LONDON SOUTH BANK UNIVERSITY

Decision Support, Crime Linkage and Analysis

Dr Don Casey [1] Prof. Phillip Burrell [1] D.C.I Nick Sumner [2]

{ caseyd4@lsbu.ac.uk phillb@lsbu.ac.uk nick.sumner@met.police.uk }

[1]London South Bank University

[2]Metropolitan Police Service (U.K.)

Research Motivation

The ultimate goal is to create a computer-based screening system that will allow routine and systematic comparison of serious offences on a national basis , selecting cases on the basis of their behavioural similarity that are appropriate for more detailed attention by detectives or crime analysts

Linking Serious Sexual Assault through Behaviour (Grubin, Kelly, & Brunsdon, 2000)

The development and test of theories and implementation of findings into computer-based, decision-support systems … has to be the proper basis for any professional derivation of inferences about offenders.

Offender profiling and criminal differentiation (Canter, 2000)

LONDON SOUTH BANK UNIVERSITY

Forging the links: Rape investigation and prosecution. **(HMIC/HMCPSI, 2012)**

Recommendation 1

We recommend that forces should initially consider every *'stranger' rape to be part of a pattern of serial offending, so that investigating officers consider the wider links to other crimes.

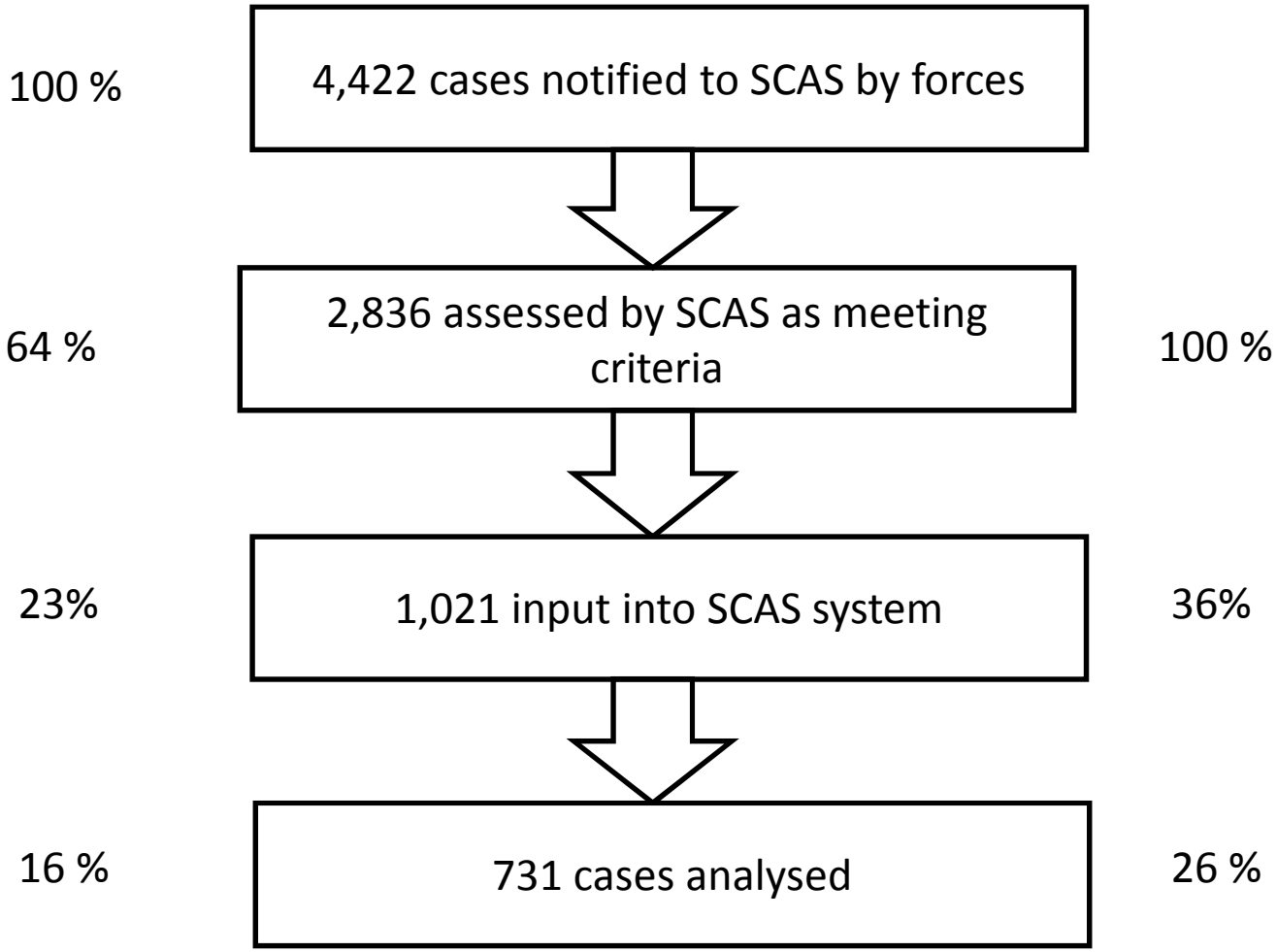
Also …
The core activities of **SCAS (Serious Crime Analysis Section) depend on computer-based gathering and analysis of information regarding crimes and individuals, often in large volume. This makes IT and system issues critical

