DATA MINING APPLICATIONS FOR COUNTER TERRORISM

SATISH.C.PANDEY, DR.A.K.SHARMA, K.S.PANDEY

Abstract: Following the terrorists attacks on America and India on 11 Sept 2001 and 26 Nov 2008, respectively, many countries became more vigilant about their national security. In the light of these attacks, several means have been adopted to assist the law enforcement agencies to identify terrorists and terror related activities. One of such means is the use of computer technology and computer analysis for effective detection of terrorist's activities. Various Data mining techniques can be applied by the law enforcement agencies to analyze information and to track the nefarious activities of terrorists. In this paper we have discussed some Data mining techniques to be adopted by the law enforcers in tracking the activities of terrorist and their criminal activities. Finally the limitations of Data mining in fighting terrorism have also been discussed.

Keywords: Data mining, Terrorism, Data warehouse, Artificial Intelligence, Prediction, Crime prevention

Introduction: Data mining deals with discovery of unexpected patterns and new rules that are " hidden " in large databases .It serves as an automated tool that uses multiple advanced computational techniques, including artificial intelligence(the use of computers to perform logical functions),to fully explore and characterizes large data sets involving one or more data sources, identifying, significant, recognizable patterns trends and relationships not easily detected through traditional analytical techniques alone. This information then may help with various purposes, such as the prediction of future events or behaviors. [1]

Fayyad et al [2] defined Data mining as a process in the Knowledge Discovery Data base (KDD) which is a nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data. Data mining Techniques have been applied in many applications such as Marketing [3], Healthcare Organizations[4], Pharmaceutical Industries [5], Education[6], and Counterterrorism. [7, 8]

Related Work in Counterterriosm: Today, due to sophistication in technology terrorists are carrying out their nefarious activities throughout the world on larger scale. They seldom operate in a vacuum but interact with one another to carry out their nefarious activities. They have a high degree of associative relationship among members of the network. For message gathering, information leaking and in the execution of crime their networks are linked to one another.

Detection of terrorists and their groups and solving the crime is not an easy task and is a challenge for law enforcers .However with the increase in sophistication in technology computer system are now being utilized in tracking terrorists and their activities and computer data analysis have helped the law enforcers and detectives to enhance the process of detection of terrorists. Cate [9] has defined Data mining as a promising tool in the fight against terrorism. It plays several important roles in terrorist's detection including location of suspects, identification and tracking suspicious financial and other transactions,

According to Scifert [10] Data mining is a key feature in the fight against terrorism and crime.

Thraisinghanuman [11] has pointed out that Data mining can be used to detect unusual patterns, terrorist activity and fraudulent behavior. According to DeRosa [12] Data mining and automated data analysis techniques are powerful tools for intelligence and law enforcers who fight against terrorism.

Okonkwo et al [13] has discussed how Data mining techniques can be adopted by law enforcement agencies in the tracking the activities of terrorists and their criminal activities. They have also pointed out the limitations of Data mining in fighting the crimes.

Elovici et al [14] has presented in his research paper an innovative knowledge based methodology for terrorist detection by using web traffic contents. Their proposed

methodology learns the typical behavior (profile) of terrorists by applying a Data mining algorithm to the textual content of terror related websites.

DeRosa [12] in his paper has narrated that automated data analysis techniques can be useful tools for counterterrorism in a number of ways and one initial benefit of the data analysis process is to assist in the important task of accurate identification.

In the present paper Data mining techniques in detecting terrorists and terror related activities have been discussed with simple example.

Main Steps of Data Mining Process: The process of Data mining is simply the collection of data into a single repository where Data mining algorithms are applied for knowledge discovery and pattern recogniton.Fig1 shows the simple process of Data mining in which data from various sources are firstly gathered together in a repository commonly known as data warehouse. After that Data mining techniques are applied for pattern evaluation, recognition and analysis.

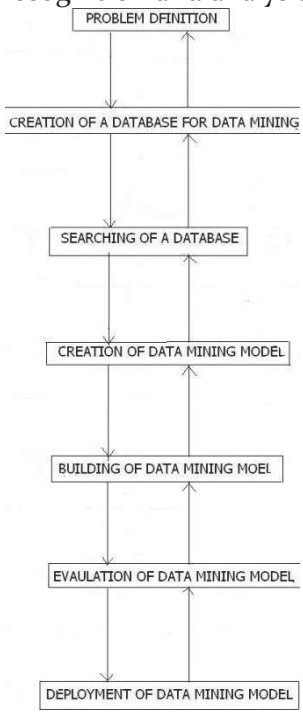

Fig-1 Main Steps of Data Mining Process

Crime Combating by Data Mining Techniques: Moreover Data mining applications also use a variety of parameters to examine the data .According to Thraisinghanuman [11] there are the methods to apply Data mining techniques in terror or crime detection:

One top down reasoning where we start with a hypothesis and then determine whether the hypothesis is true.

