The Potential of Social Network Analysis in Intelligence

https://www.e-ir.info/2014/01/09/the-potential-of-social-network-analysis-in-intelligence/

KRISTAN J. WHEATON AND MELONIE K. RICHEY, JAN 9 2014

The legality of the National Security Agency's (NSA's) use of US citizens' metadata to identify and track foreign intelligence organizations and their operatives is currently a subject of much debate. Less well understood (and consequently routinely misreported) are the capabilities and limitations of social network analysis, the methodology often used to evaluate this metadata.

One of the first causes of confusion is definitional. Social network analysis is often linked to an inappropriate degree with social media. True, social media such as Facebook and Twitter are frequently used as rich data sources for social network analysis, but understanding the importance of networks in the affairs of states has been around at least since Machiavelli.[1]

In addition, the first modern version of what would come to be called social network analysis was developed not by an intelligence agency or computer scientist but by Columbia professor and psychosociologist, Jacob Moreno, in 1934. These "sociograms," as Moreno called them were used to graph individual preferences or relations within a small group.

Little did Moreno suspect that his method for understanding the relationships between people, when combined with graph theory and the processing power of computers, would allow for the detailed analysis of thousands of people or organizations with hundreds of thousands of connections between them (See Fig. 2). [2]

Figure 2 – Modern social network analysis uses powerful computers and graph theory to map out the relationships between thousands of nodes and hundreds of thousands of links. Shown here is the network of the over 6000 Twitter users who follow the American Nuclear Society along with their over 200,000 connections. (Image Source: Melonie Richey)

The Potential of Social Network Analysis in Intelligence
Written by Kristan J. Wheaton and Melonie K. Richey

Along with the undeniable power of this type of analysis comes the inevitable (and justified) concerns for privacy and constitutionality. But just how powerful is social network analysis? What can intelligence agencies actually glean from the exabytes of data they are purportedly collecting?

Social Network Analysis, as an analytic method, has inarguable applicability to the field of intelligence and is progressively reshaping the analytic landscape in terms of how analysts understand networks. For example, analysts currently use SNA to identify key people in an organization or social network, develop a strategic agent network, identify new agents and simulate information flows through a network. Beyond this, SNA can be easily combined with other analytic practices such as Geographic Information Systems (GIS), gravity model analysis or Intelligence Preparation of the Battlefield (IPB) to create robust, predictive analyses.

Identifying Key People/Organizations in a Network

The most obvious use of SNA is its ability to identify key actors and entities within a network. Centrality measures within a network are means for measuring a node's relative importance within the network. [3] It is well-accepted that "the ability to measure centrality in social networks has been a particularly useful development in social network analysis." What is more interesting, however, is the number of centrality measures that social network analysts use to reveal different things about how key actors interact within a network. [4] For example, a node with a high degree centrality is connected to many other nodes. In Figure 3 below, it is unsurprising that the American Nuclear Society (ANS) has the highest degree centrality in its own Twitter network. However, a node with a high betweenness centrality is one that connects the cliques in the network. Figure 4 shows the same ANS network, reconfigured and revisualized with an emphasis on betweenness, with a new node, Nuclear.com, emerging as the most important.

Figure 3 (Image Source: Melonie Richey)

The Potential of Social Network Analysis in Intelligence
Written by Kristan J. Wheaton and Melonie K. Richey

Figure 4 (Image Source: Melonie Richey)

For example, by analyzing the network in accordance with different centrality measures and establishing filtering criteria (and using Carnegie Mellon's ORA software), [5] we were able to reduce a network representing the entire nuclear energy and non-proliferation communities on Twitter (6000+ nodes and 200,000+ links) to the 19 most influential individuals within that network (See Figure 5). These individuals are the nodes that would be able to disseminate information to the majority of the network within a matter of hours.

Figure 5 (Image Source: Melonie Richey)

The Potential of Social Network Analysis in Intelligence

Written by Kristan J. Wheaton and Melonie K. Richey

Identifying New Agents

Another traditional intelligence activity that could benefit from SNA is identifying potential new "agents" – people or organizations who might be willing or able to provide information to an intelligence agency.

