

Framework for Validating Geographic Profiling Using Samples of Solved Serial Crimes

By
Wilpen L. Gorr
July 26, 2004
Revised August 12, 2004

Abstract

This short paper was prepared for the *NIJ Roundtable for Developing an Evaluation Methodology for Geographic Profiling Software* (August 10 and 11, 2004) on approaches for validating geographic profiling (GP) methods. The paper presents a framework for validating any GP method or software package using solved serial crimes including data on crime locations and criminal residences or other anchor points (e.g., work location, girl friend's residence, etc.). Findings from the literature and through analysis include: 1) the appropriate performance measure for GP (that matches policing needs and as extended in this paper) concerns prioritizing relevant areas for investigation; 2) future work should correct the performance measure of GP by excluding irrelevant areas from consideration such as rivers, lakes, cemeteries, etc. (past studies apparently did not do this); 3) additional model parameters may be able to be estimated in empirical studies such as the amount to expand the search area for a serial criminal beyond the minimum rectangle or other boundary enclosing crime sites; and 4) future validation studies for GP should compare alternative models, including simple models for benchmarking, and use holdout samples in a resampling scheme for validating performance.

Acknowledgements

The revision of this short paper incorporates knowledge gained from discussions with participants and organizers of the NIJ Roundtable for Developing an Evaluation Methodology for Geographic Profiling Software: Sean Bair, Patricia Brantingham, Keith Harries, Jay Lee, Derek Paulsen, Shari Pfleeger, Thomas Rich, Michael Shively, Debra Stoe, Ronald Wilson, and Robin Wilfong.

Introduction

Geographic Profiling (GP) is “an investigative methodology that uses the locations of a connected series of crimes to determine the most probable area of offender residence” (Rossmo 2000, p. 1). Other anchor points of routine activities, such as places of work and recreation, are also of interest, but generally are not collected in crime series data. Hence, all past validation studies have used criminal residence (or residences) as the anchor point for validation. Nevertheless, the primary performance measure for validation of GP, the hit score and its corresponding optimal search area, is very likely to identify more than one anchor point. More on this follows below.

Define “total search area” to be the geographic area that has high potential for containing the residence (or other anchor point) of the serial criminal. Often the total search area is a rectangle that includes the locations of the connected series of crimes. I suggest that the total search area have sub-areas coded that are irrelevant; for example, rivers, cemeteries, etc. Such sub-areas should be excluded from consideration for location of the criminal’s anchor points.

Total search area is generally a function of time within a given crime series, with the total search area increasing as time passes and additional serial crimes are committed. Of course there are many possible complexities for search areas, such as a serial criminal attempting to elude capture by changing his hunting ground or residence, hunting areas that are patchy and distance from the criminal’s residence like red-light districts, criminals that do not have fixed residences, total search areas that lie along a route connecting two anchor points, serial criminals that bury victims on the premises of their residences, etc.

I agree with statements by Rossmo (2000, p. 2 and 201) and Canter et al. (2000, p. 458) to the effect that GP has as its purpose to prioritize investigation efforts to focus on sub-areas of the total search area. Examples of investigative efforts include follow-up on tips, interview of suspects, analysis of police records, neighborhood canvasses, information request mailings, etc. The purpose of GP cannot be to pinpoint the criminal’s residence because the underlying theories, data, and measures are too crude. Promising areas, rather than points, are the relevant outputs of GP. Clearly the overall performance and value of GP depends on two separable parts: 1) the quality of optimal search areas identified by objective methods and 2) the judgment and additional data sources and methods brought to bear by crime analysts in interpreting and further refining optimal search areas.

GP estimates a risk index that varies by location within the total search area. The greater the risk index, the higher the probability that the serial criminal resides within an area. Generally, the risk index is a sum of surfaces. There is one sub-surface for each serial crime based on a distance decay function, assuming that a criminal does not travel far from his anchor point to commit crimes. Parameters of the distance decay function need to be estimated from a sample of serial criminals and their crimes. Generally, researchers have made the risk surface be a discrete function by partitioning the total search area with grid cells and estimating one value of risk for each grid cell.

Suppose that there are data for a solved serial crime series with crime locations and residence location of the serial criminal known. The corresponding optimal search area, as I call it, is the aggregate of all grid cells within the total search area that have risk index equal to or greater than the risk index of the grid cell containing the serial criminal's residence (Rossmo, p 201). Optimal search specifies that investigation start with the highest risk grid cell first for investigation and proceed down the list of grid cells sorted by risk index until reaching the criminal's residence area. Researchers do not expect that police will actually conduct such a search, but rather that police will identify all areas with a risk index exceeding some threshold value to be priority areas for investigation.

