TGPM: Terrorist Group Prediction Model for Counter Terrorism

Abhishek Sachan
Computer Science
Maulana Azad National Institute of Technology
Bhopal, India

Devshri Roy
Computer Science
Maulana Azad National Institute of Technology
Bhopal, India

ABSTRACT
Prediction of terrorist group using historical data of attackshas been less explored due to the lack of detailed terrorist data which contain terrorist group'sattacks and activities. The reasons may be its confidentiality & sensitivity. In this paper we have shown a terrorist group prediction model (TGPM) to predict the terrorist group involved in a given attack. This model initially learns similarities of terrorist incidentsfromvarious terrorist attacks to predict the responsible group. The model has been validated with the experimental results. The overall performance of the model shows a fair degree of the accuracy.

Keywords
Terrorist groups; matching and predicting terrorist attacks; terrorism; counter-terrorism; group detection.

1. INTRODUCTION
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. 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. This paper discusses a model for the identification of responsible groups for an attack based on the available information.

Security is an important aspect that has been given toppriority by all political and government worldwide andare aiming to reduce crime incidence [1, 2].Intelligence analysismight be applied to any of several recognizedintelligence sources like Signals Intelligence (SIGINT), Imagery Intelligence(IMINT), and OpenSource 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 [4, 5, 6, 7] where required information is provided from non-crime data. 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]. In this paper, we have developed a perdition model using historical data to predict the terrorist group involved in a given attack. Our database includes terrorist attacks in India from year 1998 to 2008.

2. PREDICTING TERRORIST GROUP
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 willbe 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 [4, 8]. This database can be used to detect terrorist group responsible for an attack.We have developed a 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.

3. TERRORIST GROUP PREDICTIONMODEL (TGPM)
TGPM is developed to detect the responsible terrorist group by using historical data. TGPM uses the concept of Crime Prediction Model [2, 8], Group Detection Model (GDM) [10] 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.

Data preprocessing is an important step in which missing values are filled up, redundancies are removed and filtering is performedso that the database will be ready for use. We have filled the missing values by using various terrorist databases available over internet for terrorism research [12, 13, 14]. After pre-processing of database, percentageof attacks ofeach group is calculated based on input parameters.Each parameter is assigned a weight based on its impact over the incident. The group weight iscalculated 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.

