

Dynamic Analysis for Geographical Profiling of Serial Cases Based on Bayesian-Time Series

Guanli Huang

Beijing Vocational College of Electronic Science, Beijing, China
huangguanli@dky.bjedu.cn

Guanhua Huang

School of Social Development and Public Police, Beijing Normal University, Beijing, China
cherryhenry@163.com

Abstract— The analysis of spatial information has long been considered valuable for police agency within the criminal investigative process. This is especially true for serial crime cases where criminologists and psychologist apply geographical profiling to model criminal mobility distribution and behavior patterns in order to estimate a criminal's likely residence. In recent years the availability of advanced computational mathematic tools ensured us to establish some mathematical models to replace traditional empirical method. However, as a new technology, current geographical profiling models are still fundamental and impractical. In this article, based on existing frameworks, we establish three new methodologies, namely, Bayesian—Factor analysis model, Time series analysis model and GIS(Geographic Information System)—Decay model, to study geographical profiling problems. Then, we test and compare their accuracy, efficiency, sensitivity and robust according to 11 historical serial crime samples and Monte Carlo simulations. Finally, we discuss the advantage and disadvantage of each model and provide executive guidelines about how to synthetically apply these models for real cases.

Index Terms— geographical profiling, time series analysis, bayesian analysis, geographic information system

I. INTRODUCTION

Can we tell where an offender lives from where he or she commits crimes? Can we predict where he or she will commit next crime? An excellent police officer may answer, 'Yes, but sometimes.' Indeed, some police officers can make correct judgments for some cases according to their experience and intuition. But hardly any of them can tell which one is their 'lucky case' or why their decisions are correct. That is why we need to develop a scientific and systemic methodology to instead our traditional empirical operations. Due to the

TABLE I.
CHOICES FOR THE DECAY FUNCTION

Choices for the decay function	
Linear	$f(r) = A + Br$;
Negative exponential	$f(r) = Ae^{-Br}$;
Normal	$f(r) = A(2\pi S^2)^{-1/2} \exp[-(r-\bar{r})^2 / 2S^2]$;
Lognormal	$f(r) = A(2\pi r^2 S^2)^{-1/2} \cdot \exp[-(\ln r - \ln \bar{r})^2 / 2S^2]$;
Truncated negative exponential	$f(r) = Br$ if $r < C$ and $f(r) = Ae^{-\beta r}$ if $r \geq C$.

development of computational mathematics, there are many recently developed frameworks about this issue. Most famous one of them is called geographical profiling.

A. The past theories

Geographic profiling is a criminal investigative methodology that analyzes the locations of a connected crime series to make a prediction about where the offender is most likely to live. It is primarily used as a suspect prioritization and information management tool [1]. The main process for geographic profiling is using crime scene information (e.g. the locations of crime) to infer personal characteristics of the responsible criminal (e.g. the offender's residence). Recently, there is academic debate over the appropriate method for geographical profiling. Some researchers arguing for simple geometric and statistical methods [2][3], whereas others for complex computer-based algorithms[4].

The theory and conceptual framework of geographical profile is built upon environmental criminology [5]. Environmental criminology involves a number of theoretical concepts, such as Routine Activity, Crime Pattern and Rational Choice.

The Routine Activity theory [6], suggests that crime depends on the intersection of three elements (offender, target, and environment). In detail, when an individual

The 2010 Scientific Fund of Beijing Education Commission (ITEM NO. KM201000002002)China.

who want to offend in an environment that is appropriate for criminal activity (e.g. in the absence of surveillance) encounters a suitable target, crime will occur. Geographical profiling is based on the premise that crime will happen only in the places that during criminal routine activities. Therefore the area around an offender's home is a most likely place to commit crime.

The second theory, Crime Pattern Theory [5], recommends that crime will occur in places where an offender's awareness space intersects with perceived suitable targets. So just like the routine activity theory, criminal routine activities are also important in determining the locations of crime in this theory. So, the two theories provide the theoretical basis for geographical profiling.

According to Rational Choice Theory [7], the decision to commit crime is the outcome of a decision-making process which weighs the expected cost, efforts, and rewards. If the costs and efforts low enough and the rewards are high enough, crime will occur. The key element of this theory for geographical profiling is that locations and targets that are easily accessible to the offender will be perceived as concerning few costs, little effort and greater rewards. So it is the locations around the home predicted to be the areas in which criminals are likely to offend. But the potential costs of crime around the home location can be high (because of the risk of being recognized) and traveling long distances may gain great rewards. In sum, the area around the home location is by no means the location a criminal will choose to offend.

