Social Networks 72 (2023) 22–34

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Assessing harmfulness and vulnerability in global bipartite networks of terrorist-target relationships

Alessandro Spelta [a,*], Nicolò Pecora [b], Paolo Pagnottoni [a]

[a] Department of Economics and Management, University of Pavia, Via San Felice 5, 27100 Pavia, Italy
[b] Department of Economics and Social Sciences, Catholic University, Via Emilia Parmense 84, 29122 Piacenza, Italy

ARTICLE INFO

Keywords:
Networks
Centrality measures
Null models
Bipartite networks
Terrorism network

ABSTRACT

In this paper we leverage a massive attack-focused terrorist database and we design a dynamic bipartite network analysis to examine the structural evolution of terrorists-targets relationships. We introduce two novel measures to jointly assess harmfulness and vulnerability of terrorist and target groups, both at local and global level. Statistical validation using null models provides evidence that the information contained in these measures is new and not included in other variables, thus emphasizing the usefulness of these topological indicators. Finally, a policy-support experiment designed as network dismantling-like drill is proposed to assess the effects of potential attack preventive strategies.

1. Introduction

Global terrorism is a serious source of concern for national governments, supranational organizations and law enforcement authorities all over the world. To illustrate, deaths from terrorism reached a recent peak in 2014 and then fell for the fifth consecutive year in 2019 to 13,826 deaths. The global economic impact of terrorism in 2019 amounts to US$26.4 billion (see GTI, 2020). Ever since the September 11, 2001 attacks, researchers from a wide variety of fields have manifested a strong interest in terrorism. In the last decades, a burgeoning literature investigating the structure and the relationships among terrorist groups has been flourishing (see Clauset et al., 2007; Sharma et al., 2017; Li et al., 2019; McMillan et al., 2020). Several other papers have instead focused on the main drivers of transnational terrorism, on its economic impact and consequences for the development, and on the effectiveness of counterterrorism policies (see Blomberg et al., 2004; Sandler and Enders, 2004; Abadie, 2006; Abadie and Gardeazabal, 2008; Khusrav et al., 2013; Goel et al., 2017; Brodeur, 2018).

Moreover, statistical modelling of global terrorism has received a considerable attention over recent times (see Kyung et al., 2011; Aitkin et al., 2017; Liu et al., 2018, among others). Despite that, the simultaneous modelling of terrorist-target relationships and their activities is still imperative, other than challenging, given the large amount of subjects involved, the intricate relationships that exist among members of different groups and, ultimately, the need to identify the most influential actors. In these regards, social network methods have been increasingly employed to investigate the terrorism activities, given their ability to study such phenomena beyond the individual level (see e.g. Matthew and Shambaugh, 2005; Husslage et al., 2015; McMillan et al., 2018; Basu and Sen, 2021). A number of contributions have focused on using centrality measures for the identification of the key players in social networks (see Katz, 1953; Freeman, 1978; Brin and Page, 1998; Bonacich and Lloyd, 2001; Borgatti, 2006; Nasirian et al., 2020, among others), so that their removal could lead to severe disruption of network connectivity and, thus, of its potential harmfulness (see Berzinji et al., 2012).

Law enforcement agencies and, in general, police services, are interested in monitoring criminal activities in terrorist-related networks, with the aim of identifying the most influential actors or groups, or the "key players". Identifying key players as the most "central" nodes in a network is an intrinsically complex task, and this complex task assumes an even greater relevance when dealing with phenomena that can exert dramatic societal impacts, such as the occurrence of terrorist attacks. Although studies of terrorist network structures are inherently concerned with the evolution of these collectives, only a few investigations have considered the network evolution with the aim of detecting the terrorist groups that act as key players (or hubs) in the network, and their vulnerable targets (see, e.g., Lindelauf et al., 2013; Gialampoukidis et al., 2016; Husain et al., 2020; Malang et al., 2020).

In this paper, we propose a statistically validated bipartite network methodology to jointly examine the structural evolution of terrorist-target relationships, motivated by the use of a large attack-focused global terrorism database. The investigation of influential nodes, i.e. terrorists or targets, is achieved by representing terrorist-target relationships as bipartite networks. In this class of networks there are two

* Corresponding author.
E-mail addresses: alessandro.spelta@unipv.it (A. Spelta), nicolo.pecora@unicatt.it (N. Pecora), paolo.pagnottoni@unipv.it (P. Pagnottoni).

https://doi.org/10.1016/j.socnet.2022.08.003

Available online 1 September 2022

different types of nodes, and the edges between nodes may occur only if the nodes belong to different sets. Although several studies have been proposed in the field (see Hidalgo et al., 2007; Tacchella et al., 2012; Flori et al., 2019), to the best of our knowledge the evaluation of node importance in the bipartite (two-modes) global terrorism network is still not well defined.

We design a node evaluation technique based on the *method of reflection* (see Hidalgo and Hausmann, 2009), which allows us to retain the bipartite network representation of the relationships between terrorists and targets, and to pin down influential nodes. Differently from mainstream approaches where bipartite networks are usually compressed by one-mode projection resulting in less informative graphs, we have opted to study directly the bipartite network derived from terrorists-targets relationships. We make use of the method of reflection since this metric is informative on both the local and the global importance of each node in the network. This measure is similar to the recursive centrality developed for monopartite network. In particular, it is akin to the hub-authority centrality developed through the Hits algorithm by Kleinberg (1999) or to the PageRank algorithm (see Brin and Page, 1998). The rationale behind these metrics is that a node importance is determined by the centrality of its peers in a recursive manner. Indeed, all the measures are called feedback centrality since they exploit information of the nearest-neighbour of a node to determine its importance.

Specifically, we propose two indicators of harmfulness and vulnerability borrowing from a well-known set of measures developed in economic complexity by Hidalgo et al. (2007) and Hidalgo and Hausmann (2009), which assess the topological structure of an economic system in the global "product space". In the present context, we investigate how terrorist groups take actions in the "target space". Through our procedure, we synthesize node importance through two global indicators that we name *Harmfulness Index* (HI) and *Vulnerability Index* (VI) associated to terrorists and targets, respectively. Such measures consider both local and global properties of the network topology in order to consistently rank nodes according also to their capability to harm (of being harmed by) groups which are systematically over-targeted (over-attacked).[1]

We examine whether the information contained in the HI and VI indicators is not embedded in other simpler network measures, such as nodes degree. In other words, we test the statistical significance of such measures against a null hypothesis, which is represented by the HI and VI values computed on an ensemble of appropriately randomized networks. Finally, we propose a policy experiment designed as a network dismantling-like drill, where we reduce link weights (i.e. the number of attacks perpetrated) according to a policy function, to study how the indicators react as long as an increasing number of attacks is avoided by a hypothetical authority. This exercise has the potential to increase the power of decision analysis in enhancing decision support and advice policy makers in better dealing with the ever-increasing threat imposed by terrorism.

We find that the size of the terrorist-target network has increased non-monotonically during time, displaying an increasing hub-like structure. This phenomenon mirrors the growing transversal roles mainly played by terrorist groups which are involved into several attacks of different objectives. Terrorist groups such as the Abu Sayyaf Group (ASG), the New People's Army (NPA), the Kurdistan Workers' Party (PKK) and Taliban exhibit a linear expanding evolution of both levels of the Harmfulness Index meaning that these groups have increased the number of attacks pointing also to more vulnerable targets. Moreover, statistically validated null models provide evidence that the information contained in the higher-order measures is new and not included in the other variables, thus emphasizing the usefulness of these topological indicators. Finally, from a policy perspective, we show that a centralization of resources targeting terrorist groups, which massively attack a few highly vulnerable targets, turns out to be more efficient than simply focusing on the prevention of many attacks perpetrated by different groups against targets with a low vulnerability.

The remainder of the paper is organized as follows: Section 2 introduces the dataset and presents the methodology; Section 3 describes the results of the analysis while Section 4 provides some policy indications for attack prevention. Finally, Section 5 concludes.

2. Data and methodology

2.1. Data description and pre-processing

Bipartite terrorist networks are obtained by extracting information from the Global Terrorism Database (GTD). The GTD is a dataset created and updated by the National Consortium for the Study of Terrorism and Responses to Terrorism (START).[2] The information on terrorist activities is extrapolated from several sources including media articles, electronic news, books and journals, as well as legal documents. The GTD currently encompasses data on terrorist attacks that occurred around the world from 1970 to 2019. The GTD is a large-scale dataset and, accordingly, a data pre-process is instrumental for building the networks before their investigation. Indeed, attacks recorded in the GTD are often ambiguous between terrorism, crime, or political violence.

The raw dataset contains nearly 201,108 observations with 135 attributes such as perpetrator information, target or victim information, number of casualties, consequences, target nationality, type of weapon used for the attack, and other information on the attacks. On the other hand, in order to extract unambiguous information, we proceed with the deletion of all the incidents that START suggested to be unclear, such as those labelled with "Unknown Terrorist" and "Unknown Target". Moreover, we exclude the attacks that have been attempted but turned out to be unsuccessful. We also remove all incidents that targeted "multinations". The data pre-processing is aimed for specifically maintaining relationships among terrorist attacks and particular nations. After the data cleaning process, the resulting dataset contains 96,672 observations.

2.2. Methodology

We construct a temporal bipartite weighted directed network by aggregating the information contained in the pre-processed dataset on a yearly basis. In bipartite graphs, nodes are separated into two disjoint sets such that links only connect nodes in different partitions. In our case, the primary set of nodes V represents the terrorist groups, while the secondary set U refers to the targets which constitute the objectives of the attacks. In particular, to create the latter set of variables, we merge information concerning target types, such as Police, Military or Government institutions, just to cite a few, and their nationalities. We generate the terrorist-target bipartite network where links E represent the total number of yearly attacks perpetrated by a terrorist v on a certain target type u. For a broader view of the overall terrorist-target relationships, we also build networks where links represent the total number of yearly fatalities induced by the attacks. We dedicate the online Supplementary Material to the results of the analysis conducted on this network.

The terrorist-target network G can be represented by its weighted adjacency matrix $G \Leftrightarrow W(G)$ in which the entry $W(v,u)$ is equal to the number of attacks perpetrated by a terrorist v against a target u. Moreover, since the network in question is dynamic, the graph can

[1] Other works have also proposed nonlinear versions of such indicators to measure centrality in bipartite networks. The interested reader can refer e.g. to: Tacchella et al. (2012, 2013), Cristelli et al. (2013), Morrison et al. (2017), and Alshamsi et al. (2018), among others.

