British Journal of Mathematics & Computer Science
13(3): 1-12, 2016, Article no.BJMCS.22192
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

Bayesian Approach to Predict Offender's Probable Anchor Point Using Geographic Profiling System in Akure, Nigeria

B. O. Afeni[1*], O. Olabode[2] and N. O. Oluwaniyi[1]

[1]Department of Computer Science, Joseph Ayo Babalola University, Ikeji - Arakeji, Nigeria.
[2]Department of Computer Science, The Federal University of Technology, Akure, Nigeria.

Article Information

DOI: 10.9734/BJMCS/2016/22192
Editor(s):
(1) Kai-Long Hsiao, Taiwan Shoufu University, Taiwan.
Reviewers:
(1) Anonymous, Rivers State University of Science and Technology, Nigeria.
(2) Moamin A. Mahmoud, Universiti Tenaga Nasional, Malaysia.
(3) Alaa El-Halees, Islamic University of Gaza, Palestine.
Complete Peer review History: http://sciencedomain.org/review-history/12568

Method Article

Received: 22 September 2015
Accepted: 20 November 2015
Published: 05 December 2015

Abstract

One of the vital clues offenders leave behind when they commit crime is the crime location. In recent years, several algorithms have been used to forecast the home site of unknown serial offenders on the basis of crime locations that have been linked to one offender. These have developed from spatial typologies to softwares that can provide direct support to crime investigations. Geographic profiling is an investigative methodology used in criminology that analyses the locations of a linked series of crimes to decide the most probable location for where the offender lives. This work establishes Bayesian Dirichlet Process Mixture Model (DPM) for crime hotspots analysis and as a geographic profiling system. It expands physiological profiling approach by integrating the idea of distance decay and buffer zone. The model was then implemented and tested with 119 spatial data of report serial theft cases in Akure, Nigeria. GPS Garmin was used to collect the data. The reported crime locations were visited to gather the data and were pre-processed by converting it into the machine readable format. The final output of the analysis (geoprofile) using the model was developed that depicts the most probable area of criminal(s) anchor point. A probability score was calculated for every point within the study area to indicate the likelihood that it contained the offender's residence. The model was implemented in R. The model provides a practical tool for criminologist in targeting interventions and a more efficient use of resources for serial crime investigation. It can assist law enforcement agencies in decisions and policies making.

Keywords: Bayesian approach; criminology; Dirichlet process mixture; geographic profiling; geoprofile.

Corresponding author: E-mail: babajideafeni@gmail.com;

1 Introduction

The trend in crime analysis called geographic profiling began and became relatively common in the 1980's. Computerized mapping of crime occurrences began to substitute and improve upon "pin maps" which had been utilized for decades to analyse geographic trends in crime [1]. Recent studies into crime data have led to the development of softwares that can be used to foretell the residence of a serial offender by looking at the pattern of offenses committed. Many of these software packages require a series of at least five crimes in order to be effective [2]. Investigations of serial crime typically involve too many, rather than too few crime suspects; for example, the investigation into the Yorkshire Ripper murders in the UK between 1975 and 1980 generated 268,000 names [3]. In criminology, geographic profiling system techniques use spatial data concerning the locations of connected crime sites to create a surface of search priority that is overlaid on a map of the study area to produce a geoprofile, which in turn allows the police to prioritize investigations by systematically examining suspects associated with locations in descending order of the height on the geoprofile [4].

Geographic profiling model is based on the assumption that offenders are more likely to select their victims and commit a crime which would be centred near their home address (or place of work or another address where they spend a lot of time such a friend's address, club, and recreational centres etc.). Generally, geographic profiling system is used in cases of serial murder, or serial rape and also arson, theft, bombing, robbery, and other crimes. In addition to determining the offender's most likely area of residence, an understanding of the spatial pattern of a crime series and the characteristics of the crime sites can tell detectives other useful information, such as whether the crime was devious and the degree of offender knowledge of the crime location. Geographic profiling is a branch of offender or criminal profiling (the inference of offender characteristics from offence characteristics). Consequently, it is related to psychological or behavioural profiling. Psychological profiling is about "who" that committed the crime while geographic profiling is about "where" the crime was committed. Work establishes Bayesian Dirichlet Process Mixture Model (DPM) for crime hotspots analysis and as a geographic profiling system.

2 Materials and Methods

In the section, we have discussed most of the methods used in this research work.

