

Centralization Problem for Opinion Convergence in Decentralized Networks

Yiping Liu¹
yliu823@aucklanduni.ac.nz

Jiamou Liu¹
jiamou.liu@auckland.ac.nz

Bakhadyr Khoussainov²
bmk@uestc.edu.cn

Miao Qiao¹
miao.qiao@auckland.ac.nz

Mengxiao Zhang²
mengxiao.zhang@uestc.edu.cn

Bo Yan³
yanbo@bit.edu.cn

¹ School of computer science, University of Auckland, Auckland, New Zealand

² School of computer science, University of Electronic Science and Technology of China, ChengDu, China

³ School of Computer Science, Beijing Institute of Technology, BeiJing, China

Abstract—This paper presents a novel perspective on the relationship between decentralization, a prevalent characteristic of multi-agent systems, and centralization, which involves imposing central control to achieve system-level objectives. Specifically, within the context of a networked opinion dynamic model, we introduce and discuss a framework for centralization. In this framework, a decentralized network consists of autonomous agents and a dynamic, unknown social structure. Centralization involves appointing specific agents in the network as access units, responsible for providing information and exerting influence within their local environments. We focus on centralization for the DeGroot model of opinion dynamics, aiming to achieve opinion convergence with the minimum number of access units. To accomplish this, we demonstrate that selecting access units to form a dominating set is crucial. Moreover, we propose algorithms based on a new local algorithmic framework called *prowling* to facilitate this process. Through systematic experiments conducted on both real-world and synthetic networks, we validate our algorithm and show its superiority over benchmark methods.

Index Terms—Social network, dynamic network, partially-known network, opinion dynamics, dominating set.

I. INTRODUCTION

Opinion dynamics aims to illustrate how opinions of individuals evolve in a crowd [10], so as to provide insights on important challenges in multi-agent systems such as group decision making [13]. In an opinion dynamics model, every agent holds an opinion towards a certain issue. The agents interact by revealing their opinions to others, and in turn updating their opinions as a result of such interactions. Such a model operates in a *decentralized* manner, characterized by two key aspects. Firstly, the agents' states, including their opinions and social interactions, evolve without the control

of a higher-level entity or a “system controller”. Secondly, in systems with a large number of agents, any potential system controller only has *partial information* about these states, as direct access to individual opinions and social interactions is limited. However, in such decentralized systems, achieving consensus among the agents, a primary objective of the opinion dynamics process, may be challenging. There is a need to impose some coordinated control.

Centralized control on decentralized systems has garnered significant attention in various fields, e.g., networking and political science [14], [15]. In line with this concept, we formulate *decentralized networks* to capture the type of opinion dynamics models described above. Specifically, a decentralized network consists of agents whose repeated interactions form a social network, incorporating the following key properties: (1) **Dynamic topology**: The social network's structure is not static but evolves over time. The evolution makes influence on opinion dynamics [22]. (2) **Unknown topology**: The social network's initial structure is unknown, mainly due to its large scale and the localized knowledge of agents. Consequently, discovering and understanding the network topology necessitates exploration methods [17], [23].

In a decentralized network, we define the process of *centralization* that employs two functionalities: **monitoring** and **influencing**. Monitoring involves appointing a subset of agents as “access units” responsible for gathering information about a specific local region. Influencing refers to the ability of these access units to exert control or influence over other agents based on directions from the system controller. This setup captures two extrema: a fully decentralized system where no agent serves as an access unit, and a fully centralized system where all agents act as access units. The centralization process involves gradually appointing access units, starting from the decentralized extreme and moving towards the centralized extreme, until an appropriate level of control is achieved. The challenge lies in devising suitable strategies to explore the unknown and dynamic social network and appoint access units effectively along the way.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

© 2023 Association for Computing Machinery.

ACM ISBN 979-8-4007-0409-3/23/11 \$15.00

<http://doi.org/10.1145/3625007.3627291>

This paper focuses on the centralization process for opinion convergence. We adopt a decentralized version of the well-established *DeGroot model* to drive agents' opinions [11]. In this model, each agent's opinion is represented as a real value in the range $[0, 1]$, and it is updated based on the average opinion of its neighboring agents. The objective is to ensure that all agents' opinions converge to 1. One paradigm to achieve this, as proposed in [22], is to deploy "committed agents" whose opinion is fixed at 1. We model these committed agents as the access units in our centralization processes. We model the centralization problem within the decentralized DeGroot model, where the goal is to converge all agents' opinions to 1 using the minimum number of access units. Our contributions can be summarized as follows:

- (1). We define decentralized networks and formulate the centralization problem over decentralized DeGroot model. These are presented in Sections III and IV. These concepts can have broader implications and applications beyond opinion convergence, such as influence maximization or norm emergence [21], [25], thereby triggering fruitful new research.
- (2). We establish a connection between finding a dominating set in network theory and the centralization problem for opinion convergence. We provide theoretical results and algorithms that address the centralization problem by finding a minimum dominating set. See Sec. IV and V. These form the foundation of our experimental analysis in the last section.
- (3). To solve the centralization problem, we explore local algorithms and propose a novel algorithmic framework called *prowling*. Prowling algorithms probe unknown networks and select access units effectively. We present three prowling algorithms, including an exterior-based algorithm and two B-set prowling algorithms. See Sec. V. These algorithms ensure opinion convergence under general conditions. We validate the effectiveness of our algorithms through extensive experiments on dynamic real-world and synthetic datasets. See Sec. VI.

II. RELATED WORK

This paper is related to the field of *local algorithms* which aims to identify global properties of a graph using only local information [7]. One relevant work in this area is presented by [5], where they propose an algorithm for exploring unknown networks. The algorithm selects a "known" node greedily and then makes random moves to unknown neighboring nodes. This method guarantees coverage of a static graph using a small set of nodes, with an approximation factor of $O(\log(\Delta))$, where Δ is the maximum degree of the graph. If we interpret the nodes visited by the local algorithm as access units, it aligns with the framework studied in our paper. Wilder et al. utilize local algorithms for influence maximization [23], where they design a *probing* mechanism to collect information, i.e., walks in the unknown part of the network. The structural statistics collected by repeated probes direct the selection of seed nodes, which can also be considered as access units. In our paper, we adopt a different probing method. Instead of using probes as means to sample information, we implement

a search procedure through the unknown part of the network to locate the next access unit. We call such types of algorithms *prowling* and consider this procedure as a less costly and adaptive alternative to the exploration strategy used in [23].

