Comput Math Organ Theory (2018) 24:224–242
https://doi.org/10.1007/s10588-017-9255-3

MANUSCRIPT

Mining online communities to inform strategic messaging: practical methods to identify community-level insights

Matthew Benigni[1] · **Kenneth Joseph**[2,3] · **Kathleen M. Carley**[1]

Published online: 20 June 2017

Abstract The ability of OSNs to propagate civil unrest has been powerfully observed through the rise of the ISIS and the ongoing conflict in Crimea. As a result, the ability to understand and in some cases mitigate the effects of user communities promoting civil unrest online has become an important area of research. Although methods to detect large online extremist communities have emerged in literature, the ability to summarize community content in meaningful ways remains an open research question. We introduce novel applications of the following methods: ideological user clustering with bipartite spectral graph partitioning, narrative mining with hash tag co-occurrence graph clustering, and identifying radicalization with directed URL sharing networks. In each case we describe how the method can be applied to social media. We subsequently apply them to online Twitter communities interested in the Syrian revolution and ongoing Crimean conflict.

Keywords Social networks · Online extremist community · Online extremism · Social media · Twitter · Hashtags · Terrorism · ISIS · Euromaidan

✉ Matthew Benigni
mbenigni@cs.cmu.edu

[1] Center for Computational Analysis of Social and Organizational Systems (CASOS), Institute for Software Research, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

[2] Network Science Institute, Northeastern University, 177 Huntington Ave, Boston, MA 02115, USA

[3] Institute for Software Research, Carnegie Mellon University, Pittsburgh, PA, USA

1 Introduction

Social media's growing role in the shaping of public opinion has been observed in a variety of political (Loader and Mercea 2011) and geo-political (Herrick 2016; Juris 2012) settings. Although the true significance of social media's role in actual political change resulting from this rise in use remains in question (Howard and Parks 2012; Nanabhay and Farmanfarmaian 2011; Hussain and Howard 2013; Dewey et al. 2012), the emergence of social media as a means to at least motivate and expose desire for change has been recognized by scholars. The Arab Spring (Wei et al. 2015; Wolfsfeld et al. 2013; Starbird and Palen 2012; Tufekci 2014; Lotan et al. 2011) and the ongoing conflict in Crimea (Barbera 2014; Szostek 2014), have both highlighted the emergent role of social media, and online social networks (OSN) more specifically, as facilitators of social activism.

Initially many viewed social media's role in the Arab Spring and Euromaidan Movement as positive examples of free speech; however, there also exists a downside to the ability of OSNs to act as platforms of mobilization. More specifically, the rise of ISIS has been largely propagated and highly publicized through OSNs (Veilleux-Lepage 2014, 2015). Noting this, Western governments have begun attempts to mitigate the impacts of ISIS' propaganda approaches. However, they have found it challenging to participate in and influence online communities which show signs of extremism. In the United States, the recently-formed Global Engagement Center leads the State Department's effort 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" (Dozier 2016). Mr. Michael Lumpkin, the group's director, recently spoke to 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 by Renee Montagne, February 1, 2016

A logical follow-up question to Mr. Lumpkin's statement would be "Expose to whom, and how?" We propose that quantitative analysis of large online extremist communities (OECs) could offer insight into the populations most susceptible to radicalization and could be used to inform strategic messaging or assess ongoing diplomatic or military efforts. Although methods to detect large online extremist communities have emerged in literature (Benigni et al. 2017; Benigni and Carley 2016; Johnson et al. 2016), the ability to summarize community content in meaningful ways remains an open research question.

Online social networks now play in important role in helping people share information, particularly in times of unrest. As seen during the London riots of 2011 (Glasgow and Fink 2013) as well as post-earthquake information dissemination in Japan (Sakaki et al. 2010), online communities often organize around trending hash tags with short half lives. Political campaigns have shown similar patterns (Weber et al. 2013). Furthermore, organization around these hash tags often coalesces over time and is an important factor influencing information diffusion through social networks (Chang 2010). URL sharing and the ability to mention other users have become common attributes of many online social networks as well. In this paper we introduce three applications of existing methods to mine relevant content from large, online communities by taking advantage of tokens like hash tags, URLs, and mentions. We discuss the following methods:

- Ideological user clustering with bipartite spectral graph partitioning.
- Narrative mining with hash tag co-occurrence graph clustering.
- Identifying radicalization with directed URL sharing networks.

In each instance, we describe the data mining method in detail, present illustrative examples from online communities that exhibit varying levels of extremism, and subsequently discuss limitations and recommend future research. Our manuscript is organized as follows: we first describe the online communities we will use in Sect. 2, we then introduce the aforementioned methods in Sects. 3, 4, and 5. Finally we summarize our findings in Sect. 6.

2 Data

As case studies we will present two *online extremist communities* , which are composed of one or more *OEC members*. We define these terms 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 state stability or human rights.
OEC member a Twitter user who's timeline shows unambiguous support to the OEC of interest. For example, if the user positively affirms the OEC's leadership or ideology, glorifies its fighters, or advocates its talking points.

It is important to note that a member's *support* is relative and in many cases not in violation of local law or Twitter's terms of use. In fact, these "passive supporters" appear to be an essential to the diffusion of online propaganda and therefore represent an important element of radicalization efforts (Veilleux-Lepage 2015). In each presented case study we instantiate an n-hop snowball sampling strategy (Goodman 1961) with known members of a desired community. We then remove non-OEC members via supervised learning as presented in (Benigni et al. 2017; Benigni and Carley 2016). For this work we present two detected communities as described below.

2.1 Case study 1: the Euromaidan Twitter community

The Euromaidan movement started as a series of protests in November 2013, where large numbers began to call for the removal of then President Viktor Yanukovych. These protests reached their peak in February 2015, ultimately leading to the removal of many of Ynukovych's senior officials. These events were soon followed by the Russian occupation of Crimea, and despite the installation of a new government, a substantial online activist community continues to oppose Russian influence in the Ukraine. We will refer to this community as the Euromaidan Twitter community (ETC). Here, although strong negative sentiment toward the current Ukrainian government is observed, the online activism seen largely advocates change through legitimate government processes. Thus, while we acknowledge that little "extremism" exists in this community, we choose to examine this community due to its relevance to ongoing geopolitical events in the region. This community was extracted from a two step snowball sample of eight known Euromaidan movement members' mention ties from March 2014 to September 2015. The search resulted in 92,295 Twitter accounts, and subsequent OEC detection returned 1221 accounts actively supporting the movement. We have two collections from this community, one in March, 2016 and one in October, 2016.

