

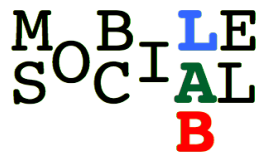
# Lecture 13: Infect (3/4)

How do epidemic and gossip reach people?  
(i.e., how computer viruses spread?)

COMS 4995-1: Introduction to Social Networks

Part II: Cascade

Tuesday, October 25<sup>th</sup>



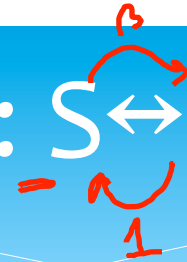
# Epidemic model #2: $S \xrightarrow{\beta} I \xrightarrow{1} R$

- \* Thm: Assuming  $\beta\rho < 1$ ,  $E[|Y(\infty)|] \leq C \sqrt{N} / (1 - \beta\rho)$ 
  - $\rho(G)$ : largest eigenvalue of  $G$ 's adjacency matrix
  - $C = \sqrt{\#\{\text{initial infected population}\}}$
- \* If  $\beta\rho < 1$  and  $C = o(\sqrt{N})$ , negligible fraction removed
- \* Examples:
  - $G$  is  $d$ -regular :  $\rho(G) = d$
  - Can be applied to bound unif. random graphs

# Outline

- \* Continuous epidemics, “logistic model”
- \* Discrete epidemics, “graph”
  - Adjacency matrix
  - SI, SIR model
  - / – SIS
- \* Epidemic algorithms

# Epidemic model #3: $S \leftrightarrow I$



- \* Nodes follow neighbor contamination / recovery
  - Node  $u \in V$  infectious ( $X_u = 1$ ) or susceptible ( $X_u = 0$ )
  - Node  $u$  becomes infected with rate  $\beta \cdot \sum_{v \in N(u)} X_v$
  - Node  $u$  recovers with rate  $\gamma=1$

- \* In a finite graph, all nodes eventually recover
  - Because  $(X_u = 0 \ \forall u \in V)$  is the only absorbing state
  - Different on infinite graphs (e.g. lattices, trees)

$2^N$



# Epidemic model #3: $S \leftrightarrow I$

$$X(t) = (X_v(t))_{v \in V}$$

$$e^{t(\beta\rho-1)} = \frac{1}{n^2}$$

\* Can we recover fast from an epidemic?

\* Thm:  $P[X(t) \neq (0, \dots, 0)] \leq C \sqrt{N} \exp(t \cdot (\beta\rho - 1))$

–  $\rho(G)$ : largest eigenvalue of  $G$ 's adjacency matrix

–  $C = \sqrt{\#\{\text{initial infected population}\}}$

\* Corollary: If we have  $\beta\rho < 1$

–  $E[\text{extinction time}] \leq (1 + \ln(n)) / (1 - \beta\rho)$

\* Bottom line: goes to zero very fast if  $\beta\rho < 1$

– complete graph:  $\rho(G) = n - 1$       $\beta < \frac{1}{n-1}$

– uniform random graph:  $\rho(G) \approx (n-1)p$  (if  $np = \omega(\log n)$ )

