Co-evolution of Popularity and Content Spread in Mobile Opportunistic Networks

Srinivasan Venkatramanan
Joint work with Prof. Anurag Kumar

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

12 December 2012

Outline

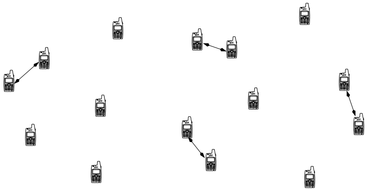
1. Content Spread on Mobile Opportunistic Networks
2. Evolution of Content Popularity
3. Co-evolution of Popularity and Content Spread
4. An Optimization Problem

Outline

1. Content Spread on Mobile Opportunistic Networks
2. Evolution of Content Popularity
3. Co-evolution of Popularity and Content Spread
4. An Optimization Problem

Mobile Opportunistic Networks (MON)

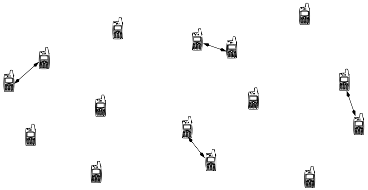
- Proliferation of personal mobile devices
  - with peer-to-peer radio interfaces (Bluetooth, and WiFi)



- Growing demand for high-bandwidth content on mobile devices
  - Music, images and video clips

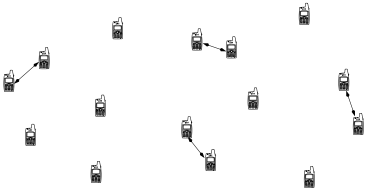
Mobile Opportunistic Networks (MON)

- Proliferation of personal mobile devices
  - with peer-to-peer radio interfaces (Bluetooth, and WiFi)



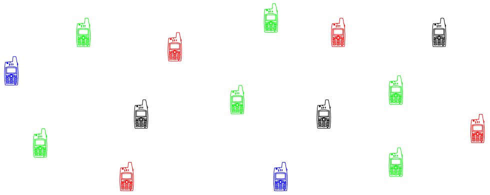
- Growing demand for high-bandwidth content on mobile devices
  - Music, images and video clips
- A single item of content may be of **interest** to several co-located users
  - e.g., slides of a conference keynote/ course lecture
  - e.g., a just released popular video song spreading in a campus setting

Mobile Opportunistic Networks (MON)



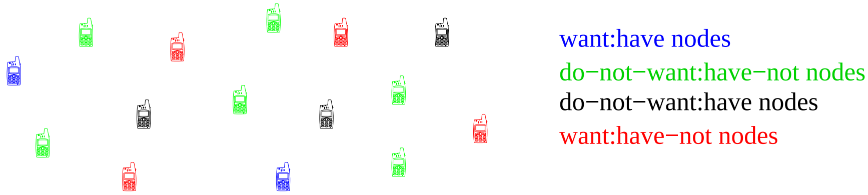
- Proliferation of personal mobile devices
  - with peer-to-peer radio interfaces (Bluetooth, and WiFi)
- Growing demand for high-bandwidth content on mobile devices
  - Music, images and video clips
- A single item of content may be of **interest** to several co-located users
  - e.g., slides of a conference keynote/ course lecture
  - e.g., a just released popular video song spreading in a campus setting
- The item can be forwarded between devices when their peer-to-peer radio interfaces make contact
  - Akin to **epidemic** spread
  - A multi-hop opportunistic mobile network
- Can provide an approach for *delay tolerant* networking

Content Popularity and Dissemination 1/2

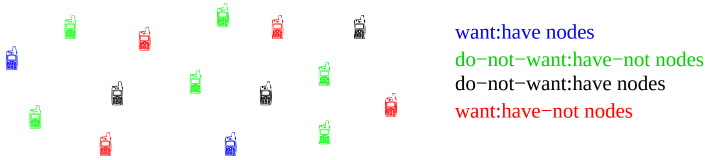


- A fixed population of nodes; size N
- A single item of content needs to reach certain *interested* nodes
  - Also called *want* nodes, or *destination* nodes
  - Some want nodes are given the content initially
- The set of want nodes grows, as more nodes express their interest in the content
  - Infrastructure might broadcast the number of nodes who want the content
  - Meetings with want nodes can help to spread the popularity

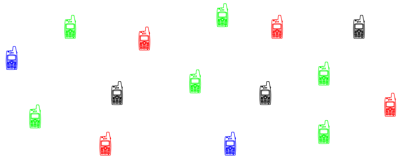
Content Popularity and Dissemination 1/2



- A fixed population of nodes; size N
- A single item of content needs to reach certain *interested* nodes
  - Also called *want* nodes, or *destination* nodes
  - Some want nodes are given the content initially
- The set of want nodes grows, as more nodes express their interest in the content
  - Infrastructure might broadcast the number of nodes who want the content
  - Meetings with want nodes can help to spread the popularity
- *Do-not-want* nodes could help in forwarding
  - Also called *relay* nodes

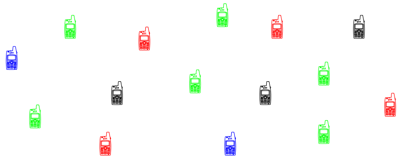


- Random meeting instants between pairs of nodes
  - Instantaneous copying is assumed at a meeting instant



want:have nodes
do−not−want:have−not nodes
do−not−want:have nodes
want:have−not nodes

- Random meeting instants between pairs of nodes
  - Instantaneous copying is assumed at a meeting instant
- At a meeting instant where a have node meets a have-not node
  - A "copy" or "do-not-copy" decision has to be made



want:have nodes

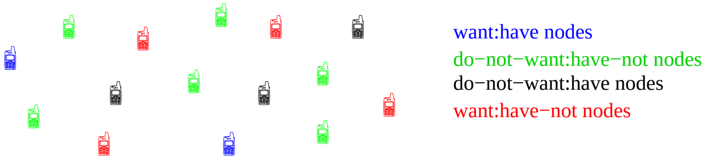
do−not−want:have−not nodes

do−not−want:have nodes

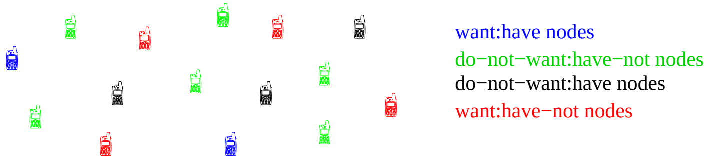
want:have−not nodes

- Random meeting instants between pairs of nodes
  - Instantaneous copying is assumed at a meeting instant
- At a meeting instant where a have node meets a have-not node
  - A “copy” or “do-not-copy” decision has to be made
- Delay-energy tradeoff
  - Copying to relays (do-not-wants) (i) promotes early delivery to destinations (wants), and (ii) they could become wants themselves

