A Mechanism Design Approach For Influence Maximization

Michael Levet[1] and Siddharth Krishnan[2]

[1] Dept. of Mathematics & Dept. of Computer Science and Engineering, University of South Carolina,
mlevet@email.sc.edu
[2] Discovery Analytics Center & Dept. Of Computer Science, Virginia Tech,
siddkris@cs.vt.edu

Abstract. With the proliferation of online social networks (OSNs), the characterization of diffusion processes and influence maximization over such processes is a problem of relevance and importance. Although several algorithmic frameworks for identifying influential nodes exist in literature, there is a paucity of literature in the setting of competitive influence. In this paper, we present a novel mechanism design approach to study the initial seeding problem where the agents, represented by vertices in the social network, are economically rational. The principals compete for influence in the network by setting price and incentives to illicit high degree initial subscribers, which in turn profit by infecting their neighbors. We restrict attention to equilibrium strategies and comparative statics for the agents.

Key words: social networks, mechanism design, influence maximization, agents, game theory

1 Introduction

The widespread adoption of online social media and networks, blogs, and internet shopping has transformed the web into a rich complex network structure. The web has therefore become a medium for diffusion of influence and information propagation. These spread processes have several applications including adoption of innovations [9], viral marketing [1], spread of rumors [3], and online recruitment. Information cascades resulting from such spreading processes have been widely observed in online social networks such as Facebook and Twitter. The algorithmic characterization of such processes is presented in two widely accepted models - the Linear Threshold and Independent Cascade [7, 5]. The influence maximization process over these two models was first presented in [8], where a principal seeks to find the optimal set of nodes to seed. Recent literature has also focused on a Competitive Contagion model in which two or more principals compete for influence in the network [2, 6] by seeding nodes in the network. An uninfected node changes state based on a stochastic process [6], rather than economic motivators, which is a major limitation in most algorithms that identify nodes of influence and agents of rapid propagation. We aim to bridge that gap in this contribution. In our work, we cast the problem as one of a referral game (described later) over a social network. The primary contribution of this work is the construction equilibrium strategies for the agents. Based on the importance of high-degree nodes in propogating influence [10], our second contribution provides conditions under which an agent will prefer seeding from one principal over another based on principal incentives and the agent's neighborhood size. Additionally, this model does not suffer from issues of complexity that plague the algorithmic approach [8, 10].

2 Model

We examine a dynamic game of imperfect information, which we call the *referral game*, over a social network. The players of this game are partitioned into two disjoint, finite sets. The first set $\mathcal{P}$ consists of the principals, and the second set $\mathcal{A}$ consists of the agents. The set of players in this game is $N := \mathcal{P} \dot{\cup} \mathcal{A}$. The social network is represented by an undirected graph $G(V, E)$ where $V(G) := \mathcal{A}$. Two agents $i, j \in \mathcal{A}$ are related if $ij \in E(G)$. We postulate that each agent is only aware of its neighborhood set. That is, each agent v believes that with probability 1 the social network is $K_{1,|N(v)|}$, where $N(v)$ denotes the open neighborhood of v. Furthermore, we postulate that each principal is only aware that G exists, but not of any properties of G.

Each principal $p \in \mathcal{P}$ provides a membership option for each agent at price $c_p \in \mathbb{R}_+$. This membership provides a valuation $v_p \in \mathbb{R}_+$ to the member, which is exogenously given for principal p. For every principal p, both the price c_p and valuation v_p are public information. Initially, each agent does not hold a membership, and an agent can hold at most one membership which cannot be revoked or changed later.

