Optimal Strategies for Targeted Influence in Signed Networks

Basak Guler* Burak Varan* Kaya Tutuncuoglu* Mohamed Nafea*
Ahmed A. Zewail* Aylin Yener* Damien Octeau**
*Wireless Communications and Networking Laboratory
Department of Electrical Engineering
The Pennsylvania State University, University Park, PA 16802
{basak, varan, kaya, msn139, aiz103}@psu.edu yener@ee.psu.edu
** Department of Computer Science and Engineering
The Pennsylvania State University, University Park, PA 16802
octeau@cse.psu.edu

***Abstract*—Online social communities often exhibit complex relationship structures, ranging from close friends to political rivals. As a result, persons are influenced by their friends and foes differently. Network applications can benefit from accompanying these structural differences in propagation schemes. In this paper, we study the optimal influence propagation policies for networks with positive and negative relationship types. We tackle the problem of minimizing the end-to-end propagation cost of influencing a target person *in favor of* an idea by utilizing the relationship types in the underlying social graph. The propagation cost is incurred by social and physical network dynamics such as frequency of interaction, the strength of friendship and foe ties, propagation delay or the impact factor of the propagating idea. We extend this problem by incorporating the impact of message deterioration and ignorance. We demonstrate our results in both a controlled environment and the Epinions dataset. Our results show that judicious propagation schemes lead to a significant reduction in the average cost and complexity of influence propagation compared to naïve myopic algorithms.**

***Index Terms*—Socially aware physical systems, network propagation for social media, signed networks, recommender systems.**

I. Introduction

Social networks have grown into a major platform for spread of information, thanks to the proliferation of smart devices and portable computers [1]. Online communities often exhibit highly complex relationship structures, ranging from like-minded friends to ideological foes. However, conventional social network analysis often considers solely monolithic relationship types, one that treats all relations as *friendly* [2]–[5]. The need for identifying multiple relation types in networks has recently been emphasized in various works [6]–[10].

Consider an online recommendation process for an upcoming election between two candidates A and B. A recommender is suggesting one candidate to each person based on the information available about her interests and friendship structures. *Alice* has two neighbors, *Bob* and *Eve*. Bob shares the same world interpretation with Alice, whereas Eve is in complete opposition. The recommender knows that both Bob and Eve *support* B. It can then provide one of the following suggestions to Alice, "Bob supports B, do you want to choose B, too?" or "Eve supports B, do you want to choose B?". In the first case, Alice is likely to support B since Bob is an ideological ally. In the latter case, when Alice sees that Eve supports B, she will have a negative opinion about B since Eve is an ideological foe. Depending on the network interests, the recommender can make a choice between the two in order to influence Alice in the desired direction.

Influence propagation has been extensively studied in the field of social networks with monolithic relationships [4], [5], [11]–[15]. Many applications focus on identifying the users that maximize the spread of influence, or limit misinformation [16], by means of probabilistic or optimization-based methods and efficient heuristics. Despite broad interest and a vast literature in influence models, the significance of relationship types has only recently been discovered [17]. Focusing on influence diffusion for opposing ideas in signed networks, [17] has observed that taking relationship types into consideration yields notable changes in diffusion patterns. However, current literature has not addressed influencing policies, or considered key metrics such as propagation costs.

In this paper, we aim to bridge this gap, by taking a new perspective on the influence propagation problem in a network with positive and negative links as in Fig. 1. The signs serve to differentiate one's like-minded neighbors from the ones with an opposite world view. Our approach is based on the principle of *homophily* [18], namely that people are influenced by their friends and foes differently. In effect, persons tend to agree with their neighbors who are ideologically similar or share similar interests, and oppose to the opinions of their ideological foes [6]. That is, people tend to take sides *in favor of* or *against* an idea (product, candidate, opinion) based on the observations made available to them. An interesting phenomenon occurs when a neighbor with an antagonistic world view is *against* an idea. In this case, one is likely to go against the neighbor, which results in a *positive* disposition towards the original idea. This phenomenon has been widely observed in the real world, including the historical details of the European alliances before World War I [19], [20].

We posit that each social link incurs a cost of propagation, which has various social and physical interpretations such as the frequency of interaction, propagation delay, the strength of friendship/foe ties or the impact factor of the propagating idea. This allows us to demonstrate the optimal policies in terms of a policy-free measurement metric. We note that the right metric, which is often a weighted combination of multiple social and physical factors, depends on the design goal of the network. One important physical metric we address in this study is the end-to-end delay, which is essential for delivering the fastest

ASONAM 2014, August 17-20, 2014, Beijing, China

network experience to the user. Based on this discussion, we address the problem of minimizing the end-to-end propagation cost to influence a target node *positively* with a given idea. By doing so, we stipulate the target node to support the idea within minimum end-to-end cost incurred through the network. In case the cost metric is end-to-end delay, this refers to the fastest policy for influencing a target node positively. The optimal policies can be integrated with routing schemes with different design goals as required by the network provider.

