Social and Information Networks

CSCC46H, Fall 2022
Lecture 11

Prof. Ashton Anderson
ashton@cs.toronto.edu

Today

A4 due next week
A4 tex posted on Quercus

Today

Missed a blog post? Finish it by next Friday, Dec 2 @ 5pm and email the TAs Richard and Conroy to let them know.

Grade reduction will apply, but you can avoid a 0.

Today

Epidemics and Contagion
Voting

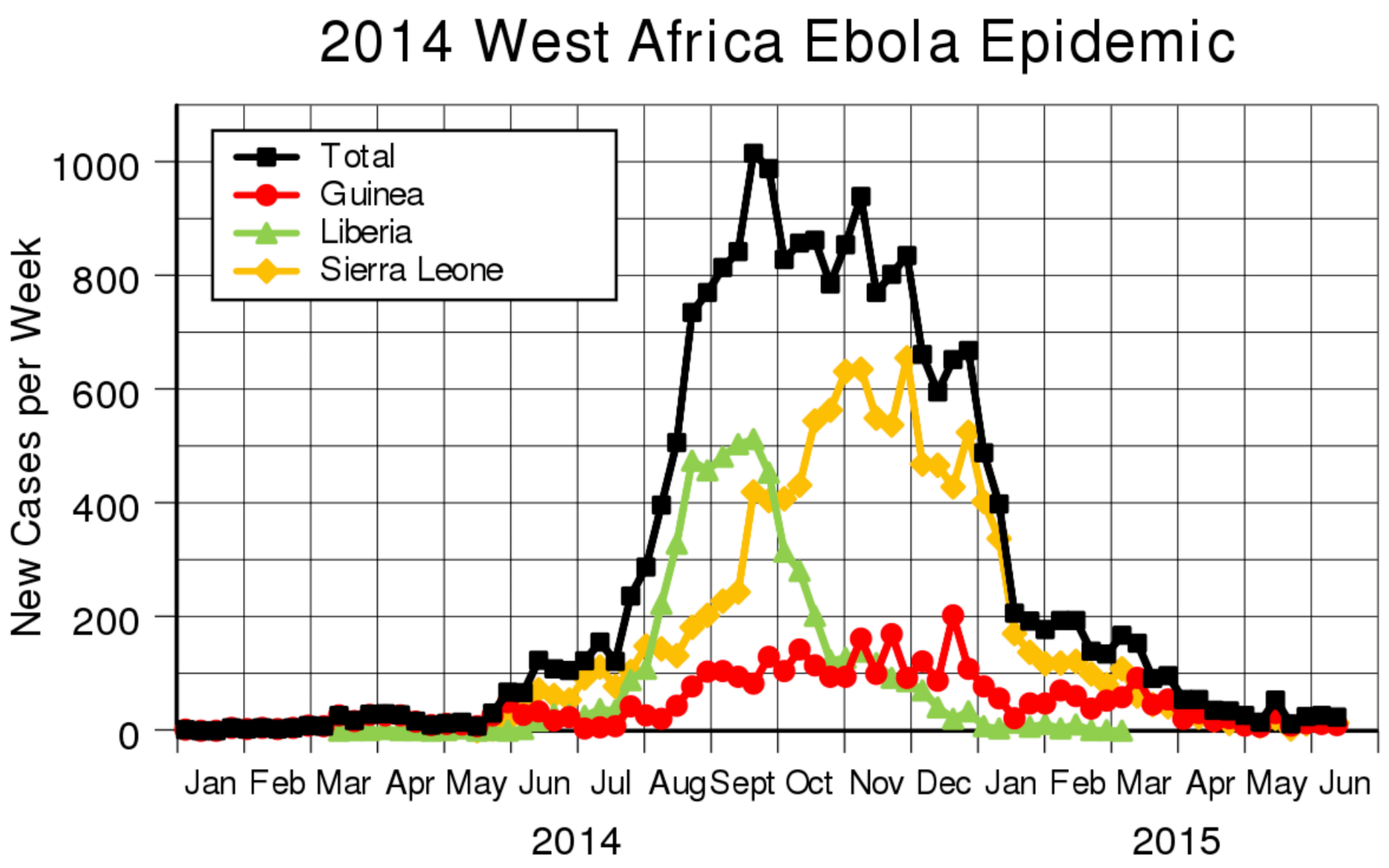
Epidemics

Epidemics

Why study epidemics in a computer science class?

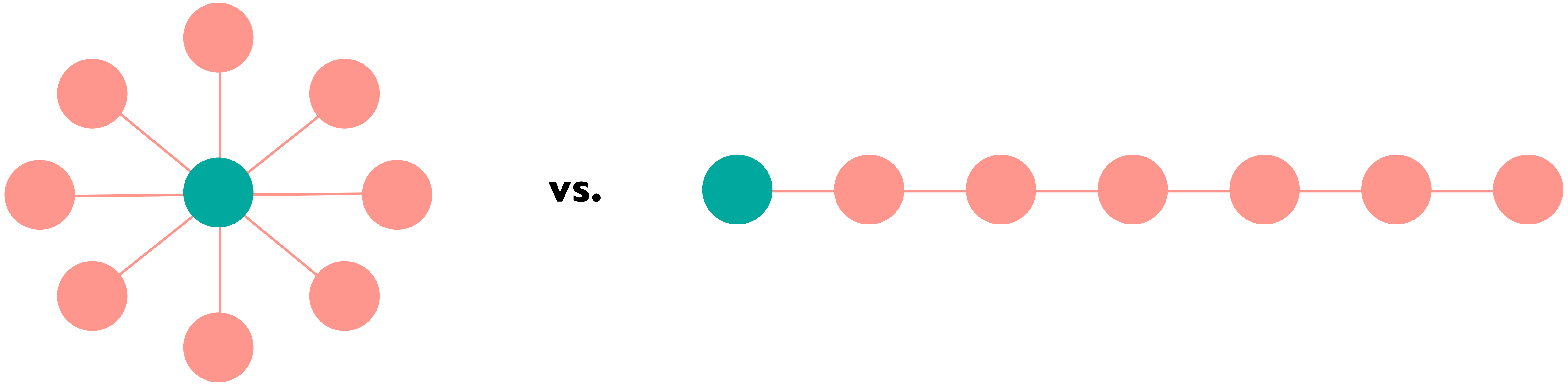
Epidemics are diseases that **travel socially**

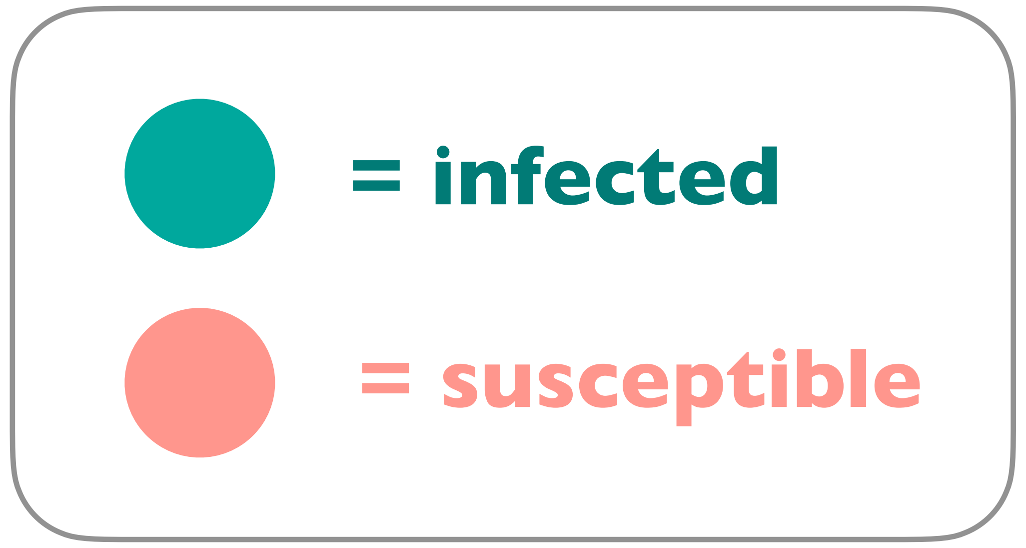
The structure of social interaction networks determine the spread of disease



Epidemics

Which outbreak is more dangerous to the population?

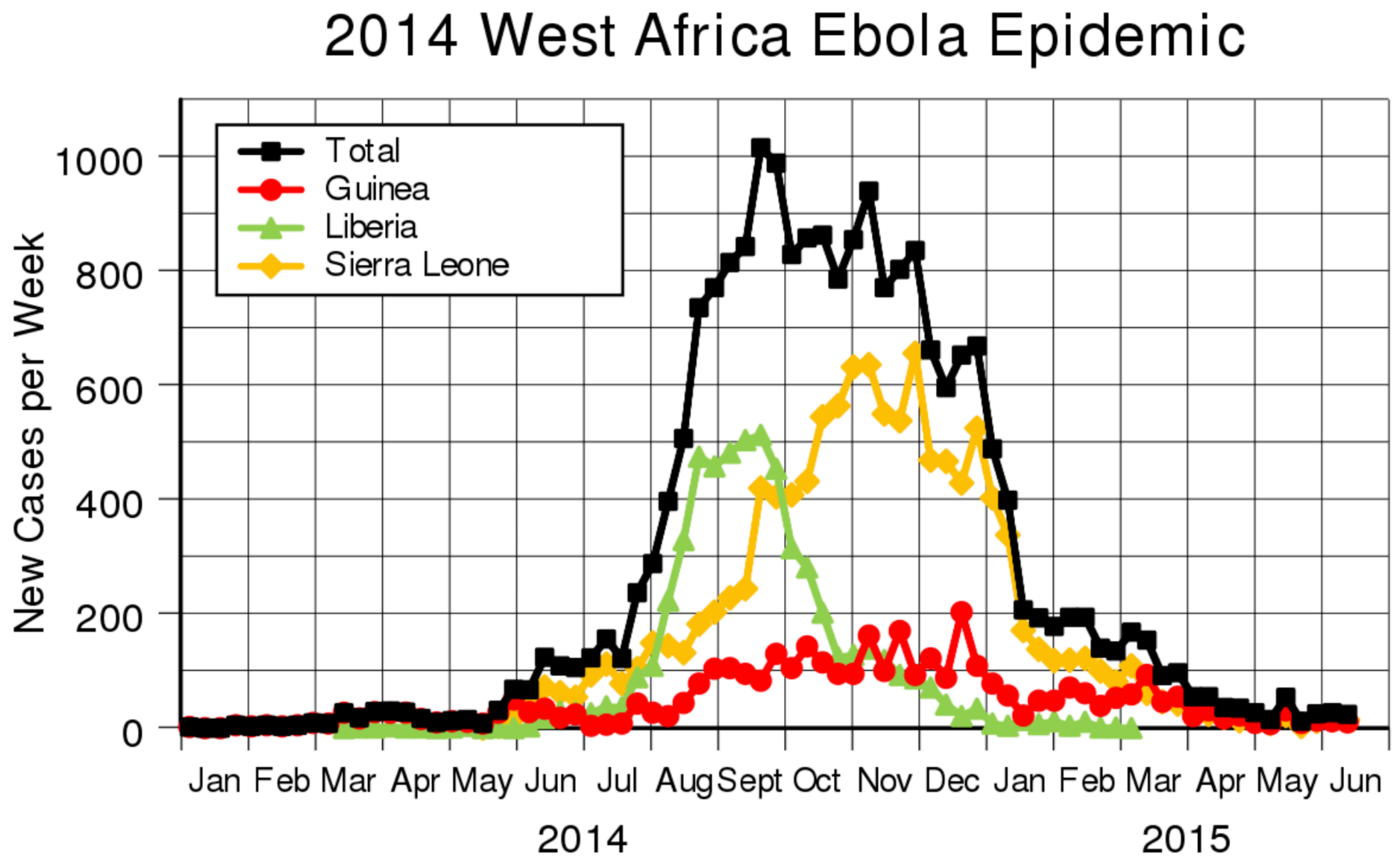




Epidemics

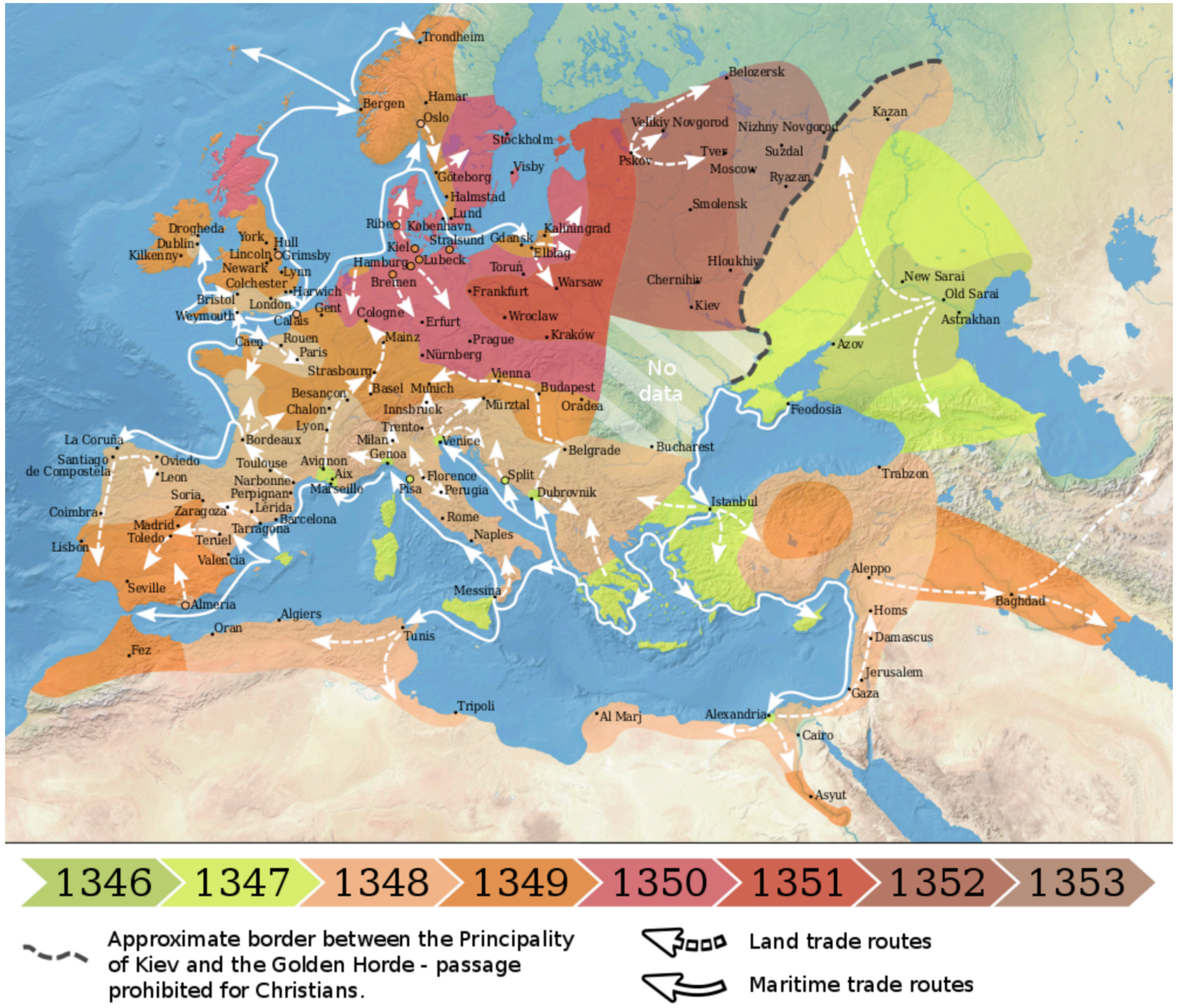
Types of epidemic diffusion:

- **Explosive spread** through a population
- **"Slow burn"** persistence over long periods of time
- Wave-like **cyclical patterns**



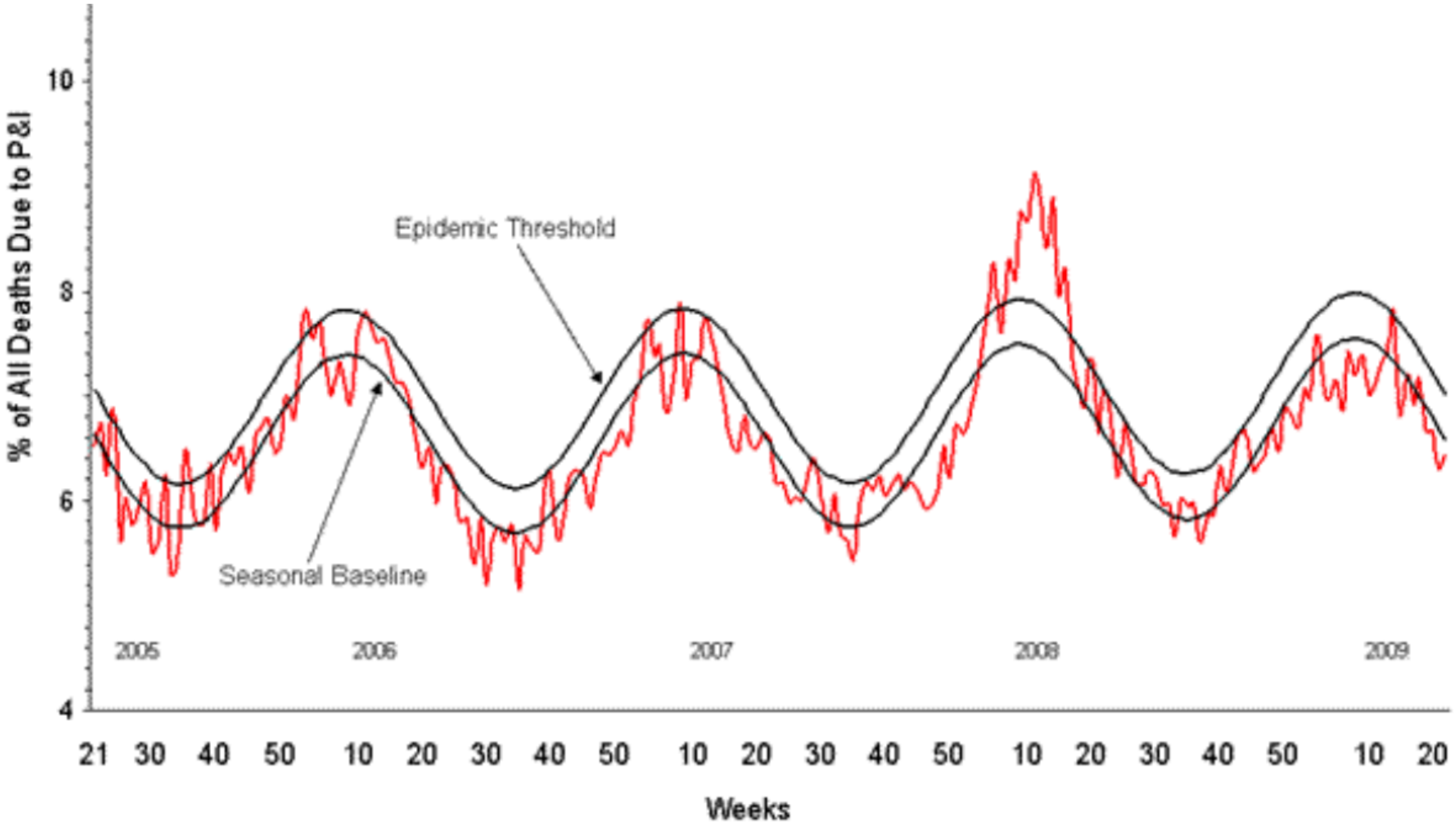
Epidemics

Explosive spread: Bubonic Plague (the “Black Death”): wiped out ~50% of the population in Europe (~150 million people) in 7 years



Epidemics

Other epidemics are cyclical

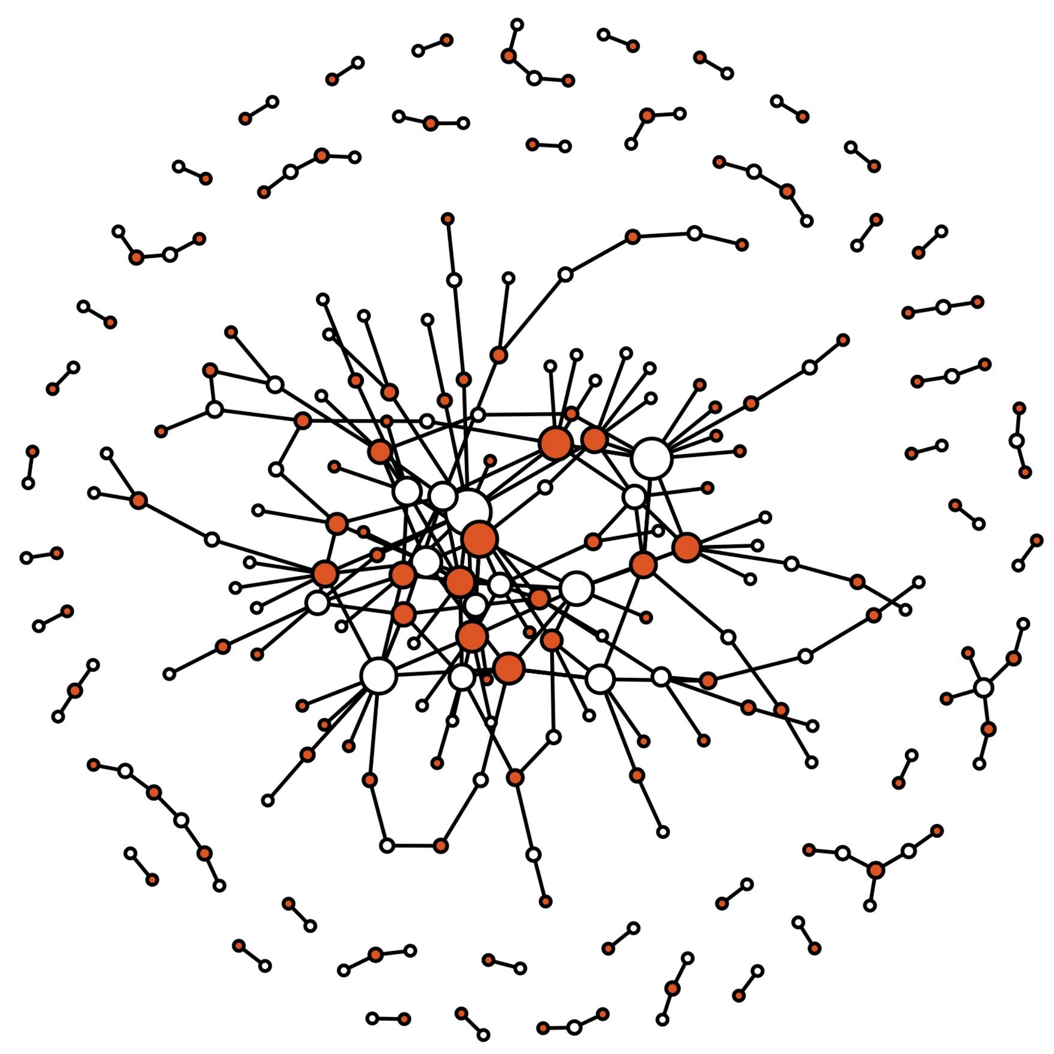


Epidemics

What determines how an epidemic might spread?

