An Approximately Optimal Bot for Non-Submodular Social Reconnaissance

J. David Smith
University of Florida
Gainesville, Florida
emallson@ufl.edu

Alan Kuhnle
University of Florida
Gainesville, Florida
kuhnle@ufl.edu

My T. Thai
University of Florida
Gainesville, Florida
mythai@cise.ufl.edu

ABSTRACT

The explosive growth of Online Social Networks in recent years has led to many individuals relying on them to keep up with friends & family. This, in turn, makes them prime targets for malicious actors seeking to collect sensitive, personal data. Prior work has studied the ability of *socialbots*, i.e. bots which pretend to be humans on OSNs, to collect personal data by befriending real users. However, this prior work has been hampered by the assumption that the likelihood of users accepting friend requests from a bot is non-increasing – a useful constraint for theoretical purposes but one contradicted by observational data. We address this limitation with a novel curvature based technique, showing that an adaptive greedy bot is approximately optimal within a factor of $1 - 1/e^{1/\delta} \approx 0.165$. This theoretical contribution is supported by simulating the infiltration of the bot on OSN topologies. Counter-intuitively, we observe that when the bot is incentivized to befriend friends-of-friends of target users it out-performs a bot that focuses on befriending targets.

CCS CONCEPTS

• **Networks** → **Online social networks**; • **Theory of computation** → **Discrete optimization**;

KEYWORDS

social networks; privacy; discrete optimization; adaptive algorithms

ACM Reference Format:
J. David Smith, Alan Kuhnle, and My T. Thai. 2018. An Approximately Optimal Bot for Non-Submodular Social Reconnaissance. In *HT '18: 29th ACM Conference on Hypertext and Social Media, July 9–12, 2018, Baltimore, MD, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3209542.3209553

ACKNOWLEDGMENTS

This work was supported by NSF CCF-1422116, NSF CNS-1443905, and NSF EFRI 1441231.

1 INTRODUCTION

Online Social Networks (OSNs) have seen explosive growth in recent years, rapidly becoming the largest repositories of personal information on the Internet. This leads to the question: how difficult is it for an attacker to steal users' personal information from popular OSNs? This may be done by, for instance, befriending users with *socialbots* that can exfiltrate normally private data via their friendship relations. Boshmaf *et al.* showed that Facebook is vulnerable to such attacks [5], while Freitas *et al.* more recently showed that Twitter is vulnerable to similar attacks [10].[1] This information can then be sold on the black market, used to enhance spearphishing attacks, or to crack password recovery systems – and thus indirectly used to reduce the security level of the rest of our infrastructure.

Li *et al.* recently used the observational data of Boshmaf *et al.* [5] to estimate the rate at which a socialbot could extract private data from an OSN [19]. In doing so, they found that a socialbot using an adaptive greedy approach would obtain at least $(1 - 1/e)$ times as much benefit as the optimal and that no algorithm can do better than $(1 - 1/e)$ unless $P = NP$. However, this approximation guarantee demands a strong assumption: the expected benefit of befriending users must be *submodular*. In the deterministic case, this is often formalized as

$$\forall S \subseteq T : f(S \cup \{e\}) - f(S) \geq f(T \cup \{e\}) - f(T)$$

Semantically, this means that the benefit has diminishing returns as more users are befriended. However, this condition does not hold due to the impact of acceptance probability on the expectation: the acceptance probability increases as more users are befriended [5], leading to non-submodular behavior. Although Li *et al.* study this setting, their guarantees do not hold without submodularity [18].

While the performance bound of submodular problems has been studied since 1978 [20], such study of monotone non-submodularity has only begun very recently [2, 27]. However, these recent results are not readily applied to the reconnaissance attack application owing to the fact that it is necessarily *adaptive*. Due to the massive size of modern OSNs, obtaining accurate knowledge of the entire network topology at once is infeasible. Therefore, the bot must explore the network as it crawls, revealing parts of the topology by befriending users and observing with whom they are friends, and then using this information to inform future steps. This property of making decisions based on the outcome of previous ones is the defining trait of adaptive algorithms. In particular, the adaptive stochastic nature of the problem makes current results inapplicable and necessitates new solutions.

In this work, we address the above limitation by introducing novel mathematical techniques to theoretically analyze the performance bound of adaptive greedy algorithms for non-submodularity. At the heart of our techniques is a key proof bounding on the effect

[1] While stealing personal information from public Twitter profiles makes little sense, the embedding of socialbots on a network has other nefarious applications such as spreading misinformation e.g. in a way that evades containment [21].

HT '18, July 9–12, 2018, Baltimore, MD, USA

ACM ISBN 978-1-4503-5427-1/18/07…$15.00
https://doi.org/10.1145/3209542.3209553

of adding a node to a solution later rather than earlier. We accomplish this through a new measure of the rate of change of a function, the *primal curvature*. A bound on this measure is shown to be both necessary & sufficient to obtain an approximation guarantee in the general case. In our specific case, we exploit the structure of the problem to obtain an approximation ratio of $1 - 1/e^{1/\delta}$ with δ a constant depending on user behavior. The generalized techniques provided in this paper advance the research front of several applications, where both adaptivity and non-submodularity are required, such as adaptive viral marketing in OSNs [12, 16, 25]. As the first work established the rigorous proofs for adaptivity and non-submodularity, this paper opens the way for the development of adaptive approximation algorithms on domains where external factors – such as human behavior – prevent common assumptions like submodularity from holding.

Our contributions can be summarized as follows:

- We provide the first theoretical study of reconnaissance attacks under a realistic model of friend request acceptance. We obtain a bound of $1 - 1/e^{1/\delta}$ with $\delta = O(1)$.
- We provide the first technique to theoretically bound the approximation quality of non-submodular adaptive approximation algorithms, which generalizes the $1 - 1/e$ ratio for adaptive submodular maximization via the greedy algorithm.
- We delve deeper into the behavior of a socialbot under this realistic model, finding that the added term rewarding a bot for improving friend request probability adds needless complexity and that, paradoxically, encouraging the bot to become friends-of-friends with targeted users actually results in a greater fraction of targets being befriended.

Related Work. *Reconnaissance Attacks.* Reconnaissance attacks on OSNs have been shown to be an effective method of extracting private information from OSN users [5, 10]. The method of attack is conceptually simple: a single "socialbot" is created on the OSN with a realistic user profile, and automatically befriends users with the goal of extracting as much private data as possible (e.g. for sale on the black market or use in breaking "secret question" password recovery schemes for further attacks). Ryan & Mauch showed that such fake profiles can be effective in obtaining access to the personal feeds of high-ranking government and corporate officers [23]. In a similar vein, Varol *et. al.* studied the presence of bots on Twitter and found that human users befriended more human-like users than bots [26], indicating that for reconnaissance to be successful it must be undertaken by human-like bots–a.k.a. socialbots.