Despite the argument on geographical profiling method, it is simple and undisputed that the basic assumptions underlying the process of geographical profiling. There are four major theoretical and methodological assumptions required for geographic profiling [1]: Left- and right-justify your columns. Use tables and figures to adjust column length. On the last page of your paper, adjust the lengths of the columns so that they are equal. Use automatic hyphenation and check spelling. Digitize or paste down figures. The case involves a series of at least five crimes, committed by the same offender. The series should be relatively complete, and any missing crimes should not be spatially biased (such as might occur with a non-reporting police jurisdiction). The offender has a single stable anchor point over the period of the crimes. The offender is using an appropriate hunting method. The target backcloth is reasonably uniform.

However, some of those assumptions are not necessary in our model which combines different schemes that are suitable for various situations (e.g. a series of four crimes, the offender has two bases, the criminal changes his hunting strategy during serial crimes). In other words, the reduced level of accuracy will be controlled in the model when some information is missing.

B. Our Dynamic Analysis models

Established three different geographical profiling models: GIS—Decay model, Bayesian—Factor analysis model, and Time series analysis model. These three

models are respectively accomplished in dealing 'short' (less than 6 offends), 'medium' (6-8), and 'long' (more than 8) serial crimes.

The first model is GIS—Decay model. Geographic Information System is professional software captures, measures and analyzes geographic features. With the help of GIS, we develop a travel metric function which includes travel distance, travel cost and time cost instead of Euclidean (straight line) or Manhattan (travel distance) metric that are widely used in former articles. This model has an obvious advantage for cases with small data set, but a relatively low efficiency than other two models. So this model can give useful information to police agency even at the beginning of a serial crime.

The second model is Bayesian—Factor analysis model. This method is based on Bayesian statistics that treat each crime location as independent random variables obey some certain distribution and then calculate the distribution of anchor point. The distribution of each crime location is decided by some factors that may influence the criminal's spatial decisions. This model has moderate accuracy and efficiency but a strong robust. This model is suitable for the case with medium data set and the case that we have some information about criminal's characteristic.

The third model is Time series analysis model. In this model, we treat crime locations as a correlated ordering sequence other than independent points. By deduce autoregressive model and solve Yule-Walker equations, we can not only find anchor point but also predict next latent crime site. The performance of this model improves significantly while data set getting large. And this model can also deal some complicated situations like 'multi anchor points' and 'nomadic criminal'.

The main purpose of our models is to aid in police agency's investigations of serial criminals with helpful prediction. The input to our model is the locations of crimes from a series and the output from our model will be some possible areas with different probability for the location of offender's home and next crime. And the local geographical information and the sequence of crimes will be used in our model. In our model, we use Geographic Information System (GIS) software to measure and calculate detailed geographical features of the map. This application guaranteed our model enjoy a better accuracy than other models simply on plane map. Also, most factors that may significantly influence spatial decision-making will be considered in our model. Father more; our models can work on some complicated situations such as: the criminal has two bases or the 'nomadic' criminal.

II. GIS-DECAY MODEL AND BAYESIAN-FACTOR ANALYSIS MODEL

A. Geographical Information System

when we solve geographical profiling problems, time is more than just money. We cannot wait the data set getting large enough for our model. That means more victims suffering from the criminal. We need a model that can work well with a small data set. So this means

sometimes a simple model works better than a too complicated one. Suppose helping the police agency investigating a new started serial crime. All the information we have so far is just a few crime locations (usually less than five), and some hypothesis of criminal's identity (gender, age, travel method, etc.). By this insufficient information, we may hard to give a precise forecasting for the location of anchor point and next crime site. However, still make full use of these resources and give some useful guidelines to help police agency to carry on searching.

Geographical Information System are cluster of tools used for collection, storage, management, processing, analysis, display and description of geographical data. They're highly efficient under the appropriate support of computer science. Subjects of GIS include spatial data, raster data, remote sensing data, etc. Development of GIS has greatly improved the methodology of geographic sciences, as well as provided chances of solving sophisticated problems on planning, decision making and managing. GIS can not only be applied on study and research, but more 'daily' issues as well, such as the case of criminology. In fact, GIS has proved its value in much larger scale recently. Some even use GIS during design of 500 m2 plaza and receive considerably convincing result. Test the efficiency and reliability of different schemes with the help of GIS. Actual application include following operation: Vectorization of spatial data; Display of geographic information; Comparison between historical sample and algorithm result.