** Rape where there is little or no previous relationship between victim and offender*

***The Serious Crime Analysis Section (SCAS) is a national service supporting investigations into serious sexual offences by strangers in order to identify the potential emergence of serial rapists at the earliest stage of their offending*.

LONDON SOUTH BANK UNIVERSITY

Throughput of cases from Forces to final SCAS Analysis (2010/2011)
(HMIC/HMCPSI, 2012)



LONDON SOUTH BANK UNIVERSITY

The SCAS System

Violent Crime Linkage Analysis System

Development by Canadian Police in 1990s of the original Violent Crime Apprehension Program (ViCAP) established by F.B.I. in the 1980s

Serious Crimes Analysis Section employ a variant of the ViCLAS database used throughout Europe , N. America and Australasia

Essentially a repository of crimes that brings together crimes in a standardised format

Development was practitioner-led with little input from Computer Science/ A.I.

'*Once they have conducted their background research they will draw upon their experience and expertise by conducting various structured queries on ViCLAS.* ***Each specialist will have his or her own approach to this process***'

RCMP ViCLAS website

Decision Support Systems

Widely used in industry, commerce

Interactive computer-based system which helps decision-makers utilise data and models to solve unstructured problems

fuzzy, complex problems for which there are no cut and dried methods

(Gorry & Scott-Martin 1971)

These systems designed to support and assist the decision making of the user not to replace it.

Not automated decision-makers

Original research by Grubin et al. posited 4 'behavioural domains' that are influential in identifying serial stranger rapes:

Control: behaviours directed towards gaining control of the victim, including the way in which the victim was approached and targeted, so that the sexual aspect of the attack can take place;

Sex: behaviours associated with the sexual component of the attack;

Escape: behaviours associated with leaving the crime scene or avoiding capture (some of which, such as wearing a mask or gloves, may actually relate to events that occur before the attack); and

Style: behaviours that reflect the offender's personality or offence style, but are not directly necessary for the success of the attack; many of these will be equivalent to "signature" behaviours described in the FBI research.

