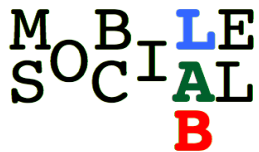


Lecture 16: Infect (3/4)

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

COMS 4995-1: Introduction to Social Networks
Thursday, November 8th



Epidemic model #2: $S \rightarrow I \rightarrow R$

$$\begin{matrix} \lambda & \gamma \\ X_0(t) & Y_0(t) \end{matrix}$$

- * 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 (same degree): $\rho(G) = d$
 - Can be applied to bound unif. random graphs

Proof: recap of step1

* By counting all chains of infection from v to u

$$\mathbb{E}[|Y(\infty)|] = \sum_{u \in V} P[Y_u(\infty) = 1] \leq \sum_{u \in V} \left(\sum_{v \in V} X_v(0) \cdot \sum_{t \geq 0} \beta^t A_{v,u}^t \right)$$

- Because $A_{v,u}^t$ is # sequences $v=u_0, u_1, \dots, u_t=u$
- The chance that each sequence succeeds is β^t
- And probability of union event \leq sum of probability

Proof: rewriting

$$\sum_v X_v(0) \cdot \left(\sum_{t \geq 0} \beta^t A_{v,u}^t \right)$$

$$\mathbb{E}[|Y(\infty)|] = \sum_{u \in V} P[Y_u(\infty) = 1] \leq \sum_{u \in V} \sum_{v \in V} X_v(0) \cdot \sum_{t \geq 0} \beta^t A_{v,u}^t$$

* Rewrite as $\mathbb{E}[|Y(\infty)|] \leq \left\langle e_1, \left(\sum_{t \geq 0} (\beta A)^t \right) X(0) \right\rangle$

– Same using vector/matrix notation

– Where $\langle x, y \rangle = \sum_{i=1}^N x_i \cdot y_i$ and $e_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

Some definition on norms

* We introduce norms

– vectors $\|v\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$ Note it implies $\langle x, y \rangle \leq \|x\|_2 \cdot \|y\|_2$

– and matrices: $\|A\|_2 = \max_{x \in \mathbb{R}^n} \frac{\|A \cdot x\|_2}{\|x\|_2}$

– This implies $\|A \cdot x\|_2 \leq \|A\|_2 \cdot \|x\|_2$

– This also implies $\|A\| = \rho(A)$

$$\|AA^{-1}\| = \|I\| = 1$$
$$\|A\| \times \|A^{-1}\|$$

Proof: completing

$$\mathbb{E}[\|Y(\infty)\|] \leq \left\langle e_1, \sum_{t \geq 0} (\beta A)^t X(0) \right\rangle$$

* We deduce $\mathbb{E}[\|Y(\infty)\|] \leq \|e_1\| \times \left\| \sum_{t \geq 0} (\beta A)^t \right\| \times \|X(0)\|$

* Note that $\sum_{t \geq 0} (\beta A)^t = (\text{Id} - \beta A)^{-1}$ $\sum = 1 + \beta A + (\beta A)^2 + \dots$
 $\frac{1}{1-r}$

– First, the series converge as $\beta \rho(A) < 1$

– Second, we can verify it is the inverse of

$$(\text{Id} - \beta A) \cdot \sum_{t \geq 0} (\beta A)^t = 1 + \beta A + (\beta A)^2 + \dots - \beta A - \beta A^2 - \beta A^3$$

Conclusion

$$\|X(0)\| = \sqrt{\sum_{v \in V} X_v(0)^2}$$

* We hence have

$$\mathbb{E}[\|Y(\infty)\|] \leq \|e_1\| \times \|(\text{Id} - \beta A)^{-1}\| \times \|X(0)\|$$

– From there, we can conclude the theorem

– $\mathbb{E}[\|Y(\infty)\|] \leq C \sqrt{N} / (1 - \beta \rho)$ where $C = \sqrt{\text{initial inf. pop.}}$

$$\|I - \beta A\| \geq 1 - \beta \rho$$

$$\|(\text{Id} - \beta A)^{-1}\| \leq \frac{1}{1 - \beta \rho}$$

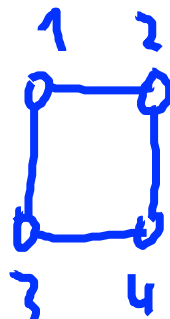
Epidemic model #2: $S \rightarrow I \rightarrow R$

- * Thm: Assuming $\beta\rho < 1$, $E[|Y(\infty)|] \leq C \sqrt{N} / (1 - \beta\rho)$
 - $\rho(G)$: largest eigenvalue of G 's adjacency matrix
 - $C = \sqrt{\#\{\text{initial infected population}\}}$

- * Examples of application

- G is d -regular (same degree)?
- G is a complete graph?
- G is a star network?
- G a uniform random graph?

