*Prediction of future crime using Data Analytics techniques*

Aarthi Vishwanathan Bhargava Bodas

Department of Computer Science and Engineering Department of Computer Science and Engineering

PES University PES University

Bangalore, India Bangalore, India

Chandini Velilani Harshit

Department of Computer Science and Engineering Department of Computer Science and Engineering

PES University PES University

Bangalore, India Bangalore, India

***Abstract*—**Concern about worldwide security has increased after the recent blasts and terrorist attacks.In this project , we look at the use of clustering,decision trees and support vector machines to help predict crime patterns and use it to help solve crimes. We apply these techniques to real crime data to identify crime patterns. We use supervised learning techniques for gleaning knowledge from previous crime records and use them to help increase the predictive accuracy.The results of this solution could be used to increase people’s awareness regarding potential locations for crime and for agencies to predict future crimes in a specific location within a certain time.

***Keywords***—Crime-patterns,clustering,data mining, law-enforcement

# Introduction

The inexorable increase in crime, in today’s civilised world has been lamented in the media and public.

Crime has affected mankind for many generations surpassing region , culture and age. It has been looked upon as one of the most menial forms of torment that can ever be abetted or committed.The root cause of crime lies in an individual’s inclinations,biases to certain frames of thought,and societal or cultural influences.

Studies on crimes bring out their various types.Starting with murders or other forms of psychological,physical

Or emotional torture in the yester years to the cyber and youth crimes on digital media of today,it has always been an indelible issue to be addressed mainly due to its

Evolving nature.

Working on solving the issue may involve many types of

Strategies , many methods.The conventional way would be to manually keep record and mundanely sit meetings with the related to unearth more and more information and come up with relations to events by intuition and experience.

But with the advent of information technology and humungous amounts of data and sophisticated analysis tools and procedures it has been possible to mimic human judgement in a faster and more reliable way.

For many years, the research conducted in the area of crime and criminology have been used towards the mitigation of crime and public safety.People have used crime statistics, research, and media stories as a basis to conjecture about the changing nature and scope of crime.Security is considered to be one of the major concerns and the issue is continuing to grow in intensity and complexity. It is given top priority by all political and government organizations worldwide and all of them are working together to reduce the incidence of crime.

Criminology is an area that focuses on analysis of crime and criminal behavior and aims to identify crime characteristics. It is one of the most important fields where the application of analytics can produce important results.The large volume of previous data and also the relationships between these kinds of data have made criminology an appropriate field for applying data analytics.Crime characteristics have to be identified first. Knowledge gained from data mining can be very helpful for government and police departments..In the present scenario, crimes are increasingly influenced by the role of technology .Therefore, police needs such a crime analysis tool to catch criminals in order to win the race between criminals and law.The police should make use of the wide variety of crime analysis technology that can give them an edge over criminals.There has to be sufficient records available pertaining to crime to help the police achieve their goal.Police organizations everywhere have to deal with this barrage of information.

An ideal crime analysis tool should be able to identify crime patterns quickly and in an efficient manner for future crime pattern detection and action.However, we do face certain challenges :

* Wide variety of crime data needs to be stored and processed efficiently
* The data available is inconsistent and incomplete.
* Due to increase in complexity, investigation and analysis of crime takes time.

All the above challenges motivated us to focus on providing solutions that can enhance the process of crime analysis for reducing crime rates.The main focus is to develop a crime analysis tool that assists the police in :

* Analyzing and identifying crime patterns
* Using these to help predict future crime patterns

# PREVIOUS WORK

# Reference 1:

The paper ‘An enhanced algorithm to predict a future crime using data mining’ authored by Malathi. A (Assistant Professor Post Graduate and Research in the Dept. of Comp Sc. , Government Arts College,Coimbatore, India). and Dr. S. Santhosh Baboo (Reader,Post Graduate and Research,Dept. of Comp Sc. ,DG Vaishnav College Chennai, India) proposes the use of a mix of of data mining techniques for developing such a crime analysis tool. For this purpose, the following specific approach were formulated:

* To create a cleaning algorithm that cleans the given dataset by deleting unwanted data and fill missing values with various techniques.
* Use of clustering algorithms

to search for crime patterns in historical data.

* Use of classification algorithms

in order to predict future crime

behaviour based on previous crime trends.

The C4.5 decision tree algorithm was used to predict the crime trends for the subsequent year. Experimentally, results have proven that the technique used for prediction is

speedy with high accuracy.

