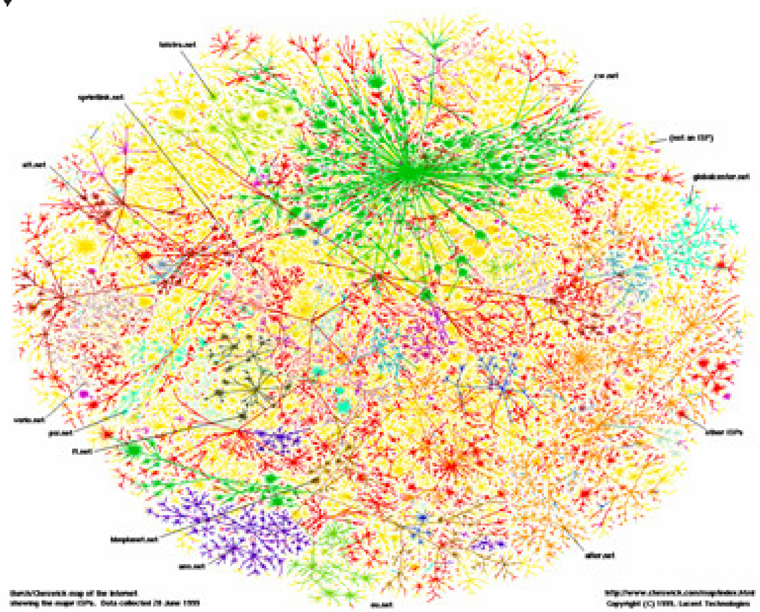
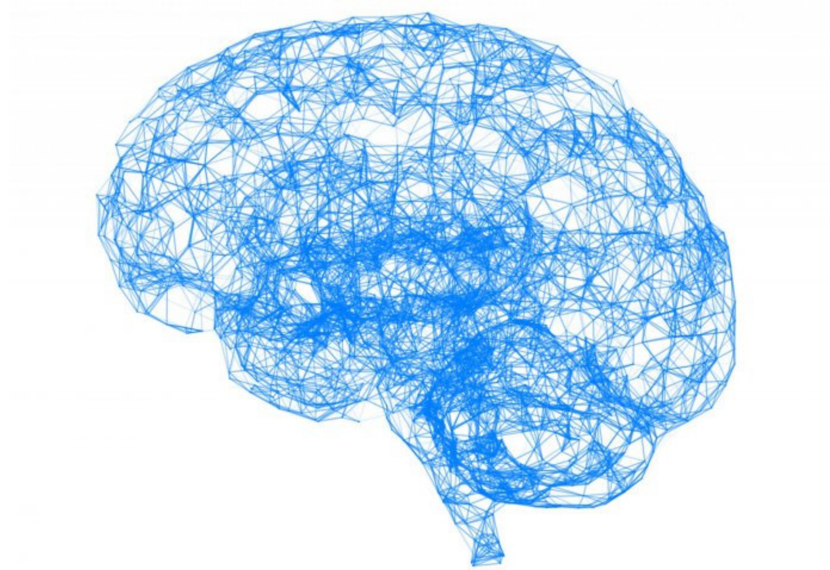
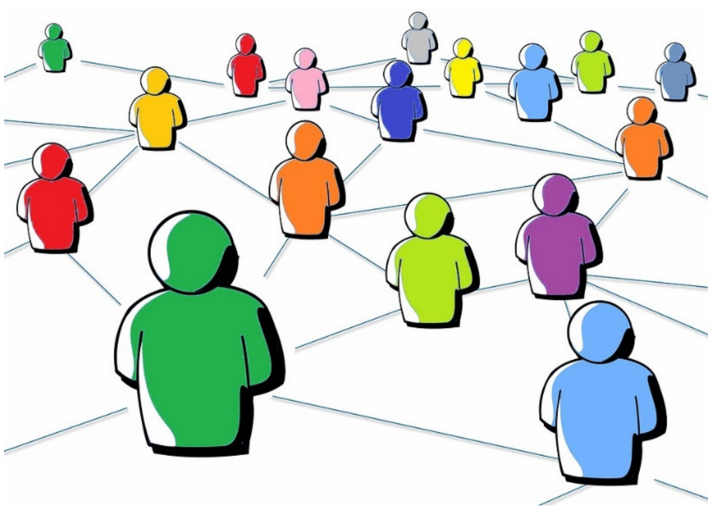
Epidemic Thresholds in Networks: Impact of Heterogeneous Infection Rates and Recovery Rates

