Malaysian Journal of Mathematical Sciences 7(1): 125-145 (2013)

Mathematical Modelling in Criminology

[1]Feroz Shah Syed, [1]Didong Li,
[1]Xun Zhang and [2]Zhenhong Guo

[1]*Department of Mathematics, Beijing Institute of Technology, Beijing 100081, P.R.China*

[2] *Department of Information Engineering, Beijing Institute of Technology, Beijing 100081, P.R.China*

E-mail: sf_shah@hotmail.com

ABSTRACT

This paper, determines the possible position of the serial criminals by solving the geographical outline problem using our developed mathematical model. Under a series of reasonable assumption, we build some models for different types of cases, including the optimization model with the minimal distance, the GM (1, 1) model, the probability contour model and the escape model. We use two schemes to generate two geographical profiles, and then give an efficient prediction by combining them together. The predictions contain the locations where the murderer is most likely to live with its approximate range and the possible time of the next crime. Furthermore, the degree of confidence of the prediction is also given. In addition, we give an executive summery which introduces our technique briefly and provide a broad overview of the potential issue. Applying the optimization model with the minimal impact, we firstly determine the geographical profile where the criminal may hide. Moreover, we estimate the next crime time with the GM (1,1) model. Then we use the probability contour model to predict the profile of the next possible crime locations. Taking this all into consideration, we give a useful and efficient scheme and test it for some cases. Also, we built an escape model for the situation in which the murderer still commits crimes in escape. Finally, a summary is made, in which we provide suitable and unsuitable cases for each model. Besides, the strengths and weaknesses of the model as well as the possible way to improve the models are discussed.

Keywords: serial crimes, escape model, degree of confidence.

1. INTRODUCTION

In recent times, a serial crime refers to commit crimes three times or more in series in a short time. This not only infringes on the rights of victim, but also keeps people in a panic. The serial murderer kills lots of victims, making a terrible influence on the society. So we hope to arrest the murderer as soon as possible. We give an example of the criminal rate, as discussed by Yan *et al.* (2009) in their studies.

Generally, the serial crime has its regularity in terms of time and space. By studying and analyzing those regularities, we can narrow the range of searching and crack the case faster. Peter Sutcliffe is taken as a sample, who was convicted of thirteen murders and number of other crimes in 1998. One of the methods to narrow the search for Peter was to find a "center of mass" of the location of crime. By this technique, it is know that the suspect happened to live in the same town. Since then, many of the sophisticated techniques have been developed to determine the "geographical profile" of a suspected serial criminal based on the locations of the crimes. Bonnie *et al* (1990) has discussed the issue of the legal representation of developmentally disabled defendants. The possibility that those deemed developmentally disabled may also suffer from a psychiatric disorder as discussed by Borthwick-Duffy and Eyman (1990).

Of all the existing models, the analysis method put forward by mathematician Kim Rossmo was best known. However, it also has some weaknesses, so we manage to come up with some new ideas. Whereas, we use two different schemes to generate a geographical profile. By combining the results of different schemes, we develop a method to detect where the murderer is most likely to live and predict the approximate range of location and time of the next crime. At last, we will use several specific cases to test the model, and check the strengths and weaknesses of the model as well as the possible way to improve of the model.

Criminals with a fixed location

According to the criminal psychology, if the murderer is ensured relatively safe, he is always likely to select the region closer to his location to commit crimes. Therefore, we need to find an area not far from every crime spot. Then we use the minimum distance model to find the profile of the murderer's location.

In terms of the times of every case of serial of crimes, the individual crime motivation presents a characteristic of cyclical change. So, we can use the GM(1,1) model to estimate the possible time of next crime.

After a murder case, the region will strengthen the police force, so the murderer would tend to commit crime twice or more at a same place. Therefore, while taking other factors into consideration, we should also consider the next place influenced by the former crime spots. The next spot of crime couldn't be too far from his location, or it will be more dangerous and the cost could be larger. So within a range of a certain distance, the murder is likely to go to the place not be influenced much by the former case. Therefore, with the critical values given at first, we would like to use the Probability contour model to predict the possible place of next crime. This approach is based on the theory of Probability and Statistics.

Combining the two approaches above together, we can predict that where the murder may be at a possible time. Hence, we can try to arrest the murder according to the time and other factors. For example, we can go to his fixed location to arrest him, or strengthen the police force at the place the murder may go to next time.

Criminals commit crime when escaping

According to the criminal psychology, the suspects will hardly retrace one's steps, so he has a general direction of escape. Every time the angle drifted won't be very large, and the distances between every two murder places won't differ much. So with the Escape model, based on the former murder spots, can help us to estimate the escape direction, the route, and the possible spot of next crime.