We expect our findings to be useful in various social network applications in which relationship types cause a significant impact on network benefits. Our contributions are:

- We develop the first model that addresses the optimal propagation policies in signed networks. This framework is extensible in multiple ways.
- We evaluate these algorithms using both a controlled setting and real-world data from the Epinions website.
- We provide interesting insights on influence propagation in the Epinions data. We find that randomly selected sources can positively influence randomly selected destinations in over 87% of the cases.

II. Related Work

Social trust, influence, friendship relations and their impact on information flow, have been investigated in various studies [2]–[10], from connecting people with trust scores [21], to characterizing trust and distrust by signs [22], [23], or using relationships in software design to assist recommenders [24].

Associating positive and negative links with social relations dates back to balance and status theories in social psychology [25], [26]. These theories investigate the cognitive relationships between living beings and provide a graph-theoretic description of dynamics of balanced structures in organizational networks. In the context of social media, signed links represent positive and negative relationships in human interactions [22], [23]. The evolution of the signed link structures over time is analyzed in [22]. The problem of predicting positive and negative relationships in online network data is studied in [23].

Much of the previous work on influence propagation is focused on triggering common behavior in the social network, namely *influence maximization*. The problem of selecting the most influential nodes for a large fraction of the network to adopt a new product is posed in [4]. Two major propagation models, independent cascade and linear threshold, are studied in [5] with provable approximation guarantees. Many heuristics are proposed to improve the efficiency of these algorithms [11]–[14], some of which depend on probabilistic methods. Independent cascade has recently been applied to signed networks in [17] to identify the optimal seeds for short and long-term influence maximization. In particular, [17] addresses a model in which the recommender has no impact on the propagation of influence once the key users have been seeded. In contrast, our model identifies the optimal strategies for message propagation to influence a target node positively, whereas in [17], each node randomly picks one of its outgoing neighbors and adopts her opinion. Unlike [17], we utilize the relationship types to *steer* the opinion in the desired direction, and address how to optimize the network propagation costs.

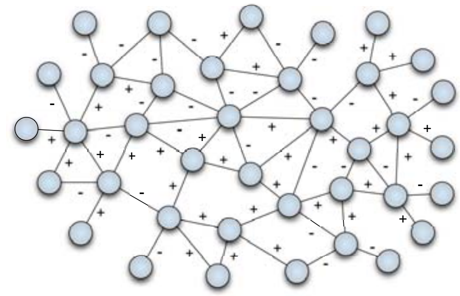

Fig. 1. Positive and negative links in a signed social network.

III. Social Network Model

We first consider a directed acyclic graph. Section VI generalizes the propagation scheme to cyclic graphs. Let $G = (V, E)$ be a graph with $|V|$ nodes representing the social network. A directed edge exists from node u to node v if $(u, v) \in E$. The coordinates of a node $u \in V$ are represented by the tuple (u_x, u_y). Throughout the paper, we represent a node by its index and its coordinates interchangeably. Each edge (u, v) is assigned a sign $s_{u,v} \in \{-1, 1\}$ to represent the relationship type between persons u and v. In effect, $s_{u,v}$ reflects the *attitude* of one person towards the other.

Initially, we activate a source node via a message from an external source, news or a promotion. This node then passes the message to one of its neighbors which results in a positive or negative influence. Propagation continues throughout the network until the message reaches the target node. Alternatively, this model can be interpreted as a recommendation network, with a recommender making suggestions to individual persons on a path, based on the preferences of their contacts. A person is likely to be positively influenced by the recommender if the previous contact is supporting an idea (candidate, product) and is a friend, whereas if the previous contact is a foe, the person is likely to oppose the idea.

We represent the cost of influence propagation between two persons by a nonnegative weight. An example of such a cost is the *propagation delay* between the two parties. The delay variable has both social and physical interpretations. From a physical standpoint, it is an important metric for assessing the QoS (quality of service) of multi-hop sensor networks, and may depend on various quantities such as the bandwidth, load, and physical distance between the travelled links. From a social perspective, it represents the strength of the actions of one person on influencing the others, either positively or negatively, with a smaller delay representing a quicker response. It may also indicate the frequency of interaction between the two persons. We provide a formal definition of the influence propagation problem in the sequel.

