Thesis Proposal: Online Extremist Community Detection, Analysis, and Intervention

Lieutenant Colonel Matthew Curran Benigni

June 2016

Societal Computing Program
Institute for Software Research
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Dr. Kathleen M. Carley
Dr. Zico Kolter
Dr. Daniel Neil
Dr. Randy Garrett

*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy.*

Keywords: Covert Network Detection, Community Detection, Annotated Networks, Multilayer Networks, Heterogeneous Networks, Spectral Clustering, Socialbots, Botnets

Abstract

The rise of the Islamic State of Iraq and al-Sham (ISIS) has been watched by millions through the lens of social media. This “crowd” of social media users has given the group broad reach resulting in a massive online support community that is an essential element of their public affairs and resourcing strategies. Other extremist groups have begun to leverage social media as well. Online Extremist community (OEC) detection holds great potential to enable social media intelligence (SOCMINT) mining as well as informed strategies to disrupt these online communities. I present Iterative Vertex Clustering and Classification (IVCC), a scalable analytic approach for OGTSC detection in annotated heterogeneous networks, and propose several extensions to this methodology to help provide policy makers the ability to identify these communities as scale, understand their interests, and shape policy decisions.

In this thesis, I propose contributions to OEC detection, analysis, and disruption:

- efficient identification of positive case examples through semi-supervised dense community detection
- monitoring dynamic OECs through a repeatable search and detection methodology
- gaining influence within OECs through topologically derived $@mention$ strategies
- an extended literature review of methods applicable to detection, analysis, and disruption of OECs

The contributions proposed in this thesis will be applied to four large Twitter corpora containing distinct online communities of interest. My goal is to provide a substantive foundation enabling follow on work in this emergent area so critical to counter-terrorism and national security.

Contents

Chapter 1

Introduction

1.1 Overview

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 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" [26]. In a recent interview, Mr. Lumpkin expressed the need for new approaches:

> "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 [11, 61] and could be the audience organizations like the Global Engagement Center need to focus on. The size and density of Twitter's social network has provided a topology enabling extremist propaganda to gain a global audience, and has become an important element of extremist group resourcing strategies[8, 11, 39]. Gaining understanding of this large population of unaffiliated sympathizers and the narratives most effective in influencing them motivates this thesis. I call these social networks online Extremist communities (OEC) and define them as follows:

> **Online Extremist Community (OEC):** a social network of users who interact within social media in support of causes or goals posing a threat to national security or human rights.

My goal is to provide a theoretical framework and methods to detect, analyze, and disrupt these communities, but to do so effectively requires significant contributions. In fact this research area will likely require ongoing collaboration from academia, industry, and government to develop effective methods to counter OEC messaging.

The importance of understanding extremist movements' use of online social networks is essential to counter-messaging and has motivated a great deal of research [6, 9, 43]. The ability for extremist groups, or more generally "threat groups," to generate large online support communities has proven significant enough to require intervention, but few methods exist to detect and understand them. The size and dynamic nature of these Online Extremist Communities (OECs) requires tailored methods for detection, analysis and intervention. Within this thesis, I will provide an extended literature review of novel methods for OEC detection, analysis, and disruption, and provide a framework enabling researchers and practitioners to effectively focus future research efforts in this important area. I will also present methods for detecting and monitoring OECs. Finally, I will highlight methods used to gain influence in OECs.

Rigorous OEC detection and analysis methods are needed to understand these communities and develop effective intervention strategies. I propose the following research questions to address current gaps in capability:

1 **How can one effectively search for and detect Online Threat Group Supporting Communities?** In the fall of 2014 ISIS launched a social media campaign at a scale never before seen. A lack of rigorous methods designed to detect OECs lead to varied estimates of the community's size [6, 9]. In this work I will present Iterative Vertex Clustering and Classification (IVCC) a classification-based community detection strategy tailored specifically for OECs.

2 **How can one account for changes in OEC membership over time?** OECs can be viewed as a form of online activism. As such, they compete for support with governments, organizations, or other activist groups. The evolution of the ISIS OEC provides a powerful example of competing stakeholders. Online activists like *Anonymous* and companies like Twitter have begun to disrupt the ISIS OEC [1, 15, 32], and suspensions appear to have been marginally effective [10]. However, this has lead to a predator-prey like relationship where these communities have shown increasing levels of adaptability and resilience. The result is that these communities are highly dynamic, and monitoring them requires repeatable methodologies to search and detect new members.

3 **What technical methods are used to generate user influence and promote narratives within these communities?** Finally I must be able to identify key users and topics. To do so requires an understanding of how to identify and account for automated accounts, bots, as well as users who have greater influence within the network. Quantitative methods to identify key users and narratives will enable us to identify the methods used to promote them, and standard measures of centrality are biased by highly followed, but unrelated accounts. Therefore tailored metrics are needed to identify key users. Similar extensions are needed to identify the narratives that catalyze discussion with in the community.

1.1.1 Outline

This thesis aims to introduce OEC detection, analysis, and intervention research in a manner that enables effective collaboration between researchers, practitioners, and industry, as well as present methodologies addressing the three research questions listed in the previous section. The goal is to provide a toolchain and framework that moves this research community of interest towards being able to develop effective, informed interventions in OECs. In Section 1.2 I will provide detailed overviews of each dataset used in subsequent chapters, I then present a detailed overview of related work associated with social media intelligence (SOCMINT) as well as the strengths and limitations of current methods available for OEC detection, monitoring, and mining in Chapter 2.

In Chapter 2 I will argue the need for a framework enabling researchers, practitioners, and industry to develop research needs as well as an extended literature review of relevant work. In Chapter 3 I will present Iterative Vertex Clustering and Classification (IVCC), a supervised method developed to accomplish the following methodological task:

> **Methodological Task 1 (MT1):** *Given a large meta-network with annotated nodes that has an embedded community of interest and a set of labelled training data, perform a bipartite partition of the network to identify a large proportion of the community of interest.*

Although IVCC provides strong results, it has limitations that must be addressed. It is impractical to assume practitioners will consistently have access to large amounts of training data, and I have found the feature space utilized in [6] to perform poorly at unsupervised and semi-supervised tasks. In Chapter 5, I will propose an unsupervised ensemble method to address these shortcomings with the following methodological task:

> **Methodological Task 2 (MT2):** *Given a large annotated heterogeneous network and limited training data, extract identifiable clusters of embedded OEC members.*

The ability to accomplish **MT1** and **MT2**, establishes a foundation that would facilitate a toolchain of methodologies allowing intelligence practitioners to monitor dynamic OECs over time, and Chapter 5 will address the following methodological task:

> **Methodological Task 3 (MT3):** *Given a large dynamic OGTSC, maintain understanding of group activity and interests.*

MT3 will require a methodology that is robust to major changes to group membership similar to Twitter's counter-ISIS suspension campaign. Monitoring such groups will require:

- A repeatable search and detection framework
- An active-learning framework to ensure robustness against shocks to group membership and structure
- An understanding of the uncertainty associated with classification methods.

Each of which will be addressed in Chapter 5.

Chapters 6 will focus on how OEC network topology is used to gain influence by members and could inform intervention by security practitioners. In Chapter 6 I will propose research related to botnet structures and community resilience tactics identified in the **ISIS14,CCMC,** and **CJTC** datasets described in Section 1.2. Sophisticated use of $@mentions$ are used in each dataset in a manner that appears to increase following ties and promote specific accounts; however, little research exists with respect to this behavior and its effect on social influence. I theorize that by mentioning accounts that are highly central within an OEC, one could gain social influence within it. Such research would provide an important step towards employing successful intervention strategies within OECs.

1.2 Data

I will introduce each of the datasets to be used within this dissertation and briefly describe them. However, I will refer to them in greater detail in subsequent chapters as to how they will be used specifically to address my research questions and evaluate the methodological tasks outlined in this proposal.

To develop each of my datasets, I instantiate an n-hop snowball sampling strategy [33] with known members of my desired network. Snowball sampling is a non-random sampling technique where a set of individuals is chosen as “seed agents.” The k most frequent friends of each seed agent are taken as members of the sample. This technique can be iterated in steps, as I have done in my search. Although this technique is not random and prone to bias, it is often used when trying to sample hidden populations [9].

The snowball method of sampling presents unique and important challenges within OSNs. Users’ social ties often represent their membership in many communities simultaneously [56]. At each step of my sample, this results in a large number of accounts that have little or no affiliation with a OEC of interest. The core problem of then involves extracting a relatively small OEC embedded in a much larger graph. In order to do so, I require rigid definitions of account types which will be used for the remainder of this proposal. I define three types of user:

member: A Twitter user who’s timeline shows unambiguous support to the OEC of interest. For example, if the user positively affirmed the OEC’s leadership or ideology, glorifies its fighters, or affirms its talking points. It is important to mention that a member’s*support* is relative and in many cases not in violation of local law or Twitter’s terms of use. However, the volume of these “passive members” appears to be an essential element of OECs ability to reach populations prone to radicalization [61]

non-member: A user whose tweets are either clearly against or show no interest in the OEC of interest.

official user: I label vertices as *official users* if they meet any of the following criteria: the user’s account identifies itself as a news correspondent for a validated news source; the account is attributed to a politician, government, or medium sized company or larger, or accounts with greater than k followers. This third is necessary to account for OEC members’ dense ties to news media, politicians, celebrities, and other official accounts. Such accounts are interesting in that there higher follower counts and mention rates tend to make them appear highly central even though they do not exhibit any ISIS supporting behaviors. *Official users* must be identified and removed for accurate classification of ISIS-supporting, thus illustrating the utility of an iterative methodology. This will be discussed in detail in Chapter 3

I will now describe in detail the datasets I will use to evaluate each of the aforementioned methodological tasks in subsequent chapters.

