Forty-Eighth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 29 - October 1, 2010

Influence Maximization in Social Networks: An Ising-model-based Approach

Shihuan Liu*, Lei Ying*, and Srinivas Shakkottai†
*Department of Electrical and Computer Engineering, Iowa State University
Email: {liush08, leiying}@iastate.edu

†Department of Electrical and Computer Engineering, Texas A&M University
Email: sshakkot@tamu.edu

***Abstract*—The past few years have seen increasing interest in understanding social networks as a medium for community interaction. A major challenge has been to understand various fundamental properties of social networks that form the basis for the formation and propagation of opinions across such networks. The main hurdle has been the absence of plausible models that specify the correlations between different members of a social network, which could then be used for algorithm design. This paper studies an *influence maximization* problem using an Ising-model-based approach. We first validate the credibility of the ferromagnetic Ising model in predicting opinion formation in social networks using cosponsorship data from the US Senate proceedings. We then develop a greedy placement algorithm that can efficiently find an appropriate subset of network members, "bribing" whom can efficiently propagate a particular opinion in the network. We use simulations to confirm the effectiveness of the greedy placement algorithm.**

I. Introduction

The rapid and global emergence of online social networks over the past few years, and their meteoric adoption by millions of Internet users has seen renewed interest in the study of the properties of social networks as a whole. Indeed, one can envision a future society in which communication, reputation, marketing and the very molding of societal opinions transpire on online social networking platforms. A major challenge has been to understand the properties of social networks that might allow for the harnessing of this new medium to attain desirable outcomes. A basic question is related to the propagation of influence in a social network, i.e., how does a person's opinion change the opinions of other people in the network? Also, if there are two possible views on a particular issue, which is the dominant one? This is a problem with important practical applications, especially for firms or organizations who want to promote their products or philosophies. The solution could help such firms or organizations decide the most effective places to advertise their products/philosophies.

In this paper, we consider an *influence maximization* problem. Specifically, suppose there is a particular issue about which members of a social network could have two possible opinions, and assume that we can "bribe" a fixed number of members, the question is *whom should we bribe to cause the dominance of the opinion of our choice?* Problems relating to such influence maximization were first formulated in [1], and have been studied in [2]–[5]. These results, however, are established based on the independent cascade model [6] wherein a member is activated (convinced by the opinion) with a probability depending on the state of her/his neighbors. Further, the models usually assume that the members either have a particular opinion or are neutral, but not that they might be trying to propagate opposite opinions.

In this paper, we consider a social network and assume there are two competing opinions on a specific subject. We define the value of a member's opinion to be "+1" if the member is in favor of the subject and "-1" if the member is against that subject. In order to represent the correlation of opinions between friends and acquaintances, we adopt the ferromagnetic Ising model [7], which was first proposed in statistical physics. In particular, we adopt the concept that the ground state (the configuration that emerges when the temperature goes to zero) of the network is the most likely opinion formed because the ground state minimizes the magnitude of conflicts in the social network (in physical systems, the ground state is the state with minimum energy). This model was shown to be effective in identifying community structure in social networks [8]. Note that the Ising-based model differs from the independent cascade model in two aspects: *(i)* the Ising-model-based approach allows both positive and negative influences so can be used to model two completing opinions while the independent cascade model (except the one in [5]) only allows one opinion in the network; and *(ii)* the "concept" of ground state models the "self-optimizing" nature of the social network. In other words, the network tries to minimize the degree of conflict and settles down in the minimum conflict state, which does not exist in the independent cascade model.

We develop an Ising-model-based approach wherein link-weights between nodes model their degree of correlation. We assume that subsets of nodes are irretrievably fixed with a positive or negative opinion, while others' opinions are caused due to the propagation of influence of these *seed nodes*. In order to validate this model we use data from the records of bill cosponsorship in the US Senate, wherein we fix link weights between senators (nodes) based on the number of bills that they cosponsored [9]. We fix some nodes as known to be "Democrat" or "Republican", and show that ground state of the system automatically and correctly identifies the party affiliations of the remaining nodes. We then attempt

978-1-4244-8216-0/10/$26.00 ©2010 IEEE

to answer the general question of how to identify the most appropriate nodes to "bribe" in order to cause a particular opinion to dominate. Thus, we are given a budget of nodes whose opinions we can fix, and must decide which nodes should be chosen. We develop a greedy algorithm to choose nodes, and show that it outperforms an algorithm that makes choices solely based on the degree of nodes.

The paper is organized as follows. In Section II, we introduce the Ising model based on a Markovian random field, and then describe the connection between the ground state of the Ising model and the min-cut of the corresponding graph. This connection allows us to compute the ground state in an efficient way. We then validate the Ising model and the ground state concept on a Senate network in Section III. We propose a greedy placement algorithm in Section IV that finds m positive seeds that can maximize the number of positive nodes at the ground state. We then present simulation results in Section V to compare the performance of the greedy algorithm with a degree-based placement algorithm. We conclude in Section VI.

