



SOCIAL NETWORK ANALYTICS

Branching Process, SIR & SIS epidemic models

Prakash C O

Department of Computer Science and
Engineering

SOCIAL NETWORK ANALYTICS

Branching Process – Epidemic Model

Prakash C O

Department of Computer Science and Engineering

➤ The simplest model of contagion the branching process works as follows.

- **(First wave.)**
 - A person carrying a new disease enters a population, and transmits it to each person he meets independently with a probability of p .
The infected person meets k people; let's call these k people the first wave of the epidemic.
 - Based on the random transmission of the disease from the initial person, some of the people in the first wave may get infected with the disease, while others may not.

➤ The simplest model of contagion the branching process works as follows. Cont..

- **(Second wave.)**

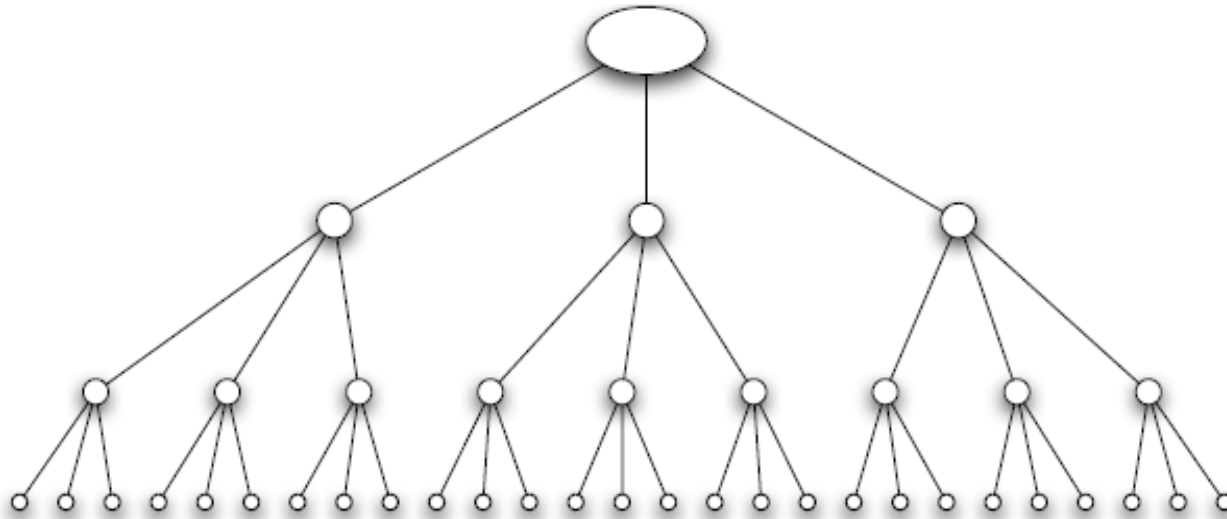
- Now, each person in the first wave goes out into the population and meets k different people, resulting in a second wave of $k \cdot k = k^2$ people.
- Each infected person in the first wave passes the disease independently to each of the k second-wave people they meet, again independently with probability p .

- **(Subsequent waves.)**

- Further waves are formed in the same way, by having each person in the current wave meet k new people, passing the disease to each independently with probability p .

Branching process

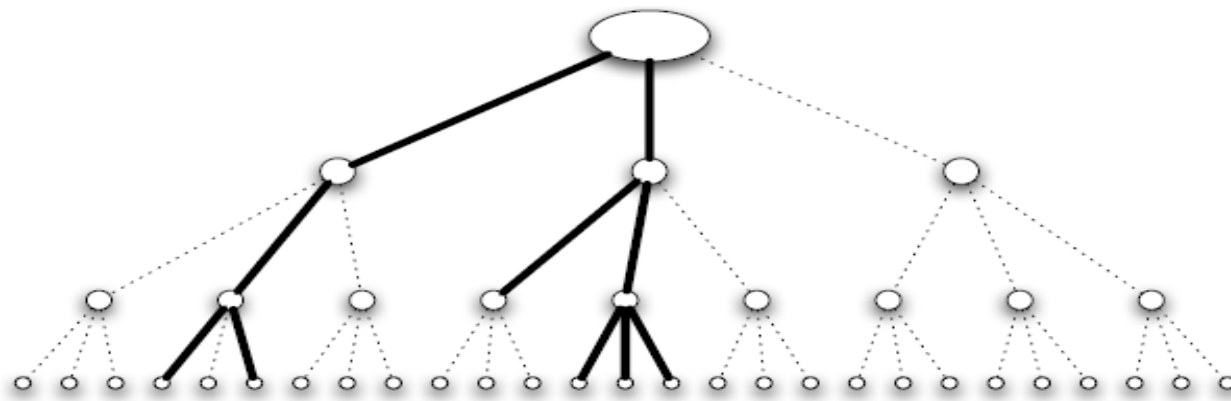
- The **contact network** for this epidemic can be drawn as in Figure 1(a) (with $k = 3$ and only the first three waves shown).
- The tree that forms the contact network for the branching process is in fact infinite, since we continue defining waves indefinitely.



(a) *The contact network for a branching process*

➤ What is the behavior of an epidemic in this model?

- The spread of the epidemic is indicated by highlighting the edges of the contact network on which the disease passes successfully from one person to another — each of these infections happens independently with probability p .
- Figure 1(b) shows an **aggressive epidemic** that infects two people in the first wave, three in the second wave, five in the third wave, and presumably more in future waves (not shown in the picture).

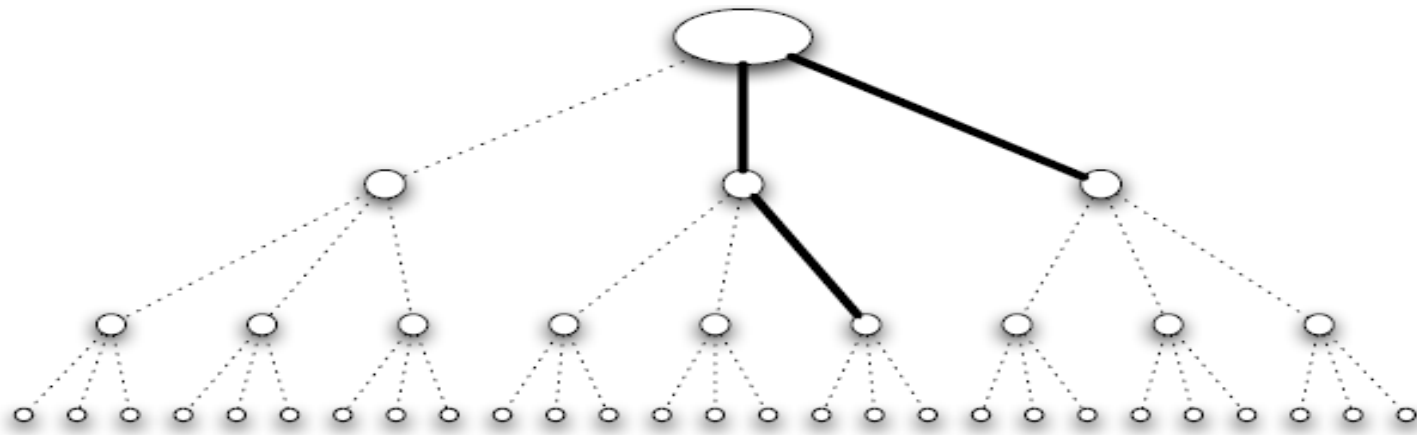


(b) *With high contagion probability, the infection spreads widely*

Branching process

- Figure 1(c), shows a much **milder epidemic** (for a less contagious disease, with a **smaller value of p**): of the two people infected in the first wave, one doesn't infect anyone else, and the other infects only one further person who in turn doesn't pass it on.