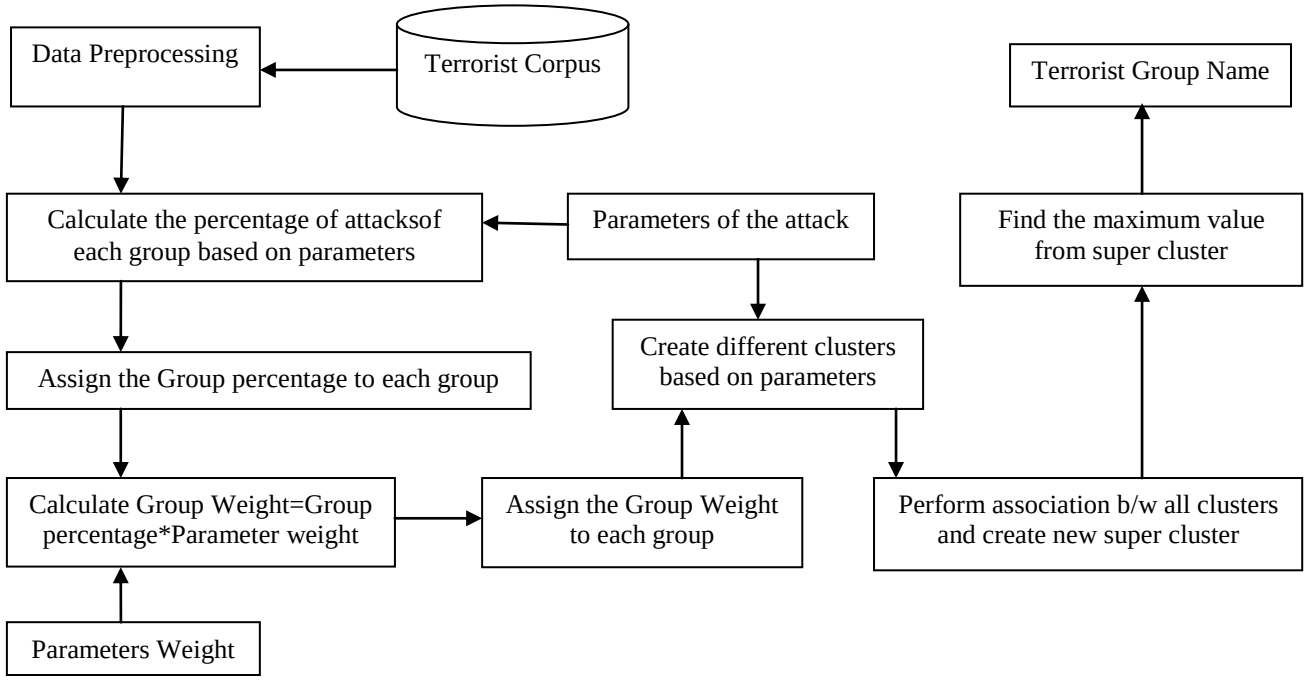

Figure 1. Terrorist Group Prediction Model (TGPM)

We have used the following mathematical equations to perform various major steps.

The percentage (P) of the total attacks in which each group involved is estimated using Eq. (1).
Here,
G = No. of time a group involved in attacks.
n = no. of unique groups retrieved from dataset.

$$P_j = \frac{G_j}{\sum_{i=1}^{n} G_i} \quad \text{where } j=1, 2, 3\ldots n \qquad (1)$$

Equation (2)is used to estimate the percentage of the total attacks in which each group involved based on parameters.
Here value of n will depends on the retrieved records from dataset based on parameters.
m = no. of parameters.

$$P_{kj} = \frac{G_{kj}}{\sum_{i=1}^{n} G_{ki}} \quad \text{where } j=1, 2, 3\ldots n \text{ and } k=1, 2, 3\ldots m \qquad (2)$$

Equation (3) is used to calculate the group weight based on each group probability and parameters weight.
Here, GW = Group Weight.
β= Parameter Weight.

$$GW_{kj} = P_{kj} * \beta_k \quad \text{where } k=1, 2, 3\ldots m \qquad (3)$$

Equation (4) will associate the different clusters values based on parameters.
Here CL is a super cluster created after association.
The value of n will be the total no. of groups in all clusters.

$$C_L = \sum_{k=1}^{m} GW_{kj} \quad \text{where } L=1, 2,3\ldots n \qquad (4)$$

Equation (5) will determine the highest value from the super cluster that will give the group name as result.

$$R = \text{Max}\{C_L\} \qquad (5)$$

4. OPTIMAL PARAMETERS WEIGHT ESTIMATION

Evaluation of TGPM is performed on standard Global Terrorism Database (GTD) [12].Data of the GTD needs to be processed because some of the values of different fields of the database were missing. Database size is reduced according to the requirement. We have used terrorist incident data from 1998 to 2008.Data related to the Naxalites, Maoist Communist Center (MCC), Maoists, People's War Group (PWG), Communist Party of India- Marxist-Leninistare removed as these group got merged to form a new entity, the Communist Party of India-Maoist (CPI-Maoist) on September 21, 2004 [13, 15]. So there parameters will not be identical. We have used top 11 groups involved in 590 incidents to evaluate our system and worked on the concept of 3:1 ratio of training and testing. It means that we have used 447 incidents as training data and 143 incidents as testing data.

The different parameters used in our model are attack type, location, target type, weapon type, hostage/kidnapping and suicide attack.When all the parameters are given equal weightage, the overall performance of the system found to be 65.73%. To improve the performance, different parameters should have different weights. To determine the optimal values of the weight for different parameters, several experiments have been conducted. The graphs are plotted by varying the weight of only one parameter and keeping all others parameters weight equal to 1.The graphs in fig.2, fig.3, fig.4, fig.5, fig.6 and fig.7areshowing the effect of parameters weight over the result of system. The performance of the system varies with the parameters weight. The overall performance 79.72% of the system is achieved when location parameter weight is more than 2.Although the overall performance of the system increases but the performance of the system to identify small groups decreases.

In this experiment we also changed the various parameters weight so that the performance of the system can be evaluated with different weight conditions. It is also showing that when we are increasing the parameter weight of different parameters of system except location parameter weight, the performance will decrease.This is due to the fact that the percentage weight of location reduces when other parametersweight increases. Based on the findings, we can say that the location isa major factor when we are discussing about the terrorist groups working in a country,because each group is having its own workingregion. We have assigned the parameters weight values ranging from 0 to 3.We identified that, when parameters weight becomes greater than 2.5, the performances of the system varies negligibly.

