Maximizing the Spread of Influence through a Social Network

Kempe, Kleinberg, and Tardos (2003)
Paper Overview

- Marketing Strategies
- Non-Progressive Processes
- Generalization
- Experimental Results
  - Efficient Approximation
  - Example Network
  - Diffusion Models (Assumptions)
- Approach
- Paper Contributions
- Diffusion Models
- History
- Introduction
Introduction

Problem: Need to try to convince a people to adopt a product/behavior

Constraints: Let's say you have a limited budget

Approach:
Customers don’t necessarily make decisions in a vacuum.

Traditional Methods:

- Mass Marketing
- Direct Marketing
- Traditional Methods
Diffusion of ideas

Social networks play a fundamental role as a medium for the spread of information.

Dynamics of adoption is important.

Prior research work in diffusion processes:
- "Viral marketing" effects in the success of new products.
- Cascading failures in power systems.
- Adoption of various strategies in game theoretic settings.

Social networks play a fundamental role in the spread of information.

Diffusion of Ideas.
Activation of Nodes (Users) in a Directed Social Network Graph (G)

Initial Assumption: Progressive

• Linear Threshold
• Independent Cascades

Models Considered:

Diffusion Models
Diffusion models: linear threshold

Granovetter and Schelling

Model parameters:

- Edge weight \( w \) : Neighbor influence
- Threshold \( \theta \) : Tendency to adopt innovation

Diffusion proceeds in discrete time steps:

- Select initial set of active nodes \( A^0 \)
- Nodes active at time step \( t \) are also active at time step \( t + 1 \)
- At time step \( t + 1 \), check activation condition and update state \( A^t \)

Activation condition:

\[
\sum_{n \in \text{neighbor of } v} \theta^v \leq w_{v,n} \Rightarrow v \in A^{t+1}
\]
Stop!

Diffusion Models: Linear Threshold

Inactive Node
Active Node
Active neighbors
Threshold
Diffusion proceeds in discrete time steps:

- Select initial set of active nodes $A^0$
- Nodes active at time step $t+1$ are also active
  
  - Active node $v$ given a single chance to activate neighbor $w$
  
  - Multiple neighbors attempted, sequenced in random order:
    - $p_{1,2}$
    - $p_{1,3}$
    - $p_{3,4}$

Diffusion proceeds in discrete time steps.

Model Parameters:

- Edge weight $p_{v,w}$: Activation success probability

Multiple neighbors attempted, sequenced in random order:

- Activate neighbor $w$
Diffusion Models: Independent Cascades

Stop!

Unsuccessful attempt
Successful node
Newly active
Active Node
Inactive Node
Influence Maximization Problem

Domingos-Richardson framework: Find a k-node subset $A^0$ of maximum influence

Questions:
- Why is this important?
- What is a submodular function?
- How is influence defined?

THIS IS NP-HARD!
Paper Contributions

A general approach for reasoning about the performance guarantees of algorithms for these types of influence problems in social networks.

- First provable approximation guarantee to within a factor of \( \frac{1}{1 - \frac{1}{\epsilon}} \)
- Experimental validation: Comparison with popular heuristics
- Extensions:
  - Non-Progressive
  - Complex Marketing Actions
  - General Framework

Extensions:

- Experimental validation: Comparison with popular heuristics
- First provable approximation guarantee to within a factor of \( \frac{1}{1 - \frac{1}{\epsilon}} \)
Approximation Guarantees

We want to find a k-element set $A^*$ for which $\sigma(A^*)$ is maximized.

**Theorem:** For a non-negative, monotone submodular function $\sigma$, let $S$ be a set of size $k$ obtained by selecting elements one at a time, each time choosing an element that provides the largest marginal increase in the function value. Let $S^*$ be a set that maximizes the value of $\sigma$ over all k-element sets. Then $\sigma(S) \geq (1 - 1/e) \cdot \sigma(S^*)$; in other words, $S$ provides a $(1 - 1/e)$ approximation.

(Nemhauser, Wolsey, Fisher, 78)

Influence function $\sigma$ submodularity proof?
Approximation Guarantees

Assumptions:

- \( \sigma \) is non-negative
- Submodular function:
  \[
  \sigma(S \cup \{v\}) - \sigma(S) \geq \sigma(T \cup \{v\}) - \sigma(T)
  \]
- Monotone:
  \[
  \sigma(S \cup \{v\}) \geq \sigma(S)
  \]

\( \sigma \) is non-negative

Approximation Guarantees
Submodularity: Independent Cascades

Pre-flip coins for each pair of nodes (v, w) using $p^{v \cap w}$.

