Hiding from Centrality Measures: A Stackelberg Game Perspective

Marcin Waniek, Jan Woźnica, Kai Zhou, Yevgeniy Vorobeychik, Tomasz P. Michalak, Talal Rahwan

Abstract—Centrality measures can rank nodes in a social network according to their importance. However, in many cases, a node may want to avoid being highly ranked by such measures, e.g., as is the case with terrorist networks. In this work, we study a confrontation between the seeker—the party analyzing a social network using centrality measures—and the evader—a node attempting to decrease its ranking according to such measures. We analyze the possible outcomes of modifying, i.e., adding or removing, a single edge by the evader, showing that even without complete knowledge about the network, the effects of the modification on the evader's ranking can often be predicted. We study the computational complexity of finding a set of modifications that reduce the evader's centrality ranking in an optimal way, proving that these decision problems are NP-complete. Moreover, we provide a 2-approximation for the degree centrality, and logarithmic approximation boundaries for the closeness and betweenness centralities. Finally, we define and investigate a Stackelberg game between the seeker and the evader, providing a Mixed Integer Linear Programming formulation of finding an equilibrium. Altogether, we provide a thorough analysis of the strategic aspects of hiding from centrality measures in social networks.

Index Terms—Centrality measure, Stackelberg game, social network, complexity analysis.

1 Introduction

Ever since the dawn of the Internet Age, a rapidly growing amount of information about our daily lives is uploaded to the Web. A plethora of this data, such as our conversations, our likes and dislikes, and even our relationships can be represented using network structures. Simultaneously with this process, we can observe the development of an increasing number of social network analysis tools and techniques capable of inferring various information from the data publicly available online. This raises a privacy-related concern, as members of social networks are no longer able to keep their sensitive information private.

One of the most widely-used social network analysis tools are centrality measures [1], [2]. A centrality measure is an algorithm that estimates the relative importance of nodes in a network. In other words, with the use of centrality measures, it is possible to identify the key players in a network, where the exact notion of importance depends on the centrality measure of choice. However, there exist situations in which such key players might not want to be identified. We have already mentioned the issues pertaining to the privacy of Internet users. At first glance, the fact that an average social media user might prefer to evade analysis performed with centrality measures may seem unimportant. However, when we consider the situation of opposition bloggers in authoritarian regimes, the consequences of being identified as the most important node in the network may be much more dire. These circumstances allow us to sympathize with a member of the social network who wishes to avoid being detected by a centrality measure. Nevertheless, one can present a scenario in which an ability to evade centrality measures presents a grave danger to public safety. In particular, centrality analysis is often used to pinpoint the leaders of criminal [3] and terrorist [4] organizations. In such situations we would like to diminish the probability that the actual ringleader of the network avoids detection.

Given this ethical asymmetry of possible real-life scenarios, in our analysis we consider an abstract model of hiding from centrality measures in a social network. More specifically, we consider a confrontation between the *seeker*—a third party who uses centrality measures to pinpoint the most important nodes in the network—and the *evader*—a member of the social network who wishes to avoid being detected by the seeker. In this work, we assume that the seeker and evader are aware of each others' existence (see Section 5.1 for the comparison with previous works on hiding from centrality measures where this assumption was not in place). This turns our model into a game-theoretic setting, where both players try to achieve their goals by applying carefully selected strategies. The set of strategies of the seeker consists of centrality measures that can be used to find the most important nodes in the network. On the other hand, the strategies of the evader take the form of adding or removing some of the edges to mislead the centrality analysis performed by the seeker.

To be more precise, in our work we consider a number of research questions. First, *assuming limited knowledge about the network structure, is it possible to predict the effect of adding or removing a given edge on the evader's ranking?* We answer this question theoretically in Section 4 and empirically in Section 7.2. Second, *is it possible for the evader to find an optimal way of hiding?* We resolve this issue in Section 5 by analyzing the computational complexity of the problem faced by the evader. Third, *what is the outcome of the confrontation between the seeker and the evader?* To address this question, we formally define a Stackelberg game between the seeker

- *M. Waniek and T. Rahwan are with Computer Science, Science Division, New York University Abu Dhabi.*
- *M. Waniek, J. Woźnica, and T. P. Michalak are with University of Warsaw.*
- *K. Zhou is with Hong Kong Polytechnic University.*
- *Y. Vorobeychik is with Washington University in St. Louis.*
- *T. P. Michalak is with IDEAS NCBR.*

and the evader in Section 6 and study its equilibria in Section 7.3. Altogether, our work presents an exploration of the strategic aspects of hiding from centrality measures in a social network.

The motivation for our work is twofold. From the perspective of the seeker, the results of our work could give law enforcement agencies a fresh insight into the possible ways in which members of criminal and terrorist organizations avoid detection. This insight is particularly crucial, as centrality measures are one of the key tools for analyzing covert networks. From the perspective of the evader, our work could be of use to the members of communities that are discriminated against, such as specific ethnic groups in authoritarian regimes. Careful rewiring of the social network structure might help the leaders of such communities elude harsh repressions.

A preliminary version of this work was published in the Proceedings of the 20th Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2021) [5]. The new results added in this extended version are:

- Theoretical analysis of whether the addition or removal of a given edge can increase or decrease the ranking of the evader, and what can be the magnitude of this change (Section 4).
- The approximation version of the problem and the analysis of its computational complexity for different centrality measures (Section 5.2).
- A formal proof that the MIQP and MILP formulations of the seeker-evader game equilibrium are equivalent (Section 6.2).
- An empirical analysis of the sensitivity of centrality measures (Section 7.2).
- Experiments with the seeker-evader game in the Ambassador network, as well as Newman and Prüfer random network generation models (Section 7.3).

2 Related Work

Our article is part of the literature on using edge perturbations to manipulate network properties. The purpose of these interventions could be minimizing the average distance between nodes [6], promoting health-related behaviors [7], and reducing the number of small dense subgroups [8]. The body of literature that is most relevant to our setting considers the task of strategically manipulating centrality measures. Some works focus on the problem of adding edges to increase the closeness centrality of a given node [9], or to increase multiple centralities at the same time [10].

In this work, we attempt not to increase, but rather to decrease the value of centrality, in the hope of hiding a selected node. This problem was considered for both the centrality value of a selected evader [11], as well as for the ranking position of a group of network's leaders [12], [13], [14]. Another facet of the problem that was analyzed in the literature is the axiomatic characterization of the centrality measures that are resilient to manipulation [15]. While most works considers standard network structure, some examine networks that consist of multiple layers [16], or temporal networks where edges exists only at specific moments [17].

TABLE 1
Summary of the notation used in the article.

Symbol	Meaning
V	The set of network nodes
E	The set of network edges
$N(v_i)$	The set of neighbors of v_i
$\kappa(v_i)$	The degree of v_i
$\Pi(v_i, v_j)$	The set of shortest paths between v_i and v_j
$d(v_i, v_j)$	The distance between v_i and v_j
$v^\dagger$	The evader
$\bar{V}$	The set of all nodes other than $v^\dagger$
$\ddot{V}$	The set of all nodes other than $v^\dagger$ and its neighbors
$\zeta_1^\dagger$	The set of all edges incident with $v^\dagger$
$\zeta_2^\dagger$	The set of all edges between the neighbors of $v^\dagger$
$\zeta_3^\dagger$	The set of all edges not belonging to $\zeta_1^\dagger$ or $\zeta_2^\dagger$
$c_{dg}(G, v_i)$	The degree centrality of v_i in G
$c_{cl}(G, v_i)$	The closeness centrality of v_i in G
$c_{bt}(G, v_i)$	The betweenness centrality of v_i in G
$c_{bt}(G, v_i)$	The eigenvector centrality of v_i in G
$\iota(G, v_i)$	The influence of v_i over G
b	The number of edges that can be added or removed
$\hat{A}$	The set of edges that can be added to the network
$\hat{R}$	The set of edges that can be removed from the network
δ	Required nodes with centrality greater than $v^\dagger$
ϕ	The type of the evader
U_e^R	The evader's utility coming from the centrality ranking
U_e^I	The evader's utility coming from the influence value
U_e	The aggregated utility of the evader

A related body of literature concerns itself with hiding from other types of social network analysis tools. These types of evasion techniques are often motivated by the need of privacy protection [18]. Social media users who do not wish some of their undisclosed relationship to be uncovered, might be interested in heuristic solutions designed to mislead link prediction algorithms [19], [20], [21]. Others might want to counter the analysis performed using node similarity measures [22]. A group of people might wish to avoid being identified as a closely-knit faction by community detection algorithms [11], [23], [24]. Yet another class of techniques allow to hide the identity of the source of network diffusion from the source detection algorithms [25]. Some techniques have also been propose to prevent the inference of an edge type in signed networks where each relation is tagged as either positive or negative [26].

In an even wider perspective, our study is a part of the literature on adversarial attack and defense in networks [27], [28]. Many of the works are focused on attacking machine learning methods processing the network data, either by manipulating the data they are trained on (poisoning attack) [29], [30], [31] or by manipulating the input to an already trained algorithm (evasion attack) [32], [33], [34]. Another example of such adversarial setting is a confrontation between an attacker trying to spread a diffusion process in a network and a defender trying to stop it [35].

3 Preliminaries

In this section, we present the basic network notation and concepts that will be used throughout the article. For the convenience of the reader, Table 1 provides a summary of the notation used in the article.

