Sindh Univ. Res. Jour. (Sci. Ser.) Vol.43 (1) 01-08 (2011)

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Prediction of the location of the crime, using statistical Model

F. S. SYED, S. HUANG, J. CHU AND T. CAI

Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China

Corresponding Author: S. F. Shah, Email: sf_shah@hotmail.com. Ph. No.008615201378069

Received 02nd March 2011 and Revised 04th May 2011

Abstract: It has been a crucial issue to study the psychology of a criminal. Continuous sequence of crimes becomes very harmful to the societies. Many of the researchers have been studying the psychology of the offender in order to know the crime nature, its timing and where the next crime could happen according to the psychology of the offender. In this paper, our aim is to accurately predict the location of next crime and arrest the offenders. We carry out research on the psychology of the offenders, and set a more accurate scheme, using the Theory of Hypothetical Circle of crime. This describes two types of crime path, the linear one and the circular one. The model has progressively been refined by adding more factors. On one hand, keeping in view the different geographical and cultural factors, we introduce the concept of Point of Attraction and Area of Attraction and quantify them as well to calculate their influence on the criminal possibility. On the other hand, according to the offenders' records, we introduce the concept of Concentration Index, which helps to quantify the degree of the data's concentration and thus we specify the crime location into a certain area of attraction.

Keywords: Serial killer, Crime location, Area of attraction, Point of attraction.

INTRODUCTION

The serial killer is always not willing to leave any clues and will try to attempt in different time and places, location and even the way he attack to mislead the police. However, in the deep sense, the type of his crimes is always the same. According to Mary Buck's Theory of Hypothetical Circle of crime, we postulate the following assumptions as studied by Barlow D.H *et al* (2011). In the first attack, the serial killer tends to choose a familiar place or the places round his residence to commit a crime as lacking of experience. As the serial killer attacks more, he will accumulate more experience and always try to carry out the attack much further away from his residence. Due to the Transportation effect, the serial killer will not extend the crime location to infinity and there will be a limit. He always tries to change to another direction after many attacks. The criminal area will turn to a sector and then a circle instead of a straight line. Thus we have two types of criminal area. One is attacking along a straight line; the other is within a sector. Combining these two types, the criminal area becomes a circle.

METHODOLOGY AND MODELING

Determination of the Crime Location:

(Li. 2008) observed that at the beginning, the killer tends to attack round his residence, so his residence is probably near the center *S;* (**As shown in Fig 1**).

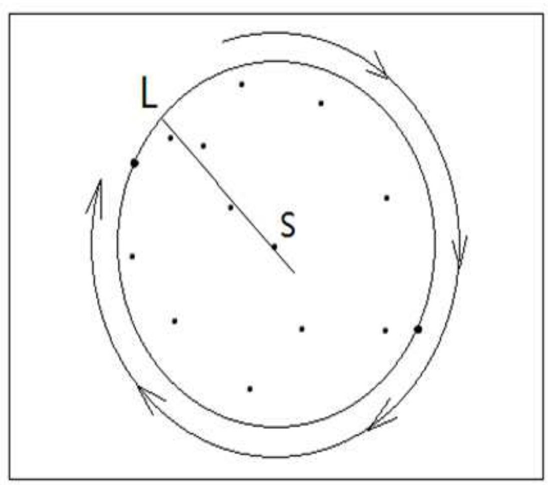

Fig 1: Rough living area

The initial attacks locate in the straight line L *with* the killer's residence. Since more experience has been accumulated, the killer's crime locations will be gradually around his residence and become further at the same time (As shown by arrows in fig 1). It is further assumed that (1) the time the killer uses to commit a crime is much shorter than the interval between two crimes. (2) When two successive crimes are in different cities, a city can be represented by a point on the map. (3) When two successive crimes are in the same cities, the landmarks can be considered into a point in the map. (4) The criminal rate and economical development level are given.

Four Procedures to Determine the Crime Location:

When the killer conducts continuous crimes, the places he chooses are always associated with the places he commits the former crimes. Based on this experience, we determine the crime location in this way.

(A) ***Preliminary analysis of the data***

We mark the former crime places with a point on the map. We give a three-point example as follows: (**Fig 2**).

