Which Centrality Metric for Which Terrorist Network Topology?

Imen Hamed[1(✉)], Malika Charrad[1,2], and Narjès Bellamine Ben Saoud[1]

[1] Univ. Manouba, ENSI, RIADI LR99ES26,
Campus Universitaire Manouba, 2010 Manouba, Tunisia
hamedimen@gmail.com, malika.charrad@riadi.rnu.tn,
narjes.bellamine@ensi.rnu.tn

[2] ISIMed, Université de Gabes, Gabès, Tunisia

Abstract. Recently, an exponential growth in the use of social network analysis (SNA) tools has been witnessed. SNA offers quantitative measures known as centralities which allow the identification of important nodes in a given network. In fact, determining such nodes in terrorist networks is a way to destabilize these cells and prevent their criminal activities. Identifying key players is highly dependent on structural characteristics of nodes. Therefore, many approaches rely on centrality metrics to propose various disruption strategies. Indeed, knowledge of these measures helps in revealing vulnerabilities of terrorist networks and may have important implications for investigations. It is debatable how to choose the suitable centrality measure that helps effectively to destabilize the terrorist network. In this paper, we aim to answer this question. We first provide an analytical study where we identify 6 topologies of terrorist networks and discuss the appropriate metrics per topology. Secondly, we provide the performed experimental analysis on five data sets (with 5 different topologies) to prove our analytical conclusions.

Keywords: Terrorism · Network topology · Centrality metric

1 Introduction

Terror is the calculated use of violence or the threat of violence to attain political or religious ideological goals through intimidation, coercion, or instilling fear [21]. This strategic and tactic crime is considered as a very complex phenomenon due to its secretive nature. Different strategies have been proposed to reveal the secrecy of terrorist networks. Social Network Analysis (SNA) is prominent among them. This latter consists in transforming the set of terrorists into network structures where the nodes represent attackers and the links are the connections between them. SNA provides deeper insights about the nodes and their interactions. Different works applying SNA on terrorist attacks were proposed such as modeling dynamic covert networks [14], analyzing links between individuals [19], subgroup detection and key players identification [14]. In this paper, we focus mainly on key players identification. In fact, the disruption of terrorist cells requires the isolation of important nodes. To do so, it is fundamental to measure centrality metrics.

P. Diaz et al. (Eds.): ISCRAM-med 2016, LNBIP 265, pp. 195–208, 2016.
DOI: 10.1007/978-3-319-47093-1_17

These latter characterize a node's position in the graph. Centrality metrics have been successfully involved in terrorist networks destabilization methods. Thus, there is certainly a need for an accurate choice of centrality indices to effectively identify influential nodes in networks. The goal of this paper is not to perform a traditional social network analysis but rather to evaluate the validity of different centrality measures according to the topology of the network by conducting an empirical study on real-world terrorist data sets. Throughout this paper, we first present background about terrorist networks properties and their different topologies (Sect. 2). Then we discuss proposed destabilization approaches using centrality metrics (Sect. 3) and we review consequently the measures used in this purpose (Sect. 4). Then, we provide a matching between different terrorist networks topology and different centrality metrics (Sect. 5). To prove our theoretical analysis, we distinguish five different terrorist data sets (Sect. 5) where we apply commonly used centralities on them and compare the results (Sect. 6). This is followed by conclusion and future work section.

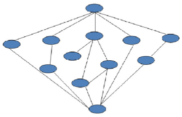
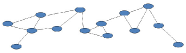
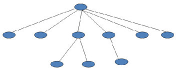
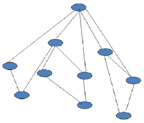
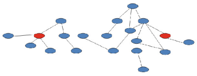
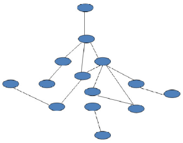
2 Terrorist Networks

Terrorist networks are described as amorphous, invisible, resilient, dispersed... [6]. So, it is difficult to visualize them. The problem with such networks is that it is highly covert and most of the time it is incomplete. Three main problems encounter the terrorist cells analysis [7]:

- Incompleteness: Some nodes or links may be missing in the investigated network.
- Fuzzy boundaries: The relationship between criminals is always unclear.
- Dynamic: These networks are changing continuously.

