From Tweets to Intelligence: Understanding the Islamic Jihad Supporting Community on Twitter

Matthew Benigni(✉) and Kathleen M. Carley

Center for Computational Analysis of Social and Organizational Systems (CASOS),
Institute for Software Research, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburg, PA, USA
{mbenigni,kathleen.carley}@cs.cmu.edu
http://www.casos.cs.cmu.edu/

Abstract. ISIS' ability to build and maintain a large online community that disseminates propaganda and garners support continues to give their message global reach. Although these communities contain trained media cadre, recent literature suggests that large numbers of "unaffiliated sympathizers" who simply retweet or repost propaganda explain ISIS' unprecedented online success [1,2]. Tailored methodologies to detect and study these online threat-group-supporting communities (OTGSC) could help provide the understanding needed to craft effective counter-narratives however continued development of these methods will require collaboration between data scientists and regional experts. We illustrate the potential of this partnership using two ongoing projects at the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University. First we present the CASOS Jihadist Twitter Community (CJTC), an online community of over 15,000 Twitter users that support one or more of the Islamic extremist groups engaged in the ongoing conflicts in Northern Iraq and Syria. We briefly discuss the methods used to detect and monitor these communities and highlight forms of information that can be extracted from them. We then present an active social botnet that attempts to elevate the social influence of users supportive to Jabhat al-Nusra's agenda. In each case we highlight the ability of these methods to incorporate regional expertise for better performance and recommend future research.

Keywords: Threat network detection · Community detection · Social media intelligence · Online social networks · Social bots · ISIS · Radicalization

1 Introduction

Extremist groups' powerful use of online social networks (OSNs) to disseminate propaganda and garner support has motivated intervention strategies from industry as well as governments however early efforts to provide effective

K.S. Xu et al. (Eds.): SBP-BRiMS 2016, LNCS 9708, pp. 346–355, 2016.
DOI: 10.1007/978-3-319-39931-7_33

counter-narratives have not produced the results desired. Mr. Michael Lumpkin, the director of the United States Department of State's Center for Global Engagement, is charged with leading efforts to "coordinate, integrate, and synchronize government-wide communications activities directed at foreign audiences in order to counter the messaging and diminish the influence of international terrorist organizations" [3]. In a recent interview, Mr. Lumpkin expressed the need for a new approach:

> So we need to, candidly, stop tweeting at terrorists. I think we need to focus on exposing the true nature of what Daesh is.
>
> Mr. Michael Lumpkin, NPR Interview March 3, 2016

A logical follow-up question to Mr. Lumkin's statement would be "Expose to whom?" Recent literature suggests that "unaffiliated sympathizers" who simply retweet or repost propaganda represent a paradigmatic shift that partly explains the unprecedented success of ISIS [1,2] and could be the audience organizations like the Global Engagement Center need to focus on. Gaining understanding of this large population of unaffiliated sympathizers and the narratives most effective in influencing them motivates methods to detect and extract information from large online threat-group-supporting communities (OTGSC). However, detecting, monitoring, and data-mining targeted OTGSs requires novel methods, and development must include both data science and regional expertise. We define data science as a set of fundamental principles that support and guide the principled extraction of information and knowledge from data, and in this paper we present the CASOS Jihadist Twitter Community (CJTC), a online community of over 15,000 Twitter users who support one or more of the radical groups engaged in the ongoing conflicts in Northern Iraq and Syria. We describe how large OTGSCs can offer unique insights into the unaffiliated supporters who appear critical to ISIS' success. We then provide an example of one method used to excite and grow these OGTSCs in the form of an active social botnet. The botnet attempts to elevate the social influence of users supportive to Jabhat al-Nusra's agenda, while encouraging following ties amongst botnet followers. Our goal is to present two novel examples of social computing applied to counterterrorism, and motivate the continued interdisciplinary collaboration required to gain understanding of large online communities and effectively counter extremist propaganda.