II. Problem description and Background

We consider a social network represented by an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ is the set of nodes, and $\mathcal{L}$ is the set of links. Denote by $\mathbf{W} \in R_+^{N \times N}$ the link weight matrix, where $W_{ij} \geq 0$ is the weight of link l_{ij}. The value of W_{ij} represents the strength of the social connection between node i and node j. Larger values of W_{ij}, represent stronger social connections between node i and node j.

In this paper, we study opinion formation in a social network and propose a placement algorithm that will be defined rigorously in Section IV, to maximize influence of a desired opinion in the social network. We define O_i to be the opinion of node i on a specific subject, and $O_i = +1$ if the node is in favor of the subject and $O_i = -1$ otherwise. We denote by $\mathbf{O} = [O_1, \ldots, O_n]$ and $\mathbf{o}$ to be a realization of $\mathbf{O}$. We assume that an opinion of node i is a function of the opinions of the neighbors of node i, i.e.,

$$O_i = F_i \left(\{O_j\}_{j:ij \in \mathcal{L}} \right).$$

Further, we assume that the network contains nodes with pre-determined opinions, whose opinions are not affected by their neighbors' opinions. We name a node with pre-determined positive opinion, i.e., $O_i = +1$, as a *positive seed*, and a node with pre-determined negative opinion, i.e., $O_i = -1$, as a *negative seed*. The subset of positive seeds is denoted by Ψ^+ and the subset of negative seeds is denoted by Ψ^-. Now, given Ψ^+, Ψ^- and the functions F_i, we can compute the following two quantities:

$$N^+(\Psi^+, \Psi^-) = \mathbf{E}\left[\sum_{i \in \mathcal{N}} 1_{O_i = +1} \middle| \Psi^+, \Psi^- \right]$$
$$N^-(\Psi^+, \Psi^-) = \mathbf{E}\left[\sum_{i \in \mathcal{N}} 1_{O_i = -1} \middle| \Psi^+, \Psi^- \right],$$

i.e., N^+ is the expected number of $+1$ in the network, and N^- is the expected number of -1s in the networks, and the randomness arises due to probabilistic correlations between opinions of neighbors.

Now, assume that m extra positive seeds can be placed in the network. For example, a company is willing to pay m extra users on the online social networks to express a positive opinion. The question is *where should these extra positive seeds be placed?* Mathematically, the problem can be written as

$$\max_{\tilde{\Psi}^+ : |\tilde{\Psi}^+| = m} N^+(\tilde{\Psi}^+, \Psi^+, \Psi^-) = \mathbf{E}\left[\sum_{i \in \mathcal{N}} 1_{O_i = +1} \middle| \tilde{\Psi}^+, \Psi^+, \Psi^- \right]. \tag{1}$$

Note that to solve the placement problem (1), we first need to define the function F_i that defines the relationship between the opinions of neighboring nodes. In this paper, we adopt a Markovian Random Field Model, the Ising model, to represent such relationships. We will validate the model using Senate voting data in Section III and then propose a placement algorithm for influence maximizing based on the Ising model.

A. Markovian Random Field and Ising Model

The probability that the opinion formed $\mathbf{o}$ is assumed to be

$$\Pr(\mathbf{o}) = \frac{1}{Z} \exp\left(\sum_{ij \in \mathcal{L}} \frac{W_{ij}}{T} o_i o_j \right),$$

where W_{ij} is a parameter indicating the correlation between i and j, T is a parameter that indicates the time remaining for a decision to be made; when $T = 0$ opinions get fixed, and $Z = \sum_{\mathbf{o}} \exp(\sum_{ij \in \mathcal{L}} \frac{W_{ij}}{T} o_i o_j)$ is the normalizing factor. Therefore,

$$\Pr\left(o_i | \mathbf{o} \backslash \{o_i\}\right) = \frac{\exp\left(\sum_{j:ij \in \mathcal{L}} \frac{W_{ij} o_i o_j}{T} \right)}{\exp\left(\sum_{j:ij \in \mathcal{L}} \frac{W_{ij} o_j}{T} \right) + \exp\left(- \sum_{j:ij \in \mathcal{L}} \frac{W_{ij} o_j}{T} \right)}.$$

So this graphical model is a Markovian random field (MRF). This graphic model in exponential form is the *Ising model* from statistical physics [7]. In statistical physics, the ferromagnetic Ising model represents atoms that form a network, and every atom i is associated with a spin variable $O_i = \pm 1$. The energy of the Ising model (the Hamiltonian) is defined to be

$$H(\mathbf{O}) = -\frac{1}{2} \sum_{ij \in \mathcal{L}} W_{ij} O_i O_j.$$

The *ground state* of the Ising model is defined to be the lowest energy configuration, i.e., a configuration $\mathbf{o}_g$ such that

$$\mathbf{o}_g \in \arg\min_{\mathbf{o}} H(\mathbf{o}), \tag{2}$$

which in statistical physics, is the configuration at zero temperature.

