Evaluation of Extremist Cohesion in a Darknet Forum Using ERGM and LDA

Mohammed Rashed
Whiting School of Engineering
Johns Hopkins University
Laurel, MD, USA
mrashed1@jhu.edu

John Piorkowski
Whiting School of Engineering
Johns Hopkins University
Laurel, MD, USA
jpiorko2@jhu.edu

Ian McCulloh
Accenture
Washington, DC, USA
ian.mcculloh@accenturefederal.com

Abstract— **ISIS and similar extremist communities are increasingly using forums in the darknet to connect with each other and spread news and propaganda. In this paper, we attempt to understand their network in an online forum by using descriptive statistics, an exponential random graph model (ERGM) and Topic Modeling. Our analysis shows how the cohesion between active members forms and grows over time and under certain thread topics. We find that the top attendants of the forum have high centrality measures and other attributes of influencers.**

Keywords— network analysis, topic modeling, darknet forums, exponential random graph modeling.

I. Introduction

Social Network Analysis (SNA) has shown ways to detect online extremist communities from popular platforms like twitter [1]. However, realizing the overtness of social media, terrorist groups have been shying away from popular social media sites while seeking out platforms in the darknet. The Surface Web may be considered too risky for anonymity-seeking terrorists: they could be monitored, traced and found [2]. The Terrorism Research & Analysis Consortium (TRAC) began to witness a massive migration from other social media sites, most notably Twitter [3]. Following the attacks in Paris in November 2015, ISIS has turned to the darknet to spread news and propaganda in an apparent attempt to protect the identities of the group's supporters and safeguard its content from hacktivists [2]. Even though forums are comparatively low tech, lacking features like retweet, likes, and mentions, extremists can still use them to connect with each other. For radical, extremist, and other ideologically "sensitive" groups and organizations in particular, Internet forums are a very efficient and widely used tool to connect members, inform others about the group's agenda, and attract new members [3]. Studying the underlying social network structure within the forums in the darknet is the focus of this study. We pose three research questions:

— RQ1: Do the active forum members form cohesion in popular threads?

— RQ2: How does the cohesion among active members differ between discussion topics?

— RQ3: Can we isolate the influential members?

These questions are explored using SNA and Topic Modeling on two sets of forum data from the darknet obtained via a third-party vendor.

II. BACKGROUND

A. Related Work

Researchers Stewart and Abidi [4] published an elaborate study to understand knowledge sharing behavior in an online discussion forum. They used statistical analyses and visualization along with SNA to delve deeper into the social network, identifying the most active members of the community, isolating the potential core members of the social network, and exploring the inter-group relationships that exist across institutions and professions [4]. They represented the forum as a 2-mode networks, in which there are 2 classes of nodes, and the edges go from one class to another - the discussion forum members, and the threads in which they communicate. An edge indicates that a specific member has communicated on a specific thread. With the 2-mode network, they were able to measure centrality to identify the most influential members. To find a central group of community members, they used Core-periphery analysis which is a clustering algorithm that assumes that there is a core set of nodes at the center of the social network, and a periphery set of nodes that connect to that core. Our work attempts to expand the scope of this analysis by using a 1-mode network. The darknet forum data does not clearly show user to comment relationships in a thread, thus there is no differentiation between thread author and reader in the data. We are therefore unable to use a 2-mode network methodology as Stewart and Abidi [4] and restrict our analysis to single mode data. There is relatively little information on the individual users and the researcher has to

ASONAM '19, August 27–30, 2019, Vancouver, BC, Canada

ACM ISBN 978-1-4503-6868-1/19/08...$15.00
https://doi.org/10.1145/3341161.3343532

find methodological ways to effectively make use of the huge amount of data [5]. Instead, we will enhance our SNA research with Topic Modeling.

B. SNA with ERGM

Exponential-family random graph modeling (ERGM) [6] [11] is a statistical modeling approach used to characterize social networks. An observed network is one of many possible network states. The ERGM attempts to determine which parameters or structural characteristics of the network contribute to formation of the observed state. These influential factors may include homophily, transitivity, reciprocity, amongst others [13]. These parameters are used to create a model. A probability distribution is formed to represent all possible network states and their respective probabilities of occurring under the model. This distribution is represented in ERGM as:

$$P(Y = y \mid \theta) = \frac{\exp(\theta^T) s(y)}{c(\theta)} \qquad (1)$$

where Y is a network, s(y) is a vector of network parameters, θ is a vector of coefficients, and c(θ) is a normalizing constant [6].