[2] https://www.start.umd.edu/gtd/.

be decomposed over time by looking at the connections that are in place at a specific period t. The resulting sub-graphs are indicated by $G_t(V_t, U_t, E_t)$. In our specific case, the aggregation of terrorist attacks on a yearly basis generates 49 sub-networks that represent terrorist actions from 1970 to 2019. We exclude data regarding the year 1993 since more than 80% of attacks are labelled as unreliable.

In order to detect the influential nodes in the terrorist-target network, we define two indices which reflect the relative importance of each actor. In particular, as concerns terrorist groups, we introduce the Harmfulness Index (HI), which accounts for their ability of harming targets at different levels. Such index ranks terrorist nodes by identifying not only terrorists that attack the most (level-1), but it also discloses cells that point to targets offended by a huge share of terrorist groups (level-2), as opposed to those attacking only specific categories of targets. On the other hand, we also define the Vulnerability Index (VI) which measures how many times a specific target has been involved in an attack (level-1) and whether a target node is hit by terrorists aiming to attack a huge share of objectives or by terrorist groups that only focus on a few specific targets (level-2). In other words, the measure evaluates how diversified terrorists are when attacking target nodes. Level-2 measures are known, in network theory, as the average nearest neighbour strength.

As in Hidalgo et al. (2007) and Hidalgo and Hausmann (2009), we consider the temporal bipartite networks $G_t(V_t, U_t, E_t)$ described by the adjacency matrices W_t. Dropping the temporal subscript t, we define the above mentioned indices of level-N as:

$$HI_{v,N} = \frac{1}{HI_{v,1}} \sum_u W(v,u)\, VI_{u,N-1}$$

$$VI_{u,N} = \frac{1}{VI_{u,1}} \sum_v W(v,u)\, HI_{v,N-1}$$

for $N \geq 2$, with initial conditions given by the strength, or sum of link weights, of terrorists and target nodes, i.e.:

$$HI_{v,1} = \sum_u W(v,u)$$

$$VI_{u,1} = \sum_v W(v,u)$$

In a nutshell, $HI_{v,1}$ represents the number of attacks perpetrated by terrorist v, i.e. the level-1 harmfulness of the cell. $VI_{u,1}$ is the level-1 vulnerability of target u, i.e. the number of attacks suffered by u. In other words, while $HI_{v,1}$ defines the diversification of the objectives, $VI_{u,1}$ describes their commonality among different terrorist groups. Recursively, the variable $HI_{v,2}$ is the average commonality of targets hit by terrorist v, while $VI_{u,2}$ represents the average objective diversification of the terrorist groups pointing target u. Thus, $HI_{v,2}$ differentiates among terrorist groups aiming at specific targets with respect to those attacking transversally different types of objectives. On the other hand, $VI_{u,2}$ discriminates among targets pointed by few specific cells as opposed to those that are commonly attacked by many terrorist groups.

2.3. Statistical validation of HI and VI

The detection of relevant nodes in our terrorist-target network requires the distinction of the topological properties that are statistically significant. That is, we need to discriminate which higher-order properties can be directly traced back to the local node features and which are, instead, the result of the complex interplay among nodes. On this respect, it is worth to mention the works by Squartini et al. (2011a,b), Alvarez-Rodriguez et al. (2021), Robins et al. (2007), Yin et al. (2018) and Yuvaraj et al. (2021) that emphasize how higher-order network properties naturally account for structured group interactions, wherein a group is simply made up of all players that are connected by a so-called hyperlink, which is the higher-order analogous of the link. Sampling random graphs with given properties is a key step in the analysis of networks, as random ensembles represent basic null models required to identify patterns such as communities and motifs. Therefore, in our work we opt for a solution to the problem applying the Maximize and Sample method (Squartini et al., 2015) to correctly sample ensembles of networks where the constraints are soft, i.e. realized as ensemble averages. The method is based on exact maximum-entropy distributions and is therefore unbiased by construction, even for strongly heterogeneous networks.

In particular, we investigate whether nodes degrees are sufficient statistics for determining level-1 centrality measures. Hence, we first build the binary adjacency matrix A induced by W as:

$$A(v,u) = \begin{cases} 1 \text{ if } W(v,u) > 0 \\ 0 \text{ otherwise} \end{cases}$$

Then, we compute the terrorist degree as the number of different targets which a terrorist points to, i.e. $k_v = \sum_u A(v,u)$, and the target degree as the number of terrorists which hit a target, i.e. $k_u = \sum_v A(v,u)$. Secondly, we aim at discriminating whether the level-2 indices are fully determined by the level-1 measures. Accordingly, we construct null-models by preserving the degrees and the strength of both types of nodes. Such randomized counterparts of the terrorist-target network are adopted to identify whether, by accounting for k_v and k_u, in the first case, and by considering $HI_{v,1}$ and $VI_{u,1}$ in the second, the indices $HI_{v,1}$ and $VI_{u,1}$, and $HI_{v,2}$ and $VI_{u,2}$ are significantly over- (or under-) represented in the real network compared to the null-models.

To achieve this goal, we adopt a family of randomized benchmarks, i.e. ensembles of graphs where the local heterogeneity is the same as in the real network, while the topology is random in any other respect. Nontrivial relevant nodes can then be identified in the form of empirical deviations from the theoretical expectations of the null model (see Garlaschelli and Loffredo, 2008, 2009; Squartini and Garlaschelli, 2011; Saracco et al., 2015). For building graph ensembles, we use the maximum likelihood approach where model parameters, e.g. nodes degree and strength, are fixed such that their expected values match the empirically observed ones.

Let $P(G \mid \vec{\theta})$ be the conditional probability of occurrence of a graph G, depending on a set of parameters $\vec{\theta}$. For a realization of the graph $G = G'$, $P(G' \mid \vec{\theta})$ is the likelihood that G' is generated by the parameter choice $\vec{\theta}$. Therefore, for fixed G', the optimal choice for $\vec{\theta}$ is the value $\vec{\theta}^*$ maximizing $P(G' \mid \vec{\theta})$ or equivalently $\lambda(\vec{\theta}) \equiv \ln P(G' \mid \vec{\theta})$.

Considering our bipartite weighted directed networks, in the case of degree preserving, each set of nodes is characterized by a degree vector whose components are $k_v \equiv \sum_u A(v,u)$ for terrorists and $k_u \equiv \sum_v A(v,u)$ for targets. The model constrains the degree sequences to those of the real network A.

The unknown parameter vectors $\vec{\theta}_v$ and $\vec{\theta}_u$ can be determined by maximizing the log-likelihood function

$$\begin{aligned} \lambda(\vec{\theta}_v, \vec{\theta}_u) \equiv \ln P\left(A \mid \vec{\theta}_v, \vec{\theta}_u\right) = \\ = \sum_v k_v \ln \vec{\theta}_v + \sum_u k_u \ln \vec{\theta}_u \\ + \sum_v \sum_u \ln\left(1 - \vec{\theta}_v \vec{\theta}_u\right) \end{aligned}$$

where

$$P(A \mid \vec{\theta}_v, \vec{\theta}_u) = \prod_v \prod_u p_{v,u}^{A(v,u)} \left(1 - p_{v,u}\right)$$

and $p_{v,u} \equiv \vec{\theta}_v \vec{\theta}_u$. Once solutions are found, we can draw weighted links from vertex v to vertex u using the geometric distribution. Similarly, we can constrain the null model in order to reproduce the empirically observed nodes strength $HI_{v,1}$ and $VI_{u,1}$. In this case, the unknown parameter vectors $\vec{\theta}^s_v$ and $\vec{\theta}^s_u$ can be determined by maximizing:

$$\lambda(\vec{\theta}^s_v, \vec{\theta}^s_u) = \sum_v HI_{v,1} \ln \vec{\theta}^s_v + \sum_u VI_{u,1} \ln \vec{\theta}^s_u + \sum_v \sum_u \ln\left(1 - \vec{\theta}^s_v \vec{\theta}^s_u\right)$$

A centrality value associated to each node is considered statistically significant if such value, computed on the original graph, is exceptionally higher compared to its mean on random networks under the null