2 The CASOS Jihadist Twitter Community (CJTC)

On November 13, 2015 much of the world watched as terrorist launched a series of coordinated attacks in Paris killing 130 people. In near real-time social media erupted with support for the victims of these attacks, but some online communities viewed the attacks as cause for celebration. In fact, passive supporting but unaffiliated social media users have become an essential element of groups like ISIS and Jabhat al-Nusra's recruiting strategy, possibly aid the motivation and resourcing for attacks like those seen in Paris [1]. Large online social networks like Twitter offer a means to generate large online communities, and many of the members appear to be "unaffiliated supporters." In fact, Twitter has suspended over 125,000 ISIS-supporting accounts from August to December of 2015 [4]. As ISIS recruiters identify community members who show increasing levels of radicalization, small teams of social media cadre have been observed lavishing attention on these recruitment targets and subsequently move the conversation to more secure online platforms [2]. Less secure but large open platforms like Twitter enable extremist groups propaganda to gain broad reach. Denying this *key terrain* requires novel methods designed specifically to identify and analyze threat-group-supporting communities embedded in OSNs. Information like key users, powerful narratives, and advanced dissemination methods can all be extracted from OTGSCs to inform messaging and intervention strategies. Benigni et al. present Iterative Vertex Clustering and Classification [5], a novel method to detect large, ideologically organized online communities, using both agent level attributes and network structure. We briefly present the methodology, introduce the CJTC, provide illustrative analysis of the network, and share ongoing research goals in this section.

2.1 Background: From Community Detection to Threat Network Detection

The application of network science to counter-terrorism has a long history [6,7]; however, the rise of social media and online social networks (OSNs) has motivated methods to apply network science theory to networks at much larger scale. Community detection attempts to identify groups of vertices more densely connected to one another than to other vertices in a network, but networks extracted from OSNs present unique challenges due to their size and high clustering coefficients. Furthermore, an individual's social network also often reflects his or her membership to many different social groups. Thus in many instances algorithms that use only network structure do not provide the precision needed to identify OTGSCs [5]. A sub-class of community detection methods has emerged that attempts to leverage node attributes and network structure called community detection in annotated networks. These methods have been shown to perform well with OSNs because of their ability to account for a great variety of vertex features like user account attributes while still capitalizing on the information provided by the structure of the graph; they also perform well at scale [8,9]. However, we find that effective OTGSC detection requires information from users' following, mention, and hashtag behaviors as well. Benigni et al. present IVCC, an community detection method designed to extract OTGSCs by modeling users within a heterogeneous graph structure with annotated nodes [5] (Fig. 1).

2.2 Iterative Vertex Clustering and Classification

Iterative Vertex Clustering and Classification (IVCC) is conducted two phases, and often iteratively. In Phase I, unsupervised clustering methods like Newman and Louvain grouping are used to both identify positive cases labels and

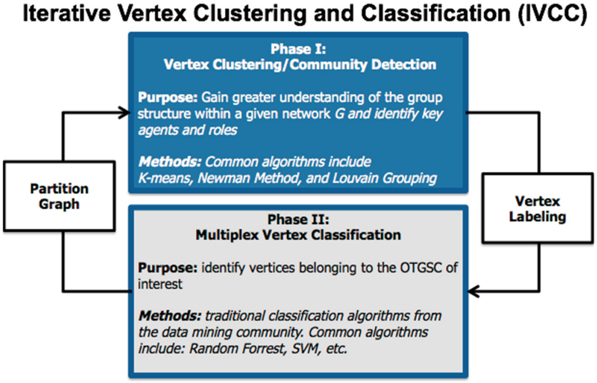
Iterative Vertex Clustering and Classification (IVCC)
Phase I:
Vertex Clustering/Community Detection
Purpose: Gain greater understanding of the group structure within a given network G and identify key agents and roles
Methods: Common algorithms include K-means, Newman Method, and Louvain Grouping
Phase II:
Multiplex Vertex Classification
Purpose: identify vertices belonging to the OTGSC of interest
Methods: traditional classification algorithms from the data mining community. Common algorithms include: Random Forrest, SVM, etc.
Partition Graph
Vertex Labeling