Dr. Zesheng Chen
Department of Computer Science
Purdue University Fort Wayne

Epidemic Spread in Complex Networks







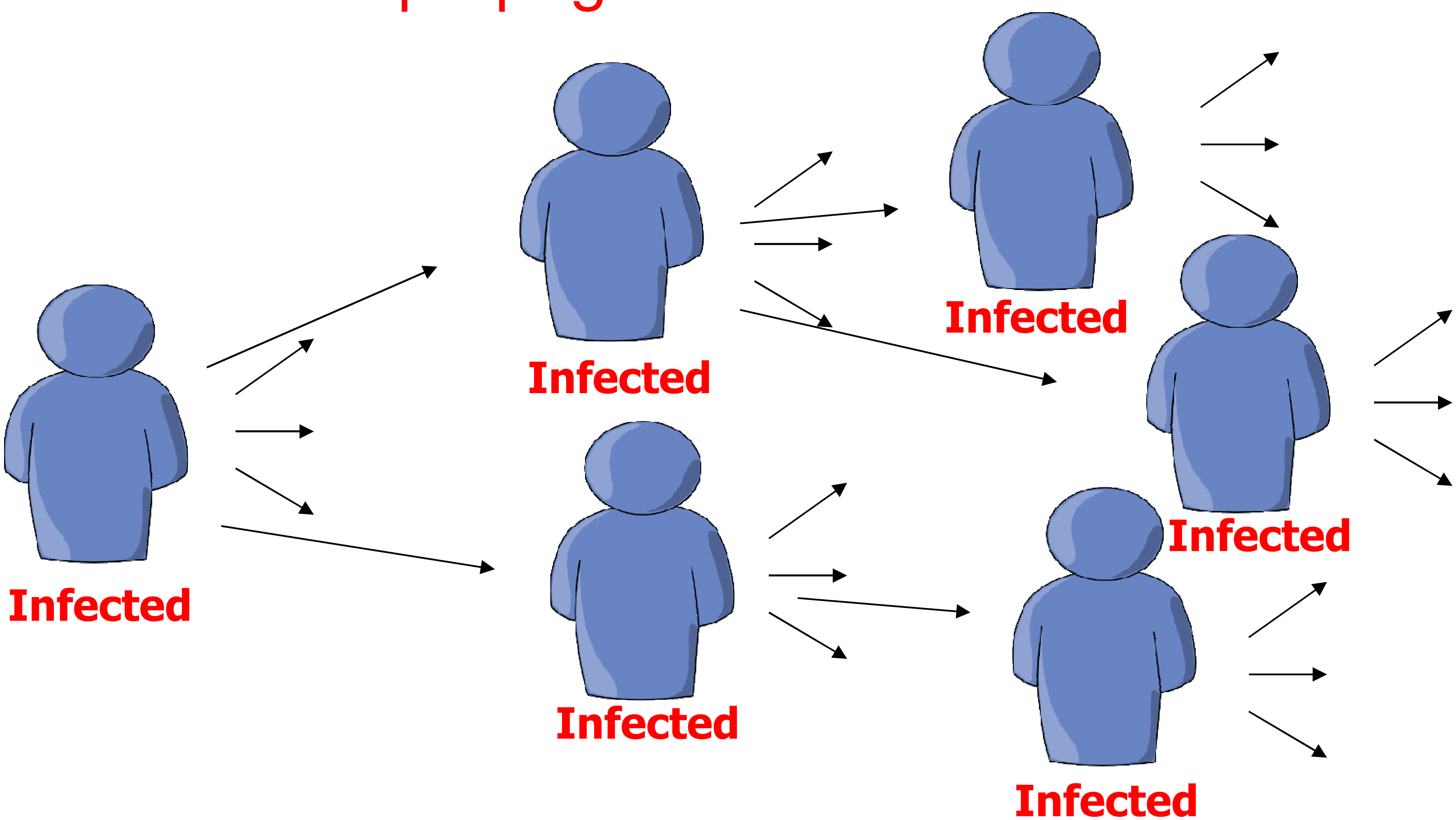
Information and Influence Diffusion in Online Social Networks

- Viral marketing (“word-of-mouth”)
- Blog information cascading
- Rumor spreading
- ……

Bear a resemblance to epidemic process!