For example, by using Twitter's list feature, which allows users to establish lists of people to follow for particular purposes, and some simple cross-referencing techniques, we were able to identify 50 new, highly reputable individuals and organizations talking about strategic mining and minerals on Twitter. [6]
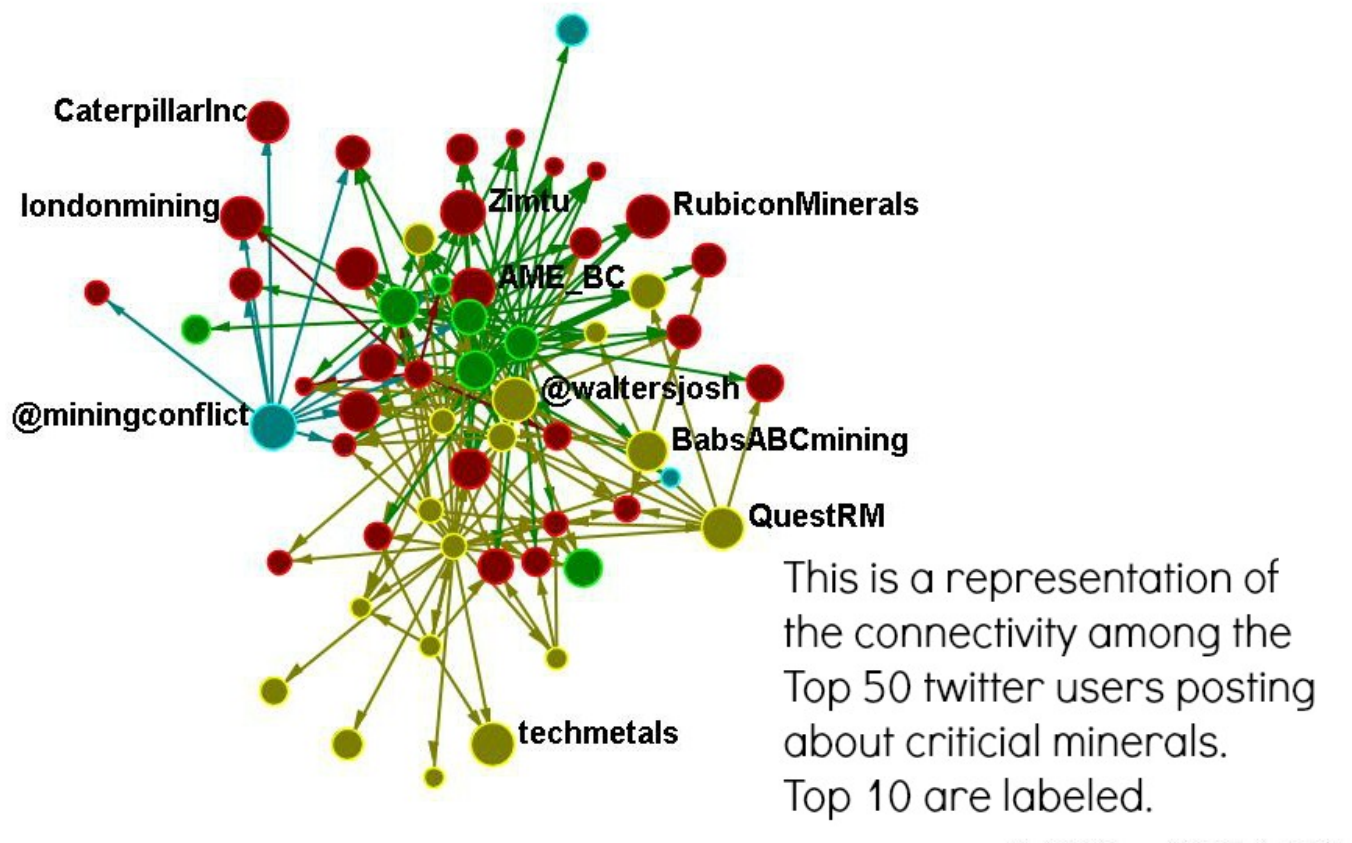

Figure 6 (Image Source: Melonie Richey)

While such a use by intelligence agencies may seem Orwellian, it is similar to techniques currently used in business to identify potential customers. Likewise, a similar algorithm likely supports various friend/colleague recommendation engines such as LinkedIn's "People You May Know" feature.

Simulating Information Flows

Of all the capabilities of SNA, simulations are likely one of the most useful. Carnegie Mellon's ORA, for example, provides four main kinds of simulations in order to demonstrate how money, information, disease or technology would move through a network. Pathway simulations locate the most direct or indirect routes from one node to another. Still other simulations also indicate how a network would react to the removal of any particular node or set of nodes (for example, how a decentralized terrorist network such as the Taliban would function if the leaders from two key cells were killed).

