



## A Review of Literature on: Superfluous Themes on Data Mining for Felon Pattern Discovery

### KEYWORDS

Crime, Data mining, Classification, Regression, Clustering, Supervised and Unsupervised Crime, GIS, hotspot, spatial pattern

**Apexa Lalitbhai Joshi**

Research Scholar, School of Science, R K University, Rajkot & Assistant professor JVIMS-MCA, Jamnagar

**Dr. Suresh M. B**

Professor & Head, Department of ISE, East West Institute of Technology, Viswanedam, Bangalore

**ABSTRACT** *This review presents and discusses an overview of the superfluous themes on data mining for felon pattern discovery for managing crime free society. When low enforcement agencies are not able to recognize various crimes in easy and faster way due to various problem like manual FIR registering, no centralized database, no data sharing among various low enforcement agencies, no mining of data, poor model of police force, No Geo graphical data monitoring, No Crime predication, no Milestones and progress report monitoring that's why felon pattern discovery of crime comes into the picture. Our manual and existing system is working well but if we apply novel approach of mining then will get more accurate and faster result. Here we discuss supervised and unsupervised technique and algorithms for classification, prediction and clustering of crime and criminal activities.*

### INTRODUCTION

Crime in the any country growing more dangerous day by day. The reason behind these crimes is the increase in population, lack of conviction continues to remain abysmal and heavy migration from other states. Due to easy availability of information and technology any destructive activities are easily performed by criminals.

All security agencies CBI, FBI and police departments are putting their effort very well though ratios of criminal activities are increased. Because the deficiencies in the current system.

With the help of modest technology we can reduce the loop holes in the system and achieve the facility like Online FIR, Novel model of police force, Geo graphical data monitoring, Crime predication, Justice also combined with the model ,Advocate & cases details, Milestones and progress report monitoring, Location monitoring. To achieve this goal we need to apply data mining concepts classification, clustering and prediction algorithms.

In this review article I review the Current police departmental working structure and advance themes of data mining and related algorithms.

### THE INDIAN POLICE SYSTEM

By analyzing current Indian police system it forms as per the police act of 1861. With the IPC criminal procedures and evidence act are new modules implanted but outdated the police system in India. The duties of police departments are riot control, Traffic control, uphold routine law & order, VIP protection and guard of state assets.

The tree structure is followed in the decision making is centralized with a few high ranking police officers. (Ncrb official records).

A state is divided into administrative police units. State is divided into several zone, ranges and district. Further district are separated into sub-divisions, circles and police stations. The authority for duties are DGP, IGP, and SSP, ASP, inspector, SI&ASI and constable. (Committees on police reforms, 2005)

No information sharing among various police station and no online track of documents and reporting system is offline and poor. No advance technology used anywhere.

### LITERATURE REVIEW

#### • Supervised and unsupervised Learning

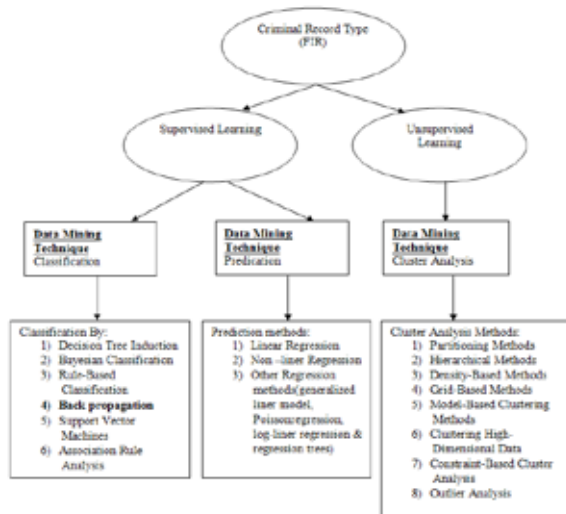
Data and knowledge mining is learning from data recorded at police station in the form of FIR. This learning from criminal data comes into two flavors Supervised learning – labels are given and try to learn how to predict them (e.g. already existing crime performed like robbery, burglary-etc.). Or unsupervised learning – labels are not given and you try to extract information in general out of your data (e.g. new type of crime is performed 1<sup>st</sup> time like years ago there no crime like cybercrime). With the help of this methodology prediction classification and cluster analysis is possible to detect of criminal data (N chambers et al.-2009, K-Funkumizu et al.-2004, DL Bean et al.- 2004, S Nair et al.-2010, J Gao et al.-2006, Y Kou et al.-2004, B-Fritzke et al.-1994, ciro donalek-2011).

#### • Data mining classes for crime

However data mining not only summary of larger datasets but actually it is expanding the scope of analytical processing beyond the fetching data from warehouses by utilizing advanced techniques for data understanding( Han et al.-2001).

