpted from [11].

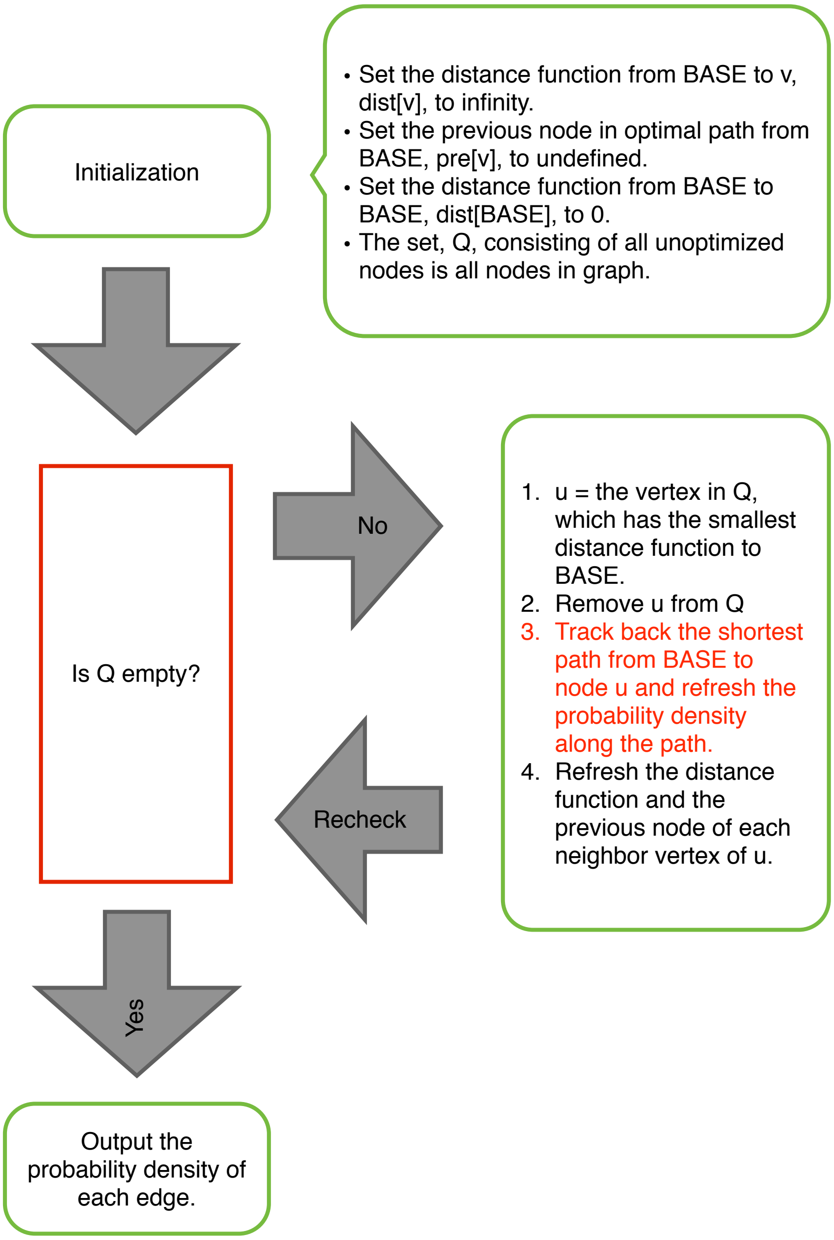

Figure 15: The flow chart of the revised Dijkstras single-source shortest paths algorithm.

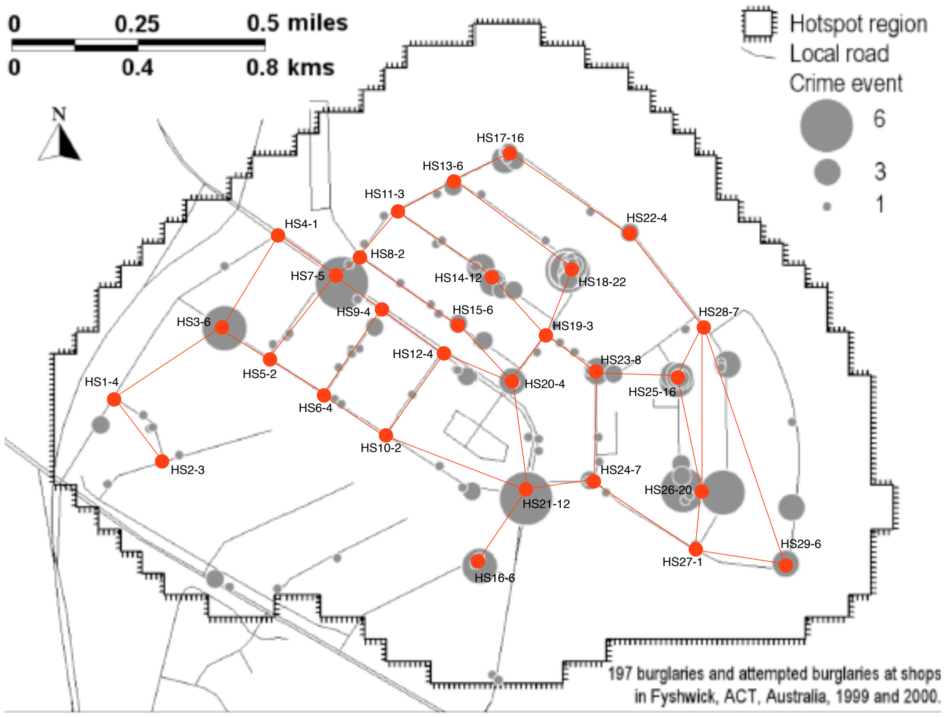

Figure 16: The abstracted graph presenting the nodes and streets of Fyshwick.

We abstract a graph, illustrated in Figure [16] and number each nodes of the city and assigned the value of PDCS to them, labeled on it in the format "HS No. - PDCS". The value of PDBase are assumed in the table.

HS NO	3	5	12	13	15	17	26	27
PDCS	7	9	3	2	2	8	3	6

Table 2: The PDBase of unlisted nodes are 0.

To simplify computation, we choose Euclidean distance as the Network Distance. First, let us start with the HS3 whose PDBase is positive 7. According to the algorithm above, we will find the optimal path from HS3 to HS2 and increase the probability density of along the path. That is to say, the probability density of the edge connecting HS3 and HS1 and the one connecting HS1 and HS2 will be increased by $PDBase[3] \times PDCS[2]$, say 7×3. We repeat this process for each possible BASE, and gain the probability density of each edge that the criminal will pass through. In Figure [17], we mark the edges with different colors to indicate the level of the probability.

These information would help the police to distribute the surveillance monitors on the streets.

11.2 Computational Complexity

In section 11.1, the implementation we store the graph $G(V, E; w)$ in an adjacent matrix. The running time would be estimated as $O(|V|^2)$. To lower the time complexity further, more advanced techniques should be applied. Noticing that the graph is planar and sparse, we can use a binary heap to implement extracting minimum from the distance function, $dist[u]$. The improved algorithm would require $O((|E| + |V|)log|V|)$ time, which is dominated by $O(|E|log|V|)$([24]). The total cost of time would be, at most, $O(|B||E|log|V|)$, in which B is set of all possible BASEs of the criminal.