- Properties of the disease
- Structure of the network

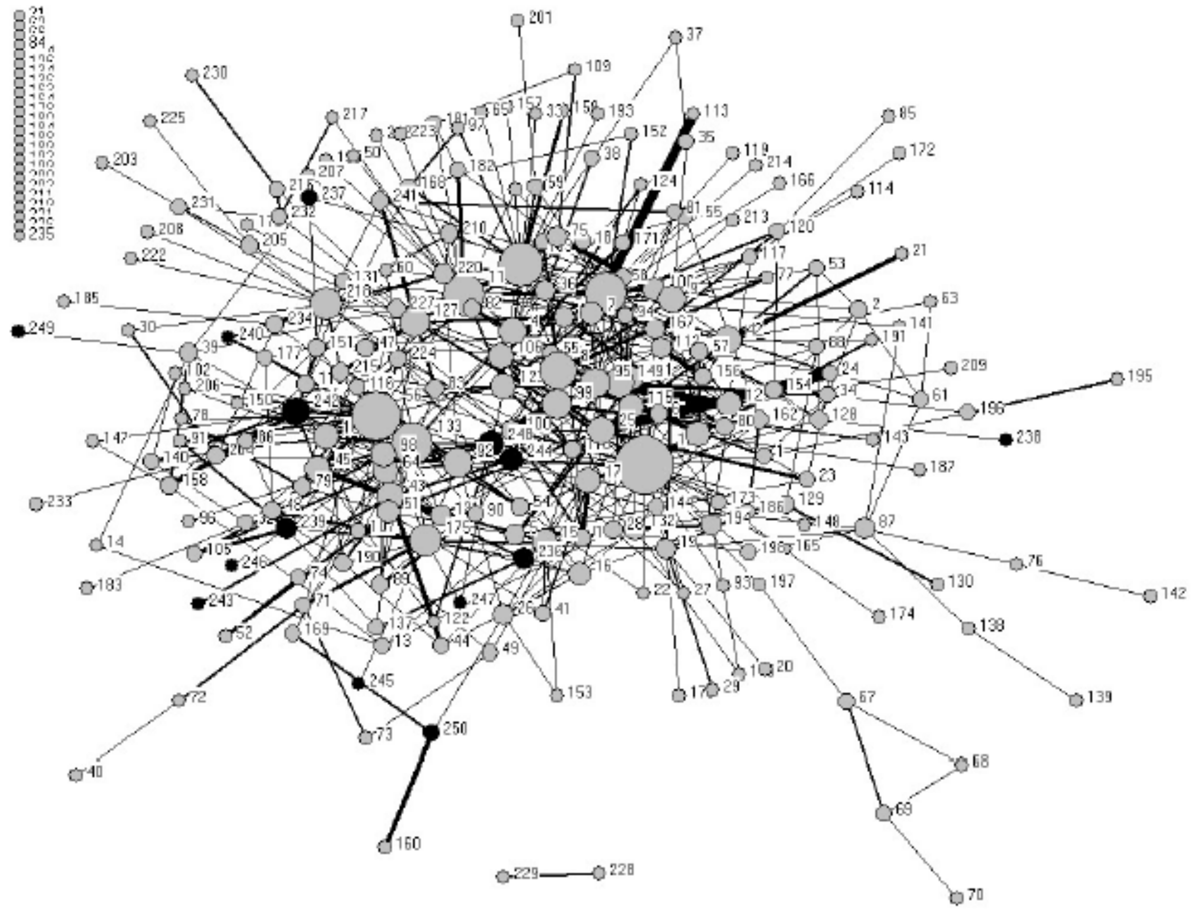
What network?



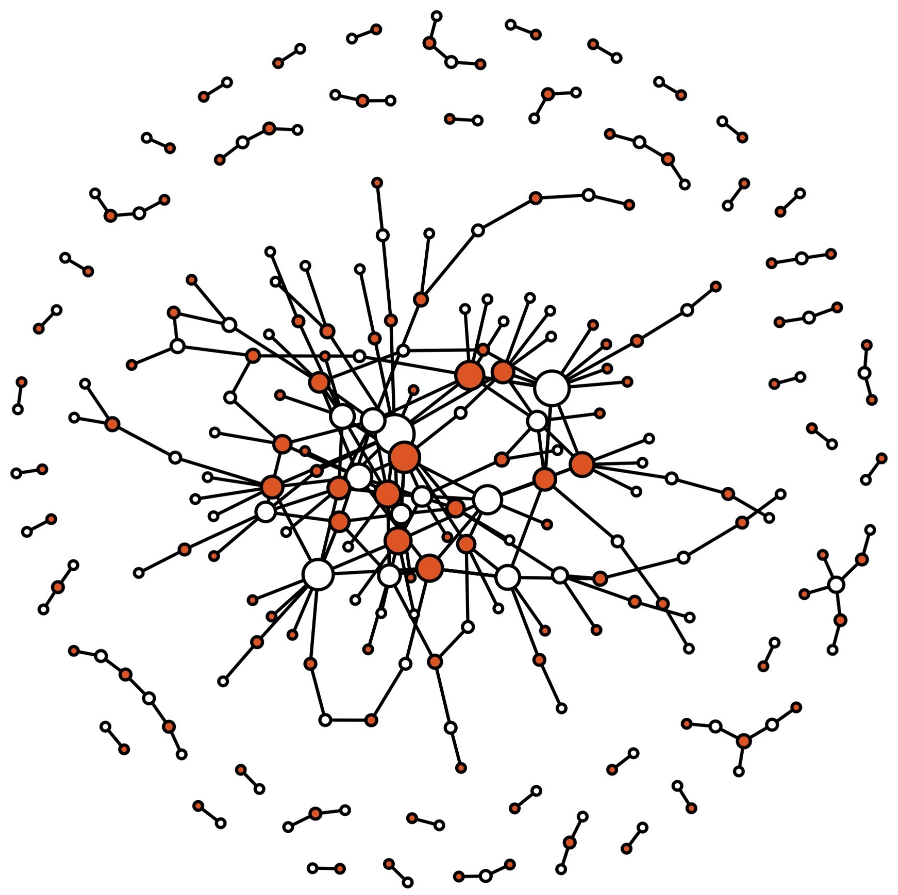
Contact Networks

Node for each person

Edge if two people come into contact with each other in a way that makes it possible for a disease to spread



UK fish farm exchanges



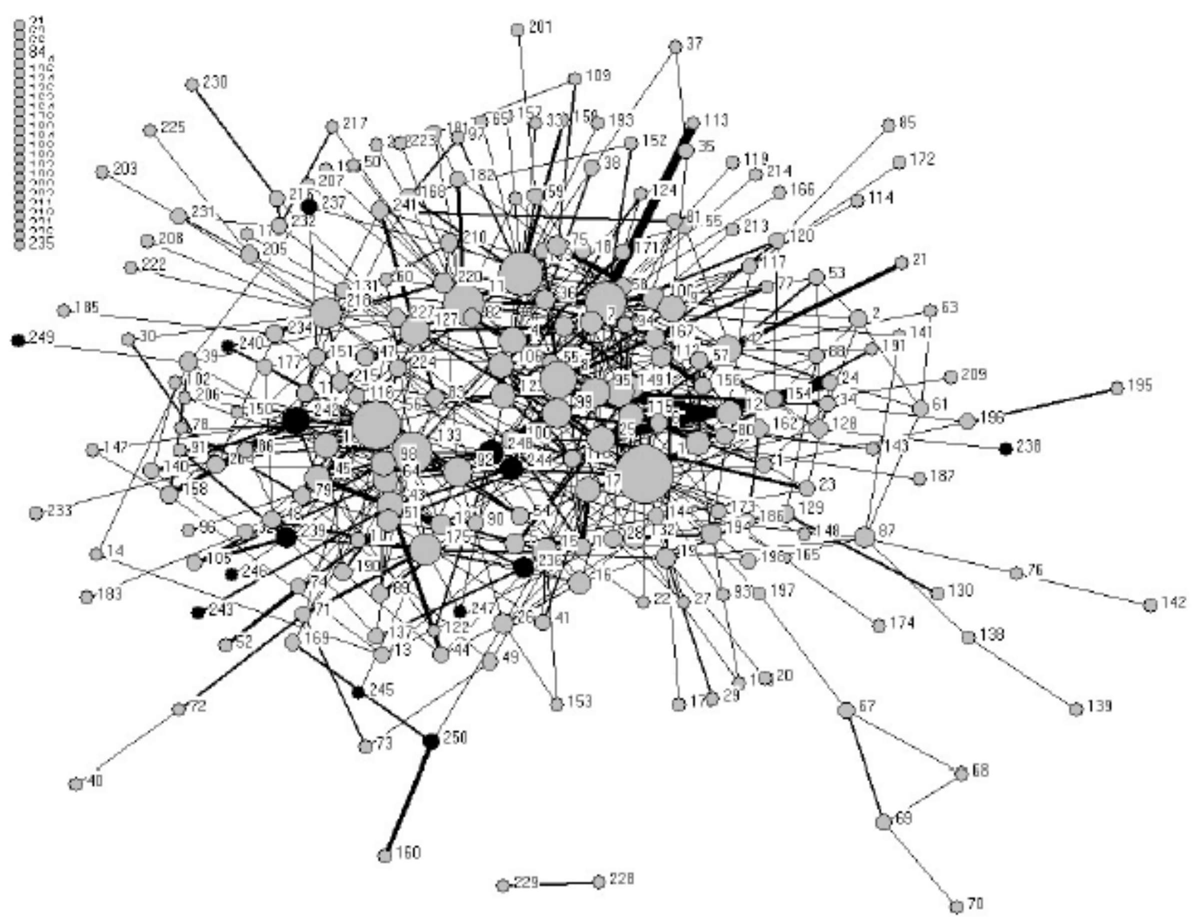
CH sexual contact network

Contact Networks

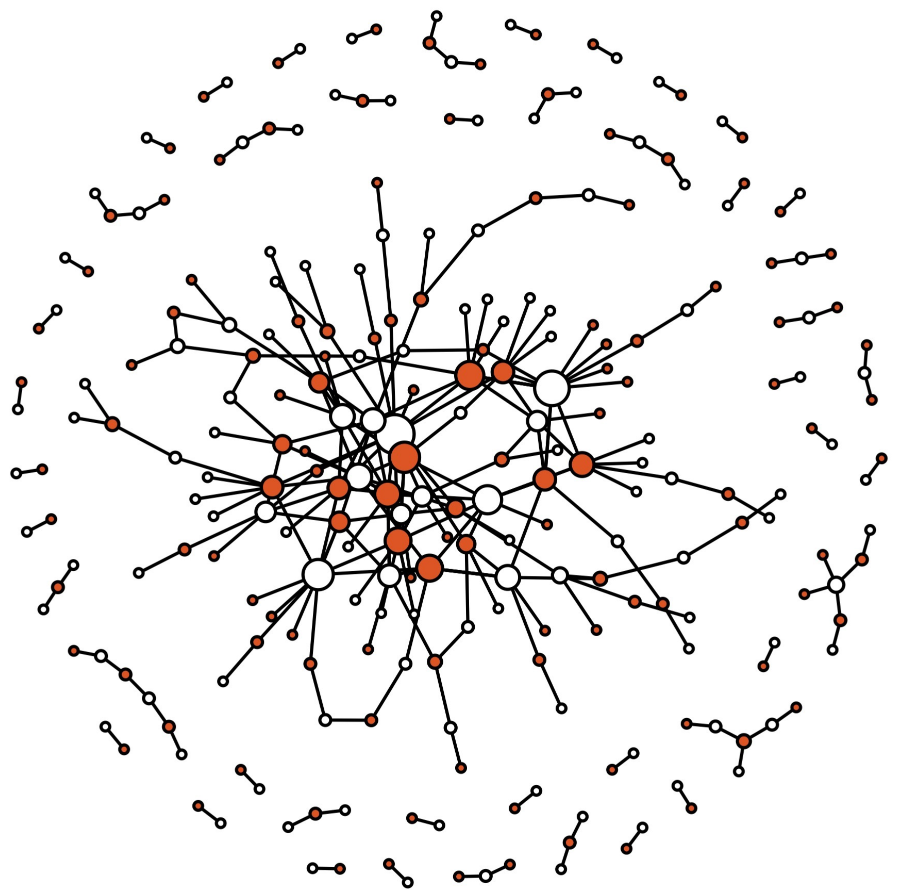
Once you've got through the laborious process of mapping out the contact network, can you us disease?

No! Definition of "contact" depends on the disease

- Airborne transmission: edge between everyone who was in the same car, etc.) — many edges
- Close contact / sexual transmission: sparser graph



UK fish farm exchanges

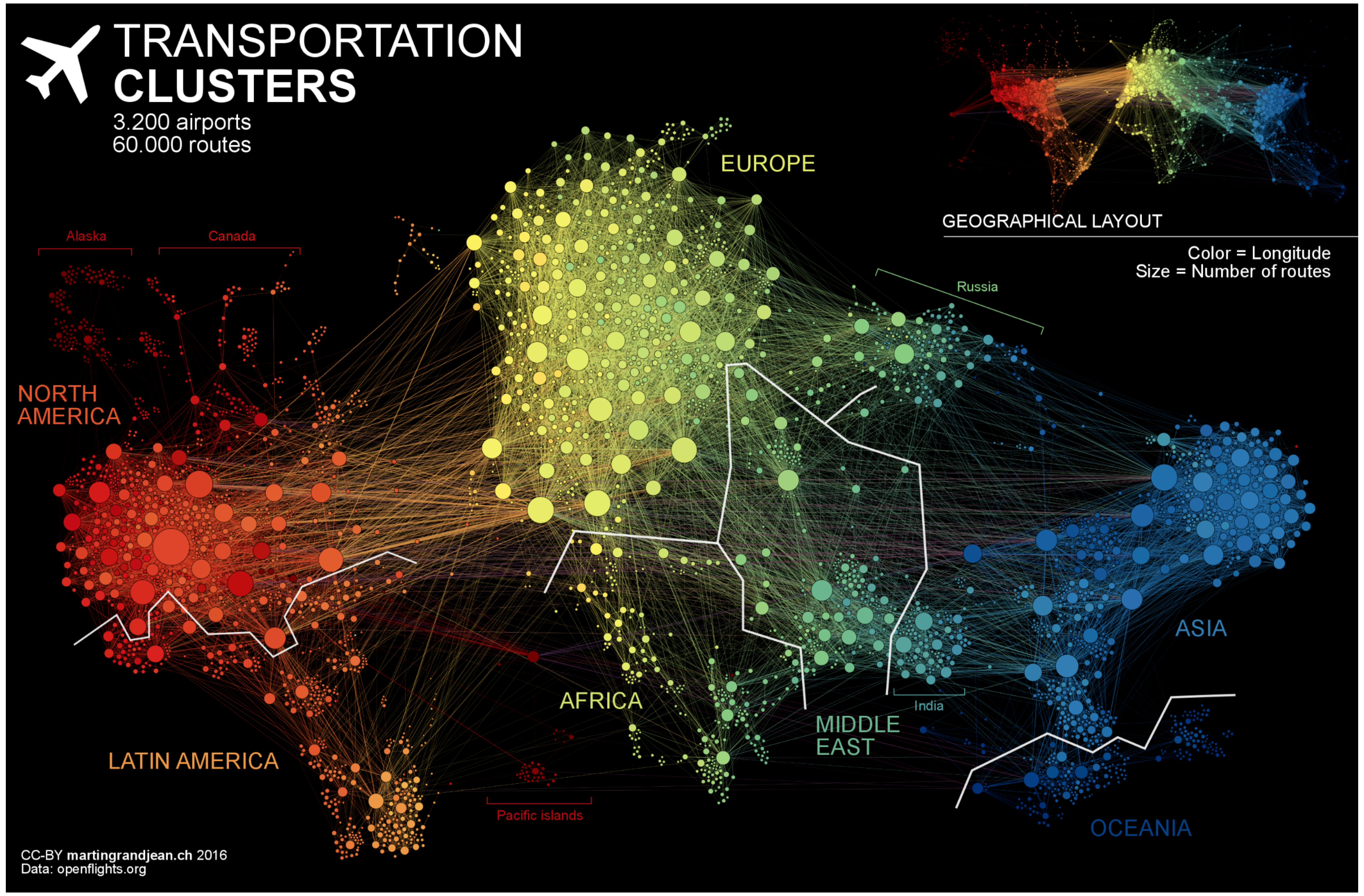


CH sexual contact network

Contact Networks

Big part of real-world epidemic research is constructing contact networks

Lots of work on travel patterns in cities, the worldwide airline network, etc. to understand how diseases can spread in today's world

