Recognizing Information Spreaders in Terrorist Networks: 26/11 Attack Case Study

Imen Hamed[1](✉) and Malika Charrad[1,2,3]

[1] RIADI Lab, La Manouba University, Manouba, Tunisia
imen.hamed@outlook.com
[2] University of Gabes, Zrig Eddakhlania, Tunisia
[3] MSDM Team, Cedric Lab, CNAM, Paris, France
malika.charrad@riadi.rnu.tn

Abstract. Terrorism is a man-made hazard characterized by its uncontrollability and unpredictability. In fact, terrorist cells are covert networks where secrecy is the prime concern during the operation. To disrupt these inhuman operations, it is crucial to reveal this secrecy and identify the responsible key actors. Therefore, a new research area emerges. Investigative Data Mining (IDM) is the study of terrorist networks using Social Network Analysis (SNA). It involves graph theory to analyze networks. Among analysis techniques, network metrics defined as centrality measures have been successfully involved in terrorist networks destabilization methods. In this paper, we propose another disruption strategy of terrorist network using the percolation centrality metric. This measure allows to conduct a dynamic analysis of terrorist network on one hand. On the other hand, it identifies information spreaders in the network. We experiment on the Mumbai 26/11 attack data set, the proposed approach recognizes the information spreaders involved in this incident.

Keywords: Social network analysis · Terrorist network · Percolation centrality

1 Introduction

The world has witnessed many terrorist attacks in the past few years such as Casablanca, Morocco (May 16, 2003), Russia (September 1–3, 2004) and Tunisia (Chaambi mount terror operations and the Bardo museum attack on March 2015). The consequences of these attacks in terms of number of lives lost and the infrastructure damage appeal for an urgent remedy. Thinking the crisis management at macro level, it is necessary to isolate terrorist group components and prevent or (at least) limit the impact of such crisis.

Several important works propose different techniques to disrupt the terrorist cells. SNA techniques are prominent among them. SNA can be used to identify, measure, visualize and analyze the ties among people, groups and organizations.

N. Bellamine Ben Saoud et al. (Eds.): ISCRAM-med 2015, LNBIP 233, pp. 27–38, 2015.
DOI: 10.1007/978-3-319-24399-3_3

Information about terrorists is usually transformed into network structure in which nodes represent the terrorists and the links are the connections between these individuals. SNA can provide useful information about these terror groups through the study and analysis of the network evolution, the node position in the graph and the connections between different individuals.

Terrorist networks are complex adaptive systems [5]. In fact, they are composed of dynamic autonomous cells which are widely dispersed and typically covert. Given that individuals play different roles in their cells [6], the illegal activities of the terrorists are split among them. Therefore, isolation of terrorist cells requires the identification of important actors and respectively their different roles.

To analyze terrorist networks, it is fundamental to measure network density and centrality. Various behaviors of individuals are significantly influenced by their positions in networks [1]. Therefore, it is required to determine the dominant roles and to reveal the key players in the network through centrality measures. Throughout this paper, we review different destabilization algorithms based on centrality measures (Sect. 3). Then, we introduce a new network destabilization method based on percolation centrality. We analyze the terror attacks in Mumbai on November 26, 2008 to illustrate the feasibility of our approach (Sect. 4). Our contributions may be summarized as follows:

- We visualize the network of the 26/11 hijackers using the statistical and mining tool R. The produced graph helps to do further interpretation and analysis.
- We explain how the percolation centrality metric can be incorporated in terrorist cells destabilization strategy. Traditional disruption methods rely on the identification of central node. We illustrate the flaws of these methods. Thereby, we introduce our strategy based on identifying information spreaders in the network.

2 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. For any finite network G of N vertices, we denote the sets of vertices and edges by V(G) and E(G) respectively:

$$V(G) = v_i | i = 1...N, |V(G)| - \text{size of network G- N, .} \tag{1}$$

$$E(G) = e_{ij} | i, j = 1...N, e(G) - |E(G)| - \text{ number of network edges, .} \tag{2}$$

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 [3].

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge connecting } v_i \text{ to } v_j \\ 0 & \text{otherwise.} \end{cases} \qquad (3)$$

A host of centrality measures have been proposed to analyze complex networks. A centrality measure of a vertex or edge gives a numerical qualification of that element's relative network importance [3]. Betweenness centrality is prominent among different centrality measures.

