Summary

Faced with serial crimes, we usually estimate the possible location of next crime by narrowing search area. We build three models to determine the geographical profile of a suspected serial criminal based on the locations of the existing crimes. Model One assumes that the crime site only depends on the average distance between the anchor point and the crime site. To ground this model in reality, we incorporate the geographic features G, the decay function D and a normalization factor N. Then we can get the geographical profile by calculating the probability density. Model Two is Based on the assumption that the choice of crime site depends on ten factors which is specifically described in Table 5 in this paper. By using analytic hierarchy process (AHP) to generate the geographical profile. Take into account these two geographical profiles and the two most likely future crime sites. By using mathematical dynamic programming method, we further estimate the possible location of next crime to narrow the search area. To demonstrate how our model works, we apply it to Peter's case and make a prediction about some uncertainties which will affect the sensitivity of the program. Both Model One and Model Two have their own strengths and weaknesses. The former is quite rigorous while it lacks considerations of practical factors. The latter takes these into account while it is too subjective in application. Combined these two models with further analysis and actual conditions, our last method has both good precision and operability. We show that this strategy is not optimal but can be improved by finding out more links between Model One and Model Two to get a more comprehensive result with smaller deviation.

Key words: geographic profiling, the probability density, anchor point, expected utility

Executive Summary

Nowadays, in a serial crime, the spatial distribution of crime sites is arousing the more and more attention of the criminologists and the geographers. The Serial criminals, in a spirit of defiance, have endangered public security, gravely infringed on the citizens' personal safety, lives and property, and are abhorred by the people across the country. Since the offenders usually have no stable residence, it is very difficult for the police to find out and arrest them. Therefore, in order to help the police solve the crime as soon as possible to maintain social security and stability, a more sophisticated technique is in urgent need to be developed to determine the "geographical profile" of a suspected serial criminal based on the locations of the crimes.

This paper presents three methods, especially a new mathematical method combining the advantages of the two previous models, which can generate a useful prediction for law enforcement officers about possible locations of the next crime.

Based on Bayesian statistical methods, Model One makes explicit connections between assumptions on offender behavior and the components of the mathematical model. It also takes into account local geographic features that either influence the selection of a crime site or influence the selection of an offender's anchor point. What's more, with rigorous inference formulas, this model has both precision and operability.

Model Two uses analytic hierarchy process (AHP). It takes full account of a variety of factors relevant to the crime sites, assisting the police to take measures adapted to local conditions so as to improve our work.

The last method is composed from Model One and Model Two. Taken the expected utility and other practical factors into consideration, it further estimate the geographic profile generated by the previous two models.

To conclude, we suggest the policemen put this mathematical method about geographical profiling into practice for it will be of significant assistance to law enforcement.

The technical details are as follows:

First of all, enhance the consciousness of the public social security and its improvement in the potential criminal area and inform the local people that a series of crimes have occurred recently, reminding them to keep vigilant and not to go into remote areas alone.

Second, the police departments should focus their activities, geographically prioritize suspects, and concentrate saturation or directed patrolling efforts in those zones where the criminal predator is most likely to be active.

Third, after arresting the criminal, the police need to make experiential analysis on all these kinds of serial crimes to prevent a similar case. What's more, the police departments set up an enhanced intelligence exchange network, especially in the area which is the next crime site according to our prediction.

Finally, search the suspect in the predicted criminal location or the likely

residence of the offender.

Generally speaking, our method has good maneuverability and practicability. However, there are various uncertain factors in the situation which cannot be predicted such as the weather condition and the traffic conditions. Additionally, the determination and analysis of weight-coefficients on the various factors is very subjective. As a result, the actual site scene of next crime may be outside of our geographical profiling. Therefore, the police deployments must be based on the analysis of local actual condition instead of applying our model blindly.

Furthermore, our prerequisite that the offender has the only one stable anchor point differs from the actual conditions. Since the offender is very likely to change his residence, the police had better search the suspect according to the latest information.

Maintaining social stability and ensuring the safety of residents will bring well-being and peace to all mankind. Every one of us should make effort to make our society become better. Hope that this paper will be of some help to prevent the criminal behaviors.

1. Introduction

Clues derived from the locations connected to violent repeat criminal offenders, such as serial murderers, arsonists, and rapists, can be of significant assistance to law enforcement. Such information helps police departments to focus their activities, geographically prioritize suspects, and to concentrate directed patrolling efforts in those zones where the criminal offender is most likely to be active. By examining spatial data connected to a series of crime sites, this methodological model generates a probability map that indicates the area most likely to be the locations of the next crime.

This paper presents two mathematical models to illustrate how geographical analysis of serial crime conducted within a geographic information system can assist crime investigation. Techniques are illustrated for determining the possible residence of offenders and for predicting the location of the next crime based on the time and locations of the existing crimes.

First, we present a mathematical survey of some of the algorithms that have been used to solve the geographic profiling problem. The geographic profiling problem is the problem of constructing an estimate for the location of the anchor point of a serial offender from the locations of the offender's crime sites.

The approach that we develop will make use of at these two different schemes to generate a geographical profile. What's more, we develop a third technique to combine the results of the two previous schemes and generate a useful prediction for law enforcement officers. The prediction provides some kind of estimate or guidance about possible locations of the next crime based on the time and locations of the past crime scenes. Our method will also provide some kind of estimate about how reliable the estimate will be in a given situation, including appropriate warnings. The executive summary will provide a broad overview of the potential issues. It will also provide an overview of our approach and describe situations when it is an appropriate tool and situations in which it is not an appropriate tool.

The purpose is to apply geographical analysis to serial crime investigations to predict the location of future targets and determine offender residence.

1.1 Restatement of the Problem

In order to indicate the origin of geographical profiling problems, the following background is worth mentioning.

In 1981 Peter Sutcliffe was convicted of thirteen murders and subjecting a number of other people to vicious attacks. One of the methods used to narrow the search for Mr. Sutcliffe was to find a "center of mass" of the locations of the attacks. In the end, the suspect happened to live in the same town predicted by this technique. Since that time, a number of more sophisticated techniques have been developed to determine the "geographical profile" of a suspected serial criminal based on the locations of the crimes.

Our team has been asked by a local police agency to develop a method to aid in their investigations of serial criminals. The approach that we develop will make use of at least two different schemes to generate a geographical profile. We also develop a technique to combine the results of the different schemes and generate a useful prediction for law enforcement officers. The prediction will provide some kind of estimate or guidance about possible locations of the next crime based on the time and locations of the past crime scenes. Our method will also provide some kind of estimate about how reliable the estimate will be in a given situation, including appropriate warnings.

1.2 Survey of Previous Research

Existing Methods

To understand how we might proceed let us begin by adopting some common notation:

- A point x will have two components $x = (x^1, x^2,)$.
  - These can be latitude and longitude
  - These can be the distances from a pair of reference axes
- The series consists of n crimes at the locations $x_1, x_2, \cdots, x_n$.
- The offender's anchor point will be denoted by z .
- Distance between the points x and y will be $d(x, y)$.