3.1 Basic Network Notation

Let $G = (V, E)$ denote a *network*, where $V = \{v_1, \ldots, v_n\}$ is the set of n *nodes* and $E \subseteq V \times V$ is the set of *edges*. We denote by (v_i, v_j) an edge between the nodes v_i and v_j. In this work we focus on *undirected* networks, i.e., we do not discern between edges (v_i, v_j) and (v_j, v_i). We also assume that networks do not contain self-loops, i.e., $\forall_{v_i \in V}(v_i, v_i) \notin E$. We denote by $N_G(v_i)$ the set of *neighbors* of v_i, i.e., $N_G(v_i) = \{v_j \in V : (v_i, v_j) \in E\}$. We denote by $\kappa_G(v_i)$ the *degree* of v_i, i.e., $\kappa_G(v) = |N_G(v)|$. We denote by $N_G(v_i, v_j)$ the set of *common neighbors* of v_i and v_j, i.e., $N_G(v_i, v_j) = \{v_k \in V : (v_i, v_k) \in E \wedge (v_j, v_k) \in E\}$. To make the notation more readable, we will often omit the network itself from the notation whenever it is clear from the context, e.g., by writing $N(v_i)$ instead of $N_G(v_i)$. This applies not only to the notation presented thus far, but rather to all notation in this article.

A *path* in (V, E) is an ordered sequence of nodes, $p = \langle v_{i_1}, \ldots, v_{i_k} \rangle$, in which every two consecutive nodes are connected by an edge in E. The *length of a path* is equal to the number of edges therein. For any pair of nodes, $v_i, v_j \in V$, we denote by $\Pi(v_i, v_j)$ the set of all shortest paths between v_i and v_j, and we denote by $d(v_i, v_j)$ the *distance* between v_i and v_j, i.e., the length of a shortest path between v_i and v_j.

We will often focus on a particular node $v^\dagger \in V$, called *the evader*. Let $\bar{V}$ denote the set of all nodes other than $v^\dagger$, i.e., $\bar{V} = V \setminus \{v^\dagger\}$. Furthermore, let $\ddot{V}$ denote the set of all nodes other than $v^\dagger$ and the neighbors of $v^\dagger$, i.e., $\ddot{V} = V \setminus (\{v^\dagger\} \cup N(v^\dagger))$. Next, we define three classes of edges denoted by $\zeta_1^\dagger$, $\zeta_2^\dagger$ and $\zeta_3^\dagger$. The first class, $\zeta_1^\dagger$, consists of every edge incident with $v^\dagger$, i.e., $\zeta_1^\dagger = \{v^\dagger\} \times \bar{V}$. The second class, $\zeta_2^\dagger$, consists of every edge whose ends are both neighbors of $v^\dagger$, i.e., $\zeta_2^\dagger = N(v^\dagger) \times N(v^\dagger)$. The third class, $\zeta_3^\dagger$, consists of every remaining edge, i.e., $\zeta_3^\dagger = \bar{V} \times \ddot{V}$. Notice that although the elements of each such class will be referred to as "edges", they may or may not be present in any given network. This is unlike the elements of E, which are the edges that are present in the network $G = (V, E)$.

3.2 Centrality Measures

A *centrality measure* is a function, $c(G, v_i)$, that expresses the relative importance of any given node v_i in the network G [1]. In this work we consider four fundamental centrality measures, namely degree, closeness, betweenness, and eigenvector.

Degree centrality [36] quantifies the importance of a node based on the number of its neighbors. Formally, the normalized degree centrality of a node $v_i \in V$ in a network G is:

$$c_{dg}(G, v_i) = \frac{|N(v_i)|}{n-1}.$$

Closeness centrality [37] assigns the importance of a node based on an average distance to all other nodes. Formally, the normalized closeness centrality of a node $v_i \in V$ is:

$$c_{cl}(G, v_i) = \frac{n-1}{\sum_{v_j \in V} d(v_i, v_j)}.$$

Betweenness centrality [38], [39] measures the importance of a given node in the context of network flow. The normalized betweenness centrality of a node $v_i \in V$ is:

$$c_{bt}(G, v_i) = \frac{2}{(n-1)(n-2)} \sum_{v_j, v_k \in V \setminus \{v_i\}} \frac{|\{p \in \Pi(v_j, v_k) : v_i \in p\}|}{|\Pi(v_j, v_k)|}.$$

Eigenvector centrality [40] quantifies the importance of a given node based on the importance of its neighbors. Formally, the eigenvector centrality of a node v_i is:

$$c_{eg}(G, v_i) = \chi_i^*$$

where χ^* is the eigenvector corresponding to the largest eigenvalue of the adjacency matrix of the network G.

3.3 Influence Models

The propagation of influence in a network can be described in terms of node activation. At the beginning of the process only a selected set of nodes (known as the *seed set*) is activated. Inactive nodes can become activated when they are sufficiently influenced by their neighbors. Assume that the process consists of discrete rounds. We then denote by $I(t) \subseteq V$ the set of active nodes in round t, where $I(1)$ is the seed set. The influence model under consideration determines the exact conditions of a node becoming active. In this work we consider two models of influence: independent cascade and linear threshold.

In the *independent cascade* [41] model, every pair of nodes is assigned an activation probability, $p : V \times V \to [0, 1]$. In every round $t > 1$ every node $v_i \in V$ that became active in round $t-1$ activates each inactive neighbor $v_j \in N(v_i) \setminus I(t-1)$ with probability $p(v_i, v_j)$. The process ends when there are no newly activated nodes, i.e., $I(t) = I(t-1)$.

In the *linear threshold* [42] model, every node $v_i \in V$ is assigned a *threshold value* t_{v_i} sampled from the set $\{0, \ldots, |N(v_i)|\}$ according to some probability distribution. In every round, $t > 1$, every inactive node v_i becomes active if $|I(t-1) \cap N(v_i)| \geq t_{v_i}$. The process ends when there are no newly activated nodes, i.e., when $I(t) = I(t-1)$.

In either model, the *influence* of a node, v_i, on another node, v_j, is denoted by $\iota(G, v_i, v_j)$ and is defined as the probability that v_j gets activated given the seed set $\{v_i\}$. We assume that $\iota(G, v_i, v_i) = 0$ for all $v_i \in V$. We define the influence of v_i over the entire network as $\iota(G, v_i) = \sum_{v_j \in V} \iota(G, v_i, v_j)$. When referring to the influence of a given node, we mean the influence over the entire network.

4 Possible Changes in Centrality Ranking

We first focus on the question of how adding a specific edge to the network or removing a specific edge from the network can affect the centrality c ranking of a given node $v^\dagger \in V$? Can the centrality ranking both increase and decrease after a given network modification? And what about the magnitude of this change, can it be arbitrarily large, or is it strictly limited? Importantly, we focus on the *ranking* of $v^\dagger$ according to centrality c, rather than on the centrality *value* of v according to c. The ranking is the position of $v^\dagger$ in the list of all nodes, sorted according to their centrality values. We assume that nodes with the same centrality value have the same ranking.

TABLE 2
Summary of our results concerning possible ranking changes. For any given evader $v^\dagger \in V$, we study three classes of edges: $\zeta_1^\dagger$, $\zeta_2^\dagger$ and $\zeta_3^\dagger$, and three centrality measures: degree, closeness and betweenness. For every class and every measure, we investigate the potential impact of adding or removing an edge from that class on the centrality ranking of $v^\dagger$. The "$\Uparrow k$" (resp. "$\Downarrow k$") mark indicates that the increase (resp. decrease) in ranking can be arbitrarily large, while the "$\Uparrow 2$" (resp. "$\Downarrow 2$") mark indicates that the ranking can only increase (resp. decrease) by at most two positions. The "$\Uparrow$ ✗" (resp. "$\Downarrow$ ✗") mark indicates that the ranking increase (resp. decrease) is impossible.

		Adding e to the network		Removing e from the network	
$e \in \zeta_1^\dagger$	Degree	$\Uparrow k$	$\Downarrow$ ✗	$\Uparrow$ ✗	$\Downarrow k$
	Closeness	$\Uparrow k$	$\Downarrow k$	$\Uparrow k$	$\Downarrow k$
	Betweenness	$\Uparrow k$	$\Downarrow k$	$\Uparrow k$	$\Downarrow k$
$e \in \zeta_2^\dagger$	Degree	$\Uparrow$ ✗	$\Downarrow 2$	$\Uparrow 2$	$\Downarrow$ ✗
	Closeness	$\Uparrow$ ✗	$\Downarrow k$	$\Uparrow k$	$\Downarrow$ ✗
	Betweenness	$\Uparrow k$	$\Downarrow k$	$\Uparrow k$	$\Downarrow k$
$e \in \zeta_3^\dagger$	Degree	$\Uparrow$ ✗	$\Downarrow 2$	$\Uparrow 2$	$\Downarrow$ ✗
	Closeness	$\Uparrow k$	$\Downarrow k$	$\Uparrow k$	$\Downarrow k$
	Betweenness	$\Uparrow k$	$\Downarrow k$	$\Uparrow k$	$\Downarrow k$

In our analysis we divide the edges that can be added to or removed from the network into three classes: edges incident with the evader $\zeta_1^\dagger$, edges between the neighbors of the evader $\zeta_2^\dagger$, and the remaining edges $\zeta_3^\dagger$ (all three classes are formally defined in Section 3.1). The reason for this division is the fact that the evader who would like to strategically control their centrality ranking has varying levels of control over different edges. The edges on which the evader can exercise the greatest control are those of which the evader belongs to, i.e., the edges from the class $\zeta_1^\dagger$. The addition of this type of edge can be interpreted as performing a telephone call with someone, while the removal of this type of edge can represent removing someone from a list of friends on a social media platform. We can assume that the evader has a smaller amount of control over edges between their neighbors, i.e., the edges from the class $\zeta_2^\dagger$. The addition of this type of edge can be interpreted as introducing two friends to each other, while the removal of this type of edge can represent asking two associates to cease contacts with each other. Finally, we can assume that the evader has the least amount of control over edges outside of their direct network vicinity, i.e., the edges from the class $\zeta_3^\dagger$. The addition of this type of edge can be interpreted as inviting two strangers to the same event, while the removal of this type of edge can represent deleting data about a certain connection from a database.