$$\beta\rho < 1 \quad \left( \frac{2 \ln n}{n p} < 1 \right)$$

$$P[\tau \geq t] \leq C \sqrt{N} e^{t(\beta\rho-1)}$$

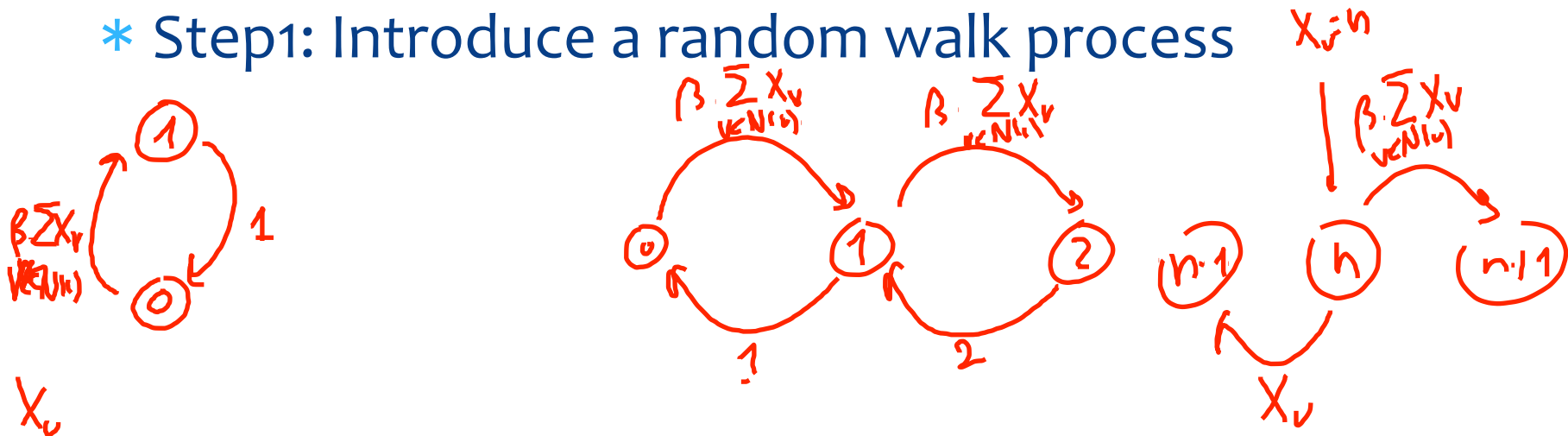
$$\tau = \inf \{ t \geq 0 \mid X(t) = (0, 0, \dots, 0) \}$$

$$E[\tau] = \int_0^\infty P[\tau \geq s] ds$$

The effect of network topology on the spread of epidemics,  
A Ganesh, L Massoulié, D Towsley, IEEE InfoCom (2005)

# Proof:

\* Step1: Introduce a random walk process



\* Intuitively we have  $P[X(t) \neq 0] \leq P[Z(t) \neq 0]$

– This statement can be made precise by coupling

$$e^z = \sum_{k \geq 0} \frac{z^k}{k!} \quad e^{tA} = \sum_{k \geq 0} \frac{(tA)^k}{k!}$$

# Proof:

$$P[X(t) \neq 0] \leq$$



$$(AX)_u = \sum_{v \in V} A_{uv} X_v = \sum_{v \in N(u)} X_v$$

$$(XA)_u = \sum_{v \in V} X_v A_{vu} = \sum_{v \in N(u)} X_v$$

\* We have  $P[Z(t) \neq 0] \leq \sum_v P[Z_v(t) \neq 0] \leq \sum_v E[Z_v(t)] \leq \text{Des. ineq.}$   
 – Using Markov Inequality

\* We can rewrite evolution of  $Z(t)$

$$A = \begin{cases} A_{uv} = 1 & \text{if } uv \in E \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{d}{dt} E[Z_u(t)] = \beta \sum_{v \in N(u)} E[Z_v(t)] - E[Z_u(t)]$$

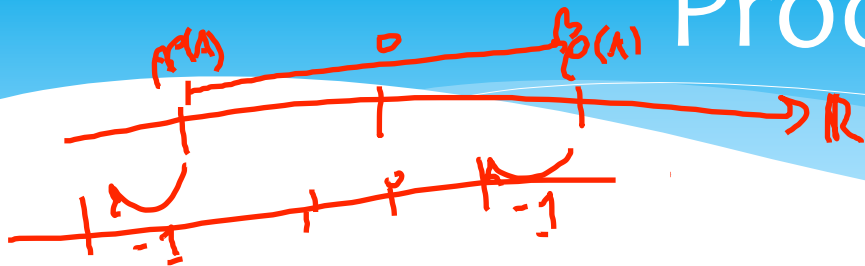
$$\frac{d}{dt} \vec{z}(t) = \beta \sum_{v \in N(u)} \vec{z}_v(t) - \vec{z}(t) = (\vec{z} \cdot \beta A)_u - \vec{z}(t) = \vec{z} (\beta A - I)$$