2.1 The Geographic Profiling System

The existing system of crime investigation using geographic profiling system makes use of some common notation [5]. A point x will have two components x = a, b. These can be latitude and longitude, or distances from a fixed pair of perpendicular reference axes. It is assumed there are series of n linked crimes, and the crime sites under study are labelled $x_1, x_2, \ldots, x_n$. Also, the offender's anchor point is represented with symbol z. The anchor point can be the offender's home, work place, or some other location of significance to the offender. The first and most vital issue to consider in a geographic profiling system is how to measure the distance between points which can be either Manhattan or Euclidean distance. Euclidean distance is a common method which makes use of the usual notion of distance. Here, the distance d(x, y) between two points x and y is given by $d_2(x, y)$ where

$$d_2(x, y) = \sqrt{[x^{(1)} + y^{(1)}]^2 + [x^{(2)} + y^{(2)}]^2} \tag{2.1.1}$$

Another means is to use the Manhattan distance. In this instance the spacing between x and y is given by $d_1(x,y)$ and

$$d_1(x, y) = \left|x^{(1)} + y^{(1)}\right| + \left|x^{(2)} + y^{(2)}\right| \tag{2.1.2}$$

There are other alternatives that can be used. For example the total time to make the trip while following the local road network, or the total street distance following the local road network. One thing to note between these distances is that the Euclidean distance gives the same results irrespective of the choice of the coordinate axes; specifically, rotating a pair of points around a third does not change the distance between the pair. This however does not hold for the Manhattan distance, and so the coordinate axes need to be chosen carefully when using the Manhattan distance.

Geographic profiling algorithm strategy can be divided into two general categories. They are: spatial distribution strategies and probability distance strategies. A spatial distribution strategy involves making use of Centroid (mean centre) of reported crime locations to predict offender(s) location. The simplest of this strategy is to estimate the offenders anchor point (z) by the centroid $C_{centroid}$ of the serial crime. The centroid, also known as the center of mass is defined to be

$$C_{centroid} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.1.3}$$

There is another method of estimating the anchor point by the center of minimum distance C_{cmd}. This is chosen to be the value of y of which the sum of the distances from that site to the location of the offence.

$$D(y) = \sum_{i=1}^{n} (x_i, y) \tag{2.1.4}$$

Here diverse choices of the distance function *d* lead to different choices for the center of minimum distance. Unlike the centroid, there is no simple formula that gives the value of C_{cmd}; however, there are a number of efficient algorithms that can approximate it to any desired accuracy [6].

2.2 Geographic Profiling Predictive Model

Geographic profiling describes the patterns observed in crime distributions and the mental process that led to their formation. Given a serial offender's home location and anchor points, crime detectives could identify which areas the offender would likely offend and also his anchor point. Researchers took this descriptive model and inverted it to create the predictive models that make up geographic profiling [4,7]. A number of models were developed, ranging from the very simple to the very complex, and are covered below. The models are divided into two categories based on their overall strategy.

2.2.1 Spatial strategies

A study found that 87% of a sample of serial rapists lived inside a circle with a diameter equal to the distance between the two farthest crime scenes [7]. However, this pattern, which they called the "circle hypothesis," barely narrows the search area. Moreover, investigators were already in the practice of looking within the "convex hull," the polygon created by connecting the outermost crime scenes, for the offender's residence.

The technique was modified to make a single-point prediction by designating the center of the circle as the anchor point. Because the circle hypothesis incorporates only a fraction of the knowledge about serial offender behaviour, it remains the most simplistic model in geographic profiling and is used more often as a control method [6].

2.2.2 Probability strategies

The strength of spatial models in offering a single-point prediction is also their greatest weakness: they only provide information for one coordinate of the map. If the predicted anchor point is not the actual anchor point, the model leaves investigators on their own to interpret how that single point influences the search of the surrounding area. Probability strategies solve this problem by combining a center of gravity method with a distance decay function to provide information for every coordinate of the map.

For each point in a reference area containing the crime distribution, a likelihood score, also called a density or hit score, is calculated by a two-step process. First, the distance between itself and each crime scene is evaluated by the distance decay function. Second, the resulting values are summed to find the likelihood score. The higher the score, the more likely it is to be the serial offender's primary anchor point [8]. Buffer zones are sometimes incorporated into the probability strategy. The reference area is then separated into a contoured density map to create prioritized search areas. The top contour containing the highest likelihood scores is called the "top profile region".