Content Popularity and Dissemination 2/2



- Random meeting instants between pairs of nodes
  - Instantaneous copying is assumed at a meeting instant
- At a meeting instant where a have node meets a have-not node
  - A “copy” or “do-not-copy” decision has to be made
- Delay-energy tradeoff
  - Copying to relays (do-not-wants) (i) promotes early delivery to destinations (wants), and (ii) they could become wants themselves
  - Incurs additional transmission energy; may need to pay an incentive to relays



- Random meeting instants between pairs of nodes
  - Instantaneous copying is assumed at a meeting instant
- At a meeting instant where a have node meets a have-not node
  - A "copy" or "do-not-copy" decision has to be made
- Delay-energy tradeoff
  - Copying to relays (do-not-wants) (i) promotes early delivery to destinations (wants), and (ii) they could become wants themselves
  - Incurs additional transmission energy; may need to pay an incentive to relays
- Objective: Quickly spread content to a large fraction of want nodes while keeping the residual number of do-not-want:have nodes small

Co-evolution of Popularity and Content Spread

- Model of the increase in the popularity of the content
  - Models for influence spread; used in viral marketing
  - The Linear Threshold Model[1]
- Model for content spread
  - Controlled epidemic model for the spread of a content in a mobile opportunistic network[2]
- Here we combine the two models
  - The popularity spread evolves autonomously
    - being driven only by the population of *want* nodes
  - The controlled epidemic model for content spread has the popularity model as "time varying parameters"

---

[1] David Kempe, Jon Kleinberg and Éva Tardos, "Maximizing the Spread of Influence in a Social Network", in Proceedings of KDD 2003

[2] Chandramani Singh, Anurag Kumar, Rajesh Sundaresan and Eitan Altman, "Optimal Forwarding in Delay Tolerant Networks with Multiple Destinations", in Proceedings of WiOpt 2011, and the references therein.

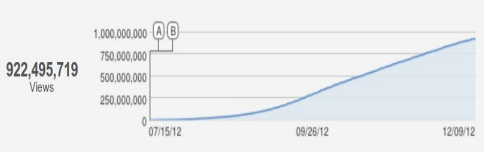
Outline

1. Content Spread on Mobile Opportunistic Networks
2. Evolution of Content Popularity
3. Co-evolution of Popularity and Content Spread
4. An Optimization Problem

Viral Media on the Internet

Viral Media on the Internet

PSY – Gangnam Style (#1 Most Viewed on Youtube)



Rebecca Black – Friday
(167 million views, 3.1 million dislikes)

Why this Kolaveri di (65 million views)

Source: YouTube Trends

Product/Service Adoption

Smartphones

Browsers

Social Networks

Diffusion of Innovation

- Diffusion of Innovation - proposed by Everett Rogers in 1962
- Threshold models of Collective Behavior - Mark Granovetter, 1978.
- Malcolm Gladwell - Tipping Point (2000)

Cui bono (To whose benefit?)

Cui bono (To whose benefit?)

Content creators

- Content creation - no longer restricted to media powerhouses
- To reach out to a wider audience
- To gain popularity and establish the brand

Cui bono (To whose benefit?)

Content creators

- Content creation - no longer restricted to media powerhouses
- To reach out to a wider audience
- To gain popularity and establish the brand

Content curators

- e.g. Yahoo! homepage, Google Reader
- To identify trending topics and serve their customers

Cui bono (To whose benefit?)

Content creators

- Content creation - no longer restricted to media powerhouses
- To reach out to a wider audience
- To gain popularity and establish the brand

Content curators

- e.g. Yahoo! homepage, Google Reader
- To identify trending topics and serve their customers

Content distributors

- Hosts of user-published content - e.g. Blogger, Soundcloud, Youtube
- To identify the user demand
- Appropriately adapt caching and delivery mechanisms

Cui bono (To whose benefit?)

Content creators

- Content creation - no longer restricted to media powerhouses
- To reach out to a wider audience
- To gain popularity and establish the brand

Content curators

- e.g. Yahoo! homepage, Google Reader
- To identify trending topics and serve their customers

Content distributors

- Hosts of user-published content - e.g. Blogger, Soundcloud, Youtube
- To identify the user demand
- Appropriately adapt caching and delivery mechanisms

Advertisers

- "Backbone of the Internet economy"
- To identify contents to "piggyback" their advertisements onto

Models of Influence spread

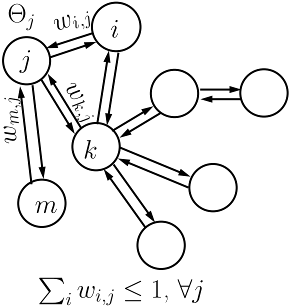
Models of Influence spread

Two most common influence spread models: Linear Threshold (LT) model and Independent Cascade (IC) model [KKT'03]

Two most common influence spread models: Linear Threshold (LT) model and Independent Cascade (IC) model [KKT'03]

Linear Threshold model

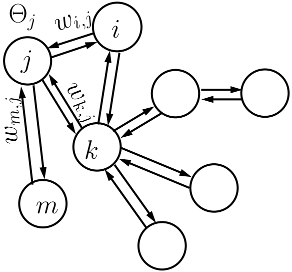
- The social network is modeled by a directed weighted graph $\mathcal{N} = (V, E)$, with edge weights $w_{i,j}$, $\sum_i w_{i,j} \leq 1$



Two most common influence spread models: Linear Threshold (LT) model and Independent Cascade (IC) model [KKT'03]