Existing algorithms begin by first making a choice of distance metric d ; they then select a decay function f and construct **a** hit score function $S(y)$ by computing

$$S(y) = \sum_{i=1}^{n} f(d(x_i, y)) = f(d(x_1, y)) + \cdots + f(d(x_n, y))$$

Regions with a high hit score are considered to be more likely to contain the offender's anchor point than regions with a low hit score. In practice, the hit score $S(y)$ is not evaluated everywhere, but simply on some rectangular array of points $y_{jk} = (y_j^1, y_k^2)$ for $j \in \{1, 2, \cdots, J\}$ and $k \in \{1, 2, \cdots, K\}$ giving us the array of values $S_{jk} = S(y_{jk})$ [4].

Rossmo's method, as described in (Rossmo, 2000, Chapter10) chooses the Manhattan distance function for d and the decay function

$$f(d)=\begin{cases}\dfrac{k}{d^h} & \text{if } d<B\\ \dfrac{kB^{g-h}}{(2B-d)^g} & \text{if } d\geq B\end{cases}$$

Other Studies

In the study of Crime Analyst's task, Bryan Hill considers that the use of the "probability grid method" (PGM) can narrow the search as it pertains to tactical crime analysis in the ArcView Geographic Information Systems (GIS) environment. The main point of his theory will be that any current statistical method of predicting the next hit location in a crime series is operationally ineffectual when the suspect covers a large geographic area. When the analyst combines several statistical methods and intuitive, logical thought processes into a combined "grid" score, the analysis product can be made more operationally effective. This new grid surface allows the analyst to make a better prediction of the next hit in a crime series and is useful in isolating specific target locations for law enforcement deployment efforts. This easy to apply "PGM method" allows the analyst to use sound statistical methods, as well as their experience and knowledge of a crime series to narrow the focus and potential hit area.

In addition, when the crime series has sufficient suspect information, journey to crime analysis using the CrimeStat software can be used to provide investigators with a list of probable offenders from law enforcement records available to the analyst[4].

2. Model Overview

1、Model One

In Model One, we assume that the crime sites depend on the distance between the anchor point and the crime site. Generate a hot zone model according to probability formula, then combine the geographic features G with the decay function D and add a normalization factor N into the model .At last, we can get the "geographical profile".

2、Model Two

In Model Two, we take account of other factors relevant to crime site location, we have studied various literature to summary two aspects denoted by U(Utility from crimes) and P(the probability of success) specifically including the following ten factors: the responding speed of the police, public security situation, resistances' diathesis, density of registered inhabitants, the advantageous position, the number of offenses, the distance from the anchor point of the offender, the time required for committing the crime, number of target persons, offender's mental satisfaction from the crime.

Then, we use Analytic Hierarchy Process (AHP) to get weighted factors. Finally, according to the formula we can draw E of some areas. A wide range of criminal area can then be divided into several small areas. Then, we give the proper scores for each small area according to the actual condition. When the area is small enough, we can get the geographical profile. Sites with higher scores have high probability to be crime areas.

3、The two models can be synthesized as one method

For their different emphasis (The theory of Model One is quite rigorous, but it ignores the factors such as the geographic features and criminal motivation while these factors is important for selecting criminal crime sites. On the other hand, Model Two takes these into account, while it is too subjective in practical application which may easily cause the deviation), the geographical profiles we get from two models must be different. By using mathematical dynamic programming method to establish the model, this method explains the common rules of the offender's choice on the crime sites. Therefore, when we get several geographical profiles from the previous two models, we can use this method to predict the next offence site.

To sum up, we can generate the geographical profile in the investigations of serial criminals, providing some kind of estimate or guidance about possible locations of the next crime and narrow the search.

3. Model One

Symbols:

symbols	meaning
$P(x)$	probability density function
x	the crime sites
z	the location of the offender's anchor point
a	The average distance that the offender is willing to travel to offend.
D	the effect of distance decay
G	the geographic features
N	a normalization factor

Hypothesis：

1、 We assume that our offender chooses potential locations to commit crimes randomly according to some unknown probability density function P.

2、 We assume that P depends upon z and a.

3、 We suppose that that the values of the anchor point z and average offense distance a are unknown, but the form of the distribution P is known.

4、 The offender has only one residence which is unchanged.

3.1 A Mathematical Approach

In order to looking for an appropriate model, we start with the simplest possible situation. Since we know nothing about the offender, we assume that our offender chooses potential locations to commit crimes randomly according to some unknown probability density function P. For any geographic region R, the probability that our offender will choose a crime location in R can be found by adding up the values of P in R, giving us the probability

$$\iint_R P(\mathrm{x})\, d(x)^1 d(x)^2$$

Upon what sorts of variables should the probability density $P(x)$ depend?

The fundamental assumption of geographic profiling is that the choice of an offender's target locations is influenced by the location of the offender's anchor point z. Therefore, we first assume that P depends upon z. (Provided that the offender has a single anchor point and that it is stable during the crime series.) A second important factor is the distance the offender is willing to travel to commit a crime from their anchor point.(Different offender's have different levels of mobility- an offender will need to travel farther to commit some types of crimes than others) Let a denote the average distance that the offender is willing to travel to offend. So a varies between offenders and crime types[4].

Let us suppose that that the values of the anchor point z and average offense distance a are unknown, but the form of the distribution P is known. Then the problem can be stated mathematically as:

Given a sample $x_1, \cdots, x_n$ (the crime sites) from the distribution $P(x|z,a)$ with parameters z and a to determine the best way to estimate the parameter z (the anchor point).

One approach of this mathematical problem is the theory of maximum likelihood. First, construct the likelihood function[4]:

$$L(\mathrm{y}, a) = \prod_{i=1}^{n} P(x_i | y, a) = P(x_1 | y, a) \cdots P(x_n | y, a)$$

In order to get the best choice of z, make the likelihood as large as possible by maximizing the log-likelihood:

$$\lambda(\mathrm{y}, a) = \sum_{i=1}^{n} lnP(x_i | y, a) = ln P(x_1 | y, a) + \cdots + ln P(x_n | y, a)$$

This approach is rigorous; however, it is unsuitable as simple point estimates for the offender's anchor point are not operationally useful. Therefore, we continue our analysis by using Bayes Theorem[4].

Bayes Theorem then implies:

$$P(z,a|x)=\frac{P(z,a|x)\pi(z,a)}{p(x)} \quad (1)$$

Here $\pi(z,a)$ is the prior distribution.

It represents our knowledge of the probability density for the anchor point z and the average offense distance a before we incorporate information about the crime.

If we assume that the choice of anchor point is independent of the average offense distance, we can write:

$$\pi(z,a)=H(z)\pi(a) \quad (2)$$

$H(z)$ is the prior distribution of anchor points, and $\pi(a)$ is the prior distribution of average offense distances.

