A REVIEW: CRIME ANALYSIS USING DATA MINING TECHNIQUES AND ALGORITHMS

ChhayaChauhan*[1],Research Scholar
Department of Computer Science and Engineering
Amity University Uttar Pradesh, India.
chhaya1chauhan@gmail.com

SmritiSehgal[2], Assistant Professor
Department of Computer Science and Engineering
Amity University Uttar Pradesh, India.
ssehgal@amity.edu

Abstract--Crime analysis is a methodical approach for identifying and analyzing patterns and trends in crime. With the increasing origin of computerized systems, crime data analysts can help the Law enforcement officers to speed up the process of solving crimes. Using the concept of data mining, we can analyze previously unknown, useful information from an unstructured data. Predictive policing means, using analytical and predictive techniques, to identify criminal and it has been found to be pretty much effective in doing the same. Because of the increased crime rate over the years, we will have to handle a huge amount of crime data stored in warehouses which would be very difficult to be analyzed manually, and also now a day's, criminals are becoming technologically advance, so there is need to use advance technologies in order to keep police ahead of them. In this paper, the main focus is on the review of algorithms and techniques used for identify the criminals.

Keywords--Data Mining, crime analysis, Naïve Bayes Classifiers, Predictive approach

I. INTRODUCTION

A crime rate has become a topic of major concern certainly to limit the development of good governance and increasing day by day. Crimes are neither systematic nor random otherwise crime cannot be analysis. When crimes like robbery, firebombing etc. have been decreased, crimes like murder, sex abuse, gang rape etc. have been increased. We cannot analyze the victims of crime but can analyze the place where crime occurred or happened. It is difficult to analyze the data to detect crime patterns or predict future crimes by intelligence agencies or local lawenforcement agencies. So, there is a need of an effective analyzing tool which can analyze crime data efficiently and quickly to give some useful crime patterns.

Predictive policing means, using analytical and predictive techniques to identify criminal and it has been found to be pretty much effective in doing the same. Because of the increased crime rate over the years, we will have to handle a huge amount of crime data stored in warehouses which would be very difficult to be analyzed manually, and also now a day's, criminals are becoming technologically advance, so there is need to use advance technologies in order to keep police ahead ofthem. The rest of this paper is organized as follows: In Literature Survey section, we provide a brief reviewtowards some related work on crime analysis using data mining techniques and algorithms. In Review Analysis section, includes analyzes five year crime. In Conclusionand Future Worksection concludes the paper and includes future work.

II. LITERATURE SURVEY

Mugdha Sharma et al. proposed advanced ID3 algorithm for presenting importance-attribute significance on the attributes which has less values but higher importance, rather than the attributes with more values and lower importance as well as solve the classification defect to choose attributions with more values. The analysis of the experimental data shows that the advanced ID3 algorithm gets more reasonable and more effective classification rules. In this Z- crime tool was also proposed to analyze the criminal activities through e- mailcommunication.

SushantBharti et al. proposed hidden link algorithm to detect hidden links of the networks of co- offenders which show the possible future crime partner and different network beyond the real network. This paper also analyzes the centrality of node. This analysis describes the importance of node of the network. This is used to discover the strongest person, power of the person and role of the person in the network. This paper gave future approaches i.e. predictive approach in crime analysis which helps in stopping the crime before it occurs and also analyze the network of Co-offenders in India and predict the possible future network of offenders.

ShijuSathyadevan et al. proposed Apriori algorithm to identify the trends and patterns in crime. This algorithm is also used to determine association rules highlighting general trends in the database. This paper has also proposed the naïve Bayes algorithm to create the model by training crime data. After testing, the result showed that Naive Bayes algorithm gave 90%accuracy.

Prashant K. Khobragade et al. proposed Forensic Tool Kit 4.0 which provides remote data investigation and visualization analysis. In remote data, investigation includes to analyze process information, service information, driver information, network device, network

information. This tool generates the file and analyzes the data. This tool is also used to analyze the victim system where the attack is occurring. With the help of crime investigation, the physical and logical memory data areanalyzed.

K. ZakirHussain et al. used data mining techniques for analysis of the criminal behavior. This paper proposed criminal investigation analysis tool (CIA). This tool was used within the law enforcement community to help solve violent crimes. It was based on a review of evidence from the crime scene and from witnesses and victims. The analysis was done from both an investigative and a behavioral perspective. It provided insight into the unknown offender as well as investigative suggestions and strategies for interviews andtrial.

