Crime Data Mining for Indian Police Information System

Manish Gupta[1]*, B. Chandra[1] and M. P. Gupta[1]

ABSTRACT

There has been an enormous increase in the crime in the recent past. Crime deterrence has become an upheaval task. The cops in their role to catch criminals are required to remain convincingly ahead in the eternal race between law breakers and law enforcers. One of the key concerns of the law enforcers is how to enhance investigative effectiveness of the police. There is need for user interactive interfaces based on current technologies to give them the much needed edge and fulfil the new emerging responsibilities of the police. The paper highlights the existing systems used by Indian police as e-governance initiatives and also proposes an interactive query based interface as crime analysis tool to assist police in their activities. The proposed interface is used to extract useful information from the vast crime database maintained by National Crime Record Bureau (NCRB) and find crime hot spots using crime data mining techniques such as clustering etc. The effectiveness of the proposed interface has been illustrated on Indian crime records. An interactive interface as crime analysis tool has been designed for this purpose.

Keywords: E-governance, Information System, Crime Data Mining and Police System.

1. Introduction

India is a vast country with more than one billion populations, and has a police force of 1.5 million. Police is a critical component of civil administration in India. It has created its own executive apparatus to discharge assigned responsibilities. Indian constitution assigns responsibility for maintaining law and order to the states and territories, and almost all routine policing, including apprehension of criminals, is carried out by state-level police forces. The police functioning have remained a constant area of governmental concern and efforts to improve it upon further and further (Chaudhary, 2003; Krishnamorthy, 2003). Way back in 1986, the Government of India created National Crime Record Bureau (NCRB). To give right impetus to the National Crime Record Bureau, State Crime Record Bureaux (SCRBx) at States and District Crime Record Bureaux (DCRBx) at Districts followed.

In the background of enormous environmental changes and challenges before the police, harnessing of information became next to impossible. The exchange of information among police agencies became very time consuming and therefore, not available in time of need. There were neither a staffs nor time for entering data in records manually. This happened over a period of time all over the country in varying degree. In this scenario, therefore, it became increasingly difficult to coordinate information and come to any meaningful crime analysis. Investigation took a back seat and the image of police suffered.

[1] Indian Institute of Technology Delhi, Hauz Khas, New Delhi - 110 016, India
* *Corresponding Author:* (Email: manish.iitdelhi@gmail.com, Telephone: +91-11-26591493)

In order to make use of the information technology, Government of India approved the design, development and implementation of a 'Government to Government (G2G), model called the Crime Criminal Information System (CCIS). The CCIS was designed to create computerised storage, analysis and retrieval of crime criminal records. The Crime Criminal Information System today is in operation in all the States. In CCIS, the information is collected at district level not at basic unit of police administration i.e. the police station. Common Integrated Police Application (CIPA) was developed with objective of automation the processes (workflow) at police station and to provide inputs for building CCIS. Till date, CCIS is only collecting the information and creating a huge crime database and there is no analytical tool for analysing huge building database. Absence of crime analysis tool made it somewhat 'standalone' system. Therefore, there is need of support systems (Chen, et al., 2003; Amarnathan, 2003) as crime analysis tool based on current technologies to meet and fulfill the new emerging responsibilities and tasks of the Police (Chaudhary, 2003). To understand the Indian police system, an overview is given in the subsequent section of the paper. Section 3 highlights the basic features of existing police information systems CCIS and CIPA along with present status and limitations. Section 4 gives the details of the proposed crime analysis tool for CCIS along with need of the crime analysis tool for CCIS. Implementation of the proposed crime analysis tool is shown in the section 5 of the paper. Concluding remarks are given in the last section of the paper.

2. Indian Police System: An Overview

To propose any intelligent system (Gupta, et al., 2006; Michelson et al., 2006; Tuchindaet al., 2007) as crime analysis tool for police, it is required to understand Indian Police structure, responsibilities of the police, key changes and challenges the police is forcing (Krishnamorthy, 2003).

