Identifying Covert or Hidden Groups in Twitter

Lieutenant Colonel Matthew Benigni (matthew.c.benigni@mail.mil)
& Kathleen M. Carley (kathleen.carley@cs.cmu.edu)
Carnegie Mellon University

Social media has become an important lens through which to understand ongoing wars of ideas in areas like Syria and Crimea. Large communities now organize in Online Social Networks (OSNs) and possibly play a significant role in geopolitical conflict [1]. Furthermore, initial governmental efforts to influence this important media have had modest results [2]. The ability to identify and understand the online communities we hope to influence is critical to developing more effective interventions. The purpose of this workshop is to present and discuss the value of techniques for identifying and analyzing covert or hidden groups in Twitter.

Two core methodologies will be demonstrated: 1) Iterative Vertex Clustering and Classification (IVCC) [3], and 2) social network analysis of Twitter data using ORA [4]. IVCC is an iterative, semi-supervised methodology to detect large, ideologically clustered, online communities. ORA is a network analysis and visualization toolkit with special metrics for Twitter data. ORA is used to analyze and visualize the groups detected vis IVCC. We illustrate these methods using two case studies: the ISIS-supporting community on Twitter, as well as a group of Ukrainian nationalists focused on the ongoing Crimean conflict. The goal for these methods is to enable researchers, policy makers and strategists to gain novel insights through the identification and analysis of large online communities. Thus, in addition to presenting these methodologies and tools, we will also discuss what kinds of information and insights can be gained by examining this type of data.

ORA is a powerful network analysis and visualization tool. ORA supports the assessment of standard social network data, organizational network data, high-dimensional network data, meta-network data, geo-spatial network data, and dynamic network data. Relatively unique features include trail and network visualization, fuzzy grouping algorithms, multi-mode network assessment, built in network simulators, JSON and CSV importers, specialized twitter analytics, two mode metrics, and powerful visualizer with data entry and mark-up capabilities. The professional version is capable of handling large 10^6 networks, and can run under the PC, Mac or linux operating system.

Who Should Attend? Those who are interested in assessing social media data and groups as networks. The material and its delivery is suitable for researchers and practitioners, alike. This is designed to be a non-technical workshop, however, by its very nature, the material will involve some mathematics, although this will be minimized as the delivery is driven towards forming a practical understanding of the tactics, techniques and procedures.

Topics Include:

- Twitter analytics
- Community Detection
- Covert group identification
- Social Network Analysis
- Multi-mode, multi-link, high dimensional network metrics
- Controlling for bots
- Twitter Analytics
- Network visualization
- IVCC
- *ORA software

- Social Media Intelligence (SOCMINT)
- Online Social Networks (OSNs)

Computer Equipment:

The software will be screen-projected to the group as a live walk-through demonstration. Time permitting, a small section of the workshop will include some hands-on training. All such components will be through the web or using the windows based operating system. The software presented in this tutorial is Windows operating system based. Participants with windows emulators should pre-load and test the *ORA software from the CASOS website – http://www.casos.cs.cmu.edu/projects/ora/. Participants should bring their own laptops to workshop. Participants not able to bring a Windows-based laptop computer to the sessions are welcome to participate, and will still fully benefit from the workshop. Participants will be eligible for an SBP-BRiMS tutorial discount on the ORA software.

Maximum Number of Attendees: Unlimited

[1] Y. Veilleux-Lepage, "Paradigmatic Shifts in Jihadism in Cyberspace: The Emerging Role of Unaffiliated Sympathizers in the Islamic State's Social Media Strategy," 2015.

[2] M. Mazzetti and M. R. Gordon, "ISIS Is Winning the Social Media War, U.S. Concludes," *The New York Times*, 12-Jun-2015.

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

[4] Kathleen M. Carley, 2014, ORA: A Toolkit for Dynamic Network Analysis and Visualization, In Reda Alhajj and Jon Rokne (Eds.) Encyclopedia of Social Network Analysis and Mining, Springer.

Keywords: Social Media Analytics, Network Analysis, ORA, Twitter,