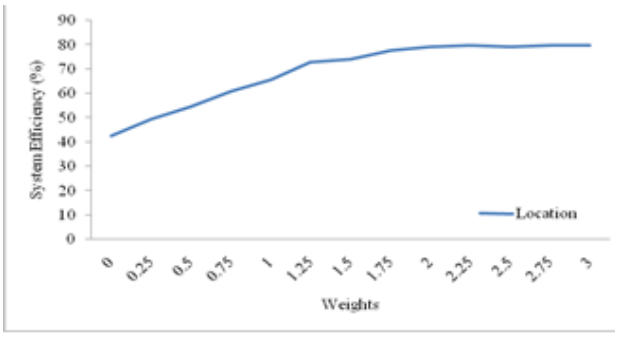

Figure 2. Location parameter weight graph

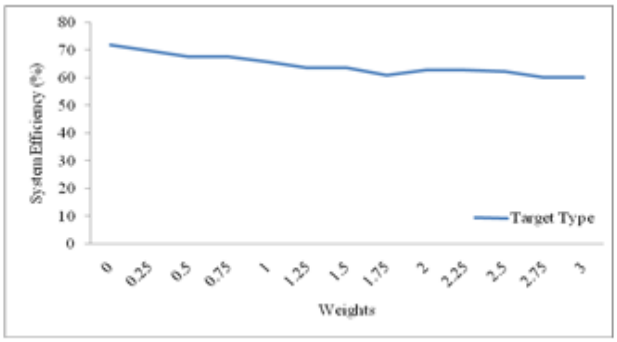

Figure 3. Target-Type parameter weight graph

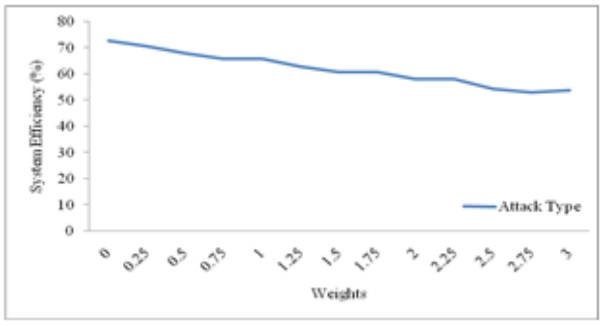

Figure 4. Attack-Type parameter weight graph

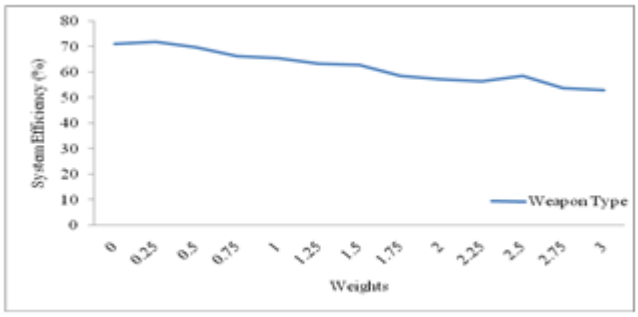

Figure 5. Weapon-Type parameter weight graph

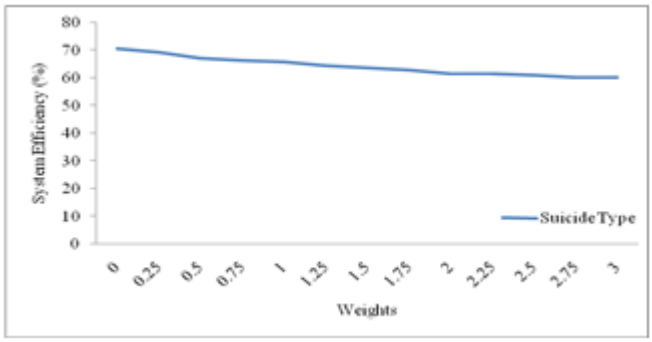

Figure 6. Suicide-Type parameter weight graph

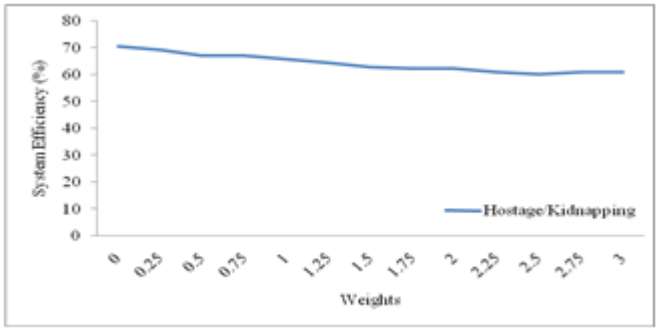

Figure 7. Hostage Kidnapping parameter weight graph

We can conclude that, except location weight all other parameters weights are showing almost the same effects over the performance of the system.