2.1 Police Structure

Superintendence over the police force in the State is exercised by the State Government. The head of the police force in the State is the Director General of Police (DGP). States are divided territorially into administrative units knows as districts. A group of districts form a range, which is looked after by an officer of the rank of Deputy Inspector General of Police (DIGP). Some States have zones comprising two or more ranges, under the charge of an officer of the rank of an Inspector General of Police (IGP). A Senior Superintendent of Police (SSP)/Superintendent of Police (SP) is the head of the district police administration and is assisted by an Assistant Superintendent of Police (ASP) and few Deputy Superintendents of Police (DSP). A district may have many Police Stations that are manned by Inspectors, Sub Inspectors, Assistant Sub Inspectors, Head Constables and Constables. Police Station is the basic unit of police administration through which both crime and non crime duties are discharged. Police Stations are the places where complaints and First Information Reports (FIRs) are lodged. Police Stations also serve as the window of 'citizen interface' for the police. Common people approach Police Stations for assistance. Therefore, public expectations from police stations are more direct, pressing and at times extremely demanding. Operationally, police stations are at the nucleus of all policing activities. All important operational duties – be it duties to the State or services to other government departments or citizens- are executed and coordinated through police stations.

2.2 Role of Indian police

With time, the role of the police has expanded and is expanding steadily to areas having larger public interface. The major existing roles of police are prevention of criminality, repression of crime, apprehension of offenders, regulation of non-criminal conduct and recovery of stolen property etc. Whereas, the emerging police roles are to protect life and property, to reduce the opportunity to commit crimes, to maintain social order and to protect the individual freedom and privacy etc. The concept of proactive policing is fast catching up the imagination in modern times across the globe. This includes emphasis on community policing and problem solving policing

2.3 Police E-Governances

It dates back to 1986 when National Crime Record Bureau (NCRB) was created with the aim to help the investigating agencies by providing them with extensive and updated information on crime and criminal data at State and National level such as modus operandi, personal data, finger print, photograph, criminal history and details of property which may be subject matter of crime. It is a pioneer organization for providing information technology training to senior police officers and law enforcement agencies. The Crime Criminal Information System (CCIS) is the flagship crime software in Indian police used across the country. Developed under the aegis of the National Crime Record Bureau (NCRB), the Crime Criminal Information System (CCIS) is a Government to Government (G2G) model that aims at creating computerized storage, analysis and retrieval of crime criminal records. The details of CCIS is given in section 3 of the paper.

The National Crime Record Bureau (NCRB) has developed following e-governance applications systems:

- Police Station Management System
- Prison Statistics
- Jail Management Softwares
- Prosecution Branch System
- National Bomb Squad System and Forensic Science Laboratory System.
- Motor Vehicle Information Counters (MVIC

At the level of State police, State Crime Record Bureaux (SCRBx) are playing key role in pushing the e-governance agenda. Most of State police administrations have appointed senior level officer for implementation of e-governance. Apart from the Crime Criminal Information System (CCIS), State police computerization programmes include many local initiatives mainly for internal operations such as Village Information System and Sewa-100 etc.

3. Existing Indian Police Information Systems

3.1 Crime Criminal Information System (CCIS)

CCIS is a national project of sharable database on crime and criminals at district, state and national level for assisting investigating and supervising officers and police planners to formulate crime-control strategies. It has been upgraded to CCIS MLe in the year 2005 as multi-lingual application with facility for different local languages i.e. Marathi, Gujarati, Tamil, Kannada and Gurmukhi besides English and Hindi. The application has been web-enabled so that field level investigating and supervisory officers can access the CCIS databases at national and State levels through internet anywhere – anytime.