1.2.1 ISIS OEC on Twitter (ISIS14) (Search Date: November 2014)

I developed this dataset in November of 2015 by seeding a two hop snowball sample of influential ISIS propagandists'[19] following ties. Step one of my search collected user account data for my 5 seed agents' 1345 unique following ties. Step 2 resulted in account information for all users followed by the 1345 accounts captured in step 1. My search resulted in 119,156 user account profiles and roughly 862 million tweets. This network is multimodal, meaning that it has two types of vertices, and multiplex, because it has multiple edge types. I represent this set of networks, or metanetwork[18], G with two node classes: users and hashtags, and four types of links: following relationships, mention relationships, hashtags used in the same tweet, and user-hashtag links.

1.2.2 The Plane Spotting Twitter Community (PSTC) (Search Date: March 2016)

In the second half of the twentieth century, the aviation community began to recognize a large community of aviation enthusiasts who would take and share photographs of planes. Members of this community are often called *plane spotters*, but are also referred to as *aircraft spotters* or *tail spotters*. The community's adoption of novel technologies like point and shoot cameras, DSLR cameras, mobile phones, and social media significantly changed the hobby. With internet websites, such as FlightAware and Airliners.net, and spotters now track and locate specific aircraft from all across the world. Spotters to upload their shots or see pictures of aircraft spotted by other people from all over the world. It is no surprise that Twitter hosts a large community of these spotters, and their desire to organize around a mutual interest, and share multimedia content makes their behavior similar to the threat group supporting networks discussed within this thesis. However, this group is far easier to evaluate with respect to ground truth because their membership is often self identified and the behaviors are easier to quantify.

I developed this dataset by conducting a two hop snowball sample of 12 popular plane spotters' mention ties from 1 September, 2014 to 1 September, 2015. The resultant yield was 518,410 user accounts' profile information and timelines resulting in a corpus of over 851M tweets. My intent is to use this dataset as *ground truth* when evaluating both clustering and classification techniques.

1.2.3 Crimean Conflict Movement Communities (CCMC) (Search Date: September 2015)

In an attempt to identify the anti-Russian, Ukrainian separatist movement on Twitter. I conducted a two step snowball sample of 8 known members' mention ties from March 2014 to September 2015. The search resulted in 92,295 Twitter accounts. Preliminary results enabled me to identify 3,895 accounts actively distributing anti-Soviet propaganda.

1.2.4 The CASOS Jihadist Twitter Community (CJTC) (Search Date: periodic)

This dataset is developed based on periodic samples of previously classified jihad supporting accounts. With each sample, I define known jihad supporting accounts as seed agents and a time interval, t. I then conduct a two hop snowball sample of my seed agents' mention ties.

Chapter 2

Related Work

2.1 Applied Network Science and Counter-Terrorism

Krebs [37, 38] was the first to cast large-scale attention on network science-based counter-terrorism analysis with his application of network science techniques to gain insight into the September 11, 2001, World Trade Center Bombings. Although similar methods were presented years earlier [17], the timeliness of Krebs' work caught the attention of the Western world and motivated a great deal of further research[16, 18, 24, 36, 40, 48, 59]. Much of this work focused on constructing networks based on intelligence and using the network's topology to identify key individuals and evaluate intervention strategies. The rise of social media has introduced new opportunities for network science-based counter-terrorism, and some foresee social media intelligence *(SOCMINT)* as being a major source in the future [34]. This presents a fundamentally different counter-terrorism network science problem. Roughly, as opposed to using information about individuals to build networks, I now use networks to gain insight into individuals. Typically, I am also trying to identify a relatively small and possibly covert community within a much larger network. Such a change requires methodologies optimized to detect covert networks embedded in social media.

2.2 Community Detection

The problem of community detection has been widely studied within the context of large-scale social networks and is well documented in works like Fortunato [29], Papadopoulos et al. [47]. Community detection algorithms attempt to identify groups of vertices more densely connected to one another than to the rest of the network. Social networks extracted from social media however present unique challenges due to their size and high clustering coefficients [31]. Furthermore, ties in online social networks like Twitter are widely recognized as having high social dimension, in that users ties represent different types of relationships [14, 41, 63]. This often requires community detection algorithms designed for specifically for online social networks (OSNs) to model user interactions as a multiplex graph or model user characteristics by annotating the nodes. For this reason, community detection has been explored in the dynamic multiplex case in Bazzi et al. [4], Mucha et al. [44], as well as the dynamic multimode case in Sun et al.

[51]. Although I will not explore the dynamic case in this proof of concept, it is worthy of follow-on research.

The Louvain Grouping algorithm presented in Blondel et al. [13] is widely used for community optimization within the network science community. Louvain grouping uses a similar objective function as the Newman-Girvan algorithm [45], but is more computationally efficient. In community optimization algorithms, the graph is partitioned into k communities based on an optimization problem that centers on minimizing inter-community connections where k is unspecified. Both Newman and Blondel find these communities by maximizing *modularity*. The modularity of a graph is defined in Equation 2.1. In Equation 2.1, the variable $A_{i,j}$ represents the weight of the edge between nodes i and j, $k_i = \sum_j A_{i,j}$ is the sum of the weights of the edges attached to vertex i, c_i is the community to which vertex i is assigned, $\delta(u, v)$ is the inverse identity function, and $m = \frac{1}{2}\sum_{i,j} A_{i,j}$.

$$Q = \frac{1}{2m}\sum_{i,j} = [A_{i,j} - \frac{k_i k_j}{2m}]\delta(c_i, c_j), \tag{2.1}$$

A problem with current community detection methods is their inability to account for high social dimension and return results that are precise enough to be useful in OEC detection [57] . To effectively cluster users in OECs the algorithm must leverage a user's friend ties, mention ties, and hashtag use to effectively cluster these ideologically grouped communities as found in [6]; however, in the absence of labelled cases neither IVCC nor Louvain grouping provides enough precision efficiently detect communities or effectively explore large OECs. Therefore community detection methods on multimode and multiplex graphs will be of interest. This work will contribute to community optimization across multiple graphs in a *meta-network* to facilitate vertex classification and detect a targeted covert community.

2.2.1 Annotated Networks

In recent years, another sub-class of community detection methods has emerged, community detection in annotated networks. This body of work attempts to effectively incorporate node level attributes into clustering algorithms to account for noisiness of social networks embedded in social media. Vertex clustering originates from traditional data clustering methods and embeds graph vertices in a vector space where pairwise, Euclidian distances can be calculated [46]. In such approaches, variety of Eigen space graph representations are used with conventional data clustering and classification techniques such as k-means or hierarchical agglomerative clustering, and support vector machines. These methods offer the practitioner great flexibility with respect to the types of information used as features. Vertex clustering and classification methods have been shown to perform well with social media because of their ability to account for a great variety of vertex features like user account attributes while still capitalizing on the information embedded in the graph; they also perform well at scale [57, 63]. [63] introduces a vertex clustering framework, *SocioDim*, which detects communities embedded in social media by performing vertex clustering where network features are represented spectrally and paired with user account features. Very similar methods are also presented in [12]. [57] then applies *SocioDim* to classification, which is analogous to a binary partition of the graph.

2.2.2 Multilayer and Heterogeneous Networks

These methods show clear promise with respect to covert network detection in social media as illustrated by [41]. Eigenspace methods have been shown to adequately model multiplex representations of various types of social ties in social media [58], and early studies of simulated networks indicate they would perform well on threat detection in social media [41]. I hypothesize that eigenspace representations of heterogeneous network representations of OSNs, when paired with user account features and node level features will provide a more powerful means to detect threat groups embedded in social media. My work utilizes community optimization across multiple graphs in a *meta-network* to facilitate vertex classification and detect a targeted covert community. In sum, I have found that each of the methods listed above offer useful information for classification, but a combination of these techniques must be used to effectively detect covert networks embedded in social media. Chapter 3 iterative vertex clustering and classification, a method to leverage multiplex, multimode, annotated graphs to conduct network bipartition as a classification task. Jiawei Han's group at University of Illinois at Urbana-Champaign has presented clustering methods on large heterogeneous networks with promising results[35, 50, 52, 53, 54, 55]. In each of these papers the authors use a combination or ranking and clustering to identify clusters of multiple node classes, which could be used to detect OECs through user and hash tag ties.

2.3 Social Bot Detection

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 [2]. Identification and removal of bots is important to measure opinion within OECs, but bot detection also helps identify narratives that extremist groups are trying to promote. ISIS' use of bots has been well documented [7], 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 [27] and abroad [28], the use of social bots has been documented in the MENA region [2, 3], and ISIS use of them motivated a DARPA challenge to develop detection methods [49].

2.4 Statement of Work

The work presented in this thesis will lay a foundation for future work, and I argue that the most useful knowledge extraction from OECs will require collaboration among data scientists, social scientists, and regional experts. Both a taxonomy for continued research and extended literature review of related works would help researchers and practitioners motivate theoretically rigorous and operationally useful objectives. As a result I propose my background chapter as an extended literature consisting of applicable research related to OEC detection, analysis, and disruption directed at two groups:

- Researchers from non-quantitative fields such as Political Science, Security, and Radicalization
- The Intelligence Community, Policy Makers, and Senior Leaders within government

The review will be structured with a brief introduction and subsequent sections on detection, analysis, and disruption of OECs. Each section will introduce related research from the past 10 years and emphasize current limitations. Within each section I will also make recommendations for future research.