B. Distance metric

One basic concept for geographical profiling problem is average distance (α) a criminal is willing afford to commit a crime. However, how to measure this distance depends on the choice of distance metric. Typically, there are many widely used choices for this metric including Euclidean metric and Manhattan metric. Both of these metrics treat the hunting area as a uniform and isotropic 2-dimensional plane, which means they ignore the detailed geographical features of the region. As we know, driving along the road is much faster than traveling through a farmland; and taking a ferry or train may cost extra travel expenses. According to psychological theory, most criminal will calculate the travel cost against his/her desire for the target. So the region with high travel cost may less likely suffer from this criminal. By GIS, we can set some travel cost functions about the geographical features. For example, in our models, we set different travel speed for different landforms; the specific settings

TABLE II.
AVERAGE TRAVEL SPEED IN DIFFERENT LANDFORMS

Landform	Average travel speed (km per hour)
Highway	100
Road	70
City area	40
Farmland	20
Mountain area	15
Forest area	10

are shown in the table I. Then we get a travel cost metric by separate the path between two points into small intervals, calculate the time cost for each intervals and integral them to get a total time cost. Then based on the average available time of a criminal, we can get the distribution of the average distance α on map.

Decay functions: When decided the distance α , we can describe the criminal's spatial decision making strategy by a decay function. The idea of this function is criminals are most likely commit crimes at the area around α , the possibility decreases while both increasing and reducing travel distance. Several typical choices for decay function are shown in table II.

Score functions: After making a choice of distance metric d and decay function f , we can construct a score function $S(y)$ by summing f for all available points $x_1 \dots x_n$

$$S(y) = \sum_1^n f(r(x_i, y)) = f(r(x_1, y)) + \dots + f(r(x_n, y)). \quad (1)$$

Areas which satisfy a high score are considered to be more likely to contain the offender's anchor point, vice versa.

C. Bayesian analysis

Recently, a new method using Bayesian analysis has been developed to solve geographic profiling problem. Firstly, let us consider the simplest (may also the worst) situation for a serial crime. Suppose the police agency know no information (gender, age, behavior habits, etc.) about the offender. All they know is this criminal has offended a serial crime, and the location of each criminal scene is available. Thus, we assume that this criminal commits each offend randomly according some particularly probability density $P(x)$. So, by statistics knowledge, we know the probability that the criminal offend a crime in a given area Ω can be calculated as $\iint_{\Omega} P(x)d\Omega$.

According to our basic assumptions, though there is no information about the criminal, we do know there are at least two factors may influence his/her target location choice. First one is the criminal's anchor point z . Another one is the average distance this criminal is willing to travel to commit a crime, this distance is denoted by α . For each given pair of z and α , the conditional probability density of x is $P(x|z, \alpha)$. By Bayes' Theorem, the conditional probability density of z for a given pair x and α can be presented as

$$P(z, \alpha | x) = \frac{P(x|z, \alpha)\pi(z, \alpha)}{P(x)}. \quad (2)$$

The term $P(x)$ is the marginal distribution which is independent with z and α . Since we do not care the absolute value of the probability density, we can ignore $P(x)$ term and rewrite (2) as

$$P(z, \alpha | x) \propto P(x|z, \alpha)\pi(z, \alpha). \quad (3)$$

$\pi(z, \alpha)$ is the factor function which characterizes the criminal's spatial decision making strategy.

So for a serial crime locations $x_1 \dots x_n$, the conditional probability density of z can be denoted as $P(z, \alpha | x_1 \dots x_n)$, then we can get the multivariate formation of (3)

$$P(z, \alpha | x_1, \dots, x_n) \propto P(x_1, \dots, x_n | z, \alpha) \pi(z, \alpha). \quad (4)$$

For the mathematical simplicity consideration, an assumption that all of the offence sites are independent is widely used. Then we get the decomposition

$$P(x_1, \dots, x_n | z, \alpha) = P(x_1 | z, \alpha) \dots P(x_n | z, \alpha). \quad (5)$$

So

$$P(z, \alpha | x_1, \dots, x_n) \propto P(x_1 | z, \alpha) \dots P(x_n | z, \alpha) \pi(z, \alpha). \quad (6)$$

Many articles still assume the independency between z and α . But in the following sections, we will illustrate this assumption is not only unnecessary but also incorrect in many cases.

At last, since we only care the location of the anchor point z , we can integral (6) and get the conditional probability density of z independent of α .

$$P(z | x_1, \dots, x_n) \propto \int P(x_1 | z, \alpha) \dots P(x_n | z, \alpha) \pi(z, \alpha) d\alpha. \quad (7)$$

The expression $P(z | x_1 \dots x_n)$ figures out the anchor point probability density of the criminal who have already committed crimes at the locations $x_1 \dots x_n$. Then naturally this probability density provides us a rigorous search area with high probability to find the anchor point of the offender.

