Information Dissemination in Social Networks under the Linear Threshold Model

Srinivasan Venkatramanan and Anurag Kumar

Department of Electrical Communication Engineering,
Indian Institute of Science,
Bangalore 560 012.
{vsrini,anurag} @ece.iisc.ernet.in

3 January, 2011

Outline of Tutorial

Outline of Talk

Social Networks

- A collection of individuals with pair-wise relationships between them
  - Naturally modelled as a graph
- Recently much interest in the context of information and communication technologies
  - World-wide-web link networks
  - Online social networks, email-contact networks
  - Citation networks and coauthorship networks
  - Forwarding in mobile opportunistic networks
- An important application area for "Network Science"

Mobile Wireless Opportunistic Networks

- Disruption (or Delay) Tolerant Networks (DTN)
  - People or vehicles carry mobile communication devices
  - Links form when such nodes come in "contact"
  - Data moves by *Store-and-Carry-Forward* (much like a postal system)
  - Suitable for some services, e.g., asynchronous messaging, content distribution (e.g., news)
  - "Pocket Switched Networks (PSN)"
- Mobility models
- Contact models
  - Not every physical contact may lead to link formation
  - Communication link formation can depend on existence of a *social link*
- Forwarding models
  - Not every link formation should lead to message transfer
  - Message forwarding should depend on who has been met and his/her relation to the destination in a social network
  - Recall: greedy geographical forwarding!
- Hence, need to bring in the social network model above the traditional DTN model

Influence Spread, Information Dissemination, Web Mining

- One of the interesting problems to consider is modeling and analyzing spread of information among the nodes in the network
- Applications
  - Identifying most influential nodes in the network for viral marketing
  - Characterize the spread of epidemics across the network
  - Propose suitable vaccination and quarantine schemes to impede the spread of epidemics
- Recent study facilitated by massive quantities of data available for empirical analysis, from social media sites.
- The recent DARPA Network Challenge [1] tested the ability of such massive ad-hoc teams to solve real-world problems, with applications to disaster management.

Role of Network Science: Characterising Graph Structure

- To characterize information spread, it is essential to know the underlying graph structure.
- In most scenarios, it is not practically feasible to obtain/work with the entire network data.
- Random graphs were used to model such networks in the past, but in reality, there are imperceptible rules rather than randomness that drives the network structure.
- Network science allows us to incorporate random graphs along with models of graph evolution (based on local interaction) to obtain a more realistic model of the network.

Outline of Talk

1. Social Networks and Applications
2. Mathematical Model
3. Recursive Analytical Expressions
4. Examples
5. The HILT Model: ODE Approximation
6. Final Remarks

The Social Network Model

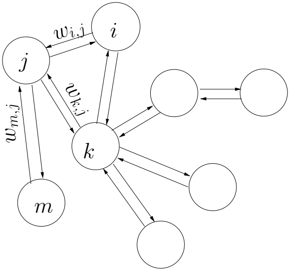
- The social network is modeled by a directed weighted graph $\mathcal{N} = (V, E)$
- Edge weights $w_{i,j}$ give a measure of influence of node i on node j
- Activation process
  - We begin with an initial set of active nodes $\mathcal{A}_0$
  - At each step k, new nodes are added to the active set A_k by the influence of their neighbours
  - Nodes once activated cannot become inactive (*progressive case*)
  - This goes on until we reach a *terminal set* A_S.

Activation Model: The Linear Threshold Model

- We require that $\sum_{i \neq j} w_{i,j} \leq 1$
- At the beginning, each node j randomly chooses a threshold Θ_j distributed uniformly over [0,1]
  - The thresholds are independent across the nodes
- At step k, a node j gets activated if, it had been inactive until step $k-1$ and

$$\sum_{i \in A_{k-1}} w_{i,j} \geq \Theta_j$$

  - i.e., the total influence of the active nodes into j exceeds the threshold of j



