Journal of applied science and engineering innovation Vol.1 No.5 2014
ISSN (Print): 2331-9062
ISSN (Online): 2331-9070

Prediction Model for Law Enforcement Agencies on the Serial Criminals

Jiatai Gang, Qinwu Dai, Shujuan Li

College of Information Engineering, Dalian University, Dalian, China

gjt1960@126.com

Keywords: Serial criminals; Geographical Profiling; Crime site; Simulated Annealing Algorithm

Abstract: Due to the rapid increase of serial crime and diversity of criminal methods, sophisticated techniques have been developed to determine serial crime though "Geographical Profiling. The combination of the geographic information technology and mathematical methods creates a new way to study the problem. This paper develops a model using the "Circle-Covering", "Center of Mass" and "Strength of Information" schemes respectively to generate different Geographical Profiles. It is useful for allowing law enforcement to find the residence of offenders more efficiently. Furthermore, in order to stop serial crimes and arrest the offenders quickly, the model is optimized by integrating "Center of Mass" and "Strength of Information" schemes to predict the next crime site based on a Simulated Annealing algorithm, and a corresponding search objective function is proposed. Finally, the optimized model is tested by the case named "The Yorkshire Ripper", in which final values of the search objective function calculated do not change sharply, indicating that the model optimization is reasonable.

Introduction

Due to the rapid development of science and technology, frequencies of crime, especially serial crimes, are also rapidly increasing with crime methods more diverse than ever before. This result in cases which are not only difficult for police to crack, but consume enormous amounts of police resources, costing millions of dollars in the process. Some serial killers treat murder as a hobby [1,2], which makes great threat for people's property and life security, and leaves them in a panic. The police deploy every available resource to stop the killings and bring them to justice, but because of the paucity of relative data on criminals and their unique geographic behavior the work for law enforcement officers are more difficult. Statistical prediction methods is widely used for other types of serial crimes such as burglaries [3] or auto thefts [4], but are less useful for finding serial killers, as these methods generate predictions which rely on a large amount of initial data. Thus, the less information containing criminal data and their unique geographic behavior to investigate, the more difficult is the serial killer's future behavior and location to track. Data on the next crime sites is urgently needed by the law enforcement agencies to help them find the residence of the offenders quickly and bring them to arrest.

There are some connections between serial criminals which provides the direction of analysis and forecasting. Crime mapping is a profiler based on the geo-spatial information, operation and handling of criminal data that can be display visually and output for analysis [5]. There is rarely research on this study of serial crimes which combines geographic information technology and mathematical modeling together. This paper studied the serial criminals based on The Time Series, and a prediction is developed to investigate residence of the offenders by three different schemes, which are named the "Circle-Covering" scheme, "Center of Mass" scheme and "Strength of Information" scheme. Then, the paper combines the "Center of Mass" Scheme and "Strength of Information" scheme together, and Simulated Annealing algorithm [6] is used to predicted locations of the next site based on the location information and the crime's time from the previous attack. In order to verify the rationality of the model, the study used the case named "The Yorkshire Ripper"

[1] by computer simulation. Several tests were demanded to increase accuracy of the results which the study predicted.

Geographical Profiling

The model is a new detection technology named Geographical Profiling based on the theories of environmental crime and associated research. It tries to obtain the offender's geographic behavior or location of crime-related geographical context in the process of crime analysis, it's usually used to speculate the possibility of criminals residence, or even re-estimate the time and location of crime by the main venue for a series of detailed analysis of the crime. It tries to rebuild offenders and victims of crime before and after the feasible distance of crime, when the activities of criminals found a suitable space for the murder of an object passing these places is often the most crime-prone position. Investigation of this new assistive technology will cover indices such as the distance decreasing, buffer zone, and an average distance committing the crime. In order to generate the Geographical profiling, there are three models which are given below.

The feasible distance of crime. In the simplest case, offenders' residences are located in the centre of their crime patterns and can be found through spatial means. Experience shows that serial killers tend to operate near their homes, in areas which they are familiar with [1]. For example, one study of 126 U.S. serial killers found 89% of them lived within a circle drawn around the location where they disposed of bodies [7], the inclination of a serial killer to work close to home base or anchor point often leads to a spatial clustering of their crime sites [1].

