Crime Analysis and Management System for Sri-Lanka Police

S.Sutharsan[#1] and Dushyanthi Vidanagama[#2]

 # Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka

[#1] sskavithan@gmail.com

Abstract— The prediction of the grave crimes shows that the security agencies and police departments must bear a high burden of the criminals in larger numbers than the past, because of the increasing number of crimes and criminals in Sri-Lanka. The continuous trend of crime management was very tedious and manual crime recording and investigating system has resulted inefficient crime analysis. Therefore, the analysis of crime data has become a huge burden for the crime officers and investigators who handle the records and investigates. This paper presents an efficient method to analyse and manage the crime records which are collected from the web-users, historical records, websites and news sites. Using the concept of data mining we can extract previously unknown, useful information from an unstructured data. Here we have a clear approach among computer science and criminal justice to develop a procedure of data mining that can help to solve crimes faster.   Crime Analysis and Management System (CAMS) will provide better prospective for the enhancement of security agencies regarding the quality and transparency. Finally, this paper to show the interest of not only to those working on related systems but also to those with a general interest in how advanced software technologies can be applied to national security agencies for their challenging application areas.

*Keywords— **Grave crimes, Generate Reports, Crime Analysis and Management System (CAMS), Datamining***

I.INTRODUCTION

From recent years there is an increase of computer usage within the Police Forces of Sri-Lanka, not just in general administration (e.g. accounts management, human resource management, e-mail, general office suites etc.), also as an important tool to support the primary business activity of policing (preventing and solving) crime. Increasing crime rates and (greater awareness of modern criminals) puts a strain on existing methods and techniques. No security agencies are immune to crimes.   In 2018 alone, there were 36355 reported crime incidents, and 3287 of those were victimless crimes and some are through web application flaws.(Fig.1) There is a clear need for the application of advanced software technologies to make the best use of limited resources in Sri-Lanka.

Presently in many fields' usage of the desktop applications such as Microsoft Excel were adopted,

thereby causing the approach to become both manual and slightly computerized system in Sri-Lanka. Later, this method of record keeping results in inconsistencies, improper use of disk space and poor management and coordination of data. The occurrence of this method resulted from reasons such as Insufficient funding, Data
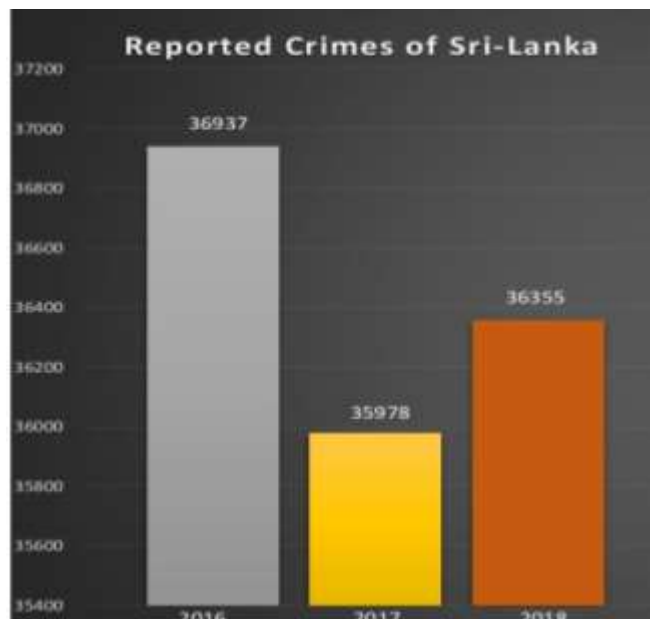
duplication and extortion, Absence of necessary skills and resources for the statistical and information management.

Figure.1: Reported total crimes in Sri-Lanka

Considerably increasing crime rate cannot be predicted neither systematic nor random. According to Sri-Lanka Police Crime statistics; crimes like theft, extorsions, arson etc have been decreased while manufacture of dangerous drugs, sex abuse, robbery etc have been increased. Even though the prediction of criminals cannot be done but the prediction of trends and patterns can help to identify the probability of the crimes and it's occurrence.

