Strategic Social Network Analysis

Tomasz P. Michalak[1,2] and **Talal Rahwan**[3] and **Michael Wooldridge**[2]
[1]Institute of Informatics, University of Warsaw, Poland
[2]Department of Computer Science, University of Oxford, United Kingdom
[3]Masdar Institute of Science and Technology, United Arab Emirates

Abstract

How might individuals, groups, or subnetworks manage their social connections so that they are less exposed to the workings of social network analysis tools? To put it differently, how can one disguise his or her standing in a network to subvert such analysis? These fundamental questions have received very little attention in the literature to date, as most research has focused on developing ever more elaborate social network analysis techniques, rather than studying how these techniques can be evaded. Addressing these questions may not only help the general public to better protect their privacy, but may also help oppressed activist groups to better conceal their existence, or even help law-enforcement and security agencies to better understand how terrorists escape detection. In this paper, we outline a new paradigm for social network analysis, whereby the strategic behaviour of network actors is explicitly considered.

Introduction

Many research problems in social network analysis (SNA) have received considerable attention in recent years across various disciplines, including Mutli-Agent Systems (Sabater and Sierra 2002; Nguyen, Kowalczyk, and Chen 2009) and, more generally, Artificial Intelligence. Scientists, developers, and users have focused on improving the performance of various SNA tools, such as *centrality* measures (Koschützki et al. 2005), *community-detection* algorithms (Orman and Labatut 2009) or *link-prediction* algorithms (Getoor and Diehl 2005), just to name a few. However, for all their benefits, the widespread use of such tools raises legitimate privacy concerns that are likely to grow in the future.

To tackle such privacy issues, various countermeasures have been proposed, ranging from strict legal controls (EU: 2015), through algorithmic solutions (Kearns et al. 2016), to market-like mechanisms that allow participants to monetize their personal information (Lane et al. 2014). However, to date, few such countermeasures have been implemented, leaving the privacy issue largely unresolved, e.g., as is evident from the very recent release of Facebook's "*Global Government Requests Report*" (Facebook 2016), which revealed a global increase in government requests to secretly access user data. Furthermore, it is unlikely that effective legal mechanisms will be introduced in countries with authoritarian regimes, where social networking sites and other internet content is policed, and anti-governmental blogs and activities are censored (King et al. 2013; 2014).

However, the relative ineffectiveness of centrally-managed mechanisms does not necessarily leave the general public with no options to protect their privacy. In fact, although such privacy issues have attracted little attention in the first years of the "digital society" revolution, nowadays the general public is becoming increasingly aware of various forms of "digital surveillance".

With this in mind, we ask the following question: *Could members of a social network strategically manipulate their on-line data in an attempt to evade SNA tools?* Addressing this question may not only help the general public to better protect their privacy, but it may also help oppressed activist groups to better conceal their existence, or even prove to be vital to law-enforcement and security agencies. In particular, recent findings on covert organizations—especially with respect to the tech-savvy ISIS—clearly demonstrate their ability to neutralize counter-terrorism efforts by the authorities. The known evasion techniques used by ISIS range from changing aliases and keeping personal profiles private (Nordrum 2016) to using encrypted communication platforms (such as Telegraf (Khayat 2015)) and staging the disappearance of an entire group from social media only to pop up again in a different place under alternative aliases (Nordrum 2016). In fact, it is believed that the evasion capabilities of ISIS significantly increased after Edward Snowden's disclosure of classified information on the SNA techniques used by US intelligence (Scarborough 2014).

Unfortunately, we have neither sufficient understanding of such evasion techniques nor do existing SNA tools have the ability to adequately deal with them. This is because most SNA tools were built around the assumption that individuals or groups in a network do not act strategically to deliberately evade those tools. Even the more advanced techniques that are especially designed for analysing covert networks typically assume that the network under investigation is not subject to strategic manipulation. Given this, we believe that the literature has now reached a point where serious attention should be directed towards the *strategic evasion of SNA tools*.