Definition 1 (Infection) *An agent v is said to be infected v holds a membership from some principal i. The agent v is said to be uninfected otherwise.*

The principals each seek to maximize influence over a graph. Define the membership function $\text{mem} : \mathcal{A} \to \mathcal{P} \cup \{\mathcal{U}\}$ which takes an agent and returns the principal from whom the agent is infected, or $\mathcal{U}$ if the agent is uninfected. Each principal i has utility function: $u_i(\mathcal{A}) = |\text{mem}^{-1}(i)|$. In order to maximize influence, each principal sets its price and incentives. Each principal's strategy set is $S := \mathbb{R}_+ \times [0, 1] \times \mathbb{R}_+$. Principal i's strategy $(c_i, \alpha_i, \beta_i) \in S_i$ is interpreted as follows: c_i is the principal's price for infection, α_i is the price discount for agents who subscribe at $T = 0$ (that is, each agent pays $\alpha_i c_i$), and β_i is the amount an agent receives for each additional agent it refers. Furthermore, agents who are referred to principal i at some time $T > 0$ also receive β_i (but must still pay the undiscounted price c_i). Each principal's strategy constitutes a mechanism with which to infect willing agents. The collection of these mechanisms induces the referral game. For the purpose of this paper, principals' strategies are taken as fixed. We restrict attention to the comparative statics with respect to the agents rather than the principals' strategies.

The game operates in discrete time steps starting at time $T = 0$. Prior to the start of the game, each principal fixes its strategy. At $T = 0$, any agent may become infected by at most one principal. After this initial joining period, uninfected agents can only become infected by principal i through referral from an agent already infected by i. An uninfected agent can only receive such a referral from one of its neighbors. At each discrete time $T > 0$, each infected agent may submit proposals simultaneously to any subset of its uninfected neighbors. The uninfected nodes that received proposals may accept at most one of the referrals or remain uninfected.

We now define each agent's utility function. Denote $\mathcal{U}$ as the option of remaining uninfected. Each agent v has the partial utility function of the form:

$$u_v : (\mathcal{P} \cup \{\mathcal{U}\}) \times (N(v) \cup \mathcal{P} \cup \{\emptyset\}) \times 2^{N(v)} \to \mathbb{R}_+ \quad (1)$$

The component $(\mathcal{P} \cup \{\mathcal{U}\})$ denotes the principal by whom v is infected, or whether v is uninfected. The component $N(v) \cup \mathcal{P} \cup \{\emptyset\}$ describes the player referring v. If the player is a principal, this indicates that v subscribed to the principal at time $T = 0$. The $\emptyset$ option denotes that no such referral has been made, and v is uninfected. Finally, the element from $2^{N(v)}$ denotes the set of neighbors which v successfully referred. We define the following cases:

1. $u_v(\mathcal{U}, \emptyset, \emptyset) = 0$. That is, an agent experiences no utility for remaining uninfected.
2. $u_v(i, i, s) = v_i - \alpha_i c_i + |s|\beta_i$ for any $i \in \mathcal{P}$ and any $s \in 2^{N(v)}$. This case indicates that agent v subscribed at time $T = 0$ to principal i.
3. $u_v(i, j, s) = v_i - c_i + (|s| + 1)\beta_i$ for any $i \in \mathcal{P}$, any $j \in N(v)$, and any $s \in 2^{N(v)\setminus\{j\}}$. This case indicates that agent v was referred by its neighbor j to principal i at time $T > 0$, and then v referred an additional $|s|$ of its uninfected neighbors.

3 Analysis

The *referral game* is a dynamic game of imperfect information. The solution concept we use is the Perfect Bayesian Equilibrium. We construct a Perfect Bayesian Equilibrium for this game using backward induction. At every time $T > 0$, each infected agent can only attempt to infect uninfected neighbors. Uninfected neighbors can only accept or reject referrals when received. We first show that an agent infected at time T need only attempt to infect its neighbors at time $T + 1$. This determines the Nash equilibria for terminal subgames.