Epidemic Process

- Epidemic process is a process that information or infection self-propagates across networks

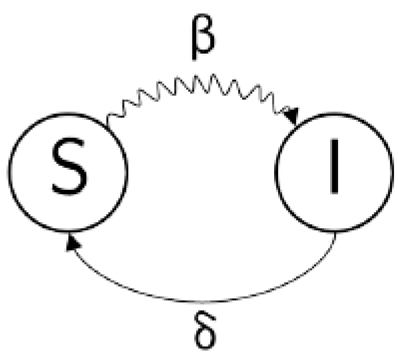


Epidemiological Models

- Susceptible-infected-recovered (SIR) model for independent cascading influence spread
- Susceptible-infected-susceptible (SIS) model for blog information cascading
- Susceptible-infected-cured (SIC) model for rumor and anti-rumor propagation

Epidemic Thresholds

- Fundamental metric used to evaluate epidemic spread
- Condition on which an information/infection will either die out or become epidemic
- In SIS model
  - Infection rate / birth rate
  - Recovery rate / death rate
  - Ratio between infection rate and recovery rate
    - $\beta / \delta >$ epidemic threshold, become epidemic
    - $\beta / \delta <=$ epidemic threshold, die out



Previous Works

- Focused on finding the epidemic threshold when both infection rates and recovery rates are homogeneous.
- Studied the transient behavior of an epidemic when the infection rates are heterogeneous.

The impact of both heterogeneous infection rates and heterogeneous recovery rates on the epidemic threshold has not been studied systematically yet.

Key Question

- Compared with the homogeneous case, how do heterogeneous infection rates and recovery rates affect epidemic thresholds?

Outline

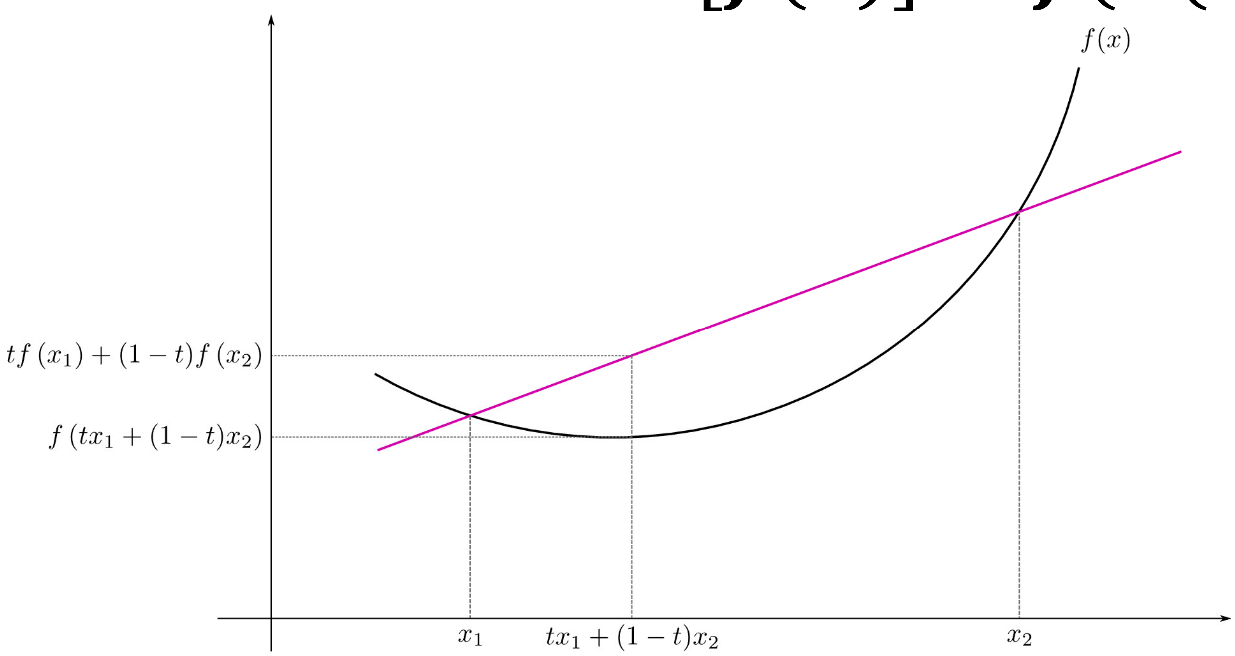
- Intuition
- Theoretical Analysis
- Simulation Verification