Opinion dynamics, with its application to group decision making, has been a subject of extensive research [1], [12], [13], [16]. However, most existing work in this field assumes that the social network is known and static, which allows for the direct application of graph theory and combinatorial optimization algorithms. Singh et al. [22] explore the co-evolution of network structures and individual opinions and demonstrate that the presence of committed agents can expedite opinion convergence. Meanwhile, He et al. [17] focus on opinion maximization by selecting seed agents when only partial information about edge weights is available. However, their work overlooks the crucial aspect of information collection, which holds significant importance in real-world applications.

While there is a lack of research specifically investigating opinion dynamics in unknown and dynamic social networks, there exist several studies that focus on the related topic of *influence maximization* in such networks. Influence is typically modeled using cascading models, which give rise to submodular functions. In the absence of the social network structure, it is common to assume that sampled influence pairs are available to extract information about the network structure [2], [3]. These works utilize submodular function optimization, which is not directly applicable to models of opinion dynamics. Wu et al. [24] assume that the network topology is known while the edge weights are unknown in a cascade model. Zhuang et al. [26] study networks that are initially known but undergo structural changes that can only be observed through probing. Their approach involves selecting a set of probing nodes to reveal neighborhood topology and subsequently selecting seed nodes accordingly. Although our setting bears similarities to [26] in that we consider probing and selection of access units, there are key differences. Our network is initially unknown, and we focus on opinion dynamics rather than information diffusion. Therefore, we tackle unique challenges in exploring and controlling the opinion convergence process in unknown and dynamic social networks.

Another distinctive aspect of our model, which sets it apart from the aforementioned work, is the iterative selection of access units as the system evolves, making centralization an *adaptive* process. This bears resemblance to the setting of *adaptive influence maximization*, where seed nodes are iteratively selected while observing network evolution [25].

III. DECENTRALIZED NETWORK

Assume a fixed (possibly infinite) universal set of *agents* U . Let Q be a set of *states* of the agents.

Definition 1. A network instance is $G = (V, E, C, S)$ where $V \subseteq U$ represents a finite set of nodes, E represents edges $E \subseteq [V]^2$, $C: V \rightarrow Q$ assigns node states, and $S \subseteq V$ is a set of access units.

Definition 2. A decentralized network is a sequence of network instances $\mathfrak{G} = (G_1, G_2, \dots, G_t, \dots)$ where $t = 0, 1, 2, \dots$ and G_t is a network instance at time t .

In our decentralized network model, the access units S_t represent the means for a system controller to monitor and influence the agents. Each access unit $v \in S_t$ has an associated accessible region $AC_t(v) \subseteq V_t$ that includes v .

The *accessible region* of set S is $AC_t(S) = \bigcup_{v \in S} AC_t(v)$. The accessible region plays two roles: (1) Scope of influence. $AC_t(v)$ determines the extent of influence that node v can exert. When a node u is accessible to two distinct access units, its opinion acts in the same manner regardless of which specific access region it belongs to. (2) Information Accessibility. Only nodes in $AC_t(S_t)$, their states, as well as all edges attached to $AC_t(S_t)$ are known by the system controller; any other information about the network is not accessible, nor influenced by the system controller. Note that, with these assumptions, the system controller can also tell whether there is a node $v \notin AC_t(S_t)$ that is connected to a node in $AC_t(S_t)$.

The rest paper models a decentralized network as a decentralized DeGroot model. We restrict ourselves to *incremental networks* where any edge or nodes, once appeared, will not be deleted, and thus $V_t \subseteq V_{t+1}, E_t \subseteq E_{t+1}$. The opinion of any node $v \in U \setminus V_t$ is irrelevant ($v(t) = 0$ for simplicity).

Write $\text{dist}_t(u, v)$ for the length of a shortest path in E_t for $u, v \in V_t$. The *accessible region* of node v is parametrized by a *radius* r : $AC_t^{(r)}(v) = \{u \in U \mid \text{dist}_t(u, v) \leq r\}$. As r is fixed, we write $AC_t(\cdot)$ for $AC_t^{(r)}(\cdot)$. In principle, r can be set to any positive integer, we usually assume that an access unit can only influence its close proximity and thus $r \leq 2$.

Following DeGroot model [11], an opinion is a value in $Q = [0, 1]$, and agents align their opinion with their direct information sources. The system controller imposes control as follows: (1) Each $u \in S_t$ is a “committed agent” with a fixed opinion at value 1. (2) Agents not in the accessible region $AC_t(S_t)$ update their opinion at time $t + 1$ as the average of their neighbors’ opinions. (3) Agents within $AC_t(S_t)$ are influenced by S_t and treat all committed neighbors as a single information source. At time $t + 1$, their opinion is determined by averaging the opinions of all their information sources, i.e. non-committed neighbors and a single committed neighbor.

This definition closely resembles the DeGroot model with committed agents [1], [20]. The distinction is that in our model, the effect of multiple access units on an agent is treated as equivalent to the effect of a single unit, as we consider the committed agents as a collective entity. In the original DeGroot model, all agents are treated equally without such a distinction.

IV. THE CENTRALIZATION PROBLEM

Centralization is a computational process performed by the system controller that selects access units in a decentralized network. As the network is initially unknown, selecting an access unit requires a form of “information gathering”. Starting with an initial network instance (V_0, E_0, C_0, S_0) , where

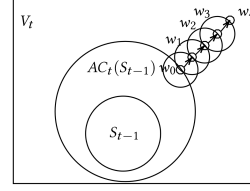


Fig. 1: A 4-step prowling: Starting from w_0 in the accessible region, the prowling produces the node w_4 as next access unit.

$S_0 = \emptyset$, the process repeatedly executes a *local algorithm* that selects access units. When executed at time $t > 0$, the algorithm uses information provided in $AC_t(S_{t-1})$ (states and edges of nodes) and returns a node $u_t \notin S_{t-1}$, which is the next access unit, i.e., $S_t = S_{t-1} \cup \{u_t\}$.

We now propose a type of local algorithms called *prowling*. Let the *closed neighborhood* of $v \in V_t$ be $N_t(v) = \{u \in V_t \mid \text{dist}(u, v) \leq 1\}$. The 1^+ -neighborhood $N_t^+(v)$ of node v is a data structure containing the states of all $v \in N_t(v)$ and their degrees in G_t .