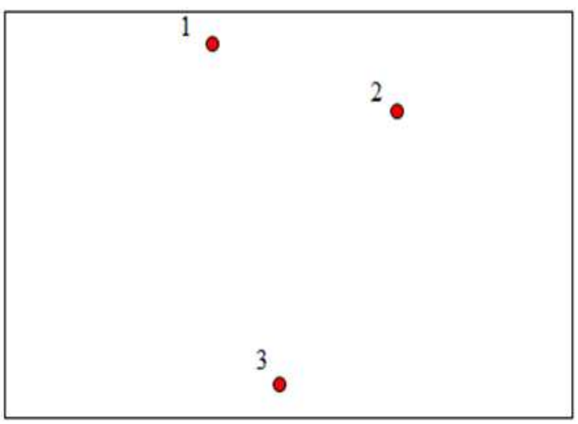

Fig 2: A three-point example

Then we make a circle with Point 2 being the centre and the distance between Point 1 and Point 2 being the radius. In the same way, let Point 3 be the centre, and the distance between Point 2 and Point 3 be the radius to make another circle. Also we have to make a circle with Point 1 the centre, the distance between Point 1 and Point 3 the radius .We will obtain a map as follows: (**Fig 3**).

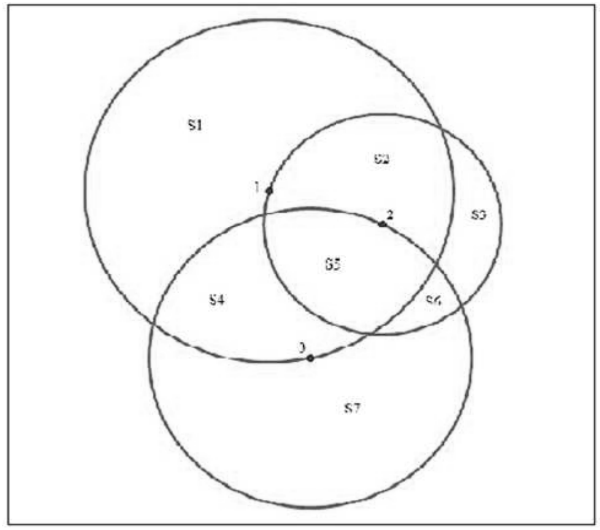

Fig 3: A rough crime area in Procedure One

Here, we suppose that the intersection of these three circles is not an empty.

According to this map, we predict that the killer will commit another crime within the union of areas $S_1, S_2, S_3, S_4, S_5, S_6$ and S_7

Proof: The proof is based on the theory of Hypothetical Circle of crime. Supposing that the killer has made two attempts, we make a circle with the distance between Crime Location 1 and Crime Location 2 being the diameter and the midpoint of them being the center. Obviously this circle is the subset of Circle 2(As is shown in the (**Fig 3**). So it is valid to say that killer lives within Circle 2. Respectively, we can obtain that the killer also may live in Circle 1 and Circle 3.Thus the killer should live in the intersection of these circles. Then we predict the next crime location according to the statistical data. We take P_1 as the possibility of the crime been committed in S_1 , in the same way P_2 the possibility of S_2, P_3 the possibility of S_3, and so on. Then we find that the area S_5 is covered by three circles, areas S_2, S_4, S_6 are covered by two circles, while the others are just covered by one circle.

Hence, we conclude that $P_1 = P_3 = P_7$, $P_2 = P_4 = P_6 = 2P_1$, $P_5 = 3P_1$

Since the crime is committed in the union of the ranges of areas S_1 to S_7, it shows that $P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 = 1$

Combining the above equations, we get the values of P_1, P_2 upto P_7

If we are informed about the previous *n* crime locations, we can get an area containing the (n+1)th crime location in the same way. At the same time, we obtain the possibility that the (n+1)th crime occurred in each partition.