\[
(\forall) x^v \cdot \mathbb{P}[v] \bigcap_{x^v} = (\forall)^v
\]

\[
(1)^x - (\{\} \cup S)^x \geq (s)^x - (\{\} \cup S)^x
\]

$X'$: Set of nodes with live-edge paths from x

$X \cap R$ = $(s)^x - (\{\} \cup S)^x$

$R^x$ = set of edges (live/blinked)

$X$: State of edges (live/blinked)

$R^x$: Set of nodes with live-edge paths from x

(\forall)^x$ - (\{\} \cup T)^x \geq (s)^x - (\{\} \cup S)^x$
NP-Hardness: Independent Cascades

Optimization is NP-Hard

Set of $k$ nodes $A$ $\sigma(A) \geq n+k$

Edge between $x'$ and $u_j$ if $u_j$ in $x'$

Nodes $u_1, \ldots, u_n$

Sets $X_1, \ldots, X_m$

Reduce to set-cover problem

NP-Hardness: Independent Cascades
Claim: Trigger graph model = Linear threshold model

- $R(v, X)$: Set of nodes with live-edge paths from $x$
- $X$: State of edges (live/blocked)

Trigger Graph Model: Pick a live incoming edge for each node $v$ using $b_v$.

Slightly more complicated: Activation dependent on aggregate

Submodularity Proof: Linear Threshold
Submodularity Proof: Linear Threshold

State of model is the pair \((A_{t-1}, A_t)\)

Show that transition probabilities between states are same in both models:

- In both cases it is the same:
  - \(Pr(v \text{ becomes active at time } = t+1)\)
  - Distribution over states is always identical
  - Both models start in same state
  - State of model is the pair \((A_{t-1}, A_t)\)
NP-Hardness: Linear Threshold

Show to be equivalent to vertex-cover problem

Optimization is NP-Hard

NP-Hardness: Linear Threshold
Experimental Results

Run on co-authorship network from the complete list of papers in the high energy physics theory section of the e-print arXiv (2003)

10748 nodes (researchers), and 53000 edges (co-authorship)

Random selection - Greedy Algorithm VS
- Degree Centrality
- Distance Centrality
- Structural measures

Run on co-authorship network from the complete list of papers in the high energy physics theory section of the e-print arXiv (2003)
Experimental Results

- Linear Threshold Model: multiplicity of edges as weights. Weight of edge $(u \rightarrow v) = \frac{C_{uv}}{d_u}$.
- Independent Cascade Model:
  - Case 1: uniform probabilities $p$ on each edge (parallel edges?)
  - Case 2: edge from $u$ to $v$ has probability $\frac{1}{d_v}$ of activating $v$.

For $\sigma(A)$: Simulate the process 10000 times for each targeted set, re-choosing thresholds or edge outcomes pseudo-randomly from $[0, 1]$ every time.
Greedy algorithm outperforms
Degree centrality node
by over 40%
Distance centrality heuristic
by about 18%
Independent Cascade Model: Case 2
Independent Cascade Model: Case 1
Generalizations of Diffusion Models

- Generalized threshold
  - Define monotone threshold function \( f^v(S) \) such that node \( v \) is activated for \( f^v(S) \geq \theta^v \)

- Linear Threshold:
  - Define success probability \( p^v(u, S) \)
  - Independent cascade \( p^v(u, S) \) independent of \( S \)

These can be shown to be equivalent NP-hard to approximate in general

- General cascade
  - Independent cascade \( p^v(u, S) \)

- Linear Threshold:
  - \( f^v(S) \) such that node \( v \) is activated for \( f^v(S) \geq \theta^v \)

- Generalized threshold
Non-progressive processes

Nodes can switch back to inactivity

Can be reduced to progressive case

Nodes can switch back to inactivity

Theorem: The non-progressive influence maximization problem on $G$ over a time horizon $T$ is equivalent to the progressive influence maximization problem on the layered graph $G_T$. Node $v$ is active at time $t$ in the non-progressive process if and only if it is activated in the progressive process.
General Marketing Strategies

- Different nodes may respond to marketing actions in different ways.
- Marketing strategy vector $\mathbf{x}$
- Probability that node $v$ is activated by strategy $\mathbf{y}$

$(\mathbf{y})^\mathbf{a} - (\mathbf{v} + \mathbf{y})^\mathbf{a} \leq (\mathbf{x})^\mathbf{a} - (\mathbf{v} + \mathbf{x})^\mathbf{a}$

Non-decreasing

$m$: #marketing actions
Maximize this using a hill-climbing algorithm:

\[ ((x)^{n_{\eta}} - 1)^{\prod_{a}^{n_{\eta}}} \prod_{a}^{\eta} \prod_{x}{\bigwedge_{A}} = (x)^{6} \]

Expected Revenue from Final Activated Set \( A \)
Questions?