2.2 Case study 2: the Syrian revolution Twitter community

The Syrian Revolution Twitter Community is an updated set of users based on the ISIS-supporting OEC presented by Benigni et al. (2017). By using mention activity of non-suspended, previously detected users and active learning, we update the SRTC based on the recent community activity. The instance presented in this work was collected in March of 2016 and contains 8718 members. We define a member 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 nearly all other anti-Assad factions are present.

Both the ETC and SRTC present a large community of Twitter users who's collective activity is of interest. In each case, group substructure, current interests, and information operations are all of interest, yet methods for mining such information remain immature.

3 Ideological user clustering with bipartite spectral graph partitioning

Algorithms designed to find clusters of highly connected users within real world social networks is often referred to as community detection and has been studied extensively (Fortunato 2010). In fact, Papadopoulos et. al. (2012) address the topic specifically with respect to social media. Though many relevant methods exist to identify highly connected sets of users, it has been shown that users' simultaneous membership in multiple social groups often makes community detection methods based exclusively on social ties imprecise (Tang et al. 2009; Benigni et al. 2017). We and others (DeMasi et al. 2016) have found that identifying user clusters based on shared beliefs can be a useful alternative to traditional methods when searching for ideologically homogeneous user groups. Within detected OECs, hash tag use often provides user clusters consistent with ideological substructure within the community. One could view our method as a means to efficiently find community structure based on hash-tag-inferred social ties. For example, within a large Sunni extremist community like the SRTC, one can identify distinct clusters of support for Jabhat Al-Nusra, the Free Syrian Army, and many other competing groups. Distinct regional interests are observed as well. We have found distinct news sharing communities focused on operations in Syria, Northern Iraq, Yemen and Palestine. To identify these user groups we apply bipartite spectral graph partitioning to co-cluster Twitter users and hash tags. Although we have applied these methods exclusively to investigate large online extremist communities, it is possible they could be used for other online settings such as targeted advertising.

Often in network science bipartite graphs can be transformed into a one mode projection by multiplying the bipartite adjacency matrix by its transpose (Zweig and Kaufmann 2011). However, in the case of online social networks and bipartite graphs of users and hash tags specifically, we have found these projections often become dense. In the case of users and hashtags, sets of approximately 100 thousand users often generate millions of unique hashtags. A one mode projection can result in a large, dense adjacency matrix which proves costly with respect to memory. For example we trimmed the user by hash tag bipartite graph extracted from the Euromaidan Twitter Community search and retained only hash tags used by five or more unique users. The resultant graph consisted of 92,295 users, 352,120 unique hash tags, and just over 16 million edges requiring 219 Mb of memory in sparse matrix format. The one mode projection requires over 60G of memory. We often find this type of increase in terms of edges with this type of matrix making one mode projections somewhat inconvenient. Furthermore a great number of these edges are quite close to zero and of little value to our task of clustering users. On option would be to simply retain the n largets edges, however “duality” between users and hash tags exists. Hash tags are used to expose content to specific groups, and groups generate their own hash tags. Dhillon et al. present bipartite spectral graph partitioning as a means to co-cluster words and documents (Dhillon 2001) and argue the method obtains more interpretable clusters than one mode projections because of the “duality of word and document clustering”. In other words, they assert that word clustering induces document clustering and document clustering induces word clustering. We assert the same duality holds true for users and hash tags. To co-cluster words and documents, the authors generate a word-document matrix, and use left and right singular vectors to project words and documents into the same euclidian space. They subsequently use k-means (MacQueen 1967) to find relevant clusters of documents and words. In our case, we claim duality in user and hash tag clustering. Such a method is notably similar to the most recent advancements in word embedding approaches, which focus on matrix decomposition of term matrices rather than focusing on developing neural models for embedding (Pennington et al. 2014; Hamilton et al. 2016). In our model, communities influence hash tag popularity, and hash tag popularity helps organize user communities.

3.1 Bipartite spectral multi-partitioning

To explain Dhillon’s Bipartite Spectral Multi-Partitioning algorithm we introduce the following notation. Lower case letters will represent column vectors. Capitol letters will denote adjacency matrices. We construct the graph $A_{m\times n}$ where an edge (or matrix cell) $e_{i,j}$ represents the number of times user i tweeted hash tag j where $i \in 1, 2, \ldots, m$ and $j \in 1, 2, \ldots, n$. We can represent this bipartite graph as a square undirected graph as follows:

$$M = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \quad (1)$$

The authors begin by explaining a spectral bipartition algorithm based on their proof illustrating the second eigenvector of the generalized eigenvalue problem $Lz = \lambda Dz$ provides a relaxation to the minimum normalized cut problem. Where L is the graph Laplacian defined as the nxn symmetric matrix, with one row and column for each vertex, such that:

$$L_{i,j} = \begin{cases} \sum_k E_{ik}, & \mathrm{i = j} \\ -E_{ij}, & \mathrm{i \neq j} \quad \text{and there is an edge}\{\mathrm{i,j}\} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Furthermore, $L = D - M$ where D is the diagonal "degree" matrix of adjacency matrix M. This allows us to express L as follows:

$$L = \begin{bmatrix} D_1 & A \\ A^T & D2 \end{bmatrix}. \quad (3)$$

We can then express the second eigen vector z_2 of L in terms of the second eigen vectors u_2 and v_2 of the left and right matrices of the singular vector decomposition of A as follows:

$$z_2 = \begin{bmatrix} D_1^{1/2} u_2 \\ D_2^{1/2} v_2 \end{bmatrix}. \quad (4)$$

One can then approximate the optimal bipartition by assigning the elements of z_2 to bimodal values m_j $(j = 1, 2)$ based on the following minimization:

$$\sum_{j=1}^{2} \sum_{z_2(i) \in m_j} (z_2(i) - m_j)^2 \quad (5)$$

which corresponds to the same objective function minimized by the k-means algorithm (Lloyd 1982). The authors then present the following bipartitioning algorithm.

