Selection of a spatial hypothesis in geographic profiling using graph theory

Marie Trotta
University of Liège/ Geomatics Unit
17 Allée du 6 août
Liège, Belgium
Marie.Trotta@ulg.ac.be

Abstract

Geographic profiling is based on the spatial hypothesis of a distance decay effect around the offender's anchor point and can be very useful for delineating prior search areas for individual investigations such as DNA testing. However, this spatial hypothesis fails to model a significant part of Belgian serial sexual offenders. In previous study, a complementary hypothesis, the minimisation of the variance (**MOV**) for the distances between the offender's anchor point and the crime site, exploiting recent literature observation, is proposed to model unexplained behaviours. A systematic comparison of the distance decay and the **MOV** hypotheses show that their overall capacities to provide an effective geoprofile are similar, but their performances may differ, depending on the geometric pattern of crimes. Here, we use graph theory to describe the favourable patterns for the **MOV**, in order to select which spatial hypothesis is best suited for a new investigation. While star graphs represent patterns for which both hypotheses could be validated, wheels with preferred directions are only explained by the **MOV** hypothesis. Real patterns on the road network are a-posteriori evaluated thanks to the generation of the shortest path between the best solution of the **MOV** and all the crime locations thanks to the Dijkstra algorithm.

Keywords: geographic profiling, spatial analysis, behavioural hypotheses, graph theory.

1 Geographic profiling, its classical spatial hypothesis and the Belgian context

Spatial analysis has been widening its scope to various research domains. Crime mapping is undoubtedly a field of application oriented towards important societal issues, analysing the pattern of crimes in order to improve prevention or provide new clues for investigations.

Among its sub-disciplines, geographic profiling (**GP**) is defined as a methodology of investigation that uses the locations of a series of connected crimes to determine the criminal's most probable area of residence [4]. **GP** is generally based on the spatial hypothesis of a uniform distance decay effect around the offender's anchor point [7] and can be very useful for delineating prior search areas for DNA testing.

Belgium is characterized by small inter-city distances so that **GP** needs to be very accurate to be operational. When analysing Belgian data on serial sexual offenders, however, we observed a long mean travelled distance, around 15 km, with high inter-offenders variations. Thus, the classical distance decay hypothesis seems inappropriate for many Belgian crime series.

2 The minimisation of variance (MOV) as a complementary hypothesis

Because the distance decay did not suit to describe Belgian serial sexual offenders, new research methodologies are needed to implement **GP** in this territory. In [6], a complementary heuristic and its associated methodology are proposed: a serial offender, by his repetitive behaviour, tends to minimise the variances in his journeys between his anchor point and the crime sites; this is called the **MOV** hypothesis. It exploits the recent observations of a small "intra-offender" variance compared to that of the "inter-offender" one [3] and has the advantage to focus on the specificities of the individual.

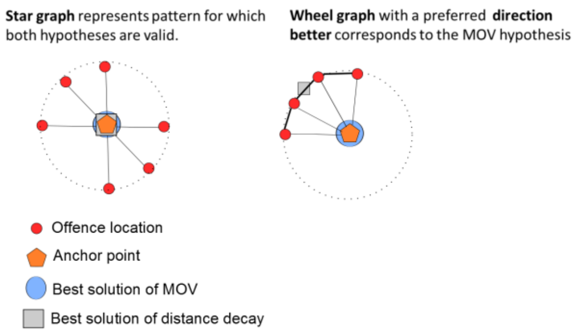
When applied on a set of the Belgian series, the distance decay and the **MOV** modelled a similar proportion of the behaviours. The objective of this paper is to help investigators choosing the right method for their case by precisely describing the spatial configurations where the **MOV** hypothesis and its associated method should be preferred to the classical distance decay functions. In order to have fewer constraints, we compare a linear distance decay function with the MOV. Other methods, depending on the distance decay hypothesis, such as the journey-to-crime and Bayesian journey-to-crime available in CrimeStat III [2], would provide other profiles. However those require a subjective choice of the distance decay function or a calibration with solved data, which is not in line within an operational purpose.

2.1 Graph theory to discern the patterns

Graph theory provides an interesting way for describing the pattern. The star graph [5] (Figure 1a) for which the central node corresponds to the offender's anchor point illustrates patterns where both assumptions may provide a satisfying solution for the geographic profile.

