# Adversarial Influence Maximization

Justin Khim
Wharton Department of Statistics
University of Pennsylvania
Philadelphia, PA 19103
Email: jkhim@wharton.upenn.edu

Varun Jog
Department of Electrical & Computer Engineering
University of Wisconsin-Madison
Madison, WI 53706
Email: jog@wisc.edu

Po-Ling Loh
Department of Statistics
University of Wisconsin-Madison
Madison, WI 53706
Email: ploh@stat.wisc.edu

***Abstract*—We consider the problem of influence maximization in fixed networks for contagion models in an adversarial setting. The goal is to select an optimal set of nodes to seed the influence process, such that the number of influenced nodes at the conclusion of the campaign is as large as possible. We formulate the problem as a repeated game between a player and adversary, where the adversary specifies the edges along which the contagion may spread, and the player chooses sets of nodes to influence in an online fashion. We establish upper and lower bounds on the minimax pseudo-regret in both undirected and directed networks.**

*A full version of this paper is accessible at:*
https://arxiv.org/abs/1611.00350

## I. Introduction

Many data sets in contemporary scientific applications possess some underlying network structure [20]. Popular examples include data collected from social media websites such as Facebook and Twitter [1], [18], or electrocortical recordings gathered from a network of firing neurons [23]. An important application of network science arises in marketing, where researchers have studied the importance of word-of-mouth advertising for decades [13]. More recently, empirical studies have suggested that word-of-mouth marketing has a significant effect in online social networks [2], [22]. At the same time, computer scientists have analyzed the problem of viral marketing from an optimization-theoretic perspective [7], [10], [17], where the goal is to select an optimal set of influencers to encourage product adoption in an online social network. This has led to rigorous theoretical guarantees that hold for stochastic models of word-of-mouth advertising inspired by physics and epidemiology, and the scope of the spread is quantified using a notion known as influence [14]. In social networks, edges represent potential interactions between individuals, and the problem of influence maximization corresponds to identifying subsets of individuals on which to impress an idea so that information spreads as widely as possible.

Formally, the influence of a subset of nodes is defined as the expected number of influenced individuals in a network at the conclusion of a spread, starting from an initial configuration where only the specified nodes are influenced. Even when the influence function is assumed to be computable for any subset using a black-box method in unit time, it is not clear that influence maximization may be performed (exactly or approximately) in polynomial time, since searching over all subsets of $k$ nodes is exponential in the number of nodes. Accordingly, the body of work in theoretical computer science has mostly focused on specific spreading models that give rise to nice properties such as submodularity, implying that a greedy algorithm for influence maximization leads to a constant-factor approximation of the optimal set [4], [14], [15]. A significant shortcoming in the analysis of stochastic spreading models is the fact that the parameters characterizing the spread of influence are generally assumed to be known, allowing for approximate evaluation of the influence function. However, such an assumption is not always practical, and one might even question a scientist's prior knowledge of the precise network structure.

To address these issues, some authors have studied the question of accurately learning the influence function itself in a stochastic spreading model based on observing multiple rounds of infection [12], [16], [19]. Another approach involves a notion of "robust influence maximization," where the parameters are only specified to lie in fixed confidence sets [8], [11]. A third approach is to frame influence maximization as a stochastic semi-bandit problem over many rounds [9], [21], [25], [26]. Such methods may also be model-dependent, meaning that an algorithm designed for the independent cascade model may lead to a significantly suboptimal solution if the influence spread actually follows linear threshold model.

In this paper, we take a rather different approach toward the problem of unknown spreading parameters that also avoids assumptions about a particular spreading mechanism. As discussed in more detail in Section II, we only assume knowledge of an underlying fixed graph representing the paths along which a influence may spread, where the case of no prior knowledge corresponds to a complete graph. We formulate the influence maximization problem as an online game, where a "player" must make sequential decisions about the next seed set to choose based on observing the behavior of the spread in previous "rounds" of the game. Here, a round represents a particular instance of an influence process initialized from the specific seed nodes from beginning to end. We allow an "adversary" to choose the path of influence on each round in a completely arbitrary manner, as long as the process may only spread along edges of the graph—in particular, this setting subsumes the stochastic models usually adopted in the influence maximization literature, while allowing for much more general spreading mechanisms. Thus, instead of

978-1-5386-9291-2/19/$31.00 ©2019 IEEE

simply trying to maximize the aggregate number of influenced vertices across all rounds, we seek to develop player strategies that bound the "regret" of the player. Such notions are taken from the literature on multi-armed bandits and online learning theory [3], [6], and adapted to the present setting.