Definition 3. A (k -step) prowling algorithm $\text{prowl}(S_{t-1})$, given the set S_{t-1} , implements a search strategy at time t :

- Start at an agent w_0 from $AC_t(S_{t-1})$.
- For $0 \leq i < k$, if $w_i \in AC_t(S_{t-1})$, move from w_i to a neighbor w_{i+1} of w_i ; if $w_i \notin AC_t(S_{t-1})$, then first inquire about the 1^+ -neighborhood $N_t^+(w_i)$ and then move from w_i to a neighbor w_{i+1} of w_i based on the obtained information.
- The last agent w_k is chosen as output.

The prowling algorithm resembles a *name-generation process* used in social network analysis [8], where an interviewer asks a participant (an agent in the network) to reveal the names of those who came into direct contact with the participant, then those whose names are revealed are asked the same question and so on. Note that every step of the prowling algorithm can make use of the degree of the current node. The length k is a fixed parameter. Fig. 1 illustrates a 4-step prowling.

Initially the opinion values of all agents are 0: $v(0) = 0$ for $v \in U$. The desired outcome of a centralization process is when the opinions of all agents in V_t converge to 1.

Definition 4. We say that the opinions converge in a decentralized network \mathfrak{G} if $\forall \varepsilon > 0, \exists t > 0$, such that $\forall v \in V_t: v(t) \geq 1 - \varepsilon$. In this case, the $(1 - \varepsilon)$ -centralization cost of \mathfrak{G} is $|S_t|$ where t is the first time when the condition above holds.

The network centralization problem asks for a prowling algorithm that explores a decentralized network \mathfrak{G} and selects access units to achieve opinion convergence with the smallest centralization cost.

Remark The choice of prowling as our approach for network centralization is motivated by several factors. Firstly, traditional local algorithms, as described in [5], only explore the network within the known region, meaning that the selected access units are always restricted to $AC_t(S_{t-1})$. While this approach adheres to the principle of accessing only local information, it may result in selecting too many access units

when applied to large networks. Secondly, the probing algorithms proposed in [23] are designed for sampling paths in unknown networks, allowing for the selection of access units beyond the currently accessible region. However, this method can be costly when extensive repeated sampling is required to choose a single access unit. Furthermore, this algorithm assumes a static network structure, which may not hold in dynamic scenarios. Prowling strikes a balance between these two approaches. It enables the selection of nodes outside of the accessible region as the network evolves, without incurring the high cost associated with repeated sampling. This adaptability makes prowling a more practical and efficient alternative.

A. From stabilized to unit-growth networks

We first consider the case when the network *stabilizes* at some time $\ell > 0$, i.e., $e(t) = e(\ell)$ for all $e \in [U]^2$ and $t > \ell$.

Theorem 1. *If a decentralized network stabilizes at a time ℓ with $S_\ell \neq \emptyset$, then opinions converge.*

Proof. Let $\iota = \min\{t \mid S_t \neq \emptyset\}$. For convenience and w.l.o.g., we prove the case when $\iota = \ell$, $S_\ell = \{s\}$ for some $s \in V_\ell$ and $v(\iota) = 0$ for all $v \in V_\ell \setminus \{s\}$. We show that $\forall v \in V_\ell \forall t \geq \iota$, $v(t+1) \geq v(t)$. Indeed, for any $v \in V_\ell$, let $t_v = \min_t \{t \in \mathbb{N} \mid v(t+1) < v(t), t \geq \iota\}$ (where $\min_t \emptyset = \infty$). Pick w from $\text{argmin}_v \{t_v \mid v \in V_\ell \setminus \{s\}\}$. Suppose $t_w < \infty$. Note that $t_w > \iota$, as clearly for any $v \in V_\ell \setminus \{s\}$, $v(\iota+1) \geq v(\iota)$. Since $w(t_w+1) < w(t_w)$ and our model is a DeGroot model, there is a neighbor $x \in N_{t_w}(w)$ such that $x(t_w) < x(t_w-1)$. This means that $t(x) < t(w)$, contradicting the def. of w .

Thus, the opinion of each $v \in V_\ell$ is a monotonic sequence $v(\ell), v(\ell+1), \dots$ bounded by 1. Let v^* be the limit of this sequence. Let y be a node in $\text{argmin}_v v^*$ and assume $y^* < 1$. Then take such a y that has the closest distance from s . Since our model is a DeGroot model, among the neighbors of y , there is an $x \in N_\ell(y)$ with $x^* > y^*$, and all other neighbors z of y have $z^* \geq y^*$. This means that the sequence $v(\ell), v(\ell+1), \dots$ will eventually surpass y^* . \square

Corollary 1. *If a decentralized DeGroot model stabilizes, then one access unit suffices for opinion convergence.* \square

When the network does not stabilize, the Thm.1 no longer holds. For example, when each time t involves two (or more) nodes entering V_t , one of the them could always have opinion 0 at time t . Nevertheless, when we assume that the network $\mathfrak{G} = \{G_t\}_{t \in \mathbb{N}}$ has *unit-growth*, that is, V_{t-1} expands by at most one node $v \in U \setminus V_{t-1}$ at each time t , it is possible to guarantee convergence of opinions. In this case, a *dominating set*, a key notion in graph theory, plays a crucial role. The set S_t is called (*distance- r*) *dominating* if $\text{AC}_t^{(r)}(S_t) = V_t$.

Definition 5. *A prowling algorithm is domination-driven if for any unit-growth decentralized network \mathfrak{G} there is a time t at which the algorithm produces S_t that is dominating for G_t .*

Algorithm 1 is a centralization process that makes use of a domination-driven prowling algorithm. Note that the network evolution (Lines 4 & 9) is beyond our control. The **while**-loop

repeatedly executes the prowling algorithm until S_t dominates. From then on, the algorithm will select new access units whenever V_t expands.

