On New Characterizations of Social Influence in Social Networks

Makan Fardad, Fu Lin, Xi Zhang, and Mihailo R. Jovanović

Abstract— We propose new characterizations of social influence, which quantify both the transient and the steady-state propagation of beliefs across society. These characterizations are used to optimally choose a desired number of agents in a social network to serve as social leaders with maximal social impact. We then consider a framework for optimally creating new social links subject to resource constraints, in order to improve the influence of designated agents or social leaders. We show that the formulated optimization problems are convex with respect to the individual elements of the optimization variables. This motivates the use of the coordinate descent method, a simple but efficient algorithm well-suited to large-scale optimization problems. Finally, using demonstrative examples, we compare the ability of our proposed characterizations of social influence in identifying the most influential agents with that of other measures of influence developed in the social networks literature.

Index Terms— Betweenness centrality, consensus, convex relaxation, coordinate descent, leader selection, optimization, social influence, social networks, sparsity, stochastic matrices.

I. INTRODUCTION

Influencing public opinion, in order to enforce a desired consensus in the beliefs of individuals, is of paramount importance in social networks. A natural question then is how to find the most influential group of agents that can most effectively sway public opinion. To answer this question, one first needs to develop a measure of social influence. The present work is motivated by the papers [1]–[4], which consider the problem of quantifying the influence of agents in a social network. Similar problems have also been investigated in [5], [6].

In this work we use techniques from systems theory to develop alternative characterizations of social influence, and employ them to formulate pertinent social network design problems. To this end, we assume that the opinion of each individual can be represented by a real scalar. Such a description has been noted to be a good approximation of individuals' opinions [1], [4]; for example, it is argued in [1] that an individual's opinions on a wide range of social and political issues can be characterized by how liberal or conservative he/she is.

We describe the discrete-time evolution of beliefs using (row) stochastic matrices, such that at every time step each agent updates its belief by taking a weighted average of the beliefs of those agents it socially interacts with. This is a widely-used belief evolution scheme, often referred to as the DeGroot model [7]. We point out that although the DeGroot model is simpler than those employed in some recent work on social networks [8]–[10], it allows for more constructive results by permitting the definition of performance measures and the formulation of optimization problems for network design.

We exploit the property that social networks are positive systems [11], [12] to introduce multiple new characterizations of social influence that capture the efficiency of a given set of nodes in propagating its opinion through society. We then use these characterizations to formulate the problem of optimally choosing an *a priori* determined number of agents to serve as social leaders. Furthermore, given a set of nodes, we address the problem of how to optimally increase their social influence through the creation of new social interactions, and the strengthening of existing social interactions, subject to resource constraints.

Motivation

Our interest in developing alternative characterizations of social influence is in part motivated by the observation that some common definitions of social influence [1], [2], [4] can render a poor characterization of an agent's true influence. We next elaborate on these observations.

Consider a social network with n agents described by the evolution $x(k + 1) = T x(k)$, where T is a matrix with nonnegative entries whose row entries sum to one for every row. This describes the averaging mechanism by which agents update their beliefs over time. The papers [1], [4] characterize the social influence of agents in terms of the entries of the left eigenvector μ of T corresponding to the eigenvalue one, $\mu^T T = \mu^T$, where the *relative* magnitudes of the (positive) entries of μ characterize the relative influence of agents on public opinion. To understand the rationale behind such a characterization, we note that using the properties of the matrix T it can be shown that as $k \to \infty$

$$x(k) \to \frac{1}{\mu^T \mathbb{1}} \left(\mu^T x(0) \right) \mathbb{1},$$

with $\mathbb{1}$ denoting the column vector of all ones. The interpretation of this result is that all agents converge to the same opinion $(\mu^T x(0))/(\mu^T \mathbb{1})$, and that the amount of influence of the ith agent on the consensus opinion is determined by the value of the ith entry μ_i of μ.

The shortcoming of using such an eigenvalue analysis to measure social influence is that it focuses only on the steady-state condition of the social network. In contrast to the eigenvalue characterization, in this paper we propose new measures of social influence that additionally account for transient behavior and the speed of convergence to the consensus opinion. We next use a simple example to demonstrate the difference between these alternative measures of social influence and the utility of our results.

