Co-evolution of Content Popularity and Delivery in Mobile P2P Networks

Srinivasan Venkatramanan and Anurag Kumar

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

March 26,2012

Outline

1. Motivation
2. System Model
3. Fluid limits
4. Optimization Problems
5. Concluding Remarks

Outline

1. Motivation
2. System Model
3. Fluid limits
4. Optimization Problems
5. Concluding Remarks

Mobile P2P Networks

- A scalable approach to content delivery for mobile nodes

Mobile P2P Networks

- A scalable approach to content delivery for mobile nodes
- Why Mobile P2P?
  - Increased penetration of smart phones

Mobile P2P Networks

- A scalable approach to content delivery for mobile nodes
- Why Mobile P2P?
  - Increased penetration of smart phones
  - Growing demand for high-bandwidth content on mobile devices

Mobile P2P Networks

- A scalable approach to content delivery for mobile nodes
- Why Mobile P2P?
  - Increased penetration of smart phones
  - Growing demand for high-bandwidth content on mobile devices
  - A single item of content (news,ringtone,etc.) may be of interest to several users in a region

Mobile P2P Networks

- A scalable approach to content delivery for mobile nodes
- Why Mobile P2P?
  - Increased penetration of smart phones
  - Growing demand for high-bandwidth content on mobile devices
  - A single item of content (news,ringtone,etc.) may be of interest to several users in a region
- Can be studied under the Mobile Opportunistic Networks paradigm

Example setup: Highlights of Infocom Keynote Speech

Example setup: Highlights of Infocom Keynote Speech

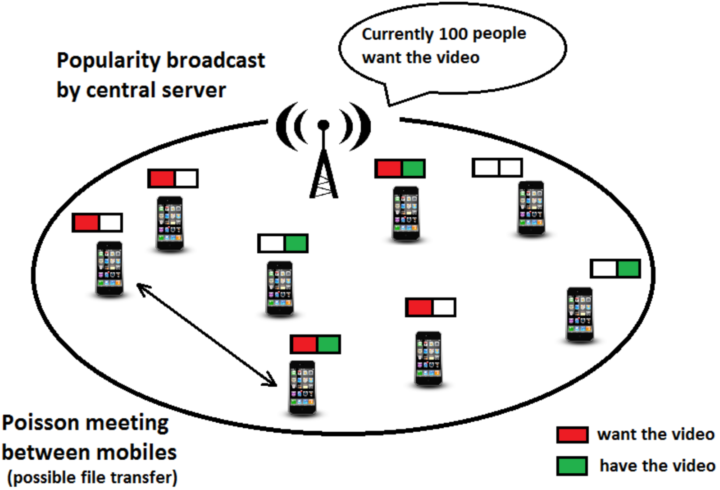


Figure: Mobile P2P Network - An Example

Evolution of Content Popularity

- Creators: Increase the popularity of the content [1]

[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", WiOpt 2011

Evolution of Content Popularity

- Creators: Increase the popularity of the content [1]
- Distributors: Optimally control delivery mechanisms [2]

[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", WiOpt 2011

Evolution of Content Popularity

- Creators: Increase the popularity of the content [1]
- Distributors: Optimally control delivery mechanisms [2]
- Influence spread models (from viral marketing) can be employed

[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", WiOpt 2011

Our Contributions

Our Contributions

- Set up Markov models for the co-evolution of content popularity and delivery

Our Contributions

- Set up Markov models for the co-evolution of content popularity and delivery
- Derive o.d.e. limits (*fluid limits*) from the Markov models

Our Contributions

- Set up Markov models for the co-evolution of content popularity and delivery
- Derive o.d.e. limits (*fluid limits*) from the Markov models
- Show that the popularity evolution model under study, is equivalent to the SIR epidemic model

Our Contributions

- Set up Markov models for the co-evolution of content popularity and delivery
- Derive o.d.e. limits (*fluid limits*) from the Markov models
- Show that the popularity evolution model under study, is equivalent to the SIR epidemic model
- Use the fluid limits to solve various optimization problems for both the content creator and the content distributor

Outline

1. Motivation
2. System Model
3. Fluid limits
4. Optimization Problems
5. Concluding Remarks

System Model

Spread of *single* item of content among a homogeneous population of mobile *nodes*

System Model

Spread of *single* item of content among a homogeneous population of mobile *nodes*

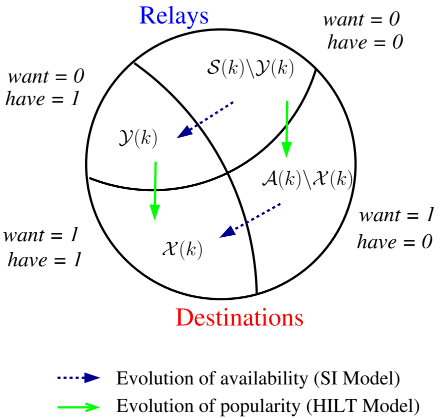


Figure: Various states and the transitions in the Co-evolution model

System Model

Spread of *single* item of content among a homogeneous population of mobile *nodes*