Chapter 3

Completed Work: OEC Detection via Classification

> **Methodological Task 1 (MT1):** *Given a large meta-network with annotated nodes that has an embedded community of interest and a set of labelled training data, perform a bipartite partition of the network to identify a large proportion of the community of interest.*

3.1 Model Overview

In this section, I describe Iterative Vertex Clustering and Classification, a supervised method to extract OECs from large OSNs. Detecting OECs requires identifying a relatively small subgraph within a large, meta-network; However, distinguishing between *members* and *non-members* is a challenge. Although one would expect to find ideologically organized communities within a sample with high levels of interconnectedness, or modularity, it is difficult to distinguish between the community of interest and other communities captured within one's sampling strategy. Furthermore, accounts with excessively high mention or following counts like celebrities, news media, or politicians often need to be systematically removed to attain useful results. I refer to these as *official accounts*. This complication requires an iterative process where multiple classifiers are trained and applied.

To explain my methodology I introduce the following notation. Let $G = (V_1, V_2, .., V_n, E_1, E_2, .., E_m)$. Where G is a directed, weighted graph with vertex sets $V_1...V_n$. Each contains vertices $v_{n,1}..v_{n,j}$ with one or more edge types $E_1, E_2, .., E_m$. I define a subset of targeted vertices $A_t \subseteq V_t$ and denote its complement as $\tilde{A}_t$. My goal is to accurately classify each vertex in V_t as members of either A_t or $\tilde{A}_t$. The challenge lies in discriminating between the two. In practice, one will often have partial knowledge of the targeted group and its members, and will leverage as much information as possible to identify vertices in A_t. My approach is conducted in two phases as depicted by Figure 3.1. In phase I, community optimization algorithms and a priori knowledge are used to gain insight into the larger social network and facilitate supervised machine learning in phase II. Phase II partitions vertices to remove vertices not belonging to A_t and find the targeted covert community.

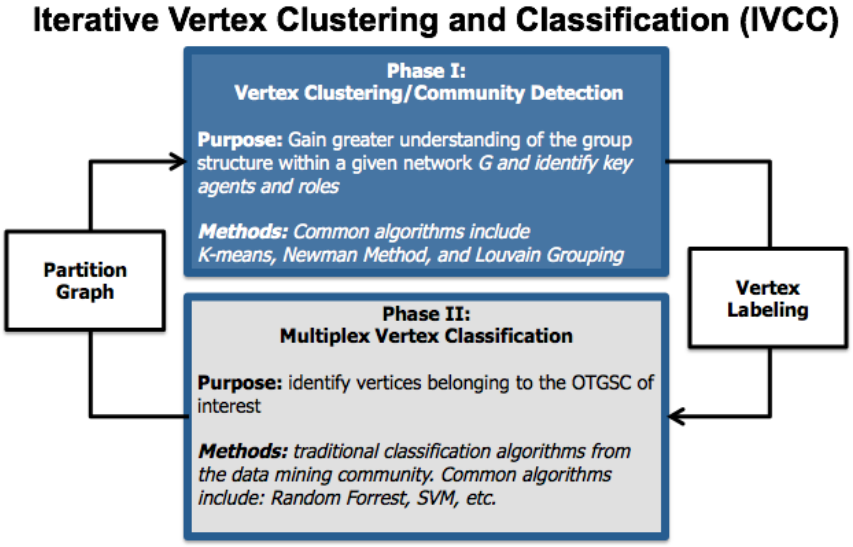
Iterative Vertex Clustering and Classification (IVCC)
Phase I:
Vertex Clustering/Community Detection
Purpose: Gain greater understanding of the group structure within a given network G and identify key agents and roles
Methods: Common algorithms include K-means, Newman Method, and Louvain Grouping
Partition Graph
Vertex Labeling
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.

Figure 3.1: I present an iterative methodology conducted in two phases. In phase I either community optimization or vertex clustering algorithms are used to remove noise and facilitate supervised machine learning to partition vertices in phase II.

3.1.1 Model

Phase I: Vertex Clustering and Community Optimization
Although community optimization and vertex clustering methods will often fail to accurately partition my networks into A_t and $\tilde{A}_t$ [41], one can often look for community structure within the network to gain insight into A_t. For example, if a subset of vertices from A_t is known, community optimization can identify clusters containing a large proportion of those known vertices belonging to A_t. Community optimization can also identify vertices that are clearly members of $\tilde{A}_t$. The insights gained from community optimization help provide necessary context with respect to algorithm selection and case labels for vertex classification in Phase II of my methodology.

Phase II: Multimode Multiplex Vertex Classification
Like [57] I classify $v_{t,1}...v_{t,j}$ using a set of features extracted from the users' social media profiles and spectral representations of the multiplex ties between V_t. I denote these spectral representations as $U_{V_t \times V_t; E_i}$, where $i = 1, ..., m$. To develop spectral representations of the meta-network, I symmetrize the graphs $W = G_{V_n \times V_n; E_m}$ for $\forall E_m$. These symmetric graphs also leverage the strength of reciprocal ties, which have been shown to better indicate connection in social networks embedded in social media [20, 30, 42]. I then extract the eigenvectors of the graph Laplacian associated with the smallest two eigenvalues as highlighted in [62], and concatenate them as presented in [58]. A graphical depiction of this feature space is provided in Figure 3.2. This

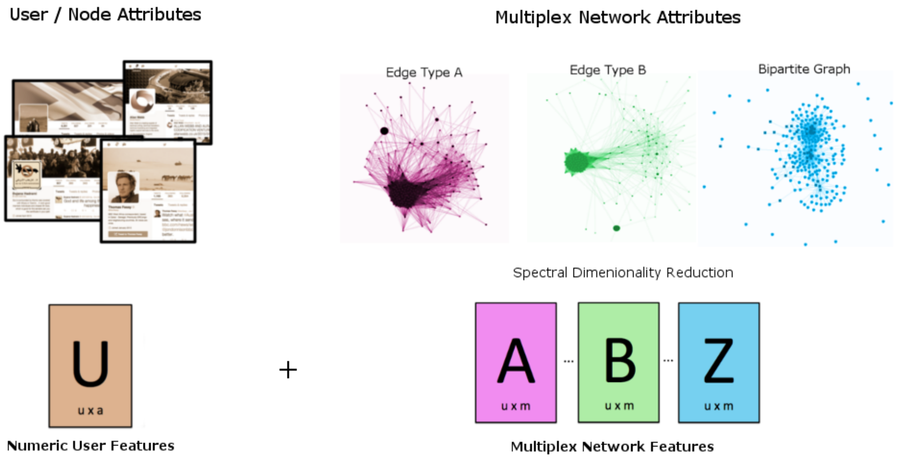

Figure 3.2: in phase II I incorporate node level and network level features by extracting lead Eigen vectors from various network representations of social media ties.

enables one to effectively capture the distinct ties represented in many types of social media, as well as node level metrics of each graph and user account features.

Users often use topical markers, like hash tags in Twitter, and these can be used to cluster users with similar topical patterns. This results in bipartite graphs, $G_{V_t \times V_n, E_m}$, where users and topical markers represent differing node sets that can prove useful to co-cluster users. To do so I implement bispectral clustering as introduced by [22] as a document clustering method. In this case, instead of co-clustering documents based on word frequency, one co-clusters users based on hashtag frequency within their tweets. I develop $W_{V_t \times V_n}$, where $w_{i,j} \in W_{V_t \times V_n}$ represents the number of time vertice $v_{n,j}$ appears in the twitter stream of $v_{t,i}$. To co-cluster $v_{t,1} ... vt, n$ I follow the biparitioning algorithm provided in [22], which results in Eigen vector features similar to those I defined in the previous paragraph.

The combination of user account attributes, node level metrics from the larger network G, and spectral features explained above provide a rich feature space. Paired with a reasonably sized set of labeled vertices, one can detect an OEC using any number of classifier algorithms. I have found the decision trees and support vector machines perform well. Again, in IVCC, one conducts classification iteratively. The first classifier removes *official accounts*, while the second classifier identifies members of the OGTSC of interest. Although many classification techniques can be used within my framework, due to performance I find the *Random Forest* [**?**] and Support Vector Machine classifiers [57] most useful in my preliminary results.

3.1.2 Results

When applied to the **ISIS NOV14** dataset the resultant classifier yielded accuracy of 91.3 % and a Kappa score of 75.8 % significantly outperforming methods introduced by Tang and Liu [57], Tang et al. [58] as illustrated by Figure 3.3.