M. Fardad and X. Zhang are with the Department of Electrical Engineering and Computer Science, Syracuse University, NY 13244. F. Lin and M. R. Jovanović are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455. E-mails: makan@syr.edu, fu@umn.edu, xzhang49@syr.edu, mihailo@umn.edu.
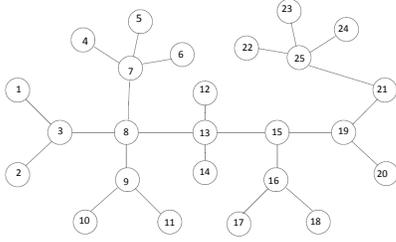
Fig. 1: The social network example of [13].

Consider the graph of a social network given in Fig. 1. We assume that the matrix T is such that at every time instant every agent updates its belief by taking the average of its own belief with that of its neighbors'. It can be shown [1] that for such a system μ_i is proportional to the number of neighbors of node i (where it is assumed that every node is also a neighbor to itself). For network in Fig. 1, this implies that agents $7, 8, 13, 25$ all have the same (and the largest) social influence as they all have the same (and the largest) number of neighbors.

On the other hand, using the measures of social influence that will be introduced in Section II, it can be shown that

influence of $13 >$ influence of $8 >$ influence of $7 >$ influence of 25.

Note that for this example, the above results agree with our intuition of which agents are more influential: Based on their location, one would expect agent 13 to be more influential than agent 25, as it takes much longer for the opinion of agent 25 to propagate through the network compared to that of agent 13.

Another measure of social influence proposed in the literature is *betweenness centrality* [2]. The betweenness centrality of node i is related to the fraction of shortest paths, between pairs of nodes in the network, that pass through node i. Betweenness centrality is more sophisticated than node degree, and counts the number paths that will become longer if a particular node is removed.

However, betweenness centrality too has shortcomings for measuring social influence. First, betweenness centrality is only defined for single nodes and it is not immediate how it should be used to measure the social influence of a group of nodes. Furthermore, betweenness centrality depends only on the interconnection topology of the graph and disregards edge directions and edge weights. It is therefore not possible to distinguish between nodes that strongly influence others versus those that are easily influenced by others. Finally, betweenness centrality ignores the temporal dynamics of belief propagation through the network.

II. New Characterizations of Social Influence

Consider a social network composed of n agents whose beliefs evolve according to [7]

$$x(k + 1) = T x(k), \qquad (1)$$

where $x(k)$ is a column n-vector composed of nonnegative values that represent the beliefs of the agents with regards to a particular social issue at time k. T is a (row) stochastic

matrix with nonnegative entries that satisfies $T \mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ is the column vector of all ones. The justification for T being a stochastic matrix is as follows: At every time step, each agent updates its belief by taking a weighted average (equivalently, a convex combination) of the beliefs of those agents it socially interacts with. This implies that every row of T is composed of nonnegative entries that sum to one. Note that T is not necessarily a symmetric matrix, which means that the weight agent i puts on the beliefs of agent j is not necessarily equal to the weight agent j puts on the beliefs of agent i.

In what follows, we develop a framework for quantifying the social influence of a designated agent, or a group of agents, in a social network. To this end, we assume that these agents always hold the belief zero and refuse to update it. We will henceforth refer to these as 'forceful' agents (FAs), and to the remaining nodes as 'regular' agents (RAs). By communicating with the rest of the network, the FAs propagate their opinion through society. If the network is strongly connected, in the sense that any given node is affected by the opinion of some FA after a sufficient number of time steps, then it can be shown that eventually the beliefs of all agents in the network will converge to zero.

Let ϕ be an n-vector whose ith element is equal to one if agent i is an FA and zero otherwise, and let

$$D_\phi = \mathrm{diag}(\phi).$$

The refusal of the FAs to update their beliefs means that the evolution of the entire social network is described by

$$x(k + 1) = (I - D_\phi) T x(k).$$

After a reordering and partitioning of the states, it can be shown that the above equation becomes

$$\begin{bmatrix} \breve{x}(k+1) \\ \tilde{x}(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ V_\phi & T_\phi \end{bmatrix} \begin{bmatrix} \breve{x}(k) \\ \tilde{x}(k) \end{bmatrix},$$

where $\breve{x}$ is a column n_ϕ-vector corresponding to the beliefs of the FAs, and $\tilde{x}$ is a column $(n - n_\phi)$-vector corresponding to the beliefs of the RAs. T_ϕ is the $(n - n_\phi) \times (n - n_\phi)$ principal submatrix of T obtained by eliminating all rows and columns corresponding to the FAs, and V_ϕ is the $(n - n_\phi) \times n_\phi$ submatrix of T obtained by eliminating all rows corresponding to the FAs and all columns corresponding to the RAs. Finally, since the beliefs of the FAs are always equal to zero, the time evolution of the RAs is described by

