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 can individuals and communities protect their privacy against social network analysis tools? How do criminals or terrorists organizations evade detection by such tools? Under which conditions can these tools be made strategy proof? These fundamental questions have attracted little attention in the literature to date, as most social network analysis tools are built around the assumption that individuals or groups in a network do not act strategically to evade such tools. With this in mind, we outline in this paper a new paradigm for social network analysis, whereby the strategic behaviour of network actors is explicitly modeled. Addressing this research challenge has various implications. For instance, it may allow two individuals to keep their relationship secret or private. It may also allow members of an activist group to conceal their membership, or even conceal the existence of their group from authoritarian regimes. Furthermore, it may assist security agencies and counter terrorism units in understanding the strategies that covert organizations use to escape detection, and give rise to new strategy-proof countermeasures.

Introduction

Many problems in social network analysis (SNA) have received considerable attention in recent years across various disciplines, including Artificial Intelligence and Multi-Agent Systems (Sabater and Sierra 2002; Nguyen, Kowalczyk, and Chen 2009). Scientists, developers, and analysts 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.

Unfortunately, while such tools have many legitimate applications, they can be also used to invade privacy or even undermine security of individuals and groups. For instance, by analysing Facebook's topology as well as the attributes of some users, it is possible to infer attributes of other users (Mislove et al. 2010). Furthermore, such an "attribute inference attack" (Zheleva and Getoor 2009) can be strengthened if it is preceded by a "link prediction attack" which aims at revealing the links that appear to be missing—either deliberately or otherwise—from the topology of the social network.

Naturally, various countermeasures supporting privacy and security that have been proposed including strict legal controls (EU: 2015), algorithmic solutions (Kearns et al. 2016), and market-like mechanisms with which the participants are able to monetize their personal information (Lane et al. 2014). They are, however, difficult to implement in practice, especially at a global scale. For instance, it is highly unlikely that legal privacy-protection mechanisms will be enforced by authoritarian regimes who typically censor internet content including social media (King et al. 2013).

Against this background, we ask the following question: *How can members of a social network strategically manipulate their online data in an attempt to evade SNA tools?* Addressing this question may help members of the general public to better protect their online privacy and security. It may also help activists groups in avoiding censorship. Furthermore, it may assist security agencies and counter-terrorism units in understanding the strategies that covert organizations use to escape detection. 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, neither do we have sufficient understanding of such evasion techniques nor do existing SNA tools have the ability to internalize them. This is because most SNA tools were built around the assumption that individuals or groups in a network do not act strategically to evade those tools. Even the more advanced tools that are especially designed for analysing covert networks (Perliger and Pedahzur 2011) 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 and building new strategy-proof tools*.

[*] Both first authors contributed equally to this paper.

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 a particular problem. Take for example a *link prediction* algorithm, which is an SNA tool whose goal is to predict, based on the current structure of the network, which connections are most likely to be added to the network in the near future (Getoor and Diehl 2005). A link prediction algorithm may also be used in scenarios where only part of the network is observable, and the goal is to identify the edges that appear to be missing, but are in fact present in the actual network (Liben-Nowell and Kleinberg 2007). Such an algorithm, like any other SNA tool, does not consider the possibility that members of the social network may act strategically in an attempt to mislead the tool.

In contrast, we propose to *frame SNA problems as strategic games*. To this end, we propose what we call the *Seaker-Evader game* which (in its basic form) is defined by:

- A set of players, N, which includes the *Seeker(s)*, as well as the *Evader(s)*, be they nodes, edges, and/or any other network entity (such as a subgraph, for example). While, in principle, there may be multiple Seekers in our model, 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 they pure or mixed). For instance, the set of strategies of an Evader i can consist of evasion algorithms that are built around the following actions: (a) creating a connection with node j (i.e., "befriending" a node); or (b) cutting an existing connection (i.e., "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, evasion 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 available to the Seeker may consist of certain SNA tools (from which the Seeker can choose a single one, or perhaps an arbitrary subset). Those SNA algorithms can also 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 evasion strategies that are available to the Evader(s).
- 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(s) of the Seeker.

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 lose 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?

We now 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: degree, closeness, and betweenness (Koschützki et al. 2005).

