Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(3):1-7

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Criminal incident prediction based on geographical profile

Sun Liping[1], Yan Shaohong[2], Peng Yamian[2] and Yang Aimin[2]

[1]Mathematics and Information Science College, Zhengzhou University of Light Industry, Zhengzhou, Henan, China
[2]Hebei United University, Tangshan, Hebei, China

ABSTRACT

Methods in committing serial crimes differ but not entirely random. By generating "Geographical profile", we tend to assist in crime investigation. By adopting model based on distance-decay principle using CGT and modified Centro graph principle, the article adopts analytical hierarchy process (AHP) to determine the living area of criminals and certifies the accuracy of the model through Peter Sutcliffe case. Eventually, it indicates the prediction coincides with the actual situation well. Based on the criminal mentality, the article predicts the possibility of next crime scene from space tendency (both the distance between the living area of criminal and the crime scene and the adjacent crime scene) .As the prediction is far away from being admitted, we choose another perspective to generate a transition density model to solve the same problem in a more effective way. The model indicates the Bayesian process. With the estimate of the parameters, we gain some useful information to forecast the future crime based on the past criminal locations and the time. In the end, the article verifies the efficiency of the model through a specific situation.

Key words: distance-decay principle, space tendency, transition density mode

---

INTRODUCTION

Police investigations of serial murder, rape and arson can be assisted by a geographic perspective on the spatial behavior that led to the crime scene. Law enforcement officers narrow down their search of criminals through "Geographical profile" and predict the scene of next crime to prevent it from happening [1, 2].

Geographical profiling: Geographical profiling is an information management strategy in investigating serial violence crimes. It is widely used in the investigation of serial murder, rape, arson etc. to determine the possible living areas of criminals based on crime scenes [3-5].

The motivations of serial criminals can be classified as affection and profit. The possibility of the first kind of crime becomes lower as the criminal commits the crimes farther from his/her living space. While the possibility of the second kind of crime becomes higher and then lower as the criminal commits the crimes farther from his/her living space. Seen from the spatial distribution of crime scenes, crime scenes can be classified as clustering, dispersion and mixture styles [6-9].

The crime process is a dynamic one. At present, the calculations on criminal movement are as follows: (1) average movement distance; (2) medial circles; (3) triangle areas; (4) distance-decay principle; Clues derived from crime location and place can be of significant assistance to law enforcement in the investigation of repetitive offenses. The probable spatial behavior of the offender can be derived from information contained in the known crime site locations (e.g., encounter/apprehension sites, murder scenes, body/property dump sites), their geographic connections, and the characteristics and demography of the surrounding neighborhoods. Determining the probability of the offender residing in various areas, and displaying those results through the use of choropleth or isopleth maps, can assist police efforts to apprehend criminals. This information allows police departments to focus their investigative activities, geographically prioritize suspects and concentrate saturation or directed patrolling efforts in those zones where the criminal predator is most likely to be active. Such investigative approaches have been termed geographic profiling (Rossmo, 1995) or geoforensic analysis (Newton and Newton, 1985) [10, 11].

The article basically solves the following problems:
(1) Propose two schemes based on distance-decay principle and refined centrograph
(2) Predict the living area of criminals and generate geographical profile applying analytical hierarchy process
(3) Generate possibility model of next crime scene based on criminal mentality and spatial tendency
(4) Verify the result of the model by using specific cases
(5) Evaluation and refine of the model
(6) An executive summary

THE MODEL

The effort to develop such a model has led to the creation of a computerized spatial profiling model called criminal geographic targeting(CGT).Criminal geographic targeting, by analyzing the spatial information associated with a series of linked crimes, attempts to determine the most probable areas in which the offender's residence might be located.

