Strategic Evasion of Centrality Measures

Marcin Waniek
New York University Abu Dhabi
and University of Warsaw
mjwaniek@nyu.edu

Jan Woźnica
University of Warsaw
jan.jakub.woznica@gmail.com

Kai Zhou
Hong Kong Polytechnic University
kaizhou@polyu.edu.hk

Yevgeniy Vorobeychik
Washington University in St. Louis
yvorobeychik@wustl.edu

Talal Rahwan
New York University Abu Dhabi
talal.rahwan@nyu.edu

Tomasz P. Michalak
University of Warsaw
tpm@mimuw.edu.pl

ABSTRACT

Among the most fundamental tools for social network analysis are centrality measures, which quantify the importance of every node in the network. This centrality analysis typically disregards the possibility that the network may have been deliberately manipulated to mislead the analysis. To solve this problem, a recent study attempted to understand how a member of a social network could rewire the connections therein to avoid being identified as a leader of that network. However, the study was based on the assumption that the network analyzer—the seeker—is oblivious to any evasion attempts by the evader. In this paper, we relax this assumption by modelling the seeker and evader as strategic players in a Bayesian Stackelberg game. In this context, we study the complexity of various optimization problems, and analyze the equilibria of the game under different assumptions, thereby drawing the first conclusions in the literature regarding which centralities the seeker should use to maximize the chances of detecting a strategic evader.

KEYWORDS

Social Network; Centrality Measure; Computational Complexity; Stackelberg Game

ACM Reference Format:
Marcin Waniek, Jan Woźnica, Kai Zhou, Yevgeniy Vorobeychik, Talal Rahwan, and Tomasz P. Michalak. 2021. Strategic Evasion of Centrality Measures. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 9 pages.

1 INTRODUCTION

Social network analysis tools have attracted significant attention in the literature [6, 9, 13]. Such tools are typically used under an assumption that the members of the network are not strategic, i.e., they do not manipulate the topology of the network to their advantage. However, as argued by Michalak et al. [15], this assumption does not hold in many situations, ranging from privacy-savvy users of social media platforms [14], through political activists [23], to the members of criminal and terrorist organizations whose primary concern is to evade attention of security agencies [12].

An attempt to fill this gap in the literature was carried out by Waniek et al. [20], who considered how one could evade popular centrality measures, such as degree, closeness, and betweenness. More specifically, the authors studied how a member of the network—called the *evader*—can rewire the network (by adding or removing edges) in order to optimally decrease the value of her centrality while maintaining her influence over other members of the network. The authors proved that, even without taking influence into consideration, the problem of decreasing the value of either closeness or betweenness centrality is NP-complete, while for the degree centrality the problem is in P. Nevertheless, the study has a number of limitations. Firstly, in their complexity analysis, the authors considered the problem of decreasing the *value* of the evader's centrality, which is insufficient if the evader is concerned with decreasing her *position* in the centrality-based ranking of all nodes, i.e., decreasing her centrality *relative* to that of other nodes in the network. Secondly, the complexity analysis assumed that the evader is able to add and remove edges in the *entire network*. This seems unrealistic in many settings such as social media platforms, where members are unable to view, let alone modify, any edge in the network. Finally, the authors assumed that the party using the social network analysis tools—the *seeker*—is *not strategic*, i.e., she is unaware of the evasion efforts made by the evader. While this assumption may hold in some settings, there are many others in which the seeker expects the evader to go to great lengths in order to mislead any analysis, as is the case with covert networks.

In this paper, we address all of the above limitations, and present the first analysis of evading centrality measures in settings where *both parties act strategically*. We start by analyzing the complexity of decreasing the evader's *position* in the centrality-based ranking, as opposed to decreasing the *value* of the evader's centrality. More specifically, we require that the evader decreases her ranking by at least d positions, and allow the evader to add or remove edges only *locally*, i.e., in her immediate neighbourhood. We prove that this problem is NP-complete not only for closeness and betweenness centralities but also for degree centrality. Table 1 presents the main theoretical contributions of this paper.

We then model the interaction between the seeker and the evader as a *Bayesian Stackelberg game* [8, 10, 17], whereby the strategy set of the seeker consists of degree, closeness, betweenness, and eigenvector centralities, while the strategy set of the evader consists of all possible sets of changes in her network neighbourhood. Our extensive experimental analysis of this game draws the first conclusions in the literature regarding which centralities the seeker should use to maximize the chances of detecting a strategic evader.

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.

Centrality	Disguising Centrality [20]	Hiding Leader [19]	Local Hiding (this paper)
Degree	P	NP-complete	**NP-complete**
Closeness	NP-complete	NP-complete	**NP-complete**
Betweenness	NP-complete	unknown	**NP-complete**

Table 1: Comparing our complexity results to the literature.

2 PRELIMINARIES

Let $G = (V, E) \in \mathbb{G}$ denote a network, where V is the set of n nodes and $E \subseteq V \times V$ is the set of edges, and let $\mathbb{G}(V)$ denote the set of all possible networks whose set of nodes is V. We denote by (v, w) the edge between nodes v and w. We restrict our attention to undirected networks, and thus we do not discern between edges (v, w) and (w, v). We also assume that networks do not contain self-loops, i.e., $\forall_{v \in V} (v, v) \notin E$. We denote by $N(v)$ the set of neighbours of v, i.e., $N(v) = \{w \in V : (v, w) \in E\}$.

A *path* in (V, E) is an ordered sequence of nodes, $p = \langle v_1, \dots, v_k \rangle$, in which every two consecutive nodes are connected by an edge in E. The length of a path equals the number of edges therein. For any pair of nodes, $v, w \in V$, we denote by $\Pi(v, w)$ the set of all shortest paths between these two nodes, and denote by $d(v, w)$ the *distance* between the two, i.e., the length of a shortest path between them.

A *centrality measure* is a function, $c\colon \mathbb{G}(V) \times V \to \mathbb{R}$, that expresses the importance of any given node in the network [2]. We consider four fundamental centrality measures, namely degree, closeness, betweenness, and eigenvector.

Degree centrality [18] of node v is proportional to its degree: $c_{degr}(G, v) = |N(v)|$. *Closeness centrality* [3] assigns the highest importance to the node with the shortest average distance to all other nodes: $c_{clos}(G, v) = \frac{1}{\sum_{w \in V} d(v,w)}$. *Betweenness centrality* [1, 7] of node v is proportional to the percentage of shortest paths between every pair of other nodes that go through v: $c_{betw}(G, v) = \sum_{w \neq w' \neq v} \frac{|\{p \in \Pi(w,w') : v \in p\}|}{|\Pi(w,w')|}$. *Eigenvector centrality* [4] evaluates each node based on the importance of its neighbours. Formally, $c_{eig}(G, v) = x_v$, where x is the eigenvector corresponding to the largest eigenvalue of the adjacency matrix of G.