2.3 Geographic Profiling Using Bayesian Approach

The problem of geographic profiling is, given the locations of a series of crimes committed by an offender, how can the offender's anchor point be estimated? This technique is able to account for geographic features that affect the selection of the crime sites, distribution of potential anchor points, differences in travel distances of different offenders, and certain demographic characteristics of the offender [9].

As stated in [5], assuming the offender chooses potential locations to offend randomly according to some unknown probability density function $P(x)$. The use of a probability distribution here does not imply that the offender is selecting targets randomly, though there is a possibility of that happening. Instead it represents inadequate knowledge to make educated guess about the decision making process of the offender.

It was also assumed that $P(x)$ depends on just two factors; first is the anchor point z of the offender. Second is the average distance that the offender is willing to travel to offend. It is factual to say that different offenders have different willingness to travel and that the travel patterns of a twenty five year old will be different that those of a seventy year old offender. For this reason, it was explicitly allowed for the possibility that the offender's average travel distance may affect the choice of targets by the offender i.e the crime location.

Thus, the initial mathematical model is that the offender with anchor point z and average offense distance α chooses to offend at the location x according to an unknown probability distribution $P(x \mid z, \alpha)$. The elements of the crime series $x_1, x_2, \ldots, x_n$ then represent a sample from this unknown distribution. Hence, suppose that the form of $P(x \mid z, \alpha)$ was specify; then the geographic profiling issue becomes a parameter identification issue for the unknown parameters z and α. This issue was tackled using Bayesian approach to try to find a probability distribution for the unknown parameters. The standard Bayesian rule states that the posterior distribution of z and α given a series of crimes $x_1, x_2, \ldots, x_n$ is

$$P(z, \alpha | x_1, x_2, \ldots, x_n) = \frac{P(x_1, x_2, \ldots, x_n \,| \mathrm{z}, \alpha) \pi(\mathrm{z}, \alpha)}{P(x_1, x_2, \ldots, x_n} \tag{2.3.1}$$

Here $P(x_1, x_2, \ldots, x_n \mid z, \alpha)$ is the model of offender behaviour; it specifies the probability density that the offender will offend at all of the locations $x_1, x_2, \ldots, x_n$ given that they have anchor point z and average offense distance α. The factor (z, α) is the prior distribution of anchor point and offense distance; it represents what is known about the offender before taking into account the information from the crime series itself. The factor $P(x_1, x_2, \ldots, x_n)$ is the marginal distribution; since it does not depend on either z or α, this can be removed by replacing the equality by a proportionality, so

$$P(z, \alpha | x_1, x_2, \ldots, x_n) \propto P(x_1, x_2, \ldots, x_n \,| \mathrm{z}, \alpha) \pi(\mathrm{z}, \alpha) \tag{2.3.2}$$

It was then anticipated that the different crime sites were selected autonomously of one another. This was accepted because it is the simplest way the probability distribution $P(x_1, x_2, \ldots, x_n \mid z,)$ can be related to the model of offender behaviour $P(x \mid z, \alpha)$. However, there is some indications that criminal offense sites are not independent, which would require a more robust model for offender behaviour. Proceeding with the simplest model of independence though, one can then say that

$$P(x_1, x_2, \ldots, x_n \,|\mathrm{z}, \alpha) = P(x_1|\mathrm{z}, \alpha)P(x_2|\mathrm{z}, \alpha)\ldots., P(x_n|\mathrm{z}, \alpha) \tag{2.3.3}$$

The simplest approach to the previous $\pi(z, \alpha)$ is to assume that anchor points and offense distances are independent; then we can write

$$\pi(z, \alpha) = H(z)\pi(\alpha)$$

Combining this, it seems and reveals that there is a relationship

$$P(z, \alpha|x_1, x_2, \ldots, x_n) \propto P(x_1|\mathrm{z}, \alpha)P(x_2|\mathrm{z}, \alpha)\ldots., P(x_n|\mathrm{z}, \alpha)H(z)\pi(\alpha) \tag{2.3.4}$$

Since we are only interested in the distribution of the location of the anchor point(s), we can then take the conditional distribution to have

$$P(z, |x_1, x_2, \ldots, x_n) \propto \int_0^\infty P(x_1|z, \alpha)P(x_2\,|\mathrm{z}, \alpha)\ldots P(x_n|\mathrm{z}, \alpha)H(z)\pi(\alpha)d\alpha \tag{2.3.5}$$

This gives the probability distribution that the offender has anchor point z, given that they have committed crimes at location x_1, x_2, . . . , x_n. If this method will be effective, there is need to specify a model for offender behaviour P(x | z, α). We take the approach that this has the form

$$P(x|\mathrm{z}, \alpha) = D(d(x, z), \alpha)G(x)N(z, \alpha) \tag{2.3.6}$$

Here above, $D(d(\mathrm{x}, \mathrm{z}), \alpha)$ represents the distance decay behaviour of the offender and $d(\mathrm{x}, \mathrm{z})$ is the distance between the crime points (x) and anchor point (z).