We can state the question that we intend to investigate as follows: *Given a centrality measure c, a network $G = (V, E)$, an evader $v^\dagger \in V$, and a class of edges $\zeta^\dagger \in \{\zeta_1^\dagger, \zeta_2^\dagger, \zeta_3^\dagger\}$, can the addition or removal of an edge $e \in \zeta^\dagger$ to the network increase or decrease the ranking of $v^\dagger$ according to c?* Our findings on this matter are summarized in Table 2. Due to space limitations, the proofs of our results can be found in Section S1 of the supplementary materials.

TABLE 3
The summary of our computational complexity results.

Centrality	Local Hiding	Minimum Local Hiding
Degree	NP-complete	We show a 2-approximation
Closeness	NP-complete	Cannot be approximated within $(1-\epsilon)\ln\lvert\hat{A}\cup\hat{R}\rvert$ for any $\epsilon > 0$
Betweenness	NP-complete	Cannot be approximated within $(1-\epsilon)\ln\lvert\hat{A}\cup\hat{R}\rvert$ for any $\epsilon > 0$

5 Computational Complexity Analysis

Having analyzed the possible outcomes of adding or removing a single edge from the network, we now analyze the computational complexity of a problem of selecting the best subset of edges to hide the evader from centrality measures. Table 3 summarizes our results.

5.1 Decision Version of the Problem

We first formally define the computational problem faced by the evader who can perform only local changes.

Definition 1 (Local Hiding)**.** *This problem is defined by a tuple $(G, v^\dagger, b, c, \delta, \hat{A}, \hat{R})$, where $G = (V, E)$ is a network, $v^\dagger \in V$ is the evader, $b \in \mathbb{N}$ is a budget specifying the maximum number of edges that can be added or removed, c is a centrality measure, $\hat{A} \subseteq N(v^\dagger) \times N(v^\dagger)$ is the set of edges allowed to be added, $\hat{R} \subseteq \{v^\dagger\} \times N(v^\dagger)$ is the set of edges allowed to be removed, and $\delta \in \mathbb{N}$ 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 δ nodes with centrality c greater than that of the evader.*

As can be seen from the definition, we focus on two kinds of network modifications: removing edges incident with the evader, and adding edges between the neighbors of the evader. This choice is informed by the results presented in Section 4. As can be seen in Table 2, when it comes to the edges incident with the evader, i.e., edges belonging to the class $\zeta_1^\dagger$, only the removal operation can decrease the evader's ranking according to all three centrality measures (notice how the addition cannot affect the degree centrality in a beneficial way). Similarly, when considering edges between the evader's neighbors, i.e., edges belonging to the class $\zeta_2^\dagger$, only the addition of such edges has a chance of making the evader more hidden from all three measures (the removal can hide the evader from neither the degree nor the closeness centrality). We could also consider adding edges outside of the direct network vicinity of the evader, i.e., edges belonging to the class $\zeta_3^\dagger$, as this operation can also result in decreasing the ranking of the evader according to all three centrality measures. However, as discussed in Section 4, the evader typically has the least amount of control over such edges. Hence, for the sake of realism of the problem, we focus on the modifications of edges belonging to the first two classes.

Let us now comment on the practical aspects of executing network modifications as part of the evader's strategy. In most cases, the evader may remove edges that they

are a part of relatively easily, e.g., by ceasing contact with a specific acquaintance. On the other hand, the addition of edges between neighbors might be more demanding. In fact, forming connections with friends of friends (i.e., triadic closure) is one of the driving mechanisms of social network formation [43], [44]. What is more, research has shown that two-thirds of Facebook users are willing to accept friend requests from complete strangers [45], suggesting they might be even more likely to accept invitations from friends of friends. Nevertheless, to accommodate for situations in which some of the edge additions or removals are impossible to implement, we introduce sets $\hat{A}$ and $\hat{R}$ that precisely designate which network changes the evader is able to perform.

We now discuss the key differences between the above problem of *Local Hiding* and the problem of *Disguising Centrality* studied by Waniek et al. [11]. First, instead of seeking the optimal way of decreasing the value of the evader's centrality (which may not provide sufficient cover, especially if they are 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 δ. Second, we assume that the evader is only capable of rewiring edges within their network neighborhood—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. We also comment on the key differences between our *Local Hiding* problem and the problem of *Hiding Leaders* studied by Waniek et al. [12], [14] in the context of constructing covert networks. First, the authors divide the nodes into leaders and the followers, where the changes in the network are allowed only among the followers. Second, they only allow edges to be *added* among the followers, meaning that no edge can be removed from the network.

Below we present the proof of one of our results. Due to space limitations, the remaining proofs can be found in Section S2 of the supplementary materials.

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

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 (an example of this construction is presented in Figure 1):

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

Now, consider an instance $(H, v^\dagger, b, c, \delta, \hat{A}, \hat{R})$ of the Local Hiding problem where $H = (V', E')$ is the network we just constructed, $v^\dagger$ is the evader, $b = \frac{k(k-1)}{2}$, c is the degree centrality measure, $\delta = k$, $\hat{A} = E$, and $\hat{R} = \emptyset$.

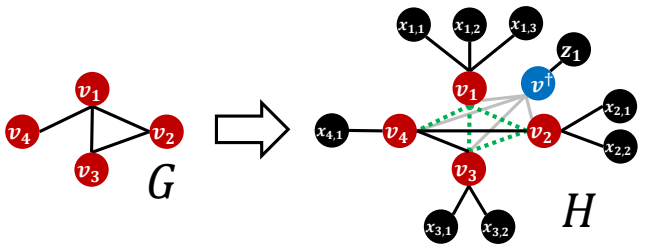


Fig. 1. An example of the construction used in the proof of Theorem 1 for $k = 3$. Some edges are printed grey for better readability. Green dotted lines correspond to the edges allowed to be added.

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 these edges, we need to choose a subset, $A^* \subseteq E$, as a solution to the Local Hiding problem (as $\hat{R} = \emptyset$, we are not allowed to remove any edges). In what follows, we will show a correspondence between a solution to the constructed instance of the Local Hiding problem and a solution to the given instance of the Finding k-Clique problem.

First, note that $v^\dagger$ has the highest degree in H, which is $n + k - 2$. Thus, in order for A^* to be a solution to the constructed Local Hiding problem instance, 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 $n + k - 1$ (notice that the addition of A^* only increases the degrees of nodes in V, since we already established that $A^* \subseteq E$). Since the degree of every node v_i in H 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^*.

Assume that there exists a solution V^* to the given instance of the Finding k-Clique problem, i.e., a subset $V^* \subseteq V$ of size k forming a clique in G. We will show that $V^* \times V^*$ is a solution to the constructed instance of the Local Hiding problem. Since the nodes in V^* form a clique in G, we have that $V^* \times V^* \subseteq E$. Since $\hat{A} = E$, we also have that $V^* \times V^* \subseteq \hat{A}$. Finally, since the nodes in V^* form a clique in G, the addition of $V^* \times V^*$ to H increases the degree of each of the k nodes in V^* by exactly $k - 1$. We showed that if there exists a solution to the given instance of the Finding k-Clique problem, then there also exists a solution to the constructed instance of the Local Hiding problem.

Assume that there exists a solution to the constructed instance of the Local Hiding problem, i.e., $A^* \subseteq \hat{A}$ the addition of which to H increases the degree of at least k nodes in V by at least $k - 1$. However, since the budget is $b = \frac{k(k-1)}{2}$, then the only possible choice of A^* is the one that increases the degree of exactly k nodes in V by exactly $k-1$ each. Hence, the edges in A^* induce a clique of size k in $\hat{A}$. However, since $\hat{A} = E$, the same edges also induce a clique of size k in G. We showed that if there exists a solution to the constructed instance of the Local Hiding problem, then there also exists a solution to the given instance of the Finding k-Clique problem

We proved that a solution to the given instance of the

Finding k-Clique problem exists if and only if there exists a solution to the constructed instance of the Local Hiding problem. □

5.2 Approximation Version of the Problem

Having discussed the decision version of the Local Hiding problem, let us now define its approximation version.

Definition 2 (Minimum Local Hiding)**.** *This problem is defined by a tuple $(G, v^\dagger, c, \delta, \hat{A}, \hat{R})$, where $G = (V, E)$ is a network, $v^\dagger \in V$ is the evader, c is a centrality measure, $\hat{A} \subseteq N(v^\dagger) \times N(v^\dagger)$ is the set of edges allowed to be added, $\hat{R} \subseteq \{v^\dagger\} \times N(v^\dagger)$ is the set of edges allowed to be removed, and $\delta \in \mathbb{N}$ 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 $|A^*| + |R^*|$ is minimal and the resulting network $(V, (E \cup A^*) \setminus R^*)$ contains at least δ nodes with centrality c greater than that of the evader.*

Notice that while in the Local Hiding problem we asked whether or not there exists a solution within a certain budget, in the Minimum Local Hiding problem we are looking for a solution that is as small as possible (hence, we are accepting solutions that are not of the optimal size). This key difference results in distinct way of analyzing this class of problems, as we will see in the proofs below.

Due to space limitations, the proofs of our results can be found in Section S2 of the supplementary materials.

6 The Seeker-Evader Game

Having analyzed the computational complexity of the problem faced by the evader, we now move to defining the confrontation between the seeker and the evader as a game.

6.1 The Game Definition

The game takes place between two players: *the seeker* who is a party analyzing a social network, and *the evader* who is one of the nodes of the social network analyzed by the seeker. The seeker uses a centrality measure to identify the most important node of the social network, while the evader wishes to avoid being pinpointed as the most important node.

We model this confrontation as a Stackelberg game [46]. A Stackelberg game is a game between two players, *a leader* and *a follower*. The leader moves first, selecting one of their strategies. This move is observed by the follower, who then select one of their strategies as a response. In our case, the leader player is the seeker, whose set of strategies C_S consists of the centrality measures that can be used to analyze the network. The follower player is the evader, who observes the centrality measure used by the seeker and selects a strategy from the set Ξ_E. Each strategy of the evader consists of removing some edges from the network and adding some edges to the network.