C. Topic Modeling

A topic model is a probabilistic model that can be used to discover latent topics in a corpus of documents [7]. Most of the recent research in the field of topic modeling is based on the Latent Dirichlet Allocation (LDA) [7]. LDA assumes that a text document has some probability distribution over "topics," and each such topic is associated with a distribution over words. Topics are not observed as input, rather they are inferred. Topic models are unsupervised models; they can be thought of as automatically clustering words into topics and associating documents with those topics [10].

III. Dataset

A. Data Collection

The hacking group Anonymous declared in 2008 that it will launch a cyber-attack on ISIS [8]. Since then, the group has successfully hacked into ISIS servers and made their forum data available to the public. Different vendors obtained these datasets, tidied them and made them available for purchase. We obtained the dataset via Flash Point. Each observation in the dataset shows a generated user id, IP (TOR exit node), forum name, room title, thread title, date/time, country, latitude/longitude and other server related data. There are two forums in our dataset – MNBR and Shamik. For the scope of our study, we combined our dataset into one dataset. Each observation in the dataset contains a TOR exit node in the form of an IP address which we used as a unique identifier of a member and since the user id is not interpretable.

B. Network Creation

To develop the network in the forum we uncovered the most popular threads and the most frequent visitors using descriptive statistics. We assumed the key influencers of the forum are among the top five visitors (Table 1). We also determined that number of visits to the forum vary with time (Fig. 1). Thus, we developed a threshold by taking the

Table 1: Top five members

ip_address	visit_count
85.225.248.131	78
213.140.72.79	65
216.75.62.36	64
90.231.127.39	64
197.231.221.211	61

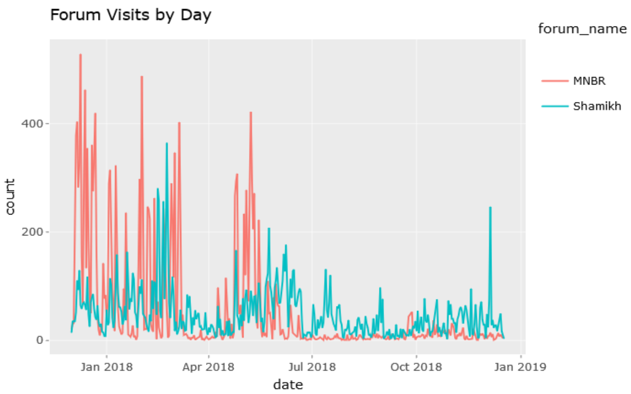

Fig 1. Forum Visit by Date

expected mean (EM) of visitors of all threads, weighted by the number of days visited. Using the threshold, we narrowed down our data to the top 14 threads and 49 members. We determined that two members appearing in k (where k>1) number of threads form a relationship, resulating in a bipartite matrix graph. We compressed this graph by taking unweighted one mode projection of one side vertices [12] and obtained a n x n matrix, where n is the member IP address. We then tested for $k=2$ and $k=4$.

IV. Member Cohesion Analysis

With regards to RQ1, ERGM was used to calculate an active member cohesion metric [6]. The ERGM projected relationship between forum members in a matrix for each month [6]. The result represents the log odds of an edge forming between two members visiting at least two different threads (for $k=2$) or at least four different threads ($k=4$). The corresponding probability of formation of this edge can be calculating through the inverse logit of the coefficients in the resulting model [6]:

$$P = \frac{\exp(coeff)}{(1+\exp(coeff)} \qquad (2)$$

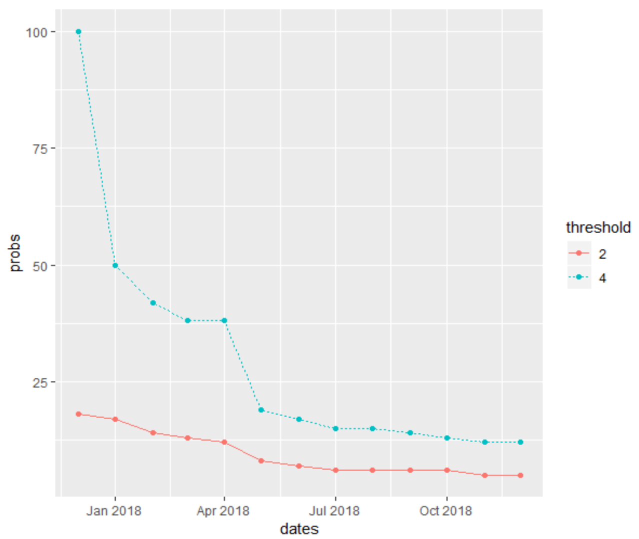

Fig. 3: Cumulative cohesiveness of members

This corresponding probability will be considered the main member cohesion metric and is analyzed in this paper [6]. Fig. 3 shows the cohesivenss of the members in the 14 threads over time (RQ1). Since threshold 4 discards nodes with fewer visits, it shows higher cohesiveness than threshold 2. The downward trends in the chart for both k values indicates that the cohension among the members falls over time. The spikes in the number of visitors (Fig. 1) coincide with this trend and explains the higher cohesiveness in the beginning of the year. Although more members are joining the network, they are visiting fewer of the popular threads, hence the declining cohesion. Fig. 4 is a depcition of the cumulative network plot showing the five most active members in red. We found high centrality measures for two of the red nodes and strong transivity among all. These two members are likely influencers in the forum (RQ3) based on our ERGM analysis.

V. TOPIC BASED COHESION ANALYSIS

We expanded our model to include understanding of the thread titles. We used LDA topic modeling to discover and infer the general topics of these thread titles by scanning the words and their distribution probabilities within documents [9]. The R package “tm” was used to construct the corpus for text mining. The “tm” package removes spaces, stop words, numbers, spaces, and punctuation, converting the words to lower case and roots to construct a term-document matrix, which allows analysis of individual words in the corpus [9]. In addition to the tm stop words, we applied custom list of regional stop words that carry insignificant contextual meaning. We used the R packages “topicmodels” and “tidytext” to calculate the term frequency, construct the inverse document matrix, remove the uncommon terms, find the most common words for individual topics and group the documents (thread rooms) by generated topics [9]. Fig. 5 shows the 6 topics identified. The colors are trivial and used to differentiate each model.

We grouped forum rooms by the their best matching topic. We then performed ERGM analysis of the forum data for every topic. We removed the visitor threshold (*from IV*)

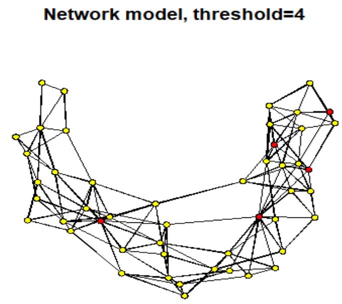

Fig 4: ERGM network model in most popular threads

Fig 5: Topics of forum thread titles

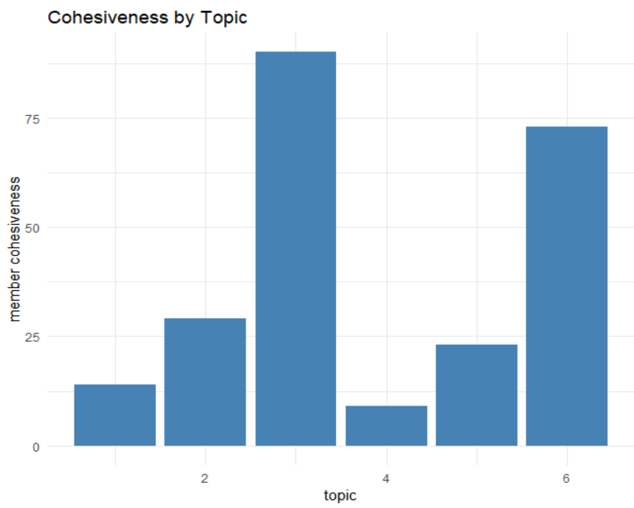

Fig 6: Cohesiveness by topic

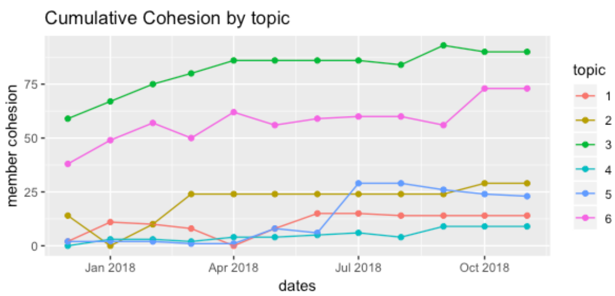

Fig 7: Cumulative cohesion by topic

Fig 8: Network Model for Topic 3

for each thread. Because the number of threads changes with each topic, we applied a varying *k (instead of a constant k in IV)* – members have to visit at least 5% of the threads in the topic group to be counted in the network. The 5% threshold drops less connected members and optimizes the links. Figures 6 and 7 show the comparision of each topic.