The approaches above can be applied to these different situations, but there are still some situations, that we have not discussed in our model. For example, employing for killing or being forced to murder, the motivation to murder is something else but not be decided by murderer himself. For these situations, the regularity of time and space is not so evident.

2. PROBLEM FORMULATION

Basic assumptions

1. Neglect the influence of natural disasters such as earthquakes, tsunamis, floods, typhoons, etc.
2. Neglect the social control (the control of turnover in the period of epidemic etc.)
3. Neglect the pressure from chasing, public opinion and other aspects.

Rossmo's formula

The renowned mathematician Kim Rossmo produced a geographic profiling formula to predict the location of a serial criminal. Consider a map with an overlaying mesh of small squares (sectors). A sector $S_{i,j}$ is the square on row i and column j, located at coordinates (x_k, y_k). The probability of the position of the killer being in a specific sector is specified as:

$$p_{i,j} = \alpha \sum_{k=1}^{N} \left[\frac{\lambda}{\left(\left|X_i - x_k\right| + \left|Y_j - y_k\right|\right)^{\xi}} + \frac{(1-\lambda)\left(\beta^{\eta-\xi}\right)}{\left(2\beta - \left|X_i - x_k\right| + \left|Y_j - y_k\right|\right)^{\eta}} \right]$$

The summation is over N past crimes located at coordinates (x_k, y_k) and it consists of two terms. The idea of decreasing probability with increasing distance is described in first term and the second term deals with the concept of a buffer zone. The variable λ is used to put more weight to one of the two ideas. The variable β defines the radius of the buffer zone. The constant α is empirically determined.

This formula expresses that the probability of crimes increases as one moves through the buffer zone away from the hot zone, but decreases afterwards. The variable ξ is chosen to work best on data of past crimes. The variable η also works on the same idea. The distance is calculated by the Manhattan distance formula.

This method is comparatively reasonable, and many problems have been solved by this method actually. Using the method, the police have cracked several serial crimes. However, it still has some weaknesses at some extent, because it merely takes the influence of the distance into account. In all the sections, we choose the only one point which has the maximum p, thus we ignore the section whose probability is close to maximum.

Then we develop new models, using two different schemes to generate a geographical profile. We finally get the expected calculation after combining the results of two different schemes.

According to the reference, we find that except the thirteen murders Sutcliffe had many attempted murders and intentional injuries. When we analyze the law of his committing crime, we shouldn't ignore them. Now we integrate them together as in Figure 2.

Scheme 1

Scheme 1 and 2 are all for the case in which murderers have a fixed hiding place. Generally, the murders won't escape when the police haven't notice them (e.g. Peter William Sutcliffe's case). For mobile criminals and other situations will be discussed later in this paper. Based on the locations of the former crime spots, we estimate the geographical profile of the murderer's anchor point.

Firstly, we establish a coordinate system at the region of crime site, and then we find the coordinates of the past crime spots in the coordinate system. On the basis of criminal psychology, the criminal is likely to choose the nearer place to commit a crime on the condition that safety is guaranteed. In this case, our objective is to find a place where we can make the distance from it to those spots reach the minimum.

Making the difference of the property and geographic environment per crime into consideration, we put a weight per distance that modulus are multiplied by each distance. We assume that the serial crime is happened in area D. There has happened N crimes and the k^{th} crime spot is C_k, which is marked by (x_k, y_k). $d(P,C_k)$ denotes the distance between the point p and point C_k. $f(C_k)$ is the buffer function at the area where C_k is located and the function has something with the surroundings of point C_k such as topography, population density, etc. The problem is equal to the equation group as follows:

$$\begin{cases} \min\left\{ \left[\sum_{k=1}^{N} \left[f(C_k) d(P,C_k) \right]^2 \right]^{1/2} \right\} \\ P \in D \end{cases} \tag{1}$$

where, $d(P,C_k)$ is the Manhattan distance which is affected by topography, landform and other factors. $f(C_k)$ should be determined according to the specific conditions because its value is affected by many different factors.

Killing and theft influence the criminal and his mentally in varying degrees. As common sense, killing puts more pressure on the criminal, so he won't choose a further place to commit a crime, so its value of function should be smaller. Correspondingly, theft has larger value of function. Besides, different crime site has different influence on the criminal. The Chicago school of geography of crime put forward concentric circle theory, suggesting dividing the city into five concentric circles with distinctly bounds. The first circle is the most smaller circle in the most inner layer of the city, or the central commercial district. The second circle is the buffer district or middle zone that is in downtown. The third circle is the residential area of workers. The forth circle is the residential section of upper-and-middle class. The outermost fifth circle is the suburb of the city or satellite city. According to the research, the crime often happen in the middle of the clear zone, between the central commercial district at the first circle and the residential area of workers at the third circle. In his theory, there are most crimes in this transition zone, and the farther away from the central commercial district there are the fewer crimes. The crime rate presents "island effect" near large shopping malls and industrial centers, so the crime rate increases. Meanwhile, the crime decreases as the distance increases. On this assumption, the value of function of this transition where the crime rate is highest reaches the maximum.