12 Model Analysis

12.1 Reliability

Consider the previous formula

$$\sum_{u,v \in V} \delta(e; u, v) \times PDBase[u] \times PDCS[v].$$

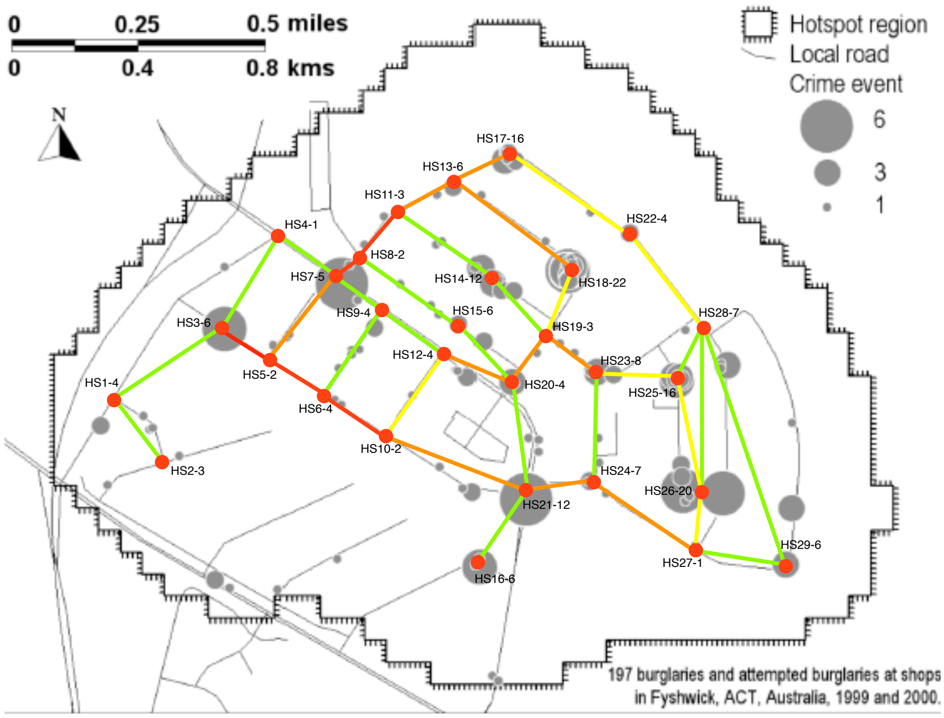

Figure 17: Green edges stand for the ones whose probability density is 0-500, yellow 500-1000, orange 1000-1500 and red 1500-2000.

When both the two factors $PDBase[u]$ and $PDCS[v]$ change by a small quantity, say, $\Delta(u)_1$ and $\Delta(v)_2$, the total change

$$\sum_{u,v\in V} \delta(e;u,v) \times (PDBase[u] + \Delta(u)_1) \times (PDCS[v] + \Delta(v)_2) - \sum_{u,v\in V} \delta(e;u,v) \times PDBase[u] \times PDCS[v]$$

is little, which means that this model is reliable.

12.2 Strengths

- The street network system is abstracted to a weighted undirected connected graph in our model. We only name a few arteries in the Fyshwick example. But in reality, the street network system can be far more complex than this example and special attention on a few suspected streets saves the police force much effort.
- As mentioned above, the street control system has been widely applied to reduce crime out of the police's experience in investigation. Our Street Control Model lays the theoretical foundation for this practice.
- Compared with the most strategies in Geographical Profiling, which generate the probable location of the criminal's home base, the advice this Street Control method provides is much easier to handle. In other words, the awareness of the criminal's probable location area is not enough to grasp the criminal. A detailed search for the residents living in this area is necessary and even a search warrant is required. Since many US cities followed suit and began installing the police surveillance cameras in the early 1990s, television monitors simply need to watch the videos taken by cameras in the suspected streets, which is much more convenient.
- The Street Control method not only supplies information to assist the investigation of a certain case, but also benefits in the preventing and control of crime.

12.3 Weaknesses

- The distribution of hotspot locations is not the unique factor and other factors should also be taken into consideration in the decision-process of the lawbreakers. The previous crime site has little tendency to be offended again and around the criminal's home there is probably a safe zone to reduce the risk, to name a few factors.

- Alarmed by more police force put into street control, the criminal might change his or her strategy, which makes the attempt worthless.

12.4 Warnings

We assume there exists a home base for the criminal, which implies that home is the anchor point. Thus our model is improper applied to the serial crime incidents of which the criminal is a commuter.

13 Further Discussion

The example of Fyshwick does not distinguish among different crime types, while in reality, the hotspot regions for different types of crime can be quite different. For example, robberies pay attention to some certain buildings such as the bank or the jewelry exhibition halls and murderers take more interest in the places where his or her target type of persons frequent. Diverse traffic situations confuse our choice of the shortest path. That is why we use Network Distance instead of simply using Euclidean Distance to take street network and traffic condition into account. Besides, traffic condition varies with time and the right moment for different types of crime should be considered.

Part V

Mixed Strategy

In order to capture the criminal, we combine Criminal Spotlight Model and Street Control Model together to help the Law Enforcement Officers to narrow down the range of the suspect.

We take the Fyshwick case as an example again. Suppose the local police discovered a serial burglar in the city. After applying Criminal Spotlight Method, they found the BASE of the criminal might be around the center of the city, as shown in Figure [18].

Thus, they can modify the PDBase of each vertex. The probability density of the nodes being the BASE should be set to 0 except for HS12 and HS15. After calculating the probability of each edge the criminal would

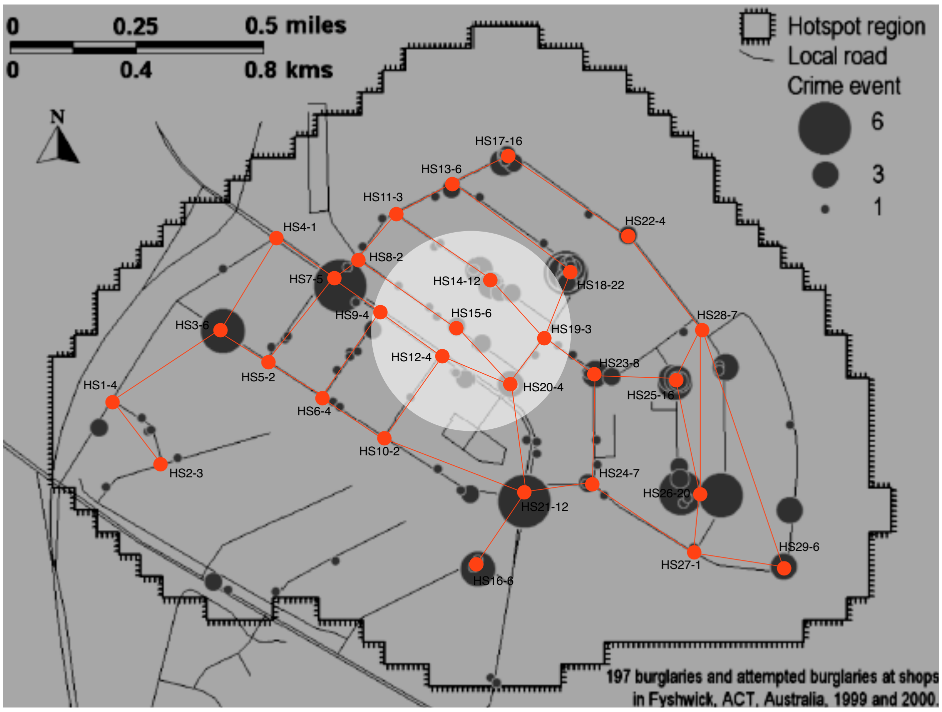

Figure 18: Criminal Spotlight Model highlighted the neighborhood of the criminals base.

pass through by Street Control Method, we can rank the streets from highest probability that the criminal would appear to the lowest. Part of the ranking information is revealed in the following table.

Streets	19-20	12-20	15-20	19-23	23-25	···
PD	452	414	278	267	227	···

We mark these 5 streets in Figure [19] with broader line.