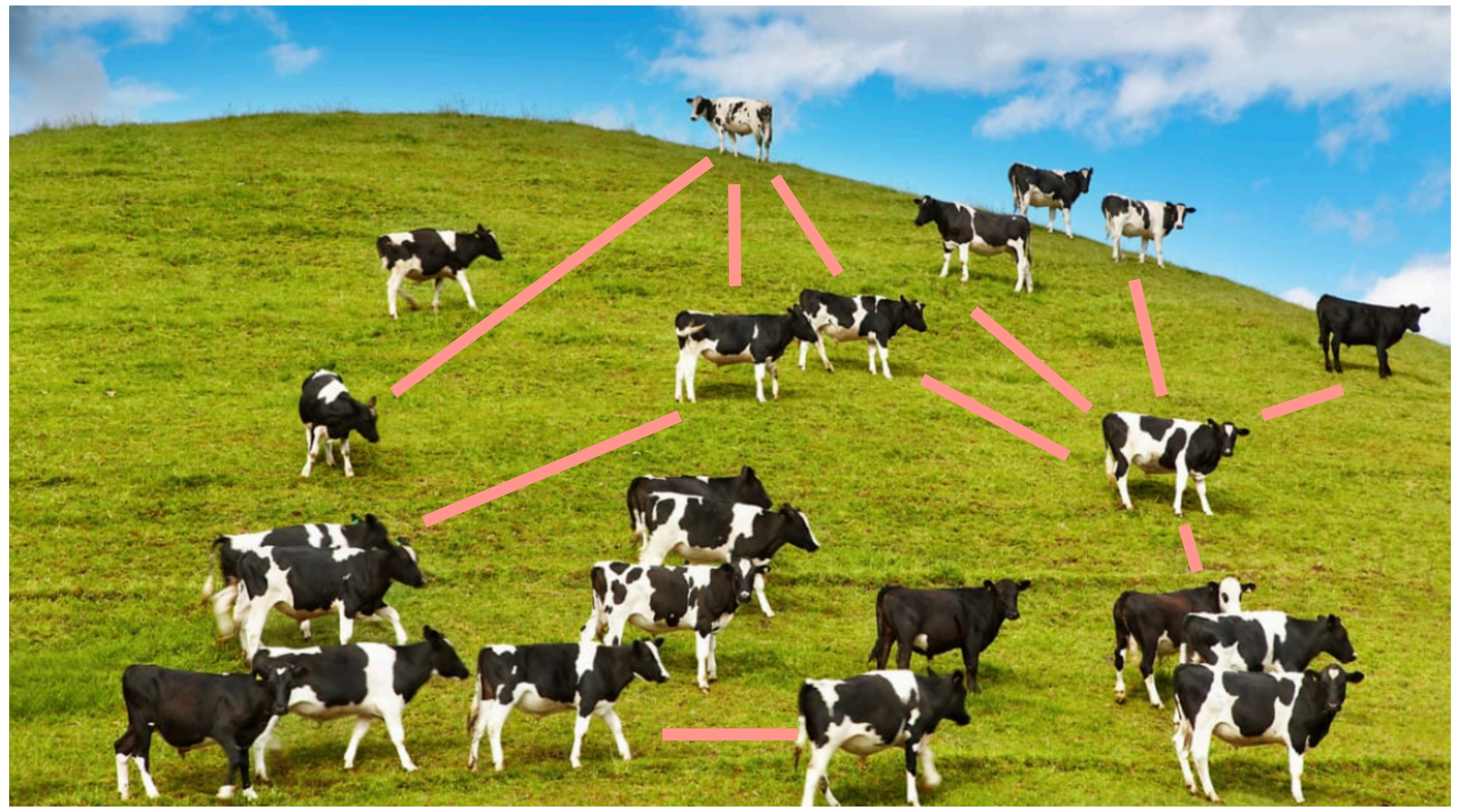


http://www.martingrandjean.ch/connected-world-air-traffic-network/

Contact Networks

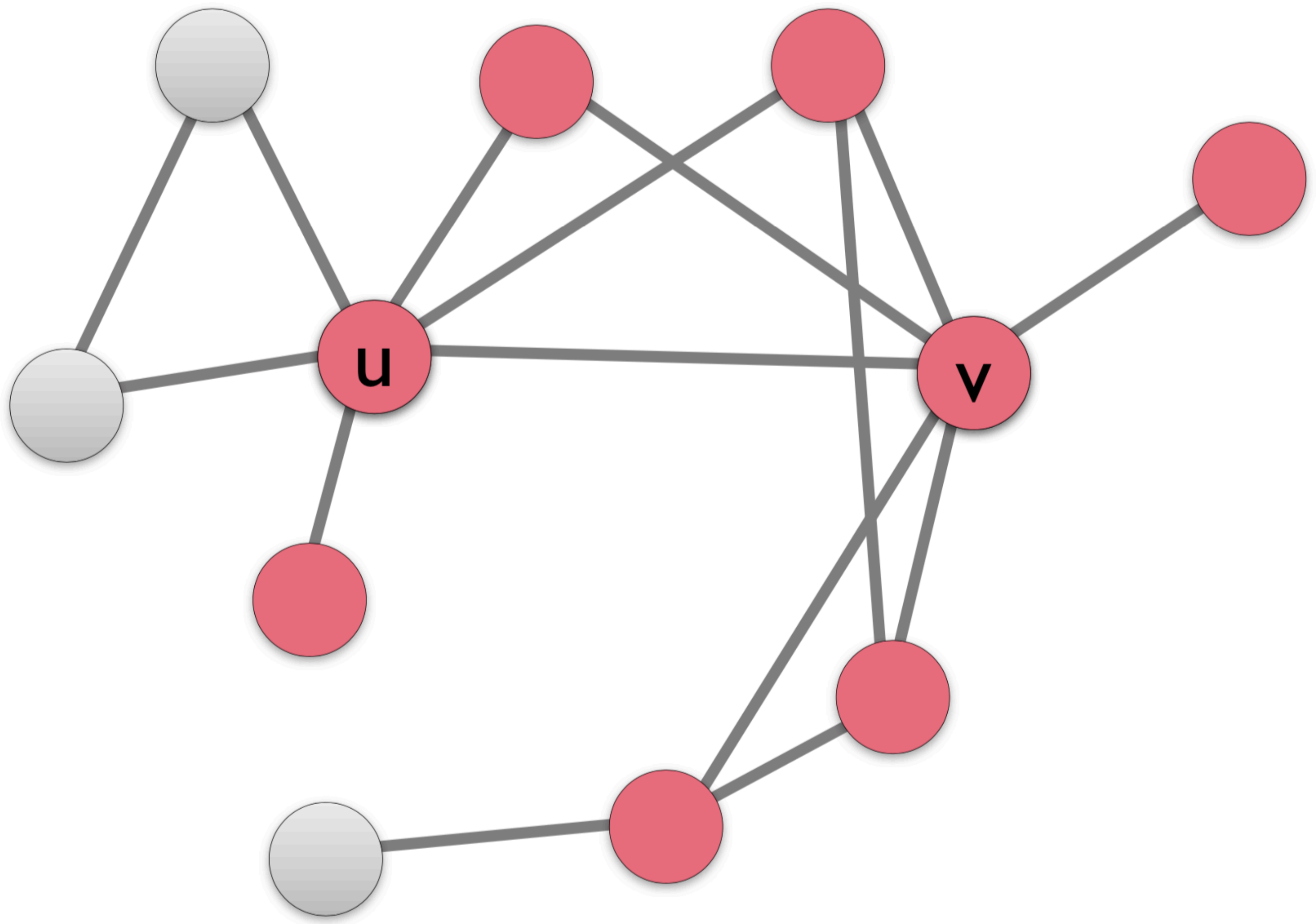
Not just human contact networks

Animal/livestock networks and plant networks

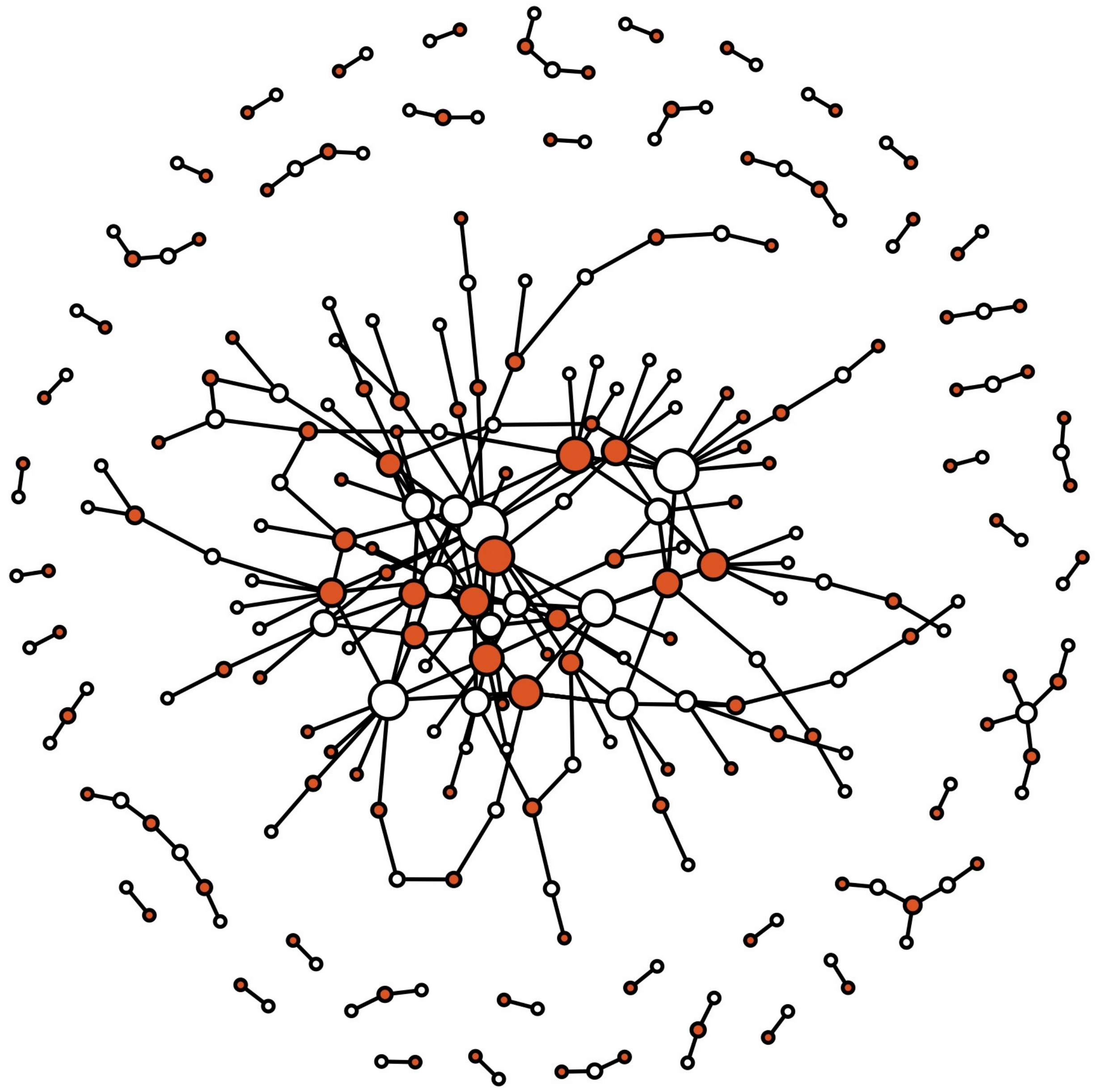


Behavioural vs. Biological Contagion

Biological/epidemic diffusion: **no decision-making!**



Decision cascade



High school contact network

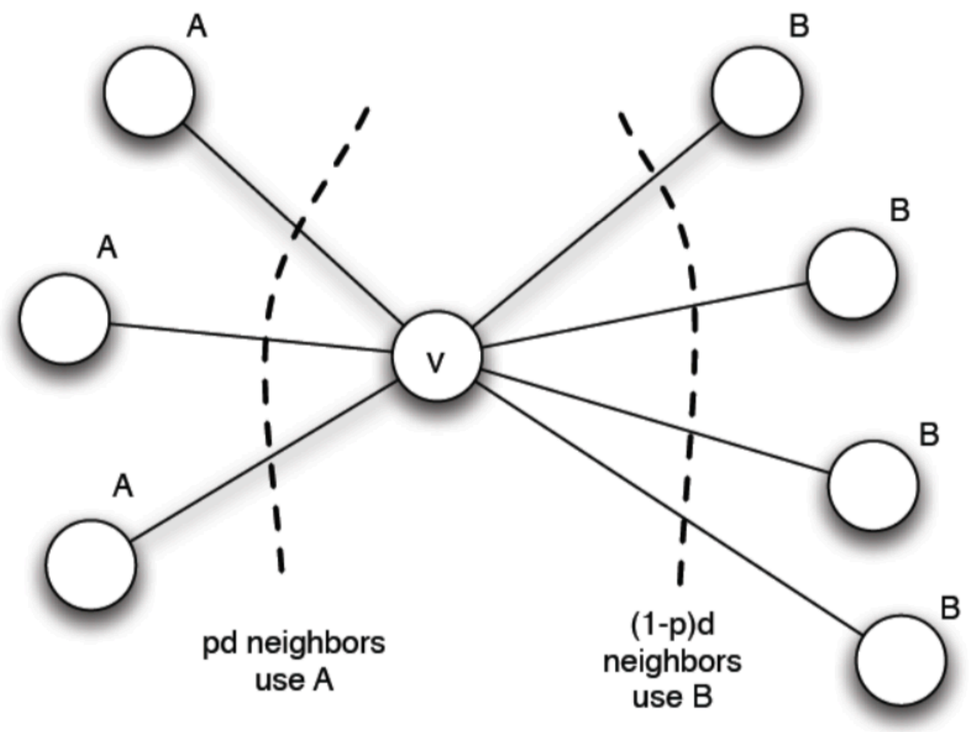
Modeling Epidemic Diffusion

Biggest difference: model transmission as **random**

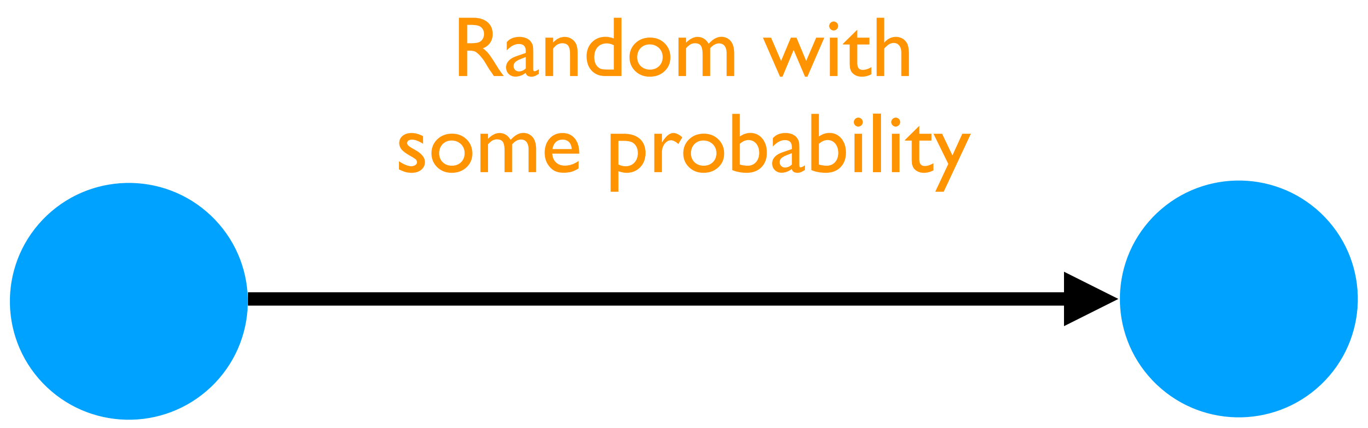
No decision-making, but also the processes by which diseases spread from one person to another are so complex and unobservable at the individual level that it's most useful to think of them as random

Use randomness to abstract away difficult biological questions about the mechanics of spread

Behaviour (last class):



Epidemics (today):

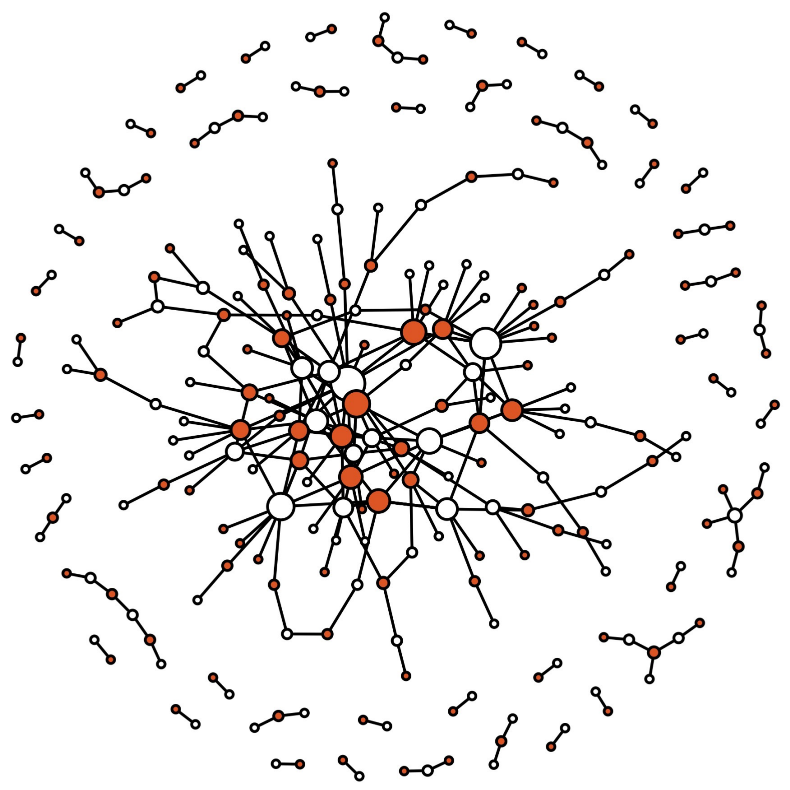


Modeling Epidemic Diffusion

Branching Process

Basic structure of epidemic diffusion:

- Someone gets infected
- Then they infect some number of people
- Those people infect others



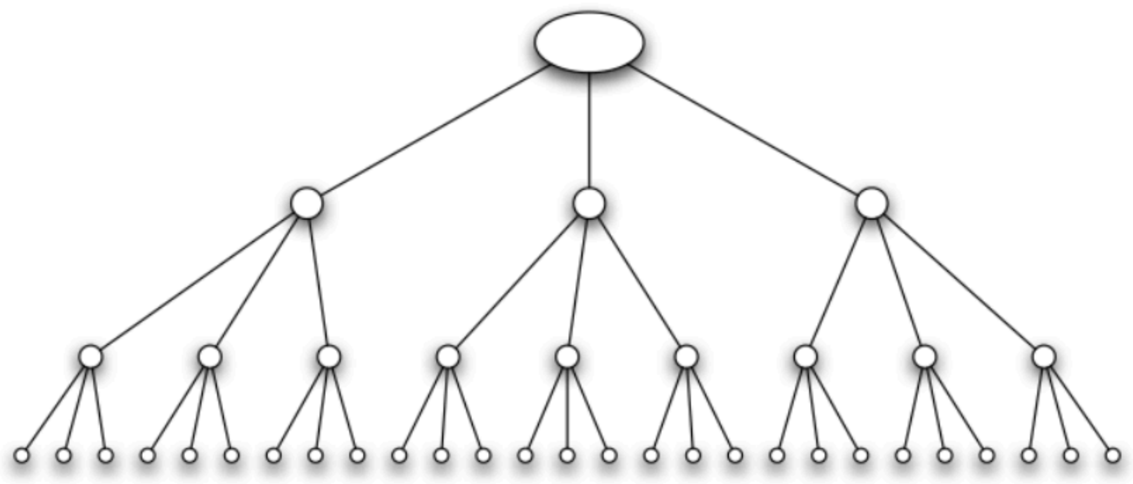
Branching Process

Model as a random process on a tree:

Wave 1: First person infected, infects each of k neighbors with independent probability p

Wave 2: For each infected person, they infect each of k neighbors with independent probability p

Wave 3+: repeat for each infected person



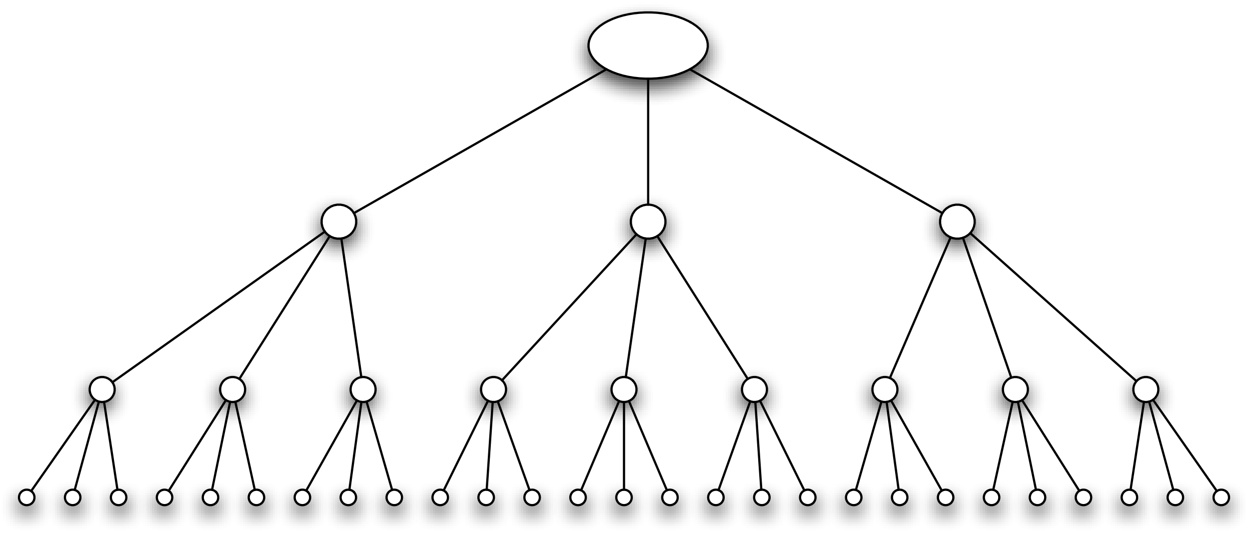
Here k=3

Extends infinitely below

Branching Process

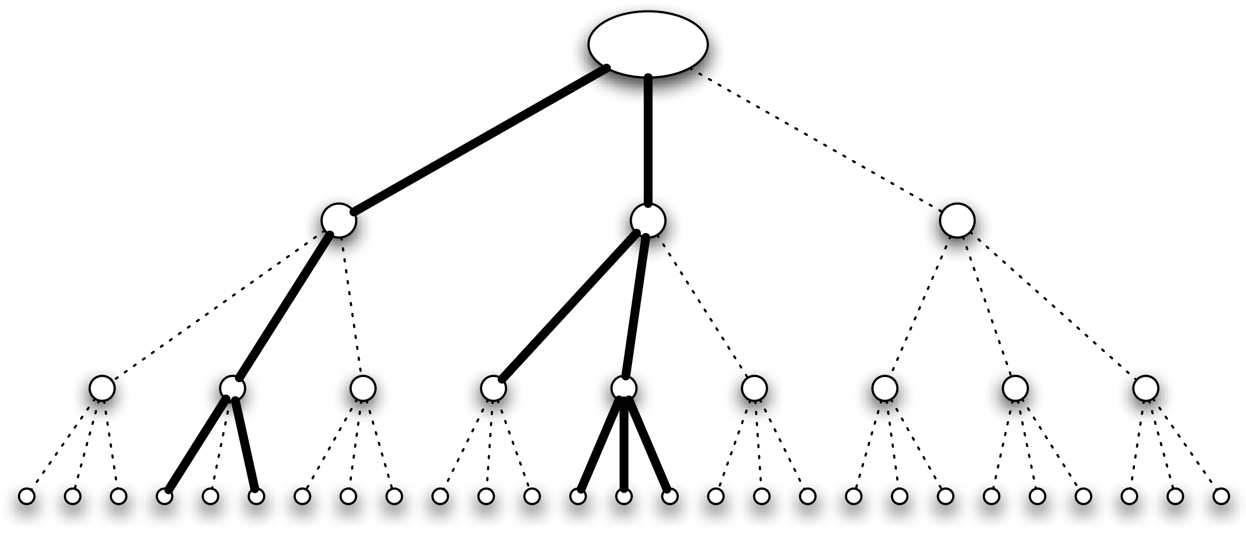
Model parameters:

k: number of individuals each person can possibly infect:



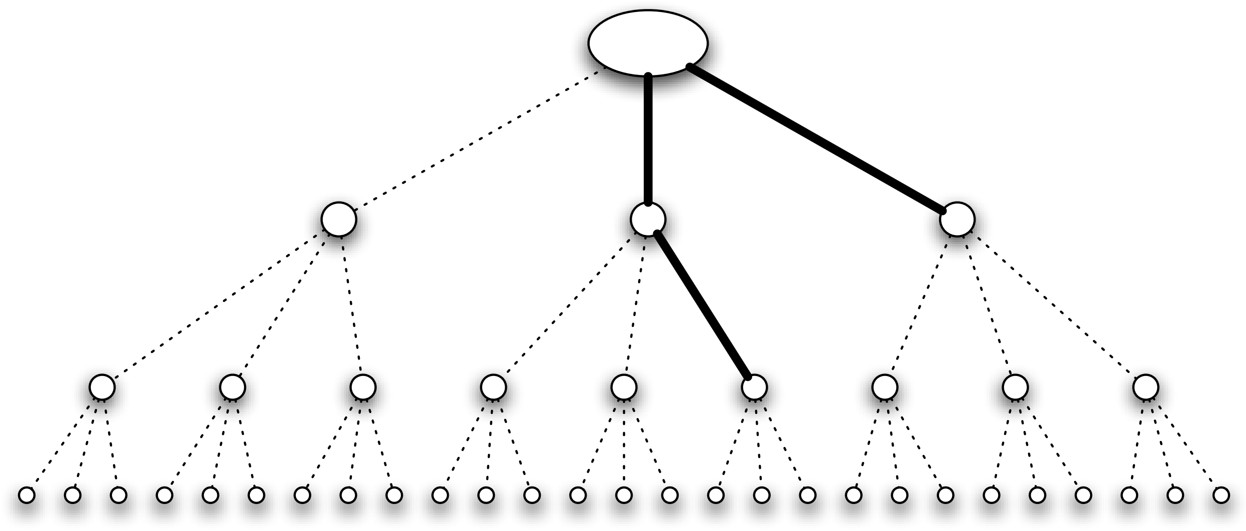
(a) *The contact network for a branching process*

Higher transmission probability *p*:



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

Lower transmission probability *p*:



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

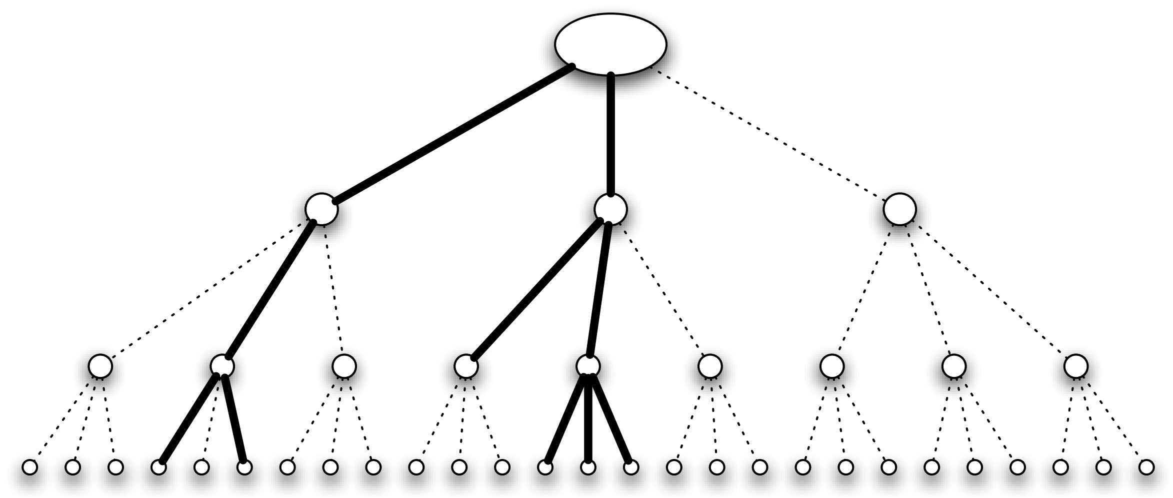
Branching Process: Outcomes

Only two possibilities in the long run: **blow up** or **die out**

How does it die out?

