Journal of Defense Management
ISSN: 2167-0374

Pilley and Sikchi, J Def Manag 2014, 4:1
DOI: 10.4172/2167-0374.1000115

Research Article Open Access

Review of Group Prediction Model for Counter Terrorism Using CLOPE Algorithm

Pawan H Pilley* and Sikchi SS

Amravati University, St Gadge Baba Amravati University, SRPF Colony, Amravati, Maharashtra-444602, India

Abstract

In the present scenario terrorist attacks are biggest problem for the mankind and whole world is under constant threat from these well-planned, sophisticated and coordinated terrorist operations. Now every country is focusing on counter-terrorism. Counter-terrorism is the practices, tactics, strategies and techniques that governments, militaries, police and security agency uses to prevent or in response to terrorist threats. This paper will focus on terrorist group prediction techniques for counter terrorism. Also analysis of predicting the responsible group using CLOPE algorithm.

Keywords: Terrorist group; Predicting terrorist attack; CLOPE algorithm; Counter–terrorism

Introduction

Prediction of terrorist group using historical data of attacks has been less explored due to the lack of detailed terrorist data which contain terrorist group's attacks and activities. The reasons may be its confidentiality & sensitivity. Intelligence agencies are having large amount of data. They are continuously monitoring terrorist activities. But they are not having enough trained officers to process bulk data in very less time period for the purpose of decision-making about terrorist attacks. In counter terrorism first step after any incident/attack is to find the group names that were involved and to make strategy to catch them.

Data Mining for Counter Terrorism

Security is an important aspect that has been given top priority by all political and government worldwide and are aiming to reduce crime incidence [1,2]. Intelligence analysis might be applied to any of several recognized intelligence sources like Signals Intelligence (SIGINT), Imagery Intelligence (IMINT), and Open Source Intelligence (OSINT) [3].

Current terrorism informatics, which aims to help security officials using data mining techniques, is mainly focused on using social network analysis (SNA) for structural and positional analysis of terrorist networks where required information is provided from non-crime data [4-7]. Prediction of terrorist group using historical data of an attack has very less work; this is because of lack of detailed terrorist data which contain terrorist group's attacks and activities [8]. The use of data mining technologies in counter terrorism has been flourishing since the U.S. Government encouraged the use of information technologies [9]. This paper, focus on developing a prediction model using historical data to predict the terrorist group involved in a given attack. Our database includes terrorist attacks in different countries from year 1970 to 2011.

Group Prediction Technique

Prediction of terrorist group after an attack is one of the most important steps for counter terrorism. As soon as we are able to find the involved group name, we will be able to make strategies to catch the culprits. Generally, the terrorist group responsible for an attack is detected by using email, telephone signal information, terrorist web sites, social network analysis etc., Terrorist activities occurred in past are available in criminal data/historical database. This database can be used to detect terrorist group responsible for an attack. Frequency of terrorist attacks can also be analyzed.

Terrorist Group Prediction Model (TGPM)

Terrorist group prediction model (TGPM) which learns the pattern of terrorist attacks from the available historical data and make an association between terrorist group and previous attacks. Every terrorist group can be differentiated based on the style of attack, targets like police, private organizations; public property etc. so by analyzing these patterns TGPM will predict the group that may be involved in a given incident.

TGPM is developed to detect the responsible terrorist group by using historical data. TGPM uses the concept of Crime Prediction Model, Group Detection Model (GDM) and Offender Group Detection Model (OGDM) [10,11]. TGPM uses various parameters like attack type, location, target type, weapon type, hostage/kidnapping and suicide attack etc. TGPM uses terrorist corpus, parameter's value and parameters weight as input.

GDM (Group Detection Model)

GDM is a general detection model based on co-offending clustering. It works by linking co-offenders with inner join query using unique crime reference id numbers. These links collected together to create criminal networks as graphs, then Strongly Connected Components (SCC) algorithm [12] is applied so that each connected component can be treated as an individual criminal network. SCC algorithm is a way of dividing big graphs into smaller chunks of sub graphs or components. A directed graph is called strongly connected if for every pair of vertices there are paths towards each of them. SCC components of a directed graph are its maximal strongly connected sub graphs. In GDM, each component represents a unique criminal network and one criminal can only belong to a single criminal network. GDM simply finds and links all criminals each other who commit the same crime with the same friends.