Assumption

Without outer interference (preventative forces from crime prevention, the strengthened crime awareness from residents, preventing and controlling strategies from the police), the style of criminal does not change. There must have been no more than two offenders involved in the case. More than one offender can bring the influence of two residences, two activity spaces, and two mental maps into play, confusing the analysis. While it was possible to address this difficulty for two killers, three different offenders creates potentially too complex a situation. Criminals search their target from a certain point and follow the distance-decay principle. There is a buffer zone between the living space and the crime scene. Within the buffer zone, as the distance becomes farther, the possibility of crime becomes higher. Out of the buffer zone, the possibility of crime becomes lower. The distance between two points is Manhattan distance. Generally speaking, streets of a town are north-south and east-west, so the most reasonable distance between two points is Manhattan distance. The murder must have commenced in 1960 or later. killer from earlier time periods might have behaved in geographic manner differ from what is commonly seen today. Advances in transportation, rapid transit, and automobiles, and other factor are likely to have influenced spatial behavior. Furthermore, many cities have significantly changed and grown over the years and it would be difficult to obtain accurate period street maps. There must have been at least 5 different location of the same type associated with a single offender residence. This condition is related to criterion number 2, but while the stipulation of a minimum of 5 victims usually translates into 5 different locations, there are several cases in which the murderer attacked more than one person at a time. The serial killer must not have been nomadic; in other words, the offender must have had a residence. There must have been suitable and sufficient information available on the addresses or locations of the crimes, and the offender's activity space, for a spatial analysis. This requirement is related to criterion number 5. Criminals generally commit crimes in areas which are familiar or similar to the neighborhood where they live. Criminals live in a certain place, and their area of activities can be predicted. There must have been a minimum of 5 murder victims. The random fluctuations caused by a small number of crimes affect the predictive power of any statistical model designed to locate offender residence.

The two schemes generating a geographic profile

Scheme one:Geographical profile based on distance-decay principle. Brantingham and Brantingham developed the spatial analysis of crime conceptually. According to them, an officer is not likely to commit a crime very near to the home base creating a so-called buffer zone resulting in increasing likelihood of a crime being committed until a range is reached where the probability of crime is the highest. After this zone, the number of crime trips would decrease as a function of distance creating a distance-decay function. Based on distance-decay principle, we tend to rely on a resultant score function to predict the location of criminals, the higher the resultant score, the greater the probability that the point contains the offender's home or workplace.

The resultant score function is presented by ROSSMO'S FORMULA:

$$P_{ij}^{D} = k\sum_{c=1}^{T}\left[\phi/(|x_i - x_c| + |y_j - y_c|)^f + (1-\phi)(B^{g-f})/(2B - |x_i - x_c| - |y_j - y_c|)^g\right]$$

in the formula, The first part of the formula expresses the idea that the probability of crime locations decreases as the distance increases, once outside of the buffer zone. And the second fraction indicates smaller answers as the distance increases. (Solving Crime with Mathematics—Keith Devlin, Gary Lorden) Where:

$$|x_i - x_c| + |y_j - y_c| > B \supset \phi = 1$$
$$|x_i - x_c| + |y_j - y_c| \le B \supset \phi = 1$$

And: P_{ij}^D is the resultant probability for point ij; ϕ is a weighting factor; k is an empirically determined constant; B is the radius of the buffer zone; T is the total number of crime sites; f is an empirically determined exponent; g is an empirically determined exponent; x_i, y_j are the coordinates of point ij ; x_c, y_c are the coordinates of point ij ; Model verification: We verifies the Peter Sutcliffe murder case below:

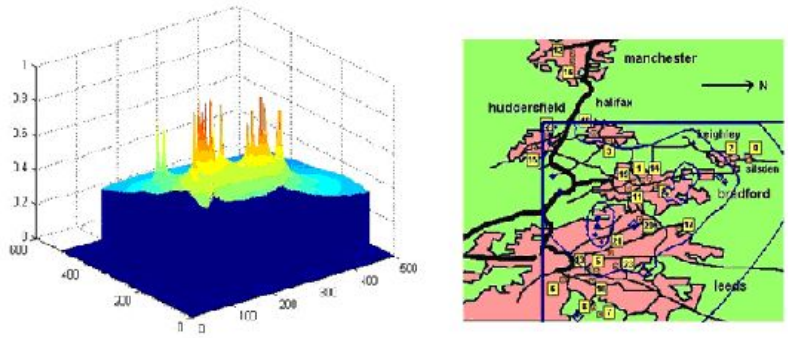

Figure 1: Isopleths and Chloroplast Map-Peter Sutcliffe murder

The possibility of Peter Sutcliffe's living area can be two places pointed in Figure 1. From the information collected, he usually committed crimes near his living area. His living area can be two partly because he has moved in 1977, or the accuracy of model parameter, or lack consideration of relevant factors.