$$\tilde{x}(k + 1) = T_\phi \tilde{x}(k). \qquad (2)$$

We now use this framework to characterize the influence of the FAs. The basic idea is that, assuming there are nonzero initial beliefs of the RAs propagating across the social network, the faster the FAs can bring the beliefs of the network back to zero the more influential they are. Let e_i denote the ith basis vector in $\mathbb{R}^{n - n_\phi}$, and let $\tilde{x}(0) = e_i$. This can be interpreted as a social network in which all the RAs have zero initial belief except for the ith RA, whose belief is equal to one. From (2) it follows that the propagation of this initial belief through society is described by

$$\tilde{x}(k) = T_\phi^k e_i.$$

Since all entries of T_ϕ are nonnegative, $\tilde{x}(k)$ has *nonnegative* entries for all k. Furthermore, since T_ϕ is a principal submatrix of the stochastic matrix T associated with a strongly connected network and therefore its eigenvalues reside in the open unit disk [14, Theorem 11.4], we have $T_\phi^k e_i \to 0$ as $k \to \infty$ and it can be shown that $\sum_{k=0}^{\infty} T_\phi^k e_i =: y_i$ converges and

$$y_i = (I - T_\phi)^{-1} e_i.$$

Note that the entries of y_i show the *cumulative effect* (in time) of the opinion of agent i on the entire network. Therefore the total (in both time and space) effect of the initial condition e_i on society is equal to $\mathbb{1}^T y_i$. Clearly, the FAs, who maintain the belief zero at all times, are responsible for the decay in $T_\phi^k e_i$ as k increases; the more influential the FAs are the faster the decay will be, and also the smaller the cumulative effect of e_i on society.

Lastly, letting i range over all RAs in the network, we define

$$J_1(T_\phi) = \sum_{i=1}^{n-n_\phi} \mathbb{1}^T y_i = \sum_{i=1}^{n-n_\phi} \mathbb{1}^T (I - T_\phi)^{-1} e_i$$
$$= \mathbb{1}^T (I - T_\phi)^{-1} \mathbb{1}. \tag{3}$$

Thus $J_1(T_\phi)$, which is equal to the sum over all entries of the matrix $(I - T_\phi)^{-1}$, captures the social influence of the FAs on the network. Indeed, the smaller $J_1(T_\phi)$ is the larger the social influence of the FAs.

Using similar reasoning, we can define $J_2(T_\phi) = \sum_{i=1}^{n-n_\phi} y_i^T y_i$ as a measure of social influence, and we have

$$J_2(T_\phi) = \sum_{i=1}^{n-n_\phi} \text{trace}(y_i y_i^T) = \text{trace}\big((I - T_\phi)^{-T}(I - T_\phi)^{-1}\big).$$

Finally, motivated by the notion of the $\mathcal{H}_2$ norm of linear dynamical systems, we can define $J_3(T_\phi) = \sum_{i=1}^{n-n_\phi} \sum_{k=0}^{\infty} (T_\phi^k e_i)^T (T_\phi^k e_i)$ as a measure of social influence, and we have

$$J_3(T_\phi) = \sum_{i=1}^{n-n_\phi} \text{trace}(e_i e_i^T + T_\phi^T T_\phi e_i e_i^T + \cdots) = \text{trace}(P),$$

where P satisfies the discrete-time algebraic Lyapunov equation $P = T_\phi^T P T_\phi + I$.

To compare the three measures J_1, J_2, J_3, we consider the optimal selection of FAs for the network shown in Fig. 1. We use an exhaustive search to find the globally optimal selection of FAs and report the results in Table I. Note that all three measures result in a similar selection of FAs for the number of agents $n_\phi = 1, \ldots, 7$.

III. FA Selection Problem

We now turn to the problem of choosing n_ϕ forceful agents that, as a group, have the maximal amount of social influence. This problem can be formulated as

$$\underset{\phi}{\text{minimize}} \quad J_1(T_\phi)$$
$$\text{subject to} \quad \mathbb{1}^T \phi = n_\phi, \quad \phi_i \in \{0,1\}, \quad i = 1, \ldots, n,$$

where the vector ϕ is the optimization variable, and $\phi_i = 1$ if agent i is an FA and $\phi_i = 0$ otherwise. This is equivalent to the optimization problem

$$\underset{\phi}{\text{minimize}} \quad f(\phi) = \mathbb{1}^T(I - (I - D_\phi)T(I - D_\phi))^{-1}\mathbb{1}$$
$$\text{subject to} \quad \mathbb{1}^T \phi = n_\phi, \quad \phi_i \in \{0,1\}, \quad i = 1, \ldots, n. \tag{4}$$