Apart from the above factors, there are many other factors which affect the value of function. We should analyze according to specific cases, and sometimes we even evaluate depending on experiences.

After we fix the centre, we need to enclose the scope of search, that is to say we should ascertain the radius of search. We assume that the radius of the searching area is r , the longest distance is R . Then we have equations as follows:

$$\begin{cases} R = \max\limits_{1 \le k \le N} \left\{ d\left(P, C_k\right) \right\} \\ r = \alpha R \qquad \alpha \in \left(0,1\right) \end{cases} \tag{2}$$

where, parameter α is determined by the situation of the police force and the progress in detection of crime, for example the increase of police force will lead to extending the range of searching area, along with which the value of parameter α will increase, or vice versa.

Next, we apply the method above to solve the criminal case about Peter William Sutcliffe. In this case, many references such as landform, topography, weather etc. can't be obtained, so we can't determine the value of $f\left(C_k\right)$.

In order to simplify the problem, we define the $f\left(C_k\right)$ as number 1. $d\left(P, C_k\right)$ also cannot be determined, so we choose Euclidean distance to calculate. The result of prediction is shown in Figure 4. Then we compare the result with the actual location (signed with red point) where the murderer was arrested, which is shown in Figure 5.

Scheme 2

Based on the locations of the past crime scenes, we predict the place and the time where and when the next crime is likely to happen.

Time predicted

According to the criminal psychology, in terms of the time of the serial crime events, individual crime motivation shows the characteristic of the periodic variation, which represents the satisfaction of bad needs of the criminal. Therefore, it is the reflection of the fluctuating of psychological development of crime. Although the periodic variation of crime motivation is the common feature of psychological development of crime, it shows personalities on different individuals. Making further research of the periodic variation of crime motivation can allow us to get further understanding of the change of criminal's mind and the occurrence of crime, especially for the study of the serial crime of individual.

We use GM(1,1)(Gray Model) to predict the time of the next crime. We assume that the time of the k^{th} crime is $T^0(k)$ then we obtain the original series $T^0 = (T^0(1), T^0(2), ..., T^0(N))$. Adding up the numbers, we can get a new array of numbers:

$$T^1 = \left(T^0(1), \sum_{m=1}^{2} T^0(m), \cdots, \sum_{m=1}^{N} T^0(m)\right) = \left(T^1(1), T^1(2), \cdots T^1(N)\right)$$

Then we give the Albino equation $\dfrac{dT^1}{dk} + aT^1 = b.$

We denote $Y_N = \begin{bmatrix} T^0(2) \\ T^0(3) \\ \vdots \\ T^0(N) \end{bmatrix}$ $B = \begin{bmatrix} -\frac{1}{2}\left(T^1(1)+T^1(2)\right) & 1 \\ -\frac{1}{2}\left(T^1(2)+T^1(3)\right) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}\left(T^1(N-1)+T^1(N)\right) & 1 \end{bmatrix}$ $\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix}$.

We hope to make the residual equation reach the minimum (the thought of least-square method). Finally we obtain the equation as follows

$$\hat{a} = \left(B^T B\right)^{-1} B^T Y_N.$$

We denote the time response function by $\hat{T}^1(x)$ including unknown parameter, and then insert the initial value to figure out the unknown parameter. After reverting the function through derivation, we get $\hat{T}^0(x)$, then insert $x = N+1$ to obtain the time of the next crime.

In the same way, we apply this method to the crime case about Peter William Sutcliffe. We record the time of the first crime scene as number 0, and in days list the time of per crime in order.

Figure 6 shows the time of each crime. In the figure, periodicity is not shown clearly and the figure has large fluctuations, so we do an accumulation of the data to obtain the Figure 7.

In Figure 7, the curve is smooth and has obvious rules. Using the gray model, we obtain the ultimate function after substituting the data into it.

$$\hat{T}^0(x) = 2160.3e^{0.030487x}.$$

Inserting $x = 16$, we predict the time of the next crime (the 16th) is 3517, the relative error is $\frac{(3517-3488)}{3488} = 0.83\%$, which is relatively accurate.