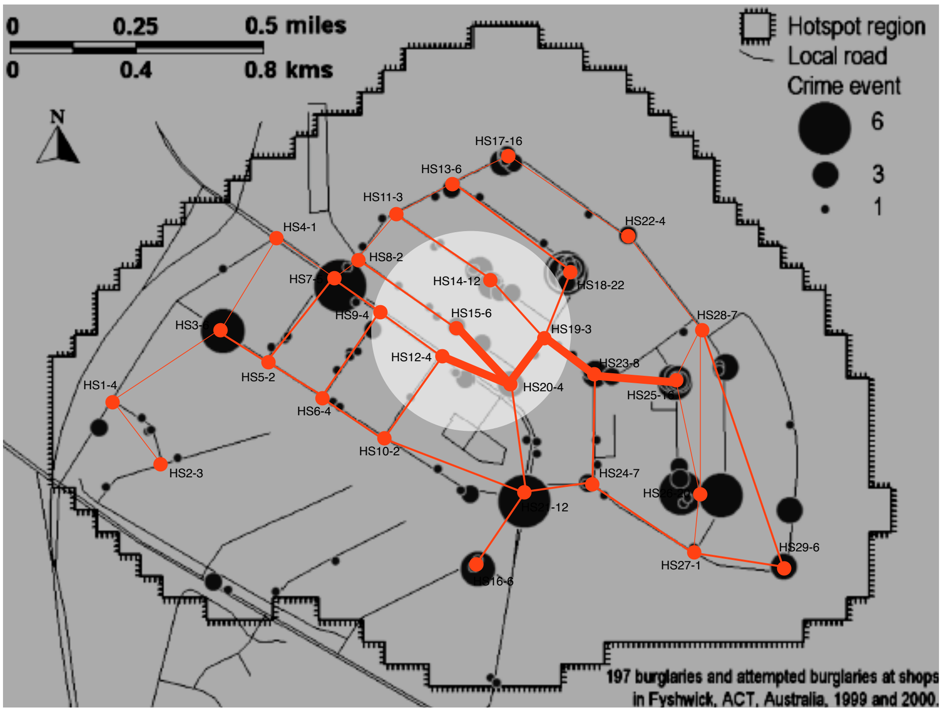

Figure 19: Top 5 suspect roads are marked by broader lines.

Hence, the police could examine the record of the monitors on these roads to find clues about the criminal. As we can see, the combination of the two schemes, Criminal Spotlight Model and Street Control Model, would help the local police to capture the criminal and collect more evidences of the serial crime.

Part VI
Conclusion

We formulated and tested two methods for aiding the police in their investigation of serial criminals.

The Criminal Spotlight Model makes an estimation of the location of the criminal's BASE. We implemented this method on serial murder cases data ranging from 1960s to 1980s and obtain preferable results.

The Street Control Model focuses on the street network with originality, rather than traditionally attempts to find out where the criminal lives. In proper positions setting up street control measures will help the police in a large extent.

Combination of these two methods will make them function better.

References

[1] Brantingham, L.P. & Brantingham, J.P. *el al.*(2008). *Crime Analysis at Multiple Scales of Aggregation: A Topological Approach* Chapter 4. Units of Analysis in Spatial Crime Research. Springer, New York

[2] Brantingham, L.P. & Brantingham, J.P.(1993). Environment, Routine, and Situation: Toward a Pattern Theory of Crime. *Routine Activity and Rational Choice* 5(2): 259-94

[3] Chainey, S. & Ratcliffe, J.(2005). *GIS and Crime Mapping.* ISBN: 9780470860984

[4] Canter, D. & Coffey, T. *et al.*(2000). Predicting Serial Killers' Home Base Using a Decision Support System. *Journal of Quantitative Criminology* 16(4)

[5] Comber, Se. L. S. *el al.*(2006). *Geographic Profiling and Animal Foraging.* Available online at http://www.sciencedirect.com

[6] Cormen, H. T. *et al.*(2001). *Introduction to Algorithms, Second Edition.* the MIT Press.

[7] Godwin, M. & Canter, D.(1997). Encounter and Death: The Spatial Behavior of US Serial Killers. *Policing: An Internationale Journal of Police Strategies and Management* 20: 24-38

[8] Crime Rate Map in US. http://geoggeol.mansfield.edu/media/images/crime-rate-map.jpg

[9] *CrimeStat III.* http://www.icpsr.umich.edu/icpsrweb/CRIMESTAT

[10] Dzemydiene, D. & Rudzkiene, V.(2002). Multiple Regression Analysis in Crime Pattern Warehouse for Decision Support. *Lecture Notes in Computer Science.* Springer.

[11] Jerry H.Ratcliffe.(2004). The Hotspot Matrix: A Framework for the Spatio-Temporal Targeting of Crime Reduction. *Police Practice and Research* 5(1): 05-23

[12] Laukkanen, M. & Santtila, P.(2005). *Predicting the Residential Location of A Serial Commercial Robber.* Available online at http://www.sciencedirect.com

[13] Levine, N., & Associates.(2000). *CrimeStat: A Spatial Statistics Programme for the Analysis of Crime Incident Locations*. Washington, DC: National Institute of Justice.

[14] Lu, Yongmei. (2005). Approaches for Cluster Analysis of Activity Locations along Streets: from Euclidean Plane to Street Network Space. [journal??]

[15] Phillips, P. & Lee, I.(2006). Mining Positive Associations of Urban Criminal Activities Using Hierarchical Crime Hot Spots. *Lecture Notes in Computer Science*

[16] Richard, N.K. & Harvey, J.I.(1997). An Analysis of Spatial Patterns in Serial Rape, Arson and Burglary: The Utility of the Circle Theory of Environmental Range for Psychological Profiling. *Psychiatry, Psychology, and Law*. 4 Psychiatry Psychol. & L.:195-206

[17] Rigel, the professional geographic profiling system, on the website of Environmental Criminology Research Inc.(ECRI). `http://www.ecricanada.com`

[18] Harries, K.(1999). *Mapping Crime: Principle and Practice*. US. Department of Justice Office of Justice Programs. `http://www.ncjrs.gov/html/nij/mapping/index.html`

[19] Robert, J. M. *Serial Murder: Multi-Disciplinary Perspectives for Investigators*. `http://www.fbi.gov/publications/serial_murder.htm`

[20] Rossmo, K. D.(1987). *Geographic Profiling: Target Patterns of Serial Murders*. Simon Fraser University.

[21] Snook, B. & Zito, M. *et al.*(2005). On the Complexity and Accuracy of Geographic Profiling Strategies. *Journal of Quantitative Criminology* 21(1)

[22] Snook, B., Taylor J.P., & Bennell, C.(2004). Geographic Profiling: the Fast, Frugal and Accurate Way. *Applied Cognitive Psychology* 18:105-21

[23] Stuart S.Kind.(1987). Navigational Ideas and the Yorkshire Ripper Investigation. *Journal of Navigation* 40:385-93

[24] Wikipedia. 2010a. *Dijkstra's algorithm.* `http://en.wikipedia.org/wiki/Dijkstra%27s_algorithm`

[25] Wikipedia. 2010b. *Geographic Profiling*. http://en.wikipedia.org/wiki/Geographic_profiling