Proposition 1 *Let x be an agent infected by principal i. Let $N_u(x)$ denote the set of uninfected neighbors of x. Then x will propose to each neighbor $y \in N_u(x)$ exactly once.*

Proof. Fix $y \in N_u(x)$. It is a weakly dominant strategy for x to propose to y. If y accepts a proposal from one of its infected neighbors, then we are done. Suppose instead at time $T_k > 0$, y rejects each of its infected neighbors' proposals. Recall that under y's belief system, the social network is $K_{1,|N(y)|}$, with y at the center. We thus have $\mathbb{E}[u_y(i, x, N_u(y))] = v_i - c_i + (|N_u(y)| + 1)\beta_i < 0$. As no infected agent can become uninfected, $|N_u(y)|$ is non-increasing as the game progresses. So $\mathbb{E}[u_y(i, x, N_u(y))] < 0$ for every time $T > T_k$. □

We now use Proposition 3.1 to show that the game is finite, which implies the existence of an equilibrium. [4]

Proposition 2 *Denote the set of agent states $K = \mathcal{P} \cup \{\mathcal{U}\}$, where u denotes a vertex remaining uninfected and each element $i \in \mathcal{P}$ denotes infection by principal i. The vertex states converge to a steady state equilibrium in $K^{|V|}$. Furthermore, this equilibrium is reached in $O(|V|)$ time steps.*

Proof. Once an agent becomes infected, its state is fixed. From Proposition 3.1, an agent v that remains uninfected at time $T = 0$ receives at most $|N(v)|$ referrals, of which v can select at most one. It follows that the vertex states converge to a steady state equilibrium in $K^{|V|}$. The bound is tight, as in the case of the graph G is a path on $|V|$ vertices where one endpoint agent becomes infected at $T = 0$. Then at most one additional agent is infected by referral at each subsequent time step, implying that the vertex states will converge to equilibrium in at most $|V|$ time steps. □

Corollary 1 *There exists a Perfect Bayesian Equilibrium. [4]*

The following proposition will construct an explicit Perfect Bayesian Equilibrium in mixed strategies for the case when $G \cong K_2$. This result will be used to construct the Perfect Bayesian Equilibrium for the general case.

Proposition 3 *Suppose the social network $G \cong K_2$. Let $j = \arg\max_{i \in \mathcal{P}} v_i - \alpha_i c_i$. Define: $Q_1 = \{i \in \mathcal{P} : v_i - c_i + \beta_i < v_j - \alpha_j c_j\}$ and let $Q_2 = \mathcal{P} \setminus Q_1$. Then there exists a symmetric Perfect Bayesian Equilibrium in mixed strategies, where each agent considers at most three principals.*

Proof. As this game is symmetric, there exists a Perfect Bayesian Equilibrium in mixed strategies where each agent employs the same strategy. [4] Consider the following cases:

Case 1: Suppose $Q_1 = \mathcal{P}$ and $Q_2 = \emptyset$. Then no player benefits from infection by referral. Therefore, each agent will always choose infection from j at time $T = 0$ if and only if $u(j, j, \emptyset) \geq 0$. Otherwise, each agent will remain uninfected. By construction, this strategy maximizes each player's payoff.

For the rest of the proof, assume there exists at least one principal x such that $v_x - c_x + \beta_x \geq 0$. For if no such principal x exists, then each agent will remain uninfected as the equilibrium strategy.

Case 2: Suppose $Q_1 = \emptyset$ and $Q_2 = \mathcal{P}$. Define:

$$k = \arg\max_{i \in \mathcal{P}} v_i - \alpha_i c_i + \beta_i \tag{2}$$

$$\text{s.t. } v_i - c_i + \beta_i \geq 0 \tag{3}$$

By construction of Q_2, such a k always exists. The game has exactly one stage if and only if both agents play the same strategy of either subscribing to principal k or remaining uninfected at time step $T = 0$. Otherwise, the uninfected agent at $T = 0$ will accept the referral at time $T = 1$. We now solve for the symmetric mixed strategies equilibrium. Suppose player 2 chooses infection at time $T = 0$ with probability p and chooses to remain uninfected at $T = 0$ with probability $1 - p$. Then player 1's expected payoffs from becoming infected and remaining uninfected at $T = 0$ respectively are:

$$\mathbb{E}[u_1(k, k, N(v_1))] = p \cdot (v_k - \alpha_k c_k) + (1 - p) \cdot (v_k - \alpha_k c_k + \beta_k) \tag{4}$$

$$\mathbb{E}[u_1(k, v_2, \emptyset)] = p \cdot (v_k - c_k + \beta_k) \tag{5}$$

Setting $\mathbb{E}[u_1(k, k, N(v_1))] = \mathbb{E}[u_1(k, v_2, \emptyset)]$ and solving for p yields:

$$p = \frac{v_k - \alpha_k c_k + \beta_k}{v_k - c_k + 2\beta_k} \tag{6}$$

Case 3: Suppose $Q_1 \neq \emptyset$ and $Q_2 \neq \emptyset$. Define k as in case 2. If $k \in Q_2$, then we reduce to case 2. Otherwise, suppose $k \in Q_1$. Define: $m = \arg\max_{i \in Q_2} v_i - \alpha_i c_i + \beta_i$. That is, m is an agent's preferred principal for infection at time $T = 0$ such that its neighbor will accept the proposal if uninfected. From the definition of j and the fact that $k \in Q_1$, we have $v_i - c_i + \beta_i \geq 0$ for all $i \in Q_2$. By construction, $v_k - \alpha_k c_k + \beta_k > v_m - \alpha_m c_m + \beta_m$. Observe the strategy of choosing infection from principal k at time $T = 0$ weakly dominates infection from any of the other principals of $Q_1 \setminus \{j\}$ at $T = 0$. Similarly, the strategy of choosing infection from principal m at time $T = 0$ weakly dominates infection to any other principal in $Q_2 \setminus \{j\}$ at time $T = 0$. We are thus left with four viable strategies for each agent: choose infection from principal $p \in \{m, j, k\}$ at $T = 0$ and then propose to its uninfected neighbor; and remain uninfected at $T = 0$, accepting any proposal where the expected payoff is non-negative. We now solve for a symmetric mixed strategies equilibrium.

Agent i mixes his strategies such that agent $-i$ is indifferent between between the four viable strategies. Let p_j, p_k, and p_m denote the frequencies in which each player at time $T = 0$ chooses infection from princiapls j, k and m respectively; and let p_u denote the frequency in which each player chooses to initiallly remain uninfected. For initially subscribing to a principal $y \in \{i, j, k\}$, each agent has the expected payoff:

$$\mathbb{E}[u_i(y, y)] = (1 - p_u) \cdot u(y, y, \emptyset) + p_u \cdot u(y, y, N(i)) = v_y - \alpha_y c_y + p_u \beta_y \tag{7}$$

And for opting to remain uninfected at time $T = 0$, each agent has the expected payoff:

$$\mathbb{E}[u_i(\mathcal{U})] = \sum_{x \in \{j,k,m\}} p_x \cdot u_i(x, -i, \emptyset) = \sum_{x \in \{j,k,m\}} p_x \cdot (v_x - c_x + \beta_x) \quad (8)$$

We solve for the mixed strategies equilibrium by setting (7) equal to (8), where we consider (8) for each $y \in \{i, j, k\}$. The following linear program yields such a mixed strategy equilibrium, with constraitnt (10) denoting the condition that (7) and (8) are equal in equilibrium.

$$\max_{p \in \Delta} \sum_{x \in \{j,k,m\}} p_x \cdot (v_x - c_x + \beta_x) \text{ s.t.} \quad (9)$$

$$\sum_{x \in \{j,k,m\}} p_x \cdot (v_x - c_x + \beta_x) = v_i - \alpha_i c_i + p_u \beta_i; \forall i \in \{j, k, m\} \quad (10)$$

□

Under an agent v's belief system, each of its neighbors believe $G \cong K_2$. As v believes $G \cong K_{1,|N(v)|}$, v believes each of its neighbors behaves independently and symmetrically in equilibrium. We use the mixed strategies equilibrium from Proposition 3.3 to construct v's equilibrium strategy at $T = 0$ based on $\mathbb{E}[|N_u(v)|]$.

