Lasso: Linkage Analysis of Serious Sexual Offences

Don Casey, Phillip Burrell

Lasso : Linkage Analysis of Serious Sexual Offences A Decision Support System for Crime Analysts and Investigators

Don Casey [1, 2], Phillip Burrell [1]

[1] London South Bank University, Borough Rd, London SE1 0AA, U.K
[2] Metropolitan Police Service, London, U.K
{ Don Casey caseyd@lsbu.ac.uk , Phillip Burrell phillb@lsbu.ac.uk }

Abstract. One of the most important considerations when investigating a serious sexual offence is to find if it can be linked to other offences. If this can be done then there is a considerable dividend in terms additional evidence and new lines of enquiry. The central problem is the construction of a satisfactory typology of these crimes, but little progress has been made. It is the authors' contention that difficulties arise from the inadequacy of the adoption of the classical or 'crisp set' paradigm. Complex events like crimes cannot be described satisfactorily in this way and it is proposed that fuzzy set theory offers a powerful framework within which crime can be portrayed in a sensitive and perceptive manner that can enhance the search for associations between offences

Keywords: fuzzy systems, decision support, crime linkage

1 Introduction

The most influential crime classification system has been that proposed in the Crime Classification Manual [1], which is the work of senior F.B.I agents and advances the notion of an organized / disorganized dichotomy in serious offences and was developed from interviews with offenders [2]. The basis of this approach is that crimes can be differentiated by the level of planning associated with them. The authors extend this to assert that the dichotomy can be applied to the offender so that organized and disorganized crimes are committed by individuals who can be divided into discrete groups with distinct characteristics.

Very serious objections have been made to the methodology employed by the F.B.I: only 36 offenders were interviewed, no attempt was made to ensure this group was representative and the interviews conducted were not structured or consistent. An evaluation of the typology by distinguished psychologists working in the field [3] applied to 100 serial murderers provided no support for it.

In the most comprehensive research programme into the linkage of serious sexual offences by Grubin et al [4], the authors propose

Our starting premise is that rape attacks can be organized into distinct types

It is certainly the aim of investigation into any field to initially classify the objects contained within it but it is the contention of this paper that that although rapes can be organized into types that these will be far from 'distinct'. And that the attempt to discriminate between crimes in this way is likely to be not only barren but actively misleading in that they will be forced in to mutually exclusive groups that will misrepresent their complexity; a view arrived at after many years of research by Canter[5] one of the area's foremost investigators

The actions of any individual criminal may therefore be thought of as a subset of all the possible activities of all criminals. Some of this subset overlaps with the subsets of many other criminals, and some with relatively few. It therefore follows that assigning criminals or crimes to one of a limited number of 'types' will always be a gross oversimplification

Canter and his associates who are identified with Investigative Psychology have published numerous studies [6], [7] on sexual assault, homicide and other serious crimes but have been unable in any of them to construct a satisfactory typology with even the most relaxed rules of assignment [8], [9].

Grubin is forced to propose a highly redundant 256 element taxonomy of serious sexual assault. A classification system in which many, if not most, elements will never occur cannot be satisfactory. The assumption of the crisp set paradigm in this research appears to be the cause of the problems relating to these difficulties. This can be illustrated by a simple description of a crime such as 'a very violent assault on a middle-aged woman by a young man' which cannot be properly expressed in terms of crisp sets. It can lead to either the misallocation of fundamentally different offences to the same place or to crimes that bear strong resemblances to each other being regarded as entirely unconnected, a phenomenon referred to as linkage blindness [4] of which researchers are fully aware but have been unable to address.

2. Applicability of Fuzzy Systems

2.1 Geographic Profiling as a Fuzzy System

There has been only one area of research into crime clustering that has been widely acknowledged to have been successful in linking crimes and it is instructive that this can be regarded as relying on fuzzy sets, although this has not as yet been acknowledged.