To see this, note that with a reordering and partitioning of the states we have $(I - D_\phi)T(I - D_\phi) = \begin{bmatrix} 0 & 0 \\ 0 & T_\phi \end{bmatrix}$, and thus

$$(I - (I - D_\phi)T(I - D_\phi))^{-1} = \begin{bmatrix} I_{n_\phi \times n_\phi} & 0_{n_\phi \times p} \\ 0_{p \times n_\phi} & (I_{p \times p} - T_\phi)^{-1} \end{bmatrix},$$

where $p = n - n_\phi$. It follows that

$$f(\phi) = J_1 + n_\phi.$$

Problem (4) is nonconvex, in part due to the Boolean constraints $\phi_i \in \{0,1\}$. Instead of solving (4), we relax the Boolean constraint set to its convex hull by taking $\phi_i \in [0,1]$, and consider the optimization problem

$$\underset{\phi}{\text{minimize}} \quad f(\phi) + \gamma \mathbb{1}^T \phi$$
$$\text{subject to} \quad \phi_i \in [0,1], \quad i = 1, \ldots, n, \tag{5}$$

where $\mathbb{1}^T \phi = \sum_i \phi_i$ is the ℓ_1 norm of the nonnegative vector ϕ, and γ is a positive scalar that determines the relative importance of the two terms in the objective function. By incorporating the ℓ_1 norm of ϕ in the objective, we encourage an optimal solution that is sparse. When $\gamma = 0$, the global solution to (5) is $\phi = \mathbb{1}$, i.e., all agents are forceful agents. As γ increases, the number of nonzero elements in the optimal ϕ decreases while the value of f increases. Since we have relaxed the Boolean constraints to $\phi_i \in [0,1]$, the solution of (5) is not integer-valued in general. We thus choose the FAs to be the nodes corresponding to the largest n_ϕ entries of ϕ.

We next examine the convexity property of the objective function f.

Proposition 1: Let ϕ be a nonzero n-vector with its entries belonging to $[0,1]$. Then the objective function $f(\phi)$ in (4) is convex with respect to each element of ϕ, that is, $f(\phi_i)$ is a convex function of ϕ_i when other elements of ϕ are held fixed.

Proof: See Appendix. ∎

Proposition 1 establishes the *elementwise convexity* of f, which is a much weaker result than the convexity of f with respect to the entire vector ϕ. Nevertheless, Proposition 1 allows the use of the coordinate descent method for problem (5), a simple but efficient algorithm which we describe in Section III-A.

A. Coordinate Descent Method

The coordinate descent method has been widely used in sparse signal processing and machine learning [15]–[18]. Due to its simple implementation, this method is often used for large-scale problems [19], [20]. Furthermore, this method has the desirable property of being applicable to, and convergent for, nonsmooth problems [21], [22].

Instead of minimizing the objective function in (5) with respect to the entire vector ϕ, the coordinate descent method holds fixed $n - 1$ variables ϕ_j for $j \neq i$ and minimizes the objective function with respect to the scalar variable ϕ_i.

TABLE I: The globally optimal selection of FAs obtained using an exhaustive search for the measures J_1, J_2, J_3, for the network shown in Fig. 1.

n_ϕ	J_1	forceful agents	J_2	forceful agents	J_3	forceful agents
1	1104.0	13	7356.0	13	108.1	13
2	334.0	8, 19	1195.1	8, 21	61.6	8, 25
3	173.5	8, 15, 25	534.4	8, 16, 25	47.8	8, 16, 25
4	129.5	7, 8, 15, 25	362.4	7, 8, 16, 25	40.1	3, 7, 16, 25
5	88.3	3, 7, 9, 15, 25	217.9	3, 7, 9, 16, 25	32.5	3, 7, 9, 16, 25
6	55.5	3, 7, 9, 13, 16, 25	117.1	3, 7, 9, 13, 16, 25	27.1	3, 7, 9, 13, 16, 25
7	34.1	3, 7, 9, 13, 16, 19, 25	65.6	3, 7, 9, 13, 16, 19, 25	23.2	3, 7, 9, 13, 16, 19, 25

Specifically, for $i = 1, \dots, n$, we solve the subproblem

$$\begin{aligned} \underset{\phi_i}{\text{minimize}} \quad & f(\phi_i) + \gamma\, \phi_i \\ \text{subject to} \quad & \phi_i \in [0, 1], \end{aligned} \qquad (6)$$

where we have abused notation by using $f(\phi_i)$ to denote $f(\phi)$ with ϕ_i being the only optimization variable. We then update the ith entry of ϕ with the obtained solution, increment i, and solve (6) again. This update sequence starting from $i = 1$ to $i = n$ is then repeated in a cyclic fashion until convergence.

