A Survey on Social Network Analysis for Counter-Terrorism

Poorva Singhania [1] , Akhilesh Tiwari [2]
Department of Computer Science & Engineering
[1,2] Madhav Institute of Technology & Science, Gwalior.

Abstract- *Fear monger Activities worldwide has moved toward the advancement of different high-finished strategies for dissecting psychological militant gatherings and systems. Existing examination found that Social Network Analysis (SNA) is a standout amongst the best and persistent technique for countering fear mongering in interpersonal organizations. The review inspected different SNA measures for foreseeing the key players/principle performers of fear based oppressor arrange regarding worldwide and additionally Indian viewpoint. Similar review among SNA instruments showed the pertinence and achievability for on the web and disconnected informal communities. It is prescribed to fuse transient examination utilizing information mining techniques. It can upgrade the capacity of SNA for giving element conduct of online interpersonal organizations. This paper plots an assortment of measures and systems utilized as a part of SNA for counter-psychological oppression exercises. Segment 2 traces the survey of research done utilizing different SNA measures in the field of counter-fear based oppression examination. Segment 3 gives itemized correlation of different apparatuses and procedures for on the web and disconnected SNA. It additionally highlights the SNA measures in the field of counter-terrorism analysis.*

Keywords- Social Network Analysis, Terrorist Networks, Counter-Terrorism, Centrality, Online and Offline Social Networks.

I. INTRODUCTION

Informal community Analysis (SNA) is the use of system hypothesis to break down interpersonal organizations as far as social connections. It involves hubs (performers, people, associations and so on.) inside the system and ties speaking to connections (fellowship, family relationship, discussion, money related exchange and so forth.) among the hubs. Social connections might be as certifiable disconnected informal organizations (like companionship, family relationship, correspondence, exchange and so forth.) or it might be online interpersonal organizations (like Facebook, Twitter and so forth.). Different SNA measures has been utilized for speaking to association among on-screen characters, inspecting solid or feeble ties, recognizing key/focal players and subgroups in system, discovering topology and quality of system. As of late SNA has showed up as a practice in different areas. It is fundamentally connected in Information Science, Political Science, Organizational Studies, Social Psychology, Biology, Communication Studies, Business Analysis, Economics and Intelligence Analysis. Facebook, Twitter and couple of more informal communication destinations utilize different measures of SNA to create systems and approaches for clients. Concentrates in light of SNA in Counter-Terrorism got to be distinctly prevalent soon after the assaults of 9/11. SNA has broadly been utilized by the knowledge and law authorization offices for comprehension the structure of fear monger systems and creating procedures to upset them by perceiving pioneers and concealed examples in the criminal and psychological oppressor systems. Some normal uses of SNA in Counter-Terrorism are Key-Player Identification, Community Discovery, Link (connection) Analysis, Node (on-screen character) Discovery and Dynamic Network Analysis and so on

II. RELATED WORK

In this section, we present a general idea of research work done in the field of SNA with a more broad set of prior research focused on counter-terrorism and analysis of criminal networks.

2.1 Social Network Analysis and its Measures

Different measures have been created after some time for breaking down the informal organization in term of recognizing key-players, group location, discovering design in system, hub and connection revelation and so on. Centrality is one of the ordinarily considered ideas in informal organization investigation for distinguishing key players. A few measures have been produced for centrality, including degree, closeness, eigenvector centrality, data centrality, stream betweenness [8], Katz's impact [9] measure and so forth. Freeman [1] proposed three distinctive natural idea of centrality in particular degree, betweenness and closeness centrality, which are for the most part utilized as a part of recognizing key players in the informal organization. Bonacich [2] [3] proposed eigenvector centrality for discovering relative significance of a hub, which is for the most part utilized for characterizing the impact of a hub on its

neighboring hubs. Everett and et al. [4] in their work, developed the centrality measures of degree, closeness and betweenness to apply to gatherings and classes and in addition people. UCINET, an apparatus for dissecting interpersonal organization information was discharged then by Everett and et al. [5] with the vast majority of the regular SNA measures for investigation. Borgatti [6] proposed a typology of system streams relying upon two measurements of variety. Everett [7] concentrated informal organizations containing negative ties, for example, despise, dodging, and strife and proposed another centrality measure, PN centrality for both positive and negative ties.