III. REVIEW ANALYSIS

The review is purely based on the crime analysis using data mining. The main aim is to decrease the crime using advance technologies in order to keep police ahead of them. One of the approaches in which crime rates has been calculated from cases registered and persons arrested data during five years. Here, it is quite visible that the cases registered during five year are high. No doubt the persons arrested are low but the aim here is to reduced crimes.

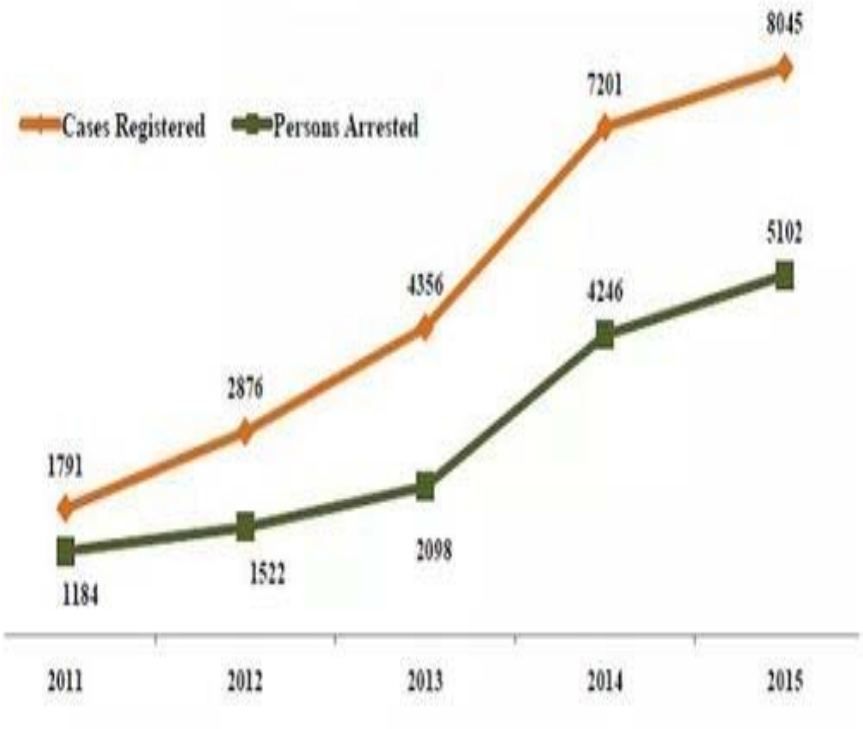

Fig. 1. Cases Registered Versus Persons Arrested

Table I. Various techniques implemented for crime analysis using data mining techniques

P. No.	FOCUS	METHODS/TOOL	ADVANTAGES	FUTURE WORK
[1]	**Focus is to predict the criminal activities and minimize them.**	•Z-CrimeTool •ID3 Algorithm	Highlight the data mining technology to predict criminal activities.	•To differentmethods for feature selection and classification. •For mobile environmentautomatic reply option will be added to the e- mails.
[2]	**Focusis on detailed analysis of the network of co-offenders, collaboration and dissolution of organized crime groups.** **To identify various relevant crime patterns, hidden links, link prediction and statistical analysis of crime data**.	•Hidden link detection algorithm	•PredictiveApproach incrime analysis can also help in stopping the crime before it occurs. •Visualandintuitive criminal and intelligenceinvestigation techniques can be developed for crime patterns ofcity.	The future work will be used predictive approach in crime analysis which helps in stopping the crime before itoccurs.
[3]	**Focus is on find the accuracy of classification and analyze them with different test set**	•NaïveBayes Classifiers •Apriorialgorithm	•Classificationgives more than 90% accuracy by using Bayes theorem	The problem will take as future work.
[4]	**Focus is on data collection from different websites.**	Forensic Tool Kit 4.0 tool	•The tool is used discussing for memory forensic analysis andremote systemforensic.	To used other tools for memory forensic analysis.

[5]	**Focus is to determine the criminalbehavior by using data mining techniques**	Criminal Investigative Analysis (CIA) method was used	• Criminal Behavioranalysis ofviolent crimes is analyzed and modeledby using datamining. • Speedofanalysisas well as the depth increased by data mining techniques	-
[6]	**Focus is to identify the doubtful words from the cyber messages and trace the doubtful criminal.**	Suspicious pattern detection algorithm was used.	Proposed algorithm was dynamic detection, complex to design, support ontology and also support Database & Data Mining support, moderate efficiency than existing algorithm.	The future work will be to Integrate with HADOOP to solve Big Data problems.

