Communicating in a Socially-Aware Network: Impact of Relationship Types

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*—Communication networks are linked to and influenced by human interactions. Socially-aware systems should integrate these complex relationship patterns in the network design. This paper studies the impact of friendship and antagonistic relationships between individuals on optimal network propagation policies. We develop a network propagation model for signed networks, and determine the optimal policies to influence a target node with an opinion while minimizing the total number of persons against it. We also provide extensions to this problem to elaborate on the impact of network parameters, such as minimum-delay propagation, while limiting the number of persons influenced against the idea before reaching the target. We provide numerical evaluations in a synthetic setup as well as the Epinions online social dataset. We demonstrate that propagation schemes with social and influence-centric constraints should take into account the relationship types in network design.**

I. Introduction

Social networks have received growing attention due to the increasing popularity and capabilities of online networking enabled by smart devices [1]. Friendship, trust, and influence relationships, often reflected by the changes of user behavior [2], have been the focus of considerable research effort [3]–[10]. Studies often concentrate on networks with monolithic relationship types, in which all relationships are friendship relations [2]–[5]. However, as pointed in [6]–[10], real societies exhibit complex relationship structures and it is important to identify relationship types in social networks analysis.

Influence spreading in networks with a single relationship type (friendship) has been considered in [4], [5], [11]–[14]. The importance of identifying positive and negative relationship structures for influence spread has recently been pointed out [15]. Reference [15] identifies effective seeds for maximum influence spread in signed networks from a random influence diffusion perspective. By contrast, optimal propagation policies for networks with influence-centric constraints, and positive and negative relationship types, have not been studied in the existing literature. This is the goal of this paper.

We consider the influence propagation problem in a signed network as in Fig. 1. In this graph, positive and negative labels represent the friendship and antagonistic relationships, respectively. We utilize the principle of *homophily*, which states that persons favor the decisions of the people that share similar interests with them, while opposing the ideas of their foes [6], [16]. By this token, when a person observes that an ideological foe is against an idea, she will go against this

This research is sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-2-0053.

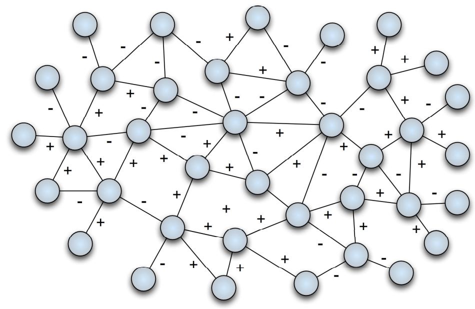


Fig. 1. Signed social network.

which means that she will support the original idea. This property has been observed in the formation of European alliances before World War I [17], [18].

In this paper, we develop algorithms to minimize propagation costs as well as the number of negatively influenced users. We assign a cost of propagation to each social link due to various social and physical causes such as the interaction frequency, propagation delay, strength of the relationship (social tie). This metric, which is often a combination of various factors, is specified by the system designer. For instance, we consider the metric of end-to-end delay, and investigate the problem of influencing a target node *positively* with minimum total propagation delay, which corresponds to the fastest policy among the other possibilities. Our main contributions in this paper are:

- We study the minimum cost network propagation problem for positive influence while ensuring that the number of negatively influenced users is below a given threshold.
- We determine the optimal propagation policies for minimizing the number of negatively influenced users while influencing a target person positively.
- We design algorithms to tackle cyclic as well as acyclic graphs.
- We demonstrate the results in an artificially created grid network in addition to the online Epinions dataset.

We observe that in most cases it is possible to find a short path between two nodes without influencing any node *against* an idea. In a companion paper [19], we study the cost minimization problem for influencing a target node positively without taking into account the negative influences. On the other hand, this paper considers the negatively influenced users and their impact on the influence propagation policies, both for acyclic and cyclic graphs. These results can be integrated to a number of social network applications in which relationship types have a significant impact on system benefits.

