

Online Social Networks and Media

Epidemics and Influence

Epidemics

Understanding the spread of viruses and epidemics is of great interest to

- Health officials
- Sociologists
- Mathematicians
- Hollywood

The underlying **contact network** clearly affects the spread of an epidemic

Model epidemic spread as a **random process** on the graph and study its properties

- Main question: will the epidemic take over most of the network?

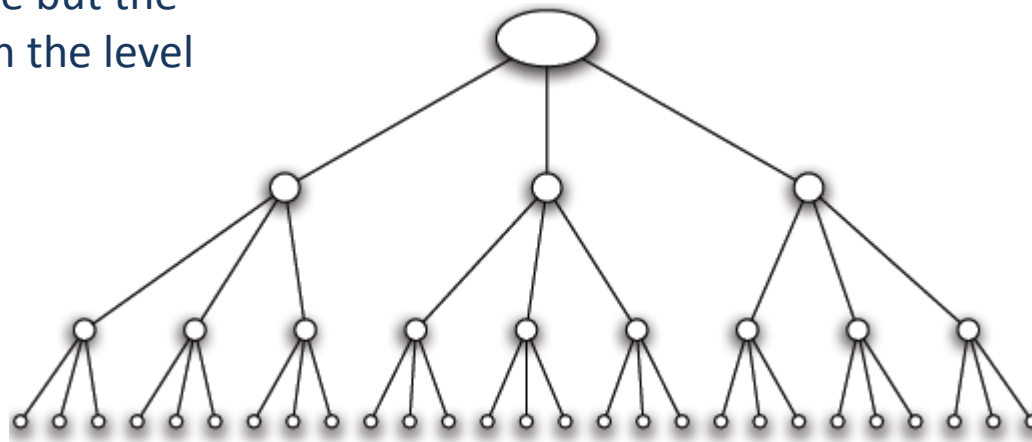
Diffusion of ideas and the **spread of influence** can also be modeled as epidemics



Branching Processes

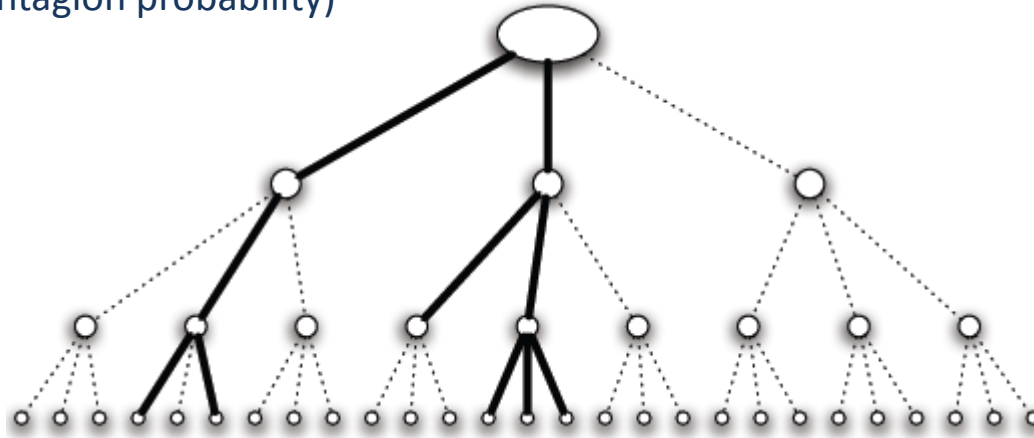
- A person transmits the disease to each people she meets *independently with a probability p*
 - Meets *k people* while she is contagious
1. A person carrying a new disease enters a population, first *wave* of k people
 2. Second wave of k^2 people
 3. Subsequent waves

A contact network with $k = 3$
Tree (root, each node but the root, a single node in the level above it)