The Potential of Social Network Analysis in Intelligence

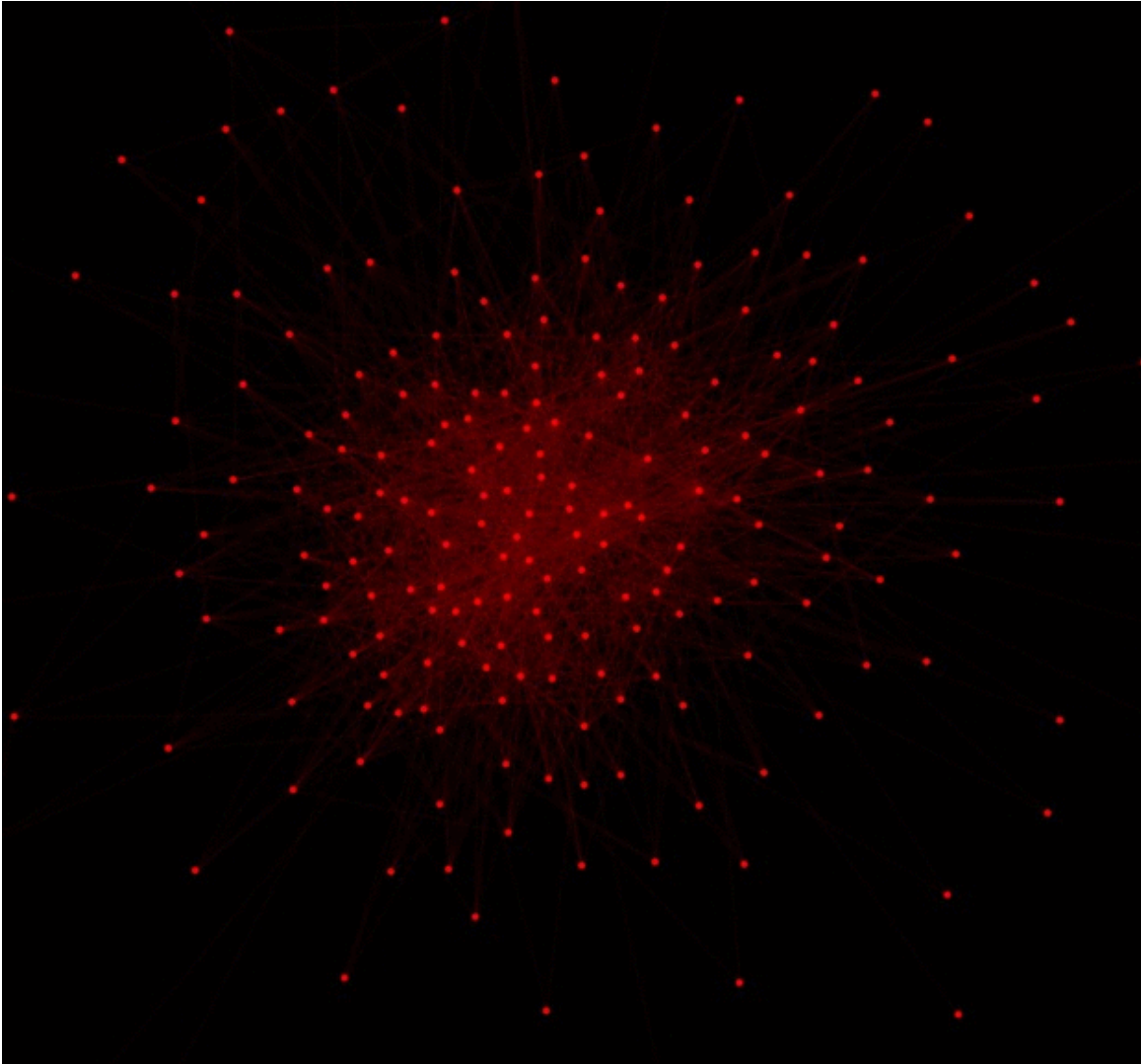Written by Kristan J. Wheaton and Melonie K. Richey

Figure 7 (Image Source: Melonie Richey)

As an example of this feature, Figure 7, shows the effect of providing a highly relevant piece of information to the 19 individuals identified in the Twitter network of nuclear specialists discussed above. The dots, representing individuals and organizations on Twitter, get larger and change color as the information flows throughout the system. Variables within the simulation allow researchers to alter the level of interest the network likely has to a particular piece of information (the information's "virality").

Combining SNA with Other Methods

These simulations and other features of SNA provide idealized analyses that can then be combined with other techniques, such as GIS. Networks within ORA and many other SNA tools can be visualized geospatially if coordinates are provided for each node. Running simulations through these networks can then be represented on a map much like the simulation of Syrian refugee population movement throughout Turkey shown in Figure 8. This, in turn, allows for powerful predictive analytics. Figure 9 reflects the outcome of the simulation in Figure 8; not only does the image represent reality (the known locations of Syrian refugees according to the UN), [7] it also predicts where refugees are likely to move within the next 12 to 24 months. This analysis employed SNA as the cornerstone analytic technique in conjunction with GIS and even includes ideas from the more traditional intelligence methodology of Intelligence Preparation of the Battlefield.