***Corresponding author:** Pawan H Pilley, Amravati University, St Gadge Baba Amravati University, SRPF Colony, Amravati, Maharashtra 444602 , India, Tel: 7709598957, E-mail: pawanhpilley@gmail.com

Received December 27, 2013; **Accepted** January 28, 2014; **Published** February 13, 2014

Citation: Pilley PH, Sikchi SS (2014) Review of Group Prediction Model for Counter Terrorism Using CLOPE Algorithm. J Def Manag 4: 115. doi:10.4172/2167-0374.1000115

OGDM (Offender Group Detection Model)

Just like GDM uses co-offending feature, OGDM uses similarity of three crime features which are criminals' choice of crime location, date, and modus operandi by assuming that similarly behaving criminals would inevitably meet or know each other since they are likely minded. OGDM steps include creation of Spatial, Temporal, and Modus Operandi links, followed by clustering for detecting criminal networks. OGDM is similar to GDM for creating links although GDM counts links between criminals as link weights, whereas OGDM calculates link weights counted and normalized against all links.

Dataset and Collection Methodology

The Global Terrorism Database (GTD) is an open-source database including information on terrorist events around the world from 1970 through 2011 (with additional annual updates planned for the future). Unlike many other event databases, the GTD includes systematic data on domestic as well as transnational and international terrorist incidents that have occurred during this time period and now includes more than 104,000 cases. For each GTD incident, information is available on the date and location of the incident, the weapons used and nature of the target, the number of casualties, and--when identifiable--the group or individual responsible. Statistical information contained in the Global Terrorism Database is based on reports from a variety of open media sources.

Characteristics of the GTD

- Contains information on over 104,000 terrorist attacks.
- Currently the most comprehensive unclassified data base on terrorist events in the world.
- Includes information on more than 47,000 bombings, 14,000 assassinations, and 5,300 kidnappings since 1970. Includes information on at least 45 variables for each case, with more recent incidents including information on more than 120 variables.
- Over 3,500,000 news articles and 25,000 news sources were reviewed to collect incident data from 1998 to 2011 alone.

The Global Terrorism Database (GTD) was developed to be a comprehensive, methodologically robust set of longitudinal data on incidents of domestic and international terrorism. Its primary purpose is to enable researchers and analysts to increase understanding of the phenomenon of terrorism. The GTD is specifically designed to be amenable to the latest quantitative analytic techniques used in the social and computational sciences.

Definition of Terrorism

The collectors of the database aimed to record every known terrorist event within and across countries and over time, as identified in multi-lingual news sources, for the purpose of performing risk analysis for U.S. businesses. Incidents were collected according to the following definition of terrorism:

- "The threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation".
- It is well-recognized that divergent definitions of terrorism abound and that the nature and causes of terrorism are hotly contested by both governments and scholars.
- "The violent act was aimed at attaining a political, economic, religious, or social goal";
- "The violent act included evidence of an intention to coerce, intimidate, or convey some other message to a larger audience (or audiences) other than the immediate victims;" and
- "The violent act was outside the precepts of International Humanitarian Law".

These criteria-which continue to be employed by data collectors in post-2007 collection efforts--were constructed to allow analysts and scholars flexibility in applying various definitions of terrorism to meet different operational needs.

The criteria for incident inclusion and the coding scheme used in GTD were developed by a START Advisory Board, which consisted of recognized experts in terrorism and data collection (Figure 1).

Data preprocessing is an important step in which missing values are filled up, redundancies are removed and filtering is performed so that the database will be ready for use. Missing values can be filled by using various terrorist databases available over internet for terrorism research [13-15]. After pre-processing of database, percentage of attacks of each group is calculated based on input parameters. Each parameter is assigned a weight based on its impact over the incident. The group weight is calculated by using the percentage of attacks of each group and the parameters weight. Different clusters are created. Association between these clusters is being performed and highest value from these associations is obtained. Group name corresponding to the highest value may be the most probable responsible terrorist group.

Selection of the Clustering Algorithm

Most clustering algorithms define some criterion functions and optimize them, maximizing the intra-cluster similarity and the inter-cluster dissimilarity. The criterion function can be defined locallyor globally. CLOPE is specific to categorical attributes, which forms a large part of our database. At each level of the hierarchy; it groups instances within clusters, again with respect to a distance function. Finally, CLOPE calculates the extent to which categorical attributes match with each other, for two given instances. If the level of this matching is greater than a certain threshold, the instances are clustered together. Considering limitations, CLOPE was the only algorithm which clustered data with missing values. Also, it doesn't require much tuning to efficiently obtain good clusters. Finally, it's designed specifically for categorical attributes, and its performance is comparable to other algorithms for categorical clustering. Therefore, we selected CLOPE for our experiments.