We consider two influence models: *independent cascade* and *linear threshold*. Both models can be described in terms of spreading the "activation" of nodes across the network. The process starts with an *active* subset of nodes called the seed set. The activation then propagates through the network in discrete time steps, whereby nodes become influenced by their previously-activated neighbours.

Formally, let $I(t)$ denote the set of nodes that are active at round t, with $I(1)$ being the seed set. In the independent cascade model, an activation probability $p : V \times V \to \mathbb{R}$ is assigned to each pair of nodes. For every round $t > 1$ each node that became active in round $t - 1$ has a single chance to activate each of her inactive neighbours w with probability $p(v, w)$. In our experiments we assume that for every pair of nodes, v, w, we have: $p(v, w) = 0.15$. As for the linear threshold model, every node, v, is assigned a threshold, t_v, sampled from the set: $\{0, \dots, |N(v)|\}$. Then, in every round $t > 1$, each inactive node becomes activated if $|I(t-1) \cap N(v)| \geq t_v$. In our experiments, the threshold of a node, v, is sampled from the set $\{1, \dots, |N(v)|\}$ uniformly at random. Notice that this variant is slightly different than the standard linear threshold model [11], in which edges are assigned random weights. We use this variant to stay consistent with the previous literature on the topic [19, 20].

In both models, the process ends when there are no new active nodes, i.e., when $I(t-1) = I(t)$. The influence of v is then measured as the expected number of active nodes at the end of the process, when starting with $\{v\}$ as the seed set. Computing the exact influence requires exponential computations under both models, which is intractable even for relatively small networks. Thus, in our experiments we approximate the influence using Monte Carlo sampling, stopping the process when the improvement over the last $1,000$ iterations is smaller than 0.00001. Note that even approximating the influence of a node becomes challenging when the number of nodes reaches thousands or more.

3 COMPLEXITY OF LOCAL HIDING

We now formally define the main computational problem of our study, and analyze its computational complexity.

Definition 1 (Local Hiding). *This problem is defined by a tuple $(G, v_e, b, c, \hat{A}, \hat{R}, d)$, where $G = (V, E)$ is a network, $v_e \in V$ is the evader, $b \in \mathbb{N}$ is a budget specifying the maximum number of edges that can be added or removed, $c\colon \mathbb{G}(V) \times V \to \mathbb{R}$ is a centrality measure, $\hat{A} \subseteq N(v_e) \times N(v_e)$ is the set of edges allowed to be added, $\hat{R} \subseteq \{v_e\} \times N(v_e)$ is the set of edges allowed to be removed, and $d \in \mathbb{N}$ is the safety margin. The goal is to identify a set of edges to be added, $A^* \subseteq \hat{A}$, and a set of edges to be removed, $R^* \subseteq \hat{R}$, such that $|A^*| + |R^*| \leq b$ and the resulting network $(V, (E \cup A^*) \setminus R^*)$ contains at least d nodes with centrality c greater than that of the evader.*

As mentioned in the introduction, the two key differences between the above problem of *Local Hiding* and the problem of *Disguising Centrality* studied by Waniek et al. [20] are as follows. Firstly, instead of seeking the optimal way of decreasing the value of the evader's centrality (which may not provide sufficient cover, especially if she is still ranked among the top nodes in the network), we want the position of the evader in the centrality-based ranking of all nodes to drop below d. Secondly, we assume that the evader is only capable of rewiring edges within her network neighbourhood—an assumption that holds in many realistic settings, e.g., the evader is able to disconnect herself from any of her friends, or even ask two of them to befriend one another, but is unable to connect to a complete stranger at will, or ask two strangers to befriend or unfriend one another. Notice that we do not allow to add any edges incident to the evader, as in case of most centrality measures such operation can only increase the ranking of the evader.

We also comment on the key differences between our Local Hiding problem and the problem of *Hiding Leaders* studied by Waniek et al. [19] in the context of constructing covert networks. Firstly, the authors divide the nodes into leaders and the followers, where the changes in the network are allowed only among the followers. Secondly, they only allow edges to be *added* among the followers, meaning that no edge can be removed from the network.

Theorem 1. *The problem of Local Hiding is NP-complete given the degree centrality measure.*

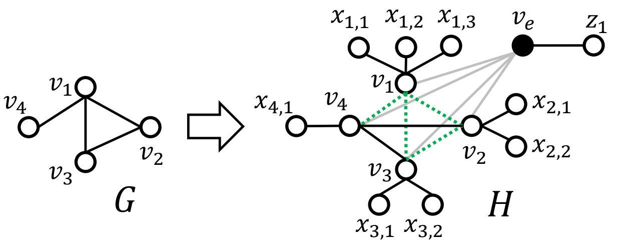


Figure 1: Network used in the proof of Theorem 1 for $k = 3$.

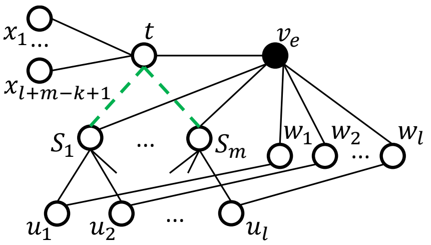


Figure 2: Network used in the proof of Theorem 2.

Proof. The problem is trivially in NP, since after the addition of a given set of edges A^* and the removal of a given set of edges R^* it is possible to compute the degree centrality of all nodes in polynomial time. Next, we prove that the problem is NP-hard. To this end, we give a reduction from the NP-complete problem of Finding k-Clique, where the goal is to determine whether there exist k nodes in G that form a clique. Given an instance of the problem of Finding k-Clique, defined by $k \in \mathbb{N}$ and a network $G = (V, E)$, let us construct a network, $H = (V', E')$, as follows:

- $V' = \{v_e\} \cup V \cup \bigcup_{v_i \in V} \bigcup_{j=1}^{|N(v_i)|} \{x_{i,j}\} \cup \bigcup_{i=1}^{k-2} \{z_i\}$,
- $E' = \bigcup_{v_i \in V'} \{(v_i, v_e)\} \cup \bigcup_{x_{i,j} \in V'} \{(v_i, x_{i,j})\} \cup \bigcup_{z_i \in V'} \{(z_i, v_e)\} \cup \bigcup_{(v_i,v_j) \notin E} \{(v_i, v_j)\}$.