3.1.1 Bipartite spectral bipartitioning

1. Given A form $A_n = D_1^{1/2} A D_2^{1/2}$.
2. Compute the second singular vectors of A_n, u_2 and v_2 and form the vector z_2 as in (4).

3. Run the k-means algorithm on the $z_2$to obtain the bipartitioning.

The authors then generalize this to the multipartitioning case by using the $l = \lceil \log(k) \rceil$ sigular vectors of of $A_n, u_2, ..., u_{l+1}$ and $v_2, ..., v_{l+1}$ to obtain a k-wise partition. To do so, they form the matrix:

$$Z = \begin{bmatrix} D_1^{1/2} U \\ D_2^{1/2} V \end{bmatrix} \tag{6}$$

where $U = [u_2, u_3, ..., u_l]$ and $V = [v_2, v3, ..., v_l]$. The k-wise partition can be minimized by the following equation:

$$\sum_{j=1}^{2} \sum_{z_2(i) \in m_j} ||Z(i) - m_j||^2 \tag{7}$$

Like Equation (5), and (7) can be minimized by classical k-means. The algorithm can be described as follows.

3.1.2 Bipartite spectral multi-partitioning

1. Given A form $A_n = D_1^{1/2} A D_2^{1/2}$.
2. Compute the $l = \lceil log(k) \rceil$ singular vectors of $A_n, u_2, ..., u_{l+1}$and $v_2, ..., v_{l+1}$ and concatenate them row wise to form Z.
3. Run the k-means algorithm on the $l - dimensional$data Z to obtain the desired k-way multipartitioning.

3.2 Ideological user clustering

We find that clustering the bipartite graph H of users and hash tags where an edge $e_{i,j}$ is defined as the number of times user i posts hashtag j within our corpus of tweets. To do so we cluster based on algorithm [REFERENCE].

To illustrate the utility of bipartite spectral partitioning we co-cluster users and hash tags within the SRTC. Due to the prevalence social bots within this community Benigni et al. (2017); Berger and Morgan (2015), we set a threshold with respect to the minimum number of unique users who posted a specific hash tag. This enables us to remove hash tags with high frequency that are not necessarily indicative of a sub-group's interests. In this case we set the minimum unique user threshold $\gamma_u = 5$ roughly reducing the number of unique hash tags observed within the community by 80%. We then co-cluster the 8718 users and 39,137 hashtags using k-means. In this case we choose $k = 50$ sub-groups, but acknowledge that selecting k requires some trial and error by the user. If k is too small, the algorithm returns few large groups; however, when k becomes 'large enough' more interesting sub-groups of users can be found. Figure 1 depicts the size of user groups, sorted from largest to smallest, from the SRTC data with $k = 25$ and $k = 50$ in the left and right panels respectively. The plot highlights the method's propensity to first partition small groups of hash tags and users and highlights how nearly 75% of the users in our dataset are clustered into one of two groups when $k = 25$. Only when k is sufficiently large do we observe interesting clusters. In the case of the SRTC we start to identify clusters talking about distinct conflict zones in the middle east. For example we find distinct clusters discussing ongoing conflict in Syria, Iraq, Egypt and Yemen.

The hash tags co-clustered with the two largest user groups depicted in the right panel of Fig. 1 are summarized in Fig. 2. Co-cluster A (left panel) consists of 2173 users and 3573 hash tags, while co-cluster B (right panel) consists of 1640 users and 3910 hash tags. We calculate hash tag uniqueness within the cluster by comparing the relative frequency of each hashtag within the group to the hash tag's frequency within the entire community. Hash tags are colored by relative frequency in Fig. 2 where darker font indicates higher within group frequency. For example *Army of Conquest* (translated from: جيش الفتح) has a much higher relative frequency within co-cluster B. The term refers to Jaish al-Fatah, an alliance of Islamist rebel faction active in Idlib and supported by Saudi Arabia and Turkey. The relative frequency and count of these terms within the sub-group indicate a common interest among users with respect to ongoing operations in Idlib and Jaish al-Fateh's role in them. Co-cluster A appears focused on denouncing ISIS which is highlighted by the hash tags *werejectisis* and *Abu Kamal Under Fire*. Abu Kamal is a Syrian town once held by ISIS that was highly targeted by US Coalition air strikes in February of 2016 Wood (2016). Identification of such sub-structure can provide novel insight to inform strategic messaging or operational assessments. For example, manual inspection of users' Twitter timelines in co-cluster A highlighted an effective

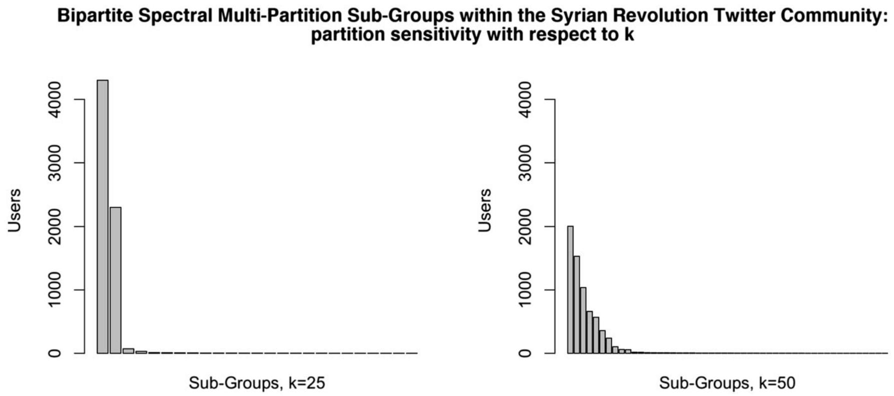

Fig. 1 Depicts the size of user groups, sorted from largest to smallest, from the SRTC data with $k = 25$ and $k = 50$ in the *left* and *right panels* respectively. The *left* highlights the algorithm's tendency to partition small sub-groups leaving one or two relatively large groups when k is not sufficiently large. The *right panel* corresponds to a selection of $k = 50$ which in this case provides interesting sub-structure

Fig. 2 Depicts the 30 most frequent hash tags co-clustered with two sub-groups within the SRTC. In each panel the hash tags are translated using Google Translate. *Color* depicts the relative frequency of the hash tag within the subgroup when compared to the rest of the community. *Darker font* connotes higher relative frequency within the subgroup when compared to the entire community

messaging theme employed by Jabhat al-Nusra that highlights ISIS' killing of Muslims. Mining these communities for effective propaganda themes could be used to inform strategic messaging, and be used to identify users with high social influence within specific topics of discussion.