A Results

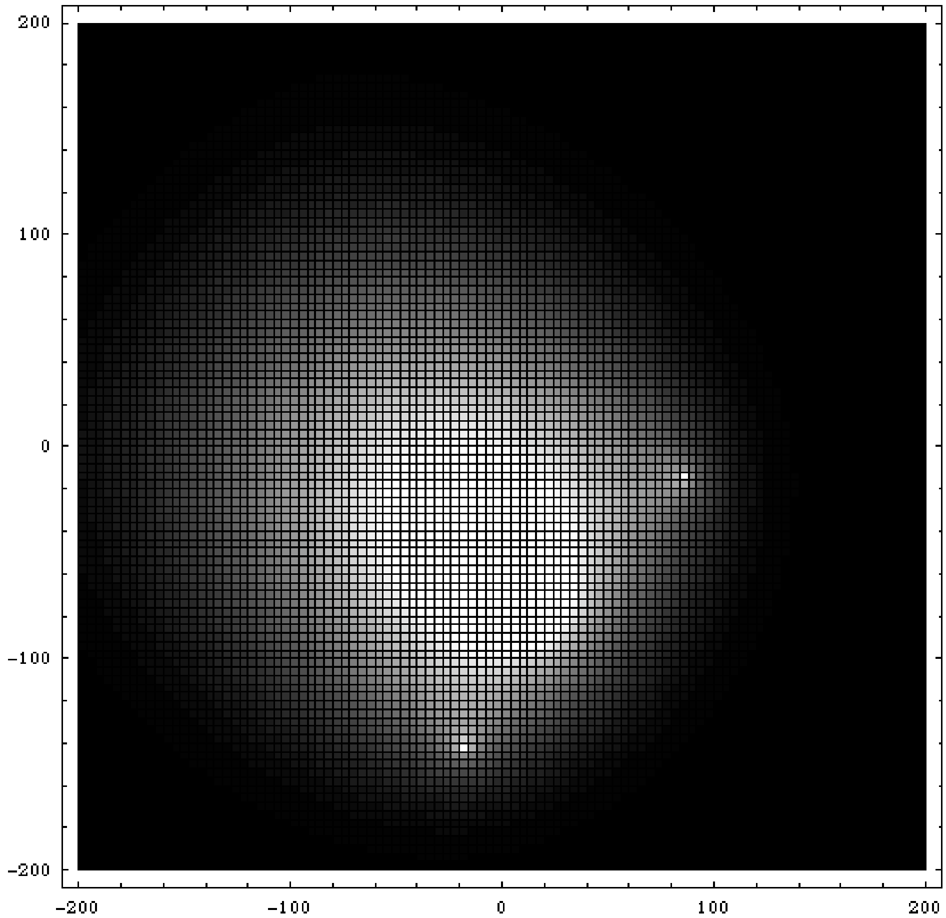

Figure 20: Richard Chase from Sacromento, CA, USA. The coordination of CSs are (-279,1), (-86,224), (-18,-143), (16,-78), (87,-15).

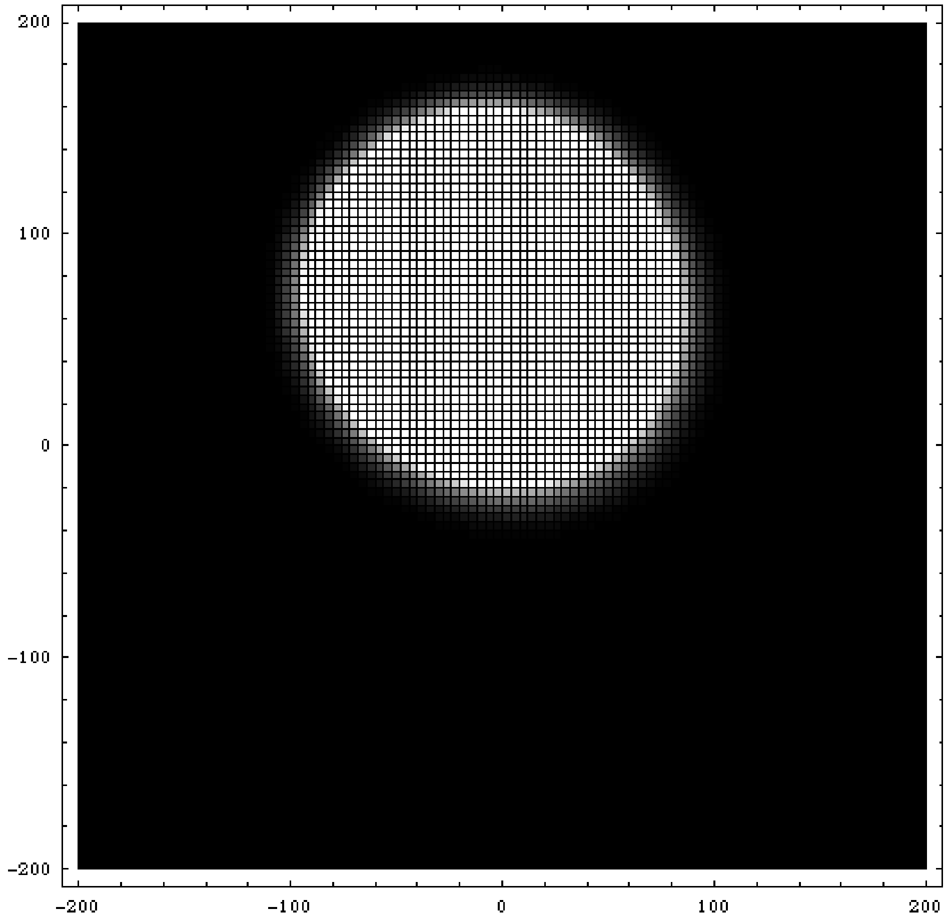

Figure 21: Albert DeSalvo from Boston, MA, USA. The coordination of CSs are (-66,91), (-48,62), (-29,88), (-18,84), (-17,94), (-9,91), (-7,132), (2,65), (-2,74), (1,83), (105,-51), (154,-99).

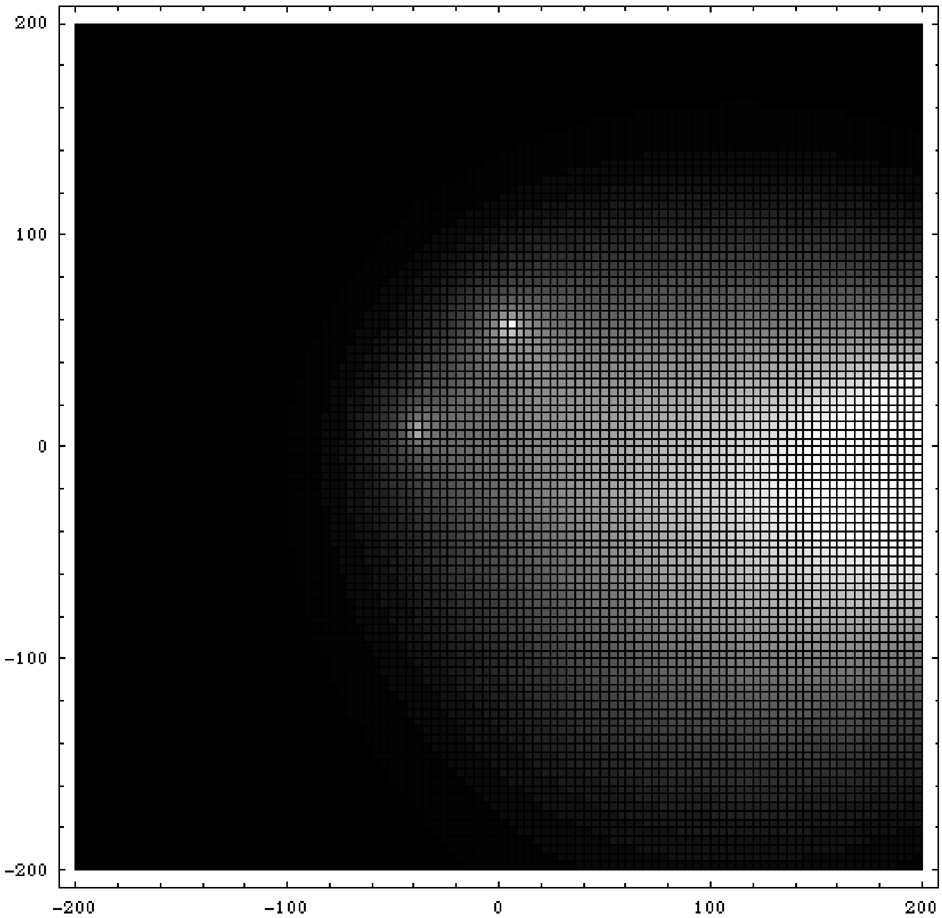

Figure 22: Clifford Olson from British Columbia, Canada. The coordination of CSs are (-39.5,8), (-10.5,-332), (4.5,58), (205.5, 15), (241.5,-43), (274.5,-6).

