ISSN:2320-0790

An International Journal of Advanced Computer Technology

Survey on counter Web Terrorism

Dr. W. Jeberson, Lucky Sharma

Department of CSE & IT
SSET, SHIATS, Allahabad
UP, India

Abstract: Terrorist activities is a big issue for the whole humanity, and as the whole world is moving rapidly by utilizing IT innovations, similarly terrorist groups are also using IT technologies as powerful tool for destructive activities. Today every terrorist group has strong presence over the web.

In this paper we have surveyed various techniques and methods to analysis and prevent terrorist related activities. We found that that a timely identification of terrorist activities could help to prevent a large-scale public spread of communication exchange pertaining to the suspects/criminals' ideas, messages, and connections. Paper also consists of an idea for identification & destabilization of terror groups in social network.

1. INTRODUCTION

In criminal networks, there may exist groups or teams, within which members have close relationships. One group also may interact with other groups to obtain or transfer illicit goods. Moreover, individuals play different roles in their groups [2]. For example, some key members may act as leaders to control activities of a group. Some others may serve as gatekeepers to ensure smooth flow of information or illicit goods and some act as outliers in a group. To analyse such criminal networks, investigators must process large volumes of crime data gathered from multiple sources. This is a non-trivial process that consumes much human time and effort. Current practice of criminal network analysis is primarily a manual process because of the lack of advanced, automated techniques. When there is a pressing need to untangle criminal networks, manual approaches may fail to generate valuable knowledge in a timely manner.

In the modern day globalized world, the ease of terrorist network information exchange is characterized by contact moments through the internet and an absence of time and location restrictions [11]. The amount of information available to police forces is continuously increasing and many police forces are not ready for handling data amounts of this size. As a consequence, pro-actively observing potential threats to our national security becomes increasingly difficult As social network data is becoming more publicly available due to the advance of online social networking, social network analysis techniques are attracting increasing research interest in both academia and industry. In intelligence and security informatics domain, social network analysis techniques have been widely used to support intelligence and law enforcement force to identify suspects, gateways, and extracting communication patterns of terrorist or criminal organizations.

Social network analysis is useful for extracting the complex relationships between social actors. However, in practice, this capability is greatly compromised to deal with the limitation of partial data or anonymized data [12]. It is understandable that the social network analysis techniques are not able to extract the essential knowledge by using partial data of a terrorist or criminal social network. For example, each law enforcement unit has its own criminal social network collected by its own agencies but this network is only part of the global criminal social network. Mining on an incomplete criminal social network may not be able to identify the bridge between two criminal subgroups. Ideally, information sharing can resolve the problem. However, due to the privacy policy, different organizations are not allowed to share the sensitive information of their social network data. As a result, they have to anonymize the social network data before publishing or sharing it. But, even the modest privacy gains require almost complete destruction of the data-mining utility, which means it is hard to make a balance between privacy and utility in network data publishing.

2. LITERATURE SURVEY:

Researchers has done lots of concrete efforts in the field of MANET security one of the most appreciate work was carried by Muhammad Akram Shaikh, Wang Jiaxin in the paper titled "Investigative Data Mining: Identifying Key Nodes in Terrorist Networks" [2], in this novel work they presented an overview of investigative data mining with its basic framework and tried our best to shed some light on the issues. They believe that investigative data mining has a promising future for increasing the effectiveness and efficiency of counter terrorism and intelligence analysis. In addition to this they have also discussed the few available approaches and their shortcomings for destabilizing terrorist networks.

Another significant work was done by Zhaobo Luo, Xiaohong Jiang, Yi Feng and Chaolun Xia as "Dynamic Core Detection in Social Network Analysis" [3] as they

presented a dynamic core detection algorithm. Based on core features, this method is capable to use the information drifting over time and generate core parts in dynamic networks. Furthermore, carried out an experiment with terrorist attack data, and got results that are accordant with the true situation to a very great extent. The experiment shows that the dynamic core detection method could generate some meaningful results in social network analysis. The dynamic core approach is as follow.

In a network, nodes mostly represent persons, organizations or some other entities. Here they consider the features as what can transmit from one node to another node. Hence, can dig out clusters in a dynamic network by finding out the transmission pattern. Of course, not all features work well, and discussed how to choose the proper features called core features. Also, it should be noted that the dynamic network changes over time. Therefore, the aspect of transferring in time dimension should be taken under advisement. Intuitively, the latter activities would be affected by the former activities in time dimension. Consequently, the transferring of features from latter to former depicts the carrying forward knowledge, technology and even viewpoint, which, in a sense, implies the existence of relationship. In addition, the strength of potential relationship varies inversely as the length of the space of time.