Linear Threshold model

- The social network is modeled by a directed weighted graph $\mathcal{N} = (V, E)$, with edge weights $w_{i,j}$, $\sum_i w_{i,j} \leq 1$
- At the beginning, each node j independently chooses a random threshold $\Theta_j \overset{dist.}{=} U[0,1]$



Models of Influence spread

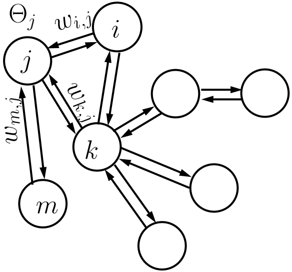
Two most common influence spread models: Linear Threshold (LT) model and Independent Cascade (IC) model [KKT'03]

Linear Threshold model

- The social network is modeled by a directed weighted graph $\mathcal{N} = (V, E)$, with edge weights $w_{i,j}$, $\sum_i w_{i,j} \leq 1$
- At the beginning, each node j independently chooses a random threshold $\Theta_j \stackrel{dist.}{=} U[0,1]$
- At step k, a node j gets activated if, it had been inactive until step $k-1$ and

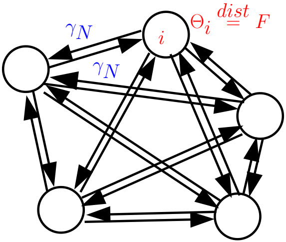
$$\sum_{i \in A_{k-1}} w_{i,j} \geq \Theta_j$$

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



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

HILT: A Special Case of the Linear Threshold Model[3]

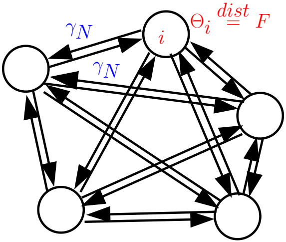


$\gamma_N = \frac{\Gamma}{N-1}$, $\Gamma \leq 1$

- Homogeneous Influence Linear Threshold model (HILT Model)

[3] David Kempe, Jon Kleinberg and Éva Tardos,"Maximizing the Spread of Influence in a Social Network", in Proceedings of KDD 2003

HILT: A Special Case of the Linear Threshold Model[3]



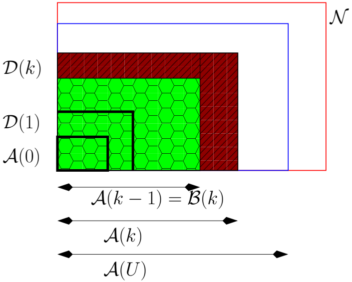
$\gamma_N = \frac{\Gamma}{N-1}$, $\Gamma \leq 1$

- Homogeneous Influence Linear Threshold model (HILT Model)
- Complete directed graph on $\mathcal{N}$, $N = |\mathcal{N}|$
- Influence edge weights $\gamma_N = \frac{\Gamma}{N-1}$
- Random independent thresholds $\Theta_i \stackrel{dist}{=} F$, $i \in \mathcal{N}$
  - Models "susceptibility" to influence
  - Chosen at the beginning by each user

[3] David Kempe, Jon Kleinberg and Éva Tardos,"Maximizing the Spread of Influence in a Social Network", in Proceedings of KDD 2003

Evolution of the HILT Model

- *Destinations (Wants)* $\mathcal{A}(k)$: Nodes interested in the content
- *Relays (Do Not Wants)* $\mathcal{N} \backslash \mathcal{A}(k)$: Nodes not yet interested



- $\mathcal{A}(0)$: Initial set of destinations
- A relay $j \notin \mathcal{A}(k-1)$ becomes a destination at time k, if

$$\gamma_N |\mathcal{A}(k-1)| \geq \Theta_j$$

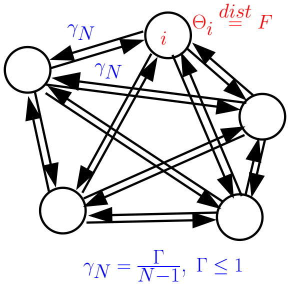
- $U := \arg\min\{k : \mathcal{A}(k-1) = \mathcal{A}(k)\}$
- $\mathcal{A}(U)$ is the terminal set of destinations

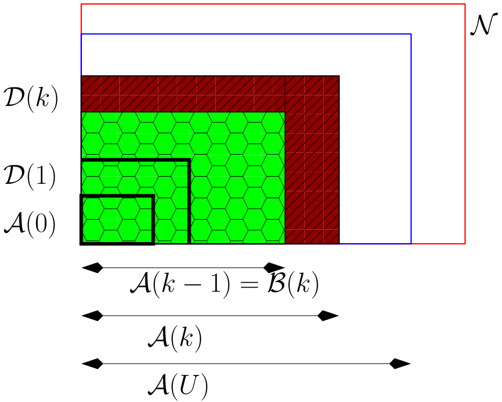
Analysis of the HILT model

- Computation of the expected terminal influence in the LT model has been shown to be NP-hard
- For the HILT model, we have provided a closed form analytical expressions for the expected terminal influence[4]
- We wish to study the evolution of demand
  - Hence, we wish to model the expected *trajectory of influence*
- We resort to taking a fluid limit of the HILT Markov chain
  - by using Kurtz's theorem[5]: Convergence of scaled Markov chains to their fluid limits

[4] Srinivasan V, Anurag Kumar, "Information Dissemination in Socially Aware Networks Under the Linear Threshold Model", NCC'2011

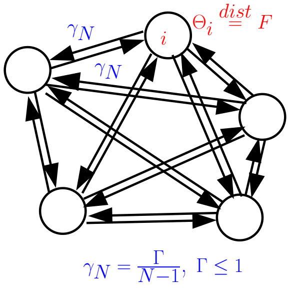
[5] Thomas G. Kurtz, "Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes", J. Appl. Prob. 7, 49-58 (1970)

