Social Network Analysis of Terrorist Networks

Ashlesha S. Nagdive, Rajkishor Tugnayat, Atharva Peshkar

Abstract: Terrorist Activities worldwide has led to the development of sophisticated methodologies for analyzing terrorist groups and networks. Ongoing and past research has found that Social Network Analysis (SNA) is most effective method for predictive counter-terrorism. Social Network Analysis (SNA) is an approach towards analyzing the terrorist networks to better understand the underlying structure of a network and to detect key players within the network and their links throughout the network. It is also need of the hour to convert available raw data into valuable information for the purpose of global security. Comparative study among SNA tools testify their applicability and usefulness for data gathered through online and offline social sources. However it is advisable to incorporate temporal analysis using data mining methods, to improve the capability of SNA tools to handle dynamic social media data. This paper examine various aspects of Social Network Analysis as applied to terrorism, taking empirical data, and open source data based studies into account. This work primarily focuses on different types of decentralized terrorist networks and nodes. The nodes can be classified as organizations, places or persons. We take help of varied centrality measures to identify key players in this network.

Keywords: Social Network Analysis, Terrorist Networks, Counter-Terrorism, Centrality, Investigative Data mining.

I. INTRODUCTION

Social Network Analysis is a tool for understanding the pattern or dynamics of terrorism and terrorist networks. The Intelligence Analysts Training Manual of the Metropolitan Police (Scotland Yard, London) on its front bears the statement.

“Analysis is the key to the successful use of information; it transforms raw data into intelligence.”

Due to this importance of intelligence analysis, the objective of this paper is to survey the tools, datasets and methods available for analysis of social network, specifically terrorist networks. Social network analysis [SNA] is the mapping and measuring of relationships flow between people, groups, organizations, URLs, and other informative entities. The nodes in the network are the people and their groups while the links show relationships between the nodes. It is a process of quantitative and qualitative analysis of a social networks that maps, measures and visualize the relation between nodes in the network.

SNA can help to pinpoint crucial nodes in a network who should be targeted in order to disrupt organizational activities. This is explored here by first considering various practical obstacles, followed by an empirical test of how centrality measures perform against known behaviour of an actual terrorist network. The analysis suggests that measures of centrality were at least superficially able to identify individuals in key network positions and also tended to highlight particular cells at times of operational importance.

Terrorism, by means of social media, has become one of the most pressing issues in the modern world. There is interplay between home-grown terrorist groups and international terrorist organisations which is playing a central role in accelerating the situations. Terrorist organisations are resorting to social media platforms with the aim of recruiting, training and communicating with their followers, supporters, donors, as it is an effective method of communication. Social media apps and file-sharing platforms, mostly used is Ask.fm, Facebook, Instagram, WhatsApp, PalTalk, kik, viper, JustPaste.it, and Tumblr. Encryption software like TOR is used in communications with journalists to obscure location information. But circumstances contribute to make Twitter the most prominent application. Most social media platforms require either 3G or Wi-Fi access but Twitter can function in the absence of either.

II. TERRORISM AND STRUCTURE OF TERRORIST NETWORKS

A. Definition of Terrorism

Terrorism [16] is unlawful violence aimed at human objects, considering that:

- It was instituted with the objective of reshaping or preserving a commonly regarded norm in a particular population.
- It had clandestine features allowing the actors to cloak their personal identity and their prospective locations.
- It wasn’t initiated for the long-lasting defence of some area;
- It was not traditional warfare due to the concealed personal identity, future location, threats, spatial mobility, and lesser susceptibility of participants to traditional military action.
The participants recognized the act as supporting the standard goal by instilling fear in individuals besides the target of the planned violence by promulgating a cause.

Some of the Foreign Terrorist Organizations (FTO) as designated by U.S. Department of State are [5]:

- Islamic State of Iraq & the Levant (ISIL) (previously al-Qaeda in Iraq)
- Tehrik-e Taliban Pakistan (TTP)
- al-Qaeda
- Jaysh Rijal al-Tariq al Naqshabandi (JRTN)
- al-Shabaab
- Harakat ul-Jihad-i-Islami (HUJI)

B. Structure of Terrorist Organizations

Terrorist organizations can be considered graphs with the actors or sub-groups (cliques) being the nodes. The general working and operational structure of an extremist organization was presented by the Iraqi counter terrorism unit [8]. The report inferred that the organization of modern terrorist groups is decentralized. [Fig. 1] The network leader generally functions as the mentor, motivator and goal-setter for the organization. While the finance manager being the de-facto leader of the organization is responsible for achieving the goals by carrying out the actual activities.