An example of such a network H is illustrated in Figure 1. Now, consider the instance $(H, v_e, b, c, d, \hat{A}, \hat{R})$ of the problem of Local Hiding where $H = (V', E')$ is the network we just constructed, v_e is the evader, $b = \frac{k(k-1)}{2}$, c is the degree centrality measure, $d = k$, $\hat{A} = E$, and $\hat{R} = \emptyset$.

From the definition of the problem we know that the edges to be added to H must be chosen from E, i.e., from the network in the Finding k-Clique problem. Out of those edges, we need to choose a subset, $A^* \subseteq E$, as a solution to the Local Hiding problem. In what follows, we will show that a solution to the above instance of the Local Hiding problem in H corresponds to a solution to the problem of Finding k-Clique in G.

First, note that v_e has the highest degree in H, which is $n + k - 2$. Thus, in order for A^* to be a solution to the Local Hiding problem, the addition of A^* to H must increase the degree of at least k nodes in V such that each of them has a degree of at least $n+k-1$ (note that the addition of A^* only increases the degrees of nodes in V, since we already established that $A^* \subseteq E$). Now since in H the degree of every node v_i equals n (because of the way H is constructed), then in order to increase the degree of k such nodes to $n + k - 1$, each of them must be an end of at least $k - 1$ edges in A^*. But since the budget in our problem instance is $\frac{k(k-1)}{2}$, then the only possible choice of A^* is the one that increases the degree of exactly k nodes in V by exactly $k - 1$. If such a choice of A^* is available, then surely those k nodes form a clique in G, since all edges in A^* are taken from G. □

Theorem 2. *The problem of Local Hiding is NP-complete given the closeness centrality measure.*

Proof. The problem is trivially in NP, since after the addition of a given A^*, and the removal of a given R^*, it is possible to compute the closeness centrality of all nodes in polynomial time. Next, we prove that the problem is NP-hard. To this end, we propose a reduction from the NP-complete 3-Set Cover problem. Let $U = \{u_1, \ldots, u_l\}$ be the universe, and let $S = \{S_1, \ldots, S_m\}$ be the set of subsets of the universe, where for every S_i we have $|S_i| = 3$. The goal is then to determine whether there exist k elements of S the union of which equals U. Given an instance of the 3-Set Cover problem, let us construct a network, $G = (V, E)$, as follows:

- $V = \{v_e, t\} \cup \bigcup_{S_i \in S} \{S_i\} \cup \bigcup_{u_i \in U} \{u_i, w_i\} \cup \bigcup_{i=1}^{l+m-k+1} \{x_i\}$,
- $E = \{(t, v_e)\} \cup \bigcup_{x_i \in V} \{(x_i, t)\} \cup \bigcup_{w_i \in V} \{(w_i, v_e), (w_i, u_i)\} \cup \bigcup_{S_i \in V} \{(S_i, v_e)\} \cup \bigcup_{u_j \in S_i} \{(S_i, u_j)\}$.

An example of the resulting network, G, is illustrated in Figure 2. Now, consider the following instance of the problem of Local Hiding, $(G, v_e, b, c, \hat{A}, \hat{R}, d)$, where G is the network we just constructed, v_e is the evader, $b = k$ (where k is the parameter of the 3-Set Cover problem), c is the closeness centrality measure, $d = 1$, $\hat{A} = \{(t, S_i) : S_i \in S\}$, and $\hat{R} = \emptyset$.

From the definition of the problem, we see that the only edges that can be added to the graph are those between t and the members of S. Notice that any such choice of A^* corresponds to selecting a subset of $|A^*|$ elements of S in the 3-Set Cover problem. In what follows, we will show that a solution to the above instance of Local Hiding corresponds to a solution to the 3-Set Cover problem.

First, we will show that for every $v \in V \setminus \{t, v_e\}$ and every $A^* \subseteq \hat{A}$ we either have $c(G', v) < c(G', t)$ or have $c(G', v) < c(G', v_e)$, where $G' = (V, E \cup A^*)$. To this end, let $D(G', v)$ denote the sum of distances from v to all other nodes, i.e., $D(G', v) = \sum_{w \in V \setminus \{v\}} d(v, w)$. Note that $D(G', v) = \frac{n-1}{c(G',v)}$. We will show that the following holds:

$$\forall_{v \in V \setminus \{t, v_e\}} \forall_{A^* \subseteq \hat{A}} \left(D(G', v) > D(G', t) \vee D(G', v) > D(G', v_e) \right).$$

Let d_t denote $\sum_{u_i \in U} d(t, u_i) + \sum_{S_i \in S} d(t, S_i)$. Notice also that $k \leq m$. Next, we compute $D(G', v)$ for the different types of node v:

- $D(G', v_e) = 5l + 3m - 2k + 3$;
- $D(G', t) = 3l + m - k + 2 + d_t$;
- $D(G', x_i) = 6l + 3m - 2k + 3 + d_t > D(G', t)$;
- $D(G', w_i) = 8l + 5m - 3k + 2 > D(G', v_e)$;
- $D(G', u_i) \geq 9l + 4m - 3k + 2 > D(G', v_e)$ as $\sum_{S_j \in S} d(u_i, S_j) \geq m$;
- $D(G', S_i) \geq 7l + 4m - 2k - 4 > D(G', v_e)$ as $d(S_i, v_e) \geq 1$.

Based on this, either t or v_e has the highest closeness centrality, therefore $A^* \subseteq \hat{A}$ is a solution to the problem of Local Hiding if and only if $D(G', t) < D(G', v_e)$. This is the case when $d_t < 2l + 2m - k + 1$. Let $U_A = \{u_i \in U : \exists_{S_j \in S} u_i \in S_j \wedge (t, S_j) \in A^*\}$. We have that $d_t = |A^*| + 2(m - |A^*|) + 2|U_A| + 3(l - |U_A|)$ which gives us $d_t = 3l - |U_A| + 2m - |A^*|$. Since by definition $|U_A| \leq l$ and $|A^*| \leq k$, it is possible that $d_t < 2l + 2m - k + 1$ only when $|U_A| = l$ and $|A^*| = k$, i.e., $\forall_{u_i \in U} \exists_{S_j \in S} u_i \in S_j \wedge (t, S_j) \in A^*$. This solution to the problem of Local Hiding corresponds to a solution to the given instance of the 3-Set Cover problem, which concludes the proof. □

Theorem 3. *The problem of Local Hiding is NP-complete given the betweenness centrality measure.*

The proof can be found in Waniek et al. [21].