The main objective of the CCIS are to

- Computerize crime and criminal information collected by Inquiry Officer
- Link a crime to criminal and property
- Link a criminal to crime
- Link a property to crime
- Link unidentified dead bodies, kidnapped persons, missing etc.
- Generate vital reports from database
- Reduce manual effort and increase efficiency, of police
- Restoration of stolen property to legitimate owner

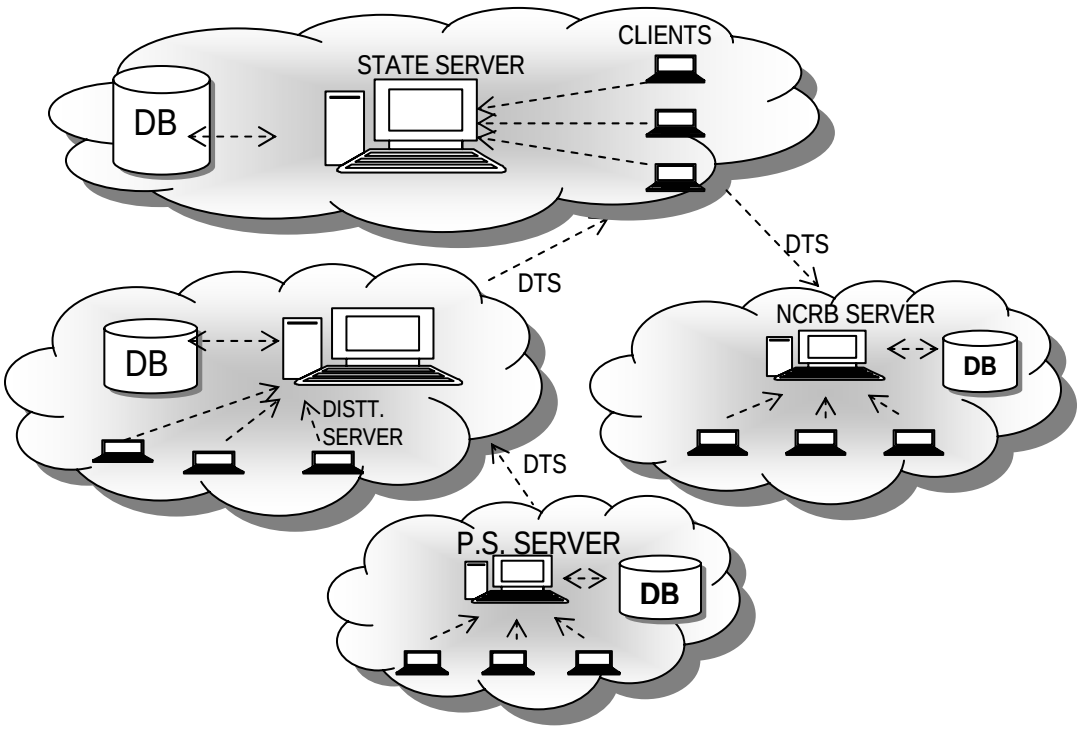

Figure 1: Architecture of CCIS

Present Status: CCIS is perhaps one of the biggest police application in the world implemented at 35 States and UTs, 727 police districts and at national level. In CCIS, the information is collected at district level not at primary source of information i.e. the police station. Common Integrated Police Application (CIPA) was developed with objective of automation the processes (workflow) at police station and to build a Crime & Criminal Information System. CIPA is described in the subsequent section of the paper. Absence of intelligent decision support system made it somewhat 'standalone' system and no analytical tools for analysing huge building database.

3.2 Common Integrated Police Application (CIPA)

Common Integrated Police Application (CIPA) aims at automation of all functions carried out at primary source of information itself the police stations. CIPA software has been designed and developed by NIC in English language with multilingual interface developed for Indian languages. This project is aimed at building in a planned manner infrastructure and mechanism to provide the basis for evolution for Crime and Criminal Information System (CCIS) which is uniform across the country from police station level onwards. CIPA was developed for planned induction of Information Technology towards better management of public order and criminal activities.

