Matching and Predicting Crimes

Dr. G.C. Oatley[1], Prof. J. Zeleznikow[2] & Dr. B.W. Ewart[3]

1. School of Computing and Technology, University of Sunderland, UK

2. School of Information Systems, Victoria University, Australia

3. Division of Psychology, University of Sunderland, UK

Abstract

Our central aim is the development of decision support systems based on appropriate technology for such purposes as profiling single and series of crimes or offenders, and matching and predicting crimes.

This paper presents research in this area for the high-volume crime of Burglary Dwelling House, with examples taken from the authors' own work a United Kingdom police force.

Discussion and experimentation include exploratory techniques from spatial statistics and forensic psychology. The crime matching techniques used are case-based reasoning, logic programming and ontologies, and naïve Bayes augmented with spatio-temporal features. The crime prediction techniques are survival analysis and Bayesian networks.

1. Introduction

The statutory requirement under the *Crime and Disorder Act (1998)* for United Kingdom police and local partnerships to undertake crime and disorder audits and produce strategies based on these audits, has provided a powerful stimulus to the mapping and analysis of crime data. [1] makes the point that the 'recent shift within British policing towards a more decentralised, proactive style has shifted the analytical focus onto analysts and intelligence officers at the police divisional level who are now expected to be the hub of the local intelligence gathering effort. For high volume crime, this has left an analytical void.'

The authors work is in this area is with the high volume crime of Burglary from Dwelling Houses (BDH) though collaboration with West Midlands Police (WMP). Software [2] was developed to interrogate the database of recorded crimes in order to explore the temporal and spatial characteristics of BDH across the entire operational command unit. The objectives were

to allow police to test their beliefs about BDH trends and patterns against the empirical realities, thereby providing a more substantive empirical foundation for the development and implementation of preventative and detection strategies. Furthermore, the data provided to the project (representing the time period 1997-2001) was used to develop diverse decision support tools. Sophisticated matching and predictive abilities were developed, based on analysis of the data through techniques from statistics and artificial intelligence guided with concepts from forensic psychology and criminology.

Following a presentation of the data, and some background to initial analysis of this data (Section 2), this paper then discusses crime matching by case-based reasoning (CBR), logic programming and ontologies, and naïve Bayes augmented with spatio-temporal features (Section 3), and the crime prediction techniques of survival analysis and Bayesian networks (Section 4).

2. Data and Initial Analysis

The dataset used by the authors represented approximately three and a half years of data, containing the age and sex of the victim and offender (if known). Each crime was represented by a list of stolen property, the features shown in Table 1, and also various Home Office classifications and internal references. Of particular importance were the date stamp, and grid references (Ordnance Survey 12-number grid system), and the Boolean modus operandi (behavioural) features shown in Table 2.

There are a wealth of approaches that can be taken to this kind of data, from Geographical Information Systems (GIS), spatial statistics, statistics, artificial intelligence and forensic psychology and criminology.

An important first step with geocoded data is the 'pins in maps' approach of plotting in a GIS (for good review papers of crime mapping see: [3,4]) - the authors integrated their own mapping functionality in the final decision support system.

1. Alcohol, cigarettes, foodstuffs 2. Antiques, paintings, china, silverware 3. Audio, radio 4. Auto accessories, fuel 5. Books 6. Building material 7. Cards, cheques 8. Cash, till, safe 9. Clothing 10. Computer 11. Cosmetics, toiletries 12. Documents 13. Domestic appliance 14. Drugs, medical 15. Fancy goods 16. Firearms 17. Furniture non antiques including carpets 18. Garden 19. Jewellery 20. Keys 21. Luggage 22. Office, stamps, stationery 23. Optical, personal 24. Other, toys, musical, pushchair, mail, tack saddle etc., animal, metals 25. Pedal cycle 26. Photographic 27. Purse, wallet 28. Sporting 29. T/v, video 30. Telecom 31. Tools, hardware, plant and equipment 32. Vehicle

Table 1 Stolen property features.

GROUPING	VARIABLES
LOCATION OF	1. Wall 2. Adjoining Property 3. Below 4. Front 5. Rear 6. Side 7.