4 THE SEEKER-EVADER GAME

Player strategies: We model the problem of strategically hiding in a network as a game between two players: the *evader* and the *seeker*. In particular, the seeker analyzes the network using a set of strategies, T_s, consisting of the fundamental centrality measures: degree, closeness, betweenness, and eigenvector. On the other hand, the goal of the evader is to decrease her position in the centrality-based ranking of all nodes, while maintaining her influence within the network (notice that the theoretical problems presented in Section 3 are focused on providing safety to the evader by lowering her ranking position, while here we additionally allow the evader to take into consideration her influence in the network). To this end, she utilizes a set of strategies, T_e, consisting of combinations of edge modifications in her neighbourhood, with the maximum number of permitted modifications being specified by a budget, b.

In our experiments, we pay particular attention to the only available evader strategy in the literature, namely ROAM (Remove One Add Many) [20]. In particular, the ROAM heuristic involves two steps. *Step 1:* Remove the edge between the evader, v_e, and its neighbour of choice, v_0; *Step 2:* Connect v_0 to $b - 1$ nodes of choice, who are neighbours of v_e but not of v_0. This simple heuristic has been shown to be rather effective in practice.

Utility functions: For any given pair of strategies, (t_s, t_e), such that $t_s \in T_s$ and $t_e \in T_e$, the utility of the evader is:

$$U_e(\phi, t_s, t_e) = \phi U_e^R(t_s, t_e) + (1 - \phi) U_e^I(t_e)$$

where:

- $U_e^R(t_s, t_e) \in \mathbb{R}$ is the evader's utility from the change in her rank according to the centrality measure t_s chosen by the seeker, when the evader plays strategy t_e,
- $U_e^I(t_e) \in \mathbb{R}$ is the evader's utility from the change in her *influence* within the network when she plays strategy t_e,
- $\phi \in \left\{\frac{1}{m+1}, \ldots, \frac{m}{m+1}\right\}$ represents the evader's evaluation of $U_e^R(t_s, t_e)$ relative to $U_e^I(t_e)$, we will refer to ϕ as the *type* of the evader, with m being the number of types.

Next, we specify how $U_e^R(t_s, t_e)$ and $U_e^I(t_e)$ are calculated (Figure 3 depicts both functions). Let $r_e(t_s, t_e)$ be the evader's ranking when she plays strategy t_e and the seeker plays strategy t_s. Then, $U_e^R(t_s, t_e)$ is calculated as follows:

$$U_e^R(t_s, t_e) = \frac{1}{\alpha\left(1 + e^{-k(r_e(t_s, t_e) - d)}\right)} - \frac{\beta}{\alpha},$$

where e is Euler's number, k is the curve steepness, d is the inflection point, $\beta = \frac{1}{1 + e^{-k(1-d)}}$ and $\alpha = (1 - 2\beta)$. This formula has the following desirable properties:

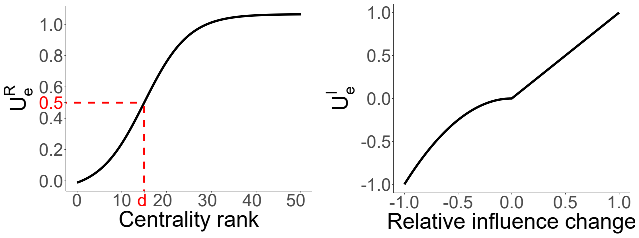


Figure 3: The evader's utility functions for $d = 15$ and $k = \frac{3}{d}$.

- The evader's utility is 0 when ranked first, i.e., fully exposed. Formally, $U_e^R(t_s, t_e) = 0$ when $r_e(v) = 1$.
- The evader's utility increases when she becomes more hidden. Formally, $U_e^R(t_s, t_e)$ increases with $r_e(t_s, t_e)$.
- $U_e^R(t_s, t_e)$ is *convex* for $1 \leq r_e(t_s, t_e) \leq d$, meaning that the marginal gain in utility increases with ranking drop, as long as the evader does not reach position d.
- $U_e^R(t_s, t_e)$ is *concave* for $d \leq r_e(t_s, t_e) \leq n$, i.e., dropping beyond position d produces diminishing returns to the evader.

Finally, note that $U_e^R(t_s, t_e) \longrightarrow 1 + \frac{\beta}{\alpha}$ when $r_e(t_s, t_e) \longrightarrow n$. Having specified how $U_e^R(t_s, t_e)$ is calculated, we now move to $U_e^I(t_e)$. Recall that the evader's influence is measured according to either the *independent cascade* model or the *linear threshold* model [3, 18]. Regardless of which model is used, let $\Delta_e(t_e)$ denote the relative change in the evader's influence when she plays strategy t_e, i.e., $\Delta_e(t_e) = (I_e(t_e) - I_e^0)/I_e^0$, where $I_e(t_e)$ is the evader's influence when she plays strategy t_e, and I_e^0 is the evader's initial influence before playing. Then, $U_e^I(t_e)$ is calculated as follows:

$$U_e^I(t_e) = \begin{cases} \Delta_e(t_e), & \text{if } \Delta_e(t_e) > 0 \\ -\Delta_e(t_e)^2, & \text{if } \Delta_e(t_e) \leq 0 \end{cases}$$

This formula has some desired properties. Firstly, $U_e^I(t_e)$ is concave when $\Delta_e(t_e) \leq 0$, meaning that the marginal loss in utility grows with the loss in influence (this is intuitive in scenarios where the evader does not mind a negligible drop in influence in return for a better disguise, but strongly opposes a significant drop in influence). Secondly, when $\Delta_e(t_e) \geq -1$, we have $U_e^I(t_e) \geq -1$, and as $\Delta_e(t_e)$ increases, $U_e^I(t_e)$ reaches a similar order of magnitude as that of $U_e^R(t_s, t_e)$, meaning that the equilibrium is not dominated by any of those two utilities.

Let us now turn our attention to the utility of the seeker. In our analysis we consider two different versions of the game: *zero-sum game* and *non-zero-sum game*. In the zero-sum version of the game we assume that the seeker is interested in minimizing the total utility of the evader, i.e., the seeker's utility is $U_s = -U_e$. In the non-zero-sum version in the game we assume that the seeker is interested solely in identifying the evader, i.e., the seeker's utility is $U_s = -U_e^R$. Notice that in the latter version of the game the seeker completely disregards any utility that the evader might gain from the change in her influence. We assume that the payoffs and the distribution of evader types are common knowledge, while the actual evader's type is private.

Network	Network size	All strategies	Undominated strategies
WTC	36	14190	60
Bali	17	280840	7
Madrid	70	45760	5
Scale-Free	30	61365	17
Small-World	30	902	36
Erdos-Renyi	30	4122	47

Table 2: The number of possible strategies vs. the number of undominated strategies (for random networks, the number is taken as the average over 100 such networks).