Many researchers recognized that data mining is a key research topic in both the domains data mining and machine learning. By knowledge discovery in databases finding interesting datasets, patterns, relations & relevant information and viewing that extracted information with different angles there by it is reliable and rich for knowledge generation and verification. Majority industries like cellular network, all selling product industry are inserted in customer pattern analysis so we can say that opportunity of major revenue generation (Fayyad et al.-1996, Piatetsky-shapiro et al.-1991, Silber schatz et al.-1995).

In the below figure classification and advance themes of data mining techniques and algorithms are listed by categorization.



**Figure 1: Themes of data mining**

To analyzed unsupervised crime self-organizing map is used only SOM is not efficient to analyze. With this method nearest neighbor search is used to improve result. Socio-economic condition and various criminal attributes are studied with the help of scatter counters. The study utilize 50 countries data set is taken for analysis. For refinement, searching, clustering accuracy uses the scatter counter, nearest neighbors & decision tree with SOM for comparison of socio-economic factors between countries with diversified criminal phenomena. (Xingan Li et al.-2013).

Why killing crimes are performed if we want to find reason behind that then study of 40 cases between 2009 to 2010 in turkey done by algorithms. In this study identifying the habits of peoples who perform the criminal activities. Clustering algorithms detected similar cluster and fitting that data into decision tree model we are able to know that majority killing crime perform in night & with guns by relatives and mostly from the same family and young people. (Fatih OZGUL et al.-2012).

Finding criminal network we need to analyze the activities performed in that network to do that k-means K-main Routes algorithms are useful. By finding shortest path we plot the output on geometry object like ellipses to study the criminal network and relation between various crime related events. (Shashi shekhar et al.-2013).

Data mining business application mapping is done by berry and linoff. They listed various data mining techniques for marketing, sales & customer support. Standard statics, Genetic algorithms, link analysis, cluster analysis, Market basket analysis decision tree, neural network are techniques the choice of technique depends on the data available for analysis and outcome required. (Berry & Linoff-1997).

For supervised and already performed crime can be solve by clustering and classification methodology. By using existing evidence where security agencies are not having any clues and witness. Then clustering mechanism used to segment data Naïve Bayesian classification is applied the it is possible to find suspect for crimescene. (Kaumalee Bogawatte et al.-2013).

## CONCLUSIONS

As a result, with the combination of supervised and unsupervised learning methodology crime can be managed in a more effective manner. These techniques must be encouraged as a part of the improved crime investigation technique efforts. This advanced technology is integrated with securities agencies give best result and the rate of crime is going to be reduced.

With help of pattern finding we can work it out with the finding intelligent patterns which is useful for police department to take intelligent decisions based on the results.

To implement an intelligent crime prediction model which is useful for the security agencies. Patrons of crime plan to forecast the time, number of crimes, places and crime types to get precautions. In this report the approach of GIS Spatial disaggregation approach for spatio-temporal crime prediction model. With the advance of crime forecasting, spatial and temporal predictions of crime are used to make long and short term planning. In the situation of getting accurate predictions, it is possible to manage security resources efficiently. Police give attention on assign resources, stressed regions, target patrols, and convey out other police involvements to prevent crime.

Police use strategic, tactical and administrative policies that assist to take precautions before an occurrence of a criminal activity. To make effective policies and improve prevention techniques, police should make use of criminological theories and crime analysis.

## Acknowledgement

Our sincere thanks to all family members and friends. We are thankful to our department for their co-operation. Authors acknowledge the immense help received from the scholars whose articles are cited & included in references of this manuscript. The authors are also grateful to authors / editors / publishers of all those articles / journals & books from where the literature for the article has been reviewed & discussed.

