An Outlier-Based Intention Detection for Discovering Terrorist Strategies

Article *in* Procedia Computer Science · October 2017

DOI: 10.1016/j.procs.2017.09.006

Salih Tutun
Binghamton University
27 PUBLICATIONS 46 CITATIONS

Ömer Biyikli
Erciyes Üniversitesi
4 PUBLICATIONS 1 CITATION

Mohammad T. Khasawneh
Binghamton University
155 PUBLICATIONS 637 CITATIONS

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 114 (2017) 132–138

Procedia Computer Science

www.elsevier.com/locate/procedia

Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems, CAS
October 30 – November 1, 2017, Chicago, Illinois, USA

An Outlier-Based Intention Detection for Discovering Terrorist Strategies

Salih Tutun[a,*], Murat Akça[b], Ömer Bıyıklı[b], Mohammad T. Khasawneh[a]

[a]*Department of Systems Science and Industrial Engineering, Binghamton University, New York, 13850, USA*
[b]*Department of Industrial Engineering, Gazi University, Ankara, 06420, Turkey*

Abstract

Terrorist groups (attackers) always strive to outmaneuver counter-terrorism agencies with different tactics and strategies for making successful attacks. Therefore, understanding unexpected attacks (outliers) is becoming more and more important. Studying such attacks will help identify the strategies from past events that will be most dangerous when counter-terrorism agencies are not ready for protection interventions. In this paper, we propose a new approach that defines terrorism outliers in the current location by using non-similarities among attacks to identify unexpected interactions. The approach is used to determine possible outliers in future attacks by analyzing the relationships among past events. In this approach, we calculate the relationship between selected features based on a proposed similarity measure that uses both categorical and numerical features of terrorism activities. Therefore, extracting relations are used to build the terrorism network for finding outliers. Experimental results showed that the comparison of actual events and the detected patterns match with more than 90% accuracy for many future strategies. Based on the properties of the outliers, counter-terrorism agencies can prevent a future bombing attack on strategic locations.

Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.

Keywords: Outlier Detection; Similarity Function; Link Formation; Network Analysis; Counter-terrorism

* Corresponding author.
E-mail address: stutun1@binghamton.edu

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.
10.1016/j.procs.2017.09.006

1. Introduction

Terrorism is a new kind of war that is increasingly characterized with uncertainty. In this war, terrorist groups (attackers) often change their strategies in an effort to surprise and shock defenders (counter-terrorism agencies) for more successful attacks. Defenders are always under pressure to learn new strategies in order to have a strong counter-terrorism strategy [1]. Moreover, terrorism has significantly increased after the September 11 attack because the uncertainties associated with such events make their prevention a very complex effort to manage [2]. Defenders need to know how to create strategies to prevent this kind of attacks, and they need to adopt more accurate approaches to investigate terrorist activities [3]. Intelligence gathering is the cornerstone through which uncertainty is reduced.

Current literature suggests that terrorism has an evolutionary nature and attackers change their behavior according to defenders' counter-terrorism policies. The behavior of attackers evolves over time, and they often copy the behavior of other attacks [4]. For instance, each attacker learns tactics from past attacks whether they were successful or not. After learning certain tactics, they seek to shock defenders through attacks that are unexpected when compared with past events. Only when defenders have the ability to predict unexpected future events is the prevention of terrorism plausible.

In the literature for understanding strategies of terrorism, network-based approaches are used to understand complex interactions [5, 6]. These approaches are becoming increasingly popular [7] because they are proving to be effective methods for understanding terrorism [8]. Moreover, many researchers have studied the behavior of people (attackers) to find the leader of attackers (with their leader). Therefore, existing network-based approaches in the literature focused on prosecution instead of prevention [9, 10]. In this research, we focus on the finding relationships between different attacks instead of connections and relationships between people [11].

This research aims to propose a new approach by analyzing relations of attacks to develop predictive capabilities. The network of attacks is modeled in the approach to understand future strategies. More specifically, a new outlier-based similarity function is proposed to find relations that will help construct a network for events. Furthermore, this similarity function is used to estimate relationships among interactive events by using non-similar attacks [11, 12]. This method extracts attacker interaction from network properties to obtain a better understanding of the attacker activity. The results could potentially help in the understanding of future attacks and enable counter-terrorism agencies to propose proactive strategies [11, 13].