Our main contribution is to derive upper and lower bounds on the pseudo-regret for various adversarial and player strategies. We study both directed and undirected networks, where in the latter setting, contagion is allowed to spread in both directions when an edge is chosen by the adversary. Furthermore, we derive information-theoretic lower bounds for the minimax pseudo-regret when the underlying network is a complete graph. Our upper and lower bounds match up to constant factors in the case of directed networks. Notably, the bounds also agree with the usual rate for pseudo-regret in multi-armed bandits, showing that no new information is gained by the player by exploiting network structure. On the other hand, a gap exists between our upper and lower bounds for undirected networks, leaving open the possibility that the player may leverage the additional information from the network to incur less regret. Additionally, the constant factor on the upper bound may be slightly improved, providing further evidence that graph structure may be exploited. Finally, we demonstrate how to extend our upper bounds to the setting where the player is allowed to choose multiple source vertices on each round based on a general online greedy algorithm [24].

The remainder of our paper is organized as follows: In Section II, we provide some important background on online learning theory and formally define the adversarial spreading model and notions of regret to be studied in our paper. In Section III, we present upper and lower bounds for pseudo-regret in the adversarial setting. We conclude the paper with a selection of open research questions in Section IV. All proofs, as well as a more technical discussion of related work, is contained in the appendices of the full version of the paper.

**Notation.** For a set $A$, let $2^A$ denote the power set of $A$. When we want to specify that we are taking the expectation with respect to a particular distribution $p$ of some random variable $X$, we write $\mathbb{E}_{X\sim p}$. In particular, we often write $\mathbb{E}_{\mathcal{S}\sim p}$ to mean the expectation taken over the player's actions for a fixed set of adversarial actions, which is the same as the conditional expectation with respect to the adversary's actions. Similarly, we write $\mathbb{E}_{\mathcal{A}}$ to indicate the conditional expectation with respect to a fixed set of player actions.

## II. Background and preliminaries

We begin by formally defining the repeated game between the player and adversary and the types of strategies we will analyze in our paper. Next, we introduce the notions of regret we will study, and then connect our setting to related work in the learning theory literature.

### A. Adversarial repeated games

Consider a fixed graph $G = (V, E)$ on $n$ vertices, which may be directed or undirected. The adversarial influence maximization problem may be described as follows: Repeatedly over $T$ rounds, the player selects an influence seed set $\mathcal{S}_t \subseteq V$, with $|\mathcal{S}_t| = k$, for $t = 1, \ldots, T$. At the same time, the adversary designates a subset of edges $\mathcal{A}_t \subseteq E$ to be "open." A node is considered to be influenced at time $t$ if and only if it is an element of $\mathcal{S}_t$ or is reachable from $\mathcal{S}_t$ via a path of open edges. Note that in the context of influence spreading, the open edges correspond to ties over which influence propagates in that round—importantly, influence only has an opportunity to be transmitted between individuals that interact in the network, but may not necessarily spread over a particular connection on a specific round. In the case when $G$ is an undirected graph, designating an edge to be open allows an influence campaign to spread in both directions. Furthermore, in the directed case, edges may exist in both directions between a given pair of nodes, in which case the adversary may designate both, one, or neither of the edges to be open. For an open edge set $A \subseteq E$ and influence seed set $S \subseteq V$, we define $f(A, S)$ to be the fraction of vertices in the graph lying in the influenced set.

To connect our model to the canonical setting of influence maximization, note that [15] proposed a very general class of influence models called triggering models, which include the independent cascade and the linear threshold models as special cases. At the beginning of the influence campaign, each node chooses a random "triggering" subset of neighbors according to a particular rule, and the incoming edges from those neighbors are designated to be "active." A vertex becomes influenced during the course of the process if and only if a path of active edges exists connecting that vertex to a vertex in the seed set. Thus, triggering models correspond to a special case of our framework, in which the edge sets are chosen in an i.i.d. manner from round to round, and the probability distribution over the edges is determined by the probability rule through which edges are assigned to be active (e.g., according to the linear threshold or independent cascade models).