978-1-4799-7088-9/14/$31.00 ©2014 IEEE

Algorithm 1 Forward Induction Dynamic Programming for Limiting the Number of Negative Influences

procedure LIMITNEGATIVE(V, E, s, d, u_o, u_d, Q)
    Assign the boundary conditions from (4)
    for all u from u_o to u_d in topological order **do**
        for $q := 0$ **to** Q **do**
            Compute $S(u, q, 0)$ and $S(u, q, 1)$ using (2)-(3)
            Record the decisions (next hop) $\pi(u, q, 0)$ and $\pi(u, q, 1)$ that lead to $S(u, q, 0)$ and $S(u, q, 1)$
    return MAKEPATH(π, s, u_o, u_d, Q)
procedure MAKEPATH(π, s, u_o, u_d, q) ▷ Compute a path from u_o to u_d in the graph with edge signs s using the π decisions, for a maximum of q negatively influenced users
    $P^* := [u_d]$; $u := u_d$; $parity := 0$
    while $u \neq u_o$ **do**
        $temp := \pi(u, q, parity)$
        if $parity = s_{temp,u}$ **then** $parity := 0$ **else** $parity := 1$
        if $parity = 1$ **then** $q := q - 1$
        $u := temp$
        Prepend u to P^*
    return P^*

II. System Model

Let $G = (V, E)$ be a graph with $|V|$ nodes representing the social network. Initially, we consider a directed acyclic graph. The model is generalized to cyclic graphs in Section V. 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 other, positively or negatively, with a smaller delay representing a quicker response. From yet another perspective, it may be the frequency of interaction between the two persons. We provide a formal definition of the influence propagation problem in the sequel.

III. Limiting the Number of Negative Influences

The propagation model studied in the previous sections required the destination node to be positively influenced while the dispositions of intermediate nodes were not of concern. Real-life scenarios, on the other hand, often necessitate one to avoid a path through which a large number of intermediate nodes are influenced negatively. To this end, we study the problem of how to influence the destination positively while controlling the number of negatively influenced intermediate nodes. Note that, assessing the influence type from the source node to each of the intermediate nodes requires a forward induction formulation, for which we develop a dynamic program. Let P_u denote the fragment of the path P that ends at node u. That is, P_u is a path from u_o to node u with the condition that if $(u', v') \in P_u$, then $(u', v') \in P$. The problem can be stated as:

$$\begin{aligned} \min_{P \in \mathcal{P}} \quad & \sum_{u,v:\,(u,v)\in P} d_{u,v} \\ \text{s.t.} \quad & \prod_{u,v:\,(u,v)\in P} s_{u,v} = +1 \\ & \left| \left\{ u : \prod_{u',v':\,(u',v')\in P_u} s_{u',v'} = -1 \right\} \right| \leq Q \end{aligned} \tag{1}$$

where Q is the maximum number of negatively influenced intermediate nodes, and we focus on deterministic costs. We define $S(u, q, 0)$ as the value of the minimum-cost even-parity path connecting the source node u_o with node u when the number of negatively influenced intermediate users are no more than q. Similarly, $S(u, q, 1)$ is the minimum-cost for the odd-parity path between u_o and u with at most q negatively influenced intermediate users. We then define the recursive relations for the even and odd-parity paths as follows:

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

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

for $u \in V$ and $0 \leq q \leq Q$. The boundary conditions are:

$$S(u_o, q, 0) = 0, S(u_o, q, 1) = \infty, \ \forall q \in \{0, 1, \ldots, Q\} \tag{4}$$

$S(u_d, Q, 0)$, can be determined using Algorithm 1.

Algorithm 2 Minimize the Total Number of Negatively Influenced Persons

procedure MINNEGATIVE(V, E, s, u_o, u_d, N)
    Set the boundary conditions from (8)
    for all u from u_o to u_d in topological order **do**
        for $n := 1$ **to** N **do**
            Compute $S(u, n, 0)$ and $S(u, n, 1)$ using (6)-(7)
            Record the decisions (next hop) $\pi(u, n, 0)$ and $\pi(u, n, 1)$ that lead to $S(u, n, 0)$ and $S(u, n, 1)$
    return MAKEPATH(π, s, u_o, u_d, N)

IV. Minimizing the Number of Negative Influences

We next consider a variation of the influence propagation problem in Section III. Specifically, we focus on *minimizing the number of negatively influenced users* subject to a maximum number of hops allowed before reaching the destination.

$$\begin{aligned} \min_{P \in \mathcal{P}} \quad & \left| \left\{ u : \prod_{u',v':\,(u',v')\in P_u} s_{u',v'} = -1 \right\} \right| \\ \text{s.t.} \quad & \prod_{u,v:\,(u,v)\in P} s_{u,v} = +1, \quad |P| \leq N \end{aligned} \tag{5}$$