Regarding these problems, modeling terrorist networks becomes a hard task. In fact, gathering data is challenging if the terrorists are not arrested. And even if the information exists, mapping ties between individuals is difficult. Indeed, secrecy is a prime concern in these cells. Covert networks often do not behave like normal social networks [6]. Ties between individuals are invisible. Besides, it is not possible to cluster nodes based on these ties because they are not their real connections but rather intermediaries between them and the other actors. The authors in [11] claim that these networks are separated by larger than normal degrees of distance between their participants which adds the possibility of mapping them as distributed networks. Another important property of terrorist networks is that they are purposive [11], they differ from each other by their formal properties. Thus, modeling terrorist cells becomes of great interest. In fact, the authors in [12] distinguished different models of terrorist organizations based on many factors. Structural properties of these dark cells differ according to their needs to be hidden, to get protected and to maximize their profit. Basically, there are 6 categories of terrorist networks structures. The following table summarizes these different topologies providing the correspondent shape of each category and its properties also proposing the best destabilization strategy (Table 1).

Table 1. Terrorist networks topologies, related properties and suggested disruption strategies

Topology	Shape	Properties	Disruption strategy
Corporate based network E.g. Irish Republican Army (IRA)		The network is composed of subgroups with different goals (propaganda subgroup, finance subgroup...)	Detecting the leaders of the subgroups as important nodes in the network
Politburo based network E.g. Red Army Faction (RAF)		Central committee which decides all the strategy of the network	Detecting the members of the central committee as key players in the network
Shura based network E.g. Turkish Al Qaida November 2003		members are of equal importance	Detecting the leader who has more connections
Multi-cell based network E.g. 9/11 attack data set		Cells are connected with key players	Detecting the nodes connecting the cells
Brokerage based network E.g. Ergenekon data set		The brokerage members are fully trusted by the leaders of the cells	Detecting the brokerage members as important nodes
Lonely wolf based network		The wolf plans, supplies and attacks in the hand of one terrorist	Detecting the wolf member as important node

3 Related Work

A branch of centrality metrics has been proposed to study the terrorist networks. The most used centrality indices are: degree, betweenness, closeness and eigenvector centralities [8–10]. The degree centrality is the number of connections a node holds. Betweenness of a node measures the number of shortest paths passing through this node while the closeness is its inverse. The eigenvector retrieves the node that allows the maximum of flow to pass through it. These traditional measures have been incorporated in various terrorist networks disruption strategies. Once these nodes are recognized and removed, it becomes easier to destabilize the network. The authors in [8] propose an algorithm that relies on three centrality measures: degree centrality, closeness centrality and betweenness centrality to retrieve the financial manager who is considered as the most important node in the network to be isolated. The same metrics were used in addition to the eigenvector centrality in [9,10] to deduce the hierarchy of the network and recognize then the most influential nodes. The researchers in [5] introduce a new metric "influence index" to detect influential nodes in the terrorist network. This measure relies basically on the shortest paths between nodes and the rule of influence which consists in three degree of influence. i.e. a node is influenced by other nodes that lie at three degree of separation but not by those beyond. Detecting the nodes with highest importance in the studied network has been the goal of different destabilization approaches. The works proposed in [1] address the issue of node's global importance in the network retrieved by traditional betweenness centrality. The authors aim to find the important nodes to a given node. The dependency of a node on other nodes is measured by the importance of these nodes and the trust between these nodes. So, the authors propose the reliance measure as a new metric to measure the importance and the trust and identify then the important nodes to a given node. The authors in [15] propose to use a recently developed metric to identify influential nodes namely the percolation centrality. It has been developed in the past to identify important nodes in the flow of information, spreading rumors or contagious diseases in a network. The authors apply this measure to the scenario of terrorist networks to retrieve the information spreaders.

All these works are considered as qualitative approaches and do not provide any quantitative analysis of terrorist networks. Furthermore, these approaches assess the importance of a node by focusing on the role played by this node but fails to capture any positive or negative synergy between different groups. Considering these limits, the authors in [2] propose to consider the terrorist network as a coalition and to adopt centralities proposed in game theory such as Meyrson value and Shapely value based centralities [3,4]. These latter are a weighted average quantifying a node's marginal contribution to the coalition. Hence, it becomes possible to quantitatively identify important nodes and the synergy between them. (e.g. the bomb expert, the funding terrorists).