$\Sigma_i w_{i,j} \leq 1, \forall j$

Literature Survey

- Threshold models to explain collective behaviour were first put forward by Granovetter in [2], where he discussed the spread of binary decisions, among a group of rational agents (e.g. voting models)
- Domingos and Richardson [3] studied information diffusion under the viral marketing framework, and proposed the combinatorial optimization problem of finding the most influential nodes
- Kempe et al. [4] studied the influence spread problem under the LT and IC (independent cascade) model
  - Studied the problem of obtaining the optimal initial set of size K (in the sense of maximising the expected size of the final active set)
  - Showed that the problem is NP-hard
  - The objective function is monotone and submodular, and hence the greedy algorithm has an approximation factor of $(1 - \frac{1}{e})$
- Recent works include a general framework for cost effective outbreak detection [5], where they exploit the submodularity property to propose a lazy forward algorithm to achieve near greedy algorithm performance

Our Contributions

- We consider only the LT activation model
- Recursive analytical expressions for the expected information spread achieved by an initial set under the Linear Threshold model
- Interpretation in terms of acyclic paths in a certain DTMC
- Insights into information spread in some simple network topologies, such as the star, ring, and certain special mesh networks
- Homogeneous Influence LT model (HILT): by using Kurtz's theorem, we show that, an appropriately scaled version of the influence process, converges to a system of ODEs
  - Provide explicit formulas for trajectories followed by the influence process
  - Simple design formulas can be obtained

Notation

$\mathcal{N}$: weighted directed graph of the entire social network

$w_{i,j}$: edge weights of $\mathcal{N}$ indicating influence from i to j

$\mathbf{W}$: influence matrix with $w_{i,j}$ as entries

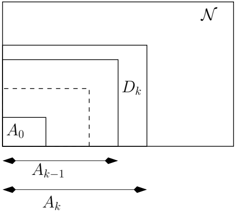
Θ_j: random threshold chosen by j uniformly from $[0,1]$

$b_j(A)$: $= \sum_{i \in A} w_{i,j}$, total influence into node j from set A

$\mathcal{A}_0$: Initial active set

A_k: Set of all active nodes at time step k

D_k: Set of infectious nodes at time step k



S: $= \arg\min_k \{A_k = A_{k-1}\}$

$g_j^{(\mathcal{N},\mathcal{A})}(k)$: $= \mathbb{P}^{(\mathcal{N})}(j \in D_k \,|\, \mathcal{A}_0 = \mathcal{A})$

$g_j^{(\mathcal{N},\mathcal{A})}$: $= \mathbb{P}^{(\mathcal{N},\mathcal{A})}(j \in A_S)$

$\sigma^{(\mathcal{N},\mathcal{A})}$: $= \mathbb{E}^{(\mathcal{N},\mathcal{A})}[|A_S|]$

Outline of Talk

Activation Probabilities $g_j^{(\mathcal{N},\mathcal{A})}(k)$ 1/3

- The following key lemma writes the activation probabilities $g_j^{(\mathcal{N},\mathcal{A}_0)}(k)$ in recursive form
- The i.i.d. uniformly distributed activation threshold plays a key role in the proof

Lemma

1. $j \in \mathcal{A}_0$,
   (a) $g_j^{(\mathcal{N},\mathcal{A}_0)}(0) = 1$
   (b) $g_j^{(\mathcal{N},\mathcal{A}_0)}(k) = 0$, *for all* $k > 0$
2. $j \notin \mathcal{A}_0$,
   (a) $g_j^{(\mathcal{N},\mathcal{A}_0)}(0) = 0$
   (b) $g_j^{(\mathcal{N},\mathcal{A}_0)}(k) = \sum_{l \in \mathcal{N}\setminus\{j\}} g_l^{(\mathcal{N}\setminus\{j\},\mathcal{A}_0)}(k-1)\, w_{l,j}$, *for all* $k > 0$

Proof.