Proposition 4 *Let $n \in \mathbb{Z}_{++}$ and suppose the social network $G \cong K_{1,n}$. Define j, Q_1, Q_2, and p^* as in Proposition 3.3. Denote p_u^* to be the component of p^* corresponding to remaining uninfected at time $T = 0$. Let v be the center vertex of G. Define: $m_1 = v_j - \alpha_j c_j$ and $m_2 = \max_{i \in \mathcal{P}} v_i - \alpha_i c_i + |N(v)| \cdot p_u^* \beta_i$. In equilibrium, agent v's expected utility is:* $\max\{m_1, m_2, 0\}$.

Proof. Recall that each node is only aware of its neighbors. We assume that each vertex in $N(v)$ plays the equilibrium strategy described in Proposition 3.3. Let X be the binomial random variable associated with the number of vertices in $N(v)$ which remain uninfected at time $T = 0$. Agent v seeks to maximize his expected utility. Suppose $p_u^* > 0$ and suppose there exists such a principal i such that $v_i - c_i + \beta_i \geq 0$. Now suppose v plays the strategy of subscribing to some principal i at time $T = 0$, then proposing to each uninfected neighbor at time $T = 1$. Then: $\max_{i \in \mathcal{P}} \mathbb{E}[u_v(i, i, N(v))] = m_2$. If agent v instead opts not to propose to each of its neighbors at time $T = 1$, then his maximum utility is $m_1 = v_j - \alpha_j c_j$, which is obtained by subscribing to agent j at time $T = 0$. Agent v's third option is to remain uninfected at time $T = 0$. Agent v chooses the best of these three strategies in equilibrium. □

The following theorem specifies and verifies a Perfect Bayesian Equilibrium.

Theorem 1 *Let G be a simple, undirected graph. Suppose each agent v plays the strategy specified by Proposition 3.4. This constitutes a Perfect Bayesian Equilibrium.*

Proof. Let v be an agent. For each time $T > 0$, v's strategy consists of proposing to each uninfected neighbor if v is infected; or if v is uninfected, it accepts a referall from its neighbor infected by principal j, which we denote x_j, if $x_j \in \arg\max_{i \in N(v)} u_v(i, x_i, N_u(v))$. By Proposition 3.1, this strategy induces a Nash equilibrium at each subgame for every $T > 0$. Recall that agent v believes with probability 1 that $G \cong K_{1,|N(v)|}$. It follows that the strategy at time $T = 0$ specified by Proposition 3.4 constitutes a Nash equilibrium. From this and Proposition 3.1, this strategy is sequentially rational. Consistency follows immediately from the fact that v believes $G \cong K_{1,|N(v)|}$ with probability 1. □

Finally, we examine the comparative statics, deriving sensitivity results for an agent's preferences for infection from a specific principal. We relate the size of an agent's neighborhood in the network to perturbations of the α and β parameter's in a principal's incentives package.

Theorem 2 *Let G be a graph, and let x be an agent. Suppose $\mathcal{P} = \{i, j\}$, $v_i = v_j$, and $c_i = c_j$, which we denote as v and c respectively. Then x strictly prefers infection from principal i rather than principal j at time $T = 0$ if one of the following conditions hold:*

1. $\alpha_i \leq \alpha_j$, $\beta_i \geq \beta_j$ *and at least one of the inequalities is strict.*
2. $\alpha_i > \alpha_j$, $\beta_i > \beta_j$ *and either:* $\beta_i - (1 - \alpha_j)c > 0$*; or* $|N(x)| > \dfrac{(\alpha_i - \alpha_j)c}{\beta_i - \beta_j}$.