We highlight the distinct difference between bipartite spectral partitions and standard user grouping algorithms like Louvain grouping Blondel et al. (2008). We do not claim that bipartite spectral clustering produces better groups as ground truth in real world networks can rarely be inferred Peel et al. (2016). However, Fig. 3 highlights the difference between clusters generated with bipartite spectral partitioning and Louvain grouping using the SRTC dataset. The black nodes on the right hand side of the plot depict clusters derived by bipartite spectral partitioning, while the white nodes depicted on the right side of the plot depict clusters derived by Louvain grouping. Node size depicts cluster size. Edges depict the number of users shared between the two cluster types. The figure highlights the differences between the two clustering methods. In practice we observe clusters that appear more homogeneous with respect to user content using bipartite spectral partitioning, but we acknowledge these observations are largely qualitative. Using both methods as an ensemble to find users who mention one another and user similar hash tags could prove interesting as well, though we will leave such questions for future research.

Although bipartite spectral multi-partitioning of users and hash tags offers unique insight into shared user activity, the method is not without limitations. The relative size of groups is highly sensitive to the researcher's selection of *k*, and an exploration of alternative clustering techniques is worthy of research Steinbach et al. (2000). Specifically, k-means implementations which incorporate a priori knowledge could be useful to cluster more complex graph representations of large social networks Wagstaff et al. (2001). It is also likely that additional information in

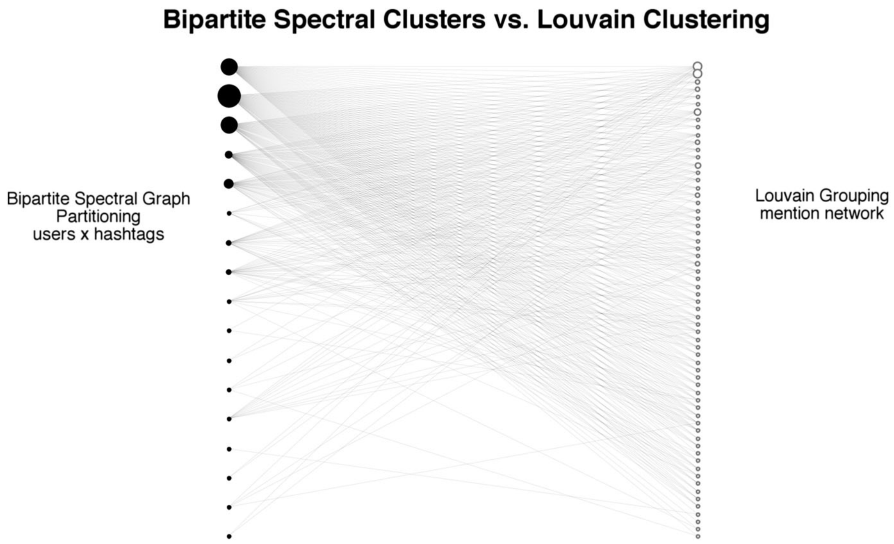

Fig. 3 Depicts the relationship between clusters generated with bipartite spectral partitioning and Louvain grouping using the SRTC dataset. The *black nodes* on the *left hand side* of the plot depict clusters derived by bipartite spectral partitioning, while the *white nodes* depicted on the *right side* of the plot depict clusters derived by Louvain grouping. Node size depicts cluster size. Edges depict the number of users shared between the two cluster types

user timelines could be used to cluster users. Keywords within tweets or URLs for example could be incorporated for more informed clusters. Finally, large sets of hash tags are difficult to interpret. Word clouds like the ones depicted in Fig. 2 do not necessarily imply sentiment. However, sentiment mining techniques could be applied to provide greater understanding, and as those methods become more mature with respect to non-English text we foresee them being highly useful.

4 Narrative mining with hash tag co-occurrence graph clustering

In this section we would like to extract trending narratives from online communities in order to gain understanding of interests and topical connections. Again, we find this particularly informative in large extremist communities, but acknowledge it could be useful in other large online communities as well. Currently, many tools summarize social media content by using naive methods like frequency. However, in online communities participating in political activism frequency alone often leads to predictable results. Table 1 highlights the redundancy in trending hashtags over different time periods within the Euromaidan Twitter Community. Of the top 10 translated hashtags with respect to frequency, only 10% occur uniquely and are highlighted in bold font. We define a narrative as a subset of online discussion organized around an identifiable event or set of events within an online community. We use hash tag co-occurrence in tweets to identify clusters of terms which are often quite interpretable to an end user. To do so, we construct a temporally-constrained hash tag co-occurrence graph and use community detection to extract community narratives.

We are interested in characterizing community narratives within an arbitrarily selected time period T, and thus start by identifying the set of hash tags which appear more frequently within T. Twitter limits collection of a user's timeline to their last 3200 tweets REST Twitter (2016), therefore the number of active users on any given day can vary significantly. Many users' accounts go dormant as well. This forces us to normalize hash tag rates based on active users within our dataset. To do so, we construct a vector $\mathbf{u}_a$ of length T where element i of $\mathbf{u}_a$ is the number of active users within our dataset at time interval i. We define an active user collected tweets span time i. Figure 4 depicts $\mathbf{u}_a$ for the ETC. We then normalize hashtag rates for a given time interval $T = [t_1, t_2]$ as follows

$$d_T = \sum_{i=t_1}^{t_2} \mathbf{u}_a$$

where d_T is the number of active user days associated with time interval T. We define d_g as the total active user days within our sample where

$$d_g = \sum \mathbf{u}_a \, .$$

We then construct $\mathbf{h}_T$ and $\mathbf{h}_G$. Both are vectors of length n, where n is the number of unique hash tags collected within the community of users U. Each entry in $\mathbf{h}_T$ and $\mathbf{h}_G$ represents the normalized counts of each hashtag j divided by the d_T and d_G respectively. λ_T, a vector of length n, represents the change in rate of hash tag j when compared to the global rate and is defined as follows:

$$\lambda_T = \mathbf{h}_T \odot \mathbf{h}_G^{-1}$$

We then define our set of trending hash tags h as those having $\lambda_T > \phi$. In our case we set $\phi = 2$ or hash tags who's rate was twice as high as their global rate. It is worth mentioning that this parameter needs to be selected with careful consideration