We will assume that the offender's choices of crime sites are mutually independent, so that

$$P(x_1,\cdots,x_n|z,a)=P(x_1|z,a)\cdots P(x_n|z,a) \quad (3)$$

Suppose that an unknown offender has committed crimes at $x_1,\cdots,x_n$, and that

The offender has a unique stable anchor point z.

The offender chooses targets to offend according to the probability density $P(x|z,a)$ where a is the average distance the offender is willing to travel.

The target locations in the series are chosen independently.

The prior distribution of anchor points is $H(z)$, the prior distribution of the average offense distance is $\pi(a)$ and these are independent of one another.

Then the probability density that the offender has anchor point at the location z satisfies

$$P(z|x_1,\cdots,x_n)\propto\int P(x_1|z,a)\cdots P(x_n|z,a)\ H(z)\pi(a)da \quad (4)$$

3.2 Simple Models for Offender Behavior

What's more, we need to be able to construct reasonable choices for our model of offender behavior, if our fundamental mathematical result is to have any practical or investigative value.

One simple model is to assume that the offender chooses a target location based only on the Euclidean distance from the offender's anchor point to the offense location and that this distribution is normal. In this case we obtain

$$P(x|z,a) = \frac{1}{4a^2}\exp\left(-\frac{\pi}{4a^2}(x-z)^2\right) \tag{5}$$

If we make the prior assumptions that the average offense distance and the anchor point are unchanged, and the offender commits n crimes at the crime site locations $x_1, \cdots, x_n$, then

$$P(z|x_1, \cdots, x_n) = (\frac{1}{4a^2})^n \exp\left(-\frac{\pi}{4a^2}\sum_{i=1}^{n}(x_i - z)^2\right)$$

We see that the anchor point probability distribution is just a product of normal distributions; the maximum likelihood estimate for the anchor point is simply the mean center of the crime site locations. We also mention that in this model of offender behavior, this is also the mode of the posterior anchor point probability distribution[4].

Another reasonable model for offender behavior is to assume that the offender still chooses a target location based only on the Euclidean distance from the offense location to the offender's anchor point, but now the distribution is a negative exponential so that

$$P(x|z,a) = \frac{2}{\pi a^2}\exp\left(-\frac{2}{a}|x-z|\right) \tag{6}$$

Once again, because of our prior assumptions that the average offense distance and the anchor point are unchanged, and the offender commits n crimes at the crime site locations $x_1, \cdots, x_n$, then we have

$$P(z|x_1, \cdots, x_n) = (\frac{2}{\pi a^2})^n \exp\left(-\frac{2}{a}\sum_{i=1}^{n}|x_i - z|\right)$$

We see that this is just a product of negative exponentials centered at each crime site. Further, the corresponding maximum likelihood estimate for the offender's anchor point is simply the center of minimum distance for the crime series locations[4].

This preceding analysis was predicated on the prior assumptions that the average offense distance and it is known in advance. Similarly, the existing methods mentioned in this paper all rely on decay functions f with one or more parameters which also need to be determined in advance. Unlike those methods, our method does not require that we make a choice for the parameter in advance.

3.3 Realistic Models for Offender Behavior

What would a more realistic model for offender behavior look like?

Consider a model in the form：

$$P(x|z,a) = D(d(x,z),a)G(x)N(z) \tag{7}$$

▲ D models the effect of distance decay using the distance metric d(x, z)

1、We can specify a normal decay, so that

$$D(d,a) = \frac{1}{4a^2}\exp\left(-\frac{\pi}{4a^2}d^2\right)$$

2、We can specify a negative exponential decay, so that

$$D(d,a) = \frac{2}{\pi a^2}\exp\left(-\frac{2}{a}d\right)$$

Any choice can be made for the distance metric (Euclidean, Manhattan, et.al)

▲ G models the geographic features that influence crime site selection

High values for G(x) indicate that x is a likely target for typical offenders;

Low values for G(x) indicate that x is a less likely target

▲ N is a normalization factor, required to ensure that P is a probability distribution

$$N(z) = \frac{1}{\iint D(d(y,z),a)G(\mathrm{y})d(y)^1\, d(y)^2}$$

N is completely determined by the choices for D and G.

G models the geographic features that influence crime site selection, with high values indicating the location was more likely to be targeted by an offender.

Then, how can we calculate G?

Use available geographic and demographic data and the correlations between crime rates and these variables that have already been published to construct an appropriate choice for G(x).

Different crime types have different etiologies; in particular their relationship to the local geographic and demographic backcloth depends strongly on the particular type of crime. This would limit the method to only those crimes where this relationship has been well studied.

Some crimes can only occur at certain, well-known locations, which are known to law enforcement For example, gas station robberies, ATM robberies, bank robberies; liquor store robberies .This does not apply to all crime types- e.g. street robberies, vehicle thefts.

We can assume that historical crime patterns are good predictors of the likelihood that a particular location will be the site of a crime.

Suppose that historical crimes have occurred at the locations c_1, c_2. . . c_n.

Choose a kernel density function

$K(y|\lambda)$

λ is the bandwidth of the kernel density function.

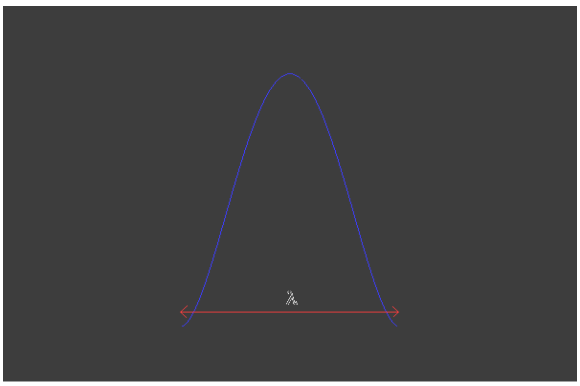

Figure 1

Calculate

$$G(\mathrm{x}) = \sum_{\mathrm{i}=1}^{N} K(\mathrm{x}\text{-}\mathrm{c}_i | \lambda) \tag{8}$$

The bandwidth λ can be e.g. the mean nearest neighbor distance.

We have assumed:

Each offender has a unique, well-defined anchor point that is stable throughout the crime series

The function H(z) represents our prior knowledge of the distribution of anchor points before we incorporate information about the crime series.

Suppose that anchor points are residences- can we estimate H(z)?

● Population density information is available from the U.S. Census at the block level, sorted by age, sex, and race/ethnic group.

1、We can use available demographic information about the offender

2、Set $H(z) = \sum_{i=1}^{N_{blocks}} = p_i K(z - q_i | \sqrt{A_i})$

3、Here block i has population pi, center q_i, and area A_i.

● Distribution of residences of past offenders can be used.

Calculate $H(z)$ using the same techniques used to calculate G(x).