Hit ratio is the performance measure that many authors use for evaluating GP. It is the ratio of optimal search area divided by the total search area. In one study, this ratio averages around 10 to 20 percent (Canter et. al 2000, p 472). An example of a performance measure that I would not recommend is the distance from the peak risk index location to the actual serial criminal's residence (Snook et. al 2004, p111). Such a measure does not match the investigative needs of police (to prioritize investigations by geographic area) nor is it sensitive to the many limitations of GP.

Serial Crime

Given an arrested serial criminal, let:

- $i = 1, \dots, m$ crimes in time sequence
- t_i = times of crimes (day number)
- $X_i = (x_{1i}, x_{2i})$ be crime-scene coordinates (e.g., body dump sites)
- $Y = (y_1, y_2)$ be coordinates of the serial criminal's residence.

We can easily add another index to represent a sample of serial criminals if needed at some point.

Question: what other variables should be routinely coded? Crime type(s), day of week, time of day, etc.? Locations of other anchor points?

Total Search Area

Total search area evolves and expands as a crime series unfolds, so its useful to consider the total search area at the i th crime in a series:

$$S_i = f(X_1, X_2, \dots, X_i) \quad i = 3, \dots, m$$

An often-used total search area is as depicted in Figure 1, a rectangle oriented along a street network (usually north/south and east/west), constructed from the minimum

rectangle, M_i , that includes all crime sites plus an extension, E , in case the criminal resides nearby but outside the minimum rectangle.

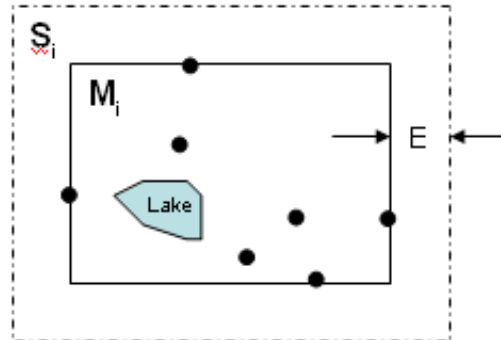


Figure 1. Total Search Area

Note that the lake in Figure 1, and similar irrelevant areas, should be subtracted from M_i and S_i .

Question: Can E be estimated empirically? Is it a function of crime type, population density, smaller patterns in the space and time series of points? When is $P(E) \neq 0$? Can this probability be estimated conditionally?

Question from the literature (e.g., Rossmo 2000 p 181-184): Does S_i increase over time, due to more than poor estimates given small sample sizes? Can we use inter-point distances to test this question?

Grid System

I think that a grid system, rather than continuous risk functions, provides some useful simplifications. Let the grid system for S_i be:

$j = 1, \dots, n_i$ be an index of grid cells

A_j = area of grid cell j (all grid systems used to date have been uniform, but they could be made to be variable in cell size)

$C_j = (c_{1j}, c_{2j})$ be the coordinates of the centroid of grid cell j

$j \in J_i$ = the set of relevant grid cells in S_i

$A_{S_i} = \sum_{j \in J_i} A_j$ be the total relevant search area (excluding lakes, etc.) from S_i

Question: What is an appropriate size for grid cells? Obviously, if too large then the approximation is too gross, if too small then there is a waste of computer processing time. Grid cells should probably be of pixel size, small enough to give the appearance of

a smooth surface when viewing S_i . Should they be constant across sets of serial crimes and jurisdictions?

Risk Index

Travel-to-crime research (e.g., Capone and Nichols 1975, Rossmo 2000 pp 99-111) suggests using a distance decay function for each serial crime i such as the following:

$$r_i(p) = F_i e^{-\beta d_i} \quad (1)$$

where

p = point in S_i

F = normalizing constant, chosen so that $0 \leq r \leq 1$

β = parameter to be estimated

$d_i = |X_i - p|$ the distance from serial crime i to point p

The distance function may be straight line distance, Manhattan block, or street network optimal travel distance (or time). The function may include a buffer zone surrounding X_i within which the risk index is zero, assuming the serial criminal takes precaution not to commit crime too near to home.

If the cross-sectional sample of crimes come from much different areas, like the sample of 79 U.S. serial killers analyzed by Canter et. al (2000), then it is necessary to model or otherwise account for differences in scale in order to apply distance decay functions. Canter et. al (2000) normalize using the average distance between offenses or the QRange function that assumes a linear structure to S_i .

The total risk index for the i th serial crime (assuming that the series is still unfolding) for point p should be the average of the $r_j(p)$:

$$R_i(p) = \sum_{k=1}^i r_k(p)/i \quad (2)$$

If a sum were used, then R would arbitrarily increase over time at all points, and thus potentially be confusing in practice. Alternatively, $R_i(p)$ could be scaled so that its maximum is always a constant, say 100.

