Maximizing Influence in Competitive Environments: A Game-Theoretic Approach

Andrew Clark and Radha Poovendran
Network Security Lab (NSL)
University of Washington, Seattle, WA 98195-2500
E-mail: RP3@uw.edu
Phone: (206) 221-6512
Outline

• Competing ideas in social networks
• Relevant properties of submodular functions and connected components
• Model for Dynamic Influence in Competitive Environments (DICE)
• Influence maximization games using DICE
  – Leader-follower game
  – Simultaneous-move game
• Simulation results with Wiki-vote
Competing Ideas in Social Networks

• Ideas spread through social networks
  – Word-of-mouth, blogs, online social networks,...

• Multiple ideas may propagate through same network

Rival advertisements  
Political campaigns  
Rumor spreading

• Possible scenarios
  – One idea introduced first, followed by others
  – All ideas introduced simultaneously
  – Reinforcing ideas introduced over time **(not considered today)**
Modeling Spread of Ideas

• Spreading ideas lead to following questions:
  – How does a set of individuals adopting an idea changes over time?
  – What fraction of population will eventually adopt each idea?
  – Who should be initially seeded with an idea to maximize its spread?

• Our Contribution: A tractable analytical model for answering the above questions within a game-theoretic framework for two ideas
Background: Submodular Functions

• Let $V$ denote a finite set, and let $f : 2^V \rightarrow \mathbb{R}$ map subsets of $V$ to real numbers.

• Function $f$ is submodular if for any $S \subseteq T \subseteq V$ and $v \notin T$,

$$f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T)$$

• Intuitively, this is a diminishing returns property.

• Properties of submodular functions:
  
  – Nonnegative weighted sum of submodular functions is submodular:
    
    $$(\alpha f_1 + \beta f_2)(S \cup \{v\}) - (\alpha f_1 + \beta f_2)(S) = \alpha f_1(S \cup \{v\}) - \alpha f_1(S) + \beta f_2(S \cup \{v\}) - \beta f_2(S)$$
    
    $$\geq \alpha(f_1(T \cup \{v\}) - f_1(T)) + \beta(f_2(T \cup \{v\}) - f_2(T))$$
    
    $$= (\alpha f_1 + \beta f_2)(T \cup \{v\}) - (\alpha f_1 + \beta f_2)(T)$$

  – If $f$ is submodular, then choosing a set $S$ with $|S| = k$ that maximizes $f(S)$ can be approximated in polynomial time up to a bound of $(1-1/e)$. 

Examples of Submodular Functions

• **First example:** Let $S$ represent a set of colored balls, and let $f(S)$ represent the number of colors in $S$

  ![Example 1](image1.png)

  - Number of colors increased by one
  - No effect on number of colors

• **Second example:** Let $S$ be a set of vectors, and let $f(S)$ be the rank of $S$
  - Adding a vector to $S$ increases the rank by one, unless it is in the span of $S$
Background: Connected Component

- A directed graph (digraph) is strongly connected if there is a path between any two vertices with correct edge orientations.
- A strongly connected component is a maximal strongly connected subgraph of a digraph.
Propagation of Ideas: Related Work

- Threshold model [Granovetter 1973]
- Cascade model [Goldenberg 2001]
- Generalized cascade model [Kempe 2003]
- Markovian models [Boyd 2005, Acemoglu 2010]
Maximizing Influence: Related Work

• Maximizing spread of a single idea
  – Goal: Choose a set $S$ of individuals to initially adopt and propagate the idea
  – Number of individuals eventually holding idea is submodular function of $S$ under generalized cascade model [Kempe 2003]

• Propagation of competing ideas [Bharathi 2007, Kostka 2008]
  – Extends generalized cascade model to multiple ideas
  – Individuals cannot switch between ideas
  – Formulate leader-follower games in which one player chooses a set $S_1$ and other player chooses a set $S_2$ after observing $S_1$
  – Reduces to facility location games, which are NP-hard to approximate
Our Competitive Influence Propagation Model

• We introduce a model of Dynamic Influence in Competitive Environments (DICE)

• Set of individuals (nodes) $V = \{1,...,n\}$, set of social relationships $E$ (represented by links), set of ideas $I = \{I_1, ..., I_m\}$

• At each time $t$, each individual $i$ interacts with one of its neighbors
  - Interacts with $j$ with probability $d_{ij}$
  - i.e., $i$ meets $j$ in person or reads $j$’s blog

• With probability $d_{ii}$, $i$ does not meet any neighbor; $d_{ii} = 1 - \sum_{j \in N(i)} d_{ij}$
Dynamic Influence in Competitive Environments (DICE)

- Set of individuals (nodes) $V = \{1,...,n\}$, set of social relationships $E$ (represented by links), set of ideas $I = \{I_1, ..., I_m\}$
- If $i$ has adopted idea $I_k$ and $j$ has adopted $I_r$:
  - Then $i$ switches to idea $I_r$ with probability $P_{ij}(k,r)$
  - Otherwise, $i$ adds $I_r$ to a list of known ideas that $i$ may adopt later
- If node $i$ does not meet any neighbor at time $t$, $i$ independently switches to a different known idea (i.e. $I_s$) with probability $P_i(k,s)$
Steady-state Distribution of Ideas

• **Question:** How many nodes will adopt each idea asymptotically?