However the group comprises of other managers with finite roles like formulating media propaganda, handling matters pertaining to security of the agents and arranging equipment for belligerent operations. All the operatives with supervisory roles report to the finance manager. Moreover, the finance manager also has the authority and contacts necessary to directly command the operatives engaged in militant operations. Thus a finance manager has a direct contact with most nodes in the organization, which signifies that the finance manager is the most central and active node in the group. Also, the finance manager is the only actor, the leader of the group has a direct contact with. [4]

Incompleteness: Terrorist network data is inexorably incomplete; which refers to certain existing nodes and links that will remain unobserved and unidentified.

Fuzzy Boundaries: The bounds of extremist networks are obscure, and hence there’s no specific criterion so as to which actors should be included or excluded.

Dynamic: Terrorist networks are dynamic in nature of their operations. The connection between any two nodes cannot be expressed in a binary form, existent-nonexistent or strong-weak, but rather has dissemination over time.

III. SOCIAL NETWORK ANALYSIS AND ITS APPLICATIONS

Social Network Analysis is the application of network theory and graph with the objective of investigating the synergy between the distinct elements of a social network. A social network consists of nodes (individuals, institutions etc.) present in the network, and ties exhibiting liaison between them [Fig. 2]. Liaison may be in the form of real world offline social networks (like friendship, kinship, communication, transaction etc.) or may be online social networks (like Social Media sites, Discussion Forums, Chat Rooms, Gaming Apps that allow players to communicate while playing etc.).

Although SNA finds numerous applications in the fields of Information Science, Political Science, Organizational Studies, Economic and Business Analysis, Social Psychology, Biology, Communication Studies, Intelligence Analysis. While business giants use various measures of SNA to develop strategies and policies according to the customers. The application of SNA by the intelligence and counter terrorism units became popular after the tragic 9/11 attacks. Numerous SNA measures have been used for studying the reciprocal action among actors, inspecting the strength of ties among elements, identifying principal players and subgroups in network, finding distribution of nodes in the network. It is essential that all the available factors should be considered while studying the extremist organizations, so that more effective decisions can be made. Visual depiction of networks is the finest method for conveying complex information, representing structural properties through quantitative analysis.
A. Gathering the data for Social Network Analysis

- **Offline Sources**: Collecting data for SNA from offline sources, that consists of publicly available data such as news articles, official documents released by national investigative agencies, textual analysis, court records, tax records, real estate and rental records, vehicle registration records etc. Few publicly available free services and database include GTD [9] and GDELT [3]. The Global Terrorism Database (GTD) is an open-source database with organized information on terrorist activities globally, consisting of 180,000+ cases. The GDELT Project monitors the world's broadcast, print, and web news from almost every country in about 100 languages and identifies the people, locations, organizations and events driving the global society. Using the above stated services in addition to manual data collection, can yield more relevant data for modelling the terrorist networks.

- **Online Sources**: Data collection from online social networks include the extraction of public and private data of users, groups and pages, which contain posts, tweets, likes, comments, photos, videos etc. A number of tools and APIs are available for extracting data from various online social networks like Facebook, Twitter, YouTube and few more. Facebook graph API, Twitter API, Netvizz [26] and NodeXL [6] are such tools for extracting social network data and further analysing most of the criminal and terrorist activities using online social networks. Owing to the rise and adoption of technology in the 21st century by everyone, even the extremist organizations have shifted a significant amount of operations to the internet. A significant source of recent and mostly genuine information can be through the posts by the publicity wings of the extremist organization on various social media websites. The posts, pages, hash tags and accounts related to these organizations. An effective way of extracting information is Social Media Data Mining.