Environmental criminology establishes a framework within which crimes mobilize research, centre diagram, and other geographic principles may be combined to create a method for determining offender residence from crime locations. There are two different feasible distances about the offenders' residence based on these results and theories above: the invariable distance and the changeable distance that showed by the diagram in fig.1 and fig.2.

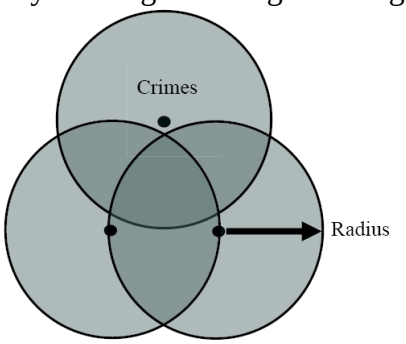

Fig.1 the invariable distance of the crime's journey

Figure 1 shows a Venn diagram for three hypothetical serial arsons. The medial circles surrounding each crime location are defined by a radius equal to crimes mobilize distance *d*, within the range of which percentage ρ of the offender's arsons occur (d=2 miles; ρ=0.70). The probability of the offender's residence lying within the area circumscribed by a single circle is therefore also ρ. Because the crimes are connected, the half-moon areas between any two circles are more likely to contain the offender's residence. The highest probability is in the middle region where all three of the circles intersect.

The other figure shows that the offenders' activities radius is changeable. Actually, their activity radius is difficult to determine because of the crimes mobilize is difficult to be forecast by the law enforcement agencies. To solve this problem, an algorithm takes the present crime site as the center, the distance from the present crime site to the site before it as radius to draw a circle. Start from the location of the crime, analogously, we can get a map full of circles, the areas overlap most has highest probability to be the target we look for. It's implemented as follows:

- Seek the second crime site base on the principle of Time Series on the map.
- Take the second crime site as the center, the distance to first crime site as the radius.

- Draw a circle and then move to next site (the third).
- Return to step 2.

The results are showed in fig. 2 following this rule:

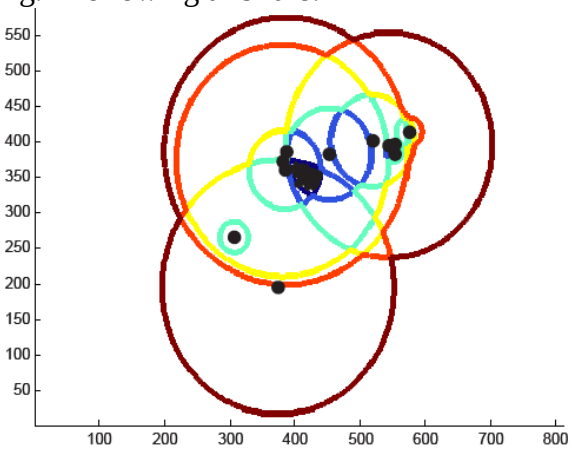

Fig.2 the changeable distance of the crimes mobilize

To test our model, the example is a case named "The Yorkshire Ripper" which consistent with the rule above. The diagram based on the map given on the internet and a MATLAB program shows the Geographical Profiling in fig.3. And the shadow area should be draw our attention to ponder and analysis for the scheme to predict the serial killer's future behavior and the location data on his next crime site or the residence of the killer's.

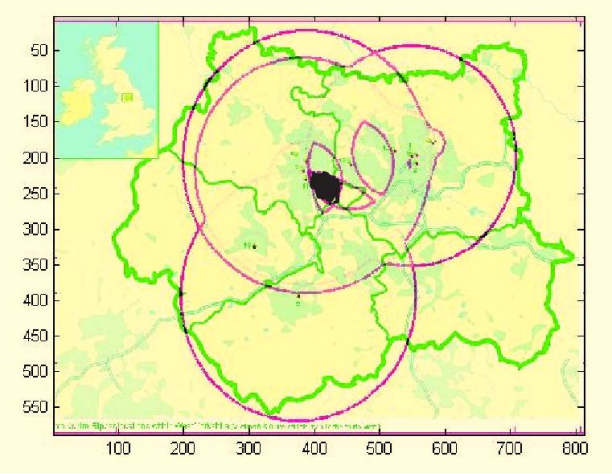

Fig.3 Geographical Profiling of "The Yorkshire Ripper"

Center of Mass. 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 [8]. Here we attempt to generate Geographical Profiling with this method. The definition we set to "Center of Mass" is a point which the sum of the Manhattan Distance [9] from it to each crime site is minimum.