Next, we describe the classes of strategies $\mathcal{A} = \{\mathcal{A}_t\}$ and $\mathcal{S} = \{\mathcal{S}_t\}$ available to the adversary and player. We assume that the adversary is *oblivious* of the player's actions; i.e., at time $t = 0$, the adversary must decide on the (possibly random) strategy $\mathcal{A}$. We use $\mathscr{A}$ to denote the set of oblivious adversary strategies and $\mathscr{A}_d$ to denote the set of deterministic adversary strategies. Turning to the classes of player strategies, we allow the player to choose his or her action at time $t$ based on the feedback provided in response to the joint actions made by the player and adversary on preceding time steps. Although the player knows the edge set $E$ of the underlying graph, we assume that the player only observes the status of edges $(i, j)$ such that either $i$ or $j$ is in the reach of $\mathcal{S}_t$ (in the undirected case), and the player observes the status of every edge $(i, j)$ such that $i$ is in the reach of $\mathcal{S}_t$ (in the directed case). In other words, whereas the player cannot observe the subset of all edges that *would have* propagated influence in the network, he or she will know which edges transmitted influence if reached by the influence cascade initialized using his or her seed set.

Formally, we write $\mathscr{I}(\mathcal{A}_t, \mathcal{S}_t)$ to denote the set of edges with status known to the player (i.e., all edges in the

subgraph induced by $\mathcal{A}_t$ belonging to connected components containing nodes in $\mathcal{S}_t$), and we denote $\mathscr{I}^t = (\mathscr{I}(\mathcal{A}_1, \mathcal{S}_1), \ldots, \mathscr{I}(\mathcal{A}_t, \mathcal{S}_t))$. If $\mathcal{A}_t$ is chosen via a stochastic model such as the independent cascade model with discrete time steps for influence campaign $t$, our setup technically allows the player knowledge of the status of an edge between two vertices $u$ and $v$ if both were actually influenced by some other vertex $w$. Realistically we would not want the status of edge $(u, v)$ to be returned as feedback, and we could enforce this by positing a model of how each influence campaign proceeds. However, this distinction does not affect our results or algorithms, and so we do not further restrict the feedback $\mathscr{I}(\mathcal{A}_t, \mathcal{S}_t)$.

The player can only make decisions based on the feedback observed in previous rounds, so any allowable player strategy $\{\mathcal{S}_t\}$ has the property that $\mathcal{S}_t$ is a function of $\mathscr{I}^{t-1}$ (possibly with additional randomization). We denote the class of all player strategies by $\mathscr{P}$, and denote the subclass of all deterministic player strategies by $\mathscr{P}_d$, meaning that $\mathcal{S}_t$ is a deterministic function of $\mathscr{I}^{t-1}$. Note that strategies $\mathcal{S}_t \in \mathscr{P}_d$ may still be random, due to possible randomization of the adversary, but *conditioned* on $\mathscr{I}^{t-1}$, the choice of $\mathcal{S}_t$ is deterministic.

### B. Minimax regret

The player wishes to devise a strategy that maximizes the aggregate number of influenced nodes up to time $T$. Using the notation from the previous section, we define the *regret* of the player to be

$$R_T(\mathcal{A}, \mathcal{S}) = \sum_{t=1}^{T} f(\mathcal{A}_t, \mathcal{S}_*) - \sum_{t=1}^{T} f(\mathcal{A}_t, \mathcal{S}_t), \tag{1}$$

where

$$\mathcal{S}_* = \operatorname*{arg\,max}_{S:|S|=k} \sum_{t=1}^{T} f(\mathcal{A}_t, S)$$

is the optimal fixed set that the player would have chosen in hindsight with full knowledge of the adversary's strategy.