The Stackelberg game: Our model allows for mixed strategies. More specifically, let $p_s(t_s)$ be the probability that the seeker plays pure strategy $t_s \in T_s$. Moreover, let $p(\phi)$ be the probability that the evader type is ϕ, and let $p_e^{\phi}(t_e)$ be the probability that an evader of type ϕ plays pure strategy $t_e \in T_e$. Now since the evader moves second, i.e., she knows the strategy of the seeker, then we can restrict her available strategies to only pure ones. Hence, the probability that an evader of type ϕ plays pure strategy $t_e \in T_e$ is $p_e^{\phi}(t_e) \in \{0, 1\}$. The seeker's objective is to maximize her expected payoff. This optimization problem can be formulated as a Mixed-Integer Quadratic problem:

$$\begin{aligned} \max \quad & \sum_{\phi \in \Phi} \sum_{t_s \in T_s} \sum_{t_e \in T_e} p(\phi) p_e^{\phi}(t_e) p_s(t_s) U_s(\phi, t_s, t_e) \\ \text{s.t.} \quad & \sum_{t_s \in T_s} p_s(t_s) = 1 \\ & \sum_{t_e \in T_e} p_e^{\phi}(t_e) = 1 \\ & \lambda \geq \sum_{t_s \in T_s} p_s(t_s) U_e(\phi, t_s, t_e) \\ & \lambda \leq (1 - p_e^{\phi}(t_e))\eta + \sum_{i \in T_s} p_s(t_s) U_e(\phi, t_s, t_e) \end{aligned}$$

The first and second constraints correspond to the probability distributions over the sets of strategies available to the players. As for $\eta \in \mathbb{R}$, it is an arbitrarily large number. This way, the third and fourth constraints ensure that, by solving the problem, we get:

$$\lambda = \max_{t_e \in T_e} \sum_{t_s \in T_s} p_s(t_s) U_e(\phi, t_s, t_e).$$

This is because, when η is arbitrarily large, $(1 - p_e^{\phi}(t_e))\eta$ reflects the fact that the evader will play the strategy that maximizes her expected payoff. Finally, in order to solve the problem efficiently, we linearize it by substituting variables: $z^{\phi}(t_s, t_e) = p_e^{\phi}(t_e) p_s(t_s)$. We use the linearization procedure described by Paruchuri et al. [16].

5 EMPIRICAL ANALYSIS

For each network, following the work by Waniek et al. [20], the evader is chosen as the node with the smallest sum of centrality ranks (based on Degree, Closeness, Betweenness and Eigenvector); ties are broken uniformly at random. The evader type ϕ is sampled uniformly at random from the set $\{0.2, 0.4, 0.6, 0.8\}$. See Waniek et al. [21] for the description of the network datasets we consider in our simulations. All results for random networks are presented as an average over 100 samples.

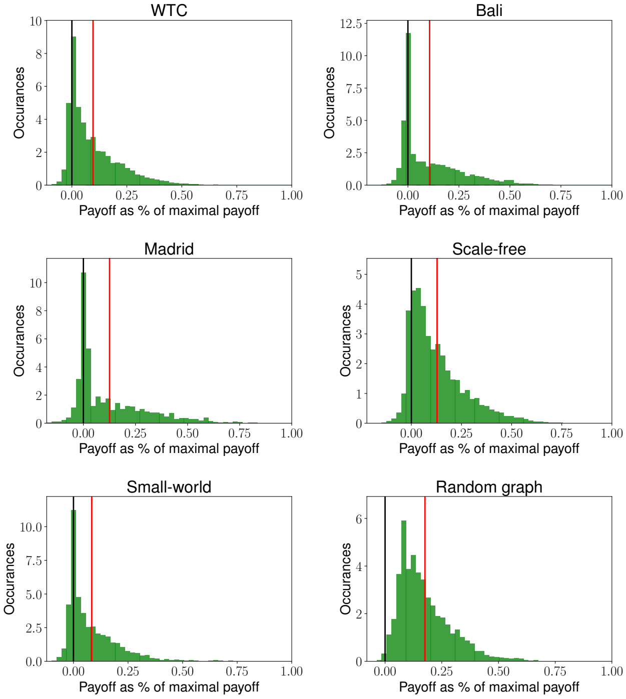


Figure 4: The distributions of the evader's payoffs for budget 3. Values are provided for evader type $\phi = 0.5$ and averaged over the seeker's equilibrium strategies. For each network, the red and black lines denote the average payoff and 0, respectively.

While the number of pure strategies of the seeker is rather small (we assume them to be the four main centrality measures), the number of pure strategies of the evader is much larger, since every possible way of rewiring the evader's neighbourhood may be considered a unique strategy. This very quickly becomes computationally challenging even for small networks and small budgets. For instance, in the case of the WTC network, the number of the evader's strategies for budget $b = 3$ is $14, 190$, for $b = 4$ it is $148, 995$, and for $b = 5$ it is $1, 221, 759$.

With this in mind, to study the evader's entire space of possible strategies, we focus first on a version of the game that is more computationally feasible. More specifically, we analyze the *zero-sum* version of the game, where the seeker's gain equals the evader's loss. This implies that the seeker is not only interested in the evader's centrality (as in the aforementioned model), but is also interested in the evader's influence (this is implied by the fact that the evader's utility does not only depend on her centrality but also on her influence). Importantly, this version of the game can be formulated as a linear program; hence, it is much easier to solve. By analyzing the zero-sum version of the model, we aim to