Several approaches were proposed to defeat terrorist cells using a branch of centrality metrics. However, it still lacks a methodology or an approach that guides the choice of the correspondent centrality metric to effectively disrupt the terrorist cells. Nevertheless, the works in [12] may be considered as a first step towards this approach. The authors present a classification of terrorist networks according to their ideology and characteristics that affect their shape. Accordingly, six categories of terrorist network topologies with different characteristics have emerged. The authors also give an example of real world data set for each category. In this paper, we aim to pursue the works in [12] and propose for each category the correspondent centrality metric in the disruption strategy. We provide experiments to justify our choices.

4 Centrality Metrics

4.1 Preliminaries

We design the terrorist graph as G. G consists of a pair (N,E) where N is the set of nodes and E is the set of edges that connect different nodes.

An edge e_{ij} represents opportunities for flow between vertices i and j. A path between two nodes is the set of edges connecting those two nodes. Once this set is minimized, the path is called the shortest path. This latter may also be called the geodesic distance between given nodes. The Adjacency matrix, $A \in M_{nn}(\Re)$, of network G is defined such that each matrix element, a_{ij}, indicates if G contains an edge e_{ij} connecting vertex v_j to v_i [19].

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge connecting } v_i \text{ to } v_j \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

4.2 Centrality Metrics in Terrorist Network Destabilization

As it is presented in related work section, terrorist networks destabilization rely on different centrality metrics mainly degree, betweenness, closeness and eigenvector centrality.

The degree centrality [13] is used to measure the number of connections each node holds in the network.i.e. the number of neighbors for a given node. For an oriented graph, the degree centrality is composed of indegree and outdegree. Accordingly, indegree is a count of the number of ties directed to the node and outdegree is the number of ties that the node directs to others. The number of these connections may be considered as an indicator of the importance of a node in a given network. Formally the degree centrality is:

$$CD(v) = \sum_{j \in G} \frac{(e_v j)}{n-1} \tag{2}$$

Although this centrality is simple to compute, it is a local measure which ignores the global structure of the network.

The betweenness centrality [13] is prominent among centrality indices. In fact, it provides an insight of the intermediary important nodes in a graph. So the node's importance is proportional to the number of shortest paths which pass the node among all node pairs. Mathematically, this measure is defined as follows:

$$BC(v) = \frac{1}{(N-1)(N-2)} \times \sum_{s \neq v \neq t} \frac{\sigma_(s,t)(v)}{\sigma_(s,t)}. \tag{3}$$

So, it is expressed as the fraction of shortest paths between source node s and target node t that pass through a given node v: $\sigma_{s,t}$ (v), averaged over all pairs of node in a network $\sigma_{s,t}$. N is the number of nodes in the network.

Unlike the degree centrality, this metric considers the whole network. It is applicable to networks with disconnected components. It is also efficient in case of information flow. However, for some nodes which do not lie on shortest paths between any two other nodes, the betweenness centrality turns to 0 while those nodes may be important.

The closeness centrality [13] measures the average shortest path length between the node and all other nodes in the graph. Therefore the node's importance is inverse-proportional to the sum of all shortest-paths (denoted here as $\mathrm{dist(v,t)}$) to other nodes. Formally, the closeness centrality is defined as follows:

$$Cc(v) = \sum \frac{1}{dist(v,t)} \tag{4}$$

The main limit of this metric is its inapplicability to networks with disconnected components.

The eigenvector centrality [19] uses the adjacency matrix A to retrieve the node that allows the maximum flow to pass within. Its mathematical presentation is as following where α is a parameter.

$$\alpha \times v = A \times v \tag{5}$$

This metric is suitable for studying spreading phenomena. However it may not scale well in case of networks with homogenous communities.

The group centrality [16] is the combination of all these metrics. It identifies the most central group in the network rather than nodes. This metric is useful in detecting communities or groups of nodes in a given graph.