Note that the regret $R_T(\mathcal{A}, \mathcal{S})$ may be a random quantity due to randomness in both the adversary's or player's strategies. Accordingly, we will seek to control the *pseudo-regret*

$$\overline{R}_T(\mathcal{A}, \mathcal{S}) := \max_{S:|S|=k} \mathbb{E}_{\mathcal{A},\mathcal{S}} \left[ \sum_{t=1}^{T} f(\mathcal{A}_t, S) - \sum_{t=1}^{T} f(\mathcal{A}_t, \mathcal{S}_t) \right], \tag{2}$$

where the expectation in equation (2) is taken with respect to potential randomization in both $\mathcal{A}$ and $\mathcal{S}$. As in the standard learning theory literature [5], recall that the expected regret and pseudo-regret are generally related via the inequality

$$\overline{R}_T(\mathcal{A}, \mathcal{S}) \leq \mathbb{E}[R_T(\mathcal{A}, \mathcal{S})],$$

although if $\mathcal{A} \in \mathscr{A}_d$, we have $\overline{R}_T(\mathcal{A}, \mathcal{S}) = \mathbb{E}[R_T(\mathcal{A}, \mathcal{S})]$. Our interest in the pseudo-regret rather than the expected regret is purely motivated by the fact that the former quantity is often easier to bound than the latter and that this simplification is common in the literature on bandits.

Finally, we introduce the *scaled regret*

$$R^\alpha(\mathcal{A}, \mathcal{S}) = \alpha \sum_{t=1}^{T} f(\mathcal{A}_t, \mathcal{S}_*) - \sum_{t=1}^{T} f(\mathcal{A}_t, \mathcal{S}_t), \tag{3}$$

and the analogous quantity

$$\overline{R}^\alpha_T(\mathcal{A}, \mathcal{S}) = \max_{S:|S|=k} \mathbb{E}_{\mathcal{A},\mathcal{S}} \left[ \alpha \sum_{t=1}^{T} f(\mathcal{A}_t, S) - \sum_{t=1}^{T} f(\mathcal{A}_t, \mathcal{S}_t) \right].$$

Note that $\alpha = 1$ corresponds to the unscaled version. Our interest in the expression (3) is again for theoretical purposes, since we may obtain convenient upper bounds on the scaled pseudo-regret in the case $\alpha = 1 - \frac{1}{e}$ using an online greedy algorithm. Note that when $k > 1$, the benchmark greedy algorithms used for influence maximization in the stochastic spreading setting are also only guaranteed to achieve a $\left(1 - \frac{1}{e}\right)$-approximation of the truth, so in some sense, the scaled regret (3) only requires the player to perform comparably well in relation to the appropriately scaled optimal strategy.

## III. Main results

In this section, we provide upper and lower bounds for the pseudo-regret. Specifically, we focus on the quantity

$$\inf_{\mathcal{S} \in \mathscr{P}} \sup_{\mathcal{A} \in \mathscr{A}} \overline{R}^\alpha_T(\mathcal{A}, \mathcal{S}),$$

where the supremum is taken over the class of adversarial strategies, and the infimum is taken over the class of player strategies based on the feedback model we have described. In other words, we wish to characterize the hardness of the influence maximization problem in terms of the player's best possible strategy measured with respect to the worst-case game.

A rough outline of our approach is as follows: We establish upper bounds by presenting particular strategies for the player that ensure an appropriately bounded regret under all adversarial strategies. For lower bounds, the general technique is to provide an ensemble of possible actions for the adversary that are difficult for the player to distinguish in the influence maximization problem, which forces the player to incur a certain level of regret.

### A. Undirected graphs

We begin by deriving regret upper bounds for undirected graphs. We initially restrict our attention to the case $k = 1$. The proposed player strategy for $k > 1$, and corresponding regret bounds, builds upon the results in the single-source setting.

*1) Upper bounds for a single source:* Consider a randomized player strategy that selects $\mathcal{S}_t = \{i\}$ with probability $p_{i,t}$. The paper [5] suggests a method based on the Online Stochastic Mirror Descent (OSMD) algorithm, which is specified by loss estimates $\{\ell_{i,t}\}$ and learning rates $\{\eta_t\}$, as well as a Legendre function $F$. Here, we comment on the losses, and in order to avoid excessive technicalities, we defer additional details of the OSMD algorithm to the appendix.

