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
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
* 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 \( v = u_0, u_1, \ldots, u_t = u \)
− The chance that each sequence succeeds is \( \beta^t \)
− And probability of union event \( \leq \) sum of probability
Proof: rewriting

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

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

- 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 $\|A\| = \rho(A)$
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

\[ \sum_{t \geq 0} (\beta A)^t = (\text{Id} - \beta A)^{-1} \]

  \[ \sum_{t \geq 0} \approx \frac{1}{1 - \gamma} \]

  – First, the series converge as \( \beta \rho(A) < 1 \)
  – Second, we can verify it is the inverse of

\[ \sum_{t \geq 0} (\beta A)^t = 1 + \beta A + (\beta A)^2 + \ldots \]

\[ = \frac{1}{1 - \beta A} \]

\[ = \frac{1 + \beta A + (\beta A)^2 + \ldots}{1 - \beta A} \]

\[ = \frac{1}{1 - \beta A} \]

\[ = \frac{1}{1 - \beta A} \]

\[ = \frac{1}{1 - \beta A} \]
We hence have

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

- From there, we can conclude the theorem
- $$E[ |Y(\infty)| ] \leq C \sqrt{N} / (1-\beta \rho)$$ where C=\text{initial inf. pop.}
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)?
  
  – $G$ is a complete graph?
  
  – $G$ is a star network?
  
  – $G$ a uniform random graph?
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)
Can we recover fast from an epidemic?

Thm: \( P[X(t) \neq (0,\ldots,0)] \leq C \sqrt{N} \exp(t \cdot (\beta \rho - 1)) \)
- \( \rho(G) \): 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 Info\( \tilde{c} \)om (2005)
Step 1: Introduce a random walk process $Z_u(t)$

- $X_u(t) = 0, 1$
- $Z_u(t) = 0, 1, ?, \ldots$

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

- This statement can be made precise by coupling
- Note: $P[Z(t) \neq 0] \leq \sum_v P[Z_v(t) \neq 0] \leq \sum_v E[Z_v(t)]$

Proof:

$Z_u(0, 0, 0, \ldots, 0) \neq (0, 0, 0, \ldots, 0)$
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 $P[Z(t) \neq 0] \leq \|e_1\| \|\exp(t \cdot (\beta A - \text{I})) X(0)\|$
Finally, we can apply the same bounding technique
- \( P[X(t) \neq 0] \leq \|e_1\| \|\exp(t \cdot (\beta A - I)) \cdot X(0)\| \)
Discrete epidemics: summary

Follow processes of infection

- Initial conditions:
  small set infected nodes

Outcomes generally trivial

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

<table>
<thead>
<tr>
<th>Type</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>S→I</td>
<td>Everyone infected</td>
</tr>
<tr>
<td>S↔I</td>
<td>No infectious nodes</td>
</tr>
<tr>
<td>S→I→R</td>
<td>No infectious node</td>
</tr>
</tbody>
</table>