- 1(a), 1(b) and 2(a) are obvious
- For 2(b), since $\mathcal{A}_0 \subset \mathcal{N}\backslash\{j\}$,

$$g_j^{(\mathcal{N},\mathcal{A}_0)}(k) = \mathbb{P}^{(\mathcal{N}\backslash\{j\},\mathcal{A}_0)}\Big(b_j(A_{k-2}) < \Theta_j \leq b_j(A_{k-1})\Big)$$

Since $D_{k-1} = A_{k-1} \backslash A_{k-2}$, and Θ_j is chosen uniformly from [0,1], we can write,

$$\begin{aligned} g_j^{(\mathcal{N},\mathcal{A}_0)}(k) &= \mathbb{E}^{(\mathcal{N}\backslash\{j\},\mathcal{A}_0)}[b_j(D_{k-1})] \\ &= \mathbb{E}^{(\mathcal{N}\backslash\{j\},\mathcal{A}_0)}\Big[\sum_{l\in\mathcal{N}\backslash\{j\}} \mathrm{I}_{\{l\in D_{k-1}\}} w_{l,j}\Big] \\ &= \sum_{l\in\mathcal{N}\backslash\{j\}} g_l^{(\mathcal{N}\backslash\{j\},\mathcal{A}_0)}(k-1)\ w_{l,j} \end{aligned}$$

□

Activation Probabilities $g_j^{(\mathcal{N},\mathcal{A})}(k)$ 3/3

- The previous result yields

$$g_j^{(\mathcal{N},i)}(k) = \sum_{l_1 \neq i,j} \sum_{l_2 \neq i,j,l_1} \cdots \sum_{l_{k-1} \neq i,j,l_1,l_2,\ldots l_{k-2}} w_{i,l_1} w_{l_1,l_2} \ldots w_{l_{k-1},j}$$

  i.e., starting with $A_0 = \{i\}$, the probability that j is activated at Step k is the sum of the product of path-weights along all k-hop acyclic paths starting at i and ending at j

- Further, $g_j^{(\mathcal{N},\mathcal{A})}(k)$ is just the sum of the weights of such paths that *avoid* the set $A_0 = A$ (except for the initial node)

Total Expected Activation due to $\{i\}$: $\sigma^{(\mathcal{N},i)}$

Theorem

Given a social network $\mathcal{N}$, with influence matrix $\mathbf{W}$, the total influence of any node i in the network under the LT model is given by

$$\sigma^{(\mathcal{N},i)} = 1 + \sum_{j \in \mathcal{N} \setminus \{i\}} w_{i,j} \sigma^{(\mathcal{N} \setminus \{i\}, j)} \tag{1}$$

Proof.

Note that,

$$\sigma^{(\mathcal{N},i)} = \sum_{k=0}^{\infty} \sum_{j \in \mathcal{N}} g_j^{(\mathcal{N},i)}(k)$$

In the above expression, by substituting using the previous lemma recursively and rearranging summations, we get the theorem. □

Total Expected Activation due to Set $\mathcal{A}_0$: $\sigma^{(\mathcal{N},\mathcal{A}_0)}$

Theorem

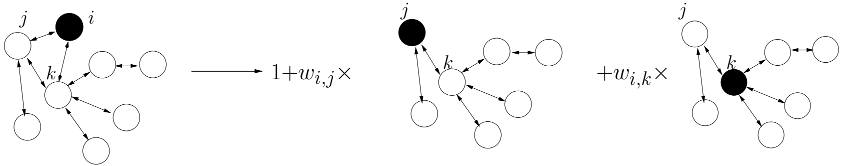
Given a network $\mathcal{N}$ with influence matrix $\mathbf{W}$ *and an initial set $\mathcal{A}_0$, for all $i \in \mathcal{A}_0$, define sub-networks $\mathcal{N}_i^{\mathcal{A}_0} = \{\mathcal{N} \backslash \mathcal{A}_0\} \cup \{i\}$.*
Then the expected influence of the initial set $\mathcal{A}_0$ is given by,

$$\sigma^{(\mathcal{N},\mathcal{A}_0)} = \sum_{i \in \mathcal{A}_0} \sigma^{(\mathcal{N}_i^{\mathcal{A}_0},i)} \tag{2}$$

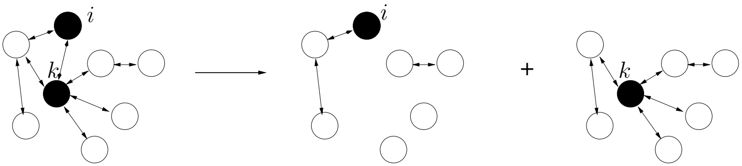
Proof.