• Define random variable $X_i(t) = $ idea held by node $i$ at time $t$; let $X(t) = (X_1(t), ..., X_n(t))$

• By definition of DICE, $X = \{X(t): 0 \leq t < \infty\}$ forms a Markov chain

• If stationary distribution exists, then it gives asymptotic behavior of $X$
Existence of Stationary Distribution

- Define $V_k = \text{set of nodes adopting idea } l_k \text{ at } t=0$
- The stationary distribution of $X$ exists and depends only on $V_1, ..., V_m$ if:
  - The probability $d_{ii} > 0$ for each $i$ in $V$ (every node avoids its neighbors with nonzero probability)
  - The probability $P_i(k,l) > 0$ for every $i, k, \text{ and } l$ (every node changes its mind with some probability)

- **Irreducibility:** $d_{ii} > 0$ and $P_i(k,l) > 0$ means nonzero probability of transitioning between any two states at each step

- **Aperiodicity:** $d_{ii} > 0$ and $P_i(k,k) > 0$ implies that state remains unchanged at each step with nonzero probability
Utilities Based on Stationary Distribution

• Define $\pi_i(k|V_1,\ldots,V_m)$ to be the probability that node $i$ adopts idea $I_k$ in the stationary distribution.

• The utility of the owner of $I_k$ is then the average number of individuals holding $I_k$ in steady-state, given as

$$U_k(V_1,\ldots,V_m) = \sum_{i\in V} \pi_i(k|V_1,\ldots,V_m)$$

• Uncertainty in social relationships $E$, represented as a distribution $P(E)$ over subsets of $E_0$, leads to average utility

$$U_k(V_1,\ldots,V_m) = \sum_{E\subseteq E_0} \sum_{i\in V} P(E) \pi_i(k|V_1,\ldots,V_m,E)$$

• How should the owners of $I_1,\ldots, I_m$ select $V_1,\ldots,V_m$ in order to maximize their utilities?
Leader-Follower Influence Game

• Two players, each with one idea (m=2)
• One player (leader) introduces idea first by choosing \(V_1\)
• Second player (follower) observes \(V_1\) and chooses \(V_2\)
• Adding a node to \(V_1\) or \(V_2\) incurs a cost \(c\)
• Leader’s utility function given by:

\[
U_1(V_1, V_2) = \sum_{E \subseteq E_0} \sum_{i \in V} \mathcal{P}(E) \pi_i(1|V_1, V_2, E) - c|V_1|
\]

• Follower’s utility function:

\[
U_2(V_1, V_2) = \sum_{E \subseteq E_0} \sum_{i \in V} \mathcal{P}(E) \pi_i(2|V_1, V_2, E) - c|V_2|
\]

• What are the optimal strategies for each player?
Maximizing Follower Utility

- Follower’s goal: Given $V_1$, choose $V_2$ to maximize $U_2(V_1, V_2)$
- Observation: Stationary distribution $\pi_i(k|V_1, V_2)$ depends only on ideas known to $i$ in steady state
- Node $i$ will know idea $I_k$ if and only if it is connected to a node in $V_k$
- Hence benefit of adding node $v$ to $V_2$ depends only on which nodes are connected to $v$

- This implies that $U_2(V_1, V_2)$ is submodular as a function of $V_2$
Maximizing Follower Utility – Derivation

• Follower’s goal: Given $V_1$, choose $V_2$ to maximize $U_2(V_1,V_2)$
• Observation: Stationary distribution $\pi_i(k|V_1,V_2)$ depends only on ideas known to i in steady state
• This depends on which connected components intersect $V_2$
• Benefit of adding a node $v$ from component $G_l$ to $V_2$ is therefore given by

$$U_2(V_1, V_2 + v) - U_2(V_1, V_2) = \begin{cases} \sum_{i \in G_l} \pi_i(2|V_1, v) - c, & V_2 \cap G_l = \emptyset \\ -c, & \text{else} \end{cases}$$

• Benefit is positive iff $V_2 \cap G_l \neq \emptyset$. Since $T \cap G_l \neq \emptyset$ implies $S \cap G_l \neq \emptyset$ when $S \subseteq T$, $U_2$ satisfies diminishing returns property and hence is submodular.
• In case of probabilistic topology, $U_2$ is nonnegative weighted sum of submodular functions, therefore still submodular
Solution Algorithms for Follower