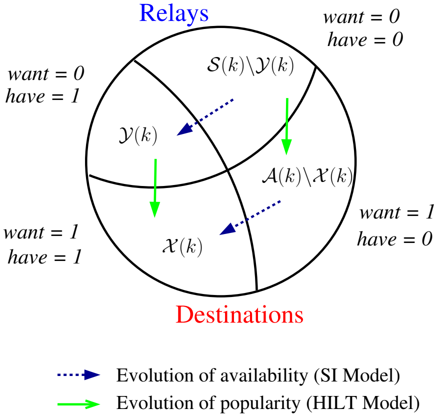


Figure: Various states and the transitions in the Co-evolution model

Central server to broadcast popularity information at regular intervals

Network setup for Popularity

Homogeneous Influence Linear Threshold model (HILT Model)[3]

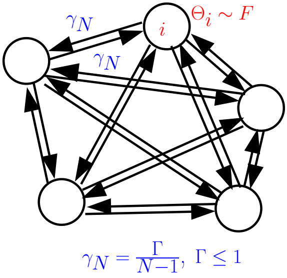


Figure: HILT Model

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

Network setup for Popularity

Homogeneous Influence Linear Threshold model (HILT Model)[3]

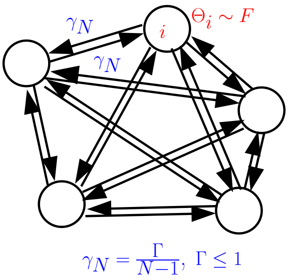


Figure: HILT Model

- Complete network on $\mathcal{N}$, $N = |\mathcal{N}|$

---

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

Network setup for Popularity

Homogeneous Influence Linear Threshold model (HILT Model)[3]

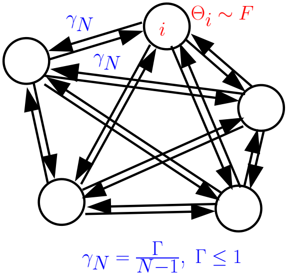


Figure: HILT Model

- Complete network on $\mathcal{N}$, $N = |\mathcal{N}|$
- Influence edge weights $\gamma_N = \frac{\Gamma}{N-1}$

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

Network setup for Popularity

Homogeneous Influence Linear Threshold model (HILT Model)[3]

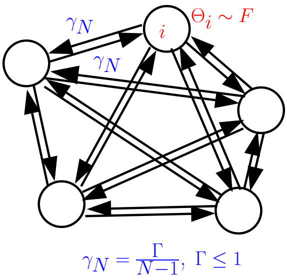


Figure: HILT Model

- Complete network on $\mathcal{N}$, $N = |\mathcal{N}|$
- Influence edge weights $\gamma_N = \frac{\Gamma}{N-1}$
- Random threshold $\Theta_i \sim F$, $i \in \mathcal{N}$

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

Popularity Evolution

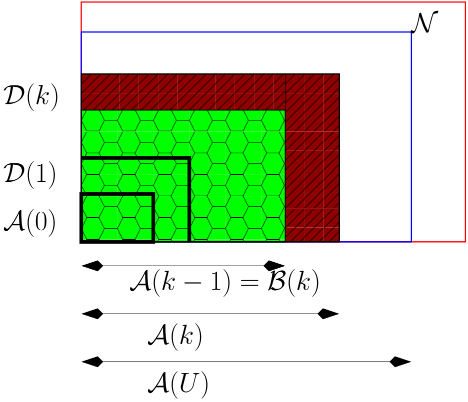
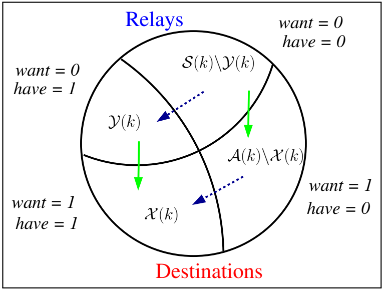


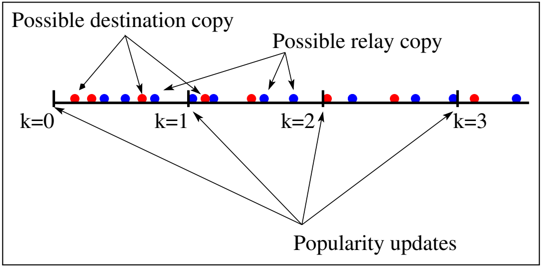
Figure: Spread of Influence in the HILT Model