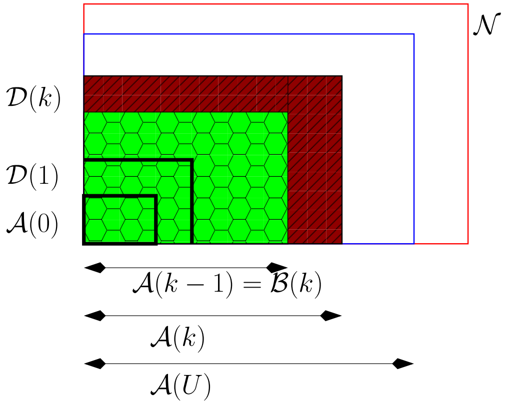




- The influence on nodes is viewed as building up cumulatively
- $\mathcal{D}(k)$: freshly influenced; hence *infectious*
  - will exert their influence in the next step

The HILT Markov Chain

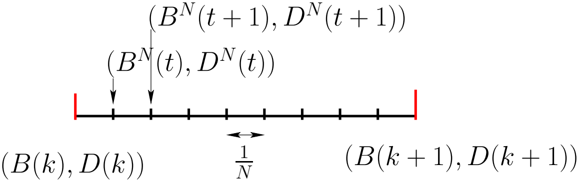




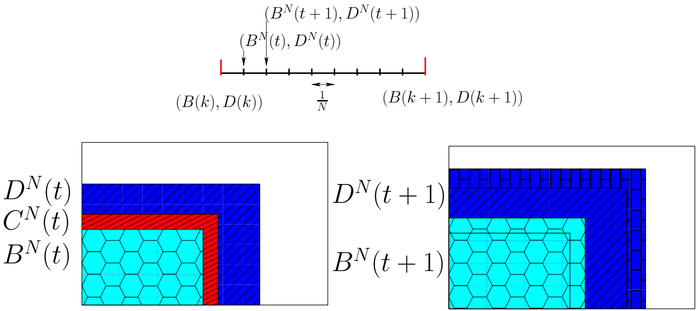
- The influence on nodes is viewed as building up cumulatively
- $\mathcal{D}(k)$: freshly influenced; hence *infectious*
  - will exert their influence in the next step
- $\mathcal{B}(k)$: have already exerted their influence; viewed as having *recovered*
- Let $A(k) = |\mathcal{A}(k)|$, $D(k) = |\mathcal{D}(k)|$ and $B(k) = |\mathcal{B}(k)|$
- $(B(k), D(k))$ can be shown to be a Markov chain

Scaling the Markov Chain: Speeding up Time

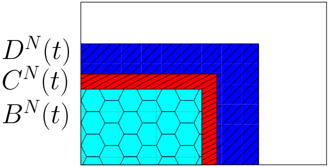
- $(B^N(t), D^N(t))$ evolves over mini-slots of width $\frac{1}{N}$

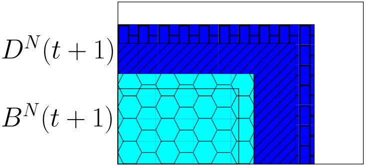


Scaling the Markov Chain: Slowing Down Growth



- Each node in $\mathcal{D}^N(t)$ chooses to influence the relays with probability $\frac{1}{N}$
  - This yields $\mathcal{C}^N(t)$, the nodes that choose to exert their influence
- Each node in $\mathcal{C}^N(t)$ contributes its influence of $\frac{\Gamma}{N-1}$,
  - and moves to $B^N(t+1)$,
  - while causing some new nodes to get influenced
- $D^N(t+1) = (\mathcal{D}^N(t) \backslash \mathcal{C}^N(t)) \cup$ newly influenced nodes





$$C^N(t) = \frac{D^N(t)}{N} + Z_b^N(t+1)$$

$$B^N(t+1) = B^N(t) + C^N(t)$$

$$D^N(t+1) = D^N(t) + \mathbb{E}\left[\frac{F(\gamma_N(B^N(t)+C^N(t))) - F(\gamma_N B^N(t))}{1 - F(\gamma_N B^N(t))}\right] \times \left(N - B^N(t) - D^N(t)\right) - C^N(t) + Z_d^N(t+1)$$

HILT Model: The ODE Limit

- Define $\tilde{B}^N(t)$, $\tilde{C}^N(t)$ and $\tilde{D}^N(t)$ as the fractional processes

Theorem

Given the interest evolution Markov process $(\tilde{B}^N(t), \tilde{D}^N(t))$, *for the threshold distribution with density* $f(\cdot)$, *with bounded* $f'(\cdot)$ *and hazard function* $h_F(x) = \frac{f(x)}{1-F(x)}$, *with* $\tilde{D}^N(0) \stackrel{N\to\infty}{\to} d_0$, *we have for each* $T > 0$ *and each* $\epsilon > 0$,

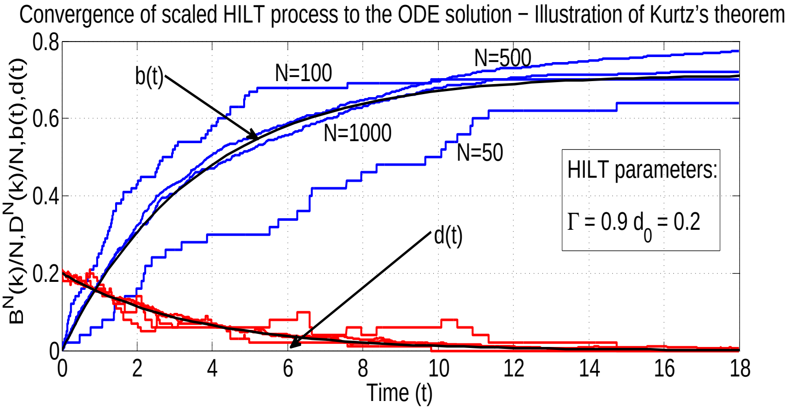
$$\mathbb{P}\left(\sup_{0 \leq u \leq T} \left|\left|(\tilde{B}^N(\lfloor Nu \rfloor), \tilde{D}^N(\lfloor Nu \rfloor)) - (b(u), d(u))\right|\right| > \epsilon \right) \stackrel{N\to\infty}{\to} 0$$

where $(b(u), d(u))$ *is the unique solution to the following o.d.e. with* $(b(0) = 0, d(0) = d_0)$.