2.2 Social Network Analysis in Counter Terrorism

2.2.1 Global Perspective

After unfortunate 9/11 assaults, different reviews included the utilization of informal community investigation in countering psychological warfare and wrongdoing. Key-player distinguishing proof, group disclosure, undercover system examination, dynamic system investigation, and disturbance of fear based oppressor systems are the most widely recognized reviews among them.

Krebs [10], considered and mapped the 9/11 Al-Qaeda fear monger arrange by get-together freely accessible data on 19 thieves of Al-Qaeda and applying essential SNA centrality and group measures with the assistance of SNA instruments to recognize the key p layers and pioneers in the system. This examination gives some vision for further work and research into the fear monger systems investigation.

Carley [11] concentrated informal organization examination and multi-specialist models for destabilizing systems. It clarify the difficulties in upsetting systems that are extensive, circulated and dynamic in nature. It likewise proposed multi-operator demonstrate for specialists, learning and undertakings in light of different meta-systems (like: interpersonal organization, information arrange, data arrange, task organize and so forth). Carley [12] proposed an approach for assessing destabilization methodologies for incognito fear based oppressor arranges by utilizing information in freely accessible records, daily paper reports, and expert diaries.

Jennifer Xu [13] sorted existing criminal system investigation methodologies and devices for recognizing subgroups, finding their examples of cooperation, discovering key players, and revealing hierarchical structures.

Memon [14] proposed a calculation for making chain of importance of the secret systems and for better comprehension the structure of the casual or irregular associations, keeping in mind the end goal to disturb these systems. The calculations are checked by utilizing open source dataset.

A review of the historical backdrop of informal organization examination in counter-fear based oppression research is given by Ressler [15]. The review was centered around different scholarly, barrier and government exercises for information gathering and demonstrating for fear mongering systems.

Everton [16] proposed to consider general geography of interpersonal organization before molding systems for their interruption by utilizing SNA centrality measures for distinguishing key players. Estimate, number of assaults and system versatility must be incorporated for adequately processing the viability of psychological militant systems.

Roberts [17] investigated motor and non-active ways to deal with disturbing dim systems. The active approach includes forceful and hostile measures to dispose of or catch arrange individuals, while the non-motor approach includes the utilization of SNA for battling dull systems. To clarify both these methodologies they utilized system information from of Noordin Top's fear monger arrange (operational and put stock in system) of South East Asia.

Everton [18] examined and arranged the Noordin Top's psychological oppressor organize from 2001-2010 with a specific end goal to identify changes after some time in system structure (i.e. thickness, centralization) and viability (i.e. Enlistment of individuals, nature of individuals). This review proposed that, thick yet decentralized psychological oppressor gatherings are more compelling and more hard to disturb.

Klausen [19] concentrated the interpersonal organization of YouTube record holders connected with al-Muhajiroun's posts of jihadist and fierce substance on YouTube by applying different SNA measures.

Utilizing interpersonal organization examination and media based confirmations, Wu attempted to locate the conceivable pioneer after Bin Laden by inspecting the Al-Queda psychological militant system by considering system position as far as social capital as an imperative component for recognizing the new pioneer.

Sarvari [21] made a vast scale interpersonal organization chart through an arrangement of freely spilled email locations of a few crooks and by recognizing Facebook profiles associated with these email addresses. By applying

different informal community investigation measures they perceived profiles of high rank culprits, different criminal groups and open gatherings on Facebook.

2.2.2 Indian Perspective

Saxena [22] concentrated the system of psychological oppressor associations, working effectively in State of Jammu and Kashmir of India by distinguishing their relations, and after that by applying different techniques for examination to recognize the key players among these associations.

With the assistance of open or insight information and applying strategies for interpersonal organization examination Basu [23] concentrated real fear monger bunches working in India and their linkages to global psychological militant gatherings.

Azad [24] investigated the system of 26/11 fear assaults of Mumbai (India), in view of the correspondence between ten aggressors and their handlers in Pakistan. Utilizing different system investigation measures they recognized the key pioneers and sub bunches inside the fear monger organize.

Karthika [25] contemplated and analyzed the examination work done in the field of informal organization investigation for clandestine systems and arranged different methodologies in interpersonal organization investigation like recognizing key player and subgroups, destabilizing fear monger and incognito systems, dynamic system investigation etc.

III. TOOLS AND TECHNIQUES USED FOR SNA IN COUNTER-TERRORISM

A social network is an architecture which consists of set of nodes (as actors or organizations) and relationships between these nodes. These social network may be online as well as offline in nature. Online social networks are basically in the form of online social networking sites (like Facebook, Twitter, YouTube, Google+, LinkedIn etc.), which allow users to interact with other users through sending messages, posting information, images, videos, likes and comments on them. On the other hand offline social networks are the real life social networks based on the relations like friendship, kinship, communication, financial transactions, locations, events etc.