2.4 Bayesian Approach to Serial Crime Analysis

For the classic circumstance, to estimate probability density for the anchor point z given crime series x_1, x_2, ., x_n, Bayes' Theorem implies

$$P(z, \alpha|x) = \frac{P(x\,|z,\ \alpha)\pi(z, \alpha)}{P(x)} \tag{2.4.1}$$

Again

- $P(z, \alpha|x)$ is the posterior distribution giving the probability density that the offender has anchor point z and average offense distance α given offender has committed offenses at x_1; : : : ; x_n.
- P(x) is the marginal distribution, since it is independent of z and x it can be ignored.
- $\pi(z, \alpha)$ can be factored again into with same definitions of H and π made above.

$P(x_1, x_2, \ldots, x_n|\mathrm{z}, \alpha) = \mathrm{P}(x_n|\mathrm{z}, \alpha)$ can be reduced assuming that the offence sites are independent.

Substituting we now have,

$$P(z, \alpha|x_1, x_2, \ldots, x_n) \propto P(x_1|\mathrm{z}, \alpha)\ldots P(x_n|\mathrm{z}, \alpha)\mathrm{H(z)}\pi(\alpha) \tag{2.4.2}$$

Finally since the main objective of this work is to determine the most probable anchor point z of the offender, the conditional distribution in (2.3.5) was taken to obtain

$$P(z|x_1, x_2, \dots x_n) \propto \int_0^\infty P(x_1\,|\mathrm{z}, \alpha)\dots P(x_n|\mathrm{z}, \alpha)\mathrm{H}(\mathrm{z})\pi(\alpha)\mathrm{d}\alpha \tag{2.4.3}$$

Where $P(z|x_1, x_2, \dots x_n)$ gives the probability density that the offender has anchor point z given offenses at locations x_1, x_2, . . . x_n.

This is the basic framework for geographic profiling that allowing for many possible choices of P (x|z,α). This also simply lets the addition of more parameters or removes parameter α without significantly changing the model. Geographic profiling works with two fundamental assumptions. The first one is that **z** is independent of α; average distance the offender is willing travel is independent of offender's anchor point. While the second one is that offender's choice of crime sites are pairwise autonomous.

2.5 A Simple Models for Offender Behaviour

The simple model for offender behaviour incorporates most models in the current literature. The following models make the assumption that all offenders have the same average offense distance α known in advance [5]. If we assume that an offender chooses a target location based only on the Euclidean distance from offense location to offender's anchor point, then we result in a normal bivariate distribution is given by

$$P(x|\mathrm{z}, \alpha) = \frac{1}{4\alpha^2}\exp\left(-\frac{\pi}{4\alpha^2}|x - z|^2\right) \tag{2.5.1}$$

Assuming that all offenders have the same average offense distance α and all anchor points are equally likely, then we obtain a product of normal distributions

$$P(z|x_1, x_2, \dots, x_n) = (\frac{1}{4\alpha^2})^n \exp\left(-\frac{\pi}{4\alpha^2}\sum_{n-1}^{n}|x_i - z|^2\right) \tag{2.5.2}$$

Alternative model of offender behaviour is the maximum likelihood estimate for the anchor point as the mean centre of the crime site locations. This is the centrography which is the basis of Rossmo's Formula; this is also the mode of the posterior anchor point probability distribution.

Another behaviour model is that the offender chooses a target location based only on the Euclidean distance from the offense location to the offender's anchor point, but as a (bivariate) negative exponential so that

$$P(x|\mathrm{z}, \alpha) = \frac{2}{\pi\alpha^2}\exp\left(-\frac{2}{\alpha}|x - z|\right) \tag{2.5.3}$$

Supposing that all offenders have same average offense distances and all anchor points are equally likely, then the maximum likelihood estimate for the offender's anchor point is simply the centre of minimum distance for the crime series locations which can be stated as

$$P(z|x_1, x_2, \dots, x_n) = (\frac{2}{\pi\alpha^2})^n \exp\,(-\frac{2}{\alpha}\sum_{i=1}^{n}|x_i - z|) \tag{2.5.4}$$