D. Factor functions

Most researches on geographic profiling only consider the main factor of the distance between criminal's anchor point and crime location. At meantime they may ignore the factors that potentially influence the spatial decision making, Lundrigan et al. [8] pointed out that all criminal spatial decisions are mediated by social, economic, and cognitive factors. Those factors that may influence the location of the criminal's home include: the development of the series; the age, intellectual capability, employment status, marital status, and motive; the mode of transportation that they use; and the type of the crime. Snook [9] gave some brief explanations for the factors influence serial murderers' spatial decisions. So there are several factor that may influence criminal's spatial decision making. In this model, all the factors we take into consideration are shown in the Table III. A simple model of criminal's spatial decision making strategy can be described as a function (π) of his/her anchor point location (z) and average distance (α) he/she is willing afford to commit a crime, i.e. $\pi(z, \alpha)$. Also, this function should be influenced by all factors in the table. Thus we can get (8)

$$\pi(z, \alpha) = g(TC, MT, IC, AG, GE) \quad (8)$$

The form of this formula can be diversity since it is merely an empirical function. In our article, we just consider the simplest situation that all factors are

TABLE III.
CONSIDERING FACTORS

Factor	Denotation
Time coverage	<i>TC</i>
Mode of transportation	<i>MT</i>
Intellectual capability	<i>IC</i>
Age	<i>AG</i>
Gender	<i>GE</i>

independent with each other and obey normal distributions.

$$\pi(z, \alpha) = A(2\pi)^{-1/2} \exp(-\{(1 - \frac{\overline{TC}}{TC})^2 + (1 - \frac{\overline{MT}}{MT})^2 + (1 - \frac{\overline{IC}}{IC})^2 + (1 - \frac{\overline{AG}}{AG})^2 + (1 - \frac{\overline{GE}}{GE})^2\}) \quad (9)$$

All the factors have already been normalized before taken in to (9).

Another approach of factor function assumes that all these factors only influence the average criminal distance α , i.e.

$$\alpha = g(TC, MT, IC, AG, GE) \quad (10)$$

Since we usually assume α is a negative exponentially decay function, a possible form of α can be represented as

$$\alpha = A \exp(-Br(\frac{TC}{TC} \cdot \frac{MT}{MT} \cdot \frac{IC}{IC} \cdot \frac{AG}{AG} \cdot \frac{GE}{GE})) \quad (11)$$

The above discussion just showed two basic example of factor function. The form of factor function for a specific case still greatly depend on the experience of local police agency and former samples of that region.

III. TIME SERIES ANALYSIS MODEL

A. Introduction of time series analysis

In geographic profiling study, the available criminal locations can be treated as a sequence of data points, which can also be called a time series in statistics. Time series analysis is the subject that aims to find the methods for analyzing time series data in order to discover meaningful statistics and relationships of the data. The Bayesian analysis method we discussed in section III only regards the location information but ignored the ordering of these crimes. The temporal ordering information of a serial crime is often as significant as the location information. And the Bayesian analysis method also requires the assumption that all locations are independent of each other in order to make the model mathematically simple in calculation. However, as we know, one important property of the serial crime is that there exist some relationships among all individual crimes. So the locations may also be related and this relationship often contains meaningful information for depicting criminal's behavior. For example, when a criminal commit a serial crime, he/she may gain experience from the previous crimes and adjust his/her strategy in the location choosing for the next one. Actually, Bayesian analysis as a widely used static

Multivariate statistics tool is inadequate in many time variable applications. On the contrary, time series analysis is naturally a tool for the temporal ordering and correlated data. Therefore, we have good reason to apply time series analysis in solving geographical profiling problems.

B. Autoregressive models

First, we label the available crime locations by their sequence as $x_1 \dots x_n$, where $1 \dots n$ is the natural temporal order of the sequence. And denote the latent new crime location as x_{n+1} . Consider most general situation, each x_n is decided by some function of t points before it and some error term, where $0 < t < n$. So we can present x_n as $x_n = f(x_{n-1}, x_{n-2}, \dots, x_{n-t}, \varepsilon_n)$, where ε_n is a white noise error term usually obey a zero mean normal distribution, i.e. $\varepsilon_n \sim WN(0, \sigma^2)$.

There are many classes of time series models which can have different stochastic presentations. One of them with practical importance is the autoregressive (AR) model. A typically autoregressive model of order t can be defined as

$$X_n = \gamma_1 X_{n-1} + \gamma_2 X_{n-2} + \dots + \gamma_t X_{n-t} + \varepsilon_n, \quad \varepsilon_n \sim WN(0, \sigma^2). \tag{12}$$

By replace x_n with $x_{n-1} \dots x_{n-t+1}$, we can get a set for formulas

$$\begin{cases} X_n = \gamma_1 X_{n-1} + \dots + \gamma_t X_{n-t} + \varepsilon_n \\ X_{n-1} = \gamma_1 X_{n-2} + \dots + \gamma_t X_{n-t-1} + \varepsilon_{n-1} \\ \vdots \\ X_{n-t+1} = \gamma_1 X_{n-t} + \dots + \gamma_t X_{n-2t+1} + \varepsilon_{n-t+1} \end{cases} \tag{13}$$