$\rho(G) = d$
 $\rho(G) = \sqrt{N-1}$

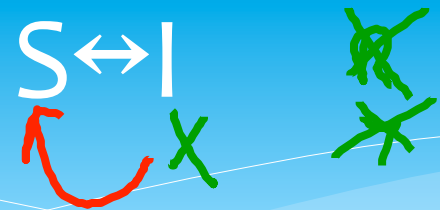


$$\begin{matrix}
 & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\
 \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}
 \end{matrix}$$

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)

Epidemic model #3: $S \leftrightarrow I$

- * 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 $\beta\rho < 1$, choosing $t = \ln(n) / (1 - \beta\rho)$ we can prove
 - $E[\text{extinction time}] \leq (1 + \ln(n)) / (1 - \beta\rho)$
- * Bottom line: goes to zero very fast if $\beta\rho < 1$
 - complete graph: $\rho(G) = n - 1$
 - uniform random graph: $\rho(G) \approx (n - 1)p$ (if $np = \omega(\log n)$)

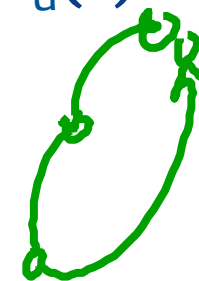
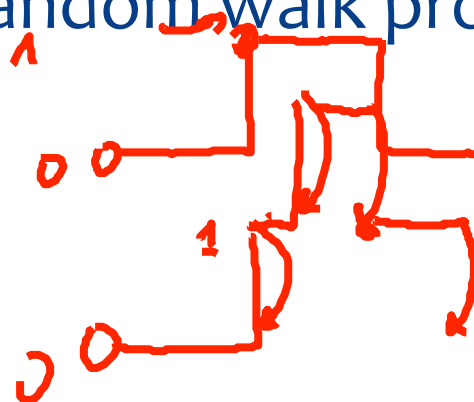
The effect of network topology on the spread of epidemics,
A Ganesh, L Massoulié, D Towsley, IEEE InfoCom (2005)

Proof:

* Step1: Introduce a random walk process $Z_u(t)$

$$X_u(t) = 0, 1$$

$$Z_u(t) = 0, 1, 2, \dots$$



> 0

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

– This statement can be made precise by coupling

– Note: $P[Z(t) \neq 0] \leq \sum_v P[Z_v(t) \neq 0] \leq \sum_v E[Z_v(t)]$

$$Z(t) \neq (0, 0, 0, \dots, 0)$$

$$\leq \sum_v P(Z_v(t) \geq 1)$$

Method

Proof:

- * How does $Z(t)$ evolve?

$$\frac{d}{dt} Z_u(t) = \sum_v \beta A_{u,v} Z_v(t) - Z_u(t)$$

- This is a linear evolution!

- * In expectation, it is $\frac{d}{dt} \mathbb{E}[Z_u(t)] = \sum_v \beta A_{u,v} \mathbb{E}[Z_v(t)] - \mathbb{E}[Z_u(t)]$

- This is a linear deterministic evolution (N dimension)

- Which is $\mathbb{E}[Z(t)] = e^{t(\beta A - \text{Id})} Z(0) = e^{t(\beta A - \text{Id})} X(0)$

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

Proof:

* Finally, we can apply the same bounding technique

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

Handwritten annotations in green:

- A bracket under $\|e_1\|$ is labeled \sqrt{n} .
- A bracket under $\|\exp(t \cdot (\beta A - I)) X(0)\|$ is labeled \leq .
- Below the main expression, it is written: $\leq e^{\gamma \cdot \beta \rho(n) \cdot t} \|X(0)\|$.
- To the right, it is written: $\leq C \cdot \sqrt{\ln n}$.

Handwritten notes in green:

- $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$
- $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$

Discrete epidemics: summary

Type	Outcomes
$S \rightarrow I$	Everyone infected
$S \leftrightarrow I$	No infectious nodes <i>inf. graph</i>
$S \rightarrow I \rightarrow R$	No infectious node

Follow processes of infection

- Initial conditions:
small set infected nodes

finite graph Outcomes generally trivial

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