Most blocks, towns, cities and states are arranged as approximate square. In order to be similar to the real world, the distance we took is Manhattan Distance:

$$D(i)=\left|X-x(i)\right|+\left|Y-y(i)\right| . \quad (1)$$

Where $x(i)$ and $y(i)$ are the Coordinate of ith site, X, Y are the Coordinate of "Center of Mass". $D(i)$ is the Manhattan Distance from "Center of Mass" to *ith* site.

To search for the "center of mass", we have:

$$D\_sum\sum_{i=1}^{i=n} D(i) . \quad (2)$$

Then X and Y can be worked out easily through Lingo11. Then the town contains the "Center of Mass" has high probability to be the offender' residence. So this method could be implemented as

follows:

1. Collect coordinates each crime site.
2. Build the equation to minimize the total distance.
3. Work out the Coordinate "Center of Mass".

We also test this method on the case of "The Yorkshire Ripper":

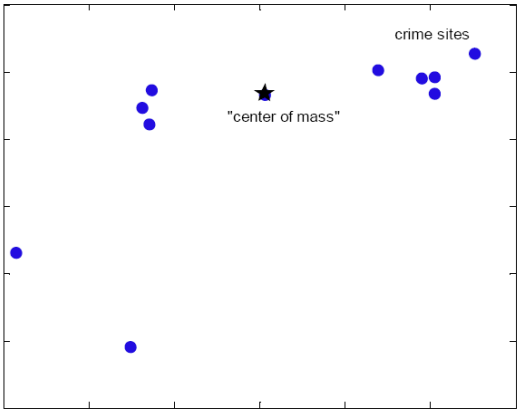

Fig.4 "Center of Mass" measured by Manhattan Distance

Figure 4 shows that the town is near the area Circle-Covering Method provided and the possible of the residence of the offenders is marked by pentagonal sign based on the test case. More important, these areas that our methods provided are fit to the theory that "the more far away from the residence, the less probability a crime happened".

Crime Information Search Scheme. There are complex factors affecting the offenders' behaviors which give the law enforcement agencies a large challenge. In another point of the view the density of the potential targets to the offenders indicate a highly possibility which in a certain extent reflects the offender's mental activity or criminal motives. Thus, the paper develops a scheme based on these theories which can indicate the crime information.

The large area was divided into a serial grid that represent the crime happened in our study. Then we mark each grid based on each factor. This scheme focuses on this factor: the density of the potential targets to the offenders.

And a rule corresponding to above is given following:

In the gridding, each grid will be given a value indicate the crimes' information which have disposed by digitization, the denser the potential targets is, the stronger the "thrill" to the offenders is.

There is different density in different areas that represent the cases. The study is hardly master the detailed crime information about the case, thus, the sample data given by MATLAB in the study which tested the scheme by fig.5.

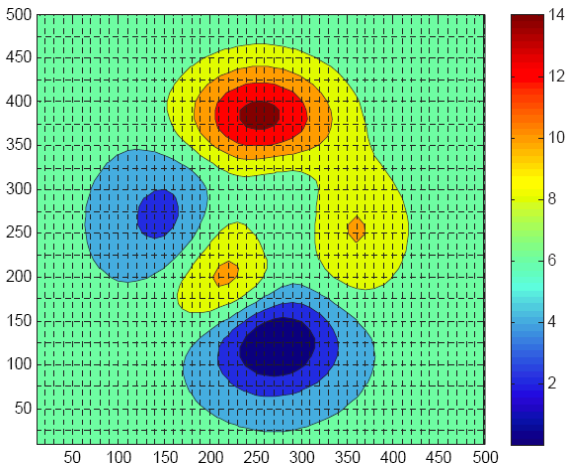

Fig.5 The gridding of Crime Information

Figure 5 shows that the deeper the color is, the higher the density of grid, and the "hot areas" should be paid attention for the law enforcement agencies to analysis how or where to seek out the offenders or predict the location information of the next crime sites and arrest them to justice. The "hot areas" based on fig.5 will be marked by a square shows in fig.6.