There are various schools of geographic profiling [10] [11] but all are based on the idea proposed by environmental criminology [12] relating to 'mental maps', i.e. that individuals, including criminals, are much more likely to conduct their activities in

areas known to them. In terms of offenders this means that they are likely to live close to, or have some other association with the locations of their crimes. This has been adapted to construct a variety of systems so that the co-ordinates of a number of offences, known or believed to be linked can be input to a function that returns a 'jeopardy surface'. This not only returns information regarding the offender's lifestyle but presents an area in which further crimes committed by him may be discovered although they are presently unlinked. Most importantly it indicates those areas where the offender is likely to live or have some other strong association with such as employment or previous address. Slightly different techniques are employed to differentiate the 'most likely' from the 'less likely' areas.

We can regard this, the most successful approach in the area as a fuzzy system in that a set of discrete values, geographical coordinates, are input to an algorithm that assigns a degree of membership of the fuzzy set 'offender has an association with' , or something similar, to other co-ordinates in the region of the offences. Typically the result will be a map that resembles a series of concentric circles or clusters that are strongly reminiscent of a family of fuzzy sets.

2.2 Crime and Fuzzy Sets

Fuzzy set theory [13] allows us to represent crimes and criminals as highly descriptive objects in the concept space and to undertake experimental procedures to discover what the most significant differentiating features are, using mathematically and logically sound methods. We have been fortunate in being successful in obtaining data on 574 serious sexual offences from the Serious Crimes Analysis Section of the U.K National Policing Improvement Agency. We have excluded those offences that do not relate to serial stranger rapes, by which we mean a set of rapes committed by a single individual, unknown to the victim. This results in a much narrower dataset (n = 110, development set n = 83, test = 27). The development set which has provided the results in this paper consists of 28 series of average length 2.96.

2.2.1 Fuzzy Similarity

We can define the universe of crimes as a data set (X) of n elements

$$X = \{ x_1, x_2, x_3, \ldots x_n \} \qquad (1)$$

Where each crime (x_i) is defined by j features

$$x_i = \{ x_{i1}, x_{i2}, x_{i3}, \ldots x_{ij} \} \qquad (2)$$

In this case they could be the characteristics identified as significant by Grubin. We can then regard a crime as a datapoint in j dimensional space. If we introduce an index crime (x_m) into this space for comparison with another crime x_k we can define a fuzzy relation that captures the concept 'x_k is close to x_m', by using a membership function that measures the Euclidean distance between crimes and divides it by some

value ‘c’, where c is a positive real number whose value is chosen as a reasonable representation of the concept ‘close’ in that application.

We have used c = d / 2 where d = the average distance between crimes. As a result the introduced crime becomes the centre of its own cluster. We can then define a number of crisp subsets around this centre by restricting membership of these sets to those elements that have a degree of membership which is greater than or equal to some value α in [0, 1]. This results in a crisp subset ${}^{\alpha}A$ of the fuzzy set A which is itself defined on the universal set X. This crisp subset is known as an α-cut of X:

$$ {}^{\alpha}A = \{ x \in X \mid A(x) \geq \alpha) \} \qquad (3) $$

Here the crisp set ${}^{\alpha}A$ contains all the elements of the universal set X whose membership degrees are greater than or equal to the value of α. In this case we would generate a number of nested sets around the index crime, membership of which would reflect their ‘closeness’ to it. This illustrates very graphically the search strategy for crime analysts and investigators when one of these very serious crimes occurs and it is required to look for other crimes committed by the same offender.

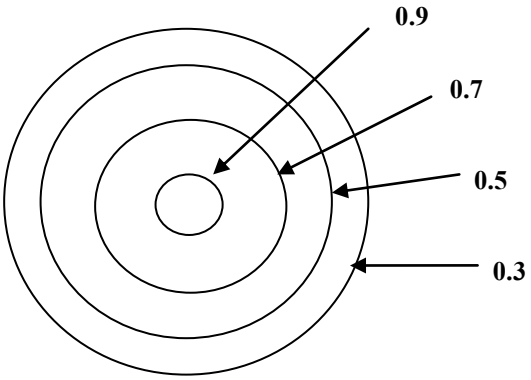

Fig. 1. closeness to the index crime

2.2.2 Fuzzy c-means clustering

Fuzzy c-means clustering [14] is the most widely used fuzzy clustering strategy and effectively addresses the problem raised by Canter of exclusive types by defining a family of fuzzy sets on the universe X so that the sum of degrees of membership of all the classes of any datapoint is unity, there will be no empty classes and no class that contains all the datapoints. This is an iterative optimisation technique of the objective function below where a degree of fuzziness $1 \leq m < \infty$ is specified and elements assigned degrees of membership of the clusters until some termination criterion has been reached.