We now discuss the utility functions of both players, starting with this of the evader. In the theoretical analysis presented so far we focused our attention on the problem of lowering the centrality ranking of the evader. Here, we introduce another factor that can motivate the evader, i.e., their influence over the network, measured using one of the models presented in Section 3.3.

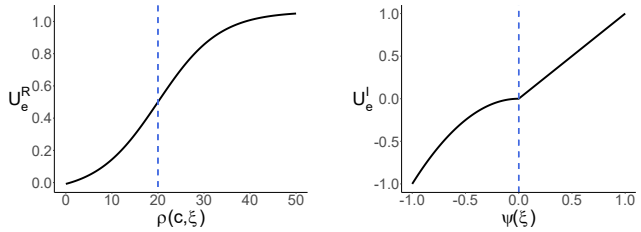


Fig. 2. The evader's utility functions given $d = 20$ and $k = \frac{3}{d}$. The dashed blue lines represent the inflection points of both functions.

Let $c \in C_S$ be the strategy selected by the seeker, and let $\xi \in \Xi_E$ be the strategy selected by the evader. We define the utility of the evader as:

$$U_e(\phi, c, \xi) = \phi U_e^R(c, \xi) + (1 - \phi) U_e^I(\xi)$$

where:

- $U_e^R(c, \xi)$ is the evader's utility coming from their ranking position according to the centrality measure c selected by the seeker, in the network resulting from introducing network modification ξ,
- $U_e^I(\xi)$ is the evader's utility coming from the change in her influence over the network after introducing network modification ξ,
- $\phi \in \left\{\frac{1}{m+1}, \ldots, \frac{m}{m+1}\right\} = \Phi$ is the type of the evader (with m being the number of types) determining whether the evader is more focused on their ranking position or their influence.

Next, we discuss the formulas of $U_e^R(c, \xi)$ and $U_e^I(\xi)$. Figure 2 presents the plots of both functions. The evader's utility based on the centrality ranking is defined as:

$$U_e^R(c, \xi) = \frac{1}{\alpha\left(1 + e^{-k(\rho(c,\xi)-d)}\right)} - \frac{\beta}{\alpha},$$

where e is Euler's number, $\rho(c, \xi)$ is the evader's position in the ranking of c after executing network modification ξ, k is the curve steepness, d is the inflection point, $\beta = \frac{1}{1+e^{-k(1-d)}}$, and $\alpha = 1 - 2\beta$. In our simulations we use the values of $d = 5$ and $k = \frac{3}{d}$.

Notice that the function defined this way has a number of desirable properties. First, if the evader is ranked first, i.e., $\rho(c, \xi) = 1$, their utility is equal to zero. Second, as the evader becomes more hidden their utility increases, i.e., $U_e^R(c, \xi)$ increases with $\rho(c, \xi)$. Third, the function is convex for $\rho(c, \xi) \leq d$, i.e., the marginal gain in utility increases with the evader's ranking, until the evader reaches position d. Fourth, the function is concave for $\rho(c, \xi) \geq d$, i.e., further decreasing the evader's ranking beyond position d has diminishing returns.

The evader's utility based on the influence is defined as:

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

where $\psi(\xi)$ is the relative change in the evader's influence after executing the strategy ξ, i.e., $\psi(\xi) = \frac{\iota(\xi) - \iota_0}{\iota_0}$ with ι_0

and $\iota(\xi)$ denoting the evader's influence before and after executing the strategy ξ respectively.

Let us now comment on the properties of the function defined this way. First, $U_e^I(\xi)$ is concave for $\psi(\xi) \leq 0$, i.e., the marginal loss in utility grows with the loss in influence. Intuitively, this can be interpreted as the evader who does not mind a negligible loss of influence, but strongly opposes a significant decrease. Second, throughout its domain the value of $U_e^I(\xi)$ has a similar order of magnitude to the value of $U_e^R(c,\xi)$, meaning that the aggregated utility of the evader $U_e(\phi,c,\xi)$ is not dominated by any of those two utilities.

We now describe the utility function of the seeker. The seeker-evader game is a *zero-sum game*. Hence, the goal of the seeker is to minimize the total utility of the evader, i.e., the utility of the seeker is $U_s(\phi,c,\xi) = -U_e(\phi,c,\xi)$. We assume that the utility functions of both players and the distribution of the evader types are common knowledge, while the actual type of the evader is unknown to the seeker.

6.2 Finding the Optimal Strategies

We now formulate the problem of identifying the optimal strategies of the seeker and the evader as a mixed-integer quadratic program (MIQP). As a reminder, the set of strategies of the seeker consists of centrality measures that can be used to analyze the network, while the set of strategies of the evader consists of different ways of rewiring the network. Let $t(\phi)$ be the probability that the evader type is ϕ. Since the seeker knows the distribution of the evader's types, but not the actual type, they are likely to use a mixed strategy, trying to optimize the choice of centrality based on different types of the evader they might be facing. Let $p(c)$ be the probability that the seeker plays pure strategy $c \in C_S$. Moreover, let $q(\phi,\xi)$ be the probability that an evader of type ϕ plays pure strategy $\xi \in \Xi_E$. Notice that since the evader observes the strategy chosen by the seeker and moves second, they can restrict their choice to pure strategies, i.e., we have that $\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E} q(\phi,\xi) \in \{0,1\}$. The problem of finding the optimal strategies can now be formulated as follows.

Definition 3 (MIQP formulation)**.** *The mixed-integer quadratic program finding the optimal strategies is:*

$$\max_{p,q,a} \sum_{\phi\in\Phi}\sum_{c\in C_S}\sum_{\xi\in\Xi_E} t(\phi)p(c)q(\phi,\xi)U_s(\phi,c,\xi)$$

$$\text{subject to } \sum_{c\in C_S} p(c) = 1 \tag{i}$$

$$\forall_{\phi\in\Phi} \sum_{\xi\in\Xi_E} q(\phi,\xi) = 1 \tag{ii}$$

$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E}\ a(\phi) \geq \sum_{c\in C_S} p(c)U_e(\phi,c,\xi) \tag{iii}$$

$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E}\ a(\phi) \leq (1-q(\phi,\xi))\eta + \sum_{c\in C_S} p(c)U_e(\phi,c,\xi) \tag{iv}$$

$$\forall_{c\in C_S}\ p(c) \in [0,1]$$
$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E}\ q(\phi,\xi) \in \{0,1\}$$
$$\forall_{\phi\in\Phi}\ a(\phi) \in \mathbb{R}$$

Constraints (i) and (ii) define the probability distributions over strategies of both players. Notice that since $\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E} q(\phi,\xi) \in \{0,1\}$, the constraint (ii) implies that for a given $\phi \in \Phi$ we have $q(\phi,\xi) = 1$ for exactly one strategy ξ, and $q(\phi,\xi) = 0$ for all other strategies. Moreover, $\eta \in \mathbb{R}$ is an arbitrarily large number. It implies that if $q(\phi,\xi^*) = 1$, i.e., if $\xi^* \in \Xi_E$ is the strategy selected by the evader, we get:

$$a(\phi) = \sum_{c\in C_S} p(c)U_e(\phi,c,\xi^*).$$

Similarly, if $q(\phi,\xi') = 0$, i.e., if $\xi' \in \Xi_E$ is not the strategy selected by the evader, then constraint (iv) is automatically satisfied. Hence, constraints (iii) and (iv) imply that for a given $\phi \in \Phi$ the strategy selected by the evader maximizes their expected payoff.

Assume to the contrary, i.e., that there exist strategies $\xi', \xi^* \in \Xi_E$ such that $q(\phi,\xi^*) = 1$, $q(\phi,\xi') = 0$, and $\sum_{c\in C_S} p(c)U_e(\phi,c,\xi') > \sum_{c\in C_S} p(c)U_e(\phi,c,\xi^*)$, i.e., the strategy ξ' that is not get selected has a better expected payoff than the selected strategy ξ^*. However, this would violate constraint (iii), which requires that:

$$a(\phi) = \sum_{c\in C_S} p(c)U_e(\phi,c,\xi^*) \geq \sum_{c\in C_S} p(c)U_e(\phi,c,\xi').$$

Therefore, such strategies $\xi', \xi^* \in \Xi_E$ cannot exist, and the evader plays the strategy maximizing their expected payoff.

In order to solve the problem efficiently, we linearize it based on procedure described by Paruchuri et al. [47]. The main idea of the procedure is based on introducing the variable $z(\phi,c,\xi) = p(c)q(\phi,\xi)$. The problem can then be formulated as a mixed-integer linear program (MILP) as follows.

Definition 4 (MILP formulation)**.** *The mixed-integer linear program finding the optimal strategies is:*

$$\max_{z,q,a} \sum_{\phi\in\Phi}\sum_{c\in C_S}\sum_{\xi\in\Xi_E} t(\phi)z(\phi,c,\xi)U_s(\phi,c,\xi)$$

$$\text{subject to } \forall_{\phi\in\Phi} \sum_{c\in C_S}\sum_{\xi\in\Xi_E} z(\phi,c,\xi) = 1 \tag{1}$$

$$\forall_{\phi\in\Phi} \sum_{\xi\in\Xi_E} q(\phi,\xi) = 1 \tag{2}$$

$$\forall_{\phi\in\Phi}\forall_{c\in C_S} \sum_{\xi\in\Xi_E} z(\phi,c,\xi) = \sum_{\xi\in\Xi_E} z(0,c,\xi) \tag{3}$$

$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E} \sum_{c\in C_S} z(\phi,c,\xi) = q(\phi,\xi) \tag{4}$$

$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E}\ a(\phi) \geq \sum_{c\in C_S} U_e(\phi,c,\xi) \sum_{\xi'\in\Xi_E} z(\phi,c,\xi') \tag{5}$$

$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E}\ a(\phi) \leq (1-q(\phi,\xi))\eta + \sum_{c\in C_S} U_e(\phi,c,\xi) \sum_{\xi'\in\Xi_E} z(\phi,c,\xi') \tag{6}$$

$$\forall_{\phi\in\Phi}\forall_{c\in C_S}\forall_{\xi\in\Xi_E}\ z(\phi,c,\xi) \in [0,1]$$
$$\forall_{\phi\in\Phi}\forall_{\xi\in\Xi_E}\ q(\phi,\xi) \in \{0,1\}$$
$$\forall_{\phi\in\Phi}\ a(\phi) \in \mathbb{R}$$

The formal proof that the linearized version of the formulation indeed describes the same problem as the MIQP

TABLE 4
Characteristics of the real-life datasets considered in our simulations.