This well known measure aims to quantify a node's importance as a conduit of information flow in a network. Formally, it is defined as:

$$BC(v) = \frac{1}{(N-1)(N-2)} \times \sum_{s \neq v \neq t} \frac{\sigma_(s,t)(v)}{\sigma_(s,t)} \ . \qquad (4)$$

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.

3 Related Work

SNA offers a branch of techniques to study terrorist networks. Different works emerge to propose methods to disrupt terrorist cells.

The authors in [10] propose a new centrality measure: network flow centrality Load. The objective of their work is to identify vertices to remove from the network in such a way as to force more flow through a critical vertex. Thus, the information flow is forced to pass through this vertex. Hence, it becomes possible to measure the activity of that node through quantifying how much flow must pass through it.

The authors in [13] introduce a new method to destabilize terrorist networks. This method consists in comparing different centrality values of different nodes to recognize nodes that are powerful, influential or worthy to neutralize.

The works in [11] analyzes the attacks of 26/11 using a set of centrality measures namely degree, betweenness, eigenvector and closeness. The degree of a node is the number of neighbors it has. The closeness measures the average shortest path length between the node and all other nodes in the graph. Eigenvector defines the influence of a node on its neighboring nodes. Based on these indices, the authors deduce the hierarchy of the terrorist network and recognize the most influential node. In [13], the authors propose two steps approach to achieve the destabilization. First, they define an algorithm that converts the undirected graph to a directed one using the degree and eigenvector centralities. Then, the second step is to construct a tree from the dependence centrality measure. So, the destabilization is reached.

As an attempt to improve the previous approach, the authors in [12] used two measures Katz centrality and PageRank centrality. Instead of running the two algorithms separately, the authors introduce a single step approach. Katz centrality can be viewed as a variant of eigenvector centrality. The aim of Katz is to measure the influence of a given node on the rest of nodes. Indeed, it counts the number of walks starting from a node or ending on a node, providing penalties to longer walks. The PageRank metric defines the influence of a node. It is an enhanced version of in-degree centrality. The computing of PageRank and Katz centralities reduce the time and space complexities.

The algorithm proposed in [9] relies on three centrality measures: degree centrality, closeness centrality and betweenness centrality. The authors claim that the financial manager is the most central node which is closest to other nodes. So these dark cells may be disrupted when the financial manager is isolated.

The authors in [8] propose a new centrality measure for destabilization purposes namely influence index. This measure is based on three degree of influence rule: A node is influenced by other nodes that lie at three degree of separation but not by those beyond. The influence index method consists of internal influence (itself) and external influence (from others'). The approaches presented above adopt traditional centrality measures such as betweenness, eigenvector and closenes. They define an influential node as a topologically central node holding multiple connections.

In this paper we propose a single step approach consisting in the computation of a single metric and adopting a different definition of influential nodes.

4 Proposed Approach

4.1 What Is Percolation

The percolation aspect in complex networks occurs in many scenarios: Viral content or rumors spread over contact networks. In a network of towns, the spread of disease and the contagious infections are considered as typical scenarios of percolation. Besides, computer viruses can divide over computer networks (Fig. 1).

The aforementioned schema illustrates the percolation process. At t = 1, the state of node A changes. Given that node A is related to node B, the state of node B changes also at t = 2. So, the contagion alters the state of the node as it spreads. A percolated node percolates its neighbors over time. The state of a node can be binary (such as received/not received a piece of news), discrete (susceptible/infected/recovered) or even continuous (such as the proportion of infected people in a town) as contagion spreads [4].

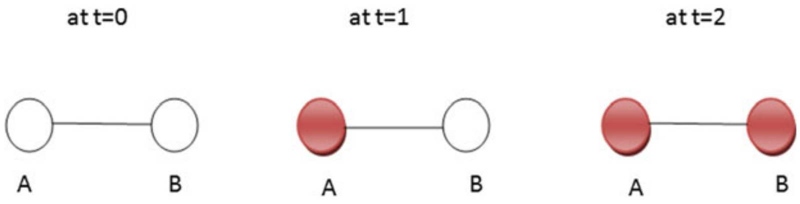

Fig. 1. The percolation process

The percolation centrality measures the importance of a node in terms of aiding the percolation through the network. While current centrality measures focus on purely topological importance of a node, the percolation measure combines the topological importance of the node in the graph and the node state (percolated/non-percolated/partially percolated). So, this measure copes very well with the network dynamics. Formally, the percolation centrality denoted PC is defined as:

$$PC^t(v) = \frac{1}{(N-2)} \times \sum_{s \neq v \neq r} \frac{\sigma_{s,r}(v)}{\sigma_{s,r}} \times \frac{x_s^t}{[\sum x_i^t] - x_v^t} \ . \tag{5}$$

x_s^t is denoted to the percolation state of the node s at time t.
x_i^t represents the percolation state of any node i at time t.
x_v^t is the percolation state for the node v to which the percolation centrality is computed.

PC determines at any time how important is the node to the overall process of percolation. Formally, the percolation centrality measure adopts the betweenness centrality measure logic since it relies on the number of shortest paths in the network. There are two extreme cases where the percolation centrality is trivially equal to the betweenness centrality: Given a single percolated node in the network; if we iterate over all possible percolated nodes one by one and then average over all the scenarios then we get:

$$PC^t(v) = \frac{1}{(N-1)(N-2)} \times \sum_{s \neq v \neq r} \frac{\sigma_{s,r}(v)}{\sigma_{s,r}} = BC(v). \tag{6}$$

So, the percolation centrality is reduced to betweenness centrality. Another scenario leads to the reduction of percolation centrality to the betweenness one is when all the nodes of the network are fully percolated:

$$\frac{x_s^t}{[\sum x_i^t] - x_v^t} = \frac{1}{(N-1)} \ . \tag{7}$$

Hence:

$$PC^t(v) = \frac{1}{(N-1)(N-2)} \times \sum_{s \neq v \neq r} \frac{\sigma_{s,r}(v)}{\sigma_{s,r}} = BC(v). \tag{8}$$

That is the percolation centrality starts as betweenness centrality and then it evolves and finally reduces to end as betweenness centrality again.

4.2 Why Percolation Centrality in Terrorist Network

The SNA defined centrality measures have been successfully incorporated in the destabilization of terrorist networks by determining the dominant role(s) from the network [12]. To sustain a successful attack operation, the terrorists tend to play different roles. A terrorist network is led by a leader who mostly acts as a mentor and only provides guidance on how to organize and motivate the group operatives [9]. The finance manager is the one who occupies the most central and active role in a decentralized terrorist network.

The majority of works focus on recognizing the leader of the terrorist cell. SNA can locate the true points of vulnerability in the network rather than simply the apparent leadership [7]. The best node to attack is not always that which has the most connections [7]. Dynamic analysis based on changing time and place can dig out more useful information that classical methods could not find [8]. It can also provide a clear picture of how information flows in the network [7].

In this paper, we aim to identify the information spreaders in the terrorist cell. We adopt a dynamic analysis based on the percolation centrality. The traditional centrality measures quantify the importance of a node in purely topological terms. The advantage of the percolation centrality is the consideration of the network dynamics. The important node is no longer the most central one, but the node which contributes more in the percolation process. Therefore, we propose a new method to destabilize the terrorist network. The nodes that spread the information in the terrorist network are important nodes. The identification of such nodes may disrupt the network. The authors in [7] claim that the leader in the network is not the most central node or the node that held many connections. Accordingly, the removal of key personnel, such as the most central node, will not necessarily collapse the network.

4.3 Terrorist Network Visualization

As a case study in this paper, we analyze the 26/11 attacks in Mumbai using the public data that were available in [11]. As a first experimentation, we considered a static terrorist network. In future work, we aim to consider dynamic network. So, we perform further analysis.

Figure 2 represents the hijackers of the Mumbai attacks on 26/11/08. Three among the thirteen terrorists were handling the operation simultaneously from Pakistan. The others were dispersed in different regions in Mumbai. The network depicted in Fig. 2 is not complete. In fact, some links are missing probably or a part of the network is still opaque. The terrorist number 12 and 13 are involved in the attack operation. However, they do not yield any connection with the other members. Despite the incompleteness of the network, we still may make some interesting inferences. Node 2 seems to be an important node. He has many connections with nodes in different regions in Mumbai. As shown in Table 1, the node 2 has the highest betweenness centrality in the network.