(B) ***The Prediction of Linear Crime***

In order to describe this procedure, we give an example with 5 known points. Based on the Theory of Hypothetical circle of crime, we find that at the beginning, the killer is more likely to attack in the same direction, but as he is more and more afraid of the police, he will attempt in further and further places. We comprehensively analyze the previous five crime locations and postulate that the intersection of the five circles is not empty region. Repeat the procedure one, we will have five circles, we number every circle with the number of its center and their intersection is where the killer lives. The whole map is divided into different partitions by these circles. Let Point 5 be the vertex of the smallest angular domain which exactly contains the certain small area. Calculate the average distance of the five points and we consider it as the distance between the next crime location and Point 5. We keep the maximum and minimum distance with respect to the longest and shortest radii. We will obtain a fan-shaped region which shares the vertex with the angular domain mentioned above. The next crime will be absolutely committed in this region.(**As is shown in the Fig 4**).

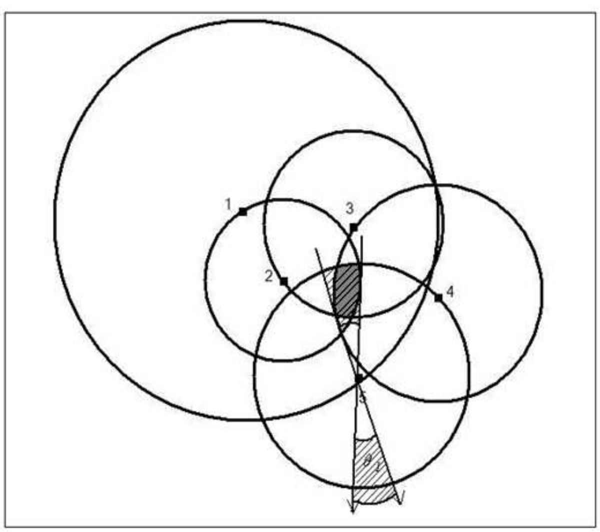

Fig 4: The crime area in Procedure One

We assign every fan-shaped region with a vector $[a, b]$, where 'a' represents the times this area is covered and b represents the sum of the numbers of the boundaries. First, we compare the value of the a. Bigger is the a means the more important is that area. When a does not change, we compare the value of b. The bigger is the b means the more important is that area.

Arithmetic: The area I is more important than area II (**as is shown in the Fig 5**).

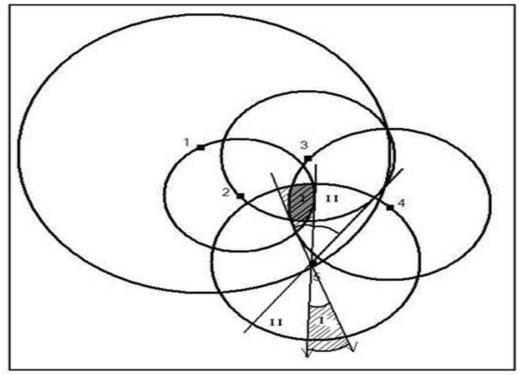

Fig 5: The example of Procedure One

Afterwards, we arrange all the fan-shaped regions according to the method mentioned above and the next location will only be in the first three regions. The possibilities that next crime occurs in the area I, II, III are 50%, 30%, 20%.The killer may also attack in the opposite direction. We find the symmetry point of the fifth crime location, which we call 5' (**shown in Fig 6**), and we can find the next crime location in the same way as above.

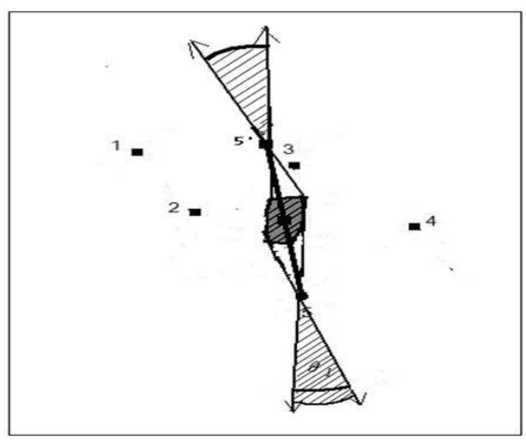

Fig 6: The crime area in the opposite direction.