- **Data Extraction of Social Media**: As data become more prominent and readily available, the temptation to analyze them and make sense of the world through specific analytics methods algorithms grows. In the past decade, agencies like the CIA and the NSA have institutionalized big data through the development of dedicated analytics units and research and development projects focusing on the analysis of online data such as YouTube videos and social media posts. We identify five ways big data analytics supports national security decision-making, anomaly detection, Association, Classification and clustering, Link analysis etc. There are predominantly two ways to extract data from social media websites - using official API and Building a crawler. There is a package for extracting Facebook data using R and Facebook API called Rfacebook Package, which basically provides an interface to the Facebook API. One can find the manual from CRAN repository. Data mining techniques such as clustering, classification, association rule mining, and visualization is been used for web mining. Classification and clustering can be used to create different classes of users. In classification classes are predefined (supervised) and in clustering they are not predefined (unsupervised). Association rule mining technique is used to discover direct or indirect relationships between web entities. Visualization is a special technique to present data and information in graphical, understandable manner and plays an important role in web structure mining.

IV. METHODS FOR SOCIAL NETWORK ANALYSIS

The objective of performing SNA is to pigeonhole and determine the roles played by the nodes within the network. The analysis assesses the direct or indirect connections between the nodes present in the network, with the purpose of capturing the most significant nodes and get an insight into the flow of information within the network.

Time and again, centrality, one of the vital properties of a social network has been employed for studying the distribution of terror cells within a network. There are numerous measures of centrality that decide the significance and influence of a node within the network considering various facets of its relations with the remaining nodes. Frequently used centrality measures are degree centrality, betweenness centrality and closeness centrality.

A. Degree Centrality

Degree Centrality [18] is a measure of the significance of the node in a network depending on how many other nodes does the subject node has direct relationships with. A node with a higher value of degree centrality can be treated as a focal point in the mainstream of information flow within the network. In terror networks, it helps identify the hubs of information and the people that these particular nodes can lead to. However in the cases of highly organized covert networks, it highly unlikely for nodes with a high degree centrality to be the leader.

\[
C_D = \sum_{i=1}^{n} a(y_i,x)
\]

\[
C_D' = \frac{C_D(x)}{n-1}
\]

Where, ’n’ is the no. of nodes in the network.
\(a(y_i,x) = 1\) if there’s a direct relation between node x and node y , else it is 0.

’\(C_D\)' is the count of the degree of direct relationships that node ‘x’ has with other nodes within the network.

\(C_D(x)\) is defined as relative centrality of node ‘x’ in the network.

\((n-1)\) is the normalizing factor.

B. Betweenness Centrality

Betweenness Centrality [18] is a measure for determining the nodes that possess a greater probability of being positioned on the geodesic path between other distinct nodes, in other words it is an index of the potential of point for control of communication in a network Nodes with greater magnitude of betweenness centrality, often act as bridges and regulate the flow of...
information among various groups of nodes in the network. In terrorist networks, nodes with high betweenness value are usually the brokers or middlemen in the network. Such nodes usually hold a privileged position in the network and can exercise control over the propagation of information inside the network. However, it also serves as a bridge-with the risk of being the potential sole point of failure, causing the network to collapse. Disconnecting such nodes can effectively cause a breakdown in the internal communications within the network.

\[
C_B(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{g_{ij}(x)}{g_{ij}}
\]  

(3)

\[
C_B(x) = \frac{C_B(x)}{(n-1) \cdot (n-2)}
\]  

(4)

Where, ‘n’ is the no. of nodes present in the network.

\('g_{ij}(x)\) is the no. of geodesic (shortest) paths between nodes \(i,j\) that contain node ‘x’ as an intermediary.

\('g_{ij}\) is the no. of geodesic paths between nodes \(i,j\)

\('C_B(x)\) is the probability of node x lying on a random path linking nodes \(i,j\)

\(C_B(x)\) is defined as betweenness centrality of node x in the network, relative to other nodes.

\((n-1)(n-2)\) is the normalizing factor.