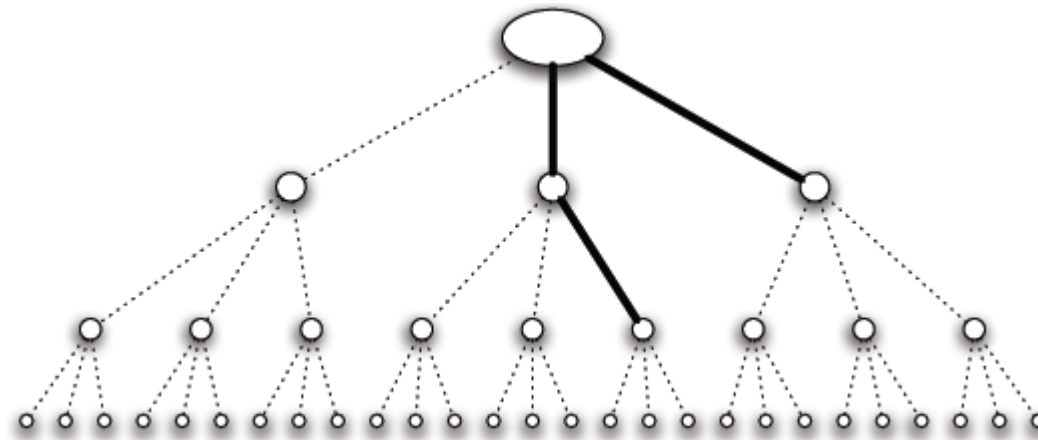
Branching Processes

Aggressive epidemic (high contagion probability)



Mild epidemic (low contagion probability)

- If it ever reaches a wave where it infects no one, then it dies out
- Or, it continues to infect people in every wave infinitely



Branching Processes: Basic Reproductive Number

Basic Reproductive Number (R_0): the expected number of new cases of the disease caused by a single individual

Claim: (a) If $R_0 < 1$, then with probability 1, the disease dies out after a finite number of waves. (b) If $R_0 > 1$, then with probability greater than 0 the disease persists by infecting at least one person in each wave.

$$R_0 = pk$$

- (a) $R_0 < 1$ -- Each infected person produces less than one new case in expectation
Outbreak constantly trends downwards
- (b) $R_0 > 1$ – trends upwards, and the disease persists with positive probability
(when $p < 1$, the disease can get unlucky!)

A “knife-edge” quality around the critical value of $R_0 = 1$

Branching process

- Assumes no network structure, no triangles or shared neighbors

The SIR model

- Each node may be in the following states
 - **Susceptible**: healthy but not immune
 - **Infected**: has the virus and can actively propagate it
 - **Removed**: (Immune or Dead) had the virus but it is no longer active
- probability of an Infected node to infect a Susceptible neighbor

The SIR process

- Initially all nodes are in state S(usceptible), except for a few nodes in state I(nfected).
- An infected node stays infected for t_I steps.
 - Simplest case: $t_I = 1$
- At each of the t_I steps the infected node has probability p of infecting any of its susceptible neighbors
 - p : Infection probability
- After t_I steps the node is Removed