Implementation of CLOPE Algorithm

Like most partition-based clustering approaches, we approximate the best solution by iterative scanning of the database. Our implementation requires a first scan of the database to build the initial clustering. After that, a few more scans are required to refine the clustering and optimize the criterion function. If no changes to the clustering are made in a previous scan, the algorithm will stop, with the final clustering as the output. The output is simply an integer label for every transaction, indicating the cluster id that the transaction belongs to.For categorical attributes, it can be expected that those instances can be clustered which contain matching values for one or more attributes. CLOPE attempts to increase the intra-cluster overlapping of categorical values by increasing the height-to-width ratio of the cluster histogram. It also uses a parameter called repulsion, to control the tightness of the

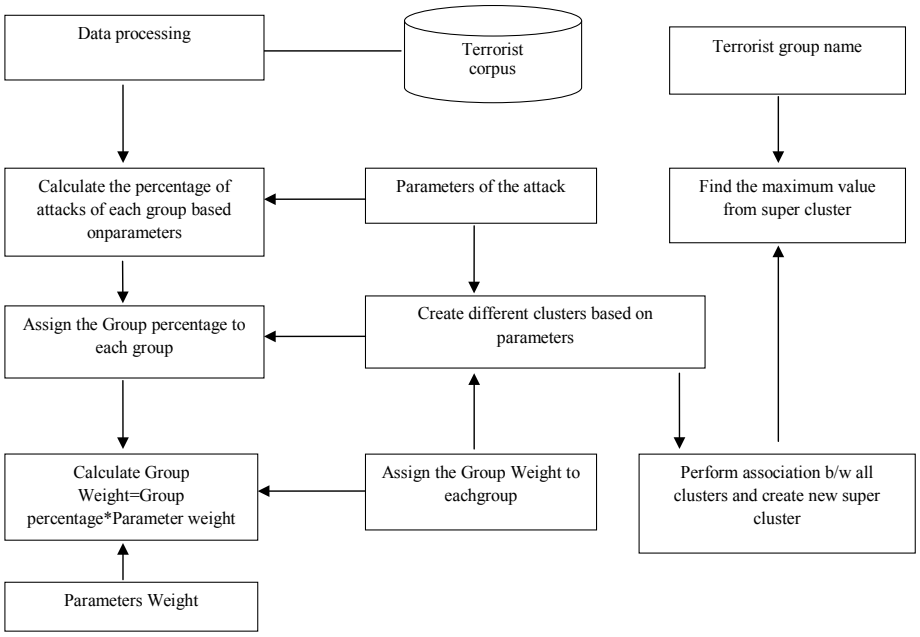

Figure 1: Terrorist Group Prediction Model (TGPM).

cluster. Different number of clusters can be obtained by varying this parameter.

Conclusion

In this study, it can be concluded that by using historical terrorist data it is possible to predict the group involved in the given attack. However, further study can be carried out to detect a terrorist group using historical data. For further research point of view it is suggested to use different artificial intelligence techniques of group detection and include more parameters like phone calls, email data etc. so that more accurate results can be obtained.

References

1. David G (2006) Globalization and International Security: Have the Rules of the Game Changed? In Annual meeting of the International Studies Association, California, USA.
2. Malathi A, Santhosh BS (2011) Evolving Data Mining Algorithms on the Prevailing Crime Trend–An Intelligent Crime Prediction Model.International Journal of Scientific & Engineering Research.
3. Chen H, Denning D (2011) The Dark Web Forum Portal: From multi-lingual to video. Intelligence and Security Informatics (ISI), IEEE conference.
4. Coffman TR, Marcus SE (2004) Pattern Classification in Social Network Analysis: A case study. In 2004 IEEE Aerospace Conference, Austin, USA.
5. Nooy WD, Mrvar A (2205) Exploratory Social Network Analysis with Pajek. Cambridge University Press, New York, USA.
6. Scott J, Peter J (2005) Social Network Analysis. SAGE Publications, London.
7. Wasserman S, Faust K (1994) Social Network Analysis: Methods and Applications.
8. Faith O,Zeki E, Bowerman C (2009) Prediction of Unsolved Terrorist Attacks Using Group Detection Algorithms. In LNCS 5477: 25-30.
9. Taipale KA (2003) Data mining and domestic security: connecting the dots to make sense of data. Columbia Sci Tech Law Rev 5: 1-83.
10. Ozgul F, Bondy J, Aksoy H (2007) Mining for offender group detection and story of a police operation. Sixth Australasian Data Mining Conference .Australian Computer Society Conferences in Research and Practice in Information Technology (CRPIT), Gold Coast, Australia.
11. Ozgul F, Erdem Z, Aksoy H (2008) Comparing Two Models for Terrorist Group Detection: GDM or OGDM. ISI Workshops 2008. LNCS 5075: 149-160.
12. Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) Introduction to Algorithms. Second Edition.
13. Global Terrorism Database (2012).
14. South Asia Terrorism Portal (2012) Incidents and Statements involving CPI-Maoist: 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014.Left-wing Extremist group.
15. http://cri-portal.dyndns.org/portal/LoginForm.action