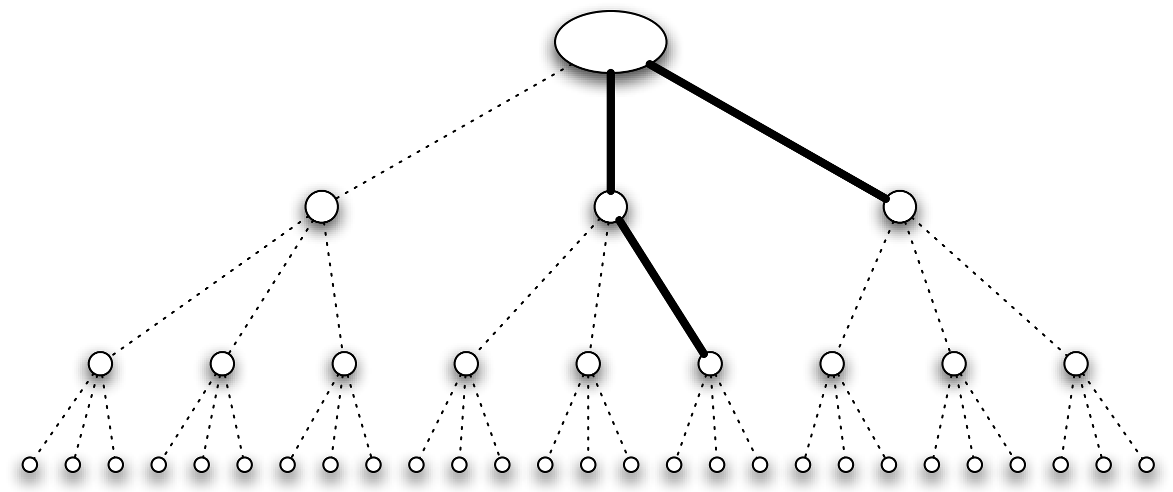
- Dies out if and only if none of the nodes on a given level are infected

Disease might blow up:



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

Disease has already died out:



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

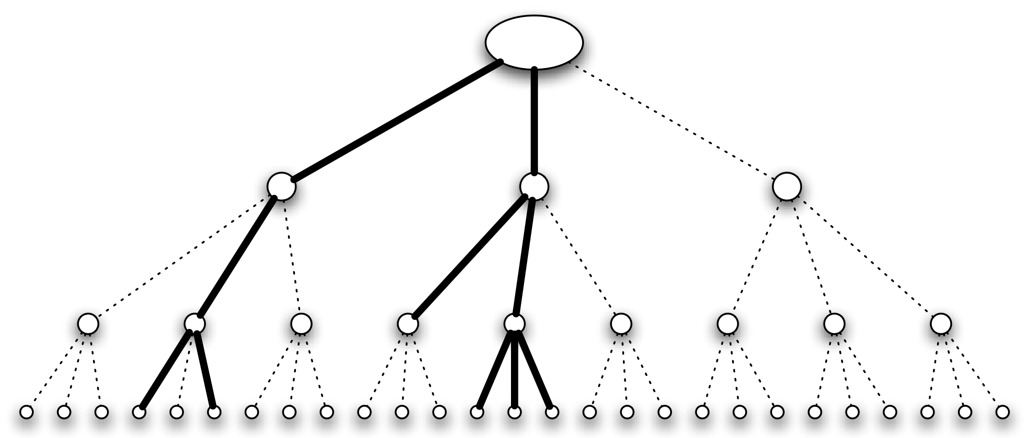
Branching Process

Only two possibilities in the long run: **blow up** or **die out**

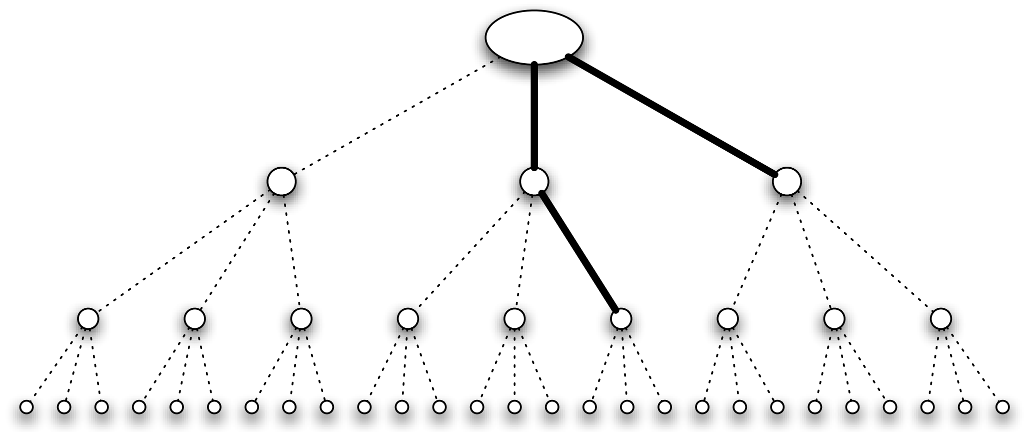
How does it die out?

- Dies out if and only if none of the nodes on a given level are infected

Define **Basic reproductive number R_0:** the number of expected new cases caused by an individual



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



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

Branching Process

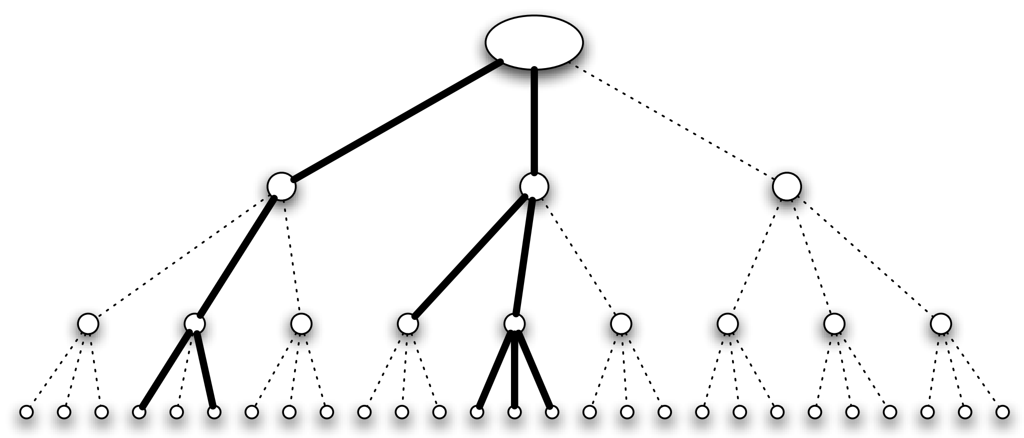
Only two possibilities in the long run: **blow up** or **die out**

How does it die out?

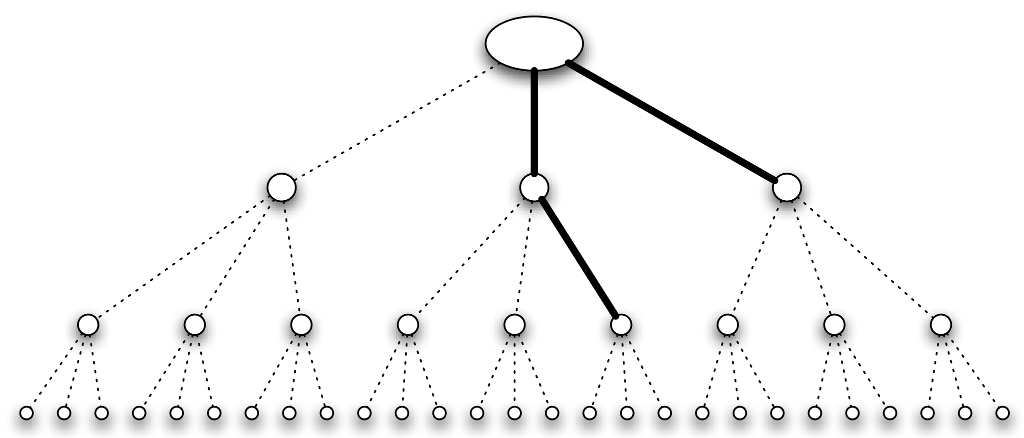
- Dies out if and only if none of the nodes on a given level are infected

Define **Basic reproductive number R_0:** the number of expected new cases caused by an individual

$$R_0 = pk$$



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



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

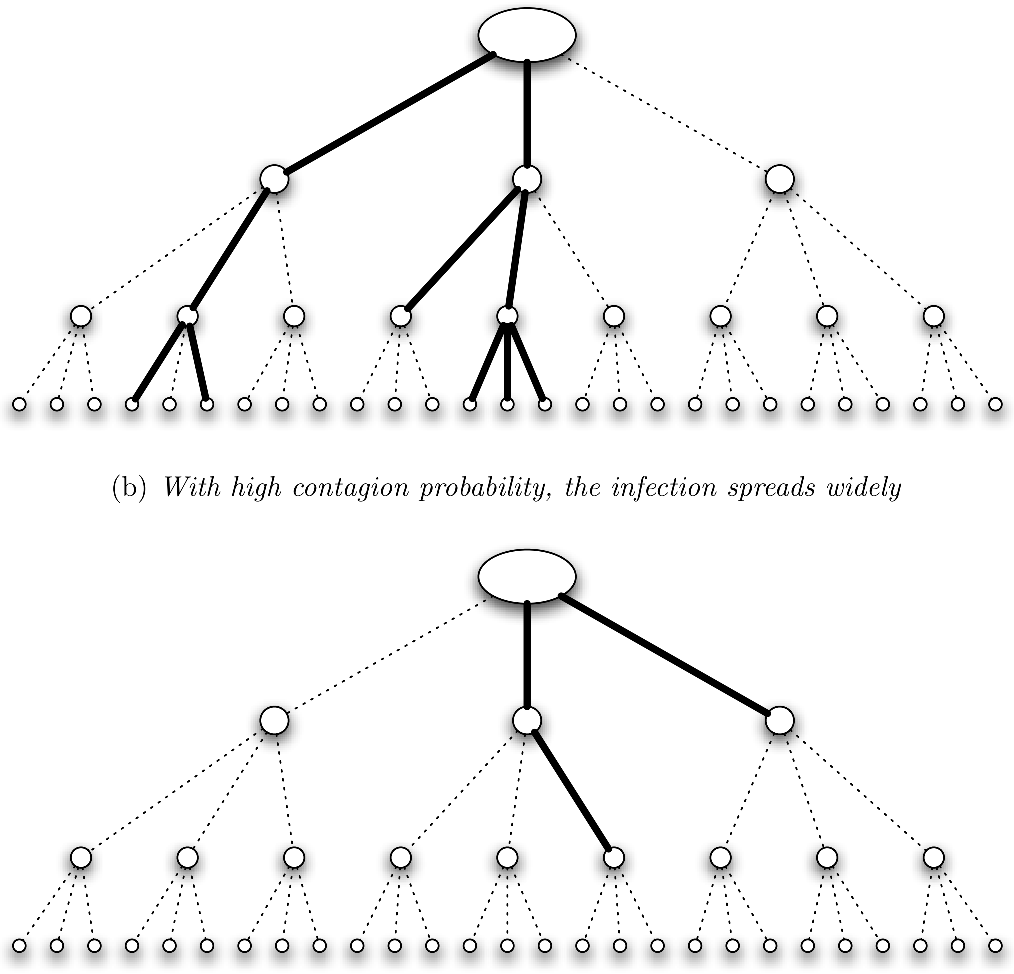
Branching Process: R_0

Claim: Epidemic spread in the branching process model is entirely controlled by the reproductive number R_0 :

- If $\mathbf{R_0 < 1}$ then with probability 1 the disease dies out after a finite number of steps.
- If $\mathbf{R_0 > 1}$ then with probability > 0 the disease persists by infecting at least one person in each wave.

"Go big or go home."

$R_0 = pk$



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

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

Branching Process: R_0

$R_0 = pk < 1$:

With probability 1 the disease dies out after a finite number of steps

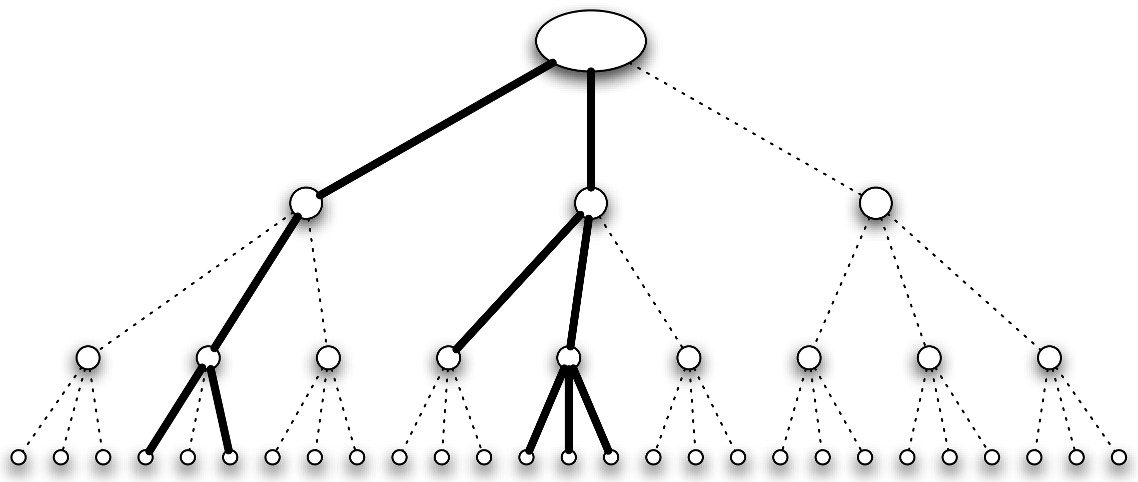
Below replacement; disease isn't able to replenish itself.

Even if it grows momentarily, it trends downward.

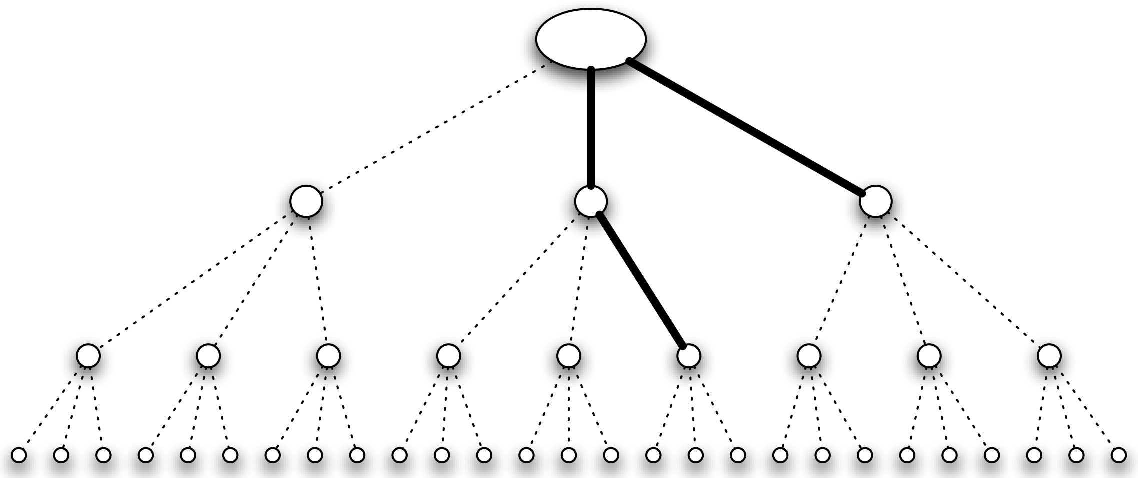
$R_0 = pk > 1$:

with probability > 0 the disease persists by infecting at least one person in each wave

Always trending upward. Could still get "unlucky" and die out, but there's a non-zero chance it runs forever.



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



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

Branching Process: R_0

$R_0 = pk < 1$:

With probability 1 the disease dies out after a finite number of steps

Below replacement; disease isn't able to replenish itself.

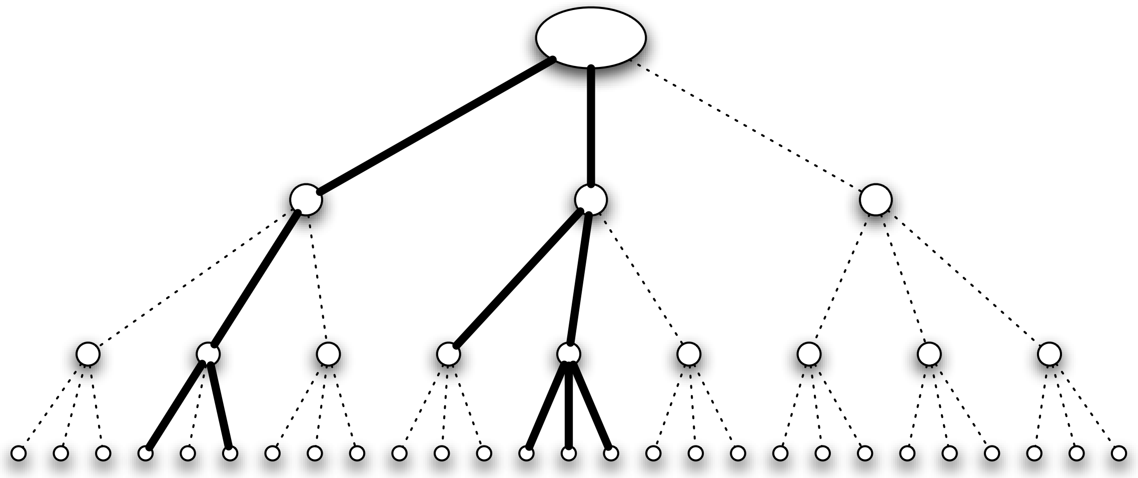
Even if it grows momentarily, it trends downward.

$R_0 = pk > 1$:

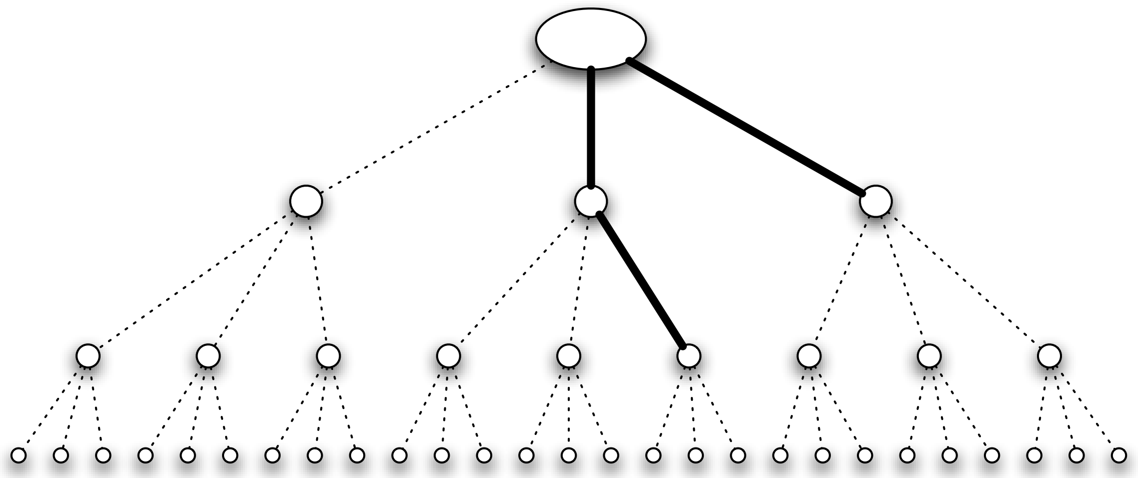
with probability > 0 the disease persists by infecting at least one person in each wave

Always trending upward. Could still get "unlucky" and die out, but there's a non-zero chance it runs forever.

What happens when **p** or **k** change near **pk=1**?