3.1 Methods for Data Collection

In recent works the data is collected from online as well as offline social networks.

3.1.1 Data Collection from Online Social Networks

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 [27] are such tools for extracting social network data and further analyzing most of the criminal and terrorist activities using online social networks.

3.1.2 Data collection from offline social networks

Data collection for offline social network analysis involves the public and open source data from various news articles, reports, phone call records, textual analysis etc. For offline social network analysis of terrorist and criminal networks, few services and databases are available like GTD [28] and GDELT [29].

Global Terrorism Database (GTD) is service which provides open source data for various terrorist and criminal activities around the world. GDELT is a project which monitors and extracts the network of individuals, events, organizations and locations based on the news media and other open source information. Using these services with manual data collection, more relevant data can be gathered for modelling and analyzing terrorist and criminal networks.

3.2 Methods for Analysis

SNA provides various measures for analysis of central node within the network, overall importance of a node in the network, highly connected node in network, groups and sub-groups in the network and flow of information within the network.

The measure of importance and position of a node in the network is often defined by centrality. In analysis of terrorist network following types of centrality is commonly used:

3.2.1 Degree Centrality

Degree centrality [1] is defined as the number of direct links or connections to a node. A node with higher degree centrality value is often considered as a hub and an active entity in the network. In terrorist and criminal networks it helps in identifying number of persons that can be reached directly from particular person. It is not necessary that highest degree

centrality node is the leader in the network.

Generally social network G is represented as adjacency matrix A. Degree Centrality of node i in the network can be defined as,

$$D_i = \frac{\sum_{j \epsilon G} A_{ij}}{N-1}$$

3.2.2 Betweenness Centrality

Betweenness centrality [1] is a measure for identifying a node, which act as a bridge to make connections to other groups or communities in the network. It can be defined as the number of shortest paths between any pair that pass through a node. A node with higher betweenness centrality considered as powerful node with great amount of influence.

In terrorist and criminal networks it helps in finding a person may be a potential broker (having maximum information) between two groups or communities in that network.

Betweenness centrality of node i the network can be defined as:

$$B_i = \frac{\sum_{j,k \epsilon G} \frac{g_{jk}(i)}{g_{jk}}}{(N-1)(N-2)}$$

3.2.3 Closeness centrality

Closeness centrality [1] is a measure of how fast one can reach from a node to all other nodes in the network. It can be defined as the mean length of all shortest paths between a node and all other nodes in the network. A node with high closeness centrality is much closer and can quickly access the other nodes in the network.

In terrorist networks, it may be useful in identifying the person which can quickly access other persons in the network.

$$C_i = \frac{N-1}{\sum_{j \epsilon G} d_{ij}}$$

3.2.4 Eigenvector centrality

Eigenvector centrality [2] is a measure of relative importance in terms of influence of a node to its neighboring nodes in the network. It is generally used for finding the most central node in the network globally. A node with high eigenvector centrality is generally considered as a more central node with more influence over other nodes and act as a leader in the network.

In terrorist network analysis, it helps in identifying the person, well connected to other well connected persons.

$$X_i = \frac{1}{\lambda} \sum_{j \epsilon N b_i} A_{ij} . X_j$$

3.2.5 PageRank

PageRank [30] is a measure for computing the relative importance and ranking the nodes in the social network. Using PageRank in terrorist network analysis overall importance of a person can be determined based on its position in the network. Page rank can be defined as:

$$PR(a) = \sum_{b \epsilon N b_a} \frac{PR(b)}{L(b)}$$

3.3 CONTENT-BASED TERRORIST DETECTION METHODOLOGY

In the content-based detection methodology, the terrorist are detected by the Web traffic
content by monitoring all ISPs traffic. The prior knowledge about the terrorists is maintained as the content for training the detection algorithm. This detection should be carried out in real time.

3.3.1 Intrusion Detection System

The detection of the content from the existing sites and known terrorist traffic on the web is done by using Intrusion Detection System (IDS).The IDS constantly monitors the various activities within the network traffic to estimate the possible hostile attacks.

The IDS could be a computer, or computers within the network or the network itself. The IDS analyzes the various activities information gained and evaluates the possibility of the intruders within the network.