The authors in [11] conclude that he is the responsible for the attack operation. Given that the terrorist organizations have quite organizational roles to conduct successful operations, computing betweenness centrality does not provide a clear insight into the hijackers' roles since it reveals the most central node in the network. We wonder what are the roles of the individuals with 0 betweenness centrality score. Therefore, we propose to apply the percolation centrality metric.

The first step in our approach consists in computing the percolation centrality for each node at different time steps then iterating over the 13 nodes at different

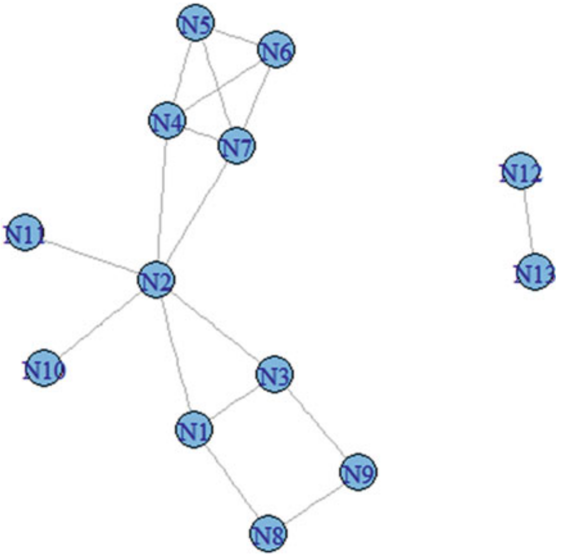

Fig. 2. 26/11 terrorist network

Table 1. Nodes betweenness centrality

Nodes N	Betweenness centrality
N1	7.5
N2	33
N3	7.5
N4	7
N5	0
N6	0
N7	7
N8	0.5
N9	0.5
N10	0
N11	0
N12	0
N13	0

time steps. We consider a binary percolation state: received/not received a piece of information. In addition, the percolation is not random .i.e. the terrorists follow a well organized strategy to spread information in the network. The second step consists in studying the impact of the percolated node on the other nodes and identifying nodes which are central in terms of their impact.

4.4 Interpretation and Discussion

To maintain clear interpretation of the experimental analysis, we divide the nodes into peripheral nodes and non peripheral nodes. A node i in a graph is called peripheral if there is a node j in the graph such that the distance d ij equals the diameter of the graph [15]. The percolation centrality illustrated in the following graphs represents the impact of the percolated node on the other nodes. The node which highly impacts the other nodes is considered as an information spreader. Due to space restriction, we can not represent a graph per node. So, we have combined the impact of 3 or 4 nodes on the other nodes in each graph.

The process of percolation is explained as follows:

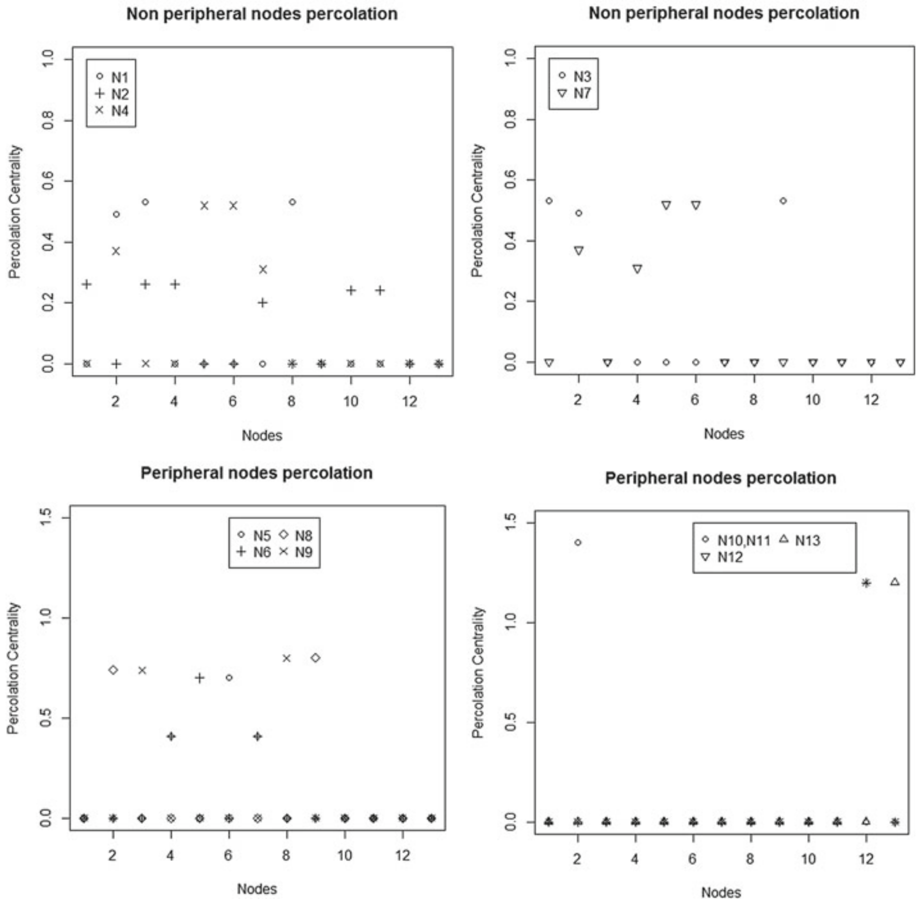

Fig. 3. Percolation centrality at t = 2

- At t = 1, one node is percolated. As discussed in previous section, when a single node is percolated in the network, the percolation centrality is equivalent to betweenness centrality.
- At t = 2, each node percolates its neighbors. Indeed, all the paths connecting the percolated node with its neighbors become percolated. The information starts to spread in the network. For example, node 1 is percolated. Consequently, it percolates its neighbors: node 2, 3 and 8. All the scenarios of percolation are presented in the following figures.
- At t = 3, the nodes continue to percolate the rest of the network as shown in Fig. 4 At t = 3, the node 2 percolates the entire network. The percolation centrality of nodes 10 and 11 decreases significantly. Nodes 12 and 13 maintain the same score given that they are isolated from the rest of the network.

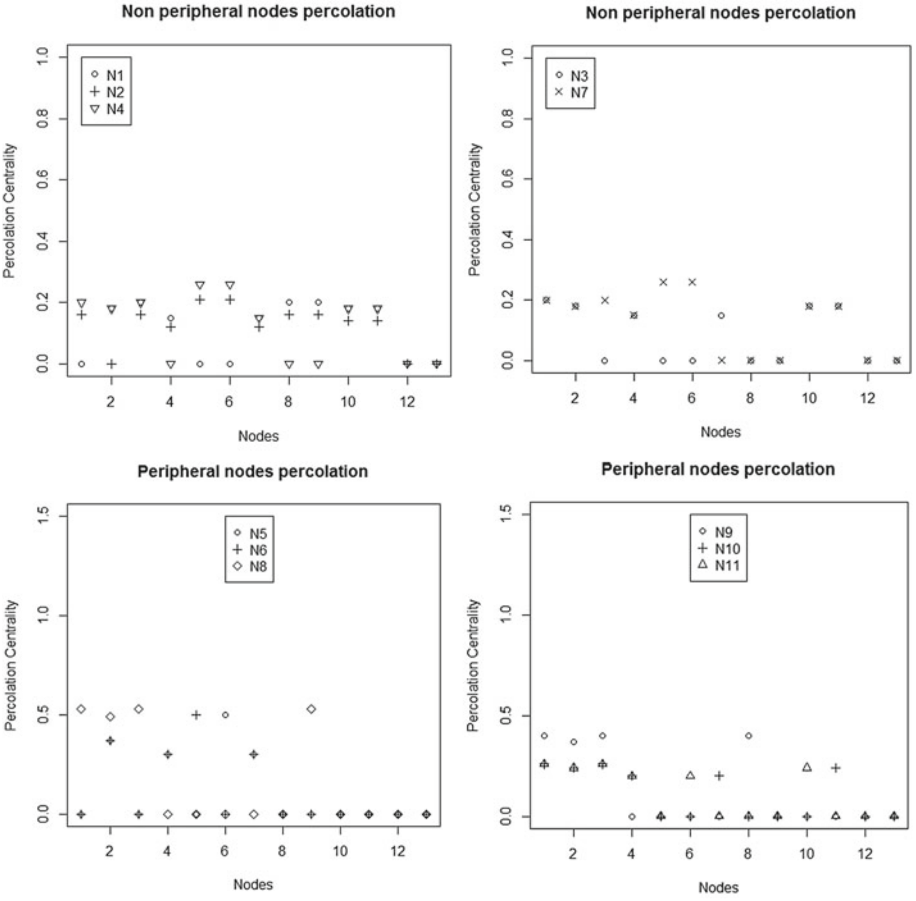

Fig. 4. Percolation centrality at t = 3

So, the percolation process for this couple of nodes stops at t = 2. These nodes preserve slightly high scores when nodes 8 and 9 are percolated.