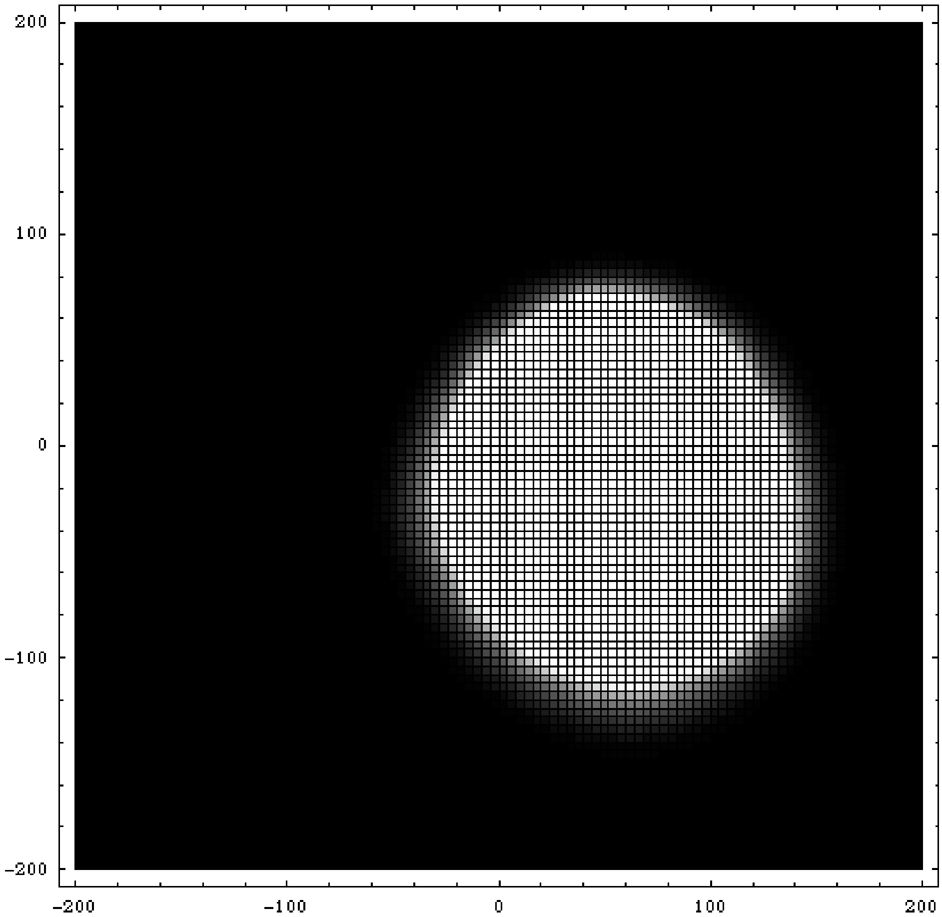

Figure 23: Angelo Buono and Kenneth Bianchi from LA, CA, USA. The coordination of CSs are (-68,-31), (18,19), (40,59), (52,49), (60, -194), (67,-2), (90,44), (123,-66), (132,-231).

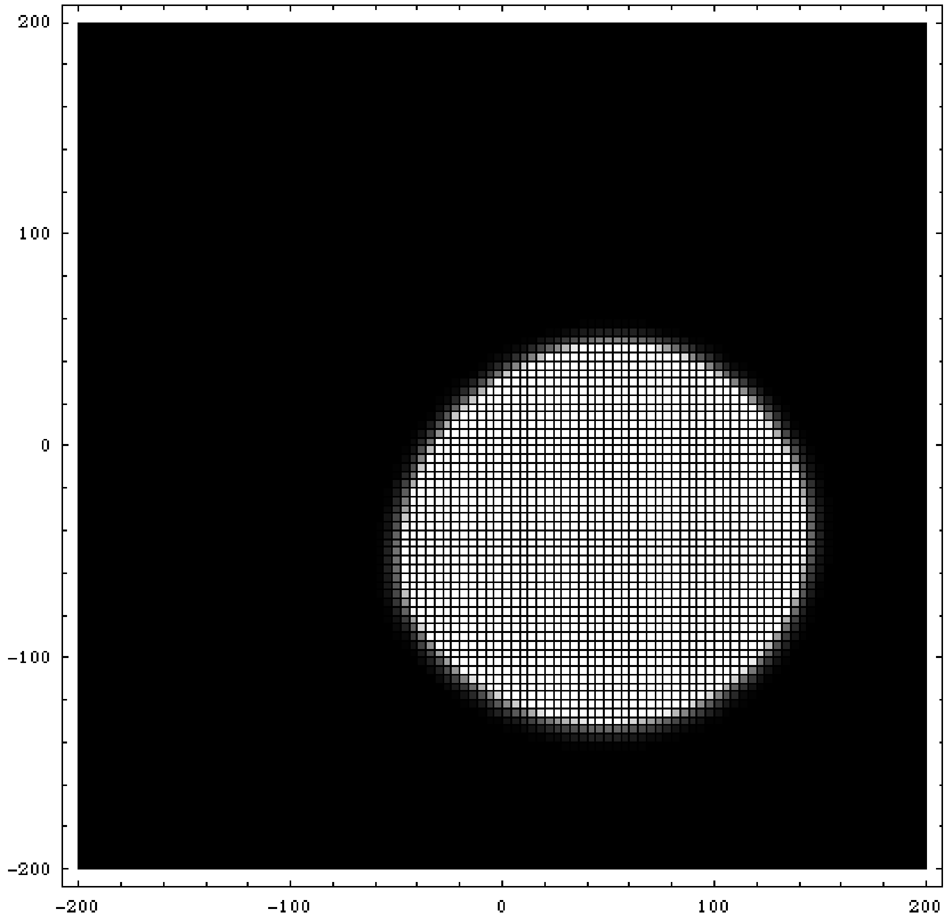

Figure 24: Richard Ramirez from LA, CA, USA. The coordination of CSs are (-132,-99), (-67,-86), (-33,-69), (-5,-36) , (-1,-50), (12,-46), (47,-8), (46,-4), (52,-8), (55,-4), (69,-5), (78,21), (72,-49), (77,-58), (77,-45), (82,-54), (86,-49), (98,-53), (185,4), (250, 224).