ENTRY	Roof 8. Window 9. Door 10. Above
ENTRY METHODS AND BEHAVIOUR	Smash 2. Cut 3. Cutting equipment 4. Duplicate Key 5. Drill 6.Force 7. Remove Glass 8. Ram 9. Insecure door/window 10. Climbed
TYPE OF DWELLING	1. Old 2. Terrace 3. Maisonette 4. Bungalow 5. Semi-detached 6. Town House 7. Flat
SEARCH BEHAVIOUR	1. Untidy Search 2. Downstairs Only 3. Many Rooms 4. Upstairs Only 5. Tidy Search 6. Search All Rooms
LOCATION OF EXIT	1. Wall, 2. Adjoining Property 3. Below, 4. Front 5. Rear 6. Side 7. Roof 8. Window 9. Door 10. Exit Same as Entry
ALARM/PHONE	1. Cut Phone 2. Tamper with Alarm 3. Alarm Activated
BOGUS OFFICIAL CRIME	Social Services 2. Bogus Official (type unknown) 3. Council 4. DSS 5. Home Help 6. Gardener 7. Other 8. Water 9. Police 10. Distraction Bogus Crime

Table 2 Modus operandi features.

Additions to simple plotting of crime locations were visual inspection of the data using Gaussian (bell-shaped) *kernel-based approximation* for the probability density function (see [5]), and the forensic psychology idea of '*prevalence, incidence and concentration*' [6].

Another very useful algorithm used was the spatial statistics check for *complete-spatial-randomness* (CSR) using *quadrat methods*. As the name suggests, CSR determines whether a set of spatial data is completely random and is the standard model against which a spatial point pattern is compared. The reasons for beginning an analysis with a test for CSR are that the rejection of CSR is a prerequisite for any serious attempt to model an observed pattern. CSR acts as a dividing hypothesis between regular and clustered patterns. Our experiments across offender data resulted in a CSR value ranging between 1.7 to 2.2 for all reasonable set quadrat sizes, indicating a high degree of clustering. In comparison the *k-function* on the same data indicates clustering at a low level but not at higher levels.

The interested reader is directed to the freely available *CrimeStat*[1] [7] software, which provides an introduction to many spatial data algorithms. Also, the Bayesian software *WinBUGS* is able to import spatial data (maps) in three common GIS formats (ArcInfo, Splus and EpiMap).

The data mining algorithms of association rules, CART trees and classification rules (from decision trees) - with or without dimensionality reduction - indicated either no significant relationships or relationships that

[1] *CrimeStat*® is a spatial statistics program for the analysis of crime incident locations, developed by Ned Levine & Associates under grants from the National Institute of Justice (grants 1997-IJ-CX-0040 and 1999-IJ-CX-0044).

are too complex to determine using these methods. However, feature selection for Kohonen neural networks used to cluster offenders, flagged the following variables as discriminating: 'entry-' and 'exit point', 'entry method', 'victim sex', and 'search strategy'. At face value these would be expected to be important. What may appear to be surprising is that there are no significant property stolen types listed. In a similar vein an analysis of Sunderland West Police Force data by the authors [8] showed that there was little evidence for the theory that offenders wait for certain items of property to be replaced by insurance companies and then strike again.

The statistical method of Multidimensional scaling (MDS) was postponed until a more restricted number of features had been determined – however interesting work using this approach can be found in [9].

3. Matching Crimes

The ability to link or match crimes is important to the Police in order to identify potential suspects. Pease [10] asserts that 'location is almost never a sufficient basis for, and seldom a necessary element in, prevention or detection', and that non-spatial variables can, and should be, used to generate patterns of concentration. To date, little has been achieved in the ability of 'soft' forensic evidence (e.g. the burglar's modus operandi) to provide the basis of crime linking and matching, and this section describes recent work [11] which investigates crime matching based upon the combinations of locational data with behavioural data.

3.1 Single crimes with similarity-based retrieval

A retrieve-and-propose CBR system [12] incorporated four different similarity metrics – *k-nearest neighbour* (KNN), *Tversky's contrast model* (TC), and the *cosine* and *modified-cosine rule matching functions* (CR and MCR) – see [13]. The cosine rule includes the benefit of *local* weighting for attributes, where the weight is dependent upon each particular case. Tversky's model extends the nearest neighbour algorithm and cosine rule approaches and includes a contrast between attributes that match and those that do not. [13] proposed a modification to the cosine similarity metric, which not only includes local weighting, but also computes a contrast.