Intuition

$$E[f(x)] \neq f(E(x))$$

- Jensen's inequality
  - If *f(x)* is a convex function

$$E[f(x)] \geq f(E(x))$$



From: Wikipedia page of Jensen's inequality

Outline

- Intuition
- Theoretical Analysis
- Simulation Verification

Impact of Heterogeneous Recovery Rates

- δ_i: Recovery rate for node *i*, random variable with mean δ and variance σ_d^2 ($\sigma_d^2 \geq 0$).
- T_d: Time that it takes for node *i* to be recovered.

$$P(T_d = k \mid \delta_i) = \delta_i(1 - \delta_i)^{k-1}, \qquad k = 1, 2, \ldots$$

$$E[T_d|\delta_i] = \frac{1}{\delta_i}$$

$$E[T_d] = E[E[T_d|\delta_i]] = E\left[\frac{1}{\delta_i}\right]$$

$$E[T_d] = E\left[\frac{1}{\delta_i}\right] \geq \frac{1}{E[\delta_i]} = \frac{1}{\delta}$$

according to the Jensen's inequality

Impact of the Degree of Heterogeneity in Recovery Rates

- Use the Taylor expansion on $f(x) = 1/x$ at point a:

$$f(x) = \frac{1}{a} - \frac{x-a}{a^2} + \frac{(x-a)^2}{a^3} + H$$

- Set $x = \delta_i$ and $a = E[\delta_i] = \delta$:

$$\frac{1}{\delta_i} \approx \frac{1}{\delta} - \frac{\delta_i - \delta}{\delta^2} + \frac{(\delta_i - \delta)^2}{\delta^3}, \qquad E\left[\frac{1}{\delta_i}\right] \approx \frac{1}{\delta}\left(1 + \frac{\sigma_d^2}{\delta^2}\right)$$

- It is statistically easier for the epidemic to survive with a higher degree of the heterogeneity in recovery rates

Impact of Heterogeneous Infection Rates

- β_{ij}: Infection rate for node *i* to infect node *j*, random variable with mean β and variance σ_b^2 ($\sigma_b^2 \geq 0$).
- *K:* (Random) number of neighbors of node *i* that are infected by node *i* during *t* time steps.

$$X_j = \begin{cases} 1, & \text{if neighbor } j \text{ is infected by node } i \text{ during } t \text{ time steps;} \\ 0, & \text{otherwise.} \end{cases}$$

$$P\left(X_j = 1 \,\middle|\, \beta_{ij}\right) = 1 - \left(1 - \beta_{ij}\right)^t$$

$$K = \sum_{j \in N_i} X_j \qquad E\left[K \middle| \beta'_{ij}s\right] = \sum_{j \in N_i} E\left[X_j \middle| \beta_{ij}\right] = m - \sum_{j \in N_i} \left(1 - \beta_{ij}\right)^t$$

$$E[K] = E\left[E\left[K \middle| \beta'_{ij}s\right]\right] = m - \sum_{j \in N_i} E\left[\left(1 - \beta_{ij}\right)^t\right]$$