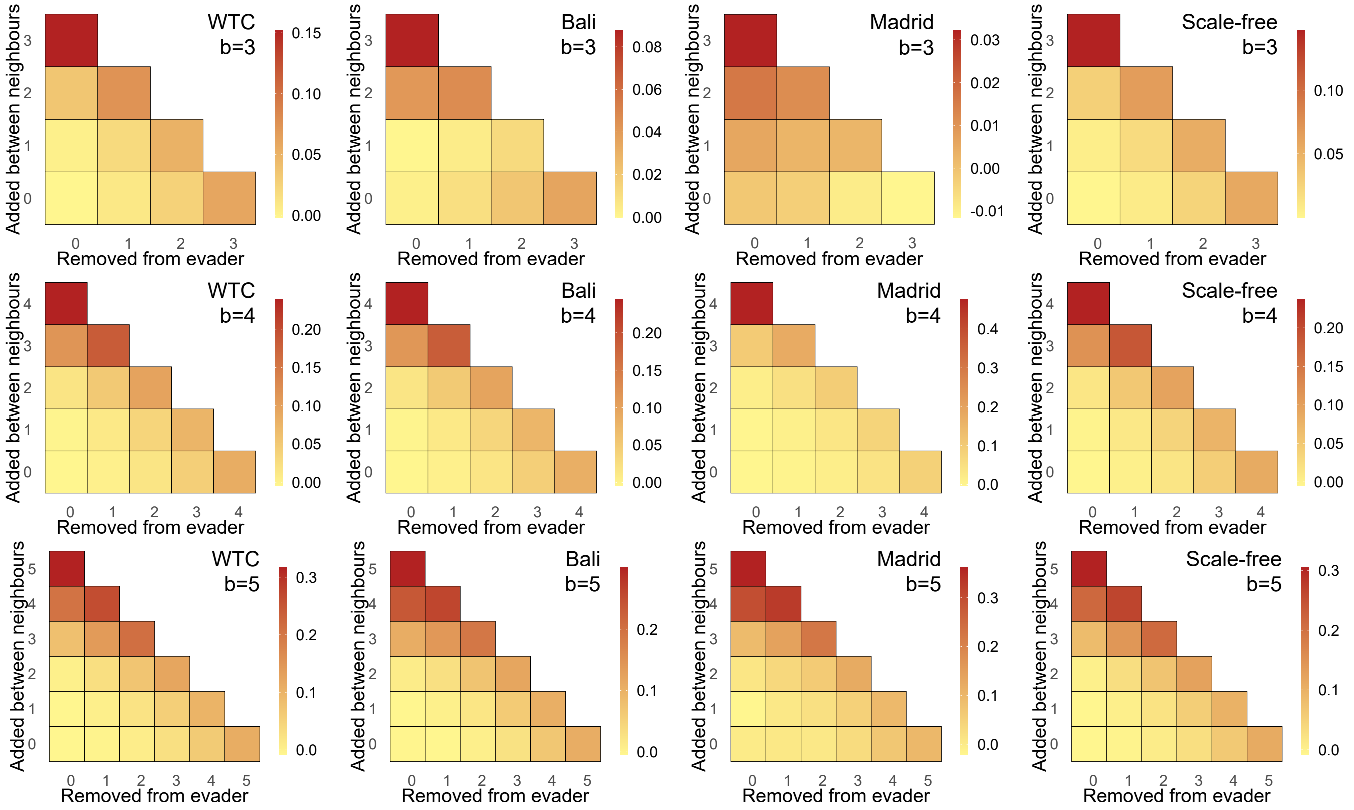


Figure 5: The evader's average payoff given the three terrorist networks (WTC, Bali, and Madrid), and scale-free networks, and given budgets 3, 4, and 5. The x-axis represents the number of neighbours the evader is disconnected from, while the y-axis represents the number of edges added between the evader's neighbours. The color intensity of each cell represents the evader's average payoff for given strategy.

understand the properties of the evader's most rewarding strategies. This understanding will help us identify effective heuristics for the evader, which in turn would enable us to study the original, more computationally-challenging version of the game.

5.1 The Zero-Sum Version

For each network we generated the payoff matrices corresponding to budgets 3 and 4 and both influence measures. We were also able to consider 10% of the strategies corresponding to budget 5 (except for the WTC network, where we considered 100%). Our main observations regarding the strategies are threefold.

Firstly, *most of the evader's strategies are strongly dominated, regardless of the evader's type.* Specifically, given different networks, Table 2 specifies the number of all strategies as well as those that are undominated. As shown, less than 1% of strategies are undominated, and this percentage is even smaller for larger networks.

Secondly, for any given equilibrium strategy of the evader, the difference in the seeker's payoff between her optimal strategy and other strategies is minimal (less than 1%). This suggests that, for the zero-sum game, the seeker could, in principle, use any centrality measure to analyse the network, without compromising much efficiency. Conversely, for any given equilibrium strategy of the seeker, the difference in the evader's payoff between her optimal strategy and other strategies is much more pronounced (more than 100%, see Figure 4). Hence, *the outcome of the game relies heavily on the evader's choice of strategy, while the seeker's choice of centrality measure has negligible impact.*

Thirdly, the strategies that yield similar payoffs seem to involve rewiring the network in similar ways; see Figures 5 and 6. Interestingly, *the ROAM heuristic of Waniek et al. [20] is often among the evader's most rewarding strategies.*

Based on these observations, we next analyze the non-zero-sum version of the game when the evader uses the ROAM heuristic.

5.2 The Non-Zero-Sum Version

In this version of the game, we assume that the evader's strategies are instances of the ROAM heuristic. More specifically, the evader's total budget b is used to repeatedly run ROAM. We write ROAM(x), where x is the number of added between the evader's neighbours. The budget of a single iteration is between 1 and $\frac{b}{2}$, i.e., there are at least two iterations. The evader repeatedly run ROAM, until the entire budget b is spent. For example, for $b = 10$, we have the following set of evader strategies: {ROAM(1) repeated 5 times,

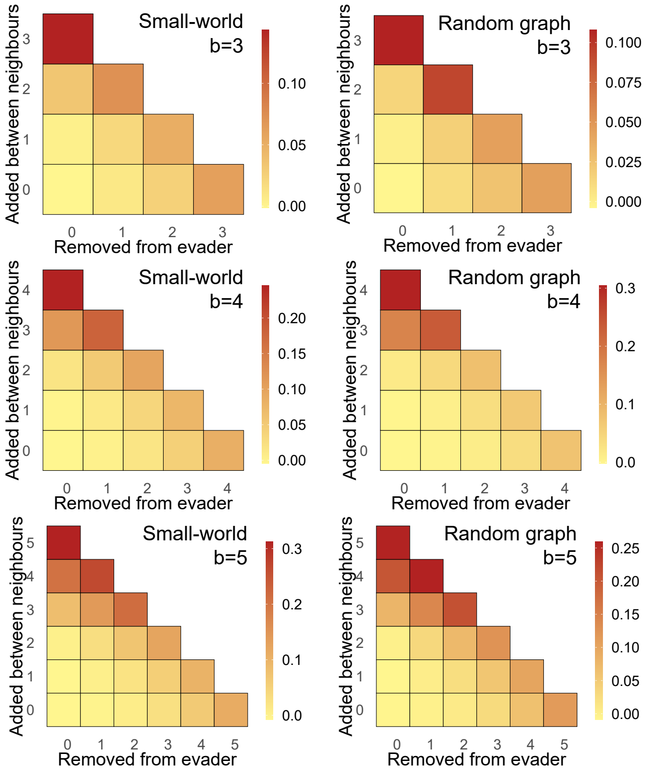


Figure 6: Same as Figure 5, but for small-world and random-graph networks.