A relay $j \notin \mathcal{A}(k-1)$ becomes a destination at time k, if

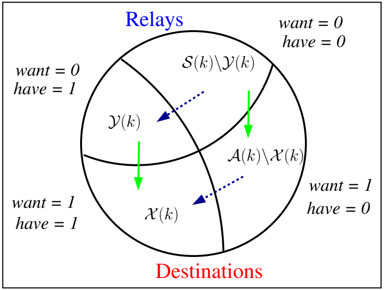
$$\gamma_N |\mathcal{A}(k-1)| \geq \Theta_j$$

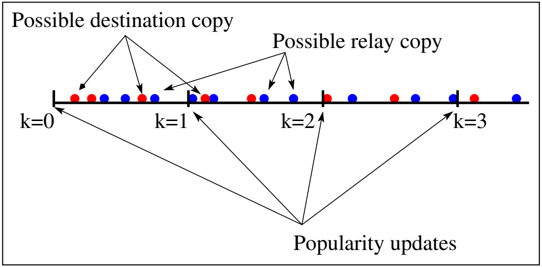
Content Spread (SI model)





Content Spread (SI model)





Each pair of nodes meet at points of a Poisson process with rate λ_N

- **Possible destination copy**: $i \in \mathcal{A}(k)\backslash\mathcal{X}(k)$ meets $j \in \mathcal{X}(k) \bigcup \mathcal{Y}(k)$
- Copy occurs with probability α
- **Possible relay copy**: $i \in \mathcal{S}(k)\backslash\mathcal{Y}(k)$ meets $j \in \mathcal{X}(k) \bigcup \mathcal{Y}(k)$
- Copy occurs with probability σ

Outline

1 Motivation

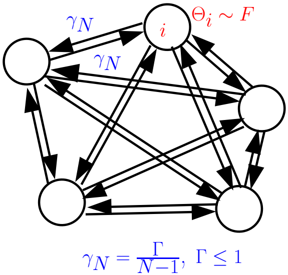
2 System Model

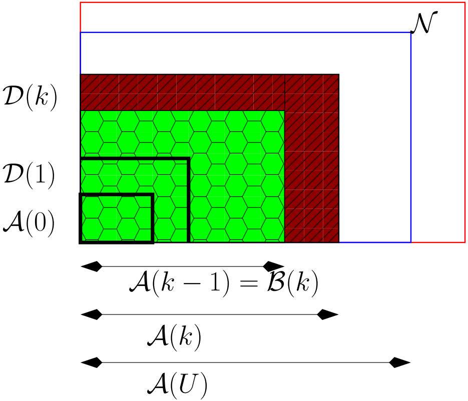
3 Fluid limits

4 Optimization Problems

5 Concluding Remarks

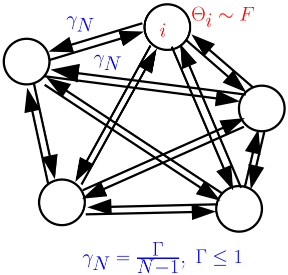
HILT model

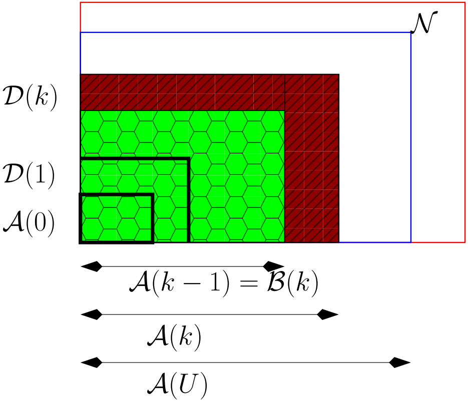




[4] Thomas G. Kurtz, "Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes", J. Appl. Prob. 7, 49-58 (1970)

HILT model

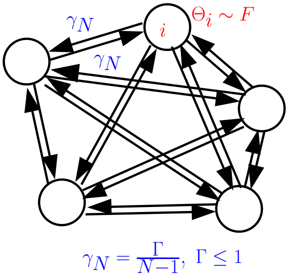


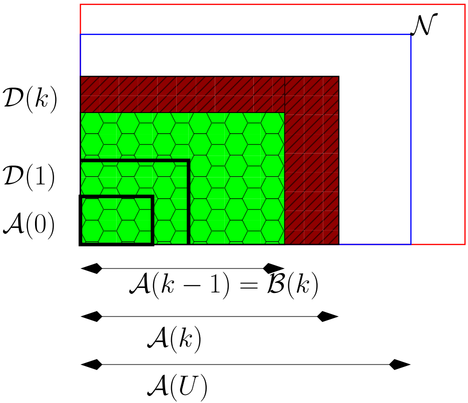


- Let $A(k) = |\mathcal{A}(k)|$, $D(k) = |\mathcal{D}(k)|$ and $B(k) = |\mathcal{B}(k)|$