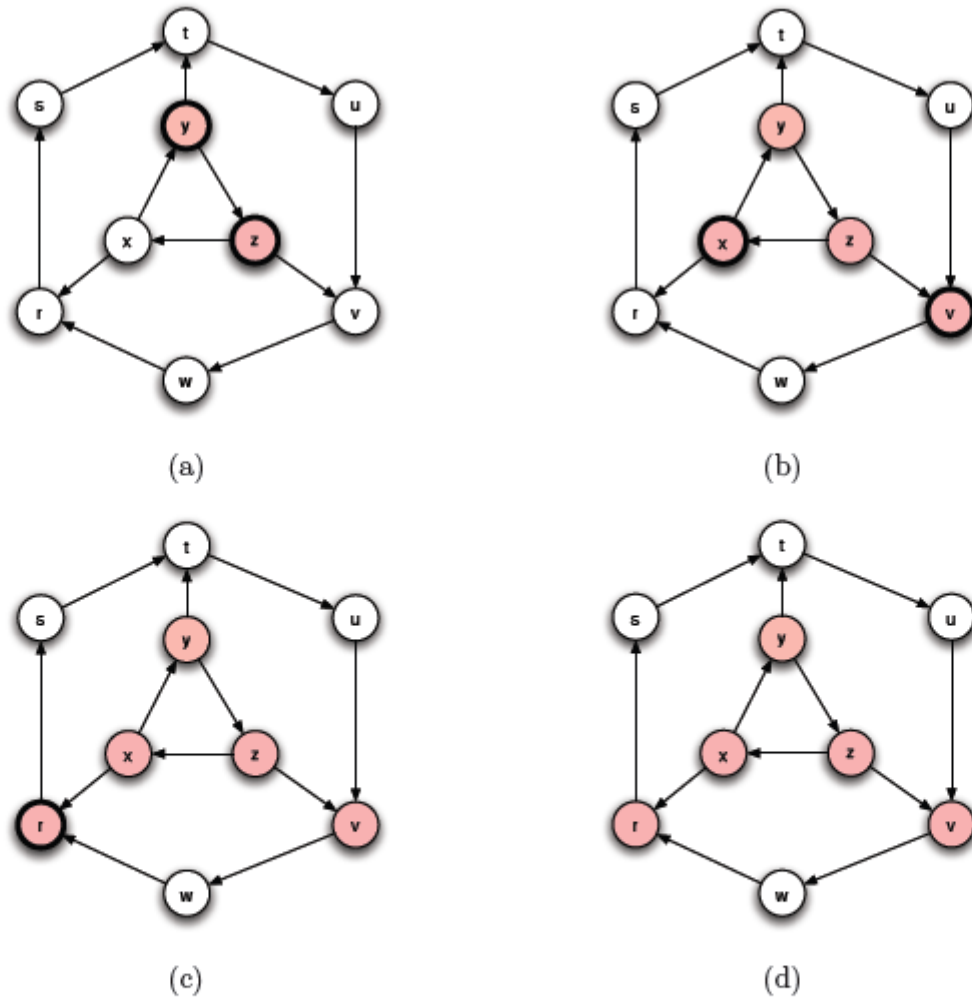


Figure 21.2: The course of an SIR epidemic in which each node remains infectious for a number of steps equal to $t_I = 1$. Starting with nodes y and z initially infected, the epidemic spreads to some but not all of the remaining nodes. In each step, shaded nodes with dark borders are in the Infectious (I) state and shaded nodes with thin borders are in the Removed (R) state.

SIR and the Branching process

- The branching process is a special case where the graph is a tree (and the infected node is the root)
- The basic reproductive number is not necessarily informative in the general case

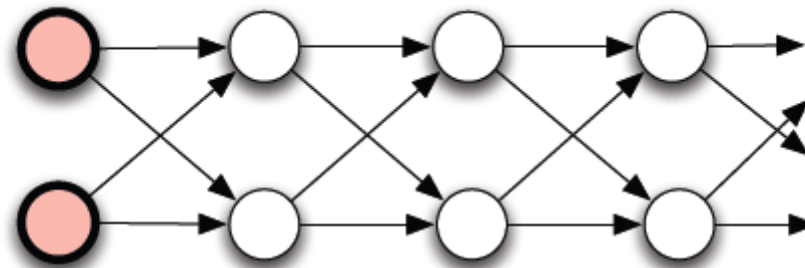


Figure 21.3: In this network, the epidemic is forced to pass through a narrow “channel” of nodes. In such a structure, even a highly contagious disease will tend to die out relatively quickly.