Although it appears IVCC performs well at a supervised learning task, it does not provide adequate results phase II is conducted as an unsupervised or semi-supervised learning task. Fig-

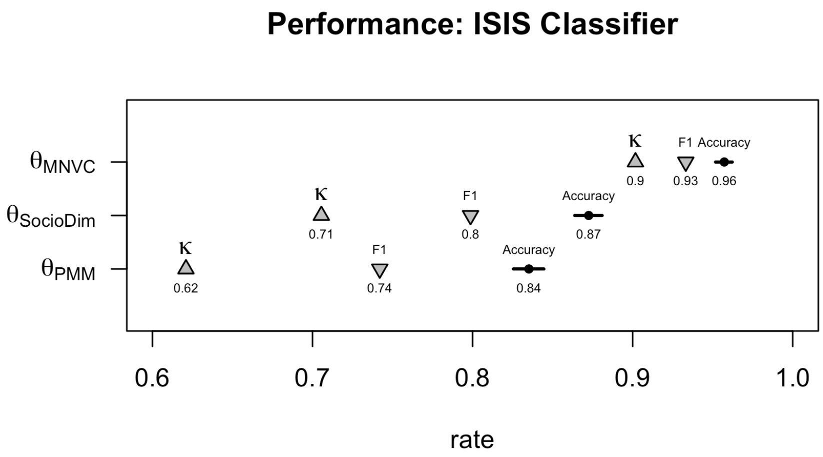

Figure 3.3: This plot graphically depicts classifier performance for the three trained ISIS classifiers. Performance was estimated using a 60% / 40% train / test split.

ure 3.3 depicts the relative performance of various subsets of the IVCC feature set when applied using unsupervised approach to the **ISIS NOV14**. The plot depicts recall vs. false detection or ROC curves of the spectral clustering results. Recall is defined in each curve as proportion of accounts classified as ISIS supporting in the aforementioned work with only clusters of size greater than 100 are depicted. Clusters were found where k was arbitrarily selected as 1000 and using the $l = \lceil log_2 k \rceil$ [23] lead eigenvectors of each respective graph's Laplacian. The relative performance of Φ_{s,F_r} (red), Φ_{s,M_r} (black), $\Phi_{b,U_{u \times ht}}$ (green) graphs highlight the relatively poor performance of each feature set with respect to the learning task. The full feature set, Φ_{IVCC} (cyan) illustrates the limitations of IVCC as a semi-supervised method, and highlights the importance of my second research objective.

3.2 Limitations

IVCC offers a scalable, annotated network analytic approach for OEC detection which outperforms existing approaches on the classification task of identifying ISIS supporting. Unsurprisingly, the limitations of these results motivate additional research. The following limitations are directly related to the research questions presented in Chapter **??** and motivate the proposed work presented in subsequent chapters of this document:

Limitation 1: IVCC requires large training sets to achieve high recall in some cases. Figure 3.4 depicts the precision, recall, and Kappa scores acheived randomly selected training sets of varying size. Although IVCC acheives over 80% precision with less than 200 labeled instances, it does not achieve over 80% recall until the training set is over 4000 labeled instances. It is unlikely that large, low-cost training sets similar to those presented in [6] would be available for OEC detection. Furthermore, the 10,000 instances used to train the classifier in [6] achieved an estimated performance of over 90%. For these rea-

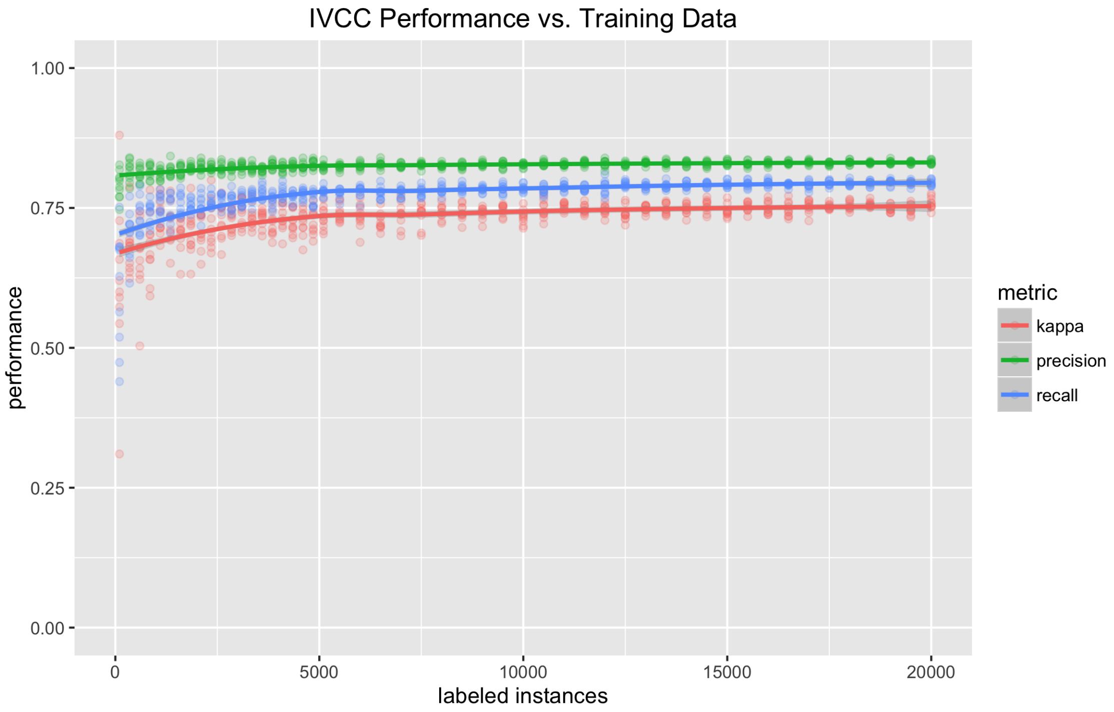

Figure 3.4: depicts recall, precision, and Cohen's Kappa for training sets of varying size. 15 training sets are generated at varying sizes (depicted on the x-axis), and each performance metric is calculated based using [6] as ground truth. The plot highlights both the need for large amounts of training data to gain adequate recall, and the value of an active learning framework as random labeling shows only minimal improvement after 5000 instances.

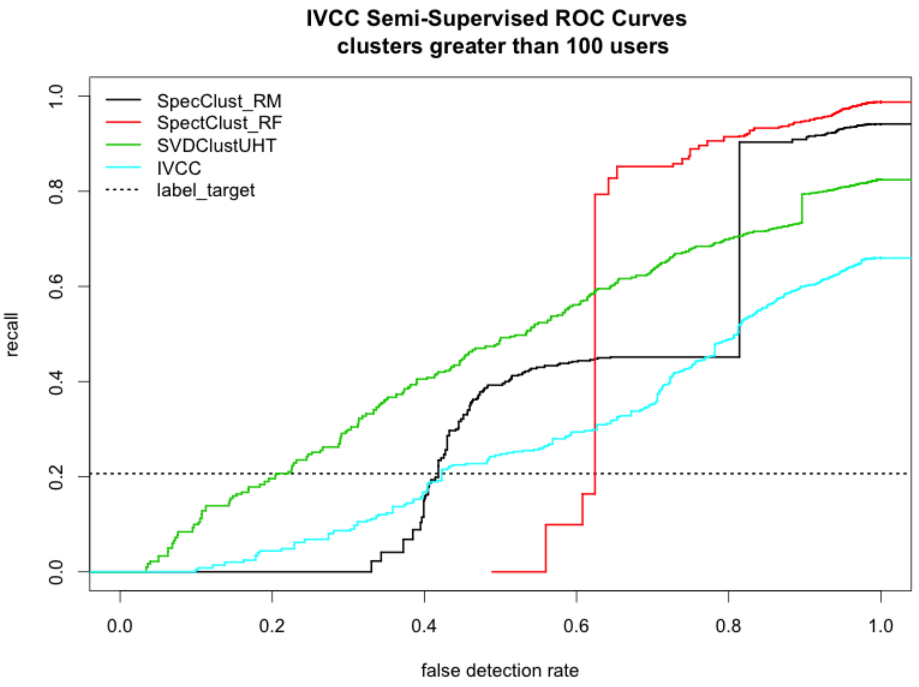

Figure 3.5: depicts recall vs. false detection or ROC curves of spectral clustering results associated with subsets of the feature space presented in [6] and using the authors' supervised results as ground truth. The plot highlights Iterative Vertex Clustering and Classification's limitations when applied to a semi-supervised learning task.

sons, I propose an active learning framework which will be discussed in greater detail in Chapter 5.

Limitation 2: The feature space developed for Multimode Multiplex Vertex Classification performs poorly as a clustering technique, as do most network clustering approaches. Figure 3.5 highlights the shortcomings of both methods by presenting ROC curves for each method using the November 2014 CASOS ISIS Search Database and the results presented in [6] as ground truth. Clearly these limitations motivate extensions to enable OEC detection when training data is scarce and proposed work will be presented in Chapter 4.

Limitation 3: Standard evaluation methods do not appear to provide strong estimates of performance. In practice, I have found that model evaluation requires both standard evaluation metrics like F1 score or Cohen's Kappa, but these metrics must also be viewed in concert with random sampling from model output. Model output appears highly sensitive to feature space selection, but more formal analysis needs to be completed. IVCC utilizes both random sampling for negative case instances as well as greedy algorithms. A detailed analysis of uncertainty assocated with both inputs is needed and will be addressed in Chapter 5.

Limitation 4: Finally, extracting information from OECs is a research area unto itself. In this case study I identify a community of nearly 23,000 members and provide illustrative intelligence extractions. However, it is likely that additional extractions are possible. This emergent research area will require collaboration among teams of experts possising technical and regional expertise, as well as an understanding of the information needs of policy makers. This diverse community would be well served by an extended literature review and framework for detection, analysis, and disruption of OECs. One specific extraction of value, would be an understanding of the techniques used within OECs to gain social influence and manipulate discussion. To address the challenge of extracting intelligence from OECs, I will formulate the background section of my thesis as a research framework for detecting, analyzing, and disrupting OECs and provide an extended literature review.