LONDON SOUTH BANK UNIVERSITY

DATA

- A dataset of stranger rapes was made available by the Serious Crime Analysis Section (SCAS) of the National Policing Improvement Agency (NPIA).
- The data related to 545 serious serial offences committed by 164 convicted offenders.
- There were 112 serial stranger rapes committed by 40 offenders and 22 unlinked crimes.
- This was divided into a development set of 83 offences in 29 series and a test set of 29 offences in 11 series.
- Feature selection was based upon Grubin's original 4 domain model.
- 'Style' excluded as least stable, least consistent

LONDON SOUTH BANK UNIVERSITY

DATA

Development and test datasets

Inititial dataset: 112 linked stranger rapes, 22 single stranger rapes

	n =	number of series	mean length of series	
Development 1	83	28	2.96	83 linked crimes
Test 1	29	11	2.64	29 linked crimes
Test 2	51			29 linked, 22 unlinked
Development 2	112	39	2.87	112 linked crimes
Test 3	134			112 linked crimes, 22 unlinked

3 Dimensional Crime Representation

CONTROL	SEX	ESCAPE
Administered drug	Kiss Face	Covered eyes
Bound victim	Kiss Chest	Gloves
Threatened victim	Kiss Other	Gagged
Sneaked up	Hugs	Covered mouth
Victim sleeping	Fondles	Told not to look
Asked assistance	Masturbate Self	Destroyed forensics
Offered ride	Fellatio	Condom
Engaged conversation	Vaginal Penetration Digital	Car used
Lay in wait outdoors	Vaginal Penetration Penile	False name
Location: park/recreation area	Vaginal Penetration Rear	Disguise
Location: retail business	Anal Penetration Penile	Blindfold
Location: residential	Ejaculate on body	
Force: some on resistance	Masturbates other	
Force: some no resistance	Cunnilingus	
Weapon: not involved		
Weapon: not displayed but threatened		
Weapon: displayed – not used		
Weapon: used		

LONDON SOUTH BANK UNIVERSITY

Testing Regime: Dimensions (domains)

Three dimensions (3D)

Sex, Control, Escape – original research

Four Dimensions (4D2C)

Divided Control domain into overtly violent actions (force, weapon etc. - C1) and engaging (approach, location etc. – C 2).

Four Dimensions (4D2S)

Sex domain also divided (rape actions – S1), other sexual actions (kissing etc. – S2) .

Five Dimensions (5D)

All 5 Dimensions (Control1, Control 2, Sex 1, Sex 2, Escape)

Crime and criminals

Complex area of behaviour

Concepts are imprecise

Relations not well defined

Cause significant problems for conventional classification systems , analysis and inferencing

Fuzzy mathematics … a methodology for dealing with phenomena that are vague, imprecise, or too complex or too ill defined to be susceptible of analysis by conventional mathematical means