The proof follows by substituting the $g_j^{(\mathcal{N},\mathcal{A}_0)}(k)$ expressions and noting that the edge weights $\{w_{i,j}, j \in \mathcal{A}_0\}$ do not have any effect on $\sigma^{(\mathcal{N},\mathcal{A}_0)}$. This allows us to split the problem of evaluating influences sub-problems each involving only one node from $\mathcal{A}_0$. □

Illustrations of the Recursive Calculations for $\sigma^{(\mathcal{N},\mathcal{A})}$



Singleton Initial Set $\{i\}$

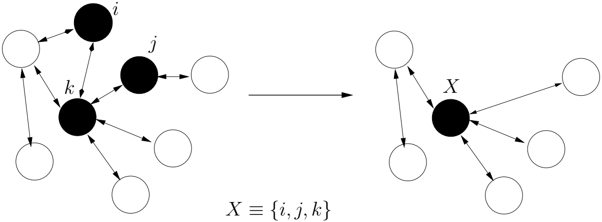


Initial Set $\mathcal{A}_0 = \{i, k\}$

Replacing an Initial Set with a "Supernode"

Theorem

Given a network $\mathcal{N}$, with influence matrix $\mathbf{W}$, an initial set $\mathcal{A}_0$ can be replaced by a supernode X as shown below. Then, $\sigma^{(\mathcal{N},\mathcal{A}_0)}$ is equal to the influence of X in the modified network, with X being counted as $|X|$ nodes instead of 1.



$$X \equiv \{i, j, k\}$$

$$\forall v \notin X,$$
$$w_{X,v} = w_{i,v} + w_{j,v} + w_{k,v}$$
$$w_{v,X} = w_{v,i} + w_{v,j} + w_{v,k}$$

Replacing the initial set with a supernode

A Discrete Time Markov Chain

- $\mathbf{P} := \mathbf{W}^T$ is a row substochastic matrix with zero diagonal terms
  - Make stochastic by additing $p_{j,j}$ values, appropriately
- Interpret $\mathbf{P}$ as the transition matrix of a DTMC $\{X_k\}$, obtained by reversing the edges in the social network
- Define

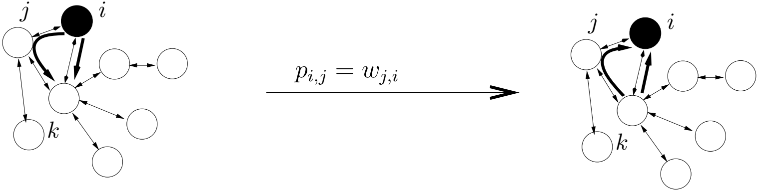
$$\begin{aligned} c_k(j \to \mathcal{A}_0) \\ := \quad \mathbb{P}(\{X_m \in \mathcal{N} \backslash \mathcal{A}_0, 0 \leq m < k, X_k \in \mathcal{A}_0, X_0 \neq X_1 \neq \cdots \neq X_k\} \big| X_0 = j) \end{aligned}$$

  the probability of the DTMC reaching $\mathcal{A}_0$ for the first time via a k step acyclic path, given that $X_0 = j$
- Also, define

$$c(j \to \mathcal{A}_0) := \sum_{k=0}^{\infty} c_k(j \to \mathcal{A}_0)$$

  i.e., given that we start from state j, the probability of hitting the set $\mathcal{A}_0$ for the first time through an acyclic path