The other bottom up reasoning where we start with sample and then come up with a hypothesis. In the case of terrorist detection the first method i.e. top down approach is adopted by asking or stating the hypothesis who committed the crime/attack or it is certain person who carried out the attack.

The next step is to start searching as to the likely reasons of the attack and the individuals who might have responsible attack, we have already pointed out that terror investigation is the prerogative for the law enforcers concern while computer and computer analysis is useful for solving detection.

Classification: Wies at al [15] has pointed out that classification is a Data mining techniques which produces the characteristics to which a population is divided based on the characteristics .According to Thraisinghanuman [11] classification divides the dataset based on certain predefined condition. In case of crime classification assumes that we have some idea of the individuals (suspects) based on the predefined criteria.

Example: let us suppose that the law enforcement agencies have reported a kidnapping case then they may try to form the idea of the kidnappers, say 4 males between 25 and 30 of age. Muslim speaks Urdu frequently between 5' and 7'tall.Then the classification has been completed and it becomes imperative to place all males meeting the above criteria under suitable observation. The algorithm which is to be used will be such that the population will be divided into two clear parts male and females.

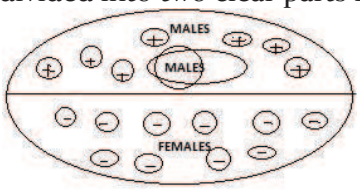

Fig-2 Classification algo depicts the segments of a dataset based on some data criteria

The algorithm will also segment the males according to their suspected criteria.Fig-2 indicates a typical segmentation scenario where males are segmented from females. The male data set could be further segmented based on our criteria. As a next step we scan the historical dataset (Data Warehouse) for concerned

matches. If the searches do not produce positive result we could narrow the search by again classifying the classified dataset. Assuming that we were not able to obtain a class the new criteria will then be used to update out historical database (Data Warehouse) if then another related crime occurs it becomes easier to form a match.

Fig-2 shows a classification algorithm .The data warehouse is classified here based on the criteria for suspected individual's .In our example the data set is classified into males and females hence this classification is not adequate. Therefore the process continuous until we have a dataset which can be matched against the criteria. Our classified database is now compared against the criteria if we have a related match then it is placed under surveillance otherwise the process continuous until there are no more matches.

Link Analysis: Link analysis is also a Data mining techniques which is very advantages in extracting valid and useful patterns .The theoretical structure of link analysis is based on the fact that incidents are connected to one another and is therefore mutually exclusive.

Data Mining Limitations: The limitations of Data mining areas under: Firstly, while Data mining products can be very powerful tools they are not self sufficient applications. For its success Data mining requires skilled technical and any analytical specialists who can perform the analysis and interpret the output that is generated. Consequently the limitations of Data mining are primarily data or personal related rather than technology related.

Secondly the Indian law enforcement agencies do not have already at hand the data on individuals and corporate data which are needed to be used to track the potential terrorists and crime perpetrators. Although some data might be available here and there but they do not have a comprehensive historical database where they can analyze and extract information.

Thirdly, Data mining techniques are very sensitive in the quality of data input; data which we find are often incorrect, error prone and above all a large quantity of dataset need to be dealt with in order to provide useful data which can help the detectives and law enforcement agencies.

Finally Data mining undoubtedly erodes our privacy.

Recommendations: Indian Government agencies should set up Data mining agencies within the law enforcement agencies where various data can be consolidated and mined. Individual data such as voters identity card, national identity such as Aadhar Identity Card, population census information etc should be linked together to profile the identity of an individual. Bank accounts, phone numbers and other related activities could easily be searched to any individuals.

Institutions and companies in India should corporate more with law enforcement agencies by reporting cases of irregular transaction patterns to the government agencies from time to time. They should not wait for the law enforcement agencies to request for the same. Moreover the law enforcement agencies should be equipped well in the field of computer technology and computer analysis in order to crack some of the grueling Data mining techniques and be able to link cases to suspected individuals .Proper training should be given to the staff members of the law enforcement agencies.

Government and corporate bodies together should set up data ware housing and mining institutes where academics, professional law enforcement agencies can interact to generate better Data mining algorithms which may be capable of detecting terrorists and crime activities very easily.