IV. Minimum-Cost Positive Influence Propagation

We study in this section influence propagation with minimum expected end-to-end cost. Let u_o and u_d represent the source and destination nodes, respectively. We seek the optimal path and policy to influence the target person (destination) positively with minimum total cost. The propagation cost from node u to its neighbor v is denoted by $d_{u,v} \geq 0$. Any direction with no edge is assigned an infinite cost. The sign of the relationship between node u and v is given by $s_{u,v}$. The set of

Algorithm 1 Backward Induction Dynamic Programming for Minimum-Cost Influence Propagation

1: **procedure** MINCOST(V, E, s, d, u_o, u_d) ▷ s and d are the arrays of all edge signs and costs
2: Set the boundary conditions from (4)
3: **for all** u from u_d to u_o in reverse topological order **do**
4: Compute $S(u,0)$ and $S(u,1)$ using (2)-(3)
5: Record the decisions (next hop) $\pi(u,0)$ and $\pi(u,1)$ that lead to $S(u,0)$ and $S(u,1)$
6: $P^* := [u_o]$; $u := u_o$; $parity := 0$
7: **while** $u \neq u_d$ **do**
8: $temp := \pi(u, parity)$
9: **if** $parity = s_{u,temp}$ **then** $parity := 0$ **else** $parity := 1$
10: $u := temp$
11: Append u to P^*
12: **return** the optimal path P^* and the minimum cost $S(u_o, 0)$

all possible paths from u_o to u_d is given by $\mathcal{P}$. The minimum-cost positive influence propagation problem is given as:

$$\min_{P\in\mathcal{P}} \sum_{u,v:\,(u,v)\in P} d_{u,v} \quad \text{s.t.} \prod_{u,v:\,(u,v)\in P} s_{u,v} = +1 \tag{1}$$

where the objective function stands for the total cost of path P, and the multiplicative constraint ensures that the destination is *positively* influenced by the intended idea. (1) is a dynamic program which we can solve via backward induction. Without loss of generality, we label the node indices in topological order. That is, for every edge $(u,v) \in E$, $u \leq v$. Such an ordering is feasible for any directed acyclic graph [27].

We define an optimal value function $S(u,z)$ that quantifies the minimum total cost of the optimal path from node u to the destination, in terms of a parity variable $z \in \{0,1\}$. The case $z = 0$ implies a path for which the product of the signs from u to the destination is equal to $+1$, termed as *even-parity path*. Similarly, $z = 1$ refers to a path for which the product of the signs from u to the destination is -1, termed as *odd-parity path*. The value functions for the even and odd-parity paths from node u to the destination are given as follows:

$$S(u,0) = \min_{v:(u,v)\in E} \{d_{u,v} + \delta(s_{u,v}-1)S(v,0) + \delta(s_{u,v}+1)S(v,1)\} \tag{2}$$

$$S(u,1) = \min_{v:(u,v)\in E} \{d_{u,v} + \delta(s_{u,v}-1)S(v,1) + \delta(s_{u,v}+1)S(v,0)\} \tag{3}$$

where $S(u,0)$ is the even and $S(u,1)$ is the odd-parity path. The delta function is given as $\delta(0) = 1$ and $\delta(x) = 0$ for all $x \neq 0$. The minimum total cost for influencing the target node positively is then determined from $S(u_o, 0)$. The boundary conditions are given as follows:

$$S(u_d, 0) = 0, \qquad S(u_d, 1) = \infty \tag{4}$$

The pseudo-code for solving (1) is presented in Algorithm 1.

V. Influence Propagation with Message Deterioration and Ignorance

An idea propagating through a social network often *distorts* as it is repeated, also known as the “Telephone” effect [28]. In effect, persons’ individual interpretations or subjective priority assessments may alter the content of the message (news, idea) reaching a target person. In this section, we quantify the *impact* of *message freshness* on influence propagation. We assume that persons may *ignore* a received message based on the link strength between the nodes and message freshness. In this case, the recommender has to *reactivate* the node with an additional cost to refresh the message either by an advertisement or a special promotion. The recommender may also activate a node on the path solely with the purpose of refreshing the message, even without being ignored, with a cost. We determine the optimal path with minimum expected cost and the activation sequence, i.e., the nodes to activate in case of no ignorance. In case a person ignores a message, which may or may not occur, reactivation is necessary.