From the perspective of MRF, the ground state configuration is the configuration that emerges with probability one when $T \to 0$ (assuming that the ground state is unique). Considering the question of identification of "Democrats" versus "Republicans" in Section III, value T can be viewed as the remaining time to decide party affiliations given one's social network. When T is large, the final decision of a member is probabilistic, but he/she has to make a decision when $T = 0$. If two neighbors i and j have different party affiliations, W_{ij} can be viewed as the amount of energy created by the conflict between i and j.

B. Ground State and Min-Cut

The Ising model can be used to understand social networks because of the following two reasons: *(i)* The Ising model defines a Markovian random field, which is consistent with our assumption that in social networks, the opinion of a node is a function of the neighboring nodes, and *(ii)* In physical systems, the ground state is the state with minimum energy. In a social network, the ground state can be viewed as the opinion formation that minimizes the magnitude of conflicts.

Given an opinion configuration $\mathbf{o}$,

$$H(\mathbf{o}) = \left(-\frac{1}{2} \sum_{ij \in \mathcal{L}} W_{ij} \right) + \sum_{ij : o_i \neq o_j} W_{ij}.$$

Since $\left(-\frac{1}{2} \sum_{ij \in \mathcal{L}} W_{ij} \right)$ is a constant independent of $\mathbf{o}$, the optimization (2) can be written as:

$$\mathbf{o}_g \in \arg\min_{\mathbf{o}} \sum_{ij : o_i \neq o_j} W_{ij}. \quad (3)$$

Given $\mathbf{o}$, define the source set to be the set of nodes with $o_i = +1$, and the sink set to be the set of nodes with $o_i = -1$. From (3), it is not difficult to verify that the ground-state configuration $\mathbf{o}_g$ of a social network with positive and negative seeds is the minimum cut of graph $\mathcal{G}$ such that the source and sink nodes are on opposite sides of the cut.

From the discussions above, *(i)* we hypothesize that the ground state of the Ising model indicates the most likely opinion formed in the given social network, and *(ii)* the ground state can be found by computing the mincut of the graph. Therefore, the Ising model could potentially be used as a quantitative tool for predicting opinion formation in social networks, and to study the influence maximization problem (1).

III. Senator Network: An Example

The Ising model has earlier been used to identify community structure of social networks, e.g., Karate club network, and coauthor network [8]. We now validate the Ising model in the network of US Senators, and show that it can correctly predict their party affiliations.

A. Senator network construction

We construct a network amongst US Senators based on data for the year 2004 [10] as follows. The senator network consists of 100 nodes (senators). The network is constructed based on the cosponsorship of bills as defined in [9]. Every bill is sponsored by one senator and can be cosponsored by multiple other senators. For a bill l, let n_l denote the number of cosponsors of the bill. Define $a_{ijl} \in \{0, 1\}$ to be a binary function such that $a_{ijl} = 1$ if and only if bill l is sponsored by senator i and cosponsored by senator j, and $a_{ijl} = 0$ otherwise. The weight of link (i, j) is defined to be

$$W_{ij} = \sum_l \frac{a_{ijl}}{n_l} + \frac{a_{jil}}{n_l}.$$

We adopt this approach because the cosponsorship reflects social connection between two senators. The network we obtained is shown in Figure 1. In the figure, blue nodes represent Democratic (or independent) senators, red node represent Republican senators. Blue links are the links between Democratic senators, red links are links between Republican senators, and green links are links across party lines. The number of links between Democratic senators is 201, between Republican senators is 180, and between a Democratic senator and a Republican senator is 196.

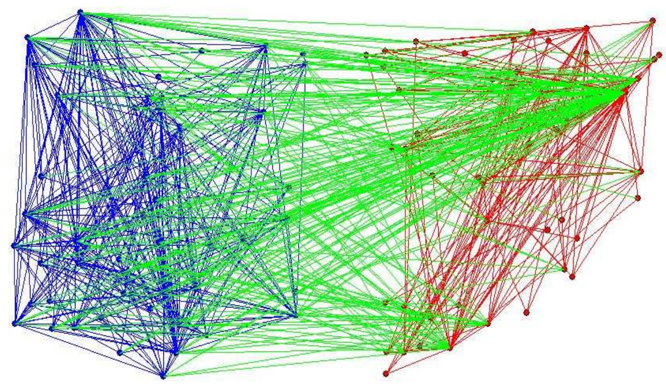


Fig. 1. The US Senator network. Links indicate cosponsorship of bills by senators, while the link colors indicate party affiliations.