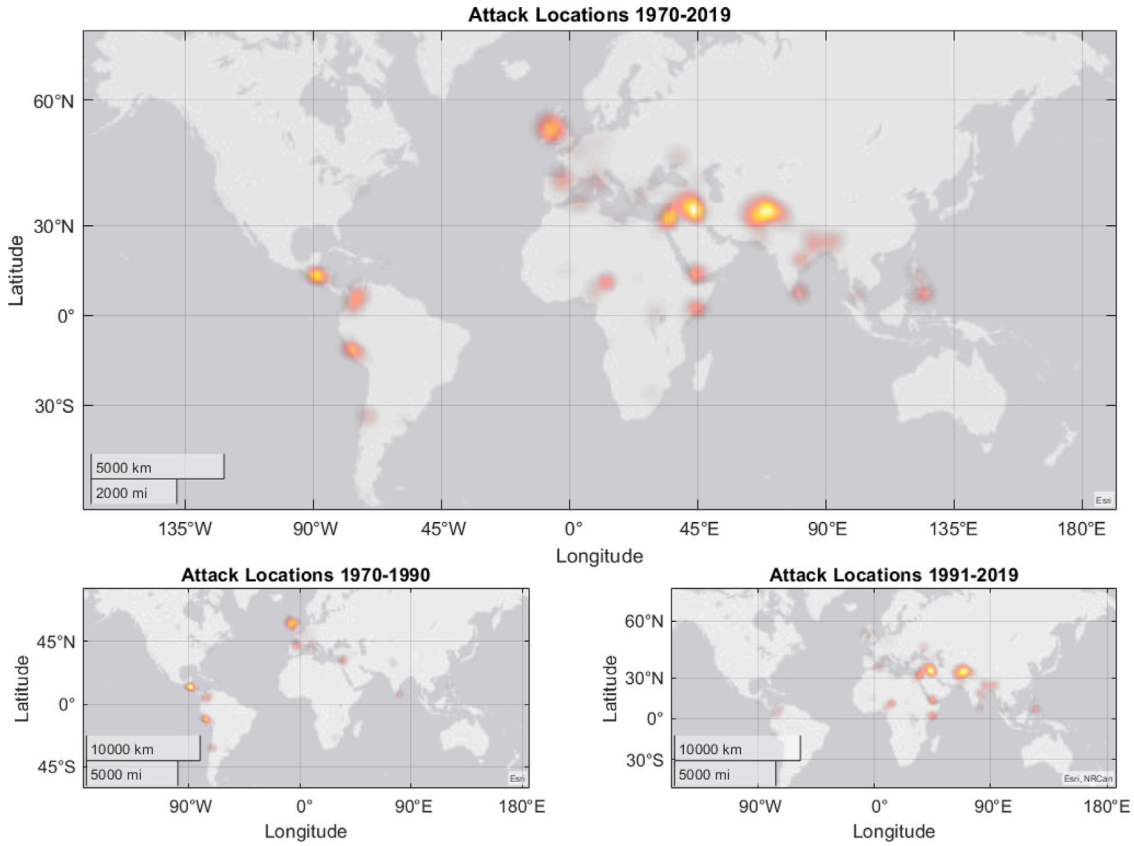

Fig. 1. Heatmap of the geographic distribution of terrorist attacks. The figure reports the geolocalization of the terrorist attacks for three different periods. In the upper panel, the subplot refers to the cumulative number of attacks occurred worldwide from 1970 to 2019. The bottom panels show the geography of the attacks for two sub-periods, 1970–1990 (left) and 1991–2019 (right). The different colours indicate the cumulative number of attacks on a certain area, the brighter the colour the higher the number of attacks that has been perpetrated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

model. To assess such exceptionality, the statistical significance of the indices is measured by means of standard inference (see Milo et al., 2002; Squartini et al., 2013). The test statistics can be defined as the difference of the value assumed by the index $\mathcal{F}(x)$ in the real network, where $x \in \{HI, VI\}$, and its mean value $\overline{\mathcal{F}}(x)_r$ in a sufficiently large set of randomized networks, divided by the standard deviation $\sigma_r(x)$ of the values for the randomized networks, namely:

$$Z(x) = \frac{\mathcal{F}(x) - \overline{\mathcal{F}}_r(x)}{\sigma_r(x)}.$$

Standard inference allows to achieve a statistical validation of terrorist-target networks, and particularly for the proposed quantitative characterizations of harmfulness and vulnerability.

3. Results

3.1. Preliminary analysis

We start our analysis by providing an overall picture of the terrorist incidents that have been carried out from 1970 to 2019. In particular, Fig. 1 reports the geographic distribution of terrorist attacks for the aggregated sample (1970–2019) and for two sub-samples (1970–1990 and 1991–2019), where the colours are associated with the cumulative number of attacks perpetrated in the area. The figure shows that areas more prone to suffer a terrorist attack are located in the Central and South America, in the horn of Africa and the Niger delta, in the Middle-East, in the regions between Afghanistan, Pakistan and India, and in the South-East Asia, while in Europe the North of Spain and the North of Ireland are the primary locations. Moreover, from the bottom panels a shift in the localization of terrorist attacks is clearly visible. While the first 20 years of our sample highlights how the attacks have been committed in the American continent and in Europe, in recent years, most of the terrorist actions have been involving mostly Middle-East countries. This shift mirrors a neat change in the motivations that drive terrorist attacks, that is from political to religious. Since the end of the 1980s, instead, there has been an increase in terrorist activity motivated by religion. Incidents and fatalities from Islamic terrorism have been concentrated in six Muslim-majority countries (Iraq, Afghanistan, Nigeria, Pakistan, Somalia, and Syria), while four Islamic extremist groups (ISIS, Boko Haram, the Taliban, and al-Qaeda) have been responsible for the majority of deaths (see GTI, 2020). Asia, however, presents a different landscape. While violent Islamist groups operate in Central, South, and South-East Asia, other organizations prioritize nationalist goals, such as those fighting for a Muslim homeland in the Philippines, the liberation of Kashmir, independence in Thailand, and recognition and rights in Burma and China.

Fig. 2 reports the yearly cumulative number of terrorist accidents discriminating among continents. From the figure, it clearly appears that most of the attacks have involved Asian countries, and in particular from the mid 2000s onward. On the other hand, the seventies constitutes a rather tranquil phase, while a surge in the attacks is especially visible during the eighties and early nineties, involving countries of the American continent. In this respect, among others, several ambushes, shootings and clashes in Colombia during the 1988, the 1992 attack on Israeli embassy and the 1994 AMIA bombing in Buenos Aires, and the Oklahoma City bombing of 1995 took place within this period. Moreover, terrorist attacks involving European countries have been perpetrated up to the nineties and have seen a resurgence from 2015 with actions mostly in France and Germany.

Fig. 3 sheds light on the mortality associated with terrorist attacks. In particular, the top panel of Fig. 3 reports the log–log distribution of

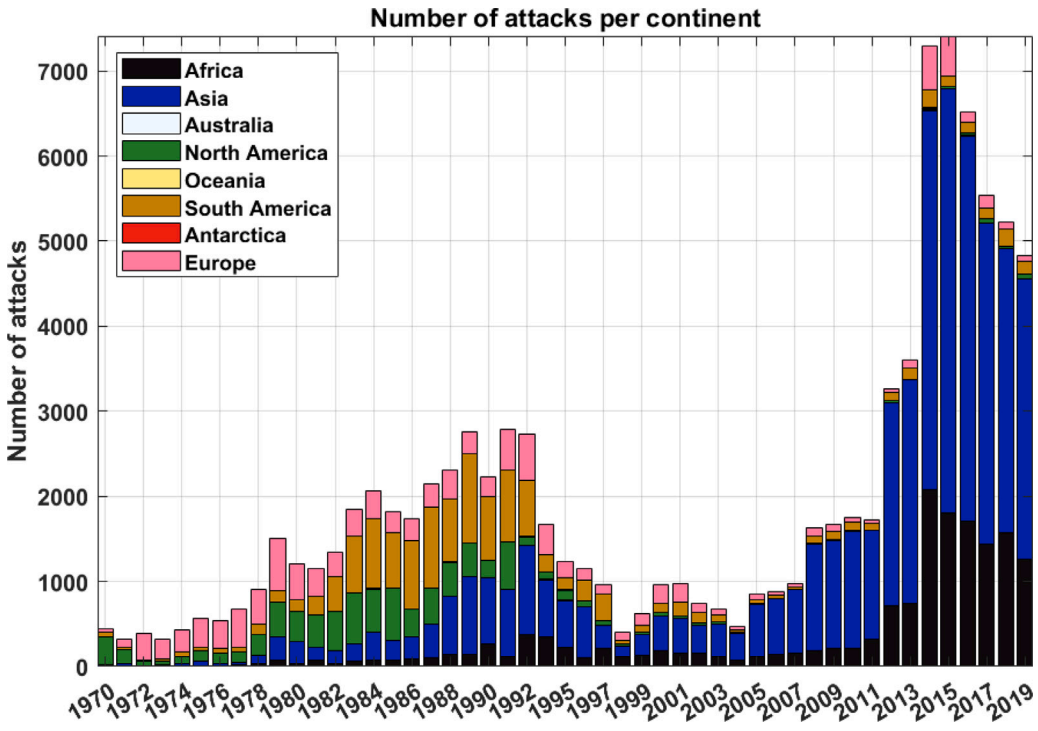

Fig. 2. Barplot of the number of attacks per continent. The figure reports the yearly cumulative number of terrorist accidents perpetrated in each continent.

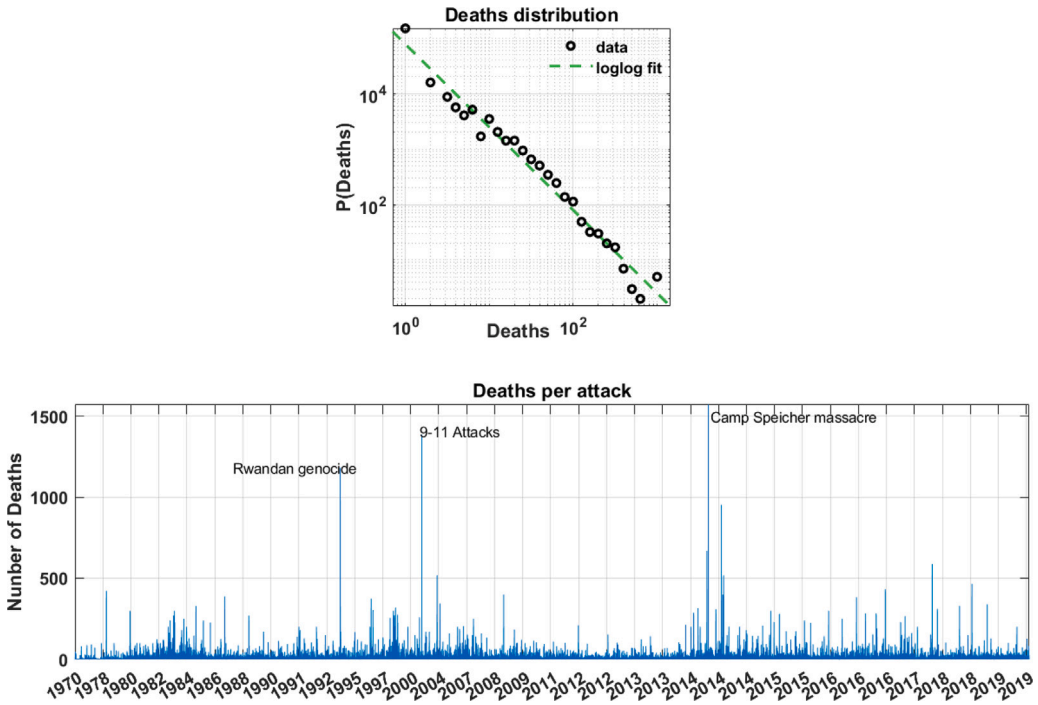

Fig. 3. Distribution of the number of deaths and number of casualties per terrorist attack. The top panel shows the probability density function of the number of deaths, reported in log–log, together with the power-law fitting. The bottom panel displays the number of casualties for each terrorist attack.

the number of deaths together with the power law fit (with exponent $\beta = -1.3943$), while the lower panel shows the time series of the number of casualties provoked by each terrorist event. From the figure it clearly emerges an heterogeneity in the number of casualties per terrorist attack. Indeed, while most of the attacks caused only a few casualties, a small number of episodes generated a disproportional amount of deaths. For instance, in April 1994, the Rwandan genocide began with several attacks (e.g. the Nyarubuye massacre) that followed one another until July, causing thousands of victims; in 2001, during the September 11 attacks the victims have been almost 3000; in June 2014, during the Camp Speicher massacre, when the Islamic State of

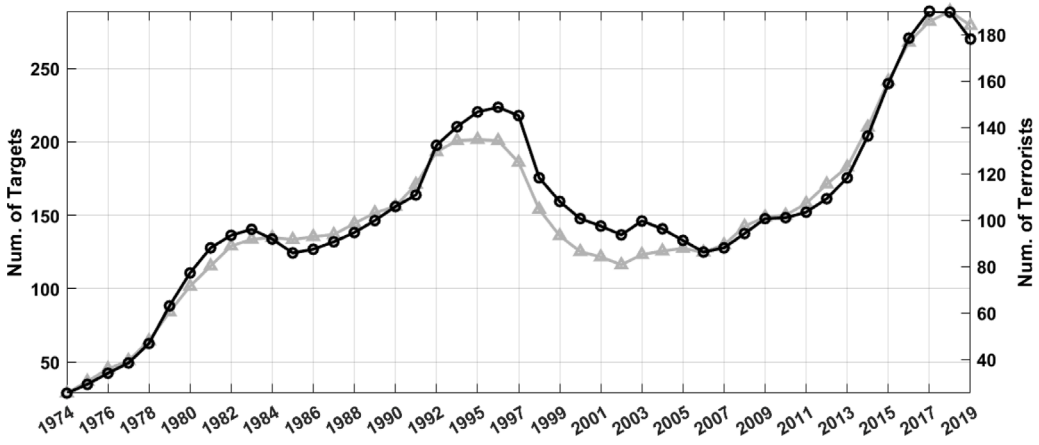

Fig. 4. 5-years moving window evolution of the number of nodes. The figure reports in black the dynamics of the terrorists (right y-axis) and in grey the targets (left y-axis) nodes along time by applying a 5-year moving window.

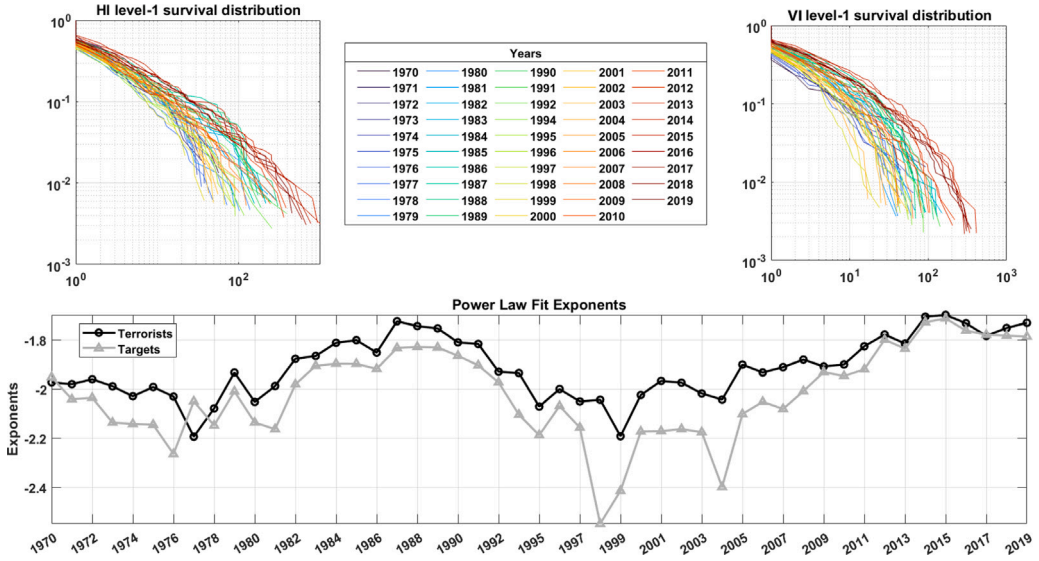

Fig. 5. Survival distribution of the level-1 Harmfulness and Vulnerability indices. The top panels report the yearly log–log survival functions associated with the level-1 HI (left) and VI (right) measures. The lower panel displays the dynamics of the power law exponents associated with the probability functions for both terrorists and targets.

Iraq and the Levant (ISIL) killed around 1700 Iraqi people in an attack in Tikrit, Iraq.

3.2. Network analysis for terrorist identification

From the beginning of the seventies to the nineties, the size of the terrorist-target network has substantially increased, although in a non monotonic way, as Fig. 4 suggests. In 1996, indeed, the number of nodes in the network reaches a maximum and the increasing trend starts again few years later. Interestingly, the last part of the sample, approximately from 2017, sees a decrease of both the number of terrorists and targets forming the network.

Besides the growth of the number of nodes composing the terrorist-target network, Fig. 5 reports the log–log survival distributions of level-1 Harmfulness and Vulnerability indices, together with the respective power law fitted exponents in the bottom panel. Similarly, Fig. 6 reports the log–log survival distributions of the HI and VI for level-2. In all the cases, we observe a general upward shift of the curves, highlighting an increasing hub-like structure of the network. In other words, influential terrorist nodes tend to point to a multiplicity of targets. This phenomenon is the direct reflection of the transversal role mainly played by terrorist groups which are involved into several attacks of different objectives. The main difference between the shape of the HI and VI distributions at the two levels relies on the magnitude of the power law exponents, which describes the evolution of the measures through time. At level-1 indeed, the power law exponent related to the terrorist nodes assumes higher values compared to the one associated to targets, while for level-2 measures, the opposite is true. This hints a lower heterogeneity of the HI level-1 measure, with respect to its VI counterpart. For the level-2 measures, the power law exponents of the terrorist nodes suggest that the average commonality of targets hit by terrorists is less evenly distributed with respect to the average objectives diversification.

For a deeper inspection of the role played by terrorist-target nodes in the network, in Figs. 7 and 8 we report the bar plots of the top ranked terrorists and targets. In particular, for both HI and VI measures, we show the terrorist and target nodes which have been ranked among the top five nodes at least three times along the whole time sample. Fig. 7 (left panel) illustrates with blue bars the top ranked terrorists according to the level-1 HI index, while the red bars (right panel) show the top terrorists according to the level-2. While Taliban and IRA (Irish Republican Army) are ranked in the first positions in terms of level-1 Harmfulness, followed by the FARC (Revolutionary Armed Forces of Colombia), the SL (Shining Path) and the FMLN (Farabundo Marti

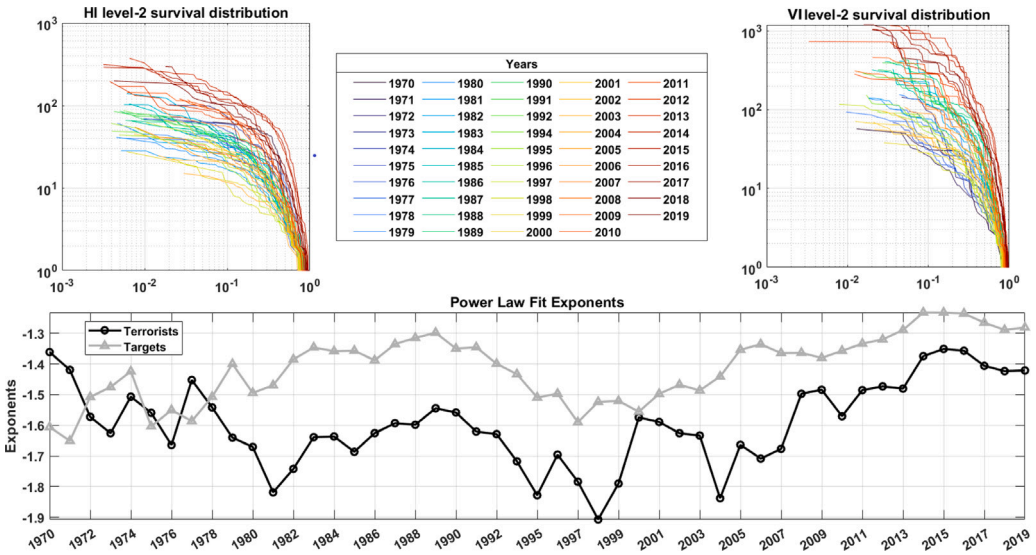

Fig. 6. Survival distribution of the level-2 Harmfulness and Vulnerability indices. The top panels report the yearly log–log survival functions associated with the level-2 HI (left) and VI (right) measures. The lower panel displays the dynamics of the power law exponents associated with the probability functions for both terrorists and targets.

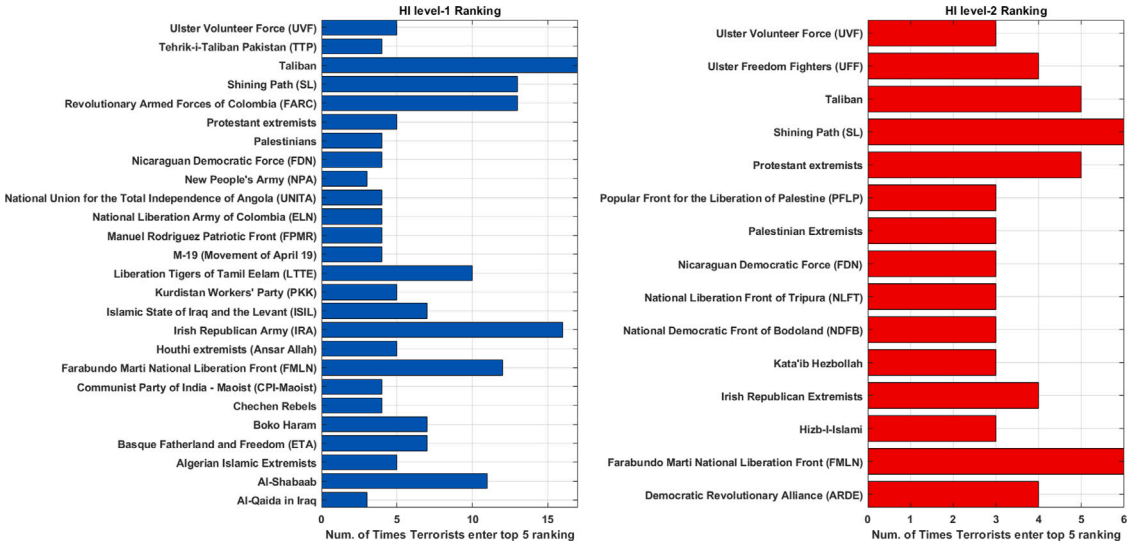

Fig. 7. Bar plot of top ranked terrorists. The left panel reports the groups that have been ranked, at least three times, among the top 5 terrorist cells according to HI level-1 over the whole sample period. The right panel illustrates the same information computed on the HI level-2 measure.

National Liberation Front) groups, the level-2 suggests an inversion of such roles, with SL and FMLN being the top terrorist nodes. This means that Taliban and IRA have executed the majority of the attacks but SL and FLMN, despite having attacked less objectives, have hit more fragile targets, i.e. with a higher value of the VI level-1 index. This occurrence is reported in Fig. 8, which displays with blue bars (left panel) the most vulnerable targets according to the level-1 VI measure. In fact, the bar plot shows that private citizens in India have suffered the highest number of attacks, followed by targets in South America and Afghanistan. Finally, the right panel of Fig. 8 reports the top ranked targets according to the level-2 VI measure, by showing the vulnerability of educational institutions and airport infrastructures in Afghanistan. This result is due to the high number of terrorist groups active in the area.

3.3. Quantitative relationship between level-1 and level-2 measures

To shed light on the dynamical relationships among HI level-1 and level-2 measures for all terrorist groups active in the last fifteen years, we report in Fig. 9 the scatter plot between the yearly HI level-1 and level-2 values. We observe two different general trends. The first refers to the behaviour of Abu Sayyaf Group (ASG), the New People's Army (NPA), the Kurdistan Workers' Party (PKK) and Taliban which exhibits a linear increasing evolution of both measures meaning that they have increased the number of attacks pointing also to more vulnerable targets. The second set of terrorists, instead, displays a non linear trend.

Next, we present the results obtained from null models. These findings are instrumental to understand whether the observed topological HI and VI measures can be explained relatively simply in terms of the observed heterogeneity of terrorist-target vertices. Comparing the indices with the predictions of a null model allows us to determine which observed structural properties are not simply explained by the constraint specifying the null model itself. Indeed, our most informative findings will correspond to a deviation, rather than an agreement, with null models. Fig. 10 shows the evolution of the Z-score distributions for the HI and VI level-1 measures, by comparing the terrorist (left panel) and target (right panel) scores with those obtained by a degree

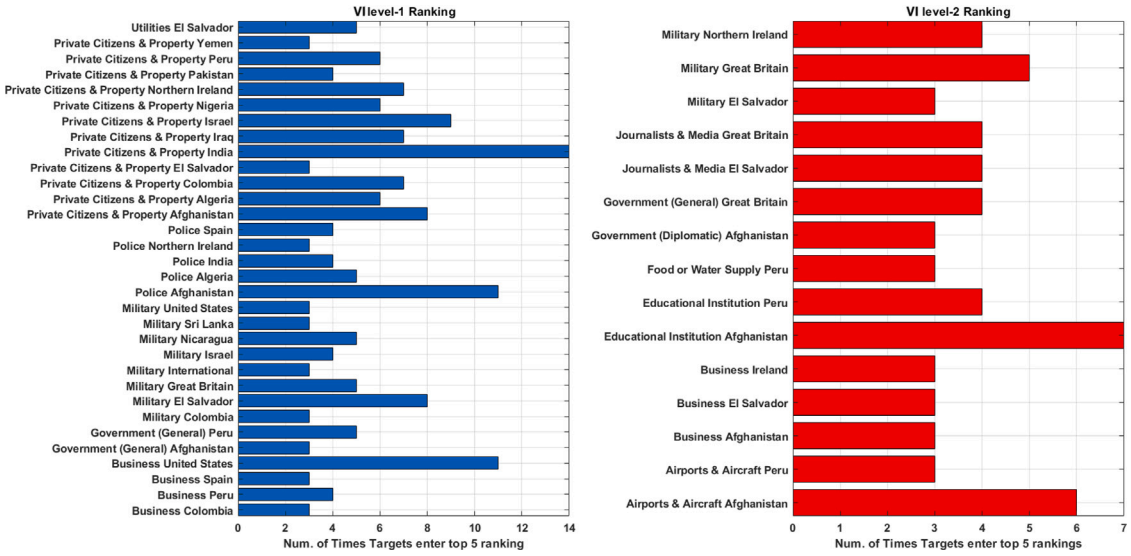

Fig. 8. Bar plot of top ranked targets. The left panel reports the objectives that have been ranked, at least three times, among the top 5 attacked targets according to VI level-1 over the whole sample period. The right panel illustrates the same information computed on the VI level-2 measure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

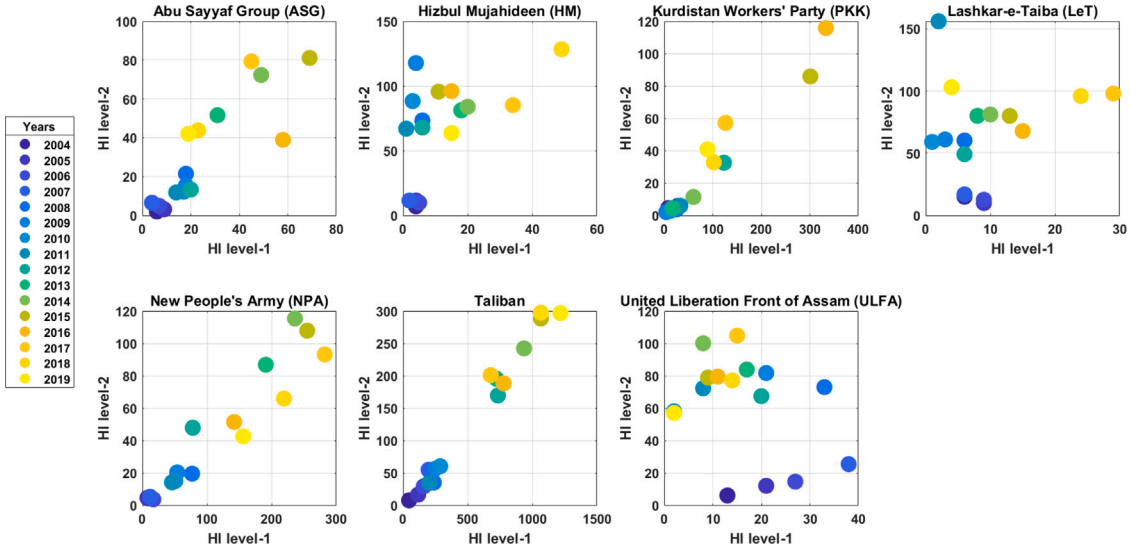

Fig. 9. Evolution of the HI level-1 and level-2 measures. Each panel identifies HI level-1 and level-2 relationships involving a particular terrorist cell. Different colours highlight the harmfulness dynamics for the terrorist groups active from 2004 to 2019. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

preserving null model. We observe heavy positive tails of both HI and VI measures for most of the years meaning that level-1 Harmfulness and Vulnerability are not accounted by nodes degree. In other words, the number of attacks cannot be predicted by simply counting the number of targets pointed by a certain terrorist.

To identify the terrorists and targets associated with the tales of the Z-score distributions we compute, as an illustrative measure, the cumulative sum of the Z-scores greater than 1.645. Figs. 11 and 12 report the over-represented terrorists and targets, respectively. Moreover, we highlight the top 15 nodes reporting their names. This analysis unveils that the most representative terrorist nodes exhibit abnormal Z-score values during the most recent years, namely from about the beginning of 2000, with the exception of the terrorist groups involved in the North Ireland fights during the eighties. On the other hand, over-represented targets can be split into two general groups: one containing targets related to the Middle-East with the highest Z-score values assumed approximately from the first decade of 2000s, and another one comprising South-American targets which suffered attacks during the eighties.

Finally, we focus on the level-2 measures by showing results related to statistical validation of null models constrained to nodes strength. In so doing, we discriminate whether the level-2 indices are fully determined by the level-1 measures. Fig. 13 reports the Z-score distributions for the HI and VI level-2 measures, by comparing the terrorist (left panel) and target (right panel) scores with those obtained by a strength preserving null model. Differently from the previous case, we observe Gaussian like distributions of both HI and VI for most of the years. This means that the most active terrorists also point to the most vulnerable targets. On the other hand, in some periods, the null model fails to reproduce the empirically observed indices, since we observe long tails of the distributions around the seventies and at the end of the nineties. This feature is confirmed by Figs. 14 and 15 which depict the over-represented terrorists and targets, respectively. Indeed, the stacked bars exactly refers to the aforementioned periods, highlighting

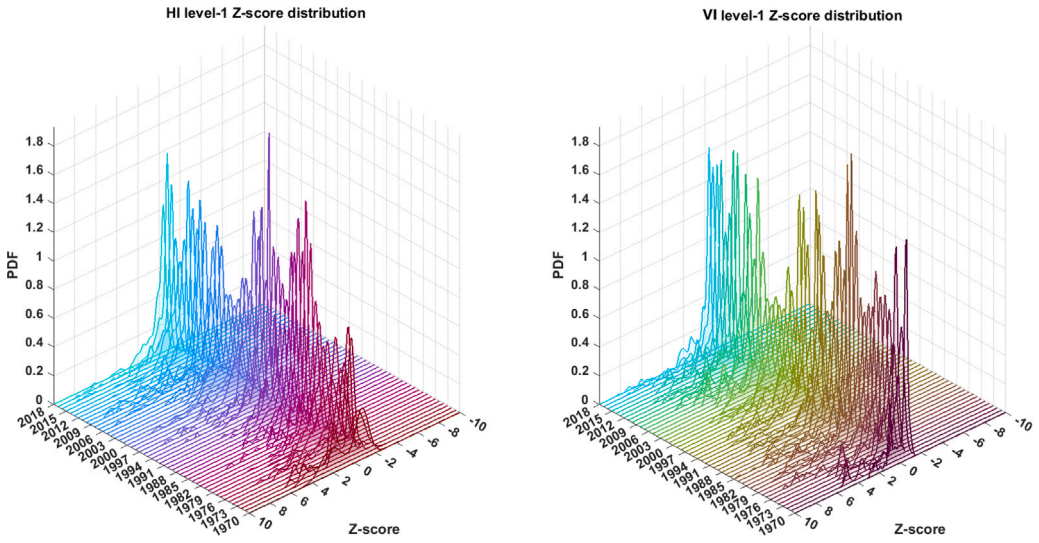

Fig. 10. Yearly Z-score distributions. The left panel shows the yearly Z-score probability density functions related to the HI level-1 index. The right panel reports the yearly Z-score distributions related to the VI level-1 index.

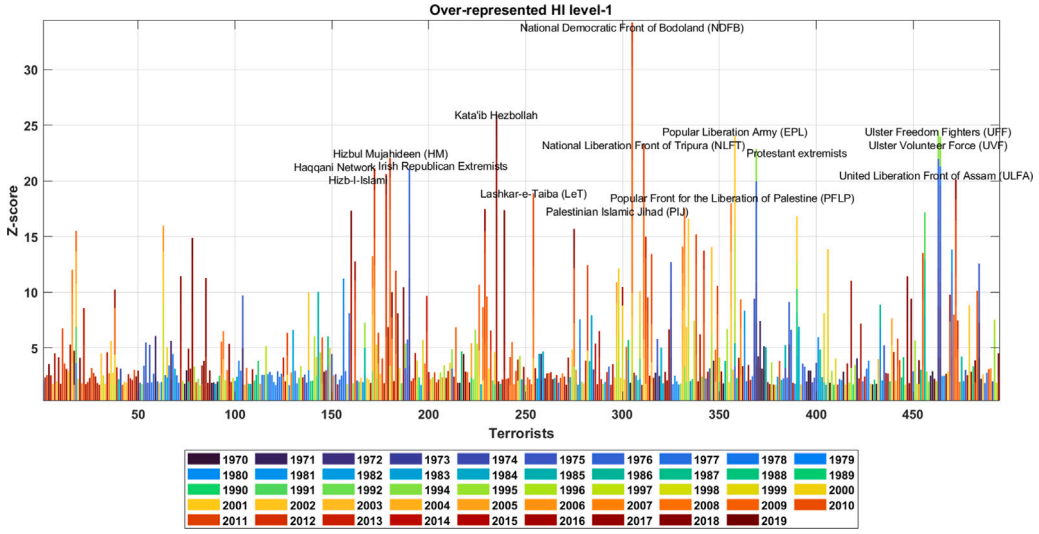

Fig. 11. Over-represented terrorist groups according to HI level-1. The figure shows the over-represented terrorist groups defined as the cells with Z-scores higher than 1.645. Colours are associated with different years, while names corresponds to the top 15 terrorists. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the role played by terrorist cells in the Middle-East, in the American continent and in Great Britain.

4. Attack preventing policies

In this Section we put ourselves in a hypothetical law enforcement agency shoes, and we carry out a scenario analysis to model a possible attack preventive strategy derived from the interplay between: (a) the total resources allocated to decrease the number of attacks perpetrated by terrorist groups; (b) a distribution scheme which targets some particular terrorist groups in order to prevent their chance of hitting a target. Our modelling strategy is designed as a network dismantling-like exercise similar to Borgatti (2006). This framework is closely related to percolation theory in statistical physics which describes the behaviour of a network when nodes or links are added and which can be leveraged to single out influential spreaders in networks (see Radicchi and Castellano, 2016).

More precisely, we evaluate the impact of different strategies by simulating the moments of the Harmfulness and Vulnerability distributions which would result after implementing hypothetical preventive measures. We first fix the total amount of resources dedicated to preventive purposes that translates into a percentage reduction in the number of attacks by considering different thresholds $Q = \{1, \dots, 99\}\%$. Secondly, we define a reduction scheme according to a sigmoid function to generate a vector which contains values weighting the network links. Formally, we first sort the network links $e \in \mathbb{N}$ in ascending order, then we use the positional indices associated with the sorted links and collected in the vector $\mathbf{j}$ to assign a weight α to each link according to the following function:

$$\alpha(j_i) = \frac{1}{1 + exp(-S(-j_i - \langle j \rangle))},$$

where $\langle j \rangle$ is the average value of the position vector $\mathbf{j}$ and the parameter $S > 0$ defines the shape of the sigmoid function. Let us recall that a link between a terrorist v and a target u defines the yearly number of attacks perpetrated by v against u. Then let us suppose to decrease the total number of attacks by a certain percentage and by targeting some particular links. This can be achieved by removing several links with a low weight, i.e. to prevent attacks that aim at different targets, or by addressing a few links with a high weight, i.e. to avoid many attacks by

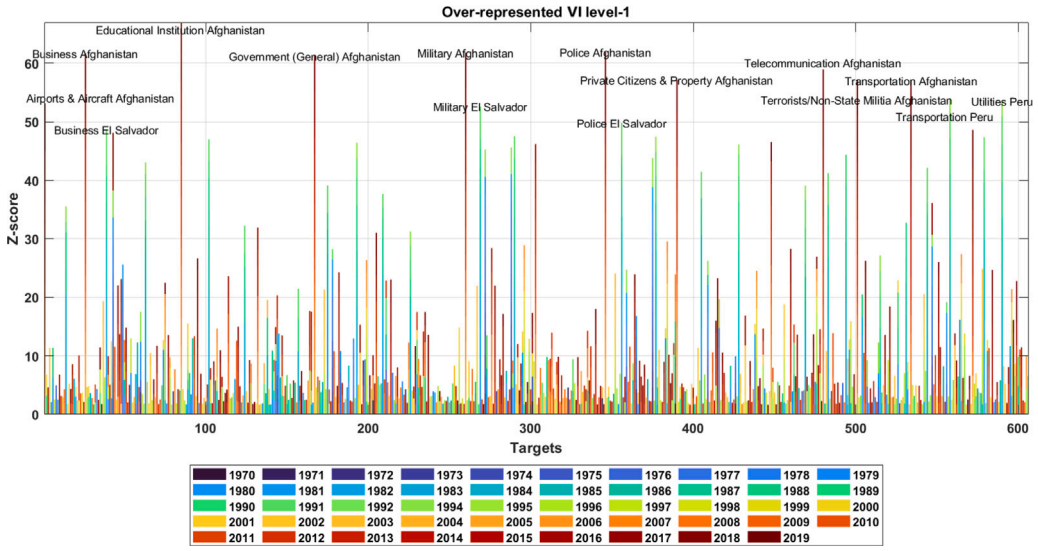

Fig. 12. Over-represented targets according to VI level-1. The figure shows the over-represented targets defined as the nodes with Z-scores higher than 1.645. Colours are associated with different years, while names corresponds to the top 15 targets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

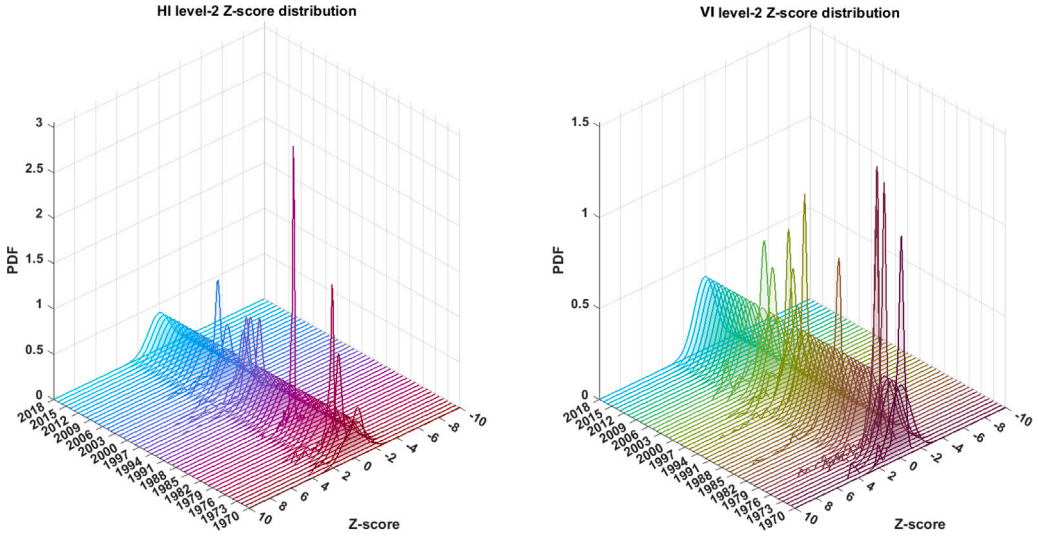

Fig. 13. Yearly Z-score distributions. The left panel shows the yearly Z-score probability density functions related to the HI level-2 index. The right panel reports the yearly Z-score distributions related to the VI level-2 index.

the same terrorist group which points to only a few targets. These two extreme cases are mirrored by high and small values of the parameter S in the policy function, as reported in the online Supplementary Material.

The weight reduction $\tilde{e}_i$ of the ith link, i.e. the attack prevented, given the value assumed by parameters Q and S, can then be computed as:

$$\tilde{e}_i = e_i - \frac{Q}{100} \sum_i e_i \frac{\alpha(j_i)}{\sum \alpha(j_i)} \tag{1}$$

where e_i is the original number of attacks perpetrated by a terrorist group versus a particular target, and $\tilde{e}_i$ is the newly reduced link weight as suggested by the application of our dismantling policy. We evaluate the proposed policy on a grid $Q \times S$, recording for each iteration the predicted evolution of the mean and the standard deviation of both HI and VI. Without loss of generality, we illustrate the results related to the terrorist-target network for the year 2019.

Results are shown in Fig. 16, where the upper panels refer to the mean and standard deviation of the HI level-2 index while the bottom panels display the outcomes for the VI level-2 index.[3] We observe a general decrease of both moments with respect to the empirically observed measures. Moreover, this trend is non-linear as reported by the green contour lines, meaning that for a fixed percentage of the total attacks prevented (as indicated by Q), targeting links with high weights is able to greatly reduce nodes' harmfulness and vulnerability with respect to a more uniform policy. In other words, this means that in order to obtain the same level of reduction of the HI and VI, an hypothetical law enforcement agency should increase the total number of attacks prevented while shifting the focus from few high-weight links to many low-weight links. Overall, a centralization of resources targeting those terrorist groups which massively attack few targets turns out to be more efficient than focusing on the prevention of many attacks perpetrated by different groups on targets with low VI values.

[3] We consider only level-2 measures since level-1 proportionally decrease with Q by construction.

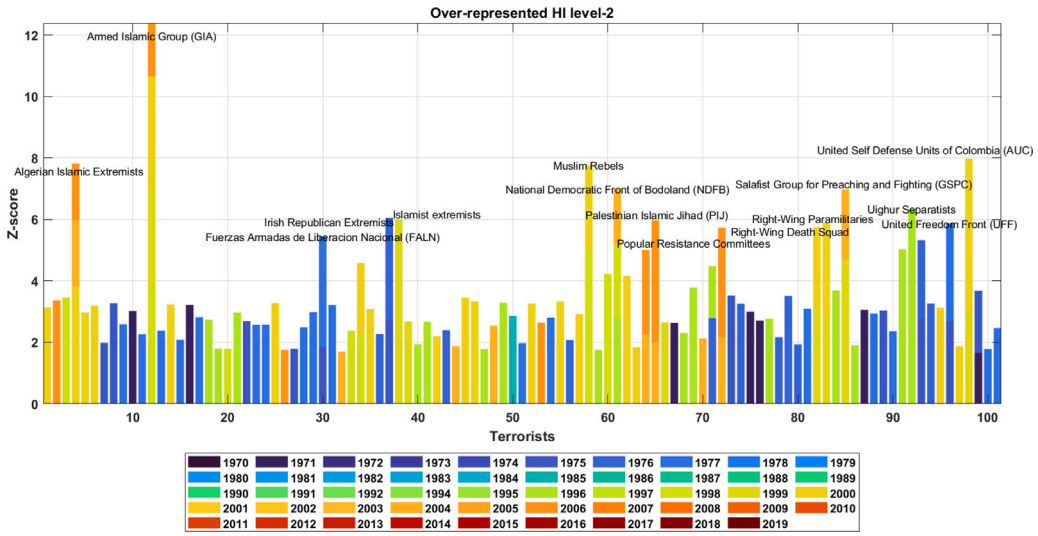

Fig. 14. Over-represented terrorist groups according to HI level-2. The figure shows the over-represented terrorist groups defined as the cells with Z-scores higher than 1.645. Colours are associated with different years, while names corresponds to the top 15 terrorists. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

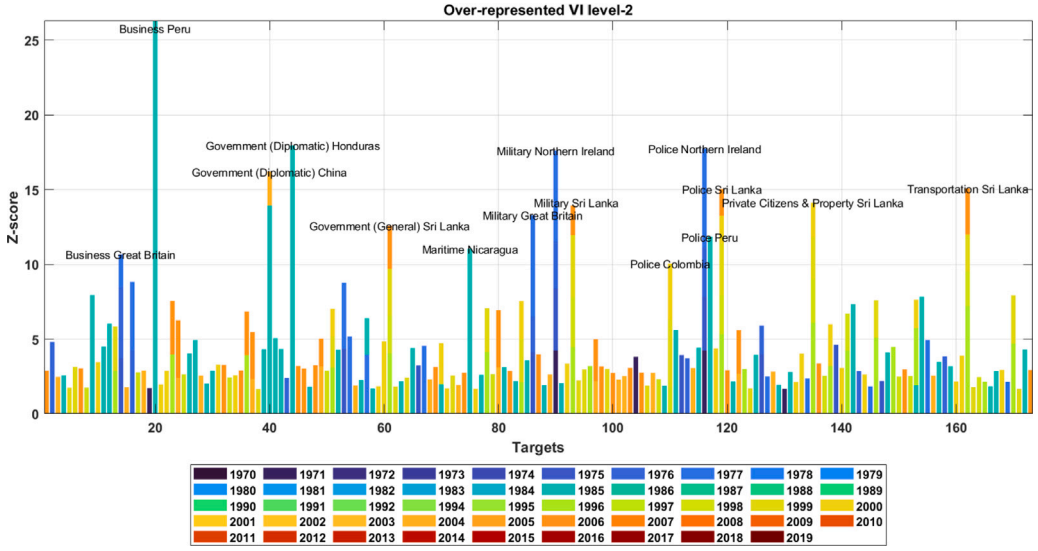

Fig. 15. Over-represented targets according to VI level-2. The figure shows the over-represented targets defined as the nodes with Z-scores higher than 1.645. Colours are associated with different years, while names corresponds to the top 15 targets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Concluding remarks

In this paper, we have proposed a dynamic bipartite network approach to examine the evolution of terrorist-target relationships and to detect relevant nodes. We have defined two measures, the Harmfulness and the Vulnerability indices, that synthesize the relative importance of terrorists and targets by considering both local and global properties of the network topology. The proposed measures are statistically validated by means of a maximum likelihood approach, where node degree and strength enter the objective function, which enable standard inference on the network parameters. We have found that terrorist groups involved into several attacks of different vulnerable objectives are the influential nodes of the network. Groups such as the Abu Sayyaf Group (ASG), the New People's Army (NPA), the Kurdistan Workers' Party (PKK) and Taliban exhibit a linear increasing evolution of their harmfulness at both local and global level, meaning that they have increased the number of attacks pointing also to more vulnerable targets. These results are also confirmed by the analysis based on the null models which highlighted the relation between the two levels of the two indices, and thus confirming the importance of including higher-order topological characteristics for a full understanding of the terrorist-target relationships. Finally, we have performed a network dismantling-like exercise to provide some policy insights, and we have shown that a centralization of resources targeting terrorists which massively attack a few targets turns out to be more efficient than focusing on the prevention of many attacks perpetrated by different groups against targets with a low vulnerability.

Future research could extend the adopted weighted bipartite network approach to consider, besides the frequency of attacks, their severity – e.g. in terms of casualties – as in related literature on operational risk measurement. For instance, the frequency of severity could be modelled through a Monte Carlo convolution of the frequency with the severity according to Giudici and Bilotta (2004) or via copulae as in Dalla Valle et al. (2008). We further mention that there are other sources of data such as the Advisen database on cyber attacks – see, for instance, Palsson et al. (2020) and Aldasoro et al. (2022) –

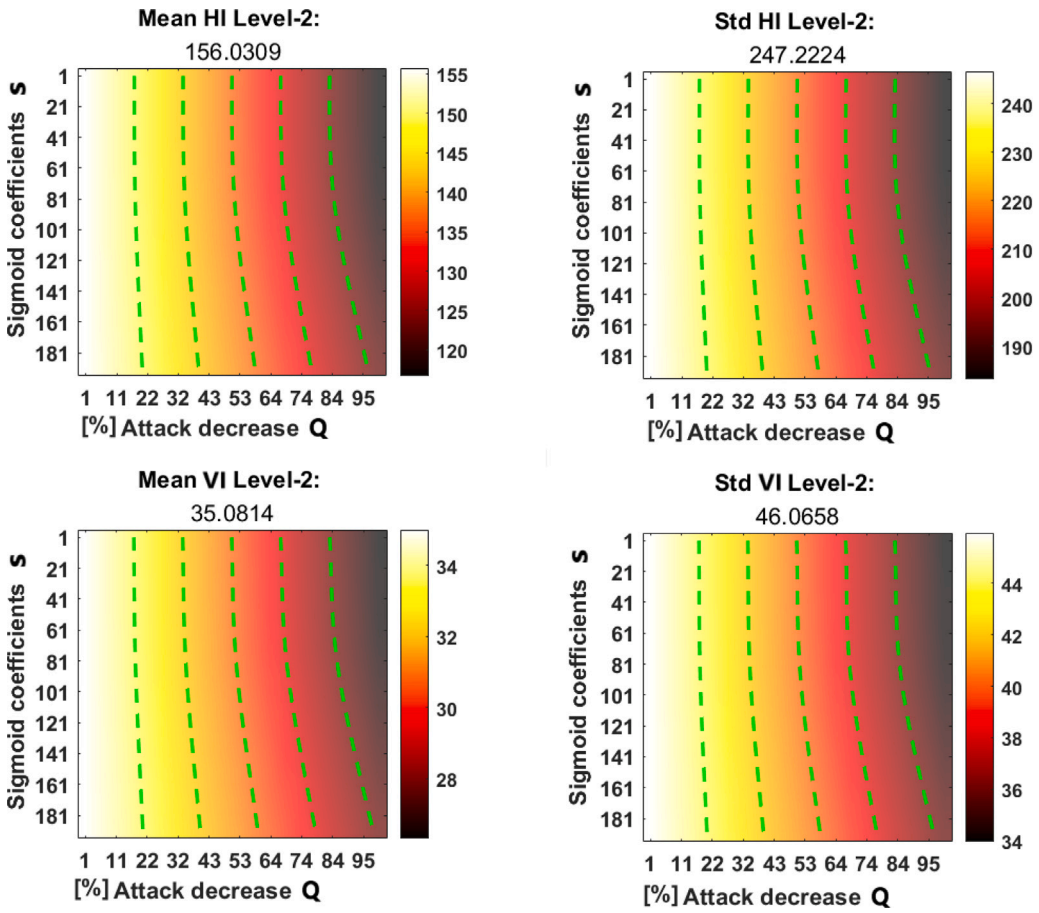

Fig. 16. Simulations of HI and VI level-2 moments in different prevention scenarios. The figure shows the mean (upper left panel) and the standard deviation (upper right panel) of the HI level-2 distribution computed by decreasing links weight, i.e. the number of attacks from 1% to 99%, and by a decreasing scheme assigned by the normalized sigmoid values. The colorbar maps colours into HI level-2 distribution moments. The lower panels report the mean (left) and the standard deviation (right) of the VI level-2. The title of each subplot reports the original value of the statistics computed on the real network. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

which could be fruitfully analysed by means of our proposed method. Finally, an additional and interesting line of research would be the adoption of alternative policy functions based on centrality measure such as in Borgatti (2006) applied to bipartite graphs with the aim of identifying key players.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank three anonymous referees for their useful comments and suggestions that led to a substantial improvement of the paper.

The research is also funded by the European Union, Horizon 2020, research and innovation program "PERISCOPE: Pan European Response to the ImpactS of COVID-19 and future Pandemics and Epidemics", grant agreement No. 101016233, H2020-SC1-PHE- CORONAVIRUS-2020-2-RTD.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.socnet.2022.08.003.

References

Abadie, A., 2006. Poverty, political freedom, and the roots of terrorism. Amer. Econ. Rev. 96 (2), 50–56.

Abadie, A., Gardeazabal, J., 2008. Terrorism and the world economy. Eur. Econ. Rev. 52 (1), 1–27.

Aitkin, M., Vu, D., Francis, B., 2017. Statistical modelling of a terrorist network. J. R. Stat. Soc.: Ser. A (Stat. Soc.) 180 (3), 751–768.

Aldasoro, I., Gambacorta, L., Giudici, P., Leach, T., 2022. The drivers of cyber risk. J. Financ. Stab. 60, 100989.

Alshamsi, A., Pinheiro, F.L., Hidalgo, C.A., 2018. Optimal diversification strategies in the networks of related products and of related research areas. Nature Commun. 9 (1), 1–7.

Alvarez-Rodriguez, U., Battiston, F., de Arruda, G.F., Moreno, Y., Perc, M., Latora, V., 2021. Evolutionary dynamics of higher-order interactions in social networks. Nat. Hum. Behav. 5 (5), 586–595.

Basu, K., Sen, A., 2021. Identifying individuals associated with organized criminal networks: a social network analysis. Social Networks 64, 42–54.

Berzinji, A., Kaati, L., Rezine, A., 2012. Detecting key players in terrorist networks. In: 2012 European Intelligence and Security Informatics Conference. IEEE, pp. 297–302.

Blomberg, S.B., Hess, G.D., Orphanides, A., 2004. The macroeconomic consequences of terrorism. J. Monetary Econ. 51 (5), 1007–1032.

Bonacich, P., Lloyd, P., 2001. Eigenvector-like measures of centrality for asymmetric relations. Social Networks 23 (3), 191–201.

Borgatti, S.P., 2006. Identifying sets of key players in a social network. Comput. Math. Org. Theor. 12 (1), 21–34.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Comput. Netw. ISDN Syst. 30 (1–7), 107–117.

Brodeur, A., 2018. The effect of terrorism on employment and consumer sentiment: Evidence from successful and failed terror attacks. Am. Econ. J.: Appl. Econ. 10 (4), 246–282.

Clauset, A., Young, M., Gleditsch, K.S., 2007. On the frequency of severe terrorist events. J. Confl. Resolut. 51 (1), 58–87.
Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., Pietronero, L., 2013. Measuring the intangibles: A metrics for the economic complexity of countries and products. PLoS One 8 (8), e70726.
Dalla Valle, L., Fantazzini, D., Giudici, P., 2008. Copulae and operational risks. Int. J. Risk Assess. Manag. 9 (3), 238–257.
Flori, A., Lillo, F., Pammolli, F., Spelta, A., 2019. Better to stay apart: asset commonality, bipartite network centrality, and investment strategies. Ann. Oper. Res. 1–37.
Freeman, L.C., 1978. Centrality in social networks conceptual clarification. Social Networks 1 (3), 215–239.
Garlaschelli, D., Loffredo, M.I., 2008. Maximum likelihood: Extracting unbiased information from complex networks. Phys. Rev. E 78 (1), 015101.
Garlaschelli, D., Loffredo, M.I., 2009. Generalized bose-fermi statistics and structural correlations in weighted networks. Phys. Rev. Lett. 102 (3), 038701.
Gialampoukidis, I., Kalpakis, G., Tsikrika, T., Vrochidis, S., Kompatsiaris, I., 2016. Key player identification in terrorism-related social media networks using centrality measures. In: 2016 European Intelligence and Security Informatics Conference (EISIC). IEEE, pp. 112–115.
Giudici, P., Bilotta, A., 2004. Modelling operational losses: a Bayesian approach. Qual. Reliab. Eng. Int. 20 (5), 407–417.
Goel, S., Cagle, S., Shawky, H., 2017. How vulnerable are international financial markets to terrorism? An empirical study based on terrorist incidents worldwide. J. Financ. Stab. 33, 120–132.
GTI, 2020. Global terrorism index 2020: Measuring the impact of terrorism.
Hidalgo, C.A., Hausmann, R., 2009. The building blocks of economic complexity. Proc. Natl. Acad. Sci. 106 (26), 10570–10575.
Hidalgo, C.A., Klinger, B., Barabási, A.-L., Hausmann, R., 2007. The product space conditions the development of nations. Science 317 (5837), 482–487.
Husain, S.S., Sharma, K., Kukreti, V., Chakraborti, A., 2020. Identifying the global terror hubs and vulnerable motifs using complex network dynamics. Physica A 540, 123113.
Husslage, B., Borm, P., Burg, T., Hamers, H., Lindelauf, R., 2015. Ranking terrorists in networks: A sensitivity analysis of Al Qaeda's 9/11 attack. Social Networks 42, 1–7.
Katz, L., 1953. A new status index derived from sociometric analysis. Psychometrika 18 (1), 39–43.
Khusrav, G., Todd, S., Donggyu, S., 2013. Common drivers of transnational terrorism: Principal component analysis. Econ. Inq. 51 (1), 707–721.
Kleinberg, J.M., 1999. Authoritative sources in a hyperlinked environment. J. ACM 46 (5), 604–632.
Kyung, M., Gill, J., Casella, G., 2011. New findings from terrorism data: Dirichlet process random-effects models for latent groups. J. R. Stat. Soc. Ser. C. Appl. Stat. 60 (5), 701–721.
Li, G., Hu, J., Song, Y., Yang, Y., Li, H.-J., 2019. Analysis of the terrorist organization alliance network based on complex network theory. IEEE Access 7, 103854–103862.
Lindelauf, R.H., Hamers, H.J., Husslage, B., 2013. Cooperative game theoretic centrality analysis of terrorist networks: The cases of jemaah islamiyah and al qaeda. European J. Oper. Res. 229 (1), 230–238.
Liu, J., Piegorsch, W.W., Grant Schissler, A., Cutter, S.L., 2018. Autologistic models for benchmark risk or vulnerability assessment of urban terrorism outcomes. J. R. Stat. Soc.: Ser. A (Stat. Soc.) 181 (3), 803–823.
Malang, K., Wang, S., Phaphuangwittayakul, A., Lv, Y., Yuan, H., Zhang, X., 2020. Identifying influential nodes of global terrorism network: A comparison for skeleton network extraction. Physica A 545, 123769.
Matthew, R., Shambaugh, G., 2005. The limits of terrorism: a network perspective. Int. Stud. Rev. 7 (4), 617–627.
McMillan, C., Felmlee, D., Braines, D., 2020. Dynamic patterns of terrorist networks: Efficiency and security in the evolution of eleven islamic extremist attack networks. J. Quant. Criminol. 36 (3), 559–581.
McMillan, C., Felmlee, D., Osgood, D.W., 2018. Peer influence, friend selection, and gender: How network processes shape adolescent smoking, drinking, and delinquency. Social Networks 55, 86–96.
Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U., 2002. Network motifs: simple building blocks of complex networks. Science 298 (5594), 824–827.
Morrison, G., Buldyrev, S.V., Imbruno, M., Doria Arrieta, O.A., Rungi, A., Riccaboni, M., Pammolli, F., 2017. On economic complexity and the fitness of nations. Sci. Rep. 7 (1), 1–11.
Nasirian, F., Pajouh, F.M., Balasundaram, B., 2020. Detecting a most closeness-central clique in complex networks. European J. Oper. Res. 283 (2), 461–475.
Palsson, K., Gudmundsson, S., Shetty, S., 2020. Analysis of the impact of cyber events for cyber insurance. Geneva Pap. Risk Insur.-Issues Pract. 45 (4), 564–579.
Radicchi, F., Castellano, C., 2016. Leveraging percolation theory to single out influential spreaders in networks. Phys. Rev. E 93 (6), 062314.
Robins, G., Snijders, T., Wang, P., Handcock, M., Pattison, P., 2007. Recent developments in exponential random graph (p*) models for social networks. Social Networks 29 (2), 192–215.
Sandler, T., Enders, W., 2004. An economic perspective on transnational terrorism. Eur. J. Political Econ. 20 (2), 301–316.
Saracco, F., Di Clemente, R., Gabrielli, A., Squartini, T., 2015. Randomizing bipartite networks: the case of the World Trade Web. Sci. Rep. 5 (1), 1–18.
Sharma, K., Sehgal, G., Gupta, B., Sharma, G., Chatterjee, A., Chakraborti, A., Shroff, G., 2017. A complex network analysis of ethnic conflicts and human rights violations. Sci. Rep. 7 (1), 1–7.
Squartini, T., Fagiolo, G., Garlaschelli, D., 2011a. Randomizing world trade I. A binary network analysis. Phys. Rev. E 84 (4), 046117.
Squartini, T., Fagiolo, G., Garlaschelli, D., 2011b. Randomizing world trade II. A weighted network analysis. Phys. Rev. E 84 (4), 046118.
Squartini, T., Garlaschelli, D., 2011. Analytical maximum-likelihood method to detect patterns in real networks. New J. Phys. 13 (8), 083001.
Squartini, T., Mastrandrea, R., Garlaschelli, D., 2015. Unbiased sampling of network ensembles. New J. Phys. 17 (2), 023052.
Squartini, T., Van Lelyveld, I., Garlaschelli, D., 2013. Early-warning signals of topological collapse in interbank networks. Sci. Rep. 3 (1), 1–9.
Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., Pietronero, L., 2012. A new metrics for countries' fitness and products' complexity. Sci. Rep. 2 (1), 1–7.
Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., Pietronero, L., 2013. Economic complexity: conceptual grounding of a new metrics for global competitiveness. J. Econom. Dynam. Control 37 (8), 1683–1691.
Yin, H., Benson, A.R., Leskovec, J., 2018. Higher-order clustering in networks. Phys. Rev. E 97 (5), 052306.
Yuvaraj, M., Dey, A.K., Lyubchich, V., Gel, Y.R., Poor, H.V., 2021. Topological clustering of multilayer networks. Proc. Natl. Acad. Sci. 118 (21), e2019994118.