The measures evaluated by the IDS include accuracy, completeness, performance, efficiency, fault tolerance, timeliness and adaptivity.

3.3.2 Vector-space Model

The information evaluated by the IDS is represented in the textual content of Web pages. The document D is represented by the n-dimensional vector V= (v1, v2, v3…vn) where vi represents the frequency-based weight of term i in D. The similarity between the two documents represented as vectors may be computed by using one of the known vector

distance measuring methods such as Euclidian distance or Cosine. The cosine similarity measure is commonly used to estimate the similarity between an accessed Web page and a given set of terrorists' topic of interest [9].

3.3.3 Clustering Techniques

In the clustering technique, we perform cluster analysis-process of partitioning data objects (records, documents etc) into meaning groups or clusters so that the objects within the cluster have similar characteristics but are dissimilar to objects (data) in other clusters [5]. The unsupervised clustering is performed for classification of patterns. The clustering is performed on the Web documents by clustering them into documents of similar interest. A centroid is calculated for each collection of cluster and represented by vector space model.

3.4 Comparison of SNA Tools

Various tools and libraries are commonly used by researchers for social network analysis and visualization of terrorist and criminal networks. We considered following commonly used tools for comparison with respect to their functionality, platform, license type and file-formats respectively, based on the features like network visualization, computation of node-level and network-level measures (centrality, community, power and information flow) and handling of large networks with many nodes

Some basic tools used in SNA:

3.4.1. UCINET [5]: It Analysis social network data, It apply clustering and community algorithm. It has facility of exporting and importing various data format. It uses Window OS as a Platform. It is available in both academics and commercial license.
Limitatation: Platform dependent, Inefficient in dynamic Network analysis, Non scalable.

3.4.2.Gephi [31]: It analysis social network and visualization. It provides various layout for visualization. It also provide various measures of centrality. It also support time line features and plugin support
Limitatation: Not suitable for very large datasets

3.4.3.-ORA [32]: It provide various visualization layouts, also provide dynamic network analysis, it also have power clustering, it can report generation of analysis *Limitatation:* Platform dependent, Limited number of nodes.

3.4.4.-NodeXL [27]: It include templates for MS-Excel, it can visualize and anayalze the network, it can import data directly from few online social networks, it can export data in various format
Limitatation: Incapable of handling and visualizing large datasets

3.4.5.-Pajek [34]: It include the social network analysis and visualization for large networks, it provide algorithm for centrality and graph layouts.
Limitatation: Limited number of nodes, less customization.

Today's requirements of SNA tools for counter-terrorism research include the high-level visualization of terrorist networks, temporal analysis of networks over time, ability to analyze two-mode and multi-relational networks, dynamic network analysis of terrorist groups and use of data mining and big data analysis techniques for analyzing very large networks. concern.

Most of the times data tends to be incomplete with lot of missing and fake nodes and relations, which often leads to incorrect analysis result of such terrorist network. Also the terrorist networks are highly decentralized and dynamic in nature with continuously changing nodes and their relations. In case of online social network analysis the privacy of individual while crawling personal data is also very sensitive issue.

IV. CONCLUSIONS

SNA can be considered as one of the intense device for investigating the psychological militant and criminal systems. Different calculations of SNA measures like: centrality, group identification, bunching, and data stream are exceptionally compelling for finding the key players or pioneer, their example of correspondence and afterward building up the disturbing procedure for psychological oppressor systems. These measures are extremely useful in investigation of incognito, decentralized and huge fear monger arrange. However, likewise there are difficulties connected with the information accumulation and investigation of fear based oppressor systems. The information gathering for psychological oppressor systems is exceptionally troublesome for specialists as it is exceedingly private and secret in nature. Thus, specialist regularly utilizes information which is open source on open stage though the unwavering quality of this information is again a genuine concern. The majority of the circumstances information has a tendency to be fragmented with parcel of lost and fake hubs and relations, which regularly prompts to wrong examination consequence of such psychological militant system. Additionally the psychological militant systems are

profoundly decentralized and dynamic in nature with persistently changing hubs and their relations

REFERENCES

[1] Freeman, Linton C., (1979) "Centrality in social networks conceptual clarification", Social networks, Vol. 1, No. 3, pp215-239.

[2] Bonacich, Phillip, (1972) "Factoring and weighting approaches to status scores and clique identification", Journal of Mathematical Sociology, Vol. 2, No.1, pp113-120.