Chapter 4

Semi-Supervised OEC Detection and Exploration

Methodological Task 2 (MT2): *Given a large annotated heterogeneous network and limited training data, extract identifiable clusters of embedded OEC members.*

4.1 Model Overview

As presented in Chapter 3, semi-supervised OEC detection remains an open problem. One limitation of existing community detection methods is their inability to account for high social dimension and return results that are precise enough to be useful in OEC detection [57] . To effectively cluster users in OECs the algorithm must leverage a user's friend ties, mention ties, and hashtag use as found in [6]; however, in the absence of labeled cases neither IVCC nor modularity-based clustering techniques provide enough precision efficiently detect communities or effectively explore large OECs. Figure 3.5 highlights the shortcomings of both methods by presenting ROC curves for each method using the November 2014 CASOS ISIS Search Database where I use the results presented in [6] as ground truth. For IVCC to be more generalizable to situations when training data is scarce, stronger unsupervised or semi-supervised methods are needed.

As stated in the previous paragraph, neither modularity-based clustering techniques nor clustering of feature spaces similar to those presented in [6] provide adequate precision. In fact, with respect to growing a training set, precision becomes far more important than recall, yet few unsupervised methods have been developed to emphasize precision. An unsupervised method to detect *core communities* could be implemented in Phase I of IVCC to grow the training set, as implemented in [5]. Novel research presenting clustering methods on large heterogeneous networks could provide promising results[35, 50, 52, 53, 54, 55], or an alternative to standard community detection strategies could be to only look for what I call *core communities*.

core community: an identifiable subset of larger online community of activists.

To detect these core communities I propose an ensemble of clustering techniques that leverage users' friend, mention, and hashtag patterns. To present preliminary results I introduce the following notation. Let $\Phi_{A,G}$ be a clustering algorithm A applied to graph G where:

$A \in \{$ l = Louvain Grouping, s = Spectral Clustering, b = Bi-spectral Co-clustering $\}$

$G \in \{F_{red}, M_{red}, U_{u \times ht}\}$ as defined in Table 4.1

Then, $\vec{\theta}_{A,G}$ is a vector of length n and represents the cluster assignments of $\Phi_{A,G}$, where n is the number of nodes in G, and each entry $\theta_{A,G,i} \in \{1, .., k\}$ is equal to the respective node's cluster assignment.

Table 4.1: Depicts G, the resultant heterogeneous network constructed from the ISIS14 dataset.

Network	F	F_{rec}	M	M_{rec}	$U_{u \times ht}$
Description	Following	Reciprocal Following	Mention	Reciprocal Mention	User Shared Hashtag
From Node	User	User	User	User	User
To Node	User	User	User	User	Hash Tag
Link Type	directed, binary	undirected, binary	directed, weighted	undirected, weighted	directed, weighted
Nodes	119 k	119 k	109 k	109 k	106 k x 4 M
Links	23.1M	3 M	14.6 M	1.1 M	27.4 M
Density	0.00163	0.000425	0.00123	0.00018	0.000065
Isolates	0	10888	291	30,047	0
Dyads	0	104	6	425	188
Triads	0	19	0	50	33
Larger	1	8	2	7	6

To develop community assignments based on users' mention and friend ties I use Blondel et al. [13], which has become a benchmark community detection algorithm in large graphs [29]. The algorithm maximizes modularity (Eq. 2.1) by first sequentially calculating the modularity gain associated with adding vertex i to its nearest neighbor j's community, and always selects the individual assignment which provides the greatest gain. In its second step communities are replaced by super-vertices, and two super-vertices are connected if there is at least an edge between vertices of the corresponding communities. These steps are repeated recursively until modularity no longer increases. In my ensemble I will develop pairs of community assignments for each user by running this algorithm on the reciprocal mention and reciprocal following networks. However, Blondel et al. cannot be used on the user x hashtag graph and other methods must be employed.

I develop community assignments for each users' hash tag use via bipartite graph partitioning [23]. Although there is no benchmark method for co-clustering in large heterogenous OSNs, I find Dhillon [23] outperforms latent semantic analysis [21]. Dhillon et. al. used spectral clustering for topic modeling of documents by forming matrices composed of the lead eigenvectors of both the left and right matrices of the singular value decomposition of a bipartite graph, concatenate them row-wise, and then cluster the nodes using the k-means algorithm. In their case, both documents and words define the kmeans clusters. In this case one clusters users and hashtags.

I extract core communities by creating sets of users where all three models are in agreement. In other words, I look for communities that mention one another, follow one another, and use similar hash tags. Formally I define $C_k \subseteq V$ as one of k core communities consisting of users $v_i, i = 1, .., n$ where each model $\Phi_{A,G}$ assigns v_i and v_j to the same communities or $C_k = \{v_i, v_j | \theta_{A,G,i} = \theta_{A,G,j} \; \forall \; A, G\}$. I then define the model $\Phi_{ensemble}$ where Louvain grouping is used to cluster the reciprocal mention and reciprocal following networks, and bispectral co-clustering is used to cluster users with the user x hashtag graph.

To explore the utility of a simple semi-supervised ensemble to extract core communities from a large OSN I test the simple model $\Theta_{ensemble}$, and evaluate its performance by assuming the results presented in Chapter 3 to be ground truth. Figure **??** the ROC curves associated with:

Φ_{l,M_r}: Louvain Grouping run on the symmetrized mention network (green line)

Φ_{l,F_r}: Louvain Grouping run on the symmetrized friend network (blue line)

$\Phi_{b,U}$: Bispectral Co-clustering with k=100 run on the user, hashtag bipartite graph (black line)

$\Phi_{ensemble}$: the ensemble of each of the above models as defined in Section **??** (red line)

The black dotted horizontal line depicts the number of labelled accounts used in [6]. $\Phi_{ensemble}$ returned 14,714 of 119k users clustered into 63 core communities. 12 of core communities contained a combined 2514 accounts where 97.175 % of the users were previously classified as *ISIS Supporting*. Although [6] constructed a training set with nearly 5,000 positively labeled cases, it is possible a smaller training set could have achieved similar results. Furthermore, the results quickly identify core communities that can be used as negative case examples. [6] use random sampling to label negative cases which could induce overfit due to differences in the expected density of within positive and negative case examples in my graphs . The result would be a classifier that simply looks for communities. Core communities reduce such bias. Figure **??** illustrates the precision,ξ_k, for each of the 63 core communities detected. In fact if I use $\alpha = .05$ and label only cases where where $\xi_k \ni [\alpha, 1-\alpha]$ then the result is C_+, a set of 2514 positive case examples containing 97.175 % ISIS supporters, and C_-, 5869 negative case labels at containing .008 % ISIS Supporters. It is likely that some additional random sampling of negative cases would be optimal which I offer as worthy of future research. Figure **??** illustrates the utility of core communities to an intelligence analyst. The precision of class within many core communities imply they would be qualitatively identifiable. The 3 account profiles depicted in the top half of the figure were randomly selected from C_+. The bottom three accounts were selected from C_-, yet this community displays a common activist cause as each member appears to actively promote Scottish independence from the United Kingdom. The output of $\Theta_{ensemble}$ certainly appears to offer utility with respect to my second research objective.

4.2 Statement of Work

Although I find the preliminary results $\Theta_{ensemble}$ encouraging, more rigorous exploration of alternatives is required. Some of the findings developed in Mucha et al. [44] could be extended by applying the methods for modularity maximization on weighted, bipartite graphs in Dormann and Strauss [25]. Dormann and Strauss [25] extends Newman Girvan further by accounting for weighted bipartite graphs with their presentation of the QuaBiMo algorithm and could provide a stronger alternative to Dhillon [23]. Similarly the rich body of literature on heterogeneous community detection shows promise as well. I propose two models for future examination:

1 Independent Heterogeneous Modularity Maximization (IHM2): The IHM2 model would independently detect communities using modularity maximization for each edge type in the multiplex graph as well as for the source node class bipartite graphs. The algorithm would then return communities where each detection strategy was in agreement, similar to the methods illustrated in $\Theta_{ensemble}$. The results of this algorithm would likely be useful for OEC detection when labeled data is scarce.

2 Weighted Multiplex Multimode Modularity Maximization(WHM2): The WHM2 model would a linear combination of the modularity functions for each edge type in the multiplex graph as well as for bipartite modularity functions as well. The results of this algorithm would likely be most useful for exploratory data analysis of large OECs.

Evaluation of these models will require at least one more dataset where ground truth is known. For this task I have developed the Twitter Plane Spotters dataset. Aircraft spotters are an active community on Twitter. This global community shares high quality photographs of planes and shares this content for various purposes. Some just share photographs, while some attempt to recreate flight networks by sharing data globally. Some focus exclusively on military airfields. Though none of this behavior is criminal, the community shares content in ways quite similar to the groups mentioned earlier. Furthermore, because their behavior is not criminal they often self identify which could be used as ground truth. A 2-step snowball sample of 8 popular aircraft spotters' friend ties yielded 300k users. I am in the process of determining how many of these users self identify as plane spotters in their account description, but expect that there will be enough examples to train a supervised classifier and achieve a ground truth dataset similar to the November 2014 ISIS Search dataset presented here. Furthermore, I will examine IHM2 and WHM2 performance with RankClust [51] and PathSim [53] as baseline models.