\* So that  $P[Z(t) \neq 0] \leq \|e_1\| \|\exp(t \cdot (\beta A - I)) X(0)\|$

$$\frac{d\vec{z}}{dt} = \vec{z} (\beta A - I) \Leftrightarrow \vec{z} = \exp((\beta A - I)t) \vec{z}(0)$$

$$\frac{dz}{dt} = z \cdot c \Rightarrow z = e^{ct}$$

# Proof:



$$z(0) = \mathbb{E}[z(0)] = X(0)$$

\* Finally, we can apply the same bounding technique

$$\frac{dz}{dt} = z \cdot (BA - Id) \Rightarrow z(t) = \exp(t \cdot (BA - Id)) X(0)$$

$$P[X(t) \neq 0] \leq \sum_v \mathbb{E}[z_v(t)] = \langle e_1, z_v(t) \rangle =$$

$$\leq \|e_1\| \times \|\exp(t \cdot (BA - Id))\| \cdot \|X(0)\|$$

$$\sqrt{N} \rho(\exp(t \cdot (BA - Id))) \leq (1)$$

$A: \begin{cases} 1 & \text{if } v, v \text{ are } 0 \\ 0 & \text{otherwise} \end{cases}$   
 $(0, 1)$

$$A: \rho(A) = \{\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n\}$$

$$(BA - Id): \rho(BA - Id) = \{(\beta \lambda_1 - 1), (\beta \lambda_2 - 1), \dots, (\beta \lambda_n - 1)\}$$

$$\exp(t \cdot (BA - Id)) = \rho(BA - Id) e^{t(\beta \lambda_1 - 1)} e^{t(\beta \lambda_2 - 1)} \dots$$

$$\exp(t \cdot (BA - Id)) = \sum_{k \geq 0} \frac{A^{k+1}}{k!} t^k = \sum_{k \geq 0} \frac{A^{k+1}}{k!} t^k e^{t \cdot (BA - Id)}$$



# Discrete epidemics: summary

Type	Outcomes
$S \rightarrow I$	Everyone infected
$S \leftrightarrow I$	No infectious nodes
$S \rightarrow I \rightarrow R$	No infectious node

Follow processes of infection

- Initial conditions:  
small set infected nodes

Outcomes generally trivial

- Speed or span depend on  
graph topology  
(e.g. spectral analysis)

# Outline

- \* Continuous epidemics, “logistic model”
- \* Discrete epidemics, “graph”
- \* Epidemic algorithms

# Epidemic Algorithms

- \* Replicated database maintenance
  - Different versions, many locations
  - How to handle communication? failures ?
- \* 1987 “Epidemic alg., rumor spreading, gossip”
  - Do not maintain fixed communication topology
  - Contact a node unif., spread if one node has a copy
- \* How many rounds  $S_n$  before rumor spreads to all
  - $S_n = (1+1/\ln(2)) \log(n) + O(1)$  in probability

On spreading a rumor, B. Pittel, SIAM J. Appl. Math. (1987)

Epidemic algorithms for replicated database maintenance,  
A Demers et. al, ACM PODC. (1987)

# How gossip compares to optimal?

- \* A binary tree:

- Also takes time  $O(\log(n))$ , using  $O(n)$  messages
- Seems optimal in both ways, but prone to failure

- \* Gossip:

- Time  $O(\log(n))$  (optimal) and  $O(n \log n)$  messages
- In fact, unif. gossip requires at least  $\omega(n)$  messages, and  $\Omega(n \log \log(n))$  if no addresses are kept (the latter can be attained)

Randomized rumor spreading,  
R Karp and C Schindelhauer and S Shenker and B Vocking, FOCS. (2000)