$$\begin{aligned} \dot{b} &= d \\ \dot{d} &= h_F(\Gamma b)\Gamma d(1 - b - d) - d \end{aligned}$$

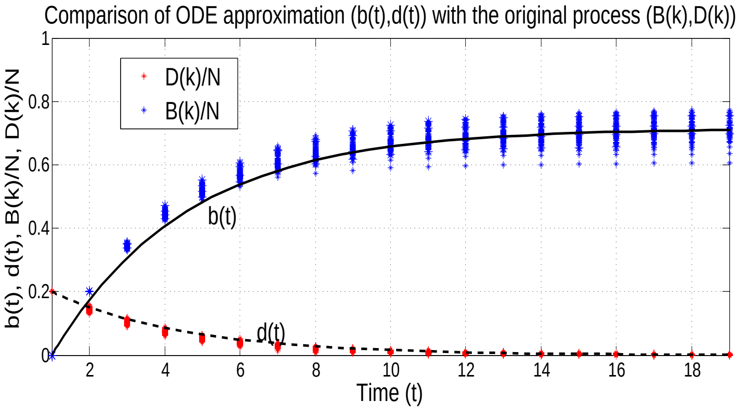
- $h_F(\Gamma b)\Gamma d$: expected rate of conversion of the uninfluenced population
  - given that the uninfluenced population already has the total influence from a fraction b of the influenced nodes

Convergence of the Scaled HILT Markov Chain to the ODE



- Influence threshold distribution is Uniform$[0, 1]$

Comparison of ODE with Unscaled HILT Markov Chain



- $N = 1000$
- We can conclude that the approximation works well for populations of a few 1000s

HILT Model: Effect of Threshold Distribution

- **Uniform[0,1] threshold distribution:** $h_F(x) = \frac{1}{1-x}$
  - Let $r = 1 - \Gamma(1 - d_0)$ then the o.d.e. has an explicit solution

$$
\begin{aligned}
b(t) &= \frac{d_0}{r} - \frac{d_0}{r} e^{-rt} \\
d(t) &= d_0 e^{-rt}
\end{aligned}
$$

- **Exponential distribution with parameter η:** $h_F(x) = \eta$
  - The influence is memoryless
  - The fluid limit is the solution to the o.d.e

$$
\begin{aligned}
\dot{b} &= d \\
\dot{d} &= \eta\Gamma d(1 - b - d) - d
\end{aligned}
$$

  - This is equivalent to the SIR epidemic model with infection rate $\eta\Gamma$ and recovery rate of 1
  - The final fraction of destinations is given by solving

$$b_\infty = 1 - (1 - d_0)e^{-\eta\Gamma b_\infty}$$

Outline

1. Content Spread on Mobile Opportunistic Networks
2. Evolution of Content Popularity
3. Co-evolution of Popularity and Content Spread
4. An Optimization Problem

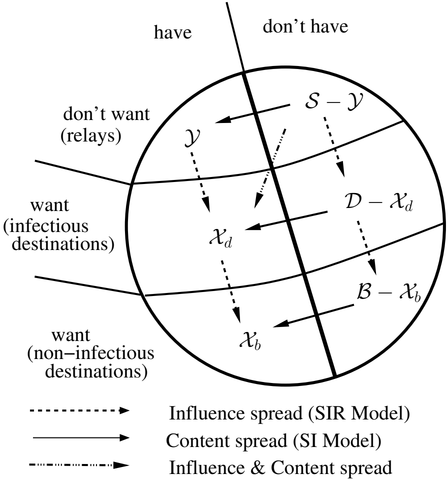
SIR-SI Model: Popularity Spread: SIR; Dissemination: SI

- Spread of a *single* item of content among a homogeneous population of mobile *nodes*
  - co-evolving with growth in interest in the content
- Influence and content transfers occur at pairwise meetings of the nodes
  - Pairwise meetings: independent Poisson point processes

SIR-SI Model: Popularity Spread: SIR; Dissemination: SI

- Spread of a *single* item of content among a homogeneous population of mobile *nodes*
  - co-evolving with growth in interest in the content
- Influence and content transfers occur at pairwise meetings of the nodes
  - Pairwise meetings: independent Poisson point processes
- SIR model motivated by the o.d.e. limit of the HILT model under exponential distribution of influence threshold
- Each node has two states
  - *want*/*do not want* (*destination*/*relay*): evolves according to the SIR model
    - Transition: do not want $\rightarrow$ want
  - *have*/*have not*: evolves according to the SI model
    - Transition: have not $\rightarrow$ have

SIR-SI States and Evolution



Snapshot of partition of the population over the states

- $\mathcal{X}(t)$: destinations that have the content
- $\mathcal{X}_d(t) = \mathcal{X}(t) \cap \mathcal{D}(t)$
  - Node that became destinations,
  - are still infectious in the popularity spread model,
  - and have the content
- $\mathcal{X}_b(t) = \mathcal{X}(t) \cap \mathcal{B}(t)$
  - Node that became destinations,
  - are no longer infectious in the popularity spread model,
  - and have the content

SIR-SI Model: Parameters for Population of Size N

- λ_N: Poisson meeting rate for each pair of nodes; $\lambda_N = \frac{\lambda}{N}$
- β_N: Recovery rate of infectious destinations; $\beta_N = \beta$
- Γ: Influence parameter (recall from the HILT-model)
- σ: Copying probability to a relay (control)
    - We first take σ to be fixed over time
    - A destination (want node) is always copied to

SIR-SI Markov Chain: Transitions at Meeting Epochs

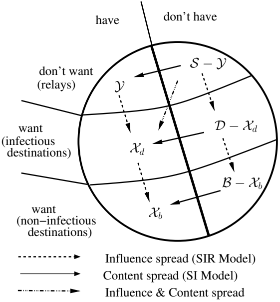
- Let $k = 0, 1, 2, \cdots$ index the meeting epochs at times t_k
- System state: $Z(k) = (B(k), D(k), X_b(k), X_d(k), Y(k))$
- Transitions at meeting epochs where non-zero state update δ_k occurs