(C) Fan-shaped prediction

As the far places are difficult to arrive, the killer will change his direction and we can use a sector to outline all the crime location. We still use the point to mark the crime locations as before. After adding up the information we find that the next point usually exists in the circle determined by the first points. So we use three points to predict the area next crime will be committed. As it is shown (**Fig 7**) in the following maps:

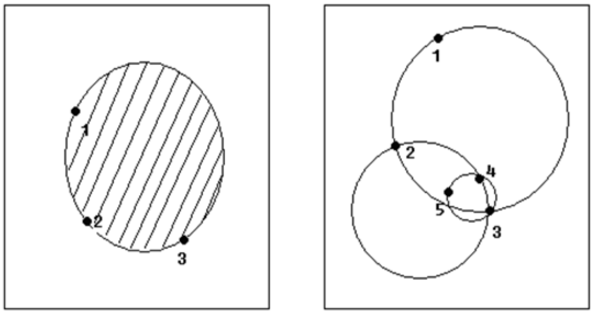

Fig 7: The crime area predicted by three points

The Point 4 is likely to exist in the shaded area, so we consider that the possibility is 1.Then we have to narrow the shaded area in order to obtain the exact location of Point 4. Area S_1 is the killer's living area obtained according to Procedure 1. The killer will change to another direction within the shaded area to attack according to the fan-shaped crime theory. Point 3, 4 and 5 are the recent crime locations and the killer will not attack in the non-shaded area in the right side. (**Fig 8**)

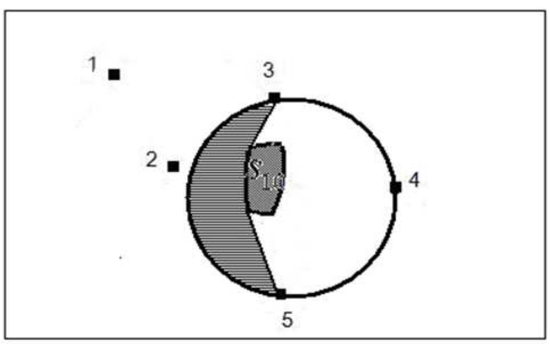

Fig 8: The fan-shaped crime area

(D) ***Calculate the Criminal possibility***

Every Scheme mentioned above has its own advantages, so in order to create a best method; we integrate the above schemes with each other. We make it based on the prediction accuracy of the method.

Firstly, we define *T* as the crime density of a certain scheme, which can be obtained from the following equation:

$$T = \frac{100\%}{S}$$

For every scheme, we observe that the bigger the predicting criminal area is, the lower the precision of the accurate location of next crime is. So the crime density can be a good criterion of the prediction accuracy of each method. Secondly, we use the weighted average method to integrate two schemes. Every scheme has a prediction of the next criminal area and every partition corresponds to different probabilities. Of course the result of the scheme with higher prediction accuracy will take a greater weight in the equation. So we consider the crime density of a scheme as its weight.

Arithmetic**:** In order to explain it more clearly, we give the following example. For a certain partition, the probabilities of three schemes are P_1 , P_2 , P_3 . The corresponding crime densities are $T_1, T_2, T_3.$ Then we will get the integrated probability of this area on the basis of the equation described below.

$$P = P_1 \times \frac{T_1}{T_1 + T_2 + T_3} + P_2 \times \frac{T_2}{T_1 + T_2 + T_3} + P_3 \times \frac{T_3}{T_1 + T_2 + T_3}$$

Thus we conclude the final prediction area and the crime probabilities of different partitions. In the following instance, we will predict the sixth crime location on the basis of the first five crime locations. Firstly, the first 5 crime locations are shown in the following map (**Fig 9**).

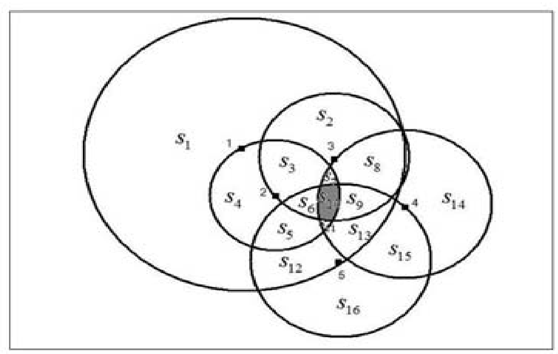

Fig 9: The map of five crime locations

We obtain the division of the sixth criminal area and calculate the probabilities of each partition according to the Scheme One (as is shown in the following map).