The remainder of the paper is organized as follows. In Section 2, data analysis and collection are explained, and the methods used in the new approach are presented. A detailed description of how the proposed approach is used to understand complex interactions is also presented. In Section 3, experimental results that show the proposed approach works to understand attacker activities efficiently are presented. Finally, Section 4 presents a discussion to highlight the improvement in modeling terrorism and the contribution of the research.

2. Materials and Methods

Terrorist attacks listed in the Global Terrorism Database (GTD) are used in this research. The data includes various events between 1970 and 2015 [2]. The data is prepared by removing missing values and incorrect events. The following section provides details of the proposed approach. Moreover, bombing (with explosives weapons) attacks, as seen in Fig. 1 and Fig. 2, are used against defenders' agencies (e.g., Military, Police, etc.). This type of attack was chosen because they constitute half of all attacks [11].

In the collected dataset, the variable names are explained as follows: Extended incident (extended) is defined as yes (1) if there is an extension for more than 24 hours or no (0). Doubt of terrorism proper (doubtterr) is defined as yes (1) or no (0). Part of multiple incidents (multiple) is determined as yes (1) or no (0). Location of events is defined using countries, regions, state, and city. Vicinity (vicinity) is used as yes (1) if the event happens near the city or no (0) if it is in the city center. Specificity is determined at the geospatial resolution of the latitude and longitude areas with five different categories. Attack type (attackttype1)is defined as a Bombing/Explosion attack. Successful Attack (success) is defined based on whether the event is successful (1) or not (0). Weapon type (weaptype1) is defined as which weapons are used for attacks. Target type (targettype1) is determined by which targets the attackers pursue. The number of killings (nkill) means the number of people killed in the attack. Hostage victims (ishostkid) is defined that the victim was taken a hostage or not. International (int-log) means that the attack was international or domestic [2]. As seen in Fig. 1, half of all attacks based on the dataset are of the bombing and explosion category. Moreover, 40% of all attacks used explosive weapons (as seen in Fig. 2). Therefore, this paper focuses on these attacks to implement the proposed approach.

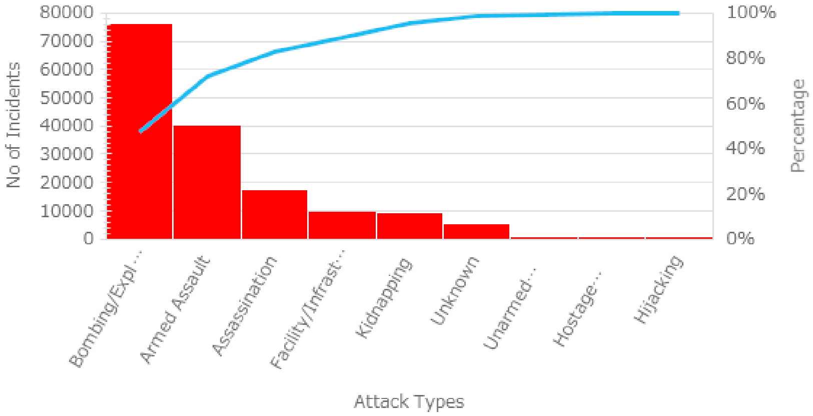


Fig. 1. Attack types for collected data.

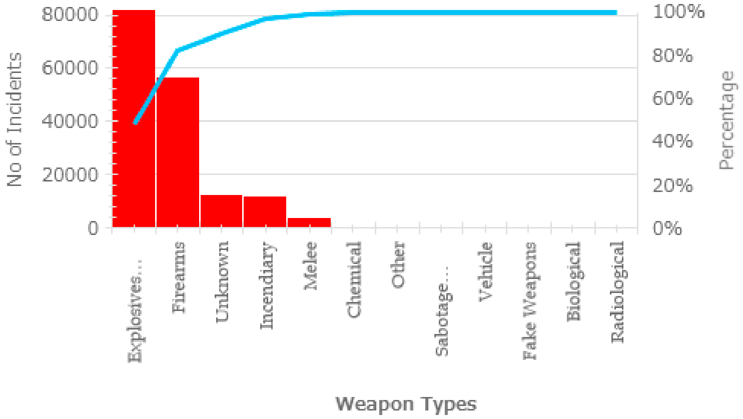


Fig. 2. Weapon types for collected data.