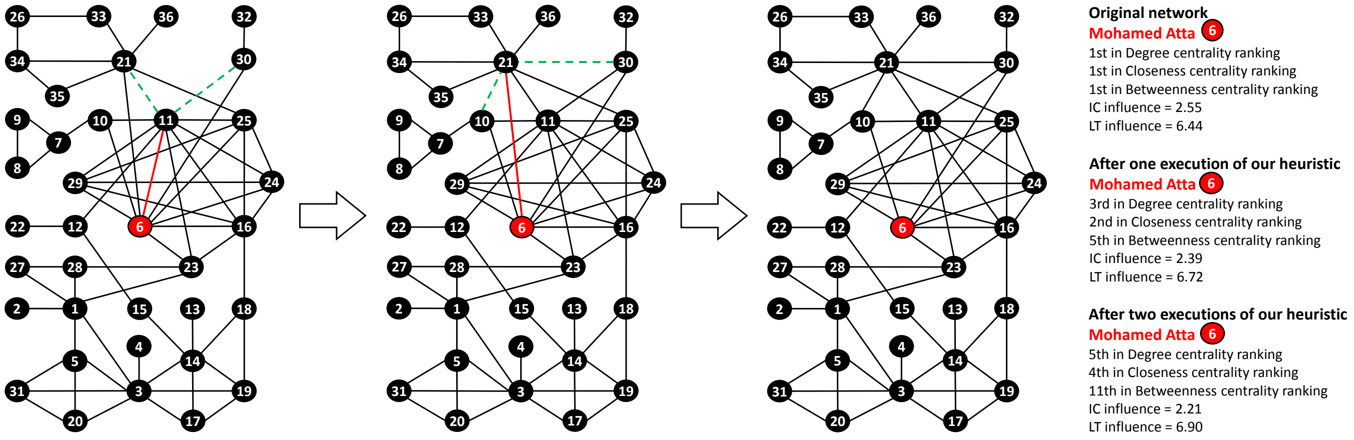
Original network
Mohamed Atta 6
1st in Degree centrality ranking
1st in Closeness centrality ranking
1st in Betweenness centrality ranking
IC influence = 2.55
LT influence = 6.44
After one execution of our heuristic
Mohamed Atta 6
3rd in Degree centrality ranking
2nd in Closeness centrality ranking
5th in Betweenness centrality ranking
IC influence = 2.39
LT influence = 6.72
After two executions of our heuristic
Mohamed Atta 6
5th in Degree centrality ranking
4th in Closeness centrality ranking
11th in Betweenness centrality ranking
IC influence = 2.21
LT influence = 6.90

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 link is the one removed by the algorithm, and the dashed green links are the ones added.

The leader is defined as the member of the social network with the *highest influence*. Here, two established mathematical models of influence were considered: the *Independent Cascade* model and the *Linear Threshold* model (Kempe, Kleinberg, and Tardos 2003). The game is degenerate as it is assumed that the Seeker is unaware of any 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 rewire the network so that the centrality of their leader is decreased, without compromising leader's influence over the network.[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 NP-hard. 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, the 2002 Bali attack, and the 2004 Madrid train bombings, respectively; and (ii) Anonymized fragments of *Social networks*: Facebook, Twitter and Google+, taken from SNAP (Leskovec and Mcauley 2012).
- three well-known classes of randomly-generated networks: (i) *Scale-free* networks, i.e., the Barabasi-Albert model; (ii) *Small-world* networks, i.e., the Watts-Strogatz model; and (iii) *Random graphs*, i.e., the Erdös-Rényi model.

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 the Barabasi-Albert model with 100 nodes and 3 edges added for each node). The subplots in the first three columns depict the *ranking* of the leader, whereas those in the latter two columns depict the *relative* influence value of the leader, compared to the *original* influence value of the leader before executing the heuristic altogether. As can be seen, the ROAM heuristic turns out to be effective in decreasing the leader's ranking, and its efficiency depends on the size of the budget. As for the influence, with higher budget the heuristic often maintains (or even increases) the leader's influence.

The work of Waniek et al. can be also seen as an extension of the line of research that analyses sensitivity of centrality measures (Correa, Crnovrsanin, and Ma 2012). However, while such analyses from the literature usually focus on the effects of random network alterations, Waniek et al. focus on alterations that are strategic in nature.

[1] Alternatively, this setting can be interpreted as seeking a balance between two measures of a node importance: its centrality and its influence.