Situation predicted

On the basis of some theory of criminal psychology, the region suffering one murder will strengthen alertness, so generally the criminal won't choose the same place to commit crimes twice or more. So when we take into account the other factors, we consider that the location of the next crime is affected by the locations of the former crime spots. But the location can't be too far away from the place he lives, because the further away the location is, the higher the danger is. So within a certain range of the hiding-place, the criminal is most likely to choose the location where isn't influenced much by the former crimes.

In this district, we make each of the former crime spot as the center and respectively generate two-dimensional normal distribution function. All of the added up and it looks like some 'hills' after plotting, as shown in the figure 8. The figure 8 intuitively expresses the impact of the former crimes. The impact on the district with lower height in the figure is smaller. So the crime is most likely to happen in the district with the lowest height, and the crime also probably occur in the district whose height is approximate to the lowest. Obviously, the height of zone which is infinitely far away this district is zero, but this phenonena is against the condition that the location shouldn't be too far away of the hinding place. So it is reasonable to search the minimum point in definite profile but not in whole plane. In scheme 1, we figure the max distance R. The location of the next crime has larger possibility to be in the circular region whose centre is the hiding place, whose radius is R.

Therefore, we should search the location of which the impact of the former crimes is slighter in this section (denoted by D_R).

Besides, the crime is impossible to happen in some districts such as mountains, rivers, desert, etc. some criminals tend to murder in some specific places such as town and country, so search should be optional.

$g\left(C_k\right)$ is the buffer function which is used to adjust the height of the peaks. We denote the two-dimensional normal distribution function which is generated by the location of the k^{th} crime by $N_k\left(P\right)$. Denoting the district where the crime is impossible to occur by I, we can obtain the following equation.

$$\begin{cases} F(P)=\begin{cases} \sum_{k=1}^{N} g(C_K)N_K(P) & P\notin I \\ +\infty & P\in I \end{cases} \\ \min\{F(P)\} \\ P\in D_R \\ m=\min\{F(P)\} \\ F(P)\le(1+\beta)m \\ \beta\in[0,1] \end{cases} . \tag{3}$$

Where the parameter is generally small. Because the district attained from the odds has the highest rate of crime, the bigger the β parameter is, the larger the district is, and accordingly the total probability will bigger. The definite value of the β parameter is determined by the situation of local police force and other factors. Within the limits of the police force, we suggest magnifying the value of the parameter β. If the police force is not enough, we suggest reducing the value of the β parameter and search the most possible location. Generally, the solution of the inequality isn't an interconnected section, but consists of many separated districts.

Similarly, we apply this method to the criminal case about Peter William Sutcliffe, for example depending on the past 23 crimes to predict the location of the 24th crime. In this case, the offender kills the floozy specially and the locations of the crimes are all in residential area and commercial zone (donated by β),so the function $F(P)$ needed to be adjusted considering that the offender is impossible to commit a crime in other district. Then we have the following equations

$$F(P)=\begin{cases} \sum_{k=1}^{N} g(C_K)N_K(P) & P\in B \\ +\infty & P\notin B \end{cases} .$$

Similar to the scheme 1, $f(C_k)=1$; We regard $d(P,C_k)$ as Euclidian distance, $\beta=0.1$, $N=21$, to obtain the minimum point. As Figure 9: Substituting 0.1 into β , then we get a region with a relatively higher probability. Its intersections with B is the zone we search for, as in Figure 10.

Combination and prediction

After predicting the hiding palce of the murderer and the location, the time of the next crime, we can assert that in the period from the last crime to the next crime, the murderer is most likely to stay in the district between the hiding place and the location of the next crime. With the approching of the next expected crime, the murderer will move to the predicted point. So based on the predicted time and other specific conditons, we carry out arrest timely .

Schemes 1 and 2 have predicted the 22^{nd} possible crime position and time for the Sutcliffe's case. Therefore, during this period of time (after the 21^{st} and before the 22^{nd} committing a crime), the murder may be in the blue area in Figure 11.

The above model apply to arresting the kind of criminal who has fixed hiding place and has not been suspected by the police, beause generally he tends to come back home. But in other cases, we need to new model to predict.

Application

Figure 12 (left) shows the locations of a serial criminal spots. By using Predator Geographical Profiling system, geographical profile of the crime was produced. The wedge shaped pattern is drawn by the computer program. This wedge shaped pattern suggests that the most likely home base area for the killer is located within the wedge and probably the home base and/or a work area has a directional bias towards Northern Virginia. This prediction is based on our analysis of 54 American serial killers' spatial behavior. Finally, the movement of the criminal is preddicted which is different from the result of Rossmo. As Figure 12 (right, red), we use the method of scheme to predict the location of the next crime. Inserting data and solving the equation. The result is in Figure 13, we match the location to the corresponding distrcit on the map, and compare our result to the one of Doc.Godwin, as the Figure 12 (right, blue). The result we predict is approximate to Doc.Godwin's, indicating that our model has consistency with the theory of Doc.Godwin to some extent.