By solving these equations, we can fix all parameters $\gamma_1 \dots \gamma_t$. Then we can predict X_{n+1} based on this model,

$$X_{n+1} = \gamma_1 X_n + \gamma_2 X_{n-1} + \dots + \gamma_t X_{n-t+1} + \varepsilon_{n+1} \tag{14}$$

For a real n point geographical profiling problem, we can carefully choosing the model order t to guarantee that $2t + 1 < n$ in order to make our model resolvable. The model will be a little more complicated if we take the anchor point z into consideration. The number of anchor point and the criminal's behavior will both affect the model form. In this sub section, we will start our discussion with the simplest situation. We assume that there is only one anchor point, and the criminal must return this anchor between any two consecutive crimes. By this basic assumption, we get the new time series which contained anchor point z as $z, x_1, z, x_2, \dots, z, x_n$.

Thus, for any given n and t , we can update our model as

$$\begin{aligned} X_n &= \gamma_1 X_{n-1} + \phi_1 Z + \gamma_2 X_{n-2} + \phi_2 Z + \dots + \gamma_t X_{n-t} + \phi_t Z + \varepsilon_n \\ &= \gamma_1 X_{n-1} + \gamma_2 X_{n-2} + \dots + \gamma_t X_{n-t} + (\phi_1 + \phi_2 + \dots + \phi_t) Z + \varepsilon_n \end{aligned} \tag{15}$$

Since $\phi_1, \phi_2, \dots, \phi_t$ are all parameters, we can combine them as a single parameter ϕ , then we can rewrite (15) as

$$X_n = \phi Z + \gamma_1 X_{n-1} + \gamma_2 X_{n-2} + \dots + \gamma_t X_{n-t} + \varepsilon_n \tag{16}$$

Then by choosing a proper order t ($2t < n$) and replacing x_n with $x_{n-1} \dots x_{n-t}$, we can get a new set for formulas,

$$\begin{cases} X_n = \phi Z + \gamma_1 X_{n-1} + \dots + \gamma_t X_{n-t} + \varepsilon_n \\ X_{n-1} = \phi Z + \gamma_1 X_{n-2} + \dots + \gamma_t X_{n-t-1} + \varepsilon_{n-1} \\ \vdots \\ X_{n-t} = \phi Z + \gamma_1 X_{n-t-1} + \dots + \gamma_t X_{n-2t} + \varepsilon_{n-t} \end{cases} \tag{17}$$

After solving these equations, we can fix $\phi, \gamma_1, \dots, \gamma_t$, and get the expression of anchor point Z ,

$$Z = (X_n - \gamma_1 X_{n-1} - \dots - \gamma_t X_{n-t} - \varepsilon_t) / \phi \tag{18}$$

C. Nomadic criminal and Multi anchor points situations

Sometimes criminal may commit crimes continuously without going home. Or some criminal commit crimes while traveling. However, it is always hard to judge when and how often a criminal will return his/her anchor point during this series or whether this criminal is traveling into this area to commit crimes. As so far there is no published evidence that police officers are able to answer above questions with any accuracy. But if we need this kind of assumption to help our investigation, we can simply replace the time series that matches our evidence and use the method in Section III to solve it. For example, $A, X_1, A, X_2, \dots, X_n$ shows the situation that the criminal just returned his/her anchor point after first crime and traveling into another area to commit continuous crimes.

TABLE IV.
DESCRIPTION OF CASES

Case No.	The number of crimes	Criminal Nationality	The type of offender
C01	4	China	Murder
C02	5	China	Murder, Raper
C03	5	China	Murder, Raper
C04	5	China	Murder
C05	5	U.S.A.	Murder
C06	6	China	Murder
C07	6	U.S.A.	Murder
C08	6	China	Murder
C09	7	China	Robber
C10	10	British	Murder
C11	12	China	Murder, Raper

As the proverb goes "The mouse that has but one hole is quickly taken." Though the single anchor point assumption perfectly fits many cases, some extreme crafty criminals may have multi anchor points to hide themselves. So for some long-serial criminals, it is necessary to doubt they have more than one hideout. That is why and when we need to take a multi anchor model into account. In the following discussion, we only deduce the model for two anchor points. The idea for the model with anchor points more than two is just the same, but more complicated in representation.

Assume a criminal committed a serial criminal x_1, \dots, x_n has two anchor points A and B . And the average distance this criminal is willing to travel to commit a crime is α . By the basic decay model and consider the stochastic factors, each crime location can be described as

$$X_n = \alpha_n A + \beta_n B + R \exp(-\varepsilon_n + i\theta_n), \quad (19)$$

Where ε_t and θ_t are white noises. The complex function $R \exp(-\varepsilon_t + i\theta_t)$ is the stochastic decay term, which implies a random vector with a negative exponentially distributed radius and a uniformly distributed polar angle. And R is the mean of its radius.