Percolation

- **Percolation**: we have a network of “pipes” which can carry liquids, and they can be either **open** with probability p , or **close** with probability $(1-p)$
 - The pipes can be pathways within a material
- If liquid enters the network from some nodes, does it **reach** most of the network?
 - The network **percolates**

SIR and Percolation

- There is a connection between SIR model and percolation
- When a virus is transmitted from u to v , the edge (u,v) is activated with probability p
- We can assume that all edge activations have happened **in advance**, and the input graph has **only** the **active edges**.
- Which nodes will be infected?
 - The nodes **reachable** from the initial infected nodes
- In this way we transformed the **dynamic SIR process** into a **static** one.

Example

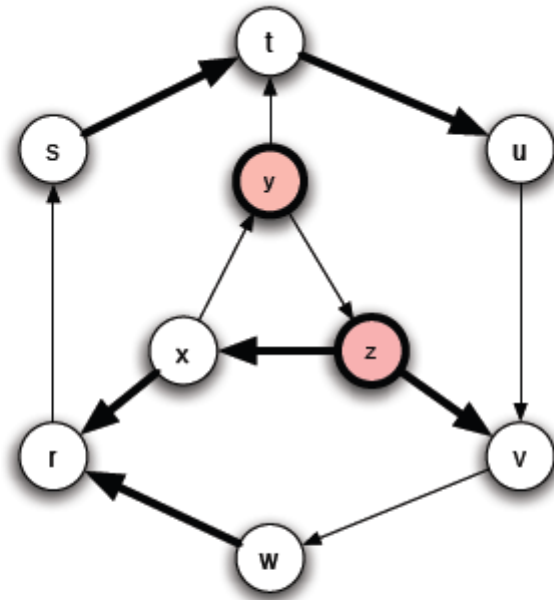


Figure 21.4: An equivalent way to view an SIR epidemic is in terms of *percolation*, where we decide in advance which edges will transmit infection (should the opportunity arise) and which will not.

The SIS model

- Susceptible-Infected-Susceptible
 - Susceptible: healthy but not immune
 - Infected: has the virus and can actively propagate it
- An Infected node infects a Susceptible neighbor with probability p
- An Infected node becomes Susceptible again with probability q (or after t_I steps)
- Nodes alternate between Susceptible and Infected status

Example

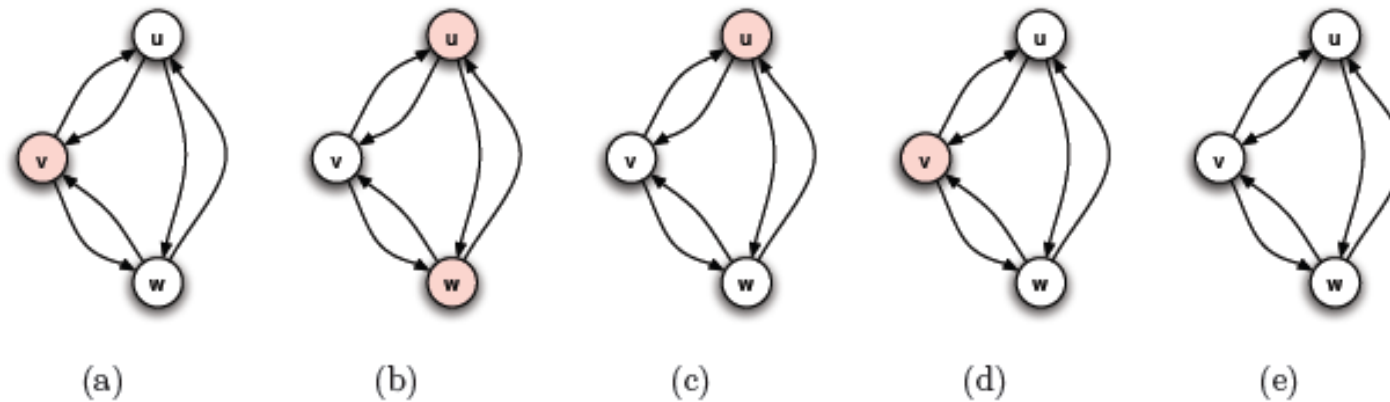


Figure 21.5: In an SIS epidemic, nodes can be infected, recover, and then be infected again. In each step, the nodes in the Infectious state are shaded.