Conclusions: For combating terror attacks, criminal activities and terrorism the proper attention of the government is required. The law enforcement agencies should be better equipped in the present information age and should be highly trained to use computer and computer analysis so that they may track the nefarious activities of the terrorists/hoodlums. Corporate sectors particularly banks should play vital part in the fight against terrorisms.

Data mining techniques used in terrorist detection solely depends on the situation at hand. Many cases demand the combinations of two or more techniques used together. For example classification and link analysis techniques can be used to complement each other.

Data mining is not all to counterterrorism because there are various short comings which included the issue of skilled manpower, insufficient investment in IT infrastructure and tell communication.

References

1. Rezafadaei-tehrari,Themas M. Green,(2002)" Crime and society" International Journal of Social Economics Vol 29,MP 10,pp 781-795.
2. *Biswapati Jana,* Survey on Cheating and Prevention Techniques in Visual Cryptography; Mathematical Sciences international Research Journal ISSN 2278 – 8697 Vol 3 Spl Issue (2014), Pg 961-976
3. Usman Fayyad,Gregory piatetsky-shapiro and padhraic smyth(1996)."From Data mining to knowledge discovery in databases".
4. R. Agrawal, A.Ghosh, T. Jmielinski, B.Lyer, & A.Swami(1992). An interval classifier for database mining applications. In proceedings of the 18th conference on very large databases,Morgan KaufinanPubs (Los Altos CA),Vancouver
5. *B. Kotaiah, Dr. Raees Ahmed Khan , D. Siva Rama Krishna,* A Neural Network Methodology for Software Reliability ... ; Mathematical Sciences International Research Journal ISSN 2278 – 8697 Vol 2 Issue 1 (2013), Pg 101-103
6. A.Milley(2000), Healthcare and Data mining, Health Management Technology, 21(8), 44-47.
7. J.Ranjan,"Applications of dta mining techniques in pharma ceutical industry"journal of theortiocal and applied information technology 2005-2007,p61-67
8. Q.A.Aal Radaideh et al.Mining student data using decision trees'The2006 International Arab Conference on information Technologyp1-5.
9. Mrs. Pratiksha P. Bhawalkar,Dr. Mrs. Meenakshi M. Sagdeo, A New Comparison Criterion for Hybrid Rules; Mathematical Sciences International Research Journal ISSN 2278 – 8697 Vol 3 Issue 1 (2014), Pg 39-45
10. M.Ingram (2001) .Internet privacy threatened following terrorist attacks on VS,URL :
11. www.wsws.org/articles/2001/Sep2001/isps24shtm.
12. H.Debar, M. Dacier, A. Wespi, (1999) towards taxonomy of intrusion-detection systems, Computer Networks, 31, pp. 805-822.
13. *A.M. Sagir,* A Class of 3 – Step Block Method for Solving Ordinary ; Mathematical Sciences International Research Journal ISSN 2278 – 8697 Vol 2 Issue 2 (2013), Pg 136-140
14. Cate H.fred(2008),"Legal standards for Data mining"rcrived from the internet on 12-03-2011
15. http://www.hunton.com/files/tbl s47 details/filesupload 265/1250/cate fourth A mindment.pdf
16. Seifert Jeffery w.(2004),or Data mining report,CRS report for congress", order code RL31789.
17. The tuhutramsinghamgam bhavani (2010)" Data mining for counterterosism" The mitre cooperation, retrived from the interation 18-03-20911.
18. A.P.Dhana Balan, C.Santhi, Rm.Sivagama Sundari, Soft Feebly Separation Spaces and Soft Feebly Continuous; Mathematical Sciences international Research Journal ISSN 2278 – 8697 Vol 3 Issue 2 (2014), Pg 575-577
19. DeRosa mary (2004)." Data mining and data analysis for counter terrorism",CSIS report-2004.
20. R.O. okonkwo and F.O. Enem",Combating crime and terrorism using Data mining techniques"Nigeria computer Society:10th international conference-july 25-29,2011.
21. Y.Elovici et al"using Data mining techniques for detecting terror related activities on the web.
22. M.Gary Wiss and D.Brain Dausison (2010)."Data mining", retrieved from the internet on 18-03-2011.

* * *

[1]Research Scholar , Sunrise University, Alwar (Raj.)
pandey.satishchandra@gmail.com
[2]Reader , Deptt of Mathematics, RR College, Alwar (Raj.)
[3]Sr.Lecturer, Deptt of Computer Application's. PG. College, Basti (UP)