New Paradigm

In this section we outline a new paradigm in social network analysis, whereby the strategic behaviour of network actors is explicitly considered. As such, it lies at the intersection of social network analysis and *game theory* (Maschler, Solan, and Zamir 2013).

Typically, an SNA tool is an algorithm designed to solve particular problem. For instance, the *link prediction problem* (Liben-Nowell and Kleinberg 2007; Lü and Zhou 2011) is to predict, based on the current structure of the network, which connections are most likely to occur in the near future. Link prediction algorithms try to solve this problem typically by evaluating the probability that every not-yet-discovered edge in the network will form (Getoor and Diehl 2005). Here, it is assumed that members of the social network do not act strategically to mislead any link prediction algorithms, nor do these algorithms account for such a possibility.

In contrast, we propose to *frame SNA problems as strategic games*. We call such a game *the Seaker-Evader game*. Its definition (in a basic form) is composed of:

- the set of players, N, that involves *the Seeker(s)*, and *Evaders*, be them nodes, edges, and/or any other network entity (such as a group of nodes). While, in principle, there may be a multiplicity of the Seekers in our model, to focus attention, in what follows we assume that there is only one Seeker.
- the set of available strategies S_i for each player $i \in N$ (be them pure or mixed). For instance, the set of strategies of an Evader i can consist of evasion algorithms which are built around the following actions: (a) creating a connection with node j ("befriending" a node); or (b) cutting an existing connection ("unfriending" a node). In more sophisticated settings, it can involve such moves as "covering-up" the true nature of an existing connection, or hiding some of its characteristics. Furthermore, evading algorithms can be parametrized, in which case the strategy spaces of the Evaders involve the corresponding parameter ranges. On the other hand, the set of strategies of the Seeker may consist of some SNA tools (from which the Seeker can choose a single one or, perhaps an arbitrary subset). Furthermore, also some SNA algorithms can be parametrized, in which case the strategy space of the Seeker involves the corresponding parameter ranges. Finally, the strategy space can allow the Seeker to develop completely new algorithms that are better adjusted to the evading strategies available to the opponents.
- the set of utility functions (or, alternatively, preference relations) represent the players' attitudes to the outcomes that result from the different choices of actions; and
- the "knowledge functions" which define "who knows what". For instance, it can be assumed that the Seeker is completely oblivious to the evasion techniques available to, or used by, the Evaders.

We assume that players in a Seeker-Evader game act rationally in the furtherance of their preferences, each accounting for the rational behaviour of others, and then we look for an *equilibrium* of the system. Importantly, due to the above definition of the strategy spaces of players, the *equilibria* of our model will be the combination of:

- the evasion techniques of the Evaders; and
- the SNA tool of the Seekers.

Hence, the SNA tools (possibly adjusted to the evasion techniques) and the evasion techniques (possibly adjusted to the SNA tools) will *emerge as the equilibria* of our model. Furthermore, while the above model is defined as a non-cooperative game, it can be straightforwardly extended to incorporate cooperative behaviour, e.g., among the Evaders.

While the variables defining the above Seeker-Evader game can be set in any constellation, we envisage that the models should be built gradually, from the easiest to the most complex ones. We suggest the following steps:

- **Step 1:** The analysis of potential evasion techniques against the existing SNA tools, where it is assumed that the Seeker does not act strategically and is not aware of potential evasion efforts of the Evaders.
- **Step 2:** The analysis of potential evasion techniques against the existing SNA tools, where it is assumed that all the parties are strategic, though the strategy spaces are limited to basic evasion techniques and known SNA algorithms.
- **Step 3:** The analysis of potential evasion techniques against the existing and possibly novel (adapted to the new setting) SNA tools. Here, it is assumed that all the parties are strategic and that the strategy spaces can take more complex forms.