The aim then was to compare the differing treatment of local and global features, and also the importance of a contrast between differing feature representations.

The case base for our experiments consisted of modus operandi, property stolen, time and location features - for each crime where the offender was known (1140 cases). In this way we propose an offender for an unsolved crime by retrieving the most similar previous crime. Four feature selection

algorithms were used - Best-First, Forward Selection, Decision Table and Genetic Search – using the WEKA data mining tool.
The rank position of each offender was determined against their own crime (ideally they should be retrieved with rank position 1), and the average rank across all offenders was the measure by which to compare the similarity performance – this is admittedly a very simple metric, which does not consider the bias introduced by offenders committing varying numbers of crimes.
The results were that CR and MCR produced very similar similarity ratings. However, MCR discriminates better with ranking - average mean and median rank was always higher for all feature sets. Surprisingly KNN and TC produced the same average similarity and rank, and thus perform equally well, and they always (across all feature sets) produced higher average similarity ratings and rankings than CR and MCR. The data consisted solely of globally weighted Boolean features, and each case was measured across the same features in query and case base (no contrast). It is expected that the modified methods (especially MCR) would be more useful when the data representation is more complex. The features selected are under review also, as is the most appropriate way to weight the features.
An important consideration in this domain is that time stamping crime is difficult as the time of occurrence, for instance of a burglary or motor vehicle theft, may not be known exactly. Therefore, it is useful to view crime events as singularities of variable length along the time line. Approaches from the forensic psychology literature include the 'aoristic' [14] approach for two events X and Y, with operators such as [`XbeforeY|XequalsY|XmeetsY|XstartsY|XendsY|XoverlapsY|XduringY`]. This has been seen previously in the artificial intelligence literature as Aamodt's implementation of Allens' [15] temporal system in the long-established Creek semantic network-based case-based reasoning system (see [16]).

3.2 Naïve Bayes classifier and spatio-temporal features

For these experiments modus operandi (MO), temporal and geographic information on the detected burglaries (n=966) attributable to specific offenders (n=306) was used. The ability of three algorithms to match a target crime to the actual offender is evaluated. The first (RCPA) uses only MO information, the second (RPAL) only temporal and geographic data and a third algorithm (COMBIN) is a combination of the two.

The RPAL algorithm represents spatio-temporal information and is the product of *recency*, *prolificness* and *actual location* data on each crime. *Recency* is the distance in time between any previous crime and the considered crime, while *prolificness* represents the amount of previous crimes committed by an offender before the considered crime. *Actual*

Comment [SU1]: Please cite references

location is the Euclidean distance between any previous crime and the considered crime.

Each RPAL equation included parameters that were assigned values drawn from empirical findings within the literatures of forensic psychology and criminology, but were also the subject of optimisation experiments using a genetic algorithm[2]. For instance in Equation 1 for *recency*, there are three parameters that can be optimised. Similar equations for *prolificness* and *actual location* containing parameters that were optimised can be found in [24]. The objective functions used to determine the fitness of a solution considered the number of times an offender occurs in the top 30 for his/her own crime. The value 30 was based upon the value of 29 determined by previous studies [17] and represents a 'reasonable' number of offenders to search through.

$$\text{Recency} = \begin{cases} PARAM\_RECENT\_SCALE \Big/ t_{crime} - t_{last\_offence\_before\_crime}; & t_{crime} \neq t_{last\_offence\_before\_crime} \\ & \textit{Crime occurring today} \\ PARAM\_RECENT\_CLOSEST; & \textit{No last crime} \\ PARAM\_RECENT\_TINY; & \end{cases} \quad [1]$$

Where:

PARAM_RECENT_SCALE=28*;*

PARAM_RECENT_CLOSEST=56;

PARAM_RECENT_TINY=0.0001.

Empirical or *naïve Bayes* approach (see [18]) was used for the classifier using the MO and property stolen features. Naïve Bayesian classifiers assume the effect of an attribute value on a given class is independent of the other attributes. This assumption is made to simplify computations – hence the use of the word naïve.