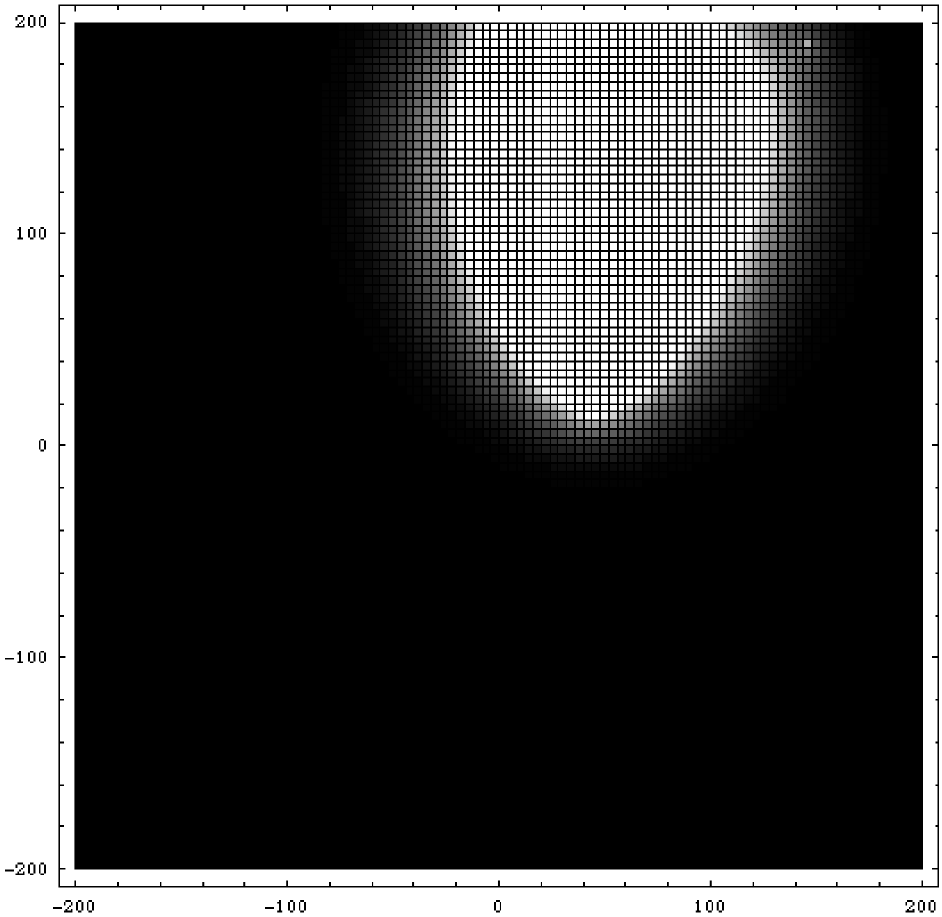

Figure 25: David Berkowitz from New York, NY, USA. The coordination of CSs are (-122.5,335.5), (19.5,214.5), (25.5,208.5), (37.5,57.5), (43.5,20.5), (50.5,32.5), (49.5,57.5), (62.5,148.5), (93.5,166.5), (148.5,191.5).

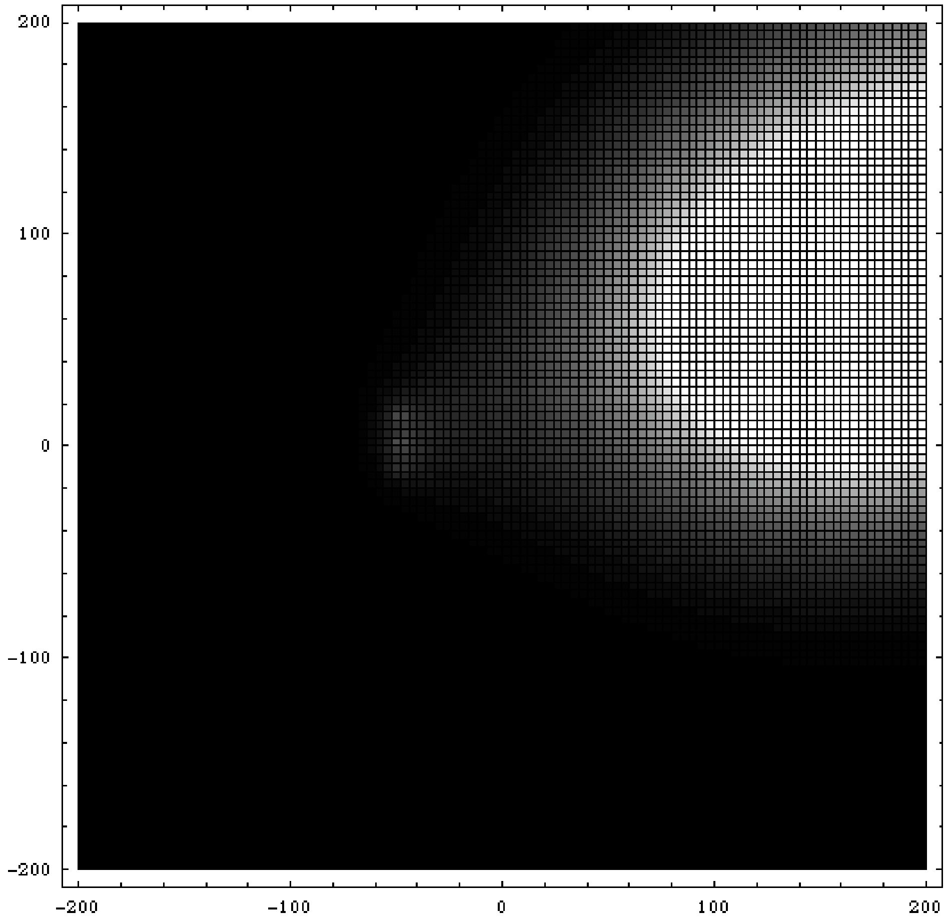

Figure 26: Jeffrey Dahmer from WI, USA. The coordination of CSs are (-49,-12), (-49,1), (-49,15), (134,51), (147,50), (343,50), (357,50), (357,196), (357,209), (357,296).

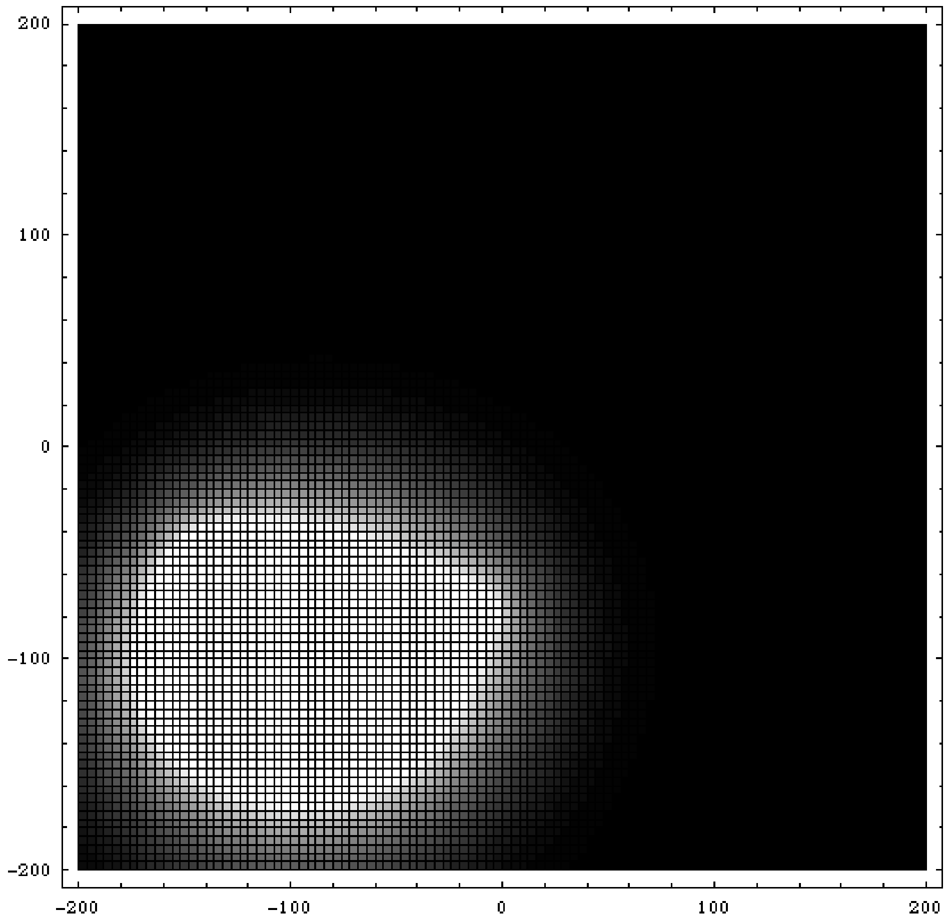

Figure 27: John Collins from MI, USA. The coordination of CSs are (-284,-352), (-137,-114), (-137,-100), (-143,-54), (-65,-123), (-10,-77), (178,0).