Note that this attack is distinct from the well-studied Sybil attack, and due to the absence of the bots creating large sub-graphs, it is unlikely to be detected by Sybil detection schemes [3]. Traditionally, literature on reconnaissance attacks has been primarily experimental in nature and lacked rigorous theoretical guarantees [5, 10, 22]. More recently, a greedy socialbot was shown to collect at least a $1 - 1/e$ fraction of the information collected by an optimal bot using adaptive techniques (see Golovin & Krause [11] for a full treatment on adaptivity) [18]. This was extended to a ratio of $1 - 1/e^{-(1-1/e)}$ when the bot is allowed to make multiple simultaneous friend requests [19].

However, these ratios *do not hold* without submodular benefit and friend request acceptance models. While the benefit model is

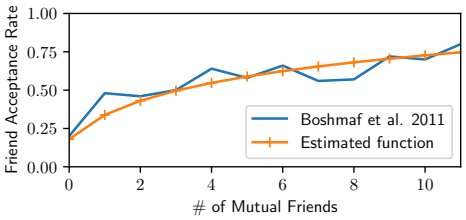


Figure 1: The acceptance probability function proposed by Li *et al.* [18] using the data of Boshmaf *et al.* [5].

under the control of the attacker and can be easily constructed to be submodular, observational data indicates that the friend request acceptance model is strongly non-submodular [5, 18].

Non-Submodular Optimization. Submodular optimization has been the subject of intense study. Perhaps the most-used work to come out of this is the tight $1 - 1/e$ ratio of Nemhauser *et al.* [7, 20], which is fundamental to the guarantees of many applied works. Quite recently, a number of works have also begun to study *non-submodular* optimization, which cannot exploit the useful "diminishing returns" of submodularity to obtain approximation guarantees. However, to obtain these guarantees constraints must be imposed on the problem (a proof of the necessity of one such constraint is contained in Sec. 3.1). Even in the case of functions that are approximately submodular and violate submodularity only due to noise, this problem requires strong constraints to obtain meaningful guarantees [13, 14]. Das & Kempe proposed the *submodularity ratio* γ as a means of quantifying the magnitude by which submodularity is violated [6]. Using this, they obtain a ratio of $1 - 1/e^{\gamma}$ for their specific problem. Recently, Bian *et al.* extended this by incorporating the *generalized curvature* α to obtain a ratio of $\frac{1}{\alpha}(1 - e^{-\alpha\gamma})$ in general for greedy maximization subject to cardinality constraints [2]. Wang *et al.* took an alternative approach, obtaining a ratio for the greedy algorithm via the *elemental curvature* [27].

However, none of these apply to adaptive stochastic optimization, which is necessary the modeling of non-deterministic systems – such as the reconnaissance attack. The problem of non-submodular adaptive maximization has not yet seen study.

Organization. We begin in Section 2 by giving a semantic description of the socialbot reconnaissance attack and presenting an algorithmic description of the socialbot in terms of our formal model. Next, we describe our measure of curvature and derive the $1 - 1/e^{1/\delta}$ ratio in Section 3. While we focus on our particular application, this ratio extends to any problem with a finite curvature bound. This is followed by our experimental evaluation of the socialbot in Section 4. Finally, we conclude with a discussion of the implications and potential future work in Section 5.

2 PROBLEM FORMULATION & ALGORITHM

Before formally defining our problem, we first describe the semantics of the socialbot attack that informs it. Consider a social network such as Facebook. Users have a significant amount of personal information, much of which is locked behind privacy controls. The default (and most common) setting for content on Facebook is "Only Friends," which allows only direct friends to see your posts and friends. While link prediction can give an estimate as to the probability of friendships existing (e.g. [1, 8, 9]), there remains significant

incentive for attackers to befriend users for the information in their profiles. We term users under such attacks *targets*. However, users may not accept a friend request from a bot. Thus, the bot must consider the probability of acceptance – and ideally take advantage of human behavior to maximize it. To improve the acceptance probability, the bot may first seek to befriend friends of the targets. The reconnaissance process then unfolds round by round.

Boshmaf *et al.* observed that acceptance probability on Facebook seems to be dominated by the number of mutual friends – likely a result of the *Triadic Closure Principle* [5], which states that if a and b are friends with c, then a and b are also likely to be friends. In this case, the bot can boost the likelihood of requests being accepted by first befriending more vulnerable mutual friends. This raises the question of how to make the critical first few friends. It has been observed that users with abnormally high number of friends (so-called "high-degree users") have a larger chance of accepting friend requests without critical examination of the requester [4]. This allows the bot to bootstrap by sending initial requests to high-degree users, then crawling along the network–taking advantage of triadic closure to keep acceptance probabilities high.

Li *et al.* fit a model of acceptance probability to the observational data of Boshmaf *et al.* [18]. The exact function they give is

$$\alpha(u \mid \psi) = \rho_1 \log(\mathbb{E}[|N(u) \cap N_{in}(s)|] + 1) + \rho_0 \tag{1}$$

with $\rho_1 = 0.22805837$ and $\rho_0 = 0.18014571$ and s representing the bot (the expectation is taken with respect to ψ, which will be defined in the next subsection). This function is shown in Fig. 1. $N(u)$ and $N_{in}(s)$ are the sets of outgoing and incoming neighbors of u and s, respectively. Taking expectations under this model results in a non-submodular objective since the probability of a user u accepting a request may *increase* after another user v is befriended without a corresponding drop in the benefit of befriending u.

2.1 Formal Definitions

In sum, this leads us to an *adaptive* model of the problem [11]. Our model incorporates two pieces of uncertainty: the possibility that edges may not exist, and that friend requests may be rejected. The former are represented as a set of random variables (RVs) $X_e \in \{0,1\} \forall e \in E$ where the OSN is represented as a digraph $G = (V, E)$, where V is a set of users and E is the set of friendships; $X_e = 0$ iff the edge e does not exist. Note that edges are added on successful friend requests by the bot. We model friend requests with two sets of RVs $Y_v \in [0,1)$ and $Z_v \in \{0,1\}$. These Y_v represent thresholds for the acceptance probability, with v accepting a request if $\alpha(v \mid \psi) \geq Y_v$, where $\alpha(v \mid \psi)$ is the acceptance probability of v under partial realization ψ (defined below). The Z_v's are induced variables representing the status of the bot's friend request to v, with $Z_v = 1$ iff $\alpha(v \mid \psi) \geq Y_v$ *at the point where s made the request.*[2]

In the adaptive framework, there exists a set of possible *(total) realizations* Φ, which encode all potential states of the random variables X_e, Y_v, and Z_v described above. An adaptive policy π makes decisions based on a partial realization ψ, which encodes the values of the random variables in a system that are currently known. The domain of a partial realization, denoted $\mathrm{dom}(\psi)$, is the set of random variables revealed in ψ. We write $F(\psi) = \{u \mid Z_u \in \mathrm{dom}(\psi)\}$. A partial realization is said to be consistent with a total realization, denoted $\psi \sim \phi$, if they are equivalent everywhere in $\mathrm{dom}(\psi)$. We will denote the adaptive greedy policy selecting k elements π_k and the optimal policy selecting k elements π^*_k. When the choice of k is clear, we drop the subscript for notational simplicity. We slightly abuse notation and denote the final partial realization produced by policy π were it to run on a realization ϕ as $\pi(\phi)$.

