

## Identifying fraudulent behavior and its types using clustering techniques

Ms.V.Vinodhini<sup>[1]</sup>, M.Hemalatha<sup>[2]</sup>

### ABSTRACT

Classification and Clustering, plays a major role in tracking of multiple humans, this is a demanding problem in composite situations in which extended occlusion, shadow and/or reflection exists. This method includes the stages, analyzing, detecting, validating and tracking. The proposed work is concentrated using Density based k-means clustering algorithm, where after detecting and tracking, the level of crime zone is analyzed, since criminal is clustered over one zone to another they are tracked using filtering method. Density based algorithm works strongly in very demanding Sequences. Information stored in database is analyzed using data mining techniques pattern discovery, evolution analysis.

### General Terms

Human Tracking system, clustering, k-means, classification, pattern discovery, evolution analysis, Density based algorithm.

### I. INTRODUCTION

Crime is the study of crime victims, illumination banned and deviant behavior, the research work solving crime against streets, police stations, courts, educational sector, financial sector, prisons, medical etc pay for itself. The significance of crime or the miscellaneous ways of

measuring criminal activity becomes an impact rate behind criminology. There are various types of study in criminologist that is penology which is a study of prisons and its system, biocriminology which relates to biological basis of criminal behavior, feminist which gives brief study of women and crime, criminality's are the study of detection in crime. Nowadays crime against people become the significant feature, where attempting to cause bodily harm or threat of bodily harm.

Assault is a crime against a person that is attempting to cause physical harm to another person. The aggravated offense is that, the level of crime is enhanced and become cruel (seriously injured and even involves deadly weapons like gun and knife). One example against this type of crime would be the incident which happened in Coimbatore on Dec 25<sup>th</sup>, the girl and her father leads to sadistic dead. Domestic violence is a crime against family or household member; this can be of psychological level of harm. This type of domestic violence also booming due to immaturity level among family members, even nowadays it is committed between siblings, between parent and children's.

Stalking is a type of crime concentrated on unwanted pattern of contact with another person. The most affected persons in this stalking are the youngsters due to their more usage of social websites and more deviation in their regular life.

<sup>[1]</sup>(Research Scholar), Karpagam University and Assistant Professor, Dr.N.GP Arts and Science College, Coimbatore - 48.

<sup>[2]</sup> Department of Computer Science, Karpagam University, Coimbatore.

The work is concentrated on cluster (of crime), that is concentrated on particular geographical area. The densely populated area are taken or marked as hot spots. Here we could have a detailed study on stalking type of crime, where most adolescent have a deviant type of behaviors. Due to this type of activity every so often leads to murdering, which murders their future also. This type of deviant behavior leads to wrong relationship, negative thoughts, and disobedience etc. The age groups mainly ranges from 16 to 18 yrs. The technique used is Critical ratio in order to cram the consequence difference between the groups with reference to their suicidal ideation. The tool used is Suicidal Ideation in order to categorize the number of suicides and reasons for this type of crime. Even more terrorism exhaust by this type of stalking crime. So in order to avoid this type of obstacle behind adolescent counseling method can be used in order to trace, monitor, behavioral record, treatments and reports. Classification play a major role where the type of crime is classified, the reason for occurrence is classified, monitored and counseled. In classification i.e., predictive model the popularity in crime zones are identified using evolution analysis. Ideal crime analysis tool is used to identify the crime patterns.

## II. DATA MINING PROGRESSION

Data mining can be used to model crime detection and crime prevention problems. Crimes are a social nuisance that is against law. The upcoming researches are faster which helps in solving crime activities. In data mining data's are collected, processed and applied for knowledge

discovery using data mining algorithms. The most data mining algorithms are pattern recognition, statistics.

Here Clustering and classification based algorithms are concentrated. The progression of data mining are of collection of datasets, modeling, and processing, this we called as KDD (Knowledge Discovery Database). Mining operations include Classification, Clustering, Sequence Mining, Itemset Mining.

The Data Mining Process includes problem definition, data gathering, model building evolution, and knowledge deployment.

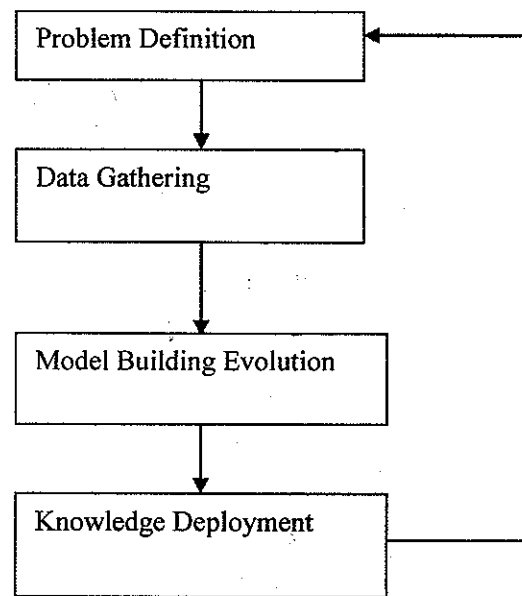


Figure 1: Data Mining Process

### 2.1 Clustering techniques and general idea

Data mining plays a major role in clustering techniques. Clustering techniques have an important role in class