The most basic loss estimate, which follows from standard bandit theory and ignores all information about the graph, is

$$\widehat{\ell}^{\text{node}}_{i,t} = \frac{\ell_{i,t}}{p_{i,t}} \mathbf{1}_{\mathcal{S}_t = \{i\}}, \tag{4}$$

where $\ell_{i,t} = 1 - f(\mathcal{A}_t, \{i\})$ is the loss incurred if the player were to choose $\mathcal{S}_t = \{i\}$. Importantly, $\widehat{\ell}^{\text{node}}_{i,t}$ is always computable for any choice the player makes at time $t$ and is an unbiased estimate of $\ell_{i,t}$.

On the other hand, if $\mathcal{S}_t = \{i\}$ and another node $j$ is influenced (i.e., in the connected component formed by the open edges of $\mathcal{A}_t$), the player also knows the loss that would have been incurred if $\mathcal{S}_t = \{j\}$, since $f(\mathcal{A}_t, \{i\}) = f(\mathcal{A}_t, \{j\})$. This motivates an alternative loss estimate that is nonzero even when $\mathcal{S}_t \neq \{i\}$. In particular, we may express

$$\ell_{i,t} = \frac{1}{n} \sum_{j \neq i} \ell^t_{i,j},$$

where $\ell^t_{i,j}$ is the indicator that $i$ and $j$ are in different connected components formed by the open edges of $\mathcal{A}_t$. We then define

$$\widehat{\ell}^{\text{sym}}_{i,t} = \frac{1}{n} \sum_{j \neq i} \ell^t_{i,j} \frac{Z_{ij}}{p_{i,t} + p_{j,t}},$$

where $Z_{ij} = \mathbf{1}_{\mathcal{S}_t \cap \{i,j\} \neq \emptyset}$. Furthermore, the estimator $\widehat{\ell}^{\text{sym}}_{i,t}$ is always computable by the player, since the value of $\ell^t_{i,j}$ is known by the player whenever $\mathcal{S}_t$ is known. We call $\widehat{\ell}^{\text{sym}}_{i,t}$ the *symmetric loss*. Now, we state the following regret bounds:

**Theorem 1** (Symmetric loss, OSMD)**.** *Suppose the player uses the strategy $\mathcal{S}^{sym}_{OSMD}$ corresponding to OSMD with the symmetric loss $\hat{\ell}^{sym}$ and appropriate parameters. Then the pseudo-regret satisfies the bound*

$$\sup_{\mathcal{A} \in \mathscr{A}} \overline{R}_T(\mathcal{A}, \mathcal{S}^{sym}_{OSMD}) \leq 2^{\frac{1}{4}} \sqrt{Tn}.$$

**Remark 1.** *It is instructive to compare the result of Theorem 1 with analogous regret bounds for generic multi-armed bandits. When the OSMD algorithm is run with the loss estimates (4), standard analysis establishes an upper bound of $2^{\frac{3}{2}}\sqrt{Tn}$. Thus, using the symmetric loss, which leverages the graphical nature of the problem, produces slight gains.*

*2) Lower bounds:* We now establish lower bounds for the pseudo-regret in the case $k = 1$. This furnishes a better understanding of the hardness of the adversarial influence maximization problem.

The intrinsic difficulty of online influence maximization may vary widely depending on the topology of the underlying graph, and methods for deriving lower bounds may also differ accordingly. In the case of a complete graph, we have the following result:

**Theorem 2.** *Suppose $G = \mathcal{K}_n$ is the complete graph on $n \geq 3$ vertices. Then the pseudo-regret satisfies the lower bound*

$$\frac{2}{243}\sqrt{T} \leq \inf_{\mathcal{S} \in \mathscr{P}} \sup_{\mathcal{A} \in \mathscr{A}} \overline{R}_T(\mathcal{A}, \mathcal{S}).$$

**Remark 2.** *Clearly, a gap exists between the lower bound derived in Theorem 2 and the upper bound appearing in Theorem 1. It is unclear which bound, if any, provides the proper minimax rate. However, note that if the lower bound were tight, it would imply that the proportion of vertices that the player misses by picking suboptimal source sets is constant, meaning the number of additional vertices the optimal source vertex influences is linear in the size of the graph. This differs substantially from the pseudo-regret of order $\sqrt{n}$ known to be minimax optimal for the standard multi-armed bandit problem (and arises, for instance, in the case of directed graphs, as discussed in the next section).*

*3) Upper bounds for multiple sources:* We now turn to the case $k > 1$, where the player chooses multiple source vertices at each time step. As discussed in Section II, we are interested in bounding the scaled pseudo-regret $\overline{R}^{\alpha}_T(\mathcal{A}, \mathcal{S})$ with $\alpha = 1 - \frac{1}{e}$, since it is difficult to maximize the influence even in an offline setting, and the greedy algorithm is only guaranteed to provide a $\left(1 - \frac{1}{e}\right)$-approximation of the truth.

