Two Steps to Target the Serial Criminal

Summary

Geographical profiling is an investigative methodology to analyze the locations of serial murder’s residence. Models of this article can predict the criminal residence and the next crime site in a series criminal.

Two steps prediction technology is developed to predict the possible crime site, a. In first step, the model is used to locate the possible criminal residence; in the second step, the model is used to target the next possible crime site.

Criminal Geographic Targeting (CGT) I & II model is built to target the criminal residence. “Hit ratio”, the ratio of optimal search area divided by the total search area, is an important index to evaluate CGT models. The results of cases show that the both of our model have a “hit ratio” below 22%. Then an “If-then” technique is developed to combine the results of two different schemes.

Two models are built to predict the next crime site based on past crime site information. Risk Index (RI) Model searches the area that has the most significant impact on the estimated residence area. Then, criminal incidents are analyzed as spatial choice processes. To describe the criminal’s decision making process, we derive a Criminal Spatial Choose (CSC) Model from the Gravity Model and Feature Selection Model. Therefore, our approaches not only focus on objective model (i.e. geographical profiling), but also integrate the subjective model (i.e. psychological profiling) into criminal profiling.

Then the models are tested by a series of real cases, the sensitivity analysis illustrates that CGT model and RI model are stable while the number of crime sites is above 6.

Key words: risk index, decay function, feature selection, gravity model, decision making process, grid system
Content

Summary ........................................................................................................................................... 1
1. Introduction .................................................................................................................................... 3
2. The proposed models for generate the geographical profiling .................................................... 3
   2.1 Assumptions ......................................................................................................................... 3
   2.2 Criminal Geographic Targeting I (CGT I ) Model ............................................................. 3
3. Models of predicting the next crime scenes ................................................................................ 6
   3.1 Criminal Spatial Choice (CSC) Model .................................................................................. 6
   3.2 Risk Index Model .................................................................................................................. 8
4. Testing and Results ....................................................................................................................... 9
   4.1 CGT I & II Evaluation ......................................................................................................... 9
   4.2 RI Model Evaluation .......................................................................................................... 13
   4.3 CSC Model Evaluation ........................................................................................................ 15
5. Strengths and Weaknesses ......................................................................................................... 16
6. Future Work .............................................................................................................................. 16
1. Introduction

Geographical profile (GP) is defined as a methodology to analyze the serial murder’s residence (http://en.wikipedia.org/wiki/Geographic_profiling), and many techniques have been developed to determine it. We get two schemes to calculate the risk index that varies by location in total search area, the greater the index, the higher the probability that the area containing the residence of serial criminal. After that, we combine the results of two schemes to minimize the optimal search area.

Risk Index (RI) Model and Criminal Spatial Choose (CSC) Model are built to predict the next crime scene based on the GP. The first one is to calculate the increased risk index of the residence area (which is estimated by the CGT models) of each area within the total search area. The area brings the greatest change is regarded as the place with high possibility to be the next crime scene. CSCGM is aimed at searching for an area that the criminal can maximum their utility to meet the objectives. We rank the probability that the criminal is interested in visiting this area in the grid and then select the greatest one.

The proposed models that used to provide an optimal search area for the local police officer is presented in Section 2, the proposed models and method for predicting the next crime scenes are presented in Section 3, application to the case study is discussed in Section 4, sensitivity analysis of the models is presented in Section 5 and conclusions are made in Section 6.

2. The proposed models for generate the geographical profiling

2.1 Assumptions

1. Every criminal only has an anchor point (residence).
2. All the crime sites are in the same city, state or shire.
3. Police could find out and arrest the criminal as long as they search the anchor point area.

2.2 Criminal Geographic Targeting (CGT I ) Model

Risk index, also called hit score, is a way to evaluate the geographical profile. The greater the risk index is, the higher the probability that the serial criminal resides within an area (Wilpen L. Gorr, 2004). The purpose of this model is to calculate the risk of each area and find the place with the highest risk index.