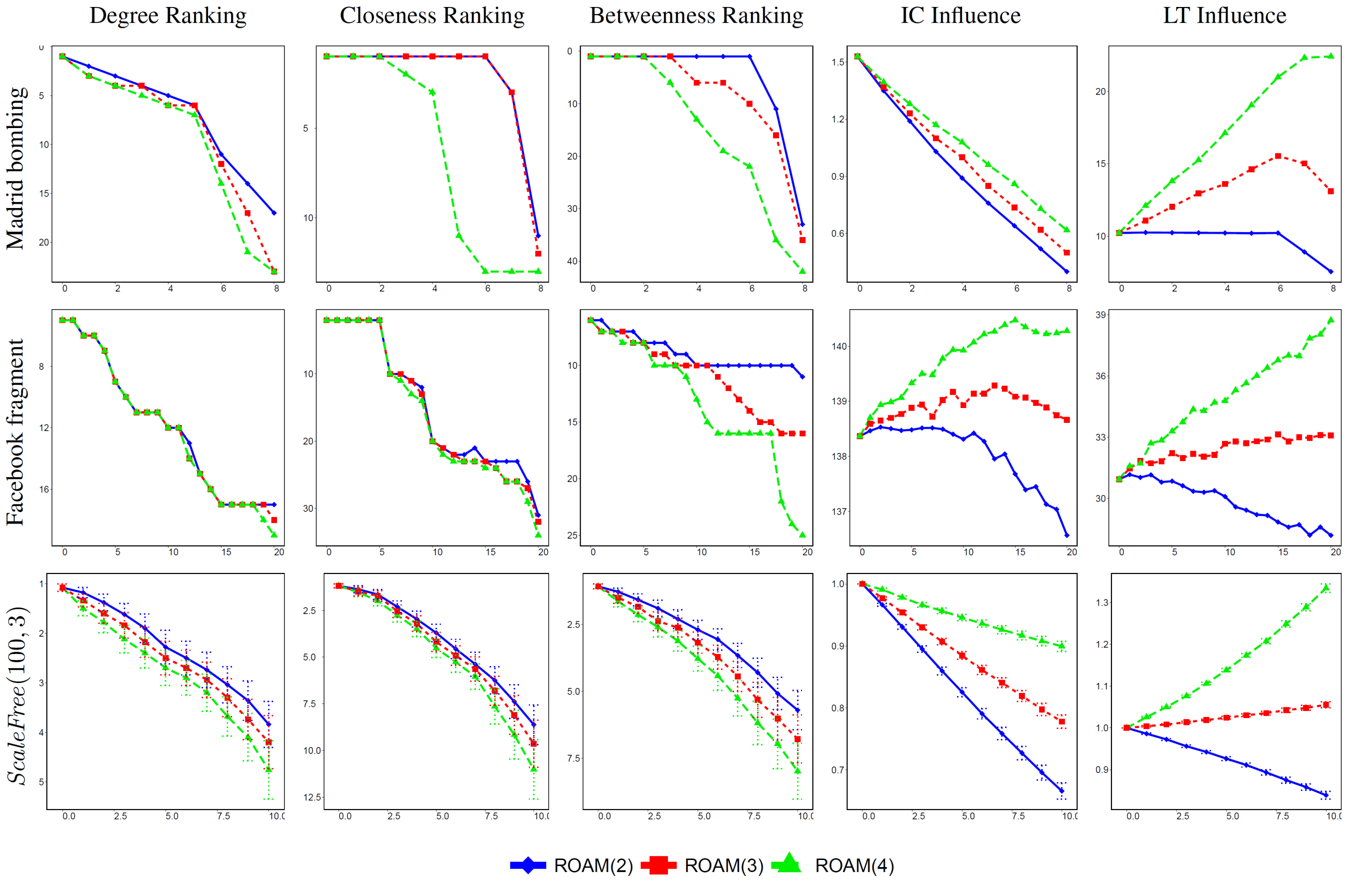

Figure 2: (Figure 3 from Waniek et al, 2016a) Executing ROAM multiple, consecutive times, where the x-axis represents the number of executions. The subfigures show the source node's ranking (according to different centrality measures), and the relative change in its influence value (according to different 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.

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 on endogenous network formation (Jackson 2005). The other one is the literature on mechanism design for social networks (Singh, Jain, and Kankanhalli 2011). Also the literature on online network threats (such as Sybil attacks, Danezis and Mittal 2009, and link reconstruction attacks, Fire et al. 2012) typically assumes strategic behaviour of some actors. Game theory is also used as a backbone of some SNA algorithms such as game-theoretic community detection algorithms (Chen et al. 2010; McSweeney, Mehrotra, and Oh 2014) or game-theoretic centrality measures (Grofman and Owen 1982; Michalak et al. 2013; Szczepański et al. 2016). Nevertheless, we believe that our proposal to explicitly and thoroughly consider the strategic behaviour of actors in social network analysis expands on the current state of the art.

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 certain items and another party then seeks to find those items. Also relevant is the work on *epistemic game teory* (Aumann and Brandenburger 2016), which tries to understand and make explicit (typically through the use of epistemic logic) the assumptions about "who knows what", which 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)—is relevant as it also considers the strategic interactions of actors within a security context.

Acknowledgements

We thank all the reviewers for their helpful comments. 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.

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.

Danezis, G., and Mittal, P. 2009. Sybilinfer: Detecting sybil nodes using social networks.

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

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.

Fire, M.; Katz, G.; Rokach, L.; and Elovici, Y. 2012. Links reconstruction attack using link prediction algorithms to compromise social networks privacy.

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

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

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

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.

Koschützki, D.; Lehmann, K. A.; Peeters, L.; Richter, S.; Tenfelde-Podehl, D.; and Zlotowski, O. 2005. Centrality indices. In *Network Analysis*, LNCS. Springer. 16–61.

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

Kumar, A., and Rathore, N. 2016. Improving attribute inference attack using link prediction in online social networks. In *Recent Advances in Mathematics, Statistics and Computer Science*. 494–503.

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.

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.

Mislove, A.; Viswanath, B.; Gummadi, K. P.; and Druschel, P. 2010. You are who you know: Inferring user profiles in online social networks. WSDM '10, 251–260.

Nguyen, N. T.; Kowalczyk, R.; and Chen, S.-M. 2009. Computational collective intelligence. semantic web, social networks and multiagent systems. *LNCS* 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.

Perliger, A., and Pedahzur, A. 2011. Social network analysis in the study of terrorism and political violence. *PS: Political Science & Politics* 44(01):45–50.

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.

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.

Zheleva, E., and Getoor, L. 2009. To join or not to join: The illusion of privacy in social networks with mixed public and private user profiles. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, 531–540. New York, NY, USA: ACM.