Definition 6. *A decentralized network $\mathfrak{G} = \{G_t\}_{t \in \mathbb{N}}$ is consistent with prowling if each S_t is produced by Algorithm 1.*

Algorithm 1 A domination-driven centralization process

PARAMETERS radius r , num of steps k of prowling

- 1: Decentralized network starts at G_0 at time $t \leftarrow 0$
 - 2: $S_0 \leftarrow \emptyset$
 - 3: **while** $V_t \setminus \text{AC}_t(S_t) \neq \emptyset$ **do**
 - 4: The network evolves to G_{t+1} and set $t \leftarrow t + 1$
 - 5: $v \leftarrow \text{prowl}(S_{t-1})$ $\triangleright k$ -step prowling
 - 6: $S_t \leftarrow S_{t-1} \cup \{v\}$
 - 7: **repeat**
 - 8: The network evolves to G_{t+1} and set $t \leftarrow t + 1$
 - 9: **if** $V_t \setminus \text{AC}_t(S_{t-1}) \neq \emptyset$ **then** \triangleright A new node enters V_t
 - 10: Select $u \in V_t \setminus \text{AC}_t(S_{t-1})$ \triangleright The only new node
 - 11: $S_t \leftarrow S_{t-1} \cup \{u\}$
 - 12: **until** Termination \triangleright It runs infinitely as G_t evolves.
-

Theorem 2. *If decentralized network \mathfrak{G} is consistent with a domination-driven prowling and \mathfrak{G} has the unit-growth property, then in \mathfrak{G} opinions converge.*

Proof. Suppose S_ℓ is dominating to G_ℓ at time $\ell > 0$. At the first time $t > \ell$ when there is a node $v \in V_t \setminus \text{AC}_t(S_{t-1})$, the algorithm selects the only node v from $V_t \setminus \text{AC}_t(S_{t-1})$. Thus $S_t = S_{t-1} \cup \{v\}$ is dominating to G_t . Repeating this procedure, one can prove inductively that for any $t' > \ell$, $S_{t'}$ forms a dominating set to $G_{t'}$ that contains all nodes in $V_{t'} \setminus V_\ell$. Using a similar argument as in Thm.1, we can show that opinions in the sub-network restricted to nodes in V_ℓ converge, and hence the theorem holds. \square

The *domination cost* $\delta(\mathfrak{G})$ of a domination-driven prowling on a unit-growth network $\mathfrak{G} = \{G_t\}_{t \in \mathbb{N}}$ consistent with prowling is the least t at which S_t dominates. The theorem above links the centralization and the domination costs: For any $0 < \epsilon < 1$, let t_ϵ be the least number of steps for all agents in $G_{\delta(\mathfrak{G})} = (V_{\delta(\mathfrak{G})}, E_{\delta(\mathfrak{G})}, C_{\delta(\mathfrak{G})}, S_{\delta(\mathfrak{G})})$ to reach an opinion value of at least $1 - \epsilon$. The $(1 - \epsilon)$ -centralization cost is then no more than $\delta(\mathfrak{G}) + t_\epsilon$. We will verify in our experiments that the value t_ϵ is negligible (close to 0) even for very small ϵ .

V. THE PROWLING ALGORITHMS

Assuming unit growth, the following designs domination-driven prowling algorithms to minimize domination cost.

A. Exterior-based Prowling

We exhibit a simple type of domination-driven prowling algorithm. Define the *exterior set* as $X_t = V_t \setminus \text{AC}_t(S_{t-1})$.

Definition 7. *A prowling algorithm $\text{prowl}(S_{t-1})$ is exterior-based if the algorithm returns a node $u_t \in X_t$.*

The exterior-based prowling algorithms are domination-driven for radius $r \geq 2$. Assume that at each time $t > 0$, a new node $v_t \notin AC_t(S_{t-1})$ enters V_t , i.e., $V_t = V_{t-1} \cup \{v_t\}$. We need the following. At time t , the *access neighborhood* is $N_t = \bigcup_{s \in S_t} N_t(s)$, and *anti-neighborhood* is $\bar{N}_t = V_t \setminus N_t$. We have $N_{t-1} \subseteq N_t$ and $\bar{N}_t \subseteq \bar{N}_{t-1} \cup \{v_t\}$. Here, $\bar{N}_t = \emptyset$ implies $N_t = V_t$, i.e., S_t is a (distance-1) dominating set.

Lemma 1. For any $t > 0$, $|\bar{N}_{t-1}| - |\bar{N}_t| \geq |N_t \setminus N_{t-1}| - 1$.

Proof. $|\bar{N}_{t-1}| - |\bar{N}_t| \geq |\bar{N}_{t-1} \cup \{v_t\}| - |\bar{N}_t| - 1$. By $\bar{N}_t = (\bar{N}_{t-1} \cup \{v_t\}) \cap \bar{N}_t$, it equals to $|\bar{N}_{t-1} \cup \{v_t\}| - |(\bar{N}_{t-1} \cup \{v_t\}) \cap \bar{N}_t| - 1$, which is $|\bar{N}_{t-1} \cup \{v_t\}| - |\bar{N}_t| - 1$. By $\bar{N}_{t-1} \cup \{v_t\} = V_t \setminus N_{t-1}$, we have $|(V_t \setminus N_{t-1}) \setminus \bar{N}_t| - 1$, which is $|(V_t \setminus (V_t \setminus N_t)) \setminus (N_{t-1} \setminus (V_t \setminus N_t))| - 1$. It equals to $|N_t \setminus (N_{t-1} \cap N_t)| - 1 = |N_t \setminus N_{t-1}| - 1$. \square

Theorem 3. Any exterior-based prowling algorithm with the radius $r \geq 2$ is domination-driven. Moreover, the domination cost is no more than $t + |\bar{N}_t|$ for any $t > 0$.

Proof. Let u_t be the node returned by the prowling algorithm $\text{prowl}(S_{t-1})$ when it is executed at time t . By assumption $u_t \in X_t$. Take a shortest path from u_t to a node $s \in S_{t-1}$ and let w be the node adjacent to u_t on that path. It is clear that $w \notin AC_t(S_{t-1})$ and $w \in AC_t(\{u_t\})$. This means that $\{u, w\} \subseteq N_t$ while $\{u, w\} \cap N_{t-1} = \emptyset$. Hence $|N_t \setminus N_{t-1}| \geq 2$. By Lemma 1, $|\bar{N}_{t-1}| - |\bar{N}_t| \geq |N_t \setminus N_{t-1}| - 1 = 1$ and thus $|\bar{N}_t| < |\bar{N}_{t-1}|$. Therefore for some $t' > 0$, $\bar{N}_{t'} = \emptyset$ implying $S_{t'}$ is dominating. Moreover, at any time t , the number of times before the anti-neighborhood becomes \emptyset is at most $|\bar{N}_t|$. \square