3.4 Future Offense Prediction

Given a series of crimes at the locations $x_1, \cdots, x_n$ committed by a single serial offender, estimate the probability density $P(x_{next} | x_1, \cdots, x_n)$, that X_{next} will be the location of the next offense[4]. The Bayesian approach to this problem is to calculate the posterior predictive distribution:

$$P(x_{next} | x_1, \cdots, x_n) = \iiint P(x_{next} | z, a) P(z, a | x_1, \cdots, x_n) d(z)^1 d(z)^2 d(a)$$

We can use the method above to simplify, and so obtain the expression:

$$P(x_{next} | x_1, \cdots, x_n) \propto \iiint P(x_{next} | z, a) P(z, a | x_1, \cdots, x_n) H(z) \pi(a) d(z)^1 d(z)^2 d(a)$$

This approach makes the same independence assumptions about offender behavior as our fundamental result.

4. Model Two

Symbols:

symbols	Meaning
w	Weight
E	Total score
r_{B1} , $\cdots r_{B10}$	Each score
m_{B1} , $\cdots m_{B10}$	Various factors
P	The probability of success
U	The utility from crimes

Assumption:

The selection for crime sites depends on ten factors that demonstrated in Table 5 in the paper.

4.1 Analysis of models

The selection for crime sites depends on various factors instead of only one factor that the distance from the anchor point to the crime site. Therefore, Combine a large number of previous studies and related papers with our own thinking, we have summed up 10 key factors and a grading system about how to make the decisions on where the criminal locations will be (see Table 5), which are listed as follows: the responding speed of the police, public security situation, resistances' diathesis, density of registered inhabitants, the advantageous position, the number of offenses, the distance from the anchor point of the offender, the time required for committing the crime, number of target persons, offender's mental satisfaction from the crime. Classify these 10 index factors into two categories P and U, based on the actual situation. Let the former six factors belong to P and the latter four belong to U.

Next, we use Analytic hierarchy process (AHP) to get weighted factors (Table 4). Then we give the proper scores for each small area according to the actual condition. (Table 5). (There are some data which is difficult for us to achieve but they may be easily achieved by the local police .) Finally, we can work out the total score of E:

$$E = w_{P1} \times r_{P1} + \cdots + w_{P6} \times r_{P6} + w_{U1} \times r_{U1} + \cdots + w_{U4} \times r_{U4}$$

4.2 Use AHP to get the weighted factors

Then we can get a comparison matrix as follows:

The probability of success (P) is as important as the expected utility (U), so $E_{PU}=1$, $E_{UP}=1$

$$E = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Ui and the judgment matrix of U:

$$U = \begin{bmatrix} 1 & 7 & 5 & 2 \\ \frac{1}{7} & 1 & \frac{1}{3} & \frac{1}{7} \\ \frac{1}{5} & 3 & 1 & \frac{1}{5} \\ \frac{1}{2} & 7 & 5 & 1 \end{bmatrix}$$

Pi and the judgment matrix of P:

$$P = \begin{bmatrix} 1 & 3 & 7 & 5 & 4 & 6 \\ \frac{1}{3} & 1 & 6 & 4 & 2 & 3 \\ \frac{1}{7} & \frac{1}{6} & 1 & \frac{1}{2} & \frac{1}{7} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{4} & 2 & 1 & \frac{1}{6} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{2} & 7 & 6 & 1 & 3 \\ \frac{1}{6} & \frac{1}{3} & 5 & 5 & \frac{1}{3} & 1 \end{bmatrix}$$

According to all levels of comparison matrix, let's obtain its maximum of eigenvalue, coincidence indicator and the weight corresponding to all factors by using Matlab. The results are shown in the table below:

comparison matrix	Maximum eigenvalue $\lambda_{\max}$	coincidence indicator C_I	the weight of evaluation indexes w
E	$\lambda_{\max}=2$	0	[0.5000; 0.5000]
U	$\lambda_{\max}=4.1341$	0.0447	[0.4971;0.0495;0.1016;0.3519]
P	$\lambda_{\max} = 6.5633$	0.1127	[0.4319;0.2105;0.0298; 0.0448;0.1798;0.1033]

Table 1

The coincidence indicator can be computed by the formula

$$C_I = \frac{\lambda_{\max} - n}{n-1}$$

The results are shown in the form below. Find the average random coincidence indicator R_I. The average random coincidence indicator (exponent number is within 15) is as follows:

exponent number	1	2	3	4	5	6	7	8
R_I	0	0	0.52	0.89	1.12	1.26	1.36	1.41
exponent number	9	10	11	12	13	14	15	
R_I	1.46	1.49	1.52	1.54	1.56	1.58	1.59	

Table 2

Based on $C_R = \frac{C_I}{R_I}$, we can work out coincidence ratios:

C_{Rw}	C_{RU}	C_{RP}
0	0.0502	0.0894

Table 3

C_{Rw}、C_{RU}、C_{RP} are all less than 0.1 within the scope of consistency, which shows that the index factors we set meet the requirements, that is, they have some credibility within the error range.

Combination weight for all levels. Calculate the weight of the index factors *Ui*, *Pi*.

For example, w_{U1} =0.4971×0.5000=0.24855

In the same way, we can work out w_{Pi} (i=1, 2, …6) ,and w_{Ui} (i=1, 2, …4).

See Table 4:

w_{P1}	w_{P2}	w_{P3}	w_{P4}	w_{P5}	w_{P6}
0.21595	0.10525	0.0149	0.0224	0.0899	0.05165
w_{U1}	w_{U2}	w_{U3}	w_{U4}		
0.24855	0.02475	0.0508	0.17595		

Table 4

4.3 Scoring system for factors

The probability of success	The responding speed of the police	$r_{P1}=\begin{cases}1 & m_{P1}=0\\ 1-\frac{m_{P1}}{10} & 0<m_{P1}<10\\ 0 & m_{P1}=10\end{cases}$	m_{P1} stands for the responding speed of the police. 10 represents very fast, 6 represents fast,3 represents slow, 0 represents very slow.
	public security situation	$r_{P2}=\begin{cases}1 & m_{P2}\geq 50\\ 1-\frac{m_{P2}}{50} & 0<m_{P2}<50\\ 0 & m_{P2}=0\end{cases}$	m_{P2} stands for the region's monthly criminal records. If criminal record is 0, then the public security situation is good; if Criminal record is larger than 50, then the public security situation is bad.
	Resistances' diathesis	$r_{P3}=\begin{cases}1 & m_{P3}=0\\ 1-\frac{m_{P3}}{10} & 0<m_{P3}<10\\ 0 & m_{P3}=10\end{cases}$	m_{P3} stands for resistances' diathesis,10 stands for very good, 0 stands for very bad.
	density of registered inhabitants	$r_{P4}=\begin{cases}1 & m_{P4}=10\\ \frac{m_{P4}}{10} & 0<m_{P4}<10\\ 0 & m_{P4}=0\end{cases}$	m_{P4} stands for density of registered inhabitants.10 for the highest, 6 for the general, 3 for few people in that area and 0 for no one there.
	the advantageous position	$r_{P5}=\begin{cases}1 & m_{P5}=10\\ \frac{m_{P5}}{10} & 0<m_{P5}<10\\ 0 & m_{P5}=0\end{cases}$	m_{P5} stands for the extent of favorableness. 10 for the quite favorable, 6 for general, 3 for not going well, 0 for not executable.
	The number of crimes	$r_{P6}=\begin{cases}1 & m_{P6}\geq 3\\ \frac{m_{P6}}{3} & 0<m_{P6}<3\\ 0 & m_{P6}=0\end{cases}$	m_{P6} stands for the number of crimes, greater than 3 times, for the very skilled and unskilled times for 0