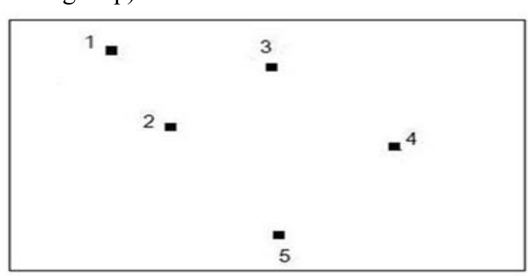

Fig 10: The sixth crime partition

Secondly, we calculate the crime probabilities of each partition according to the Scheme Two.

Thirdly, we calculate the crime probabilities of each according to the Scheme One.

Finally, the probabilities of each partition, the crime will be committed in, will be reached, if we combine the above results. Supposing that the probability of Partition I from three schemes are P_1, P_2, P_3 and the crime densities are T_1, T_2, T_3, the final probability of Partition I will be determined by the following equation:

$$P = P_1 \times \frac{T_1}{T_1+T_2+T_3} + P_2 \times \frac{T_2}{T_1+T_2+T_3} + P_3 \times \frac{T_3}{T_1+T_2+T_3}$$

Then we apply this method to every partition, the outcome describes the final prediction crime location and its probability.

Special attractions

Brantingham, P. J (1981) described that some landmarks will effect the criminal location, which we call Special Attraction.

Positive Attraction

After analyzing Peter's case, we find that he always committed a crime in cities. For other kinds of crimes we also find that some special landmarks have always something to do with the criminal location chosen by killers. For example, normally robberies occur in the bank or near to it, so the possibility that a robbery happen near the bank is much higher. In the map, a bank is indicated by a point, so we call it Attraction Point.

The positive attraction point is in the external area.

As shown in the following map, we have had already a possible raw criminal area .We take its geometric center O as the center to find the smallest circle which must cover this area. Then, we triple the radius and get a bigger circle, which we call the criminal boundary. Finally, we compare all the cities within the boundary and select the nearest three ones.

Firstly, we take a city among the above three and find its geometric center S. Connect O and S; and prolong it. Secondly, we evenly divide the part of line where the criminal area covers it and we get 5 points. Finally, make a perpendicular line passing through each point, thus, we divide the criminal area into five partitions.

The shortest distance between Point S and the crime boundary is d_0 and every two adjacent points share a distance of d. Since Point S has a positive attraction to the killer, we consider that earer to S, exhibits a higher possibility of committing a crime. According to the order from near to far, we take the five bar areas as A_1, A_2, A_3, A_4, A_5

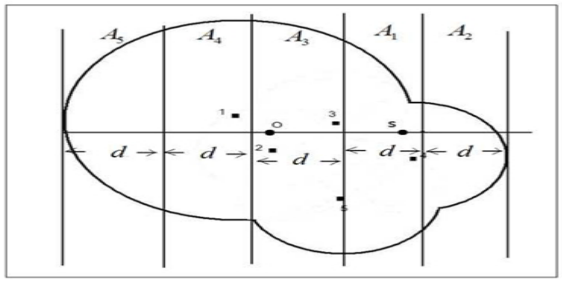

Fig 11: The division of the crime area (five bar areas)

Their corresponding original criminal possibilities are P_{10}, P_{20}, P_{30}, P_{40}, P_{50}, which are changed because of the effect of point *S*. we have that:

$$P_i' = P_{i0} \times \frac{1}{d_0 + id} \qquad (i = 1,2,3,4,5)$$

$$P_i = \frac{P_i'}{\sum_{k=1}^{5} P_k'} \qquad (i = 1,2,3,4,5)$$

Thus the criminal possibility of nearer area will be bigger.