As seen in Fig. 3, there are interactions among attacks. The similarity function can be proposed to capture these complex interactions. This research explores the opportunities for the application of network analytic techniques to make precautions before attacks. Similarity function can be used to measure non-similarities (links) to form relations between nodes. However, computing categorical features is not straightforward because there is no explicit ordering among categorical variables. A new data-driven heterogeneous similarity function is proposed to solve this problem.

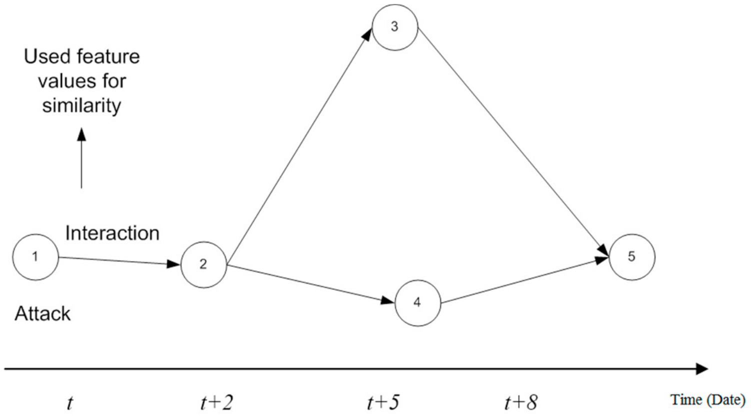


Fig. 3. Interactions and learning among terrorist attacks.

Table 1.Explanation of the data with formulas.

Attributes ID	A_1	A_2	A_3	…	A_d
n_1	x_{11}	x_{12}	x_{13}	…	x_{1d}
n_2	x_{21}	x_{22}	x_{23}	…	x_{2d}
n_3	x_{31}	x_{32}	x_{33}	…	x_{3d}
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
n_N	x_{N1}	x_{N2}	x_{N3}	…	X_{nd}
Frequency	$f_1(x)$	$f_2(x)$	$f_3(x)$	…	$f_n(x)$

For an overlap measure between categorical data, we define the notations (as seen in Table 1) as categorical dataset D that contains N objects. This dataset has d categorical features and continuous features where A_h denotes the h_{th} feature. Let the feature A_h take n_h values in the dataset D.

$$P_h(x) = \frac{f_h(x)}{N} \qquad \forall h = (1,2,3,4,...,d) \tag{1}$$

The following notations are used. The frequency of values is defined as the number of times that feature A_h taking the value x in the Ddataset (Note: if x not in A_h, $f_h(x) = 0$), and $P_h(x)$ (as seen in Eq. (1)). The sample probability of feature A_h takes value x in D dataset, as seen in above matrix [12]. The similarity value between X and Y (see Eq. (2)) that belongs to the dataset D is calculated as follows:

$$c_h = \begin{Bmatrix} P_h(x) & if\ X_h = Y_h\ as\,categorical\ features \\ 0 & otherwise \end{Bmatrix}$$

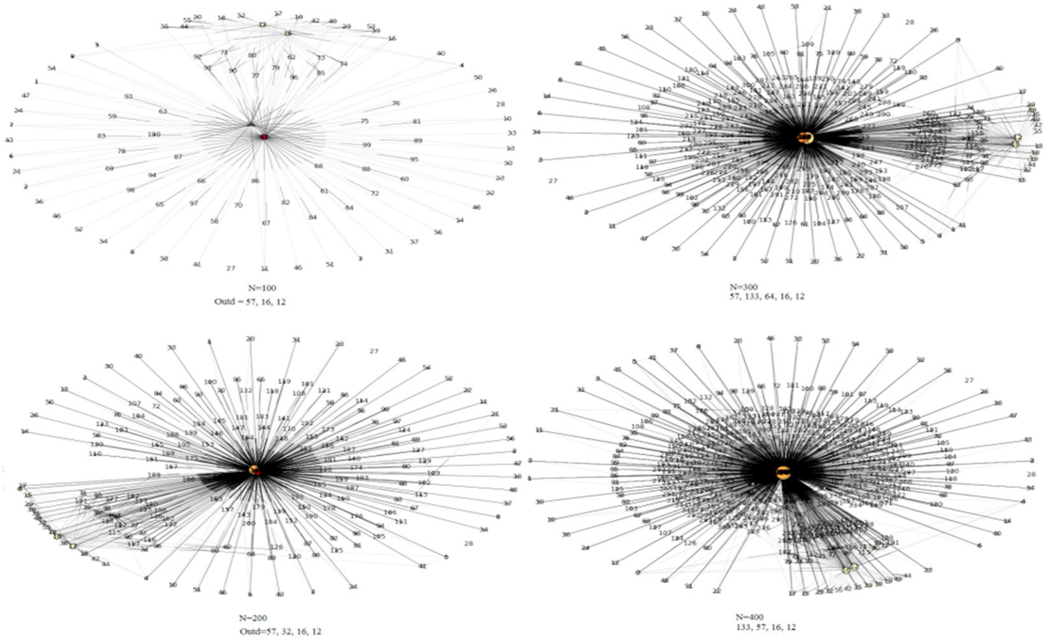
$$n_h = \begin{Bmatrix} (X_h / Y_h) & if\ X_h < Y_h\ as\,continuous\ features \\ (Y_h / X_h) & otherwise \end{Bmatrix}$$