- $U_2(V_1, V_2)$ can be maximized using a greedy approach
- Initialize $V_2 = \emptyset$
- At t-th iteration, add node $v$ that maximizes $U_2(V_1, V_2 \cup \{v\}) - U_2(V_1, V_2)$
- Algorithm terminates when $U_2(V_1, V_2 \cup \{v\}) \leq U_2(V_1, V_2)$ for all $v$

- **Theorem:** The greedy algorithm returns a set $\tilde{V}_2$ such that

$$U_2(V_1, V_2^*) - U_2(V_1, \tilde{V}_2) \leq c|V_2^*|$$

Where $V_2^*$ is the optimum selection for the follower
- **Proof:** Based on [Nemhauser 1978], Proposition 4.1
Maximizing the Leader’s Utility

- Leader chooses set of nodes $V_1$ based on follower’s best response, $V_2^*(V_1)$
- Let $\pi_i(2|I_1,I_2)$ denote the probability that $i$ holds $I_2$ in steady state, given that both $I_1$ and $I_2$ are known to $i$
- Let $G_l$ represent a strongly connected component of $V$
- For a given topology, the follower will select one node from $G_l$ and add it to the set $V_2$ if the benefit outweighs the cost:

$$\sum_{i\in G_l} \pi_i(2|I_1,I_2) > c.$$ 

- Hence the leader’s utility is given by

$$\sum \sum w_i - |V_1|c$$ 

where

$$w_i = \begin{cases} 
\pi_i(1|I_1,I_2), & \sum_{i\in G_l} \pi_i(2|I_1,I_2) > c \\
\pi_i(1|I_1), & \text{else}
\end{cases}$$
Solution Algorithms for Leader

• The incremental benefit from adding a node v from strongly connected component $G_l$ to leader set $V_1$ is then given by

$$U_1(V_1 + v, V_2^*(V_1 + v)) - U_1(V_1, V_2^*(V_1)) = \begin{cases} \sum_{i \in G_l} w_i - c, & V_1 \cap G_l = \emptyset \\ -c, & \text{else} \end{cases}$$

• If $S \subseteq T$, then $T \cap G_l = \emptyset$ implies $S \cap G_l = \emptyset$
• Hence incremental benefit from adding v to T is no more than incremental benefit from adding v to S
• $U_1(V_1, V_2^*(V_1))$ is therefore a submodular function of $V_1$, can be solved by algorithm analogous to follower solution algorithm
Simultaneous Influence Game

- Occurs when two ideas introduced at the same time
- Utility functions are the same, but $V_2$ chosen without observing $V_1$
- **Goal:** Find strategies such that neither player has incentive to deviate unilaterally (Nash equilibrium)

\[
V_1^* = \arg\max_{V_1} U_1(V_1, V_2^*)
\]

\[
V_2^* = \arg\max_{V_2} U_2(V_1^*, V_2)
\]
Finding Nash Equilibria

• Consider a network which is deterministic (no uncertainty) and strongly connected

• Strategies reduce to: introduce idea or not?

• Let \( B_k = \sum_{i \in G} \pi_i(k) \) and define H and H’ to be strategies of introducing and not introducing idea, respectively

• Game described by payoff matrix

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>H'</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>( B_1 - c, B_2 - c )</td>
<td>( n - c, 0 )</td>
</tr>
<tr>
<td>H'</td>
<td>( 0, n - c )</td>
<td>( 0, 0 )</td>
</tr>
</tbody>
</table>

• If \( B_k > c \) for \( k=1,2 \), then (H,H) is Nash equilibrium

• If \( c > n \), then (H’,H’) is Nash equilibrium

• \( B_k < c < n \), then two pure-strategy equilibria and one mixed strategy equilibrium
Simulation Study

- Simulations performed using Wiki-vote social network dataset
- Nodes represent Wikipedia users, an edge \((i,j)\) exists if user \(i\) has influenced user \(j\)
- Leader-follower game simulated
- Parameters chosen:
  - Number of ideas \(m = 2\)
  - \(d_{ii} = 0.5\) for all \(i\)
  - \(d_{ij} = 0.5/\text{degree}(i)\) for all edges \((i,j)\)
  - Probabilities \(P_{ij}(k,l)\) chosen uniformly at random from \([0,1]\)
  - Probabilities \(P_i(k,l)\) chosen uniformly at random from \([0,0.05]\)
  - Cost of choosing a seed node set to \(c=5\) or \(c=15\)
  - Number of nodes \(n\) ranges from 100 to 700
Simulation Results

- Utility of each player increases with # nodes
- $U_1 > U_2$
  - Follower must decide whether to compete for each connected component
  - Leader can choose strategy so that follower has no incentive to compete
- As cost increases, each player may choose not to enter the market (barriers to entry)
Future Work

• Modeling reinforcing ideas
  – A new idea is introduced in order to strengthen or expand influence of existing idea
• Choosing nodes to adopt ideas after the initial time step
• Other utility functions
  – Transient dynamics of influence
  – Nonlinear cost of introducing idea
Questions?