Algorithm 2 Minimum-Cost Influence Propagation with Message Deterioration and Ignorance

procedure DETERIORATIONMINCOST(V, E, s, d, u_o, u_d, K)
Assign the boundary conditions from (9)
for all u from u_d to u_o in reverse topological order **do**
for $k := 1$ **to** K **do**
Compute $S(u,k,0)$ and $S(u,k,1)$ using (6)
Record the decisions (next hop and activation) $\pi(u,k,0)$ and $\pi(u,k,1)$ that lead to $S(u,k,0)$ and $S(u,k,1)$
Compute the optimal path in a way similar to Lines 6-11 in Algorithm 1, additionally taking into account the decisions on age and activation
return the optimal path and the minimum expected cost $S(u_o, 1, 0)$

In order to model the effect of deterioration and ignorance, we take into account message freshness and the possibility that nodes may choose to ignore each other. For message freshness, we define the age of a message, k as the number of hops the message has traveled since the last activation. An activation resets the message age to 1 and is required if a node chooses to ignore the message. The recommender may choose to activate a node even if the message is not ignored, in order to reset the message age to 0, however, there is a cost c for each activation. The maximum age for the message is K; if a message is of age K, then the next node on the path has to be activated.

We denote the cost to influence node v through node u with a message of age $k_{u,v}$ by the random variable $d_{u,v}(k_{u,v})$. Every node has a growing tendency of ignoring a message as its age increases. We denote by $p_{u,v}(k_{u,v})$ the probability that node v ignores node u for a message of age $k_{u,v}$, and is a monotonically increasing function of the distance between two nodes and message age. The minimum expected cost is:

$$\begin{aligned} \min_{\substack{P\in\mathcal{P}\\ a_{u,v}}} \sum_{(u,v)\in P} \big\{ & \mathrm{E}[d_{u,v}(k_{u,v})] + c\,\delta(a_{u,v}-1)(1-p_{u,v}(k_{u,v})) \\ & + c\,p_{u,v}(k_{u,v}) \big\} \\ \text{s.t.} \prod_{(u,v)\in P} & s_{u,v} = 1 \\ & a_{u,v} \in \{0,1\},\ \forall (u,v)\in P \\ & k_{u,v} \in \{1,2,\ldots,K\},\ \forall (u,v)\in P \\ & k_{v,w} = (k_{u,v}+1)\delta(a_{u,v}),\ \forall (u,v),(v,w)\in P \\ & k_{u_o,v} = 1,\ \forall (u_o,v)\in P \end{aligned} \tag{5}$$

where we optimize over path P and the activation sequence $(a_{u,v})$. Each variable $a_{u,v}$ is 1 if node v is activated, and 0 otherwise. We solve (5) by utilizing dynamic programming. We state the recursive equations for backward induction in (6) where $S(u,k,z)$ denotes the value at node u with message age $k \in \{1,2,\ldots,K\}$ and disparity $z \in \{0,1\}$.

Algorithm 3 Minimum-Cost Positive Influence Propagation in Graphs with Cycles

1: **procedure** CycleMinCost(V, E, s, d, u_o, u_d)
2: $N_+ := \varnothing$; $N_- := \varnothing$ ▷ N_+ and N_- are the sets of nodes with positive and negative permanent labels, respectively.
3: $\pi_+[u_o] := 0$; $\pi_-[u_o] := \infty$ ▷ π_+ and π_- are temporary labels
4: **for all** $u \in V - \{u_o\}$ **do**
5: $\pi_+[u] := \infty$; $\pi_-[u] := \infty$
6: **while true do**
7: Find a node $v \in V$ such that $\pi[v] = \min_{\substack{i \in V - N_+ \\ j \in V - N_-}} \{\pi_+[i], \pi_-[j]\}$
8: **if** $\pi[v] = \pi_+[v]$ **then**
9: $\pi'_+[v] := \pi[v]$; $N_+ := N_+ \cup \{v\}$ ▷ π'_+, π'_-: permanent labels.
10: **else** ▷ $\pi[v] = \pi_-[v]$
11: $\pi'_-[v] := \pi[v]$; $N_- = N_- \cup \{v\}$
12: **if** $u_d \in N_+$ **then break**
13: **for all** $u \in V$ s.t. $(v, u) \in E$ **do** ▷ Update temporary labels
14: $t := \pi[v] + d_{v,u}$
15: **if** $s_{v,u} = +1$ **then**
16: **if** $\pi[v] = \pi_+[v]$ **and** $u \in V - N_+$ **and** $t < \pi_+[u]$ **then**
17: $\pi_+[u] := t$
18: $pred[u][+1] := v$ ▷ Predecessor of u on positive path
19: **else if** $\pi[v] = \pi_-[v]$ **and** $u \in V - N_-$ **and** $t < \pi_-[u]$ **then**
20: $\pi_-[u] := t$
21: $pred[u][-1] := v$ ▷ Predecessor of u on negative path
22: **else**
23: **if** $\pi[v] = \pi_+[v]$ **and** $u \in V - N_-$ **and** $t < \pi_-[u]$ **then**
24: $\pi_-[u] := t$
25: $pred[u][-1] := v$
26: **else if** $\pi[v] = \pi_-[v]$ **and** $u \in V - N_+$ **and** $t < \pi_+[u]$ **then**
27: $\pi_+[u] := t$
28: $pred[u][+1] := v$
29: $P^* := [u_d]$; $u := u_d$; $positive := 1$
30: **while** $u \neq u_o$ **do**
31: $temp := pred[u][positive]$
32: Prepend $temp$ to P^*
33: $positive := positive \times s_{temp,u}$
34: $u = temp$
35: **return** the optimal path P^* and the minimum cost $\pi'_+[u_d]$