Network	Nodes	Edges	All strategies	Undominated strategies
WTC	36	64	$10,902$	205
Madrid	70	98	$3,213$	124
Bali	17	63	$30,913$	30
Ambassador	16	69	$52,462$	123
Facebook	44	138	$32,567$	92

formulation can be found in Section S3 of the supplementary materials.

7 Empirical Analysis

In this section we present the results of our simulations. We first detail the network datasets and random network generation models that we use. Then, we describe the results of our experiments with the sensitivity of centrality measures. Finally, we present the experimental analysis of the seeker-evader game.

7.1 Network Datasets and Models

In our experiments we consider the following real-life network datasets (their characteristics are presented in Table 4):

- **WTC** [48]—the network of terrorists responsible for the $9/11$ attacks in 2001.
- **Madrid** [49]—the network of terrorists responsible for the 2004 Madrid train bombings.
- **Bali** [50]—the network of terrorists responsible for the 2002 Bali bombing.
- **Ambassador** [51]—the network of terrorists responsible for the Philippines ambassador Jakarta residence bombing in 2000.
- **Facebook** [52]—ego network of a student of one of the American colleges.

We also consider the following random network models:

- **Barabási-Albert networks** [53]—preferential attachment networks with scale-free degree distribution. In our experiments we add 5 links with each new node (which results in the expected degree of 10) and we set the size of the initial clique to 5.
- **Erdős-Rényi networks** [54]—networks with the structure of a random graph. In our experiments we set the expected average degree to 10
- **Watts-Strogatz networks** [55]— networks exhibiting the small world property. In our experiments we set the expected average degree to 10 and the probability of rewiring to $\frac{1}{4}$.
- **Newman networks** [56]—networks with the scale-free structure, but without the preferential attachment property, generated using the configuration model. In our experiments we set the configuration model parameter to 2.3.
- **Prüfer networks** [57]—random trees generated using Prüfer sequences. We use sequences where each element is chosen uniformly at random from set $\{1, \ldots, n\}$.

7.2 Sensitivity of Centrality Measures

In Section 4 we investigated what changes in ranking are possible after addition or removal of an edge belonging to one of the three classes: edges incident with the evader $\zeta_1^\dagger$, edges between the neighbors of the evader $\zeta_2^\dagger$, and other edges $\zeta_3^\dagger$. However, even though our analysis resolved whether the ranking change can happen or not, it remains unclear how probable it is to happen. To resolve this issue, we now perform empirical analysis.

The networks that we consider are described in Section S4 of the supplementary materials. We use the real-life networks as they are, whereas for each of the random models we generate $1,000$ networks with 100 nodes. For network under consideration $G = (V, E)$ (whether real-life or random) we compute the initial rankings of degree, closeness, betweeness, and eigenvector centralities. For each pair of nodes $v, w \in V$ we consider a network G' resulting from either adding (v, w) to G (in case $(v, w) \notin E$), or removing (v, w) from G (in case $(v, w) \in E$). We then compute the rankings of all centralities in G' and for every node in the network we record how its ranking positions changed as a result of adding or removing (v, w) (notice that for every node in the network the edge (v, w) belongs either to class $\zeta_1^\dagger$, $\zeta_2^\dagger$, or $\zeta_3^\dagger$).

Some of the results of our simulations are presented in Figure 3, the remaining results can be found in Figures S7 and S8 in the supplementary materials. As it can be seen from the figures, in the vast majority of cases we are able to predict whether the ranking will increase or decrease with high certainty. For example, the removal of an edge belonging to the class $\zeta_1^\dagger$ almost always results in a decrease in ranking, while the addition of such edge in an increase. For edges belonging to classes $\zeta_2^\dagger$ and $\zeta_3^\dagger$ the possibility that the ranking will not change at all becomes significant (indeed, it is often the most probable outcome). However, if we disregard network modifications that do not affect centrality rankings, either the decrease or the increase in ranking is much more probable than its counterpart in most cases. Hence, even without knowledge necessary to compute the centrality ranking, e.g., information about the structure of the entire network, we can usually predict how a given network modification will affect the centrality rankings.

There remains a question about the magnitude of the ranking change, i.e., even if we can predict whether the centrality ranking of the evader will increase or decrease, can we predict the number of positions by which it will change? Figure 4 presents some of our results regarding the magnitude of the ranking change, the remaining results can be found in Figure S9 in the supplementary materials. As can be seen from the figure, adding or removing edges belonging to the class $\zeta_1^\dagger$, i.e., edge incident with the evader, not only gives the greatest chance of predicting whether the ranking will increase or decrease, but also said change will have the greatest magnitude. Class $\zeta_2^\dagger$ offers significantly smaller ranking changes, with class $\zeta_3^\dagger$ being the least effective in shifting the ranking of the evader. From the point of view of the evader these results are promising, as edges belonging to the class $\zeta_1^\dagger$ are also those the evader has the greatest amount of control over.

Fig. 3. Percentage of the edge modifications that result in a given change in the evader's centrality ranking. Results for random networks are presented as an average over 1000 networks with 100 nodes and the average degree of 10 generated using each model. Labels present values rounded to the nearest percent, values below 0.5% have been omitted for readability.

7.3 Seeker-Evader Game Experiments

In this section we present an empirical analysis of the seeker-evader game. We first describe the experimental procedure, before presenting results for real-life and random networks.

For a given network $G = (V, E)$ we first select the evader $v^\dagger \in V$ as the node with the highest average position in rankings generated by the four centrality measures considered in this study, i.e., degree, closeness, betweenness, and eigenvector centralities. We then generate strategies of the evader. The exact set of strategies under consideration depends on whether we consider real-life or random networks and is described in detail below.

For each evader strategy, we then execute it on the original network and compute the evader's ranking positions according to all four centrality, as well as the evader's influence according to the independent cascade model with the activation probability $p = 0.1$, and the linear threshold model with the uniform distribution of thresholds. We consider the set of evader's types $\Phi = \{.25, .50, .75\}$. For each evader type, we compute the value of U_e using the formula presented in Section 6.1, with the influence value of the evader being taken as an average over the two influence models mentioned above.

Finally, we compute the equilibrium of the seeker-evader game by using the MILP formulation presented in Section 6.2. To this end, we utilize the PuLP library version $2.6.0$ in Python version $3.7.9$.

We now empirically analyze the seeker-evader game in real-life networks. Given the reasonable size of these networks, we are able to generate all possible strategies of the evader with the budget of at most $b = 4$ (for WTC and Madrid datasets) or at most $b = 3$ (for the other datasets). In other words, we generate all k-subsets of edges between the evader and their neighbors $\hat{R}$ (that can be removed from the network) and non-edges between

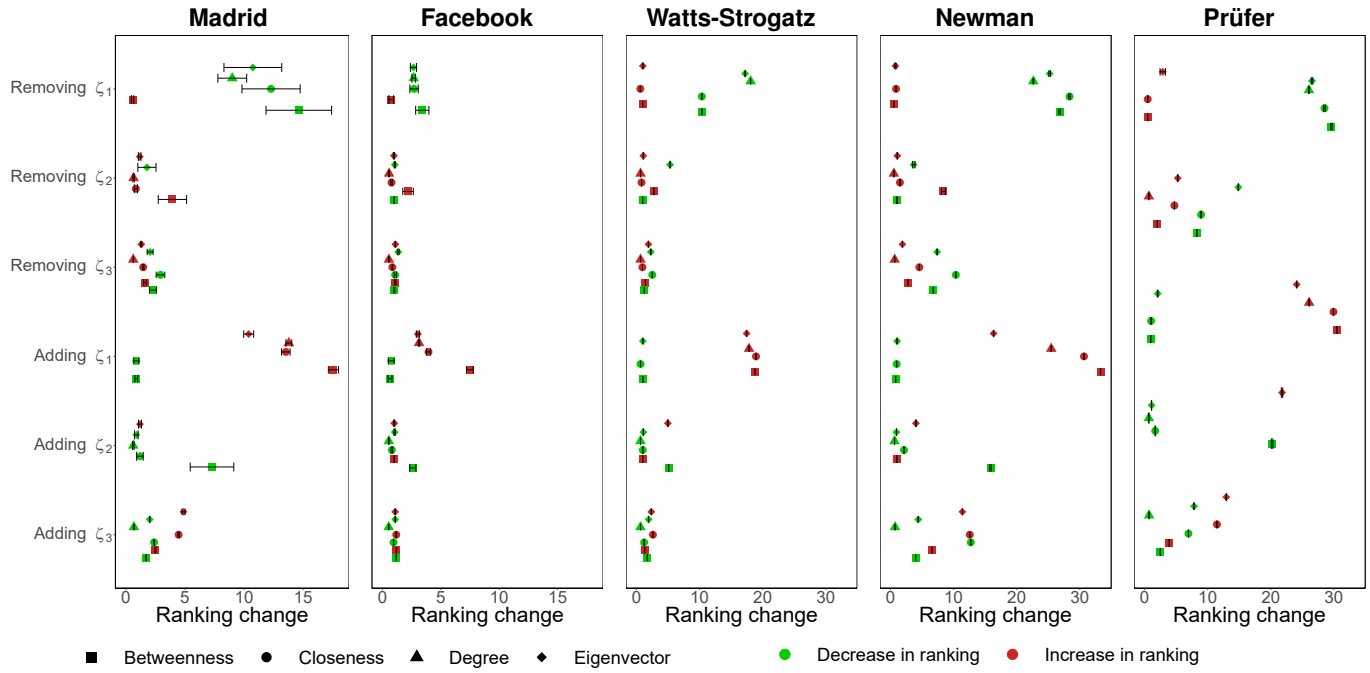


Fig. 4. Magnitude of the change in the evader's centrality ranking. The first row presents results for real-life networks. Results in the second row are presented as an average over 1000 networks with 100 nodes and the average degree of 10 generated using each model. Error bars (very narrow in some cases) correspond to 95% confidence intervals. Scales in each row are fixed for easier comparison.