This disease has completely vanished from the population after the second wave, having infected only four people in total.



(c) *With low contagion probability, the infection is likely to die out quickly*

➤ There are **only two possibilities** for a disease in the branching process model:

1. It reaches a wave where it infects no one, thus dying out after a finite number of steps; or
2. It continues to infect people in *every* wave, proceeding infinitely through the contact network.

There is a simple condition to tell these two possibilities apart, based on a quantity called the ***basic reproductive number*** of the disease.

SOCIAL NETWORK ANALYTICS

Branching process



The basic reproductive number(R_0)

- R_0 is a mathematical term that indicates how contagious an infectious disease is.
- **R_0 represents the average number of secondary infections generated from one infectious individual in a completely susceptible population.**
- For example, if a disease has an R_0 of 12, a person who has the disease will transmit it to an average of 12 other people.
That replication will continue if no one has been vaccinated against the disease or is already immune to it in their community.

- For example, in 1918 there was a worldwide outbreak of the **swine flu** that killed 50 million people.

According to a review article published in BMC Medicine, the **R_0 value of the 1918 pandemic was estimated to be between 1.4 and 2.8.**

- But when the swine flu, or H1N1 virus, came back in 2009, its **R_0 value was between 1.4 and 1.6**, report researchers in the journal Science.

The existence of vaccines and antiviral drugs made the 2009 outbreak much less deadly.

The Basic Reproductive Number for Branching Process.

➤ The basic reproductive number(R_0)

R_0 = the expected number of new cases of the disease caused by a single individual.

- In branching process model everyone meets k new people and infects each with probability p (contagion/infection probability), the basic reproductive number here is given by $R_0 = pk$.
- The outcome of the disease in a branching process model is determined by whether the basic reproductive number is smaller or larger than 1.

What do R_0 values mean?

- Three possibilities exist for the potential transmission or decline of a disease, depending on its R_0 value:
 - 1) If $R_0 < 1$, then with probability less than 1, each existing infection causes less than one new infection, the disease dies out after a finite number of waves.*
 - 2) If $R_0 == 1$, each existing infection causes one new infection. The disease will stay alive and stable, but there won't be an outbreak or an epidemic.*

What do R_0 values mean?

- Three possibilities exist for the potential transmission or decline of a disease, depending on its R_0 value: [cont....](#)

3) If $R_0 > 1$, then with probability greater than 0 the disease persists by infecting at least one person in each wave. The disease will be transmitted between people, and there may be an outbreak or epidemic.

Notice, however, that even when $R_0 > 1$, the conclusion is simply that the disease persists with positive probability, not with absolute certainty:

whenever $p < 1$, then there is always some chance that none of the first few infected people will succeed in infecting anyone else, causing the disease to die out.

In other words, even an ultra-contagious disease can simply get “unlucky” and vanish from the population before it has a chance to really get going.

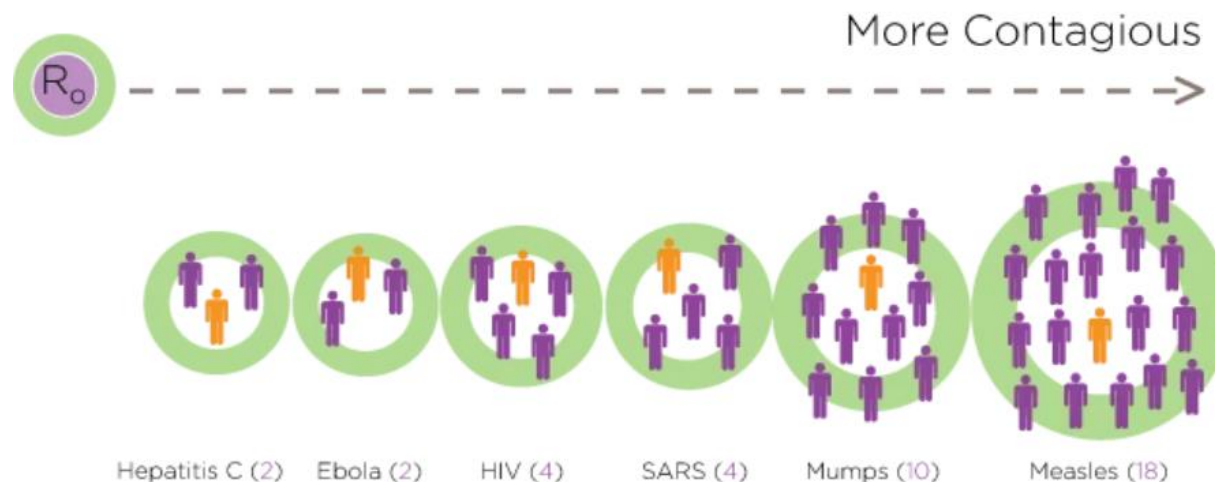
- The dichotomy expressed by this condition has an interesting “knife-edge” quality to it when R_0 is close to 1.
 - Consider a branching process where R_0 is **very slightly below 1**, and **increase the contagion(infection) probability p by a little bit**; **the result could push R_0 above 1, suddenly resulting in a positive probability of an enormous outbreak.**
 - The same effect can happen in the reverse direction as well, where **slightly reducing the contagiousness of a disease to push R_0 below 1** **can eliminate the risk of a large epidemic.**

- Since $R_0 = pk$, small changes in the number of people k that each person comes into contact with can also have a large effect when R_0 is near 1.
- All this suggests that around the critical value $R_0 = 1$, it can be **worth investing large amounts of effort even to produce small shifts in the basic reproductive number.**
- **The two basic kinds of public-health measures in terms of reductions to R_0 are**
 1. **Quarantining people, which reduces the quantity k , and**
 2. **Encouraging behavioral measures such as better sanitary practices to reduce the spread of germs, which reduces the quantity p .**

- Is airborne illnesses tend to have a higher R_0 value than those spread through direct contact?
1. YES
 2. NO

What conditions are measured by R_0 ?

- R_0 can be used to measure any contagious disease that may spread in a susceptible population.
Some of the most highly contagious conditions are measles and the common flu. More serious conditions, such as Ebola and HIV, spread less easily between people.
- This illustration shows some commonly known diseases and their estimated R_0 values.