$$S(u,k,z) = \min_v \Big\{ \mathrm{E}[d_{u,v}(k)] + p_{u,v}(k)(c + \delta(s_{u,v}-1)S(v,1,z) + \delta(s_{u,v}+1)S(v,1,\bar{z})) + (1-p_{u,v}(k)) \min\{\theta, \phi\} \Big\} \quad (6)$$

where $\bar{0} = 1$ and $\bar{1} = 0$. We denote θ and ϕ as follows:

$$\theta = \delta(s_{u,v}-1)S(v,1,z) + \delta(s_{u,v}+1)S(v,1,\bar{z}) + c \quad (7)$$

$$\phi = \delta(s_{u,v}-1)S(v,k+1,z) + \delta(s_{u,v}+1)S(v,k+1,\bar{z}) \quad (8)$$

where θ corresponds to the case in which the next node on the path is activated, whereas ϕ means that no activation takes place. The boundary conditions are given as:

$$S(u_d,k,0) = 0, \ S(u_d,k,1) = \infty, \ \forall k \in \{0,1,\ldots,K\} \quad (9)$$

The minimum expected cost is then given by $S(u_o,1,0)$. The steps of the proposed scheme are given in Algorithm 2.

VI. Minimum-Cost Influence Propagation for Graphs with Cycles

This section studies the minimum-cost influence propagation problem introduced in (1) for directed cyclic graphs. As the graphs considered in this section may consist of directed cycles, methods from Section IV cannot be applied to tackle (1) directly. Hence, we present a modified Dijkstra-type algorithm to solve (1) in Algorithm 3. It uses the same central structure as the usual Dijkstra's shortest path algorithm, which maintains and updates a list of shortest paths from the source to every other node in the graph (temporary labels π_+ and π_-). The first difference in our algorithm is that we keep track of the shortest path from the source to any node $u \in V$, for two cases. The first case accounts for the cost of a path where u is positively influenced by the source ($\pi_+(u)$), whereas the second case ($\pi_-(u)$) represents negative influence.

At each iteration, the shortest temporary label is fixed as a permanent label, represented by arrays π'_+ and π'_- (Lines 7-11). The temporary labels of the successors of the corresponding node are then updated (Lines 13-28). An important difference with the vanilla Dijkstra's algorithm is that this updating process takes the edge signs and path parity into account. We also keep track of the decisions corresponding to the optimal path at each node using the $pred$ array. After the iteration is over, the $pred$ array is used to generate the optimal path (Lines 30-34). It is important to note that the optimal path in our model may include a cycle, unlike the generalized shortest path algorithms for cyclic graphs. The intuition behind this idea lies in the fact that traversing a cycle may result in an even-parity path with a smaller cost than an acyclic path, due to a sign change through the cycle. Note that the time complexity of Algorithm 3 is asymptotically the same as Dijkstra's original shortest path algorithm.

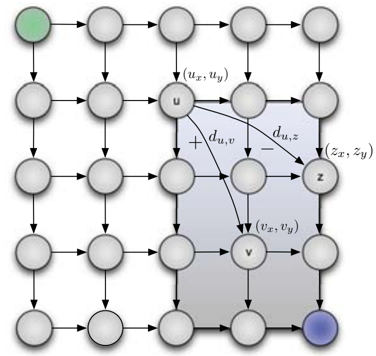

Fig. 2. Grid network structure for the signed social graph.