C. Closeness Centrality

Closeness Centrality[18] is yet another measure of centrality, which depends on the closeness of a node from rest of the nodes in the network. It is defined as the average of the geodesic path between a subject node and the other nodes directly linked to it. A high value of closeness centrality exhibits the independence of the node and its ability to propagate or receive information via other nodes in the network in the shortest possible time. In terrorist networks, it may provide useful insights into identifying messengers in the network. Capturing of targets with high closeness centrality can lead to an enormous amount of information about the other nodes and the network.

\[
C_c(x) = \frac{n-1}{\sum_{i=1}^{n} d(p_i, x)}
\]

(5)

Where, ‘n’ is the no. of nodes existing in the network.

\('d(p_i, x)\) is the distance between node \(p_i\) and node x.

\('C_c(x)\) is the closeness centrality of the node in the network.

V. RESULT

The analysis of data from the Global Terrorism Database for the terrorist attacks in the year 1993 exposes a positive correlation between the number of successful terrorist attacks (Fig 3) and the number of suicide attacks (Fig 4) throughout the world. The result is that the extremist organizations with the primary targets being India, United Kingdom and Eastern part of the African continent resort to suicide attacks in order to fulfill their agenda, with high success rates.
Table- I: Available Tools for Social Network Analysis

| Tool   | Operating System | License          | Acceptable Formats | Functionality                                                                                          | Limitations                                                                                                                                 |
|--------|------------------|------------------|-------------------|--------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| JUNG [12] | All Platforms | Open-Source | .net, .graphml   | ➔ Java API  
 ➔ Modeling library  
 ➔ Visualization and analysis of Social Networks  
 ➔ Availability of variety of algorithms for centrality flow, clustering etc.  
 ➔ Suitable for working with large datasets. | ➔ Absence of interactive interface.  
 ➔ Inappropriate for Dynamic Network Analysis. |
| Igraph [10] | Windows, Linux | Open-Source | .gml, .graphml, .net | ➔ Python, C and R libraries for network analysis.  
 ➔ Numerous algorithms for centrality, clustering etc.  
 ➔ Added support for two mode networks.  
 ➔ Capable of handling huge and complex networks. | ➔ Absence of user interface.  
 ➔ Lacks support for direct importing from Online social networks. |
| Network X [2] | Windows, Linux | Open-Source | .gml, .graphml, .net | ➔ Python support for network analysis.  
 ➔ Support for temporal and two mode networks.  
 ➔ Capability to handle large networks with ease. | ➔ No user interface.  
 ➔ No direct graph importing.  
 ➔ Less centrality measures. |
| Pajek [14] | Windows, Linux | Open-Source | net, .dl          | ➔ Visualization and SNA for huge networks  
 ➔ Supports Two Mode, temporal and multi relational networks.  
 ➔ Centrality algorithms  
 ➔ Graph Layouts | ➔ Limited number of nodes.  
 ➔ Less customization. |
| Gephi [7] | Windows, Linux | Open-Source | Databases, .gexf, .gml, .gdf, .dot, .graphml, .net, .csv, .dot, .dl | ➔ Visualization and SNA.  
 ➔ Variety of visualization layouts.  
 ➔ Availability of various centrality, modularity and clustering measures.  
 ➔ Allows ranking and partitioning the networks.  
 ➔ Plug-in support.  
 ➔ Timeline Feature. | ➔ Not suitable for very large datasets. |
| NodeXL [6] | Windows | Open-Source | matrix formats, .graphml, .dl, .net | ➔ Has a dedicated MS-Excel template  
 ➔ Various visualization layouts and SNA measures.  
 ➔ Supports direct data import from some specific Online Social Networks.  
 ➔ Data export in multiple formats. | ➔ Incapable of handling and visualizing large datasets. |
| ORA [11] | Windows | Commercial & Academic | .xml, .dynetml, .zip | ➔ Supports two mode, multi relational and dynamic network analysis.  
 ➔ Availability various visualization layouts.  
 ➔ Numerous centrality, clustering and power measures. | ➔ Platform dependent, Limited number of nodes. |
| UCINET [13] | Windows | Commercial & Academic | .dl, .net, .vna, .csv, Raw matrices | ➔ SNA  
 ➔ Added NetDraw visualization support.  
 ➔ Numerous Power, Centrality, Community and Clustering algorithms. | ➔ Platform dependent.  
 ➔ Not Scalable. |

**VI. CONCLUSION**

To draw a definite picture of a terrorist network, it is essential to map the ties between the operatives in the network. With the few high priority targets being the various managers, messengers and especially the finance manager. SNA can be considered as one of the most effective and powerful tools for in-depth analysis of terrorist and criminal networks.
Numerous algorithms of SNA measures like: community detection, clustering, and information flow and centrality are extremely productive while identifying significant nodes in the network, communication patterns and other valuable information that can be useful information while planning strategies to destabilize the network.