B. Opinion formation in the senator network

After constructing the social network, we then applied the Ising model to predict the party affiliations of senators. We choose "+1" to represent Democratic (or independent) affiliation and a "-1" to represent Republican affiliation. We randomly selected m nodes known to be Democrats as positive seeds, and m nodes known to be Republican as negative seeds. For given Ψ^+ and Ψ^-, we found the min-cut of the graph, which specifies the ground state of the system. In other words, we found the ground state labels of all senators given m known Republicans and Democrats, and compared the labels thus obtained with the actual party affiliations of the senators. Figure 2 shows the number of prediction errors (label a Democratic senator as a Republican or vice versa) as

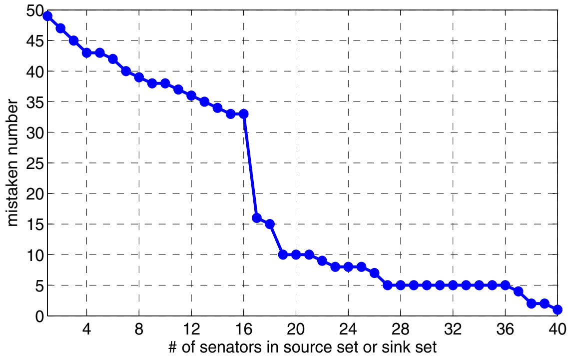


Fig. 2. The number of incorrect labels in the ground state versus the number of seeds of each type m. We see that the Ising model is a good predictor of party affiliations.

a function of m. We see that ground state of the Ising model matches real party affiliations well. Given twenty known Republicans and Democrats, the number of prediction errors is only ten (recall that we have 100 senates in the network). These results along with the earlier results of [8] indicate that the Ising model is an accurate method of representing the correlation of opinions in a social network.

IV. Maximizing Influence in Social Networks

In this section, we study efficient algorithms that solve the placement problem (1) based on the Ising model. The straightforward approach is to consider all subsets $\tilde{\Psi}^+$ such that $|\tilde{\Psi}^+| = m$ and $\tilde{\Psi}^+ \cap (\Psi^+ \cup \Psi^-) = \emptyset$, and compute the ground state. Then, we can set the $\tilde{\Psi}^+$ yielding the largest N^+. This algorithm however requires us to consider $\binom{n-|\Psi^+|-|\Psi^-|}{m}$ combinations, and is very computationally intensive.

We therefore consider a greedy algorithm that places the extra positive seeds one by one. For each positive seed, we check all possible positions and compute N^+ using the Ising model, and the select the position yielding the largest N^+. We then repeat this process until all m positive seeds are placed. This sequential placement computes the min-cut mn times instead of $\binom{n-|\Psi^+|-|\Psi^-|}{m}$ times. Further, we have the following proposition which states that it is sufficient to search the sink set instead of all nodes to place a positive seed. We define $\mathcal{N}^+(\Psi^+,\Psi^-)$ to be the source set under the min-cut given positive seeds Ψ^+ and negative seeds Ψ^-. We then have the following proposition.

Proposition 1: Assume that the graph $\mathcal{G}$ has a unique min-cut for any given Ψ^+ and Ψ^-. If $i \in \mathcal{N}^+(\Psi^+,\Psi^-)$, then

$$\mathcal{N}^+\left(\Psi^+ \cup \{i\}, \Psi^-\right) = \mathcal{N}^+\left(\Psi^+,\Psi^-\right), \tag{4}$$

and otherwise

$$\mathcal{N}^+\left(\Psi^+ \cup \{i\}, \Psi^-\right) \supseteq \mathcal{N}^+\left(\Psi^+,\Psi^-\right). \tag{5}$$

Proof: To prove this lemma, we first define $\mathcal{L}(\Psi^+,\Psi^-)$ to be the min-cut-set of links given Ψ^+ and Ψ^-, i.e.,

$$\mathcal{L}\left(\Psi^+,\Psi^-\right) = \left\{(ij) : i \in \mathcal{N}^+(\Psi^+,\Psi^-) \text{ and } j \in \mathcal{N}^-(\Psi^+,\Psi^-)\right\}.$$

We further define the weight of the min-cut as

$$W\left(\mathcal{L}\left(\Psi^+,\Psi^-\right)\right) = \sum_{ij:ij\in\mathcal{L}(\Psi^+,\Psi^-)} W_{ij}.$$

We first prove (4). We note that if $i \in \mathcal{N}^+\left(\Psi^+,\Psi^-\right)$, then $\mathcal{L}(\Psi^+,\Psi^-)$ remains a cut-set for given $\Psi^+ \cup \{i\}$ and Ψ^-, which implies that

$$W\left(\mathcal{L}\left(\Psi^+,\Psi^-\right)\right) \geq W\left(\mathcal{L}\left(\Psi^+ \cup \{i\},\Psi^-\right)\right).$$