The general approach involves creating a matrix containing all of the crimes for all offenders, with the presence or absence of the behavioural features represented as Boolean attributes. Interpreting this matrix enabled the matching of new crimes against all the offender data.

Optimisation was also carried out over the 105 modus operandi and 32 property stolen features, using the *GAlib* genome *'GA1DBinaryStringGenome'* with each bit in the string representing the presence or absence of the feature.

[2]The optimization was carried out using the C++ genetic algorithm library *GAlib* [Wall, 1996]

In Table 3, the ranks are presented by the RCPA, RPAL and COMBIN models for each of the three offender groups (SDs are in parenthesis).

The RPAL and COMBIN each achieve a perfect match for 24% of the crimes. For prolific offenders, matching using MO information alone is better than temporal and geographic data, although the best performance is achieved when in combination.

Model	RCPA	RPAL	COMBIN
Group 1 (committed 1-5 crimes)	195.96 (88.42)	406.26 (158.17)	35.07 (24.86)
Group 2 (committed 6-10 crimes)	35.10 (9.28)	181.45 (117.25)	12.62 (8.69)
Group 3 (committed over 10 crimes)	13.73 (7.92)	84.02 (84.41)	7.53 (7.57)

Table 3 Mean retrieval ranks.

Combining the algorithms RPAL and RCPA with their different properties within this forensic domain is an extemporized process. The lessons of this work should be used to guide the formal evaluation of such processes in the future.

3.3 Matching series of crimes with logic

Several commercial systems perform link analysis, for instance *COPLINK* [19] and *FLINTS* - 'Forensic Led Intelligence System' [20]. Recent work by the authors in this area uses a logic programming language (Prolog) to generate flexible matching facilities. The data were converted into Prolog facts, such as `offender('20E1/4441/98',m,14,193)`, where the first attribute is a reference to a crime, then sex, age, and unique offender identifier.

The property stolen and modus operandi data were converted into ontologies to permit generalization of matching, for instance 'cash' and 'cheques' can be generalized to 'money'. The ontologies are imported into the SWI-Prolog environment as RDFS. The benefit of using such an ontology (which is essentially a set of predicates) is that it is easy to view and adapt in an editor such as Protégé, and to engage a domain expert in its construction [21].

In this manner the logic programming paradigm gives an extremely flexible way of querying and examining the data. Consider the way that a *series* of crimes can be determined. To be a member of a series each crime in the list has to be *linked* with another by some '`location_link`', '`property_stolen_link`', '`property_description_link`',

'time_link' and 'gang_link'. An example of the former is where two crimes are within a certain geographical area. The 'gang_link' is slightly more complex, locating crimes where two or more offenders were caught, offenders who obviously know each other. Several networks of offenders can be found, who either know each other directly ('strong' offenders) or through intermediary links with other offenders ('weak' offenders).

Based on these predicates it is simple to determine a series, given the set of all the links - all that remains are that the links are ordered chronologically. This then provides the data miner the possibility of examining the features of crimes in the series, also of finding out all crimes carried out by a gang, and determining an appropriate metric for crime matching. As well as being able to relax the 'friendship' (strong_offender_friends or weak_offender_friends), it is also possible to relax the criteria for 'property_stolen_link' (i.e. move up one step in the ontology), relax the geographical search area, and relax the time period between crimes. Of course all of these 'relaxing' criteria are not equivalent, and it might be pertinent on one occasion to relax the geographic catchment area, however on another occasion relax the property stolen criteria, or gang-membership.

This work is in its early stages, and it is clear that development of this approach for exploratory purposes will need to closely tie in with development with the graphical user interface.

[22] describe the benefits in using such an inductive logic programming approach to identify complex relational patterns that might indicate terrorism threats in large amounts of relational data. They have tested their approach on 'nuclear smuggling' and 'contract killing' data to great success, although optimizing the performance is an issue that is still to be addressed with large datasets.

4. Predicting Crimes

There exists very little literature about predictive models for police decision support (see [23]). Crime prediction with the West Midlands data is motivated by the observation that using officially reported burglaries, most houses are victimised only once and most repeat victims are 'hit' twice only in a 12 month period. If police wait until the second victimisation, it appears from official statistics that it would be too late – the likelihood of a third burglary is small. The first issue described in this section is survival analysis which [24] used to explore if modus operandi distinguishes houses burgled once only from those suffering a revictimisation. This is important to the implementation strategy of both preventative and detection resources.