Influence Spread Analysis via the DTMC



$$g_k^{(\mathcal{N},i)} = w_{i,k} + w_{i,j} \times w_{j,k}$$

Social Network

$$c(k \to i) = p_{k,i} + p_{k,j} \times p_{j,i}$$

Equivalent Markov Chain

- In general, it can be shown that

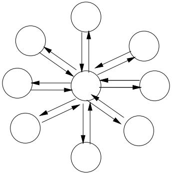
$$\begin{aligned} g_j^{(\mathcal{N},\mathcal{A}_0)}(k) &= c_k(j \to \mathcal{A}_0) \\ g_j^{(\mathcal{N},\mathcal{A}_0)} &= c(j \to \mathcal{A}_0) \end{aligned}$$

- This viewpoint leads to an alternative proof of the monotonicity and submodularity of $\sigma^{(\mathcal{N},\mathcal{A}_0)}$

Outline of Talk

Star Topology

- Consider the star topology network $\mathcal{N}$ with N nodes including the hub
  - e.g., a central authority through which all influence/information must be disseminated
- $\alpha(\leq 1)$: influence of the hub on peripheral nodes
- $\beta(\leq \frac{1}{N-1})$: influence of each peripheral node on the hub



Star topology

Star Topology: Analysis using Recursive Expressions

- What is the optimal initial set of size K?
- Let $H(K)$: set with one hub node and $K-1$ peripheral nodes.
- $\tilde{H}(K)$: set with K peripheral nodes
- The recursive expressions yield

$$\sigma^{(\mathcal{N},H(K))} = K + \alpha(N-K)$$

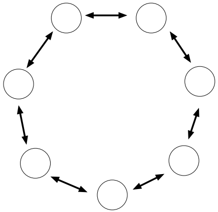
$$\sigma^{(\mathcal{N},\tilde{H}(K))} = K + K\beta\big(1+\alpha(N-K-1)\big)$$

- There exists $\alpha^\star$, such that $\sigma^{(\mathcal{N},\tilde{H}(K))} > \sigma^{(\mathcal{N},H(K))}$ for $\alpha < \alpha^\star$, i.e., it is optimal to exclude the hub from the initial set

$$\alpha^\star = \frac{K}{(N-K)\frac{1}{\beta} - K(N-K-1)}$$

Ring Topology

- Consider the ring topology $\mathcal{N}$ with N homogeneous nodes
- $\alpha \leq 0.5$ denotes the influence of any node on each of its neighbours



Ring topology

Ring Topology: Analysis using Recursive Expressions

- Let $A(K)$ be an initial set of K nodes
- $(a_1, a_2, \ldots a_K)$: Indices of the chosen nodes in the ring
- $(l_1, l_2, \ldots, l_K)$ be the number of nodes in between the chosen nodes in the ring, i.e. $l_1 = a_2 - a_1 - 1$, $l_2 = a_3 - a_2 - 1$ and so on
- Then the earlier analysis yields

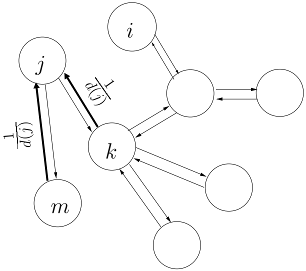
$$\sigma^{(\mathcal{N}, A(K))} = K + 2\frac{\alpha}{1-\alpha}\left(K - \sum_{i=1}^{K} \alpha^{l_i}\right)$$

- In order to maximize the influence, we see that the K nodes should be distributed uniformly over the ring
  - Argument via majorisation and Shur convexity
- Also, with $\alpha = 0.5$, for large N and small K, the influence of $A(K)$ grows as $3K$

Degree Based Model

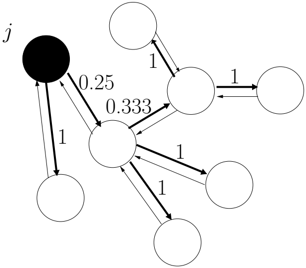
- Start with an undirected graph without cycles, whose adjacency matrix is given by **M**
- d_j : the degree of node j
- Define W as follows

$$w_{i,j} = m_{i,j}/d_j$$



Node Degree Based Model

Degree Based Model: Example Calculation of Influence



$$\sigma^{(\mathcal{N},j)} = 1 + 1 + 0.25(1 + 1 + 1 + 0.333(1 + 1 + 1)) = 3 = d_j + 1$$