COVID-19 R_0

- The R_0 for COVID-19 is a median of 5.7, according to a study published online in [Emerging Infectious Diseases](#). That's about double an earlier R_0 estimate of 2.2 to 2.7
- The 5.7 means that one person with COVID-19 can potentially transmit the coronavirus to 5 to 6 people, rather than the 2 to 3 researchers originally thought.
- Researchers calculated the new number based on data from the original outbreak in Wuhan, China. They used parameters like the virus incubation period (4.2 days) — how much time elapsed from when people were exposed to the virus and when they started to show symptoms.

COVID-19 R_0

- The researchers estimated a doubling time of 2 to 3 days, which is much faster than earlier estimates of 6 to 7 days.
The doubling time is how long it takes for the number of coronavirus cases, hospitalizations, and deaths to double.
The shorter the time, the faster the disease is spreading.
- With an R_0 of 5.7, at least 82 percent of the population needs to be immune to COVID-19 to stop its transmission through vaccination and **herd immunity**.
- The study authors say active surveillance, tracking the contacts of people who contracted the coronavirus, quarantine, and strong physical distancing measures are needed to stop the coronavirus from being transmitted.

- R_0 is a useful calculation for predicting and controlling the transmission of disease.
- Medical science continues to advance. Researchers are discovering new cures for different conditions, but contagious diseases aren't going to disappear anytime soon.
- Take these steps to help prevent the transmission of contagious diseases:
 1. Learn how different contagious diseases are transmitted.
 2. Ask your doctor about steps you can take to stop the transmission of infection. For example, wash your hands regularly with soap and water, especially before you prepare or eat food.
 3. Stay up to date on routine vaccinations.
 4. Ask your doctor what diseases you should be vaccinated against.

Video – Reproductive number (R_0)

➤ [https://upload.wikimedia.org/wikipedia/commons/f/f4/COVID19_in_numbers-R0%2C the case fatality rate and why we need to flatten the curve.webm](https://upload.wikimedia.org/wikipedia/commons/f/f4/COVID19_in_numbers-R0%2C_the_case_fatality_rate_and_why_we_need_to_flatten_the_curve.webm)

A video discussing the basic reproduction number and case fatality rate in the context of COVID-19 pandemic.

SOCIAL NETWORK ANALYTICS

SIR epidemic model

Prakash C O

Department of Computer Science and Engineering

The SIR model

- In 1927, W. O. Kermack and A. G. McKendrick created a model, which divides the (fixed) population of N individuals into three "compartments" which may vary as a function of time, t :
- $S(t)$ are those susceptible but not yet infected with the disease;
 - $I(t)$ is the number of infected individuals;
 - $R(t)$ are those individuals who have recovered from the disease and now have immunity to it.



The SIR model

- As the first step in the modeling process, we identify the independent and dependent variables.
- **The independent variable is time t , measured in days.**
- We consider two related sets of dependent variables.
- **The first set of dependent variables** counts *people in each of the groups*, each as a function of time:
 - **S or $S(t)$** is the number of susceptible individuals,
 - **I or $I(t)$** is the number of infected individuals, and
 - **R or $R(t)$** is the number of recovered individuals.

The SIR model

➤ The second set of dependent variables represents **the fraction of the total population in each of the three categories.**

So, if N is the total population, we have

- $s(t) = S(t)/N$, the susceptible fraction of the population,
- $i(t) = I(t)/N$, the infected fraction of the population, and
- $r(t) = R(t)/N$, the recovered fraction of the population.

Where $N = S+I+R$ or $N = S(t)+I(t)+R(t)$

At $t=0$, $S(0)=S_0$ $I(0)=I_0$ $R(0)=0$

The SIR model

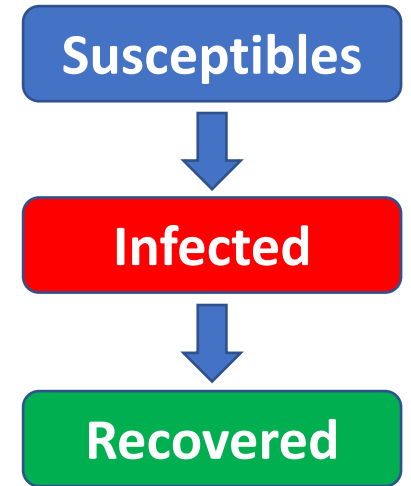
- The two sets of dependent variables are proportional to each other, so either set will give us the same information about the progress of the epidemic.
1. Under the assumptions we have made, how do you think $s(t)$ should vary with time? How should $r(t)$ vary with time? How should $i(t)$ vary with time?
 2. Sketch on a piece of paper what you think the graph of each of these functions looks like.
 3. Explain why, at each time t , $s(t) + i(t) + r(t) = 1$.



The SIR model

- The SIR model **describes the change in the population of each of these compartments in terms of two parameters, β and γ .**
- β is the effective **contact rate/transmission rate** of the disease.
- γ is the **recovery rate** of the disease and γI is the number of infected individuals that recover per unit time.
- $1/\gamma$ is the **infectious period** i.e. the average duration of time an individual remains infected.

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta SI}{N}, \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I.\end{aligned}$$



Demo using NetworkX

- The following Python code integrates these equations for a disease characterized by parameters $\beta=0.2$, $1/\gamma=10$ days in a population of $N=1000$ (perhaps 'flu in a school').

The model is started with a single infected individual on day 0: $I(0)=1$.