The analysed results of CAM system cannot be assured of 100% accuracy, but our result helps in reducing the crime rate to a certain extent by considering the trends, patterns of the crime and providing security in crime sensitive areas.

The data used by the Crime analysing system is stored in a centralized database which holds information about criminals, crimes and official and unofficial users of the system. The database is the core unit for all actions in the system and can be easily updated and used to aid in all of the system's processes, and all of the necessary information is stored in one central location and thus is

easily accessible. Furthermore, the proper handling of the system's centralized database will allow functions such as crime report generation and statistical analysis of crime data. Main objective of this paper is to provide a complete view for a software system which can help to analyse the crime and change the manual working patterns which are practiced by Police forces in Sri-Lanka. Identifying a crime patterns using the classified data records. Clustering techniques will work better to detect the newer and unknown crime patterns in future. The advantages of the proposed system are as follows effective data handling, reduced time consumption, reduction of the cost., efficient operations provide support for gathering and recording case data, prediction of crime patterns, visualized view of analysed crime data.

II. LITERATURE REVIEW

In this section let's consider about the Sri-Lankan daily practice of complaints recording, crime recording, and the operational crime analysts aims to provide a detailed task description and task analysis based on expert interviews and the literature. To elaborate upon the problems faced by crime officers, analysts and their demands touched upon in the introductory section and introduces requirements for a system that supports crime management and analysis during the complaints recording, report generation, prediction of accused and that aims to reduce errors on the duties of national security agencies. (Anon., May 2015)

The objective of crime analysis to find meaningful information in the vast amounts of available data and to disseminate this information to investigators in the field to assist in the efforts to solve crimes, apprehend criminals, and to suppress criminal activity.

A. Computerised crime systems for dealing with crimes

Most crimes around Sri-Lanka have a serious nature to the whole society according to the type of crimes that occur. Considerably, the individual crimes are of a less serious nature to the whole society (although still serious to the victims). The crop volume of those crimes creates a different type of problem for the investigator. The entering of complaints and managing them according to the chronological for their daily routine activities face a major fail as the investigators and crime officers mostly maintain a manual system for the crime management and analysis. The lack of resources for detailed investigation of any one particular crime creates problem in both solving a crime and preventing similar future crimes. (Calhoun, 2008) The challenge for advanced software is to apply modern crime management procedures, report generation and crime analysis techniques to alleviate the problem.

B. Complaints management system

This system was aimed at the implementation of managing the complaint records in a system. It is also a database system in which the police department keep the record of criminals who have been arrested, to be arrested, or escaped. This system will help the police department in improved quality of managing the information. The main existence in the whole process include; the petitioner (the person who files a First Incident Report (FIR)), victim, accused or criminal, case, and investigating officer.

The CMS keeps petitioner records, victim, accused, FIR, case and investigation officer entities. The system's strength state that it allows for storage of multiple data for a criminal. A weakness observed in the system however, is that it lacks in covering all the necessary entities required for a CMS built for the national security agencies, it has no graphical user interface; as only the database was designed, there is no proper dissimilarity between an accused and the criminal in the system, there is no generation of crime analysis and report.

C. Existing Crime Investigation and Analysis Techniques

The invited paper on towards crime investigation clearly state about the existing crime investigation and analysis technique. Collect crimes, find links between crimes, and develop profiles of Criminals and criminal network detection are some of the most common areas of data mining applied in crime analysis. Association analysis, classification and prediction analysis, and external analysis are some of the traditional data mining techniques that can be used to define patterns in Structured data. New data collection techniques help to identify patterns in both structured and unstructured data. The concept of space concept was used in COPLINK project (COPLINK is a comprehensive crime investigation and analysis system which has incorporated a collection of data mining tools to support the investigation process) for the extraction of criminal relations from accident summaries networks have been established likely suspects by measuring the combined weight of two criminals. (Schroeder, 2002)

The single-link hierarchical assembly was used to split the grid into subgroups and model the ban. The approach was used to determine patterns of interaction between subgroups central measures such as degree, it has been used between proximity and proximity to reveal key roles between each group, such as commanders and gatekeepers. Social Network Analysis (SNA) is a technique used to identify closely relevant population groupings. Therefore, it can be used for exploring criminal organizations in a huge number of criminal records in databases. (Anon., May 2015)