VII. Numerical Results

We initially implement a small-scale network to motivate the propagation model and the optimal policies. Next, we move to a large-scale network and use the online Epinions dataset to test and demonstrate the impact of our findings. To this end, we first consider a grid network with directed acyclic links given in Fig. 2. In order to prevent directed cycles, we entail the following condition on the network structure: Edge (u, v) exists only if $u_x \leq v_x$, $u_y \leq v_y$, and $u \neq v$. That is, node u can only influence the nodes in the shaded rectangle in Fig. 2. Here, the source node is at the top left corner in green and the destination node is at the bottom right corner in blue.

We consider a random graph where the probability of existence for edge (u, v) is modeled by a Bernoulli random variable with a parameter that is a monotonically decreasing function of the distance between nodes u and v. We presume that edges with a small $\|u - v\|$ refers to close neighbors such that individuals are frequently interacting with each other. We note that this differs from the traditional notion of *friendship*,

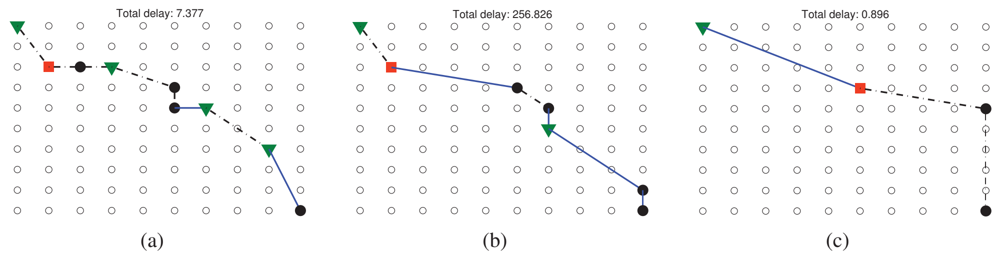

Fig. 3. Simulation results for cost minimization with message deterioration and ignorance with activation cost (a) 1, (b) 100, (c) 1 with distance-independent costs, i.e., $\alpha = 0$. Solid (dash-dotted) lines denote edges with a positive (negative) sign. The nodes visited by the optimal path are filled with black. A square with an orange filling indicates activation as a result of ignorance, whereas a triangle with a green filling indicates activation without ignorance.

as two individuals may be frequently engaging in social interactions with different ideologies such as political rivals. To this end, we posit that this information is gathered from sensor data that measures the frequency of one person *interacting* with another person through social discussions or debates. This can be obtained by various methods ranging from analyzing the conversations in which one person *mentions* another or processing the textual transactions in social media. Similarly, a large $\|u - v\|$ stands for a distant neighbor such that the two persons know each other at an acquaintance level. We model the propagation cost between nodes u and v as a random variable $d_{u,v}$ uniformly distributed over $[0, d_{max}(u, v)]$:

$$d_{u,v} \sim U(0, d_{max}(u, v)) \tag{10}$$

where $d_{max}(u, v)$ is chosen as a monotonically increasing function of the distance between the two nodes:

$$d_{max}(u, v) = \beta \|u - v\|^{\alpha} \quad \alpha, \beta \geq 0 \tag{11}$$

where $\|u-v\| \geq 1$ for all $u \neq v$. The parameter α is introduced to capture the impact of social distances on physical costs such as propagation delay. To this end, a large α increases the difference between the propagation costs incurred by close and distant neighbors. In effect, many real-world applications suggest that propagating a message through distant neighbors often takes more effort. On the other hand, when α is decreased, distant neighbors start being treated by the network as close contacts as their propagation cost approaches to those. All neighbors, whether socially distant or close, are treated as equals when α is zero. Accordingly, α is termed the *distance impact parameter*. Coefficient β is a design-specific weight parameter that is equal for all node pairs. We denote the probability of an edge having a positive sign by μ, which refers to a friendship relation between the two nodes. Accordingly, the probability of any edge having a negative label is $\bar{\mu} = 1-\mu$ in which case the two persons experience an antagonistic relationship type. The default values for our simulations are $\alpha = 1/2$, $\beta = 1$, and $\mu = 1/2$.

We now demonstrate the optimal policies for a 10-by-10 grid network. We introduce *ignorance* through an ignorance probability $p_{u,v}(k_{u,v})$ which is the probability that node v will ignore node u while u is attempting to transmit a message of age $k_{u,v}$ to v. It is defined as a monotonically increasing function of $k_{u,v}$ and the distance between the two nodes. Fig. 3(a), (b), (c) show optimal paths for cost minimization with message deterioration and ignorance following the steps in Algorithm 2. By comparing Fig. 3(a) and (b), we see that increasing the activation cost results in a lower number of activations even though older messages are more likely to be ignored. We note that in Fig. 3(a), since the activation cost is low, a large number of nodes are activated even though the message they receive has age 1. This is done in order to avoid ignorance further down the path. Fig. 3(c) shows the optimal path for the same setup except the costs do not depend on distance; therefore, the optimal path is able to make bigger jumps without incurring additional cost. However, we observe that bigger jumps are more likely to result in ignorance, and therefore a penalty for activation in the total cost.