Epoch type	Rate	State update δ_k
$\mathcal{D} - \mathcal{X}_d$ recovers	$\beta_N(D(k) - X_d(k))$	(1,-1,0,0,0)
$\mathcal{X}_d$ recovers	$\beta_N X_d(k)$	(1,-1,1,-1,0)
$\mathcal{B} - \mathcal{X}_b$ meets $\mathcal{X} + \mathcal{Y}$	$\lambda_N(B(k) - X_b(k))(X(k) + Y(k))$	(0,0,1,0,0)
$\mathcal{D} - \mathcal{X}_d$ meets $\mathcal{Y}$	$\lambda_N(D(k) - X_d(k))Y(k)$	(0,0,0,1,0) + (0,1,0,1,-1) w.p. Γ
$\mathcal{D} - \mathcal{X}_d$ meets $\mathcal{X}$	$\lambda_N(D(k) - X_d(k))X(k)$	(0,0,0,1,0)
$\mathcal{X}_d$ meets $\mathcal{Y}$	$\lambda_N X_d(k)Y(k)$	(0,1,0,1,-1) w.p Γ
$\mathcal{S} - \mathcal{Y}$ meets $\mathcal{X}_b + \mathcal{Y}$	$\lambda_N(X_b(k) + Y(k))(S(k) - Y(k))$	(0,0,0,0,1) w.p. σ
$\mathcal{S} - \mathcal{Y}$ meets $\mathcal{D} - \mathcal{X}_d$	$\lambda_N(D(k) - X_d(k))(S(k) - Y(k))$	(0,1,0,0,0) w.p. Γ
$\mathcal{S} - \mathcal{Y}$ meets $\mathcal{X}_d$	$\lambda_N(X_d(k))(S(k) - Y(k))$	(0,1,0,1,0) w.p. Γ (0,0,0,0,1) w.p$(1-\Gamma)\sigma$

SIR-SI Markov Chain: Expected Drift Rates

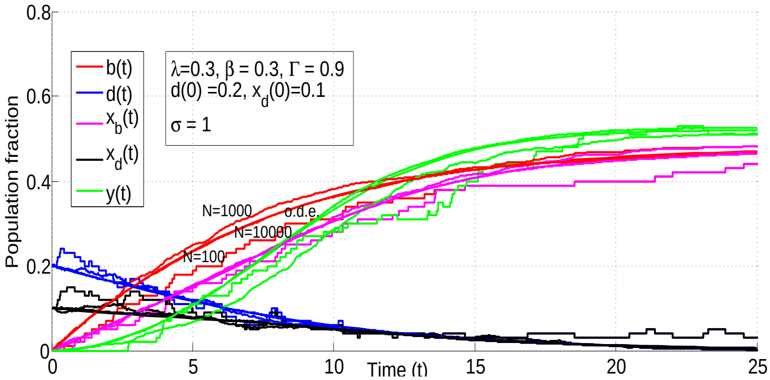
$$
\begin{aligned}
F_b^N &= \beta_N D \\
F_d^N &= -\beta_N D + \lambda_N \Gamma D S \\
F_{x_b}^N &= \beta_N X_d + \\
&\quad \lambda_N (B - X_b)(X + Y) \\
F_{x_d}^N &= \lambda_N (D - X_d) X + \\
&\quad \lambda_N (1 + \Gamma)(D - X_d) Y + \\
&\quad \lambda_N \Gamma X_d S - \beta_N X_d \\
F_y^N &= -\lambda_N \Gamma Y + \\
&\quad \lambda_N \sigma (S - Y)(X_b + Y + (1 - \Gamma) X_d)
\end{aligned}
$$

SIR-SI Markov Chain: O.D.E. Limit



$$
\begin{aligned}
\dot{b} &= \beta d \\
\dot{d} &= -\beta d + \lambda \Gamma d s \\
\dot{x}_b &= \beta x_d + \lambda (b - x_b)(x + y) \\
\dot{x}_d &= \lambda (d - x_d)(x + y) + \lambda \Gamma d y + \Gamma \lambda x_d (s - y) - \beta x_d \\
\dot{y} &= -\Gamma \lambda d y + \lambda \sigma (s - y)(x_b + y + (1 - \Gamma) x_d)
\end{aligned}
$$

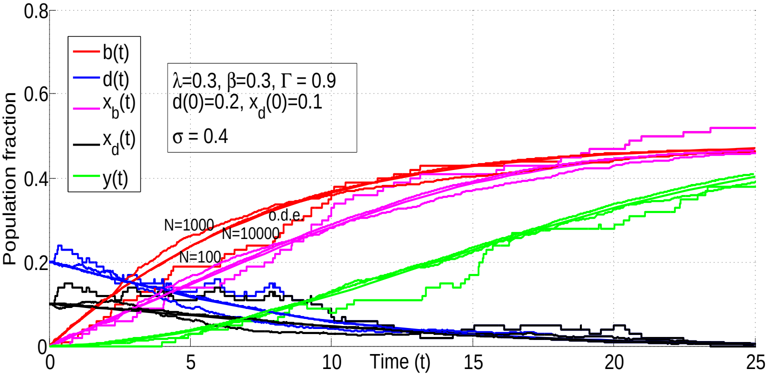
where $s(t) = 1 - b(t) - d(t)$



Simulation of SIR-SI CTMC for various N; shown with the fluid limit; $\sigma = 1$

- 95% of the "wants" have the content by $t = 17.5$
- But, by then, almost all the "do-not-wants" have the content too
  - This could be a waste, if do-not-wants get an incentive (e.g., rebate) to carry the content

Convergence of Scaled Markov Chain to O.D.E. Limit



Simulation of SIR-SI CTMC for various N; shown with fluid limit; $\sigma = 0.4$

- 95% of the "wants" have the content by $t = 18$
- At this time, just 60% of the eventual "do-not-wants" have the content

SIR-SI Model: Dynamic Control of Copying

- For each N we have a controlled continuous time Markov chain
    - Optimal control appears difficult to obtain

---

[6] N. Gast, B. Gaujal, and J. Boudec, "Mean field for Markov decision processes: from discrete to continuous optimization," Arxiv preprint arXiv:1004.2342, 2010.