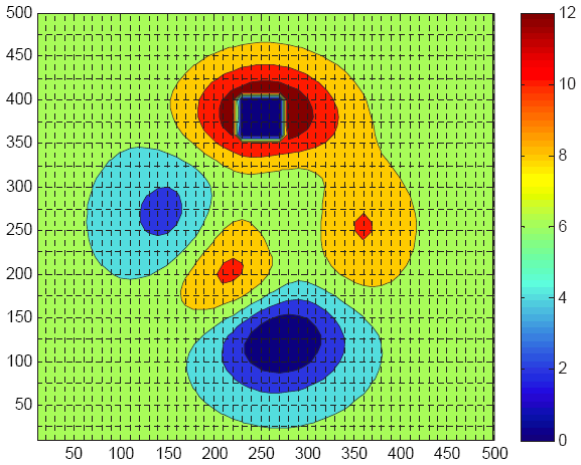

Fig.6 The "hot areas" of crime

The figure shows that the law enforcement agencies will seek out and arrest the offenders quickly by the detailed information which they have mastered combine with the scheme suggested by the study.

Prediction and Simulation

Geographical Profiling is generated by each model of the study discussed, but it cannot predict the residence of the offenders or the location of the next crime site accurately because of the law enforcement agencies who are limited about the information on the offender's mental activity or criminal motives. For example, Peter William Sutcliffe (born 2 June 1946) the protagonist of a case serial killings, no one have a view that the British serial killer who was dubbed by the press "The Yorkshire Ripper" during his crime spree. In 1981 Sutcliffe was convicted of murdering 13 women and attempting to murder seven others [1].

So a technique is needed, which combines the results of the different scheme and generate a useful prediction for law enforcement agencies. A better combination should pay attention to two or more scheme which can contain their strengths, avoid their weakness. Based on the models of the study have discussed in Geographical Profiling, the paper combined the "Center of Mass" Scheme and Crime Information Search Scheme together which turned the predict problem into a search problem. And it also combines the factors affect serial criminals together to search the possible residence of the offenders or the locations of the next crime site. And the prediction of the new model can be implemented as following:

- The crime areas are divided into grids and each grid represents a possible crime site;
- Label the crime sites which happened past based on the Time Serial;
- Work out the "Center of Mass" by Manhattan Distance;
- Collect the information of the density of the potential targets;
- Start the searching program to simulation by computer;
- Pick out the grids which represent the sites where the crime happened and plot them on the map.

The Simulated Annealing Algorithm. In the study, an idea of Heuristic Algorithm is considered into the process of simulation named Simulated Annealing Algorithm. Which is a stochastic optimization algorithm for solving stochastic optimization based on a Monte-Carlo iterative strategy, and it starting point is based on the similarity of the annealing process of solid matter in physics and

general combinatorial optimization issues [10]. The algorithm is recognized a time-varying and the probability of jumps tends to zero ultimately by giving the search process, thus can avoid local minima effectively and eventually become a global optimization algorithm in serial structure.

In this paper, the hit score is measure of Geographic Profiling search effective which is defined as the ratio of the area searched before the offender's site is found to the total hunting area. And the algorithm searches the areas of grids by computer simulation with a high "score" which is described as the following:

$$s(i) = \frac{\rho(i)}{r(i)^2}. \tag{3}$$

Here, $s(i)$ is the score of *ith* grid;

$r(i)$ is the Manhattan Distance to the "Center of Mass" of *ith* grid;

$\rho(i)$ is the density of the potential targets of *ith* grid.

Then the total "score" of a search area is:

$$S(k) = \sum(s(j)). \tag{4}$$

Here, $s(j)$ is the "score" of grid contain in the area. The maximize value of the group of the areas is the target for a high "score" to hit, and the areas' number will be determinate actual situation. And this predication model's Objectives Function as:

$$TotalScore = \sum(S(k)). \tag{4}$$

Here, $S(k)$ is the *kth* area "score".