Table 1 Depicts the top 10 translated hash tags with respect to frequency in the Euromaidan Twitter community from 10–15 January, 10–15 March, and 10–15 May 2017

10–15 January	10–15 March	10–15 May
Ukraine	Freesavchenko	Eurovision
Russia	Ukraine	Ukraine
Little	Russia	**Rukiprochotmirotvortsa**
Freesavchenko	Savchenko	**Jamala**
Russia	Savchenko	Feesavchenko
Saveuatwi	Donetsk	Russia
Ukraine	Little	News
Donetsk	Russia	Donetsk
Ato	News	Eurovision
Ukraine	**Syria**	**Crimea**

Hashtags that occur only once are in bold. The table highlights the limitations of naive methods like frequencies to summarize community discussion

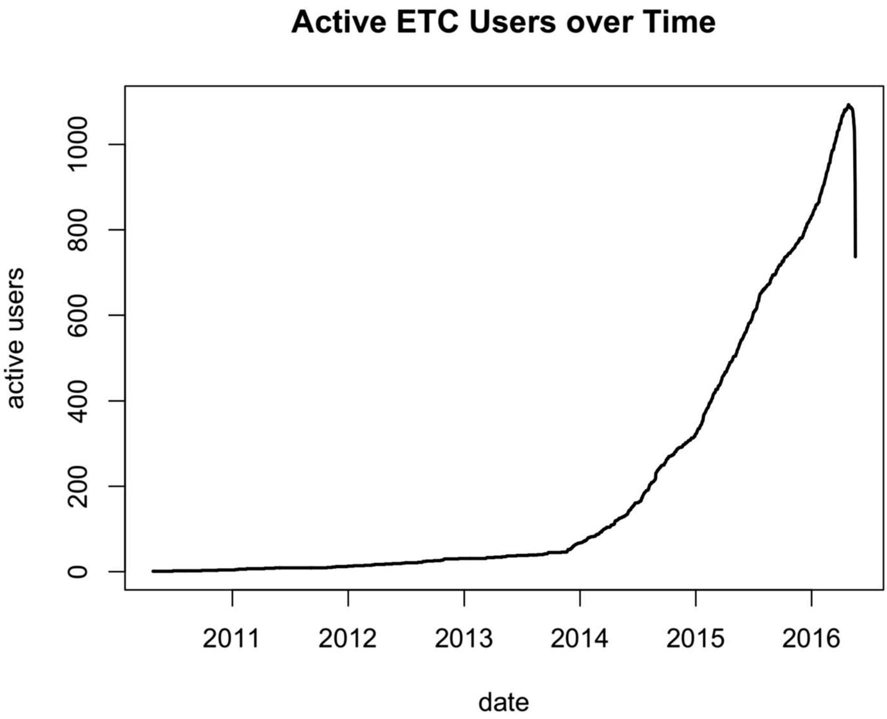

Fig. 4 The amount of time spanned by a given user's last 3200 tweets varies greatly resulting in a non-uniform number of active user's tweets captured within our dataset. To evaluate trends we need to normalize by active users per day. The *figure* above depicts active users per day in the SRTC dataset

as T gets large. Furthermore, we only select hashtags with γ_u or more unique users posting them to ensure we are capturing community narratives and account for bots as discussed in Sect. 3.

We then construct the network H_T where an edge is defined as the number of times hashtag i and hashtag j co-occur within user tweets posted within time interval T. Finally we cluster H_T using the Louvain Grouping algorithm and extract the resultant clusters to identify narratives Blondel et al. (2008). It is worth mentioning that any graph clustering approach suitable to large, weighted, undirected graphs could be used for this step. For an extensive discussion of suitable alternatives we refer researchers to Fortunato et al. Fortunato (2010).

Figures 5 and 6 depict H_T for the Euromaidan Twitter Community from 10-15 May, 2016 and 9-16 October, 2016 respectively. In each plot the left panel depicts H_T where nodes are sized by $\mathbf{h}_T$ and colored based on their membership to large Louvain clusters. The right panel summarizes the top 25 terms with respect to λ_T within a specific cluster or narratives. Each word is sized by $\mathbf{h}_T$, and colored by λ_T where gray font indicates values close to one and increased rates are colored. The four narratives depicted in the right panel of Fig. 5 center around major news story lines in Eastern Europe in early May, 2016. Narrative one, in red, centers around actions taken by the pro-Ukrainian hacker groups Falcons Flame and Trinity, who

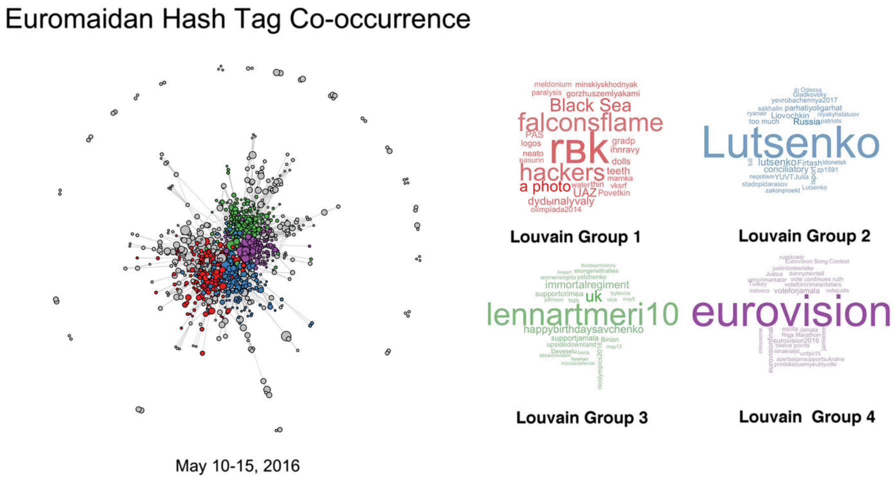

Fig. 5 Depicts community narrative extraction within the Euromaidan Twitter Community from 10–15 May, 2016