These measures prove to be suitable for understanding large terrorist and covert criminal networks. The terrorist networks function in a manner similar to business organizations. Hence, mapping the same relationships would provide us with helpful information about these networks[Table 2]. However, gathering this information due to the covert nature is a difficult task.

| Relationship          | Data Sources                                                                 |
|-----------------------|------------------------------------------------------------------------------|
| Trust                 | Family connections, school, organization, club friends or acquaintances, neighbours. Governmental or other records regarding the subject. Available in the subject’s primitive country. |
| Task                  | Accounts of voice and video calls, text messages, e-mails, instant messages, internet search history. Presence at meetings and attendance at social or community, events i.e. surveillance |
| Money & Resources     | Logs and documents of banks accounts. Patterns and geographical location of money transfer, credit, debit or ATM cards. Prior police or court documents. Surveillance – use of alternative sources of banking like Hawala. |
| Strategy & Goals      | Internet Search history, Social Media usage. Audio, video or securely encrypted disks or other objects delivered through courier. Travel history. Surveillance: Presence at social and community events. |

Also it is essential to study, what make these networks effective and improves their performance. This can be performed using the network and counter network [19] levels of analysis and functioning. The levels being:

- Resource and Recruitment – Available resources and skills of human resources
- Organizational level – Structure of the organization
- Narrative level – The information (propaganda) being spread.
- Doctrinal level – Adopted collaborative strategies and policies.
- Technological level – Technology employed.
- Social level - The personal connections that guarantee faithfulness and integrity.

Of course, it isn’t possible for a common researcher to have access to this data, due to its highly confidential and covert nature. Hence, the researchers most commonly use the open source data available, but the authenticity of such data is a major concern, except for the data collected from public court proceedings which is the best source of reliable data. Other serious issues are the large size of the networks, incomplete nature of data, fuzzy boundaries and extremely dynamic nature of relationships between nodes, which is waxing and waning from time to time. While for online data, intruding into the privacy of individuals is also a sensitive issue. Inculcating the temporal and spatio temporal factors from social media data can provide a great boost in the process.

Considering the challenges faced currently, the foremost aim is to create a strong and authentic database of various designated terrorist groups. One way for the governments, intelligence agencies and counter-terrorism units to pool in their information and consider the overlapping information as reliable. To get a clear picture of the possible danger, the sharing of information and knowledge on a global level is essential. In a journey to fight terrorism, the first step is to build a very strong knowledge sharing network, together.