Degree Based Model: General Result on Influence

Theorem

Consider an acyclic undirected network $\mathcal{N}$ with degree based influence weights (as defined). Then, for any node $i \in \mathcal{N}$,

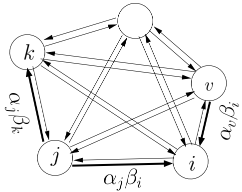
$$\sigma^{(\mathcal{N},i)} = d_i + 1$$

- Note that the local property (a node's degree) governs its global property (influence on the entire network)
- The most influential node is the one with the highest degree, and the set of top-K influential nodes has high correlation with the set of top-K high degree nodes

Uniform Influence and Susceptance Model

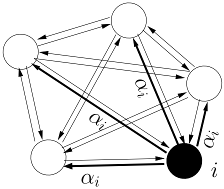
- **UISLT**: Uniform Influence and Susceptance LT Model
  - Two parameters α_i and β_i associated with the node i
  - Measures of the level of influence and susceptance of the node i
- Define $w_{i,j} = \alpha_i \times \beta_j$, for all$j \neq i$
- $\alpha_i, \beta_i, 1 \leq i \leq N$, are chosen such that, $\sum_{j \neq i} w_{j,i} \leq 1$

Uniform Models



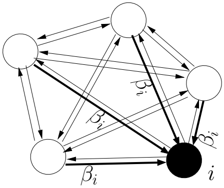
$$\Sigma_{j\neq i}\,\alpha_j \leq \frac{1}{\beta_i}\ ,\ \forall i$$

Uniform Influence and Susceptance (UISLT) model



$$\Sigma_{i\neq j}\,\alpha_i \leq 1\ ,\ \forall j$$

Uniform Influence (UILT) model



$$\beta_i \leq \frac{1}{|\mathcal{N}|-1}\ ,\ \forall i$$

Uniform Susceptance (USLT) model

USLT and UILT Models: Optimal Initial Set

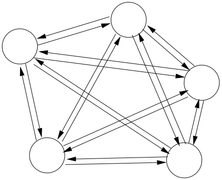
- **USLT**: The optimal $\mathcal{A}_0$ in this case, is found to be the set of K nodes with least β_i values, i.e. *the least susceptible ones*
- **UILT**: The optimal $\mathcal{A}_0$ in this case, is found to be the set of K nodes with highest α_i values, i.e. *the most influential ones*
- For these models, this turns out to be equivalent to picking the K nodes with the top-K values of the stationary p.m.f. of $\mathbf{P} = \mathbf{W}^T$

Comments on the PageRank Algorithm

- Google's PageRank algorithm, in its simplest form, ranks pages by calculating the stationary probability over the "influence" graph of web-pages, by assuming a *random surfer model*
- Thus, PageRank yields the optimum solution for UILT and USLT
- We have found general UISLT examples where PageRank can perform poorly

Outline of Talk

Homogeneous Influence Linear Threshold (HILT) Model



All edge weights $= \gamma$, $\gamma \leq \frac{1}{|\mathcal{N}|-1}$

HILT model

- Consider a network of N nodes with a complete graph
- Each edge carries an equal weight $\gamma \leq \frac{1}{N-1}$.
- Define $h_\gamma^{(N)}(k) := \sigma^{(\mathcal{N},\mathcal{A})}$, for all $\mathcal{A}$ of size k
- By using the analytical expressions derived earlier, we obtain

$$h_\gamma^{(N)}(k) = k[1 + (N-k)\gamma[1 + (N-k-1)\gamma[1 + \cdots$$

- We would like to approximate the evolution of the influence process
- We study this via a fluid limit in the large N regime