- When no **Infected** nodes, virus dies out
- Question: will the virus die out?

An eigenvalue point of view

- If A is the adjacency matrix of the network, then the virus dies out if

$$\lambda_1(A) \leq \frac{q}{p}$$

- Where λ_1 is the first **eigenvalue** of A

Multiple copies model

- Each node may have **multiple copies** of the same virus
 - \mathbf{v} : state vector : v_i : number of virus copies at node i
- At time $t = 0$, the state vector is initialized to \mathbf{v}^0
- At time t ,
 - For each node i
 - For each of the v_i^t virus copies at node i
 - the copy is copied to a neighbor j with prob p
 - the copy dies with probability q

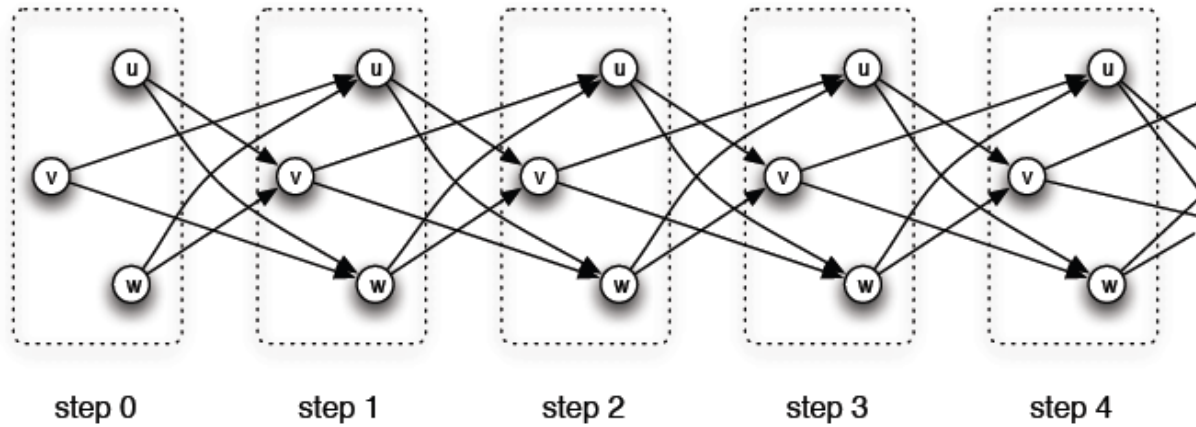
Analysis

- The expected state of the system at time t is given by

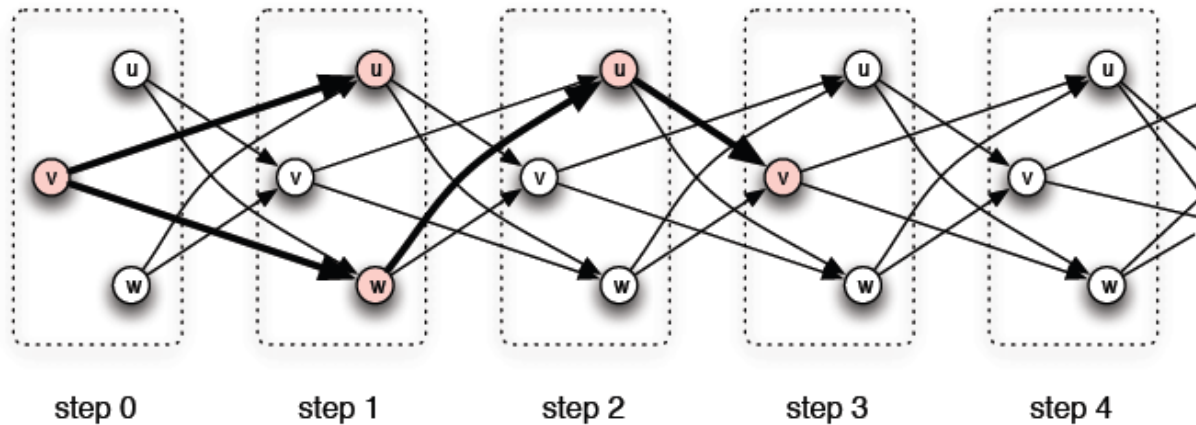
$$\overline{\mathbf{v}}^t = (\mathbf{p}\mathbf{A} + (1 - \mathbf{q})\mathbf{I})\overline{\mathbf{v}}^{t-1}$$