As we known, crime locations x_1, \dots, x_n themselves is a time series which can be depicted by an AR (t) model

$$X_n = \gamma_1 X_{n-1} + \gamma_2 X_{n-2} + \dots + \gamma_t X_{n-t} + \mu_n. \quad (20)$$

Combine these two functions, we get

$$\begin{aligned} & \alpha_n A + \beta_n B + R \exp(-\varepsilon_n + i\theta_n) \\ &= \gamma_1 (\alpha_{n-1} A + \beta_{n-1} B + R \exp(-\varepsilon_{n-1} + i\theta_{n-1})) \\ &+ \dots + \gamma_t (\alpha_{n-t} A + \beta_{n-t} B + R \exp(-\varepsilon_{n-t} + i\theta_{n-t})) + \lambda_n \end{aligned} \quad (21)$$

Due to $\gamma_1 \dots \gamma_{n-1}$ are parameters, we can replace them by $-\gamma_1 \dots -\gamma_{n-1}$ and rewrite the above function as

$$A \sum_{i=0}^t \gamma_i \alpha_{n-1} + B \sum_{i=0}^t \gamma_i \beta_{n-1} = \Phi_t, \quad (22)$$

Where

$$\Phi_t = \lambda_n + R \sum_{i=1}^t \exp(-\varepsilon_{n-i} + i\theta_{n-i}) - R \exp(-\varepsilon_n + i\theta_n). \quad (23)$$

Since A and B are two different anchor point, we can treat α_n and β_n as two independent time series. Furthermore, we need α_n and β_n be two weak stationary time series for the mathematical stringency. The mean and covariance of a weak stationary time series X_t have the following properties:

Expectation $E(X_n) = \mu < \infty$, independent of n ;

Variance $Var(X_n) = \sigma^2 < \infty$, independent of n ;

Covariance $Cov(X_n, X_{n-t}) = \eta_t$ is a function of t only.

The idea of weak stationary condition is that the behavior of the underlying process does not change with time. As a basic assumption of time series analysis, this condition can be easily satisfied by most non-dynamic time series. Also we notice that λ_n , ε_n and θ_n are all white noises independent with α_n and β_n . Thus we get

$$\begin{aligned} E(\Phi_t) &= E(\lambda_n + R \sum_{i=1}^t \exp(-\varepsilon_{n-i} + i\theta_{n-i}) - R \exp(-\varepsilon_n + i\theta_n)) \\ &= E(\lambda_n) + R \sum_{i=1}^t E(\exp(-\varepsilon_{n-i} + i\theta_{n-i})) - RE(\exp(-\varepsilon_n + i\theta_n)) \quad (2) \\ &= R(t-1) \end{aligned}$$

4)

And $Cov(\Phi_t, \alpha_n) = Cov(\Phi_t, \beta_n) = 0$.

Taking $E(\Phi_t)$, $Cov(\Phi_t, \alpha_n) \dots Cov(\Phi_t, \alpha_{n-t})$ at the right side of function (22), we can get

$$\begin{cases} (A \mu_\alpha + B \mu_\beta) \sum_{i=0}^t \gamma_i = R(t-1) \\ A(\gamma_0 \eta_1 + \gamma_1 \eta_0 + \gamma_2 \eta_1 \dots + \gamma_t \eta_{t-1}) = 0 \\ A(\gamma_0 \eta_2 + \gamma_1 \eta_1 + \gamma_2 \eta_0 \dots + \gamma_t \eta_{t-2}) = 0 \\ \vdots \\ A(\gamma_0 \eta_t + \gamma_1 \eta_{t-2} + \gamma_2 \eta_{t-2} \dots + \gamma_t \eta_0) = 0 \end{cases} \quad (25)$$

This set of linear equations is usually called Yule-Walker equation. Similarly, we can also write the Yule-Walker equation for B. By combining these two sets Yule-Walker equations, we got a set of liner equations which contains $2t+2$ variables and $2t+2$ independent equations. So far, we get the unique solution of this question.

TABLE V. THE COMPARISON BETWEEN MODELS WITH THE SHORT GROUP OF CASES

	GIS-Decay model	Bayesian-Factor analysis model	Time series analysis model
Hit	4	5	1
Close	1	0	4
Further	0	0	0
Average Search Cost	19.5%	14.5%	27.2%

IV. MODEL EVALUATION AND ANALYSIS

In this section, we will apply several tests to evaluate the accuracy, effectiveness and robustness of our model. Firstly, we define two measures— ‘Hit score percentage’ and ‘Profile accuracy’ to present the effectiveness and accuracy of geographical profiling models. Then we calculate these two indexes of our models with both historical samples and Monte Carlo simulation cases. Finally, we test the robustness of our models by ‘misabeled point’ method. Followed by each test, we analyze and compare the advantages and disadvantages among these three models.