Existing algorithms calculate the risk index by computing

\[ R(y) = \sum_{i=1}^{m} f(d(x_i, y)), \quad i = 1, 2, ..., m \] (1)

Where,

\( f(d(x_i, y)) \): the decay function.

\( i = 1, 2, ..., m \): crimes in time sequence

\( x_i \): crime sites i (including body dump sites)
$d(x_i, y)$: the distance from crime site $i$ to point $y$.

$R(y)$: the risk index of point $y$.

Many kinds of decay function are used, including linear function, negative exponential function, normal function and so on. The following function is determined as the basic function.

$$f(d) = Fe^{-\beta d}$$

where,

$F$: normalizing constant

$\beta$: parameter to be estimated

It is important to identify the value of $\beta$ in order to estimate the risk index reasonably. Actually, coefficient $\beta$ will determine the speed of decreasing of decay function, the higher the $\beta$ is, the faster the value of function decrease. Logically speaking, a criminal with serious crimes may leave the crime scene as soon as fast, and the risk index also has a strong relationship with the time from the last crime happened. The $\beta$ is defined in table 1.

<table>
<thead>
<tr>
<th>crimes conducted</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>below 5</td>
<td>0.05</td>
</tr>
<tr>
<td>5 to 10</td>
<td>0.1</td>
</tr>
<tr>
<td>above 10</td>
<td>1.2</td>
</tr>
</tbody>
</table>

And $\beta$ also could be regarded as a function of time in the form of

$$\beta = k / t, t \in (0, +\infty),$$

Where, $k$ is an empirically determined constant and $t$ is time.

As a matter of fact, how we choose $\beta$ does not affect the rank of risk index of each area, therefore it does not change the priority of the optimal search area.

After the crime scenes of a serial criminal are located, the total search area is usually a rectangular zone containing all the crime locations, oriented along north-south and east-west lines, details are defined as following:

$$y_{high} = y_{max} + \left(y_{max} - y_{min}\right) / 2(T - 1)$$

$$y_{low} = y_{min} - \left(y_{max} - y_{min}\right) / 2(T - 1)$$

$$x_{high} = x_{max} + \left(x_{max} - x_{min}\right) / 2(T - 1)$$

$$x_{low} = x_{min} - \left(x_{max} - x_{min}\right) / 2(T - 1)$$
Where,

\( y_{\text{high}} \): the y value of the northernmost boundary.

\( y_{\text{low}} \): the y value of the southernmost boundary.

\( y_{\text{max}} \): the maximum y value for any crime site.

\( y_{\text{min}} \): the minimum y value for any crime site.

\( x_{\text{high}} \): the x value of the easternmost boundary.

\( x_{\text{low}} \): the x value of the westernmost boundary.

\( x_{\text{max}} \): the maximum x value for any crime site.

\( x_{\text{min}} \): the minimum x value for any crime site.

\( T \): the total number of the crime sites.

Then the area is divided into little grids. It is suggested to start the search from the grids with the highest risk index, and then working down the list of grids sorted by risk index in descending order.

### 2.3 Criminal Geographic Targeting II (CGT II) Model

Criminal Geographic Targeting II (CGT II) Model is a classical model designed by Rossmo. It has a different total search area and decay function from the former model we built, and we will use CGT as our second schemes to approach the criminal’s residence with some changes to meet our requirements and assumptions.

The total search area will be divided into small grids (usually 10,000) as the former model, and then the Manhattan distances to each crime location could be determined.

The decay function we use is expressed as the following:

\[
p_y = K \sum_{c=1}^{T} \frac{\theta}{(\lvert x_i - x_c \rvert + \lvert y_j - y_c \rvert)^{\theta}} \frac{1-\theta}{(2B - \lvert x_i - x_j \rvert - \lvert y_j - y_i \rvert)^{\theta}}
\]

Where,

\[
\lvert x_i - x_c \rvert + \lvert y_j - y_c \rvert \leq B \Rightarrow \phi = 0 \quad \text{(9)}
\]

\[
\lvert x_i - x_c \rvert + \lvert y_j - y_c \rvert > B \Rightarrow \phi = 1 \quad \text{(10)}
\]

And

\( p_y \): the risk index we will calculate.