As official figures show most repeat victims suffer twice only, and so defining high-risk properties by waiting for the second burglary has obvious limitations. The second issue described in this section is a Bayesian belief network, which predicts the likelihood of burglary as a combination of varying kinds of forensic information.

4.1 Survival/failure time analysis

Survival/failure time analysis, developed primarily in the health sciences, deals with the issue of censored observations, which arise whenever the dependent variable of interest represents the time to a terminal event, and the duration of the study is limited in time. For example, it is possible to study the 'survival' of victims of a given crime, given the current point in time. Binary logistic regression with the *'Single victimisations'* and *'First in a series of victimisations',* as the dependent groups were executed with the set of crime scene variables of Tables 1 and 2.

The use of force, searching behaviour, type of property, place of entry, place of exit, alarm activation and use of a bogus official method of entry are discriminating features. Comparing non repeats with 'quick' repeats (i.e. within 365 days), searching behaviour, type of property, entry method and a bogus official strategy are discriminating features. In this way we are able to identify the features of a 'first' burglary which are predictive of a revictimisation. Survival analysis was employed to examine if a burglar's behaviour could be used to identify which properties within this high risk group were likely to be burgled 'sooner rather than later' Examining the proportions of the group that are revictimized within specified time intervals reveals the temporal pattern of repeat burglaries. This is represented at each time period by the *hazard rate* and the *probability density*. The former is of the *rate* of being revictimized at the midpoint of a time interval, while the latter is *probability* of being revictimized by the midpoint. [25] found the risk of a repeat was highest within a month (28 days) of the first crime. To facilitate comparison, the time intervals here are taken in increments of 56 days. Cox regression examines the association of modus operandi variables and the timing of revictimizations. The period (in days) to the second victimisation is the 'survival time' while the crime scene variables are the covariates. Separate analyses are conducted for each group of variables. Nineteen properties were revictimised on the same day (i.e. within hours of the first burglary) and so have a survival time of zero days. These would normally be dropped by SPSS as they have non positive time values. Their inclusion was achieved by giving each a time value of one day.

The highest probability of revictimization occurs within the first time interval. The probability of not surviving (i.e. being revictimized) to the

mid point of the 0-56 day interval is 0.0058. The two subsequent intervals each have a probability density of 0.0026. The lowest probability of repeat burglaries is found in the longest interval beginning 336 days from the first crime.

Interval Start Time (in days)	No. of Properties Entering the Interval	No. of Properties Suffering a Repeat During the Interval	Probability Density	Hazard Rate
0	606	196	.0058	.0069
56	410	87	.0026	.0042
112	323	87	.0026	.0056
168	236	63	.0019	.0055
224	173	66	.0019	.0084
280	107	73	.0022	.0185
336	34	34	.0010	.0357

Table 4 Probability densities and hazard rates of revictimization.

In Table 4, there are successive 56 day intervals for properties comprising the Twelve Month Repeats group. The rate of reconvictimization is greatest within the longest time intervals. The hazard rate (0.0069) for the shortest time interval is the fourth highest across the seven intervals. The interval beginning at 56 days after the first burglary has the lowest hazard rate (0.0042).

Grouping	Variables	Beta Value	Significance	Mean Time (days) to Revictimization
Entry Methods and Behaviour	7. Remove Glass	-0.57	0.03	P = 73, A=146
	8. Ram	-1.55	0.003	P= 26, A=144
Search Behaviour	6. Search All Rooms	-0.25	0.06	P = 131, A=154
Location of Exit	8. Window	0.35	0.05	P = 168, A=133

Table 5 Crime scene variables which survival analyses reveals are significantly and marginally significantly associated with the time to revictimisation.

Table 5 presents the significant crime scene covariates. The mean time to revictimization is presented for those properties where the variable is present (code P) and for properties where the factor is absent (code A). Ramming and removing glass to gain entry are strongly associated with

early revictimisation. A search of all rooms is a marginally significant indicator of relatively early revictimisation. In contrast, exit by a window is significantly associated with a longer period between the first and subsequent burglary.