Under this model, the bot is given as input a graph G with known nodes and unknown edges, along with an edge probability function $p(u, v)$ and a benefit model $\mathcal{B} = (B_f, B_{fof}, B_e)$. The bot ultimately outputs the sequence of friend requests made and the final partial realization uncovered by this sequence. At each step, an adaptive greedy bot will select the element maximizing the expected marginal gain $\Delta(v \mid \psi)$ and sends it a friend request. If it is successful, we observe $Y_v \leq \alpha(v \mid \psi) \implies Z_v = 1$, and we observe each X_e where $e \in E$ is an outgoing edge of v. On the other hand, if the request fails we observe $Z_v = 0$ and do not observe any variables X_e. Given this formulation, we write the objective $f(S, \phi)$ as:

$$f(S,\phi) = \sum_{u \in S} Z_u \left[B_f(u) + \sum_{v \in N(u)} B_e(u,v) \right] + \sum_{v \in N(S) \setminus S} \left(1 - \prod_{\substack{v \in N(u) \\ u \in S}} (1 - Z_u X_{u,v}) \right) B_{fof}(v) \tag{2}$$

where $N(u)$ is the set of nodes that may be adjacent u, $N(S)$ is the union of such over all $u \in S$, and $B_f(\cdot)$, $B_{fof}(\cdot)$, and $B_e(\cdot)$ represent the benefit assigned to a given friend, friend of friend, or edge revelation. In addition, we require $B_{fof}(v) \leq B_f(v)$ for all users v. We refer to a user u as a *target* if $B_f(u) > 0$. The expected benefit of f w.r.t. all possible realizations is $f_{\mathrm{avg}}(\pi) = \mathbb{E}[f(F(\pi(\Phi)), \Phi)]$ where Φ is a random total realization. This gives us the final piece to formally define the socialbot attack.

Problem 1 (Maximal InformatioN Extraction (MINE)). *Given a social graph $G = (V, E)$ with edge probabilities $p(u, v)$, an acceptance model $\alpha(v \mid \psi)$ that is adaptive monotone non-decreasing w.r.t. ψ, and a benefit model $f(S, \phi)$ that is adaptive monotone non-decreasing submodular w.r.t. S, find the k-element policy π that maximizes the expected benefit obtained.*

It has been shown that this problem[3] is inapproximable within $1 - 1/e - \epsilon$ for any $\epsilon > 0$ unless $P = NP$ even in the case where α is also submodular [19]. Were α submodular, this objective would be adaptive submodular [18] in addition to adaptive monotone, and the adaptive greedy policy would then have a tight ratio of $1 - 1/e$ [11]. It will be shown in the next section that – under some mild conditions on α – this ratio is preserved nearly exactly when α is allowed to be non-submodular.

2.2 Needed Properties for the Greedy Solution

As in prior work, we take a greedy approach to optimizing f. In this approach, the bot at each step chooses the user to befriend

[2] We found a definition exclusively in terms of Y_v or Z_v problematic due to the need to denote the answers to two distinct questions: (Y_v) *a request was just made–did it succeed?* and (Z_v) *a request was made in a prior step–was it successful?*. Using both together greatly simplifies presentation.

[3] This problem has seen prior study under the moniker "AReST" [18, 19].

Algorithm 1 Greedy MINE

Input: Problem instance (G, p, α, B, k)
Output: An ordered set of nodes $F \subset V$ to befriend, realization ψ.

1: $F \leftarrow \emptyset, \psi \leftarrow \emptyset$
2: **for** $i = 1 \dots k$ **do**
3: $u^* \leftarrow \arg\max_{u \in V \setminus F} \Delta(u \mid \psi)$
4: $F \leftarrow F \cup \{u^*\}$
5: Send a friend request to u^*, observing Y_{u^*}
6: **if** $\alpha(u^* \mid \psi) \geq Y_{u^*}$ **then**
7: **for** $v \in N(u^*)$ **do**
8: Observe $X_{u^*,v}$, updating ψ
9: Set $Z_{u^*} = 1$
10: **else**
11: Set $Z_{u^*} = 0$
12: Update ψ with the observed value of Y_{u^*}
13: **return** F, ψ

with highest expected marginal gain (line 3 of Alg. 1). It then sends this request and observes the result. If the user accepts the request, additional observations are made (lines 6-11). After having sent k requests, the bot returns the set of requests it made F and the partial realization resulting from those requests ψ.

Despite a similar approach to prior work, our objective function differs in the omission of a term rewarding the bot directly for improving marginal gain. We therefore prove necessary properties for the greedy algorithm to be applied. First, we prove that our objective maintains the property of *adaptive monotonicity*. Then, we derive a closed form for the expected marginal gain. The section closes by using this closed form to prove that the function is in general not *adaptive submodular*.

Adaptive Monotonicity of f. We adopt an alternative definition of adaptive monotonicity that is equivalent to the standard one. First, we require the definition of policy concatenation.

Definition 1 (Concatenation of policies [11]). *Given two policies π, π', define $\pi@\pi'$ as the policy obtained by running π to completion, and then running π' as if from a fresh start, ignoring the information gathered during the running of π.*

Definition 2 (Adaptive Monotonicity [11]). *A function $g : 2^E \times O^E \rightarrow \mathbb{R}_{\geq 0}$ is adaptive montone if for all polices π, π', it holds that $g_{avg}(\pi) \leq g_{avg}(\pi'@\pi)$.*

Lemma 2.1. *f is adaptive monotone.*

Proof. Let π, π' be policies and for a realization ϕ write $\psi = \pi(\phi), \psi' = \pi@\pi'(\phi)$. Notice that for any realization $\phi, F(\psi) \subseteq F(\psi')$. Hence it is enough to show for any $S \subseteq S'$, $f(S, \phi) \leq f(S', \phi)$. It is clear that any $B_f(u), B_e(v, w)$ in $f(S, \phi)$ is also present in $f(S', \phi)$. Furthermore, any $B_{fof}(v)$ term in $f(S, \phi)$ is absent in $f(S', \phi)$ only if it is replaced by $B_f(v)$. Since we have $B_{fof}(v) \leq B_f(v)$ for all v, the result follows. □

Closed Form of $\Delta(u \mid \psi)$. The expected marginal gain $\Delta(u \mid \psi)$ is defined by Golovin & Krause [11] as

$$\Delta(u \mid \psi) = \mathbb{E}[f(F(\psi) \cup \{u\}, \Phi) - f(F(\psi), \Phi) \mid \Phi \sim \psi] \quad (3)$$

Based on the definition given in Equation (2) and the definitions of the variables, this has the closed form

$$\begin{aligned}\Delta(u \mid \psi) &= \alpha(u \mid \psi) \Bigg[B_f(u) + \sum_{v \in N(u)} B_e(u, v) \\ &\quad + \sum_{v \in N(u) \setminus F(\psi)} (1 - I_{fof}(v)) p(u, v) B_{fof}(v) \Bigg] \\ &= \alpha(u \mid \psi) \mathcal{B}(\psi, u) \end{aligned} \quad (4)$$

where $I_{fof}(v)$ is the indicator function returning 1 if v is already a friend-of-friend of s and 0 otherwise.