We perform our large-scale simulations using the online data from the Epinions social graph [22]. Epinions is a consumer review website where users can indicate whether they trust or distrust the opinions of other users. This signed social graph has 131828 nodes and 841372 edges. Throughout our evaluations, the source and destination nodes are selected randomly. For every possible source-destination pair, we try to find the optimal path and propagation policy such that the source positively influences the destination. Since there is currently no algorithm to compute influence propagation in signed networks, we compare the results between our Dijkstra-type algorithm and a naïve myopic algorithm to find a low cost positive path. We first implement the Dijkstra-type algorithm in Algorithm 3 to find the paths that minimize the sum of the costs between source and destination nodes. We assume that the cost between nodes i and j is given by $\kappa|i - j|$. In our simulations, we select $\kappa = 0.1$. The myopic algorithm is referred to as *shortest DFS*. It is a depth first search algorithm that traverses the graph starting from the source looking for the destination. At each node, it selects the successor with lowest cost. The procedure is recursively repeated until the destination is reached. If a path from a node to the destination is not found, the algorithm selects a successor of the node with higher cost. For computational reasons, we limit the length of the paths to 1500.

We randomly select 100 sources and 100 destinations. We apply the two algorithms to compute the optimal paths between all 10000 source-destination pairs. We then repeat the process for 500 sources and 500 destinations. Finally, we compute paths between 10000 sources and 10000 destinations using Algorithm 3 and the *min negative path* algorithm. We have observed that in this case, *shortest DFS* algorithm could not terminate in a reasonable amount of computing time.

Tables I and II show statistics about the paths computed using Algorithm 3 and the *shortest DFS* algorithm. In the case with 100 sources and 100 destinations, using Algorithm 3 we find that each of the 100 sources is positively connected

TABLE I
MINIMUM DELAY ALGORITHM (ALGORITHM 3) RESULTS WITH EPINIONS.

Number of sources/dest.	Total number of paths found	Average path length	Median path length	Average path cost	Median path cost
100/100	8830	54.450	40.0	3436.569	2488.4
500/500	218499	55.148	42.0	3419.907	2363.7
10000/10000	78029370	47.024	30.0	5027.842	4145.1

to 88.3 destinations on average. The average path length is 54.45 hops with an average path cost of 3436.57. On the other hand, the positive paths found by the *shortest DFS* algorithm has an average length of 660.73 hops with an average cost of 17604.64. We observe that on average, the cost of the paths found using Algorithm 3 is less than a fifth of the cost of the paths found with *shortest DFS*. No path is found by the *shortest DFS* algorithm for 8959 of the 10000 source-destination pairs. We note that some of these pairs may in fact be unreachable as a natural result of the graph structure, i.e., the source and the destination may not be connected. However, the same analysis shows that there exists only 1170 cases in which the destination is not reachable from the source with Algorithm 3. Hence, the number of paths discovered by the *shortest DFS* algorithm is only about a tenth of the paths found by Algorithm 3. In the case with 500 sources and 500 destinations, we find from Algorithm 3 that each source is positively connected on average to 436.998 destinations. In other words, we can find a positive path for source-destination pairs in 87.4% of cases. The average path length is 55.15 hops with an average path cost of 3419.91. On the other hand, with the *shortest DFS* algorithm the average length of the paths is 726.95 hops with an average cost of 19134.18. As in the previous case, the average cost of Algorithm 3 is less than a fifth of the cost of the paths found using *shortest DFS*. We similarly observe that *shortest DFS* can find about one tenth of the paths that Algorithm 3 computes. Another significant observation is that a randomly chosen source is very likely to be positively connected to a random destination.

VIII. CONCLUSION

We have studied a signed social network in which friends and foes are identified by positive and negative signs. We have introduced a propagation model to influence a target person $infavor$ of an idea, a product, or an action. Persons are influenced in their decisions by the observations made available to them. We provide the optimal propagation policies under influence-centric constraints. We implement the proposed algorithms in a controlled environment as well as using real-world traces to understand the optimal policies in both small and large-scale networks. We expect our study to open many directions, including optimal multicast, multilevel relationship or influence types, and the impact of other personal factors on opinion forming and recommendation policies.