Fig. 1. IVCC is an online threat-group-Supporting community (OTGSC) detection methodology conducted in two phases. In Phase I either community optimization or vertex clustering algorithms are used to identify positive and negative case examples to facilitate supervised detection in Phase II.

remove noise. This pre-clustering facilitates supervised classification of OTGSC members in Phase II. At the core of the methodology is the use of both user level features and rich multiplex network structures offered by OSNs. First the authors construct $U_{u \times a}$ consisting of a numeric user attributes where u is the total number of users or nodes in the network. Examples of such attributes are follower count, number of posts, or creation date. Node attributes could also be developed from other sources of intelligence. Spectral methods are used to dimensionally reduce network data like following, mention, or hashtag behaviors. By constructing symmetric graphs of users' following F and mention M relationships, and a weighted bipartite graph H of hash tags in a user's timeline, lead eigenvectors can then be extracted from each graph and concatenated with U to form a feature space for classification. Although IVCC is presented using Twitter data [5], similar methods could be used more generally with large heterogeneous networks.

Benigni et al. collected a two-hop snowball sample of five popular ISIS propagandists presented in [10], resulting in approximately 120,000 Twitter users. With two iterations of IVCC, they removed accounts with high following counts (i.e. politicians, news media members, celebrities, etc.), and extracted a network of nearly 23,000 *ISIS supporters*. The results of this initial work form the seed accounts for the CJTC.

2.3 Threat Network Analysis: The CJTC

CASOS is currently extending IVCC to dynamically monitor threat-group-supporting online communities. By using historical results and active learning, we update the CJTC based on the recent community activity. Currently the community contains just over 15,000 supporters, where we define a supporter as a Twitter user who positively affirms the leadership, ideology, fighters, or call to Jihad of any of the known Jihadist groups engaged in ongoing operations in Northern Iraq and Syria. The majority of tweeters voice support for ISIS or Jabhat al-Nusra though other groups are present. The size of this community offers insights not easily gleaned from randomly sampled Twitter data or manually developed datasets as will be highlighted in the remainder of this section.

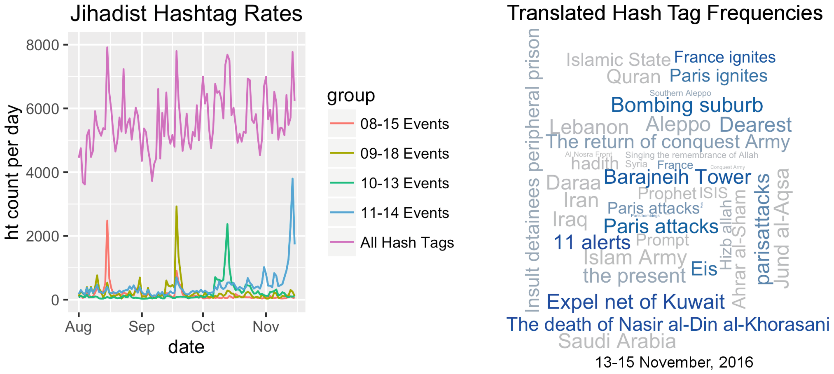

Fig. 2. The left panel depicts the volume of hashtags used within the CJTC from AUG-NOV 2015. The right panel highlights the hashtags most explanatory of the increased activiy on November 14, 2015.