On the other hand, $\mathcal{L}(\Psi^+ \cup \{i\},\Psi^-)$ is a cut-set for given Ψ^+ and Ψ^-, which implies that

$$W\left(\mathcal{L}\left(\Psi^+,\Psi^-\right)\right) \leq W\left(\mathcal{L}\left(\Psi^+ \cup \{i\},\Psi^-\right)\right).$$

Hence, we have that

$$W\left(\mathcal{L}\left(\Psi^+,\Psi^-\right)\right) = W\left(\mathcal{L}\left(\Psi^+ \cup \{i\},\Psi^-\right)\right),$$

and equality (4) holds under the assumption that the min-cut is unique.

We next prove result (5) by contradiction. We assume that

$$\mathcal{N}^+\left(\Psi^+ \cup \{i\}, \Psi^-\right) \not\supseteq \mathcal{N}^+\left(\Psi^+,\Psi^-\right). \tag{6}$$

Then we define the following four sets:

$$\begin{aligned}
\mathcal{S}_E &= \mathcal{N}^+\left(\Psi^+ \cup \{i\}, \Psi^-\right) \bigcap \mathcal{N}^+\left(\Psi^+,\Psi^-\right) \\
\mathcal{S}_A &= \mathcal{N}^-\left(\Psi^+ \cup \{i\}, \Psi^-\right) \bigcap \mathcal{N}^+\left(\Psi^+,\Psi^-\right) \\
\mathcal{S}_B &= \mathcal{N}^+\left(\Psi^+ \cup \{i\}, \Psi^-\right) \bigcap \mathcal{N}^-\left(\Psi^+,\Psi^-\right) \\
\mathcal{S}_C &= \mathcal{N}^-\left(\Psi^+ \cup \{i\}, \Psi^-\right) \bigcap \mathcal{N}^-\left(\Psi^+,\Psi^-\right).
\end{aligned}$$

Then it is easy to see that $\mathcal{S}_E \bigcup \mathcal{S}_A \bigcup \mathcal{S}_B \bigcup \mathcal{S}_C = \mathcal{N}$. We further define the following link sets:

$$\begin{aligned}
\mathcal{L}_a &= \{(v,u) \in \mathcal{L} : v \in \mathcal{S}_E, u \in \mathcal{S}_A\} \\
\mathcal{L}_b &= \{(v,u) \in \mathcal{L} : v \in \mathcal{S}_A, u \in \mathcal{S}_C\} \\
\mathcal{L}_c &= \{(v,u) \in \mathcal{L} : v \in \mathcal{S}_E, u \in \mathcal{S}_B\} \\
\mathcal{L}_d &= \{(v,u) \in \mathcal{L} : v \in \mathcal{S}_B, u \in \mathcal{S}_C\} \\
\mathcal{L}_e &= \{(v,u) \in \mathcal{L} : v \in \mathcal{S}_E, u \in \mathcal{S}_C\} \\
\mathcal{L}_f &= \{(v,u) \in \mathcal{L} : v \in \mathcal{S}_A, u \in \mathcal{S}_B\}
\end{aligned}$$

Under the assumption (6), we have that $\mathcal{S}_A \neq \emptyset$, which implies that $\mathcal{L}_a \neq \emptyset$ and $\mathcal{L}_b \neq \emptyset$. According to the definitions of $\mathcal{L}.$, we know

$$\begin{aligned}
\mathcal{N}^+\left(\Psi^+,\Psi^-\right) &= \mathcal{S}_E \cup \mathcal{S}_A \\
\mathcal{N}^-\left(\Psi^+,\Psi^-\right) &= \mathcal{S}_B \cup \mathcal{S}_C.
\end{aligned}$$

Therefore, the weight of the min-cut given Ψ^+ and Ψ^- satisfies

$$W\left(\mathcal{L}\left(\Psi^+,\Psi^-\right)\right) = W(\mathcal{L}_b) + W(\mathcal{L}_e) + W(\mathcal{L}_c) + W(\mathcal{L}_f).$$