In the year 2013 at "European Intelligence and Security Informatics Conference", Ala Berzinji, Frzand Sherko Abdullah, Ali Hayder kakei has presented a research paper on the titled "Analysis of Terrorist Groups on Facebook" [4], here they presented the way to analyze the terrorist group on FB.

According to them social media referees to the information of people in which they share create exchange and comment contents among themselves in virtual communications and networks. A social network is a social structure made up of a set of actors (such as individuals or organizations) and the relation between the actors. Most of the communications between individuals are made through Facebook. SNS (social network sites) are mainly attractive to the younger generations, and it is not uncommon for parents or grandparents to be active users. People, organizations and groups uses SNS for a variety of purposes. Here they try to monitor some terrorist groups on Facebook by human using Facebook Operation techniques and then analyze them by using algorithms to find the target. The case that is presented to collect the data on Facebook is named Facebook Operation. It is very recent, Special and unique. After getting the data of the target group the presented algorithm by us in this work will analyze the nodes and the relations among them, by using Social Network Analysis (SNA) witch it is a set of powerful techniques that can be used to identify clusters, patterns and hidden structures within social networks. The algorithm will detect the most active person in the network that try to increase the member rate of the group and the propaganda through this node will reach other nodes quickly.

FACEBOOK OPERATION:

In the platform for cyber-counterterrorism in the Facebook kingdom, the purpose of contemporary methods served the following purpose:

1- Gaining information about the user and his friends (connections) so as to construct a link chart.
2- Monitoring the target user's status, activities, together with the video and pictures published by the user.
3- Identifying areas of interest and potential target locations.
4- The target that are identified are simultaneously been worked on through Facebook.
5- Strategic and up-to-date information is collected on the target individuals and groups as the agent plays the role of a friend of the target.

At "Aberdeen Proving Ground, USA",[8] Kirk Ogaard carried very constructive research that, software can be used to automatically process and analyze data gathered from popular social media websites, such as Twitter. SNA software is often tested with small-scale data sets manually constructed by analysts to coincide with typical military scenarios. However, for SNA software to be practical for real world intelligence analysis, such software must be scalable when tested with large-scale data sets for which manual construction is too costly. An information morphing algorithm (InfoMorph) is presented. InfoMorph can generate large-scale synthetic data sets which follow specific scenarios by morphing existing large-scale real data sets. Morphing transforms the large-scale real data sets into large-scale synthetic data sets by replacing entity references according to a substitution table. In this paper we tested InfoMorph with two Twitter data sets: 1) 1,007 tweets from the Kandahar province in Afghanistan gathered in 2013 and 2) 738,717 tweets and 10,000 news articles gathered about the Egypt Unrest in 2011. Testing SNA software with Twitter data is important because tweets contain many abbreviations and acronyms which make them more challenging to parse. The first data set was morphed to coincide with part of a scenario designed for the Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR) On the Move (OTM) exercise in 2013. The second data set was morphed to the Ali Baba data set, which is a synthetic data set containing simulated text communications about a terrorist plot in London.

Some syntactic and semantic errors occurred when the InfoMorph algorithm was applied to these data sets. The syntactic errors only occurred when substitutions resulted in incorrectly formed Twitter hash tags. The majority of the semantic errors occurred with entity substitutions for locations. Morphing geographic locations is often much more complicated than morphing people due to the more complex dependencies between geographic locations and their surrounding features. Other entity types, such as organizations, are likely to have more complicated interdependencies with other features in the text as well, making morphing more difficult.

2014 IEEE Students' Technology Symposium's proceeding consist of a novel paper titled "Framework for Surveillance of Instant Messages in Instant messengers and Social networking sites using Data Mining and Ontology" [9], explain that Innumerable terror and suspicious messages are sent through Instant Messengers (IM) and Social Networking Sites (SNS) which are untraced, leading to

hindrance for network communications and cyber security. We propose a Framework that discover and predict such messages that are sent using IM or SNS like Facebook, Twitter, LinkedIn, and others. Further, these instant messages are put under surveillance that identifies the type of suspected cyber threat activity by culprit along with their personnel details. Framework is developed using Ontology based Information Extraction technique (OBIE), Association rule mining (ARM) a data mining technique with set of pre-defined Knowledge-based rules (logical), for decision making process that are learned from domain experts and past learning experiences of suspicious dataset like GTD (Global Terrorist Database). The experimental results obtained will aid to take prompt decision for eradicating cyber-crimes.