Adaptive Submodularity For the sake of completeness, we now present the definition of adaptive submodularity:

Definition 3 (Adaptive Submodularity [11]). *A function $g : 2^E \times O^E \rightarrow \mathbb{R}_{\geq 0}$ is adaptive submodular if for a pair of partial realizations $\psi \subseteq \psi'$:*

$$\forall e \in E \setminus dom(\psi') : \Delta(e \mid \psi) \geq \Delta(e \mid \psi')$$

Lemma 2.2. *f is not adaptive submodular in general.*

Proof. A trivial counter-example is $\forall u \in V : B_f(u) = 1, B_{fof}(u) = 0, \forall (u, v) \in E : B_e(u, v) = 0$. Then, we easily have non-submodularity because α is increasing w.r.t. ψ. This example will be shown non-submodular by contradiction. Suppose we have $\psi' \supset \psi$ s.t. $\alpha(u \mid \psi) \neq \alpha(u \mid \psi')$. Begin with the definition of adaptive submodularity:

$$\Delta(u \mid \psi) \geq \Delta(u \mid \psi')$$

By the closed form of Δ derived previously and the selection of B_f, B_e, B_{fof} this simplifies to

$$\alpha(u \mid \psi) B_f(u) \geq \alpha(u \mid \psi') B_f(u)$$

However, we know $\alpha(u \mid \psi) < \alpha(u \mid \psi')$ for this pair ψ, ψ'. We thus arrive at a contradiction. □

3 APPROXIMATION RATIO

Greedy methods are often chosen for their good real-world performance in addition to the strong theoretical guarantee that any solution produced is at least $1 - 1/e$ times as good as the optimal if the objective is submodular [11, 20]. However, the behavior of users observed by Boshmaf et al. [5] indicates that the objective is must be non-submodular since it incorporates the increasing acceptance function.[4] Therefore, we introduce a new technique for deriving the approximation guarantee for the greedy adaptive policy.

Wang et al. [27] were among the first to provide an approximation guarantee for general non-submodular set functions in terms of the *elemental* curvature: the maximum ratio between the marginal gain of an element i at any pair of sets S and $S \cup \{j\}$. We extend their idea to the adaptive realm with the *primal curvature*: a localized definition of curvature.

Definition 4 (Adaptive Primal Curvature). *The primal curvature of an adaptive monotone non-decreasing function f is*

$$\nabla_f(i, j \mid \psi) = \mathbb{E}\left[\frac{\Delta(i \mid \psi \cup s)}{\Delta(i \mid \psi)} \,\middle|\, s \in S(j) \right]$$

[4] We remark that any reasonable objective must incorporate the likelihood of acceptance, as doing so is a fundamental part of computing the expected value.

where $S(j)$ is the set of possible states of j and Δ is the conditional expected marginal gain [11].

Intuitively, the adaptive primal curvature measures the immediate change in the (expected) marginal value of i after j is added to the solution. In the non-adaptive case (i.e. $|S(j)| = 1$), the elemental curvature is the maximum primal curvature. However, this localization allows us to proceed in a new direction with the proof. We use the *total primal curvature*, defined below, to measure the total change between two partial realizations.

Definition 5 (Adaptive Total Primal Curvature). *Let $\psi \subset \psi'$ and $\psi \to \psi'$ represent the set of possible state sequences leading from ψ to ψ'. Then the adaptive total primal curvature is*

$$\Gamma(i \mid \psi', \psi) = \mathbb{E}\left[\prod_{s_j \in Q} \nabla'(i, s_j \mid \psi \cup \{s_1, \ldots, s_{j-1}\}) \,\middle|\, Q \in \psi \to \psi'\right]$$

The following lemma clarifies the relation between total primal curvature and the marginal gain, and the corresponding result in Corollary 3.4 directly enables our proof of the adaptive approximation guarantee.

Lemma 3.1.

$$\Gamma(i \mid \psi', \psi) = \frac{\Delta(i \mid \psi')}{\Delta(i \mid \psi)}$$

Proof. Fix a sequence $Q \in \psi \to \psi'$ of length r. Then, expanding the product we obtain

$$\frac{\Delta(i \mid \psi \cup \{s_1\})}{\Delta(i \mid \psi)} \cdot \frac{\Delta(i \mid \psi \cup \{s_1, s_2\})}{\Delta(i \mid \psi \cup \{s_1\})} \cdots \frac{\Delta(i \mid \psi')}{\Delta(i \mid \psi' \setminus \{s_{r-1}\})}$$

If we take the expectation of this w.r.t. the possible sequences Q, we obtain the same ratio regardless of Q, and therefore the claim holds trivially. □

This identity allows us to place a constant bound on the total primal curvature for the MINE problem. As we will show later, this ultimately leads to a constant approximation ratio.

Lemma 3.2. $\max_{i,\psi,\psi'} \Gamma(i \mid \psi', \psi)$ *is upper bounded by*

$$\delta = \max_{u,\psi,\psi'} \frac{\alpha(u \mid \psi)}{\alpha(u \mid \psi')}$$

Proof. For any i, ψ', ψ, we have

$$\Gamma(i \mid \psi', \psi) = \frac{\alpha(i \mid \psi')\mathcal{B}(\psi', i)}{\alpha(i \mid \psi)\mathcal{B}(\psi, i)} \leq \delta \frac{\mathcal{B}(\psi', i)}{\mathcal{B}(\psi, i)}$$

by the derivation of the closed form in Section 2.2. $\mathcal{B}(\psi', i) \leq \mathcal{B}(\psi, i)$ by definition, and therefore $\Gamma(i \mid \psi', \psi) \leq \delta$. □

Corollary 3.3. *For the ETC acceptance function, $\delta = O(1)$.*

Proof. Recall that the ETC acceptance function is defined as:

$$\alpha(u) = \rho_1 \log(\mathbb{E}[|N(u) \cap N(s)|] + 1) + \rho_0$$

Thus, for any u, $\min_\psi \alpha(u \mid \psi)$ is achieved in all partial realizations that guarantee $|N(u) \cap N(s)| = 0$ and $\max_\psi \alpha(u \mid \psi) \leq 1$. Thus, $\rho_0 \leq \alpha(u \mid \psi) \leq 1, \forall \psi$. So we have:

$$\delta \leq \frac{\max_\psi \alpha(u \mid \psi)}{\min_\psi \alpha(u \mid \psi)} \leq \frac{1}{\rho_0}$$

As ρ_0 is a constant, $\delta = O(1)$ for the ETC acceptance function. □

This leaves the task of proving a ratio in terms of this bound. While the following proofs hold for more general statements of the adaptive TPC, we will prove them w.r.t. δ instead as this dramatically simplifies our notation.