[3] Bonacich, Phillip, (1987) "Power and centrality: A family of measures", American Journal of Sociology, pp1170-1182.

[4] Everett, Martin G., & Stephen P. Borgatti, (1999) "The centrality of groups and classes", The Journal of mathematical sociology, Vol. 23, No.3, pp181-201.

[5] Borgatti, Stephen P., Martin G. Everett, & Linton C. Freeman, (2002) "Ucinet for Windows: Software for social network analysis".

[6] Borgatti, Stephen P., (2005) "Centrality and network flow", Social networks, Vol. 27, No.1, pp55-71.

[7] Everett, Martin G., and Borgatti, Stephen P., (2014) "Networks containing negative ties", Social Networks, Vol. 38, pp111-120.

[8] Freeman, Linton C., Borgatti, Stephen P., and Douglas R. White., (1991) "Centrality in valued graphs: A measure of betweenness based on network flow", Social networks, Vol. 13, No.2, pp141-154.

[9] Katz, Leo, (1953) "A new status index derived from sociometric analysis", Psychometrika Vol. 18, No. 1, pp39-43.

[10] Krebs, Valdis E., (2002) "Mapping networks of terrorist cells", Connections, Vol. 24, No. 3, pp43-52.

[11] Carley, K. M., Ju-Sung Lee, and David Krackhardt, (2001) "Destabilizing networks", Connections, Vol. 24, No. 3, pp31-34.

[12] Carley, K. M., Dombroski, M., Tsvetovat, M., Reminga, J., & Kamneva, N., (2003) "Destabilizing dynamic covert networks", Proceedings of the 8th international Command and Control Research and Technology Symposium.

[13] Xu, Jennifer, and Hsinchun Chen, (2005) "Criminal network analysis and visualization", Communications of the ACM, Vol. 48, No.6, pp100-107.

[14] Memon, Nasrullah, and Henrik Legind Larsen, (2006) "Practical algorithms for destabilizing terrorist networks", Intelligence and Security Informatics. Springer, Berlin Heidelberg, pp389-400.

[15] Ressler, Steve, (2006) "Social network analysis as an approach to combat terrorism: past, present, and future research", Homeland Security Affairs, Vol. 2, No.2, pp1-10.

[16] Everton, Sean F., (2009) "Network topography, key players and terrorist networks" Annual Conference of the Association for the Study of Economics, Religion and Culture, Washington, DC.

[17] Nancy, and Everton, Sean F., (2011) "Strategies for Combating Dark Networks", Journal of Social Structure.

[18] Nancy, and Everton, Sean F., (2011) "Strategies for Combating Dark Networks", Journal of Social Structure.

[19] Klausen, J., Barbieri, E. T., Reichlin- Melnick, A., &Zelin, A. Y., (2012) "The YouTube Jihadists: A Social Network Analysis of Al-Muhajiroun's Propaganda Campaign", Perspectives on Terrorism, Vol. 6, No. 1.

[20] Saxena, Sudhir, K. Santhanam, and Basu, Aparna, (2004) "Application of social network analysis (SNA) to terrorist networks in Jammu & Kashmir", Strategic Analysis, Vol. 28, No. 1, pp84-101.

[21] Basu, Aparna, (2005) "Social network analysis of terrorist organizations in India", North American Association for Computational Social and Organizational Science (NAACSOS) Conference.

[22] Azad, Sarita, and Arvind Gupta, (2011) "A quantitative assessment on 26/11 Mumbai attack using social network analysis", Journal of Terrorism Research, Vol. 2, No. 2.

[23] Karthika, S., and S. Bose, (2011) "A comparative study of social networking approaches in identifying the covert nodes", International Journal on Web Services Computing (IJWSC), Vol. 2, pp65-78.

[24] Rieder, B., (2013) "Studying Facebook via data extraction: the Netvizz application", 5th Annual ACM Web Science Conference, pp346-355.

[25] Smith, M., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., & Dunne, C., (2010)

[26] Leetaru, Kalev, and Philip A. Schrodt, (2013) "GDELT: Global data on events, location, and tone, 1979–2012", ISA Annual Convention, Vol. 2. Page, L., Brin, S., Motwani, R., & Winograd, T., (1999).

[27] "The PageRank citation ranking: Bringing order to the web". Stanford Digital Library Project.

[28] Bastian, M., Heymann, S., & Jacomy, M., (2009) "Gephi: an open source software for exploring and manipulating networks". ICWSM, 8, 361-362.