- **SIR_Model.py**

The SIR model

```
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

# Total population, N.
N = 1000
# Initial number of infected and recovered individuals, I0 and R0.
I0, R0 = 1, 0
# Everyone else, S0, is susceptible to infection initially.
S0 = N - I0 - R0
# Contact rate, beta, and mean recovery rate, gamma, (in 1/days).
beta, gamma = 0.2, 1./10
# A grid of time points (in days)
t = np.linspace(0, 160, 160)

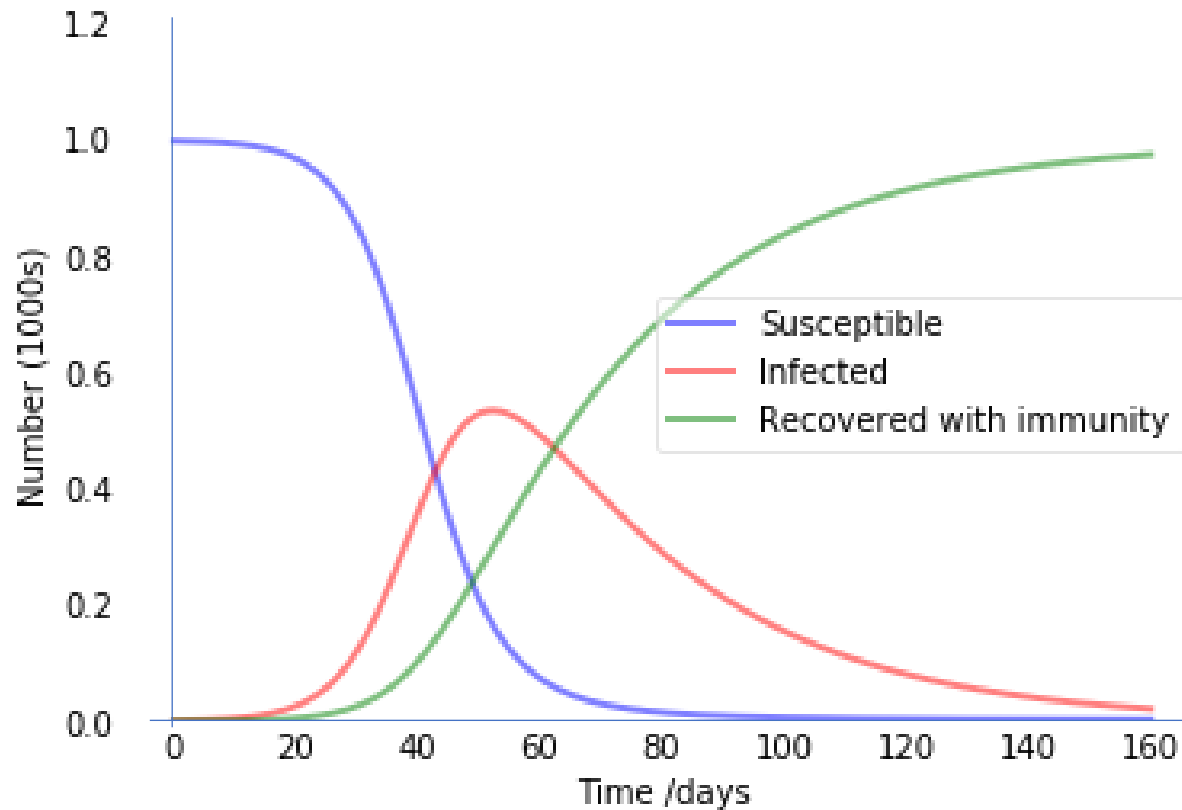
# The SIR model differential equations.
def deriv(y, t, N, beta, gamma):
    S, I, R = y
    dSdt = -beta * S * I / N
    dIdt = beta * S * I / N - gamma * I
    dRdt = gamma * I
    return dSdt, dIdt, dRdt

# Initial conditions vector
y0 = S0, I0, R0
# Integrate the SIR equations over the time grid, t.
ret = odeint(deriv, y0, t, args=(N, beta, gamma))
S, I, R = ret.T
```

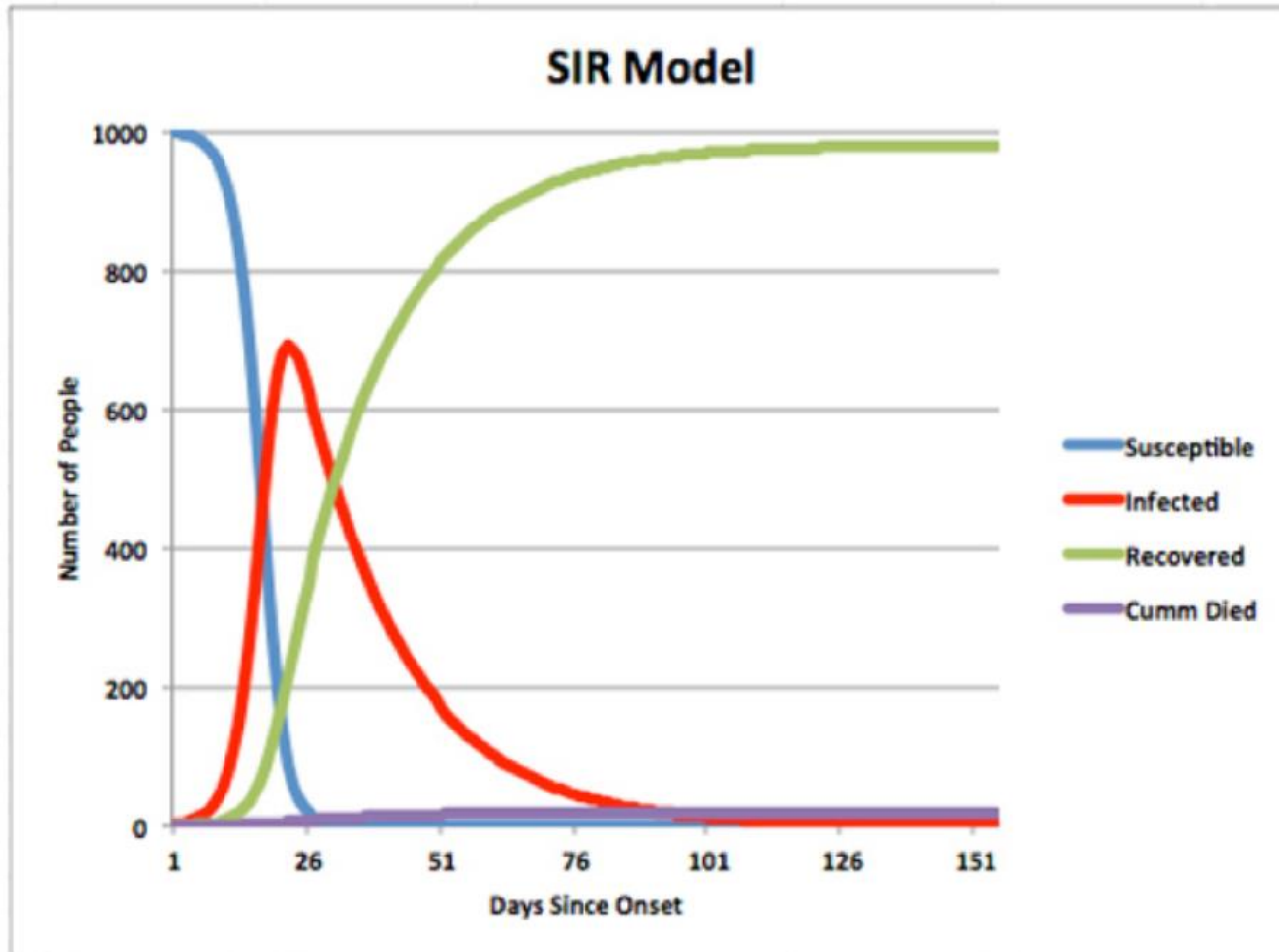
The SIR model

```
# Plot the data on three separate curves for S(t), I(t) and R(t)
fig = plt.figure(facecolor='w')
ax = fig.add_subplot(111, axisbelow=True)
ax.plot(t, S/1000, 'b', alpha=0.5, lw=2, label='Susceptible')
ax.plot(t, I/1000, 'r', alpha=0.5, lw=2, label='Infected')
ax.plot(t, R/1000, 'g', alpha=0.5, lw=2, label='Recovered with immunity')
ax.set_xlabel('Time /days')
ax.set_ylabel('Number (1000s)')
ax.set_ylim(0,1.2)
ax.yaxis.set_tick_params(length=0)
ax.xaxis.set_tick_params(length=0)
ax.grid(b=True, which='major', c='w', lw=2, ls='-')
legend = ax.legend()
legend.get_frame().set_alpha(0.5)
for spine in ('top', 'right', 'bottom', 'left'):
    ax.spines[spine].set_visible(False)
plt.show()
```