"Generalizing Terrorist Social Networks with K-Nearest Neighbor and Edge Betweeness for Social Network Integration and Privacy Preservation"[12], another research carried to solve the complexity of SNA issues, it presented that, Social network analysis has been shown to be effective in supporting intelligence and law enforcement force to identify suspects, terrorist or criminal subgroups, and their communication patterns. However, social network data owned by individual law enforcement units contain private information that must be preserved before sharing with other law enforcement units. Such privacy issue tremendously reduces the utility of the social network data since the integration of social networks from different law enforcement units cannot be fully integrated. Without integration of social network data, the effectiveness of terrorist or criminal social network analysis is diminished. Here introduce the KNN and EBB algorithm for constructing generalized subgraphs and a mechanism to integrate the generalized information to conduct the closeness centrality measures. The result shows that the proposed technique improves the accuracy of closeness centrality measures substantially while protecting the sensitive data.

Increasing number of social network data has been made publicly available and analyze. In intelligence and security informatics domain, social network analysis is very useful in investigating the terrorist and criminal communication patterns and the structure of their organizations. However, most law enforcement force and intelligence agents only have a small piece of information before integration with the information from other agents. Further, information sharing is not always possible due to privacy issue. In this paper, they propose to construct generalized graphs before sharing the social network with other parties. The generalized graph will then be integrated with the social network owned by the agent himself to conduct social network analysis such as closeness centrality. By sharing and integrating generalized information, an agent not only preserves the privacy of
data but also preserves the utility of social network data. Experiment shows that these techniques improve the closeness centrality measurements substantially. Further, the experimental results show that different subgraph generalization methods can make significant impact on the effectiveness of information integration. In this work, using edge betweenness based method to generalize social network yields consistently lower error rate in closeness computation than using k-nearest neighbor method.

In the paper "Measuring Link Importance in Terrorist Networks" authors presented the importance of the link weights according to them terrorist network is a special kind of social network with emphasis on both secrecy and efficiency. Such networks are intentionally structured to ensure efficient communication between members without being detected. A terrorist network can be modeled as a generalized network (graph) consisting of nodes and links. Techniques from social network analysis and graph theory can be used to identify key entities in the network, which is helpful for network destabilization purposes. Research on terrorist network analysis has mainly focuses on analysis of nodes, which is in contrast to the fact that the links between the nodes provide at least as much relevant information about the network as the nodes themselves. This method analyzes the importance of links in terrorist networks inspired by research on transportation networks. The link importance measure is implemented in Crime Fighter Assistant and evaluated on known terrorist networks.

Authors have proposed link importance as a new measure for destabilizing terrorist networks. The usefulness of the method was demonstrated based on analysis of the 9/11 and the 2002 Bali bombing networks. The presented work provides the following contributions:

a) Description and evaluation of a novel method for measuring link importance in terrorist networks, which is inspired by research on transportation networks. It uses the measures of secrecy and efficiency proposed by Lindelauf et al. together with the measure of link betweenness.
b) An implementation of the proposed link importance measure in CrimeFighter Assistant, which to their knowledge also provides the first implementation of the secrecy and efficiency (information performance) measures as proposed by Lindelauf et al.

3. Discussion:

This paper is a part of continuous research work carried in the field of counter terrorism over the web. Our objective of this research is to understand the working and pattern of various terrorist groups over the web especially in SNS (social network sites).

While doing the survey we found that almost every terrorist groups have strong presence over the SNS, and they are using these services as their effective communication tools. Various researcher has worked in this field using following techniques and methods such as:

a) Data Mining
b) Graph Theory
c) Link importance
d) Neural Networks
e) Artificial Intelligence and others

But still we are unable to identify the illegal groups and their activities.

4. Proposed Method:

As we are at the initial level of our research work, here we are sharing the framework design by us to counter web terrorism. Major building blocks of proposed framework are as follow:

a) Setup Control: This component will allow our program to get entry in any SNS
b) Identification Block: IB will do the identification of suspicious groups/ networks over the SNS, for this we are using algorithms based on following concepts:
    a. Graph Theory
    b. Soft computing
c) De-stabilization Block: De- stabilization of terrorism network will be carried out using the federal services established by UN in their resolutions.

5. Conclusion: Terrorism is a big problem not for few countries but for the whole world. Now the terrorist groups are also using all the technological advancements for their effective operations. The fact is that being human, terrorist are blessed with natural intelligence so they are capable of using natural soft computing techniques. So we also have to counter their activities using soft computing techniques.

