Lecture 12: Infect (2/4)

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

COMS 4995-1: Introduction to Social Networks
Wednesday, November 7th
Outline

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

* Epidemic algorithms
Infection only spreads along edges of a given graph
  - to account for connections, closeness among nodes
Initial conditions: one or several nodes infected

Challenges:
  - Can a fraction be infected in a large graph?
  - What is the speed of evolution of epidemics?
  - How does it depend on the properties of the graph?
  - What if some individuals are immune?
A graph $G=(V,E)$ with $N$ nodes

- Adjacency matrix $A$ is $N \times N$ matrix
  - $A_{uv} = 1$ if $(u,v)$ in $E$, 0 otherwise
  - $A_{uu} = 1$ (for $u \neq u$, $A_{uu} = 0$)

Prop 1: for any $k$, $(A^k)_{uv}$ is $\#\text{ paths } u \rightarrow v \text{ of length } k$
Prop2: A is symmetric, with non-negative entries

THM 1 (Perron Frobenius):
- Eigenvalues of A are all real \( \lambda_1 \geq \lambda_2 \geq \ldots \)
- \( \lambda_1 \) is the largest in absolute value (\( \lambda_1 \geq |\lambda_i| \) for any \( i \))
- \( \lambda_1 \) has a non-negative eigenvector \( x_1 \)

Generally \( \lambda_1 \) is called the spectral radius of A, \( \rho(A) \)
Outline

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

* Epidemic algorithms
Epidemic Model #1: $S \rightarrow I$

* Model 1: broadcast
  - Node infected at time $t$ infects all its neighbors in $t+1$
  - Within time $D \leq \text{diam}(G)$, all nodes are infected

* Model 2: gossip
  - Node infects each neighbor with a given rate $\beta$
  - Eventually all nodes are infected within $O(D/\beta)$ time
  - What if rates reduces with degree? (see later lecture)
Model 1: single infection attempt
- Infected node infect neighbors with probability $\beta$
- Many names: “Independent cascade model”, Reed- Frost epidemics, SIR with single slot

Model 2: Random infectious period (normalized)
- Similar (probability to spread is $\beta$) but dependencies!

Eventually: no infectious nodes, fraction removed
What is the size of the removed fraction?

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 $C = o(\sqrt{N})$, remove only negligible fraction

Proof based on expectation not on independence
- Similar results hold for model 2
\[ X_v \text{ (resp. } Y_v \text{ )} = 1 \text{ iff } v \text{ is infected (resp, inf or recov.)} \]

First, we bound \( Y_v(t) \) using adjacency matrix.
$y(0) = 0$

$x(0) = 1$ if $v$ infected, 0 otherwise.

$E[|y(\infty)|] \leq \sum_{u \in E} \sum_{v \in E} x_{v}(0) \sum_{t \geq 0} \beta A_{vu}$
Proof

* Second, we use a bounding norm technique