Furthermore, it is easy to verify that $\mathcal{L}_a \bigcup \mathcal{L}_e \bigcup \mathcal{L}_c$ is a cut-set given Ψ^+ and Ψ^-. Under the assumption that the graph has a unique mini-cut, we then conclude that

$$\begin{aligned} & W(\mathcal{L}_b) + W(\mathcal{L}_e) + W(\mathcal{L}_c) + W(\mathcal{L}_f) \\ < \quad & W(\mathcal{L}_a) + W(\mathcal{L}_e) + W(\mathcal{L}_c) \end{aligned}$$

which implies that

$$W(\mathcal{L}_b) + W(\mathcal{L}_f) < W(\mathcal{L}_a).$$

and

$$W(\mathcal{L}_b) < W(\mathcal{L}_a). \tag{7}$$

Next, according to the definitions, we have

$$\begin{aligned} \mathcal{N}^+\left(\Psi^+ \cup \{i\}, \Psi^-\right) &= \mathcal{S}_E \cup \mathcal{S}_B \\ \mathcal{N}^-\left(\Psi^+ \cup \{i\}, \Psi^-\right) &= \mathcal{S}_A \cup \mathcal{S}_C \end{aligned}$$

Therefore, the min-cut given $\Psi^+ \cup \{i\}$ and Ψ^- satisfies

$$W\left(\mathcal{L}\left(\Psi^+ \cup \{i\}, \Psi^-\right)\right) = W(\mathcal{L}_e) + W(\mathcal{L}_d) + W(\mathcal{L}_a) + W(\mathcal{L}_f).$$

Further, $\mathcal{L}_e \cup \mathcal{L}_d \cup \mathcal{L}_b$ is a cut-set given $\Psi^+ \cup \{i\}$ and Ψ^-. Under the assumption that the min-cut is unique, we have that

$$\begin{aligned} & W(\mathcal{L}_e) + W(\mathcal{L}_d) + W(\mathcal{L}_a) + W(\mathcal{L}_f) \\ < \quad & W(\mathcal{L}_e) + W(\mathcal{L}_d) + W(\mathcal{L}_b) \end{aligned}$$

which implies that

$$W(\mathcal{L}_a) + W(\mathcal{L}_f) < W(\mathcal{L}_b).$$

and

$$W(\mathcal{L}_a) < W(\mathcal{L}_b)., \tag{8}$$

which contradicts (7). Therefore, the assumption (6) does not hold, which leads to result (5). ■

From Proposition 1, we have the following two observations:

- **Observation 1:** To place one (and only one) extra positive seed given Ψ^+ and Ψ^-, we only need to check the nodes in $\mathcal{N}^-\left(\Psi^+, \Psi^-\right) \setminus \Psi^-$.
- **Observation 2:** If a node is in source set in the ground state of the graph given Ψ^+ and Ψ^-, then the node remains in the source set when extra positive seeds are placed.

From the observations above, we can see the complexity of the sequential placement algorithm can be further reduced. In the sequential placement algorithm, suppose Ψ_k^+ is the set of positive seeds after k extra positive seeds are placed. We can combine all nodes in $\mathcal{N}^+\left(\Psi_k^+, \Psi^-\right)$ as a super positive seed since these nodes remain to be positive after the extra positive seed is placed (according to observation 2). Then we select nodes in $\mathcal{N}^-\left(\Psi_k^+, \Psi^-\right) \setminus \Psi^-$ one by one, and compute the corresponding ground state. Node j^* is selected to be the extra positive seed if

$$j^* \in \arg \max_{j: j \in \mathcal{N}^-\left(\Psi_k^+, \Psi^-\right) \setminus \Psi^-} \mathcal{N}^+\left(\Psi_k^+ \bigcup \{j\}, \Psi^-\right).$$

Therefore, we not only reduce the set of candidate nodes but also reduce the size of the network. The algorithm is formally defined as follows.

Algorithm 1 Greedy Placement Algorithm

$\Psi_0^+ = \Psi^+$
for $j = 1$ to m **do**
  Combine all nodes in $\mathcal{N}^+\left(\Psi_{j-1}^+, \Psi^-\right)$ as a super node, named node τ_{j-1}, and form a new graph $\mathcal{G}_{j-1}$;
  Find a node

$$i^* = \arg \max_{i \notin \{\tau_{j-1}\} \cup \Psi^-} \mathcal{N}^+(\{\tau_{j-1}\} \cup \{i\}, \Psi^-);$$

  Let $\Psi_j^+ = \Psi_{j-1}^+ \bigcup \{i^*\}$;
end for
return $\Psi_m^+ \setminus \Psi^+$.

V. Simulation

We now use simulations to evaluate the performance of the greedy algorithm. In social network, basically, nodes with large degrees can be considered as "influential nodes", since they have more connections with other nodes and as a result have more influence on the opinion configuration of the network. Therefore, a natural approach to our problem is to choose those large-degree nodes as extra positive seeds. Specifically, we can sort all free nodes (nodes that are not seeds) in decreasing order of their degrees, and choose the first m nodes as extra positive seeds. We call this approach the *Degree-based Placement Algorithm*. In contrast to Degree-based placement, the Greedy algorithm that we developed in this paper implicitly takes multi-hop influence propagation into account while placing extra positive seeds. We expect, therefore, that it would perform better at the goal of maximizing the number of positive opinions.

