A Framework of Dynamic Visualization Tool for Terrorist Meta Crawler

R. D. Gaharwar
Assistant. Professor
G. H. Patel Department of
Computer Science and Technology,
Sardar Patel University,
Vallabh Vidyanagar, India

D. B. Shah
Professor
G. H. Patel Department of
Computer Science and Technology,
Sardar Patel University,
Vallabh Vidyanagar, India

ABSTRACT
Nowadays different types of web crawlers are available to perform variety of tasks. As internet is expanding the need of specialized web crawlers is also increasing with each passing days. Numbers of visualization tools which can be used to dynamically generate visual images are available in market. The need for visual representation of data increases in the case of counter terrorism. Terror activities can be easily and effectively represented as Terrorist Network Graphs where nodes can be terrorist organization/ terrorist and edges between these nodes can be communicational, ideological, monitorial etc relationships among them. This graph helps in deciphering the operations of terrorist camps. This paper presents the framework of dynamic Visualization Tool which is connected to Terrorist Meta Crawler and generates Terrorist Network Graphs as output. Moreover this tool will generate the output dynamically hence any topological change in the structure of terrorist networks will be effectively reflected in the Terrorist Network Graphs produced by this tool also.

Keywords
Terrorist Network Graph, Terrorist Networks, Terrorist Meta Crawler, Visualization Meta Crawler

1. INTRODUCTION
1.1 Network Analysis
Network Analysis is the study of different networks. Networks can be social where nodes are actors and links between nodes denote relationship between these nodes. [1] The study of different networks is done with the perspective to understand the characteristics of relationships of these actors. Interesting facts about network analysis is that it is indifferent towards the characteristics of individual actors and mainly focuses on the characteristics of their interactions. The larger web of social connection is created by putting together such relationships/links of every node so that later on such social web can be analyzed to extracts hidden patterns in them. Hence Network Analysis helps in explaining how an individual actor/node is related other actor/node which explains importance of that node in the network and aids in understanding network.

1.2 Terrorist Networks:
Terrorist Network is the special network where node in the network represents terrorist/ terrorist organization and link between these nodes are the interaction/ communicational relationships. [2] Activities carried out by terrorist/ terrorist organization is categorized under organized crime which is accomplished by systematic communication between them. Hence while committing any terror activity these terrorists/ terrorist organizations form a network of interaction which can be considered as terrorist networks. [3] These types of networks are studied with the help of Social Network Analysis tools to understand the structure of terrorist organizations.

1.3 Terrorist Meta Crawler (TMC)
Meta crawlers are the web spiders which refer to the links on the web pages and surfs to the other related web pages. [4] These web spiders keep on crawling from one web page to other for useful, related and appropriate information. Terrorist Meta crawlers are the specialized crawler tools designed to gather terrorist activities related information from the internet [5]. Any Terrorist Meta Crawler can have following general steps:

1. User interface
2. Relevance check
3. Filter results
4. Store results
5. User output

1.4 Visualization Meta Crawlers
There are specialized crawlers which gives visuals as output. This unique experience of search helps end users to easily understand and analyze the results. Simple crawlers explore the World Wide Web by traversing from one web page to another to output the related links in the form of simple text whereas Visualization Meta Crawlers delivers same output in graphical form. Visualization effects can be in the form of flash files, graphs, maps etc. End User can query Visualization Meta Crawlers like any other Meta Crawler but the output displayed to the end users will be in the form of visuals. This visual output increases the usefulness of the Meta Crawler. [6]

2. RELATED WORK
William H. Hsu studied the visualization techniques for quantitative data that may scale up to voluminous data sets. The authors used predictive and geospatial data but visualization techniques he described are generally applicable to large variety of different data sets.

The author featured the challenges of heterogeneous big data. Apart from high volume, big data is heterogeneous in nature. This heterogeneity arises due to the fact that information in big data might be distributed on multiple sites, multiple data models etc. Moreover meta data for such databases are also difficult to discover. Due to these special characteristics of big data, the visualization techniques used has to handle the problem of heterogeneity in data visual analytics for big data has to consider Human adaptability & interpretability of the system. Such system should be able to interact with other intelligent system operating in the environment. [7]

The author described different types of data and the visualization techniques used for them as follows:

1. Spatiotemporal Data: The data related to space scale and time scale can be considered as spatiotemporal data. Such data have both space coordinates and time coordinates. Visualization of spatiotemporal data can be done through map and timeline visualization.

