

Topic 1: Evading Social Network Analysis Tools

Tomasz P. Michalak
University of Warsaw & University of Oxford
tomasz.michalak@cs.ox.ac.uk

November 20, 2016

Many research problems in social network analysis (SNA) have received considerable attention in recent years across various disciplines, including Mutli-Agent Systems [5, 3] and, more generally, Artificial Intelligence. Scientists, developers, and users have focused on improving performance of various SNA tools, such as *centrality* measures, *community-detection* algorithms [4] or *link-prediction* algorithms [1], 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.

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.

Our three works already address the above question for three key social network analysis tools:

1. **Centrality analysis tools** — in the works by Waniek et al. [7, 8], we answer the question: *how individuals might proactively 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 would not loose their influence in the network?* We considered three most fundamental centrality measures: degree, closeness, and betweenness.
- **Community detection algorithms** — in the work by Waniek et al. [7], we answer the question: *how communities might proactively manage their social connections so that they are less exposed to the workings of community detection algorithms?*
- **Link prediction algorithms** — in the work by Waniek et al. [6], *given a “seeker” who is running a link prediction algorithms, and “evaders” who want to hide some of their connections, we study how the evaders can make those connections harder for the seeker to identify.*

All the above topics belong to a much wider research agenda recently outlined in the forthcoming Senior Member Blue Sky paper at AAAI 2017 [2]. Some of the **open research topics** include:

- **Evading other centrality measures** — while in the works by Waniek et al. [7, 8] we studied how to evade three fundamental centrality measures there are other important centrality measures that should be considered, e.g. eigenvector centrality or Katz centrality.
- **Evading influence maximization algorithms** — the influence maximization problem involves identifying a set of nodes of a given size that has the highest expected influence over the network. The question is *how a comparatively very influential node may evade being chosen to be a part of this set of nodes.*
- **Evading network topology measures** — five most important measures of network topology are centralization, density, average degree, average path length and network diameter. *How difficult is it to modify the network so the measures of network topology become significantly different?*

Note: all unpublished papers referenced in this note are available on request. Please contact me at tomasz.michalak@cs.ox.ac.uk.

References

- [1] Lise Getoor and Christopher P Diehl. Link mining: a survey. *ACM SIGKDD Explorations Newsletter*, 7(2):3–12, 2005.
- [2] Tomasz Michalak, Talal Rahwan, and Michael Wooldridge. Strategic social network analysis. In *Senior Member Blue Sky Papers, forthcoming in Proceedings of 31th National Conference on Artificial Intelligence (AAAI)*, 2017.
- [3] Ngoc Thanh Nguyen, Ryszard Kowalczyk, and Shyi-Ming Chen. Computational collective intelligence. semantic web, social networks and multiagent systems. *Lecture Notes in Computer Science*, 5796, 2009.
- [4] Günce Keziban Orman and Vincent Labatut. A comparison of community detection algorithms on artificial networks. In *Discovery science*, pages 242–256. Springer, 2009.
- [5] Jordi Sabater and Carles Sierra. 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*, pages 475–482. ACM, 2002.
- [6] Marcin Waniek, Talal Rahwan, and Tomasz Michalak. Hiding Relationships in a Social Network. mimeo, University of Oxford, available on request, 2016.
- [7] Marcin Waniek, Talal Rahwan, Tomasz Michalak, and Michael Wooldridge. Hiding Individuals and Communities in a Social Network. <https://arxiv.org/abs/1608.00375>, 2016.
- [8] Marcin Waniek, Talal Rahwan, Tomasz Michalak, and Michael Wooldridge. On the Construction of Covert Network. mimeo, University of Oxford, available on request, 2016.