5 Which Centrality Metric for Which Terrorist Network?

The main aim of this paper is to match each terrorist network topology with the correspondent centrality metric. Starting with the first category, the corporate based networks consist in different subgroups in the network: financial subgroup, propaganda subgroup, armed subgroup etc. These subgroups are led by important nodes which are the leaders. To destabilize this network, it is crucial to detect these leaders. The destabilization approach starts by identifying different subgroups in the network then to recognize the leaders of these groups. Therefore, traditional centrality metrics are not able to detect hidden groups of terrorists in the dark cell. So, we propose to use the group centrality to identify different groups. Once the groups are retrieved, we propose to use the degree centrality to identify the most central node in each group. So the destabilization approach here is two steps process.

The second category is the Politburo based terrorist network. This type of terrorist cells consists in one central committee which decides all the strategy of the network. Thus, it is important to retrieve this committee. We propose thereafter to use the group centrality mainly the kpset function [17] which identifies the most central group of players in a network. The destabilization approach here is a single step process. Retrieving the central committee is the target. No other processing is needed.

The third category is the shura based networks. “Shura” means “consultation”; the potential leader is the terrorist with more consultations: the node with more connections. So, the most important node is the most central node. Therefore, to identify key players in such networks, it is sufficient to use traditional centrality metrics: closeness, betweenness and degree.

The fourth category is the Multi-cell based network. Terrorist networks are formed of different cells. The key players are the nodes connecting these cells. To disrupt this kind of networks, it is crucial to retrieve the cells composing the network. We opt for the use of group centrality to identify different groups then we compute the betweenness centrality for the nodes connecting these cells. So, the disruption strategy is two step process using two centrality metrics: group centrality and betweenness centrality.

The fifth category is the brokerage based network. The important key players are the brokerage members. These latter are fully trusted by other members of the network. They are also considered as influential nodes. However, they are not the most central nodes in the network. So, traditional centrality measures are not able to detect these members since these nodes are characterized as “trusted” and “influential”. It is possible, thereafter to use the newly introduced metrics “the reliance measure” and/or the “influence index”. The reliance measure combines two essential aspects to detect this kind of nodes mainly “importance” and “trust”. Also, the influence index measure may be helpful in revealing these nodes. So, to detect brokerage members we propose to use the reliance measure and the influence index.

The last category is the wolf based terrorist network. A principal actor “The wolf” acts secretly in the hand of one terrorist who is the important node This node is peripheral and is connected to only one node. This latter is an important node holding many connections. We are looking for the node connected to important neighbors. Consequently, we may use the eigenvector centrality to retrieve this node. Other centralities such as: degree, betweenness and closeness are not able to find it because they reveal the most central node. The eigenvector centrality is able to find the node which is not central but connected to central ones.

6 Experimental Results

We experiment on five different real world data sets (detailed next) representing five categories: corporate based networks, politburo based networks, shura based networks, multi-cell based networks and brokerage based networks. For the last category: Lonely wolf based network, we do not have any real world data set so we may not conduct any experimental analysis. For each data set, we constructed its adjacency matrix to build the graph and process our experiments on it in R [20]. Due to space limitation, we omit the data sets representations. We present in the following sections the experimental results of the five data sets.

6.1 Corporate Based Terrorist Network: IRA Case Study

The IRA (Irish Republican Army) is a network of terrorists consisting of 55 individuals [18]. We use [18] to build our adjacency matrix and visualize the network in R. This cell is formed of different subgroups: financial, propaganda... According to Fig. 1, there are only two important key players to be isolated which are node 28 and 37. However, the network consists of more than two subgroups. Therefore, we propose to identify different subgroups in the network then to retrieve the central nodes inside these groups. So, we start by applying group centrality to this data set.

According to the group centrality, this network consists in five different subgroups which are: $g1 = V4, V5, V6, V7, V8$, $g2 = V12, V13, V14, V15, V16, V17, V18, V19$, $g3 = V20, V22, V23, V24, V25, V26, V27$, $g4 = V28, V29, V30, V31, V32, V54, V55$ and $g5 = V37, V38, V39, V40, V41, V42, V43, V44, V45, V46, V47, V48$. Once the groups are known, we apply the degree centrality to recognize the leaders of these groups. These latter are respectively: V7, V14, V22, V28 and V37. So to effectively disrupt this network, it is necessary to isolate these nodes.