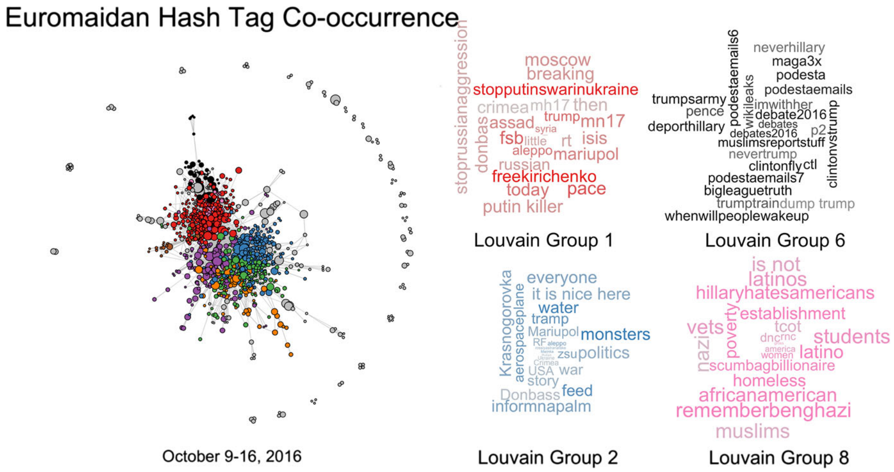

Fig. 6 Depicts community narrative extraction within the Euromaidan Twitter Community from 9–15 October, 2016

defaced the official websites of 9, Russian-backed militant groups involved in the Crimean Conflict Shamanska (2016). Narrative two, in blue, centers around the Ukrainian government's the appointment of Yuriy Lutsenko as Ukraine's attorney general UNIAN (2016). To appoint Lutsenko the government passed a special amendment removing the requirement for the country's attorney general possess a law degree, and was perceived as corruption within the ETC. Narratives two and three discuss the annual Lennart Meri conference in Tallinn Jensen (2016), and Susana Alimivna Jamaladinova's Eurovision 2016 victory Roxburgh (2016) respectively. We have shared similar figures with members of the United States

Army Asymmetric Warfare Group with extensive operational advisory experience in Ukraine which they found helpful in interpreting community interests over time. If we think of the identified narratives as components within H_T, the off diagonal densities between narratives could prove interesting as well. For example the connection between narratives associated with the 2016 United States Presidential Election and Russian aggression can clearly be seen in Fig. 6. Louvain Groups six and eight are densely connected to ongoing narrative of Russian actions in Syria and Ukraine highlighted in Louvain groups 1 and 2.

We assert that clustering and visualizing hash tag co-occurrence matrices offer researchers and analysts a quick means to distill community-level discussion in online social networks. The ability to mine quickly evolving hash tags within these communities offers insight into the complex interests driving activist discussion. Although these methods are limited to fixed time periods of interest, it is likely that they could be analyzed as dynamic networks and offer insights into changing interests over time Carley (2006). Additionally, these methods could be extended to account for URLs and hash tags to quickly find the external sources most influential within specific topic areas. Similar to the points made in Sect. 3, incorporation of sentiment analysis could provide additional value as well.

5 Identifying radicalization with directed URL sharing networks

The methods introduced in Sects. 3 and 4 illustrate how large online communities can be mined for high level understanding of community interests. In this section we provide an example where we can mine communities for social media intelligence (SOCMINT). Detecting extremist communities at scale enables information extractions that highlight tactics and techniques used for the radicalization process. For example ISIS uses Twitter for broad, general recruiting and typically transitions to more secure messaging platforms as they identify individuals worth personally targeting for radicalization (Berger and Morgan 2015; Berger 2015; Callimachi 2015). This "direct messaging" is typically done on Twitter through the @*mention*, where $user_a$ can ensure his content is included on $user_b$'s timeline by including @$user_b$ in the body of his or her tweet.

These hypothesized recruitment behavior patterns can be extracted and investigated by constructing graphs based on specific types of tweets. The Twitter REST API REST Twitter (2016) provides structure to easily extract tweets containing URLs and @*mentions*, and with this subset of content we form the following graphs:

P, a weighted, directed graph, where we define an edge $e_{i,j}$ as the number of tweets containing a URL where $user_i$ mentions $user_j$.

U_p, a weighted, bipartite graph, where we define an edge $e_{i,j}$ as the number of tweets posted by $user_i$ containing url_j. We hypothesis that propagandists and recruiters would be a subset of the users within U_p.

R_u, a weighted, bipartite graph, where we define an edge $e_{i,j}$ as the number of tweets posted containing url_i and mentioning $user_j$. We hypothesize that recruits would be a subset of the users within R_u.

A simple analysis of node centrality of users within P provides inference into possible roles. Users who send many messages with *@mentions* and urls are potential propagandists or recruiters and would have relatively high out-degree Wasserman and Faust (1994) within P. Naturally, users with high in-degree would be potential recruiting targets. Identifying both role types offers insight into more nuances analysis to recruiting techniques, materials, and could potentially help identify the user behaviors used to identify potential recruits. Both U_P and R_u can be used to further infer user roles. For example, if the URLs shared by a user in U_p often contain links to peer-to-peer messaging services like Telegram or WhatsApp, they would possibly be recruiters. If they share inflammatory news sources and videos they could be propagandists. Furthermore, the types of links received by users in R_U could offer similar insight to identify potential recruits.

Figure 7 depicts the the relationship between recruiters and their recruiting targets as well as the sites they reference in the radicalization process. To develop the plot we extract the 184,225 tweets where SRTC members use the *@mention* to share URL content between May, 2015 and May, 2016. The left panel depicts the directed mention network, where nodes are SRTC members and edges depict the number of tweets posted by user a that contain both an URL and mention of user b. Both the color and size of nodes denote the number of messages sent. Thus, small blue nodes would be likely recruiting targets, and large red nodes would likely be recruiters or propagandists. The center panel summarizes the top 35 sites shared within messages of this type, as does the bar plot in the right panel. For privacy reasons we have chosen not to publish identified recruiters in this forum, but we find dyads within P with relatively small follower counts and edge weights between 20 and 100 to often identify user to user exchanges consistent with targeted recruiting.