In order to formulate the dynamic program we define $S(u, n, 0)$ as the number of negatively influenced users through the even-parity path between the source node u_o and node u where no more than n hops are used to reach u. Similarly, we let $S(u, n, 1)$ denote the number of negatively influenced users through the odd-parity path between u_o and u with at most n hops from u_o to u. Then the

Algorithm 3 Minimizing the Number of Negatively Influenced Persons in Graphs with Cycles

```
procedure CycleMinCost(V, E, s, d, u_o, u_d)
  N_+ := ∅; N_- := ∅        ▷ N_+ and N_- are the sets of nodes with positive and negative permanent labels, respectively.
  π_+[u_o] := 0; π_-[u_o] := ∞        ▷ π_+ and π_- are temporary labels
  for all u ∈ V − {u_o} do
    π_+[u] := ∞; π_-[u] := ∞
  while true do
    Find a node v ∈ V such that π[v] = min_{i ∈ V−N_+, j ∈ V−N_-} {π_+[i], π_-[j]}
    if π[v] = π_+[v] then
      π'_+[v] := π[v]; N_+ := N_+ ∪ {v}     ▷ π'_+ and π'_- are permanent labels.
    else                                    ▷ π[v] = π_-[v]
      π'_-[v] := π[v]; N_- = N_- ∪ {v}
    if N_+ ∩ N_- = V then break
    for all u ∈ V s.t. (v, u) ∈ E do        ▷ Update temporary labels
      if s_{v,u} = +1 then
        if π[v] = π_+[v] and u ∈ V − N_+ and π[v] < π_+[u] then
          π_+[u] := π[v]
          pred[u][+1] := v                  ▷ Predecessor of u on positive path
        else if π[v] = π_-[v] and u ∈ V − N_- and π[v] + 1 < π_-[u] then
          π_-[u] := π[v] + 1
          pred[u][−1] := v                  ▷ Predecessor of u on negative path
      else
        if π[v] = π_+[v] and u ∈ V − N_- and π[v] + 1 < π_-[u] then
          π_-[u] := π[v] + 1
          pred[u][−1] := v
        else if π[v] = π_-[v] and u ∈ V − N_+ and π[v] < π_+[u] then
          π_+[u] := π[v]
          pred[u][+1] := v
  P* := [u_d]; u := u_d; positive := 1
  while u ≠ u_o do
    temp := pred[u][positive]
    Prepend temp to P*
    positive := positive × s_{temp,u}
    u = temp
  return the optimal path P*
```

recursive relations for the even and odd-parity paths are given as:

$$S(u,n,0) = \min_{v:(u,v)\in E}\{\delta(s_{u,v}-1)S(v,n-1,0) + \delta(s_{u,v}+1)(S(v,n-1,1)+1)\} \quad (6)$$

$$S(u,n,1) = \min_{v:(u,v)\in E}\{\delta(s_{u,v}-1)(S(v,n-1,1)+1) + \delta(s_{u,v}+1)S(v,n-1,0)\} \quad (7)$$

where $u \in V$ and $n = 1,\ldots,N$. The maximum number of hops allowed to reach the destination is given by N. We define the boundary conditions for this problem as follows:

$$S(u_o,n,0) = 0,\ S(u_o,n,1) = \infty,\ \forall n \in \{0,1,\ldots,N\} \quad (8)$$

Lastly, the answer that refers to the even-parity path with the minimum number of negatively influenced users upon reaching the destination with no more than N hops is given by $S(u_d,N,0)$. Algorithm 2 provides the steps of the forward induction dynamic program formulated to find the optimal path for influencing a target node positively while minimizing the number of negatively influenced persons on the path.

V. Minimum-Cost Influence Propagation for Graphs with Cycles

In this section, we consider the minimum-cost influence propagation problem introduced in (5) for directed cyclic graphs. Techniques from Section IV cannot be applied directly here since the graphs may contain directed cycles. To this end, we propose a Dijkstra-type algorithm to solve (5) in Algorithm 3. We maintain and update a list of shortest paths from the source to each node in the graph by the temporary labels π_+ and π_-. However, unlike the conventional Dijkstra's algorithm, we keep track of the shortest path from the source to any node $u \in V$ for two different situations: $\pi_+(u)$ represents the number of negatively influenced users on a path where u is positively influenced by the source, whereas $\pi_-(u)$ represents negative influence.

