Chapter 18
A Classification Framework for Data Mining Applications in Criminal Science and Investigations

Mahima Goyal
Ambedkar Institute of Advanced Communication Technologies and Research, India

Vishal Bhatnagar
Ambedkar Institute of Advanced Communication Technologies and Research, India

Arushi Jain
Ambedkar Institute of Advanced Communication Technologies and Research, India

ABSTRACT

The importance of data analysis across different domains is growing day by day. This is evident in the fact that crucial information is retrieved through data analysis, using different available tools. The usage of data mining as a tool to uncover the nuggets of critical and crucial information is evident in modern day scenarios. This chapter presents a discussion on the usage of data mining tools and techniques in the area of criminal science and investigations. The application of data mining techniques in criminal science help in understanding the criminal psychology and consequently provides insight into effective measures to curb crime. This chapter provides a state-of-the-art report on the research conducted in this domain of interest by using a classification scheme and providing a road map on the usage of various data mining tools and techniques. Furthermore, the challenges and opportunities in the application of data mining techniques in criminal investigation is explored and detailed in this chapter.

DOI: 10.4018/978-1-7998-2466-4.ch018

INTRODUCTION

In recent years, there has been an alarming increase in the volume of crimes in different applications like finance, credit card fraud, theft, violent crime, cybercrime, intrusion detection, and online fraud. It is a challenge for researchers to develop a crime analysis tool that can identify crime patterns accurately and efficiently. Thus crime science and investigation plays an important role in accounting such challenges of crime control and maintenance of public order.

Data mining is a field of extracting knowledge from huge amounts of data stored in Data warehouse and repositories. It is often used to uncover the latent truths behind huge volumes of data. It extracts interesting patterns which helps in detecting crimes and decision making. Thus, various advancements in crime data applications adopt data mining techniques to follow the task of crime investigation.

This chapter has three targets to achieve. The first target presents a framework which classifies the applications of data mining to Criminal Science and Investigations (CSI). The second is to provide a rigorous and extensive review of existing work on the applications of data mining to CSI. The third is to develop a road map for the researchers in this field. The major applications of CSI include financial frauds like credit card and money laundering, violent crime like sex and drug offense, theft and intrusion detection. Crime detection uses data mining techniques like classification, clustering, prediction, association, neural networks and genetic algorithms. The literature review shows that data mining techniques have been applied in detecting different types of crime, although, financial fraud and cyber crime are favored among the researchers.

The three subtasks of CSI involves extracting named entities from narrative reports, detecting deceptive criminal identities and identifying criminal group and key members of the group (Kevyanpour et al., 2011). The important algorithms used in 'Clustering' are k-means and AK-mode algorithm, in' Association', Apriori and Frequent Pattern (FP) growth algorithm, in 'Classification', Naïve Bayes Algorithm and Support Vector Machines (SVM). Other data mining techniques like neural networks and genetic algorithm have also made their way in appealing to the criminal investigators in the recent times.

The chapter is divided into various subtopics. The sections include an introduction, framework for data mining applications in CSI, data mining based layered framework, opportunities of data mining techniques associated with the criminal science and investigation and challenges of data mining associated with the criminal science and investigation. The framework for research includes the research methodologies used in writing the review. The framework for data mining applications includes classification of data mining techniques and classification of CSI. The analysis for CSI includes the steps performed while analyzing the classification.

RESEARCH DESIGN

Data mining techniques in Criminal Science Investigations is the field of study which has enormously gained huge popularity in a few years. It is a diverse discipline on which extensive research has been done. This field attracts a lot of researchers because of its wide applications in different domains. This chapter is dispersed across diverse journals and conferences. This diverse search comprises 16 research journals, 25 conference proceedings, two book chapters and lecture notes. The articles were reviewed with publication limit starting from 2008 to 2015. The various journal sites, which were chosen for the catalogue of the literature in this area are: