Journal of Complex Networks (2022) **2**, Advance Access Publication on 1 April 2022
https://doi.org/10.1093/comnet/cnac008

The geometrical shapes of violence: predicting and explaining terrorist operations through graph embeddings

GIAN MARIA CAMPEDELLI †
Department of Sociology and Social Research, University of Trento, Via Verdi 26, 38122 Trento, Italy
†Corresponding author. Email: gianmaria.campedelli@unitn.it

JANET LAYNE, JACK HERZOFF AND EDOARDO SERRA
Department of Computer Science, Boise State University. 777 W Main St., Boise, ID 83702, USA

Edited by: Carlo Piccardi

[Received on 26 November 2021; editorial decision on 07 February 2022; accepted on 08 March 2022]

Behaviours across terrorist groups differ based on a variety of factors, such as groups' resources or objectives. We here show that organizations can also be distinguished by network representations of their operations. We provide evidence in this direction in the frame of a computational methodology organized in two steps, exploiting data on attacks plotted by Al Shabaab, Boko Haram, the Islamic State and the Taliban in the 2013–2018 period. First, we present LabeledSparseStruct, a graph embedding approach, to predict the group associated with each operational meta-graph. Second, we introduce SparseStruct-Explanation, an algorithmic explainer based on LabeledSparseStruct, that disentangles characterizing features for each organization, enhancing interpretability at the dyadic level. We demonstrate that groups can be discriminated according to the structure and topology of their operational meta-graphs, and that each organization is characterized by the recurrence of specific dyadic interactions among event features.

Keywords: political violence; security; conflict; complex networks; graph learning; neural networks.

1. Introduction

Terrorist events are generally studied by only considering their temporal dependence employing, for instance, traditional time-series approaches [1–5]. Little is known, conversely, about the complex structure of their operational interdependencies.

The literature has highlighted how certain factors can influence the different decision-making processes leading to the stage and execution of attacks [6, 7]. This wealth of research has contributed to advancing our knowledge on the diversity and heterogeneity that characterizes terrorist actions, not only concerning direct events characteristics (e.g. who to target) but also in relation to locations and timing. Nonetheless, research on terrorism has vastly overlooked the micro-level operational mechanics arising from the study of attack streams, as well as the dynamics emerging from the analysis of attacks framed in a relational perspective. Especially when we consider groups that are to some extent structured and organized, attacks are connected because they are part of coordinated campaigns occurring in bounded timeframes. This temporal feature justifies analytic designs concerned with temporal dependencies [4, 8–10]. Yet, not only are events patterned in time but they are also linked in terms of their characteristics, unravelling patterns that exist on an operational layer. An organization does not randomly choose targets or weapons: these decisions are made in relation to a specific strategy and therefore form behavioural patterns that can be mathematically represented through complex graphs.

To shed light on the distinctive behavioural patterns characterizing terrorist organizations, we here represent weekly streams of events as meta-graphs in which nodes are event features and edges map the degree of co-occurrence among features. By doing so, we computationally derive the distinctive operational topologies of terrorist organizations, which we hypothesize unravel different behavioural patterns of such organizations. To test our hypothesis, we provide a two-step analytical procedure. First, we propose a modified version of SparseStruct, called LabeledSparseStruct [11], a graph embedding algorithm that we train and evaluate against a set of state-of-the-art graph learning algorithms to demonstrate its predictive ability and show that distinctive operational patterns can be inferred to identify perpetrators of attacks. This algorithm includes edge weights, which are invaluable to understanding our dataset, and incorporates attack characteristics in the form of node labels. Second, we present SparseStructExplanation, a new graph-designed algorithmic explainer, and we analyse its results showing that—beyond structural and topological differences—also simpler dyadic edge-level distinctive operational characteristics can be ascertained. Relying on data from the Global Terrorism Database (GTD), the present study specifically focuses on terrorist attacks plotted from 2013 to 2018 by Al Shabaab by Boko Haram, the Islamic State and the Taliban.

In order to maximize the utility of meta-graphs to represent the data, optimal node embeddings must be generated prior to classification tasks. We hypothesize that embeddings which capture the structural similarity between nodes will be most effective. Additionally, the frequency of certain attack characteristics is likely to inform a predictive model. To that end, LabeledSparseStruct [11] includes node features, and weighted edges as are found in this dataset. The node features represent fundamental characteristics of the attack, and edge weights correspond to the frequency that those characteristics were observed in a given time frame. These node labels are considered in each iteration of the exploration of a node's neighbourhood. This process allows the algorithm to consider the attack characteristics associated with a node as well as those associated with each of its neighbours. Further, we introduce an explanation algorithm that scores the importance of each edge in the graphs for the classification of each terror group. This enriches the information that can be extracted about the actions of terrorist organizations, further showcasing behavioural signatures among them.

The remainder of the present article is as follows. In Section 2, we will provide an overview of research on both terrorist operations and graph learning approaches, which constitute the two building blocks—the theoretical and the methodological ones—of the work. Concerning the former aspect, we will frame the extant research in terms of a relational perspective to terrorist operations. Then, in Section 3, we will describe in detail the data used in the analyses, the information engineering step as well as the functioning of the LabeledSparseStruct embedding approach and the SparseExplanation algorithmic explainer. Section 4 will present the prediction outcomes from the comparative analyses of the various classification approaches and the interpretation outcomes. Finally, in Section 5, we will summarize the contribution of the present work, highlighting possible future research venues and intelligence implications.

2. Background

2.1 *A relational perspective to terrorist operations*

The extensive literature that studies terrorist behaviors clearly shows how several endogenous and exogenous factors often explain organizations' behavioural profiles. In turn, terrorist organizations differ in terms of the nature and characteristics of their attacks, which are either explained by or correlated with such factors.

Among these, scholars have for instance identified ideology [12–14], resources, goals and objectives [15], organizational features [16] as well as exogenous factors such as counter-terrorism strategies and

military campaigns [17]. All these aspects contribute to shed light on the critical differences that distinguish groups perpetrating political violence and help researchers and analysts disentangle the multifaceted complexity of terrorism.

Within the scholarship on terrorism decision-making, a series of studies recently presented a novel framework to study terrorist behavioural patterns through the use of complex networks. These works proposed to represent terrorist actions and operations through the use of dynamic meta-graphs mapping interconnections between attack features—such as attacked targets, employed weapons and deployed tactics—in order to capture the hidden relations between distinct events and highlight existing behavioural mechanisms that identify terrorists' profiles.

In [18] for instance, authors show how representing terrorist actions by means of meta-graphs mapping interconnections in terms of tactics, weapons and targets provides much more contextual information on the recurring patterns of terrorist violence than traditional feature space engineering practices that only rely on simple multivariate time-series measuring the frequency of feature occurrence.

Specifically, the meta-graph approach outperforms the shallow time-series approach when the outcomes of a set of algorithmic architectures on the task of targets forecasting are compared. The authors argue that the higher success determined by the meta-graph approach is related to the fact that such a technique provides information on temporal dependencies and unravels a further layer of information that includes operational dependency. This indicates that by engineering time-series through meta-graphs, the measurements embed time-dependent structural information and help identify connections between qualitatively similar behaviours. Graphs, in this regard, flexibly formalize these multiple layers of dependencies which are overlooked in traditional time-series approaches.

In [19, 20], instead, a similar representation framework is utilized to characterize terrorist behavioural profiles to then detect clusters of groups that are found to be operationally similar. In the first work, the authors analyse events pertaining to almost 1,500 organizations active worldwide from 1997 to 2016 to construct networks picturing behavioural profiles in a cross-sectional setting, using an entropy-based approach to assess the relevance of each feature mode (i.e. targets, weapons and tactics). Each mode is treated as a separate source of information for tracking organizations' behaviours, giving more importance to modes that were less homogeneous in their composition.

The second work describes a refined approach that considers multi-modal graphs originated from the network representation of tactics, weapons and targets at the yearly level, focusing only on the 105 major terrorist organizations that plotted at least 50 attacks globally from 1997 to 2018. By concentrating on yearly multi-modal graphs, Campedelli *et al.* [20] assess the evolution in the number of clusters, in their composition, and in the drivers that explain behavioural heterogeneity over time, indicating, for instance, that ideology affinity and geographical proximity do not explain the operational similarity.

While the three works seek to respond to different research questions and are designed in different ways, they all propose the use of operational meta-graphs as a new way to represent and characterize terrorist behaviours. Such a framework can be used for forecasting tasks as well as for describing the changing dynamics associated to terrorist violence worldwide. Instead of treating events as independent from one another, or events' characteristics as independent from one another, a graph-oriented approach facilitates the investigation of recurring behavioural patterns, distinctive features, anomalies and behavioural changes.

This recent methodological line of research goes beyond the traditional use of networks in terrorism studies and criminology. The now abundant literature that encompasses a network orientation on the study of illicit and criminal networks mostly concentrates on relationships between individuals [21–23]. More recent efforts have moved beyond traditional social network analysis to employ complex networks approaches to study, for instance, organized crime [24–26] and corruption [27, 28], but most of them relied

on tangible connections measured by, among others, meeting co-attendance, telephone communication or co-offending.

Social network analysis has been an important toolbox to reveal behavioural dynamics and organizational features behind a wide spectrum of organizations engaging in political violence as well. The methodological breakthroughs in network science, in fact, benefited the study of criminal and terrorist networks by providing empirical insights that have supported old theories or contributed to the generation of new ones [29–32]. In terrorism research, the network paradigm has been applied among other things to study Islamist or jihadism organizations [33–36], alliances between actors in the global scenario [37] and support and radicalization through social media platform [38–40]. Yet, relational perspectives to the study of such social phenomena fail to go beyond the tangible connections between individuals (or groups of individuals). With very few exceptions, however, the literature on networks and crime and networks and terrorism have not considered other types of relationships, e.g., those between events or characteristics of events, which may reveal underlying knowledge structures that escape the traditional methodologies embedded in traditional Euclidean spaces generally employed to study actors or their behaviours.

The particular representation strategy outlined in this work follows this intuition and is situated in the theoretical framework of the strategic theories of terrorism. This frame originates in the work of [41] and posits that terrorist actions are instrumental choices made by a rational entity (i.e. the group) to obtain particular gains or achieve specific objectives. The strategic frame imposes to treat each group as if it was guided by collective rationality [42, 43]: a group is therefore motivated by a defined set of preferences, which in turn are transformed into events, acts and attacks. The strategic frame thus conveniently unveils the possibility that these constrained decision-making processes hold characteristics that can be formalized in a relational dimension. Proponents of the strategic frame have mostly considered terrorist actions as the consequence of neoclassical rational agents that seek to optimize a certain objective function, often identified by the groups' specific political goals [42–44]. Other competing theoretical frames such as the psychological and the organizational ones argue that alternative external and internal mechanisms influence terrorists' actions, suggesting that the strategic frame simplifies the inherent complexity of terrorism [45]. Terrorist actions might be the effect of adversarial views within the same organization, or could be determined by non-rational psychological mechanisms at the individual level. The unitary definition of terrorist groups is the nucleus around which the strategic, psychological and organizational frames wrestle. Yet, these alternative theories do not contest that actions remain strategic to some extent. Is it a rational, calculated choice to cause a deviation in standard attack strategies? Is it due to a change in leadership? Could it be explained in terms of psychological mechanisms? Regardless of the answer—which is beyond the scope of this article—the strategic frame offers a flexible description of reality. This description of reality is backed by empirical evidence demonstrating that although terrorists may not be perfectly rational agents, their actions are crucially characterized by patterns at various scales and across different dimensions in many circumstances and contexts. In light of these aspects, we model terrorist behaviours, and particularly their attacks, using graphs, which are mathematical objects that enable us to represent and capture the fundamental dependencies across events and, most importantly, across event features. The hypothesis is that the resulting structural and topological characteristics of these graphs offer a rich representation of terrorist operations, and that, given the non-random nature of their actions and the unique goals and contexts in which the groups under analysis operate, we can leverage such representations to better understand the nature of their complex behaviours.

2.2 *Graph learning: related work*

The development and application of machine learning approaches on graphs have become extremely popular in the last years, as networks are central in many domains, ranging from the natural to the social

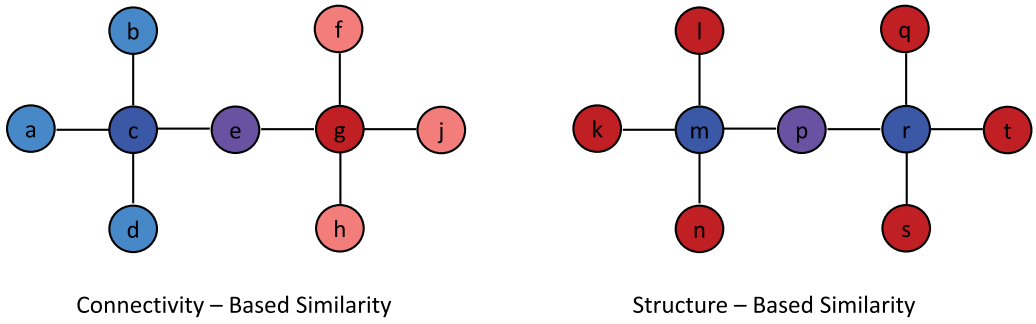

FIG. 1. Connectivity vs. structural connectivity. Node colour represents similarity. Colour for "e" and "p" indicates similarity to all other nodes in the two respective graphs.

sciences [46]. As reported by [47], the most important task of graph learning is to find ways to represent or encode graph structures in order to easily exploit this information in machine learning models. One type of graph representation learning refers to a body of approaches that convert graph data—a list of nodes and edges—into a numerical vector representation for each node. This vector then becomes the input to downstream standard machine learning tasks, such as classification. These methods aim to project the node into some low-dimensional latent space while maintaining that similar nodes will be found close to one another in said space.

Though a wide variety of methods exist, they can generally be classified into one of two groups based upon their notion of node similarity. If similarity between nodes is based upon connectivity, then nodes that share many neighbours will be proximal in the latent space. A structure-based node representation will instead project two nodes with similar neighbourhood structures close to one another.

Figure 1 shows a schematic representation of the differences between connectivity-based similarity and structural similarity. Using approaches based upon connectivity, node *c* would be proximal to *a, b* and *d*, while distant from node *g*. Node *e* would be placed between *c* and *g*, and closer to them than the leaf nodes. Notably, when a structure-based similarity is observed, one can clearly see that though nodes *m* and *r* share only one neighbour, their structural role in the network is identical. This is also true for the leaf nodes; each would be considered identical based upon their neighbourhood structure. Interestingly, node *p* would be projected between the central nodes and the leaf nodes, as it is structurally less central than *m* and *r*, but more central than the leaf nodes.

For this study, our aim is to capture information regarding the structure of the relationships between characteristics, and further to gather information about the importance of individual structures in the resulting networks. To this end, we will compare methods that are suited for structural-based proximity embeddings.

We will consider graph neural network-based methods that are expected to show good performance at capturing structural information, and particularly a graph convolutional network (GCN) and a graph isomorphism network (GIN). Like LabeledSparseStruct, both employ a method that is based upon the Weisfeiler-Lehman's graph isomorphism test, wherein a node's structural representation is iteratively updated as its neighbours are explored. Both the GCN and the GIN share a connection with convolutional neural networks, where the convolution operation is an aggregation function performed upon a node's neighbours' representations. For the GCN, the mean aggregation function is used and generates a linear mapping followed by a ReLU activation function. Mean pooling is applied to generate graph

representations [48]. GIN uses a multi-layer perceptron with a ReLU and a sum aggregation function for node representations. Graph representations are generated via sum pooling [49].

3. Materials and methods

3.1 *Data*

To study the interconnected operational features of attacks plotted by Al Shabaab, Boko Haram, the Islamic State and the Taliban, we have relied on the GTD [50]. The GTD is the most comprehensive open-access source for research on terrorist events, including more than 200,000 events. We have specifically focused on attacks plotted by the abovementioned groups in the 2013–2018 period in the current work. Although the GTD includes attacks from 1970 on, we have only considered the 2013–2018 frame because it is the most extended time window in which there is temporal overlap between the organizations. In fact, the most recent terrorist organization's first registered attack in the dataset—the Islamic State—occurred in 2013. Temporal overlap is a crucial ingredient of our prediction task, as we are particularly interested in finding distinctive operational characteristics in the topology of non-temporally separable attacks.

First, we have removed all those events of uncertain terrorist nature, as reported by the GTD, to ensure that all the events used in our study were relevant. Second, per each given group, we have kept all those attacks in which at least one of the identified perpetrators was the given group, in the time-frame of interest, which started on 1 January 2013 and ended on 31 December 2018. The GTD records up to three perpetrators for each attack: in fact, one event may be plotted by two or three allied organizations. This operation led to a total of 14,104 events. The Taliban accounted for 5,280 observations, the Islamic State 5,104, Boko Haram was responsible for 1,829 while Al Shabaab for 1,891. We have then grouped all the events by 1-week windows comprising 7 days of activity for each actor. The weekly distribution of events is provided in Fig. 2.

To generate the temporal operational graphs for each group, we have relied on three sets of variables that are available in the GTD. These are attack types (that we refer to as *tactics* henceforth), targets, and weapons. In the GTD, each event can be associated with up to three different tactics, three different targets and four different weapons. Similarly to the group variable where an event may be plotted by more than one group at a time, an attack can thus be directed to a mix of targets, using multiple tactics and multiple weapons.

The set of attacked targets in the period under consideration is given by *Airports & Aircraft, Business, Educational Institution, Food or Water Supply, Government (Diplomatic), Government (General), Journalists & Media, Maritime, Military, NGO, Police, Private Citizens & Property, Religious Figures/Institutions, Telecommunication, Terrorists/Non-State Militia, Tourists, Transportation, Unknown, Utilities and Violent Political Party*. The set of weapons instead includes *Chemical, Explosives, Fake Weapons, Firearms, Incendiary, Melee, Other, Unknown and Vehicle (not to include vehicle-borne explosives, i.e., car or truck bombs)* . Finally, possible tactics are *Armed Assault, Assassination, Bombing/Explosion, Facility/Infrastructure Attack, Hijacking, Hostage Taking (Barricade Incident), Hostage Taking (Kidnapping), Sabotage Equipment and Unarmed Assault*.

We used all the available information on the operational characteristics of each event to generate our graphs. The specific procedure is described below.

3.2 *Operational meta-graphs*

3.2.1 *Defining operational meta-graphs* Once our filtered dataset containing all the attacks plotted by each of the four groups has been prepared, we have proceeded to create the temporal operational graphs.

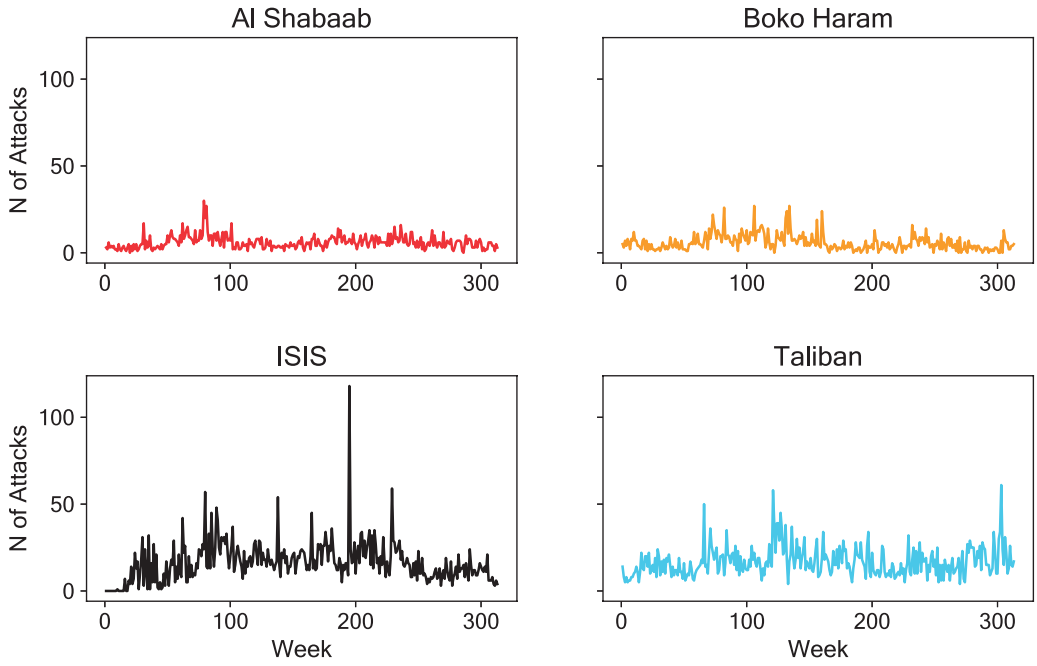

FIG. 2. Number of terrorist attacks per week in the 2013–2018 time window, per group.

However, before moving to the proper explanation of how meta-graphs are generated, it is useful to provide a general, non-technical definition to describe what they are.

An operational meta-graph is intended as a graph having as nodes specific terrorist events' features that have characterized attacks in a given time frame. In the present work, every operational meta-graph is generated using the information at the weekly level. Hence, for instance, the node-set in a meta-graph contains the targets hit by a given group in the week under consideration, as well as the tactics utilized and the employed weapons. These nodes (i.e. event features) would form connections if they co-occurred in some events in the same week. Connections are weighted, meaning that the higher the number of times a certain target has been associated with a specific weapon, the stronger the connection will be. It is relevant to note that edges do not connect together attacks, but features of attacks. We are not using the framework to connect attacks that are part of coordinated campaigns: we are interested in the relationships that map feature co-occurrence in the same time unit. Describing attacks via meta-graphs leads to a topological representation of terrorist actions that do not treat event features as independent of one another. Instead, temporally close attacks are seen in a relational perspective, in line with the theoretical proposition that strategic patterns govern terrorist decision-making. To exemplify, by analyzing attacks of a group we might find that Attacking religious institutions is put in relation to the use of firearms, which in turn are also connected to diplomatic targets, and so on. Our core hypothesis then is that this web of connections between features should hold distinctive characteristics for different groups, as a byproduct of distinct strategies.

3.2.2 *Generating operational meta-graphs* For each group o in our set O, we have represented the terrorist attacks plotted in the 2013–2018 period by separating them in weekly chunks that can be formalized as sequential bipartite matrices $\mathbf{W}_{1,2,\ldots,313}$ in the ordered list $\mathcal{W}$, where one mode represents the

days associated to a specific week, and the other mode represents the specific set of tactics, targets and weapons that characterize a group's repertoire. For a group o_i, Equation 1 visualizes the weekly bipartite matrices $\mathbf{W}_{1,2,\ldots,313}$, where the entries α are integers indicating the number of times a specific feature $\mathrm{F_n}$ has characterized attacks in a given day within $\mathbf{W}$.

$$\mathbf{W}_{1,o_i} := \begin{array}{c} \\ \text{day 1} \\ \vdots \\ \text{day 7} \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{F_1} & \cdots & \mathrm{F_n} \end{array} \\ \begin{pmatrix} \alpha_{1,\mathrm{F_1}} & \cdots & \alpha_{1,\mathrm{F_n}} \\ \vdots & \ddots & \vdots \\ \alpha_{7,\mathrm{F_1}} & \cdots & \alpha_{7,\mathrm{F_n}} \end{pmatrix} \end{array}$$

$$\mathbf{W}_{2,o_i} := \begin{array}{c} \\ \text{day 8} \\ \vdots \\ \text{day 14} \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{F_1} & \cdots & \mathrm{F_n} \end{array} \\ \begin{pmatrix} \alpha_{8,\mathrm{F_1}} & \cdots & \alpha_{8,\mathrm{F_n}} \\ \vdots & \ddots & \vdots \\ \alpha_{14,\mathrm{F_1}} & \cdots & \alpha_{14,\mathrm{F_n}} \end{pmatrix}. \end{array} \tag{1}$$

$$\vdots$$

$$\mathbf{W}_{313,o_i} := \begin{array}{c} \\ \text{day 2184} \\ \vdots \\ \text{day 2191} \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{F_1} & \cdots & \mathrm{F_n} \end{array} \\ \begin{pmatrix} \alpha_{2184,\mathrm{F_1}} & \cdots & \alpha_{2184,\mathrm{F_n}} \\ \vdots & \ddots & \vdots \\ \alpha_{2191,\mathrm{F_1}} & \cdots & \alpha_{2191,\mathrm{F_n}} \end{pmatrix} \end{array}$$

The bipartite graphs in the $\mathcal{W}$ list are then individually projected into monopartite graphs in order to obtain $F \times F$ weekly weighted graphs for each group, where nodes are the features derived from the repertoire of tactics, targets and weapons of each group. These graphs are then stored in the ordered list $\mathcal{G}$, where they are indexed by the week they refer to. For a group o_i in a week W_k, the projection into $\mathbf{G}$ is computed multiplying the transpose of the bipartite graph with the original bipartite graph itself:

$$\mathbf{G}_{W_k;o_i} = \mathbf{W}^{\mathrm{T}}_{k,o_i}(\mathbf{W}_{k,o_i}). \tag{2}$$

This step leads to a square monopartite matrix with non-zero diagonal, where entries are weights $\mathcal{W}$. Raw weights in each monopartite matrix are a measure of the number of times two features F_l and F_m co-occurred together in attacks plotted within the range of 7 days covering each week.

Then $\mathcal{G}_{o_i} = (\mathbf{G}_{W_1;o_i}, \mathbf{G}_{W_2;o_i}, \cdots, \mathbf{G}_{W_k;o_i})$ represents the ordered list of weekly operational meta-graphs for a group o_i. A sample visualization of the final outcome of the meta-graph processing phase is provided in Fig. 3.

3.3 *The embedding approach*

3.3.1 *Graph basics and notations* Let $\mathbf{G} = (V, E, L, \mathcal{W})$ be a labelled, weighted, undirected operational meta-graph (for parsimony in this section we will ignore the subscripts referring to the specific week and terrorist group an operational meta-graph is associated with). V denotes the set of all nodes in the graph, and $E \subseteq (V \times V)$ the set of edges connecting the nodes. The list of node labels and edge

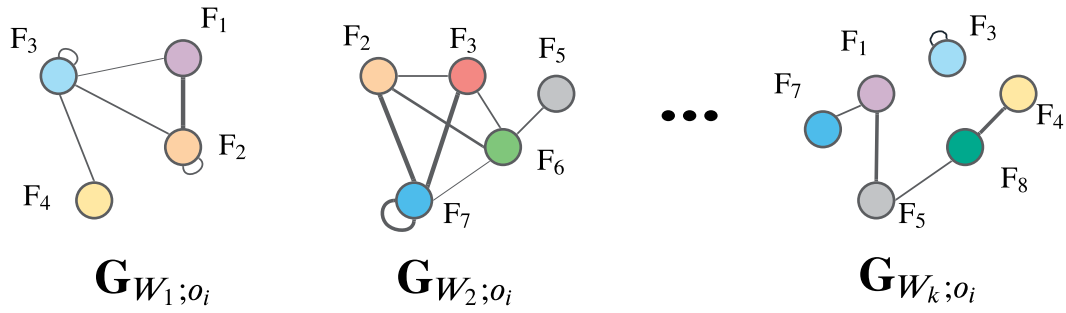

FIG. 3. Sample visualization of operational meta-graphs $\mathbf{G}$ for 3 weeks W_1, W_2, W_k for group o_i. Nodes are the features present in each weekly time window, indexed by $F_1, ..., F_8$.

TABLE 1 *Notations found throughout this work*

Notation	Description
$\mathbf{G}$	The given graph
V	The set of nodes in the graph
E	The set of edges in the graph
$\mathcal{W}$	The list of edge weights in the graph
L	The list of node labels in the graph
$\|V\|$	The number of vertices
$\|E\|$	The number of edges
u	Single node in the graph
$nbr(u)$	The set containing the neighbours of node u
$knbr(u,k)$	The set of neighbours in the k-hop neighbourhood of node u
n	The size of the representation of node u

weights are denoted as L and $\mathcal{W}$. $|V|$ and $|E|$ represent the number of nodes and edges, respectively. The neighbourhood of node $u \in V$ is given by $nbr(u) = \{v|(u,v) \in E \text{ or } (v,u) \in E\}$. For our purposes, $knbr(u,k)$ will refer to the set of all edges connecting each node in the k-hop neighbourhood of node u. Following the embedding algorithm, each node will have a representation of the same size, denoted here as n. The full list of notations is found in Table 1.

3.3.2 *LabeledSparseStruct algorithm* In [11], the authors present SparseStruct, an unsupervised method which generates structural node representations based upon Weisfeiler–Lehman's graph isomorphism test [51].

SparseStruct was found to perform at least as well many state of the art graph representation learning algorithms including GraphSage [52], Node2Vec [53], Struc2Vec [54] and GraphWave [55] on both node classification and in its ability to capture node structure. Additionally, SparseStruct shows rapid convergence of the algorithm, and efficient time complexity.

We here propose a modified version of SparseStruct, denominated LabeledSparseStruct , which incorporates weighted edges and node labels into the structural representation of each node. Prior to

generating node embeddings, the edge weights are normalized by the terror group and timestamp to which they belong. Weights are row-normalized as follows in the range [0, 1] in order to control their high variance, a byproduct of the variations in terms of number of attacks observed for the considered groups:

$$\mathcal{W}^{\text{Norm}}_{u,v_W} = \frac{\mathcal{W}_{u,v_W}}{\text{sum}(\mathcal{W} \in \text{TerrorGroup}_{u_W})}. \tag{3}$$

Row-normalization allows easier temporal comparison in relative terms among features. Trivially, the higher the weight, the higher the number of times two features have co-characterized events in the same weekly time window.

Briefly, the algorithm (Algorithm 1) consists of an iterative convergent process that explores the neighbourhood of each node and updates the node's representation at each depth. The algorithm continues until the maximal number of allowed iterations is reached or further iterations fail to generate additional novel node structures. Here, we say that the algorithm has converged. We begin the while loop with all nodes considered to have the same structure, so the starting matrix is set to a single column, all with index zero. At each iteration i, the one-hot encoded labels are concatenated to the previous iteration's node representation matrix. The *SparseMatrixGen* method assigns to all the identical rows the same progressive identifier with the use of a hash tree index (Line 17). A sparse matrix SM^i is then generated of size $|V| \times |indexID(SM^i)|$ (Line 18). This generates a column for each structure type that is known at this level of exploration. Each column is then updated by adding the weight of each neighbour matching the corresponding structure type. For node u, the value of the cell $SM^i[id(u), j]$ is the sum of the weights of edges between node u and all of its neighbours with identifier j (Line 20). The sparse matrices generated in each iteration are concatenated to the previous iteration's representations to capture information from the evolution of a node's structural representation as its neighbourhood is further explored. After the *SparseMatrixGen* function has converged (or the maximal number of iterations have been completed), the sparse matrix is condensed using a Truncated SVD to generate the final node embeddings.

A detailed description of the space and time complexity of the original version of the algorithm, along with its performance can be found in [11].

3.4 *SparseStructExplanation: explaining graph embedding results*

The embedding procedure provides us with a further layer of information, namelythe quantification of the importance of each connection between nodes. Once graph classification has been performed, a generic explainer can be used in fact to return importance scores for attributes of the node (or graph) embedding. Algorithm 2 traces a score back through the iterations of node representation parsing performed in Algorithm 1 and places it on all edges responsible for that structural feature. The final result is an edge score representing an accumulation of scores for all structural features to which it contributed.

During the iterations of the node embedding generation method of LabeledSparseStruct, information about each structural feature in the sparse matrix is stored in a hashmap under a tuple key (*node*, *featureIndex*). The value at each key is a list of all neighbours contributing to that feature, and the level at which they were explored (Algorithm 1, Line 21). This results in an easy lookup table to which we can refer to trace each attribute score to its component edges.

For this work, each node embedding for an individual terror group and timestamp was concatenated to represent a graph embedding. ExtraTrees was used to perform graph classification, followed by the LIME Explainer to obtain graph attribute importance scores [56]. Because the number of features was

Algorithm 1 LabeledSparseStruct algorithm

1. **function** SPARSESTRUCT($\mathbf{G} = (V, E), \mathcal{W}, L, explorationDepth, k$)
2. $\mathcal{W}^{\text{Norm}}_{u,v} = NormalizeEdgeWeights(\mathbf{G}, W)$
3. $LC = OneHotEncode(L)$
4. $SM, explanVal =$ SPARSEMATRIXGEN($\mathbf{G}, \mathcal{W}^{\text{Norm}}_{u,v}, L, explorationDepth$)
5. $M, svdComponents = TruncatedSVD(SM, k)$
6. **return** $M, explanVal, svdComponents$
7. **end function**
8. **function** SPARSEMATRIXGEN($\mathbf{G} = (V, E), \mathcal{W}^{\text{Norm}}_{u,v}, LC, explorationDepth$)
9. Initialize a sparse matrix $SM^0 \in \mathbb{Z}^{|V|\times 1}$ to zero
10. Initialize a list L of sparse matrices
11. Initialize a hashmap explanVal
12. $current = 0$
13. $i = 1$
14. $len = 0$
15. **while** $i <= explorationDepth$ **do** ▷ Stopping criteria
16. $SM^i = horizontalstack(SM^i, LC)$
17. $index = IndexID(SM^i)$
18. Initialize a sparse matrix $SM^i \in \mathbb{Z}^{|V|\times|index|}$ to zero
19. **for all** $(u, v) \in E$ **do**
20. $SM^i[id(u), index(SM^{i-1}[id(v), :])] += \mathcal{W}^{\text{Norm}}_{u,v}$
21. append $(id(v), i)$ to $explanVal[(id(u), index(SM^{i-1}[id(v), :]) + current)]$
22. **end for**
23. append SM^i to L
24. **if** $len = |index|$ **then**
25. break
26. **else**
27. append SM^i to L
28. $len = |index|$
29. $current += len$
30. $i = i + 1$
31. **end if**
32. **end while**
33. $SM^{tot} = horizontalStack(L)$
34. **return** $SM, explanVal$
35. **end function**

reduced using SelectKBest, scores were mapped back to the feature space generated prior to the reduction, and the scores were also decomposed from the graph representation back into its node constituents. This generates a sparse matrix for the attribute scores, and Algorithm 2 is used to obtain explanations only for edges that were relevant to the classification. For each node u, we reverse the truncatedSVD performed in LabeledSparseStruct to retrieve a list of feature scores equal to the size of the original node embedding (Line 9). This means that the value found at any index in the score matrix is the corresponding score for a feature in that same index in the node embedding. We iterate through each of these features, obtaining the list of all the neighbour nodes of u that contributed to that structural feature (Line 11). To each of the neighbours v in that list, the score is allocated according to the proportion of weight that its edge contributed to the original embedding ((Line 16). If the first level iteration is being explored, the resulting *individualScore* can simply be added to the *edgeScore* for that edge (u, v). However, recall that each iteration represents a depth of exploration of each node's neighbourhood, resulting in an updated structural representation of that node. Therefore, if we consider scores that result from later iterations, that feature

must be 'unfolded' to each of the structures attributed to it, and the score applied to each corresponding edge. To accomplish this, we gather every edge in the k-hop neighbourhood of the neighbour v, where k is the levelExplored for the feature. This represents every edge that makes up the known structure of neighbour v at a depth of exploration k. Each of these edges were considered in the parsing of node u through its edge with neighbour v, so each of their *edgeScores* is incremented by the value *individualScore*. It becomes clear that after all features for every node are input to the algorithm, each edge's score reflects its level of participation in the overlapping structures of every node's local neighbourhood.

Algorithm 2 SparseStructExplanation algorithm

function SPARSESTRUCTEXPLANATION($\mathbf{G} = (V, E)$)
  initialize a hashmap *edgeScores*
  $M, explanVal, svdComponents = SparseStruct(\mathbf{G} = (V, E), \mathcal{W}, L, explorationDepth, k)$
  EXPLAINEDGE($\mathbf{G}, explanVal, svdComponents, score$)
  return *edgeScores*
end function
function EXPLAINEDGE($\mathbf{G}, explanVal, svdComponents, score$)
  for all $(u) \in V$ **do**
    $scores = score \times svdComponents$
    for all $i \in n$ **do**
      $(nbr(u)_i, levelExplored_i) = explanVal[(id(u), i)]$
      for all $(v, levelExplored) \in (nbr(u)_i, levelExplored)$ **do**
        if $(u, v) \notin edgeScores$ **then**
          $edgeScores[(u, v)] = 0$
        end if
        $individualScore = \frac{score[i]}{\mathcal{W}^{\text{Norm}}_{u,v} / \mathcal{W}^{\text{Norm}}_{tot} \in (nbr(u)_i)}$
        if $levelExplored_i = 0$ **then**
          $edgeScores[(u, v)] edgeScores[(u, v)] + individualScore$
        else
          for all $(x, y) \in knbr(v, levelExplored)$ **do**
            $edgeScores[(x, y)] = edgeScores[(x, y)] + individualScore$
          end for
        end if
      end for
    end for
  end for
  return *edgeScores*
end function

4. Results

4.1 *Experiments*

LabeledSparseStruct was used to generate node embeddings that incorporated normalized edge weights and node labels for the dataset described above. Node labels refer to the attack characteristics outlined in the dataset description, while edge weights are proportional to the frequency of observed characteristics in the measured time period. Data were split into a 75:25 train-test set maintaining the integrity of individual subgraphs (terror groups and time-stamps). Two different splits were performed, and both maintained a chronological order between attacks (The earliest 75% of time-stamps were used for training, and the latest 25%). The first one used weekly time-stamps, the second instead used bi-weekly time-stamps for

TABLE 2 *Graph classification performance on the dataset. Means are presented plus/minus standard deviation. SparseStruct refers to the original algorithm in which no labels or edge weights are included. For all LabeledSparseStruct analyses, both labels and edge weights are included. LabeledSparseStruct 1 wk groups attacks in a weekly interval, while LabeledSparseStruct 2 wk groups is the dataset where attacks are grouped in a biweekly interval*

Model	Accuracy%	F1 score%	Recall%	Precision%
GIN (no weights/labels)	43.27 ± 5.51	35.72 ± 4.61	43.59 ± 4.57	32.04 ± 4.88
GIN (weights+node labels)	60.93 ± 5.46	60.40 ± 5.41	61.32 ± 5.38	61.21 ± 5.55
GCN (no weights/labels)	25.88 ± 4.89	10.22 ± 1.55	25.00 ± 0.00	6.47 ± 1.22
GCN (weights+node labels)	72.99 ± 5.07	65.29 ± 13.40	67.32 ± 11.76	68.37 ± 13.78
SparseStruct	53.39 ± 1.07	53.07 ± 1.06	53.84 ± 1.08	52.78 ± 1.08
LabeledSparseStruct 1 wk	75.38 ± 0.74	74.95 ± 0.77	75.50 ± 0.75	76.11 ± 0.75
LabeledSparseStruct 2 wk	74.96 ± 1.20	74.51 ± 1.20	75.03 ± 1.20	75.52 ± 1.10

comparative purposes, following previous studies that found that micro-cycles of violence for many terrorist organizations follow bi-weekly time periods [57–59]. SelectKBest was used to reduce the number of features; then the ExtraTrees Classifier was used for graph classification. Performances are measured as the mean of 10 KBest transformations each run on 10 ExtraTrees Classifiers. For comparison, embeddings were generated from LabeledSparseStruct without node labels or edge weights to demonstrate the improvement on classification performance. Hyperparameters for the TruncatedSVD, ExtraTrees Classifier and SelectKBest were determined empirically to optimize accuracy. The TruncatedSVD reduced each node's embedding to 17 features, which were then concatenated into a graph representation. This resulted in a sparse matrix with 629 possible features (17×37 possible node types). The list was further reduced using the SelectKBest method, to 130 features per graph.

For comparison, graph classifications were performed on the dataset using a GIN and a GCN. Five thousand epochs were run with a node representation size of 100, and the metrics were averaged over all after a settling period, from 1000 to 5000. Single and double-layer neural networks were tested, and results for the best-performing models are presented below.

Table 2 shows that our modifications to the SparseStruct algorithm increased the performance on the graph classification task by over 20 percentage points in terms of accuracy. Indeed, the performance of all models was significantly increased by the addition of node labels. The incorporation of both node labels and edge weights in the LabeledSparseStruct algorithm has resulted in a model that exceeds the performance of the graph neural network models. Similar performance was observed when attacks were grouped in 2-week time intervals. The LabeledSparseStruct with 1-week time-stamps has on average an accuracy that is 0.42 percentage points higher than the second-best model, which is the LabeledSparseStruct with bi-weekly time-stamps, 2.39 percentage points more than GCN with weights and node labels and 14.45 percentage points more than GIN with weights and labels. In terms of F1 score, compared to the other three methods, LabeledSparseStruct with weekly time-stamps achieves results that are 0.44, 9.66 and 14.55 percentage points higher, respectively. The difference in terms of performance remains similar in recall (+0.47, +8.18 and +14.18) and precision (+0.59, +7.74 and +14.9) compared to the other three best alternatives.

These results are relevant in two different directions. First, from an algorithmic perspective, they demonstrate the better performance of LabeledSparseStruct against other competing approaches. Given

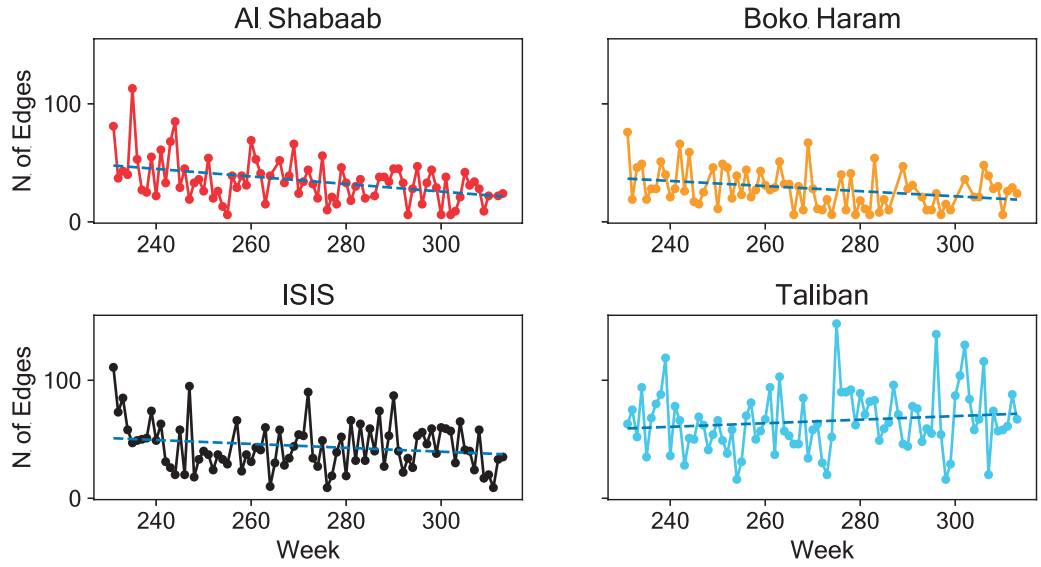

FIG. 4. Trend in weekly number of edges in each graph **G**—per each group. Dashed lines in each subplot represent polynomial fits describing the overall trend in the number of edges over the period under consideration.

the nature of the problem, which is challenging hard as these groups acted in the same historical period, and considering the relatively small amount of data at our disposal to avoid temporal separability, the outcomes achieved across the four metrics are promising.

From a theoretical point of view, the results empirically demonstrate graph-level differences in the representation of terrorist events, suggesting a sort of geometrical shape of violence characterizing each of the groups. In line with our initial intuition, terrorist groups can also be distinguished in terms of the hidden connections between features of temporally close events. Additionally, it is relevant to notice that, despite the chosen algorithmic approach, better performances have been achieved by considering also node labels and weights. This means that not only geometry is important but that also access to the specific characteristics of the elements in the graphs is critical. In fact, while the topology and geometry of violence have a role in describing distinct actions by distinct groups, it is not sufficient to rely on them alone to fully capture the complexity of terrorist behaviours. Similar topologies may be very different in terms of the features represented in each graph and the strength of the association between distinct nodes.

4.2 *Explaining terrorist signatures*

4.2.1 *Operational heterogeneity and temporal trends* Before delving into the explanation results at the dyadic level to identify terrorist signatures, we review the evolving nature of each group's operational graphs in two ways. First, in terms of trends in the number of edges composing each of the graphs (Fig. 4). Second, in terms of node heterogeneity, as a complementary measure of repertoire heterogeneity (Fig. 5). Node heterogeneity is calculated as the ratio between the total number of unique nodes in each weekly graph **G** per each group and the universe of unique nodes present at least in one graph in the test set. This measure aims at quantifying how diverse are the attacks in a given week, for a given terrorist organization.

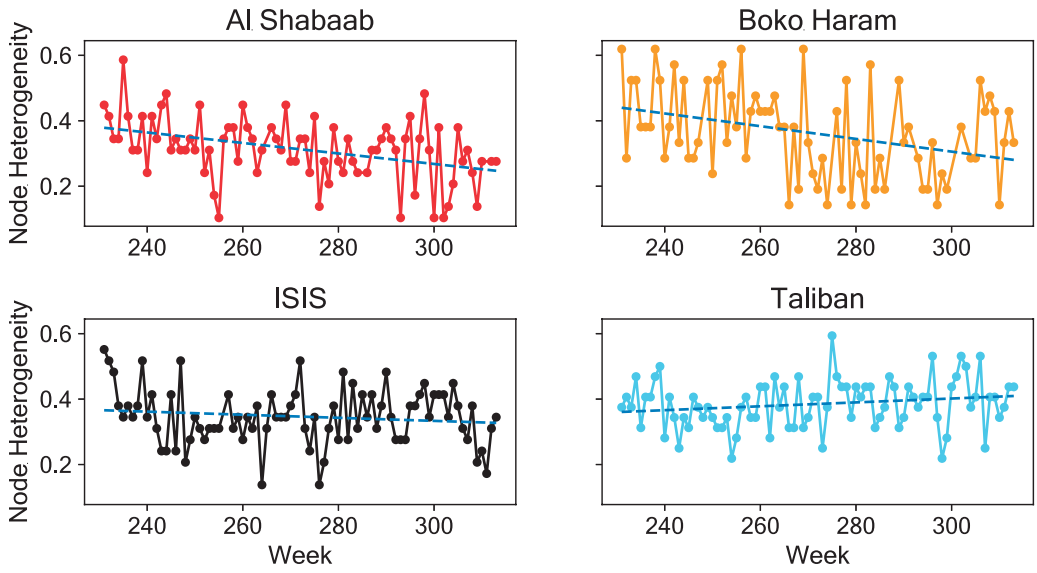

FIG. 5. Trend in weekly node heterogeneity in each graph **G**—per each group. Node heterogeneity is defined as the ratio between the number of unique nodes at each time stamp and the total number of unique nodes over the period under consideration (in the test set). Dashed lines in each subplot represent polynomial fits describing the overall trend in the number of edges over the period under consideration.

The plot reveals three sets of findings. First, the graph structures across our analysed organizations considerably vary from one group to the other in terms of number of edges. Boko Haram is the jihadist group with the overall lower number of edges, as demonstrated by the scale of the y-axis in the dedicated subplot. Conversely, the Taliban stands out compared to the others, indicating that a higher number of connections generally characterizes the graphs associated with the Afghan organization.

Second, and more importantly, the Taliban is the only group showing an increasing trend in the number of edges over time. This means that, in our test test—which maps the most recent bulk of activity spanning approximately the period that goes from June 2017 to the end of 2018—the number of edges in the Taliban operational graphs has continued to grow, indicating increasing diversity in feature connectivity. It should be kept in mind that a growing trend does not mean increasing activity, as we are here considering feature compositions of attacks rather than quantity of events. To exemplify, one group may plot 100 attacks only using one weapon, one tactic and choosing one target, and in that case the operational graph will only be composed of three feature nodes and at most three edges. Alternatively, another group may plot 10 attacks using 6 different weapons, 5 different tactics against 10 different targets. Trivially, in the latter case, the number of edges will be higher in spite of a lower level of activity compared to the former case.

Third, in terms of node heterogeneity, instead, the four groups are much more comparable, but the Taliban remains the only group displaying a slightly increasing trend. This underscores that the Taliban have widened their repertoire in the last period of the analyses, while the other three have consistently clustered their behaviours around a more restricted set of tactics, weapons and targets.

The Taliban, as anticipated, is the only one presenting increased complexity in behavioural heterogeneity. All the other groups, conversely, are characterized by diminishing operational combinations.

ISIS displays a less steep decrease, Al Shabaab and Boko Haram instead reveal instead significant edge reductions over the period under consideration.

These two results suggest macro differences exist among groups, even without focusing on micro-level qualitative distinctions in terms of behavioural characteristics.

4.2.2 *Most recurring edges* To further describe the distinctive characteristics of the four considered groups, we focus on the edge-wise composition of the temporal meta-graphs in each group's test set. In this subsection, we specifically analyse the distribution of the 10 most common dyadic edges. This allows shedding light on the most recurring operational preferences for each group, along with an assessment of the prevalence of such preferences. For each group, the 10 most common edges are visualized in Fig. 6: each edge's presence is measured in terms of the share out of the total number of time units, 83, in the test set. If a certain edge is always present in a given graph $\mathbf{G}_{\text{Test Set}}$, the share will be equal to 1.

Several indications emerge from Fig. 6. First, the 10 most recurring edges for the Taliban are all very close to the maximum of 1. This is peculiar, especially in comparison with Al Shabaab and Boko Haram, pointing out that there exist a consistent bulk of dyadic edges that always characterize Taliban operational graphs and, therefore, behaviours. The Taliban's preferences are not converging over a single dyad or very few of them but are, conversely, distributed over an extended set of edges, suggesting that the Taliban's behavioural stability is reflected in larger topological networks, rather than in consistent sub-graphs. In fact, while for Al Shabaab and Boko Haram the distribution of the most recurring edges entails variations in graph structures but stability in specific edges, the almost universal presence of the 10 most recurring edges for the Taliban demonstrates that they constitute the backbone of the operational graphs in the test set. The situation for ISIS is similar to the one commented for the Taliban.

A qualitative comparison of the edges in the four subplots allows one to further appreciate differences among these organizations. First, it is straightforward for the Taliban to note a strong tendency towards targeting police forces. This attitude is a unicum in the considered sample of terrorist groups. However, ISIS demonstrates to prefer soft targets and civilians in their decision-making of targeting choices, as highlighted by the high and diffused prevalence of private citizens and properties as a feature node in the reported most common edges. Al Shabaab and Boko Haram also share this distinctive characteristic.

Concerning weapons, three groups out of four (i.e. Al Shabaab, ISIS and the Taliban) have explosives as the preferred means for conducting attacks, and particularly the combination of two explosives. Boko Haram, conversely, prefers lighter weapon choices, such as firearms, although explosives are yet present in its most recurring edges. Interestingly, heterogeneous weapon selections (namely, edges between two distinct weapon types) do not characterize any of the recurring behavioural choices of the four groups.

In terms of tactics, they largely reflect the groups' preferences for explosives and firearms. Thus, bombing/explosion and armed assault are the most common tactics present in the core recurring structures of the organizations' operational graphs. Boko Haram behaviours, however, show uniqueness compared to the others given their peculiar recurring tendency of using facility and infrastructure attack tactics, which are present in almost 60% of the temporal graphs in the test set for the Nigerian group. This reflects a particular dimension of Boko Haram's confrontation strategy with Nigerian institutions and companies aimed at disrupting critical infrastructure, communication networks and legitimate economic sectors.

4.2.3 *Edge importance* Through SparseStructExplanation, we retrieve the edges that were important for the classification of the graphs into each terror group. Figure 7 displays the distributions in terms of feature importance scores over all the weeks in the test set, for each group. Importance scores can take both positive and negative values, with negative values indicating that a given edge, conditional on the

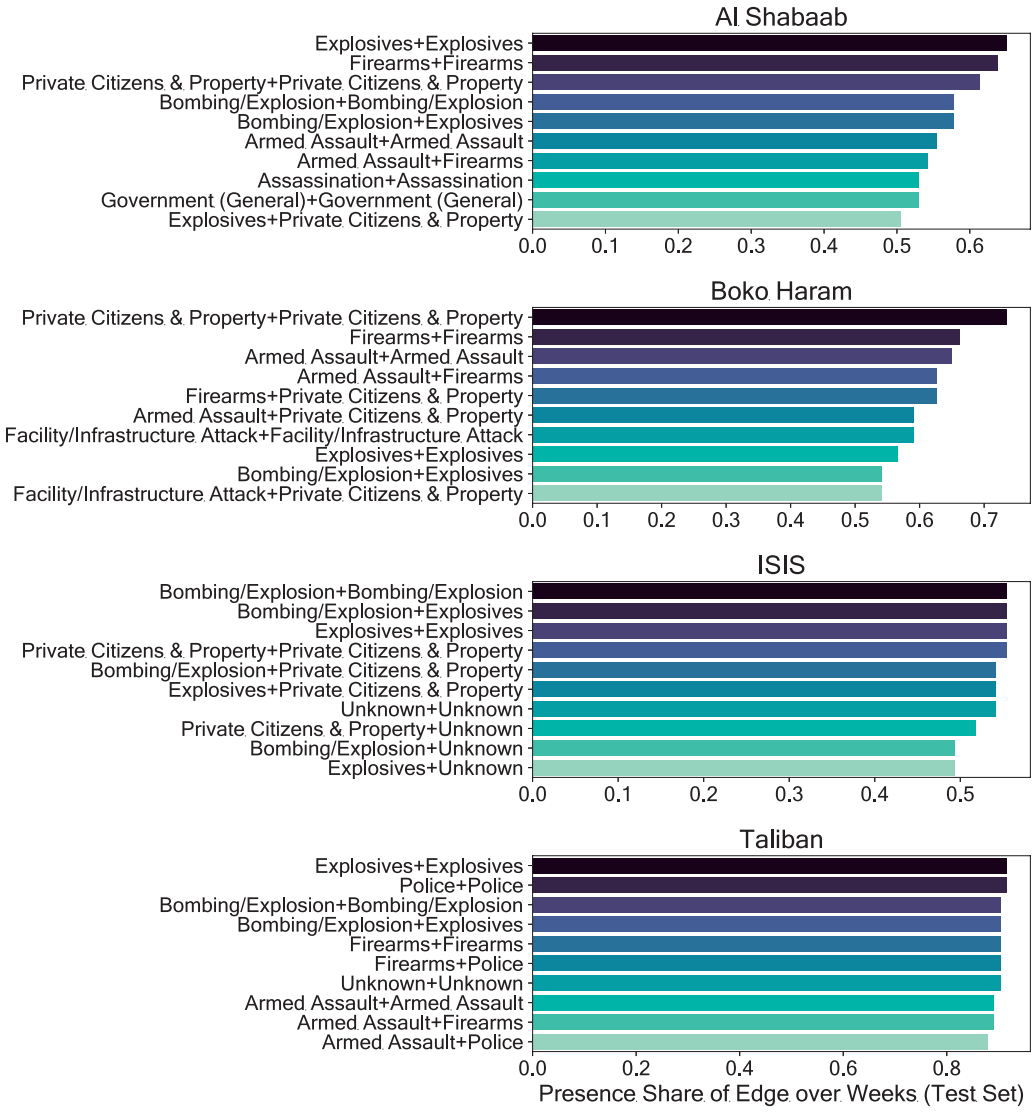

FIG. 6. Top 10 most recurring edges in the test set—per each group, measured in terms of presence share over number of weeks.

particular structure of graph **G** in which it is embedded, is detected to be anomalous. In other terms, when an importance score is lower than zero, the explanation algorithm indicates that the specific edge would have fit better within the operational topology of another group.

It should be noted that the ranking in Fig. 7 does not reflect the one seen in Fig. 6. In fact, edges are listed here based on their decreasing average importance, from the top to the bottom of each subplot. In other words, this means that if an edge is the most popular one in the test set does not mean that it is also the one that has higher importance in the explanation phase.

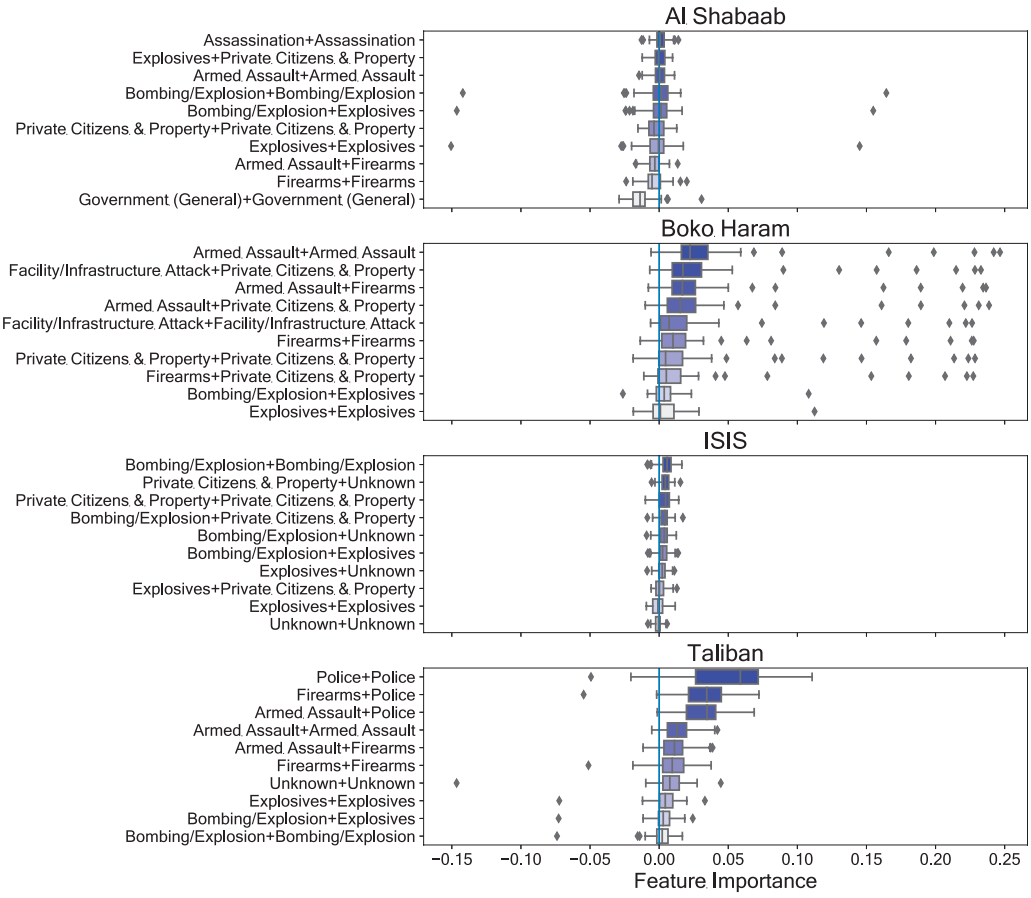

FIG. 7. Distribution of feature importance scores for the top 10 most recurring edges—per each group.

Interestingly, the highest average values are associated with the Taliban, especially concerning the top 4 edges within the 10 most recurring ones. This outcome further testifies to the strong homogeneity of behaviours of the Afghan group, characterized by recurring patterns. Particularly, edges connecting firearms weapons used against police, armed assault used again police, and police targeted in combination represent the most important Taliban's signatures, not only in terms of prevalence but also in terms of importance. Notwithstanding, these three relevant edges also report some considerable degree of variation suggesting that, while characterizing operations for most weeks in the test set, the overall topology of the operational graphs changes, slightly modifying signatures at the edge-level. This result may be explained by the considerable number of average links in each Taliban weekly meta-graph, as displayed in Fig. 4.

While the Taliban and ISIS exhibited similarities in terms of the high distributed prevalence of a restricted number of edges, the two groups display distinct situations in terms of distribution of feature importance scores. ISIS, in fact, does not have a straightforward subset of edges that definitely characterize its operations, compared to the Taliban, as magnitudes are lower and much more homogeneous across

the ten most common edges. The distributional ranges for the ten most recurrent edges are very similar, indicating a very low degree of operational variation in the weeks included in the test set for ISIS.

The situation for Boko Haram is the most peculiar one. Not only are the distributions considerably wider than those of Al Shabaab, ISIS and, on average, also the Taliban, but they also exhibit a considerable amount of outliers. At least one positive outlier is present in each of the Nigerian group's 10 most common operational edges. The magnitude of the highest importance scores is significantly different from the others. The diffused presence of outliers points toward a noticeable degree of evolution in behaviours and operations, making Boko Haram the group with the highest levels of variability. This finding likely reflects the varying strategies that the group has embraced over the last years considered in the analysis, in line with previous research depicting Boko Haram's tendency to modify tactics according to new strategies or in response to external factors [60–62]. Finally, besides the considerable presence and importance of edges involving the use of armed assault as a tactic, it is worth highlighting the significance of attacks involving facility/infrastructure attacks against private citizens and property, a highly distinctive feature of Boko Haram [63]. Facility and infrastructure attacks represent the only considerable anomalous edge across the four groups and our algorithmic approach recognized its relevance in the characterization of Boko Haram's operational graphs.

From a distributional standpoint, Al Shabaab manifests a feature importance situation that resembles the one commented for ISIS, with very restricted distribution having mean values very close to zero. The joint use of assassination tactics represents the edge with the highest average overall, another peculiarity compared to the other three groups, which did not report a tendency in using this particular tactic to carry out their attacks. Al Shabaab's increasing resort to assassination as a means of targeted terrorism was already highlighted by [64] observing data from 2007 to 2013. Most of the other links in the list are very common also to other groups, indicating a general cohesiveness in a relevant share of operational fabrics.

5. Discussion and conclusions

Social network analysis and network science in general have consistently become two valuable methodological approaches in the study of terrorism [21, 22, 65, 66]. The literature now counts dozens of studies analysing how terrorist groups are organized, evolve, adapt, and act in the perpetration of political violence. However, the vast majority of these studies mostly concentrated on networks of individuals. In this context, networks are used to map relationships between agents, such as groups' affiliates: links connecting two actors may describe communication connections, attack co-participation and other similar constructs.

These approaches have offered researchers considerable opportunities to expand our knowledge on how terrorist organizations work. Yet, the power of networks has remained almost unexplored when it comes to mapping other types of terrorist dynamics that do not regard individual-level connections. Although empirical research has now demonstrated that terrorist events cluster in space and time, scholars have only marginally investigated whether other types of less visible dependencies exist. In particular, we know very little about the possible interdependencies between event characteristics. We know that terrorist events often self-excite sequences of other attacks [4, 67, 68], and that temporal concentration also has a spatial component [57, 69, 70], but we have only recently started to investigate whether we can gain significant knowledge also from the study of connections among features describing temporally close attacks [18] or operationally similar actors [20]. Network science and computational modelling are, in this regard, crucial frameworks to investigate this possibility.

Motivated by the strategic theories of terrorism, this work has precisely sought to showcase the promise of geometrical representations of terrorist attacks in advancing our understanding of how terrorist

organizations operate. The strategic theories of terrorism posit that, in light of collective rationality, terrorist groups act strategically to reach specific objectives [45]. This strategical behaviour, in turn, generates patterned dynamics that, we hypothesized, can also be detected in the peculiar characteristics defining a group's history of attacks.

In the current work, we have specifically explored the topological characteristics of terrorist attacks of four major jihadist organizations using data obtained from the GTD [50]: Al Shabaab, Boko Haram, ISIS and the Taliban. These topological characteristics are investigated by creating operational meta-graphs representing connections among event features (i.e. tactics, weapons and targets) at the weekly level for the 2013–2018 period, accounting for more than 14,000 events.

Our computational approach has been organized into two phases. First, we were interested in verifying that graph-based differences existed between the groups, possibly indicating that different groups exhibit different behaviours also in terms of dependencies between event characteristics. Second, we aimed to capture each group's micro-level distinctive operational features to qualitatively understand the dyadic features that, beyond overall topology, help us describe each group's behaviours. This second phase sought to contribute to the increasing demand for interpretability in black-box machine learning models that are often extremely difficult to describe and understand [71, 72].

Concerning the first phase, we have proposed a modified version of SparseStruct [11], named LabeledSparseStruct, to correctly classify the group associated to each weekly operational meta-graph in a multiclass classification problem setting. SparseStruct is a fast and scalable structural representation learning approach that uses a sparse internal representation for each node in a given graph, preserving nodes' structural information. Unlike the original version, our modified algorithm incorporates node labels and weights to fully exploit the information portfolio offered by our operational meta-graphs. The performance of LabeledSparseStruct has been compared to five other algorithmic architectures: the original SparseStruct with no labels and weights, two versions of GINs (one with nodes labels and weights, and one without) and two versions of GCNs (again, with and without node labels and weights). LabeledSparseStruct with weekly time-stamps reached an accuracy performance more than 20 percentage points larger than the original SparseStruct algorithm, demonstrating the importance of weights and labels in describing terrorist behavioural and operational differences. Node labels and weights increased the performance of GINs and GCNs as well, but no approach obtained a better global performance than LabeledSparseStruct across the four considered metrics (i.e. Accuracy, F1 Score, Recall and Precision).

With regard to the second phase, we used LabeledSparseStruct—being the algorithms with the highest performance—to generate SparseStructExplanator, an approach to obtain information about the importance of dyad-level features in the prediction task. SparseStructExplanator is built upon the LIME Explainer [56] and allows to identify what links contribute more to the operational profile of the Islamist organizations in our sample. The results of our explanation technique reveal interesting differences across the groups. First, the Taliban—the only group that overall displayed an increase in the number of edges and node heterogeneity in the operational meta-graphs over time—reported a distinctive subset of edges, which appear to map recurring operations over time. Yet, we also documented a sizable degree of variation of importance scores, meaning that the Taliban's strong preference towards certain operations sometimes is coupled with alternative sets of behaviours. Second, the Taliban and ISIS show a certain similarity in their most recurring edges, although the distribution of the feature scores for such edges is very different between the two. The quantitative results for ISIS, in this case, qualitatively suggest a limited degree of variation and, in turn, a high degree of consistency of behaviours. Third, Boko Haram exhibits a peculiar situation. Outliers in the distribution of the feature importance scores of the 10 most common links point in the direction of a high degree of strategical evolution over the period

under consideration. Fourth, Al Shabaab feature importance scores distributions resemble those of ISIS, suggesting a comparable level of operational homogeneity between the two.

Our study fits in a recent line of inquiry that addressed terrorist behaviours from a relational perspective, with relationships represented in forms of operational meta-graphs [18, 19]. It has highlighted how representing streams of attacks through connections between event features provides relevant information on the distinctive characteristics associated with groups' different strategies. Our sample of compared groups revealed a substrate of topological characterizations which becomes even more interesting in light of the fact that all groups are associated with Islamist extremism. Not only historical, cultural, geographical, political differences as well as ideological nuances exist between these groups but also operational ones at various levels, as documented by both our analytic phases—the prediction and the explanation ones.

Albeit algorithmic results are encouraging, our work naturally comes with several limitations. First, we needed to use a relatively small amount of data to satisfy the condition of overlapping timeframes of activity, given the fact that the first attack recorded for ISIS available in the GTD dates back to 2013. More data would be useful to fully exploit the computational performance of LabeledSparseStruct which is specifically designed to handle vast amount of data efficiently.

Second, our sample of groups is interesting because it comprises groups that are all linked to the same broader category of Islamist extremism but do not include organizations that are overlapping also in terms of local geography. Each of the four acts in distinct countries or regions somehow limiting the practical relevance of our analysis beyond its mere research value.

Third, our explainability only considers dyadic features rather than expanding the focus to more complex sub-graphs that might distinguish groups even more sensibly. While the dyadic approach certainly offers more straightforward results to be interpreted, it may be limiting in situations where deeper topological differences are at play.

In spite of these limits, this work provided two sets of contributions. From a research standpoint, it generated insights on behavioural patterns that are not merely visible when considering aggregated data in traditional cross-sectional or spatio-temporal settings, calling for new systematic approaches that incorporate other layers of dependencies among events. The results of our classification experiments and the insights provided by the explanation phase demonstrated the existence of topological complex differences in the nature of attacks across different organizations. How these actors organize and plot terrorist events unravel the role interdependencies between features have in shaping their behavioural profiles.

In the future, our framework may also be used to study topological differences between different criminal groups or criminal behaviours, extending its applicability beyond terrorism research. The algorithmic approach here presented may reveal significant geometrical differences distinguishing different criminal organizations or gangs, enriching the set of network-oriented tools that are already available to researchers and law enforcement. Its applicability could cover both traditional social networks as well as knowledge graphs mapping, for instance, offenders' roles, skills or tasks.

From a practical counter-terrorism perspective, our computational approach may become helpful for engineering tools that can help analysts identify perpetrators of specific sets of attacks for which responsibility is uncertain or not known at all. Exploiting the relational perspective might be useful for determining perpetrators of unclaimed attacks and can facilitate counter-intervention and enhance security strategies for civilians and other sensible subjects. The contained computational cost of our models makes them particularly palatable as they do not require excessive time or resources to be computed. Further, our explainer supplies dyad-level information that can be embedded into risk assessment tools for intelligence monitoring. Scores mapping the importance of specific connections between features for a given group

might enhance security policies to disrupt such links. However, as also mentioned in the limitations, this type of policy contribution will require additional experiments in geographical contexts with overlapping actors beyond simple overlapping timeframes of activity to become effective. Future work will hence go in this direction.

Code Availability

The data and code used to conduct the current research are available for reproducibility purposes at https://github.com/janetlayne2/SparseStruct.

Funding

The National Science Foundation (NSF REU 1820685) and the Department of Excellence Initiative of the Italian Ministry of University and Research.

References

1. Weimann, G. & Brosius, H.-B. (1988) The predictability of international terrorism: a time-series analysis. *Terrorism*, **11**, 491–502.
2. Enders, W. & Sandler, T. (2000) Is transnational terrorism becoming more threatening?: A time-series investigation. *J. Confl. Resolut.*, **44**, 307–332.
3. Enders, W. & Sandler, T. (2002) Patterns of transnational terrorism, 1970-1999: alternative time-series estimates. *Int. Stud. Q.*, **46**, 145–165.
4. LaFree, G., Dugan, L., Xie, M. & Singh, P. (2012) Spatial and temporal patterns of terrorist attacks by ETA 1970 to 2007. *J. Quant. Criminol.*, **28**, 7–29.
5. Hsu, H. Y., Vásquez, B. E. & McDowall, D. (2018) A time-series analysis of terrorism: intervention, displacement, and diffusion of benefits. *Justice Q.*, **35**, 557–583.
6. Drake, C. J. M. (1998) Decision-making. *Terrorists' Target Selection* (C. J. M. Drake, ed.). London: Palgrave Macmillan UK, pp. 163–174.
7. Shapiro, J. N. (2012) Terrorist decision-making: insights from economics and political science. *Perspect. Terror.*, **6**, 5–20.
8. Kleinman, K. (2004) Kleinman et al. Respond to "Surveilling surveillance". *Am. J. Epidemiol.*, **159**, 228–228.
9. Clark, N. J. & Dixon, P. M. (2018) Modeling and estimation for self-exciting spatio-temporal models of terrorist activity. *Ann. Appl. Stat.*, **12**. doi: 10.1214/17-AOAS1112.
10. Findley, M. G. & Young, J. K. (2012) Terrorism and civil war: a spatial and temporal approach to a conceptual problem. *Perspect. Politics*, **10**, 285–305.
11. Serra, E., Joaristi, M. & Cuzzocrea, A. (2020) Large-scale sparse structural node representation. *2020 IEEE International Conference on Big Data (Big Data)*. Atlanta, GA, USA: IEEE, pp. 5247–5253.
12. Drake, C. J. M. (1998) Ideology. *Terrorists' Target Selection* (C. J. M. Drake, ed.). London: Palgrave Macmillan UK, pp. 16–34.
13. Asal, V. H., Rethemeyer, R. K., Anderson, I., Stein, A., Rizzo, J. & Rozea, M. (2009) The softest of targets: a study on terrorist target selection. *J. Appl. Security Res.*, **4**, 258–278.
14. Ahmed, R. (2018) Terrorist ideologies and target selection. *J. Appl. Security Res.*, **13**, 376–390.
15. Polo, S. M. T. & Gleditsch, K. S. (2016) Twisting arms and sending messages: terrorist tactics in civil war. *J. Peace Res.*, **53**, 815–829.
16. Asal, V. & Rethemeyer, R. K. (2008) The nature of the beast: organizational structures and the lethality of terrorist attacks. *J. Politics*, **70**, 437–449.
17. Daxecker, U. E. & Hess, M. L. (2013) Repression hurts: coercive government responses and the demise of terrorist campaigns. *Br. J. Political Sci.*, **43**, 559–577.

18. Campedelli, G. M., Bartulovic, M. & Carley, K. M. (2021) Learning future terrorist targets through temporal meta-graphs. *Sci. Rep.*, **11**, 8533.

19. Campedelli, G. M., Cruickshank, I. & Carley, K. M. (2019) A complex networks approach to find latent clusters of terrorist groups. *Appl. Netw. Sci.*, **4**, 1–22.

20. Campedelli, G. M., Cruickshank, I. J. & Carley, K. M. (2021) Multi-modal networks reveal patterns of operational similarity of terrorist organizations. *Terror. Political Violence*. doi: 10.1080/09546553.2021.2003785.

21. Asal, V. & Rethemeyer, R. K. (2006) Researching terrorist networks. *J. Security Educ.*, **1**, 65–74.

22. Perliger, A. & Pedahzur, A. (2011) Social network analysis in the study of terrorism and political violence. *PS: Political Sci. Politics*, **44**, 45–50.

23. Morselli, C. (2013) *Crime and Networks*. New York: Routledge, 2013.

24. Calderoni, F., Catanese, S., De Meo, P., Ficara, A. & Fiumara, G. (2020) Robust link prediction in criminal networks: a case study of the Sicilian Mafia. *Expert Syst. Appl.*, **161**, 113666.

25. Castellano, N. G., Cerqueti, R. & Franceschetti, B. M. (2021) Evaluating risks-based communities of Mafia companies: a complex networks perspective. *Rev. Quant. Finan. Account.*, **57**, 1463–1486.

26. Tumminello, M., Petruzzella, F., Ferrara, C. & Miccichè, S. (2021) Anagraphical relationships and crime specialization within Cosa Nostra. *Soc. Netw.*, **64**, 29–41.

27. Ribeiro, H. V., Alves, L. G. A., Martins, A. F., Lenzi, E. K. & Perc, M. (2018) The dynamical structure of political corruption networks. *J. Complex Netw.*, **6**, 989–1003.

28. Luna-Pla, I. & Nicolás-Carlock, J. R. (2020) Corruption and complexity: a scientific framework for the analysis of corruption networks. *Appl. Netw. Sci.*, **5**, 13.

29. McIllwain, J. S. (1999) Organized crime: a social network approach. *Crime, Law Soc. Change*, **32**, 301–323.

30. Morselli, C., Giguère, C. & Petit, K. (2007) The efficiency/security trade-off in criminal networks. *Soc. Netw.*, **29**, 143–153.

31. Schmidle, R. E. (2010) Positioning theory and terrorist networks. *J. Theory Soc. Behav.*, **40**, 65–78.

32. Bakker, R. M., Raab, J. & Milward, H. B. (2012) A preliminary theory of dark network resilience: a preliminary theory of dark network resilience. *J. Policy Anal. Manag.*, **31**, 33–62.

33. Krebs, V. (2002) Uncloaking terrorist networks. *First Monday*. https://firstmonday.org/ojs/index.php/fm/article/view/941/863.

34. Medina, R. M. (2014) Social network analysis: a case study of the Islamist terrorist network. *Security J.*, **27**, 97–121.

35. de Bie, J. L., de Poot, C. J., Freilich, J. D. & Chermak, S. M. (2017) Changing organizational structures of jihadist networks in the Netherlands. *Soc. Netw.*, **48**, 270–283.

36. Ouellet, M., Bouchard, M. & Hart, M. (2017) Criminal collaboration and risk: the drivers of Al Qaeda's network structure before and after 9/11. *Soc. Netw.*, **51**, 171–177.

37. Asal, V. H., Park, H. H., Rethemeyer, R. K. & Ackerman, G. (2016) With friends like these ... why terrorist organizations ally. *Int. Public Manag. J.*, **19**, 1–30.

38. Benigni, M. C., Joseph, K. & Carley, K. M. (2017) Online extremism and the communities that sustain it: detecting the ISIS supporting community on Twitter. *PLoS One*, **12**, e0181405.

39. Gialampoukidis, I., Kalpakis, G., Tsikrika, T., Papadopoulos, S., Vrochidis, S., & Kompatsiaris, I. (2017) Detection of terrorism-related twitter communities using centrality scores. *Proceedings of the 2nd International Workshop on Multimedia Forensics and Security*, MFSec '17. New York, NY, USA: Association for Computing Machinery, pp. 21–25.

40. Ferrara, E. (2017) Contagion dynamics of extremist propaganda in social networks. *Inf. Sci.*, **418-419**, 1–12.

41. Schelling, T. C. (1980) *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

42. Waterman, H. (1981) Reasons and reason: collective political activity in comparative and historical perspective. *World Politics*, **33**, 554–589.

43. Crenshaw, M. (1990) The logic of terrorism: terrorist behavior as a product of strategic choice. (W. Reich, ed.). *Origins of Terrorism: Psychologies, Ideologies, Theologies, States of Mind*. Washington, DC: Woodrow Wilson Center Press with Johns Hopkins University Press.

44. Sandler, T. & Lapan, H. E. (1988) The calculus of dissent: an analysis of terrorists' choice of targets. *Synthese*, **76**, 245–261.
45. McCormick, G. H. (2003) Terrorist decision making. *Annu. Rev. Political Sci.*, **6**, 473–507.
46. Hamilton, W. L. (2020) Graph representation learning. *Synth. Lect. Artif. Intell. Mach. Learn.*, **14**, 1–159.
47. Hamilton, W. L., Ying, R. & Leskovec, J. (2018) Representation learning on graphs: methods and applications. *arXiv:1709.05584 [cs]*. http://arxiv.org/abs/1709.05584.
48. Kipf, T. N. & Welling, M. (2017) Semi-supervised classification with graph convolutional networks. *arXiv:1609.02907 [cs, stat]*. http://arxiv.org/abs/1609.02907.
49. Xu, K., Hu, W., Leskovec, J. & Jegelka, S. (2018) How powerful are graph neural networks? https://arxiv.org/abs/1810.00826v3.
50. LaFree, G. & Dugan, L. (2007) Introducing the global terrorism database. *Terror. Political Violence*, **19**, 181–204.
51. Weisfeiler, B. & Lehman, A. A. (1968) A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsia*, **2**, 12–16.
52. Hamilton, W. L., Ying, R. & Leskovec, J. (2017) Inductive representation learning on large graphs. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17. Red Hook, NY, USA: Curran Associates Inc., pp. 1025–1035.
53. Grover, A. & Leskovec, J. (2016) node2vec: scalable feature learning for networks. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco CA, USA: ACM, pp. 855–864.
54. Ribeiro, L. F. R., Saverese, P. H. P. & Figueiredo, D. R. (2017) Learning node representations from structural identity. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Halifax, NS, Canada: ACM, pp. 385–394.
55. Donnat, C., Zitnik, M., Hallac, D. & Leskovec, J. (2018) Learning structural node embeddings via diffusion wavelets. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. London, UK: ACM, pp. 1320–1329.
56. Ribeiro, M. T., Singh, S. & Guestrin, C. (2016) "Why should I trust you?": explaining the predictions of any classifier. *arXiv:1602.04938 [cs, stat]*. http://arxiv.org/abs/1602.04938.
57. Behlendorf, B., LaFree, G. & Legault, R. (2012) Microcycles of violence: evidence from terrorist attacks by ETA and the FMLN. *J. Quant. Criminol.*, **28**, 49–75.
58. Duru, H., Onat, I., Akyuz, K. & Akbas, H. (2021) Microcycles of terrorist violence in Turkey: a spatio-temporal analysis of the PKK attacks. *Asian J. Criminol.*, **16**, 235–256.
59. Rieber-Mohn, J. H. & Tripathi, K. (2021) An investigation into microcycles of violence by the Taliban. *Security J.*, **34**, 126–147.
60. Zenn, J. & Pearson, E. (2014) Women, Gender and the evolving tactics of Boko Haram. *Contemp. Voices St Andrews J. Int. Relat.*, **5**, 46–57.
61. Weeraratne, S. (2017) Theorizing the Expansion of the Boko Haram Insurgency in Nigeria. *Terror. Political Violence*, **29**, 610–634.
62. Voronkova, A. (2017) Boko Haram's cross-border attacks: tactical manoeuvring to mitigate weakness. https://www.iiss.org/blogs/analysis/2017/07/boko-haram-tactics.
63. Onuoha, F. C. (2013) The costs of Boko Haram attacks on critical telecommunication infrastructure in Nigeria. https://www.e-ir.info/2013/11/03/the-costs-of-boko-haram-attacks-on-critical-telecommunication-infrastructure-in-nigeria/.
64. Mueller, J. C. (2018) The evolution of political violence: the case of Somalia's Al-Shabaab. *Terror. Political Violence*, **30**, 116–141.
65. Ressler, S. (2006) Social network analysis as an approach to combat terrorism: past, present, and future research. Homeland Security Affairs, 2. https://www.hsaj.org/articles/171.
66. Bouchard, M. (ed.) (2017) *Social Networks, Terrorism and Counter-Terrorism*, 1st edn. London: Routledge.
67. White, G., Porter, M. D. & Mazerolle, L. (2013) Terrorism risk, resilience and volatility: a comparison of terrorism patterns in three Southeast Asian countries. *J. Quant. Criminol.*, **29**, 295–320.

68. Tench, S., Fry, H. & Gill, P. (2016) Spatio-temporal patterns of IED usage by the Provisional Irish Republican Army. *Eur. J. Appl. Math.*, **27**, 377–402.

69. Braithwaite, A. & Li, Q. (2007) Transnational terrorism hot spots: identification and impact evaluation. *Conflict Manag. Peace Sci.*, **24**, 281–296.

70. Neumayer, E. & Plümper, T. (2010) Galton's problem and contagion in international terrorism along civilizational lines. *Conflict Manag. Peace Sci.*, **27**, 308–325.

71. Doshi-Velez, F. & Kim, B. (2017) Towards a rigorous science of interpretable machine learning. *arXiv:1702.08608 [cs, stat]*. http://arxiv.org/abs/1702.08608.

72. Molnar, C., G. Casalicchio, G. & Bischl, B. (2020) Interpretable machine learning — a brief history, state-of-the-art and challenges. *ECML PKDD 2020 Workshops*, *Communications in Computer and Information Science* (I. Koprinska, M. Kamp, A. Appice, C. Loglisci, L. Antonie, A. Zimmermann, R. Guidotti, Ö. Özgöbek, R. P. Ribeiro, R. Gavaldà, J. Gama, L. Adilova, Y. Krishnamurthy, P. M. Ferreira, D. Malerba, I. Medeiros, M. Ceci, G. Manco, E. Masciari, Z. W. Ras, P. Christen, E. Ntoutsi, E. Schubert, A. Zimek, A. Monreale, P. Biecek, S. Rinzivillo, B. Kille, A. Lommatzsch & J Atle Gulla, eds). Cham: Springer International Publishing, pp. 417–431.