Sargo Union is a mining company there are two official mining applications control. It can be used to find bomb attacks crimes committed by one convict only criminals can involve criminals. Therefore, this method is commonly

used sexual genital mutilation and sequential killing. This is the task of finding and discovering related methods / link analysis among people, weapons and organizations. The purpose of this system is crime laws seem invisible. The surface is truly interconnected. You can charge crime connections to criminals. If there are two different places belong to a specific way, they can easily attribute this wrongdoer. Classification applications: fraud detection, computer and network infiltration detection, bank failure computation, image classification. Classification techniques: Bayesian models, end trees, artificial vein networks, sub-vector engines. Compare string the techniques used to find solidarity between posts are used. This determines the database registration pairs unity among them.

D. Solutions

The proposed crime analysis and management system addresses the key issues at Sri Lanka Police by providing a clear mechanism to help administrative decisions and information of the specific crime situation. In general, the system helps crime officers to identify same techniques used by criminals to do their crimes or illegal actions. Clustering technique have been used to crime analysis. The reason for choosing this method is that we have the records of the known crimes and records around the web. Crime nature changes over time, so using the clustering technique we can be able to detect the newer and unknown patterns in future.

III. METHODOLOGY

According to the above literature, there is a vast crime data available about present and past. In manual approach of solving crimes taking into account of all related information is not practical due to the capacity of the human calculation and analytical abilities. Hence to identify the criminals who are using multiple identities, possible resources accessible to criminals and possible potential threats encountered may not be done as the demanded. Moreover, changing nature in the behaviour of criminals required dynamic models to the data analysis process.

In the security field, reducing the risk is a key goal of the industry. Risk is defined as the possibility of losing of assets, damage or destruction caused as a result of a threat utilize a specific vulnerability or alternatively as the possibility of loss, which is resulted from a threat, security incident, or event (ASIS International Guidelines commission, 2003). As illustrated in the figure 2, risk exists at the intersection of threats, assets and vulnerability. To reduce the crime and risks in the present world automated systems will be smart enough and apart from manual analysis and management process made even better with the system.

The crime analysis and management system (CAMS) will be a recording and analysing application that will be used for automating the manual process of registering and retrieving different case in police department. This system

Figure 2. Intersection factors of Crime & Risk
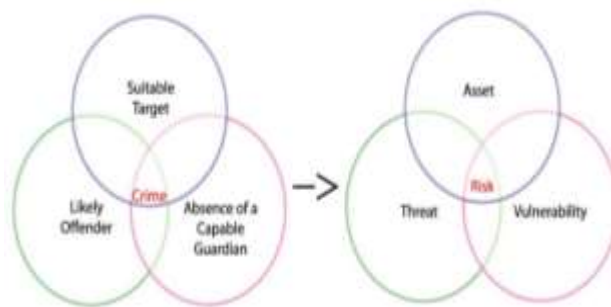
will retrieve the details through web sites and news sites and achieve all the complaints which register themselves in police department. It provides an interface to store, the crime details of different types of crimes and allows the administrator to search and view the details of records, criminals' register management, first incident report management, case history details management, analysed case detail management, visualization of analysed crime, diagrams for prediction and reports generation are also the functional features of the system.

A. System Architecture

A web-based system which can be run in the servers located at the police headquarters of Sri Lanka located at Colombo so that all the police stations which are located all around the country can access the system through the VPN (virtually private network) which can be facilitated by Sri-Lanka's network service providers. The data for the system is collected from different websites like news sites, blogs, social media etc. Since the data collected is an unstructured data, we use Mongo DB. Crime data is very complex with the number of fields, content and size of the document can differ from one document to another, so the best option is to have a schema less database. Also, the absence of joins reduces the complexity. As well the SQL database usage will allow for insertion of data without a predefined schema.

Then for the classification of data we use Naïve Bayes algorithm which is a supervised learning method as well as a statistical method for classification. Naïve Bayes is a probabilistic classifier which gives out a probability distribution of set of all classes rather than providing a single output when we give an input. And complex articles a classified into a crime type by the algorithm.