To the best of our knowledge, our recent series of papers on evading centrality measures (Waniek et al. 2016a; 2016b), community detection algorithms (Waniek et al. 2016a), and link prediction algorithms (Waniek, Rahwan, and Michalak 2016), are the first works that can be categorized under **Step 1**. In the next section we discuss some results from Waniek et al. (2016a) in more detail.

Sample Analysis

Waniek et al. (2016a) focused on the following questions:

- how key individuals might pro-actively manage their social connections so that they are less likely to be identified as important nodes by centrality measures but, at the same time, they do not loose much of their influence in the network?
- how communities might proactively manage their social connections so that they are less exposed to the workings of community detection algorithms?

In the remainder of this section, we briefly describe the model and some of the main results that concern the first question.

The model by Waniek et al. can be seen a degenerate Seeker-Evader game. It is defined as follows. The set of players consists of the non-strategic Seeker and the strategic Evaders. The latter take a joint action to disguise the *leader* of the network from three fundamental centrality measures:

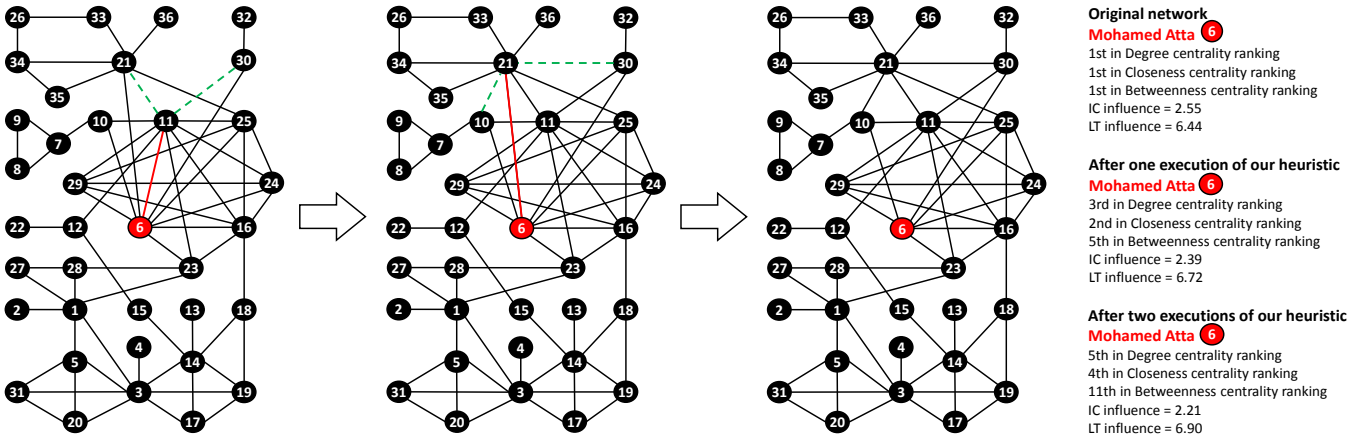

Figure 1: (Figure 1 in Waniek et al., 2016a) It is sufficient to execute the ROAM heuristic twice on the 9/11 terrorist network to hide Mohamed Atta—who is generally concsidered to be one of the ringleaders of the attack (Krebs 2002). In each step the solid red links are the ones removed by the algorithm, and the dashed green links are the ones added.

degree, closeness, and betweenness (Koschützki et al. 2005). The leader is defined as a member of the social network with the *highest influence*. Here, Waniek et al. consider two best established mathematical models of influence: the *Independent Cascade* model (Goldenberg, Libai, and Muller 2001) and the *Linear Threshold* model (Kempe, Kleinberg, and Tardos 2003). The game is degenerate as it is assumed that the Seeker is unaware of evasion efforts undertaken by the Evaders.

Since the leader is the most influential, he or she will be typically ranked among the top nodes by all three centrality measures: degree, closeness and betweenness. Hence, the objective of the Evaders is to alter the network so that the centrality of their leader is lowered, however, at the same time, the leader's influence is not compromised.[1] It is assumed that, to achieve their objective, the Evaders can rewire the links of the network, without exceeding a certain *budget*—the maximum number of links allowed to be modified (i.e., added or removed).