\( K \): the empirically determined constant, which we set it as 1.

-5-
\( B \): the radius of the buffer zone.

\( f, g \): the empirically determined exponent.

\( x_i, y_i \): the coordinates of point \( ij \), and \( x_c, y_c \) are the coordinates of the cth crime site location

Just like the former model, the choosing of \( f, g \) will determine the figure of the function, \( f=1.2 \) was selected from a gravity model formulation (Rossmo, 1987). And the radius of the buffer zone (B) was set at 0.3 miles by Rossmo. However, the variable should be changed from case to case. As we just study the cases within one city, state, or shire, according to the statistical result provided by Geographic Profiling: Target Patterns of Serial Murder (Rossmo, 1987), we chose B as 35km in the application of our model, which could be seen later.

Compared with the CGT I model, one of the advantages of CGT II is that a buffer zone is set, the risk index will: 1) if the point lies outside the buffer zone, becomes smaller the longer the distance, 2) if the point lies inside the buffer zone, becomes larger the longer the distance. (Rossmo, 1987). It is based on the fact that criminal will seldom commit a crime near the places they have done before.

### 2.4 Proposed method of generating a GP

Then “hit ratio” is used to evaluate the two models and help us find the best way to combine the results to generate a geographic profile. Risk indexes will still help us to locate the next crime scene by comparing its changing amount.

We develop an “If-then” structure method to generate a GP. The procedures of this method are described as follows.

Step1. The case study is conducted using 2 models.
Step2. If the CGT I performances better than CGT II, we choose the CGT I to predict the offender residence.
Step3. If the CGT II performances better than CGT I, we choose the CGT II to predict the offender residence.
Step4. Select a suitable model in a specific condition. Then the result of that model is determined as the GP.

### 3. Models of predicting the next crime scenes

It is not difficult to generate our first way to predict: if an area is with high possibility to be the next crime scene, it will significantly change the risk index of the murder’s residence which is located in the former model.

#### 3.1 Criminal Spatial Choice (CSC) Model

##### 3.1.1 Assumptions

(1) The criminal activity space (i.e. residence or workplace) is known or estimated.
(2) The criminal is individual.
(3) The criminal residence or workplace is stable.
3.1.2 Feature Selection Model

Within geography of crime perspective, locations of crime sites were seen to be influenced by hunting style, target backcloth, and changes in offender activity space (Donald E. Brown, Hua Liu, Yifei Xue, 2001). K.Baumgartner, S.Ferrari and G.Palermo (2008) studied the Bayesian networks for criminal profiling. Yifei Xue and Donald E.Brown (2006) used clustering to identify distinct criminals or groups of criminals who share a common preference structure.

We select the method developed by Yifei Xue and Donald E.Brown. A triplet \((F, c, s)\) is used to determine the feature selection, where \(F\) is the initial feature set, \(c\) is a criterion function defined for subsets of \(F\), and \(s\) is a subset search or selection procedure. Then we rank the scores of individual features.

The influence of additional variables, usually grouped into offender and offence characteristics, are also often explored (Darcy Kim Rossmo, 1987). Usually, it can classify into two groups as follows.

Offender characteristics: sex, race, age, prior criminal experience, and nature of home area

Offence characteristics: crime type, target area attribute, perceived level of potential

3.1.3 Criminal Spatial Choice of Gravity Model

The gravity model was proposed by Newton in 1931 and is used in various social sciences to predict and describe certain behaviors that mimic gravitational interaction. Huff (1964) proposed a Gravity Model to solve the competitive facility location problem. It’s described as:

\[
P_{ij} = \frac{V_j}{\sum V_{in}}
\]  

(11)

where,

\(P_{ij}\): the probability that customers in the district \(i\) choose the shop in the district \(j\).

\(V_j\): the utility that measures the shop \(j\) can attract the customers in the district \(i\).