TABLE II
RESULTS OF THE *shortest DFS* (MYOPIC) ALGORITHM WITH EPINIONS.

Number of sources/dest.	Number of paths found	Average path length	Median path length	Average path cost	Median path cost
100/100	1041	660.727	638.0	17604.642	16875.4
500/500	27309	726.949	727.0	19134.178	18425.9

REFERENCES

[1] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge, UK: Cambridge University Press, 2010.

[2] D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, pp. 393–440, 1998.

[3] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proc. of National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[4] P. Domingos and M. Richardson, "Mining the network value of customers," in *Proc. of ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2001, pp. 57–66.

[5] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. of ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2003, pp. 137–146.

[6] M. J. Brzozowski, T. Hogg, and G. Szabo, "Friends and foes: ideological social networking," in *Proc. of SIGCHI Conf. on Human Factors in Computing Systems*, 2008, pp. 817–820.

[7] J. Kunegis, A. Lommatzsch, and C. Bauckhage, "The slashdot zoo: mining a social network with negative edges," in *Proc. of Int. Conf. on WWW*, 2009, pp. 741–750.

[8] T. Hogg, D. M. Wilkinson, G. Szabo, and M. J. Brzozowski, "Multiple relationship types in online communities and social networks." in *AAAI Spring Symposium: Social Information Processing*, 2008, pp. 30–35.

[9] C. A. Lampe, E. Johnston, and P. Resnick, "Follow the reader: filtering comments on slashdot," in *Proc. of SIGCHI Conf. on Human Factors in Computing Systems*, 2007, pp. 1253–1262.

[10] P. Anchuri and M. Magdon-Ismail, "Communities and balance in signed networks: A spectral approach," in *Proc. of IEEE/ACM Int. Conf. on Advances in Social Networks Analysis and Mining*, 2012, pp. 235–242.

[11] M. Kimura and K. Saito, "Tractable models for information diffusion in social networks," in *Knowledge Discovery in Databases (PKDD'06)*, 2006, pp. 259–271.

[12] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proc. of ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2009, pp. 199–208.

[13] W. Chen, Y. Yuan, and L. Zhang, "Scalable influence maximization in social networks under the linear threshold model," in *Proc. of IEEE Int. Conf. on Data Mining (ICDM'10)*, 2010, pp. 88–97.

[14] A. Goyal, F. Bonchi, and L. V. Lakshmanan, "Learning influence probabilities in social networks," in *Proc. of ACM Int. Conf. on Web Search and Data Mining*, 2010, pp. 241–250.

[15] B. W. Hung, S. E. Kolitz, and A. Ozdaglar, "Optimization-based influencing of village social networks in a counterinsurgency," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 3, pp. 52:1–52:22, 2013.

[16] C. Budak, D. Agrawal, and A. El Abbadi, "Limiting the spread of misinformation in social networks," in *Proc. of Int. Conf. on WWW*, NY, 2011, pp. 665–674.

[17] Y. Li, W. Chen, Y. Wang, and Z.-L. Zhang, "Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships," in *Proc. of ACM Int. Conf. on Web Search and Data Mining*, 2013, pp. 657–666.

[18] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, pp. 415–444, 2001.

[19] W. L. Langer, *European alliances and alignments, 1871-1890*. Greenwood Press, 1977.

[20] B. E. Schmitt, "Triple alliance and triple entente, 1902–1914," *The American Historical Review*, vol. 29, no. 3, pp. 449–473, 1924.

[21] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust," in *Proc. of Int. Conf. on WWW*, NY, 2004, pp. 403–412.

[22] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proc. of SIGCHI Conf. on Human Factors in Computing Systems*, New York, NY, 2010, pp. 1361–1370.

[23] ——, "Predicting positive and negative links in online social networks," in *Proc. of Int. Conf. on WWW*, NC, 2010, pp. 641–650.

[24] F. S. F. Walter, S. Battiston, "A model of a trust-based recommendation system on a social network," *Autonomous Agents and Multi-Agent Systems*, vol. 16, pp. 57–74, 2008.

[25] F. Heider, "Attitudes and cognitive organization," *The Journal of Psychology*, vol. 21, no. 1, pp. 107–112, 1946.

[26] D. Cartwright and F. Harary, "Structural balance: a generalization of heider's theory." *Psychological review*, vol. 63, no. 5, p. 277, 1956.

[27] S. E. Dreyfus and A. M. Law, *The art and theory of dynamic programming*. Elsevier Science, 1977.

[28] S. Blackmore, *The meme machine*. Oxford, UK: Oxford University Press, 2000.