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



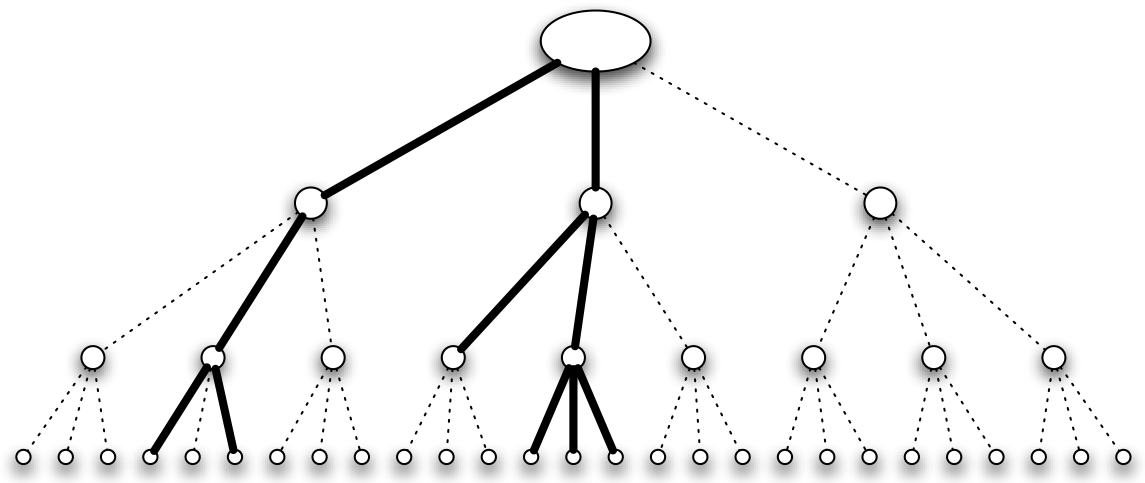
(c) *With low contagion probability, the infection is likely to die out quickly*

Sensitivity of p and k

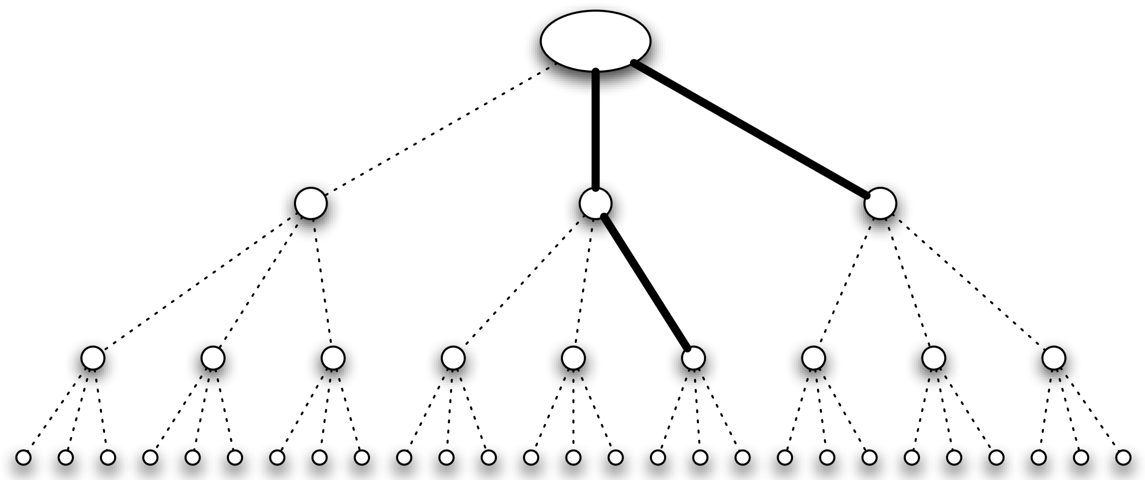
Because epidemics have a "critical threshold", it can be worth it to do a lot of work or expend resources to push **p** or **k** **down** a little bit.

Quarantine = reduce **k**

Improved sanitation = reduce **p**



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



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

Disease	Transmission	R_0
Measles	Airborne	12–18
Diphtheria	Saliva	6-7
Smallpox	Airborne droplet	5–7
Polio	Fecal-oral route	5–7
Rubella	Airborne droplet	5–7
Mumps	Airborne droplet	4–7
HIV/AIDS	Sexual contact	2–5
Pertussis	Airborne droplet	5.5[2]
SARS	Airborne droplet	2–5[3]
Influenza (1918 pandemic strain)	Airborne droplet	2–3[4]
Ebola (2014 Ebola outbreak)	Bodily fluids	1.5-2.5[5]

COVID-19: ~2

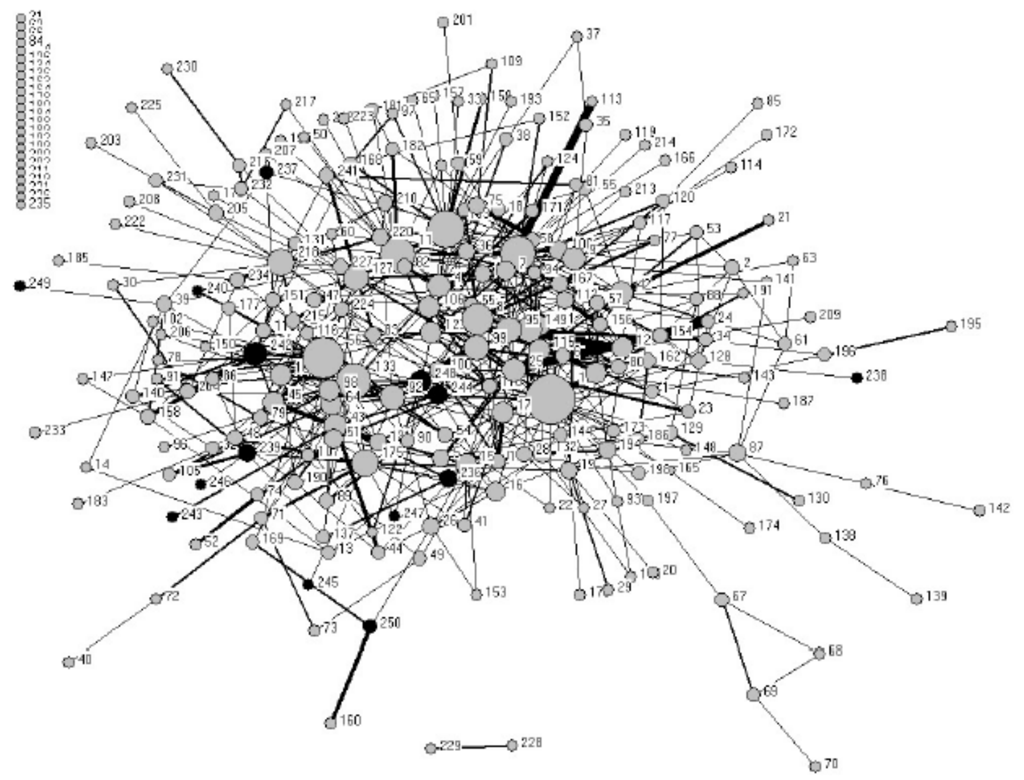
General Models of Contagion

Epidemics on General Graphs

We just studied epidemics as ideal trees

But of course real-life networks are more complicated than that

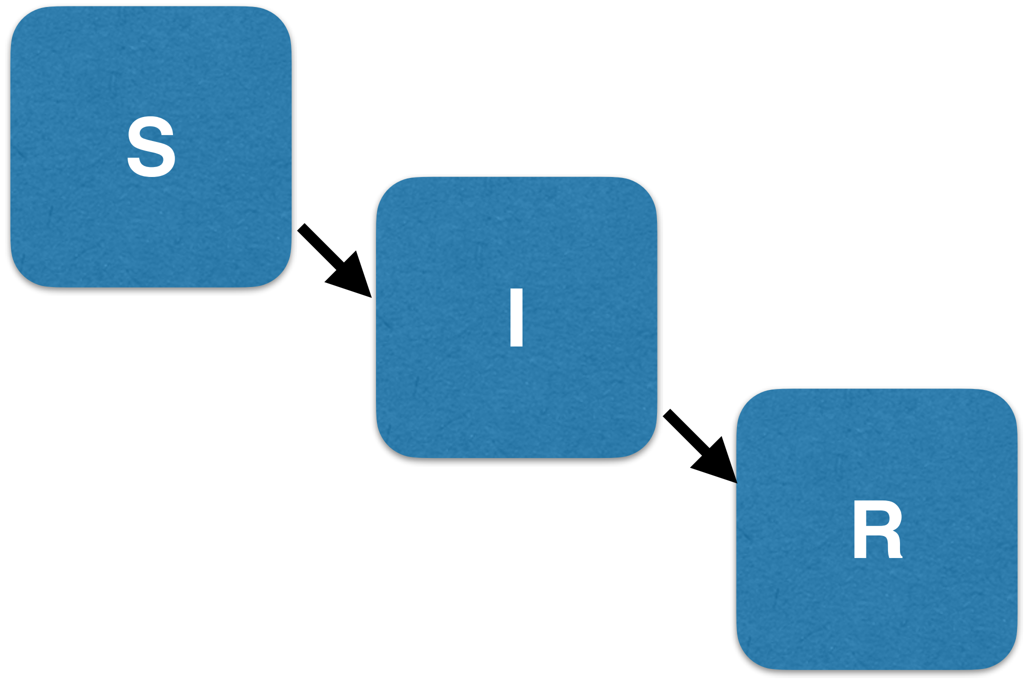
What does epidemic diffusion look like in general graphs?

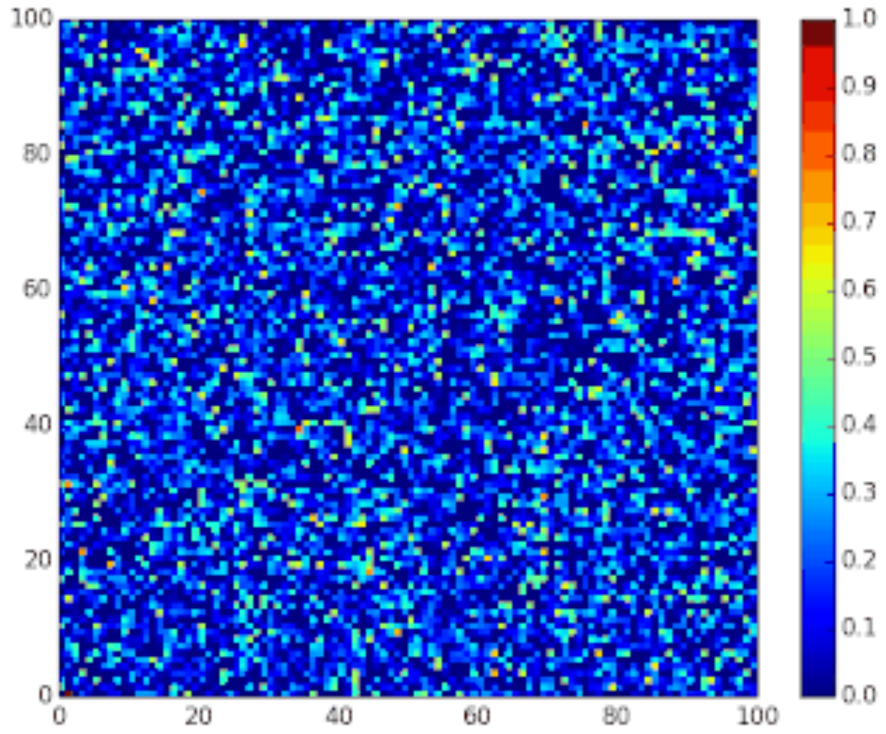


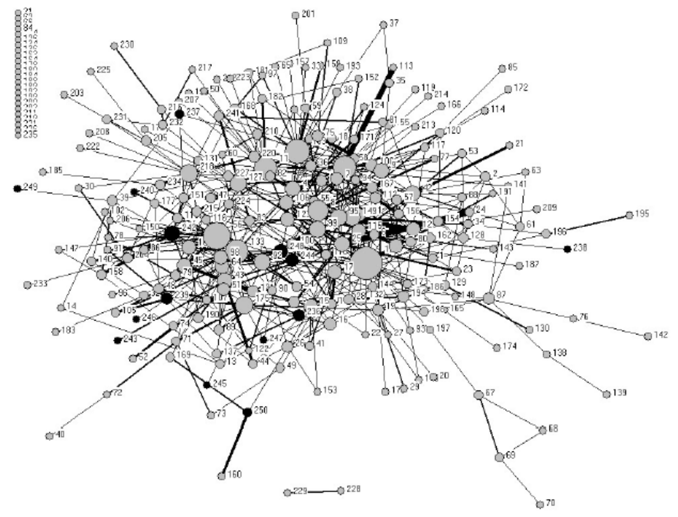
SIR Epidemic Models

Simple lifecycle model with three stages:

- **S** = Susceptible
- **I** = Infectious: node is infected and infects with prob **p**
- **R** = Removed: after $\mathbf{t_I}$ time, no longer infected or infectious







SIR Epidemic Models

S = Susceptible

I = Infectious: node is infected and infects with prob **p**

R = Removed: after $\mathbf{t_I}$ time, no longer infected or infectious

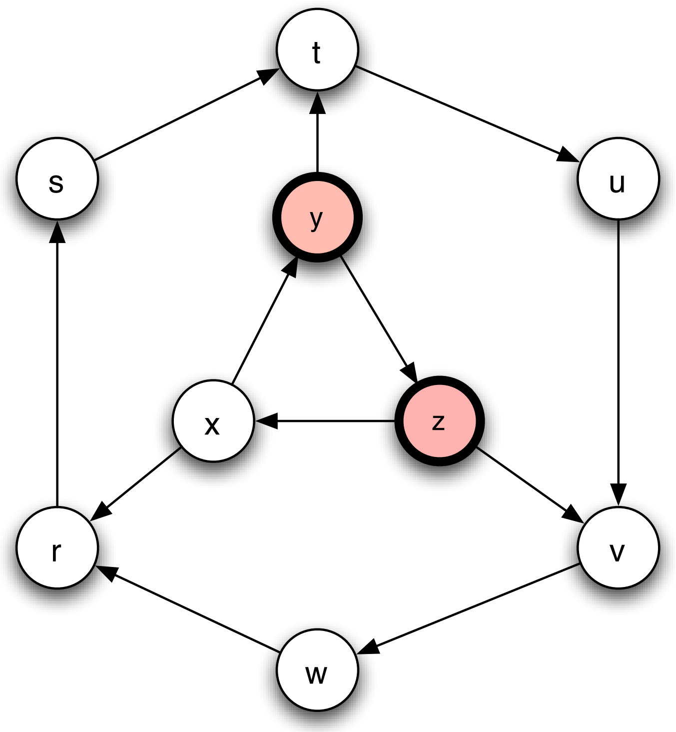
Initially some nodes in **I** state, rest in **S** state.

Each node in I state remains infected for $\mathbf{t_I}$ time steps

During each step, each node has probability **p** of infecting each susceptible neighbour

After $\mathbf{t_I}$ time steps, no longer **S** nor **I;** removed to **R**

SIR Epidemics on Networks

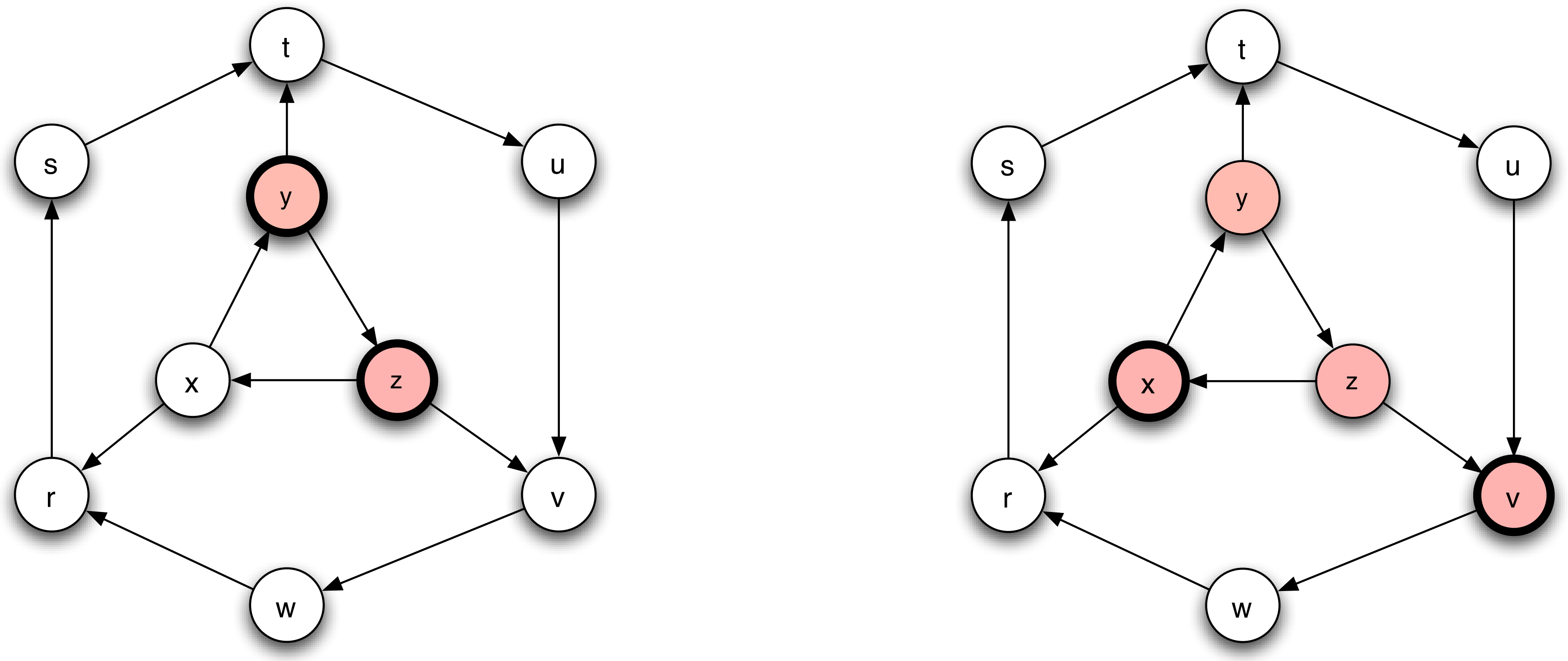


(a)

p=1/2

t_I =1

SIR Epidemics on Networks



(a) (b)