\(n\): the competitor in the district \(i\).

Criminals choose possible sites in space to commit crimes. It’s similar to the behavior that customer choose a store. As the result, we consider the series criminal as customer and the distances from custom demand site to store as the distances from criminal residence to potential crime site. It’s possible for us to apply the gravity model to the possible crime site prediction. Generally, different site in the alternative set has different attractiveness to the specific criminals.

Using the selected key feature, we derive the probabilities in every grid with the distance \(d_i\) to the criminal residence as shown below.
\[
P_i = a \left( \prod_{j=1}^{S^i} d_i^j \right) \left( \sum_{j=1}^{S^i} d_i^j \right)
\]  

(12)

where,

- \( P_i \): the probability of alternative \( i \) that the criminal chooses.
- \( S^i \): the quality value of the potential crime site which contains the key features.
- \( d_i^j \): the distance from the criminal residence to the alternative.
- \( a \): the adjusted factor that fits the model to specific crime type.
- \( j \): the adjusted factor that fits key features to specific crime type.

### 3.1.4 Method Description

The discovery of preferences in space and time is important in investigation of criminals. According to the rational choice perspective in criminology, criminal incidents, like many other human initiated events, involve a decision making and choice process (R. Clarke, D, 1985). We combine the feature selection and Criminal Spatial Choice of Gravity (CSCG) Model to generate a technology to predict the possible next crime location. The procedure of this method is described as follow.

**Figure 1** flow chart of the procedure

#### 3.2 Risk Index Model

It is not difficult to generate our first way to predict: if an area is with high possibility to be the next crime scene, it will significantly change the risk index (calculated in the way of CGT II) of the murder’s residence which is located in the CGT model. The function is:

\[
\Delta R = R_b - R_c
\]  

(13)
Where

\[ R_b \] is the risk index of residence before the supposed crime area is added

\[ R_c \] is the risk index of residence after the supposed crime area is added

However, if the area is added as a crime cite, it should not change the position of the estimated residence (that means the former area with highest risk index still is still with highest risk index.

So the increasing amount of each grid (\( \Delta R \)) is need to be calculated, to those grids that change the location of estimated residence, we just set it as zero. The area with relatively higher \( \Delta R \) will be predicted to be the next crime.

4. Testing and Results

In this section, a real-world application of our proposed model is given. The case of series killer, Peter Sutcliffe, is chosen to study. To locate the crime sites and criminal residence on the rectangular coordinate system, we transferred the geographic coordinate system to rectangular coordinate system. Considering the profile of this case, we still use the Manhattan distance to calculate the distance between two different points.

4.1 CGT I & II Evaluation

Then we locate the total search area and divide it into 1,000 grids. The risk index of each grid is given. We use the past 13 crime events as input data and then predict the offender residence. The actual offender residence is chosen as the benchmark. Figure 2 is presented as the third-dimension of risk index distribution using CGT I.
Figure 2 the risk index of the case (CGT I).

Then a two-dimension contour figure is obtained in a clear way using CGT I. Figure ??? shows the risk indexes distribution of two-dimension.

Figure 3 the risk index of the case (CGT II)
Figure 4 the risk index of the case (CGT II)

Figure 5 the contour of risk index of the case (CGT II)
In Table 2, we summarize that both of two models has a relatively low hit ratio. In other words, both of the models contribute to narrow the search areas. The police just need to search 11.2% of original total search area using CGT I model and search 21.8% of original total search area using CGT II. Furthermore, in this case, CGT I performances better than CGT II. As the tool mentioned in Section 2.3 we define CGT I as a better tool, although CGT II is set with a buffer zone which is more reasonable in theory. Because of insufficient of the data, it’s hard for us to test the model. So we still recommend give the priority to the areas with relatively higher risk index during the search.