The convergence of coordinate descent method to a stationary point of problem (5) is guaranteed by the smoothness and convexity of subproblem (6) in each coordinate [19], [21]. Furthermore, since each subproblem (6) is a smooth convex optimization problem with box constraints, it can be solved efficiently, for example using projected gradient methods [19]. Here, we provide the expression for the gradient of $f(\phi_i)$,

$$\frac{\partial f}{\partial \phi_i} = -\mathbb{1}^T M_\phi^{-1}(e_i e_i^T T(I - D_\phi) + (I - D_\phi)T e_i e_i^T)M_\phi^{-1}\mathbb{1},$$

where $M_\phi = I - (I - D_\phi)T(I - D_\phi)$.

B. An Example

We next examine the network shown in Fig. 1. The results obtained using the coordinate descent method of Section III-A are shown in Fig. 2. As Figs. 2a and 2b demonstrate, the number of forceful agents n_ϕ decreases and the social influence measure J_1 increases with γ. Additionally, we use an exhaustive search to find the globally optimal selection of FAs for $n_\phi \leq 7$. It can be seen in Table II that the coordinate descent method finds the global solution for $n_\phi = 6, 7$.

The *betweenness centrality* [2] of node i is defined as

$$\beta(i) = \sum_{j \neq i \neq l} \frac{\tau_{jl}(i)}{\tau_{jl}},$$

where τ_{jl} is the total number of shortest paths from node j to node l, and $\tau_{jl}(i)$ is the number of those paths that pass through node i. Betweenness centrality can be used to measure the importance of a node, as it quantifies the number of shortest paths that will become longer when a node is removed [2], [3].

To compare the selection of FAs based on betweenness centrality versus that obtained from solving (5), we first rank the nodes in Fig. 1 in descending order based on their betweenness centrality, and then choose the first n_ϕ nodes as FAs. We note that this selection is outperformed by the FAs obtained from solving problem (5) using coordinate descent.
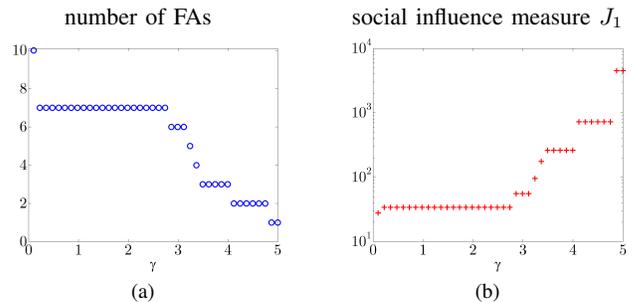
Fig. 2: The performance of the coordinate descent method for problem (5), for the network shown in Fig. 1. As γ increases, (a) the number of forceful agents n_ϕ decreases, and (b) J_1 increases.

The reason for the poor performance of the betweenness-centrality-based algorithm can be attributed to the fact that the nodes with large betweenness centrality are all clustered along the path from node 8 to node 25. In contrast, the FAs obtained from (5) spread out in the network to maximally influence the RAs.

IV. OPTIMAL INCREASE OF FAS' INFLUENCE

In this section we assume that having chosen n_ϕ nodes to serve as FAs, we now want to invest in increasing their social influence. We model this as creating new links (and, if necessary, strengthening existing links) between the FAs and the RAs.

Let us recall that in the framework considered in this work, each agent updates its belief by taking a weighted average over the beliefs of those agents it interacts with. Therefore, if an FA wants to create a link between itself and an RA it did not interact with previously (or if an FA wants to increase its influence on an RA it is already interacting with), that RA would have to 'make room' (or 'make more room') in its weighted averaging scheme for the belief of the FA.

To model this we first notice that since all FAs are identical (i.e., all have belief value equal to zero at all times), from a mathematical point of view it does not make any difference which FA establishes a link with a given RA. Therefore, we proceed by finding merely the weight of the new links between RAs and the FAs (or the added weight between RAs and FAs already interacting with them), without specifying exactly which FA each RA will be linked to. Once the optimal weights are found, one can use other metrics such as physical proximity or availability to decide how to link the FAs and RAs.