Utility from crimes	the distance from the anchor point of the offender	$r_{u1}=\begin{cases}1 & m_{u1}=10\\ \frac{m_{u1}}{10} & 0<m_{u1}<10\\ 0 & m_{u1}=0\end{cases}$	m_{u1} stands for the distance from the anchor point of the offender. Advantageous distance for 10, disadvantageous distance for 0.
	The time required for committing the crimes	$r_{u2}=\begin{cases}1 & m_{u2}=10\\ \frac{m_{u2}}{10} & 0<m_{u2}<10\\ 0 & m_{u2}=0\end{cases}$	m_{u2} stands for the time required for committing the crimes. Long time for 0, less longer time for 3, normal long for 6, short time for 8, very short time for 10
	Number of target persons	$r_{u3}=\begin{cases}1 & m_{u3}=10\\ \frac{m_{u3}}{10} & 0<m_{u3}<10\\ 0 & m_{u3}=0\end{cases}$	m_{u3} stands for the intensity of the target population. Very intensive for 10, intensive for 8, generally for 6, less-intensive 3, no goals for 0
	Offender's mental satisfaction	$r_{u4}=\begin{cases}1 & m_{u4}=10\\ \frac{m_{u4}}{10} & 0<m_{u4}<10\\ 0 & m_{u4}=0\end{cases}$	m_{u4} stands for the offender's mental satisfaction in that criminal location. Very satisfied for 10, satisfied for 8, general for 6, not very satisfied for 3,not satisfied at all for 0

Table 5

4.4 Results

We can get a wide range of criminal area and divide it into several small regions. Next, we give the proper scores for each small region according to the actual condition. Finally, using the formula

$$E = w_{P1}\times r_{P1}+\cdots+w_{P6}\times r_{P6}+w_{U1}\times r_{U1}+\cdots+w_{U4}\times r_{U4}$$

to calculate the E of each area.

Sites with higher scores have high probability to be crime areas. The police should enhance the patrols in these areas, and people should strengthen the awareness of safety.

5. The synthesized method

Up until now, we have already given two different models to show the geographic profile of next possible crime site. But both these two models have advantages and disadvantages. Model One allows us to find possible locations of the next crime as well as the offender's resistance based on only the time and locations of the past crime scenes, but it ignores some critical factors related to the crime, such as the geographical environment and the motive for the crime and other factors. Model Two takes 10 factors which affects the selection of crime sites into consideration and qualify them to work out the geographic profile, but it is too subjective and easily leads to deflection when put into practice. Because of their different emphases, the two models will not give the same geographical profile. Assume we work out n geographical profiles through Model One and m geographical profiles through Model Two. Apparently these geographical profiles will not be the same. Therefore, we come up with the following new method which can combine the two models together.

Symbols:

symbols	meaning
U	Total income
s	the number of crimes
I	income for the offender
P（x）	the probability of success for each crime
E	expected utility

Hypothesis：

1、we assume that the net income in of each crime occurred in the same region is unchanged, that is, for any $i, j,$

2、we assume in the same area, each time the probability of successfully committing a crime is also unchanged, i.e. $P_i^1 = P_j^1, P_i^2 = P_j^2$

With the method of dynamic programming, two spatial variables, the expected utility and the probability of success for each offense, are used to model the criminal's location choices. A criminal usually commits his first offence in the district which has the highest probability of success but a lowest expected utility. If an area has both higher expected utility and a higher probability of success, the criminal will commit all his offences in this place. The model also suggests that crime prevention measures should be adopted in the local conditions. "Covering" measures, such as patrolling, should be taken in the poor residential districts or delinquency districts, while more sophisticated and advanced measures should be introduced in the richer districts or the districts where career criminals haunt.

5.1 The Foundation of Model

First, let the utility function for a certain criminal be given:

$$U = U_1(I_1) + U_2(I_2) + \cdots + U_S(I_S) \tag{8}$$

In this function, s indicates the number of crimes; Ii (i=1,2,…s) indicates the net income for the ith time offender.

Then, denote Pi as the probability of success for each crime (i.e., the probability of not being caught). If the criminal is caught when he commits the hth crime, he will lose I_h. Then we can get the total net income

$$I_1 + I_2 + \cdots + I_{h-1}$$

Meanwhile, we get the expected utility when the cumulative number of crime reaches s.

$$\begin{aligned} E &= U_1 P_1(1-P_2) + (U_1+U_2)P_1P_2(1-P_3) + \cdots + (U_1+U_2+\cdots+U_s)P_1P_2\cdots P_s \\ &= \sum_{h=1}^{s}(\sum_{i=1}^{h}U_i)\prod_{i=1}^{h}P_i(1-P_{h+1}) \end{aligned} \tag{9}$$

For here $P_{s+1} \equiv 0$ so we can change Function (9) into

$$E = U_1\sum_{h=1}^{s}\prod_{i=1}^{h}P_i(1-P_{h+1}) + U_2\sum_{h=2}^{s}\prod_{i=1}^{h}P_i(1-P_{h+1}) + \cdots + U_s\prod_{i=1}^{s}P_i \tag{10}$$

For:

$$1 - P_1 + \sum_{h=1}^{s}\prod_{i=1}^{h}P_i(1-P_{h+1}) \equiv 1$$

So in Equation (10), the coefficient of U_1 can be abbreviated as P_1, and the coefficient of U_2 can be simplified as P_1P_2, the coefficient of U_h can be simplified as $\prod_{i=1}^{h}P_i$

In this way, Equation (9) can be further abbreviated as:

$$E = U_1P_1 + U_2P_1P_2 + \cdots + U_sP_1P_2\cdots P_s \equiv \sum_{h=1}^{s}(U_h)\prod_{i=1}^{h}P_i \tag{11}$$

Now assume that there are two potential criminal regions in a city, which are labeled by superscripts 1 and 2. For the sake of simplicity, further assume that the net income in of each crime occurred in the same region is unchanged, that is, for any i, j,

$$I_i^1 = I_j^1, I_i^2 = I_j^2$$

But I^1 does not necessarily equal I^2

Meanwhile, assume in the same area, each time the probability of successfully committing a crime is also unchanged, i.e. $P_i^1 = P_j^1, P_i^2 = P_j^2$

Therefore, if the offender intends to commit s crimes, he needs to make location choices for 2^s times.