The algorithm is simple and converges quicker than logistic regression. Also, it takes less space and perform well in compared to other algorithms. For each type of the crime we must train the algorithm. By training means, we have to teach them on particular inputs so that the testing can be implemented for unknown inputs. Tested results show more than the 90% accuracy since it categorises each word as token and remove the repetitive words like "the", "and", "of" etc which develops the accuracy.

With that we integrate Named Entity Recognition (NER) in the articles of crime. NER is also an Entity Extraction finds and classify elements in text into pre-defined categories such as names of person, locations, organizations, date, time etc. So that we can extract the important details in crime articles. Another concept we have included is Coreference Resolution to find the referenced entities in a text. When two or more expressions in a text refer to the same person or thing coreference occurs in linguistics.

Next for the pattern identification, we have to identify the patterns and trends in crime. For this we are using the Apriori algorithm, which is used to find the crime pattern that occurs frequently. Apriori also used to determine association rules which highlight general trends database. General crime pattern can be identified for area sensitivity, notable event, presence of criminal groups etc. Pattern recognition helps the police officials to direct the resources in an efficient and effective manner. This will help police forces to provide security, patrolling in crime prone areas, fixing burglar alarms, CCTV etc.

B. Crime Analysis for Prediction

Decision Tree concept is used here for the prediction of crimes. Concept of decision tree is very similar to graph where internal node represents the result of a test on an attribute, and each branch of the tree represents the outcome of the test. Decisions tree is simple to understand and interpret. For large data sets this concept works well and it helps to make better decisions about variables.

C. Information access and Visualization

The graphical user interfaces and the visualization tools are self-explanatory so, the target user group can easily understand and the users who does not have high computer literacy to cope with complex system scenarios. Police department have a need to store many sets of information. Therefore, the system provides a heat map which indicates level of high and low activity using the darker and lighter colours respectively. Heat maps are very useful than the other representational mechanisms. Advantages of heat maps are out of range data is automatically discarded, gradient colour range, analyse only the data we want and numeric and category-based colour images. By knowing about the relevant regions, we can be able to proceed the preventive mechanisms like night patrolling, fixing burglar alarms, fixing CCTV camera etc.

D. Technology

CAMS is planned to encrypt the sensitive data using the trustable security framework. (E.g. Apache Shiro) Information such as accused name, age, address is very sensitive information which should not be exposed outside. The passwords encryption considered using the hash algorithm using the Spring security. All the entities of all the entity sets have been incorporated with fields to store IP address of the machine used in inserting the record , record inserted user's ID, record inserted time, record inserted date, record updated user's ID, record updated time, record updated date, IP address of the machine used in updating the record, authority value to a particular tuple so that it will be validate before the record being updated by a particular user. As the information consist of high value in the means of law enforcement, when a user deletes a record, the status of that record will be changed to 'deleted' with a help of a flag field. The record will not be displayed again. Logging in by several users with same credentials from anywhere is blocked in the system. Inactive sessions are considered to get invalidated in a fixed time.

E. Deployment and Maintenance of the System

Once the functional and non-functional testing is done, the final system is deployed in the customer environment or released into the Sri-Lankan police department. The proposed system is installed in the new police environment and the transition phase from the old environment is monitored. This kind of stage involves in the training of police officers that will be given the privilege of operating system, populating the database with the already existing data and converting such data.

At the maintenance phase, issues arising on deployment the client-side problems are fixed. To fix these kinds of issues, patches are released. Also, to broaden the development of the product some better versions are release. Maintenance is done to deliver these changes in the client-side environment.

IV. EVALUATION

User interfaces have been categorized into three sections, namely Add/Edit/View Tab, Search Tab and Analysis/Reports Tab. Users are not allowed to do create, read, update and delete operations in one window. The system literally inherits the flow of events. This makes easy for the police personals to work in an easy and efficient manner. Closely related systems were helped in elaborating system requirements. The proposed system was measured in terms of authentication, real-time access and centralized storage, and it proved efficient performance and also exists inter-station communication and generation of crime and criminal reports

The data retrieval can be performed very fast than manual and this has made the processes of criminal profiling, link analysing and predicting in corresponding to a very large data set of crime information. Further, web-based implementation provides the efficient mobility and accessibility of the data to law enforcement personnel at required time. However, the exact accuracy of the result

depends on the algorithms used and the implementation of them. Finding the most suitable model can be still situational due to the judicious, dynamic criminal behaviour and vigorous nature in the criminal organizations.