- At t = 4, almost the entire network gets percolated for different scenarios.

We may notice that non peripheral nodes are speedier in percolating the network rather than the peripheral nodes. For example, node 2 percolates the entire network in 3 time steps. However, some other nodes such as nodes 8 and 9 need 5 time steps to percolate all the other nodes. In t = 4, we notice that the nodes have higher percolation centrality when nodes 5 and 6 are considered the source of percolation. However nodes 8 and 9 lose their importance in the percolation process.

When all nodes are percolated, the percolation centrality becomes the same regardless to the initially percolated node. As shown in Figs. 3, 4 and 5, nodes 5 and 6 held the higher percolation centrality at t = 4. Nodes 8 and 9 are also

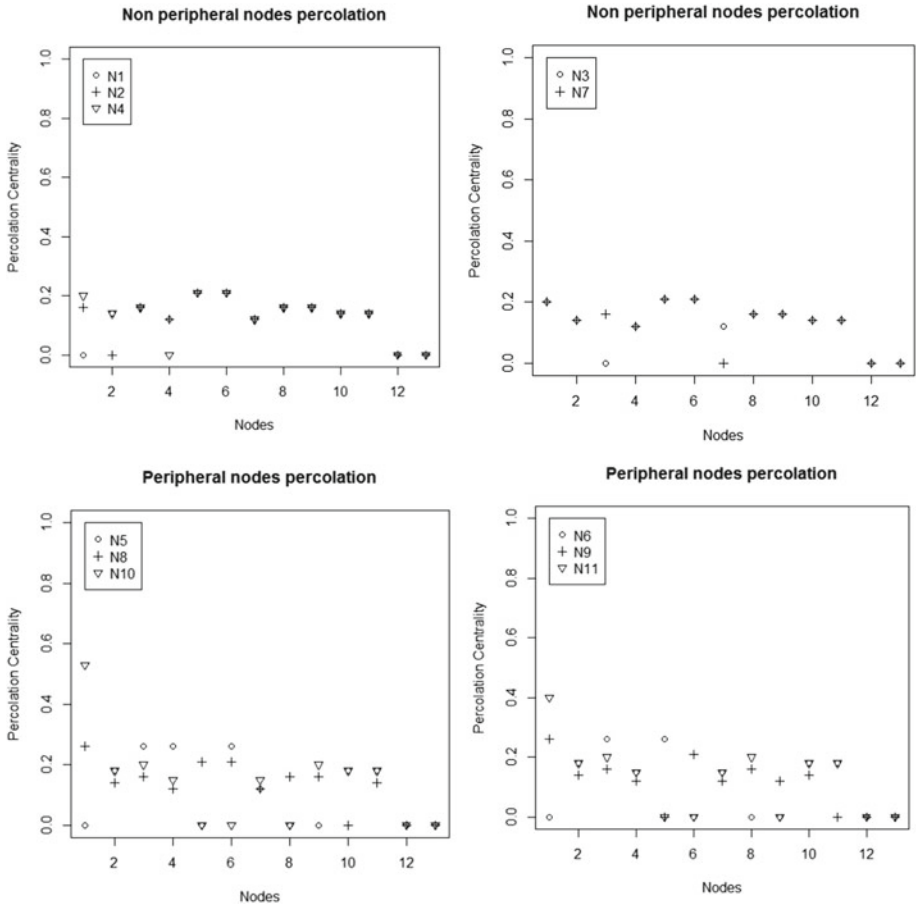

Fig. 5. Percolation centrality at t = 4

considered of higher score of percolation centrality at t = 2 and t = 3. Thus, it is clear that the peripheral nodes are the most important nodes in the process of information spread.