Our proposed player strategy is based on an online greedy adaptation of the strategy used in the single-source setting, and the full details are given in the appendix. We then have the following result concerning the scaled pseudo-regret:

**Theorem 3** (Symmetric loss, multiple sources)**.** *Suppose $k > 1$ and the player uses the strategy $\mathcal{S}^{sym,k}_{OSMD}$ corresponding to the Online Greedy Algorithm with single-source strategy $\mathcal{S}^{sym}_{OSMD}$. Then the scaled pseudo-regret satisfies the bound*

$$\sup_{\mathcal{A} \in \mathscr{A}} \overline{R}^{(1-1/e)}_T(\mathcal{A}, \mathcal{S}^{sym,k}_{OSMD}) \leq 2^{\frac{1}{4}} k \sqrt{Tn}.$$

Comparing Theorem 3 to Theorem 1, we see an additional factor of $k$ in the pseudo-regret upper bound. Similar results may be derived when alternative single-source strategies are used as subroutines in the Online Greedy Algorithm.

### *B. Directed graphs*

We now derive upper and lower bounds for the pseudo-regret in the case of directed graphs, when $k = 1$.

*1) Upper bounds:* The symmetric loss does not have a clear analog in the case of directed graphs. However, we may still use the node loss estimate for multi-armed bandit problems, given by equation (4). This leads to the following upper bound:

**Theorem 4.** *Suppose the player uses the strategy $\mathcal{S}^{node}_{OSMD}$ corresponding to OSMD with the node loss $\widehat{\ell}^{node}$ and appropriate parameters. Then the pseudo-regret satisfies the bound*

$$\sup_{\mathcal{A} \in \mathscr{A}} \overline{R}_T(\mathcal{A}, \mathcal{S}^{node}_{OSMD}) \leq 2^{\frac{3}{2}} \sqrt{Tn}.$$

**Remark 3.** *In the case $k > 1$, we may again use the Online Greedy Algorithm to obtain a player strategy composed of parallel runs of a single-source strategy. If the player uses*

*the single-source strategy $\mathcal{S}^{node}_{OSMD}$, we may obtain the scaled pseudo-regret bound*

$$\sup_{\mathcal{A}\in\mathscr{A}} \overline{R}_T^{(1-1/e)}(\mathcal{A}, \mathcal{S}^{node,k}_{OSMD}) \leq 2^{\frac{3}{2}} k\sqrt{Tn}.$$

*2) Lower bounds:* Finally, we provide a lower bound for the directed complete graph on $n$ vertices. (This refers to the case where all edges are present and bidirectional.) We have the following result:

**Theorem 5.** *Suppose $G$ is the directed complete graph on $n$ vertices. Then the pseudo-regret satisfies the lower bound*

$$\frac{1}{48\sqrt{6}}\sqrt{Tn} \leq \inf_{\mathcal{S}\in\mathscr{P}} \sup_{A\in\mathscr{A}} \overline{R}_T(\mathcal{A}, \mathcal{S}).$$

Notably, the lower bound in Theorem 5 matches the upper bound in Theorem 4, up to constant factors. Thus, the minimax pseudo-regret for the influence maximization problem is $\Theta(\sqrt{Tn})$ in the case of directed graphs. In the case of undirected graphs, however (cf. Theorem 2), we only obtained a pseudo-regret lower bound of $\Omega(\sqrt{T})$. This is due to the fact that in undirected graphs, one may learn about the loss of other nodes at time $t$ besides the loss at $\mathcal{S}_t$.

Finally, we remark that a different choice of $G$ might affect the lower bound, since influence maximization is easier for some graph topologies than others. However, Theorem 5 shows that the case of the complete graph is always guaranteed to incur a pseudo-regret that matches the general upper bound in Theorem 4, implying that this is the minimax optimal rate for any class of graphs containing the complete graph.