Corollary 3.4. *Given a partial realization ψ resulting from application of the l-element greedy policy, $\psi \subset \psi'$, $i \notin dom(\psi)$, and g_{l+1} the next element that would be selected by the greedy policy at partial realization ψ, we have:*

$$\Delta(i \mid \psi') \leq \delta \Delta(g_{l+1} \mid \psi)$$

Proof. By Lemmas 3.1 and 3.2,

$$\Delta(i \mid \psi') = \Gamma(i \mid \psi', \psi)\Delta(i \mid \psi) \leq \delta \Delta(g_{l+1} \mid \psi)$$

and thus the statement holds. □

We exploit this corollary in the following lemma to explicitly relate the difference between an arbitrary policy and the l-element greedy policy to the expected marginal gain of adding an $l+1$'st element to the greedy solution. Note that this "arbitrary policy" will, in practice, be an *optimal* policy.

Lemma 3.5.

$$f_{avg}(\pi') - f_{avg}(\pi_l) \leq k\delta \Delta_{avg}(\pi_l, \pi_{l+1}) \tag{5}$$

where π_l is the greedy policy selecting l elements with $l < k$, π' selects exactly k elements, and $\Delta_{avg}(\pi_l, \pi_{l+1}) = f_{avg}(\pi_{l+1}) - f_{avg}(\pi_l)$.

Proof. Note that

$$f_{\text{avg}}(\pi') - f_{\text{avg}}(\pi_l) \leq f_{\text{avg}}(\pi_l @ \pi') - f_{\text{avg}}(\pi_l)$$

since $f_{\text{avg}}(\pi') \leq f_{\text{avg}}(\pi_l @ \pi')$ due to the adaptive monotonicity of f. From this inequality, it is clear that the difference in the expected values of π' and π_l is bounded by the marginal gain of running π' after π_l. This involves sending at most k additional requests. By Corollary 3.4, the marginal gain of each of these requests is bounded above by $\delta\Delta(g_{l+1} \mid \psi)$ for each possible ψ. Thus, we have:

$$\begin{aligned}
&f_{\text{avg}}(\pi') - f_{\text{avg}}(\pi_l) \leq \mathbb{E}[k\delta\Delta(g_{l+1} \mid \psi) \mid \psi] \\
&= k\delta\mathbb{E}[\Delta(g_{l+1} \mid \psi) \mid \psi] \\
&= k\delta\mathbb{E}[\mathbb{E}[f(\text{dom}(\psi) + g_{l+1}, \Phi) - f(\text{dom}(\psi), \Phi) \mid \Phi \sim \psi] \mid \psi] \\
&= k\delta\mathbb{E}[f(E(\pi_{l+1}, \Phi), \Phi) - f(E(\pi_l, \Phi), \Phi) \mid \Phi] \\
&= k\delta\Delta_{avg}(\pi_l, \pi_{l+1})
\end{aligned}$$

where the second equality uses the definition of $\Delta(\cdot)$. □

Finally, we have the main theorem providing the adaptive approximation guarantee:

Theorem 3.6.

$$\left[1 - \left(1 - \frac{1}{k\delta}\right)^k\right] f_{avg}(\pi_k^*) \leq f_{avg}(\pi_k) \tag{6}$$

Proof. By Lemma 3.5, we have

$$f_{\text{avg}}(\pi_k^*) \leq f_{\text{avg}}(\pi_l) + k\delta\Delta_{avg}(\pi_l, \pi_{l+1})$$

Multiply both sides by $(1 - (k\delta)^{-1})^{k-1-l}$ and sum from $l = 0$ to $k - 1$. We directly get that the left hand side reduces to

$$k\delta\left[1 - \left(\frac{k\delta - 1}{k\delta}\right)^k\right] f_{\text{avg}}(\pi_k^*)$$

The right-hand side reduces to

$$\sum_{l=0}^{k-1} \left[f_{\text{avg}}(\pi_l) + k\delta\Delta_{avg}(\pi_l, \pi_{l+1}) \right] \left(1 - \frac{1}{k\delta}\right)^{k-1-l}$$

To simplify the below equations, we will denote $\beta = 1 - \frac{1}{k\delta}$ and use the identity $f_{\text{avg}}(\pi_l) = \sum_{i=0}^{l-1} \Delta_{\text{avg}}(\pi_i, \pi_{i+1})$.

Consider the $j+1$'st decision made by the policy π. Inside the summation, decision π_{j+1} appears in the terms

$$\begin{aligned} &\beta^{k-1-j} k\delta \Delta_{\text{avg}}(\pi, \pi_{j+1}) + \beta^{k-1-(j+1)} \Delta_{\text{avg}}(\pi, \pi_{j+1}) \\ &+ \beta^{k-1-(j+2)} \Delta_{\text{avg}}(\pi, \pi_{j+1}) + \dots \\ &= \left[\beta^{k-1-j} k\delta + \sum_{l=j+1}^{k} \beta^{k-1-l} \right] \Delta_{\text{avg}}(\pi, \pi_{j+1}) \end{aligned}$$

Applying the closed form of the geometric series, we can simplify this coefficient to

$$\beta^{k-1-j} k\delta + \frac{\beta^{k-1-j} - 1}{\beta - 1} = \beta^{k-1-j} k\delta - \beta^{k-1-j} k\delta + k\delta$$

where the right-hand of this inequality comes from noting that $(1 - 1/(k\delta)) - 1 = -1/(k\delta)$. Therefore, we have the sum

$$\sum_{l=0}^{k-1} k\delta \Delta_{\text{avg}}(\pi_l, \pi_{l+1}) = k\delta f_{\text{avg}}(\pi)$$

Rearranging terms, we arrive at the statement of the theorem. □

Corollary 3.7. *Greedy maximization of an adaptive monotone function with total primal curvature bound δ satisfies*

$$\left(1 - 1/e^{1/\delta}\right) f_{avg}(\pi_k^*) \leq f_{avg}(\pi_k) \tag{7}$$

Proof. This follows immediately by noting

$$\lim_{k \to \infty} \left(1 - \frac{1}{k\delta}\right)^k = 1/e^{1/\delta}$$ □

Thus, we have a constant approximation ratio of $1 - 1/e^{1/\delta} \approx 0.165$ for the MINE problem under the ETC model.

3.1 The Necessity of Finite Curvature

In the statement of Theorem 3.6, we use our problem-specific bound δ on Γ. We now show that it is necessary for any problem to have a finite bound δ to obtain an approximation ratio unless $P = NP$. We accomplish this by showing that any $\rho(n)$-approximation algorithm for maximizing an arbitrary monotone non-submodular g must solve a class of NP-hard problems exactly, but that the constraint $\Gamma < \infty$ excludes such problems.