Waniek et al. prove that finding an optimal solution to the above problem is hard to compute. However, they demonstrate that even a simple heuristic, whereby attention is restricted to the individual's immediate neighbourhood, can be surprisingly effective in practice. Their heuristic, called ROAM—Remove One, Add Many—is as follows. Given a budget b:

- **Step 1:** Remove the link between the leader, v^L, and its neighbour of choice, v_i;
- **Step 2:** Connect v_i to $b-1$ nodes of choice, who are neighbours of v^L but not of v_i (if there are fewer than $b-1$ such neighbours, connect v_i to all of them).

Figure 1 illustrates how this heuristic works on the WTC 9/11 terrorist network. Interestingly, ROAM is able to disguise Mohamed Atta's leading position within the network. This is achieved by rewiring a strikingly-small number of his connections and the connections between his immediate neighbours.

Waniek et al. experiment with two-types of real-life networks: (i) *Covert networks* responsible for the WTC 9/11 attacks (Krebs 2002), the 2002 Bali attack (Hayes 2006), and the 2004 Madrid train bombings (Hayes 2006), respectively; and (ii) Anonymized fragments of *Social networks*: Facebook, Twitter and Google+, taken from SNAP (Leskovec and Mcauley 2012). They also consider randomly-generated networks: (i) *Scale-free* networks, the Barabasi-Albert model (Barabási and Albert 1999); (ii) *Small-world* networks, the Watts-Strogatz model (Watts and Strogatz 1998); and (iii) *Random graphs*, the Erdos-Renyi model (Erdős and Rényi 1959).

Each of their experiments consists of a network, a budget (either 2, 3, or 4), a leader, and an influence model (either Independent Cascade or Linear Threshold). The node chosen to be a leader is the one with the lowest sum of centrality rankings (with ties broken uniformly at random). The results of some of the experiments are presented in Figure 2. They concern one covert organization (Madrid bombing), one social network fragment (Facebook fragment), and one randomly-generated network (scale-free network generated using Barabasi-Albert model with 100 nodes and 3 edges added for each node). The thirst three columns of plots report the *ranking* of the leader, whereas the latter two columns of plots report the *relative* influence value of the leader as compared to the *original* influence value of the leader before executing the heuristic altogether. The ROAM heuristic turns out to be effective in decreasing the leader's ranking, and its efficiency depends on the size of the budget. The same holds for the influence, i.e., heuristic with higher budget often manages to maintain (or even increase) the influence from before the disguising process.

The work of Waniek et al. can be also seen as an extension of the sensitivity analyses of centrality measures (Correa et al. 2012) and community detection algorithms (Orman and

[1] This setting can be also interpreted as seeking a balance between two measures of a node (leader) importance: its centrality and its influence.