The software is a total work-flow system having the following three major modules along with reports and queries viz., Registration, Investigation and Prosecution. In the first phase the FIR registration process would be computerized and after familiarity in working with computers the staff would be gradually made to computerize further processes including investigation. Implementation of CIPA Project would make all the police stations work more efficiently and effectively thereby improving accessibility, transparency and accountability in the functioning of the Police Department in the State.

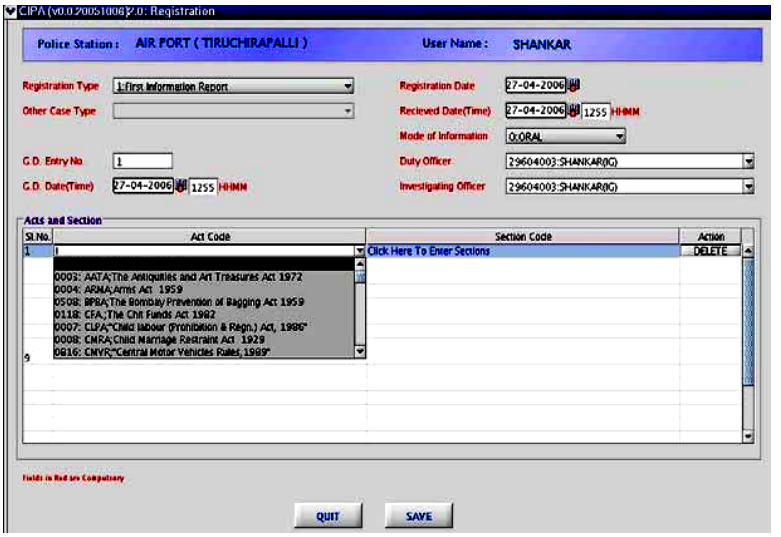

Figure 2: Registration Module of CIPA

The major benefits from the CIPA are as follows

- Significant reduction in manual register maintenance
- Elimination of duplicate and inconsistent record keeping
- Facilitate for maintenance of details of criminals
- Keep track of the Status of Cases
- Introducing element of transparency in the working of Police
- Facilitating investigating officer with availability of records
- Facilitating supervision by the Senior Officers
- Generate various reports required from time to time

Present Status: It is being implemented by NIC and coordination and monitoring is being done by NCRB.Initially only Registration Module has been taken up. Implementation of CIPA project at police stations is being done in phases. 10% Police Stations (about 1400 Police Stations) are being covered throughout the country in the first phase by 2007. 30% Police Stations (Approximately 3,700 police stations) will be covered in second phase and remaining in final phase. It has been implemented in all the Police Stations of Delhi.

4. Proposed Crime Analysis Tool based on Crime Data Mining

In today's world criminals have become techno savvy and they make maximum use of all the modern technologies and methods in committing crimes. This has facilitated them in operating over the length and breadth of the country also. If we have to effectively meet out challenges of crime control and maintenance of public order, creation of databases on crimes & criminals in digital form for sharing by all, an intelligent police information system (Chen, et al., 2003; Ozkan, 2004) cannot be neglected anymore. This section highlights the need for Crime Analysis Tool as interactive interface and describes the proposed Crime Analysis Tool for CCIS based on Crime Data Mining.

4.1 Need for Crime Analysis Tool as Interactive Interface

In the present scenario, the criminals are becoming technologically sophisticated in committing crimes (Amarnathan, 2003). Therefore, police needs such a crime analysis tool to catch criminals and to remain

ahead in the eternal race between the criminals and the law enforcement. The police should use the current technologies (Corcoran,et al., 2003; Ozkan, 2004) to give themselves the much-needed edge. Availability of relevant and timely information is of utmost necessity in conducting of daily business and activities by the police, particularly in crime investigation and detection of criminals. Police organizations everywhere have been handling a large amount of such information and huge volume of records. There is an urgent need to analysing the increasing number of crimes as approximately 17 lakhs Indian Penal Code (IPC) crime, and 38 lakhs local and Special Law crimes per year.