$$NS(X,Y) = 1 - \sum_{h=1}^{d}\left(\sqrt{(c_h)^2}\ or\ (n_h)\right) \tag{2}$$

where *NS(X,Y)* is the non-similarity between two events. This value is used to define relations between events in networks.

3. Defining Outliers (Unexpected Events) for Future Threats

In this section, we look at the non-similarity for the events because attackers will always change strategies. As seen in Fig. 4, outliers are changing dynamically based on the past attacks. In Fig. 4, some events are not in the center because they are not similar to others. At the same time, we observed these events dynamically to understand unexpected behaviors. As a result, Events 57, 16, 12 are the most unused events for first 100 events. After that, Event 52 is added when 200 events are examined. Events 133, 64 are found for the next 100 events. As a conclusion, Events 57, 52, 16, 12, 133 are outliers for future behaviors.



N=100
Outd = 57, 16, 12

N=300
57, 133, 64, 16, 12

N=200
Outd=57, 32, 16, 12

N=400
133, 57, 16, 12

Fig. 4. Defining outliers for future attacks.

As seen in Table 2, we calculated the similarity of outlier attacks for the next 100 attacks with high accuracy. Therefore, we can understand outlier behaviors for future attacks. As a result, based on the non-similar relations, we can find outlier behaviors for future attacks. When defenders focus on these actions, they can understand which behaviors are unused and have a high potential for occurrence in the future.

Table 2.Accuracies (%) of occurrence for future attacks.

extended	country	Region	specificity	vicinity	crit3
0.97%	0.92%	0.95%	0.97%	0.93%	0.98%
doubtterr	multiple	attacktype1	targtype1	guncertain1	
0.74%	0.86%	0.96%	0.62%	0.94%	
weaptype1	property	Ishostkid	int-log	nkill	success
0.98%	0.98%	0.99%	0.99%	0.88%	0.97%

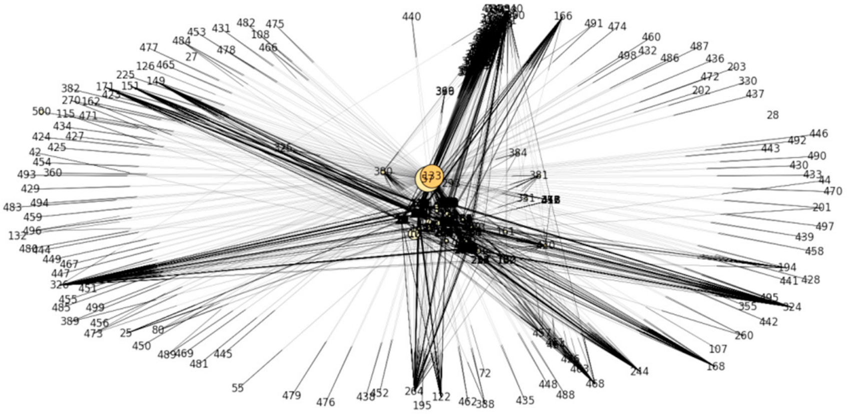


Fig. 5. Defining the most important outliers for future attacks.