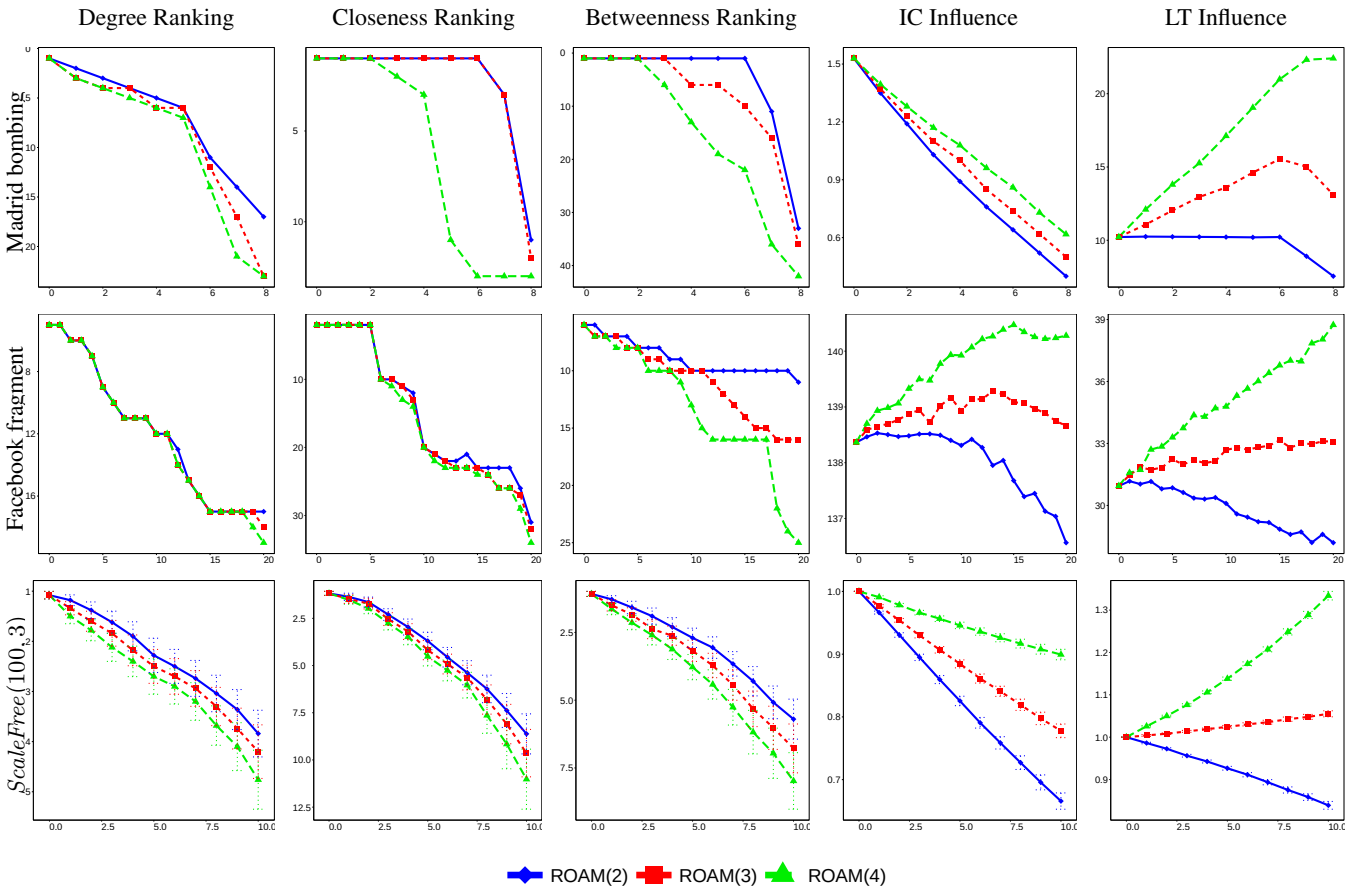

Figure 2: (Figure 3 from Waniek et al, 2016a) Executing ROAM multiple, consecutive times, where the x-axis represents the number of executions. The subfigures show the source node's ranking (according to different centrality measures), and the relative change in its influence value (according to different influence models) for the Madrid-attack network, 50 scale-free networks, and a medium-sized fragment of Facebook's network (333 nodes, 5038 edges). The size of the budget b is $2, 3$, or 4.

Labatut 2009). Nevertheless, while such analyses from the literature usually focus on the effects random network alterations, Waniek et al. studies changes that strategic in nature.

Related Work

Various, well-established, research themes are positioned at the interface of social network analysis and game theory. One example is the economic literature (Jackson 2005; 2008; 2010) on endogenous network formation. The other one is the literature on mechanism design for social networks (Shneidman and Parkes 2003; Singh, Jain, and Kankanhalli 2011). Game theory is also used as a backbone of some SNA algorithms such as game-theoretic centrality measures (Grofman and Owen 1982; Michalak et al. 2013; Szczepański et al. 2016) or game-theoretic community detection algorithms (Chen et al. 2010; McSweeney, Mehrotra, and Oh 2014). Nevertheless, we believe that our proposal of strategic social network analysis has not been yet considered in the literature.