Suppose, for example, that a new social link is to be established between an RA and an FA. And suppose that

TABLE II: For the network shown in Fig. 1, the performance of the coordinate descent method of Section III-A is compared to the globally optimal selection of FAs obtained using an exhaustive search, and is also compared to the selection of FAs based on betweenness centrality.

n_ϕ	coordinate descent		exhaustive search		betweenness centrality	
	J_1	forceful agents	J_1	forceful agents	J_1	forceful agents
1	4508.0	25	1104.0	13	1326.0	8
2	724.0	$7,25$	334.0	$8,19$	752.0	$8,13$
3	261.8	$7,13,25$	173.5	$8,15,25$	400.0	$8,13,15$
4	175.8	$7,13,16,25$	129.5	$7,8,15,25$	260.0	$8,13,15,19$
5	95.6	$3,7,13,16,25$	88.3	$3,7,9,15,25$	190.0	$8,13,15,19,21$
6	55.5	$3,7,9,13,16,25$	55.5	$3,7,9,13,16,25$	146.0	$7,8,13,15,19,21$
7	34.1	$3,7,9,13,16,19,25$	34.1	$3,7,9,13,16,19,25$	102.0	$7,8,13,15,19,21,25$

the RA is to give a 10% weight in its averaging scheme to the belief of the FA. Since the RA's existing weights sum to one, it makes room for the FA's belief by scaling all its existing weights by 0.9. (If a social link already existed between an RA and an FA, the procedure stays the same in the sense that all weights, including that corresponding to the FA, are scaled down to make room for the new weight.)

The description above corresponds to the following evolution of beliefs

$$\begin{bmatrix} \breve{x}(k+1) \\ \tilde{x}(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ SV_\phi + V_\sigma & ST_\phi \end{bmatrix} \begin{bmatrix} \breve{x}(k) \\ \tilde{x}(k) \end{bmatrix},$$

where

$$\sigma = \begin{bmatrix} \sigma_1 & \cdots & \sigma_{n-n_\phi} \end{bmatrix}^T,$$

$$0 \leq \sigma_i \leq 1 \quad \text{for} \quad i = 1, \ldots, n - n_\phi,$$

σ_i = weight agent i puts on the belief of any FA,

V_σ = matrix whose only nonzero entry in its ith row is σ_i,

$$S = I - \text{diag}(\sigma).$$

Note that the exact location in which σ_i appears in the ith row of V_σ is not important to us, as we only care about the weight the ith agent puts on the belief of *some* FA (i.e., the index of the FA is not important).

Once again, since the beliefs of the FAs are always equal to zero, the public beliefs $\tilde{x}$ evolve according to

$$\tilde{x}(k+1) = ST_\phi \tilde{x}(k).$$

Note that if $\sigma_i = 1$ then the ith node effectively becomes an FA itself, since it will stubbornly maintain the belief zero at all times.

It can be demonstrated that the eigenvalues of ST_ϕ all reside in the open unit disk. This means that ST_ϕ defines a stable evolution and therefore the social influence metrics J_i, $i = 1, 2, 3$ developed in the previous section are applicable, in which we replace T_ϕ with ST_ϕ. We keep in mind that the variable in $J_i(ST_\phi)$ is now σ rather than ϕ.

We choose the objective function of our optimization problem such that it accounts for both the social influence of the FAs, which we wish to maximize, and the total weight of the new links made to the FAs, which we would like to keep small. Thus to maximize the social influence of the FAs while keeping down the resources spent by them, we solve the problem

$$\begin{aligned} \underset{\sigma}{\text{minimize}} \quad & J_1(ST_\phi) + \gamma \mathbb{1}^T \sigma \\ \text{subject to} \quad & \sigma_i \in [0,1], \quad i = 1, \ldots, n - n_\phi, \end{aligned} \quad (7)$$

where the vector of weights σ is the optimization variable, $S = I - \text{diag}(\sigma)$, and γ is a positive scalar that determines the relative importance of increasing the social influence of the FAs versus keeping down the total weight of the new social links.

A. Coordinate Descent Method with Analytical Solution

Using an argument analogous to that in the proof of Proposition 1, it can be shown that J_1 in (7) is a convex function of individual elements σ_i of the vector σ. Thus, we can apply coordinate descent method to problem (7). Furthermore, we next show that each coordinate descent step can be solved analytically.

