ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 10, December 2012

Criminals and crime hotspot detection using data mining algorithms: clustering and classification

Sukanya.M[1], T.Kalaikumaran[2] and Dr.S.Karthik[3]
[1] II year M.E SE, Department of Computer Science and Engineering (PG)
SNS College of Technology, Sathy Main Road, Coimbatore-641035, India

[2]Professor and Head, Department of Computer Science and Engineering (UG &PG)
SNS College of Technology, Sathy Main Road, Coimbatore-641035, India

[3]Dean, Department of Computer Science and Engineering (UG&PG)
SNS College of Technology, Sathy Main Road, Coimbatore-641035, India

***Abstract*— Criminal activities have been increased tremendously. So security has become a major issue to the people. The control actions should be taken by the police department. The police department has to take corrective actions for reducing the criminal activities. To analyze the criminals' data, clustering and classification techniques are used. These data will be stored in the criminals' database. Spatial clustering algorithm and structured crime classification are used to classify the criminal activities. These algorithms help to identify the hotspot of criminal activities. Criminals will be identified based on the witness or clue at the crime spot. The identification of hotspot of criminals' activities will help the police department to provide more security to the particular area, this helps to prevent crimes in future. When we apply this concept to all areas at least we can reduce criminal activities. Completely the crimes cannot be controlled. In this paper we find the hotspot of the criminal activities and finding the criminals by using clustering and classification algorithms.**

***Index Terms*— Burglary, classification, clustering, crime analysis, crime mapping, GIS, hotspot, spatial pattern.**

I. INTRODUCTION

Security and crime forecasting activities are most important concerns for both citizens and government. Crime is an integral part of risks. We are facing it in everyday. Which are very harmful activities. Crime activities are spread throughout the world. In some of the places the crime activities are absent, which is not spread evenly in all the places. The crime activities have increased, but police department is the responsible for reducing and controlling the crime activities. Crime prevention and criminal identification are the major issues to the police department. The challenge activities are analyzing the crime and arrest the criminals. This is most difficult activity. The major disadvantage is, there are large number of data regarding crime activities and criminals, so we need better knowledge about the crime and criminals. The data about criminals will be stored in the database; these data will be available in all the police stations. To identify the criminal activities and criminals there are methodologies available to reduce it. Where more number of criminal activities have happened that location is identified as hotspot. By identifying the hotspot of the criminal activities, this will help the police department to avoid such kind of activity in future in all the location. Based on the result in which location more number of crimes occurred, for the particular location more protection will be given by the police department. Through this way we can prevent and identify the crimes and criminal activities. The data mining concept is very much helpful to analyse the crimes and criminals. Classification and clustering algorithms are applied. Before applying the clustering algorithm, classification is applied. Based on the type of crime, the crime activity will be classified, after the classification is done based on the classification result, the similar type of crime activities will be grouped together. By using GIS the hotspot of the criminal activities will be viewed.

I. PROTECTION ON INFORMATION

To regionally maps the one agency with other agencies to serve various information to substations and remote offices, the internet and intranet. By using this many number of applications will be served from a single server to more number of clients. It serves the dynamic maps and data quickly on different servers. Both data and applications

administrations are centralized. Several reports will be accessed by officers for a user specified period, number of incidents occurred in a particular address, at which time the incident is occurred. Another benefit of intranet based application is team's ability for breaking the information. The mapping functions will be distributed to all the district offices. Any officers or detective can generate the customized maps. The intranet based applications should be very secure and from the unauthorized users. To maintain the information as confidential one the system should be controlled by the passwords.

II. CRIME MAPPING

Crime mapping helps the police department to protect the people from the crime more effectively. An understanding of where and why crimes occur will help to fight against with the crime. Simple map shows where the crimes have been occurred.

A. Visualizing The Crime Location

Digital maps visualize the crime scenario in quicker manner. At which place the crimes have occurred that will be visualized. Instead of searching from the list of events, mapping is easy to visualize the crime hot spot.

B. Integrate the community characteristics

Community characteristics mean the most possible places for occurring the crime activities. For example slums, schools, parks, colleges, alcohol permit location and etc.

C. Producing the maps

At any geographical level the maps can be produced. Where the crimes have occurred that particular place will be shaded darker. The number of crime incidents percentage change will be displayed by shading the area's location.

II. CRIME ANALYSIS

Crime analysis is a set of systematic and analytical process for providing the information regarding crime patterns at the particular time. Crime investigation is an important activity for identifying the crime hotspot. It supports the number of department functions that includes patrol deployment, special operations, tactical units, investigations, planning and research, crime prevention and administrative services. Crime analysis can be divided into three categories, these are following as,

A. *Tactical*

Tactical is an analytical process for providing the information to assist operations personnel (patrol and investigative officers) for identifying the crime trends, patterns, series and hotspot. It includes at which time crime is occurred and associating the criminal activities by crime method.

B. *Strategic*

It includes the preparation of crime statistical summaries, resource acquisition and allocation studies.

C. *Administrative*

It focuses on provisioning on economic, geographic or social information to administration.

D. *Identification of clusters*

GIS identifies the areas that contains the more number of clusters (hotspot). The similar type of crime activities will be grouped together. Based on the clusters' result, which cluster contains the more number of criminal activities that will be called as crime hotspot for the particular crime.