The following dynamic programming method reveals the optimal solution of the offender's location choice of committing the crime. Here assume that during the planning period, the offender implement all his crimes. What's more, he determines all the optimal locations one by one from the back to the front. The planning period can be any time that a crime will happen.

Suppose the offender has already determined the locations of the crimes for the first $s-1$ times and attempts to optimize the location of *sth* crime.

Denote the expected utility for committing *sth* crime in district 1 and district 2 as

$$E^1 = U_1 P_1 + U_1 P_1 P_2 + \cdots + U_s^1 P_1 P_2 \cdots P_s^1 \tag{12}$$

$$E^1 = U_1 P_1 + U_2 P_1 P_2 + \cdots + U_s^2 P_1 P_2 \cdots P_s^2 \tag{13}$$

The first s-1 terms on the right side in Equation (12) and (12) has no region label, because of the assumption of they are the same in the two equations mentioned above and the only difference exits in the sth term. Therefore, the difference between the expecting utilities of committing sth crime in district1 and district 2 is

$$\Delta E = E^1 - E^2 = (U_s^1 P_1 - U_s^2 P_2)\prod_{i=1}^{s-1} P_i \tag{14}$$

Equation (14) indicates that whatever the probability of success P or the probability of being caught (1-P) is, the offender will choose the place which has the highest expected profit to commit his last crime. When the utility function is monotone, and the offender does not care about his location of crime activity, the results mentioned above also suggest that the offender will commit sth crime in the location where the expected utility is highest[3].

Generally, if the offender has already determined the crime locations of the first time, the second time, the (k-1)th time, the (k+1)th time, …the sth time, the expected utility of committing the kth crime in district1 and district2 are as follows:

$$E^1 = \sum_{h=1}^{k-1} U_h \prod_{i=1}^{h} P_i + U_k^1 P^1 \prod_{i=1}^{k-1} P_i + P^1 \sum_{h=k+1}^{s} U_h \prod_{\substack{i=1 \\ i\neq k}}^{h} P_i \tag{15}$$

$$E^2 = \sum_{h=1}^{k-1} U_h \prod_{i=1}^{h} P_i + U_k^2 P^2 \prod_{i=1}^{k-1} P_i + P^2 \sum_{h=k+1}^{s} U_h \prod_{\substack{i=1 \\ i\neq k}}^{h} P_i \tag{16}$$

The difference between Equation (15) and Equation (16) is

$$\Delta E = E^1 - E^2 = (U_k^1 P_1 - U_k^2 P_2)\prod_{i=1}^{h-1} P_i + (P^1 - P^2)\sum_{h=k+1}^{s} U_h \prod_{\substack{i=1 \\ i\neq k}}^{h} P_i \qquad (17)$$

or

$$E = \prod_{i=1}^{k-1} P_i \left[U_k^1 P^2 - U_k^2 P^2 + (P_1 - P_2)\sum_{h=k+1}^{s} U_h \prod_{i=k+1}^{h} P_i \right] \qquad (17')$$

When $k=s$, we can change Equation (17') into Equation (14). Here k is the offender's last crime during his planning period. The offender will commit this crime in the district which has the highest expected utility. Similarly, k can be replaced by $s-1, s-2, \cdots, 1$. By repeating the calculation, we can track the criminal's optimal locations of committing the crimes each time[6].

Equation (17') shows that there are several possible locations for the criminal to select. When $U_k^1 P^1 > U_k^2 P^2$, $I^1 P^1 > I^2 P^2$. If $P^1 > P^2$, the value of equation (17') is positive, which means the region with the highest expected utility, is also the most secure region. In this case, region 1 is a perfect location of crime, so the offender will commit all the crimes in this region.

For an offender who constantly changing places for committing crimes between the two regions, there is no ideal region with both high E and P. If $U^1 P^1 > U^2 P^2$ or $I^1 P^1 > I^2 P^2$, then $P^1 < P^2$. It means that areas with the highest expected profit, is also the highest risk areas, the probability of being captured (l-P) reaches its maximum in the same time. From the Equation (17') and (14) we can see, in this case, the last crime the offender planned will occur in region 1. When the offender reduces the number of crimes to s times or less, the coefficient ($(P^1 - P^2)$) of the equation (l7') will increase. The larger the coefficient is, the fewer the certain number of actually committed crimes will be as long as s is certain. Therefore, the fewer the number of actually committed crimes is, the greater the second term's absolute value in the second square brackets in equation (17') will be. Because of $(P^1 - P^2) < 0$, this term should be negative. So probably there exists a number of crimes that makes the whole equation (17') be negative. Then region 1 is not perfect for committing crimes, so the offender will change his criminal location to region 2. From then on, as the coefficient of equation (17'), i.e. $(P^1 - P^2)$ is increasing, the offender will choose Region 2 as criminal location instead of Region 1. In other word, If an offender is faced with this situation, in order to change his criminal places, the best option is to commit his initial crime in Region 2 at a lower expected return and a lower the probability of being arrested, and then go to commit the crimes in zone 1 where both the benefits and risks are higher.

5.2 Choice of multi-regional areas

We have already discussed the situation of making choices of criminal locations when there exist only two regions. However, it is very likely that there are many regions in one area. The distribution of their criminal scenes is very similar to the situation with the two regions.

Now suppose there are n potential criminal locations and the offender has determined all these areas except the *k*th criminal location. The difference between expected utility of region c and region m for committing *k*th crime is as follows:

$$\Delta E = \prod_{i=1}^{k-1} P_1 (U_k^c P^c - U_k^m P^m) \tag{18}$$

Therefore, the offender will commit the crime in the region which has the highest $U^k P$. This result works out the same as the situation of two-region area[8].

Equation (18) shows that if $U_k^n P^n > U_k^m P^m$, $P^n > P^m$, Region c is better for the offender to commit crimes than region m in all respects, so the offender will exclude region m from its location decisions. Therefore, for *i*th crime (*i* is arbitrary), the criminal location is decided by the following sequences

$$U_i^1 P^1 < U_i^2 P^2 < \cdots < U^n P^n \tag{19}$$

$$P^1 > P^2 > \cdots > P^n \tag{20}$$

In our real life, these two kinds of arrangement rarely exist at the same time; it is only a theoretical solution. Regions in Sequence (19) and (20) are all criminal locations. If the offender commits crimes in all those regions to, then there are some offsets among the various regions.

If an area is foolproof, then he will focus on here to implement all the criminal activities. In reality, criminals always tend to be concentrated in a few areas of crime. Therefore, suppose n is small, then there are large chances of Sequence (19) and (20).

According to Sequence (19) and (20), the offender will commit the last crime in region *n*.