Let $\sigma = \bar{\sigma}_i + \sigma_i e_i$, where

$$\bar{\sigma}_i = [\sigma_1, \ldots, \sigma_{i-1}, 0, \sigma_{i+1}, \ldots, \sigma_{n-n_\phi}]^T \in \mathbb{R}^{n-n_\phi}$$

and e_i is the ith basis vector of $\mathbb{R}^{n-n_\phi}$. Then, using the Sherman-Morrison formula for the inverse of a matrix with a rank-1 modification, it can be shown that

$$\mathbb{1}^T (I - ST_\phi)^{-1} \mathbb{1} + \gamma \mathbb{1}^T \sigma = \gamma \sigma_i - \frac{a_i \sigma_i}{1 + b_i \sigma_i} + c_i,$$

where

$$a_i = \mathbb{1}^T \bar{\Sigma}_i^{-1} e_i e_i^T T_\phi \bar{\Sigma}_i^{-1} \mathbb{1}, \quad b_i = e_i^T T_\phi \bar{\Sigma}_i^{-1} e_i,$$

$$c_i = \mathbb{1}^T \bar{\Sigma}_i^{-1} \mathbb{1} + \gamma \mathbb{1}^T \bar{\sigma}_i, \qquad \bar{\Sigma}_i = I - T_\phi + \text{diag}(\bar{\sigma}_i) T_\phi.$$

Thus, the minimization problem of (7) with respect to σ_i can be formulated as

$$\begin{aligned} \underset{\sigma_i}{\text{minimize}} \quad & \gamma \sigma_i - \frac{a_i \sigma_i}{1 + b_i \sigma_i} \\ \text{subject to} \quad & \sigma_i \in [0,1], \end{aligned}$$

and its solution is given by

$$\sigma_i = \begin{cases} 1, & \gamma < a_i/(1+b_i)^2 \\ \sigma_i^*, & a_i/(1+b_i)^2 \leq \gamma \leq a_i \\ 0, & \gamma > a_i \end{cases} \quad (8)$$

with $\sigma_i^* = (\sqrt{a_i} - \sqrt{\gamma})/(b_i \sqrt{\gamma})$. Since the spectral radius of the nonnegative matrix $(I - \text{diag}(\bar{\sigma}_i))T_\phi$ is less than 1, we have $\bar{\Sigma}_i^{-1} = \sum_{r=0}^{\infty} ((I - \text{diag}(\bar{\sigma}_i))T_\phi)^r$, and thus $\bar{\Sigma}_i^{-1}$ is an elementwise nonnegative matrix. Therefore a_i in (8) is a nonnegative number.

B. An example

We again consider the network shown in Fig. 1, and apply the coordinate descent method of Section IV-A to problem (7). When node 13 is selected as an FA, the tradeoff between the number of social links (between the RAs and node 13)
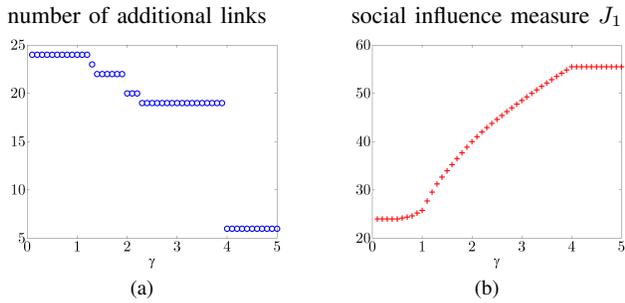
Fig. 3: The performance of the coordinate descent method for problem (7), for the network shown in Fig. 1 in which node 13 is an FA. As γ increases, (a) the number of social links between RAs and the FA decreases, and (b) J_1 increases.

and the performance J_1 is illustrated in Fig. 3. As γ increases, the number of social links decreases and J_1 increases. When γ is large ($\gamma \geq 4$), 6 new links are created between the FA at node 13 and the RAs at nodes $\{3, 7, 9, 16, 19, 25\}$. It is interesting to note that these 6 nodes together with node 13 form the globally optimal selection of FAs shown in Table II. Note that when $\gamma = 0$, the solution of the optimization problem is such that every RA communicates with some FA. For large γ ($\gamma \geq 4$), using only 6 additional social links between the RAs and the FA, the measure of social influence J_1 decreases to $J_1 = 55$ compared to $J_1 = 1104$ before additional social links were introduced.

V. CONCLUSIONS

We define several new measures of social influence and compare them to some existing ones in the literature. Using these measures, we address the problems of optimal leader selection in social networks and the optimal creation of new social interactions; also see [23], [24] for related efforts. We show that although the formulated optimization problems are nonconvex, they still possess the property of being convex with respect to individual entries of the optimization variables. This motivates the use of the coordinate descent method. We provide an example that demonstrates the effectiveness of our proposed measures of social influence and our optimization algorithms.