Question: A complication raised by Rossmo is on crimes that are close to each other. Should such crimes be considered as non-independent and be collapsed to a single crime? If there are non-independent crimes included in (2), they would otherwise increase the risk surface falsely in the vicinity of those crimes.

Performance Measures

Suppose that the serial criminal's residence, Y , is in grid cell j^* which has centroid C_{j^*} . Let $J \in J_{i^*}$ be the set of grid cells in such that:

$$R_i(C_j) \geq R_i(C_{j^*})$$

Then the optimal search area has area

$$O_i = \sum_{j \in J_{i^*}} A_j$$

and the hit score percentage (Rossmo, p 201) is

$$H_i = 100O_i/T_{S_i} \quad i = 2, \dots, m. \quad (3)$$

To accommodate a "commuter" serial criminal living outside S_i , (3) can be extended as follows:

$$H_i' = 100N_i/A_{M_i} \quad i = 2, \dots, m. \quad (4)$$

For marauders,

$$N_i = O_i \text{ as in (3)}$$

For commuters,

$$N_i = A_{D_i} = \text{Area of rectangle } D_i \text{ (center on rectangle } M_i \text{) that has the serial criminal's residence on its boundary}$$

$$A_{M_i} = \text{Area of minimum bounding rectangle } M_i$$

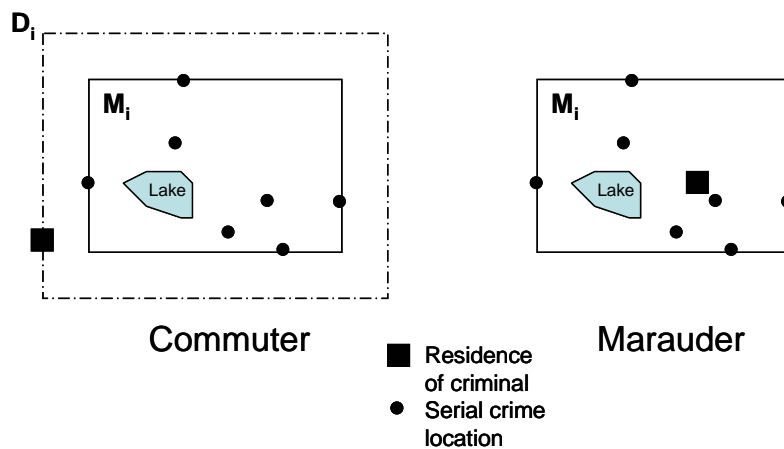


Figure 2. Commuter versus Marauder

For marauders, $H_i' \leq 100$ and has an interpretation similar to that of (3), but for commuters, $H_i' \geq 100$ and the distribution of such values has information useful for determining E , the extension of M_i in Figure 1. Figure 3 illustrates the use of the extended hit ratio to implement geographic profiling for use with an ongoing, unsolved crime series (or for use in hold-out sample validation as below).

Suppose we have a sample with many sets of serial criminals and their crimes, have applied an optimal GP procedure, and determined a distribution of extended hit ratios (H_i 's) as in Figure 3.

For marauders (with $H_i' < 100$), choose a threshold value to define the optimal search area. The threshold depicted, say with value 30, includes some acceptably large percentage of cases, P_1 . These case had optimal search areas that were 30 percent or smaller in area compared to their minimum bounding rectangles. For identifying optimal search areas in a hold-out sample case, or new case without knowing the serial criminal's residence in advance, aggregate grid cells starting with the one with the highest risk index. Add more grid cells, working down the list of grid cells sorted by risk index in descending order, until 30 percent of the minimum bounding rectangle is included.

For commuters, select a threshold greater than 100 on the horizontal axis so that an acceptably high percentage, P_2 , of commuters is included, say 120 for discussion purposes. Set the extension for optimal search to be the rectangle centered on M_i that is the multiple 1.20 in area of M_i .