5. EVALUATION AND RESULTS

The investigation was done to estimate the effect of parameters and the optimal parameter weights (β) thus obtained were used to calculate the system performance. After estimation of optimal parameters weight and analyzing them, we came to the conclusion that the location weight be 2.25, target type weight be 0.50, attack type weight be 0.50, weapon type weight be 0.25 , suicidesweight be 0.25, hostage/kidnapping weight be 0.25.By using these weights our system performance is found to be 80.41%. This model works well for big groups but some time it fails to identify small groups or sub groups which are working with a big group(because subgroup's parameters will be similar to that of big group and hence they cannot be differentiated). One more problem also associated with the system is that it fails to detect terrorist groups which are using different pattern each time or those which are new or unheard. The performance of the system is dependent on the input parameters. More the parameters, more accurate will be the

results.The performance of the system is also associated with the availability of proper data and parameters for training the system so that specific small groups can also be detected.

6. 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.

7. REFERENCES

[1] David, G., "Globalization and International Security: Have the Rules of the Game Changed?", In Annual meeting of the International Studies Association, California, USA, http://www.allacademic.com/meta/p98627_index.html, 2006.

[2] A. Malathi and Dr. S. Santhosh Baboo, "Evolving Data Mining Algorithms on the Prevailing Crime Trend – An Intelligent CrimePrediction Model", In International Journal of Scientific & Engineering Research, June 2011, Vol. 2, Issue 6.

[3] H. Chen, D. Denning et al., "The Dark Web Forum Portal: From multi-lingual to video", InIntelligence and Security Informatics (ISI), IEEE conference, 2011.

[4] Coffman, T.R., Marcus, S.E.: Pattern Classification in Social Network Analysis: A case study. In: 2004 IEEE Aerospace Conference, March 6-13, 2004.

[5] Nooy, W.d., Mrvar, A., et al.: Exploratory Social Network Analysis with Pajek. Cambridge University Press, New York, 2005.

[6] Scott, J.: Social Network Analysis. SAGE Publications, London, 2005.

[7] Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications, 1994, pp. 266.

[8] Faith Ozgul,Zeki Erdem and Chris Bowerman, "Prediction of Unsolved Terrorist Attacks Using Group Detection Algorithms," In LNCS, vol. 5477, pp. 25-30. Springer, Heidelberg, 2009.

[9] Taipale KA, "Data mining and domestic security: connecting the dots to make sense of data", In Columbia Sci Tech Law Rev 5, 2003, pp. 1–83.

[10] Ozgul, F., Bondy, J., Aksoy, H.: Mining for offender group detection and story of a police operation. In: Sixth Australasian Data Mining Conference (AusDM 2007). Australian Computer Society Conferences in Research and Practice in Information Technology (CRPIT), Gold Coast, Australia , 2007.

[11] Ozgul, F., Erdem, Z., Aksoy, H.: Comparing Two Models for Terrorist Group Detection: GDM or OGDM? In: Yang, C.C., Chen, H., Chau, M., Chang, K., Lang, S.-D., Chen, P.S., Hsieh, R., Zeng, D., Wang, F.-Y., Carley, K.M., Mao, W., Zhan, J. (eds.) ISI Workshops 2008. LNCS, vol. 5075, pp. 149–160. Springer, Heidelberg, 2008.

[12] Global Terrorism Database, http://www.start.umd.edu/gtd/, Retreived on 05/02/2012.

[13] South asia terrorism portal , http://www.satp.org/satporgtp/countries/india/terroristoutfits/CPI_M.htm, Retreived on 05/02/2012.

[14] Dark web portal, "http://cri-portal.dyndns.org/portal/Home.action", Retrieved on 19/02/2012.

[15] Council on Foreign Relations, http://www.cfr.org/india/terror-groups-india/p12773, Retreived on 05/02/2012.