Scheme two: Prediction of Centro graph and its refine scheme. Spatial mean is a single variable to evaluate the endocentric tendency of crime scene. It is mainly sued to analyze the pattern of crime scene. The center of gravity scales down the distribution of serial crime and projects the central areas. Centro graph scheme is widely used in crime investigation and detection. It was used in the investigation of San Doego Rape case and the hillside strangler case. The calculation equation of spatial mean is $SM_x = \left[\sum_{n=1}^{c} x_n\right] / C$, $SM_y = \left[\sum_{n=1}^{c} y_n\right] / C$

Where SM_x coordinate of x axis of spatial mean. SM_y coordinate of y axis of spatial mean. C sum of crime scene. x_n, y_n coordinate of nth crime scene. The calculation equation of standard distance $Sd = \sqrt{\left(\sum r_{ns}^2\right) / C}$.

Where Sd standard distance. C sum of crime scene. r_{ns} distance from the standard distance to the second crime scene.

Spatial mean proves very helpful in investigation crimes but has three difficulties.(1)It generally provide limited information. (2) Because of the uncertainty of crime, the results are not effective. (3) Highlighting information of criminal's range of activity and crime purpose; the result proves unrelated to the center of analysis. If a certain criminal's range of activity does not center on his/her living area or he/she does not commit crimes in a certain pattern, the spatial mean can be irrelevant to the living area of the criminal.

Because of deficiencies of spatial mean, the following equation was proposed by a research $y = ax + b$.

Where y is the largest distance from the criminal living area to the crime scene in miles. x is the largest distance between crime scene in miles. a is endocentric deviation from crime scene and criminal's living area(a is 0.5 if it is endocentric) researches find out different types of crimes and crime sconces can affect the value of a and b.

If the time and crime scene of serial violence crime are known, the most possible living area of the criminal can be determined by the area with the crime scene as the center and overlapping areas of square with Manhattan distance of y. S_{ij} possibility of being the living area can be determined by the times overlapped.

Verification of the model: The research towards to England murders proves the values of a and b are 0.79 and 0.61. Verification of Peter Sutcliffe serial murder case blew：

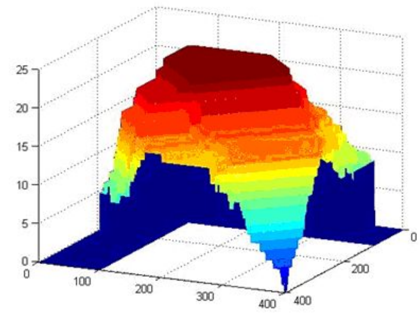

Figure 2: Probability Map-Peter Sutcliffe murder

Obviously, Predicted area of the maximum probability contains the location of residence. The uncertainty caused by large predicted areas increases the difficulty of the police investigation.

Generate a geographical

Human performance on the geographic profiling task is to predict an offender's home location from their crime locations. The above two schemes determines the living areas of criminals from different angles. Given the consideration of accuracy, the article predicts the distribution possibility of criminal's living area:

$$P_{ij} = \alpha P_{ij}^{1} + (1-\alpha) P_{ij}^{2}$$

Where α is the relative accuracy of scheme one over scheme two. P_{ij}^{1} is the possibility of criminal's living area based on scheme one. P_{ij}^{2} is the possibility of criminal's living area based on scheme two.

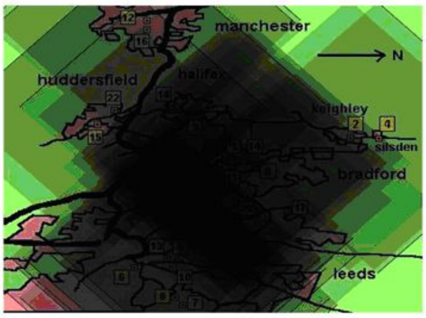

Figure 3: The probability map

From the above analysis, we can see Peter Sutcliffe usually committed crimes near his living area. He is the affective type and tends to commit crimes in focused areas.

The location prediction of the next crime

Assumption: Criminal could make rational choice when he/she committed crimes and could evaluate relevant risks. Criminal's possibility of committing next crime in neighboring communities is low. The possibility of next crime scene can be predicted by the distance of criminal's living area and last crime scene.