identification of records on a database, so it is established as one of the main topic of research in data mining. Most techniques require previous knowledge of the database. That is domain knowledge to determine the input parameters. In k-means which must be previously known for whichever database it is applied on, mostly there will be lacking of priori domain knowledge, which outputs in incomplete and undesirable results. The clustering techniques used in previous work was k-clustering, where k-clustering is fast in processing clustering result but it cannot work well in huge environment, due to this K-cores and density based techniques have been concentrated in recent work's k-cores provides or it is a simple tool for identifying well-connected structure within large graph, also k-cores are need not to be connected it is themselves well-connected. In Density based one can identify three clusters along with several points of noise. Density based clustering starts by selecting cluster randomly, if it holds minimum points then it is marked as MinPts neighbors else it is marked as noise. Expand Cluster is chosen, the main goal of expand cluster is to find all points that are densely-reachable and it is marked as unclassified.

In order to know that all points lie close to each other threshold is used. Use of clustering algorithms for data mining helps to detect the crime patterns and helps in speed up the process for solving crime. Semi-supervised type of leaning techniques is used from crime records which help in predicting crime accuracy. The Weighting scheme which helps in supervision data from out of the box (Clustering). Usually the crime data are of massive in handling due to enormous quantities of data that must be

processed and turned into useful information, data fusion and data mining is used for handling noisy backgrounds.

Crime approach is concentrated both local law enforcement and national/international law enforcement. Extraction gives an imperative role in crime approach, so entity extraction method is used in order to identify particular pattern from data such as text, images, audio/video materials, unstructured textual documents, criminal networks. This helps in identifying persons address, vehicles, and other personal characteristics from police report. The integrated type of police application is designed with the objective of automation. The automation process results in identifying the interactive interfaces behind crime analyzing tool.

### III. CRIME DATAMINING

Some of the applications in crime investigation include fraud detection-form an event online stream, Interpret insurance claims, Health insurance frauds. Crime data mining plays a vital role in clustering based techniques. Each class represents a set of techniques for use in certain types of crime analyzing; some of the types are Named-entity extraction, criminal network analysis, analysis in cyber crime, clustering techniques used in prediction and association. Social network analysis uses in association, pattern visualization. This type helps in identifying crime tribulations. Crime analyzing tool is used in order to analyze the type of crime and the modules behind the crime analyzing tool is extraction of data, data information and comparison, analyzing module.

#### 3.1 Crime types and Prevention measures

Crime becomes a criminal act which encompasses a wide range of activities from illegal parking to international mass murder. Nowadays cyber crime becomes one of the most committed crimes. Usage of internet and other social websites are booming in the world. It creates problem for people who spend long hours browsing the cyber world especially the younger minds-which is a cyber crimes. One of the main ways to avoid being a victim of cyber crimes is making use of unfathomable security that uses a unified system of software and hardware to authenticate any information. Some of the types of cyber crimes are,

- 1) Hacking
- 2) Theft
- 3) Cyber stalking
- 4) Identity Theft
- 5) Malicious Software
- 6) Child abuse

Cyber crimes are broadly classified into crime against individual (stalking type of crime), crime against property (criminal can steal and rob), crime against government (cyber terrorism).

Traffic violations, it is a local law enforcement. Traffic violations are mostly due to influence of drugs.

Sex Crime, this type of crime is becoming thriving both national and international level. Even we could see as daily information in news paper and all Medias. Sexual crime is a trafficking in women and children for sexual exploitation, including prostitution and pornography.

Violent Crime, murder, aggravated assault, armed robbery, rape, and hate crime.

Property crime committed (2013 to 2014)

- Ø Arson
- Ø Auto Burglary
- Ø Commercial Burglary
- Ø Residential Burglary
- Ø Grand theft
- Ø Petty theft
- Ø Auto theft

Here is the graphical structure which shows number of crimes committed from 2013 to 2014.

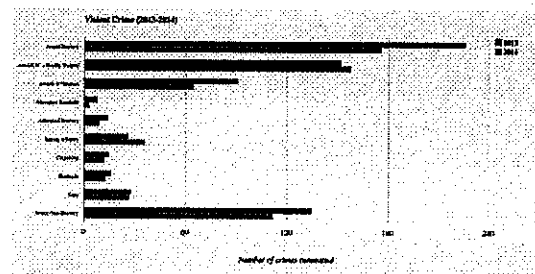


Figure 2: Number of crimes committed

These are some of the crime types concentrated in this research work.