The SIR model

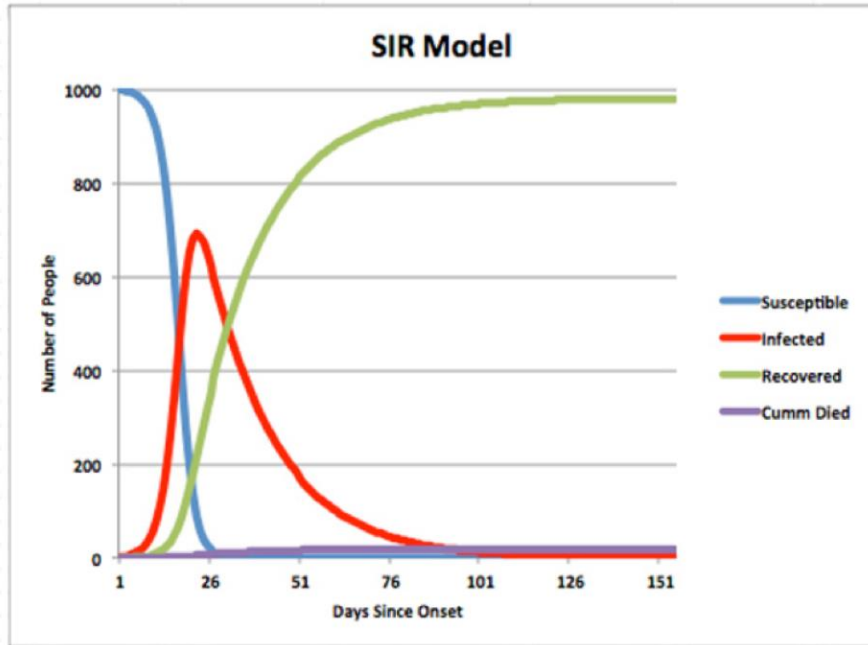


The SIR model



SOCIAL NETWORK ANALYTICS

The SIR Epidemic Model



Questions!

1. What happens if initial $I = 0$?
2. What does it mean that red line increases so rapidly?
3. What does it mean that green line also rises rapidly, but not as rapidly?
4. What does it mean that the green line reaches nearly to N ?

The SIR model

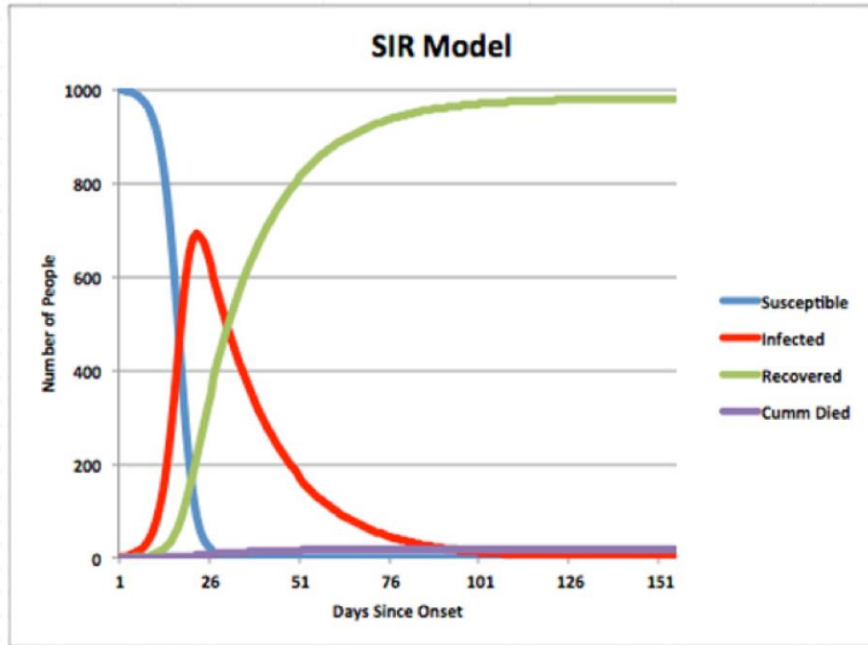
Questions!

1. What happens if initial $I = 0$?

(Answer: Nobody infected so disease can not spread, nobody gets sick)

SOCIAL NETWORK ANALYTICS

The SIR Epidemic Model



Questions!

1. What happens if initial $I = 0$?
2. What does it mean that red line increases so rapidly?
3. What does it mean that green line also rises rapidly, but not as rapidly?
4. What does it mean that the green line reaches nearly to N ?

The SIR model

Questions!

1. What happens if initial $I = 0$?

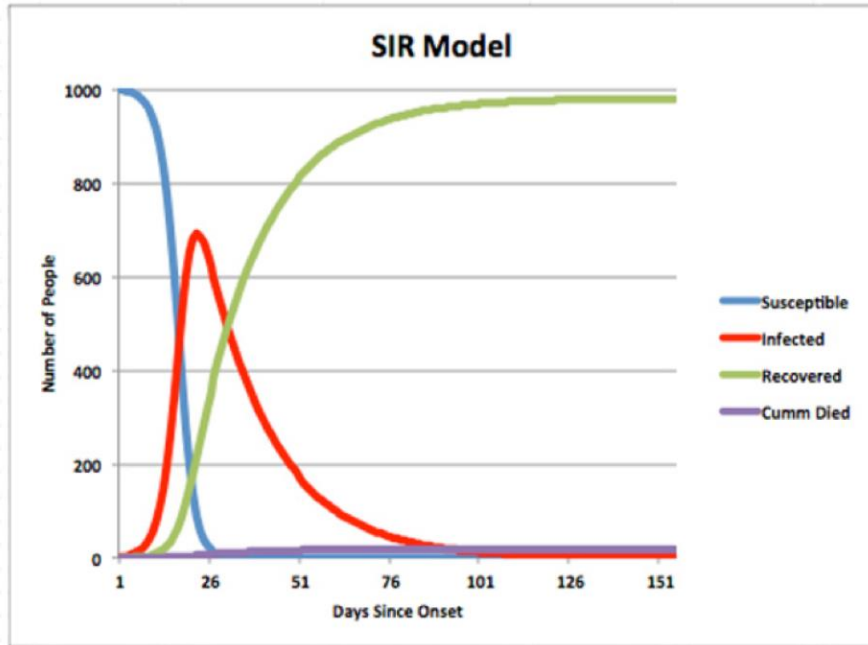
(Answer: Nobody infected so disease can not spread, nobody gets sick)

2. What does it mean that red line increases so rapidly?

(Answer: This is a very contagious disease.)