Though many demographical analyses could be useful, for conciseness we will use temporal network activity patterns to illustrate information extraction from OTGSCs. The Twitter REST API limits collection to a tweeter's last 3,200 posts which forces us to normalize daily volume. Some tweeters have more than 3,200 posts in the past 6 months, and quite a few of our tweeters have not posted in over 90 days. Identification of dormant users could provide insight into the radicalization process, but will not be analyzed or discussed in this work. We estimate the CJTC's daily volume by normalizing based on the number of tweeters in our dataset who have a collected tweet before and after any given day which often highlights current events that stimulate this community. A simple news search of events on days of increased activity often reveals operational events in Syria, Northern Iraq, or large scale terror attacks. Similar analysis of hash tag trends often provides richer insight. Figure 2 highlights temporal analysis of CJTC hash tag use. The left panel of depicts hash tag frequencies over time, while the right panel depicts trending hashtags on 13–14 November, 2015. Size in the word cloud connotes frequency, and color denotes how anomalous a particular tag's frequency was when compared to a 6 months average. The community's reaction to the 13 November, 2015 Paris attacks is illustrated with both increased volume and trending hashtags. Increased hash tag volume depicted in the left panel of Fig 2, coupled with the corresponding hash tag trends in the right panel give startling insight into the unique nature of this online community. Ongoing operations in Syria provide another example. The hash tag [illegible], translated Zabadani, increases tenfold in terms of daily frequency on 15 August and 18 September, 2015. Both dates refer the breakdown of ceasefire agreements in the region [11]. With proper subject matter and language expertise, similar analysis can identify changes in popularity of leaders, organizations, or narratives over time.

2.4 Moving Forward

As a supervised learning methodology, IVCC lends itself to leveraging regional expertise by learning patterns based on examples. Active-learning refers to supervised algorithms that iteratively select examples to be labelled by experts, and have been found substantially increase performance with far fewer labelled instances. Such methods could enable regional expertise to be incorporated into the classifier at minimal cost. Furthermore, a user-oriented, server-based interface could enable the regional expert to contribute to the set of annotated instances while conducting his or her own exploratory data analysis. As the set of annotated examples or “training set” grows new, more nuanced classifiers could be trained. Due to the size and diversity of these online communities, exploration and interpretation of results is likely a research area unto itself. One could identify the news sources or propagandists most influential within these communities, and develop more-informed counter-narratives and strategic communications strategies. The challenge in developing tools and methods to facilitate OTGSC analysis lies in the novelty of the analytical task. Regional experts cannot yet articulate exactly what they want methods to provide, and researchers are challenged to understand what information extractions are most useful to senior leader information requirements. Establishing online tools that provide illustrative analyses and capture feedback while end users to explore large communities would likely accelerate research efforts aimed at countering groups like ISIS.

3 The FiribiNome Social Botnet: Sophisticated Promotion of Propaganda to Excite a Community

While analyzing the CJTC, as well as a similar dataset focused on online dialogue focused on the Russian occupation of Crimea, we observe accounts that tweet with high daily volume, but each tweet or retweet simply contains a string of @mentions. In this section we analyze a network of social bots used to promote specific online activists or propagandists.

Social bots, software automated social media accounts, have become increasingly common in OSNs. Though some provide useful services, like news aggregating bots, others can be used to shape online discourse [12]. ISIS’ use of bots has been well documented [13], and their competitors are following suit. Social botnets are teams of software controlled online social network accounts designed to mimic human users and manipulate discussion by increasing the likelihood of a supported account's content going viral. The use of bots to influence political opinion has been observed in both domestically [14] and abroad [15], the use of social bots has been documented in the MENA region [12], and ISIS use of them motivated a DARPA challenge to develop detection methods [16]. In isolation, these accounts appear to be producing spam and relatively harmless, however they are examples of a sophisticated strategy to promote specific accounts while remaining undetected by Twitter.