Now let's find out where the (*k*-1)th crime will be committed. First, the regions can be arranged according to the following ratio

$$\frac{U^n P_{k-1} - U_{k-1}^1 P^1}{P^1 - P^n} \equiv W \quad (k = 1, 2, \cdots, n) \tag{21}$$

When (18) is negative, the offender will leave region n for another region which has the smallest value of *W*. W is the marginal expected return with the risk of committing the crime. The offender optimizes his criminal acts by selecting the area with smallest W value, such that the expected profit will reach its maximum. Other crime locations $k-2, k-1, \cdots, 1$ can also be determined by the order of the ratio of *W*.

The result turns out to be the same as the two-region situation. No matter the offender is faced up with the situation of a two-region area or a multi-region area; he will optimize his choices of the criminal locations and commit the serial crimes according to expected utility and the probability of success.

5.3 Solution and Result

With the method of dynamic programming, two spatial variables, the expected utility and the probability of success for each offense, are used to model the criminal's location choices. It shows a single offender's criminal acts in two regions and various districts, and conversions throughout the various regions. The results showed that: If a region has both high expected utility and high probability of success, then the offender will concentrate in this region committing all the crimes and will never go to other areas. Once the two regions has different expected utility and probability of success, the offender will change his location of committing crimes.

6. Application

In the case of Peter Sutcliffe, we use the Google Earth to find out the sites of 13 victims and 10 survivors. We also find out the anchor point. See Figure

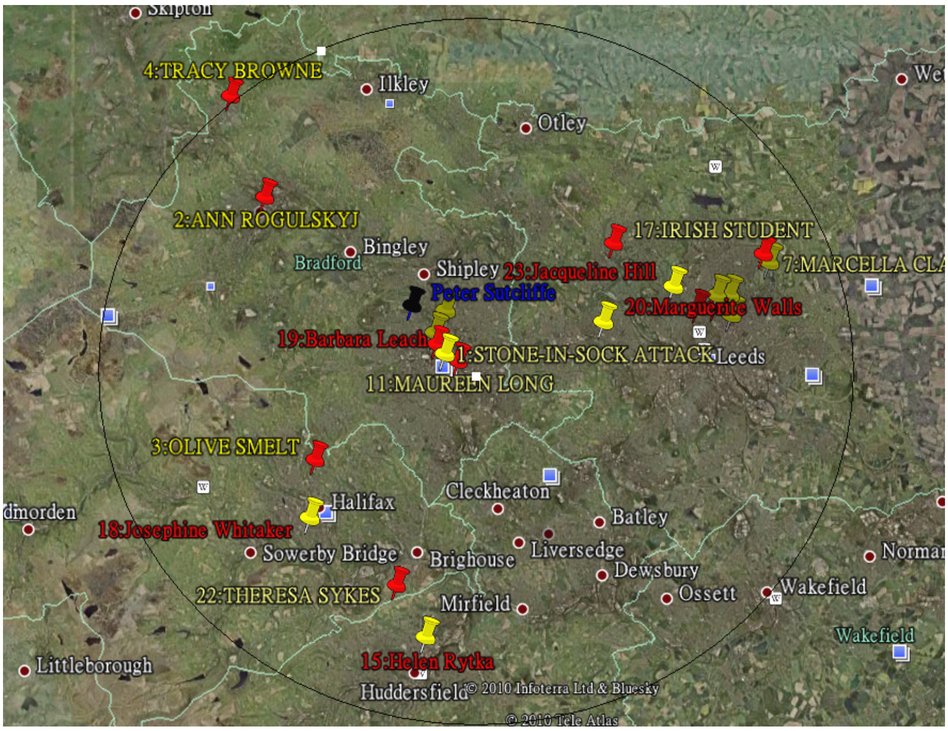

Figure 2

As we have seen, the offender's anchor point is like the mean center of the crime site locations. Then, according to the Model One, we can get the geographic profile. See figure 3：

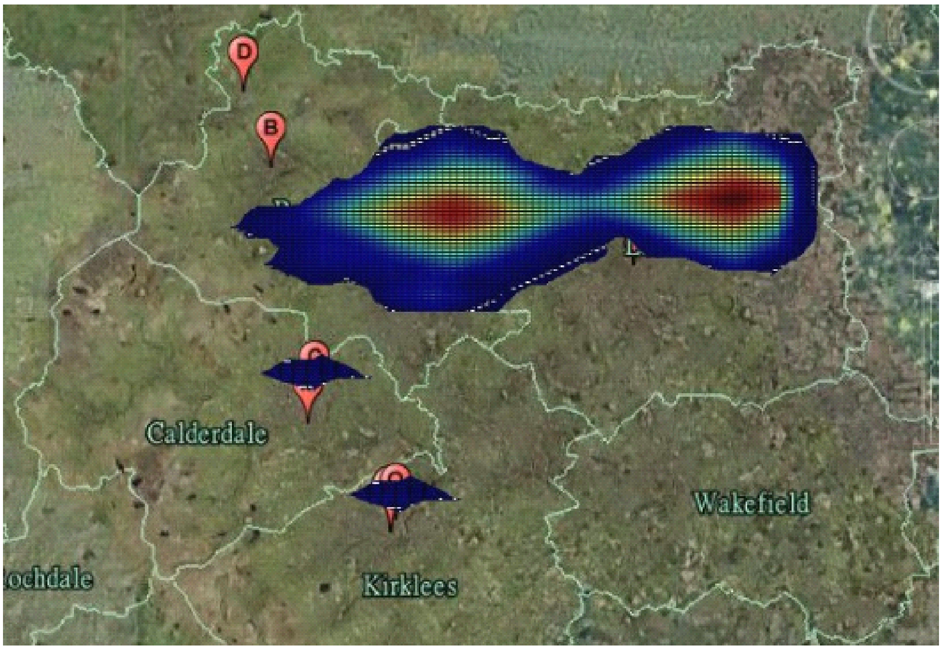

Figure 3

In Model Two, a wide range of criminal area can then be divided into several small areas. See figure 4：