By Thm. 3, a centralization process consistent with an exterior-based prowling algorithm (with $r \geq 2$) guarantees to produce a dominating set over a unit-growth decentralized network. Furthermore, the theorem also provides a sequence of upper bounds on the domination cost, i.e., $1 + |\bar{N}_1|, 2 + |\bar{N}_2|, 3 + |\bar{N}_3|, \dots$. To lower domination cost, it makes sense to minimize values in this sequence. As V_t grows at a constant rate, this is equivalent to making the access neighborhood N_1, N_2, N_3, \dots grow as fast as possible. This naturally suggests selecting access units based on their *degrees*. We thus design the following exterior-based algorithm:

Definition 8. The k -step degree prowling algorithm $\text{xdeg}(S_{t-1})$: Start from a random node w_0 in $AC_t(S_{t-1})$ that has outgoing edge, and move to a random neighbor $w_1 \notin AC_t(S_{t-1})$. From w_i ($1 \leq i < k$), the algorithm selects among neighbors outside of $AC_t(S_{t-1})$ one that (1) has a higher degree than w_i , and (2) the degree is as high as possible, i.e., set $w_{i+1} = \arg\max\{\deg(v) \mid v \in X_t, \deg(v) > \deg(w_i)\}$. This step is repeated for at most $k - 1$ steps or until it is not possible to do so within k steps.

X-set Convergence. It directly follows from Thm. 2 and 3 that if a unit-growth decentralized network \mathcal{G} is consistent with $\text{xdeg}(S_{t-1})$ with $r \geq 2$, then in \mathcal{G} opinions converge.

B. B-set Prowling

Exterior-based algorithms confine prowling to the induced subgraph of the exterior set X_t , limiting the selection of central nodes. This becomes problematic when the exterior subgraph becomes fragmented and contains isolated nodes, leading to a high domination cost as shown in Fig. 2. Enabling prowling within the accessible region allows for the exploration of more central nodes and mitigates this issue.

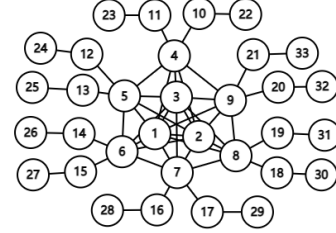


Fig. 2: A static network with 33 nodes and 51 edges with $r = 2$, the optimal solution has 6 nodes, e.g., $\{4, 5, 6, 7, 8, 9\}$. W.l.o.g, assume node 4 is the initial node for all algorithms. Then the exterior-based algorithms outputs 11 nodes, e.g., $\{4, 14, 15, 16, 17, 18, 19, 24, 25, 32, 33\}$.

We define the set $B_t = \{v \in V_t \mid \text{dist}(v, X_t) \leq r - 1\}$ when $r > 1$, and $B_t = \{v \in V_t \mid \text{dist}(v, X_t) \leq 1\}$ when $r = 1$.

Definition 9. A B-set prowling algorithm $\text{prowl}(S_{t-1})$ is one that returns a node $u_t \in B_t$.

Theorem 4. If $r \geq 2$, then a B-set prowling algorithm is domination-driven. Moreover, the domination cost is no more than $t + |\bar{N}_t|$ for any $t > 0$.

Proof. Write u_t as the node returned by the B-set $\text{prowl}(S_{t-1})$. By def., there is a node $w \in X_t$ such that $\text{dist}_t(u_t, w) \leq r - 1$.

When $u_t = w$, then $u_t \in X_t$, implying $u_t \notin N_{t-1}$. Notice that $\text{dist}_t(u_t, S_{t-1}) > r$, then for any a node $s \in S_{t-1}$, $\text{dist}_t(u_t, s) > r$. Take a shortest path between s and u_t , and let x be the node on this path that is adjacent to u_t . By triangle inequality, $\text{dist}_t(x, s) \geq \text{dist}_t(s, u_t) - 1 \geq r$. Note that s is any a node in S_{t-1} , thus $\text{dist}_t(u_t, S_{t-1}) \geq r$, meaning that $x \notin N_{t-1}$. On the other hand, $u_t \in S_t$, implying $\{u_t, x\} \subseteq N_t$. Thus $|N_t \setminus N_{t-1}| \geq 2$.

When $u_t \neq w$. In this case, we take a shortest path between w and u_t and let y be the node on this path that is adjacent to u_t . Note that y may be the same as w . Since $\text{dist}_t(w, S_{t-1}) > r$, we have $\text{dist}_t(v, w) > r - 1 \geq 1$ for any $v \in N_{t-1}$. Hence as $\text{dist}_t(u_t, w) \leq r - 1$, $u_t \notin N_{t-1}$. Also, $\text{dist}_t(y, w) \leq r - 2$ implies that $y \notin N_{t-1}$. On the other hand, $u_t \in S_t$, implying that $\{u_t, y\} \subseteq N_t$. Thus $|N_t \setminus N_{t-1}| \geq 2$.

In either case, $|N_t \setminus N_{t-1}| \geq 2$. By Lem. 1, $|\bar{N}_t| < |\bar{N}_{t-1}|$. Thus in a similar argument as in the proof of Thm. 3, there is a time $t' > 0$ where $\bar{N}_{t'} = \emptyset$ and $S_{t'}$ is dominating. And the domination cost is no more than $t + |\bar{N}_t|$ for any $t > 0$. \square

We now propose the algorithm $\text{bdeg}(S_{t-1})$.

Definition 10. The k -step B-degree algorithm $\text{bdeg}(S_{t-1})$ starts from a random node w_0 in $AC_t(S_{t-1})$ that has an edge

going out of $AC_t(S_{t-1})$, and moves to a random neighbor $w_1 \notin AC_t(S_{t-1})$. From w_i ($1 \leq i < k$), the algorithm selects among neighbors in B_t one that (1) has a higher degree than w_i , and (2) the degree is as high as possible, i.e., set $w_{i+1} = \operatorname{argmax}\{\deg(v) \mid v \in B_t, \deg(v) > \deg(w_i)\}$. This step is repeated for at most $k-1$ steps or until it is not possible to do so within $k-1$ steps.

Note that the upper bounds of the dominating costs (see Thm. 4), imply that it is preferred to make the access neighborhood N_1, N_2, N_3, \dots grow as fast as possible. Maximizing the degrees of the selected nodes does not optimally speed up the growth of N_t ; Instead we define N -degree of a node v as the number $\operatorname{ndeg}_t(v)$ of its neighbors that not in N_t .