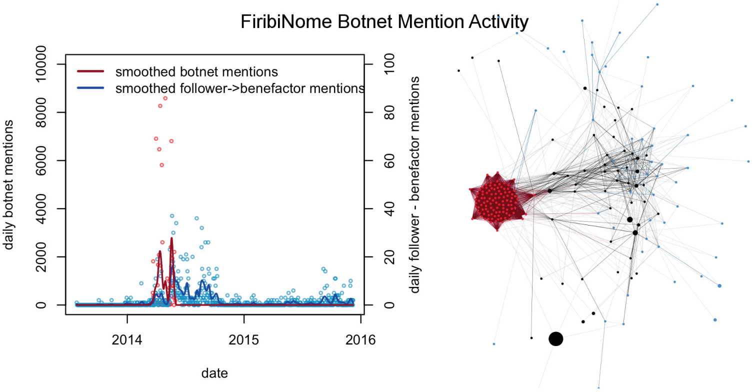

Fig. 3. Depicts mention behaviors and their effects within the FiribiNome Social Botnet. The left panel depicts two scaled time series. The red circles and smoothed trend line depict the number of daily mentions by botnet members. The blue circles and corresponding trend line depict botnet followers' mentions of benefactor accounts. The association between the two series implies the botnet was able to generate discussion about benefactor accounts among its followers. The right panel depicts the mention network of the FiribiNome social botnet. The vertices are user accounts. The plot depicts how *botnet members*, red vertices, are used to increase the social influence of *benefactors*, black vertices, by promoting them to *botnet followers*, blue vertices. Vertices are scaled by follower count (Color figure online).

3.1 CJTC Botnet Analysis

Figure 3 depicts the mention activity associated with the a Jabhat al-Nusra supporting social botnet designed to increase the social influence of a specific set of accounts and encourage following connections between Jabhat al-Nusra supporting tweeters. The botnet consists of two types of accounts. *Botnet members*, are depicted by red vertices in the right panel of Fig. 3, and consist of 74 accounts exhibiting near identical behavior. Each account follows between 116 and 134 accounts, most of which are *botnet members*. Their following counts vary from 142 to 322 accounts of which many appear to be real tweeters. They come online for 38–58 days, tweet between 71 to 170 times, then go dormant. This behavior can clearly be seen by the red trend line in Fig. 3. Their tweets consist of original posts or retweets containing strings of @mentions of other *botnet members*, but occasionally mention or retweet content from what we call *benefactor accounts* (depicted by black vertices in the right panel of Fig. 3). The *botnet* account FiribiNome20 illustrates this behavior. In isolation, these accounts appear to be producing spam and relatively harmless, however our analysis indicates the network of *botnet members* increases the social influence of *benefactor accounts*. The blue series in the left panel of Fig. 3 and corresponding blue vertices in the right panel depict the mention activity of the 843 active botnet followers as of February 2016. The left panel depicts *follower* accounts' mentions of *benefactor* accounts and the temporal relationship betwen the activity associated with each account type implying the botnet effectively promotes discussion of *benefactor* accounts. How much discussion is generated remains an open question. Due to the large number of extremist accounts suspended by Twitter, the number of *botnet followers* active in the summer of 2014 was likely much larger. This mention behavior exhibited by *botnet members* could also trigger Twitter's recommendation system to recommend following ties between *botnet followers*,or encourage *botnet followers* to follow *benefactors*.

Examples of *benefactor accounts* are depicted in Table 1; each representing a slightly different style and type of messaging commonly observed in the CJTC. Dr. Hani al-Sibai is a London-based radical Islamic Scholar cited by Ansar al-Sharia as one of five influential motivators of Tunisian terrorists [17]. *@ba8yaa* or "Daesh are the Enemy" attempts to discredit ISIS through satire and counter-propaganda and could prove informative in development of counter-narratives. There are also many accounts that present the appearance of reporting near-real-time news like *@Ghshmarjhy*, while other accounts promote third party applications like @Almokhtsar and @FiribiNome12. We have found some of these applications request permission to tweet or follow users on the tweeter's behalf. These highly followed and highly mentioned accounts each could offer insight into the sophisticated methods used to leverage social media.

3.2 Moving Forward

It is possible that botnet structures with similar mention behavior could be developed in a more sophisticated manner. Larger networks with more human-like behavior would be much more challenging to detect. The FiribiNome botnet

Table 1. Depicts four account promoted by the FiribiNome social botnet. Each account represents a slightly different style and type of messaging.

Account	Follower Count	Messaging Type
@Hanisibu	104K	Islamic Scholar
@ba8yaa	1,272	anti-ISIS satire/propaganda
@Ghshmarjhy	6,644	Syrian revolution updates
@Almokhtsar	164K	app: MENA news feed

could simply represent a proof of concept explaining its lack of activity since 2014. Although simple heuristics like average mentions per tweet enabled us to detect the botnet, more advanced detection strategies are needed to determine if more sophisticated botnets are influencing the CJTC. Methods of operationalizing this type of intelligence are worth exploring as well. It is possible that similar mention behaviors could be used to target specific online communities with counter-narratives. Again, the need for an interdisciplinary collaboration between the data scientist, regional expert, and decision maker is needed to identify opportunities for useful intelligence extraction.