SOCIAL NETWORK ANALYTICS

The SIR Epidemic Model



Questions!

1. What happens if initial $I = 0$?
2. What does it mean that red line increases so rapidly?
3. What does it mean that green line also rises rapidly, but not as rapidly?
4. What does it mean that the green line reaches nearly to N ?

The SIR model

Questions!

1. What happens if initial $I = 0$?

(Answer: Nobody infected so disease can not spread, nobody gets sick)

2. What does it mean that red line increases so rapidly?

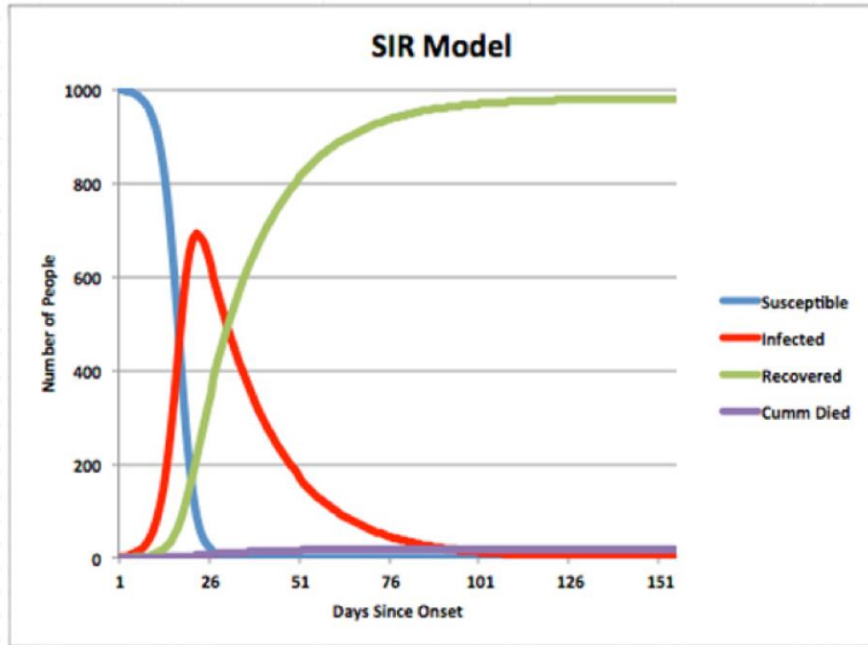
(Answer: This is a very contagious disease.)

3. What does it mean that green line also rises rapidly, but not as rapidly?

(Answer: People recover quickly, but more slowly than they become infected.)

SOCIAL NETWORK ANALYTICS

The SIR Epidemic Model



Questions!

1. What happens if initial $I = 0$?
2. What does it mean that red line increases so rapidly?
3. What does it mean that green line also rises rapidly, but not as rapidly?
4. What does it mean that the green line reaches nearly to N ?

The SIR model

Questions!

1. What happens if initial $I = 0$?

(Answer: Nobody infected so disease can not spread, nobody gets sick)

2. What does it mean that red line increases so rapidly?

(Answer: This is a very contagious disease.)

3. What does it mean that green line also rises rapidly, but not as rapidly?

(Answer: People recover quickly, but more slowly than they become infected.)

4. What does it mean that the green line reaches nearly to 1,000?

(Answer: Nearly every one recovers, disease not very deadly.)

The SIR model

- The **SIR epidemic model** applied to more general network/contact structure that **preserves the basic ingredients of the branching process model** at the level of individual nodes.

- An individual node in this model goes through three potential stages during the course of the epidemic:
 - ❖ **Susceptible:** Before the node has caught the disease, it is susceptible to infection from its neighbors.
 - ❖ **Infectious:** Once the node has caught the disease, it is infectious and has some probability of infecting each of its susceptible neighbors.
 - ❖ **Removed:** After a particular node has experienced the full infectious period, this node is removed from consideration, since it no longer poses a threat of future infection.

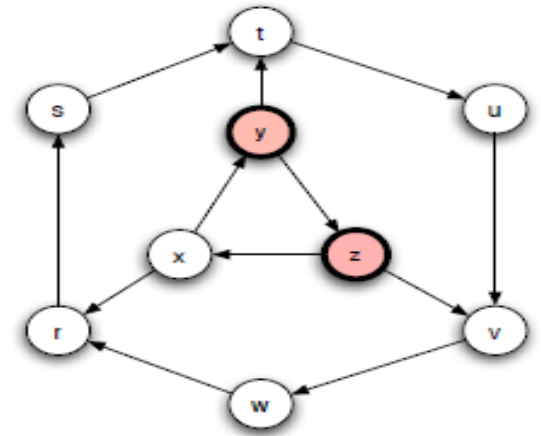
SOCIAL NETWORK ANALYTICS

The SIR Epidemic Model

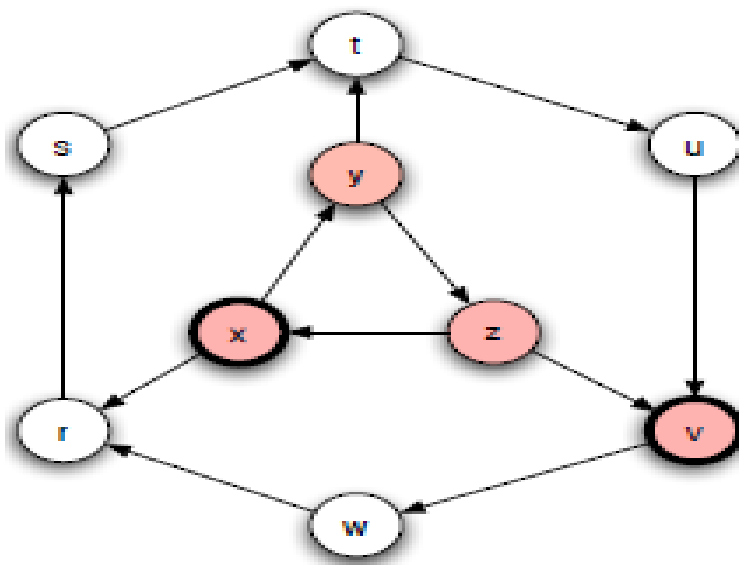
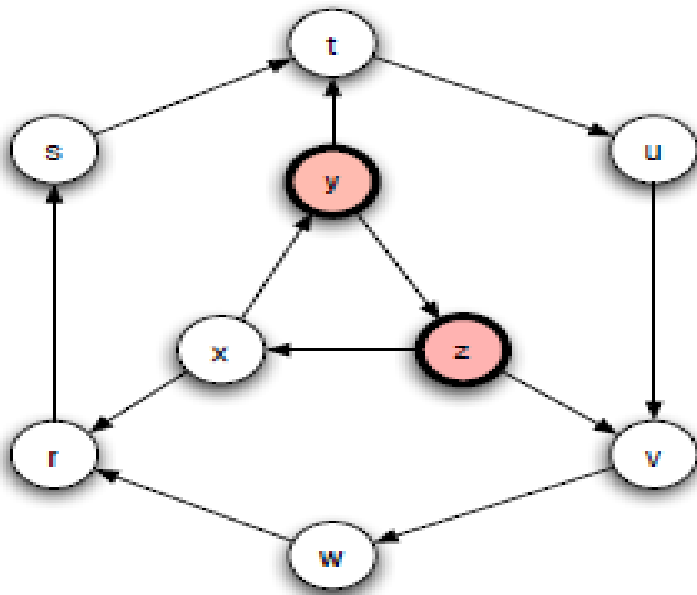
- Using the **SIR three-stage “life cycle” for the disease at each node**, we now define a model for epidemics on networks.
- Given a directed graph representing the contact network; an **edge pointing from v to w** in the graph means that **if v becomes infected at some point, the disease has the potential to spread directly to w** .
 - To represent a symmetric contact between people, where either has the potential to directly infect the other, we can put in directed edges pointing each way: both from v to w and also from w to v .



