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→I→R

x $\quad$ y

$X_v(t)$ $\quad$ $Y_v(t)$

* Thm: Assuming $\beta\rho<1$, $E[\,|Y(\infty)|\,] \le C\sqrt{N} / (1-\beta\rho)$
  - $\rho(G)$: 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

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

  – Because $A^t_{v,u}$ is # sequences $v=u_0,u_1,\ldots,u_t=u$
  – The chance that each sequence succeeds is β^t
  – And probability of union event ≤ sum of probability

Proof: rewriting

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

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

* Rewrite as $\mathbb{E}[|Y(\infty)|] \leq \left\langle e_1, \sum_{t \geq 0} (\beta A)^t X(0) \right\rangle$
  - 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: $||A||_2 = \max_{x \in \mathbb{R}^n} \frac{||A.x||_2}{||x||_2}$
  - This implies $||A \cdot x||_2 \leq ||A||_2 \cdot ||x||_2$
  - This also implies $||A|| = \rho(A)$

$||AA^{-1}|| = ||1|| = 1$

$\leq$

$||A|| \times ||A^{-1}||$

Proof: completing

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

- We deduce $\mathbb{E}[||Y(\infty)||] \leq ||e_1|| \times ||\sum_{t\geq 0} (\beta A)^t|| \times ||X(0)||$
- Note that $\sum_{t\geq 0} (\beta A)^t = (\underbrace{\mathrm{Id} - \beta A}) ^{-1}$ $\sum = 1 + \beta A + (\beta A)^2 + \dots$ $\frac{1}{1-r}$
  - First, the series converge as $\beta\rho(A) < 1$
  - Second, we can verify it is the inverse of

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

CS@CU

Conclusion

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

* We hence have

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

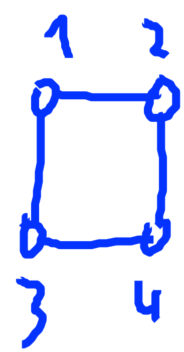
  – From there, we can conclude the theorem
  – E[|Y(∞)|] ≤ C √N / (1-βρ) where C=√initial inf. pop.

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

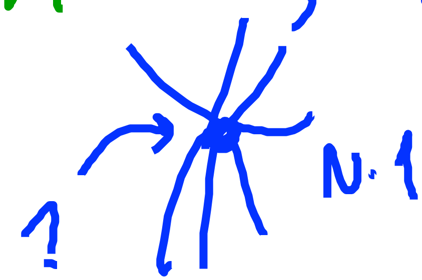
$||(Id - \beta A)^{-1}|| \leq \frac{1}{1-\beta\rho}$

Epidemic model #2: S→I→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)? $\rho(G)=d$
  – G is a complete graph? $\rho(G)=N-1$
  – G is a star network? $\rho(G)=\sqrt{N}$
  – G a uniform random graph?



$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \\ 3 & 1 & 0 & 0 & 1 \\ 4 & 0 & 1 & 1 & 0 \end{array}$$



CS@CU

Outline

* Continuous epidemics, "logistic model"
* Discrete epidemics, "graph"
  - Adjacency matrix
  - SI, SIR model
  - SIS

* Epidemic algorithms

Epidemic model #3: S↔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 γ=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↔I

* Can we recover fast from an epidemy?
* Thm: $P[X(t) \neq (0,..,0)] \leq C \sqrt{N} \exp(t \cdot (\beta\rho-1))$
  – $\rho(G)$: 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 $np = \omega(\log n)$)

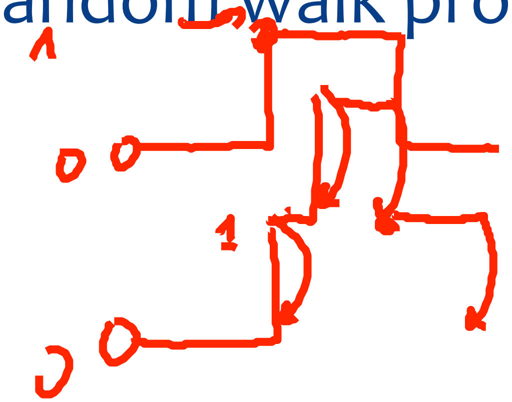
The effect of network topology on the spread of epidemics,
A Ganesh, L Massoulié, D Towsley, IEEE Infocom (2005)

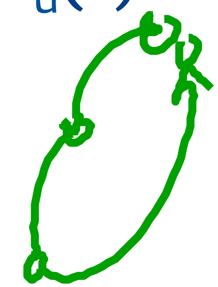
Proof:

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

$X_v(t) = 0, 1$

$Z_v(t) = 0, 1, 2, \ldots$





>0

* Intuitively we have $P[X(t) \neq 0] \leq P[Z(t) \neq 0]$
  - This statement can be made precise by coupling
  - Note: $P[Z(t) \neq \vec{0}] \leq \Sigma_v P[Z_v(t) \neq 0] \leq \Sigma_v E[Z_v(t)]$ Markov

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

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

CS@CU

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 $\mathbb{E}[Z(t)] = e^{t(\beta A - \mathrm{Id})} Z(0) = e^{t(\beta A - \mathrm{Id})} X(0)$
* 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
  – $P[X(t) \neq 0] \leq \underbrace{\|e_1\|}_{\sqrt{n}} \; \underbrace{\|\exp(t \cdot (\beta A - I))\, X(0)\|}_{\leq}$

$$\leq \underbrace{\|\exp(t(\beta A - I))\|}_{\leq e^{t \cdot \beta\rho(A) - 1}} \; \underbrace{\|X(0)\|}_{\leq C = \sqrt{\text{inf}\,\exp}}$$

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

$$e^A = \sum_{k \geq 0} \frac{A^k}{k!}$$

Discrete epidemics: summary

Type	Outcomes
S→I	Everyone infected
S↔I	No infectious nodes finite graph inf. graph
S→I→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)