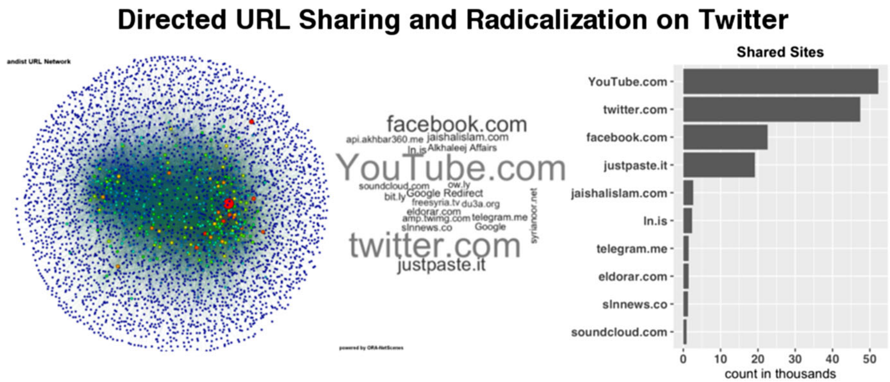

Fig. 7 The *left panel* depicts directed URL sharing network P within the SRTC. Nodes are both *colored* and sized by out degree. The *middle panel* depicts the top 25 shortened URLs posted in U_P with respect to frequency as does the *right panel*

These online dialogues often facilitate transition to more secure platforms over time as well. Conveniently, as these recruiters move discussion to peer to peer messaging platforms, they typically provide their user identification number. For example, this analysis yielded over 200 unique telegram.me accounts of likely recruiters.

The methods described above to identify recruitment and propaganda dissemination within online communities unfortunately still require a significant amount of manual exploration in order to confidently extract users of specific role types. Moreover, the example provided above merely highlights the ability to develop useful network representations based on hypothesized behaviors. Although these techniques still require a significant level of manual inspection, it is likely that more detailed NLP analysis of users' tweets could further inform automated detection methods, and active learning would provide the technical framework needed to gain adequate performance in an efficient manner Settles (2009). Furthermore, we hypothesize that similar strategies could be used to identify other user types with identifiable communication patterns.

6 Conclusion

In this paper we have provided researchers and practitioners three novel applications of existing methods in network science in order to facilitate improved data analysis of large communities in online social networks. We introduced:

- Ideologically clustering users by hashtag behavior with bipartite spectral graph partitioning.
- Narrative extraction through hash tag co-occurence graph clustering.
- User role inference from directed URL sharing networks.

In each instance, we presented the method in detail and illustrated meaningful insights from two activist online communities. We also offered useful extensions to these methods for future work. As stated in Sect. 3, we cannot claim the performance of bipartite spectral partitioning is better than that of other clustering algorithms. However, we do show that clustering users by hash tag use provides a different perspective. We see great potential in developing ensemble methods to cluster users across multiple dimensions such as following ties, mention ties, and shared hash tag use. Hash tag co-clustering also has limitations. Selection of time interval greatly influences the relative frequency of hash tags over time, and understanding the correct T to extract narratives needs further research. It is likely that some temporal decay function on link weights could be used or other methods from dynamic network analysis. Moreover, we highlight that the methods presented in this work almost exclusively mine patterns based on graph structure. This affords these methods to perform independent of the language use of community members, and it is likely that they would complement NLP-based methods to mine intelligence from OECs. Our hope is that this work enables both researchers and practitioners to draw novel insights from large online communities.

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 conclu- sions 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

Barbera P (2014) Tweeting the revolution: social media use and the #euromaidan protests | huffington post

Benigni M, Carley KM (2016) From tweets to intelligence: understanding the Islamic Jihad supporting community on Twitter. In social computing, behavioral-cultural modeling, and prediction, page to appear. Springer, New York

Benigni M, Joseph K, Carley KM (2017) Online extremism and the communities that sustain it: Detecting the isis supporting community on twitter. Under Revision to PLOS ONE

Berger JM, Morgan J (2015) The isis twitter census: defining and describing the population of isis supporters on twitter. The Brookings Project on US Relations with the Islamic World 3(20)

Berger JM (2015) Tailored online interventions: the Islamic state's recruitment strategy. Combating Terrorism Center Sentinel

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech 2008(10):P10008

Callimachi R (2015) ISIS and the Lonely Young American. The New York Times

Carley KM (2006) A dynamic network approach to the assessment of terrorist groups and the impact of alternative courses of action. Technical report

Chang HC (2010) A new perspective on twitter hashtag use: Diffusion of innovation theory. Proc Am Soc Inf Sci Technol 47(1):1–4

DeMasi O, Mason D, Ma J (2016) Understanding communities via hashtag engagement: a clustering based approach. In: Tenth International AAAI Conference on Web and Social Media

Dewey T, Kaden J, Marks M, Matsushima S, Zhu B (2012) The impact of social media on social unrest in the Arab Spring. Int Policy Prog

Dhillon IS (2001) 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, ACM, pp 269–274

Dozier K (2016) Anti-ISIS-Propaganda Czars Ninja war plan: we were never here, March 2016

Fortunato S (2010) Community detection in graphs. Physics Rep 486(3):75–174

Glasgow K, Fink C (2013) Hashtag lifespan and social networks during the london riots. In: Greenberg AM, Kennedy WG, Bos ND (eds) Social computing, behavioral-cultural modeling and prediction. Springer, Berlin, pp 311–320. doi:10.1007/978-3-642-37210-0_34

Goodman LA (1961) Snowball sampling. Ann Math Stat 32(1):148–170

Hamilton WL, Clark K, Leskovec J, Jurafsky D (2016) Inducing domain-specific sentiment lexicons from unlabeled corpora

Herrick D (2016) The social side of cyber power? social media and cyber operations. In: Cyber Conflict (CyCon), 2016 8th International Conference on, NATO CCD COE, pp 99–111

Howard PN, Parks MR (2012) Social media and political change: capacity, constraint, and consequence. J Commun 62(2):359–362

Hussain MM, Howard PN (2013) What best explains successful protest cascades? ICTs and the fuzzy causes of the Arab Spring. Int Stud Rev 15(1):48–66

Jensen DN (2016) Lennart meri and the 'new normal' | huffington post

Johnson NF, Zheng M, Vorobyeva Y, Gabriel A, Qi H, Velasquez N, Manrique P, Johnson D, Restrepo E, Song C, Wuchty S (2016) online ecology of adversarial aggregates: ISIS and beyond. Science 352(6292):1459–1463