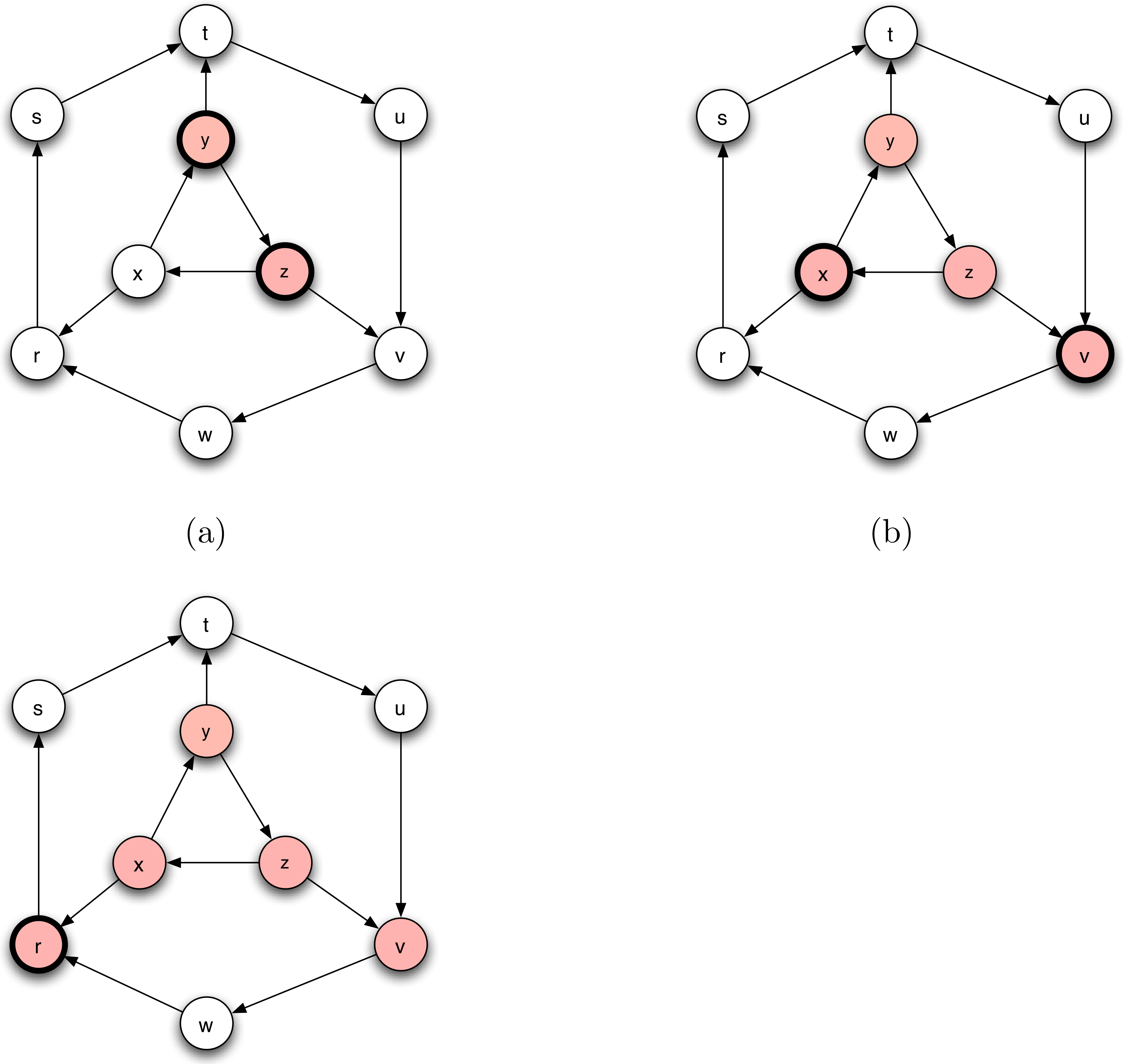
p=1/2

$\mathbf{t_I} = 1$

SIR Epidemics on Networks

p=1/2

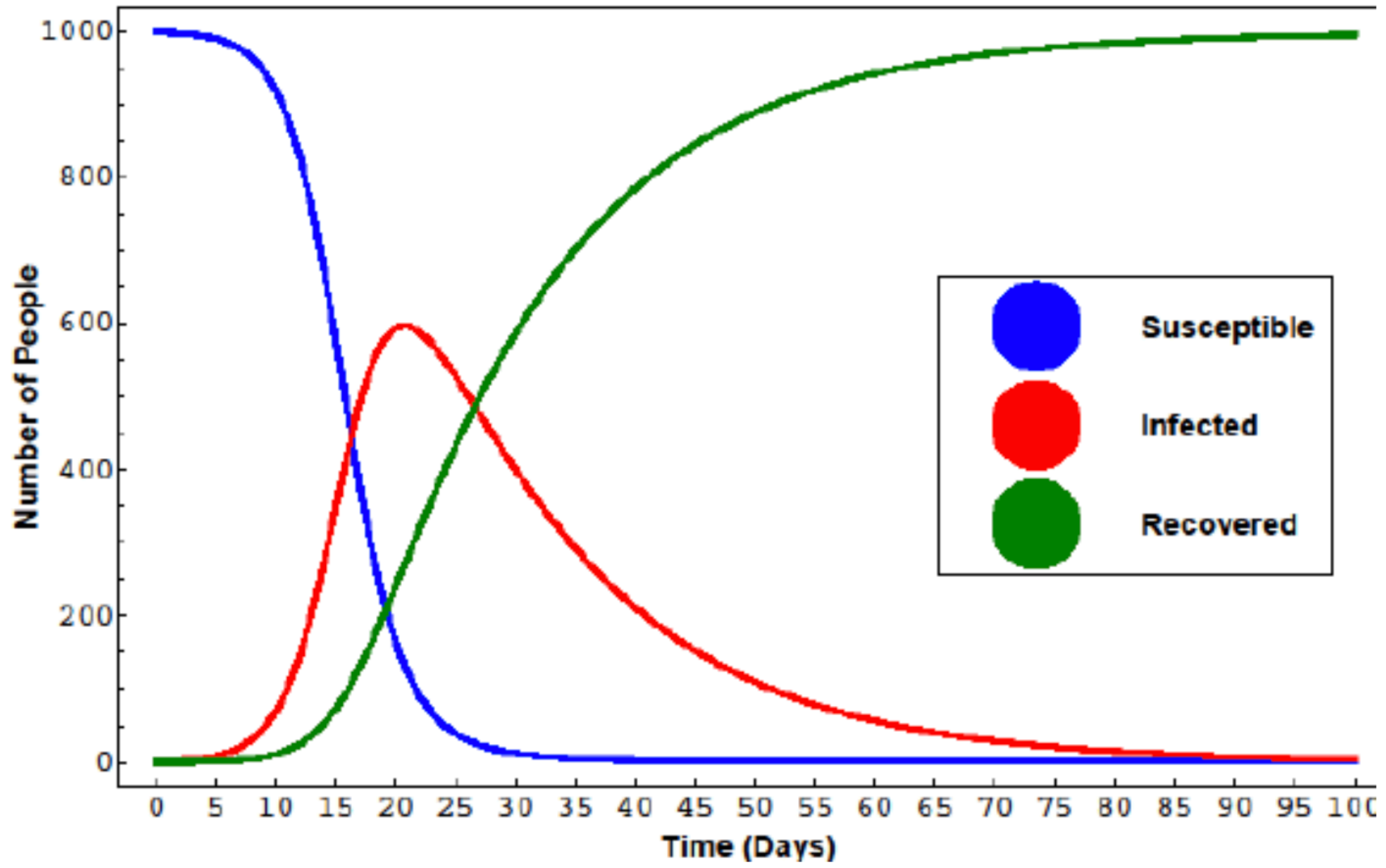
$\mathbf{t_I}$ =1



(a) (b) (c)

SIR Epidemics on Networks

Typical run of SIR on a graph representing a contact network



Big questions in epidemiology: how many will an epidemic infect?
How will the spread change with changes in parameters?
Based on that, what are best defences?

SIR Epidemic Extensions

Many extensions to accommodate different parameters

Some contacts more likely than others:

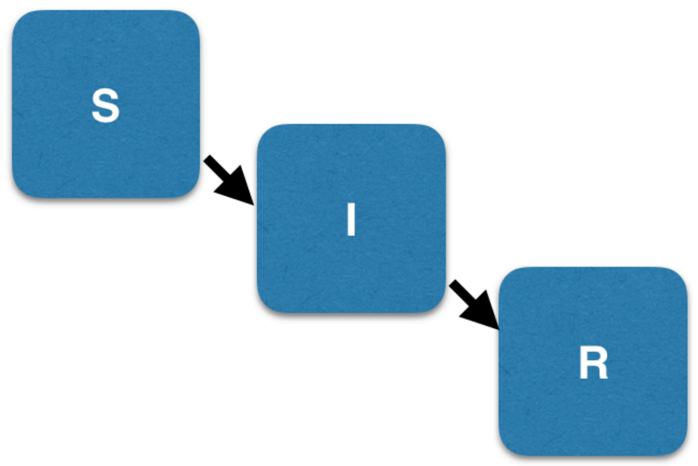
→ probability $\mathbf{p_{uv}}$ that is pair-dependent

Disease goes through different stages (infectious incubation, then less infectious symptomatic transmission):

→ **SEIR or S"III"R:** either **E**xposed state or several different infectious states (with different **p**'s or **t**'s)

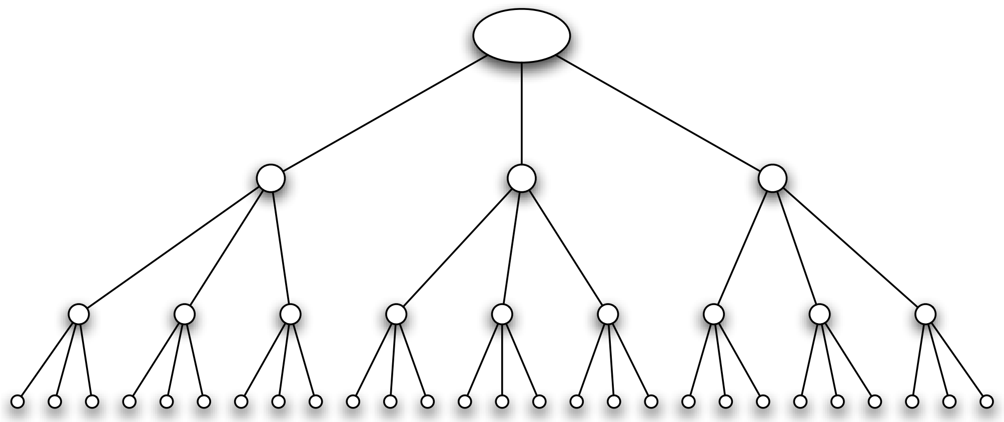
SIS: later in the lecture

Mutations (infectiousness, breaking immunity, etc)



From trees to networks

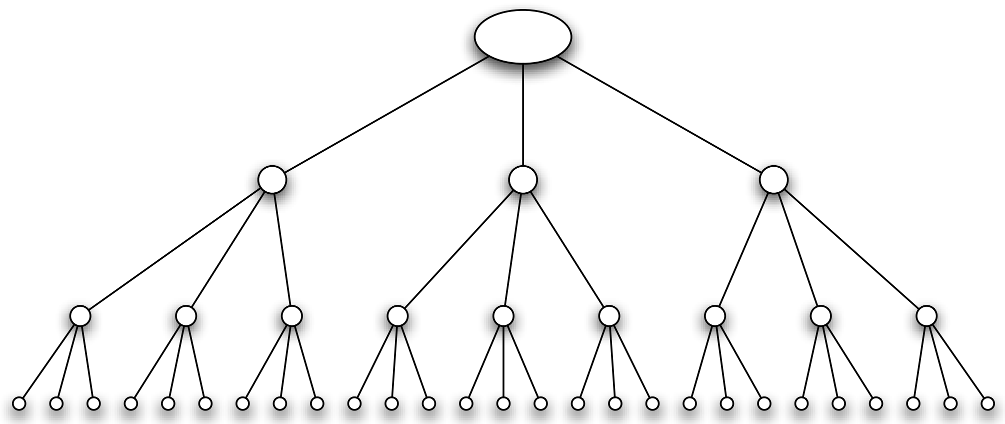
Recall that analysis of R_0 was for trees:



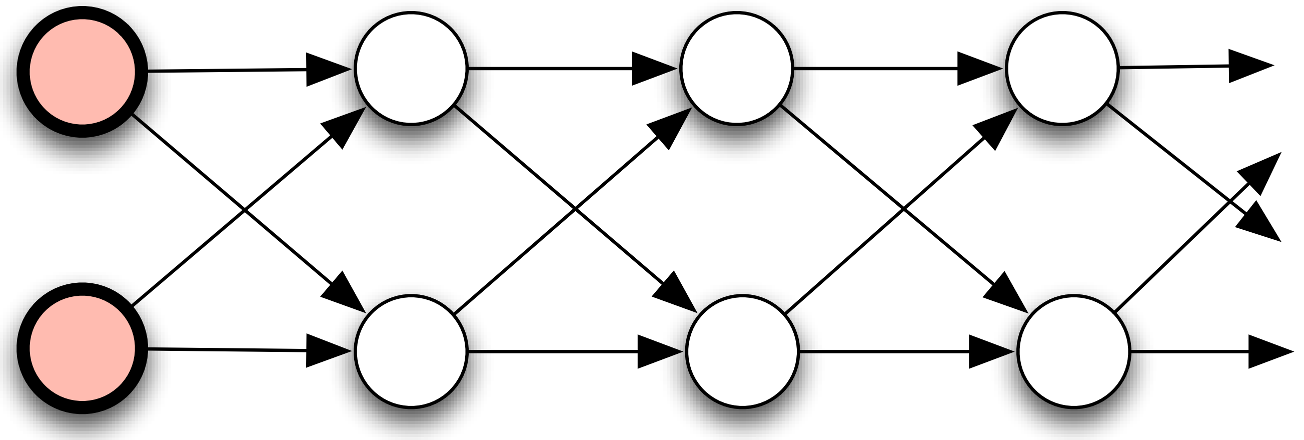
Do we have the same knife-edge R0 ~ 1 result in general graphs?

From trees to networks

Recall that analysis of R_0 was for trees:

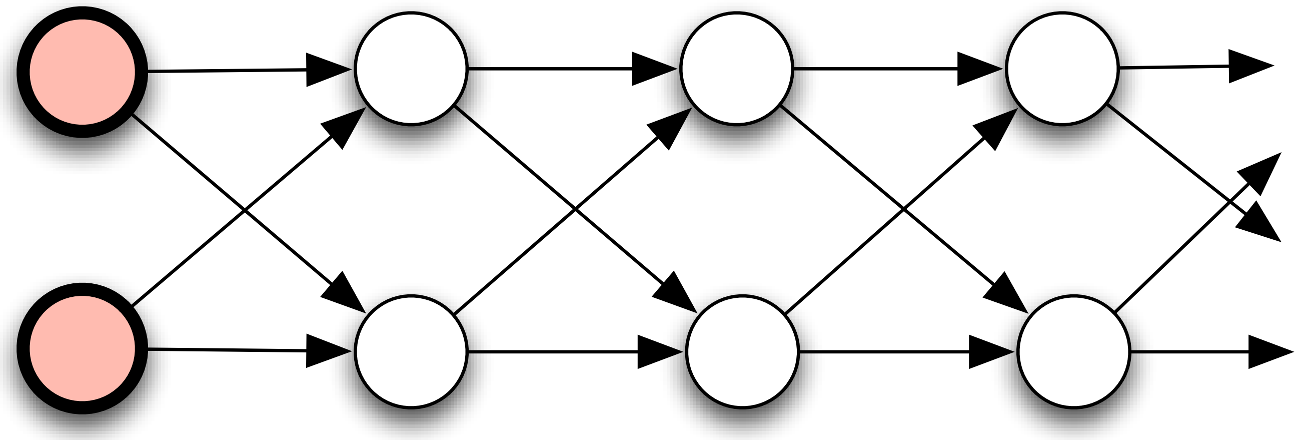


What happens on other networks? Consider **p=2/3, k=2.**



From trees to networks

What happens on other networks? Consider **p=2/3, k=2.**



Calculate R_0 as number of expected new cases per node

R0 = (2/3)*2 = 4/3 > 1

But this will almost certainly die out: $(1/3)^4 = 1/81$ chance that all four edges fail even if both nodes are infected

Prob that this happens after finite number of steps converges to 1

Now: SIS Epidemic Model

S = Susceptible

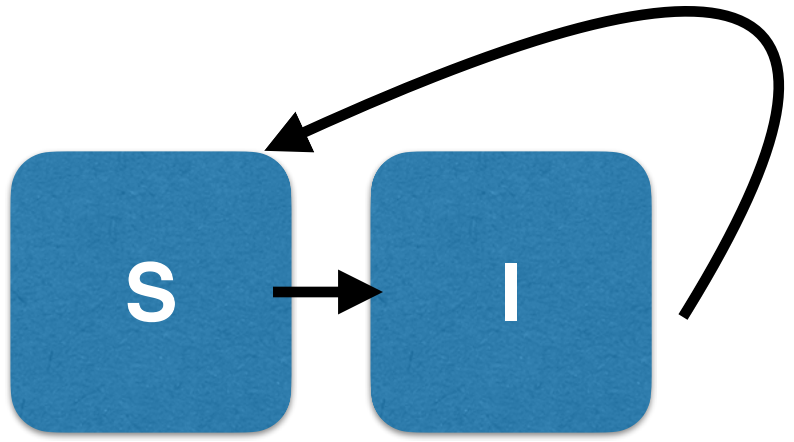
I = Infectious: node is infected and infects with prob **p**

Initially some nodes in **I** state, rest in **S** state.

Each node in I state remains infected for $\mathbf{t_I}$ time steps

During each step, each node has probability **p** of infecting all neighbors

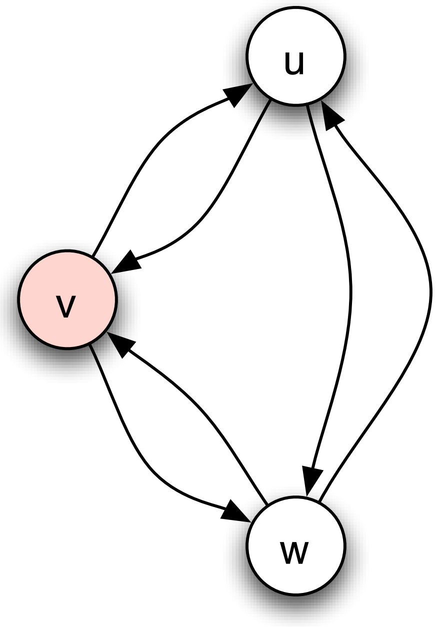
After $\mathbf{t_I}$ time steps, node **returns to S**