Theorem 3.8. *There is no polynomial time algorithm for $\rho(n)$-approximate maximization of an arbitrary monotone non-submodular function g with $\rho(n) > 0$ unless $P = NP$.*

Proof. To show this, we first construct an objective function with infinite curvature that cannot be exactly solved in polynomial time by reduction to SAT [15]. We then show that any approximation algorithm with ratio $\rho(n) > 0$ is necessarily exact. While our proof uses discrete terminology, we note that discrete problems are a subset of adaptive problems where there is only a single realization and therefore our proof extends to the adaptive case.

Suppose we are given a CNF formula $\mathbf{F}$ with C clauses each having k_l literals and containing L literals total. We show how to construct a monotone supermodular function g which returns 1 when the formula is satisfied and 0 otherwise.

First, we construct the domain of g. For every literal x_i in $\mathbf{F}$, insert two elements T_i and F_i into set N, corresponding to assigning the literal x_i 1 and 0, respectively. Then, a satisfying assignment for $\mathbf{F}$ corresponds to a set S containing either T_i or F_i for every literal x_i. For the moment, we assume that such an assignment exists.

The verifying function g is then composed of three semantic parts, each of which returns 1 when satisfied and 0 otherwise: (1) S contains a satisfying assignment, (2) S does not contain both T_i and F_i for any i, and (3) S assigns every literal a value. The latter two conditions are needed because it is possible for a formula such as $\mathbf{F} = (x_1 \vee \cdots) \wedge (x_1 \vee \cdots) \wedge \cdots$ to be given, which is satisfied by the assignment corresponding to $S = \{T_1\}$.

For each clause $c_l = x_i \vee x_j \vee \cdots \vee \bar{x_r} \vee \cdots$, define a function $C_l(S)$ verifying the clause is satisfied. For each literal x_i, define a function $A_i(S)$ verifying x_i is not simultaneously assigned 0 and 1. Finally, define a function $B(S)$ verifying that every literal is assigned a value. These can be constructed and evaluated exactly in polynomial time with the following closed forms:

$$\begin{aligned} C_l(S) &= \left\lceil \frac{1}{k_l} \left(|\{T_i\} \cap S| + |\{T_j\} \cap S| + \cdots + |\{F_r\} \cap S| \right) \right\rceil \\ A_i(S) &= 1 - \left\lfloor \frac{1}{2} |\{T_i, F_i\} \cap S| \right\rfloor \\ B(S) &= \left\lfloor \frac{1}{L} |S| \right\rfloor \end{aligned}$$

Then define g as

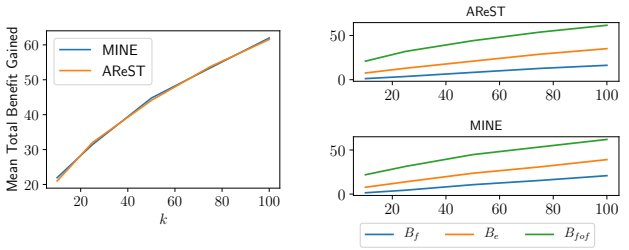
$$g(S) = B(S) \prod_{i=1}^{L} A_i(S) \prod_{l=1}^{C} C_l(S)$$

By construction, g is 1 for any S corresponding to a satisfying assignment and 0 otherwise. Further, note that $\forall |S| \leq k$, this function is monotone supermodular. To have this property everywhere, we extend it piecewise to $g'(S) = g(S) \forall |S| \leq k$ and $g'(S) = 1 \forall |S| > k$.

Now, suppose we have an $\mathbf{F}$ with exactly one satisfying assignment. Then there is exactly one $S^*, |S^*| = k$ s.t. $g'(S^*) = 1$, and $\forall S, |S| \leq k : g'(S) = 0$. Clearly, if a polynomial-time algorithm $\mathcal{A}$ can approximate the optimal solution S^* with $\rho(n) > 0$, then on this problem it must find the optimal solution. Otherwise, it would have $\rho(n) = 0$. Therefore, either $\mathcal{A}$ can solve SAT in polynomial time, implying $P = NP$, or $\mathcal{A}$ does not have a non-zero approximation ratio for the stated class of objective functions. □

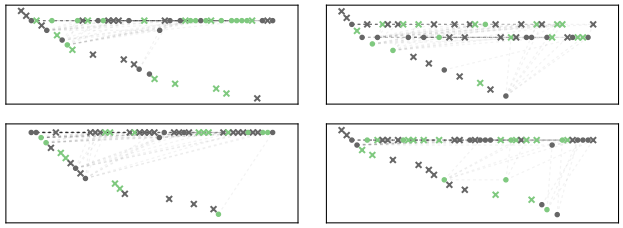
Observe that the above problem does not have a finite bound δ, since $\Gamma(i \mid S^* \setminus \{i\}, S^* \setminus \{i, j\}) = 1/0$.[5] Thus, the constraint that δ be finite is also sufficient to exclude such cases. More generally, this means that every technique for giving an approximation ratio for

[5] We abuse our notation here. $\Gamma(i \mid T, S)$ is the discrete analogue of the adaptive total primal curvature, and can be defined as the ratio of marginal gains.



(a) 100 Targets (b) Breakdown by Category

Figure 2: Benefit with and without a term rewarding boosting future acceptance chances on the Slashdot network.



(a) π_α (b) π

Figure 3: Sample traces under policies π_α and π with $k = 50$ and 100 targets on the Slashdot network. Dark nodes are targeted. $\times$ marks represent users that rejected the bot's friend request and circular marks the opposite. Time proceeds left to right. Each node is placed on the first row from the top containing a friend or is added to a new row if none exist. Lines correspond to edges in the OSN.

non-submodular maximization must also bound δ unless $P = NP$.[6]

4 EXPERIMENTAL EVALUATION

We now examine several key questions about the performance of the greedy socialbot. Extensive experiments have already established the general efficacy of the greedy policy under various acceptance models, including the one studied in this paper, and that the choice of acceptance model significantly impacts the potency of the attack [18, 19]. Thus, we instead focus on investigation of particular factors influencing the behavior of the bot. First, we establish equivalence with the bot models studied previously – a result necessitated by the omission of a term rewarding the bot for improving acceptance probability from our definition of $f(F, \phi)$. We then examine the impact of the benefit model on the bot, finding the counter-intuitive result that rewarding the bot for being merely a friend-of-friend of a target increases the fraction of targets befriended.

We adopt the benefit model given by Li *et al.* [18]: for a *target set* of users T, $B_f(u) = 1$ if $u \in T$, 0 otherwise; $B_{fof}(u) = \frac{1}{2}B_f(u)$; and $B_e(u, v) = 2^{T_u + T_v}/M$ where M is the maximum expected degree of any user on the network, $T_u = 1$ if $u \in T$ and 0 otherwise. We additionally use their "degree incentive" function to model the tendency of high-degree users to accept friend requests without critical examination. We set $\epsilon = 0.2$, which yields $\beta \leq 10$ and gives the performance ratio of roughly 10%. Our simulations are implemented in Rust.[7][8] Unless otherwise specified, we run each method 250 times per data point and plot the mean.