APPENDIX

Proof of Proposition 1

Let $z := \mathbb{1} - \phi$ where each element of the vector ϕ belongs to $[0, 1]$. Since ϕ is not identically equal to zero, then $D_z := \mathrm{diag}\,(z)$ satisfies $0 \preceq D_z \preceq I$ and $D_z \neq I$. Using Neumann series, we have $f(z) = \mathbb{1}^T (I - D_z T D_z)^{-1} \mathbb{1} = \sum_{r=0}^{\infty} \mathbb{1}^T (D_z T D_z)^r \mathbb{1}$. Therefore, to show the convexity of $f(z)$ with respect to each element $z_i \in [0, 1]$, it is sufficient to show that each entry of $W^r = (D_z T D_z)^r$ is a convex function of z_i. Let W_{pq}^r be the pqth entry of W^r with $p, q = 1, \ldots, n$. Since T is a nonnegative matrix and since z is a nonnegative vector, it follows that $W_{pq}^r(z_i)$ is a polynomial of z_i with nonnegative exponents and nonnegative coefficients. This implies that taking the second-order derivative of W_{pq}^r with respect to z_i, we again obtain a polynomial of z_i with nonnegative coefficients. Thus, $\partial^2 W_{pq}^r / \partial z_i^2 \geq 0$ for $z_i \in [0, 1]$, i.e., $W_{pq}^r(z_i)$ is a convex function over $[0, 1]$. The proof is complete by noting that the composition of a convex function $W_{pq}^r(z_i)$ with an affine function $z_i = 1 - \phi_i$ is a convex function of ϕ_i; see [25, Section 3.2.2].

REFERENCES

[1] P. M. DeMarzo, D. Vayanos, and J. Zwiebel, "Persuasion bias, social influence, and unidimensional opinions," *The Quarterly Journal of Economics*, vol. 118, no. 3, pp. 909–968, 2003.

[2] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[3] S. P. Borgatti and M. G. Everett, "A graph-theoretic perspective on centrality," *Social Networks*, vol. 28, pp. 466–484, 2006.

[4] B. Golub and M. Jackson, "Naive learning social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.

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

[6] H. Amini, M. Draief, and M. Lelarge, "Marketing in a random network," in *Network Control and Optimization (Lecture Notes in Computer Science Volume 5425)*. Springer, 2009, pp. 17–25.

[7] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, pp. 118–121, 1974.

[8] D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi, "Spread of (mis)information in social networks," *Games and Economic Behavior*, vol. 70, pp. 194–227, 2010.

[9] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1, pp. 3–49, 2010.

[10] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Mathematics of Operation Research*, 2013, to appear.

[11] T. Tanaka and C. Langbort, "KYP lemma for internally positive systems and a tractable class of distributed H-infinity control problems," in *Proceedings of the 2010 American Control Conference*, 2010.

[12] A. Rantzer, "Distributed control of positive systems," in *Proceedings of the 50th IEEE Conference on Decision and Control*, 2011, pp. 6608–6611.

[13] S. Patterson and B. Bamieh, "Leader selection for optimal network coherence," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 2692–2697.

[14] C. M. Grinstead and J. L. Snell, *Introduction to probability*. Amer Mathematical Society, 1997.

[15] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *The Annals of Applied Statistics*, vol. 1, no. 2, pp. 302–332, 2007.

[16] C.-J. Hsieh, K.-W. Chang, C.-J. Lin, S. S. Keerthi, and S. Sundararajan, "A dual coordinate descent method for large-scale linear SVM," in *Proceedings of the 25th international conference on Machine learning*, vol. 951, no. 8, 2008, pp. 408–415.

[17] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, 2009.

[18] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of statistical software*, vol. 33, no. 1, p. 1, 2010.

[19] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[20] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*. Springer, 2008.

[21] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of optimization theory and applications*, vol. 109, no. 3, pp. 475–494, 2001.

[22] P. Tseng and S. Yun, "A coordinate gradient descent method for nonsmooth separable minimization," *Mathematical Programming*, vol. 117, no. 1, pp. 387–423, 2009.

[23] M. Fardad, X. Zhang, F. Lin, and M. R. Jovanović, "On the optimal dissemination of information in social networks," in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012, pp. 2539–2544.

[24] F. Lin, M. Fardad, and M. R. Jovanović, "Algorithms for leader selection in stochastically forced consensus networks," *IEEE Transactions on Automatic Control*, 2012, conditionally accepted; also arXiv:1302.0450.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.