## REFERENCE

- [1] Escalation of crime trouble diminished via crime data mining stimulation in the International Journal of Engineering Research and Application (IJERA- June 2012). IJERA (By African Quality Centre for Journals) (ISSN: 2248-9622) | [2] Compact Structure of Felonious Crime Sets Using FP-Tree Comparable Algorithms Analysis in International Journal of Computer Science and Information Technologies (ISSN 0975 – 9646) (<http://www.ijcsit.com/docs/Volume%205/vol5issue03/ijcsit2014050304.pdf>) | [3] Finding similar case subset and hotspot detection in felonious data set using data mining algorithms: Weighted Clustering and classification in International Journal of Application or Innovation in Engineering & Management (IJAIEM - Volume 3, Issue 4, April 2014) (<http://www.ijaiem.org/volume3issue4/IJAIEM-2014-04-25-054.pdf>) | [4] <http://nccur.lib.nccu.edu.tw/bitstream/140.119/35231/8/35603108.pdf> | [5] <http://www.ijcsit.com/docs/Volume%202/vol2issue1/ijcsit2011020111.pdf> | [6] <http://www.ijcsit.com/docs/Volume%204/Vol4Issue1/ijcsit2013040128.pdf> | [7] Al-Madfai, H., Ivaha, C., Higgs, G., Ware, A., Corcoran, J. (2007). "The Spatial Dissaggregation Approach to Spatio-Temporal Crime Forecasting," International Journal of Innovative Computing, Information and Control, Vol. 3, Number 3. | [8] Www.ncrb.nic.in, E. "Using Geographic Information Systems in Analyzing the Pattern of Crime Incidents and the Relationship Between Land Use and Incidents," M.S report, Middle east Technical University, Turkey, 2005. | [9] Ball, G.H., "Data-analysis in the social sciences: What about the details?," Proceedings of the Fall Joint Computer Conference, 27, 533-559 (1966). | [10] Boba, R. (2005) Crime Analysis and Crime Mapping, Sage, USA. | [11] Bowerman, B. L. & O'Connell, R. T. (1993). Forecasting and Time Series: an Applied Approach, Pacific Grove, CA: Duxbury | [12] Brantingham, P.J and Brantingham, P.L. (Eds.) (1981). Environmental Criminology, Beverly Hills: Sage. | [13] Chainey, S. And Ratcliffe, J. (2005) GIS and Crime Mapping, John Wiley, England. | [14] Cohen, J., Gorr, W.L., and Olligschlaeger, A. (2007), "Leading Indicators and Spatial Interactions: A Crime Forecasting Model for Proactive Police Deployment" Geographical Analysis, Vol.39 (1), pp. 105–127. | [15] Corcoran J., Wilson I.D., and Ware J.A. (2003), "Predicting the Geo-Temporal Variations of Crime and Disorder", International Journal of Forecasting, Vol 19, 2003, pp 623-634. | [16] Deadman, D. (2003). "Forecasting Residential Burglary," International Journal of Forecasting, Special Section on Crime Forecasting, Vol. 19, pp. 567-578. | [17] Everitt, B. (1974) Cluster Analysis, Heinemann, London. | [18] Felson, M., & Poulson, E. (2003). Simple indicators of crime by time of day. International Journal of Forecasting, Vol.19, pp.595-602. | [19] Felson, M. and R.V. Clarke (1998). Opportunity Makes the Thief. Crime Detection and Prevention Series, Paper 98. Police research Group. London: Home Office. | [20] Gorr, W.L. and Harries, R. (2003). "Introduction to Crime Forecasting," | [21] International Journal of Forecasting, Special Section on Crime Forecasting, Vol. 19, pp. 551-555. | [22] Gorr W.L., Olligschlaeger A. (1997). "Spatio-Temporal Forecasting of Crime: Application of Classical and Neural Network Methods." Working Papers of The Heinz School, Carnegie Mellon University. Groff, E.R. and LaVigne, N.G. (2002). Forecasting the future of predictive crime mapping, NY | [23] Grubestic, T.H. (2006). "On the Application of Fuzzy Clustering for Crime Hot Spot Detection," Journal of Quantitative Criminology, Vol. 22, No. 1. | [24] Harries, R. (2003). Modelling and predicting recorded property crime trends in England and Wales—a retrospective. International Journal of Forecasting, 2003, vol. 19, issue 4, pages 557-566. | [25] Hirschfield, A. and Bowers, K. (2001) Mapping and Analysing Crime Data: Lessons from Research and Practice, Taylor and Francis, New York. | [26] Kaufman, L., and Rousseeuw, P. (1990). Finding Groups in Data: An Introduction to Cluster Analysis, John Wiley, New York. | [27] Levine, N. (2002). Crimestat: A Spatial Statistics Program for the Analysis of Crime Incident Locations, Ned Levine and Associates and the National Institute of Justice, Washington, DC. | [28] National Institute of Justice, Office of Justice Programs, U.S. Department of Justice, Thomas F. Rich, "The Use of Computerized Mapping in Crime Control and Prevention Programs.", July 1995. | [29] Openshaw, S. (1998) Building automated geographical analysis and explanation machines. In Longley et al. (eds.) Geocomputation: A Primer. Chichester: Wiley: 95-115 | [30] Olligschlaeger, A. (1997). Artificial neural networks and crime mapping, NY | [31] Pena, D. Tiao, G.C. Tsay, R.S. (2001) A Course in Time Series Analysis, John Wiley, Canada | [32] Phillips, P.C.B. and P.Perron (1988): Testing for a Unit Root in Time Series Regression. Biometrika 75, 335-346 | [33] Ratcliffe J.H. (2004) "the hot spot matrix: a framework for the spatio temporal targeting of crime reduction police practice and research," Vol.5(1) pp.5-23. | [34] Vann, B.I. and Garson, D. (2003) Crime Mapping, Peter Lang Publishing, New York. | [35] Weisburd, D., Maher, L. and Sherman L. W. (1992). Contrasting Crime General and Crime Specific Theory: The Case of Hot-Spots of Crime. Advances in Criminological Theory. Vol.4, 45-70. NJ: Transaction Press. | [36] [http://ap.loksatta.org/documents/indianpolicesystem\\_summary.pdf](http://ap.loksatta.org/documents/indianpolicesystem_summary.pdf)