6.2 Politburo Based Terrorist Network: RAF Case Study

The RAF (Red Army Faction) is a German terrorist network composed of 29 individuals [7]. We took the data from [7] and represent it in R. The topology of this data set is based on a central committee that decides all the strategy of the network. The key players in the network to be isolated are the members

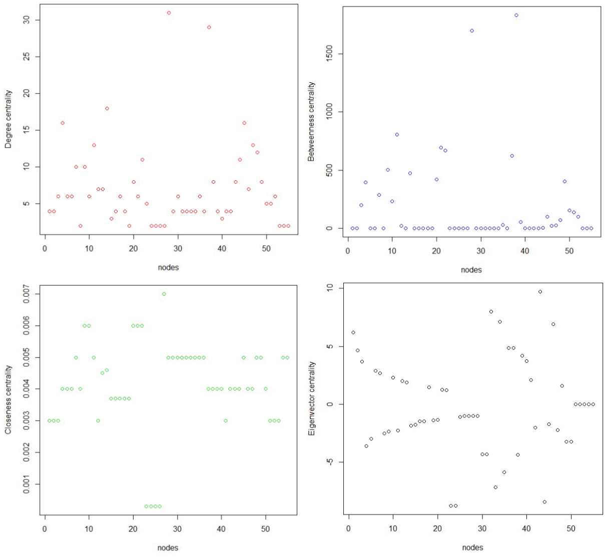


Fig. 1. IRA traditional centralities

of this central committee. This committee is close to front end members [7]. The following figure represents the corresponding values of the four traditional centrality metrics for each node.

According to Fig. 2, we may notice that important nodes with high degree centrality, have also high betweenness values and closeness ones. These nodes are most central in the network. However these nodes do not form a committee and some of them does not have any relationship with other nodes. Therefore, these metrics are not able to detect key players in this network. We propose to use the group centrality based on degree centrality. This metric identifies the most central group in a given network. Using this centrality we may identify the central committee. As we can see, the network is composed of three dense subgroups $g1 = V4, V5, V6, V7, V8, V9$, $g2 = V10, V11, V14, V15, V16$ and $g3 = V21, V22, V23, V24, V25, V26, V27$. The group of nodes that have the higher group centrality is the group $g2 = V10, V11, V14, V15, V16$ as it is indicated in Fig. 3. So, the central committee in the RAF terrorist network to be isolated is the group of nodes $V10, V11, V14, V15, V16$.

6.3 Shura Based Terrorist Network: Turkish Al Qaida, November 2003

The data set studied here represent the terrorists of the November 2003 attack in Turkey [7]. The adjacency matrix is constructed using the data in [7] to