The positive attraction point is in the area

When the positive attraction point is in the area, the killer is more likely to commit crimes in the cities ascribed in this area. As depicted in the following Fig, d_i is the distance between the geometric center of the criminal area and the each bar area. We have

$$P_i' = P_{i0} \times \frac{1}{d_i} \qquad (i = 1,2,3,4,5)$$

$$P_i = \frac{P_i'}{\sum_{k=1}^{5} P_k'} \qquad (i = 1,2,3,4,5)$$

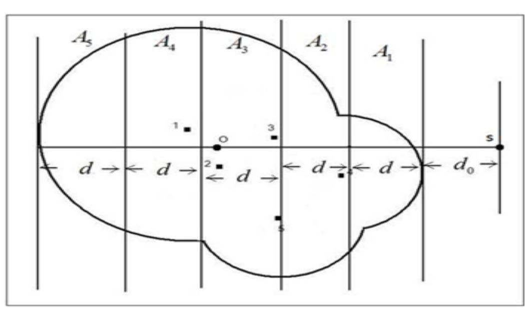

Fig.12: The division II of the crime area (five bar areas)

Negative Attraction

The above discussion is on the basis of positive attraction, which will make the criminal possibility higher than before. But the landmark like a police station has a negative attraction to the killer as he is easily arrested there. We discuss it in the same way as the positive attraction.

The negative attraction is in the external of the criminal area Divide the criminal area as the positive area. Since the negative attraction point reduces the criminal possibility of the place near it, it is valid to assume that near to the negative attraction point, the crime possibility is greatly reduced. Hence we have

$$P_i^{'} = P_{i0} \times (d_0 + id)$$

$$P_i = \frac{P_i^{'}}{\sum_{k=1}^{5} P_k^{'}} \qquad (i = 1,2,3,4,5)$$

The negative attraction is in the criminal are
In this situation, we have that

$$P_i^{'} = P_{i0} \times d_i \qquad (i = 1,2,3,4,5)$$

$$P_i = \frac{P_i^{'}}{\sum_{k=1}^{5} P_k^{'}} \qquad (i = 1,2,3,4,5)$$

Now we find the positive and negative attraction according to different kinds of crimes, which are shown in the following: (**Table-1**).

Table 1 Positive and negative "Point attraction" according to different kinds of crimes

Crime	Positive attraction	Negative attraction
Theft	Bus station, shopping mall	Police station
Murder	City	Police station
Robbery	Bank	Police station
Rape	Police station	

Area of Attraction

The economic development and the criminal possibility are greatly dependent on each other. Hence it is necessary to consider this factor also.

Here we introduce the conception of Area Attraction and its formula and we give an example of the criminal rate, as discussed by Yan, J. *et al.,* (2009) in their studies.

Step 1: We divide the map according to different districts, which is shown in the following Fig and try to find their corresponding criminal rate.

Step 2: Outline the criminal area on (**Fig 13**) and the criminal area will be divided into different partitions. As is shown in (**Fig 14**), the criminal area is divided into nine partitions, namely, S_1, S_2 S_3 . . . , S_9

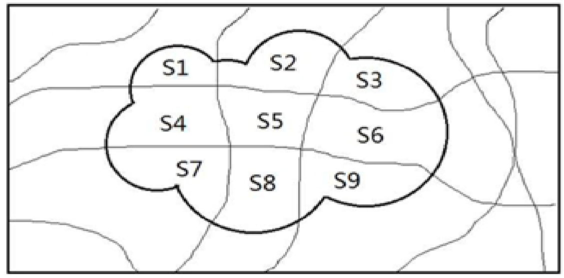

Fig 13: The division of the map according to different districts

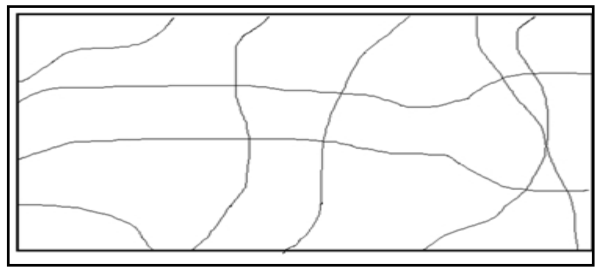

Fig 14: The division of the crime area according to different districts

Step 3: The corresponding criminal rate of partitions S_1, S_2 S_3 . . . , S_9 and U_1, U_2 , U_3 , . . . U_9 ,combining the former division and this division, we will get a new division. Thus the criminal area is divided into *n* partitions, whose corresponding criminal rate is u_i and the crime possibility is p_i