SIR-SI Model: Dynamic Control of Copying

- For each N we have a controlled continuous time Markov chain
  - Optimal control appears difficult to obtain
- Replace the constant copying probability in the ODE limit by $\sigma(t)$
  - which yields a controlled ODE
- Obtain the optimal (deterministic) control for the controlled ODE
- What value does this control have for the original controlled Markov chain?

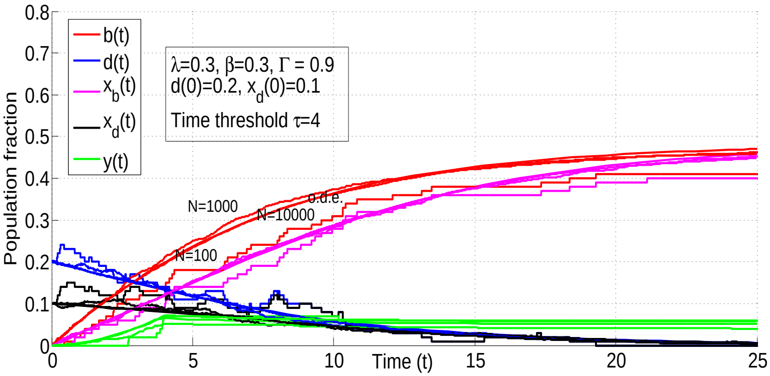
[6] N. Gast, B. Gaujal, and J. Boudec, "Mean field for Markov decision processes: from discrete to continuous optimization," Arxiv preprint arXiv:1004.2342, 2010.

SIR-SI Model: Dynamic Control of Copying

- For each N we have a controlled continuous time Markov chain
  - Optimal control appears difficult to obtain
- Replace the constant copying probability in the ODE limit by $\sigma(t)$
  - which yields a controlled ODE
- Obtain the optimal (deterministic) control for the controlled ODE
- What value does this control have for the original controlled Markov chain?
- In some situations it can be shown that this control is asymptotically optimal
  - as the size of the finite size problem increases
  - *without* needing to solve the finite size problem[6]
- We have adopted this approach
  - The proof in Gast et al. needs to be extended to accommodate our setting

[6] N. Gast, B. Gaujal, and J. Boudec, "Mean field for Markov decision processes: from discrete to continuous optimization," Arxiv preprint arXiv:1004.2342, 2010.

Dynamic Control: Comparison of SIR-SI and the O.D.E.



Simulation of SIR-SI CTMC for various N; shown with the fluid limit

- The control $\sigma(t) = 1, 0 \leq t \leq 4$, and $\sigma(t) = 0$ thereafter
- 95% of the "wants" have the content by $t = 21$
- Just 15% of the eventual "do-not-wants" have the content at this time
- Is there an *optimal control*?

Outline

1. Content Spread on Mobile Opportunistic Networks
2. Evolution of Content Popularity
3. Co-evolution of Popularity and Content Spread
4. An Optimization Problem

Optimal Control of Co-evolution: An Objective

Fix α, and, for a control functon $\sigma(t)$, let

$$T_\sigma = \inf\{t : x_\sigma(t) \geq \alpha a(\infty)\}$$

where

- $a(\infty) = b(\infty) + d(\infty)$ is the terminal fraction of destinations (SIR model)
- $x_\sigma(t)$ is the fraction of destinations that have the content at time t

The cost we are interested in optimizing is of the form

$$C_\sigma = \psi y_\sigma(T_\sigma) + T_\sigma = \psi y_\sigma(T_\sigma) + \int_0^{T_\sigma} 1 dt$$

- $y_\sigma(T_\sigma)$ denotes the number of relays that have the content at time T_σ

Optimality of a Time Threshold Control

$$
\begin{aligned}
\dot{b} &= \beta d \\
\dot{d} &= -\beta d + \lambda \Gamma d s \\
\dot{x}_b &= \beta x_d + \lambda (b - x_b)(x + y) \\
\dot{x}_d &= \lambda (d - x_d) x + \lambda (1 + \Gamma)(d - x_d) y + \Gamma \lambda x_d s - \beta x_d \\
\dot{y} &= -\Gamma \lambda d y + \lambda \sigma(t)(s - y)(x_b + y + (1 - \Gamma) x_d)
\end{aligned}
$$

where $s(t) = 1 - b(t) - d(t)$

Theorem

For the above o.d.e. system, for the cost function displayed earlier, there exists an optimal control of the form,

$$
\sigma_\tau(t) = \begin{cases} 1, & 0 < t < \tau \\ 0, & t \geq \tau \end{cases}
$$

Optimality of a Time Threshold Control

$$
\begin{aligned}
\dot{b} &= \beta d \\
\dot{d} &= -\beta d + \lambda \Gamma d s \\
\dot{x}_b &= \beta x_d + \lambda (b - x_b)(x + y) \\
\dot{x}_d &= \lambda (d - x_d) x + \lambda (1 + \Gamma)(d - x_d) y + \Gamma \lambda x_d s - \beta x_d \\
\dot{y} &= -\Gamma \lambda d y + \lambda \sigma(t)(s - y)(x_b + y + (1 - \Gamma) x_d)
\end{aligned}
$$

where $s(t) = 1 - b(t) - d(t)$

Theorem

For the above o.d.e. system, for the cost function displayed earlier, there exists an optimal control of the form,

$$
\sigma_\tau(t) = \begin{cases} 1, & 0 < t < \tau \\ 0, & t \geq \tau \end{cases}
$$

- Distributed implementation will require
  - Time stamping of the content when it is initially distributed
  - Zero skew between node clocks; or a distributed time-synch algorithm

Optimality of a Time Threshold Control: Sketch of Proof

- Define Kamke dominance: extension of Kamke condition for o.d.e.s[7]
- Here Kamke dominance holds, hence $\forall t$, $\sigma^{(1)}(t) \geq \sigma^{(2)}(t)$, and

$$(x_b^{(1)}(0), x_d^{(1)}(0), y^{(1)}(0)) \geq (x_b^{(2)}(0), x_d^{(2)}(0), y^{(2)}(0)),$$

implies that, $\forall t$,

$$(x_b^{(1)}(t), x_d^{(1)}(t), y^{(1)}(t)) \geq (x_b^{(2)}(t), x_d^{(2)}(t), y^{(2)}(t))$$

[7] H. Smith, "Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems." American Mathematical Soc., 1995.