The Potential of Social Network Analysis in Intelligence
Written by Kristan J. Wheaton and Melonie K. Richey

Figure 8 (Image Source: Melonie Richey)

Figure 9 (Image Source: Melonie Richey)

Caveat Emptor

Like all analytic techniques, SNA is imperfect and comes with a number of caveats for researchers new to the method. SNA, while widely applicable, is by no means universally applicable.

For example, in early 2013, one of the authors sought to use SNA to locate terrorists using social media. SNA and social media seemed like a good place to start, even though it seemed unlikely that many such individuals would self identify as a "radical extremist" or "Al-Qaeda affiliate." Ultimately though, the effort failed because there was just too much of what social network analysts like to call "white noise," or extraneous information picked up through a comprehensive scraping of the Internet. Our search for radical extremists returned journalists, university students of international relations and politics, and a slew of ordinary people just keeping up with current events and Tweeting about it.

The Potential of Social Network Analysis in Intelligence

Written by Kristan J. Wheaton and Melonie K. Richey

Another issue with SNA has to do with the nature of relationships. In the real world, they are often messy and convoluted. Just because two people work together and do so often, does not necessarily mean that they like each other. Similarly, the best way to describe the relationship between two businesses might not be the number of contracts the two have signed together. SNA works best, however, with clearly definable relationships and where one factor in the relationship correlates well with other factors important in a relationship. Modern intelligence problems, which often contain, political, economic, military, tribal, geographic, personal, and historical relationship data require the application of advanced SNA techniques and, even then, may yield little of real use to decisionmakers.

Finally, SNA is fundamentally a mathematical tool but is most useful in the decisionmaking process when the networks are visualized. It is, without doubt, the visualization of these networks that tends to capture the most attention from the policymakers that intelligence units typically support. This is both a blessing and a curse. While it is easy to capture attention, explaining why the charts and graphs look the way they do is an art. All too often, the initial excited reaction to these diagrams turns to boredom and confusion as analysts bog the decisionmakers down with the arcana of SNA. In addition, creating these complex visualizations often stresses even the most powerful personal computers (the images of the simulation in Figure 8 above took approximately 2 hours to produce using a powerful desktop PC with two high end graphics cards).

Like every analytic technique, SNA has great utility for the right question. Within its limits, SNA is unmatched and can be usefully applied to identify key individuals or organizations within a network, generate new leads and simulate the flows of information or money throughout a network. SNA, however, remains just *an* answer, not *the* answer. Used inappropriately or without a full understanding of the limits of the method and analysts will only be finding new and more technically sophisticated ways to fail. That, then, is the primary job of the modern day analyst: making the judgment call of which techniques to use and when. Equally as important as knowing when to use SNA is knowing when *not* to use it.

—

Kristan J. Wheaton *is an Associate Professor at Mercyhurst University who teaches the use of social network analysis and its application to traditional intelligence problems.*

Melonie K. Richey *is a graduate student at Mercyhurst University whose current research focuses on modeling and simulating the Syrian refugee crisis by combining principles of social network analysis, geospatial analysis and military techniques designed to understand how terrain impacts the movements of enemy forces.*

[1] Machiavelli, N. (1515). Why the kingdom of darius, conquered by alexander, did not rebel against the successors of alexander at his death. In *The Prince* Retrieved from http://www.constitution.org/mac/prince00.htm

[2] Rieder, B. (2012, March 19). Retrieved from http://thepoliticsofsystems.net/2012/03/

[3] Newman, M. (2009). *Networks: an introduction*. Oxford University Press, chap. 3.

[4] Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled.*Social networks*, *25*(4), 283-307.

*ORA: *Software*. (2013). Available from Carnegie Mellon. Retrieved from http://www.casos.cs.cmu.edu/projects/ora/software.php

[6] For more information on how we did this analysis (and both the strengths and weaknesses of SNA as a tool for finding "agents"), see "The New HUMINT?"

[7] "UNHCR Turkey Syrian refugee daily sitrep." *UNHCR: The UN refugee agency*, November 25, 2013.

The Potential of Social Network Analysis in Intelligence

Written by Kristan J. Wheaton and Melonie K. Richey

[8] UNHCR: UN refugee agency, "Syrian refugee camps in Turkey." Last modified: October 243, 2013. Accessed: November 25, 2013.

[9] Syria needs analysis project (SNAP), "Regional analysis Syria – Part II: Host Countries." Last modified September 26, 2013. Accessed November 25, 2013.

About the author:

Kristan J. Wheaton is an Associate Professor at Mercyhurst University whose research focuses on the intersection of games-based learning and intelligence analysis.

Melonie K. Richey is a graduate student at Mercyhurst University who is currently participating in IARPA's Sirius Project. Her current research focuses on exploring the viability of tabletop games in the intelligence classroom.