Naturally, there are various models in the game-theoretic literature upon which we can build our analysis of the Seeker-Evader game. Perhaps the most relevant model is the game of hide-and-seek (Rubinstein et al. 1997, Chapman et al. 2014), where one party hides something and another party then seeks to find these items. Also relevant is work on *epistemic game teory* (Aumann and Brandenburger 2016), which tries to understand and make explicit (often through the use of epistemic logic) the assumptions about "who knows what" that are often left implicit in game theoretic settings. Finally, *security games*—a rich research line championed by M. Tambe and his lab (Fave et al. 2015)—should be mentioned.

Acknowledgements

Michael Wooldridge and Tomasz P. Michalak were supported by the European Research Council under Advanced Grant 291528 ("RACE"). Tomasz P. Michalak was also supported by the Polish National Science Centre grant 2014/13/B/ST6/01807.

References

Aumann, R. J., and Brandenburger, A. 2016. Epistemic conditions for nash equilibrium. In *Readings in Formal Epistemology*. Springer. 863–894.

Barabási, A.-L., and Albert, R. 1999. Emergence of scaling in random networks. *science* 286(5439):509–512.

Chapman, M.; Tyson, G.; McBurney, P.; Luck, M.; and Parsons, S. 2014. Playing hide-and-seek: an abstract game for cyber security. In *Proceedings of the 1st International Workshop on Agents and CyberSecurity*, 3. ACM.

Chen, W.; Liu, Z.; Sun, X.; and Wang, Y. 2010. A game-theoretic framework to identify overlapping communities in social networks. *Data Mining and Knowledge Discovery* 21(2):224–240.

Correa, C. D.; Crnovrsanin, T.; and Ma, K.-L. 2012. Visual reasoning about social networks using centrality sensitivity. *Visualization and Computer Graphics, IEEE Transactions on* 18(1):106–120.

Erdős, P., and Rényi, A. 1959. On random graphs i. *Publ. Math. Debrecen* 6:290–297.

2015. European data protection supervisor, meeting the challenges of big data, opinion 7/2015.

Facebook. 2016. https://govtrequests.facebook.com/.

Fave, F. M. D.; Shieh, E. A.; Jain, M.; Jiang, A. X.; Rosoff, H.; Tambe, M.; and Sullivan, J. P. 2015. Efficient solutions for joint activity based security games: fast algorithms, results and a field experiment on a transit system. *Autonomous Agents and Multi-Agent Systems* 29(5):787–820.

Getoor, L., and Diehl, C. P. 2005. Link mining: a survey. *ACM SIGKDD Explorations Newsletter* 7(2):3–12.

Goldenberg, J.; Libai, B.; and Muller, E. 2001. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *Academy of Marketing Science Review* 9(3):1–18.

Grofman, B., and Owen, G. 1982. A game theoretic approach to measuring degree of centrality in social networks. *Social Networks* 4:213–224.

Hayes, B. 2006. Connecting the dots can the tools of graph theory and social-network studies unravel the next big plot? *American Scientist* 94(5):400–404.

Jackson, M. O. 2005. A survey of network formation models: stability and efficiency. *Group Formation in Economics: Networks, Clubs, and Coalitions* 11–49.

Jackson, M. O. 2008. *Social and economic networks*, volume 3. Princeton university press.

Jackson, M. O. 2010. An overview of social networks and economic applications. *The handbook of social economics* 1:511–85.

Kearns, M.; Roth, A.; Wu, Z. S.; and Yaroslavtsev, G. 2016. Private algorithms for the protected in social network search. *Proceedings of the National Academy of Sciences* 201510612.

Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *SIGKDD*.

Khayat, M. 2015. Jihadis Shift To Using Secure Communication App Telegram's Channels Service. *Inquiry & Analysis Series* (1198).