the neighbors of the evader $\hat{A}$ (that can be added to the network) for $k \leq b$. To speed up the MILP computation, we remove all evader strategies that are dominated by another strategy, i.e., we remove a strategy if another strategy is at least as good according to all four centrality measures and both influence measures. The number of remaining (undominated) strategies is presented in the last column of Table 4. Notice that increasing the size of the network causes the growth of the set of potential evader strategies, but does not affect the effectiveness of the equilibria computation once the effective strategies against each centrality measure have been identified. Here, we consider smaller networks to be able to exhaustively search the space of all strategies. Below, we use MILP to compute equilibria in networks up to $100,000$ nodes while focusing on a set of particularly effective evader strategies.

First, we investigate the utility of the evader U_e depending on the composition of the strategy, i.e., the number of removed edges from $\hat{R}$, and the number of added edges from $\hat{A}$. Figure S10 in the supplementary materials presents the value of U_e with the ranking of the evader taken as the average ranking over the four centrality measures. As it can be seen from the figure, the greatest utility of the evader is consistently achieved for strategies that focus on edge removal, as opposed to edge addition. What is more, greater utility can be achieved by the evaders focused on their centrality ranking (greater values of ϕ), rather than by the evaders focused on their influence (smaller values of ϕ).

The results regarding the equilibria of the seeker-evader game are presented in Figure S11 in the supplementary materials. As it can be seen, in most networks the mixed strategy of the seeker involves almost exclusively using a particular centrality (the only exceptions being the WTC network). However, the exact centrality used by the seeker strongly depends on the network under consideration. Similarly, the evader usually uses the same strategy in a given network, no matter their type. The exact strategy choice depends on the network, although strong preference for the strategies focused on edge removal can be observed. As for the utility of the evader, we can see that evaders with greater values of ϕ, i.e., evaders more focused on the centrality ranking rather than influence value, are able to achieve better expected utility.

We now move to the empirical analysis of the seeker-evader game in random networks. In our simulations we generate networks with $100,000$ nodes. Given the significant size of the networks, we are unable to generate all possible strategies of the evader. Instead, we consider the repeated use of the hiding heuristic ROAM (*Remove One, Add Many*) proposed by Waniek et al. [11]. A single execution of ROAM with budget k (which we will denote ROAM(k)) comprises of removing the connection between the evader $v^\dagger$ and their neighbor with the greatest degree v^*, followed by connecting v^* to $k-1$ other neighbors of $v^\dagger$ with the lowest degrees. In our experiments with random networks we assume the total hiding budget of $b = 12$, and we consider evader strategies consisting of repeatedly executing one of the strategies ROAM(1), ROAM(2), ROAM(3), or ROAM(4). In other words, we consider the following four evader strategies:

- executing ROAM(1) twelve times, which in total removes twelve edges from the network,
- executing ROAM(2) six times, which in total removes six edges from the network and adds six edges to the network,
- executing ROAM(3) four times, which in total re-

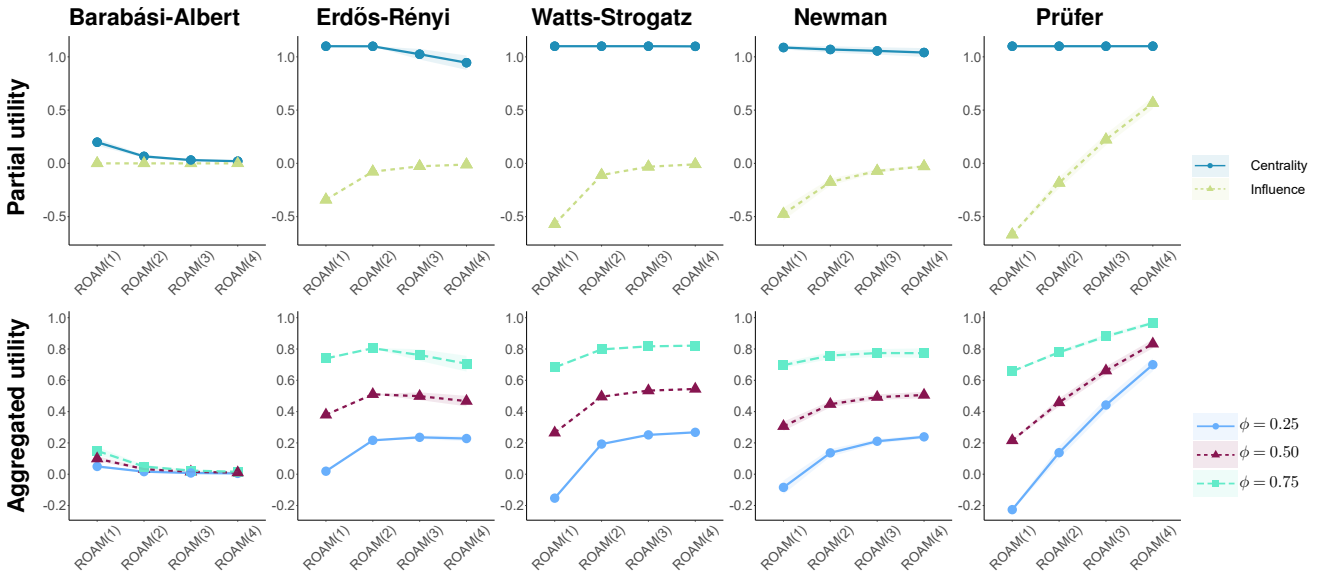


Fig. 5. Utility of the evader in random networks. Each column corresponds to a different network generation model. The first row presents the values of the utility based on centrality ranking $U_e^R(c, \xi)$ and the utility based on the influence $U_e^I(\xi)$. The second row presents the values of the aggregated utility $U_e(\phi, c, \xi,)$ for different types of evaders. In each plot the x-axis corresponds to the heuristic used by the evader, while the y-axis corresponds to the utility value. The results are presented as an average over 100 networks with $100,000$ nodes and the average degree of 10 generated for each model. The colored areas (very narrow) represent 95% confidence intervals.

moves four edges from the network and adds eight edges to the network,
- executing ROAM(4) three times, which in total removes three edges from the network and adds nine edges to the network,.

Figure 5 presents the results regarding the utility of the evader with the ranking of the evader taken as the average ranking over the four centrality measures. As can be seen, running ROAM(k) heuristic with greater values of k (i.e., focusing on edge addition, as opposed to edge removal) is slightly detrimental to the utility corresponding to centrality ranking, but it significantly improves the utility corresponding to the influence value. As for the evader types, we observe similar results to those in the real-life networks, with the evaders focused on their centrality ranking (greater values of ϕ), attaining greater utility values than the evaders focused on the influence (smaller values of ϕ). Altogether, the utility of the evader seems to be driven by their desire to maintain the influence over the network, as in terms of hiding from centrality measures, the considered strategies offer comparable performance. This is consistent with our findings regarding the equilibria of the game.

Figure 6 presents the results pertaining to the equilibria of the seeker-evader game in random networks. As it can be seen, the centrality measure used by the seeker varies significantly between the network types. As for the evader's strategy, ROAM(4) is the most commonly used, although for all network structures other ways of hiding are also in use. As for the utility of the evader, we can observe that evaders who are more focused on the centrality ranking rather than influence value achieve greater expected utility.

8 Conclusions

In this work, we analyzed the problem of strategically decreasing the evader's centrality ranking in a social network by performing local edge perturbations. First, we investigated what ranking changes are possible after adding or removing a single edge, depending on whether that edge is incident with the evader, between two of the evader's friends, or outside the evader's local neighborhood. In the case of degree centrality it is usually easy to predict both the direction of the ranking change (i.e., whether the ranking increases or decreases) and its magnitude (i.e., by how many positions does the ranking change). However, in the case of closeness and betweenness centrality measures, it is most often impossible to make such predictions. Second, we analyzed the computational complexity of the problem faced by the evader when adding or removing multiple edges, rather than a single one, to reduce the evader's ranking. We found that identifying the best possible way of hiding from a given centrality measure is most probably impossible (i.e., the corresponding decision problems are NP-complete), and that optimal solution is usually difficult to approximate (although we were able to identify a 2-approximation algorithm for the degree centrality). Third, we modeled the confrontation between the seeker and the evader as a Stackelberg game. We not only defined the strategies and the utility functions of both players, but we also showed a mixed-integer linear programming formulation of identifying an equilibrium. Fourth, we used simulations on real-life and randomly generated networks to study the effect of a single edge addition or removal on the evader's ranking. We found that even if it is impossible to predict said effect with absolute certainty, based on our results it is possible to make an educated guess in the vast majority of cases. Finally, we perform an empirical analysis of the seeker-evader game on networks. We found that while the exact strategies used by both players vary significantly between settings, the evader usually favors strategies including edge removal. Moreover, evaders who are more focused on their centrality ranking,