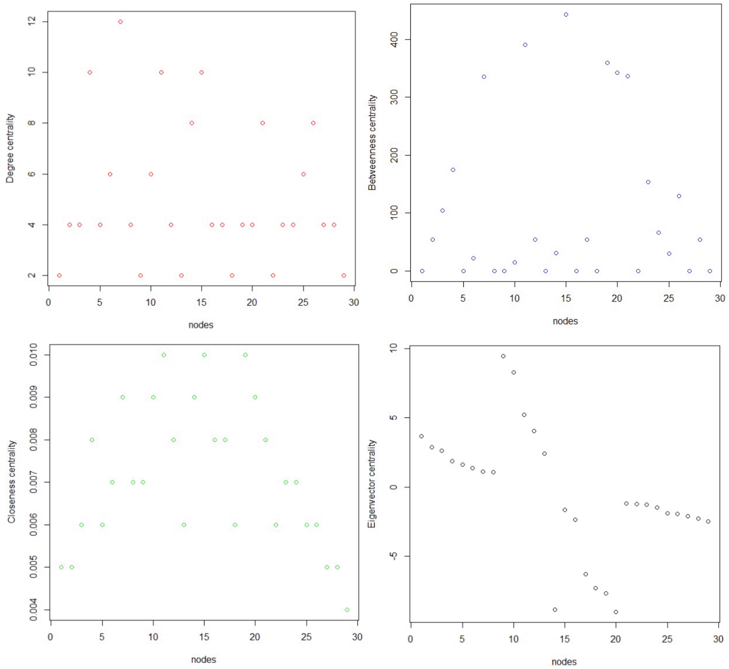


Fig. 2. Traditional centrality metrics of RAF data set

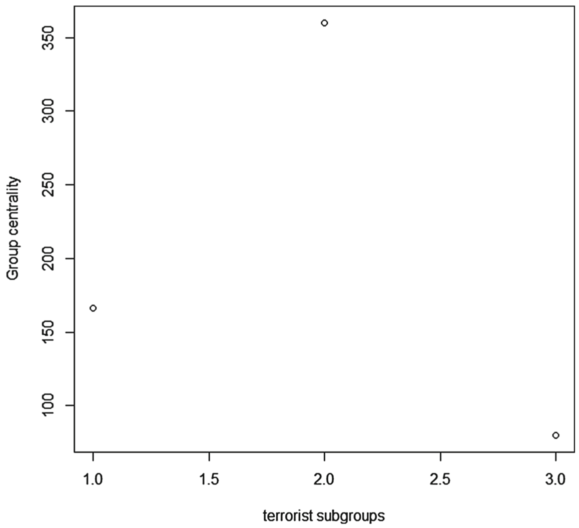


Fig. 3. Most central group in RAF

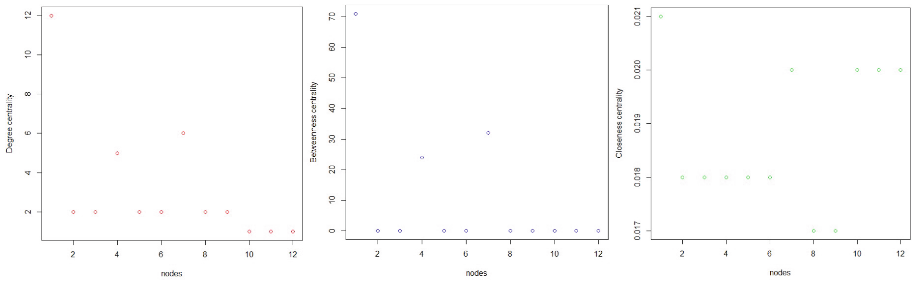


Fig. 4. Centrality metrics of Turkish Al Qaida November 2003 data set

obtain the graphic representation of the terrorist network which is composed of 12 members. The potential leader is the node with more "consultations" which is the most central node. Traditional centrality metrics in this case may retrieve the key players in this cell. So, we present in Fig. 4 the values of the degree, betweenness and closeness centrality metrics.

As it is shown in Fig. 4, the node with the highest centralities values is the node V1. To destabilize this network, it is necessary to isolate the leader which is node V1.

6.4 Multi-cell Based Terrorist Network: 9/11 Attack Case Study

This well known data set represents the 9/11 attack terrorists [2]. To obtain the terrorist network representation, we used the data in [2]. There are 19 members who participated in the attack. The topology of the network is composed of different cells. The key players are the nodes connecting these cells also characterized as intermediary nodes. In this case, the betweenness centrality may be used to retrieve these nodes after identifying the composing cells. These latter according to the group centrality are: $C1 = V1, V2, V3$, $C2 = V5, V7, V9, V11, V12$, $C3 = V13, V14, V16, V17, V18$. The intermediary nodes are: $V4, V16, V8$ and $V15$. The following figure illustrates the different values of different centralities in the network. We may notice that degree centrality and eigenvector do not retrieve the desired results. The nodes retrieved by the closeness are $V5, V7, V9$ and $V16$. However, these nodes are not the intermediary nodes. The nodes $V4$ and $V16$ are considered of high betweenness centrality. So these nodes are connecting cells and they are the key players in the network (Fig. 5).