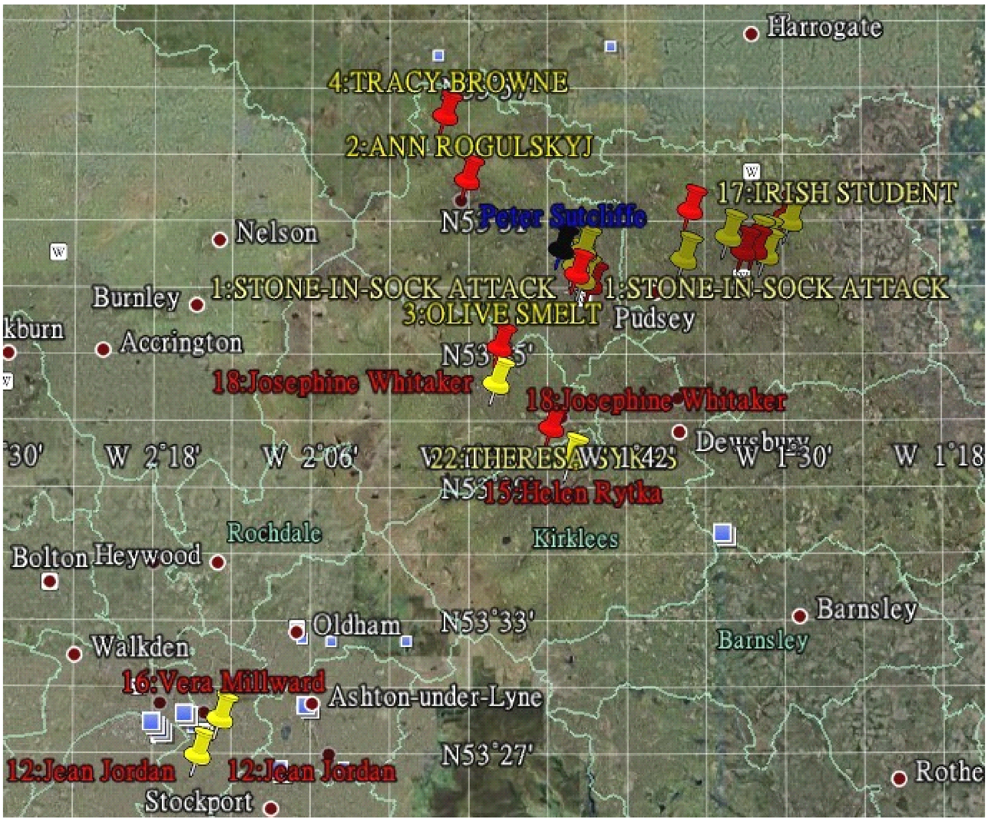

Figure 4

Next， we give the proper score for each factor according to each area’s actual situation. It’s a pity that we don’t have detailed information, so we can’t give the score .However, we believe it’s easy for local police .Now we assume that we have got the geographical profile and n hot points. Then, we use our synthesized method to analyze the n+2 hot points to predict the next crime. The detail of steps:

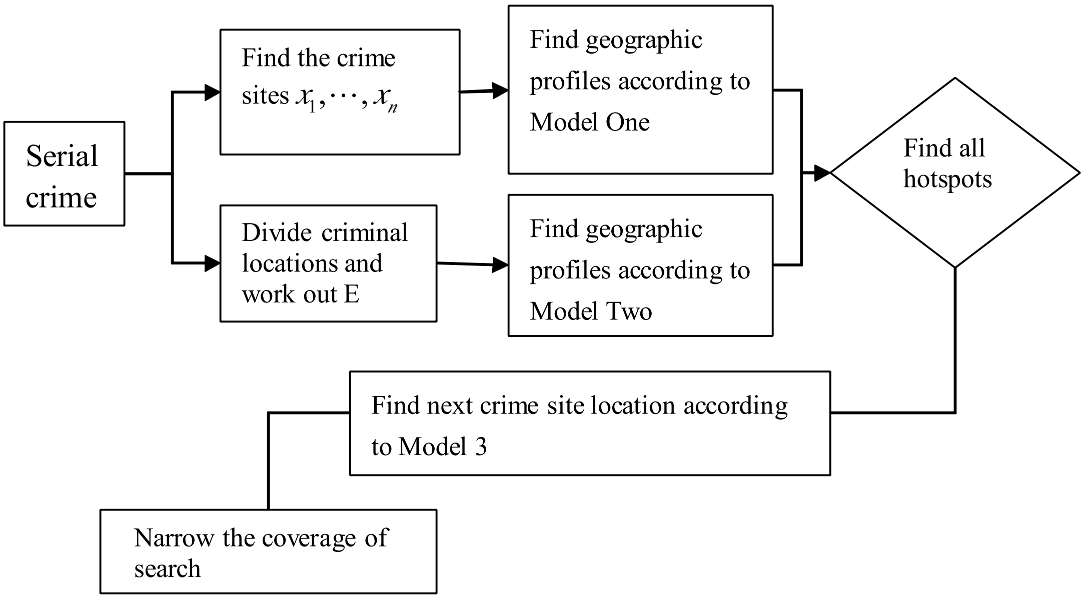

Figure 5

7. Evaluation and improvements

Model One:

Strength: Geographic profiles generated in Model One is based on the assumption and computed from the theoretical formulas, which makes it rigorous. What's more, it takes into account local geographic features, in particular, it account for geographic features that influence the selection of a crime site and geographic features that influence the potential anchor points of offenders.

Weakness: Since the offender is very likely to change his residence, the prerequisite that the offender has the only anchor point and it keeps stable differs from the actual situation. Put too much emphasis on theoretical formulas while lack of adequate thinking about practical factors, such as the geographical environment and the criminal motives and so on.

Model Two:

Strength: Model Two aims to make up for the application of the method of geographical profiling through using analytic hierarchy process (AHP). It takes full account of a variety of factors relevant to the crime site selection, quantify those factors and give the appropriate weighting factor to them in accordance with local conditions.

Weakness: Require large amounts of data which are difficult to obtain accurately. And it is difficult to make accurate ratings of various factors, which takes long time to work them out.

The synthesized method：

Strength: Formulize the factors, estimate the nest crime site location more rigorously.

Additionally, we can further estimate where the crime site location is according to the number of crimes by combining two geographical profiles generated respectively from Model One and Model Two.

Weakness: Repeatedly take the concepts of the expected utility of crime and the probability of the success into account, which may make the results close to that of Model Two which focus on these factors, so that the conclusions of Model One will be neglected.

The model is still an approximate on a large scale. This has doomed to limit the applications of it.

Further improvements:

After consolidating the previous two methods, a major improvement to the methodology is to score the factors more precisely and more objectively. It is best if there is a special program designed for analyzing data and data processing. In addition, we should find out more links between Model One and Model Two, to get a more comprehensive result with smaller deviation.

8. References

[1] Smith, C. and Guillen, T. 1991 The search for the Green River Killer, New York: Onyx.

[2] Beltrami, E. J. (1993). Mathematical models in the social and biological sciences. Jones and Bartlett Publishers.

[3] Canter, D., Coffey, T., Huntley, M., & Missen, C. (2000). Predictingserial killers' home base using a decision support system.Journal of Quantitative Criminology, 16(4), 457–478.

[4] O'Leary, M.(2009). The mathematics of geographic profiling. preprint.

[5] Rossmo, K. (2000). Geographic Profiling. CRC Press.

[6] George O. Mohler and Martin B.(2009).Short Geographic profiling from kinetic models of criminal behavior

[7] Janet Warren, Roland Reboussin, Robert R.(1998).Hazelwood,Crime Scene and Distance Correlates of Serial Rape . Journal of Quantitative Criminology.

[8] D. Kim Rossmo.(2003).A Methodological Model.

[9] Brantingham, P. L., & Brantingham, P. J. (1993). Nodes, paths and edges: considerations on the complexity of crime and the physical environment. Journal of Environmental Psychology,13, 3–28.

[10] Canter, D., & Larkin, P. (1993). The environmental range of serial rapists. Journal of Environmental Psychology, 13, 63-69.