Based on the idea of Simulated Annealing Algorithm, the simulation of the study following as:

Initialization (Current_ solution, Temperature)
Calculation of the Current_ Cost
While (Judge Method)
Select a New_ State from the Neighborhood of Current_ State
Calculation of the new_ Cost
IF Δ (Current_ cost - New_ Cost) ≤ 0
Current_ State= New_ State
ELSE
IF exp (Current_ cost-New_ cost)/ Temperature >Random (0, 1)
Current_ State=New_ state;
END
END
Decrease the temperature
EXIT When STOP_CRITERION
END

Fig.7 The computer simulation program of Simulated Annealing Algorithm

Simulation and Results

The model which combines the "Center of Mass" Scheme and Crime Information Search Scheme together is according to the assumptions following as:

1. There is only place in every case, namely, didn't consider the scene that the criminal location in one case;
2. The offender has a residence, his/her home or just a place he/she often goes to;
3. Most cities existed criminal hotspot area;
4. There are at least three crime site have being found, because this is an actual serial offenders.

The paper was tested to simulate by a group sample data from MATLAB as well as the criminal locations case of "The Yorkshire Ripper", and the Geographical Profiling is showed in Fig.8:

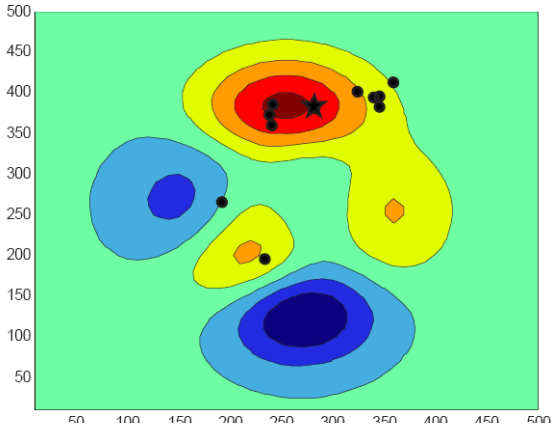

Fig. 8 Final Geographical Profiling

First, in this figure, a group of circle areas were set as the search areas in the test and the program can generate a moving process for searching which showed in Fig.9, the states that indicate the possible interrelate information of the offenders were marked with a circle.

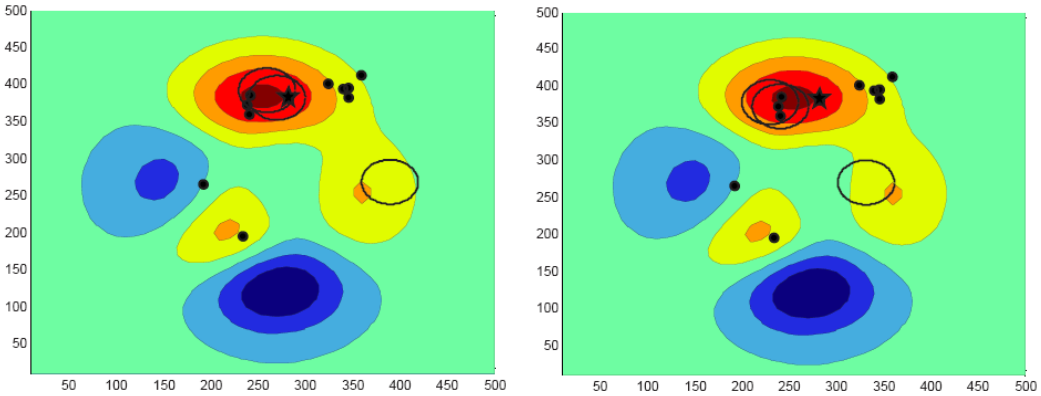

Fig. 9 Glimpse of the Searching Process

Secondly, form the figure above most search areas can be finding located near the profile with dark color. As we initialize the search points half-randomly, this search method is also effective. Finally, the values of the Objective Function were calculated 20 times continuously, and a figure was plotted to show the change of the values in fig.10, the tendency of the line suggest the information about the values of the Objective Function doesn't change acutely.