B C Code For Stochastic Simulation of 11×11 Grids

The following is the source code, written in C, simulating the 11×11-grid situation. The constant PF can be varied to simulate different police force.

```
#include <stdio.h>
#include <stdlib.h>

int main() {
    const int PF = 3;   //police force
    const int N = 100000;   //number of experiments
    int map[11][11] =
    {
        0,0,0,0,12,96,12,0,0,0,0,
        0,0,0,12,96,165,96,12,0,0,0,
        0,0,12,96,165,820,165,96,12,0,0,
        0,12,96,165,820,720,820,165,96,12,0,
        12,96,165,820,720,1020,720,820,165,96,12,
        96,165,820,720,1020,0,1020,720,820,165,96,
        12,96,165,820,720,1020,720,820,165,96,12,
        0,12,96,165,820,720,820,165,96,12,0,
        0,0,12,96,165,820,165,96,12,0,0,
        0,0,0,12,96,165,96,12,0,0,0,
        0,0,0,0,12,96,12,0,0,0,0
    };

    const int tt = 24480;   //sum of the elements in map
    double p[11][11], cp[100], sp, q, e, E = 0;
    int x, y, i, j, t, n, T;

    double WD[16] =
    {
        0, 107.408, 49.3492, 22.1727,
        9.26527, 3.57061, 1.26899, 0.416831,
        0.126899, 0.0359061, 0.00946793, 0.00233234,
        0.000537976, 0.000116433, 0.0000236896,
        0.00000453909
    };

    for (T = 0; T < N; T++)
```

```
{
    for (i = 0; i < 11; i++)
        for (j = 0; j < 11; j++)
            p[i][j] = 1;
    q = 1;
    e = 0;
    t = 0;
    while (q > 0.01)
    {
        t++;
        n = rand() % tt;
        for (x = 0; x < 11; x++)
        {
            for (y = 0; y < 11; y++)
            {
                n -= map[x][y];
                if (n < 0) break;
            }
            if (n < 0) break;
        }
        for (i = 0; i < 11; i++)
            for (j = 0; j < 11; j++)
                p[i][j] *= WD[abs(x - i) + abs(y - j)];
        sp = 0;
        for (i = 0; i < 11; i++)
            for (j = 0; j < 11; j++)
                sp += p[i][j];
        for (i = 0; i < 11; i++)
            for (j = 0; j < 11; j++)
                p[i][j] = p[i][j] * 100 / sp;
        cp[t] = 0;
        for (i = 0; i < 11; i++)
            for (j = 0; j < 11; j++)
                if (abs(i - 5) + abs(j - 5) < PF)
                    cp[t] += p[i][j];
        cp[t] /= 100;
        e += q * cp[t] * t;
        q = q * (1 - cp[t]);

    }
    E += e;
```

```
    }

    printf("%f", E / N);

    return 0;
}
```

C C Code For Stochastic Simulation of 21×21 Grids

The following is the source code, written in C, simulating the 21×21-grid situation. The user could modify the code to investigate the distribution of 10^6 experiments results.

```
#include <stdio.h>
#include <stdlib.h>

int main() {
    FILE * file = "";
    const int PF = 4;   //police force
    const int N = 100000;   //number of experiments
    const int S = 21; //scale of the map
    int map[21][21] =
    {
        0,0,0,0,0,0,0,0,0,5,25,5,0,0,0,0,0,0,0,0,0,
        0,0,0,0,0,0,0,0,5,25,17,25,5,0,0,0,0,0,0,0,0,
        0,0,0,0,0,0,0,5,25,17,19,17,25,5,0,0,0,0,0,0,0,
        0,0,0,0,0,0,5,25,17,19,57,19,17,25,5,0,0,0,0,0,0,
        0,0,0,0,0,5,25,17,19,57,108,57,19,17,25,5,0,0,0,0,0,
        0,0,0,0,5,25,17,19,57,108,280,108,57,19,17,25,5,0,0,
        0,0,0,0,0,5,25,17,19,57,108,280,125,280,108,57,19,17,
        25,5,0,0,0,0,0,5,25,17,19,57,108,280,125,233,125,280,
        108,57,19,17,25,5,0,0,0,5,25,17,19,57,108,280,125,233,
        175,233,125,280,108,57,19,17,25,5,0,5,25,17,19,57,108,
        280,125,233,175,500,175,233,125,280,108,57,19,17,25,5,
        25,17,19,57,108,280,125,233,175,500,0,500,175,233,125,
        280,108,57,19,17,25,5,25,17,19,57,108,280,125,233,175,
        500,175,233,125,280,108,57,19,17,25,5,0,5,25,17,19,57,
        108,280,125,233,175,233,125,280,108,57,19,17,25,5,0,
        0,0,5,25,17,19,57,108,280,125,233,125,280,108,57,19,
        17,25,5,0,0,0,0,0,5,25,17,19,57,108,280,125,280,108,57,
        19,17,25,5,0,0,0,0,0,0,0,5,25,17,19,57,108,280,108,57,
        19,17,25,5,0,0,0,0,0,0,0,0,0,5,25,17,19,57,108,57,19,
        17,25,5,0,0,0,0,0,0,0,0,0,0,0,5,25,17,19,57,19,17,25,
        5,0,0,0,0,0,0,0,0,0,0,0,0,0,0,5,25,17,19,17,25,5,0,0,0,
        0,0,0,0,0,0,0,0,0,0,0,0,5,25,17,25,5,0,0,0,0,0,0,0,0,
        0,0,0,0,0,0,0,0,0,5,25,5,0,0,0,0,0,0,0,0,0
    };
```

```
const int tt = 20400;   //sum of the elements in map
double p[S][S], cp[500], sp, q, e, E = 0;
int x, y, i, j, t, n, T;

double WD[50] =
{
    0, 14686.6, 6747.85, 3031.82, 1266.9, 488.233,
    173.517, 56.9961, 17.3517, 4.90968, 1.29461,
    0.318916, 0.073561, 0.0159206, 0.00323923,
    0.000620661, 0.000112176, 0.0000191524,
    0.0000030933, 0.00000046324, 0.00000006866,
    0.00000000946, 0.00000000124, 0.00000000015
};

for (T = 0; T < N; T++)
{
    for (i = 0; i < S; i++)
        for (j = 0; j < S; j++)
            p[i][j] = 1;
    q = 1;
    e = 0;
    t = 0;
    while (q > 0.01)
    {
        t++;
        n = rand() % tt;
        for (x = 0; x < S; x++)
        {
            for (y = 0; y < S; y++)
            {
                n -= map[x][y];
                if (n < 0) break;
            }
            if (n < 0) break;
        }
        for (i = 0; i < S; i++)
            for (j = 0; j < S; j++)
                p[i][j] *= WD[abs(x - i) + abs(y - j)];
        sp = 0;
        for (i = 0; i < S; i++)
```

```
                for (j = 0; j < S; j++)
                    sp += p[i][j];
            for (i = 0; i < S; i++)
                for (j = 0; j < S; j++)
                    p[i][j] = p[i][j] * 100 / sp;
            cp[t] = 0;
            for (i = 0; i < S; i++)
                for (j = 0; j < S; j++)
                   if(abs(i-(S-1)/2)+abs(j-(S-1)/2)<PF)
                     cp[t] += p[i][j];
            cp[t] /= 100;
            e += q * cp[t] * t;
            q = q * (1 - cp[t]);

        }
        E += e;
    }

    printf("%f", E / N);

    return 0;
}
```