Our experiments focus on the DBLP (317k nodes, 1M edges, 13.5k ground-truth communities) and Slashdot 2008 (77k nodes, 516k edges) datasets, both taken from SNAP [17]. As location of targets may impact the attack policy, we focus on two natural target models in our experiment. First, a simple breadth-first-search from a randomly chosen node, collecting 100 possible neighbors (i.e. the BFS progresses as if each potential edge exists). This models an adversary taking a simple topological approach to building a target list. Alternately, the attacker could obtain a ground-truth list of targets from an external source (e.g. an organizational list of employees). We model this by targeting ground-truth communities, which have been provided for the DBLP topology. These are paired with a baseline we term "Untargeted:" every user is assigned a random $b_u \in [0, 1]$ and we set $B_f(u) = b_u$. This models the attacker wishing only to collect private data, but having some idea of which users are more likely to give high return for their investment.

4.1 Equivalence to Prior Models

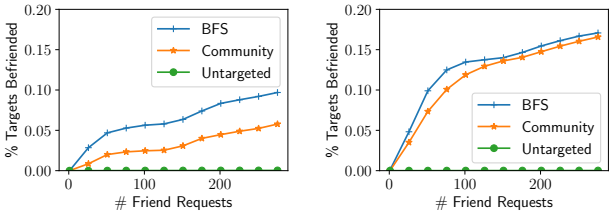
In prior work, an additional term is present in the optimization objective that directly rewards the bot for increasing the acceptance probability α of other nodes. Note that this term is absent from our formulation (compare Eqn. (4) to the equivalent in e.g. [18]). Figure 2 shows that this change does very little to alter the total benefit gained. However, it does not rule out the possibility that the corresponding policies encode different choices. Golovin & Krause observe that a policy π can be viewed as a decision tree encoding the actions to take based on the current partial realization [11]. We wish to verify whether the decision tree for the α-rewarding policy (which we will denote π_α) is fundamentally different from the policy encoded by Alg. 1 (simply denoted π).

As the problem of constructing – let alone comparing – these decision trees is quite difficult, we take the simpler approach of qualitative comparison to check for macro-level differences in behavior. Figure 3 shows a sample sets of *traces* for each policy. We see very similar patterns of behavior in both policies: early attempts to establish a foothold, followed by exploitation of that foothold to befriend other users. From these figures, it is clear that there is not a large difference in behavior caused by omitting this term, and therefore we assume that prior results apply to this formulation.

4.2 Befriending Targets by Encouraging Friend-of-Friend Relations

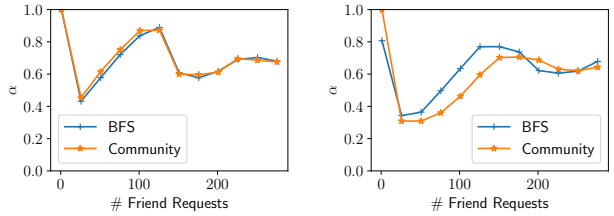
According to Li *et al.*, the benefit model they use is constructed to reward the bot based on the amount of information it may obtain from friendship: full benefit for befriending a user, part of that benefit for friend-of-friend relations, and a small additional amount for revealing an edge (present or not). We observe that this means the bot is rewarded a nontrivial amount for only becoming a friend-of-friend of targets. This leads us to ask what impact this has on the bot's success in infiltrating the target set.

[6] We remark that if one lets $g'(S) = c$, $c > 0$ rather than setting it to 0 for non-solution sets, then *any* algorithm obtains a $1/c$ approximation ratio.
[7] https://rust-lang.org
[8] Code is available at https://gitlab.com/emallson/ht2018-experiments.



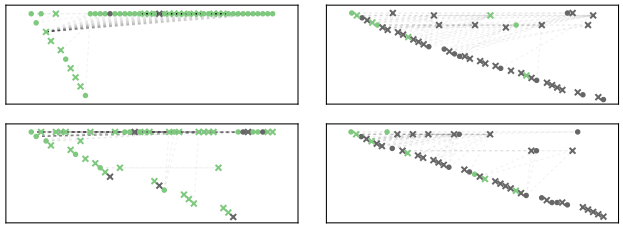
(a) $B_{fof}(v) = 0$ **(b)** $B_{fof}(v) = \frac{1}{2}B_f(v)$

Figure 4: Fraction of target set T befriended by the bot as a function of the # of requests sent on the DBLP network.



(a) $B_{fof}(v) = 0$ **(b)** $B_{fof}(v) = \frac{1}{2}B_f(v)$

Figure 5: Acceptance probability of each request sent by the bot on the DBLP network.



(a) $B_{fof}(v) = 0$ **(b)** $B_{fof}(v) = \frac{1}{2}B_f(v)$

Figure 6: Sample traces with community targets on DBLP both (a) without and (b) with FoF benefit.

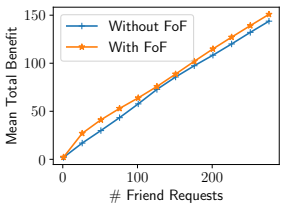


Figure 7: Mean Benefit gained with and without FoF benefit when targeting communities on DBLP.

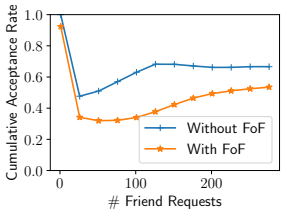


Figure 8: Overall acceptance rate over the lifetime of a bot when targeting communities on DBLP.

Figure 4 shows that the "consolation prize" of $B_{fof}(u) > 0$ actually increases the fraction of targets befriended, although as seen in Fig. 7 the amount of benefit gained remains similar. On the other hand, Figure 5 shows that this setting leads to a lower overall probability of acceptance of requests sent by the bot. Figure 6 shows sample traces covering the first 50 requests when a community is targeted on DBLP. Immediately apparent are two results: the bot having friend-of-friend benefit focuses almost all of its requests on targeted nodes, while the one without is dramatically more successful early in the process. We note that the mean acceptance probability shown in Fig. 5 is similar early in the process, so these degree of failure seen in Fig. 6b (which does not show average behavior) is likely not representative.

However, this does illuminate the change caused by removing friend-of-friend benefit. Without this benefit source, the higher probability of befriending untargeted users and for their edge benefit edges out the value of befriending target users. The addition friend-of-friend benefit causes a greater number of targets to be befriended through a pair of effects: the direct effect of greater benefit for clustered targets and the secondary effect of improving the acceptance probability of targets by befriending their neighbors.

We note that the side effect of reducing early acceptance rate may not be worth the cost, however. The use of rejection rate in bot detection was remarked upon by the developers of the Facebook Immune System [24], although the overall acceptance rate seen by Boshmaf *et al.* during the same time period was quite low (19.3% in the first 6 days, moving up to 59.1% over the entire 6-week experiment) [5]. Figure 8 shows the mean acceptance rate (defined as the rate of successful requests, as opposed to the predicted values shown in Fig. 5) over the lifetime of the bot. Notably, the acceptance rate never drops as low as the sub-20% seen in the experiments of Boshmaf *et al.* However, there is a sizable gap between the rates of a bot operating with and without friend-of-friend benefit.