Step 4: Considering the crime rate, the criminal possibility changes to p_i , which can be calculated as follows:

$$P_i = \frac{P_i U_i}{\sum_{j=1}^{n} P_j U_j}$$

After taking the criminal rate into consideration, the criminal possibility of a partition with a high criminal rate will become bigger. At the same time, the place with a low criminal rate will have a smaller criminal possibility. As the criminal possibility is positively correlated with the criminal rate, we call the criminal rate as Positive Area Attraction. On the contrary, the criminal possibility is negatively correlated with the police power, which we call the Negative Area of Attraction. We take the police power into consideration in the same way as the criminal rate. If we know the police power in a certain area, we have the new criminal possibility as follows:

$$Pi = \frac{\frac{Pi}{Vi....}}{\sum_{J=1}^{n\,Pi} \frac{pj}{Vi}}$$

Now we clarify the following factors into two categories, namely, positive and negative according to different kinds of crimes.

Table 2 Positive and negative "Area attraction" according to different kinds of crimes

Crime	Positive area attraction	negative area attraction
Theft	Population, economy development level, Ethnic conflict, religious conflict, transportation, The gap between rich and poor	social security.
Murder	Economy development level, Ethnic conflict, religious conflict, The gap between rich and poor	Police power, Education, social security.
Robbery	Economy development level, Ethnic conflict, religious conflict, The gap between rich and poor	Police power, Education, social security.
Rape	Ethnic conflict, religious conflict, The gap between rich and poor	Police power, Education

3. Model Testing

Case Review

From June of the year of 1988, a killer in Baiyin, Gansu Province in China specially chose young girls in red as his targets. Every time, he followed and observed the girls and then seeks an opportunity to sneak into their apartments and kill them, as reported by Sina News, China (2004). The crime locations and time are listed in the following (**Table-3**).

Table 3. The list of crime locations and time.

N	Time	Occasion
1	May 26th	Yongfeng Street, Baiyin District, Baiyin, Gansu Province,
2	July 27th 1994	Baiyin Power Supply Bureau, Gansu Province, China
3	Jan 16th 1998	Shengli Street, Baiyin District, Baiyin, Gansu Province, China
4	Jan 19th 1998	Shuichuan Road, Baiyin District, Baiyin, Gansu Province China
5	July 30th1998	Baiyin Computation Office, Baiyin, Gansu Province, China
6	Nov 30th 1998	Dongshan Road, Baiyin District, Baiyin, Gansu Province, China
7	Nov 20th 2000	Baiyin Spinning Factory, Baiyin, Gansu Province, China
8	May 22nd2001	Shuichuan Road, Baiyin District, Baiyin, Gansu Province, China
9	Feb 9th 2002	Taolechun Hotel, Baiyin, Gansu Province, China

The testing and prediction of the crime locations

We mark the former criminal locations on the map with points and make rectangular coordinates whose origin lies in the upper left corner (**As is shown in the following Fig-15**).

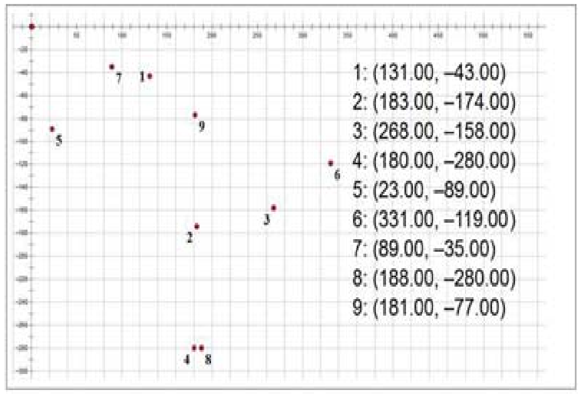

Fig 15: The criminal locations

Now, we use the second scheme mentioned in the paper to test the Point 7, Point 8 and Point 9. Since the killer has not been arrested, we will predict the location and time of his next crime to help the police. We also give the testing details next.