Chapter 5

Monitoring Dynamic OECs: An active learning framework for robust monitoring highly dynamic online communities

Methodological Task 3 (MT3): *Given a large dynamic OGTSC, maintain understanding of group activity and interests.*

5.1 Model Overview

In most cases the ultimate goal will be to disrupt of neutralize OECs and the tools used to do so could be applied in the diplomatic, industrial ,military or economic realms, and within each realm interventions could be conducted inside or outside the cyber domain. As a result, assessing network dynamics will be required. Maintaining understanding of these dynamic communities will require a principled methodology. This work will propose a procedure to periodically update OECs and address three challenges associated with this task:

- effectively sampling known OEC members' activity to identify changes in membership
- maintaining adequate training data in highly dynamic OECs
- understanding the sources of uncertainty with respect to OEC membership

The groups associated with Sunni Islamic extremism have proven highly dynamic and resilient over the past 10 years. They adapt quickly to changes in tactics employed against them, as well as changes in public opinion in areas where they garner support. If recent history is an indicator, ISIS' success in social media will be copied wether the group maintains a position of relative power or not. It is likely that OECs will change in both membership and allegiance over time, making monitoring them a significant but important challenge. Principled means to search for new or reconstituted members of the community and methods to grow adequately sized training sets are needed.

Choosing an appropriate sampling strategy will be a function of the user's understanding of the monitored OEC. For example, depending on OEC behaviors a mention-based snowball sample may provide a more fruitful search. In some cases a following-tie-based search may be more effective. The first portion of this chapter will present a rigorous discussion of sampling techniques and the OEC behaviors associated with informed strategy selection.

The most significant contribution of this chapter will be methods to develop adequate training data through active learning. As depicted in Figure 3.4 IVCC requires large training sets to reach goals with respect to recall and its precision is sensitive to false positive cases within the training set. Therefore, simply using previous model output as positive case training labels when conducting periodic updates leads to unsatisfactory results. Furthermore, highly dynamic networks like those seen in the ISIS case study make many of the researchers positive case labels invalid for updates. Active learning provides the technical framework needed to efficiently gain performance while minimizing the high costs associated with manual labeling.

I propose an active learning framework that incorporates output sampling into an uncertainty-sampling-based approach. As stated in previous chapters, performance evaluation in IVCC requires both standard evaluation techniques as well as output sampling, due to sensitivity to noise in the training set and sensitivity to feature selection. However, the feature space associated with IVCC lends itself to standard uncertainty-sampling techniques with one caveat; incorporation of blindly labelled classifier output when defining the decision-boundary and uncertainty.

As a methodology, IVCC has a great number of methods that contribute to the uncertainty with respect to its results. In almost all cases the clustering techniques used in Phase I are greedy searches, negative case labels are currently selected through random sampling, and the classifiers used are often greedy as well. How much uncertainty is associated with each of these and how dramatic is the effect of them? If researchers are to gain an understanding of network dynamics, some understanding of this uncertainty must be developed.

5.2 Statement of Work

The end result of this chapter will be a repeatable methodology to monitor an OEC through periodic search and detection. Such a methodology will require:

1 A consistent sampling methodology that is informed by observed OEC behavior

2 An uncertainty-sampling-based active learning framework which incorporates precision and recall estimates into decision boundary calculation.

3 A rigorous understanding of uncertainty with respect to community membership and its effects on analysis.

Sampling methods will be evaluated using the CJTC, CCMC, and PSTC datasets. Specific behaviors within each dataset will be compared to provide illustrative examples of appropriate sampling techniques for periodic updates. The active learning framework will be evaluated using the ISIS NOV14 dataset and assuming the results of [6] as ground truth. Finally I will assess uncertainty across each dataset with respect to clustering, sampling, and classification methods by quantifying the change in output associated with each. Finally, this chapter will also provide a case study of the CJTC with 4 periodic updates in March, July and November of 2016. The case study will be used to compare detected OEC changes over time with ongoing events in the region.

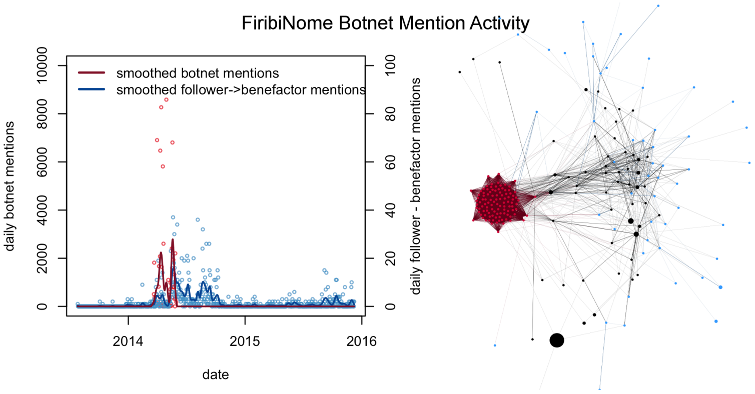

Figure 5.1: Depicts mention behaviors and their effects within the FiribiNome Social Botnet. The left panel depicts two scaled time series. The red circles and smoothed trend line depict the number of daily mentions by botnet members. The blue circles and corresponding trend line depict botnet followers' 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.

Chapter 6

Social Influence within OECs: from detection to disruption

6.1 Overview

Ideally policy makers would like to neutralize the effectiveness of extremist propaganda and recruitment within OECs, but efforts to intervene within these communities have not produced desired results [26]. One possible shortcoming to previous approaches to gaining influence within OECs is a lack of understanding of the social network topology within the OEC. Detection and analysis of OECs offers insight. By isolating a large proportion of the network of interest I can observe behaviors used to gain influence; I can also leverage social network theory to understand what users offer the most potential with respect to interventions. In the ISIS14, CJTC, and CCMC similar mention behaviors appear. Some are used within a social botnet, while some appear to be individual users, but in each case accounts will tweet with high daily volume and "spend" a large portion of their 140 characters mentioning highly central accounts within the OEC. This behavior appears to enable user to develop a large following within the OEC and in some cases increase the social influence of users that it mentions.

One sophisticated example of this type of behavior is the FiribiNome social botnet, a group of over 100 bot accounts designed to gain a large following of Jabhat al-Nusra supporters while promoting the propaganda of highly central figures within the community. 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 [2]. ISIS' use of bots has been well documented [7], 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 [27] and abroad [28], and has been documented in the MENA region [2]. ISIS' use of them motivated a DARPA challenge to develop detection methods [49]. To describe how this behavior contributes to social influence, I define three account types:

botnet workers: These accounts come online and tweet strings of @mentions where a large portion of the accounts mentioned are fellow botnet workers and a relatively small number of the accounts mentioned are benefactors. Botnet workers make no attempt to appear to be human users.

benefactors: These accounts are mentioned by botnet workers and do not mention the botnet workers. They accounts are typically highly central with large following counts within the OEC.

botnet followers: These accounts follow one or more of the botnet workers. They often follow and mention benefactors. My hypothesis is that this botnet structure somehow facilitates intra-botnet-follower connections as well as increased benefactor social influence.

Figure 5.1 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 Figure 5.1, 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 Figure 5.1. Their tweets consist of original posts or retweets containing strings of @mentions of other *botnet members*, but occasionally mention or retweet content from what I call *benefactor accounts* (depicted by black vertices in the right panel of Figure 5.1). The *botnet* account FiribiNome20 illustrates this behavior. In isolation, these accounts appear to be producing spam and relatively harmless, however my analysis indicates the network of *botnet members* increases the social influence of *benefactor accounts*. The blue series in the left panel of Figure 5.1 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 between 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 6.1; 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[60]. @*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. I 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.

I have observed similar behavior in the *reconstitutor* community and the anti-Soviet pro-

Table 6.1: 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

paganda sharing community in the Ukraine. *Reconstitutors* are CJTC accounts that have been suspended by Twitter and come back online under new usernames and try to regain centrality within the community. $D2\_\_M$ is an example of one such account. Here the user posts a combination of mentions and hash tags to regain a following of users. The $D2\_\_M$ account was created on February 20, 2016, and as of May 9, 2016 had posted 288 times and gained a following of 656 users. Within the CCMC, I observe an image sharing community with similar structure to FiribiNome. This group of users post images of women and anti-Soviet propaganda. While the accounts try to gain a following, each post contains a string of mentions of accounts highly followed with similar behavior. As the accounts gain followers, the mention activity appears to subside. The top 20 accounts in terms of mention per tweet ratio within the CCMC have a cumulative following count of over 100K users.

The behaviors observed within each of the aforementioned datasets indicate that an account can gain a targeted following by understanding the topology of the OEC and mentioning highly central members of the community. However, this behavior has not been formally studied.

6.2 Statement of Work

The main goal of this section is to validate the hypothesis that mentioning accounts highly central within an OEC can be used to gain followers within it. To accomplish this, I will complete the following:

1 develop detailed analyses of: the FiribiNome Botnet, Donbas Propaganda Network, and reconstitutors within the CJTC. Identify correlation between followers of mentioned accounts and mentioner follower ties, as well as generated mention content.

2 create a set of plane spotter bots and utilize various @mention behaviors in an attempt to statistically prove the hypothesis that this behavior grows a user's following within a targeted community.

Chapter 7

Conclusion

The ability to detect, analyze, and disrupt OECs will continue to be an important capability as neither terrorist groups nor their use of social media will abate in the near future. Furthermore, methods that can increase the understanding of the passive support structure essential to the distribution of extremist propaganda will be necessary to shape effective strategies to counter this propaganda. In essence, understanding this passive support structure means understanding the demographic that these groups are competing for. This thesis will introduce methods foundational to this area of study and establish a framework to grow a consortium of practitioners and researchers from academia, industry, and government to develop these capabilities so critical to national security and human rights.