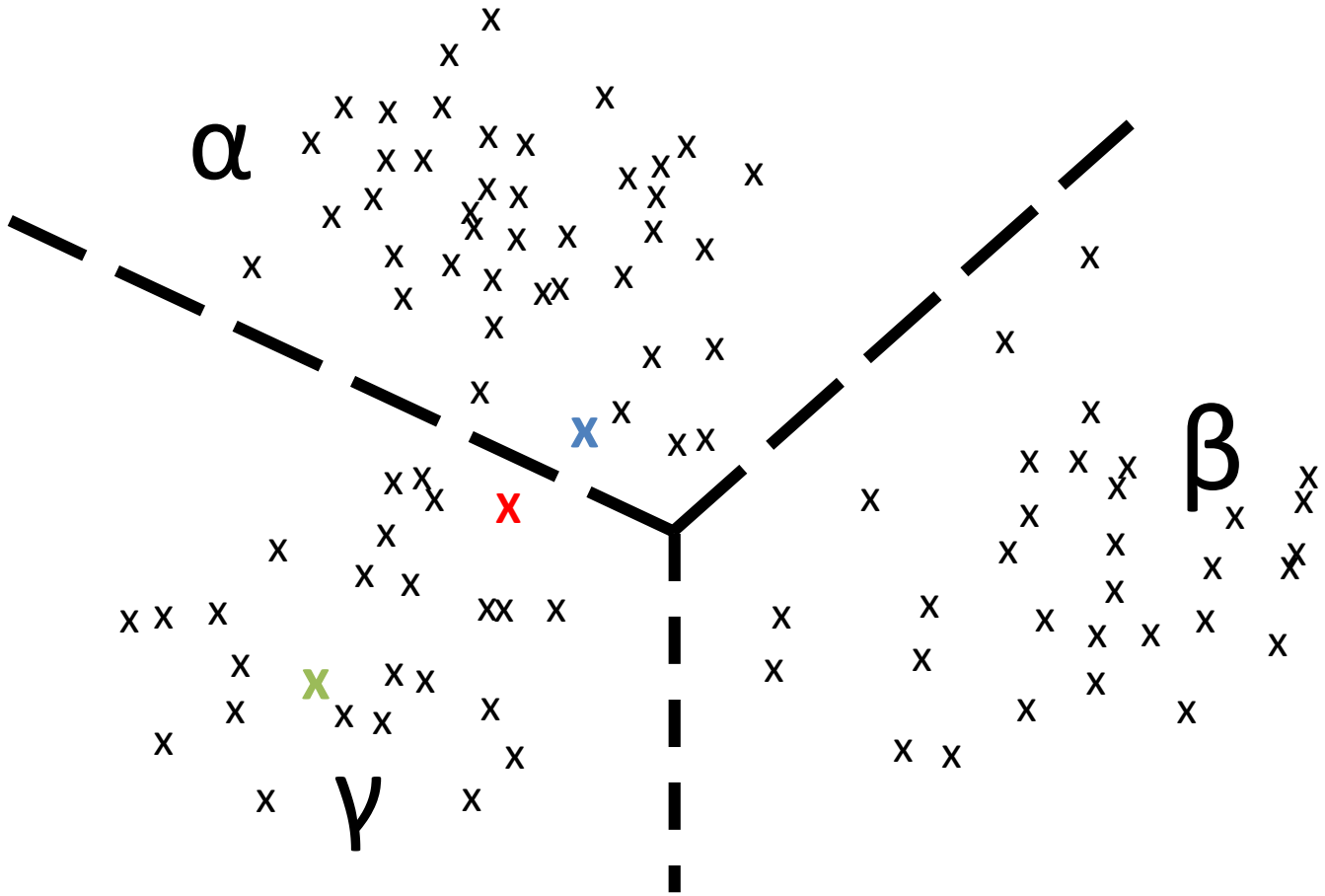
Kandel (1986)

3 Dimensional Degrees of Membership

Crime	series	CONTROL	SEX	ESCAPE
1	1	0.38	0.17	0.15
2		0.38	0.22	0.08
3		0.32	0.24	0.07
4		0.34	0.34	0.37
5		0.34	0.10	0.14
6	2	0.32	0.06	0.00
7		0.27	0.14	0.16
8		0.38	0.03	0.09
9	3	0.32	0.24	0.07
10		0.31	0.13	0.17
11		0.08	0.36	0.00
12		0.13	0.34	0.15
13		0.32	0.32	0.24
14		0.21	0.31	0.24

LONDON SOUTH BANK UNIVERSITY

Classifying crime



Crime are classified in an insensitive manner that takes little account of their positions relative to other crimes, particularly near boundaries

Data Clustering

The basic principle of the technique is "to allocate individuals to mutually exclusive groups such that individuals within a group are similar to one another while individuals in different groups are dissimilar"

Grubin et al.(2000)