Law enforcement agencies need crime forecasts to support their tactical operations; namely, predicted crime locations for the next time base on date from the previous case. Nearest neighbor distance analysis: The process of choosing criminal target is a dynamic territorial displacement process. The choice of crime scene can be affected by the daily activities of criminals, their crime style and their rational choice. Obtaining serial violence crime data record from geographic information system, the article generates geographical profile. The article tends to analyze criminal's mentality and predict next crime scene based on spatial tendency (the distance between criminal's living area and crime scene and distance between criminal scenes). It's criminal's mentality that he/she is more likely to commit crimes in areas he/she is familiar with and search target out of his/her neighbor following distance-decaying principle. That is to say the crime possibility of a certain place may change with its distance from criminal's living area accordingly. The possibility may become higher then lower as a rule. In addition, two succeeding crime scenes may not be very close because of the strengthening safety awareness of residents in the community or preventive and controlling strategies from the police.

Suppose the criminal spatial tendency style can be predicted by serial crimes and distance between criminal's living area and predicted crime scene is irrelevant to the distance between two succeeding crime scenes, the possibility of next crime scene is $P_{ij} = P_{ij}^{H} * P_{ij}^{L}$.

Where P_{ij} is the possibility of next crime scene at the place S_{ij} . P_{ij}^{H} is the possibility of next crime scene based on criminal's living area. P_{ij}^{L} is the possibility of next crime scene based on criminal's last crime scene.

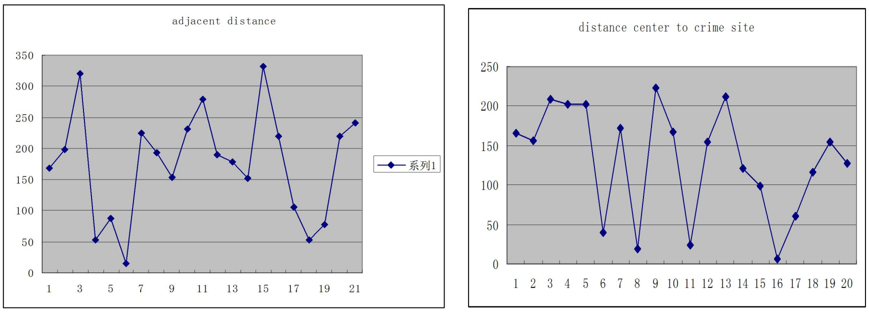

Figure 4: The distance center of crime site

Verification indicates the model is unavailable. The transition density model: Considering that the prediction above brings a bad result, we develop another model based on the transition density. To formally describe the forecast problem, we note the locations and times of criminal incidents as $(s_1,t_1),(s_2,t_2),\ldots..,$ $t_0 = 0 < t_1 < t_2 < \ldots$ where si is the two-dimensional location of incident i of a given crime type and ti is the corresponding time.

The quantity of interest is the density of the process, which formally captures the likelihood that a future criminal incident occurs within a study region and a study horizon, given the times, locations, and feature values of past incidents of the same type bounded by the same region and time range.

Let $T_n = \{t_1,t_2,\ldots, t_n\}$, $D_n = \{s_1,s_2,\ldots, s_n\}$, and $\chi_n = \{x_1, x_2,\ldots, x_n\}$ where $s_i = (s_{i1},s_{i2})$ The transition density is defined as follow:

$$\psi_n(s_{n+1} \mid D_n,\chi_n,T_n,t_{n+1}) = \alpha\ \psi_n^{(11)}(x_{n+1} \mid \chi_n)\sum_{j=1}^{C}\psi_n^{(12)}(s_{n+1} \mid D_n^{(j)},T_n^{(j)},t_{n+1})\Pr\{x_{n+1} \in \chi^{(j)} \mid \chi_n^{(j)}\}$$

Where $\psi_n^{(11)}(x_{n+1} \mid \chi_n)$ is called the first-order spatial transition density which reflects the event intensity (i.e. first-order effects)at x in the key feature space; $\psi_n^{(12)}(s_{n+1} \mid D_n^{(j)},T_n^{(j)},t_{n+1}), j = 1,2,3\ldots.C.$ are called second-order spatial transition densities which describe the interaction of a new event location s_{n+1} with past event locations in each $D_n^{(j)}$, respectively; $\Pr\{x_{n+1} \in \chi^{(j)} \mid \chi_n^{j}\}, j = 1,2,\ldots,C,$ are called spatial interaction probabilities which are the probabilities that x_{n+1} falls in the same continuum $\chi^{(j)}$ of the key feature space as α is a normalizing factor.
We postulate the following function for the second-order spatial transition density for cluster j

$$\psi_n^{(12)}(s \mid D_n^{(j)},T_n^{(j)},t) = \varphi_m(s \mid s_{1,\ldots,}s_m) = \frac{\lambda^2}{2\pi m}\sum_{i=1}^{m} e^{-\lambda\|s-s_i\|}$$