Topic 3 draws the most thread sharing and hence higher cohesion. These threads are likely sharing media information and are tagged 'urgent'. The cumulative chart in Fig. 7 shows how cohesion grows over time under similar topics. Topic modeling thus enhances our understanding of the network – members are more cohesive under certain media related topic (RQ 2). The network plot in fig. 8, designed to coincide with topic 3 (the most cohesive), also shows triadic relationship between the influential members (1st, 2nd and 4th ranking influential members from Table 1), indicating another attribute of their influence in the network [6].

VI. Discussion and Conclusion

In this paper, we use ERGM to get a better understanding of the cohesion [6] among extremist members in a darknet forum, and identify longitudinal trends with regards to overall cohesion, as well as topic-based cohesion. We were able to use results of the ERGM modeling to calculate the overall probability of two members visiting most popular threads a certain number of times. The probability that two members of the forum visited at least k number of threads (where k>1) became the metric by which we measured cohesion. We tested for k=2 and k=4 and found in both situation there was higher cohesion during spikes in visits to the most popular threats (top 14), while downward trend rest of the time. Google searches show that the spikes in visits coincided with extremist events that took place around the world. When analyzing topic-specific member cohesion, we found that, media related topics that were tagged urgent drew higher cohesion which grew with time. This indicates members tend to visit more threads that share media related information. In both analyses, we noticed that the top influencers were among the 5 most frequently visiting members in the most popular threads. This indicates that they may be key players in propagating information throughout the forum.

While this finding is not completely conclusive for identification of top influencers, it gives us a start point to the analysis of a complex and noisy network. Understanding the key influencers in the network is just one piece of the puzzle. It is a complex process to enrich the nodes with additional information as lot of the member identity is masked in the darknet.

As threats to democracy are increasingly originating online in the form of extremist groups and propaganda, it is increasingly important to explore methods to measure and understand the impact of key influencers in online space. This paper extended the application of ERGMs for measuring cohesion and demonstrated the ability to locate key influencers. Furthermore, we introduced a method to apply topic modeling and then measure within topic cohesion. In a scenario where an orchestrated campaign may target extremist content, the methods laid out in this paper provide a template for measuring the effectiveness of such a campaign. It is our hope that more papers along these lines will be published in the near future to grow the body of work for understanding the tactics, techniques and procedures of online extremist movements, as well as the methods to measure the success of intervention campaigns.

References

[1] M. Benigni, K. Joseph, and K. Carley. "Threat Group Detection in Social Media: Uncovering the ISIS Supporting Network on Twitter." Carnegie Mellon University Doctoral Thesis. May 2017.

[2] G. Weimann. "Going Darker? The Challenge of Dark Net Terrorism." Washington, DC: Woodrow Wilson International Center for Scholars. 2016.

[3] Terrorism Research and Analysis Consortium. "Massive Migration to Telegram, the new Jihadist Destination", TRAC Insight, November 4, 2015

[4] S. Stewart, and S. Abidi. "Applying social network analysis to understand the knowledge sharing behaviour of practitioners in a clinical online discussion forum." Journal of medical Internet research 14 (6), 2012.

[5] P. Holtz, N. Kronberger, and W. Wagner. "Analyzing internet forums." Journal of Media Psychology. 2012.

[6] S. Sadayappan, I. McCulloh, and J. Piorkowski. "Evaluation of Political Party Cohesion Using Exponential Random Graph Modeling." In Proc. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 298-301. IEEE, 2018.

[7] A. Ahuja, W. Wei, and K. Carley. "Topic Modeling in Large Scale Social Network Data." Carnegie Mellon University Technical Report CMU-ISR-15-108. 2015.

[8] A. Bal, K. Weidner, R. Hanna, and A. Mills. "Crowdsourcing and brand control." Business Horizons 60, no. 2 (2017) pp. 219-228.

[9] X. Kuang, H. Chae, B. Hughes, and G. Natriello. "An LDA Topic Model and Social Network Analysis of a School Blogging Platform." In Proc. 10th International Conference on Educational Data Mining, pp. 362-363.

[10] M. Paul and M. Dredze. "Discovering health topics in social media using topic models." PloS one 9 (8) 2014.

[11] D. Lusher, J. Koskinen, and G. Robins. "Exponential random graph models for social networks: Theory, methods, and applications." Cambridge University Press, 2013.

[12] S. Banerjee, M. Jenamani, and D. Pratihar. "Properties of a projected network of a bipartite network." In 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 0143-0147. IEEE, 2017.

[13] I. McCulloh, H. Armstrong, and A. Johnson. "Social network analysis with applications." John Wiley & Sons, 2013.