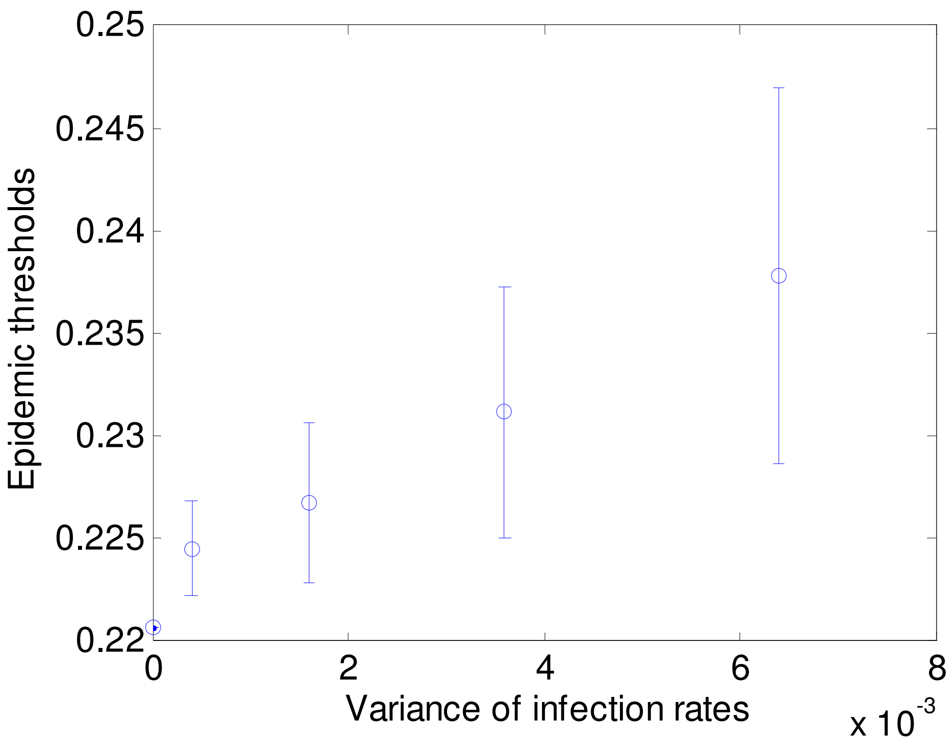
Outline

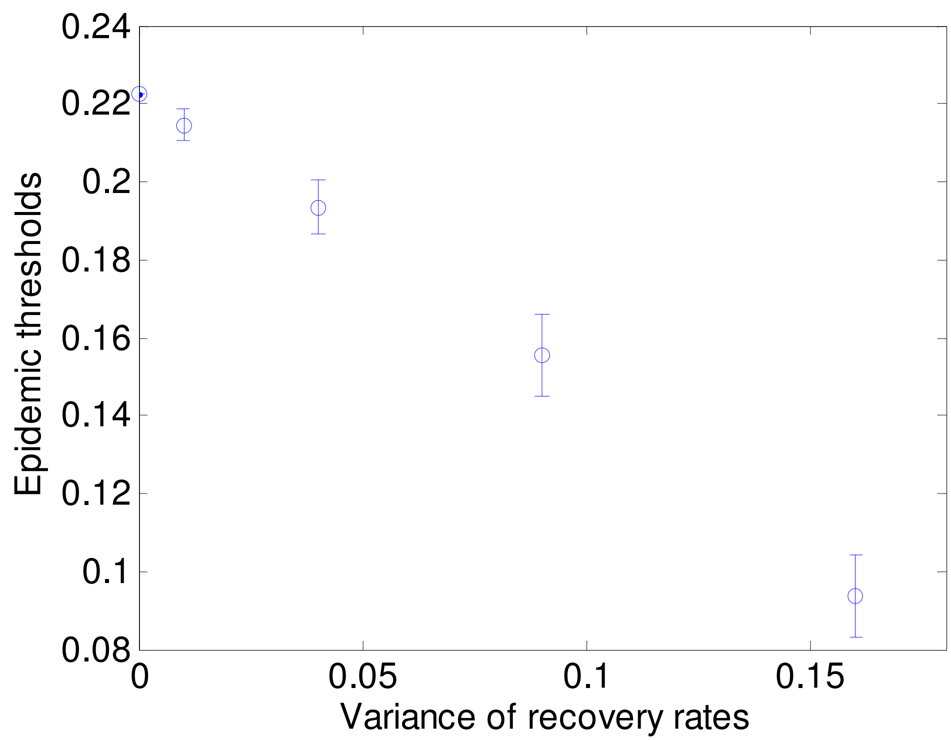
- Intuition
- Theoretical Analysis
- Simulation Verification

Simulation Setup

- Simulator is based on discrete time and random number generator
- Run 100 times for each scenario
- Run long enough so that it reaches the steady state
- Randomly select half of nodes to be infected initially
- Study epidemic thresholds in a lattice, a BA power-law topology, and a real topology

Performance Evaluation (Coauthorship Network)





Conclusion (Our Discovery)

- The heterogeneity in infection rates leads to a larger epidemic threshold than in the homogeneous case
  - Moreover, as the degree of the heterogeneity of infection rates gets higher, the epidemic threshold increases
- The heterogeneity in recovery rates generates a smaller epidemic threshold than in the homogeneous case
  - The epidemic threshold decreases as the degree of the heterogeneity of recovery rates gets higher

Thanks For Your Attention