A. Historical sample tests

TABLE VI. THE COMPARISON BETWEEN MODELS WITH THE LONG GROUP OF CASES

	GIS-Decay model	Bayesian-Factor analysis model	Time series analysis model
Hit	1	1	2
Close	1	1	0
Further	0	0	0
Average Search Cost	23%	28%	6%

By online data collection, we get 11 true historical serial criminal samples, including 8 Chinese cases, 2 American cases, and one British case. The British sample is just the Peter Sutcliffe case which was mentioned in the Problem B. All these cases are categorized into three classes by the number of crimes: less than 6, between 6

and 8, and above 8. (See Table IV) In this article, we choose ‘Hit Score Percentage’ and ‘Profile Accuracy’ as measures to evaluate the effectiveness and accuracy of models.

The Hit Score Percentage is a measure of search efficiency of the model. It is defined as the ratio of the area searched (following the geographic profiling prioritization) before the offender’s base is found, to the total hunting area; the smaller this ratio is, the better model performs. The hunting area is defined as the rectangular zone oriented along the street grid containing all crime locations. [4] Profile Accuracy is defined as a measure of whether the offender’s base is within the top profile area. In this paper, we provide a simple classification to indicate whether the estimations are ‘hit’, ‘close’, and ‘off-target.’

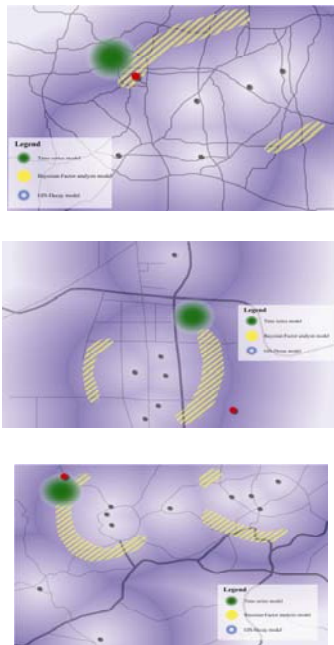


Figure1. C02, C08 and C10 case results using those three models.

The tests of Bayesian-Factor analysis and Time series models are programmed by MATLAB, while GIS-Decay model is conducted by GIS directly. The parameter values of Bayesian-Factor model are selected from the average values published in some overview articles [9]. The typical results for the three models in each class are shown as Fig. 1. Table V, VI and VII is the comparison of the models when handling the cases in short group, medium group and long group separately. From above table we can tell that these three models are respectively accomplished in dealing ‘short’ (less than 6 offends), ‘medium’ (6-8), and ‘long’ (more than 8) serial crimes. GIS-Decay model has an obvious advantage for cases with small data set, but a relatively low efficiency than other two models. So this model can give useful information to police agency even at the beginning of a serial crime. Bayesian-Factor analysis model has moderate accuracy and efficiency but a strong robust. This model is suitable for the case with medium data set and the case that we have some information about

criminal’s characteristic. And time series model, though performs poor at short serial crimes, its estimate accuracy increases greatly while data set getting large. For long serial crimes, it can give a precious result.

B. Monte-Carlo simulation tests

Due to the number of true historical serial criminal samples are limited, our model need turn to Monte-Carlo simulation for fully tests. Our idea of simulation is first fix an anchor point, and generate several data points around anchor point as crime locations by Monte Carlo algorithm. Then we run three models to check how well they can find the anchor point. In this test, we change the number of generated data points from 3 to 20, and run 1000 repeated simulations for each circumstance. The results of the test are shown in the Fig. 2. According to the figure, we can see, the result of simulation test basically matches the result of historical sample test. This proves that are suitable for common cases analyze and have good stability.

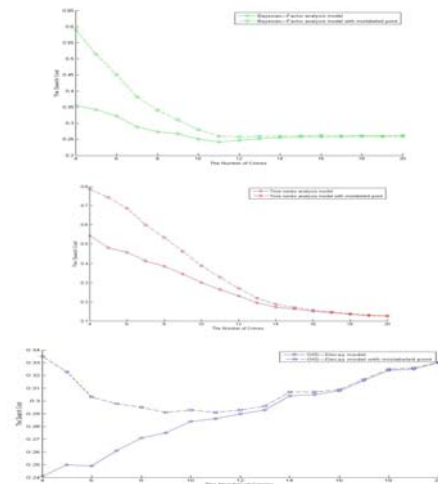


Figure2. Robustness of Those Models

C. Robust test

Sometimes, it is inevitable for police officers to mislabel an irrelevant offend into the underlying serial crimes. We concern whether our models can still give a relatively correct answer when this happened. So the robust test for models is necessary and meaningful.