2. Textual Data: The textual data extracted from research articles, journals etc can be classified into number of atomic elements using Named Entity Recognition (NER). Such data can also be visualized using tag clouds. Tag clouds show different words on bases of its occurrences in the articles but it have its own limitations.

3. Heterogeneous Data: Heterogeneous data are extracted from different web searches other social sites etc. Such data forms social networks and can serve as input to different intelligent system. Generally heterogeneous data is collected from multiple sources and classification and distribution of such data under uniform categories become difficult. Such data is visualized as graph topologies and graphical models are used for this purpose.

4. Predictive Data: The data that is used to predict the future events is called predictive data. This type of data is generally used to predict the disease outbreak and to mitigate its effects. The visualization tools used for such data are simulation based models.

The authors described various steps for Integrative Information Visualization which is as follows:

1. Collect data

2. Model development

3. Filtration

4. Collect different data set

5. Comparison

6. Set purpose

7. Integration

These are the general steps which can be used for any types of information visualization modeling.

The Lesev et al. designed a framework based on the ray tracing algorithms for interactive visualization for network data. The authors penned down main components of this visual analyzer tool as following:

1. Data logging system

2. Data log streaming

3. Filtering Data log

4. Data stream visualization

On network as data packets are transported, visualization tool will keep adding packets to the data structure created for storing data logs. This system is both multithreaded and distributed. The data collected during this phase is huge because of the inconsistency in packets arrival timings. This voluminous data need to store as a file which can be later retrieved for analyses purpose. These packets store data related to any scene like color, direction, light, position, geometry and material properties. The authors stored all the packets with same unique ID in a file which is stored on cloud and retrieved later dynamically for on line visualizations. Different types of filters are used on packets attributes like spatial filtering, light path expression filtering, time based filtering etc. These filtration techniques make it easy for end users to analyze and understand the effect of each attribute. [8] Hence the authors successfully created interactive visualization tool.

Azad studied 26/11 Mumbai terror attacks using mathematical tools like Social Network Analysis. The author studied that terrorist attackers were in constant touch with their handlers from Pakistan. Azad's study showed that there were 10 terrorist involved in these attacks. These terrorists captured 6 posh areas like Taj Mahal Hotel, Gateway of India, The Oberoi Trident Hotel, Chhatrapati Shivaji Terminal, Café Leopold and the Nariman house. The media reports showed details about who called whom during this attack. The author used this information to figure out the terrorist network of 26/11 terror attack in the form of a directed graph. In this graph 6 terror attack spots where consider as nodes and the intercepted phone call as the link between these nodes.

The author used different centrality principles like degree, closeness, betweeness etc to analyze and evaluate the role of each terrorist in 26/11 attack. [9] With the help of this information author was able to decipher 26/11 Mumbai attacks correctly and accurately.

Selberg et al. studied Terrorist networks and its topologies under the light of social network techniques. Any social network can be explained by both graphical and mathematical representations. The authors stated different properties of social network graphs which can help in evaluation of its like degree, geodesic path, closeness, bridge and hub. The authors described clustering as the tendency of nodes in the network to appear as connected triangle and assortativity as a measure of proximity of any node to the core of the network. On the bases of the above mentioned properties the authors classified complex and covert networks like criminal and terrorist networks into 3 categories as follows:

1. Random

2. Small-world

3. Scale-free

The author studied these networks and revealed that decoding of these covert networks can be done in 3 steps:

Step 1: Collecting the relevant data

Step 2: Use of existing mathematical or statistical models on the collected data

Step 3: Creating new mathematical models [10]

The basic drawback is the existing techniques are visualization tools developed so far generally focuses on big data and none of them specially focuses on terrorist networks only. Moreover the major development in the field of terrorist network visualization tools is limited to using readymade tools or manually drawing network graphs. Hence this paper purposes a framework for it which is customized for the visualization of Terrorist Network Graphs.

3. ARCHITECTURE

The current work in the field of terrorist network visualization is limited to the use to readymade tools available in market. In this paper the authors propose a framework for terrorist network visualization tool. This framework works on the data