4.2 Bayesian network

Bayesian belief networks (see [26]) specify joint conditional probability distributions. They allow class conditional independencies to be defined between subsets of variables. Bayesian belief networks provide a graphical model of causal relationships on which learning can be performed. The developed network can be seen in Figure 1.

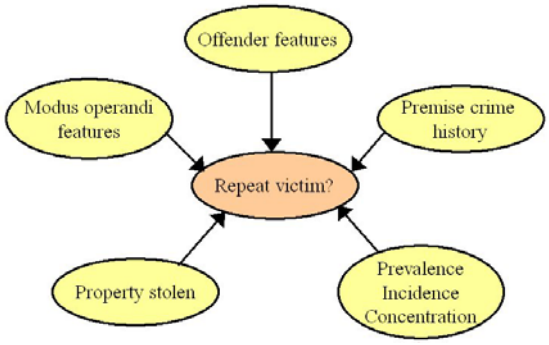

Figure 1 OVER Project Bayesian network.

'Premise crime history' is the number of previous crimes suffered by this premise. 'Prevalence, incidence and concentration' represent the status of the area. 'Property stolen' had been explored by the Police using a frequency count – with 'AUDIO/RADIO' and 'JEWELLERY' being demonstrated to be the most significant variables. The remaining feature is 'Modus operandi features', which was subjected to analysis through binary logistic regression analysis. The specific details of this approach can be found in [2].

The results of the Bayesian network (for instance calibration) are not presented, indeed this approach has never been validated - this is because it contained so many arbitrary decisions and was intended as a prototype 'proof of principle' demonstration, although considerable thought was given to how this network was embedded in the final decision support system for use by police officers [27].

5. Conclusions

Using a wide range of techniques it is possible to discover useful information to assist in crime matching, not only of single crimes, but also of series of crimes. Techniques from artificial intelligence (Kohonen networks, logic programming, CBR and feature selection, genetic algorithms and empirical Bayes) and forensic psychology have proven more or less useful, and also differ in requirements for explicit encoding of domain knowledge. Modified cosine rule matching function will be investigated more thoroughly as the case description becomes more complex.

The logic programming approach however has been crafted explicitly for this problem domain and provide great potential for exploring the data, for instance matching *series* of crimes against previous series of crimes. More thought will be given to modelling as ontologies, and the automated generation of ontologies, structures that can more easily engage domain experts.

The RPAL algorithm and offender features are not complicated, but proved very useful in the experiments. Further thought will be made regarding the inclusion of new features.

A desirable feature of an analysis is that it is robust, and none of the approaches we have considered are predictive in any robust sense except survival analysis, naïve Bayes, and Bayesian belief networks. Future work will continue with these, and also look to develop a full Bayesian model. Because of the additionally desirable features of the Dempster-Shafer model of uncertain reasoning, this will also be explored.

Acknowledgements

The authors acknowledge the support of West Midlands Police for data, and Matthew Wall of MIT for GAlib.

References

1. Hirschfield, A., 2001. Decision support in crime prevention: data analysis, policy evaluation and GIS. *In: Mapping and Analysing Crime Data – Lessons from Research and Practice*, A., Hirschfield & K, Bowers (Eds.). Taylor and Francis, 2001, pp. 237-269.
2. Oatley, G.C., & Ewart, B.W., 2003. Crimes Analysis Software: 'Pins in Maps', Clustering and Bayes Net Prediction. *Expert Systems with Applications* 25 (4) Nov 2003 569-588

3. Soomro, T.R., Naqvi, M.,R. & Zheng, K., 2001. GIS: A Weapon to Combat the Crime. In: Proceedings of SCI 2001/ISAS 2001 (International Conference on Information Systems, Analysis and Synthesis): Part I.

4. COPS 2004. Community Oriented Policing Services home page. http://www.cops.usdoj.gov. Online, accessed 2004

5. Bishop. 1995. Neural Networks for Pattern Recognition, Clarendon Press, Oxford.

6. Pease, K., 1998. Repeat Victimisation: Taking Stock. Police Research Group. Crime Detection and Prevention Series, Paper 90. London, Crown Copyright

7. Levine, N., 2002. CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 2.0). Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. May 2002.