The shortest temporary label is assigned as a permanent label at each iteration. The permanent labels are implemented by arrays $\pi_+^{'}$ and $\pi_-^{'}$. Then the temporary labels of the successors are updated, by taking into account the edge signs and path parity. The decisions at each node are stored in the $pred$ array, which is then used to generate the optimal path.

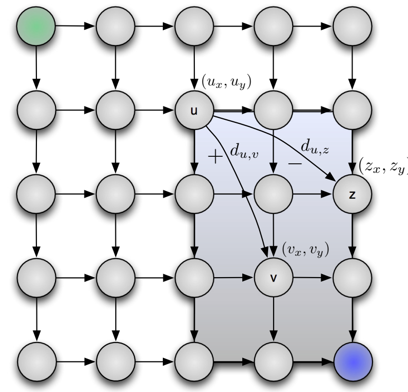


Fig. 2. Signed grid network. Source and destination nodes are indicated by green and blue colors, respectively.

VI. Numerical Results

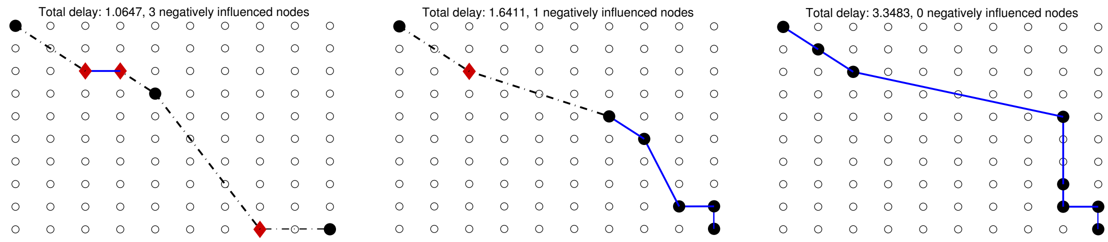
We first consider a grid network with directed acyclic links given in Fig. 2. The following condition is imposed to prevent directed cycles: Edge (u,v) exists only if $u_x \le v_x$, $u_y \le v_y$, and $u \ne v$. In other words, u can influence the nodes from the shaded rectangle in Fig. 2. The probability that an edge (u,v) exists is given by a Bernoulli random variable whose parameter is a monotonically decreasing function of the distance between u and v. Edges for which $\|u-v\|$ is small refers to close neighbors with frequent interaction. Unlike the conventional notion of *friendship*, we assume that two individuals may be frequently interacting even if they have an antagonistic relationship, such as political rivals with opposing ideologies. Accordingly, a distant neighbor or an acquaintance is represented by a large $\|u-v\|$. The propagation cost between u and v is a random variable $d_{u,v}$ uniformly distributed over $[0, d_{max}(u,v)]$:

$$d_{u,v} \sim U(0, d_{max}(u,v)) \quad (9)$$

in which $d_{max}(u,v)$ is a monotonically increasing function of the distance between u and v:

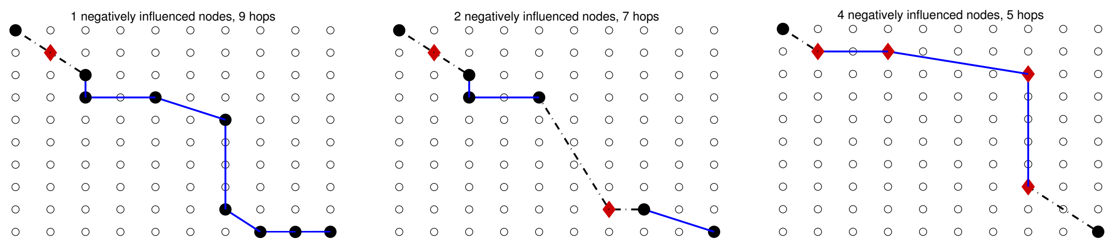
$$d_{max}(u,v) = \beta\|u-v\|^{\alpha} \quad \alpha,\beta \ge 0 \quad (10)$$

where $\|u-v\| \ge 1$ for all $u \ne v$. The impact of social distances on physical costs is indicated by the parameter α. The propagation costs for close and distant neighbors become more distinct larger values of α. Distant neighbors are treated comparatively to close neighbors as α is decreased, and are treated as equals for $\alpha = 0$. We term α as the *distance impact parameter*. We introduce a design-specific weight coefficient β. The probability of a positive label for any edge is μ, whereas the probability of a negative label is $\bar{\mu} = 1-\mu$.