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2$$

(4)

u_{ij}^{m} is the degree of membership of x_i in cluster j and c_j is the cluster centre

2.2.3 Initial results

This data is extremely rich and comprises over 370 fields with dichotomous values relating to every feature of the crime. By employing only those variables that inform the domains identified as significant by Grubin and excluding ill-defined or poorly recorded data we have reduced this to 41: sex(11), escape (11) and control (19). These reflect the offender / victim transaction that is at the heart of this serious offence. A problem that arose was that these concepts, unlike those usually identified with fuzzy membership functions like age, height etc are cumulatively or hierarchically scaled. This makes the use of a conventional membership function difficult. In order to overcome this we have proposed that the *amount* of these activities can be measured, i.e. the number of separate sexual, controlling or escape-centred actions.

The membership function we have employed derives closely from the techniques used by Canter [9] and Investigative Psychology which emphasizes the frequency of variables and their co-occurrence within crimes. Each variable is assigned a value c / n where c is the number of times it occurs in the dataset n, so a variable occurring 14 times has a value 14/83, this simple technique ensures a commonly occurring variable has a lower weighting than a more unusual action. It also means that a very simple form of learning analogous to experience is enabled as crimes are added and the distribution of variables changes. These values are then summed for all the variables for each dimension of the index crime. This value is then normalized using the highest sum in the dataset as divisor and as a result a degree of membership of each dimension can be assigned to each offence.

Our initial results have been encouraging using both c-means clustering and fuzzy similarity measures .By varying the fuzzy exponent (m) from the crispest fuzzy value of 1.25 to 3 and the number of clusters from 2 to 5 we have been able to evaluate the significance of these variables, something which has not before been done in this field. For instance using a fuzzy exponent of 1.25, 3 dimensions and 3 clusters, 15 of the 28 series of crimes where assigned a membership value ≥ 0.9 in one cluster. However it is unlikely that the greatest utility of this technique will be at this low level of fuzziness. Table 1 indicates an example of the very high level of consistency across the dimensions that can be achieved in linked series (bold underline) and the resulting strong level of clustering at m = 1.25 in the clusters we identify as A,B and C.

This differs from previous attempts to classify crimes in that it is as empty as possible of psychological precepts, as is reflected in the cluster names. As a result the assignment of crimes to classes is as a result of their positions in the concept space rather than their perceived association with wider psychological principles. If clustering is successful then there undoubtedly are psychologically meaningful alternatives for these labels but this would lie within the remit of criminal psychologists

table 1.

Dimensions				Membership		
Crime	Control	Sex	Escape	A	B	C
x53	44	16	6	0.99	0	0.01
x54	44	2	6	0.95	0.01	0.04
x50	87	2	32	0.12	0.05	0.83
x51	59	2	32	0.29	0.04	0.67
x52	87	2	32	0.12	0.05	0.83

Overall where m=1.25 88% of crimes are assigned to a set with > 0.80 degree of membership: type A = 12%, type B = 20%. Type C = 53%. 15 of the 28 series are assigned to a single set

In terms of the closeness of crimes we have looked at the overall number of possible comparisons between the 83 offences and then at the relative differences in closeness between the linked and unlinked crimes, i.e. those committed by different individuals and those committed by the same. The relatively small number of comparisons of linked offences is as a result of the high number of series of two crimes

table 2.
number of comparisons between crimes

	All	Unlinked	Linked
total	3403	3393	110

The very marked difference in the degrees of closeness between linked and unlinked crimes is encouraging, particularly as the gap between the two widens very significantly as degree of closeness increases. So at the average degree of closeness of 0.23 a linked crime is a little less than twice as likely to appear while at the highest degree of closeness, > 0.9, a linked crime is seven times as likely to figure.