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

HILT model



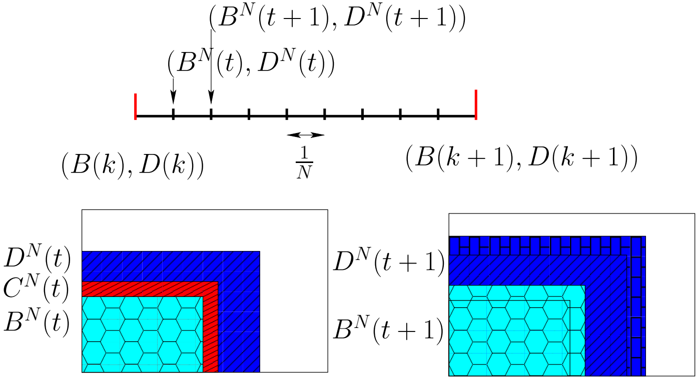


- Let $A(k) = |\mathcal{A}(k)|$, $D(k) = |\mathcal{D}(k)|$ and $B(k) = |\mathcal{B}(k)|$
- $(B(k), D(k))$ is a Markov chain
- Kurtz theorem [4]: Convergence of Markov chains to their fluid limits

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

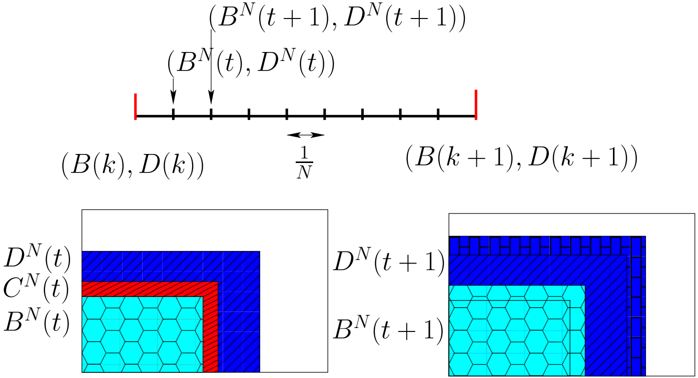
Scaling and Drift Equations

Scaled Markov process $(B^N(t), D^N(t))$ evolving over mini-slots



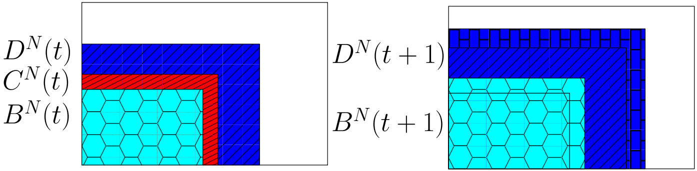
Scaling and Drift Equations

Scaled Markov process $(B^N(t), D^N(t))$ evolving over mini-slots



- Each infectious destination in $D^N(t)$ attempts to influence the relays with probability $\frac{1}{N}$
  - Success: Contributes its influence of $\frac{\Gamma}{N-1}$ and moves to $B^N(t+1)$
  - Failure: Stays in the $D^N(t+1)$

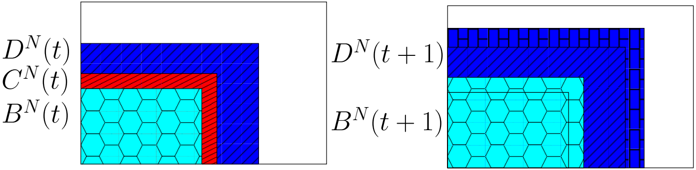
Scaling and Drift Equations



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

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

Scaling and Drift Equations



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

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

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

Fluid Limit for Interest Evolution - HILT Model

- Defining $\tilde{B}^N(t)$, $\tilde{C}^N(t)$ and $\tilde{D}^N(t)$ as the fractional processes, we can then state the following theorem

Fluid Limit for Interest Evolution - HILT Model

- Defining $\tilde{B}^N(t)$, $\tilde{C}^N(t)$ and $\tilde{D}^N(t)$ as the fractional processes, we can then state the following theorem

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* $h_F(x) = \frac{f(x)}{1-F(x)}$, *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 o.d.e.,*

$$\dot{b} = d \quad (1)$$

$$\dot{d} = h_F(\Gamma b)\Gamma d(1 - b - d) - d \quad (2)$$

with initial conditions $(b(0) = 0, d(0) = d_0)$.