Juris JS (2012) Reflections on #occupy everywhere: social media, public space, and emerging logics of aggregation. Am Ethnol 39(2):259–279

Lloyd Stuart (1982) Least squares quantization in pcm. IEEE Trans Inf Theory 28(2):129–137

Loader BD, Mercea D (2011) Networking democracy? Inf Commun Soc 14:757–769

Lotan G, Graeff E, Ananny M, Gaffney D, Pearce I, Boyd D (2011) The revolutions were tweeted: information flows during the 2011 Tunisian and Egyptian revolutions. Int J Commun 5:1375–1405
MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol 1, University of California Press, pp 281–297
Nanabhay M, Farmanfarmaian R (2011) From spectacle to spectacular: how physical space, social media and mainstream broadcast amplified the public sphere in Egypt's Revolution'. J N Afr Stud 16(4):573–603
Papadopoulos S, Kompatsiaris Y, Vakali A, Spyridonos P (2012) Community detection in social media. Data Min Knowl Discov 24(3):515–554
Peel L, Larremore DB, Clauset A (2016) The ground truth about metadata and community detection in networks
Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. Proc Emprical Methods Nat Lang Process 12:1532–1543
REST Twitter (2016) API
Roxburgh G (2016) Ukraine wins the 2016 eurovision song contest | news | eurovision song contest
Sakaki T, Okazaki M, Matsuo Y (2016) Earthquake shakes twitter users: real-time event detection by social sensors. In: Proceedings of the 19th International Conference on World Wide Web WWW '10, ACM, pp 851–860
Settles B (2009) Active learning literature survey. Computer Sciences Technical Report 1648. University of Wisconsin, Madison
Shamanska A (2016) Hackers in Ukraine deface separatist websites to mark victory day. Accessed 23 Oct 2016
Starbird K, Palen L (2012) (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. In: Proceedings of the acm 2012 conference on computer supported cooperative work, ACM, 2012, pp 7–16
Steinbach M, Karypis G, Kumar V, et al. (2000) A comparison of document clustering techniques. In: KDD workshop on text mining, vol 400, Boston, pp 525–526
Szostek J (2014) The media battles of Ukraine's EuroMaidan. Digital Icons 11:1–19
Tang L, Wang X, Liu H (2009) Uncoverning groups via heterogeneous interaction analysis. In: Ninth IEEE International Conference on Data Mining, ICDM'09, IEEE, pp 503–512
Tufekci Z (2014) Big questions for social media big data: representativeness, validity and other methodological pitfalls. In: ICWSM '14: Proceedings of the 8th international AAAI conference on weblogs and social media
UNIAN (2016) Rada adopts bill allowing lutsenko to be nominated prosecutor general. Accessed 20 Oct 2016
Veilleux-Lepage Y (2015) Paradigmatic Shifts in Jihadism in Cyberspace: the emerging role of unaffiliated sympathizers in the Islamic State's social media strategy. J Terror Res
Veilleux-Lepage Y (2014) Retweeting the Caliphate: the role of soft-sympathizers in the Islamic state's social media strategy. In: 2014 6th International Terrorism and Transnational Crime Conference, March 2014
Wagstaff K, Cardie C, Rogers S , Schrdl S, et al. (2001)Constrained k-means clustering with background knowledge. In: ICML, vol 1, pp 577–584
Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge
Weber I, Garimella VRK, Teka A (2013) Political hashtag trends. In: Serdyukov P, Kuznetsov OS, Kamps J, Agichtein E, Segalovich I, Yilmaz E (eds) Advances in information retrieval. Springer, Berlin, pp 857–860. doi:10.1007/978-3-642-36973-5_102
Wei W, Joseph K, Liu H, Carley KM (2015) The fragility of twitter social networks against suspended users. In: Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 9–16
Wolfsfeld G, Segev E, Sheafer T (2013) Social media and the arab spring politics comes first. Int J Press 18(2):115–137
Wood P (2016) IS conflict: counting the civilian cost of US-led air strikes
Zweig KA, Kaufmann M (2011) A systematic approach to the one-mode projection of bipartite graphs. Soc Netw Anal Min 1(3):187–218

Matthew Benigni Lieutenant Colonel Matthew Benigni was commissioned is an active officer in the United States Army and serves as an Operations Research Systems Analyst. His army assignments have included commanding a company of tanks and infantry in Eastern Baghdad from 2004-2005, serving as an assistant professor in the Department of Mathematical Sciences at the United States Military Academy, and providing decision support at the United States Special Operations Command, MacDill Air Force Base, FL, USA. He has a B.S. in operations research from the United States Military Academy at West Point, NY, a M.S. in applied statistics from the Colorado School of Mines, and a Ph.D in Societal Computing from the Carnegie Mellon School of Computer Science. His research interests center around the relationship between online communities and extremism.

Kenneth Joseph is currently a postdoc at the Network Science Institute at Northeastern University and a fellow at Harvard's Institute for Quantitative Social Science. He completed his graduate work in the Societal Computing program in the School of Computer Science at Carnegie Mellon University. His research focuses on obtaining a better understanding of the dynamics and cognitive representations of stereotypes and prejudice, and their interrelationships with sociocultural structure. In his work, he leverages a variety of machine learning methods, agent-based modeling strategies and socio-cognitive theories. Kenneth's work has been published in a variety of outlets, including KDD, ICWSM, WWW, CSCW and the Journal of Mathematical Sociology, and has been covered by popular outlets including 538 and the Chicago Tribune.

Kathleen M. Carley is the director of the center for Computational Analysis of Social and Organizational Systems (CASOS) which has over 25 members, including students, post doctoral fellows, research staff, and faculty. She is the founding co-editor of the journal Computational and Mathematical Organization Theory which she now co-edits with Dr. Terrill Frantz. She has co-edited several books in the computational organizations and dynamic network area. Her research combines cognitive science, social networks and computer science to address complex social and organizational problems. Her specific research areas are dynamic network analysis, computational social and organization theory, adaptation and evolution, text mining, and the impact of telecommunication technologies and policy on communication, information diffusion, disease contagion and response within and among groups particularly in disaster or crisis situations. She and her lab have developed infrastructure tools for analyzing large scale dynamic networks and various multi-agent simulation systems.