Scheme for escape

Once the crimial is suspected by the police , consequentially he won't come back home, and run away in a certain direction. During his escape, he will contiue commiting crimes and doesn't has a fixed hiding place, so there is no need to predict the hiding place. Therefore, we should use different method to predict the location and the time of crime.

Time prediction

Depending upon the criminal psychology, during the escape, the state of psychology is not steady, as is the speed of escape. The offender may run away madly or even stay still when he feels safe. Thus the time of crime hasn't regularity to follow so that it is different to predict the time.

Situation prediction

In general, the offender won't take the same path two times, so there is a approximate direction and the angle of avertence won't changes much. As the distance of escape increases, the mental pressure of the offender will reduce and the criminal motive will enhance, thus the distance between the adjacent location of crime on the way won't vary much.

We assume that the angle between the directed line segment form the k^{th} location of the crime to the $(k+1)^{th}$ and the level or vertical positive axis is $\theta_k = \theta(C_k, C_{k+1})$, the length of the line segment ,the point on the centerline of the possible district of crime is marked by letter P , and U stands for the group of city.

Then we give the following equations

$$\begin{cases} \bar{\theta} = \dfrac{1}{N-1}\sum_{k=1}^{N-1}\theta_k \\ \bar{L} = \dfrac{1}{N-1}\sum_{k=1}^{N-1}L_K \\ \left|\theta(C_N, P) - \bar{\theta}\right| \le \varphi \qquad \varphi \in \left[0, \dfrac{\pi}{2}\right] \\ \left|\dfrac{L(C_{N+1}, P)}{L_N} - 1\right| \le \delta \qquad \delta \in [0,1] \\ \dfrac{L(C_{N+1}, P)}{L_N} \le \delta' \qquad \delta' \in [0,1] \\ C_{N+1} \in U \end{cases} \tag{4}$$

C_{N+1}, we make out is the location of the next crime.

3. RESULTS AND DISCUSSION

About the problem how to ascertain the geogrphic profile of the different kinds of serial criminal, we build several model and finally obtain effective prediction. Using several existing cases to text our model, the results validate our model to the realities at a high degree. Different models have different scopes of applications. For example, scheme 1 applies to the offender having fixed hiding place, scheme 2 applies the offender who commits crime as he runs away, etc.

However , crime is subjective action of human being, so it has large uncertainty and is diffcult to predict accurately. In fact, there are many factors we need to consider and some of the factors are even complex. In the assumption, we neglect the influence of natural disasters, the social control, pressure public opinion, etc. We classify the serial crime, and as for different sort, we build different model, which reduces the error.

When predicting the hiding place of the criminal, with a relatively comprehensive consideration, we apply buffer function, combine the influence of landform, topography, density of population, the character of crime and other factors and add some modulus to adjust.

When predicting the location of the next crime, we not only find the point with maximum, but also mark out the district with the value approximate to the maximum, which avoids missing the section with high possibility.