A. Simulation settings

Previous research has shown that most social networks are small-world networks [11]. Hence, we construct a small world network to imitate a social network based on the method in [12]. Specifically, we first generate a 10×10 grid network. Then, we add 1 long link to each node according to [12], where $r = 1$. By doing this we obtain a small world network. Note that the network is undirected. We randomly place 20 negative seeds, and 10 positive seeds on this network. Now, suppose that a firm is willing to place at most 10 extra positive seeds on this network. The objective of the firm is to maximize the expected number of $+1$ in the ground state of the network.

B. Greedy placement versus Degree-based placement

In order to compare the performance of two algorithms, we vary the number of extra positive seeds from 0 to 10, and observe the number of $+1$ in the ground state of the network under two algorithms. The results are shown in figure 3. From the graph we see that under the Greedy Placement Algorithm, when we place 10 extra seeds, the number of $+1$ in the ground state almost reaches the maximum value (at most 80). However, under the Degree-based algorithm, the number of $+1$ is only 33 when we place 10 extra positive seeds. Thus, the Greedy Placement Algorithm outperforms the Degree-based algorithm significantly.

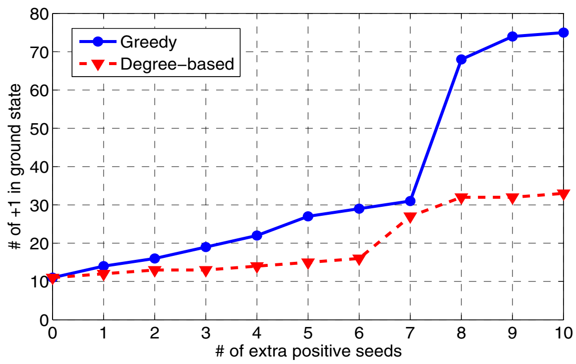


Fig. 3. Comparison between two algorithms: We observe that the Greedy Algorithm is more successful than the Degree-based Algorithm in maximizing the desired opinion.

In particular, we present the situation where we can place 10 extra positive seeds. The seed distributions under two algorithms are shown in Figure 4(a) and 4(b). In these two figures, red circles are positive seeds, blue triangles are negative seeds, and black squares are non-seed nodes. The ground state configuration in this setting are shown in Figure 5(a) and 5(b). In the figure, red circles represent $+1$ nodes, and blue triangles represent -1 nodes. Black links represent the links between two $+1$ nodes or -1 nodes, and green links represent the links between a $+1$ and a -1 node (i.e. the cut of the network). From these two graphs, it is easy to see that the Greedy Algorithm gives better results than the Degree-based Algorithm.

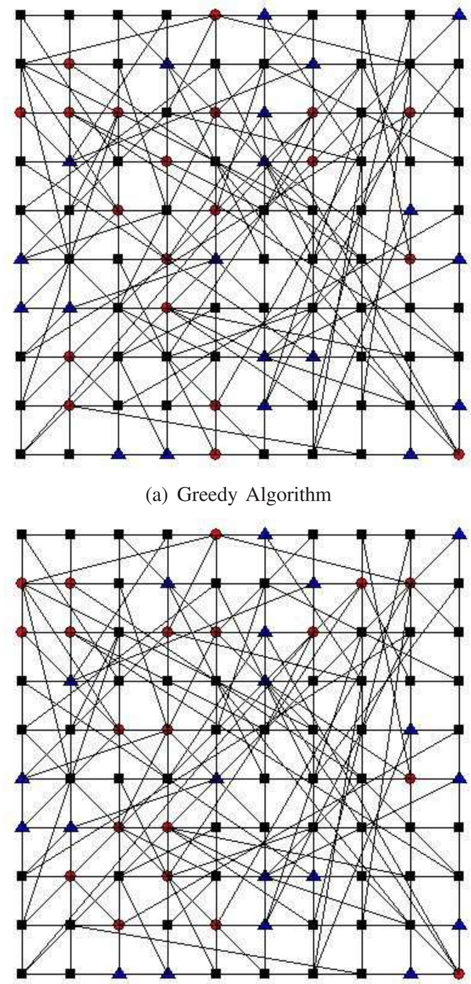
C. Greedy placement versus exhaustive search

As we know, exhaustive search is a way to find the optimal placement solution. However, due to its high computational complexity, the exhaustive search algorithm is not a practical algorithm. In this part of simulation, we use exhaustive search to find the optimal solution, and compare the result with our Greedy Algorithm. Here we assume that we can place at most 3 extra positive seeds, otherwise the computational complexity is unacceptably high.

The results are shown in Figure 6. We see from the figure that the Greedy Algorithm performs slightly worse than the optimal solution. Specifically, when we are allowed 3 extra positive nodes, the $+1$ in the ground state under the Greedy Algorithm is 19, while the optimal number is 25.