CCIS has facility only to view multi-dimensional view of data on crime, criminal and properties and not analysing crime data e.g crime hot spot identification, crime zone identification, crime trends prediction so that police planners can enforce their manpower at high density areas and crime hot spots to reduce and control the crime. Integration of Analytical Tools for providing support to decision makers are very much needed for effective use of the Crime Criminal Information System in the Police stations, Districts and States.

The police require interactive interface as crime analysis tool that assists them in

- Detecting of crime locations and carrying out crime hot spot analysis
- Providing information to formulate strategies for crime prevention and reduction
- Reducing the further occurrences of similar incidence by analyzing crime patterns
- Targeting resources for preventative and detection strategies more efficiently

4.2 Proposed Crime Analysis Tool for CCIS

The proposed crime analysis tool for CCIS is faster to implement and easier to use. The tool will give an edge to the police to respond and thereby serve the people better. The proposed crime analysis tool provides a user-friendly tool to analyze huge crime database as building up in CCIS. The traditional way of executing the queries is the extraction of records and aggregates them for every execution. This is more time consuming process because of rescanning of database every time and required a specialized user to extract the information from the database. The proposed crime analysis tool eliminates the rescanning of the database for every new query as well need of skill users. It provides the user an interactive and fast way to carry out process of identification of crime hot spots and crime zones as well as data comparison among various area of interest. Interaction of many relational tables is required for analyzing crime data since not all information required for crime analysis is stored in a single table. The proposed query interface extracts the records from these tables and aggregates them for further online querying. The crime analysis tool as an adaptive query interface has been designed to make best use of the existing CCIS. Police station level analysis is also possible through the interface.

Crime Analysis Tool consists of three major modules such as Data Extraction Module, Crime Analysis Module and Data Information and Comparison Module for carrying out analysis on CCIS. The features and utility of each module is described in the next subsection.

4.2.1 Data Extraction Module

Data Extraction module establishes connection from CCIS database to Crime Analysis Tool & also extracts records state-wise for carrying out further crime analysis. An ODBC data source is required before running this module to create a link from CCIS database. To establish database connection some inputs such as database name, username and password are required from the user. The extraction of records is also carried out through this wizard. This module is need not to run every time for carrying out crime analysis on the same data for same period, only first time it is required to run for extracting records. For example, if a user has already extracted the records of Karnataka State for period 2000-2006 then user needs not to run Data Extraction Wizard for the same state & same period. But if user would like to run a query for Punjab state

or any other state, he/she has to run the Data Extraction Wizard for extracting records of the desired state. UPDATE option will save lots of time for future coming records since it will run the query on the latest data and update the already extracted data files.

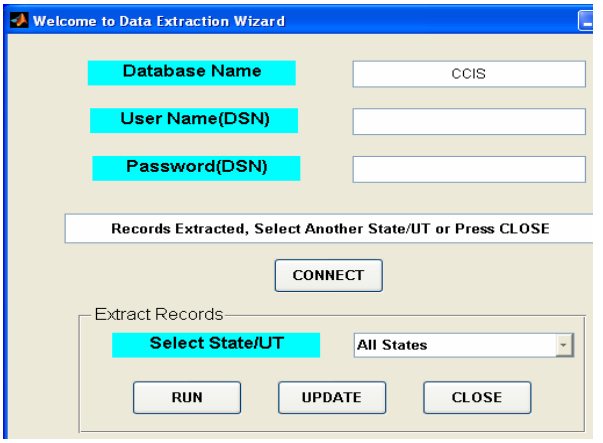

Figure 3: Data Extraction Wizard

4.2.2 Crime Analysis Module

Crime Analysis Wizard will identify crime hot spots and crime zones of a particular region on certain crime types for a specific period. This interface will provide a tool for making an online query and based on the query, crime hot spots are identified. The user can select a particular State/UT from the list. List contains only those State/UT for which data extraction has been taken place. A user can select any year as well as multiple years with the help of Year Selection Input Dialog Box, which will appear after select a year. The module has also provided the facility to select crime types whether from all 105 listed crime types by selecting Selection Type as Crime Type or from one of the major crime heads i.e. crime against body, crime against women, crime against property and kidnapping and abduction by selecting Selection Type as Crime Head. Crime Head as selection type helps the laymen user for analyzing crime hot spots and crime zones of general type of crimes without going into the specific details.