SIS Epidemic Example

$\mathbf{p} = 1/2$

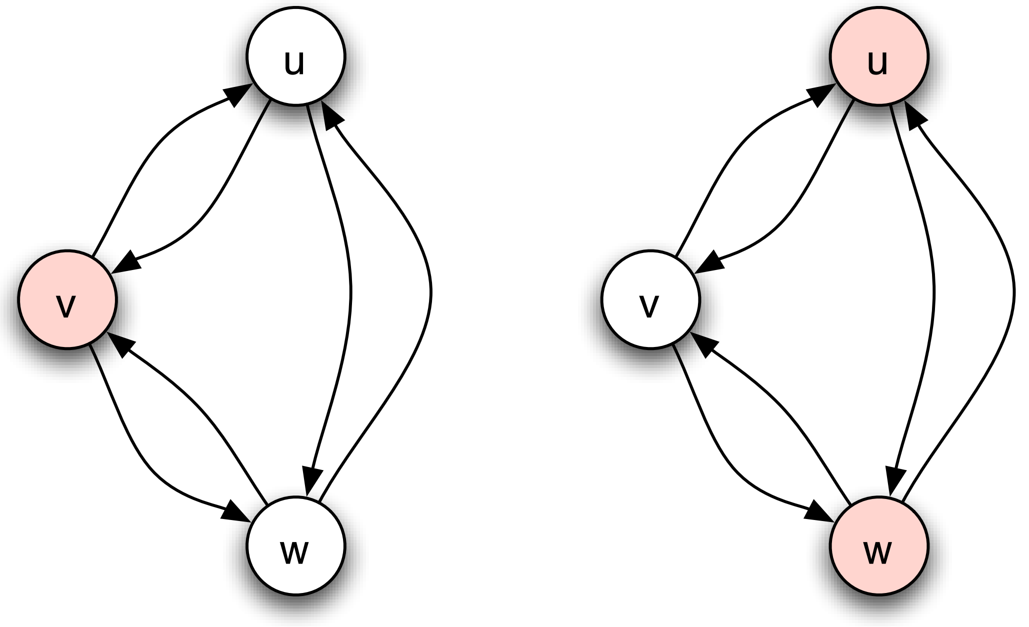
$\mathbf{t_I} = 1$



SIS Epidemic Example

p=1/2

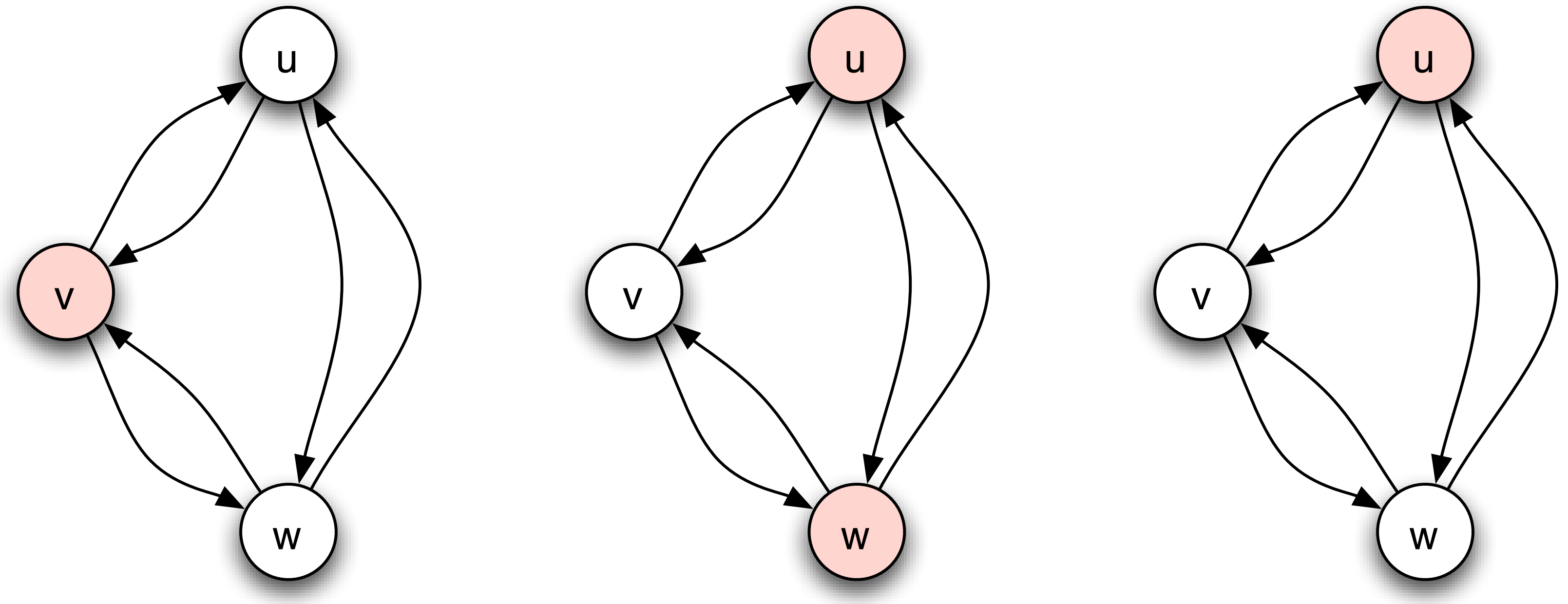
$\mathbf{t_I}$ =1



SIS Epidemic Example

p= 1/2

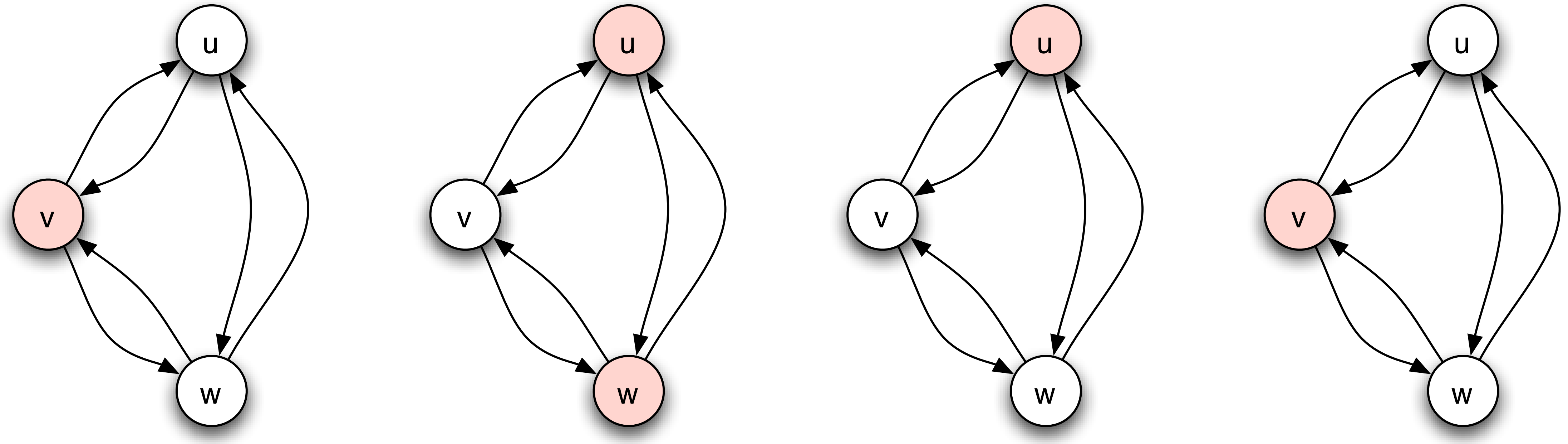
$\mathbf{t_I}$ = 1



SIS Epidemic Example

p= 1/2

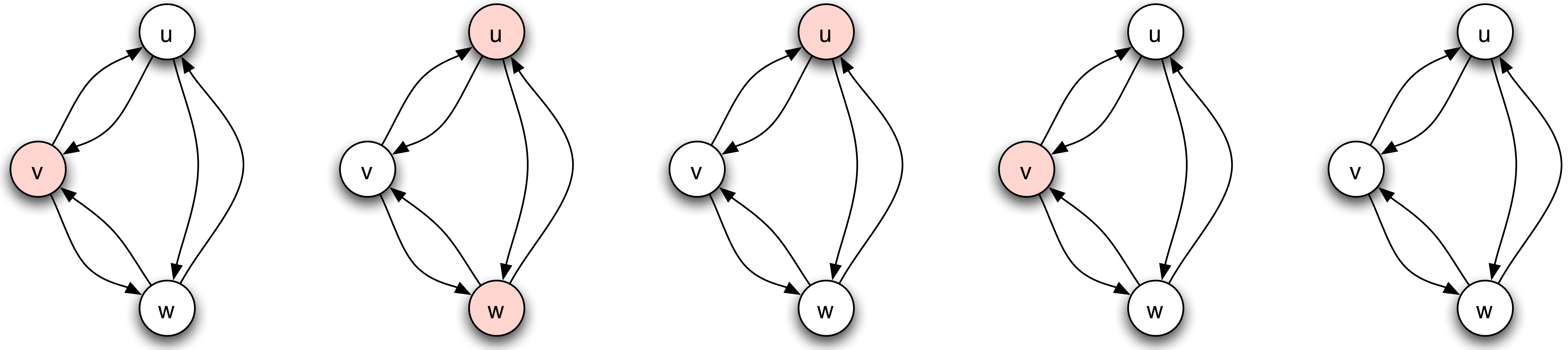
$\mathbf{t_I}$ = 1



SIS Epidemic Example

p= 1/2

$\mathbf{t_I}$ = 1

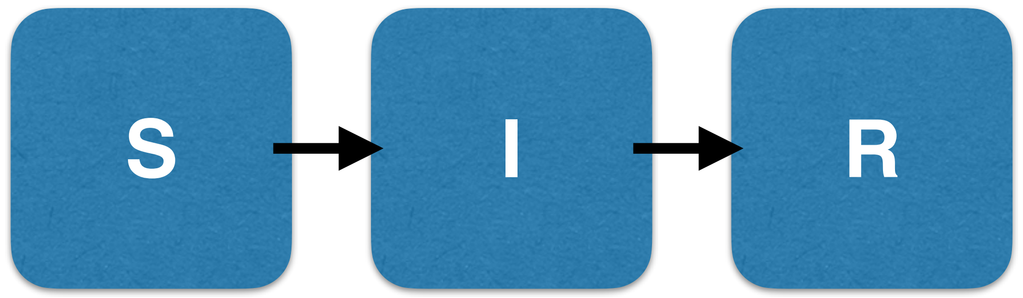


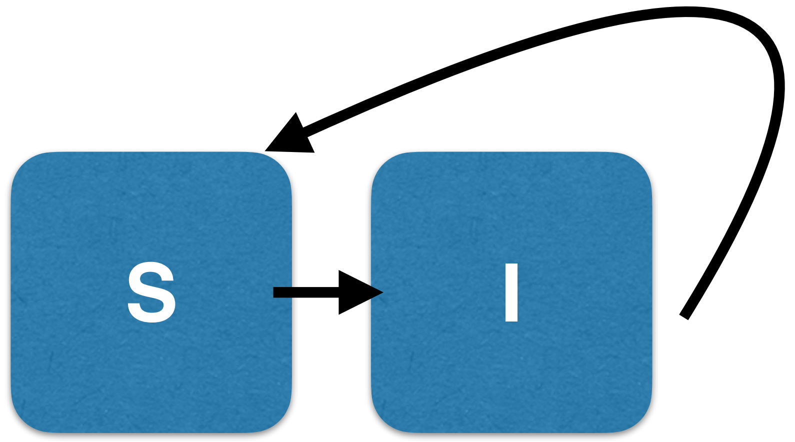
SIR vs. SIS

SIR: "burning through" a finite supply of susceptible

SIS: can run for a very long time, cycling through targets

SIS, like SIR, has a **critical threshold** ("knife-edge"/"tipping point"); trickier mathematical analysis. On non-trees both depend on more than just R_0.

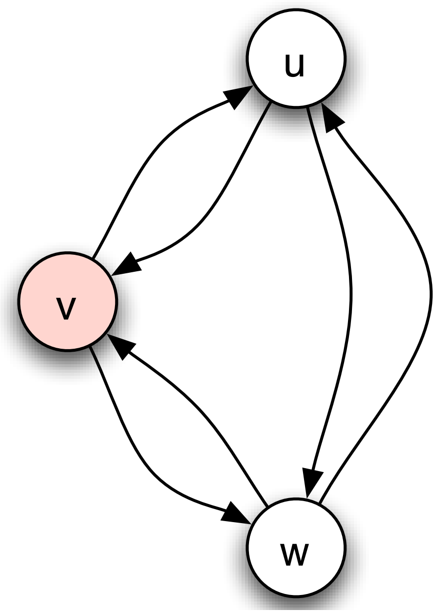




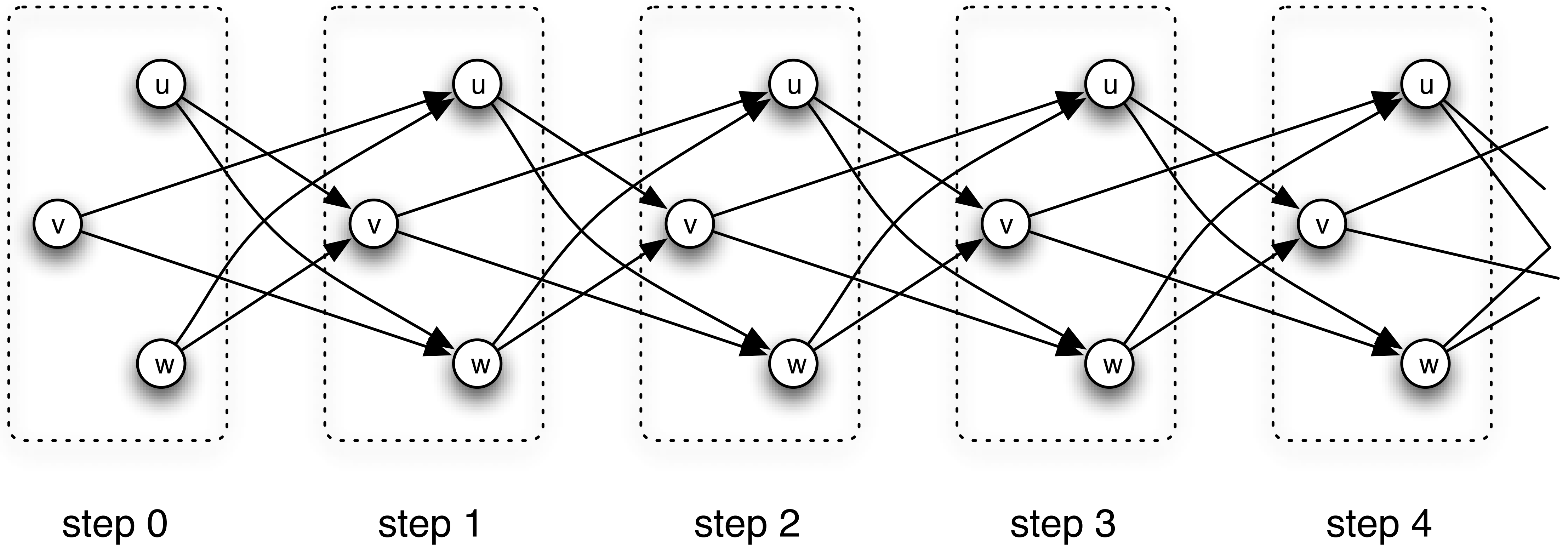
SIS as SIR on a bigger network

Consider **time-expanded** network: if u connects to v in network, have u_t connect to v_{t+1}

SIS is SIR on a time-expanded network.



(t_I = 1 here)

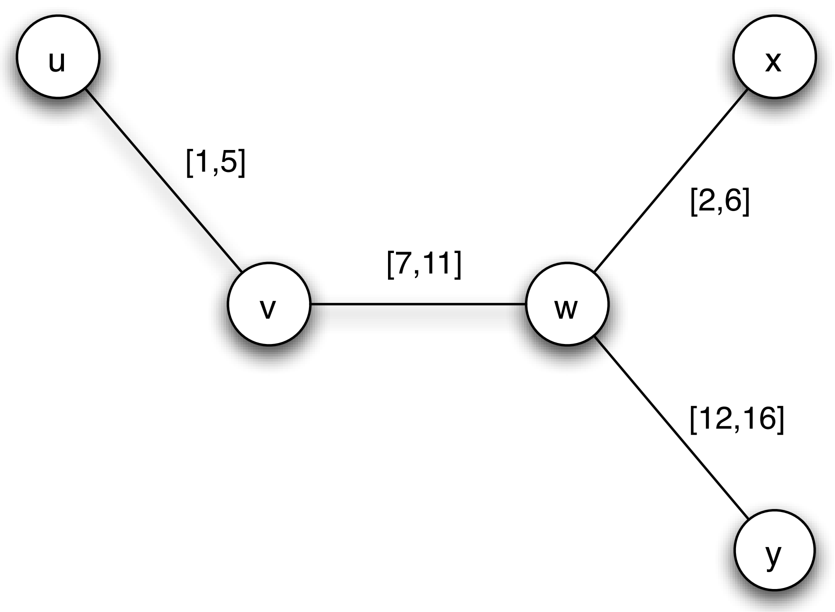


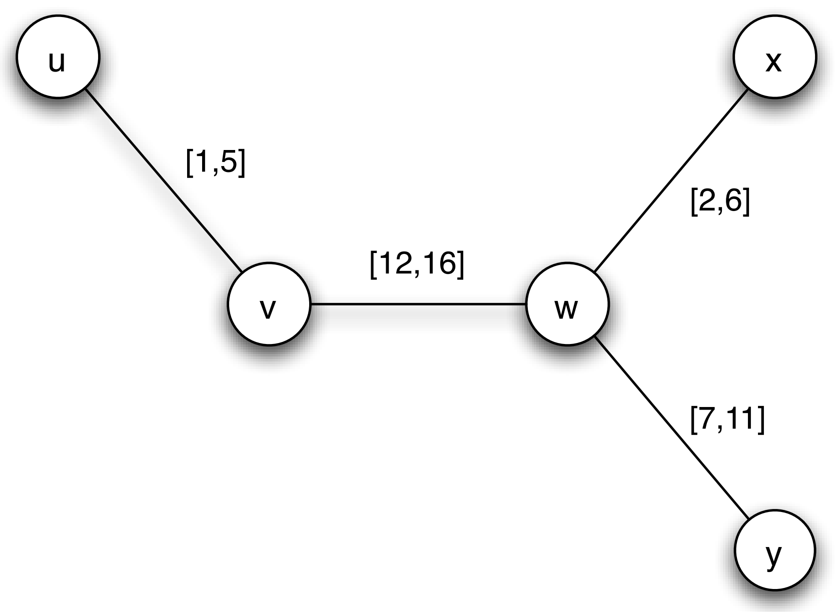
Transient Contacts & Concurrency

- So far, we've been analyzing **static** networks
- This is reasonable when the rate of transmission is **typically much faster** than edge creation/deletion
- But some epidemic diseases last for years (HIV)
- **When edges are active** becomes very important

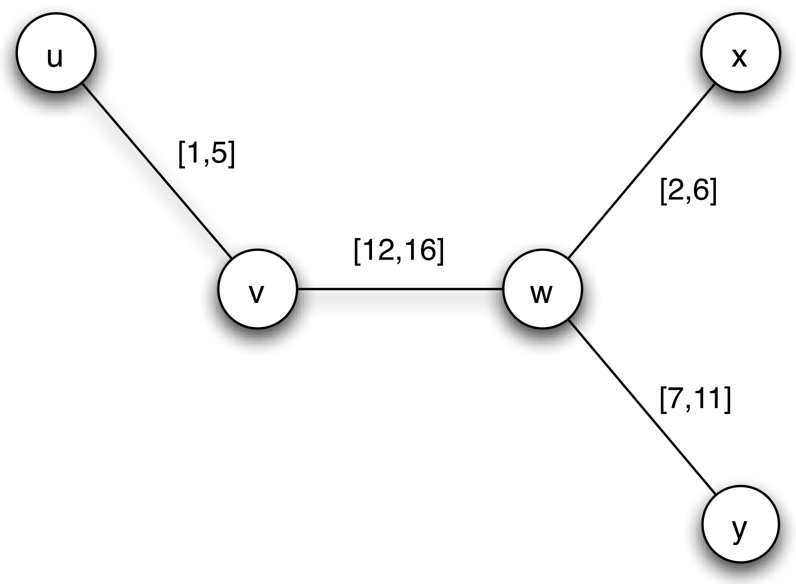
Transient Contacts & Concurrency

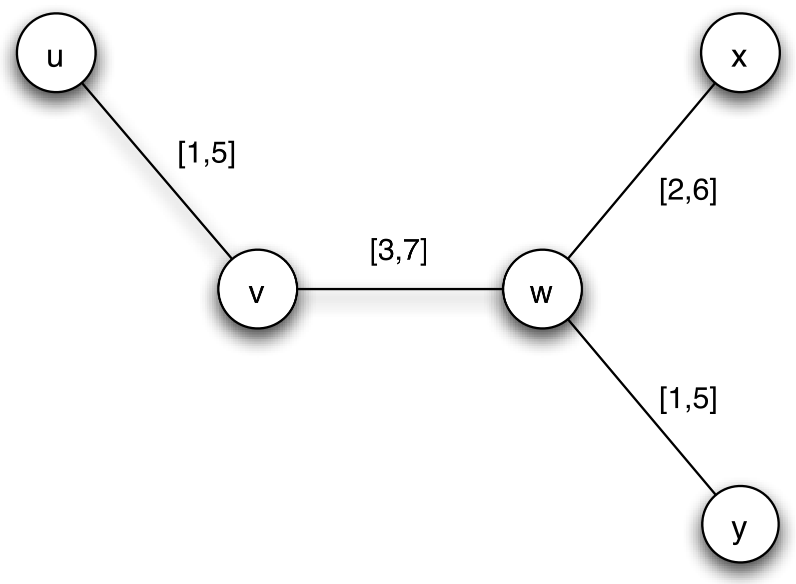
A less random model: it matters in what order contact is made in the contact network.





Concurrency: having two or more contacts at once.



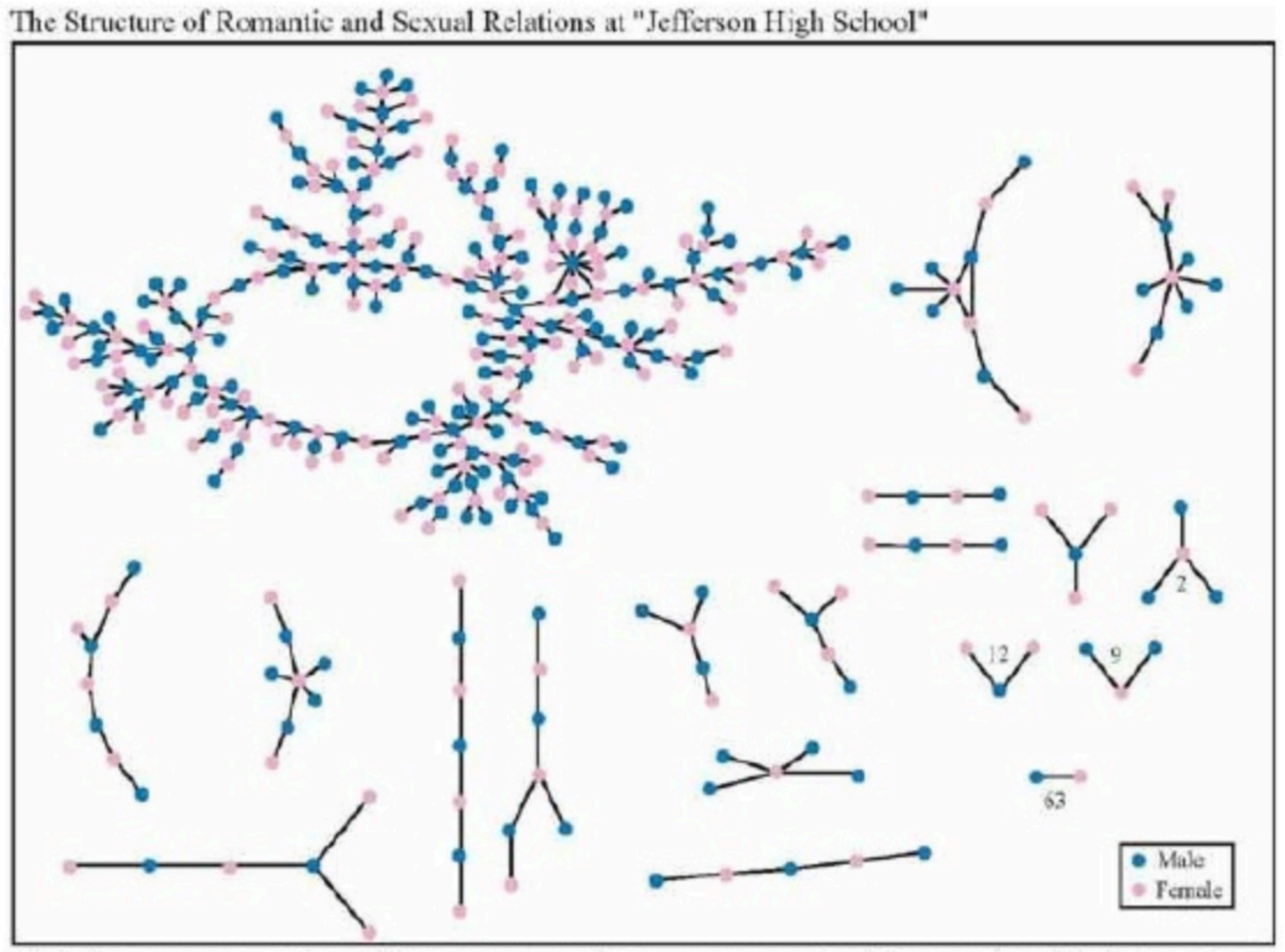


Transient Contacts & Concurrency

Small changes in times can produce **large differences in global epidemic spread**

There are **rich classes of network models** incorporating transience and concurrency

It's not enough to just know the structure



Oscillations

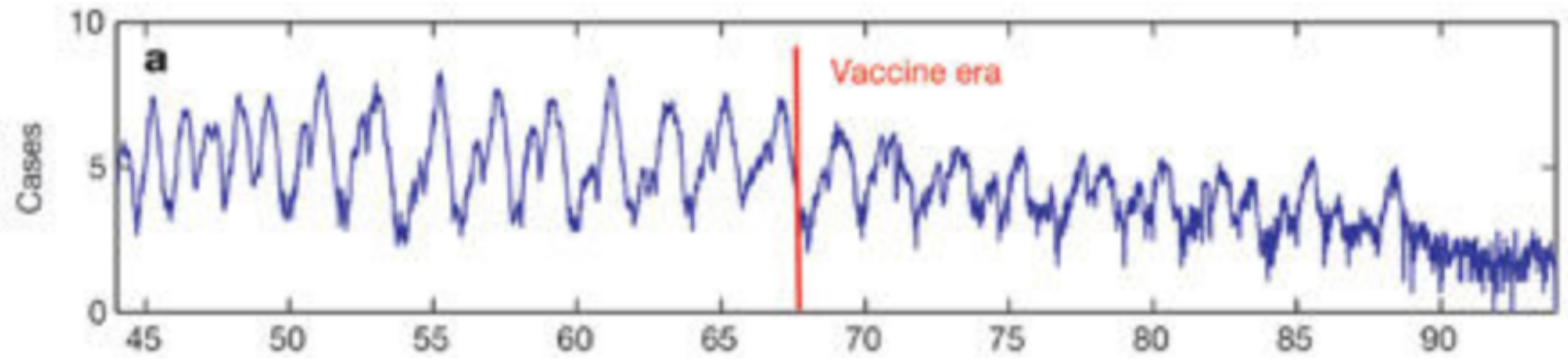
Diseases can be **cyclical** / have **oscillations** (like measles and syphilis)

To model this, vary the model so nodes have **temporary immunity**

SIRS: Susceptible, Infected for I steps, Recovered for R steps, then Susceptible again

This can **produce oscillations in very localized** parts of the network

But for large fluctuations at the global network level, need **small-world structure** (random long-range contacts)

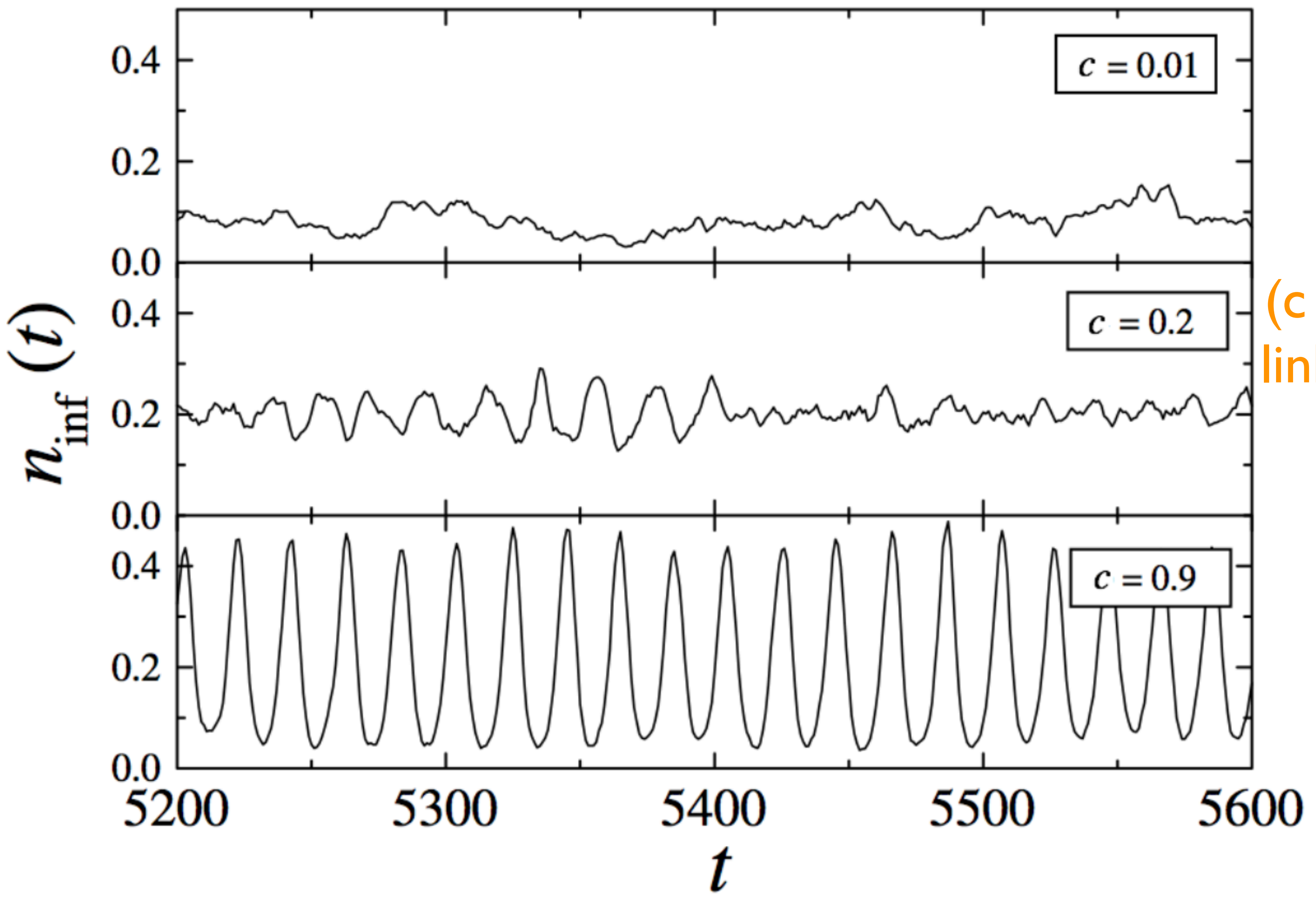


Measles

Oscillations

Diseases can be **cyclical** / have **oscillations** (like the flu)

But for large fluctuations at the global network level, need **small-world structure** (random long-range contacts)



(c = long-range link probability)

Epidemics vs. Behaviour

In epidemic models, nodes get infected from one particular other node

To model information spread, people often use epidemic models ("viral diffusion”)

But many social phenomena (behaviours, beliefs, practices, etc.) are *complex*: costly, risky, uncertain, etc.