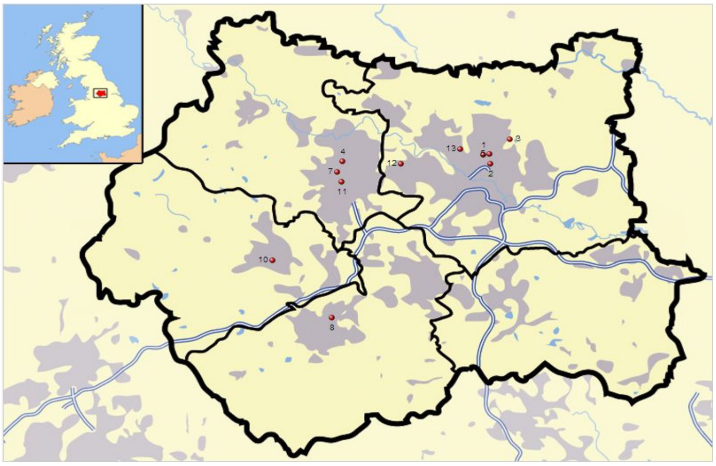

Figure 1: The information of 13 victims

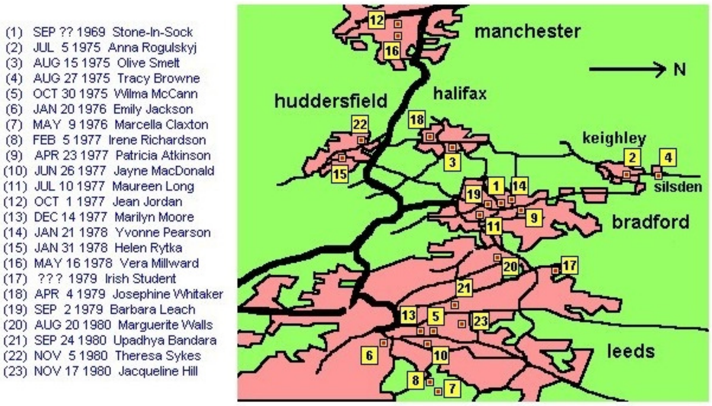

Figure 2: The attacks and murders

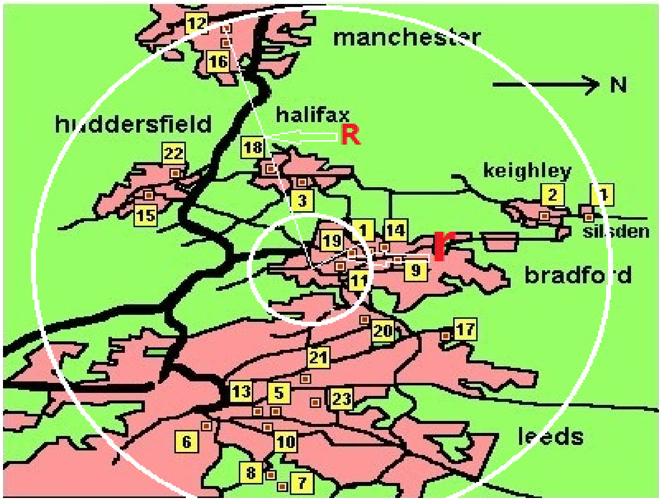

Figure 3: Two circles

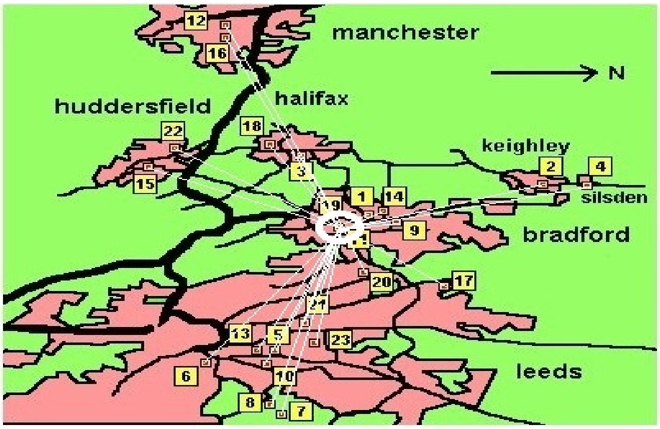

Figure 4: The refuge predicted by scheme 1

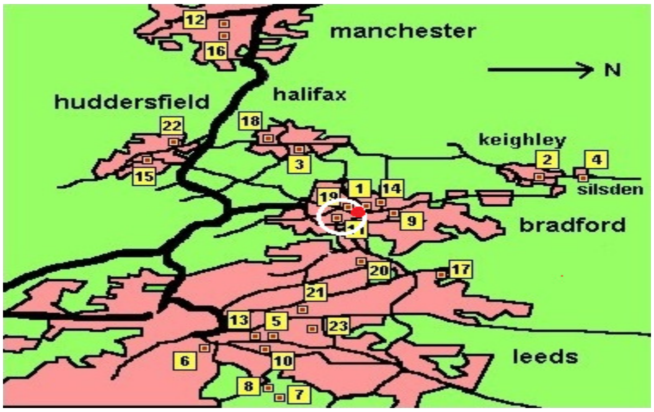

Figure 5: Comparison of theoretical and real

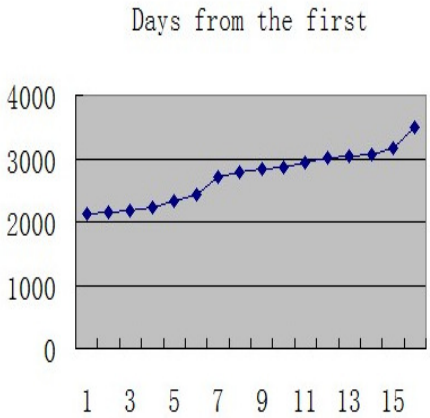

Figure 6: Days from the first

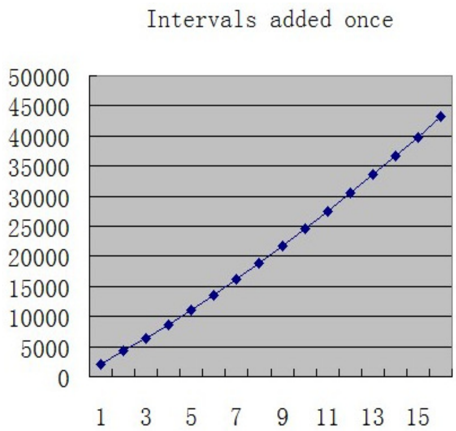

Figure 7: Intervals added once

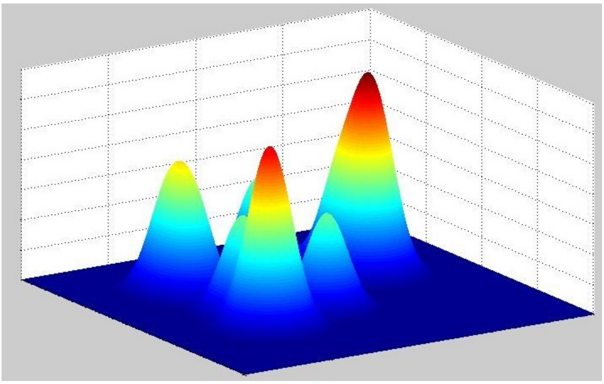

Figure 8: Probability contour map

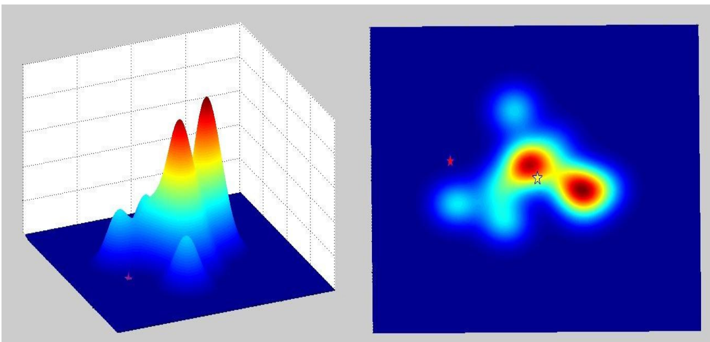

Figure 9: Solid and plane figure of scheme 2

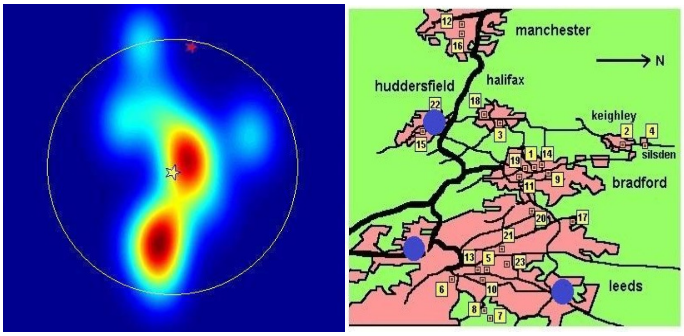

Figure 10: The situation predicted by scheme 2

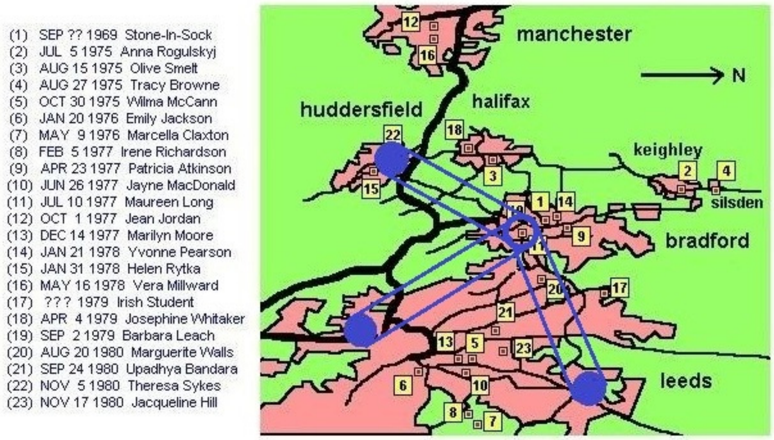

Figure 11: The scope of activities

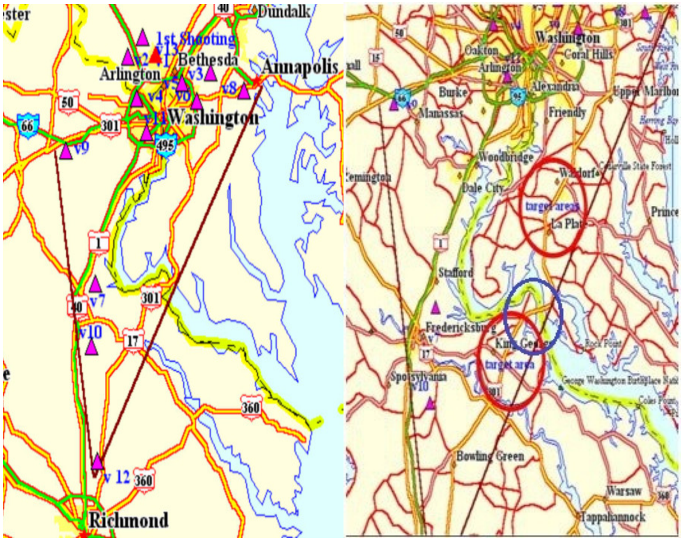

Figure 12: The situations and predictions of murder

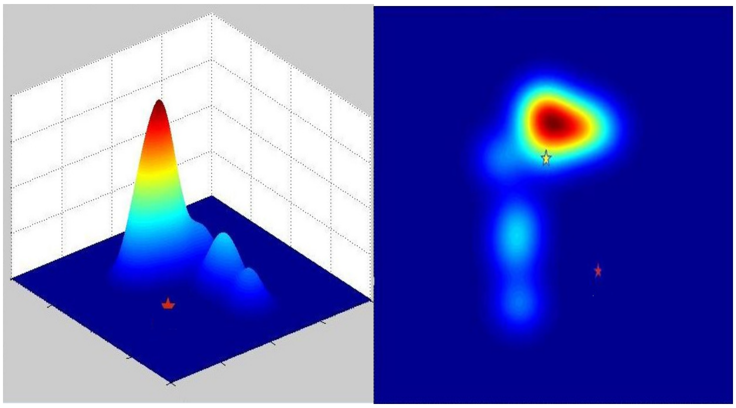