The results of query consist of crime hot spots, high crime zones, moderate crime zone and low crime zones based on the average density of these crime. Cluster center as average density of various crime zones are also given in the results of the query interface. The results obtained using proposed adaptive query interface will be helpful in identifying the crime hot spots, predicting crime trends for the crime hot spots which will ultimately help in controlling the crime. The user can view the results of the state and any of its by selecting a particular district from the options provided in the result panel. The user can also save the results of specified query into a text file and also analyze the results by using Data Information and Comparison Module which is described in the subsequent section.

4.2.3 Data Information and Comparison Module

Data Information & Comparison Wizard displays comparative crime plotting of specified query fired in crime analysis wizard. It displays plot for a particular type of crime selected from the Crime Type popup Menu. User can see the plots of the State/UT/Districts by selecting in State/UT/Districts List Box. Label List is also shown for the selected locations. User can also save the data files containing crime data information for any state and districts under various crime types. Figure 5 displays crime plots of one of the district of Karnataka state for year 2002-2006 under murder crime.

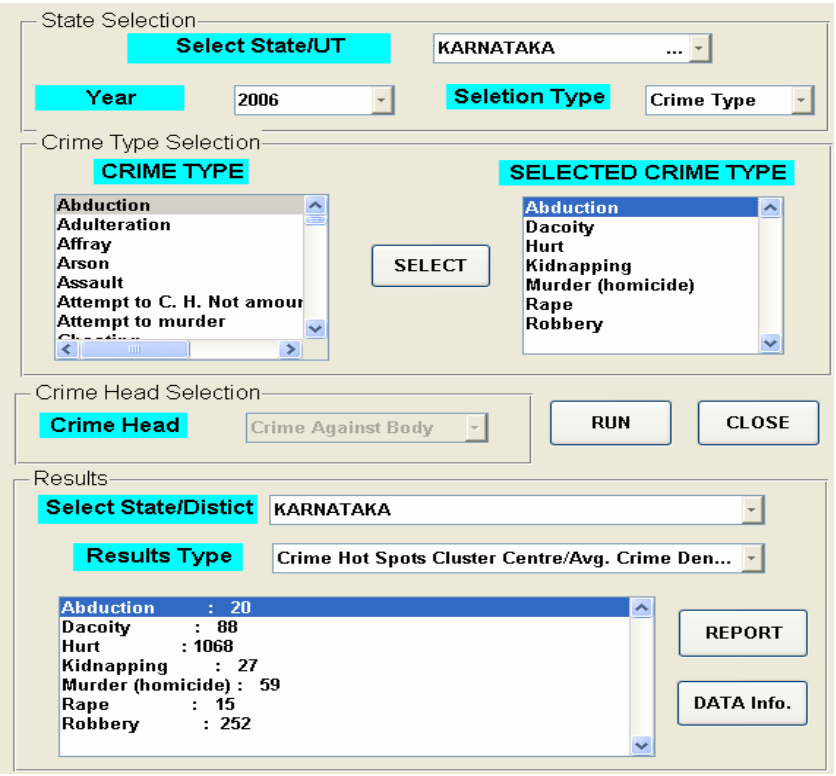

Figure 4: Crime Analysis Wizard

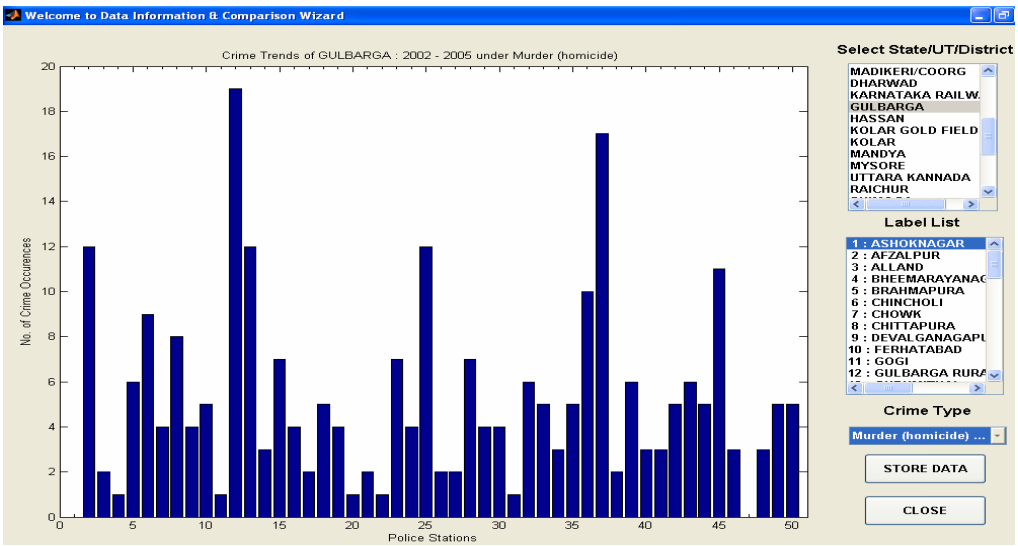

Figure 5: Data Information & Comparison Wizard

5. Implementation of Crime Analysis Tool

To demonstrate the utility of the crime analysis tool, it has been implemented on CCIS database of National Crime Records Bureau (NCRB) of India. The tool is used on crime data of Karnataka state at initial stage