PES
UNIVERSITY
ONLINE



- Now, each node has the potential to go through the **Susceptible-Infectious-Removed** cycle, where we abbreviate these three states as S , I , and R .



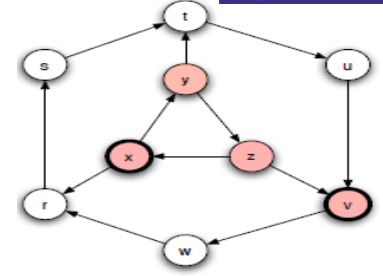
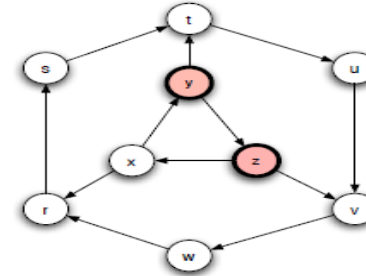
SOCIAL NETWORK ANALYTICS

The SIR Epidemic Model

➤ The **progress of the epidemic** is controlled by

- The contact network structure and
- Two-additional quantities:
 - 1) p (the probability of contagion) and
 - 2) t_i (the length of the infection).

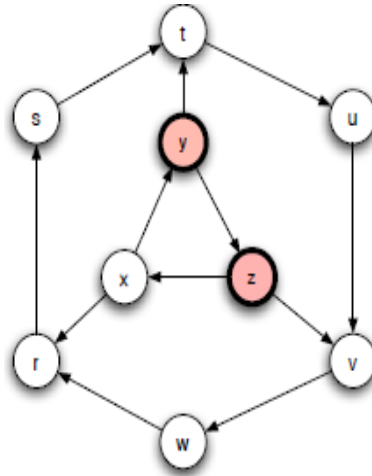
- Initially, some nodes are in the **I state** and all others are in the **S state**.
- Each node v that enters the I state remains infectious for a fixed number of steps t_i .
- During each of these t_i steps, v has a probability p of passing the disease to each of its susceptible neighbors.
- After t_i steps, node v is no longer infectious or susceptible to further bouts of the disease; we describe it as *removed* (R), since it is now an inert node in the contact network that can no longer either catch or transmit the disease.



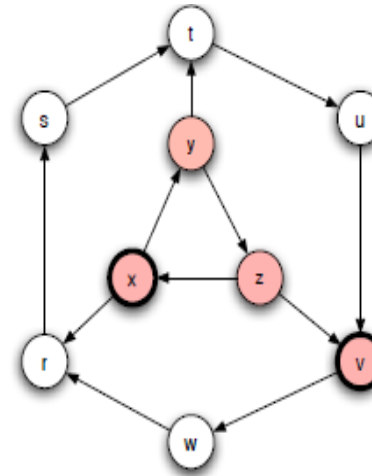
PES
UNIVERSITY
ONLINE

Figure 2: Example of the SIR model

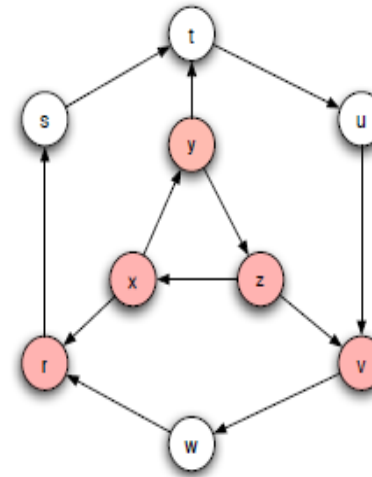
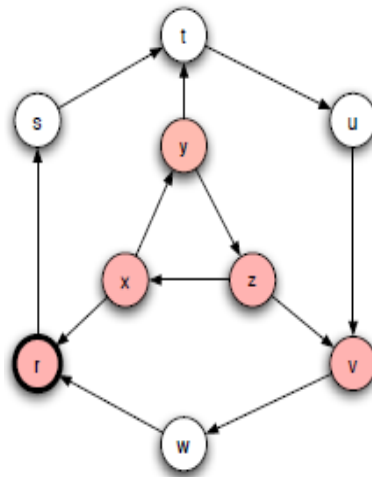
- The course of an SIR epidemic in which each node remains infectious for a number of steps equal to $t_i = 1$.
- Starting with nodes **y** and **z** initially infected, the epidemic spreads to some but not all of the remaining nodes.
- In each step, **shaded nodes with dark borders** are in the **Infectious(I)**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



(a)



(b)



- The **SIR model** is clearly most appropriate for a **disease that each individual only catches once in their lifetime**; after being infected, a node is removed either because it has acquired lifetime immunity or because the disease has killed it.
- In this model, Contagion probabilities are set to a uniform value p , and contagiousness has a kind of “on-off” property: a node is equally contagious for each of the t_i steps while it has the disease.

Extensions to the SIR model.

- Extend the model to handle more complex assumptions.
 1. First, we can easily capture the idea that **contagion is more likely between certain pairs of nodes** by assigning a **separate probability $p_{v,w}$ to each pair of nodes v and w** for which v links to w in the directed contact network.
 2. Here, higher values of $p_{v,w}$ correspond to closer contact and more likely contagion, while lower values indicate less intensive contact.
 3. We can also choose to model **the infectious period as random in length**, by assuming that an infected node has a **probability q of recovering** in each step while it is infected, while leaving the other details of the model as they are.

Extensions to the SIR model.

- More elaborate extensions to the model involve
 - separating the I state into a sequence of several states (e.g. **early**, **middle**, and **late periods of the infection**), and
 - allowing the contagion probabilities to vary across these states.
- This could be used, for example, to model a disease with a highly contagious incubation period, followed by a less contagious period while symptoms are being expressed.
- Researchers have also considered variations on the SIR model in which the disease-causing pathogen is mutating (and thus changing its disease characteristics) over the course of the outbreak.