6.5 Brokerage Based Terrorist Network: Ergenekon Network Case Study

The Ergenekon turkish terrorist organization is composed of 33 members [7]. We use data in [7] to visualize our network in R. The key players in this network are the brokerage members who are considered as influencers and fully trusted by other members. Figure 6 represents the correspondent values of traditional centrality metrics. As it is indicated in [7] the brokerage members in the Ergenekon

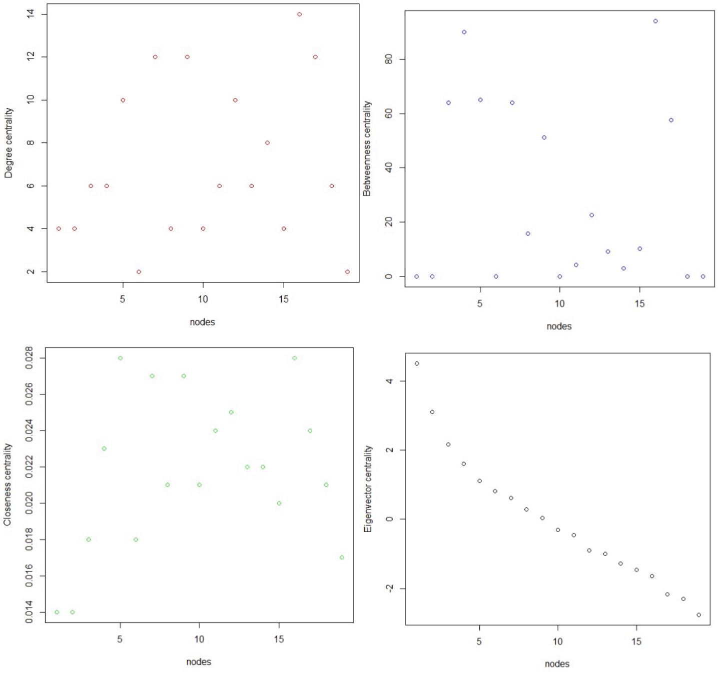


Fig. 5. Centralities of 9/11 data set

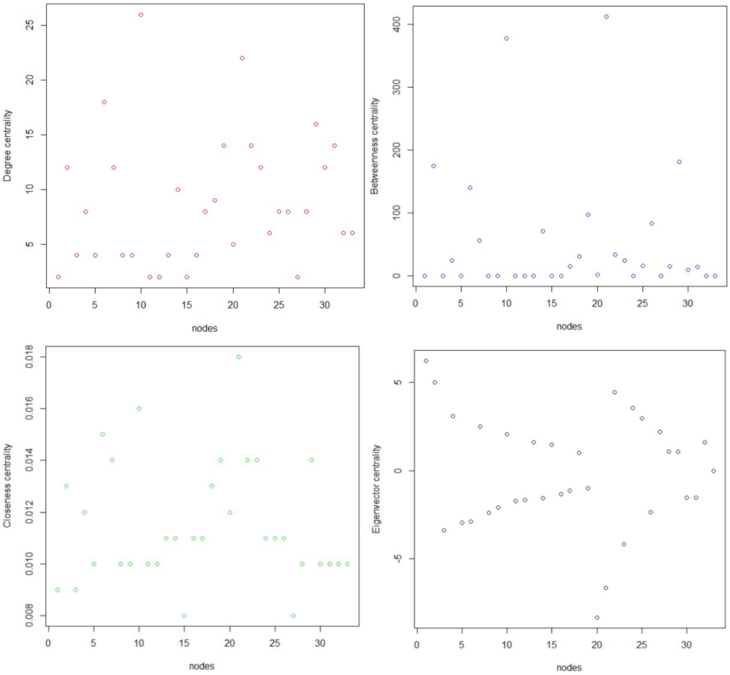


Fig. 6. Centralities of Ergenekon data set

network are nodes V2 and V30. However, as we may notice in Fig. 6 none of the four measures above identify those nodes as "important". Hence we have to opt for other measures to identify the brokerage members in this network such as the influence index and the reliance measure.

Corporate based networks, politburo and multi-cell based ones rely on an active group of people in the network. So, we apply group centrality to retrieve them. Then, we opt for degree and betweenness to find central members in these groups. A combination of degree, betweenness and closeness is necessary to identify central terrorist in Shura based networks. The definition of key players changes for the last category: the most important node is not the most central node but rather the most trusted one. So none of these metrics would reveal the key node.

7 Conclusion

In this paper, we studied different centrality metrics widely used especially in the terrorist organizations analysis. We focused on the different topologies of terrorist networks and provide for each topology the correspondent centrality metric that may identify the key players therein. This paper contributes by providing first steps towards the matching between centrality metrics and the correspondent network topology. The experiments conducted on five different data sets prove our theoretical analysis and lay the ground for further investigations.