TABLE VII.
THE COMPARISON BETWEEN MODELS WITH THE MEDIUM GROUP OF CASES

	GIS-Decay model	Bayesian-Factor analysis model	Time series analysis model
Hit	3	2	1
Close	1	2	0
Further	0	0	3
Average Search Cost	16.75%	32%	56.75%

In order to test robustness of three models, we conduct Monte Carlo method to generate data points around the fixed anchor point, and then we randomly locate a mislabeled data point into the hunting area to test if our

models can still find the anchor point. For each n ($3 \leq n \leq 20$), we run 1000 repeated simulations. Then for each model, we compare its new search result with the former one. The comparison results are shown in the following pictures. According to this result the robust of all models increases while data set getting large. We should particularly notice that time series model has a poor robust when $n < 6$.

V. CONCLUSIONS

Geographical profiling is an investigative methodology that uses the locations of a connected series of crimes to determine the most probable offender residence. In this article, we established three different geographical profiling models: GIS—Decay model, Bayesian—Factor analysis model, and Time series analysis model. These three models are respectively accomplished in dealing ‘short’ (less than 6 offends), ‘medium’ (6-8), and ‘long’ (more than 8) serial crimes.

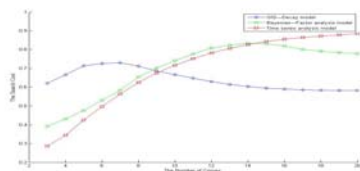


Figure3. Profile Accuracy

However, as a new developed methodology, our models are still imperfect. The form of some functions in our models still relay on artificial choices. That means the police officer who make use of these models need have a good experience of regional serial crimes and will not bring his/her own bias into model operation. Inappropriate use of geographical profiling can have serious consequences. That is all the more reason why we must develop our understanding of what introduces biases and errors into geo-behavioral decision support systems and into the cognitive processes of those who use those systems. Treating these errors as operational problems that have to remain in the hands of police officers will keep criminal investigation in the dark ages of intuition and hunch.

ACKNOWLEDGMENT

This work was supported in part by the 2010 Scientific Fund of Beijing Education Commission (ITEM NO. KM201000002002).

REFERENCES

- [1] Rossmo, D. K. Geographic profiling. Boca Raton, FL: CRC Press,2000.
- [2] Paulsen, D. J. Human versus machine: A comparison of the accuracy of geographic profiling methods. *Journal of Investigative Psychology and Offender Profiling*, 3 (2), pp.77–89,2006.
- [3] Snook, B. Individual differences in distances traveled by serial burglars. *Journal of Investigative Psychology and Offender Profiling*1, pp. 53–66. DOI: 10.1002/jip.003,2004.
- [4] Rossmo, D. K. Geographic heuristics or shortcuts to failure? Response to Snook et al. *Applied Cognitive Psychology*, 19, pp. 651–654. DOI: 10.1002/acp.1144,2005.
- [5] Brantingham, P. J., & Brantingham, P. L. *Environmental criminology*. Beverly Hills, CA: Sage Publications,1981.
- [6] Felson, M. Linking criminals’ choices, routine activities, informal control, and criminal outcomes. In D. Cornish, & R. V. Clarke (Eds.), *the reasoning criminal: Rational choice perspectives on offending*,pp. 119–128, New York: Springer,1986.
- [7] Clarke, R. V., & Felson, M. *Routine activity and rational choice*. New Brunswick, NJ: Transaction,1993.
- [8] Lundrigan, S., & Canter, D. V. Spatial patterns of serial murder: An analysis of disposal site location choice. *Behavioral Sciences and the Law*, 19, pp. 595–610,2001.
- [9] Brent Snook & Richard M. Cullen. *Serial Murderers’ Spatial Decisions: Factors that Influence Crime Location Choice*. *Journal of Investigative Psychology and Offender Profiling*. *Psych. Offender Profil.* 2: pp.147–164,2005.

Guanli Huang, female, born in 1975, Associated professor of Beijing Vocational College of Electronic Science, Master's degree with the major of computer science. Participant Leader of Beijing Quality Course on Network development and the Committee member on China Computer Federation. She published widely, such as academic articles in Computer Engineering and Applications, Computer Science etc; her research interests include algorithm design, Computer Education etc; and has won one National Patents and has published over twenty academic articles as the first author, two of which are EI Indexed. At present, Ms. Huang also undertakes some items, such as Research and Development on Dynamic dispatching GPS system of the Adaptation Road Condition of the Beijing Education Committee Scientific Project. She is also the Participant of Simulation Platform of Small Hybrid Vehicle Control Based on dSPACE of the Beijing Science and Technology Innovation Platform Project, and Practice of School-enterprise Cooperation Mechanism and Platform based on Diversification of the Beijing University Education Reform Projected. Her research Directions are data analysis, signal control, information security, education management.

Guanhua Huang, 1978, a graduate student at Beijing Normal University. His research interests include Data management and technology translation.