Fig 16: Criminal boundary and residence (the shaded area)

Step 1: Find his criminal boundary and residence (as the shaded area).

Step 2: Determine the probable criminal area.

As in the Fig above, the shaded area is the probable seventh criminal area. In fact, the seventh crime location is indeed in this area and simultaneously we have the criminal possibility that place is 44.32%.In the same way, we test the 8th and 9th crime locations are also lies in the predicted area with the criminal possibility of 35.78%, 38.34%.

We can get the criminal possibility of the dark shaded area as 39.72%.

Step 3: The prediction of the tenth crime location
We predict it in the same way, as is shown in the following (**Fig-17**).

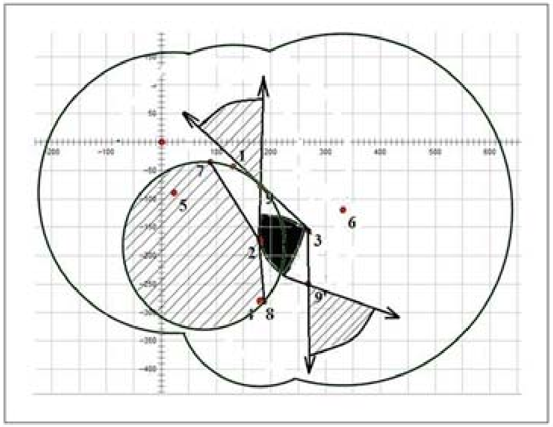

Fig 17: The prediction of the tenth crime location

RESULTS AND DISCUSSIONS

By pointing out the former criminal locations on the map and making rectangular coordinates with origin in the upper left corner, we try to find the criminal possibility of the partition with high criminal rate. We use the second scheme mentioned in this paper to test the Point 7, Point 8 and Point 9. Since the killer has not been arrested, we will predict the location of his next crime to help the police. As we have discussed the steps and shown the probable criminal boundary and residence, we also determined the probable criminal area. As in the (**Fig 17**), the shaded area is the probable seventh criminal area. In fact, the seventh crime location is indeed in this area and simultaneously we have the criminal possibility of that place is 44.32%.In the same way, we test the eighth and ninth crime locations are also lies in the predicted area with the criminal possibility of 35.78%, 38.34%. The same way, we predict the probability of the tenth crime location, which in results is found to be 39.72%.

CONCLUSION

Based on the theory of Hypothetical Circle of crime, we designed our statistical model in four stages. Initially, we have the Preliminary analysis of the data, and then we calculate the Prediction of Linear Crime. In the next stage, we design Fan-shaped prediction model and finally, we calculate the Criminal possibility location. We are quite confident, that this model can further be improved in order to predict the next crime time.

REFERENCES

Barlow, D.H. and V.M. Durand (2011) *Abnormal psycholog: An integrative approach* (6th ed.). Belmont, CA: Wadsworth, Cengage Learning.

Brantingham. P. J. (1981) *Environment Criminology*, Beverly Hills, CA: Sage,

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

Barnett, W. S. (1986)*Definition and Classification of Mental Retardation: A reply to Zigler, Balla, and Hodapp.* American J. Mental Deficiency, (**2**):111-116.

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**): 419Pp.

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

Brown, H. and J. Stein (1997). *Sexual Abuse perpetrated by Men with Intellectual Disabilities: A Comparative Study.* Journal of Intellectual Disability Research, 41(**3**): 215-224.

Calnen, T. and L. S. Blackman (1992) *Capital Punishment and Offenders with Mental Retardation: Response to the Penry Brief.* American Journal on Mental Retardation, 96 (**6**): 557-564.

Li. G. (2008) *The Application of Map Analysis Methods in Criminal Information Analysis*, Journal of Guizhou Police Officer Vocational College.

Yan. J, H. Yuan, X. Shu and S. Zhong (2009) Optimal clustering algorithm for crime spatial aggregation states analysis. Center for Public Safety Res, Deprt. of Eng. Physics, Tsinghua University.