Online Extremist Community Detection:

Detecting, Analyzing, and Disrupting Extremist Communities in Social Media

Matthew Benigni
mbenigni@cs.cmu.edu

Kenneth Joseph
kjoseph@cs.cmu.edu

Prof. Kathleen M. Carley
kathleen.carley@cs.cmu.edu

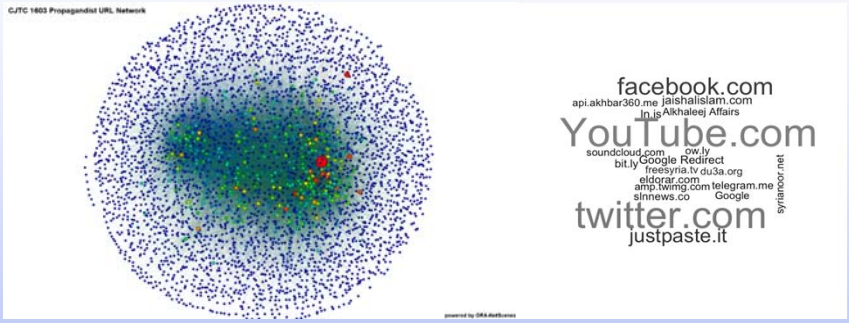

Figure 1 depicts url sharing within the CASOS Sunni Jihadist Twitter Community (CJTC). The left panel depicts the directed mention network where nodes are users and edges represent tweets containing urls. Nodes are colored by out degree, and we hypothesize that small blue nodes represent recruiting targets. The right panel illustrates the web services typically linked in these interactions. The shortened urls are sized by frequency..

Introductions Current detection methods are unable to accurately detect online extremist communities (OEC) at scale without a great deal of manual intervention. We present *Iterative Vertex Clustering and Classification*, a novel approach to identify OECs, and provide an illustrative case study of the extremist Syrian revolution supporting community on Twitter. By searching known members following ties and partitioning the resultant network, we identify over 10,000 user accounts actively supporting one or more of the declared terrorist organizations fighting against the Assad regime in Syria with 96% accuracy. Our approach utilizes unsupervised techniques in conjunction with vertex classification techniques to partitions vertices in large social network and identify a targeted covert community and overcomes the precision and recall shortcomings of existing methods. By using network structure and user account features to inform classification algorithms, we remove user accounts unlikely to be community members. The detected community offers unique insights into the online populations most susceptible to radicalization.

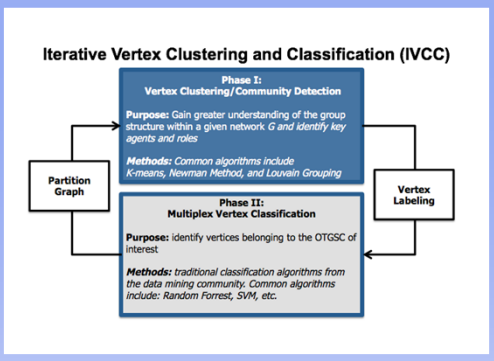

Figure 2 Depicts the iterative work flow for extremist community detection

Case Study. Our classification methods use a combination of the features and methods developed in Tang et al. [2009] and Tang and Liu [2011]. *Meta-Network Vertex Classification* (MVC), utilizes user account features as well as spectral representations of the many link types available in social media, similar to the methods presented in Tang et al. [2009], but we also include user account features similar to the methods presented in Tang and Liu [2011]. Furthermore, we capitalize on node metrics captured by each respective network, as these metrics have been proven measures of influence in social network analysis. The result is a superior classifier as depicted in Table 1.

Additional Case Studies.

- **CASOS Jihadist Twitter Community (CJTC)** - ~10K Twitter users who support one or more of the Sunni extremist groups fighting in the Syrian Revolution
- **Crimean Conflict Community (CCC)** – a group of ~2K Twitter users actively opposing Russian intervention in Crimea

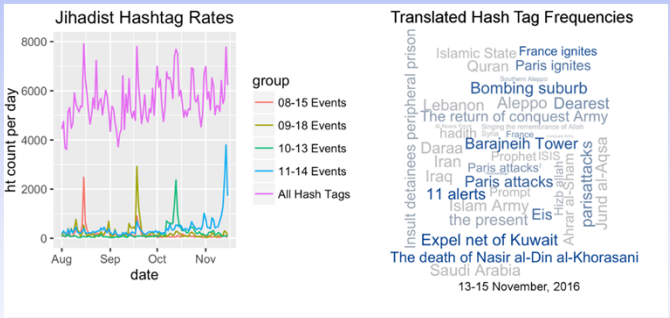

Figure 5 The left panel depicts the volume of hash tags used within the CJTC from AUG- NOV 2015. The right panel highlights the hash tags most explanatory of the increased activity on November 14, 2015.

Key Findings.

- Distinct spectral representations of following, mention, and hashtag ties improve detection.
- Twitter suspension policies have significantly impacted the size and activity of the ISIS supporting community; however, other Sunni extremist groups now have a significant presence within Twitter exhibiting behaviors similar to ISIS supporters.
- Unsupervised techniques can be used to maintain training sets and monitor highly dynamic communities over time
- Community detection enables previously unattainable intelligence extractions

Future Work.

- Unsupervised detection of dense activist communities
- An active learning framework for OEC detection
- Topology-based disruption of extremist online communities

This work was supported in part by the Office of Naval Research (ONR) through N000141512564 , MINERVA N000141310835 and by Center for Computational Analysis of Social and Organization Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research, or the U.S. government.