In Sri-Lanka as the future scope researchers wishes to generate crime maps using historical data. Also, various kinds of analysing charts generation with some analysis and update information about historical data like selecting the most wanted. If opportunities are offered, it is also very easy to the officials for automating all the divisions.

V. Discussion

The computerized platform for a crime management and analysis is an overemphasis need in Sri-Lankan police departments. It enhances the proper and efficient management of crime records and help for the effective decision making and improving reliability thus improving law enforcement operations.

In this paper we have tested the accuracy of classification and analysis for prediction based on different attempts of test sets. Bayes theorem is used for the classification of data sets which proved us a result of more than 90% accuracy. Numerous news articles are trained using the Naïve Bayes algorithm and build a model which shows better results. We are inputting some test data into the model for testing purposes. Our system takes factors/attributes of a place and Apriori algorithm produces the frequent patterns of that place. Using the pattern, model for decision tree is built.

Model is built by training on these frequent patterns, which is corresponding to the place. Patterns change over the time so; crime patterns cannot be static. System is taught by means of teaching based on some inputs. So, the system could be able to learn about the changing patterns in crime by examining the crime patterns. And the factors of crime changes by time. By sifting through the crime data, we have to identify the new factors, which cause the crime. As we are concerning about only some limited factors achieving of exact accuracy is not possible. To obtain better results in analysing we have to find more crime attributes of places instead of fixing certain attributes. Till now only certain attributes are used to train the system but we are planning to include more factors to improve the accuracy.

Our system analyses the crimes and manage the prone regions in Sri-Lanka on a particular day. More accuracy can be achieved if we consider a particular region or district. Another problem we face is we are not analysing the time in which the crime is happening. As the time is an important factor in crime analysis and management, we have to predict not only the crime prone areas but also the accurate time.

References

Anon., May 2015. Oludele Awodele. *A REAL-TIME CRIME RECORDS MANAGEMENT SYSTEM FOR NATIONAL SECURITY AGENCIES,* Volume 15, p. 12.

Calhoun, 2008. *Improving Crime Data Sharing and Analysis Tools for a Web-Based Crime Analysis.* Charlottesville, University of Virginia, pp. 40-45.

Chamikara, M., 2016. Towards Intelligent Crime Investigation. *Invited Paper,* p. 12.

Crime Reports. [Online] (2009). https://www.crimereports.com/

Falaye.A, Nov.2013. Design And Implementation Of Crime Investigation System Using Biometric Approach(Nigeria Police Force). *The pacific journal of science and technology,* p. 22.

Frey, B. J., Feb. 2007. Clustering by Passing Messages Between Data Points. *SCIENCE,* p. 12.

Govardhan, N. M., September 2010. A Comparison Between Five Models Of Software Engineering. *IJCSI International Journal of Computer Science Issues,* p. 12.

H.Chung, April 2014. Crime Data Mining: A general framework and some examples. *Computer,* Volume vol. 37, p. 8.

Henry, 1998. *United States Department of Justice, Uniform Crime Reporting,* America: s.n.

Information Technology Division Sri Lanka Police. (1998) Sri Lanka Police. [Online]. http://www.police.lk/index.php/component/content/article/190

Karim H. Vellani,   2010. *Crime Analysis for Problem Solving Security Professionals in 25 Small Steps .* Houston: s.n.

Rhodes, C. J., Dec. 2007. Social Network Topology: A Bayesian Approach. *The Journal of Operational Research Society,* p. 8.

Schroeder, J., 2002. COPLINK Connect: information and knowledge management for law enforcement. *Decision Support Systems,* p. 24.

York, U. N. N., 2011. *Criminal Intelligence Manual For Analysts,* New York: UN office on Drugs and Crimes.

Zaharchuk, T. .., 2010. Study of Police Management Information Systems. *Report,* Volume I-V, p. 66.