HILT Model: Effect of Threshold Distribution

- **Uniform threshold distribution:** $h_F(x) = \frac{1}{1-x}$. Let $r = 1 - \Gamma + \Gamma d_0$ then we have,

HILT Model: Effect of Threshold Distribution

- **Uniform threshold distribution:** $h_F(x) = \frac{1}{1-x}$. Let $r = 1 - \Gamma + \Gamma d_0$ then we have,

$$b(t) = \frac{d_0}{r} - \frac{d_0}{r} e^{-rt} \quad (3)$$

$$d(t) = d_0 e^{-rt} \quad (4)$$

HILT Model: Effect of Threshold Distribution

- **Uniform threshold distribution:** $h_F(x) = \frac{1}{1-x}$. Let $r = 1 - \Gamma + \Gamma d_0$ then we have,

$$b(t) = \frac{d_0}{r} - \frac{d_0}{r} e^{-rt} \quad (3)$$

$$d(t) = d_0 e^{-rt} \quad (4)$$

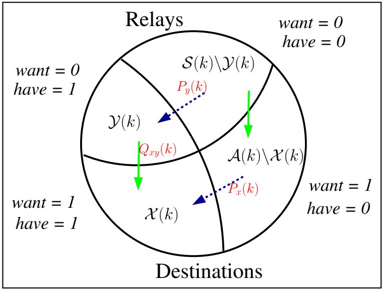
- **Exponential distribution with parameter** β**:** $h_F(x) = \beta$ (memoryless). The fluid limit is the solution to the o.d.e.,

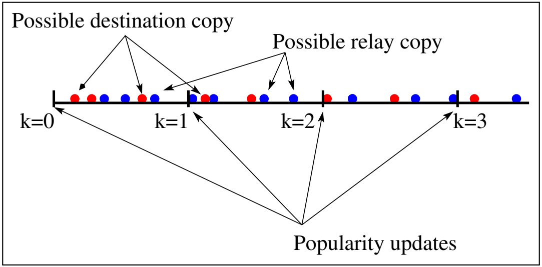
$$\dot{b} = d$$

$$\dot{d} = \beta\Gamma d(1 - b - d) - d$$

This is equivalent to the SIR epidemic model with infection rate $\beta\Gamma$ and recovery rate of 1

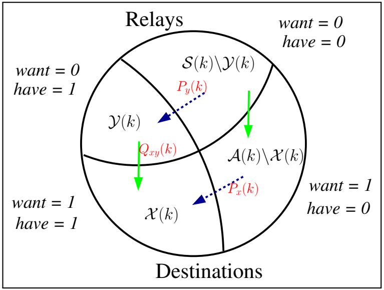
Coevolution - HILT-SI Model

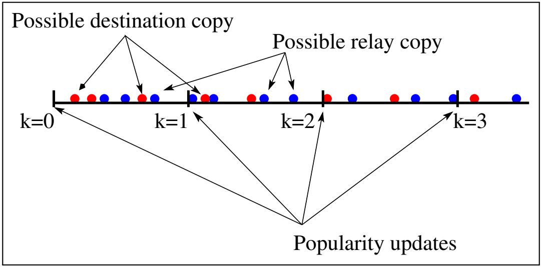




- Probability that a destination node without the content, receives it in the given time slot $= 1 - e^{-\lambda_N \alpha (X(k)+Y(k))}$ and similarly for a relay

Coevolution - HILT-SI Model





- Probability that a destination node without the content, receives it in the given time slot $= 1 - e^{-\lambda_N \alpha(X(k)+Y(k))}$ and similarly for a relay

$$P_x(k) \overset{\text{dist.}}{=} \text{Bin}\left(A(k) - X(k), \lambda_N \alpha(X(k) + Y(k))\right)$$

$$P_y(k) \overset{\text{dist.}}{=} \text{Bin}\left(N - A(k) - Y(k), \lambda_N \sigma(X(k) + Y(k))\right)$$

$$Q_{xy}(k) \overset{\text{dist.}}{=} \text{Bin}\left(Y(k), \frac{\gamma_N D(k)}{1 - \gamma_N B(k)}\right)$$

Drift Equations

$$
\begin{aligned}
X(k+1) - X(k) &= \lambda_N \alpha (X(k) + Y(k))(A(k) - X(k)) \\
&+ \frac{\gamma_N D(k)}{1 - \gamma_N B(k)} Y(k) + Z_x(k)
\end{aligned}
$$

$$
\begin{aligned}
Y(k+1) - Y(k) &= \lambda_N \sigma (X(k) + Y(k))(N - A(k) - Y(k)) \\
&- \frac{\gamma_N D(k)}{1 - \gamma_N B(k)} Y(k) + Z_y(k)
\end{aligned}
$$

where $Z_x(k)$ and $Z_y(k)$ are noise terms with zero mean conditional on the history of the joint evolution process until period k.
Obtain fluid limit by working with the fractional scaled process which evolves over mini-slots of width $\frac{1}{N}$