Definition 11. The k -step B-Ndegree algorithm $\operatorname{bnde}(S_{t-1})$ starts from a random node w_0 in $AC_t(S_{t-1})$ that has an edge going out of $AC_t(S_{t-1})$, and moves to a random neighbor $w_1 \notin AC_t(S_{t-1})$. From w_i ($1 \leq i < k$), among neighbors of w_i in B_t , do:

- If $w_i \in AC_t(S_{t-1})$, then set w_{i+1} as the neighbor of w_i that has a higher N -degree and $\operatorname{ndeg}_t(w_{i+1})$ is maximized.
- if $w_i \notin AC_t(S_{t-1})$ (in this case, the N -degree of neighbors of w_i is not accessible), then set w_{i+1} as the neighbor of w_i that has a higher degree and $\deg_t(w_{i+1})$ is maximized.

This step is repeated for at most $k-1$ steps or until it is not possible to do so within $k-1$ steps.

The following directly follows from Thm. 4 and Thm. 2.

B-set Convergence. If a unit-growth decentralized network \mathcal{G} is consistent with $\operatorname{bdeg}(S_{t-1})$ or $\operatorname{bnde}(S_{t-1})$ with $r \geq 2$, then in \mathcal{G} opinions converge.

VI. EXPERIMENTS

Algorithms. We run Algorithm 1 using the three proposed prowling algorithms: xdeg , bdeg , and bnde . To put them in perspective, we also test the following benchmarks:

- **Local random algorithm** (Iran): This is a local algorithm that picks a random node adjacent to $AC_t(S_{t-1})$ as the next access unit. It does not use prowling.
- **Local degree algorithm** (ldeg): This local algorithm resembles the one proposed in [5]. When selecting the t th access unit, it picks the node $u \in AC_t(S_{t-1})$ that has the highest N -degree. This, just like Iran , also does not use prowling.
- **Exterior random algorithm** (xran): This prowling algorithm starts by moving out of $AC_t(S_t)$ and performs a random walk (of length $k-1$) in the exterior set X_t to choose the next access unit. In this way, the algorithm can be considered exterior-based, and thus Thm 3 asserts it is domination-driven.
- **B-set random algorithm** (bran): This prowling algorithm is defined in the same way as xran except that the random walk takes place in B_t rather than X_t . In this way, the algorithm is a B-set prowling algorithm, and by Thm.4 it is domination-driven.

Experiment design. We perform three experiments. Exp. 1 compares the domination (and centralization) costs of different

TABLE I: Statistics of the real-world networks. The last six rows are properties about the last time stamp.

	Facebook	Wikitalk	Citation	Enron
# of time stamps	876,992	55,197	3,308	220,363
initial time stamp	400,000	27,000	375	100,000
initial $ V $	21,059	15,807	14,526	45,596
initial $ E $	81,473	17,765	30,815	156,596
final $ V $	63,731	40,993	34,449	87,273
final $ E $	193,493	45,682	405,823	299,219
average degree	8.24	2.23	23.56	6.86
clustering coefficient	0.11	0.02	0.28	0.12
max degree	223	29,479	846	1,728
diameter	15	7	15	14

prowling algorithms over real-world networks. We vary the radius r and steps k of prowling. For each dataset, we start from the specified initial time stamp with no accessing units, running prowling until the access units form a dominating set. At any time step, we assume that $G_{t+1} = G_t$ if no successive timestamp is provided in the dataset. Exp. 2 compares the algorithms over networks generated by dynamic network models. Here, we evaluate the domination costs with different initial network size N and average degree \bar{d} . Exp. 3 further validates the findings from the two experiments above by evaluating the opinion values as a result of different prowling algorithms. We test the average opinion values of nodes in G_t during a centralization process until this value becomes higher than 0.99. An algorithm that results in faster growth in opinion values can be considered more effective at centralization. We provide our source code: <https://github.com/Leafflowave/Centralization>. **Real-world datasets.** We utilize four real-world social networks. These datasets were collected from KONNECT and SNAP and capture the interactions of agents in social networks. They serve as suitable datasets for evaluating the performance of our algorithms. Tab. I shows datasets' statistics. Note that these datasets involve only incremental changes of the networks. To ensure that the network has a sufficient size as well as a sufficient number of time stamps, we begin the centralization process from different initial time stamps. Moreover, to approximate unit growth, we selecting an access unit for every 13,1,1,2 times stamps for Facebook, Wikitalk, Citation, and Enron, respectively.

Dynamic models. We also run four models of dynamic networks. *Dynamic Barabasi-Albert model (BA)* produces preferential attachment networks [4]. The model adds a new node at each time stamp which randomly connects with $\bar{d}/2$ nodes, where \bar{d} is the average degree. *Dynamic stochastic block model (SB)* is utilized to generate networks with three equal-size communities. At each time stamp, a new node enters a random community. It then links to nodes in the same community with a high probability p_{in} , and to nodes in other communities with a low probability cp_{in} with $c = 1/6$ to mimic real-world networks as shown in [18]. *Dynamic Jackson-Rogers model (JR)* [19] simulates stochastic friendship-making among individuals in a population. There are two parameters p and \bar{d} . We pick $p = 0.5$. *Dynamic rich-club model (RC)* develops a typical core/periphery structure [6], [9]. At each timestamp,

TABLE II: Statistics of network models with $\bar{d} = 6$ and $N = 5000$

	BA	SB	JR	RC
clustering coefficient	0.01	0.0015	0.34	0.01
max degree	202	16	52	107
diameter	8	12	10	13

the model adds a new node with probability $\alpha \in [0, 1]$ or an edge between two existing nodes with probability $1 - \alpha$, where $\alpha = 2/\bar{d}$. Tab. II summarizes key statistics of the models by setting the average degree¹ $\bar{d} = 6$ on size $N = 5000$. For these models, there is no upper bound on the number of time stamps that we can generate.

Exp. 1. On real-world datasets, all prowling algorithms produce a dominating set, and hence opinions converge in the network. Our discussions above point out that the $(1 - \epsilon)$ -centralization cost equals to the domination cost plus t_ϵ (see Sec. IV-A). In this experiment, we compute the value of t_ϵ when ϵ is roughly 0. It turns out that, over the real-world networks, for all cases considered, $t_\epsilon = 0$ for $\epsilon \sim 0$. This validates our conjecture above that domination costs directly reflects centralization costs.