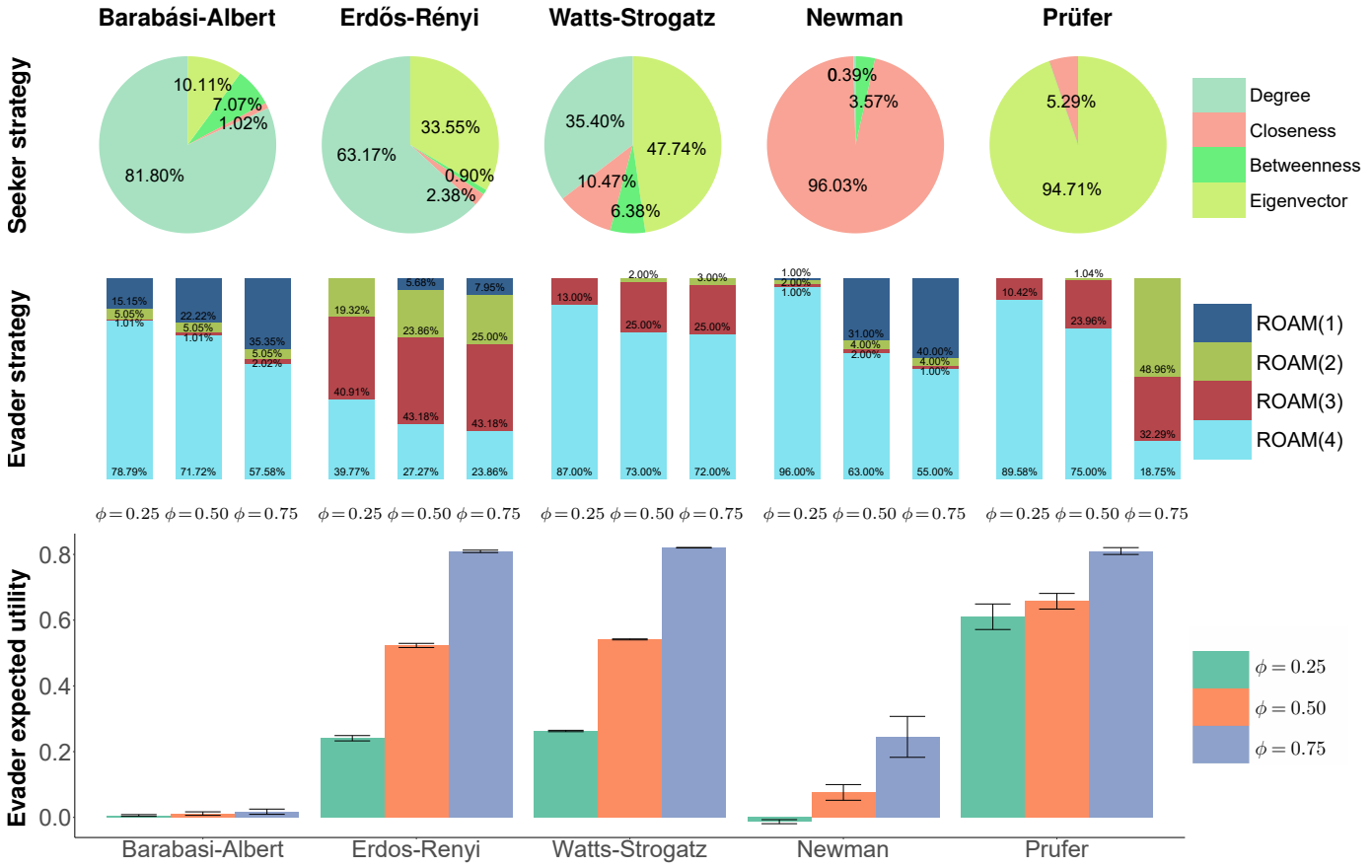


Fig. 6. The seeker-evader game equilibria in random networks. Each column corresponds to a different network generation model. Each pie chart in the first row presents the mixed strategy selected by the seeeker in an equilibrium of the seeker-evader game. Each group of bars in the second row presents the strategy the evader, with each bar corresponding to a different type of the evader. Each group of bars in the third row presents the expected utility of the evader, with each bar corresponding to a different type of the evader. The results are presented as an average over 100 networks with $100,000$ nodes and the average degree of 10 generated for each model. The error bars (very narrow in some cases) represent 95% confidence intervals.

as opposed to their influence value, can generally achieve greater utility. Altogether, our study provides a broad analysis of the strategic aspects of hiding from centrality measures in social networks.

Our work can be extended in a number of ways. First, in this study, we focus on the case of a single evader. However, one could consider a setting with multiple evaders, either cooperating with each other (i.e., wishing to hide as a group) or confronting each other (i.e., working towards exposing their opponents to the seeker while at the same time remaining hidden), with each variant of the setting posing unique challenges. In the case of cooperating evaders, the strategy space available to them would grow significantly, potentially requiring new computational methods to find effective sets of network modification, while the seeker might apply group centrality measures rather than the standard tools considered in this work. On the other hand, the case of adversarial evaders would greatly complicate the computation of equilibria, changing the decision of the evader from simply selecting the best response to taking into consideration the strategic incentives of all other evaders. Second, we considered the most popular centrality measures, as they are most widely implemented in actual software used for network analysis. However, one could study a broader portfolio of centrality measures, e.g., those based on game theory [58]. Alternatively, one could consider a similar setting in an different network class, e.g., in temporal networks [59] or multilayer networks [60]. Finally, equivalents of the seeker-evader game presented in this work can be developed for other social network analysis tools that already have hiding tools against unsuspecting seeker. Potential candidates include link prediction algorithms [19], [20], [21], node similarity measures [22], community detection algorithms [11], [23], [24], and source detection algorithms [25].

Acknowledgments

T. P. Michalak was supported by the Polish National Science Centre (grant 2016/23/B/ST6/03599). Y. Vorobeychik was partly supported by the National Science Foundation (IIS-1903207, IIS-1905558, IIS-2214141) and Army Research Office (MURI W911NF1810208). K. Zhou was supported by National Natural Science Foundation of China (No. 62106210) and Hong Kong RGC Project (No. PolyU25210821). M. Waniek was supported by the Polish National Science Centre (grant 2015/17/N/ST6/03686). This research was carried out on the High Performance Computing resources at New York University Abu Dhabi.

REFERENCES

[1] A. Bavelas, "A mathematical model for group structures," *Human organization*, vol. 7, no. 3, pp. 16–30, 1948.

[2] N. E. Friedkin, "Theoretical foundations for centrality measures," *American journal of Sociology*, vol. 96, no. 6, pp. 1478–1504, 1991.

[3] C. Morselli, "Assessing vulnerable and strategic positions in a criminal network," *Journal of Contemporary Criminal Justice*, vol. 26, no. 4, pp. 382–392, 2010.

[4] R. H. Lindelauf, H. J. Hamers, and B. Husslage, "Cooperative game theoretic centrality analysis of terrorist networks: The cases of jemaah islamiyah and al qaeda," *European Journal of Operational Research*, vol. 229, no. 1, pp. 230–238, 2013.

[5] M. Waniek, J. Woźnica, K. Zhou, Y. Vorobeychik, T. Rahwan, and T. P. Michalak, "Strategic evasion of centrality measures," in *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, ser. AAMAS '21. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2021, p. 1389–1397.

[6] M. Papagelis, "Refining social graph connectivity via shortcut edge addition," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 10, no. 2, pp. 1–35, 2015.

[7] B. Wilder, H.-C. Ou, K. de la Haye, and M. Tambe, "Optimizing network structure for preventative health." in *AAMAS*, 2018, pp. 841–849.

[8] B.-Y. Hsu, C.-Y. Shen, and X. Yan, "Network intervention for mental disorders with minimum small dense subgroups," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 5, pp. 2121–2136, 2019.

[9] P. Crescenzi, G. D'angelo, L. Severini, and Y. Velaj, "Greedily improving our own closeness centrality in a network," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 11, no. 1, pp. 1–32, 2016.

[10] H.-J. Hung, W.-C. Lee, D.-N. Yang, C.-Y. Shen, Z. Lei, and S.-M. Chow, "Efficient algorithms towards network intervention," in *Proceedings of The Web Conference 2020*, 2020, pp. 2021–2031.

[11] M. Waniek, T. P. Michalak, M. J. Wooldridge, and T. Rahwan, "Hiding individuals and communities in a social network," *Nature Human Behaviour*, vol. 2, no. 2, p. 139, 2018.

[12] M. Waniek, T. P. Michalak, T. Rahwan, and M. Wooldridge, "On the construction of covert networks," in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2017, pp. 1341–1349.

[13] P. Dey and S. Medya, "Covert networks: How hard is it to hide?" in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. Montreal, Canada: IFAAMAS, 2019, pp. 628–637.

[14] M. Waniek, T. P. Michalak, M. Wooldridge, and T. Rahwan, "How members of covert networks conceal the identities of their leaders," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 13, no. 1, pp. 1–29, 2021.

[15] T. Was, M. Waniek, T. Rahwan, and T. Michalak, "The manipulability of centrality measures-an axiomatic approach," in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. Auckland, New Zealand: AAMAS, 2020, pp. 1467–1475.

[16] M. Waniek, T. Michalak, and T. Rahwan, "Hiding in multilayer networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34. New York, USA: AAAI, 2020, pp. 1021–1028.

[17] M. Waniek, P. Holme, and T. Rahwan, "Hiding in temporal networks," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 3, pp. 1645–1657, 2022.

[18] W. Jin, Y. Li, H. Xu, Y. Wang, S. Ji, C. Aggarwal, and J. Tang, "Adversarial attacks and defenses on graphs: A review, a tool and empirical studies," 2020.

[19] S. Yu, M. Zhao, C. Fu, J. Zheng, H. Huang, X. Shu, Q. Xuan, and G. Chen, "Target defense against link-prediction-based attacks via evolutionary perturbations," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 2, pp. 754–767, 2019.