Figure 13: The results of scheme 1

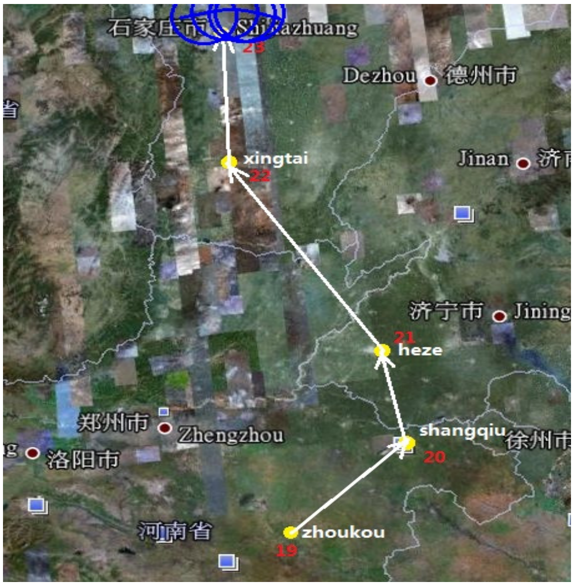

Figure 14: The escape area

4. CONCLUSION

We predict the hiding place and the location and the time of the next crime respectively, and combine them to find the district where the criminal is likely to stay in a period, thus we achieve our goal that is to make out an effective plan of arrest.

As with the offender who is running away, we not only predict the location of the next crime, but also roughly define the path of escape he takes, which benefits catching the criminal.

ACKNOWLEDGEMENTS

The financial support received in the form of the scholarship from the China Scholarship Council (CSC), Government of China is gratefully acknowledged. We are especially thankful to respected Professor Huafei Sun for his kind help and guidance.

REFERENCES

Balogh, R., Bretherton, K., Whibley, S., Berney, T., Graham, S., Richold, P., Worsley, C. and Firth, H. 2001. Sexual Abuse in Children and Adolescents with Intellectual Disability. *Journal of Intellectual Disability Research.* **45**(3): 194-201.

Bonnie, R. J. 1990. The Competence of Criminal Defendants with Mental Retardation to Participate in their own Defense. *The Journal of Criminal law and Criminology.* **81**(3): 419.

Borthwick-Duffy, S. A. and Eyman, R. K. 1990. Who are the Dually Diagnosed? *American Journal on Mental Retardation.* **94**(6): 586-595.

Fenghua, S. and Aihua, M. 2003. Study on the Theories of Criminal Geography. *Human Geography Journal.* **5**(18):70-74.

Li, G. 2008. The Application of Map Analysis. *Methods in Criminal Information Analysis*, 04-0011-05J.

Yingzhou, Z. 2006. A Research on the Periodicity of the Criminal Motive. *Journal of Henan Judicial Police Vocational College.* **3**(4): 31-34.

Yan, J., Yuan, H., Shu, X. and Zhong, S. 2009. Optimal clustering algorithm for crime spatial aggregation states analysis. *Center for Public Safety Research, Department of Engineering Physics, Tsinghua University.*