but later on extended for carrying out analysis on all India crime data. Karnataka state has more than 20 lakhs of crime records from 1991 to 2006 and more than 2 GB in size. The proposed crime analysis tool performs remarkably well to handle such database and carrying out crime analysis for Karnataka state. Crime hot spots have been identified for the state under different crime heads and for different period. Figure 4 shows Crime Type Selection and Abduction, Dacoity, Hurt, Kidnapping, Murder, Rape and Robbery crime has bee selected to carry out analysis for Karnataka state for year 2006. The result shown in the Figure 4 is average crime density of selected for crime hot spots districts of Karnataka state for year 2006. The crime records have also been analyzed using Data Information & Comparison Wizard. Crime plots of one of the district of Karnataka state under murder crime is shown in the Figure 5.

6. Concluding Remarks

In this paper, a interactive interface as crime analysis tool for CCIS based on the current decision support and data mining techniques has been proposed in order to carrying out police activities efficiently. The proposed crime analysis tool will provide an upper edge and with the use of the crime analysis tool policing can be made effective, fast and responsible in their operation. The successful implementation of the proposed crime analysis tool for India crime records depicts the utility of the tool. The effectiveness of proposed crime analysis tool has also been illustrated for crime hot spot analysis. The proposed crime analysis tool for CCIS is faster to implement and easier to use. The crime analysis tool can be integrated with latest visualization techniques such as Geographical Information System for enhancing the understanding of the results and patterns. The tool has very promising use in the current changing scenario and provides an effective tool to law enforcement agencies for crime detection and crime prevention. Hence, the proposed crime analysis tool has wider variety of application for enforcement of laws.

Acknowledgements

This paper is an outcome of a project funded by NCRB. We are highly indebted to Shri Sudhir Awasthi, Director, NCRB for funding the project. We are also thankful to him and his team members for sharing their valuable knowledge in this field.