3 The Analytical Solution to the Bayesian Dirichlet Process Mixture (DPM) Model

As supported by [10,11] in respect of the expressions for some posterior quantities of interest under the Bayesian Dirichlet Process Mixture (DPM) model. These analytical solutions are only practically useful when the number of observations is small as stated previously. To start with, let us condition on a particular partition of the data, as specified by the vector of group indices c = $c_1 \ldots c_n$ described above. Let there be n_j elements in group j (in other words $n_j = \#\{ci : c_i = j\}$), and u groups in total, so that. The prior probability of this (or any) partition under the Chinese restaurant process (CRP) can be written as follows:

$$\Pr(c \mid \alpha) = \frac{\Gamma(\alpha)}{\Gamma(n+\alpha)} \prod_{j=1}^{u} \alpha\Gamma(n_j) \tag{3.1}$$

The following hyper-prior on α was assumed throughout this research work:

$$h(\alpha) = \frac{1}{(1+\alpha)^2} \tag{3.2}$$

Rather than integrate over this hyper-prior separately for each partition, we can define the function t(u) as follows:

$$t(u) = \int_0^\infty \frac{\Gamma(\alpha)\alpha^u}{\Gamma(n+\alpha)} h(\alpha) d\alpha \tag{3.3}$$

This function can be pre-computed for u=1, . . . , n. Then the probability of any partition, integrated over the hyper-prior on α, can be written as

$$\Pr(c) = t(u) \prod_{j=1}^{u} \Gamma(u_j) \tag{3.4}$$

Before deriving the posterior distribution of a partition we must calculate the marginal probability of the data, summed (or integrated) over all possible anchor points. This can be written as

$$\Pr(x \mid c) = \prod_{j=1}^{\infty} \sum_{\omega \in \Omega} g_0(\omega) \prod_{i:c_i=j} f(x_i \mid \omega). \tag{3.5}$$

Although the first product is over an infinite number of potential anchor points, only a finite subset of these anchors (those for which $c_i = j$) have observed data associated with them. For all other values of j the summation is over the space of the prior only, returning a value of one. In fact, while this expression appears complex, in many common situations (such as normal prior and likelihood) a closed form solution can be obtained. Combining (3.4) with (3.5) we arrive at the posterior probability of a partition via Bayes' rule:

$$\Pr(c \mid x) = \frac{\Pr(x \mid c)\Pr(c)}{\sum_{\omega \in P} \Pr(x \mid c)\Pr(c)} \tag{3.6}$$

where the sum is over the set (denoted *P*) of all possible partitions of n observations into up to n groups.

As well as being useful in its own right, the information in (10) is required when computing the posterior distribution of all unknown anchor points (i.e. the Bayesian equivalent of the traditional jeopardy surface). First, the posterior distribution of the j^{th} anchor point, conditional on a particular partition, can be obtained as follows:

$$\Pr(z_j \mid c, x) = \frac{g_0(z_j) \prod_{i:c_i=j} f(x_i \mid z_j)}{\sum_{\omega \in \Omega g_0}(\omega) \prod_{i:c_i=j} f(x_i \mid \omega)} \tag{3.7}$$

We are interested in the probability of finding any anchor point location at the point z, which is given by a Boolean or operation (union) over all z_j for j 1, . . . , u. To a close approximation this is equal to the mean of $Pr(z_j \mid c, x)$ for j 1, . . . , u (when dealing with probability density rather than probability mass this result is exact). Thus, a geoprofile (jeopardy surface) can be easily defined [12], conditionally on a particular partition as:

$$\mathcal{J}(z \mid c, x) = \frac{1}{u}\sum_{j=1}^{u} Pr\left(z_j \mid c, x\right) \tag{3.8}$$

Finally, making use of (3.6), the overall jeopardy surface (geoprofile) can be defined as:

$$\mathcal{J}(z \mid x) = \sum_{c \in P}^{u} \mathcal{J}(z \mid c, x) pr(c \mid x) \tag{3.9}$$

3.1 The Crime Profiling Process

The first of the steps of data acquisition, pre-processing and analysis is to get the coordinates of crime hot spots in Akure, Nigeria.

3.1.1 Data acquisition

The Global Positioning System (GPS) device called GPS Garmin was used to acquire the 119 spatial data of the serial theft hot spots in Akure, Ondo State, Nigeria. The information is the reported cases of theft by the resident of the city. The location where the theft took place was visited and the spatial data was taken. The data gathering spans the period of six months from January to July, 2015. As reported by the residence of Akure, theft is the main type of serial crime that is of security concern.