**Result1**: Crime is either steady or dropping. The

rate of sexual harassment is the main issue ,

along with slight incidences of murder , dowry

death , dacoity , homicide.

**Result2**: Crime is either rising or is unstable due to

Rioting, cheating, counterfeiting, cruelty by family

members.

**Result3**: Crime is generally increasing. Thefts are the

primary crimes on the rise with some increase in

fire-raising. Lesser cases of the property crimes:

burglary and theft are observed.

**Result4**: Few crimes are unstable, the main ones being murder,rape and arson which are also unstable. Property crimes such as burglary and theft have less change. Issue is to demonstrate at least some characteristics of these clusters.

Experimental results prove that the tool is effective in terms of analysis speed, identifying common crime patterns and future prediction.No limitations were reported as such for the process.

Overall we perceived this as a focused approach however noticed lacuna in the predicting ability of the system to possible future crimes.

# Reference 2:

A framework for crime trends was introduced by comparing all individuals using a new distance measure and then clustering them as needed.This procedure provided identification of various criminal types and provides

a visual clustering.

Reference 3:

Work by Dr.Wayne Os Good shows how a regression model

can be built based on poisson distribution as a tool

For addressing problems in analysis of aggregate crime rates.When the aggregate unit population size is less

compared to offense rate , crime rates must be calculated

from a small number of offenses,however this is ill suited to least square analysis.

Reference 4:

Work by Abraham et al.(2006) focused on employing

Computer log files as dataset in order to find relationships

using the frequency of an incident occurring.Once a

set of results were obtained they were analyzed to produce

profiles used to chart out the criminal behaviour.

Through all of the references studied ,the lacuna we observe

is that any tool changes brought about by crime department

is superseded by even advanced technologies used by criminals ,so there has to be a dynamic mechanism to

track this advancement.There has to be even more sophisticated tools and methods which are continuously subjected to verifications in the field.

# PROBLEM STATEMENT

Build a predictor system for predicting crime patterns for

specified periods using clustering algorithm,decision trees and support vector machines in order to address the delays in manual investigations and deployments of safety measures to curb crimes in any societal layout.

# Approach

We begin by collecting data from the fbi website belonging

To the Criminal Justice Information Services Division from the California dataset containing various offences to Law

Enforcement by City in 2013.

If any NAs are observed we try to use linear interpolation based on the nearest attributes.To get a general idea about the distribution of data , we plot a visual graph such as a histogram or a bar chart. This is followed by computing basic statistical measures like the mean or median to understand where the data is centres

We also try and examine outliers and discard them after checking whether the reason for their existence is

pertinent to the problem being solved. Once this is done,

we bifurcate the data into training and testing sets and

use algorithms like the clustering algorithm,decision trees , and support vector machines on the training set.

Next we validate the model using the test data and

perform this a considerable number of times in order to reach a satisfying value of the outcome.

We also try and compare our results with the previously done work in order to benchmark our analysis.

We have observed that the most recent approach of crime analysis and prediction involves only forecasting crime trend for the next year or so. The tool is effective in terms of analysis speed, identifying common crime patterns and future predictions .

Overall the process is perceived as a focussed approach , however we noticed a lacuna in the approach - the predictor system did not have the ability to predict the next type of crime which could occur in the near future.

In this project , we try to use analytical skills and algorithms to overcome this anomaly and come up with a system that could possibly predict the next type of crime.

CITATIONS

[1]Paper on “ Prep Search: An integrated crime detection system", Intelligence and Security Informatics 2009. ISI '09” submitted for the IEEE International Conference on

Pp. 161-163,2009 by Li Ding,Brandon Dixon and Dana Steil.

[2] Paper on "An Analysis of Data Mining Applications in Crime Domain", at the *Computer and Information Technology Workshops 2008. CIT Workshops 2008 for the*

*IEEE 8th International Conference on*, pp. 122-126, 2008

by P. Thongtae and S. Srisuk,

[3] Paper on “Detecting and investigating crime by means of data mining: a general crime matching framework”,at the Procedia Computer Science, World Conference on Information Technology, Elsvier B.V., Vol. 3, Pp. 872 by

Keyvanpour, M.R., Javideh, M. and Ebrahimi, M.R.

(2010)

[4] Arunima S. Kumar and Raju K. Gopal, "Data mining based crime investigation systems: Taxonomy and relevance", *Communication Technologies (GCCT) 2015 Global Conference on*, pp. 850-853, 2015.