8. Ewart, B.W., Inglis P., Wilbert, M.N. & Hill, I., 1997. An analysis of time intervals of victimisation. *Forensic Update* 50, pp 4-9, 1997

9. Green, E. J., Booth, C. E., & Biderman, M. D. 1976. Cluster analysis of burglary M/O's. *Journal of Police Science and Administration*, 4, 382-388.

10. Pease, K., 2001. What to do about it? Lets turn off our minds and GIS. In: Mapping and Analysing Crime Data – Lessons from Research and Practice, A., Hirschfield & K, Bowers (Eds.). Taylor and Francis, London and New York, 2001, pp. 225-237.

11. Ewart, B.W., Oatley, G.C., & Burn K., 2004. Matching Crimes Using Burglars' Modus Operandi: A Test of Three Models. Forthcoming.

12. Oatley, G.C., 2004. Case-based reasoning (chapter). In: D., Addison & J., MacIntyre (Eds.), Intelligent Computing Techniques; A Review, Springer-Verlag, ISBN: 1-85233-585-8

13. Gupta, K. M., & Montazemi, A. R., 1997. Empirical Evaluation of Retrieval in Case-Based Reasoning Systems Using Modified Cosine Matching Function. Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 27(5), pp. 601-612.

14. Ratcliffe J.H., 2002. Aoristic signatures and spatio-temporal analysis of high volume crime patterns. *J. Quantitative Criminology* Vol. 18, No. 1, March

15. Allen J.F., 1983. Maintaining Knowledge about Temporal Intervals. *Communication of ACM* Vol.26, 123-154, (1983).

16. Jaere, M.D., Aamodt A., & Skalle, P., 2002. Representing temporal knowledge for case-based prediction. In: S., Craw & A., Preece (Eds.) Proceedings of ECCBR 2002, Springer Verlag, LNAI 2416, pp. 174-188

17. Yokota, K., & Watanabe, S., 2002. Computer based retrieval of suspects using similarity of modus operandi. International Journal of Police Science and Management 2002 Vol. 4, Pt. 1, pp. 5-15.

18. Carlin, J.B. and Louis, T.A. 2000. Bayes and Empirical Bayes Methods for Data Analysis (2nd edition), New York: Chapman and Hall.

19. Chen, H., Chung, W., Xu, J.J., Wang, G., Qin, Y., & Chau, M., 2004. Crime data mining: a general framework and some examples. IEEE Computer April 2004, Vol 37, No. 4

20. Leary, R.M. 2003. New Intelligence of the 21st Century: How Smart is it? *Forensic Technology News* November: 6.

21. Noy, N.F., Grosso, W. & Musen, M.A., 2000. Knowledge-Acquisition Interfaces for Domain Experts: An Empirical Evaluation of Protege-2000. Twelfth International Conference on Software Engineering and Knowledge Engineering (SEKE2000), Chicago, IL.

22. Mooney, R.J., Melville, P., Rupert, Tang, R.T, Shavlik, J., Dutra, I., Page, D., & Costa, V.S., 2004. Inductive Logic Programming for Link Discovery. In: H. Kargupta, A., Joshi, K., Sivajumar, & Y., Yesha (Eds.), Data Mining: Next Generation Challenges and Future Directions, AAAI Press, 2004

23. Oatley, G.C., MacIntyre, J., Ewart, B.W., & Mugambi, E., 2002. SMART Software for Decision Makers KDD Experience. *Knowledge Based Systems* 15 (2002) 323-333.

24. Ewart, B.W., & Oatley, G.C., 2003. Applying the concept of revictimization: using burglars' behaviour to predict houses at risk of future victimization. *International Journal of Police Science and Management* 5 (2) 2003

25. Polvi, N., Looman, T., Humphries, C., & Pease, K., 1991. The Time Course of Repeat Burglary Victimization. *British Journal of Criminology* 31 (4) 411-414

26. Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, Inc, 1988.

27. Zeleznikow, J., 2002. Designing decision support systems for crime investigation. In: *Proceedings. of the Fifth International Conference on Forensic Statistics (ICFS5)*, Isola di San Servolo, Venice, Italy, August 30 - September 2, 2002