Fuzzy Clustering

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

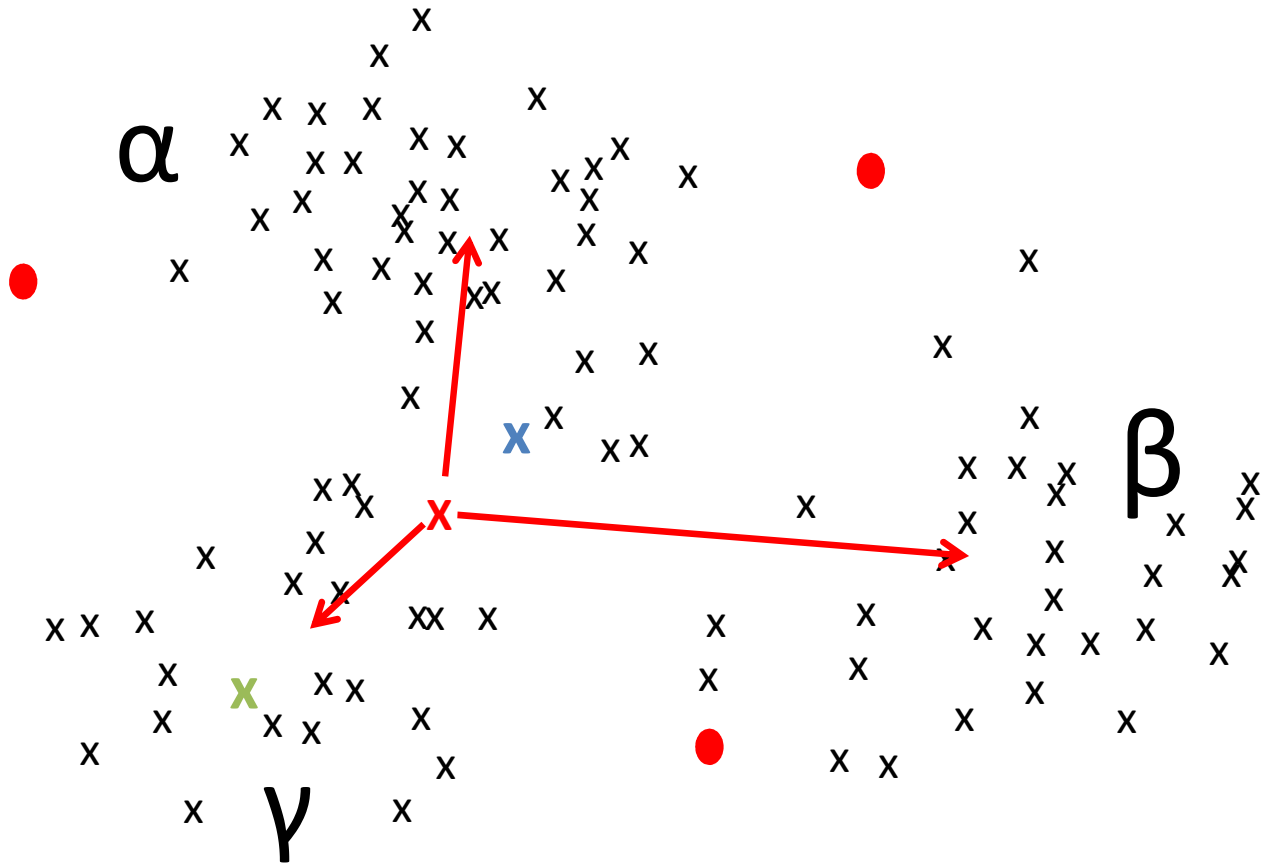
Canter (2000)

Fuzzy Clustering Algorithms

Allows objects (crimes) to belong to be placed in two or more clusters simultaneously to varying degrees. The problem of 'types' is avoided