Where $t > t_m$ is a future event's time of occurrence and $\|s - s_i\|$ the distance from that future event's location s to an older event location s_i $(i = 1,2,...,m)$. This is called an order model since only the temporal order of the events is considered. The instant model actually utilizes the values of the series $t_1, t_2,..., t_m$. Based on this model, we postulate that the second-order spatial transition density for cluster j takes on the form

$$\psi_n^{(12)}(s \mid D_n^{(j)}, T_n^{(j)}, t) = \eta_m(s \mid s_{1,......,} s_m, t_1, t_2, ..., t_m, t)$$
$$= \frac{\lambda^2}{2\pi \sum_{i=1}^{m} e^{-\tau(t-t_i)}} \sum_{i=1}^{m} e^{-\lambda\|s-s_i\| - \tau(t-ti)}$$

We can numerically solve for the maximum likelihood estimates of the parameters λ and τ.

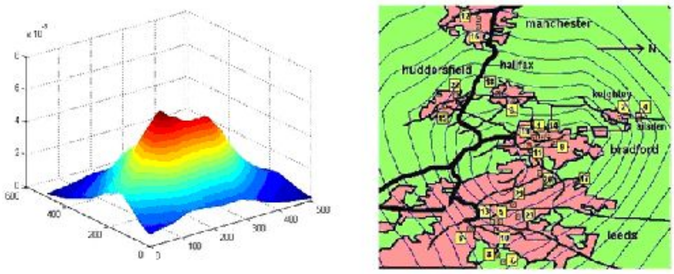

Figure 5: The possibility of 23th crime

Figure 5 shows the most possible location of the next time is near position 23, the actual position where the 23th murder happened.

STRENGTH AND WEAKNESS

Strength

Based on distance-decaying principle and crime scene rule, the article generates two schemes. The predict results are the same with reality thanks to model verification. In consideration of criminal mentality, the article adopts criminal spatial tendency and criminal site selection preference and obtain possibility model of next crime scene. The transition density model, Based on the Bayie process, makes use of locations, times and features of all incidents.

Weakness

The article uses empirical data which need large amount of statistical analysis of cases to determine. Because of the difficulty to search data, the model obtains data by computer search and selects the optimal data but the accuracy of the final results may as well be affected. There may be many personal and social factors affecting criminal's style in committing crimes. Based on the difficulty in generating the model, the article selects only some of the most representative ones. Of course, the more factors considered, the most accurate the results.

Acknowledgements

We thank anonymous reviewers for helpful comments. This research is partially supported by the National Natural Science Foundation of China (No. 61170317) and the National Natural Science Foundation of Hebei Province (No. E2013209215).

REFERENCES

[1] Fenghua Sun, Shitai Li, Liping Hua. *China academic journal electronic publishing houses*, **2006**, 21(5).
[2] Hua Liu, Donald E. Brown. *International Journal of Forecasting*, **2003**, (19), 603–622.
[3]Wan Hua-zhe. *journal Of Nanchang Junior Coll*ege. **2010**, 3, 154-156.
[4]Li Ke. *Journal of Shenyang Sport University.* **2012**, 31(2), 111-113.
[5]Zhang Shu-xue. *Journal of Nanjing Institute of Physical Education*. **1995**, 31(2), 25-27.
[6]Pan Li. *Journal of nanjing institute of physical education(natural science).* **2004**, 19(1), 54-55.
[7]Li Yu-he, Ling Wen-tao. *Journal of Guangzhou Physical Education Institute*. **1997**, 17(3), 27-31.

[8] Xu Guo-qin. *Journal Of Hebei Institute Of Physical Education*. **2008**, 22(2), 70-72.
[9] Chen Qing-hong. *China Sport Science and Technology*. **1990**, 21(10), 63-65
[10] Tian Jun-ning. *Journal of Nanjing Institute of Physical Education*. **2000**, 14(4), 149-150.
[11] Zhang B.; Zhang S.; Lu G.. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.