The implication in terms of assisting crime analysts and investigators is clear: at the highest levels of closeness the greatest possibility of a match occurs and this

declines in line with closeness. It provides the clearest and most effective search strategy for analysts in order to maximise their chances of a positive 'hit'.

table 3.
degree of closeness between crimes

	All	Unlinked	Linked
above ave	1387	1309	78
> 0.23	40.76%	38.58%	70.91%
above 0.5	674	632	45
	19.81%	18.63%	40.91%
above 0.6	464	427	37
	14%	12.58%	33.64%
above 0.7	286	256	30
	8.40%	7.54%	27.27%
above 0.8	124	110	14
	3.60%	3.24%	12.73%
above 0.9	36	29	7
	1.00%	0.85%	6.36%

3. Conclusion

The problem of rigid typology that has hampered this area of research is precisely the one that fuzzy sets avoids. Because of the nature of the area under investigation any crisp classification method is bound to fail. Either a large number of crimes elude classification as in Investigative Psychology or an enormous system that specifies 256 type of stranger rape, which is itself a small subset of rape has to be proposed. These crimes must be placed somewhere if research is to be fruitful. The answer may be that instead of belonging nowhere or in a tiny compartment of a huge structure, that they belong in several places at the same time to differing degrees.

The success of geographic profiling in modelling criminal conduct is illuminating. An algorithm that can be regarded as a membership function uses a number of geographic locations, of linked crimes, as input in order to assign degrees of membership to a larger set of geographic points and thereby construct a fuzzy set. And in so doing effectively assist crime investigators in highlighting areas in which to find the offender. In this case longitude and latitude are the relevant dimensions on

which the system operates. If one generalizes from this and is able to identify the pertinent dimensions that describe a landscape of actions rather than the physical landscape associated with crime then the achievements of geographic profiling may be possible. There is also an interesting symmetry in that the input to geo-profiling systems is a set of linked crimes and the desired output of LASSO is also a set of linked offences.

The 'set of meaningful numbers' called for in the earliest days of research [15] into this area can be achieved by using fuzzy set theory and in thus allow empirical research rather than the experiential and anecdotal or hypothetical approaches that have so dominated the field for so long and so unproductively.

References

1. Douglas, J. E., Burgess A.W, et al. (1992). *Crime classification manual: A standard system for investigating and classifying violent crime*. New York, Simon and Schuster
2. Ressler R.K and J. E. Douglas (1985). Crime Scene and Profile characteristics of organized and disorganized murderers *FBI Law Enforcement Bulletin* **54**(8): 18-25.
3. Canter D.V, Alison L.J, et al. (2004). The Organized/Disorganized Typology of Serial Murder. Myth or Model ? *Psychology, Public Policy and Law* **10**(3): 293-320.
4. Grubin D, Kelly P, et al. (2000). Linking Serious Sexual Assault through Behaviour *Home Office Research Study 215. London.*
5. Canter D (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology* **5**: 23-46.
6. Santilla, A., H. Hakkanen, et al. (2003). I*nferring the Crime Scene Characterstics of an Arsonist. Interrnational Journal of Police Science and Management* **5**(1)
7. Hakkanen, H., P. Lindof, et al. (2004). Crime Scene Actions and offender characteristics in a sample of Finnish stranger rapes. *Journal of Investigative Psychology and Offender Profiling* **1**(2): 153-167.
8. Salfati E.C and Canter. D.V (1999). Differentiating Stranger Murders: profiling offender characteristics. *Behavioural Sciences and Law* **17**(3).
9. Canter D.V, Bennell C, et al. (2003). Differentiating Sex Offences. *Behavioural Sciences and Law* **21**.
10. Rossmo D.K (2000). *Geographic Profiling*. Boca Raton Fl, CRC Press.
11. Canter D, Coffey T, et al. (2000). Predicting serial killers' home base using a decision support system. *Journal of Quantitative Criminology* **16**(4): 457-478..
12. Brantingham P.J and B. P.L (1981). *Environmental Criminology*. Prospect Heights, IL, Waveland Press.
13. Zadeh L (1965). Fuzzy Sets. *Information and Control*(8): 228-353.
14. Bezdek J.C (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, Plenum.
15. Canter D (1985) *Facet theory : approaches to social research* New York, Springer-Verlag