Epidemics vs. Behaviour

When a behaviour is risky, costly, or uncertain, you may not do it just because one of your friends is (but this is what epidemic diffusion looks like)

Social movements, health technologies, political activism, etc.

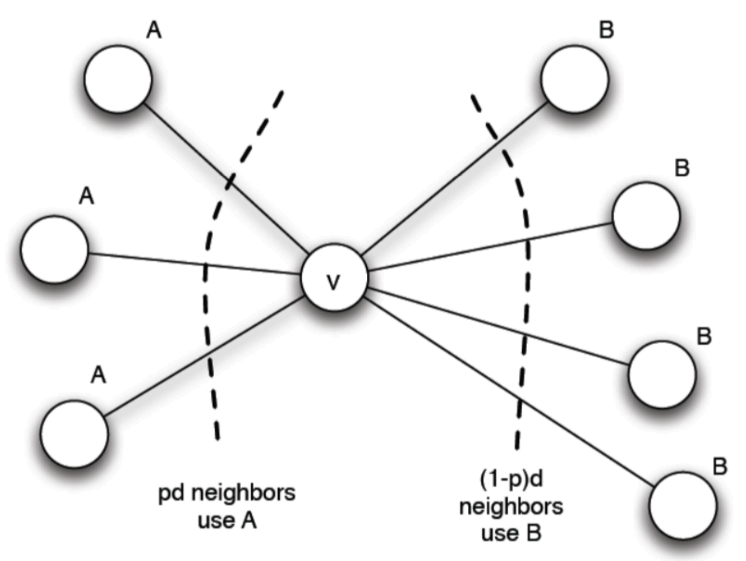
E.g.: PrEP medication is the best latest in HIV prevention → one pill a day gives 90% prevention.

But in two trials in sub-Saharan Africa, it didn't work ... because no one was taking it! (fears of discrimination, etc.)

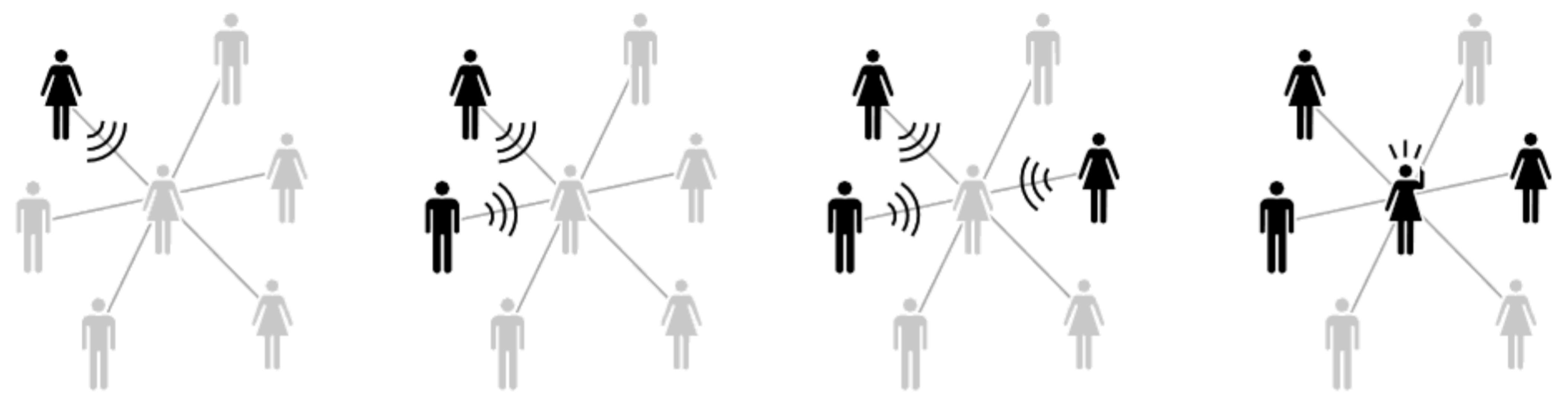
How do you get behaviour to diffuse?

Epidemics vs. Behaviour

Previously we saw a model of behaviour diffusion based on utility



This is an example of *complex* diffusion: in general, need more than one neighbour to adopt before you adopt a behaviour.

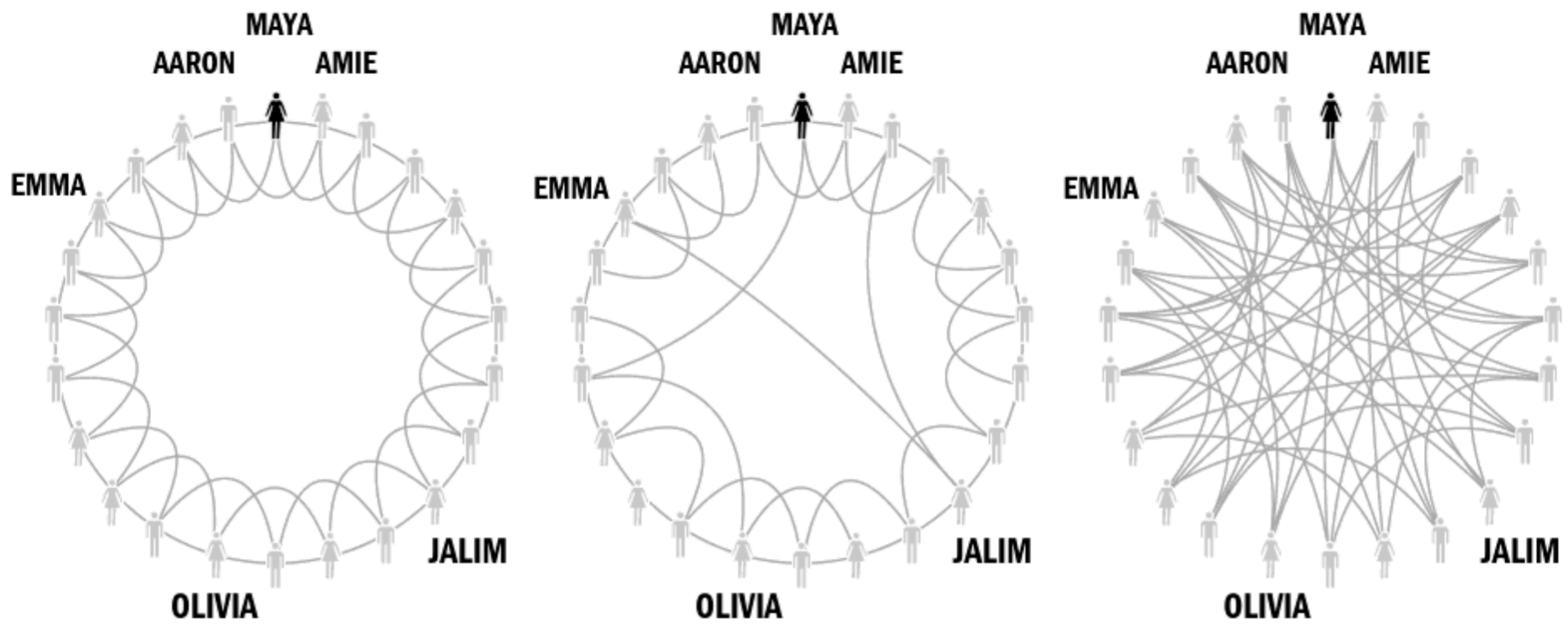


Epidemics vs. Behaviour

Simple vs. complex diffusion
Epidemics vs. behaviour

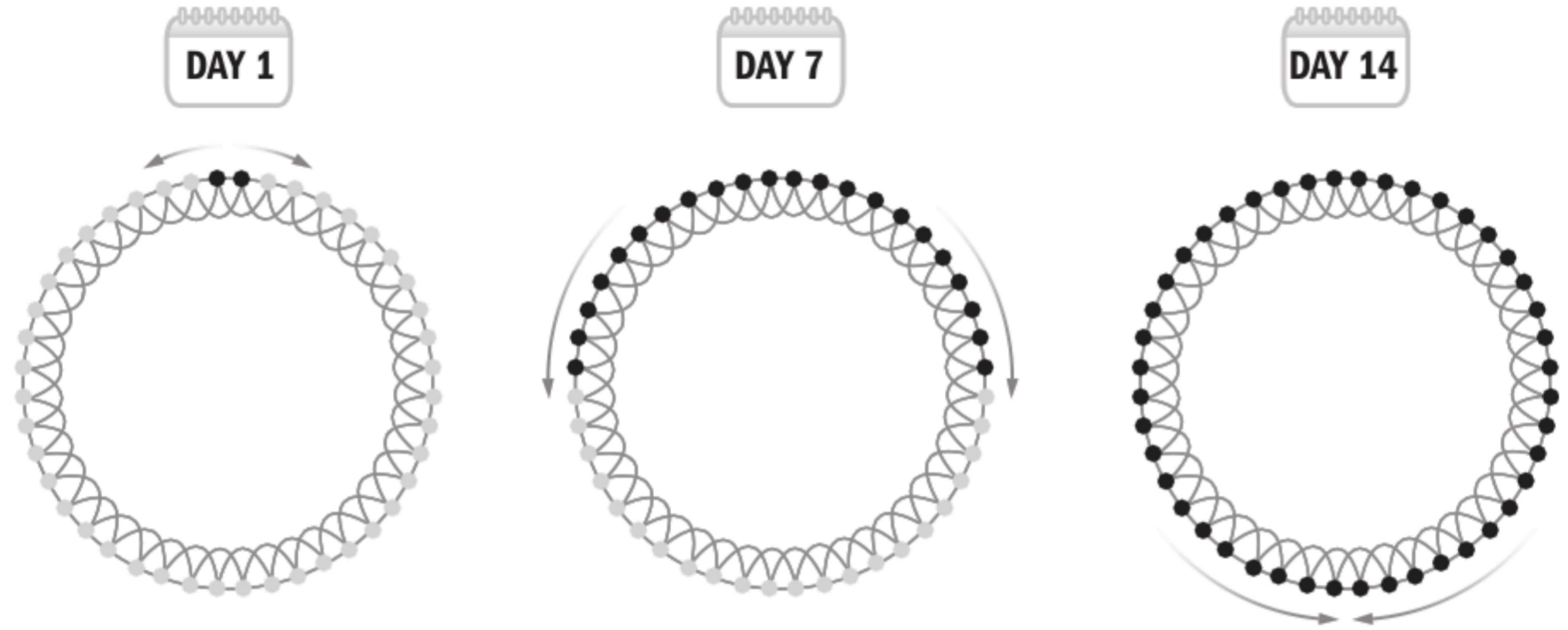
What's the difference?

Recall the small-world model

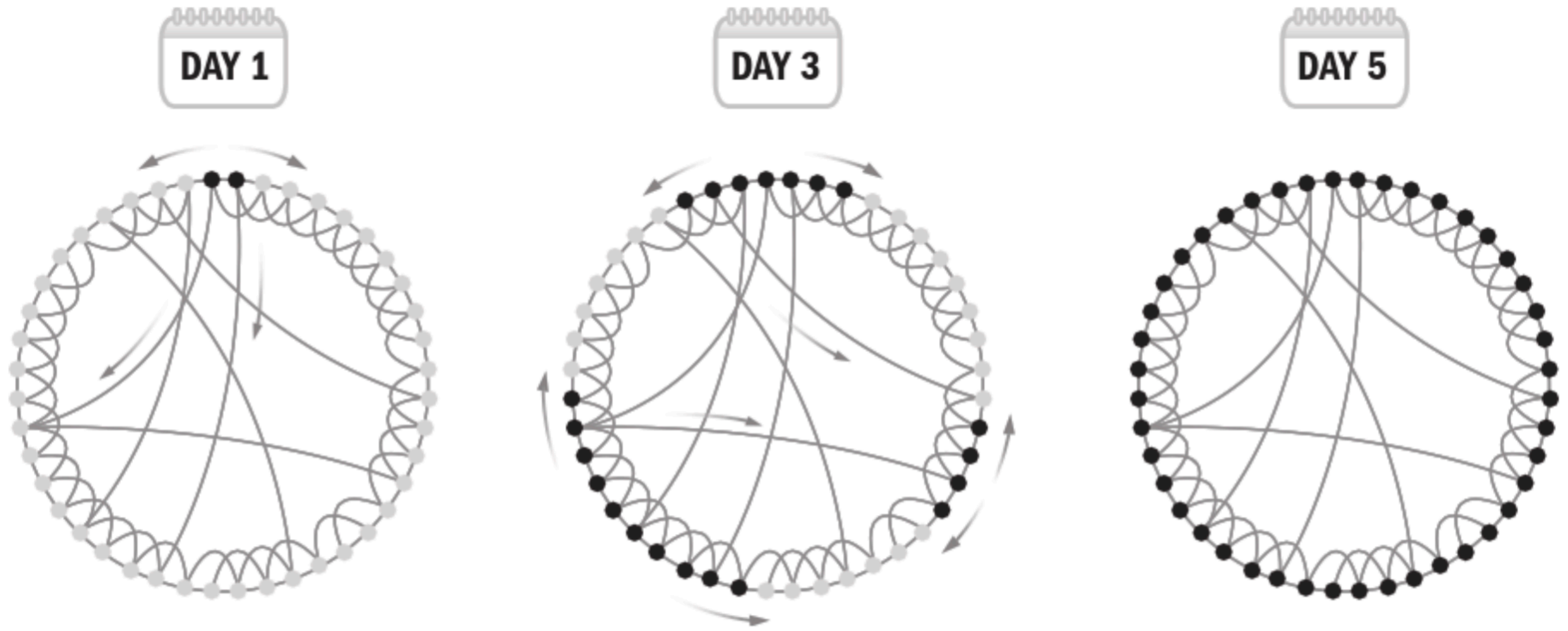


Simple Diffusion

Large world:

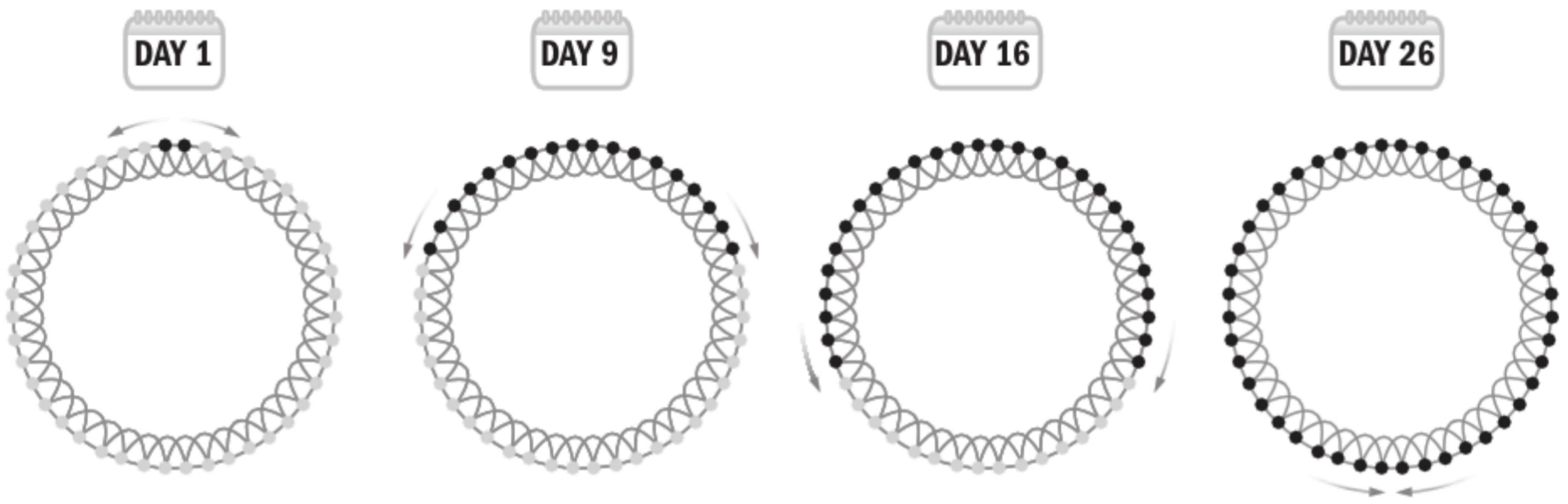


Small world:

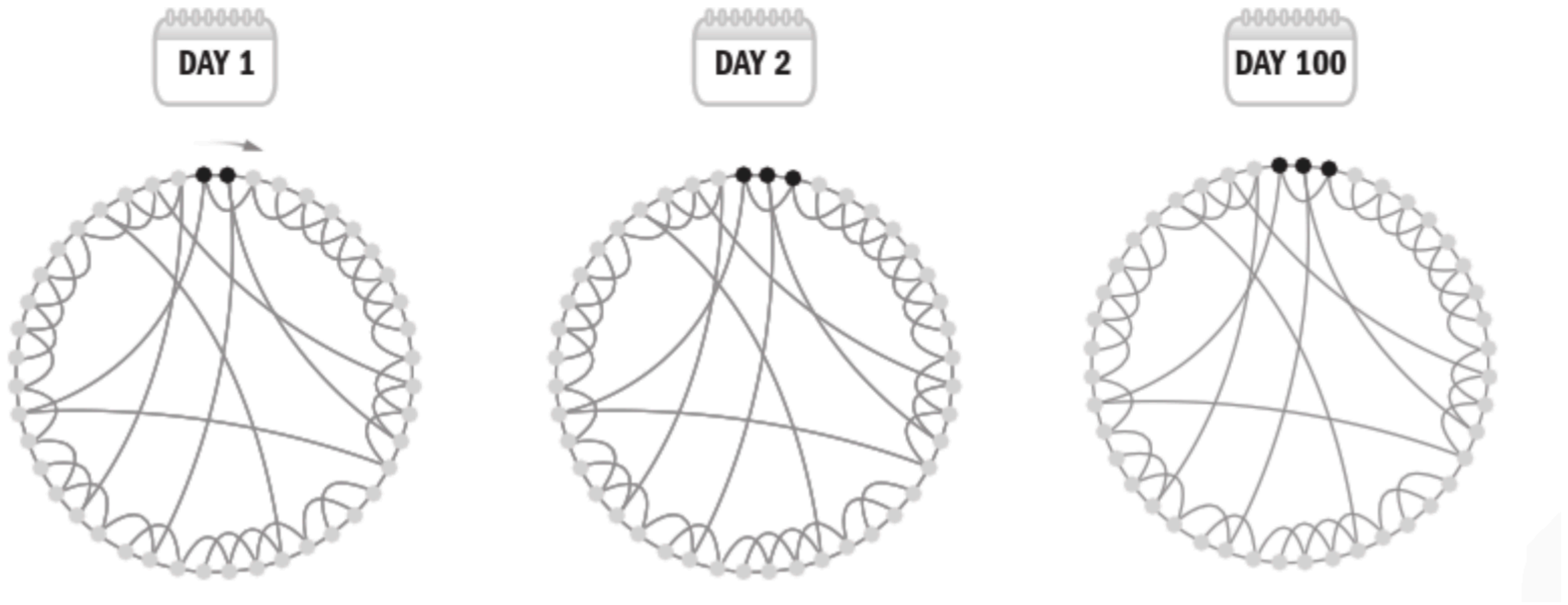


Complex Diffusion

Large world:



Small world:



Simple vs. Complex Diffusion

Weak ties are extremely useful for simple diffusion and contagion, but they inhibit complex diffusion!