The Activation Markov Chain

- Let $A(k)$ and $D(k)$ be the sizes of the active and infectious sets at step $k \geq 0$
- Define $B(k) = A(k-1)$
  - $B(k)$ is the size of the set of active nodes that have already "exercised their influence" (no longer "infectious")
- Consider the process $(B(k), D(k)), k \geq 0$, with $B(0) = 0$ and $D(0) = |A(0)|$
- Using the uniformly distributed i.i.d. activation thresholds

$$B(k+1) = B(k) + D(k)$$

$$\begin{aligned} & P(D(k+1) = l | (B(i), D(i)), 0 \leq i \leq k-1, (B(k), D(k)) = (b, d)) \\ = & \binom{N-b-d}{\ell} \left(\frac{\gamma d}{1-\gamma b}\right)^{\ell} \left(1 - \frac{\gamma d}{1-\gamma b}\right)^{(N-b-d-\ell)} \\ = & P(D(k+1) = \ell | B(k) = b, D(k) = d) \end{aligned}$$

Scaling the Process $(B(k), D(k))$

- Define a scaled Markov process $(B^N(k), D^N(k))$
  - Can be viewed as evolving over "minislots" of width $\frac{1}{N}$, whereas the original process evolves over unit steps
- In each minislot, each node in $D^N(k)$ decides to spread its influence with probability $\frac{1}{N}$ or defer with probability $1 - \frac{1}{N}$
- In the former case, it contributes influence of γ and then moves to the set $B^N(k+1)$, else it stays in $D^N(k+1)$ set
- Defining $\tilde{B}^N(k) = \frac{B^N(k)}{N}$, $\tilde{D}^N(k) = \frac{D^N(k)}{N}$, the evolution equations can be written as follows:

$$\tilde{B}^N(k+1) = \tilde{B}^N(k) + \frac{\tilde{D}^N(k)}{N} + \tilde{Y}^N(k+1)$$

$$\tilde{D}^N(k+1) = \frac{N-1}{N}\tilde{D}^N(k) + \frac{\Gamma\frac{\tilde{D}^N(k)}{N}}{1 - \Gamma\tilde{B}^N(k)}(N - \tilde{B}^N(k) - \tilde{D}^N(k)) + \tilde{Z}^N(k+1)$$

where we have asserted that γ scales with N as $\frac{\Gamma}{N}$

ODE approximation to HILT model

- The hypotheses of Kurtz's theorem [6] (uniform convergence of drift to a Lipschitz function, and condition on "noise" variance)
- Applying the theorem we conclude that for every $\epsilon > 0$ and $T > 0$

$$P\left(\sup_{0 \le t \le T} \left|\left| (\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor)) - \big(b(t), d(t)\big) \right|\right| > \epsilon \right) \overset{N \to \infty}{\to} 0$$

- where $(b(t), d(t))$ is the unique solution of the ODE (with given $(b(0), d(0))$)

$$\dot{b}(t) = d(t)$$

$$\dot{d(t)} = -d(t) + \frac{\Gamma d(t)}{1 - \Gamma b(t)}(1 - b(t) - d(t))$$

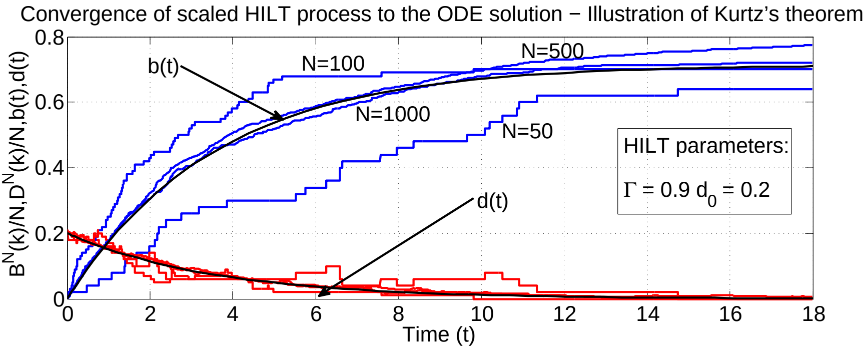
- Solving the ODE for $b(0) = 0, d(0) = d_0$, defining $r = 1 - \Gamma - \Gamma d_0$, we obtain