6. References:

[1] Alexander Semenov, Alexander Nikolaev, Jari Veijalainen, "Online Activity Traces Around a "Boston Bomber"", IEEE 2013, pp. 1050-1053
[2] Muhammad Akram Shaikh, Wang Jiaxin, "Investigative Data Mining: Identifying Key Nodes in Terrorist Networks", IEEE 2006, pp. 201-206
[3] Zhaobo Luo, Xiaohong Jiang, Yi Feng and Chaolun Xia, "Dynamic Core Detection in Social Network Analysis", IEEE 2009, pp 343-347
[4] Ala Berzinji, Frzand Sherko Abdullah, Ali Hayder kakei, "Analysis of Terrorist Groups on Facebook", IEEE 2013, pp. 221-221
[5] Fredrik Johansson, Lisa Kaati, Amendra Shrestha, "Detecting Multiple Aliases in Social Media", IEEE 2013, pp.1004-1011
[6] M. Banko, M. Cafarella, S. Soderland, M. Broadhead, and O. Etzioni, "Open information extraction from the web," Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI), pp. 2670–2676, 2007.
[7] http://apollo2.cs.illinois.edu
[8] H. Huang, Z. Wen, D. Yu, H. Ji, Y. Sun, J. Han, and H. Li, "Resolving entity morphs in censored data," Proceedings of the 51st Annual Meeting of the Association of Computational Linguistics, 2013.
[9] H. Ji, H. Deng, and J. Han, "Uncertainty reduction for knowledge discovery and information extraction on WWW," IEEE Special Issue on Web-Scale Multimedia Processing and Applications, vol. 100, no. 9, pp. 2658–2674.
[10] Q. Li, H. Ji, and L. Huang, "Joint event extraction via structured prediction with global features," Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, 2013.
[11] Q. Li, H. Li, H. Ji, W. Wang, J. Zheng, and F. Huang, "Joint bilingual name tagging for parallel corpora," Proceedings of the 21st ACM International Conference on Information and Knowledge Management, 2012.
[12] K.-W. Chang, R. Samdani, A. Rozovskaya, M. Sammons, and D. Roth, "Illinois-coref: the UI system in the CoNLL-2012 shared task," Proceedings of the CoNLL Shared Task, 2012.
[13] J. D'hondt, P.-A. Verhaegen, J. Vertommen, D. Cattryse, and J. Duflou, "Topic identification based on document coherence and spectral analysis," Information Sciences, vol. 181, pp. 3783–3797, 2011.
[14] L. Duboc, D. Rosenblum, and T. Wicks, "A framework for characterization and analysis of software system scalability," Proceedings of the 6th European Software Engineering Conference and ACM Symposium on Foundations of Software Engineering, pp. 375– 384, 2007.
[15] K. Sambhoos, R. Nagi, M. Sudit, and A. Stotz, "Enhancements to high level data fusion using graph matching and state space search," Information Fusion, in press, corrected proof, 2009.
[16] http://lucene.apache.org
[17] M. Mittrick, H. Roy, S. Kase, and E. Bowman, "Refinement of the Ali Baba data set," ARL-TN-0476, 2012.
[18] Witold Drozdzynski, Hans-Ulrich Krieger, Jakub Piskorskiand, and Ulrich Schafer, "SProUT–a general-purpose NLP framework integrating finite-state and unification-based grammar formalisms," published by springer in Finite-State Methods and Natural Language Processing (LNCS) Volume 4002, pp. 302-303, 2006,.
[19] (2012). [Online]. Available: http://www.webconfs.com/stop-words.php
[20] M.W.Du, and S.C.Chang, "An Approach to Designing Very Fast Approximate String Matching algorithms," IEEE journal, 1994.
[21] Y. Zhai and B. Liu, "Web data extraction based on partial tree alignment," in Prococeeding of ACM, 2005.
[22] Jer Lang Hong, "Data Extraction for Deep Web Using WordNet," published by IEEE Transactions on systems, man and cybernetics, 2011.
[23] Sunitha Ramanujam, and et al., "A Relational Wrapper for RDF Reification," E. Ferrari et al. (Eds.): TM 2009, IFIP AICT 300, pp. 196– 214, IFIP International Federation for Information Processing 2009.
[24] C.D. Manning, P. Raghavan, and H. Schutze, Introduction to Information Retrieval, Cambridge Univ. Press, 2008.
[25] (2013). [Online].http://www.start.umd.edu/gtd/dowloads/codebook.pdf.
[26] (2012). [Online]. Available: http://dir.yahoo.com/Society_and_Culture/Crime/Types_ of_Crime/
[27] E. Thambiraja,G. Ramesh, and Uma Rani, "A Survey on Various Most Common Encryption Techniques," published by IJARCSSE Journal volume 2 issue 7, pp. 226-233, 2012.
[28] M. Mahmood Ali, and Lakshmi Rajamani, "Framework for Surveillance of Emails to Detect Multilingual Spam and Suspicious Messages," IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions , IIT Kanpur, India, pp. 42-56, 2013.
[29] Chaditsa Poulatova, "The Media: A Terrorist Tool or a Silent Ally," published by IEEE in 2011.