As a future work, we aim first to measure the reliance measure and the influence index of the last category so we can confirm our theoretical results. Besides in order to complete our work, we will look to a data set of the lonely wolf category. The scope of this paper is on terrorist networks, a further analysis and experimentation on other data sets category and large scale graphs is needed as future work.

References

1. Magalingam, P., Davis, S.: Ranking the importance level of intermediaries to a criminal using a reliance measure. arXiv preprint arXiv:1506.06221 (2016)
2. Michalak, T.P., Rahwan, T., Skibski, O., Wooldridge, M.: Defeating terrorist networks with game theory. IEEE Intell. Syst. **1**, 53–61 (2015)
3. Michalak, T.P., Aadithya, K.V., Szczepanski, P.L., Ravindran, B., Jennings, N.R.: Efficient computation of the Shapley value for game-theoretic network centrality. J. Artif. Intell. Res. **46**, 607–650 (2013)
4. Skibski, O., Michalak, T.P., Rahwan, T., Wooldridge, M.: Algorithms for the Shapley and Myerson values in graph-restricted games. In: Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems, pp. 197–204. International Foundation for Autonomous Agents and Multiagent Systems, May 2014
5. Xuan, D., Yu, H., Wang, J.: A novel method of centrality in terrorist network. In: Seventh International Symposium on Computational Intelligence and Design (ISCID), December 2014, vol. 2, pp. 144–149. IEEE (2014)
6. Krebs, V.E.: Mapping networks of terrorist cells. Connections **24**(3), 43–52 (2002)
7. Sparrow, M.K.: The application of network analysis to criminal intelligence: an assessment of the prospects. Soc. Netw. **13**(3), 251–274 (1991)
8. Berzinji, A., Kaati, L., Rezine, A.: Detecting key players in terrorist networks. In: 2012 European Intelligence and Security Informatics Conference (EISIC), pp. 297–302. IEEE, August 2012
9. Azad, S., Gupta, A.: A quantitative assessment on 26/11 Mumbai attack using social network analysis. J. Terrorism Res. **2**(2), 1–10 (2011)
10. Nasrullah, M., Larsen, H.L.: Structural analysis and mathematical methods for destabilizing terrorist networks. In: Proceedings of the International Conference on Advanced Data Mining Applications, pp. 1037–1048 (2006)
11. Fellman, P.V., Clemens, J.P., Wright, R., Post, J.V., Dadmun, M.: Disrupting terrorist networks: a dynamic fitness landscape approach. In: Minai, A.A., Braha, D., Bar-Yam, Y. (eds.) Conflict and Complexity, pp. 165–178. Springer, New York (2015)
12. Ozgul, F., Bowerman, C.: Characteristics of terrorists networks based on ideology and practices. In: 2014 European Network Intelligence Conference (ENIC), pp. 95–99. IEEE (2014)
13. Bonacich, P.: Factoring and weighting approaches to status scores and clique identification. J. Math. Soc. **2**(1), 113–120 (1972)
14. Karthika, S., Bose, S.: A comparative study of social networking approaches in identifying the covert nodes. Int. J. Web Serv. Comput. **2**(3), 65 (2011)

15. Hamed, I., Charrad, M.: Recognizing information spreaders in terrorist networks: 26/11 attack case study. In: Bellamine Ben Saoud, N., Adam, C., Hanachi, C. (eds.) ISCRAM-med 2015. LNBIP, vol. 233, pp. 27–38. Springer, Heidelberg (2015). doi:10.1007/978-3-319-24399-3_3
16. Everett, M.G., Borgatti, S.P.: The centrality of groups and classes. J. Math. Sociol. **23**(3), 181–201 (1999)
17. http://rpackages.ianhowson.com/cran/keyplayer/
18. http://news.psu.edu/story/264519/2013/02/18/research/international-center-study-terrorism-focuses-latest-research
19. Everton, S.F.: Network topography, key players and terrorist networks (2009)
20. https://www.r-project.org/
21. http://www.inf.fu-berlin.de/lehre/WS06/pmo/eng/audio/Chomsky.pdf