3.1.2 Data pre-processing

Computer is not really comfortable handling minutes and seconds so decimal degrees (DD) are more frequently used in Geographic Information System (GIS). These express minutes and seconds as portions of whole degree (i.e. 1 minutes = 0.0166667 degree, 1 second = 0.0002778 degree).

Afterwards, the data was then pre-processed i.e the readings (the actual data) were converted to decimal degree which is the machine readable format and the result was used as the input for the system to produce the geoprofile.

3.1.3 Data analysis

The data file was saved in a text tab delimited file, with two columns only. The first column has longitude values and the second column has latitude values. There is no need for column headings. It was ensure that there is no other information in the file. The longitude and latitude data are both be in decimal format. The next thing involves the reading in the serial crime spatial data into R for the model to run on.

3.1.4 The concept of distance decay and buffer zone function

In several studies, there has been an increased emphases placed on the journeys criminals normally take to define the spatial range of criminal activity. These areas become comfort zone for predatory offenders to commit their crime with a sense of no possibility of being caught. Subsequently, criminal acts follow a particular function known as distance-decay function which have it that the further away the regular activity space of an offender is, the less likely that the person will engage in a criminal activity. Nevertheless, there is also a buffer zone where an offender will avoid committing crimes too close to their homes in the likely event that they will be identified by a neighbour. This work engaged the two basic concept as the value of sigma was set to 0.01 under ModelParameters (). This is critical, since sigma relates to the size (in degrees of lat/long) of the spatial clusters of crimes. A typical value for human movement in urban areas would be 0.01, equating to something around 1km. Sigma is actually the standard deviation of the normal distribution of movement from the anchor point - so, for this value of sigma, around 68% of crimes would happen within 900 m of the criminal's home, about 95% within 1800 m and 99% within 2700 m.

3.2 The Algorithm

Below is the algorithm of how the model runs on R.

```
Start
    LoadData(){
    // file.txt contains data
    Input file.txt
    // Assign data length, number of anchor points, etc.
    …
    }
    ModelParameters(Sigma = 0.1){
    // Sets the parameters for the model and the MCMC
    //
    …
    }
    GraphicParameters(){
    // Sets the parameters for plotting and the creation of Google maps
    …
    }
    CreateMaps(){
    …
    }
    RunMCMC(){
    …
    }
    ThinandAnalyse(){
    // Thins the MCMC samples and analyses them to construct the geoprofile
    …
    // Takes the posteriors obtained by the MCMC and thins them to remove autocorrelation
    …
    }
    PlotGP(){
    // Output geoprofile
    …
    }
End
```

4 Results and Discussion

In the section, the model was implemented and tested to analyse crime hot spots and the result was presented.

Fig. 4.1 depicts the map of the study area showing the 119 spatial data used for this study. The CreateMaps () function produced this in R. This map also gives a better understanding about the data collected as it shows each of the data (location) on the map. The black spots on the map are the observed data points which are the location where serial theft took place in Akure from the period of January to July, 2015.

Figs. 4.2a and b represents the marginal likelihood of different numbers of realized anchor points for the analyzed data. It is the output from the MCMC function. The model estimates that there are 3-6 anchors and assigns the highest likelihood to four anchors. This reflects in the geoprofile developed as the white areas seen on the geoprofile are four. This signifies that there are four probable offender's anchor points in the study area.

Fig. 4.3 is the geoprofile i.e the probability map of the offender as it relates to serial crime theft anchor points. This is the final output from the proposed system. The final and calculations completed output is the geoprofile, which is a probability surface overlaid on the study area. The probability surface displays the different probabilities of the study area in terms of anchor point likelihood. This is based on the spatial pattern of the crime cases. The probabilities are represented by surface height: the higher the surface is the greater the likelihood of having a criminal location in that area, these are the white areas. The higher parts of the surface are denoted by low hit score percentages, representing the percentage of area that must be searched before the actual anchor point is found.