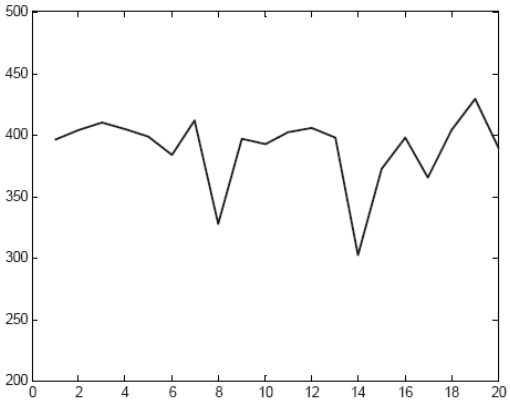

Fig.10 Change of the objective function

The Objective Function value of 20 tests is shown in Table 1, and the standard deviation of these values is 29.1742 which is acceptable. Which indicated the method is effective and reliable in a given situation based on the paper.

Table 1 The values of Objective Function

Order	1	2	3	4	5	6	7	8	9	10
Value	396.1617	403.6148	410.1070	404.7207	398.8484	383.4062	411.5952	327.9466	397.2238	392.3982
Order	11	12	13	14	15	16	17	18	19	20
Value	402.2021	405.4790	397.5546	302.6271	372.2278	397.9050	365.6077	404.0927	429.7628	389.2576

5. Conclusion

According to the characteristics of serial criminals, the paper developed the "Circle-Covering" scheme, "Center of Mass" scheme and "Strength of Information" scheme, and generates the appropriate Geographical Profiling for the law enforcement officers, which is provide a feasible method to seek out the residence of offenders quickly based on the ideas of Time Series. What's more, the model is optimized by integrating the "Center of Mass" scheme and "Strength of Information" scheme in together, which can research and analysis the possible location of the next crime site or the residence of the offenders based on the theory of Simulated Annealing method that can find the optimal solution by computer simulation. And the case named "The Yorkshire Ripper" is used to validate the model, at the same time the validation results were evaluated, which indicating the model optimization is reasonable. The paper give a new way to study serial criminals by combining the geographic information technology and mathematical principles, and the crime prevention measures will be taken for law enforcement agencies to specific populations, and also give a strong theoretical support for the arrest.

References

[1] Information of Peter Sutcliffe on http://en.wikipedia.org/wiki/Peter_Sutcliffe.

[2] F. B. of Investigation, Serial killers in the United States, 1960 to present 2003.

[3] Bowers K J, Johnson S D, Pease K. "Prospective Hot-Spotting The Future of Crime Mapping?" *British Journal of Criminology*, Vol.44, No.5, pp. 641-658(2004).

[4] Tonkin M, Woodhams J, Bond J W, et al. "A theoretical and practical test of geographical profiling with serial vehicle theft in a UK context." *Behavioral sciences & the law*, Vol.28, No.3, pp. 442-460(2010).

[5] Fenghua SUN, Aihua MAO. "Study on the theories of criminal geography." *Human Geography*, Vol.18, No.5, pp. 70-74 (2003).

[6] Zhao Xinchao. "Simulated annealing algorithm with adaptive neighborhood." *Applied Soft Computing Journal* , Vol. 11, No. 2, pp. 1827-1836(2011).

[7] S. Lundrigan and D. Canter, "Spatial Patterns of Serial Murder: An Analysis of Disposal Site Location Choice," *Behavioral Sciences and the Law*, Vol.19, No. 4, pp. 595-610 (2001).

[8] Xiao-long Liu Yuan-Biao Zhang Ji-xuan Han Sha Zhao, "Finding the locations of offenders in serial crimes". 2010 International Conference on E-Health Networking Digital Ecosystems and Technologies (EDT), Shenzhen, China, Apr. 17-18, pp. 240-243(2010).

[9] Kennedy, Dane. *Using Manhattan distance and standard deviation for expressed sequence tag clustering*. Electronic Theses and Dissertations (ETD), 2011.

[10] The principle of Simulated Annealing Algorithm on http://baike.baidu.com/link?url=R69wFYbdR5uVKJ7nxq8GYI5WFf3S2E5dmWVhz5Ytx2Vxbk228lS2rlAdjfW45ESG.