Furthermore, some events when followed all past attacks, are found as the most non-similar for using future events. As seen in Fig.5, Event 57 and Event 133 have successful strategies in the past attacks. Once attackers used these strategies, they will shock defenders with successful attacks. In order to control attackers, defenders need to analyze these events deeply. They also continue to search other events dynamically. In this way, attackers can be controlled to prevent the most dangerous attacks.

4. Conclusions

Nowadays, counter-terrorism agencies need to develop better defense strategies to combat the attackers’ tactics. This research proposes a new approach based on a similarity function. More specifically, a heterogeneous similarity function is proposed to analyze relationships between interactive events to understand how attackers seek to surprise defenders. At the same time, the proposed network approach is different because it uses attackers (as events) instead of people.

The proposed approach proves its usefulness due to the use of the proposed similarity function. We show that attacks can be prevented by learning from outlier behavior of attacks. The results prove that we can understand outlier behaviors for bombing attacks by finding patterns. The patterns identified with more than 90% accuracy show that the framework can be used to understand future attacks.

In future work, larger dynamic networks could be used to discover the patterns as a big data project for future events. Moreover, people could study a unified approach that applies pattern classification techniques to the proposed network topology to improve detection accuracy. Based on the proposed network, pattern recognition methods could be used to detect terrorism events. Also, conditional probability can be used to understand which event could lead to a future event. At the same time, the framework can be implemented in other application areas if they have interactions among terrorism-related observations for detection.

In conclusion, defenders can deter threats by using this approach. They can understand how terrorism will impact future events, and governments can control attackers' behaviors to reduce the risk of future events. After attacks occur, the defenders can understand differences between attacks. The proposed approach enables policy makers to develop precise global and/or local counter-terrorism strategies. Furthermore, this information can be extremely useful for law enforcement agencies, which allows them to propose timely reactive strategies.

References

[1] Byman, Daniel, and Jeremy Shapiro. (2014). "We Shouldn’t Stop Terrorists from Tweeting.” The Washington Post 9.

[2] National Consortium for the Study of Terrorism and Responses to Terrorism (START). (2015) "Global terrorism database." http://www.start.umd.edu/gtd.

[3] Jackson, Brian A., and David R. Frelinger. (2009) "Understanding why terrorist operations succeed or fail." RAND CORP ARLINGTON VA.

[4] Chenoweth, Erica, and Elizabeth Lowham. (2007) "On classifying terrorism: A potential contribution of cluster analysis for academics and policy-makers." *Defence& Security Analysis* **23(4)**: 345-357.

[5] Chen, Hsinchun. (2011) "Dark web: Exploring and data mining the dark side of the web (Vol. 30)."*Springer Science & Business Media*.

[6] Netzer, Michael, Karl G. Kugler, Laurin AJ Müller, Klaus M. Weinberger, Armin Graber, Christian Baumgartner, and Matthias Dehmer (2012) "A network-based feature selection approach to identify metabolic signatures in disease."*Journal of theoretical biology* **310**: 216-222.

[7] Coffman, Thayne R., and Sherry E. Marcus. (2004) "Dynamic classification of groups through social network analysis and HMMs". *In Aerospace Conference, 2004. Proceedings*. IEEE (Vol. 5, pp. 3197-3205).

[8] Bohannon, John (2009) "Counterterrorism's new tool: 'metanetwork' analysis." http://science.sciencemag.org/content/325/5939/409

[9] Xu, Jennifer J., and Hsinchun Chen. (2005) "CrimeNet explorer: a framework for criminal network knowledge discovery."*ACM Transactions on Information Systems (TOIS)***23(2):** 201-226.

[10] Krebs, Valdis E. (2002) "Mapping networks of terrorist cells."*Connections* **24 (3)**:43-52.

[11] Tutun, Salih, Mohammad T. Khasawneh, and Jun Zhuang. (2017) "New framework that uses patterns and relations to understand terrorist behaviors."*Expert Systems with Applications* **78**: 358-375.

[12] Tutun, Salih, Sina Khanmohammadi, and Chun-an Chou. (2016) "A network-based approach for understanding suicide attack behavior", in Industrial & Systems Engineering Research Conference (ISERC). Institute of Industrial Engineers (IIE).

[13] Li, Ben-xian, Jun-fang Zhu, and Shun-guo Wang. (2015) "Networks model of the East Turkistan terrorism." *Physica A: Statistical Mechanics and its Applications* **419**: 479-486.