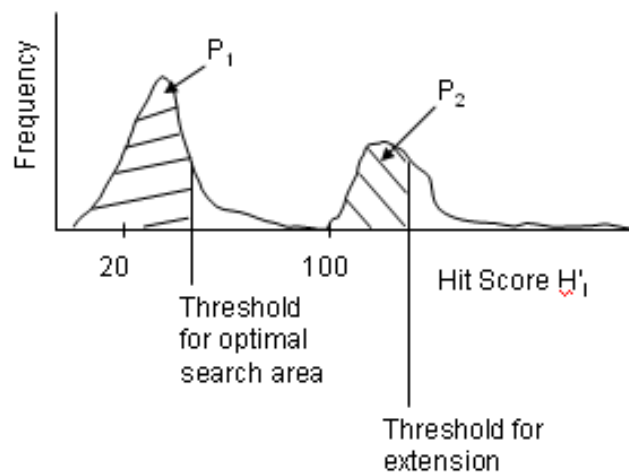


Figure 3. Example of how to use distribution of extended hit score to implement geographic

With an extension to M_i included, its possible to recode commuters and marauders, with commuters being all criminals residing in S_i (Figure 1). Any criminal outside the extended area would be a commuter.

Note that additional work on performance measures is needed for the case where the serial criminal has two or more anchor points around which he commits crimes. The hit score for such cases likely will be inflated relative to those that have a single anchor point. The reason is that multiple anchor points 1) are likely relatively widely distributed throughout a search area, and 2) the optimal search area is likely to include separated areas (local maxima) from each anchor point. Nevertheless the validation criterion only rewards finding the home residence, while also uncovering other anchor points. It may be that different threshold values for hit scores are necessary for implementing GP for single anchor points versus multiple search points.

Question: Is a model possible for H_i , a function of crime type, population density, smaller patterns in the space and time series of points, or other variables? Can such a model be estimated from a sample of serial crimes?

Samples

For low volume serial crimes, like serial criminals and perhaps rapists, it may be necessary to collect data across large regions or an entire country. Canter et al. (2000) collected 79 convicted U.S. serial criminals from 1960 to about 2000.

For relatively high volume crimes, like burglary and robbery, it may be possible to collect and analyze serial crimes within jurisdictions. For example, NYPD and the Detroit Police Departments record data on serial criminals.

Validation Studies

Next is the approach to optimize over any proposed GP methodology. It is assumed that the methodology is implemented in research-oriented software that includes empirical estimation and optimization over sub-methods and their parameters. Such software would optimize and apply a GP method across a large sample of solved serial crimes and criminals, and tabulate multiple error measures.

For a given GP methodology and data sample, it is necessary to optimize over all choices and parameter values. This can be done by writing a batch program that runs all alternatives through a sample of solved crime series (with known residences of serial criminals), via complete enumeration to minimize a measure such as H_i . It is necessary to optimize over alternative data normalization procedures, distance decay functions and their parameters, and adjustments to distance decay functions such as buffers and buffer sizes. For example, Canter et. al (2000) did a complete enumeration of β values for (1) an interval, plus did the same for additional distance decay functions. The optimal parameters are the ones that minimize the average adjusted hit score for commuters, H_i , across all sets of serial crimes from an estimation sample.

To date estimation and holdout samples have not been used in validation studies for serial criminals. Validation should use a resampling scheme, where, for example, a random subset of 90% of the sets of serial crimes is used for estimation and yielding a distribution of H_i values. This distribution is then used to select a thresholds as in Figure 3 to define optimal search areas and minimum rectangle extensions for each crime series. Some percentage of cases will have the actual criminal residence inside such threshold values, others will be outside. The research can summarize the results.

This process can be repeated several times to make the most of data samples; for example, ten times over take a new random sample of 90 percent of cases for estimation and remaining 10 percent for hold-out analysis. The forecasting literature has demonstrated the value of including simple benchmark methods and other comparisons such as judgmental methods when validating sophisticated methods. Often times simple methods perform as well as complex ones. Validation of GP should include alternative GP methods/models run in parallel.

References

- Canter, D. (2003). *Mapping Murder: The Secrets of Geographic Profiling*, Virgin Books, UK.
- Canter, D., T. Coffey, M. Huntley, and C. Missen (2000). "Predicting serial killers' home base using a decision support system," *Journal of Quantitative Criminology*, Vol. 16, No. 4, pp 457-478.
- Capone, D.L. and W.W. Nichols (1975). "Crime and Distance: An Analysis of Offender Behavior in Space," *Proceedings of the Association of American Geographers*, Vol. 7, pp 45-49.
- Kocsis, R.N. and H.J. Irwin (1997). "An analysis of spatial patterns in serial rape, arson, and burglary: the utility of the circle theory of environmental range for psychological profiling," *International Journal of the Australian and new Zealand Association of Psychiatry, Psychology, and Law*, Vol. 4, pp 195-206.
- Rossmo, K., (2000). *Geographic Profiling*, CRC Press, Boca Raton.
- Snook, B., P.J. Taylor, and C. Bennel (2004). "Geographic profiling: the Fast, frugal, and accurate way," *Applied Cognitive Psychology*, Vol. 18, pp 105-121.