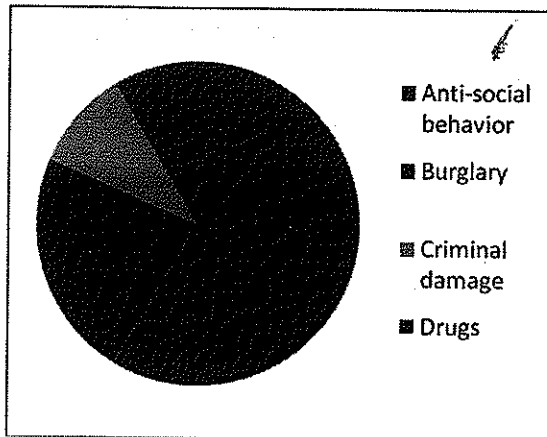


Figure 3: Types of Crime and its Measures

#### IV. CONCLUSION

In this work we have concentrated on data mining techniques used and types of crime measures. We also presented an overview of clustering techniques for a data mining approach which helps in identifying crime patterns, extraction of crime data from huge amount of reports. Here we also concentrated on some of the deviation in adolescent behavior which also helps in reducing crime among youngsters also helps to study the suicidal ideation among students. The future research can also be concentrated in hottest visualization techniques for attractive report and patterns. Data mining techniques and clustering algorithms can be applied in most cases to solve the issues of crime by identifying the activities of criminals and tracking them down. The results of this data mining could potentially be used to lessen and prevent crime for upcoming years. Crime data mining have a promising future for increasing the effectiveness and efficiency of criminal analysis.

#### REFERENCES

- [1] Chen H, Chung W, Xu JJ, Wang G, Qin Y, Chau M. Crime data mining: a general framework and some examples. *Computer* 2004;37(4):50-6.
- [2] Finkel JR, Grenager T, Manning C. Incorporating non-local information into information extraction systems by Gibbs sampling.
- [3] Hope T, Nishimura T, Takeda H. An integrated method for social network extraction. In: Proc. Of the 15th international conference on World Wide Web (WWW); 2006. p. 845-6.
- [4] Jin W, Srihari RK, Ho HH. A text mining model for hypothesis generation. In: Proc. Of the 19th IEEE international conference on tools with artificial intelligence ICTAI; 2007. p. 156-62.
- [5] Jin Y, Matsuo Y, Ishizuka M. Ranking companies on the web using social network mining. In: Ting IH, Wu HJ, editors. *Web mining applications in e-commerce and e-services. Studies in computational intelligence.*
- [6] Jörg Sander, Martin Ester, Hans-Peter Kriegel, Xiaowei Xu (1998). Density-Based Clustering in Spatial Databases
- [7] Martin Ester, Hans-Peter Kriegel, Joerg Sander, Xiaowei Xu (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.

- [8] Freisthler, B.Byrnes, H.F,& Grenewald,P.J. (2010).Alcohol outlet density,parental monitoring, adolescent deviance:A multilevel analysis.
- [9] Lin,C.-F(1991).Deviant behavior conseling and case studies.Taipei,Taiwan:Psychological Publishing Co.Ltd.
- [10] T. Abraham and O.de Vel.Investigating profiling with computer forensic log data and association rules.Proc. Of the IEEE International Conference on Data Mining (2006).
- [11] C.C.Aggarwal,P.S.Yu,outlier detection for high dimensional data,proceedings of (2001),International Conference on Management of data.
- [12] I.Jolliffe,Principal component analysis,H.Kim,S.Lee.A semi-supervised document clustering technique for information organization,proceedings of 9<sup>th</sup> international conference on information and Knowledge Management,New York(2000).



**Dr. M.Hemalatha** completed M.Sc., M.C.A., M. Phil., Ph.D (Ph.D, Mother Terasa women's University, Kodaikanal). She is Professor & Head and guiding Ph.D Scholars in

Department of Computer Science in Karpagam University, Coimbatore. Twelve years of experience in teaching and published more than hundred papers in International Journals and also presented more than eighty papers in various national and international conferences. Area of research is Data Mining, Software Engineering, Bioinformatics, and Neural Network. She is a Reviewer in several teaching and published more than hundred papers in International Journals and also presented more than eighty papers in various national and international conferences. Area of research is Data Mining, Software Engineering, Bioinformatics, and Neural Network. She is a Reviewer in several National and International Journals.

#### Author's Biography



**Ms. V.Vinodhini<sup>II</sup>**, (Research Scholar) ,Karpagam University and working as Assistant Professor in Dr.N.G.P Arts And Science College,Coimbatore-48.She has seven years of experience

in teaching. Her research area includes data mining. She has published papers both in National and International journals.