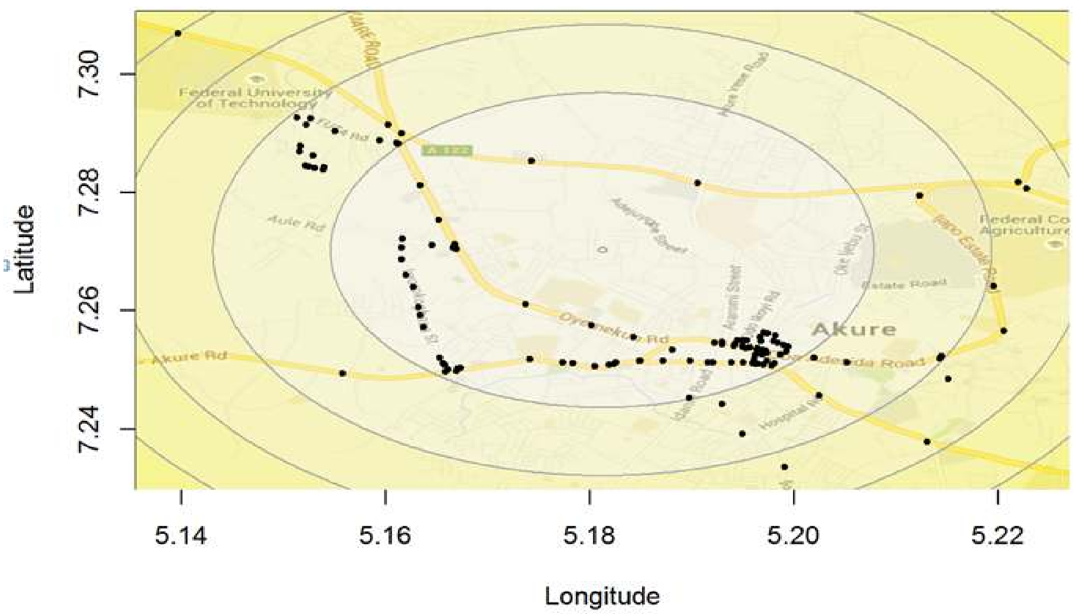

Fig. 4.1. The Map of the study area with the serial crime coordinates

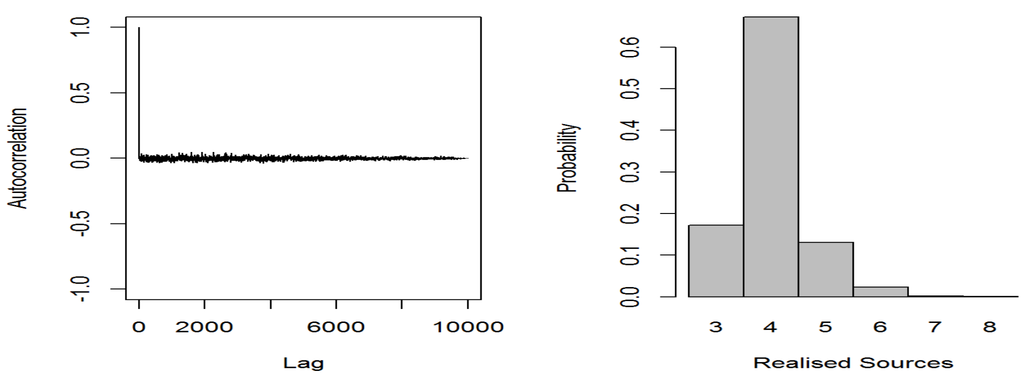

Fig. 4.2. Marginal likelihood of anchor points

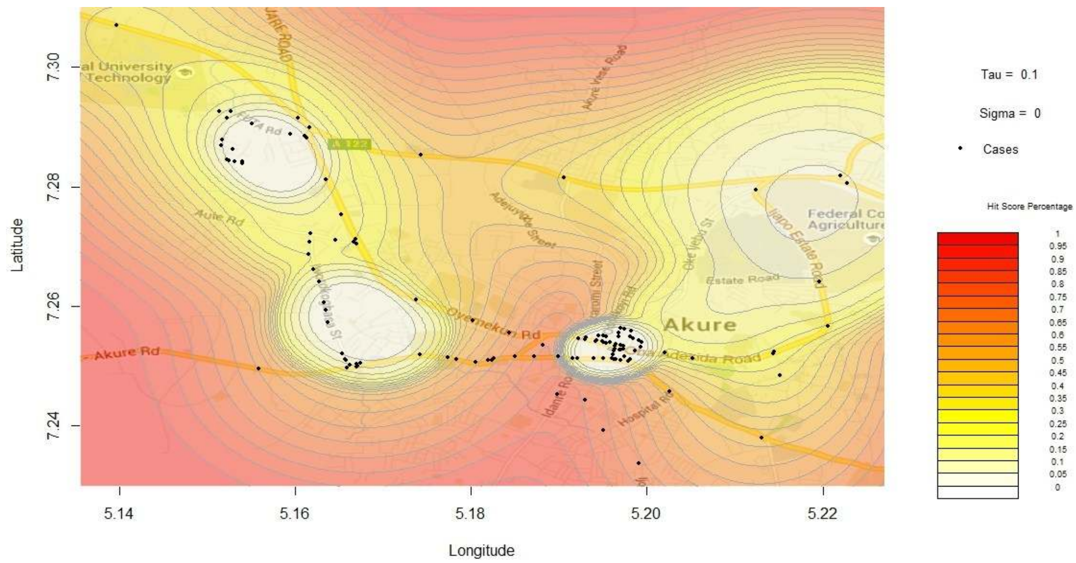

Fig. 4.3. The Geoprofile (Probability map)

5 Summary, Conclusion and Recommendation

5.1 Summary

This research work has been able to expand existing geographic profiling system by incorporating the concept of distance decay and buffer zone. The Bayesian Dirichlet process mixture (DPM) model was implemented and tested with 119 spatial data of report serial theft cases in Akure, Nigeria. The final output of the spatial crime data analysis (geoprofile) using the model was developed that depicts the most probable area of criminal(s) anchor point in the study area. A geoprofile can thus serve as a useful and practical tool for criminologist and should be used to assist police in its decisions and policy making.

5.2 Conclusion

Crime is not randomly distributed, either temporally or geographically. It is obvious that offences happen more often in certain places. This work provides the means to Identify the number of offences that are linked (e.g. carried out by the same offender or offenders) and which are not, which helps to focus investigations. The conventional way of policing is to deploy law enforcement agencies to crime scenes to arrest suspects around that area, whereas the offender might not actually be apprehended with this approach. The proposed method will enhance an in-depth analysis of the serial crimes and help predict the most probable locations where the police should focus attention on in order to swiftly apprehend the offender. This system can be packaged and improved upon to become a standard one that can be deployed for use by the police and other official bodies to target crime prevention resources in the right areas and in the right sorts of ways. To realise this however, there is a need to carry out activities such as Data test, User acceptance testing, System Review and Deployment. The documented processes in this paper are also good source of information for further crime database system development and data analysis. The foremost focus of this work is to locate the comfort zones of offenders using the geographic location of reported serial crime.