However, the **MOV** assumption will only be verified if the concept of weighted graph is introduced. In such graph, each vertex is associated with a value that could be a cost, a weight. According to our hypothesis, this value is the network distance between the offender's residence and the crime location. To obtain a small search area, the weights must be similar on each vertex, what is not required for the application of the distance decay.

Figure 1: a. Star and b. wheel graphs for describing patterns suitable to the **MOV** hypothesis



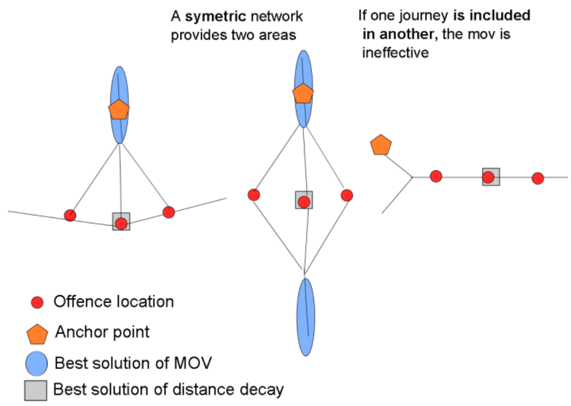
The second graph (Figure 1b) corresponds to a wheel [4] restricted to one preferred direction. A wheel on n points (the crime locations and the residence) is the union of a ring and a star [5]. It typically reflects a pattern for which only the **MOV** will give satisfying results. In this situation, the offender chooses to travel in a specific direction. Such direction may be influenced by a better knowledge of this region or the configuration of the road network.

2.2 A-posteriori evaluation of the patterns

The evaluation of patterns requires an estimation of an unknown node: the anchor point. As illustrated in Figure 2, a linear pattern for crime locations may correspond to a wheel or a star (a and b) or even other configurations for which a journey is totally included in another. Such pattern may contradict the **MOV** hypothesis depending on the location of the anchor point.

For this reason, the pattern can only be evaluated a-posteriori with the best solution determined according to the chosen hypothesis. This solution is computed with a raster approach so that the node can be located everywhere on the network (see [6] for the computational procedure). The Figure 3 illustrates the best solution provided by the **MOV** for a real case. The shortest paths between this solution and the crime locations on the road network is generated using Dijkstra algorithm [1]. They create a pattern for which the investigator may check the independence of the journeys. In Figure 3, the pattern is very close from the wheel in accordance with the **MOV** hypothesis. The residence was finally located near the best solution. In [6], a jackknife procedure is proposed to evaluate the sensitivity of such method to every crime location. In the case of great variability, the patterns created by each solution of n-1 locations can be compared in order to evaluate which is valuable for both hypotheses. The investigator may, then, choose between the generated profiles according to the additional information of the investigation.

Figure 2: A linear pattern might correspond to very different radically distinct situations.



3 Conclusions

This paper discusses the necessity to analyse the pattern of crime locations on the road network in **GP.** It provides a way for choosing between two hypotheses that have proven to model a similar share of Belgian sexual offender's spatial behaviour. It also demonstrates that both hypotheses may be validated for some typical patterns.

Acknowledgment

The research achieved by the author is funded under a Belgian F.R.S.-FNRS fellowship.

References

[1] Edsger Dijkstra, A note on two problems in connexion with graphs". *Numerische Mathematik*, 1*:* 269–271, 1959.

[2] Ned Levine, CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 3.3). Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. July. 2010

[3] Samentha Lundrigan., Sarah Czarnomski, and Marc Wilson, Spatial and environmental consistency in serial

sexual assault. *Journal of Investigative Psychology and Offender Profiling*, **7**(1):15-30, 2010.

[4] Kim Rossmo, *Geographic profiling*, CRC Press, Boca Raton, 2000.

[5] Keith J. Tinkler, An introduction to graph theoretical methods in geography. *Concepts and techniques in modern geography*,14:3-55. Norwich Geo Abstracts Ltd, 1977

[6] Marie Trotta. New hypotheses on serial offender's spatial behaviour . In Hardy Pundt and Lars Bernard editors. *Proceedings of the 1 st AGILE Phd School*. Wernigerode: Shaker Verlag, 1-9, 2012.

[7] Vere van Koppen, Henk Elffers, and Stijn Ruiter, When to refrain from using likelihood surface methods for geographic offender profiling : An ex ante test of assumptions. *Journal of Investigative Psychology and Offender Profiling*, 242-256, 2011.

Figure 3: Evaluation a-posteriori of the likelihood of the MOV thanks to the shortest paths from the best solution