Tab. III shows the domination costs for $r \in \{1, 2\}$ and $k \in \{4, 7, 10\}$ after 10 runs of each algorithm. We make the following observations: First, when comparing the local algorithms (Iran, ldeg) with the prowling algorithms, the prowling algorithms in general achieve lower domination costs. This is especially so for $r = 2$ which is consistent with the theoretical analysis above. Nevertheless, the B-set prowling algorithms in general outperform the other algorithms in all cases. Second, the prowling algorithms that are based on degrees (xdeg, bdeg, bnde) in general perform better than their random-walk counterparts. This is also consistent with our theoretical analysis as we stipulate that degrees can be used as an effective criterion for centralization. Third, between the two degree-based B-set prowling algorithms, we observe comparable performance when $r = 2$. Nevertheless, bnde performs significantly better when $r = 1$ for all datasets apart from Wikitalk. This suggests that it may be more appropriate to adopt N-degree in prowling.

A slight unexpected result is that when k increases, there does not seem to be a noticeable difference on the domination cost. This may be due to that the degree ascending paths are in general very short, or the exterior set (and B-set) are fragmented as access units are selected, resulting in a lack of long paths. In all subsequent experiments, we thus set $k = 4$.

Also, Wikitalk demonstrates noticeable differences as compared to the other datasets. This can be explained by its sparsity, which indicates a much lower average degree and clustering coefficient, while it holds a node with a disproportionately high degree (29,479).

Exp. 2. We perform three tests over networks generated by the dynamic models: BA, SBM, JR, and RC. (1) Setting $r = 2$, and $N = 5000$, we then compare domination costs with varying average degree \bar{d} . (2) Finally, setting $r = 2$, $\bar{d} = 6$,

we compare domination costs with varying initial size N . (3) Setting $N = 5000$, and $\bar{d} = 6$, we first compare domination costs with varying radius r . The results are shown in Fig.3.

In general, we can see $\text{prowl}_{\text{bdeg}}$ and $\text{prowl}_{\text{bnde}}$ outperform other algorithms. As expected, when we increase the average degree from 4 to 14 or increase the radius from 1 to 4, all prowling algorithms achieve smaller domination costs, and moreover they tend to have similar performance. On the contrary, as we increase the initial size of the network, the domination costs also tend to increase. All algorithms exhibit rather stable performance across the cases where in general, bnde and bdeg achieve the best performance. ldeg results in far higher domination costs than the other algorithm in many cases. The results also reflect differences between the networks with different structural properties. In particular, ldeg tends to perform well in the scale-free networks (BA), where as all algorithms apart from ldeg perform comparably well in SBM.

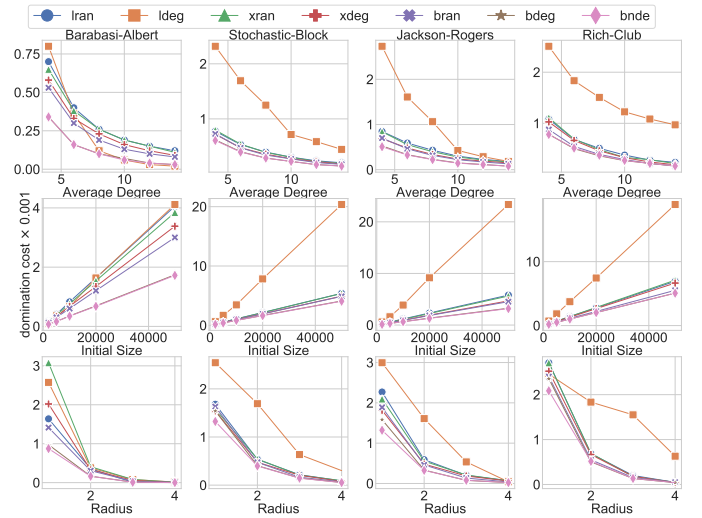


Fig. 3: Exp. 2. Domination costs of synthetic networks produced by four dynamic network models (columns) with varying average degree (top row), initial size (middle row) and radius (bottom row).

Exp. 3. For further analysis, we calculate the average opinion values during the centralization processes over the real-world networks. Fig. 4 displays changes to the average opinion starting from the initial time stamp, as more access units are chosen by different prowling algorithms (fixing $r = 2$), until the average opinion reaches above 0.8. In all datasets, the B-set prowling algorithms bdeg and bnde are the ones that lift the average opinion the fastest, and are the first ones to achieve opinion value ≥ 0.8 . It is also remarkable that this is achieved only within a small number (≤ 50) of steps. In the citation network, ldeg achieves comparable performance as the prowling algorithms, which may be explained by the high density of this dataset. In general, the degree-based prowling methods perform better than their random-walk counterparts.

VII. CONCLUSION AND FUTURE WORK

This paper investigates *centralization problem* in a *decentralized networks*. From a dynamic and unknown networks

¹80 datasets on KONECT and SNAP have average degrees 2-10.

TABLE III: Exp. 1. The domination costs of different prowling algorithms over real-world datasets for $r \in \{1, 2\}$ and $k \in \{4, 7, 10\}$.

r	Facebook						Wikitalk					
	1			2			1			2		
	4	7	10	4	7	10	4	7	10	4	7	10
Iran	8,490±80			3,831±39			14,370±4,090			187±9		
ldeg	7,868±19			5,579±16			196±0			184±0		
xran	8,333±64	8,328±42	8,322±49	3,866±33	3,831±22	3,854±18	6,890±6,119	13,372±6,024	12,183±4,806	185±20	192±17	181±21
xdeg	7,619±52	7,606±42	7,586±19	3,710±51	3,692±26	3,725±44	2,701±24	2,713±5	2,712±10	185±7	185±7	183±8
bran	7,256±102	7,566±91	7,328±63	3,358±33	3,470±15	3,388±23	204±4	173±10	210±4	155±5	171±10	157±7
bdeg	6,926±29	6,943±32	6,916±40	3,210±16	3,205±30	3,220±27	194±0	194±0	194±0	148±0	148±0	148±0
bnde	6,360±14	6,348±26	6,353±13	3,225±8	3,210±20	3,218±45	194±0	194±0	194±0	148±0	148±0	148±0
r	Citation						Enron					
	1			2			1			2		
	4	7	10	4	7	10	4	7	10	4	7	10
Iran	2,457±45			498±14			21,313±12405			824±32		
ldeg	1,748±32			870±10			2,514±9			2,059±11		
xran	2,405±48	2,471±40	2,416±31	495±15	485±12	503±23	19,697±175	20,906±395	19,593±940	811±11	811±25	808±15
xdeg	2,337±30	2,290±57	2,300±36	455±24	465±17	463±23	17,693±590	17,684±145	17,817±704	777±12	771±11	780±23
bran	17,57±38	2,105±59	1,819±35	338±19	444±25	359±16	2,397±17	12,457±311	2,556±31	741±36	755±30	720±28
bdeg	1,596±22	1,599±23	1,595±26	319±12	321±10	316±20	2,212±7	2,210±11	2,214±9	699±7	699±13	697±4
bnde	1,548±27	1,543±24	1,551±28	316±21	312±8	320±17	2,137±5	2,135±10	2,133±10	700±7	696±4	700±9