collected by Terrorist Meta Crawler (TMC). Collecting data of terrorist networks is crucial task because such covert networks are hidden and the links between related nodes do not surface easily. Here the authors considered open sources such as news papers, journal articles etc as foundation of data. The data collected from these sources are kept in Terrorist Meta Crawler (TMC) database in the form of links between related terrorist organizations. This links fetched from the database are further used for the analysis. Every terrorist organization might not have related link to other terrorist organization. In this case no node will be created for such terrorist organizations in the output graph. Hence number of points on graph will be equal to number of terrorist organizations found in database with links to other terrorist organizations. Once the links are fetched and points are place on graph, all the points are considered as terrorist organization node and a line is sketch between the related terrorist organizations representing the relationship between them. This task is repeated for every terrorist organization in the database. The output of this framework is the Terrorist Network graph. This graph is created without the use of any readymade tool hence the degree of flexibility also increases. Moreover this architecture is customized for studying and analyzing terrorist networks only.
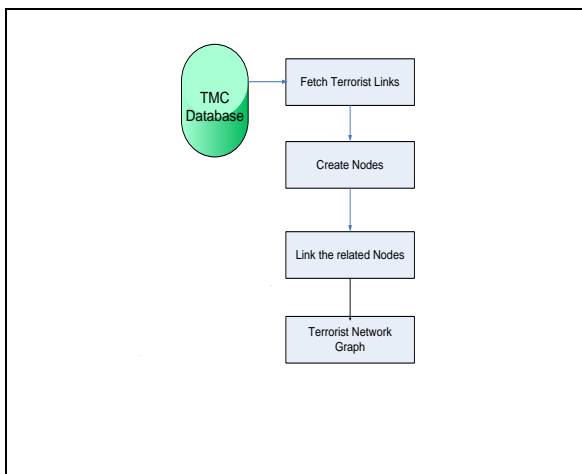
Fig 1: Visualization framework for Terrorist Meta Crawler

4. FUTURE WORK

This visualization tool aids in generating a visual representation of the terrorist network graph. This visual representation helps in understanding the relationships of terrorist organizations in the network. Later on centrality principles such as closeness, betweeness etc can be applied on it to understand role of each organization in the graph. Every terrorist organization present in the network has its special role in fulfilling the task, for example some terrorist organizations acts as a broker, some as communicator and some as apex/ leader. Removal of important nodes may help in disintegration of terrorist network.

5. CONCLUSION

This paper presents the framework for visualization tool designed to work specially in integration with Terrorist Meta Crawler. The Terrorist Meta Crawler is expected to have data related to terrorist organization and the relationship link between them. Once Terrorist Meta Crawler database is populated with such data; visualization tool starts to fetch the useful information from it. This paper describes that this visualization tool works in the following 3 steps: Fetching links, Creating nodes and linking nodes. In the end this tool gives Terrorist Network Graph as output. Terrorist Network Graph shows the visual representation of terrorist networks which is useful in understanding the role of each terrorist organization in the terrorist graph. Moreover Terrorist Network Graph generated by this visualization tool is dynamic in a way that as new links are added to the network or odd links are deleted from the network, these changes will be reflected in the output graph.

6. REFERENCES

[1] R. D. Gaharwar, D. B. Shah, and G.K.S. Gaharwar, "Terrorist Network Mining: Issues and Challenges," International Journal of Advance Research in Science and Engineering, vol. 4, no. 1, pp. 33-37, 2015.

[2] M. A. Shaikh and W. Jaixin, "Investigative Data Mining : Identifying Key Nodes in Terrorist Networks,"IEEE International Conference Multi Topic, 2006.

[3] N. Chaurasia, M. Dhakar, A. Tiwari and R. K. Gupta, "A survey on Terrorist Network Mining: CurrentTrends and Opportunities," International Journal of Computer Science & Engineering Survey (IJCSES),vol. 3, no. 4, pp. 59-66, August, 2012.

[4] Etzioni, Oren. "Moving up the information food chain: Deploying softbots on the world wide web." AI magazine 18.2 (1997): 11.

[5] R. D. Gaharwar, D. B. Shah, and G.K.S. Gaharwar, "The Study of Multi-Search Services for Terrorist Network Mining", International Journal of Computer Applications, vol. 147, no. 11, pp. 30-32, August 2016.

[6] http://www.seochat.com/c/a/search-engine-optimization-help/search-engines-and-algorithms-metasearch-engines-a-tool-for-seo/

[7] Hsu, William H. "Information Visualization Techniques for Big Data: Analytics using Heterogeneous." Packaging Digital Information for Enhanced Learning and Analysis: Data Visualization, Spatialization, and Multidimensionality: Data Visualization, Spatialization, and Multidimensionality (2013): 184.

[8] Lesev, Hristo, and Alexander Penev. "A framework for visual dynamic analysis of ray tracing algorithms." Cybernetics and Information Technologies 14.2 (2014): 38-49.

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

[10] Selberg, Erik, and Oren Etzioni. "The MetaCrawler architecture for resource aggregation on the Web." IEEE expert 12, no. 1 (1997): 11-14.