7.1 Contributions

This thesis will provide four major contributions to the study of OECs. First, I introduce iterative vertex clustering and classification in Chapter 3 and extend it in Chapter 4 by incorporating unsupervised methods to quickly gain positive case examples. I will then extend IVCC in Chapter 5 by introducing an active learning framework to more efficiently incorporate regional expertise in detection. I will also provide formal analysis of the uncertainty associated with the greedy algorithms used in the IVCC pipeline. These contributions with enable researchers to use IVCC to monitor highly dynamic OECs. I will contribute to the theoretical understanding of how propagandists manipulate community structure to gain social influence in Chapter 6. Chapter 2 will provide a much needed framework and extended literature review that will help researchers, industry, and government effectively direct future research in this important area. The case studies used in each of these chapters also will provide novel understanding of important socio-political discussion within social media. My hope is that the case studies motivate other researchers to mine these large datasets in future work.

7.2 Limitations

Of course the contributions of this thesis will not be without limitations. My ability to evaluate how precisely the methods presented can discern ideological substructure is limited due to lim-

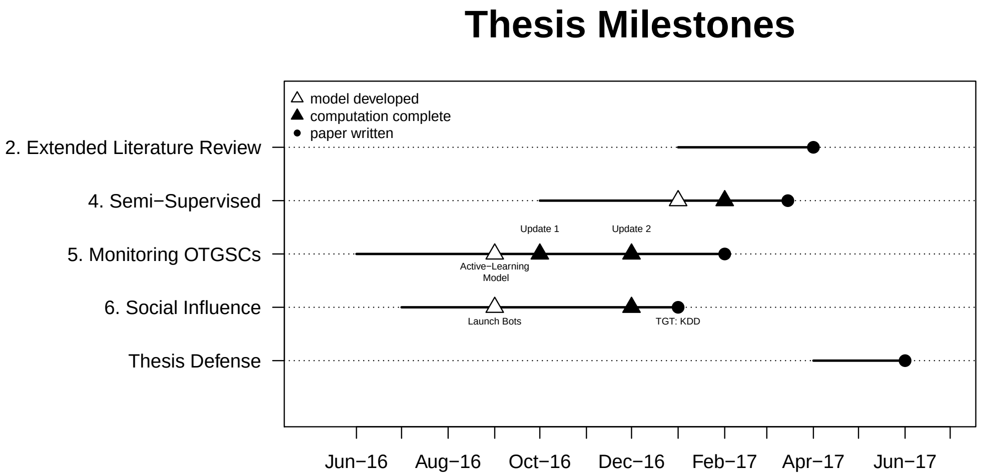

Figure 7.1: timeline of thesis milestones.

ited access to regional expertise. The uncertainty introduced by this limitation must be clearly communicated as to not overstate or understate the potential of the methods presented. I am also limited in that I will only use Twitter data. Although it is generally accepted that these groups utilize a broad range of social media, our lab's access to Twitter data makes it the most logical choice to evaluate my methods.

7.3 Timeline / Milestones

Due to the constraints placed upon my status as a full time student by the Army, I will pursue an aggressive timeline to complete the proposed work. Table **??** describes the major milestones associated with each chapter of my thesis, and they are visually depicted in Figure 7.1. I want to thank all of you for agreeing to be members of my committee. I welcome your feedback, advice, and mentorship throughout the completion of my proposed work.

Table 7.1: describes thesis milestones depicted in Figure 7.1.

Date	**Milestone**
September 2016	**Ch.5** Develop and implement active-learning framework for OCT and DEC CJTC updates **Ch.6** Launch PlaneSpotter news aggregators with $@mention$ implementation
October 2016	**Ch. 5** Update CJTC
December 2016	**Ch. 5** Update CJTC **Ch. 6** Analyze PlaneSpotter aggregators
January 2017	**Ch. 4** KDD submission written **Ch. 6** Dense community model developed
February 2017	**Ch. 5** paper complete, venue TBD **Ch. 4** results complete
March 2017	**Ch. 4** paper written, venue TBD
April 2017	**Ch. 2** extended literature review written, venue TBD
May 2017	**Thesis** write up complete
June 2017	**Thesis Defense**
July 2017	**Thesis Revisions Complete**

List of Figures

List of Tables

Bibliography

[1] Anonymous exposes US and UK companies hosting pro-Isis websites. URL http://www.ibtimes.co.uk/anonymous-exposes-us-uk-companies-hosting-pro-isis-websites-1495426.

[2] Norah Abokhodair, Daisy Yoo, and David W. McDonald. Dissecting a Social Botnet: Growth, Content and Influence in Twitter. pages 839–851. ACM Press, 2015. ISBN 978-1-4503-2922-4. doi: 10.1145/2675133.2675208. URL http://dl.acm.org/citation.cfm?doid=2675133.2675208.

[3] Samer Al-khateeb and Nitin Agarwal. Examining Botnet Behaviors for Propaganda Dissemination: A Case Study of ISIL's Beheading Videos-Based Propaganda. In *Data Mining Workshop (ICDMW), 2015 IEEE International Conference on*, pages 51–57. IEEE, 2015. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7395652.

[4] Marya Bazzi, Mason A Porter, Stacy Williams, Mark McDonald, Daniel J Fenn, and Sam D Howison. Community detection in temporal multilayer networks, with an application to correlation networks. *Multiscale Modeling & Simulation*, 14(1):1–41, 2016.

[5] Benigni, Matthew. Tutorial: Online Threat-Group-Supporting Community Detection, April 2016. URL http://dscoe.org/ABIDSTutorial/.

[6] Benigni, Matthew, Joseph, Kenneth, and Carley, Kathleen. Threat Group Detection in Social Media: Uncovering the ISIS Supporting Network on Twitter. *Submitted to Plos One*.

[7] J. M. Berger. How ISIS Games Twitter. *The Atlantic*, June 2014. ISSN 1072-7825. URL http://www.theatlantic.com/international/archive/2014/06/isis-iraq-twitter-social-media-strategy/372856/.

[8] J. M. Berger and Jonathon Morgan. Defining and describing the population of ISIS supporters on Twitter. URL http://www.brookings.edu/research/papers/2015/03/isis-twitter-census-berger-morgan.

[9] J. M. Berger and Jonathon Morgan. Defining and describing the population of ISIS supporters on Twitter, 2015. URL http://www.brookings.edu/research/papers/2015/03/isis-twitter-census-berger-morgan. 1.1, 1, 1.2

[10] JM Berger and Heather Perez. Twitter Account Suspensions Help in Curbing ISIS Rhetoric Online | Office of Media Relations | The George

Washington University. URL https://mediarelations.gwu.edu/twitter-account-suspensions-help-curbing-isis-rhetoric-online.

[11] Berger, JM. Tailored Online Interventions: The Islamic States Recruitment Strategy. *Combating Terrorism Center Sentinel*. URL https://www.ctc.usma.edu/posts/tailored-online-interventions-the-islamic-states-recruitment-strategy.

[12] Norbert Binkiewicz, Joshua T. Vogelstein, and Karl Rohe. Covariate Assisted Spectral Clustering. *arXiv preprint arXiv:1411.2158*, 2014. URL http://arxiv.org/abs/1411.2158. 2.2.1

[13] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, October 2008. ISSN 1742-5468. doi: 10.1088/1742-5468/2008/10/P10008. URL http://arxiv.org/abs/0803.0476. arXiv: 0803.0476. 2.2, 4.1

[14] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006. 2.2

[15] Krishnadev Calamur. Twitter's New ISIS Policy. *The Atlantic*, February 2016. ISSN 1072-7825. URL http://www.theatlantic.com/international/archive/2016/02/twitter-isis/460269/.

[16] Kathleen M. Carley. A Dynamic Network Approach to the Assessment of Terrorist Groups and the Impact of Alternative Courses of Action. Technical report, October 2006. 2.1

[17] Kathleen M. Carley, Jeffrey Reminga, and Natasha Kamneva. Destabilizing terrorist networks. *Institute for Software Research*, page 45, 1998. URL http://repository.cmu.edu/cgi/viewcontent.cgi?article=1031&context=isr. 2.1

[18] Kathleen M. Carley, Matthew Dombroski, Maksim Tsvetovat, Jeffrey Reminga, Natasha Kamneva, and others. Destabilizing dynamic covert networks. In *Proceedings of the 8th international command and control research and technology symposium*, 2003. URL http://alliance.casos.cs.cmu.edu/publications/resources_others/a2c2_carley_2003_destabilizing.pdf. 1.2.1, 2.1

[19] Joseph A. Carter, Shiraz Maher, and Peter R. Neumann. #Greenbirds Measuring Importance and Influence in Syrian Foreign Fighter Networks. *International Centre for the Study of Radicalization Report*, April 2014. URL http://icsr.info/wp-content/uploads/2014/04/ICSR-Report-Greenbirds-Measuring-Importance-and-Infleunce-in-Syrian-For
pdf. 1.2.1

[20] Chao-Min Chiu, Meng-Hsiang Hsu, and Eric TG Wang. Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision support systems*, 42(3):1872–1888, 2006. URL http://www.sciencedirect.com/science/article/pii/S0167923606000583. 3.1.1

[21] Scott C. Deerwester, Susan T Dumais, Thomas K. Landauer, George W. Furnas, and Richard A. Harshman. Indexing by latent semantic analysis. *JAsIs*, 41(6):391–407, 1990.