ROAM(2) repeated 3 times + ROAM(0), ROAM(3) repeated twice + ROAM(1), ROAM(4) repeated twice}.

We calculate the equilibrium strategy profiles for different networks. For each network, we consider budgets $b \in \{5, 10, 15, 20, 25, 35\}$, assuming that b is no more than 25% of all edges in the network. This cap is meant to limit the changes in the network characteristics resulting from the evader's actions.

Figure 7 illustrate the mixed strategies played by the seeker in the equilibrium for different networks and evader budgets. For each centrality, Tables 3 and 4 present the average probability of being used in different networks.

The equilibrium strategies show, on one hand, which heuristics the evader should use to minimize her centrality while maintaining as much influence as possible. On the other hand, they indicate which centrality the seeker should adopt to have the greatest chance of identifying the evader among the top nodes in the network. Our first key observation in the non-zero-sum game setting is that the choice of the strategy by the seeker has a much greater impact on her payoff than in the zero-sum game. Hence, in what follows, we will focus particularly on the strategies of the seeker, i.e., we will consider which centrality a network analyzer should use when facing a strategic evader.

Regarding the results for the randomly-generated networks, we observe clear, robust patterns, suggesting that it is possible to identify some combination(s) of centrality measures that can be used against the evader. In particular:

Network	c_{betw}	c_{clos}	c_{degr}	c_{eig}
Scale-free	0	0.04	0.94	0.04
Random graphs	0.05	0.08	0.25	0.62
Small-world	0	0	0.06	0.94

Table 3: The average probability of using each centrality given randomly-generated networks.

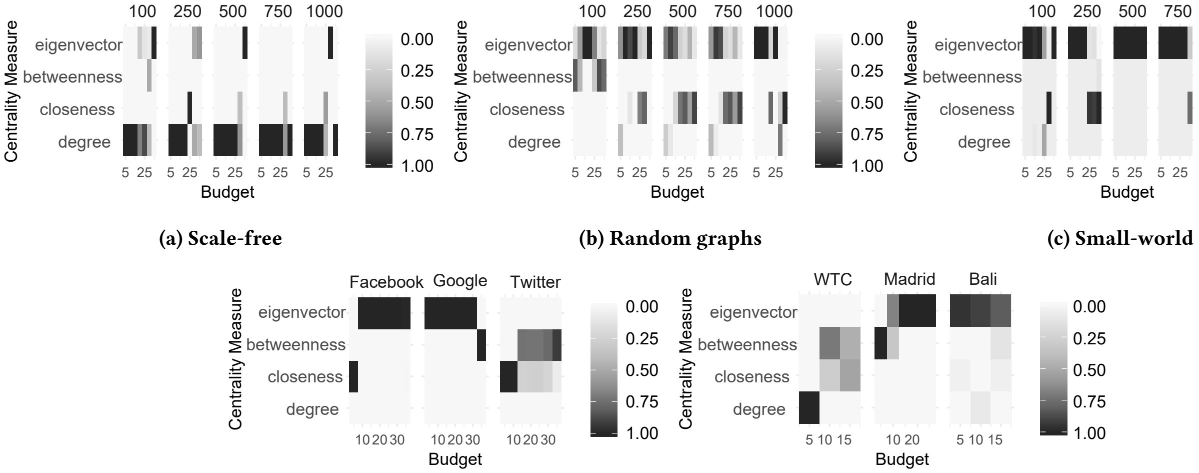
Network	c_{betw}	c_{clos}	c_{degr}	c_{eig}
WTC	0.04	0.03	0.03	0.89
Bali	0.39	0.27	0.33	0
Madrid	0.27	0	0	0.73
Overall Terrorist	0.23	0.10	0.12	0.54
Facebook	0	0.14	0	0.86
Google+	0	0.14	0	0.86
Twitter	0	0.56	0.44	0
Overall Social	0	0.28	0.15	0.57

Table 4: The average probability of using each centrality given different real-life networks.

- *Scale-free networks:* degree centrality is used almost exclusively. Due to the power-law distribution of nodes' degrees in scale-free networks, the "hubs" have extremely high degree, and the evader is most certainly one of them. As such, even with a large budget, any attempts to reduce the evader's position in the degree-based ranking have limited impact.
- *Small-world networks:* eigenvector centrality consistently proves to be most difficult to manipulate, it is played by the seeker in almost every small-world network.
- *Random graph networks:* For low values of the evader's budget, eigenvector centrality is the most effective. However, for larger budgets, it is often replaced by closeness centrality. This shift occurs when budget reaches about 15, regardless of the network size.

Regarding the results for the real-life networks, we also find regularities. Overall, for the networks with lower average clustering coefficient and lower density (Madrid and WTC attacks, Facebook, Google+), eigenvector centrality seems to be played most often. Furthermore, degree centrality is never played against the evader in larger networks. In more detail:

- *Covert organizations*: for the WTC 9/11 attack and the Madrid train attack networks, eigenvector centrality is played almost exclusively. On the other hand, for the Bali attack network, degree and betweenness centralities are chosen. This last network, in addition to being the smallest, consists of two subnetworks connected by one node—Samudra—the leader of the terrorist organization. This atypical topology of the network may be responsible for the difference. Moreover, the average clustering coefficient and the density for the Bali network are much greater than for the other networks.



(a) Scale-free

(b) Random graphs

(c) Small-world

(d) Social media

(e) Terrorist networks

Figure 7: The seeker's equilibrium strategies given the evader types $\{0.2, 0.4, 0.6, 0.8\}$, in (a) *scale-free,* (b) *random graph* and (c) *small-world* networks with 100, 250, 500, 750 and 1000 nodes, as well as in (d) social media and (e) terrorist networks. Results are presented for $d = 15$, and the independent cascade influence model. A darker color indicates that the corresponding centrality measure has a greater weight in the seeker's mixed strategy.

- *Social media*: eigenvector centrality is the most frequent choice for Facebook and Google+ networks, but for the Twitter network it is replaced by closeness and betweenness. This could be due to the former networks having a lower density and average clustering coefficient than the last one, making them more similar to small-world networks.

The above analysis of equilibrium strategies, both for real-life and randomly-generated networks, allows us to derive a number of policy recommendations:

- Eigenvector centrality should be used by the seeker in networks exhibiting small-world properties. This finding is supported by the results for both randomly generated small-world networks and real-life social media networks.
- Degree centrality should be used by the seeker in scale-free networks, as evident by the results for Barabasi-Albert networks. However, since those networks exhibit some small-world properties, eigenvector centrality can be considered as a second choice.
- For networks that resemble random graphs, eigenvector centrality proves to be useful, at least against evaders whose budget is small. As for larger budgets, closeness centrality yields superior results.
- For two of the three terrorist networks under consideration, eigenvector centrality dominates the alternatives, highlighting its potential benefits when facing covert networks.