## IV. Discussion

We have proposed and analyzed player strategies that control the pseudo-regret uniformly across all possible oblivious adversarial strategies. For the problem of single-source influence maximization in complete networks, we have also derived minimax lower bounds that establish the fundamental hardness of the online influence maximization problem. In particular, our lower and upper bounds match up to constant factors in the case of directed complete graphs, implying that our proposed player strategy is in some sense optimal.

Our work inspires a number of interesting questions for future study. An important open question concerns closing the gap between upper and lower bounds on the minimax pseudo-regret in the case of undirected graphs, to determine whether the feedback available in the influence maximization setting actually makes the online game easier than a standard bandit setting. Furthermore, our lower bounds only hold in the case of complete graphs and single-source influence maximization, and it would be worthwhile to obtain lower bounds that hold for other network topologies and larger seed sets.

## References


[1] L. A. Adamic and E. Adar. Friends and neighbors on the Web. *Social Networks*, 25(3):211 – 230, 2003.

[2] S. Aral and D. Walker. Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*, 57(9):1623–1639, 2011.

[3] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

[4] C. Borgs, M. Brautbar, J. Chayes, and B. Lucier. Maximizing social influence in nearly optimal time. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 946–957. SIAM, 2014.

[5] S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandits. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.

[6] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, New York, NY, 2006.

[7] W. Chen, L. V. Lakshmanan, and C. Castillo. Information and influence propagation in social networks. *Synthesis Lectures on Data Management*, 5(4):1–177, 2013.

[8] W. Chen, T. Lin, Z. Tan, M. Zhao, and X. Zhou. Robust influence maximization. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[9] W. Chen, Y. Wang, Y. Yuan, and Q. Wang. Combinatorial multi-armed bandit and its extension to probabilistically triggered arms. *Journal of Machine Learning Research*, 17(50):1–33, 2016.

[10] P. Domingos and M. Richardson. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 57–66. ACM, 2001.

[11] X. He and D. Kempe. Robust influence maximization. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[12] X. He, K. Xu, D. Kempe, and Y. Liu. Learning influence functions from incomplete observations. In *Advances in Neural Information Processing Systems*, pages 2073–2081, 2016.

[13] D. Katz and R. L. Kahn. *The social psychology of organizations*, volume 2. Wiley New York, 1978.

[14] D. Kempe, J. Kleinberg, and E. 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*, KDD '03, pages 137–146, New York, NY, USA, 2003. ACM.

[15] D. Kempe, J. Kleinberg, and É. Tardos. Influential nodes in a diffusion model for social networks. In *Automata, languages and programming*, pages 1127–1138. Springer, 2005.

[16] S. Lei, S. Maniu, L. Mo, R. Cheng, and P. Senellart. Online influence maximization. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 645–654, New York, NY, USA, 2015. ACM.

[17] J. Leskovec, L. A. Adamic, and B. A. Huberman. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5, 2007.

[18] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 58(7):1019–1031, 2007.

[19] H. Narasimhan, D. C. Parkes, and Y. Singer. Learnability of influence in networks. In *Advances in Neural Information Processing Systems*, pages 3186–3194, 2015.

[20] M. E. J. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

[21] J. Olkhovskaya, G. Neu, and G. Lugosi. Online influence maximization with local observations. *arXiv preprint arXiv:1805.11022*, 2018.

[22] S. Seiler, S. Yao, and W. Wang. Does online word of mouth increase demand? (And how?) Evidence from a natural experiment. *Marketing Science*, 2017.

[23] O. Sporns. The human connectome: A complex network. *Annals of the New York Academy of Sciences*, 1224(1):109–125, 2011.

[24] M. Streeter and D. Golovin. An online algorithm for maximizing submodular functions. *Technical report*, pages 1–35, 2007.

[25] S. Vaswani, L. Lakshmanan, and M. Schmidt. Influence maximization with bandits. *arXiv preprint arXiv:1503.00024*, pages 1–12, 2015.

[26] Z. Wen, B. Kveton, and M. Valko. Online influence maximization under independent cascade model with semi-bandit feedback. *Advances in Neural Information Processing Systems*, 2017.