The obtained results illustrate that the nodes are shown to have a high percolation centrality when nodes 8, 9, 10 and 11 are percolated. Also, nodes 12 and 13 are shown to have a high score as well. We may conclude that these nodes are the most important in the spreading process. At this time step, they are considered as the information spreaders.

As depicted in the presented figures, we considered all the possible scenarios. Each node is susceptible to be the information spreader. At each time step, we measure the effect of the percolated node on the rest of nodes. The percolation centrality decreases considerably when the number of neighbors increases. Peripheral nodes tend to have higher percolation centrality, although they spend more time to percolate the entire network. Percolation centrality scores start to stabilize over time.

The dynamic aspect of the percolation centrality measure provides better understanding of how information flows in the network even though the network is static. The percolation centrality may be calculated in O(NM) time. So, there is no significant increase in time complexity compared to betweenness centrality.

As indicated in [14], nodes $5, 6, 8$ and 9 were exchanging information with the three handlers in Pakistan by phone calls. These information confirm the retrieved results with the proposed approach. The destabilization strategy of this dark network can take place by isolating those node spreaders.

5 Conclusion

In this paper, we propose a dynamical analysis of the Mumbai terrorist incident. Based on percolation centrality metric, we reveal the information spreaders in the dark network through different time steps. The proposed approach may aid the law enforcement agencies to track the terrorists and develop stronger disruption strategies.

Future work consists in considering large and dynamic networks. Furthermore, we tend to combine the percolation centrality metric with other metrics to develop more robust destabilization method.

References

1. Eom, Y.H., Jo, H.H.: Generalized friendship paradox in complex networks: the case of scientific collaboration. J. Sci. Rep. **4**, Article No. 4603 (2014)
2. Krebs, V.E.: Mapping networks of terrorist cells. J. Connect. **24**, 43–52 (2002)
3. Newman, M.: Networks: An Introduction. Oxford University Press, Oxford (2010)
4. Piraveenan, M., Prokopenko, M., Hossain, L.: Percolation centrality: quantifying graph-theoretic impact of nodes during percolation in networks. J. PLoS One **8**, 53–95 (2013)
5. Ilachinski, A.: Self-organized terrorist-counterterrorist adaptive coevolutions, part 1: a conceptual design. Technical report (2005)

6. Shaikh, M.A., Wang, J., Yang, Z., Song, Y.: Advanced Data Mining and Applications. LNCS(LNAI), vol. 4632, pp. 570–577. Springer, Heidelberg (2007)
7. Lauchs, M.A., Keast, R., Le, V.: Social network analysis of terrorist networks: can it add value? J. Criminol. **3**, 21–32 (2012)
8. Xuan, D., Yu, H., Wang, J.: A novel method of centrality in terrorist network. In: 7th International Symposium on Computational Intelligence and Design (ISCID), pp. 144–149. IEEE Press, China (2014)
9. Berzinji, A., Kaati, L., Rezine, A.: Detecting key players in terrorist networks. In: European Intelligence and Security Informatics Conference (EISIC), pp. 297–302. IEEE Press, Denmark (2012)
10. Martonosi, S., Altner, D., Ernst, M., Ferme, E., Langsjoen, K., Lindsay, D., Plott, S., Ronan, A.S.: A new framework for network disruption. J. ArX. Prep (2011)
11. Azad, S., Gupta, A.: A quantitative assessment on 26/11 mumbai attack using social network analysis. J. Terror. Res. **2**(1) (2011)
12. Chaurasia, N., Tiwari, A.: Efficient algorithm for destabilization of terrorist networks. J. Inf. Technol. Comput. Sci. **5**, 21–30 (2013)
13. Memon, N., Larsen, H.L.: Structural analysis and mathematical methods for destabilizing terrorist networks using investigative data mining. In: Li, X., Zaïane, O.R., Li, Z. (eds.) ADMA 2006. LNCS (LNAI), vol. 4093, pp. 1037–1048. Springer, Heidelberg (2006)
14. Government of India: Mumbai terrorist attack, (26–29 Nov 2008) (2009)
15. Gimes, R.G., Pierce, D.J., Simon, H.D.: A new algorithm for finding a pseudoperipheral node in a graph. J. Matrix Anal. Appl. **11**, 323–334 (1990)