ODE limit of HILT-SI Model

Theorem

Given the joint evolution Markov process $(\tilde{B}^N(t), \tilde{D}^N(t), \tilde{X}^N(t), \tilde{Y}^N(t))$*, we have for each* $T > 0$ *and each* $\epsilon > 0$*,*

$$\mathbb{P}\left(\sup_{0 \le u \le T} \left|\left| (\tilde{B}^N(\lfloor Nu \rfloor), \tilde{D}^N(\lfloor Nu \rfloor), \tilde{X}^N(\lfloor Nu \rfloor), \tilde{Y}^N(\lfloor Nu \rfloor)) - (b(u), d(u), x(u), y(u)) \right|\right| > \epsilon \right) \overset{N \to \infty}{\to} 0$$

where $(b(u), d(u), x(u), y(u))$ *is the unique solution to*

$$\dot{b} = d; \qquad \dot{d} = \frac{\Gamma d}{1 - \Gamma b}(1 - b - d) - d \tag{5}$$

$$\dot{x} = \lambda\alpha(x + y)(a - x) + \frac{\Gamma d}{1 - \Gamma b} y \tag{6}$$

$$\dot{y} = \lambda\sigma(x + y)(1 - a - y) - \frac{\Gamma d}{1 - \Gamma b} y \tag{7}$$

with initial conditions $(b(0) = 0,\ d(0) = d_0,\ x(0) = x_0,\ y(0) = y_0)$*.*

Accuracy of Fluid limits (Scaled process)

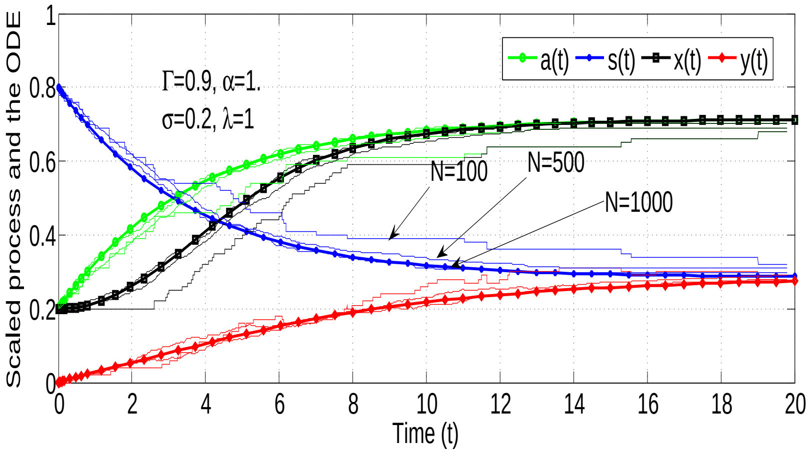


Figure: Comparison of the scaled HILT-SI process for various values of N with the corresponding fluid limit.

Accuracy of Fluid limits (Original process)

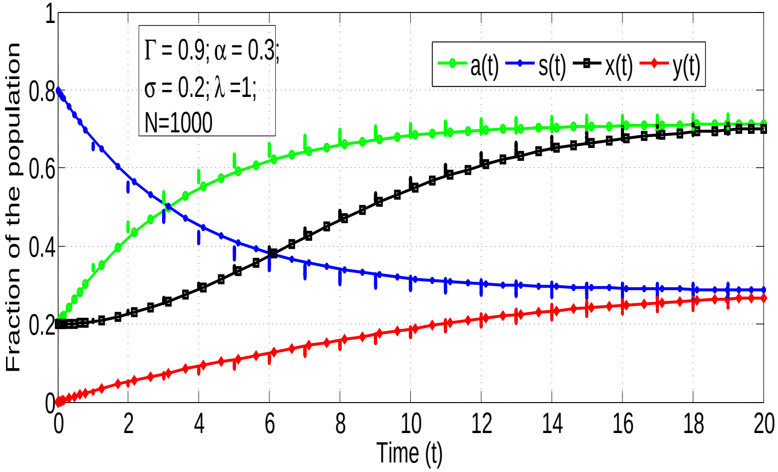


Figure: Comparison of the unscaled HILT-SI process ($N = 1000$) with the corresponding fluid limit. The solid lines indicate the evolution of o.d.e. solutions.