Table 2  the hit ratios of CGT I & II

<table>
<thead>
<tr>
<th>Model</th>
<th>Hit ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGT I</td>
<td>11.2%</td>
</tr>
<tr>
<td>CGT II</td>
<td>21.8%</td>
</tr>
</tbody>
</table>

Figure 6 the prior search area and actual offender residence (CGT I )
Figure 7  the prior search area and actual offender residence (CGT II )

We can find that the actual offender residence is near the prior search area in CGT I Model. The figure 6 and figure 7 shows the prior search area and actual offender residence using CGT I & II separately.

Then the Sensitivity Analysis is conducted with the different crime sites number, it shows that the prediction results of CGT is stable when the number of crime sites is larger than 5. The results are shown in Table 3.

<table>
<thead>
<tr>
<th>number of crime cites used</th>
<th>hit radio</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.5297</td>
</tr>
<tr>
<td>6</td>
<td>0.2248</td>
</tr>
<tr>
<td>7</td>
<td>0.2471</td>
</tr>
<tr>
<td>8</td>
<td>0.2154</td>
</tr>
<tr>
<td>9</td>
<td>0.2409</td>
</tr>
<tr>
<td>10</td>
<td>0.2065</td>
</tr>
<tr>
<td>11</td>
<td>0.2103</td>
</tr>
<tr>
<td>12</td>
<td>0.2046</td>
</tr>
<tr>
<td>13</td>
<td>0.1676</td>
</tr>
</tbody>
</table>

4.2 RI Model Evaluation
The grids of higher altitude are the places predicted to be the next crime cite.

Figure 8 risk index of the case (RI model)

Figure 9 contour distribution
Then a series of prediction is conducted with different number of the crime sites provided. The results are shown in table 4.

### Table 4 Prediction Results

<table>
<thead>
<tr>
<th>Number of crime sites used</th>
<th>Hit ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>22.30%</td>
</tr>
<tr>
<td>7</td>
<td>52.50%</td>
</tr>
<tr>
<td>8</td>
<td>98.41%</td>
</tr>
<tr>
<td>9</td>
<td>22.67%</td>
</tr>
<tr>
<td>10</td>
<td>33.05%</td>
</tr>
<tr>
<td>11</td>
<td>1.76%</td>
</tr>
<tr>
<td>12</td>
<td>14.77%</td>
</tr>
</tbody>
</table>

If we regard the hit ratio less than 30% is a successful prediction, then more than 50% of the prediction is successful. It seems that when the number of crime cites is larger than 8, so the prediction result is more acceptable.

### 4.3 CSC Model Evaluation

With the constraint of the expert knowledge about the criminology, we select a case that criminal has distinctive features. The case of Ted Bundy, a series school killer, is selected. He eventually confessed to over 30 murders, although the actual total of victims remains unknown. Estimates range from 26 to over 100, the general estimate being 35.

We select his crime data in 1978 and use them to evaluate the model. The distribution of the probability is shown in figure 10. We selected the first 5 crime events as the past crime events and the last crime events reported in 1978 as the benchmark. The result shows that this model targets the last crime site successfully.
5. Strengths and Weaknesses

Since the CGTs model has a low hit ratio, the search area of police will be significantly reduced. However, the parameters of Model CGT II is not estimated by a strict statistical result, it is still could be improved. Risk Index Model also achieves a high prediction accuracy in next crime cite prediction. But apparently it will just predict the area around the estimated residence, so it is unreliable while the estimated residence is not accurate. Criminal Spatial Choice Model integrated the criminal psychology factor and objective quantitative decision making process. But this model needs a expert system to generate correct psychology profile. If it fails to meet this requirement, it probably leads an unsolved condition.

6. Future Work

Geographic profiling is considered as the powerful tool to law enforcement such as police agency. With the development of Geographic Information System (GIS), it’s possible to analysis the criminals’ behaviors in both space and time. In addition, the data mining technology make the classification of the criminal’s key features into groups possible. Besides the traditional tools, we need to improve a simpler and more premise tool to complete the geographic profiling.

References

Yifei Xue, Donald E.Brown. Spatial analysis with preference specification of latent decision
Mike O’Leary, The Mathematics of Geographic Profiling, Department of Mathematics, Towson University.