(a) (b) (c)

Fig. 3. Optimal policies for the minimum cost with at most K negatively influenced nodes where (a) $K = 10$, (b) $K = 2$, (c) $K = 0$. Solid (dash-dotted) lines denote edges with a positive (negative) sign. The nodes visited by the optimal path are filled. A red diamond indicates a negatively influenced node.



(a) (b) (c)

Fig. 4. Simulation results for negative influence minimization with at most K hops where (a) $K = 9$, (b) $K = 7$, (c) $K = 5$. Solid (dash-dotted) lines denote edges with a positive (negative) sign. The nodes visited by the optimal path are filled, whereas a red diamond indicates a negatively influenced node.

Unless otherwise stated, our default values are $\alpha = 1/2$, $\beta = 1$, and $\mu = 1/2$. The size of our grid network is 10-by-10.

Fig. 3(a), (b), (c) show optimal paths that minimize total cost while negatively influencing no more than K nodes from Algorithm 1. As can be observed, a lower K results in a smaller number of feasible paths and the minimum total cost potentially increases. In addition, we see that the optimal path in Fig. 3(a) negatively influences only 3 nodes when it can actually influence $K = 10$ nodes, implying that it is not always optimal to negatively influence as many nodes as possible, and increasing K does not always result in a lower total cost.

Fig. 4(a), (b), (c) show optimal paths that minimize the number of negatively influenced users in at most K hops via Algorithm 2. We observe that lowering the value of K results in the elimination of some of the feasible paths and the optimal path is compelled to negatively influence more nodes.

In addition to the small-scale implementations that reflect our intuition, we also perform large-scale simulations using the online signed network topology from the Epinions consumer review website [20]. It is a graph with 131828 nodes and 841372 edges. We select the source and destination nodes randomly. Then we try to find an optimal path and propagation policy for each source-destination pair. We implement Algorithm 3 which seeks to minimize the number of negatively influenced nodes on a path between source and destination nodes. The temporary labels of each node at each iteration are updated to reflect the minimum number of negatively influenced nodes from the source to the node.

First, 100 sources and 100 destinations are selected randomly. Algorithm 3 is applied to find the optimal paths for all the 10000 source-destination pairs. These steps are then repeated for 500 sources and 500 destinations, and later for 10000 sources and 10000 destinations.

The results for the implementation of Algorithm 3, i.e., minimizing the total number of negatively influenced persons on the Epinions dataset, is given in Table I for 100, 500 and 10000 randomly selected sources and destinations. Importantly, we observe that even for a very large number of source and destination pairs, at least half of the paths have *zero* negatively influenced nodes. This is indicated by the fact that the median number of negatively influenced people is 0. The average number of hops in each path found by Algorithm 3 is less than 5. This justifies our intuition that it is actually possible to find a relatively short path from one node to another purely dominated by friendship (homophily) relations. In effect, our findings show that in general any node can influence another node *positively* within a small number of hops.

TABLE I
RESULTS FOR ALGORITHM 3 WITH EPINIONS.

Number of sources/dest.	Total number of paths found	Average path length	Median path length	Negatively influenced users (average)	Negatively influenced users (median)
100/100	8830	4.023	4.0	0.096	0.0
500/500	218499	4.097	4.0	0.057	0.0
10000/10000	78029370	4.646	5.0	0.123	0.0

VII. CONCLUSION

In this paper, we have considered a signed network in which signs represents the positive and negative relationships in a human community. Persons are influenced by the relationship structures while responding to their neighbors' decisions. We have proposed a network propagation scheme to influence a target person positively, and study how to reduce the number of negatively influenced users while doing so. We have implemented the proposed algorithms in an artificially created small dataset and using the large-scale Epinions signed network topology available online. Future directions include networks with more than two and/or fuzzy relationship types, and the extension of the point-to-point scheme to multicast influence propagation schemes.

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] 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.

[16] 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.

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

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

[19] B. Guler, B. Varan, K. Tutuncuoglu, M. Nafea, A. A. Zewail, A. Yener, and D. Octeau, "Optimal strategies for targeted influence in signed networks," in *Proc. of IEEE/ACM ASONAM International Workshop on Curbing Collusive Cyber-gossips in Social Networks*, 2014.

[20] 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.