Outline

1. Motivation
2. System Model
3. Fluid limits
4. Optimization Problems
5. Concluding Remarks

Interest evolution

Interest evolution

- Content creators: Wish to maximize the level of popularity achieved
- Do not have control over influence weight Γ or the threshold distribution F,
- Only parameter under control is d_0

Interest evolution

- Content creators: Wish to maximize the level of popularity achieved
- Do not have control over influence weight Γ or the threshold distribution F,
- Only parameter under control is d_0

Optimal d_0 to obtain given b_∞

$$d_0 = \frac{b_\infty(1-\Gamma)}{1-b_\infty\Gamma} \tag{8}$$

Interest evolution

- Content creators: Wish to maximize the level of popularity achieved
- Do not have control over influence weight Γ or the threshold distribution F,
- Only parameter under control is d_0

Optimal d_0 to obtain given b_∞

$$d_0 = \frac{b_\infty(1-\Gamma)}{1-b_\infty\Gamma} \tag{8}$$

Time to reach target β ($\beta < \frac{d_0}{r}$), given d_0

$$T(\beta, d_0, \Gamma) = \frac{1}{r}\ln\left(\frac{1-r}{1-\frac{\beta}{d_0}r}\right) \tag{9}$$

with $r = 1 - \Gamma + \Gamma d_0$.

Optimizations in Coevolution

Optimizations in Coevolution

- Delivering to as many destinations as possible by a fixed time
- Content is time-dependent and its usefulness expires by that fixed time

Optimizations in Coevolution

- Delivering to as many destinations as possible by a fixed time
- Content is time-dependent and its usefulness expires by that fixed time

Maximize target spread

$$\max_{\{\sigma : y(\tau) \leq \zeta\}} x(\tau)$$

Optimizations in Coevolution

- Delivering to as many destinations as possible by a fixed time
- Content is time-dependent and its usefulness expires by that fixed time

Maximize target spread

$$\max_{\{\sigma : y(\tau) \leq \zeta\}} x(\tau)$$

- Deliver the content to a given fraction of destinations as early as possible

Optimizations in Coevolution

- Delivering to as many destinations as possible by a fixed time
- Content is time-dependent and its usefulness expires by that fixed time

Maximize target spread

$$\max_{\{\sigma: y(\tau) \leq \zeta\}} x(\tau)$$

- Deliver the content to a given fraction of destinations as early as possible
- Define $\tau_\eta = \inf\{t : x(t) \geq \eta\}$

Minimize reach time

$$\min_{\{\sigma: y(\tau_\eta) \leq \zeta\}} \tau_\eta$$

- In both cases, the constraint is on the number of wasted copies to relays

Maximize target spread (Variation w.r.t. τ)

- τ - Target time
- ζ - Constraint on wasted copies
- σ - Copy probability to relays

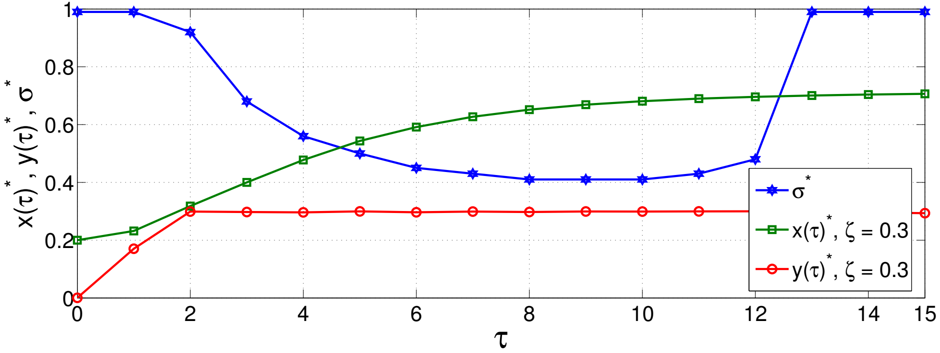


Figure: Maximize target spread: The optimal solution plotted for a fixed value of ζ and varying τ.

Maximize target spread (Variation w.r.t. ζ)

- τ - Target time
- ζ - Constraint on wasted copies
- σ - Copy probability to relays

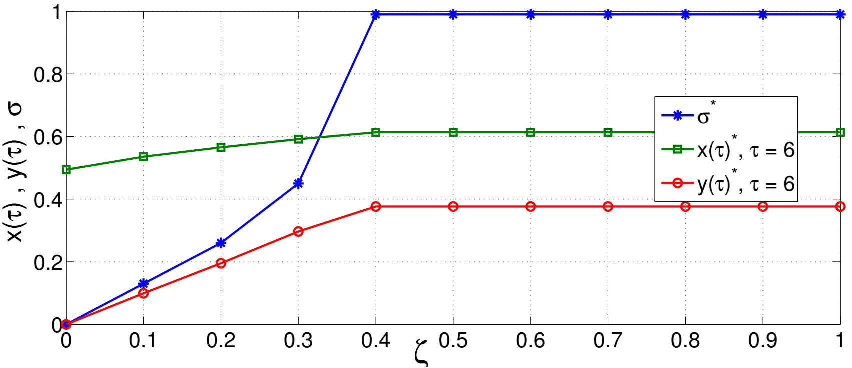


Figure: Maximize target spread: The optimal solution plotted for a fixed value of τ and varying ζ.