Reference

1. Chen H., Zeng D., Atabakhsh H., Wyzga, W. & Schroeder, J. (2003). COPLINK: managing law enforcement data and knowledge. *Communications of the ACM* 46 (1), 28–34.
2. Corcoran J. J., Wilson I. D. & Ware J. A. (2003). Predicting the geo-temporal variations of crime and disorder. *International Journal of Forecasting* 19, 623–634.
3. Chaudhary, J.N. (2003). Police in United States: Contemporary issues. *The Indian Police Journal*, July – September, 173-179.
4. Kumar, M., Gupta, A. & Saha, S. (2006). Approach to Adaptive User Interfaces using Interactive Media Systems. *Proceedings of the 11th international conference on Intelligent user interfaces*.
5. Amarnathan, L.C. (2003). Technological Advancement: Implications for Crime. *The Indian Police Journal*, April-June.
6. Michelson, M. & Knoblock, C.A. (2006). Phoebus: A System for Extracting and Integrating Data from Unstructured and Ungrammatical Sources. In *Proceedings of AAAI*.
7. Ozkan, K. (2004). Managing data mining at digital crime investigation. *Forensic Science International* 146, S37-S38.
8. Krishnamorthy, S. (2003). Preparing the Indian Police for 21st Century. *Puliani and Puliani*, Bangalore, India.
9. Tuchinda, R., Szekely P. & Knoblock, C. A (2007). Building Data Integration Queries by Demonstration. *Proceedings of the 12th international conference on Intelligent user interfaces.*

About the Authors

Manish Gupta is a doctoral student in the Department of Mathematics, Indian Institute of Technology, Delhi (IIT Delhi). He has previously worked as a Junior Research Fellow (JRF) at Department of Management Studies, IIT Delhi.

He has been awarded Junior Research Fellowship under Council of Scientific and Industrial Research (CSIR) schemes in 2001. His research interests include Intelligent Decision Support System, Data Mining, Clustering and Artificial Neural Network.

B. Chandra is a Professor in the Department of Mathematics, Indian Institute of Technology, Delhi. She belongs to the computing group of the Department.. She has published a number of research papers in reputed International journals in the area of Neural Networks, Classification, Clustering and Association rule mining.. .She has been a visiting Professor for a year at the Graduate School of Business, University of Pittsburgh, USA and worked at World Bank at Washington D.C. During 1998-1999 she has been a visiting Professor Penn State University, USA. She has also been a visiting Scientist at INRIA, France. She has been Chairman in the session on Neural Networks at various International conferences held in USA,UK, Canada ,France and Singapore. She has given invited lectures at various universities in USA: University of Maryland, University of Hawaii, University of Pittsburgh, Pennsylvania State University Kaist ,Korea, University of Uttercht, Netherlands, Ecole de Mines Paris and university of Bangor U.K. She has been a principal investigator of many research and consultancy research projects in the of Neural Networks and Machine learning. Among the recent ones are Identification of Crime Hotspots using clustering funded by NCRB and Clustering for developing efficient trading strategies for stock market data She is also the author of three books.

M. P. Gupta is Chair-Information Systems Group & Coordinator-Center for Excellence in E-gov at the Department of Management Studies, Indian Institute of Technology (IIT Delhi). His research interests lies in the areas of IS/ IT planning and E-government. Prof.. Gupta has authored acclaimed book “Government Online” and edited two others entitled “Towards E-Government” and “Promise of E-Government”, published by McGraw Hill, 2005. His research papers have appeared in National and International Journals/Conference Proceedings. He was the recipient of the prestigious Humanities & Social Sciences (HSS) fellowship of Shastri Indo Canadian Institute, Calgary (Canada) and a Visiting Fellow at the University of Manitoba. He supervised e-government portal “Gram Prabhat” which won the IBM Great Mind Challenge Award for the year 2003. He has steered several seminars and also founded the International Conference on E-governance (ICEG) in 2003 which running into fifth year. He is on the jury of Computer Society of India (CSI) E-gov Awards and also a member of Program Committee of several International Conferences. He is life member of Global Institute of Flexible Systems Management (GIFT) and Systems Society of India.