[22] Inderjit S Dhillon. Co-clustering documents and words using bipartite spectral graph partitioning. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 269–274. ACM, 2001. 3.1.1

[23] Inderjit S. Dhillon. Co-clustering documents and words using bipartite spectral graph partitioning. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 269–274. ACM, 2001. URL http://dl.acm.org/citation.cfm?id=502550.

[24] Jana Diesner and Kathleen M. Carley. Using network text analysis to detect the organizational structure of covert networks. In *Proceedings of the North American Association for Computational Social and Organizational Science (NAACSOS) Conference*, 2004. URL http://alliance.casos.cs.cmu.edu/publications/papers/NAACSOS_2004_Diesner_Carley_Detect_Covert_Networks.pdf. 2.1

[25] Carsten F Dormann and Rouven Strauss. Detecting modules in quantitative bipartite networks: the quabimo algorithm. *arXiv preprint arXiv:1304.3218*, 2013.

[26] Kimberly Dozier. Anti-ISIS-Propaganda Czars Ninja War Plan: We Were Never Here., March 2016. URL http://www.thedailybeast.com/articles/2016/03/15/obama-s-new-anti-isis-czar-wants-to-use-algorithms-to-target-jihadis.html.

[27] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *arXiv preprint arXiv:1407.5225*, 2014. URL http://arxiv.org/abs/1407.5225.

[28] Michelle Forelle, Phil Howard, Andrs Monroy-Hernndez, and Saiph Savage. Political Bots and the Manipulation of Public Opinion in Venezuela. *arXiv:1507.07109 [physics]*, July 2015. URL http://arxiv.org/abs/1507.07109. arXiv: 1507.07109.

[29] Santo Fortunato. Community detection in graphs. *Physics reports*, 486(3):75–174, 2010.

[30] Eric Gilbert and Karrie Karahalios. Predicting Tie Strength with Social Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 211–220, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518736. URL http://doi.acm.org/10.1145/1518701.1518736. 3.1.1

[31] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002. 2.2

[32] Rick Gladstone. Behind a Veil of Anonymity, Online Vigilantes Battle the Islamic State. *The New York Times*, March 2015. ISSN 0362-4331. URL http://www.nytimes.com/2015/03/25/world/middleeast/behind-a-veil-of-anonymity-online-vigilantes-battle-the-islamic-state.

html.

[33] Leo A. Goodman. Snowball Sampling. *The Annals of Mathematical Statistics*, 32 (1):148–170, March 1961. ISSN 0003-4851, 2168-8990. doi: 10.1214/aoms/1177705148. URL http://projecteuclid.org/euclid.aoms/1177705148. 1.2

[34] Jane Harman. Disrupting the Intelligence Community. *Foreign Affairs*, (March/April 2015), April 2015. ISSN 0015-7120. URL http://www.foreignaffairs.com/articles/143042/jane-harman/disrupting-the-intelligence-community. 2.1

[35] Ming Ji, Jiawei Han, and Marina Danilevsky. Ranking-based classification of heterogeneous information networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1298–1306. ACM, 2011. URL http://dl.acm.org/citation.cfm?id=2020603.

[36] Stuart Koschade. A social network analysis of Jemaah Islamiyah: The applications to counterterrorism and intelligence. *Studies in Conflict & Terrorism*, 29(6): 559–575, 2006. URL http://www.tandfonline.com/doi/abs/10.1080/10576100600798418. 2.1

[37] Valdis Krebs. Uncloaking terrorist networks. *First Monday*, 7(4), 2002. URL http://journals.uic.edu/ojs/index.php/fm/article/view/941. 2.1

[38] Valdis E. Krebs. Mapping networks of terrorist cells. *Connections*, 24(3):43–52, 2002. URL http://www.aclu.org/files/fbimappingfoia/20111110/ACLURM002810.pdf. 2.1

[39] Nelly Lahoud, Daniel Milton, Bryan Price, and others. The Group That Calls Itself a State: Understanding the Evolution and Challenges of the Islamic State. Technical report, DTIC Document, 2014. URL http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA619696.

[40] Vito Latora and Massimo Marchiori. How the science of complex networks can help developing strategies against terrorism. *Chaos, solitons & fractals*, 20(1):69–75, 2004. 2.1

[41] Benjamin A Miller, Michelle S Beard, and Nadya T Bliss. Eigenspace analysis for threat detection in social networks. In *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pages 1–7. IEEE, 2011. 2.2, 2.2.2, 3.1.1

[42] Alan Mislove, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 29–42. ACM, 2007. URL http://dl.acm.org/citation.cfm?id=1298311. 3.1.1

[43] Batsheva Moriarty. Defeating ISIS on Twitter. *Technology Science*, September 2015. URL http://techscience.org/a/2015092904/.

[44] Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *science*, 328(5980):876–878, 2010.

[45] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in

networks. *Physical Review E*, 69(2):026113, February 2004. doi: 10.1103/PhysRevE.69.026113.

[46] Symeon Papadopoulos, Yiannis Kompatsiaris, Athena Vakali, and Ploutarchos Spyridonos. Community detection in Social Media. *Data Mining and Knowledge Discovery*, 24(3):515–554, June 2011. ISSN 1384-5810, 1573-756X. doi: 10.1007/s10618-011-0224-z. URL `http://link.springer.com/article/10.1007/s10618-011-0224-z`. 2.2.1

[47] Symeon Papadopoulos, Yiannis Kompatsiaris, Athena Vakali, and Ploutarchos Spyridonos. Community detection in social media. *Data Mining and Knowledge Discovery*, 24 (3):515–554, 2012.

[48] Steve Ressler. Social network analysis as an approach to combat terrorism: Past, present, and future research. *Homeland Security Affairs*, 2(2):1–10, 2006. 2.1

[49] V. S. Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, Filippo Menczer, Rand Waltzman, Andrew Stevens, Alexander Dekhtyar, Shuyang Gao, Tad Hogg, Farshad Kooti, Yan Liu, Onur Varol, Prashant Shiralkar, Vinod Vydiswaran, Qiaozhu Mei, and Tim Huang. The DARPA Twitter Bot Challenge. *arXiv:1601.05140 [physics]*, January 2016. URL `http://arxiv.org/abs/1601.05140`. arXiv: 1601.05140.

[50] Yizhou Sun and Jiawei Han. Mining heterogeneous information networks: principles and methodologies. *Synthesis Lectures on Data Mining and Knowledge Discovery*, 3(2): 1–159, 2012. URL `http://www.morganclaypool.com/doi/abs/10.2200/S00433ED1V01Y201207DMK005`.

[51] Yizhou Sun, Jie Tang, Jiawei Han, Manish Gupta, and Bo Zhao. Community evolution detection in dynamic heterogeneous information networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, pages 137–146. ACM, 2010.

[52] Yizhou Sun, Rick Barber, Manish Gupta, Charu C. Aggarwal, and Jiawei Han. Coauthor relationship prediction in heterogeneous bibliographic networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, pages 121–128. IEEE, 2011. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5992571`.

[53] Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S. Yu, and Tianyi Wu. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *VLDB11*, 2011. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.227.9062&rep=rep1&type=pdf`.

[54] Yizhou Sun, Charu C. Aggarwal, and Jiawei Han. Relation Strength-aware Clustering of Heterogeneous Information Networks with Incomplete Attributes. *Proc. VLDB Endow.*, 5(5):394–405, January 2012. ISSN 2150-8097. doi: 10.14778/2140436.2140437. URL `http://dx.doi.org/10.14778/2140436.2140437`.

[55] Yizhou Sun, Jiawei Han, Charu C. Aggarwal, and Nitesh V. Chawla. When will it happen?: relationship prediction in heterogeneous information networks. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 663–672.

ACM, 2012. URL http://dl.acm.org/citation.cfm?id=2124373.

[56] Lei Tang and Huan Liu. Community Detection and Mining in Social Media. *Synthesis Lectures on Data Mining and Knowledge Discovery*, 2(1):1–137, January 2010. ISSN 2151-0067. doi: 10.2200/S00298ED1V01Y201009DMK003. URL http://www.morganclaypool.com/doi/abs/10.2200/S00298ED1V01Y201009DMK003.

[57] Lei Tang and Huan Liu. Leveraging social media networks for classification. *Data Mining and Knowledge Discovery*, 23(3):447–478, 2011. URL http://link.springer.com/article/10.1007/s10618-010-0210-x. 2.2, 2.2.1, 3.1.1, 3.1.2, 4.1

[58] Lei Tang, Xufei Wang, and Huan Liu. Uncoverning groups via heterogeneous interaction analysis. In *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on*, pages 503–512. IEEE, 2009. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5360276. 2.2.2, 3.1.1, 3.1.2

[59] Noordin Mohammed Top. Counterterrorism?s new tool:?metanetwork?analysis. 2009. 2.1

[60] Ansar al-Sharia Tunisias and Long Game. Dawa, hisba, and jihad. 2013.

[61] 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. 1.1, 1.2

[62] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007. URL http://link.springer.com/article/10.1007/s11222-007-9033-z. 3.1.1

[63] Xufei Wang, Lei Tang, Huiji Gao, and Huan Liu. Discovering overlapping groups in social media. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 569–578. IEEE, 2010. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5694011. 2.2, 2.2.1