Minimize reach time (Variation w.r.t. η)

- η - Target fraction
- τ_η - Time to reach target fraction
- ζ - Constraint on wasted copies
- σ - Copy probability to relays

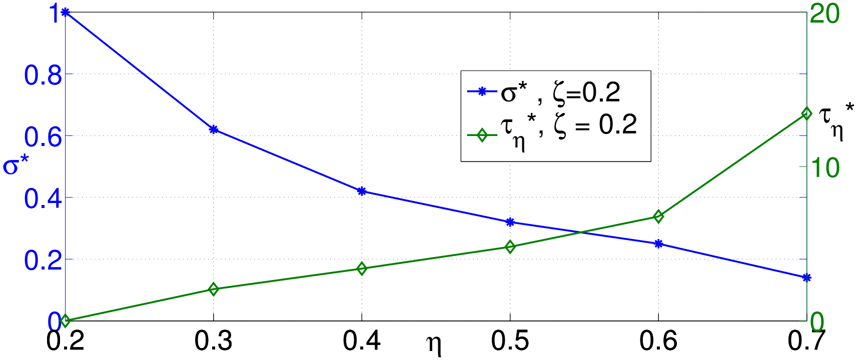


Figure: Minimize reach time: The optimal solution plotted for a fixed value of ζ and varying η.

Minimize reach time (Variation w.r.t. ζ)

- η - Target fraction
- τ_η - Time to reach target fraction
- ζ - Constraint on wasted copies
- σ - Copy probability to relays

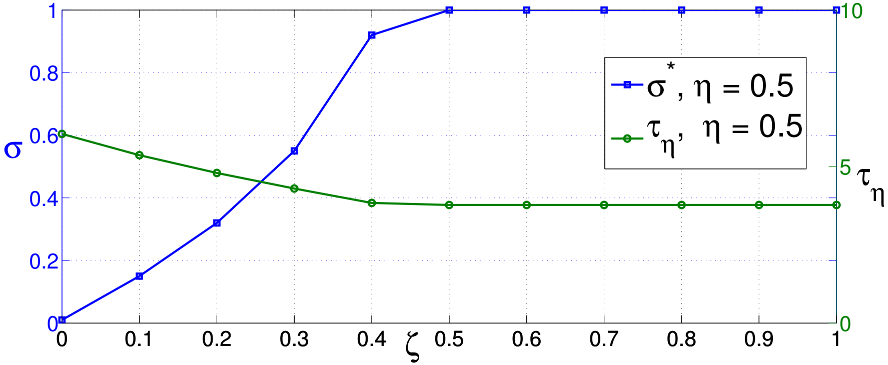


Figure: Minimize reach time: The optimal solution plotted for a fixed value of η and varying ζ.

Outline

1. Motivation
2. System Model
3. Fluid limits
4. Optimization Problems
5. Concluding Remarks

Our Contributions

- Set up Markov models for the co-evolution of content popularity and delivery
- Derive o.d.e. limits (*fluid limits*) from the Markov models
- Show that the popularity evolution model under study, is equivalent to the SIR epidemic model
- Use the fluid limits to solve various optimization problems for both the content creator and the content distributor

Future Work

- Decentralized spread of both popularity and the content

Future Work

- Decentralized spread of both popularity and the content
- Incentive schemes for the relay nodes to carry the content
- *Leecher* behavior - nodes can delete the content after consumption

Future Work

- Decentralized spread of both popularity and the content
- Incentive schemes for the relay nodes to carry the content
- *Leecher* behavior - nodes can delete the content after consumption
- Inter-mobile contact duration not sufficient for full transfer

Future Work

- Decentralized spread of both popularity and the content
- Incentive schemes for the relay nodes to carry the content
- *Leecher* behavior - nodes can delete the content after consumption
- Inter-mobile contact duration not sufficient for full transfer
- Generalize the framework for multiple P2P content

Future Work

- Decentralized spread of both popularity and the content
- Incentive schemes for the relay nodes to carry the content
- *Leecher* behavior - nodes can delete the content after consumption
- Inter-mobile contact duration not sufficient for full transfer
- Generalize the framework for multiple P2P content

Funding Organizations

- Indo-French Center for the Promotion of Advanced Research (CEFIPRA)
- Department of Science and Technology, Govt. of India

Thank you