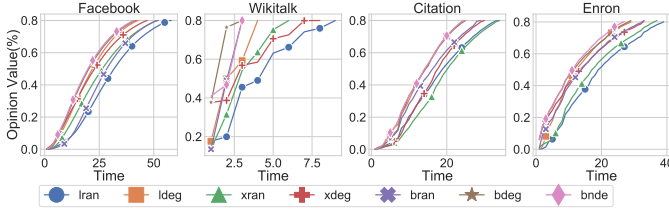


Fig. 4: Exp. 3. The average opinion values with varying time.

where agents are monitored and influenced through access units. We seek a way to impose control to a decentralized network by selecting access units. To explore the decentralized network, we investigate prowling algorithms to explore the decentralized network and to facilitate opinion convergence. Firstly, we reveal the relationship between *domination* and *centralization* and show that domination-driven prowling algorithms hold the key to opinion convergence. Then, we explore prowling algorithms in unit-growth networks that guarantee to find dominating sets. Our theoretical finding inspires us to apply degree (or N-degree)-based prowling to reduce domination costs. Experimental results on real-world and synthetic dynamic networks verify that our algorithm achieves best domination costs. The proposed framework provides a new perspective on decentralized systems such as multi-agent systems and computational social sciences.

REFERENCES

- [1] Olle Abrahamsson, Danyo Danev, and Erik G Larsson. 2019. Opinion dynamics with random actions and a stubborn agent. In ACSSC, 1486–149.
- [2] Eric Balkanski, Nicole Immorlica, and Yaron Singer. 2017. The importance of communities for learning to influence. In NIPS. 5862–5871.
- [3] Eric Balkanski, Aviad Rubinstein, and Yaron Singer. 2017. The limitations of optimization from samples. In STOC. 1016–1027.
- [4] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439, 509–512.
- [5] Christian Borgs, Michael Brautbar, Jennifer Chayes, Sanjeev Khanna, and Brendan Lucier. 2012. The power of local information in social networks. In WINE. 406–419.
- [6] Stefan Bornholdt and Holger Ebel. 2001. World Wide Web scaling exponent from Simon’s 1955 model. *Physical Review E* 64, 3, 035104.
- [7] Mickey Brautbar and Michael Kearns. 2010. Local algorithms for finding interesting individuals in large networks. In ICS. 188–199.
- [8] Kenneth KS Chung, Liaquat Hossain, and Joseph Davis. 2005. Exploring socio-centric and egocentric approaches for social network analysis. In IKMAP. 1–8.
- [9] Peter Csermely, András London, Ling-Yun Wu, and Brian Uzzi. 2013. Structure and dynamics of core/periphery networks. *J Complex Netw* 1, 2, 93–123.
- [10] Abir De, Sourangshu Bhattacharya, and Niloy Ganguly. 2018. Shaping opinion dynamics in social networks. In AAMAS. 1336–1344.
- [11] Morris H DeGroot. 1974. Reaching a consensus. *JASA* 69, 345, 118–12.
- [12] Florian Dietrich, Samuel Martin, and Marc Jungers. 2017. Control via Leadership of Opinion Dynamics with State and Time-Dependent Interactions. *TAC* 63, 4, 1200–1207.
- [13] Yucheng Dong, Quanbo Zha, Hengjie Zhang, Gang Kou, Hamido Fujita, Francisco Chiclana, and Enrique Herrera-Viedma. 2018. Consensus reaching in social network group decision making: Research paradigms and challenges. *KBS* 162 (2018), 3–13.
- [14] Sean F Everton and Dan Cunningham. 2015. Dark network resilience in a hostile environment: Optimizing centralization and density. *Criminology, Crim. Just. L, Soc’y* 16, 1.
- [15] Alexander R Galloway. 2004. Protocol: How control exists after decentralization. MIT press.
- [16] Javad Ghaderi and Rayadurgam Srikant. 2014. Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica* 50, 12, 3209–3215.
- [17] Qiang He, Xingwei Wang, Bo Yi, Fubing Mao, Yuliang Cai, and Min Huang. 2019. Opinion maximization through unknown influence power in social networks under weighted voter model. *IEEE Systems Journal* 14, 2, 1874–1885.
- [18] Paul W Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. 1983. Stochastic blockmodels: First steps. *Social networks* 5, 2, 109–137.
- [19] Matthew Jackson and Brian Rogers. 2007. Meeting strangers and friends of friends: How random are social networks? *AER* 97, 3, 890–915.
- [20] Zhihong Liu, Jianfeng Ma, Yong Zeng, Li Yang, Qiping Huang, and Hongliang Wu. 2014. On the control of opinion dynamics in social networks. *Physica A: SMA*, 183–198.
- [21] James Marchant and Nathan Griffiths. 2017. Convention emergence in partially observable topologies. In AAMAS. 187–202.
- [22] Pushpendra Singh, Sameet Sreenivasan, Boleslaw K Szymanski, and Gyorgy Korniss. 2012. Accelerating consensus on coevolving networks: The effect of committed individuals. *Physical Review E* 85, 4, 046104.
- [23] Bryan Wilder, Nicole Immorlica, Eric Rice, and Milind Tambe. 2018. Maximizing Influence in an Unknown Social Network. In AAAI.
- [24] Xudong Wu, Luoyi Fu, Jingfan Meng, and Xinbing Wang. 2019. Maximizing Influence Diffusion over Evolving Social Networks. In SocialSense. 6–11.
- [25] Amulya Yadav, Hau Chan, Albert Xin Jiang, Haifeng Xu, Eric Rice, and Milind Tambe. 2016. Using Social Networks to Aid Homeless Shelters: Dynamic Influence Maximization under Uncertainty. In AAMAS, 740–748.
- [26] Honglei Zhuang, Yihan Sun, Jie Tang, Jialin Zhang, and Xiaoming Sun. 2013. Influence Maximization in Dynamic Social Networks. In ICDM. 1313–1318.