VI. Conclusion

In this paper, we studied the influence maximization in social networks. We adopted the Ising model and validated the model using it to predict party affiliations in the US Senate network using bill cosponsorship data. We then developed a greedy placement algorithm that can efficiently find that subset of users, "bribing" which can efficiently propagate the desired opinion in the network. Our simulations confirmed the effectiveness of the greedy placement algorithm over one that purely makes use of node-degree information.



(a) Greedy Algorithm

(b) Degree-based Algorithm

Fig. 4. Seed distribution under the Greedy and Degree-based Algorithms

References

[1] P. Domingos and M. Richardson, "Mining the network value of customers," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 2001, pp. 57–66.

[2] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence in a social network," in *KDD*, Washington DC, August 2003.

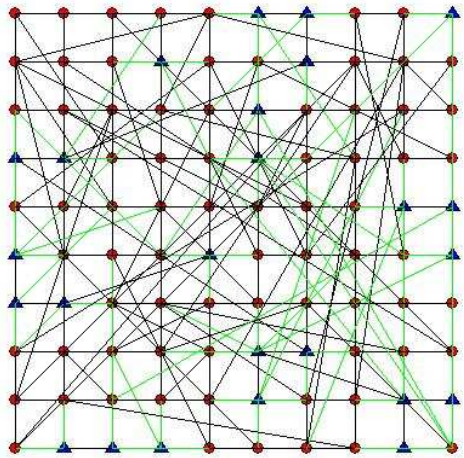
[3] D. Kempe, J. Kleinberg, and É. Tardos, "Influential nodes in a diffusion model for social networks," *Automata, Languages and Programming*, pp. 1127–1138, 2005.

[4] E. Mossel and S. Roch, "On the submodularity of influence in social networks," in *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, 2007, p. 134.

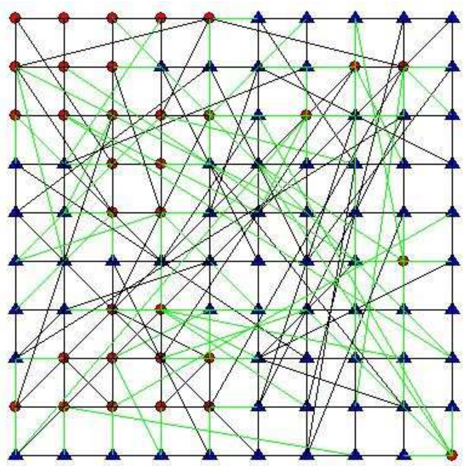
[5] S. Bharathi, D. Kempe, and M. Salek, "Competitive influence maximization in social networks," *Internet and Network Economics*, pp. 306–311.

[6] J. Goldenberg, B. Libai, and E. Muller, "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing Letters*, vol. 12, no. 3, pp. 211–223, 2001.

[7] E. Ising, "Beitrag zur theorie des ferromagnetismus," *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 31, no. 1, pp. 253–258, 1925.



(a) Greedy Algorithm



(b) Degree-based Algorithm

Fig. 5. The ground state configuration under the Greedy and Degree-based Algorithms

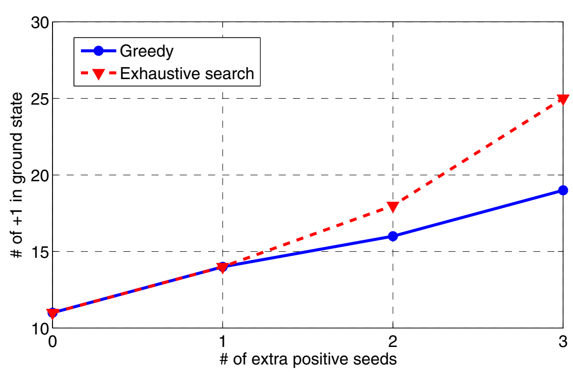


Fig. 6. Comparison between the Greedy Algorithm and exhaustive search

[8] S. W. Son, H. Jeong, and J. D. Noh, "Random field Ising model and community structure in complex networks," *The European Physical Journal B*, vol. 50, p. 431, 2006.

[9] J. Fowler, "Legislative cosponsorship networks in the US House and Senate," *Social Networks*, vol. 28, no. 4, pp. 454–465, 2006.

[10] Website of Senate, http://www.senate.gov/legislative/LIS/roll_call_lists/vote_menu_108_2.htm.

[11] M. Kochen, I. de Sola Pool, S. Milgram, and T. Newcomb, *The small world*. Ablex, Norwood, NJ, 1989.

[12] J. Kleinberg, "The small-world phenomenon: an algorithm perspective," in *Proceedings of the thirty-second annual ACM symposium on Theory of computing*. ACM, 2000, p. 170.