LONDON SOUTH BANK UNIVERSITY

Fuzzy Clustering and cluster centres



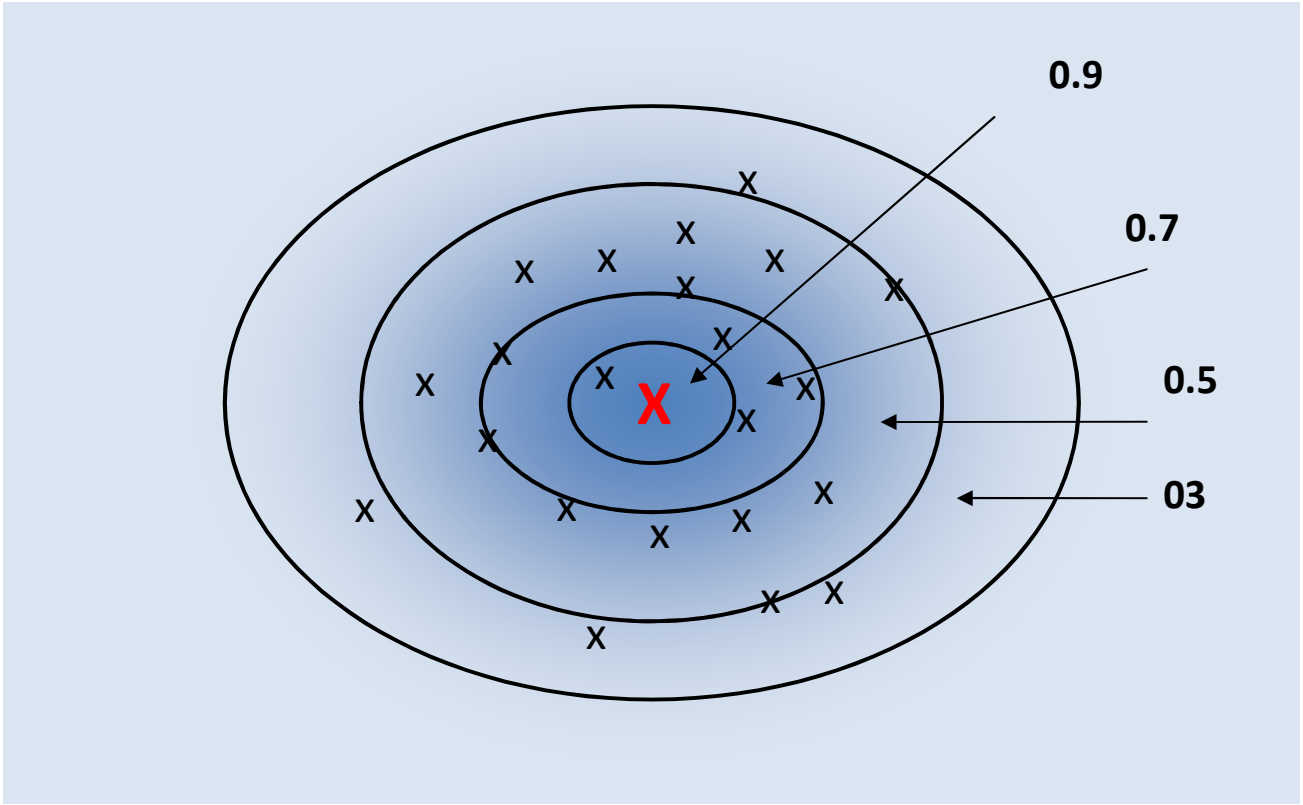
The index crime has a degree of 'membership' of each cluster which is a function of its distance from the cluster centre. Hard boundaries do not exist

Similarity and Search

A ‘similarity function’ is applied to a pairwise comparison between each offence and every other offence in the dataset .

Returns a single value that indicates the degree of similarity between crimes

The effect is to develop a structured search strategy from the crime under consideration to those other crimes that may be linked to it ordered by their similarity