In general, eigenvector centrality seems to be a reliable choice for a variety of network types. Although for some networks it is the second best choice, generally it outperforms other measures, and seems to be more resilient against strategic manipulation.

6 CONCLUSIONS

We investigated the problem of concealing the importance of an individual in a social network, where both the evader, i.e., the person who wishes to hide, and the seeker, i.e., the party analyzing the network, act strategically. We focused on settings where the evader cannot rewire edges between complete strangers, but instead can only modify connections involving her neighbours in the networks. We showed that even in this simplified setting, the problem of finding an optimal way to hide from the most fundamental centrality measures is NP-complete. In light of these hardness results, we analyzed a number of instances of the game under both the zero-sum and the non-zero-sum payoffs; this highlighted some potential policy implications for network analyzers in the face of a strategic evader.

For future work, we intend to study this setting more rigorously, e.g., by analyzing the case in which multiple evaders are acting simultaneously, and more broadly, e.g., by considering a wider range of centrality measures available to the seeker. Another interesting follow-up of this study is to analyze the problem of hiding from link-prediction algorithms [5, 22, 24] under the assumption that both the evader and the seeker act strategically.

ACKNOWLEDGMENTS

Tomasz Michalak was supported by the Polish National Science Centre (grant 2016/23/B/ST6/03599). Yevgeniy Vorobeychik was supported by the National Science Foundation (IIS-1903207, IIS-1905558) and Army Research Office (MURI W911NF1810208). Kai Zhou was supported by PolyU (UGC) Internal Fund (1-BE3U). For an earlier version of this work, Marcin Waniek was supported by the Polish National Science Centre (grant 2015/17/N/ST6/03686).

REFERENCES

[1] J M Anthonisse. 1971. The rush in a directed graph. *Stichting Mathematisch Centrum. Mathematische Besliskunde* BN 9/71 (1971), 1–10.

[2] Alex Bavelas. 1948. A mathematical model for group structures. *Human organization* 7, 3 (1948), 16–30.

[3] Murray A Beauchamp. 1965. An improved index of centrality. *Behavioral Science* 10, 2 (1965), 161–163.

[4] Phillip Bonacich. 1987. Power and centrality: A family of measures. *American journal of sociology* 92, 5 (1987), 1170–1182.

[5] Palash Dey and Sourav Medya. 2020. Manipulating Node Similarity Measures in Networks. In *AAMAS '20: Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. IFAAMAS, Auckland, New Zealand, 321–329.

[6] Santo Fortunato. 2010. Community detection in graphs. *Physics Reports* 486, 3 (2010), 75–174.

[7] Linton C Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry* 40 (1977), 35–41.

[8] Drew Fudenberg and Jean Tirole. 1991. *Game theory*. Technical Report. MIT press.

[9] Lise Getoor and Christopher P Diehl. 2005. Link mining: a survey. *Acm Sigkdd Explorations Newsletter* 7, 2 (2005), 3–12.

[10] Manish Jain, James Pita, Milind Tambe, Fernando Ordónez, Praveen Paruchuri, and Sarit Kraus. 2008. Bayesian Stackelberg games and their application for security at Los Angeles International Airport. *ACM SIGecom Exchanges* 7, 2 (2008), 10.

[11] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, New York, USA, 137–146.

[12] Michael Kenney, John Horgan, Cale Horne, Peter Vining, Kathleen M Carley, Michael W Bigrigg, Mia Bloom, and Kurt Braddock. 2013. Organisational adaptation in an activist network: Social networks, leadership, and change in al-Muhajiroun. *Applied ergonomics* 44, 5 (2013), 739–747.

[13] Dirk Koschützki, Katharina Anna Lehmann, Leon Peeters, Stefan Richter, Dagmar Tenfelde-Podehl, and Oliver Zlotowski. 2005. Centrality indices. In *Network analysis*. Springer, Berlin, Germany, 16–61.

[14] Wanying Luo, Qi Xie, and Urs Hengartner. 2009. Facecloak: An architecture for user privacy on social networking sites. In *Computational Science and Engineering, 2009. CSE'09. International Conference on*, Vol. 3. IEEE, New York, USA, 26–33.

[15] Tomasz P Michalak, Talal Rahwan, and Michael Wooldridge. 2017. Strategic Social Network Analysis.. In *AAAI 2017*. AAAI, San Francisco, USA, 4841–4845.

[16] Praveen Paruchuri, Jonathan P Pearce, Janusz Marecki, Milind Tambe, Fernando Ordonez, and Sarit Kraus. 2008. Playing games for security: an efficient exact algorithm for solving Bayesian Stackelberg games. In *AAMAS 2008*. IFAAMAS, Estoril, Portugal, 895–902.

[17] Praveen Paruchuri, Jonathan P Pearce, Milind Tambe, Fernando Ordonez, and Sarit Kraus. 2007. An efficient heuristic approach for security against multiple adversaries. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. ACM, New York, USA, 181.

[18] Marvin E Shaw. 1954. Group structure and the behavior of individuals in small groups. *The Journal of Psychology* 38, 1 (1954), 139–149.

[19] Marcin Waniek, Tomasz P Michalak, Talal Rahwan, and Michael Wooldridge. 2017. On the Construction of Covert Networks. In *AAMAS 2017*. IFAAMAS, Sao Paulo, Brazil, 1341–1349.

[20] Marcin Waniek, Tomasz P Michalak, Michael J Wooldridge, and Talal Rahwan. 2018. Hiding individuals and communities in a social network. *Nature Human Behaviour* 2, 2 (2018), 139.

[21] Marcin Waniek, Jan Woźnica, Kai Zhou, Yevgeniy Vorobeychik, Talal Rahwan, and Tomasz Michalak. 2021. Strategic Evasion of Centrality Measures. arXiv:2101.10648 [cs.SI]

[22] Marcin Waniek, Kai Zhou, Yevgeniy Vorobeychik, Esteban Moro, Tomasz P Michalak, and Talal Rahwan. 2019. How to hide one's relationships from link prediction algorithms. *Scientific reports* 9, 1 (2019), 1–10.

[23] William Lafi Youmans and Jillian C York. 2012. Social media and the activist toolkit: User agreements, corporate interests, and the information infrastructure of modern social movements. *Journal of Communication* 62, 2 (2012), 315–329.

[24] Kai Zhou, Tomasz P Michalak, and Yevgeniy Vorobeychik. 2019. Adversarial robustness of similarity-based link prediction. In *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, IEEE, Beijing, China, 926–935.