D Program for the Street Control Model

The following is the source code, written in C, calculating the probability that the criminal would appear in a given street. The user could modify the array rsd[] if more information is provided by the Criminal Spotlight Model.

```
#include <stdio.h>
#include <stdlib.h>

#define NN 29 #define SN 40

int main() {
    const int D = 200000;//infinity
    const int hsd[NN] = {
        4, 3, 6, 1, 2, 4, 5, 2, 4, 2,
        3, 4, 6, 12, 6, 6, 16, 22, 3, 4,
        12, 4, 8, 7, 16, 20, 1, 7, 6
        }; //hotspot rank
    const int hsc[NN][2] = {
        252,542, 301,604, 361,469, 416,377, 408,502,
        462,538, 474,418, 499,400, 519,452, 524,577,
        535,353, 581,496, 591,325, 629,421, 596,467,
        616,704, 647,295, 710,412, 685,477, 650,523,
        665,631, 768,376, 734,513, 732,623, 816,520,
        840,634, 834,692, 841,469, 924,707
        }; //hotspot coordination
    const int e[SN][2] = {
        1,2, 1,3, 3,4, 3,5, 4,7,
        5,6, 5,7, 6,9, 6,10, 7,8,
        7,9, 8,11, 8,15, 9,12, 10,12,
        10,21, 11,13, 11,14, 12,20, 13,17,
        13,18, 19,18, 14,19, 15,20, 16,21,
        17,22, 19,20, 19,23, 20,21, 21,24,
        22,28, 23,24, 23,25, 24,27, 25,26,
        25,28, 26,27, 26,28, 27,29, 28,29
        }; //Connection Relation
    const int rsd[NN] = {
        0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 3, 0, 0, 2, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 0, 0
        }; //Residence Density
```

```
//You can change it when Criminal Spotlight Method is applied.

double dis[NN][NN], dx, dy, min;
int i, j, k, a, b;

for (i = 0; i < NN; i++) for (j = 0; j < NN; j++)
    if (i == j) dis[i][j] = 0;
    else dis[i][j] = -1;
for (i = 0; i < SN; i++)
{
    a = e[i][0] - 1;
    b = e[i][1] - 1;
    dx = hsc[a][0] - hsc[b][0];
    dy = hsc[a][1] - hsc[b][1];
    dis[a][b] = dis[b][a] = sqrt(dx * dx + dy * dy);
}

int vote[NN][NN] = {0};
double dijk[NN], alt;
int p[NN], q[NN];

for (i = 0; i < NN; i++)
{
    if (rsd[i] > 0)
    {
        for (j = 0; j < NN; j++)
        {
            dijk[j] = D;
            p[j] = -1;
            q[j] = 1;
        }
        dijk[i] = 0;
        for (k = 0; k < NN; k++)
        {
            min = D;
            for (j = 0; j < NN; j++)
                if (q[j] == 1 && dijk[j] < min)
                    {
                        a = j;
                        min = dijk[j];
                    }
```

```
                q[a] = 0;
                for (j = 0; j < NN; j++) if (dis[a][j] != -1)
                {
                    alt = dijk[a] + dis[a][j];
                    if (alt < dijk[j])
                    {
                        dijk[j] = alt;
                        p[j] = a;
                    }
                }
                b = a;
                while (p[b] != -1)
                {
                    vote[b][p[b]] += rsd[i] * hsd[a];
                    vote[p[b]][b] += rsd[i] * hsd[a];
                    b = p[b];
                }
            }
        }
    }

    for (i = 0; i < NN; i++) for (j = i + 1; j < NN; j++)
        if (dis[i][j] != -1)
            printf("%d\t%d\t%d\n", i + 1, j + 1, vote[i][j]);

    return 0;
}
```

Appendix E: Executive Summary

For the years 1989-2008, the average number of violent crimes reported in the USA annually, according to the Uniform Reports, was 1,590,440, and the average number of property crimes was 11,163,675. According to the same resource, the overall number of murders in 2008 is 14,180. Furthermore, stranger murders account for 12 percent of all murders reported in the USA. It is suspected that many of these strange-related murders are committed by “serial killers”.

The increasing impact of serial crimes is likely to provoke the law enforcement agencies to respond to those actions. We present here our new approaches dealing with serial criminals’ geographical profiling. The following table gives a decision tree for agencies to determine which of our schemes are suitable in the given situation. We treat two kinds of crime, violent crime and property crime, separately. Our basic assumption is that the data required in the solution are accurately investigated and recorded.

<table>
<tr><th colspan="4">Serial Violent Crimes: Murder, Forcible Rape, Aggravated assault, etc.</th></tr>
<tr><td rowspan="2">The criminal focus on a special type of people. For example he or she choose the victims of a specified age range, race or sex.</td><td>Data of the distribution of this particular type of offender are available.</td><td>1.Based on the empirical behavior pattern of criminals of this type and the distribution data, Criminal Spotlight Model (CSM) will light up the neighborhood of the offender’s base.
2.The statistical information of hotspots and the distribution of this kind of offender would allow Street Control Model (SCM) to generate the probability that the criminal would pass through a given street.
3.Using the results of step 1 and 2 comprehensively, police could search around the ‘lighten-up area’ and watch the monitor record of suspect roads.</td><td rowspan="4">Optional Choice:
If the criminals tend to offend closer to their bases or further from their bases, elaborate CSM process will be taken based on these phenomenon.
Attention: if the number of the victims have exceeded the empirical value
21.97×0.57^{PF}
severely, the suitability of our model should be reconsidered. Here, PF represents the police force, defined by dividing the search radius with the median length of all journeys to crime.</td></tr>
<tr><td>Those data are missing.</td><td>1.Replace distribution data in the above situation by the distribution of crime rate in the city and run through all steps mentioned above.</td></tr>
<tr><td rowspan="2">The criminal randomly choose his or her object.</td><td>Data of the distribution of crime rate are available.</td><td>1.Based on the empirical behavior pattern of criminals of this type and the distribution data, Criminal Spotlight Model (CSM) will light up the neighborhood of the offender’s base. Police could investigate around this area or narrow down the name list of suspects.</td></tr>
<tr><td>Those Data are missing.</td><td>1.Replace distribution data in above situation.</td></tr>
</table>

Serial Property Crimes: Burglary, Larceny-theft, Motor vehicle theft, etc.	
The criminal is judged to have an anchor point, such as home, by police's experience or definite evidences.	1.Based on the given information, CSM will light up the neighborhood of the offender's base on the map. 2.SCM will generate the probability that the criminal would pass through a given street. 3.Using the results of step 1 and 2, police could investigate the 'lighten-up area' and watch the monitor record of suspect roads. After the criminal is caught, the result of step 2 will help the police to collect evidences more efficiently.
The criminal does not have an activity center.	None of our methods are suitable for this situation.

The following describes the technical details of two models mentioned above.

Criminal Spotlight Model

Find two parameters a and b in the Weibull distribution function

$$\mathrm{W}(x;a,b)=\begin{cases}\dfrac{b}{a}\left(\dfrac{x}{a}\right)^{b-1}e^{-\left(\frac{x}{a}\right)^{b}} & x\geq 0\\ 0 & x<0\end{cases}$$

according to the empirical data of criminals' behavior pattern. Suppose the crime sites are CS[1], CS[2], ..., CS[n]. Plot on the map the probability density function of Place P

$$\mathrm{Base}(\mathrm{P})\times\prod_{i=1}^{n}\frac{\mathrm{WF}\left(\mathrm{D}\left(\mathrm{P},\mathrm{CS}_i\right);a,b\right)}{2\pi\,\mathrm{D}\left(\mathrm{P},\mathrm{CS}_i\right)},$$

where BASE() stands for the distribution of a particular kind of offender or the crime rate in different situations, D(P, CS[i]) stands for the network distance between these two places. It is recommended to use computer-aided method, such as the *DensityPlot* command in *Mathematica*, to plot the function. If the criminals' behavior pattern generally has a correlation with the temporal factor, we can gain a finer result by adjust the parameter, a, in each individual item of the function above.

Street Control Model

Firstly, we use an undirected connected graph G(V, E; w) to depict the map of city. Here, V represent the set of all nodes and E all streets. The function w records the network length of each edge. Given the data of the probability density of crime sites and criminals' base on the graph, one can get the probability that the criminal would pass through a given edge, e, by the following formula,

$$\mathrm{P}(e)=\sum_{u,v\in V}\delta(e;u,v)\times\mathrm{PDB}(u)\times\mathrm{PDC}(v),$$

where the delta function equals 1 if e lies on the shortest path from u to v, and equals 0 for other cases. After revising Dijkstra's single-source shortest path algorithm, one can calculate P(e) for each edge at the cost of time of complexity O(|V|^2 log|V|).