Similarity to the index crime

LONDON SOUTH BANK UNIVERSITY

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	1.00	0.90	0.72	0.33	0.83	0.91	0.80	0.85	0.72	0.90	0.45	0.36	0.31	0.32	1
2	0.90	1.00	0.79	0.37	0.76	0.83	0.88	0.77	0.79	0.96	0.50	0.40	0.35	0.36	2
3	0.72	0.79	1.00	0.49	0.59	0.65	0.86	0.60	1.00	0.77	0.64	0.53	0.46	0.48	3
4	0.33	0.37	0.49	1.00	0.25	0.29	0.43	0.26	0.49	0.37	0.63	0.59	0.58	0.57	4
5	0.83	0.76	0.59	0.25	1.00	0.91	0.67	0.97	0.59	0.75	0.36	0.28	0.23	0.24	5
6	0.91	0.83	0.65	0.29	0.91	1.00	0.73	0.93	0.65	0.83	0.40	0.32	0.27	0.28	6
7	0.80	0.88	0.86	0.43	0.67	0.73	1.00	0.68	0.86	0.89	0.57	0.46	0.41	0.41	7
8	0.85	0.77	0.60	0.26	0.97	0.93	0.68	1.00	0.60	0.77	0.36	0.28	0.24	0.25	8
9	0.72	0.79	1.00	0.49	0.59	0.65	0.86	0.60	1.00	0.77	0.64	0.53	0.46	0.48	9
10	0.90	0.96	0.77	0.37	0.75	0.83	0.89	0.77	0.77	1.00	0.50	0.40	0.35	0.36	10
11	0.45	0.50	0.64	0.63	0.36	0.40	0.57	0.36	0.64	0.50	1.00	0.78	0.72	0.71	11
12	0.36	0.40	0.53	0.59	0.28	0.32	0.46	0.28	0.53	0.40	0.78	1.00	0.90	0.91	12
13	0.31	0.35	0.46	0.58	0.23	0.27	0.41	0.24	0.46	0.35	0.72	0.90	1.00	0.96	13
14	0.32	0.36	0.48	0.57	0.24	0.28	0.41	0.25	0.48	0.36	0.71	0.91	0.96	1.00	14
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	

Similarity between the first 3 series, 3 Dimensions, 3 clusters, degree of fuzziness 2.25

LONDON SOUTH BANK UNIVERSITY

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	0	10	31	64	19	8	24	18	31	9	50	63	65	71	1
2	4	0	25	64	29	15	6	27	25	1	50	63	65	71	2
3	25	15	0	61	41	32	10	40	0	18	47	60	62	66	3
4	66	60	50	0	79	70	57	78	50	61	25	33	24	32	4
5	16	22	31	64	0	12	25	1	31	21	50	63	65	71	5
6	13	22	31	64	12	0	25	11	31	21	50	63	65	71	6
7	19	2	10	63	38	27	0	37	10	4	49	64	65	67	7
8	16	22	31	64	2	11	25	0	31	21	50	63	65	71	8
9	25	15	0	61	41	32	10	40	0	18	47	60	62	66	9
10	6	1	29	64	28	15	8	26	29	0	50	63	65	70	10
11	56	48	29	23	73	63	38	72	29	50	0	10	7	12	11
12	59	49	26	17	79	68	40	77	26	52	9	0	4	1	12
13	65	54	30	17	79	70	46	78	30	55	9	4	0	3	13
14	63	51	26	17	79	69	41	78	26	52	9	1	5	0	14
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	

‘Closeness’ between the first 3 series, 3 Dimensions, 3 clusters, degree of fuzziness 2.25

values for similarity and closeness between crimes 1 and 2

index crime										
crime	1	69	15	54	18	82	83	30	6	10	2
similarity	1.00	0.97	0.96	0.96	0.96	0.95	0.95	0.92	0.91	0.91	0.90
closeness		1	2	3	4	5	6	7	8	9	10

crime	2	10	82	83	1	15	7	18	69	54	55
similarity	1.00	0.96	0.94	0.94	0.90	0.90	0.88	0.88	0.88	0.86	0.85
closeness		1	2	3	4	5	6	7	8	9	10

LONDON SOUTH BANK UNIVERSITY

Results Summary

	n=	lowest median	Dimension	clusters	values of m
Development set 1	83	< 50 %	3D	2	low to medium
Test set 1	29	< 35%	3D	3	all values
			4D2C	4,5	low to medium
			4D2S	4	all except lowest value
Test set 2	51	< 25%	3D	4,5	low to medium
			4D2S		all values
Development set 2	112	< 55%	3D	2	all values
				5	medium to high
Test set 3	134	50%	4D2C	4	low to medium
				5	low
			4D2S	5	medium

Other applications

Enquiries about individual crimes

e.g.