Optimality of a Time Threshold Control: Sketch of Proof

- Define Kamke dominance: extension of Kamke condition for o.d.e.s[7]
- Here Kamke dominance holds, hence $\forall t$, $\sigma^{(1)}(t) \geq \sigma^{(2)}(t)$, and

$$(x_b^{(1)}(0), x_d^{(1)}(0), y^{(1)}(0)) \geq (x_b^{(2)}(0), x_d^{(2)}(0), y^{(2)}(0)),$$

  implies that, $\forall t$,

$$(x_b^{(1)}(t), x_d^{(1)}(t), y^{(1)}(t)) \geq (x_b^{(2)}(t), x_d^{(2)}(t), y^{(2)}(t))$$

- Consider $\sigma(t)$ (any action function) and $\sigma_\tau(t)$ (a time threshold action function) such that $y_\sigma(T_\sigma) = y_{\sigma_\tau}(T_{\sigma_\tau}) = \rho$
  - We can show that $T_{\sigma_\tau} \leq T_\sigma$ and hence the time threshold action function is optimal in $\{\sigma(\cdot) : y_\sigma(T_\sigma) = \rho\}$

[7] H. Smith, "Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems." American Mathematical Soc., 1995.

Optimality of a Time Threshold Control: Sketch of Proof

- Define Kamke dominance: extension of Kamke condition for o.d.e.s[7]
- Here Kamke dominance holds, hence $\forall t$, $\sigma^{(1)}(t) \geq \sigma^{(2)}(t)$, and

$$(x_b^{(1)}(0), x_d^{(1)}(0), y^{(1)}(0)) \geq (x_b^{(2)}(0), x_d^{(2)}(0), y^{(2)}(0)),$$

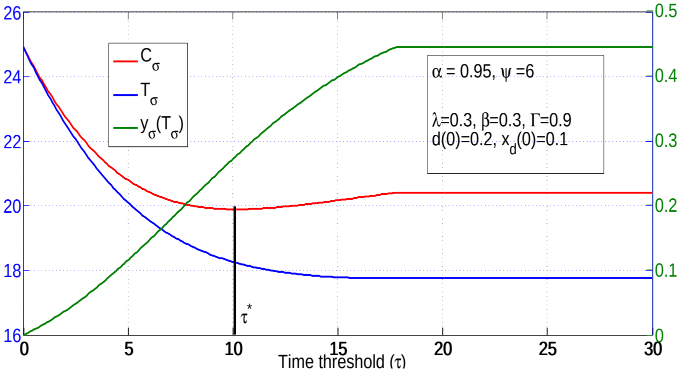
implies that, $\forall t$,

$$(x_b^{(1)}(t), x_d^{(1)}(t), y^{(1)}(t)) \geq (x_b^{(2)}(t), x_d^{(2)}(t), y^{(2)}(t))$$

- Consider $\sigma(t)$ (any action function) and $\sigma_\tau(t)$ (a time threshold action function) such that $y_\sigma(T_\sigma) = y_{\sigma_\tau}(T_{\sigma_\tau}) = \rho$
  - We can show that $T_{\sigma_\tau} \leq T_\sigma$ and hence the time threshold action function is optimal in $\{\sigma(\cdot) : y_\sigma(T_\sigma) = \rho\}$
- Define $\rho_{max} = \max\{y_{\sigma_\tau}(T_{\sigma_\tau}) : \tau \geq 0\}$
  - Consider an action function $\sigma(t)$ with $y_\sigma(T_\sigma) > \rho_{max}$
  - The threshold policy with $\tau := \sup\{t : \sigma(t) > 0\}$ has lower cost

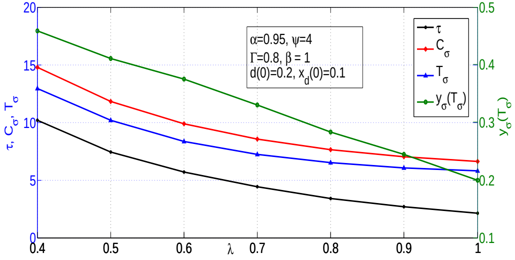
[7] H. Smith, "Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems." American Mathematical Soc., 1995.

Obtaining the Optimal Control for the Running Example

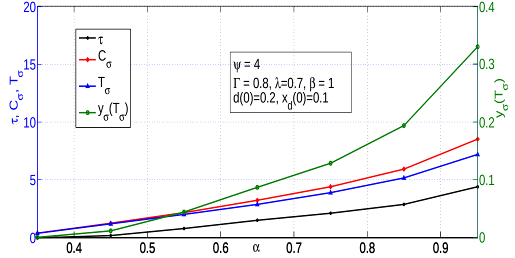


- $\alpha = 0.95$, $C_\sigma = 6y_\sigma(T_\sigma) + T_\sigma$
- The optimal control is to copy until $\tau = 10.1$ and then stop copying

Examples of the Optimal Control and Optimal Cost



Optimal policy and costs vs. meeting rate parameter λ



Optimal policy and costs vs. target spread fraction α

Conclusion

- Studied co-evolution of popularity and content spread
  - by combining a social influence model (HILT)
  - with a mobile opportunistic network (MON) model
- Included the HILT model via its o.d.e. limit
  - For exponential threshold this gave the SIR model
- Modeled the MON by its fluid limit
  - SI model driven by the SIR model for popularity spread
- Showed the existence of a time-threshold optimal control for an optimization problem under the SIR-SI model

Conclusion

- Studied co-evolution of popularity and content spread
  - by combining a social influence model (HILT)
  - with a mobile opportunistic network (MON) model
- Included the HILT model via its o.d.e. limit
  - For exponential threshold this gave the SIR model
- Modeled the MON by its fluid limit
  - SI model driven by the SIR model for popularity spread
- Showed the existence of a time-threshold optimal control for an optimization problem under the SIR-SI model
- Possible extensions:
  - Multiple items of content; communities of interest
  - Service pricing, and incentive mechanisms for relays