5 CONCLUSION & FUTURE WORK

In this work, we developed a novel technique for bounding the approximation quality of a greedy socialbot conducting a reconnaissance attack under a *realistic* model of user behavior. This was then generalized to provide a bound for the much broader class of adaptive monotone non-submodular problems given a bound on the *adaptive primal curvature* (APC) of their objective functions. We further showed that a finite bound on the APC is necessary for any approximation guarantee to hold for any algorithm. Our definition of curvature differs from those used in prior work, which indicates that more study is needed to identify the exact properties necessary to obtain an approximation ratio.

We then conducted further analysis of the behavior of a socialbot under this realistic model. Notably, we found the counter-intuitive result that rewarding the bot for becoming friends-of-friends of its targets actually improved the rate at which it befriended targets – at a small cost to overall acceptance rate. This counter-intuitive result leads us to note that the current definition of benefit may be sub-optimal. Further study should be devoted to finding the optimal scheme for assigning benefit to users to maximize particular metrics (e.g. target friending rate). We also note that our model of user acceptance only incorporates topological features. Future work may explore the impact of profile and temporal features on acceptance probability, and optimality under such features.

REFERENCES

[1] Lars Backstrom and Jure Leskovec. 2011. Supervised random walks: predicting and recommending links in social networks. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 635–644.

[2] Andrew An Bian, Joachim Buhmann, Andreas Krause, and Sebastian Tschiatschek. 2017. Guarantees for Greedy Maximization of Non-submodular Functions with Applications. In *Proceedings of the 34th International Conference on Machine Learning (ICML '17)*.

[3] Yazan Boshmaf, Konstantin Beznosov, and Matei Ripeanu. 2013. Graph-based sybil detection in social and information systems. In *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on*. IEEE, 466–473.

[4] Yazan Boshmaf, Dionysios Logothetis, Georgos Siganos, Jorge Lería, Jose Lorenzo, Matei Ripeanu, and Konstantin Beznosov. 2015. Íntegro: Leveraging Victim Prediction for Robust Fake Account Detection in OSNs. In *Proc. of NDSS*.

[5] Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu. 2011. The Socialbot Network: When Bots Socialize for Fame and Money. In *Proceedings of the 27th Annual Computer Security Applications Conference (ACSAC '11)*. ACM, 93–102. https://doi.org/10.1145/2076732.2076746

[6] Abhimanyu Das and David Kempe. 2011. Submodular Meets Spectral: Greedy Algorithms for Subset Selection, Sparse Approximation and Dictionary Selection. (2011). arXiv:cs, stat/1102.3975

[7] Uriel Feige. 1998. A Threshold of Ln N for Approximating Set Cover. 45, 4 (1998), 634–652. https://doi.org/10.1145/285055.285059

[8] Michael Fire, Rami Puzis, and Yuval Elovici. 2013. Link prediction in highly fractional data sets. In *Handbook of computational approaches to counterterrorism*. Springer, 283–300.

[9] Michael Fire, Lena Tenenboim, Ofrit Lesser, Rami Puzis, Lior Rokach, and Yuval Elovici. 2011. Link prediction in social networks using computationally efficient topological features. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*. IEEE, 73–80.

[10] Carlos Freitas, Fabricio Benevenuto, Saptarshi Ghosh, and Adriano Veloso. 2015. Reverse Engineering Socialbot Infiltration Strategies in Twitter. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 (ASONAM '15)*. ACM, 25–32. https://doi.org/10.1145/2808797.2809292

[11] Daniel Golovin and Andreas Krause. 2011. Adaptive Submodularity: Theory and Applications in Active Learning and Stochastic Optimization. 42 (2011), 427–486.

[12] Thibaut Horel and Yaron Singer. 2015. Scalable Methods for Adaptively Seeding a Social Network. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. International World Wide Web Conferences Steering Committee, 441–451. https://doi.org/10.1145/2736277.2741127

[13] Thibaut Horel and Yaron Singer. 2016. Maximization of Approximately Submodular Functions. In *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 3045–3053.

[14] Rishabh K Iyer, Stefanie Jegelka, and Jeff A Bilmes. 2013. Curvature and Optimal Algorithms for Learning and Minimizing Submodular Functions. (2013), 2742–2750.

[15] Richard M. Karp. 1972. Reducibility among Combinatorial Problems. In *Complexity of Computer Computations*, Raymond E. Miller, James W. Thatcher, and Jean D. Bohlinger (Eds.). Springer US, 85–103. DOI: 10.1007/978-1-4684-2001-2_9.

[16] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the Spread of Influence Through a Social Network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03)*. ACM, 137–146. https://doi.org/10.1145/956750.956769

[17] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data. (June 2014).

[18] Xiang Li, J. David Smith, Thang N. Dinh, and My T. Thai. 2016. Privacy Issues in Light of Reconnaissance Attacks with Incomplete Information. In *Proceedings of the 2016 IEEE/WIC/ACM International Conference on Web Intelligence*. IEEE/WIC/ACM.

[19] Xiang Li, J. David Smith, and My T. Thai. 2017. Adaptive Reconnaissance Attacks with Near-Optimal Parallel Batching. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)* (2017-06). 699–709. https://doi.org/10.1109/ICDCS.2017.130

[20] George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher. 1978. An Analysis of Approximations for Maximizing Submodular Set Functions—I. 14, 1 (1978), 265–294.

[21] Nam P Nguyen, Guanhua Yan, and My T Thai. 2013. Analysis of misinformation containment in online social networks. *Computer Networks* 57, 10 (2013), 2133–2146.

[22] Abigail Paradise, Asaf Shabtai, and Rami Puzis. 2015. Hunting Organization-Targeted Socialbots. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 (ASONAM '15)*. ACM, New York, NY, USA, 537–540. https://doi.org/10.1145/2808797.2809396

[23] Thomas Ryan and G Mauch. 2010. Getting in bed with Robin Sage. In *Black Hat Conference*.

[24] Tao Stein, Erdong Chen, and Karan Mangla. 2011. Facebook Immune System. In *Proceedings of the 4th Workshop on Social Network Systems (SNS '11)*. ACM, 8:1–8:8. https://doi.org/10.1145/1989656.1989664

[25] Guangmo Tong, Weili Wu, Shaojie Tang, and Ding-Zhu Du. 2017. Adaptive Influence Maximization in Dynamic Social Networks. 25, 1 (2017), 112–125. https://doi.org/10.1109/TNET.2016.2563397

[26] Onur Varol, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online Human-Bot Interactions: Detection, Estimation, and Characterization. In *Eleventh International AAAI Conference on Web and Social Media* (2017-05-03). https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15587

[27] Zengfu Wang, Bill Moran, Xuezhi Wang, and Quan Pan. 2014. Approximation for Maximizing Monotone Non-Decreasing Set Functions with a Greedy Method. 31, 1 (2014), 29–43. https://doi.org/10.1007/s10878-014-9707-3