Exercise 1:

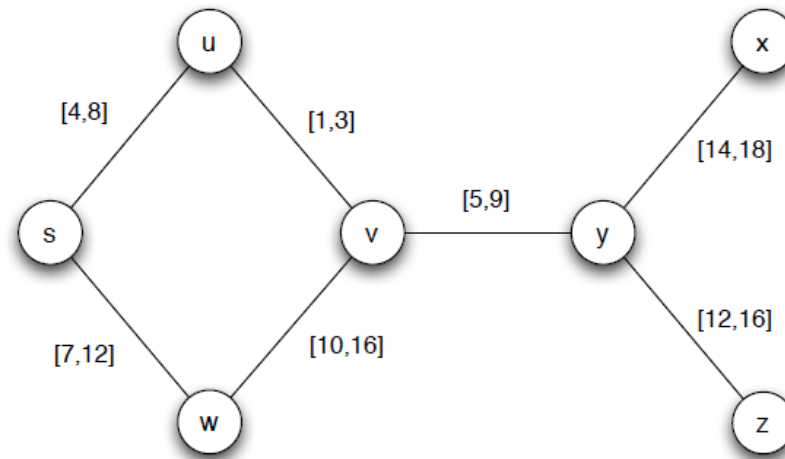


Figure: Contacts among a set of people, with time intervals showing when the contacts occurred. We assume that the period of observation runs from time 0 to time 20.

Suppose that *s* is the only individual who had the disease at time 0. Which nodes could potentially have acquired the disease by the end of the observation period, at time 20?

The Role of the Basic Reproductive Number in arbitrary network.

- In networks that do not have a tree structure, the simple dichotomy in epidemic behavior determined by R_0 does not necessarily hold.
- In fact, it is not hard to construct an example showing how this dichotomy breaks down. To do this, let's start with the network depicted in **Figure 21.3**, and suppose that these layers of two nodes at a time continue indefinitely to the right.

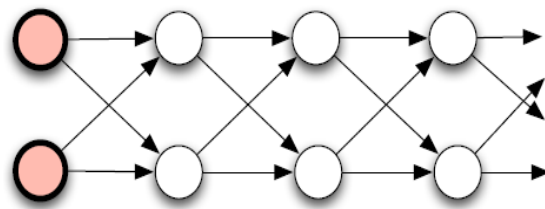


Figure 21.3: In this network, the epidemic is forced to pass through a narrow “channel” of nodes. In such a structure, even a highly contagious disease will tend to die out relatively quickly.

The Role of the Basic Reproductive Number in arbitrary network.

Let's consider an SIR epidemic in which $t_i = 1$, the infection probability p is $2/3$, and the two nodes at the far left are the ones that are initially infected.

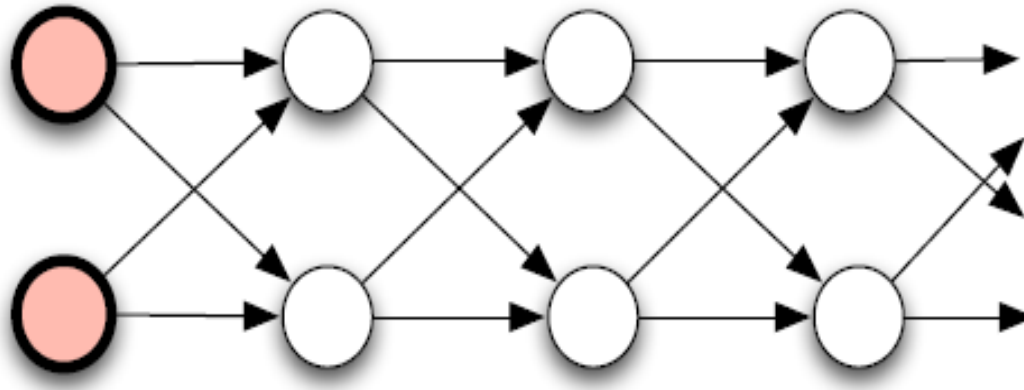
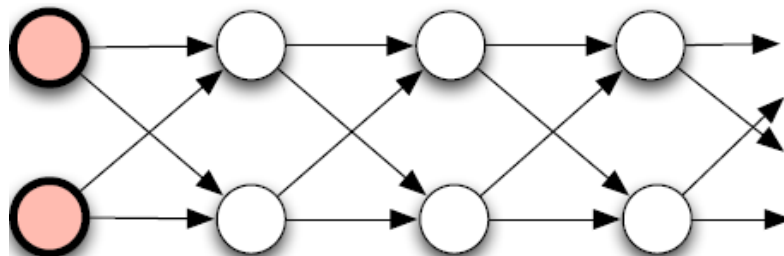


Figure 21.3: In this network, the epidemic is forced to pass through a narrow “channel” of nodes. In such a structure, even a highly contagious disease will tend to die out relatively quickly.

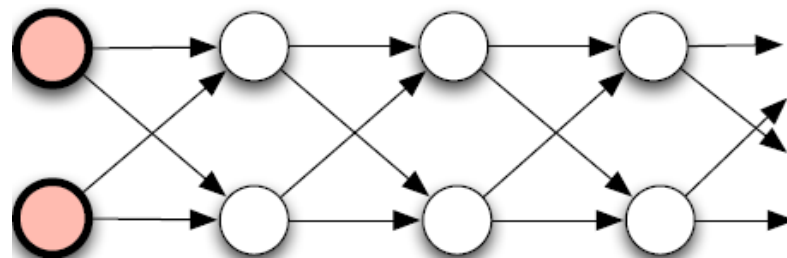
The Role of the Basic Reproductive Number in arbitrary network.

- When we don't have a tree network, we need to decide how to define an analogue of the basic reproductive number.
- In a **network as highly structure** as the one in Figure 21.3, we can work directly from the definition of R_0 as the **expected number of new cases of the disease caused by a single individual**.
- In Figure 21.3, each infected node has edges to two nodes in the next layer; since it infects each with probability $2/3$, the expected number of new cases caused by this node is $4/3$ ($2 \cdot (2/3)$).



The Role of the Basic Reproductive Number in arbitrary network.

- So in our example, $R_0 > 1$. Despite this, however, it is easy to see that the disease will die out almost surely after reaching only a finite number of steps.
- In each layer, there are four edges leading to the next layer, and each will independently fail to transmit the disease with probability $1/3$. Therefore, with probability $(1/3)^4 = 1/81$, all four edges will fail to transmit the disease — and at this point, these four edges become a “roadblock” guaranteeing the disease can never reach the portion of the network beyond them.
- Thus, as the disease moves along layer-by-layer, there is a probability of at least $1/81$ that each layer will be its last. Therefore, with probability 1, it must come to an end after a finite number of layers.



- Self study

SIR Epidemics and Percolation



THANK YOU

Prakash C O

Department of Computer Science and Engineering

coprakasha@pes.edu

+91 98 8059 1946