5.3 Recommendation

This work has been able to provide a platform for police to limit the fruitless effort of deploying its men all over the place in search of a serial offender. The system should be adopted to help track them down faster than the traditional approach of sending out policemen to immediate crime scene to arrest suspects. As revealed in the result of this work, the actual offender might not actually commit crime in his immediate neighbourhood.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Rich T, Shively M. A methodology for evaluating geographic profiling software. Cambridge, MA: Abt Associates; 2004.

[2] Levine N. Crimestat: A spatial statistics program for the analysis of crime incident locations (version 3). Washington, DC: National Institute of Justice; 2004.

[3] Doney R. The aftermath of the Yorkshire Ripper: The response of the United Kingdom police service. Serial Murder: An Elusive Phenomenon, Praeger, New York. 1990;95–112.

[4] Rossmo DK. Geographic profiling. 1st Edn, CRC Press, New York; 2000.

[5] O'Leary M. A new mathematical approach to geographic profiling. Research report submitted to the U.S. Department of Justice. 2012;16–20.

[6] Snook B, Zito M, Bennell C, Taylor PJ. On the complexity and accuracy of geographic profiling strategies. Journal of Quantitative Criminology. 2005;21:1–26.

[7] Canter D, Larkin P. The environmental range of serial rapists. Journal of Environmental Psychology. 2003;13:63–69.

[8] Canter D, Coffey T, Huntley M, Missen C. Predicting serial killers' home base using a decision support system. Journal of Quantitative Criminology. 2000;16:457–478.

[9] O'Leary M. The mathematics of geographic profiling. Journal of Investigative Psychology and Offender Profiling. 2009;6:253–265.

[10] Neal RM. Markov chain sampling methods for Dirichlet process mixture models. Journal of Computational and Graphical Statistics. 2000;9:249–265.

[11] Verity R, Stevenson MD, Rossmo DK, Nichols RA, Le Comber SC. Spatial targeting of infectious disease control: Identifying multiple. Unknown Sources; 2014.

[12] Rossmo DK. Center for geospatial intelligence and investigation. An evaluation of NIJ's evaluation methodology for geographic profiling software. San Marcos, TX: Rossmo, D.K; 2005a.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/12568