- As $t \rightarrow \infty$
 - if $\lambda_1(\mathbf{p}\mathbf{A} + (1 - \mathbf{q})\mathbf{I}) < 1 \Leftrightarrow \lambda_1(\mathbf{A}) < \mathbf{q}/\mathbf{p}$ then $\overline{\mathbf{v}}^t \rightarrow 0$
 - the probability that all copies die converges to 1
 - if $\lambda_1(\mathbf{p}\mathbf{A} + (1 - \mathbf{q})\mathbf{I}) = 1 \Leftrightarrow \lambda_1(\mathbf{A}) = \mathbf{q}/\mathbf{p}$ then $\overline{\mathbf{v}}^t \rightarrow \mathbf{c}$
 - the probability that all copies die converges to 1
 - if $\lambda_1(\mathbf{p}\mathbf{A} + (1 - \mathbf{q})\mathbf{I}) > 1 \Leftrightarrow \lambda_1(\mathbf{A}) > \mathbf{q}/\mathbf{p}$ then $\overline{\mathbf{v}}^t \rightarrow \infty$
 - the probability that all copies die converges to a constant < 1

SIS and SIR



(a) To represent the SIS epidemic using the SIR model, we use a “time-expanded” contact network

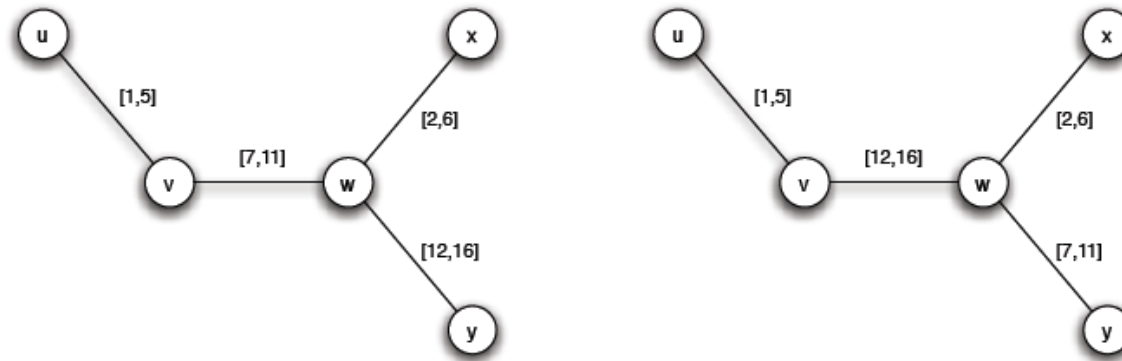


(b) The SIS epidemic can then be represented as an SIR epidemic on this time-expanded network.

Figure 21.6: An SIS epidemic can be represented in the SIR model by creating a separate copy of the contact network for each time step: a node at time t can infect its contact neighbors at time $t + 1$.

Including time

- Infection can only happen within the **active window**



(a) In a contact network, we can annotate the edges with time windows during which they existed.

(b) The same network as in (a), except that the timing of the w-v and w-y partnerships have been reversed.

Figure 21.8: Different timings for the edges in a contact network can affect the potential for a disease to spread among individuals. For example, in (a) the disease can potentially pass all the way from *u* to *y*, while in (b) it cannot.

- Importance of concurrency – enables **branching**