King, G.; Pan, J.; and Roberts, M. E. 2013. How censorship in china allows government criticism but silences collective expression. *American Political Science Review* 107(02):326–343.

King, G.; Pan, J.; and Roberts, M. E. 2014. Reverse-engineering censorship in china: Randomized experimentation and participant observation. *Science* 345(6199):1251722.

Koschützki, D.; Lehmann, K. A.; Peeters, L.; Richter, S.; Tenfelde-Podehl, D.; and Zlotowski, O. 2005. Centrality indices. In *Network Analysis*, volume 3418 of *Lecture Notes in Computer Science*. Springer. 16–61.

Krebs, V. 2002. Mapping networks of terrorist cells. *Connections* 24:43–52.

Lane, J. I.; Stodden, V.; Bender, S.; and Nissenbaum, H., eds. 2014. *Privacy, big data, and the public good: frameworks for engagement*.

Leskovec, J., and Mcauley, J. J. 2012. Learning to discover social circles in ego networks. In *Advances in neural information processing systems*, 539–547.

Liben-Nowell, D., and Kleinberg, J. 2007. The link-prediction problem for social networks. *Journal of the American society for information science and technology* 58(7):1019–1031.

Lü, L., and Zhou, T. 2011. Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and its Applications* 390(6):1150–1170.

Maschler, M.; Solan, E.; and Zamir, S. 2013. *Game Theory*. Cambridge University Press.

McSweeney, P. J.; Mehrotra, K.; and Oh, J. C. 2014. Game-theoretic framework for community detection. In *Encyclopedia of Social Network Analysis and Mining*. Springer. 573–588.

Michalak, T. P.; Aadithya, K. V.; Szczepański, P. L.; Ravindran, B.; and Jennings, N. R. 2013. Efficient computation of the Shapley value for game-theoretic network centrality. *Journal of Artificial Intelligence Research* 46:607–650.

Nguyen, N. T.; Kowalczyk, R.; and Chen, S.-M. 2009. Computational collective intelligence. semantic web, social networks and multiagent systems. *Lecture Notes in Computer Science* 5796.

Nordrum, A. 2016. Pro-ISIS Online Groups Use Social Media Survival Strategies to Evade Authorities.

Orman, G. K., and Labatut, V. 2009. A comparison of community detection algorithms on artificial networks. In *Discovery science*, 242–256. Springer.

Rubinstein, A.; Tversky, A.; and Heller, D. 1997. Naive strategies in competitive games. In *Understanding Strategic Interaction*. Springer. 394–402.

Sabater, J., and Sierra, C. 2002. Reputation and social network analysis in multi-agent systems. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, 475–482. ACM.

Scarborough, R. 2014. Islamic State using leaked Snowden info to evade U.S. intelligence.

Shneidman, J., and Parkes, D. C. 2003. Rationality and self-interest in peer to peer networks. In *International Workshop on Peer-to-Peer Systems*, 139–148. Springer.

Singh, V. K.; Jain, R.; and Kankanhalli, M. 2011. Mechanism design for incentivizing social media contributions. In *Social media modeling and computing*. Springer. 121–143.

Szczepański, P. L.; Michalak, T. P.; and Rahwan, T. 2016. Efficient algorithms for game-theoretic betweenness centrality. *Artificial Intelligence* 231:39–63.

Waniek, M.; Rahwan, T.; Michalak, T.; and Wooldridge, M. 2016a. Hiding Individuals and Communities in a Social Network. https://arxiv.org/abs/1608.00375.

Waniek, M.; Rahwan, T.; Michalak, T.; and Wooldridge, M. 2016b. On the Construction of Covert Network. mimeo, University of Oxford, available on request.

Waniek, M.; Rahwan, T.; and Michalak, T. 2016. Hiding Relationships in a Social Network. mimeo, University of Oxford, available on request.

Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of small-world networks. *nature* 393(6684):440–442.