IV. CONCLUSION

It is concluded that advanced ID3 algorithm is more reasonable and more effective classification rules during analysis of experimental data. Hidden link algorithm was todetect hidden links of the networks of co-offenders which shows the possible future crime partner and different network beyond the real network. Classification techniques gave more than 90% accuracy using Bayes theorem. The forensic kit tool generates the file and analyzes the data and also used to analyze the victim system where the attack is occurring. Criminal investigation analysis (CIA) tool is used to help solve violent crimes and the accuracy was limited.

V. FUTURE WORK

The future work is to use new tool to analyze and minimize the criminal activities. Criminal investigative Analysis tool don't used crystal ballexperience as well as cannot replace a thorough investigation and the accuracy and detail of a method is limited by the accuracy taken these two problems for future purpose. Forensic toolkit will be used to detect remote node and easily to analyze data for report presentation and to use other tools for memory forensic analysis. The future work will be to use predictive approach in crime analysis which helps in stopping the crime before it occurs and also analyze the network of Co-offenders in India and predict the possible future network of offenders.

ACKNOWLEGMENT

I thank my college Amity University, Noida for giving us an opportunity to be a part of the independent study ofresearch that leads to the development of this work. Many thanks to Ms.SmritiSehgalfor countless discussions and feedback that helped me to complete the worksuccessfully.

REFERENCES

[1] Mugdha Sharma. "Z - CRIME: A Data Mining Tool for the Detection of Suspicious Criminal Activities Based on Decision Tree", International Conference on Data Mining and Intelligent Computing, 5-6 September, 2014.

[2] SushantBharti, Ashutosh Mishra. "Prediction of Future possible offender's network and role of offender's",Fifth International Conference on Advances in Computing and Communications, 2015.

[3] ShijuSathyadevan, Devan M.S and Surya Gangadharan.S."Crime Analysis and Prediction Using Data Mining",First International Conference on Networks & Soft Computing, 2014.

[4] Prashant K. Khobragade and Latesh G. Malik."Data Generation and Analysis for Digital Forensic Application using Data mining", Fourth International Conference on Communication Systems and Network Technologies, 2014.

[5]K.ZakirHussain, M. Durairaj and G. RabialahaniFarzana. "Criminal Behavior Analysis By Using Data Mining Techniques",IEEE-International Conference on Advances in Engineering, Science and Management (ICAESM -2012), March 30-31, 2012.

[6] Mohammed MahmoodAli,KhajaMoizuddin Mohammed and Lakshmi Rajamani."Framework for Surveillance of Instant Messages in Instant messengers and Social neworking sites using Data Mining and Ontology", IEEE- Students' Technology Symposium, 2014.

[7] A. Q. Ansari ,TapasyaPatki , A. B. Patki and V. Kumar. "Integrating Fuzzy Logic and Data Mining: Impact on Cyber Security ", IEEE-Fourth International Conference on Fuzzy Systems and Knowledge Discovery, 2007.

[8] B. Chandra , Manish Gupta and M. P. Gupta. "A Multivariate Time Series Clustering Approach for Crime Trends Prediction", IEEE International Conference on Systems, Man and Cybernetics, 200*8*.

[9] Rajesh Kumar, Nikhil R. Pal, BhabatoshChanda and J. D. Sharma."Detection of Fraudulent Alterations in Ball-Point Pen Strokes using Support Vector Machines", IEEE India Conference(INDICON), 2009.

[10] Veena H Bhat , Prasanth G Rao, Abhilash R V, DeepaShenoy,Venugopal K R and L M Patnaik. "A Novel Data Generation Approach for Digital Forensic Application in Data Mining", IEEE Second International Conference on Machine Learning and Computing,2010.

[11] T.Subbulakshmi, Dr. S. Mercy ShalinieV.GanapathiSubramanian, K.BalaKrishnan,D.AnandKumar and K.Kannathal ."Detection of DDoS Attacks using Enhanced SupportVector Machines with Real Time Generated Dataset", Third International Conference on Advanced Computing (ICoAC), 2011.

[12] MayankPandey and Vadlamani Ravi. "Detecting phishing e-mails using Text and Data mining", IEEE International Conference on Computational Intelligence and Computing Research, 2012.

[13] Vinit Kumar Gunjan, Amit Kumar and ShardaAvdhanam.**"A** Survey of Cyber Crime in India",15th International Conference on Advanced Computing Technologies (ICACT), 2013.

[14] Malathi. A and Dr. S. SanthoshBaboo."An enhanced algorithm to predict a future crime using data mining", International Journal of Computer Applications, 1-6, May 2011.