4 Conclusion

We have highlighted the potential of extracting intelligence from large online threat-group-supporting communities (OTGSCs) and presented illustrative examples with a goal of motivating continued interdisciplinary collaboration. We have also presented the CJTC dataset as an example of an OTGSC to emphasize how detecting and monitoring OTGSCs can be an important tool in understanding the passive support structure essential to the distribution of extremist propaganda. Furthermore, these methods could facilitate identification of sophisticated dissemination techniques used in these communities and inform our own information operations. Our goal is to refine these methods and grow a consortium of data scientists, regional experts, and strategic decision makers by hosting, curating, and reporting on datasets like the CJTC.

Acknowledgements. This work was supported in part by the Office of Naval Research (ONR) through a MINERVA N000141310835 on State Stability. Additional support for this project was provided by the center for Computational Analysis of Social and Organizational Systems (CASOS) at CMU. The views ond conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research or the U.S. Government.

References

1. Yannick Veilleux-Lepage. Paradigmatic Shifts in Jihadism in Cyberspace: The Emerging Role of Unaffiliated Sympathizers in the Islamic State&#39;s Social Media Strategy (2015)
2. Berger, JM.: Tailored Online Interventions: The Islamic States RecruitmentStrategy. Combating Terrorism Center Sentinel
3. Dozier, K.: Anti-ISIS-Propaganda Czars Ninja War Plan: We Were Never Here, March 2016
4. Isaac, M.: Twitter Steps Up Efforts to Thwart Terrorists' Tweets. The New York Times (2016). http://www.nytimes.com/2016/02/06/technology/twitter-account-suspensions-terrorism.html
5. Benigni, M., Joseph, K., Carley, K.: Threat Group Detection in Social Media: Uncovering the ISIS Supporting Network on Twitter. Submitted to Plos One

6. Krebs, V.: Uncloaking terrorist networks. First Monday **7**(4), 215–235 (2002)
7. Carley, K.M.: A Dynamic Network Approach to the Assessment of Terrorist Groups and the Impact of Alternative Courses of Action. Technical report, October 2006
8. Tang, L., Liu, H.: Leveraging social media networks for classification. Data Min. Knowl. Discov. **23**(3), 447–478 (2011)
9. Binkiewicz, N., Vogelstein, J.T., Rohe, K.: Clustering, Covariate Assisted Spectral (2014). arXiv preprint arXiv: 1411.2158
10. Carter, J.A., Maher, S., Neumann, P.R.: #Greenbirds Measuring Importance and Influence in Syrian Foreign Fighter Networks. International Centre for the Study of Radicalization Report, April 2014
11. Syria ceasefire ends, fighting resumes. Reuters, August 2015
12. Abokhodair, N., Yoo, D., McDonald, D.W.: Dissecting a Social Botnet: Growth, Content and Influence in Twitter, pp. 839–851. ACM Press (2015)
13. Berger, J.M.: How ISIS Games Twitter. The Atlantic, June 2014
14. Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A.: The rise of social bots (2014). arXiv preprint arXiv: 1407.5225
15. Forelle, M., Howard, P., Monroy-Hernndez, A., Savage, S.: Political Bots the Manipulation of Public Opinion in Venezuela. arXiv: 1507.07109 [physics]. arxiv: 1507.07109, July 2015
16. Subrahmanian, V.S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A., Menczer, F., Waltzman, R., Stevens, A., Dekhtyar, A., Gao, S., Hogg, T., Kooti, F., Liu, Y., Varol, O., Shiralkar, P., Vydiswaran, V., Mei, Q., Huang, T.: The DARPATwitter Bot Challenge. arXiv: 1601.05140 [physics]. arxiv: 1601.05140, January 2016
17. Ansar al-Sharia Tunisias and Long Game. Dawa, hisba, and jihad (2013)