How controlled is this crime?

How forensically aware was this criminal?

How sexually demeaning was the attack?

Crime retrieval

Apart from overall similarity, which crimes have a similar level of control / sex / escape to this one ?

Cluster Centres and Descriptions of crime

4D2S 5 Clusters - cluster centre values

Cluster	Control	Sex 1	Sex 2	Escape
α	0.16	0.10	0.08	0.02
β	0.20	0.15	0.48	0.10
γ	0.70	0.15	0.20	0.81
δ	0.16	0.08	0.12	0.48
ε	0.21	0.87	0.15	0.08

LONDON SOUTH BANK UNIVERSITY

Cluster Centres and Descriptions of crime

4D2S 5 Clusters - cluster centre descriptions

Cluster	Control	Sex 1	Sex 2	Escape
α	Low	Low	Low	V.Low
β	Moderate	Low	Medium	Low
γ	High	Low	Moderate	High
δ	Low	Low	Low	Medium
ε	Moderate	High	Low	Low

LONDON SOUTH BANK UNIVERSITY

Computing with Words (Zadeh, 1999)

Introduces the possibility of using natural descriptions to process enquiries

Crime descriptions

A very violent assault on a middle-aged woman by a young man

In conventional processing this statement cannot be adequately expressed. However 'Very violent' , 'middle-aged' and 'young' can be represented by fuzzy sets and as a result this crime description can be characterised as a single set of co-ordinates

The distance between this point and other crimes in the dataset can be calculated to return a set of offences that best match the initial description

The result is that the descriptive flexibility and depth of meaning conveyed in the initial account is reflected in the search for associated crimes.

Decision Support, Crime Linkage and Analysis

LONDON SOUTH BANK UNIVERSITY

These results relate solely to linkage through behaviour

No account of other intelligence , descriptions, geographic profiling

Essentially acts as filter that could be used with other information / expertise to assist analysis

Makes expensively assembled data work with the analyst interactively so that crimes are represented more sensitively and more sophisticated enquiries are possible

Presents the possibility of developing the computer-based ‘screening’ or ‘decision-support’ system demanded by Grubin / Canter

LONDON SOUTH BANK UNIVERSITY

Research with the Metropolitan Police Service [U.K. London]

Nearly 4,000 stranger rapes over a 9 year period 2005 – 2014 have been released

By far the largest dataset to have been provided by any Police Service.

Prospect of finding truly significant patterns in the data to generate the first effective typology of stranger rape

Incorporate knowledge into a Decision Support System that will perform as an intelligent assistant to crime analysts and investigators

Offer the possibility of discovering an associated typology of stranger rapist and profiling offenders

References

Canter, D. (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology, 5*, 23-46.

Gorry, G., & Scott-Martin, M. (1971). A Framework for Management Information Systems. *Sloan Management Review, 46*(4).

Grubin, D., Kelly, P., & Brunsdon, C. (2000). *Linking Serious Sexual Assault through Behaviour Home Office Research Study 215*. London.

HMIC/HMCPSI. (2012). *Forging the links: Rape investigation and prosecution. A joint Review by HMIC and HMCPSI*: H.M. Inspectorate of Constabularies / H.M. Crown Prosecution Service Inspectorate.

Kandel , A. (1986). *Fuzzy mathematical techniques with applications*. Reading, Mass.: Addison-Wesley.

References

Woodhams, J., Bull, R., & Hollin, C. (2007). Case linkage: Identifying crimes committed by the same offender. In R. N. Koscis (Ed.), *Criminal profiling: International theory, research and practice*. Totowa, N.J: Humana Press Inc.

Zadeh, L. (1965). Fuzzy Sets. *Information and Control*(8), 228-353.

Zadeh L. (1999). From Computing with Numbers to Computing with Words—From Manipulation of Measurements to Manipulation of Perceptions. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications, 45*(1).

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—II. *Information Sciences, 8*(4), 301-357