[20] M. Waniek, K. Zhou, Y. Vorobeychik, E. Moro, T. P. Michalak, and T. Rahwan, "How to hide one's relationships from link prediction algorithms," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[21] K. Zhou, T. P. Michalak, M. Waniek, T. Rahwan, and Y. Vorobeychik, "Attacking similarity-based link prediction in social networks," in *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems*. Montreal, Canada: AAMAS, 2019, p. 305–313.

[22] P. Dey and S. Medya, "Manipulating node similarity measures in networks," 2020.

[23] J. Li, H. Zhang, Z. Han, Y. Rong, H. Cheng, and J. Huang, "Adversarial attack on community detection by hiding individuals," in *Proceedings of The Web Conference 2020*, 2020, pp. 917–927.

[24] J. Jia, B. Wang, X. Cao, and N. Z. Gong, "Certified robustness of community detection against adversarial structural perturbation via randomized smoothing," in *Proceedings of The Web Conference 2020*, 2020, pp. 2718–2724.

[25] M. Waniek, M. Cebrian, P. Holme, and T. Rahwan, "Social diffusion sources can escape detection," *iScience*, 2022.

[26] M. T. Godziszewski, T. P. Michalak, M. Waniek, T. Rahwan, K. Zhou, and Y. Zhu, "Attacking similarity-based sign prediction," in *2021 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2021, pp. 1072–1077.

[27] L. Sun, Y. Dou, C. Yang, J. Wang, P. S. Yu, L. He, and B. Li, "Adversarial attack and defense on graph data: A survey," *arXiv preprint arXiv:1812.10528*, 2018.

[28] W. Jin, Y. Li, H. Xu, Y. Wang, S. Ji, C. Aggarwal, and J. Tang, "Adversarial attacks and defenses on graphs," *ACM SIGKDD Explorations Newsletter*, vol. 22, no. 2, pp. 19–34, 2021.

[29] A. Bojchevski and S. Günnemann, "Adversarial attacks on node embeddings via graph poisoning," in *International Conference on Machine Learning*. Long Beach, USA: PMLR, 2019, pp. 695–704.

[30] B. Wang and N. Z. Gong, "Attacking graph-based classification via manipulating the graph structure," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. New York, USA: ACM, 2019, pp. 2023–2040.

[31] D. Zügner and S. Günnemann, "Adversarial attacks on graph neural networks via meta learning," in *International Conference on Learning Representations*, 2018.

[32] H. Dai, H. Li, T. Tian, X. Huang, L. Wang, J. Zhu, and L. Song, "Adversarial attack on graph structured data," in *International conference on machine learning*. Stockholm, Sweden: PMLR, 2018, pp. 1115–1124.

[33] H. Wu, C. Wang, Y. Tyshetskiy, A. Docherty, K. Lu, and L. Zhu, "Adversarial examples for graph data: deep insights into attack and defense," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 2019, pp. 4816–4823.

[34] Y. Ma, S. Wang, T. Derr, L. Wu, and J. Tang, "Attacking graph convolutional networks via rewiring," *arXiv preprint arXiv:1906.03750*, 2019.

[35] M. Waniek, T. P. Michalak, and A. Alshamsi, "Strategic attack & defense in security diffusion games," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 1, pp. 1–35, 2019.

[36] M. E. Shaw, "Group structure and the behavior of individuals in small groups," *The Journal of Psychology*, vol. 38, no. 1, pp. 139–149, 1954.

[37] M. A. Beauchamp, "An improved index of centrality," *Behavioral Science*, vol. 10, no. 2, pp. 161–163, 1965.

[38] J. M. Anthonisse, "The rush in a directed graph," *Stichting Mathematisch Centrum. Mathematische Besliskunde*, vol. 71, no. BN 9, pp. 1–10, 1971.

[39] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.

[40] P. Bonacich, "Power and centrality: A family of measures," *American journal of sociology*, vol. 92, no. 5, pp. 1170–1182, 1987.

[41] J. Goldenberg, B. Libai, and E. Muller, "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*, vol. 9, no. 3, pp. 1–18, 2001.

[42] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, USA: ACM, 2003, pp. 137–146.

[43] M. O. Jackson and B. W. Rogers, "Meeting strangers and friends of friends: How random are social networks?" *American Economic Review*, vol. 97, no. 3, pp. 890–915, 2007.

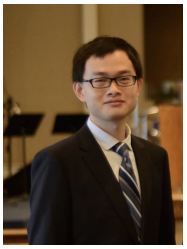
[44] G. Bianconi, R. K. Darst, J. Iacovacci, and S. Fortunato, "Triadic closure as a basic generating mechanism of communities in complex networks," *Phys. Rev. E*, vol. 90, p. 042806, Oct 2014. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.90.042806

[45] M. Adrian, M. Moreno, S. Nicodimos, E. McCauley, and A. Vander Stoep, "Research strategy for health sciences: Facebook friend request is non-differentially accepted in a diverse, young adult

population," *Nursing & health sciences*, vol. 21, no. 1, pp. 71–77, 2019.
[46] H. Von Stackelberg, *Market structure and equilibrium*. Springer Science & Business Media, 2010.
[47] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordonez, and S. Kraus, "Playing games for security: An efficient exact algorithm for solving bayesian stackelberg games," in *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, 2008, pp. 895–902.
[48] V. E. Krebs, "Mapping networks of terrorist cells," *Connections*, vol. 24, no. 3, pp. 43–52, 2002.
[49] B. Hayes, "Connecting the dots can the tools of graph theory and social-network studies unravel the next big plot?" *American Scientist*, vol. 94, no. 5, pp. 400–404, 2006.
[50] S. Koschade, "A social network analysis of jemaah islamiyah: The applications to counterterrorism and intelligence," *Studies in Conflict & Terrorism*, vol. 29, no. 6, pp. 559–575, 2006.
[51] D. Wright, "The John Jay & ARTIS transnational terrorism database," http://doitapps.jjay.cuny.edu/jjatt/, 2009, accessed: 2016-10-28.
[52] J. Leskovec and J. J. Mcauley, "Learning to discover social circles in ego networks," in *Advances in neural information processing systems*, 2012, pp. 539–547.
[53] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
[54] P. Erdős and A. Rényi, "On random graphs i." *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.
[55] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[56] M. E. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
[57] H. Prüfer, "Neuer beweis eines satzes über permutationen," *Archiv der Mathematik und Physik. 3. Reihe*, vol. 27, 01 1918.
[58] O. Skibski, T. P. Michalak, and T. Rahwan, "Axiomatic characterization of game-theoretic centrality," *Journal of Artificial Intelligence Research*, vol. 62, pp. 33–68, 2018.
[59] P. Holme and J. Saramäki, "Temporal networks," *Physics reports*, vol. 519, no. 3, pp. 97–125, 2012.
[60] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks," *Journal of complex networks*, vol. 2, no. 3, pp. 203–271, 2014.

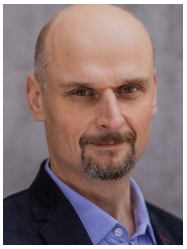
Marcin Waniek Marcin Waniek is a Post-Doctoral Associate at Computer Science Department, New York University Abu Dhabi. His research interests lie in social network analysis, graph theory, computational complexity theory, artificial intelligence, and game theory. He received his Ph.D. in Computer Science from the University of Warsaw, Poland. It earned him the Polish Artificial Intelligence Society Award for the Best Ph.D. Dissertation in Artificial Intelligence in 2017.

Jan Woźnica Jan Woźnica graduated from Faculty of Economic Sciences, University of Warsaw in 2017. He is a software developer, currently at Scanye.



Kai Zhou is an Assistant Professor in the Dept. of Computing at The Hong Kong Polytechnic University. His research interests center around security with emphasis on adversarial network analysis, adversarial machine learning, and data security and privacy. He earned his Ph.D. from the Department of Electrical and Computer Engineering at Michigan State University in 2018. He worked as a Postdoctoral Research Associate in the Department of Computer Science at Washington University in St. Louis from 2018 to 2020.

Yevgeniy Vorobeychik is a Professor of Computer Science and Engineering at Washington University in Saint Louis. Previously, he was an Assistant Professor of Computer Science at Vanderbilt University. Between 2008 and 2010 he was a post-doctoral research associate at the University of Pennsylvania Computer and Information Science department. He received Ph.D. (2008) and M.S.E. (2004) degrees in Computer Science and Engineering from the University of Michigan, and a B.S. degree in Computer Engineering from Northwestern University. His work focuses on game theoretic modeling of security and privacy, adversarial machine learning, algorithmic and behavioral game theory and incentive design, optimization, agent-based modeling, complex systems, network science, and epidemic control. Dr. Vorobeychik received an NSF CAREER award in 2017, and was invited to give an IJCAI-16 early career spotlight talk. He also received several Best Paper awards.



Tomasz P. Michalak is a lecturer at the Institute of Informatics, the Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, and the Team Leader at the IDEAS NCBR research institute. He conducted research at the Department of Computer Science, University of Oxford, School of Engineering and Computer Science, University of Southampton, Department of Computer Science, University of Liverpool, and Faculty of Applied Economics, University of Antwerp. He graduated from the Faculty of Economic Sciences, University of Warsaw and holds a Ph.D. in economics from the Faculty of Applied Economics, University of Antwerp.

Talal Rahwan is an associate professor of Computer Science and the director of the Data Science and AI Lab at New York University Abu Dhabi. He earned his Ph.D. in Computer Science in 2007 from The University of Southampton, UK. His thesis earned him the British Computer Society's Distinguished Dissertation Award, which annually recognizes the most outstanding Ph.D. thesis in the UK in Computer Science. He was selected by the IEEE Computer Society as one of the ten most promising, young Artificial Intelligence (AI) researchers in the world. His work appeared in major academic journals including PNAS, Nature Communications, Nature Human Behaviour, and Nature Machine Intelligence.