Proof. If $\alpha_i \leq \alpha_j$ and $\beta_i \geq \beta_j$, then $u_x(i, i, S) \geq u_x(j, j, S)$ for every $S \subset N(x)$. So in this case, x prefers holding infection from principal i over principal j. Now suppose instead that $\alpha_i > \alpha_j$ and $\beta_i > \beta_j$. As x prefers infection from i over j at $T = 0$, it is necessary that under x's belief system, the following conditions must hold:

1. Under x's belief system, each $y \in N(x)$ prefers referral to i rather than joining j at $T = 0$.
2. Under x's belief system, each $y \in N_u(x)$ would accept referral to j. However, $u_x(i, i, N(x)) > u_x(j, j, N(x))$.

Condition one is equivalent to $\beta_i - c > -\alpha_j c$, which implies that $\beta_i - (1 - \alpha_j)c > 0$. Condition two is equivalent to $-\alpha_j c + |N(x)|\beta_j < -\alpha_i c + |N(x)|\beta_i$, which implies that $|N(x)| > \dfrac{(\alpha_i - \alpha_j) \cdot c}{\beta_i - \beta_j}$. □

4 Conclusion and Future Work

In this paper, we constructed equilibrium strategies for agents in the referral game under the assumption that the principals had no knowledge of the social network's structure. The natural extension of this work is to utilize these results to construct principals' equilibrium strategies. We propose designing a beliefs system regarding network connectivity for each principal, allowing for the continuation of the backward induction argument. Additionally, this model prohibits awareness of any mutual relations; while in most social situations, imperfect knowledge of mutual relations. We propose the problem of determining agent equilibrium strategies when every adjacent pair of agents i and j in the social network are aware of some $S \subset (N(i) \cap N(j))$ a priori. We are interested in the symmetric case, where related i and j are aware of the same $S \subset (N(i) \cap N(j))$, as well as the asymmetric case where i is aware of some $S_1 \subset (N(i) \cap N(j))$ and j is aware of some potentially different $S_2 \subset (N(i) \cap N(j))$.

Acknowledgments. We wish to thank Brendan Avent, Éva Czabarka, Stephen Fenner, and Alexander Matros for their helpful discussions and suggestions.

References

[1] Pedro Domingos and Matt Richardson. Mining the network value of customers. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '01, pages 57–66, New York, NY, USA, 2001. ACM. ISBN 1-58113-391-X.

[2] Moez Draief, Hoda Heidari, and Michael Kearns. New models for competitive contagion, 2014. URL `http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8399`.

[3] Adrien Friggeri, Lada A. Adamic, Dean Eckles, and Justin Cheng. Rumor cascades. In *Proceedings of the Eighth International Conference on Weblogs and Social Media, ICWSM 2014, Ann Arbor, Michigan, USA, June 1-4, 2014.*, 2014. URL `http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8122`.

[4] Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.

[5] Jacob Goldenberg, Barak Libai, and Eitan Muller. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 2001.

[6] Sanjeev Goyal and Michael Kearns. Competitive contagion in networks. *CoRR*, abs/1110.6372, 2011. URL `http://arxiv.org/abs/1110.6372`.

[7] M. Granovetter. Threshold models of collective behavior. *The American Journal of Sociology*, 83(6):1420–1443, 1978.

[8] David Kempe, Jon Kleinberg, and Éva 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*, KDD '03, pages 137–146, New York, NY, USA, 2003. ACM.

[9] Everett M. Rogers. *Diffusion of innovations*. Free Press, New York, NY [u.a.], 5th edition, 08 2003. ISBN 0-7432-2209-1, 978-0-7432-2209-9.

[10] Lior Seeman and Yaron Singer. Adaptive seeding in social networks. In *Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of*

Computer Science, FOCS '13, pages 459–468, Washington, DC, USA, 2013. IEEE Computer Society. ISBN 978-0-7695-5135-7.

Supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number D12PC000337, The US Government is authorized to reproduce and distribute reprints of this work for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, DTRA, or the US Government.