REFERENCES

1. Choudhary, Pankaj, and Upasna Singh. "A survey on social network analysis for counter-terrorism." International Journal of Computer Applications 112.9 (2015): 24-29.
2. Hagberg, Aric, et al. "Networks. High productivity software for complex networks." Webóvá stránka https://networkx.lanl.gov/wiki (2013).
3. Leetaru, Kalev, and Philip A. Schrodt. "Gide: Global data on events, location, and tone, 1979–2012." ISA annual convention. Vol. 2. No. 4. Cnieser, 2013.
4. Berzinji, Ala, Lisa Kaati, and Ahmed Rezine. "Detecting key players in terrorist networks." 2012 European Intelligence and Security Informatics Conference. IEEE, 2012.
5. "Foreign Terrorist Organizations" U.S. Department of State. https://www.state.gov/foreign-terrorist-organizations/, retrieved August 12, 2019.
6. Smith, Marc, et al. "NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010." (2010).
7. Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. "Gephi: an open source software for exploring and manipulating networks." Third international AAAI conference on weblogs and social media. 2009.
8. Memon, Nasrullah, et al., "Novel algorithms for subgroup detection in terrorist networks." 2009 International Conference on Availability, Reliability and Security. 2009.
9. LaFree, Gary, and Laura Dugan. "Introducing the global terrorism database." Terrorism and Political Violence 19.2 (2007): 181-204.
10. Csardi, Gabor, and Tamas Nepusz. "The igraph software package for complex network research." InterJournal, Complex Systems 1695.5 (2006): 1-9.
11. Carley, Kathleen M., and Jeff Reminga. Ora: Organization risk analyzer. No. CMU-SRI-04-106. CARNEGIE-MELLON UNIV PITTSBURGH PA INST OF SOFTWARE RESEARCH INTERNAT, 2004.
12. O’Madadhain, Joshua, et al. "The jung (java universal network/graph) framework.” University of California, Irvine, California (2003).
13. Borgatti, Stephen P., Martin G. Everett, and Linton C. Freeman. "Ucinet for Windows: Software for social network analysis." Harvard, MA: analytic technologies 6 (2002).
14. Batagelj, Vladimir, and Andrej Mrvar. "Pajek-program for large network analysis." Connections 21.2 (1998): 47-57.
15. Sparrow, Malcolm K. "The application of network analysis to criminal intelligence: An assessment of the prospects." Social networks 13.3 (1991): 251-274.
16. Gibbs, Jack P. "Conceptualization of terrorism.” American Sociological Review (1989): 329-340.
17. Krebs, Valdis E. "Mapping networks of terrorist cells.” Connections 24.3 (2002): 43-52.
18. Freeman, Linton C. "Centrality in social networks conceptual clarification.” Social networks 1.3 (1978): 215-239.
19. Ronfeldt, David, & John Arquilla. "Networks, netwars and the fight for the future: First Monday [Online], 6.10 (2001): n. pag. Web. 14 Aug. 2019.
20. Agarwal, S., Sureka, A.: A focused crawler for mining hate and extremism promoting videos on youtube. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media, pp. 294–296 (2014).

21. Sureka, A., Agarwal, S.: Learning to classify hate and extremism promoting tweets. In: Intelligence and Security Informatics Conference (IISIC), 2014 IEEE Joint, pp. 320–320 (2014). IEEE.

22. Fisher, A.: How jihadist networks maintain a persistent online presence. Perspectives on Terrorism 9(3) (2015).

23. Yannakogeorgos, P.: Rethinking the threat of cyberterrorism. In: Chen, T.M., Jarvis, L., Macdonald, S. (eds.) Cyberterrorism, pp. 43–62. Springer (2014).

24. Rowe, M., Stankovic, M., Dadzie, A.-S. (eds.): A Topical Crawler for Uncovering Hidden Communities of Extremist Micro-Bloggers on Tumblr (2015)

25. Zhao, L., Chen, F., Dai, J., Hua, T., Lu, C.-T., Ramakrishnan, N.: Unsupervised spatial event detection in targeted domains with applications to civil unrest modeling. PLoS ONE 9(10), 110206 (2014)

AUTHORS PROFILE

Prof. Ashlesha S. Nagdive completed Bachelors of Engineering in Information Technology in 2008 from Amravati university and Masters of Engineering in Embedded Systems & Computing from G H Raisoni College of Engineering Nagpur in 2011. Currently pursuing PhD from Amravati university. Also working as Assistant Professor in Information Technology department at G H Raisoni College of Engineering Nagpur since 2010. Member of IEEE and published various papers in International Journal and conferences. Area of interest is Big Data & Hadoop, Data Analytics, data visualization.

Dr. R.M Tugnayat, Principal of Shri Shankarprasad Agnihotri college of Engineering Wardha. He has completed his PhD from Nagpur university. He has more than 20 years of teaching experience and Research Experience. He is a member of IEEE. He has publications in various International conferences and Journals. Subject of Expertise Software Engineering, BigData, Computer Networks and Image Processing.

Atharva Peshkar, Student Pursuing Bachelors of Engineering in Information Technology at G H Raisoni College of Engineering Nagpur. He is a member of IEEE Computational Intelligence Society. He has publications in International conferences and Journals. His interest and experience lie in applying Machine Learning and Computer Vision to domains ranging from Healthcare to Defence Technologies.