$$b(t) = \frac{d_0}{r} - \frac{d_0}{r} e^{-rt}$$

$$d(t) = d_0 e^{-rt}$$

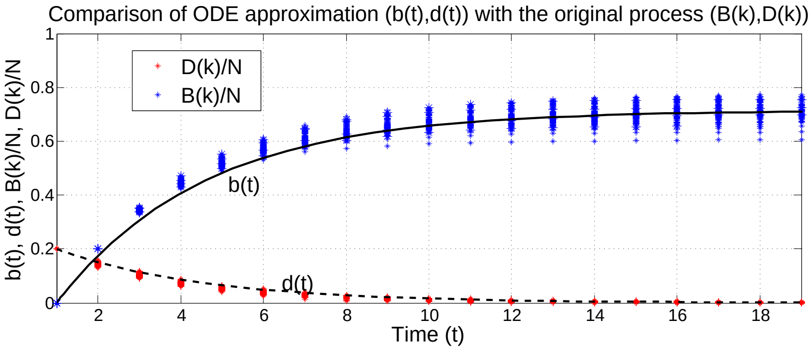
Convergence of $\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor)$ to $(b(t), d(t))$

- The trajectory of the ODE is plotted along with sample paths of the scaled process $\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor)$ for N=50,100,500,1000



Approximation of the Original Process by the ODE

- Simulation of the original process $\left(\frac{B(k)}{N}, \frac{D(k)}{N}\right)$ compared with the ODE solution $(b(t), d(t))$



HILT Network: Applications of the ODE Solution

- We have an approximation to the trajectory of the activation process
- Given that we start with d_0 fraction of nodes in the infectious set, the final fraction of activated nodes (b_∞) is $\frac{d_0}{r}$ where $r = 1 - \Gamma + \Gamma d_0$.
- Given b_∞, the optimal d_0 to be chosen is given by,

$$d_0 = \frac{b_\infty(1 - \Gamma)}{1 - b_\infty \Gamma}$$

  - For large N, with $\Gamma < 1$ we cannot influence the entire population (i.e., $b_\infty = 1$) unless we start with the entire population active (i.e., $d_0 = 1$)
  - But if $\Gamma = 1$ then $b_\infty = 1$ provided $d_0 > 0$
- We can also answer questions on time-constrained influence spread, where we are interested in the value of $a(T) = b(T) + d(T)$, given finite T
  - Time T at which $a(T) = \alpha$, given d_0 and Γ
  - Optimal d_0 to start with so that $a(T) \geq \alpha$

Outline of Talk

Other Related Work

- We have used the Acyclic Path Probabilities interpretation to provide an alternative proof to the monotonicity and submodularity of the function $\sigma^{(\mathcal{N},\mathcal{A}_0)}$
- Using the recursive expression, we have identified special cases where the PageRank algorithm is optimal for the influence maximization problem
- We have also given a heuristic algorithm (G1-sieving) for *Influence maximization* problem, based on insights from the recursive expression, which performs on par with the Greedy algorithm
- We have extended the work done on fluid limit for HILT model to multi-class scenario and derived the approximating system of ODE

Future Work

- We can consider mobile nodes (as in a Delay-Tolerant Network), where nodes are able to transfer influence only on meeting
- The problem can also be generalized to edge weights and threshold functions varying with time
- Finally, we can also study information dissemination with different activation processes, and on more generic networks to gain insights into the underlying mechanisms of information dissemination

Darpa Network Challenge 2010.
https://networkchallenge.darpa.mil/, 2010.

Mark Granovetter.
Threshold models of collective behaviour.
American Journal of Sociology, 1978.

Matt Richardson and Pedro Domingos.
Mining knowledge-sharing sites for viral marketing.
In *ACM SIGKDD*, 2002.

D. Kempe, J. Kleinberg, and E. Tardos.
Maximizing the spread of influence through a social network.
In *ACM SIGKDD*, 2003.

References II

C. Guestrin C. Faloutsos J. Vanbriesen J. Leskovec, A. Krause and N. Glance.
Cost-effective outbreak detection in networks.
In *ACM SIGKDD*, 2007.

Thomas G. Kurtz.
Solutions of ordinary differential equations as limits of pure jump markov processes, 1970.