E. *Comparison of location of crime hotspot*

The crime hotspot that have been identified over several months.

F. *Comparison of hotspot with different crime types*

The identified crime hotspot will be compared with the other type of crime hotspot. For example burglary type of crime hotspot will be compared with the murder type of crime hotspot.

III. CRIME CLASSIFICATION

To classify the crime incidents based on the similarity between the crime objects stored in the class, structure crime classification is used. Classification is the hierarchy of these attributes. These attributes are represented by classification in three ways,

1) Classification of crime place
2) Classification crime types
3) Classification of crime time

The structured crime classification algorithm is used to identify the more similar objects in the data sets.
Algorithm, to find the hotspot and coldspot from the dataset.

Input: Database DB
Output: hotspot or coldspot

1. Assign S=DB
2. Apply purification attribute Ai by Cn]
3. Repeat
   a. Find the similarity of crime attribute objects(C.Ai, C.Ai+1)

b. Find the probability of particular crime classification = Probability (C.Ai, Classification)
c. Threshold T=(Cluster Area-Sparse Area)
d. Find F(C)=classification U Pi(Ci)
e. If F(C)>positive description
Produce a hot spot
Else
Produce a cold spot

4. Go to step 3

Let S denotes a set of crime incidents. Ai be an attribute of crime incidents and Ci be a classification of each crime attribute Ai. For two elements x1, x2 in the tree of Ci, if there is a path from x1 to x2 is called the parent of x2. Further more, x1 is a generalization of x2. In the structure crime classification algorithm , the national dissipation between the events is similar and the events are more similar. Choose the crime attribute Ai in the crime class C. Find the similarity of each crime attribute of crime objects if both objects have the same similarity, join these two objects have the same crime attribute incident and put into the same class C. And finally find the F(C) based on the probability of crime incident occurring in the particular class to which it is merged. If F(C) is greater than the positive description, it produces a crime hot spot, otherwise it produces a crime cold spot.

IV. CRIME CLUSTERING

Clustering is data mining technique for grouping the similar type of crimes will be grouped together. In this paper the burglary crime will be clustered, based on the clusters' result the crime hotspot will be identied.

CONCLUSION

This paper presents the method to identify the hotspot of crime. Based on the type of crime the police department can easily identify the hotspot of the burglary crime. GIS is used to visualize the hotspot of burglary crime. Data mining concept is used to prevent and identify the crimes. Classification technique is used to classify the different crimes. Clustering technique is used to cluster the similar type of crimes together, based on the clusters' result the burglary type of crime hotspot will be identified. This result will help to reduce the burglary type crime. In future all type of crimes' hotspot will be identified, through this the crime activities will be reduced.

REFERENCES

[1] Sajendra Kumar, Ram Kumar Rana, Pawan Singh, "A Semantic Query Transformation Approach Based on Ontology for Search Engine" ISSN : 0975-3397 Vol. 4 No. 05 May 2012, International Journal on Computer Science and Engineering (IJCSE).

[2]Lizhen Li, Zhifeng Dong, Keming Xie, *Ontology of general concept for Semantic Searching*, Second International Conference on Computer Modeling and Simulation 2010.

[3]Yuangui Lei, Victoria Uren, and Enrico Motta, *SemSearch: A Search Engine for the Semantic Web*: IEEE Transactions on knowledge and Data engineering, VOL.19, NO. 2, FEBRUARY 2007.

[4]Thomas B. Passin, *Explores Guide to the Semantic Web*, Manning Publications Co. GreenWich 2004.

[5]Ronen Feldman, James Sanger, *The text mining handbook: Advanced approach in analyzing unstructured data*, Cambridge University Press USA 2007.

[6]C.P.Johnson "Crime Mapping and Analysis using GIS", Geomatics 2000, Conference on Geomatics in Electronic Governence, January 2000, Pune.

M.Sukanya was born on May 1990 at Dindigul, Tamilnadu, India, received B.E degree in Computer Science and Engineering from Angel College of Engineering and Technology, Tirupur in 2011. She is currently pursuing M.E degree in Software Engineering at SNS College of Technology, Coimbatore. She has presented three National and two International conferences. And published one paper in international journal.Her area of interests are Data Mining, Cloud Computing and Software Engineering.

T.Kalaikumaran received MCA degree from the Madras University in 1996, and the M.E. degree in Computer Science and Engineering from the Anna University, Chennai, in 2006. He is a Postdoctoral fellow at the Anna University of Technology, Coimbatore. He is interested in the research areas of data mining, spatial data mining, machine learning, uncertain data classification and clustering, pattern recognition, database management system and informational retrieval system. He is a member of CSI and IEEE.

Dr.S.Karthik is presently Professor & Dean,Department of Computer Science & Engineering, SNS College of Technology, affiliated to Anna University, Tamilnadu, India. He has received his Master Degree in Engineering from Anna University Chennai and Ph.D degree from Anna University Chennai in the area of information and Communication Engineering. His research interests include network security, web services and wireless systems. In particular, he is currently working in a research group developing new Internet security architectures and active defense systems against DDoS attacks. Dr.S.Karthik published papers in international journal and conference papers and has been involved many international conferences as Technical Chair and tutorial presenter. He is an active member of IEEE, ISTE and Indian Computer Society