Epidemic modelling using networks

Simon Dobson

School of Computer Science
University of St Andrews
Scotland UK

simon.dobson@st-andrews.ac.uk
https://simondobson.org

University of St Andrews | FOUNDED 1413

Introduction

We're experiencing one of the biggest socio-economic disruptions of all time

- How do large epidemics behave?
- What changes might we need to make to control and accommodate to the disease as it changes?

This talk

- Look at *one approach* to modelling, using network science

Why am I here?

I'm a computer scientist interested in network science, sensing, and data analytics

- A tool builder: how do we make computers useful for answering questions?
- The questions themselves are less important…

The uses of computers in generating *insight*

- To simulate *particular events* in detail
- To explore the space of *possible events* to suggest options
- To understand the general computational and mathematical *processes* involved

Acknowledgements

Collaborators (some of whom did most of the work)

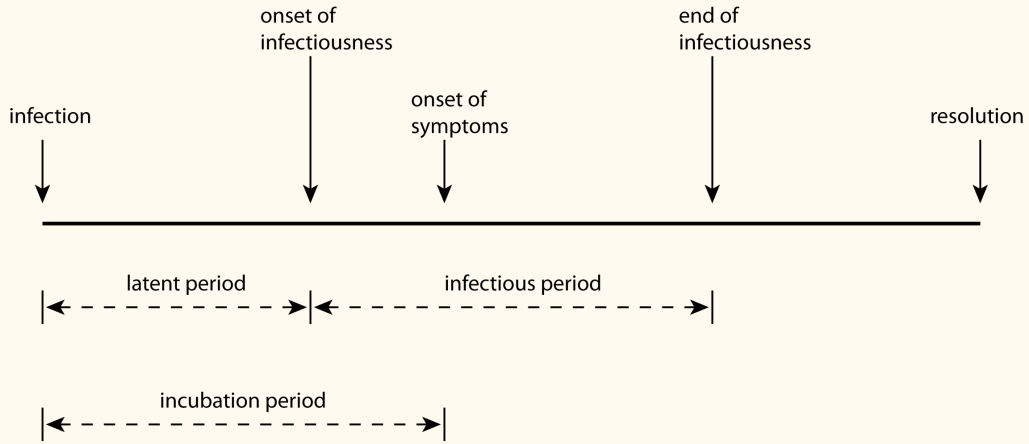
- Peter Mann, Saray Shai, Mike Pitcher, Davide Cellai
- V. Anne Smith, John Mitchell, Juan Ye, Lei Fang
- Libby Askew, Leo Pfeiffer, Martynas Noreika

Funding from UK EPSRC

EPSRC
Engineering and Physical Sciences Research Council

Structure of this talk

Real diseases – general structure

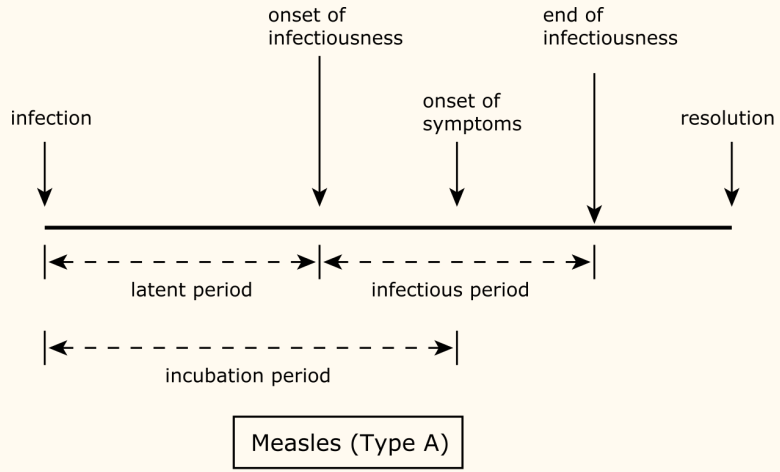


Different periods

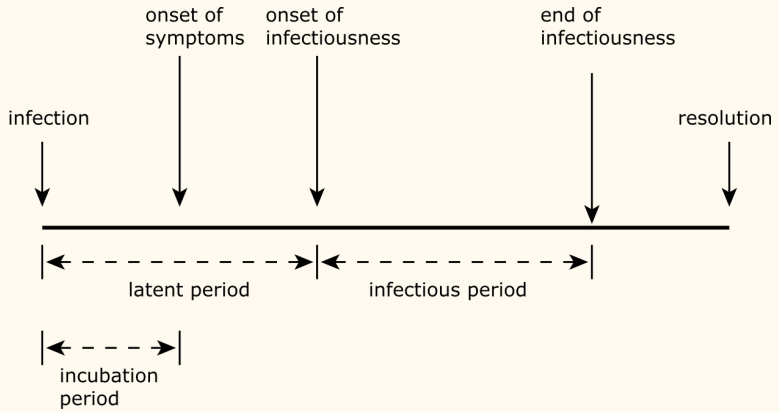
- *Incubation*: from infection to onset of symptoms
- *Latent*: from exposure to infectiousness
- *Infectious*: overlapping with symptoms (usually)

Periods defined by biology, of both disease and host

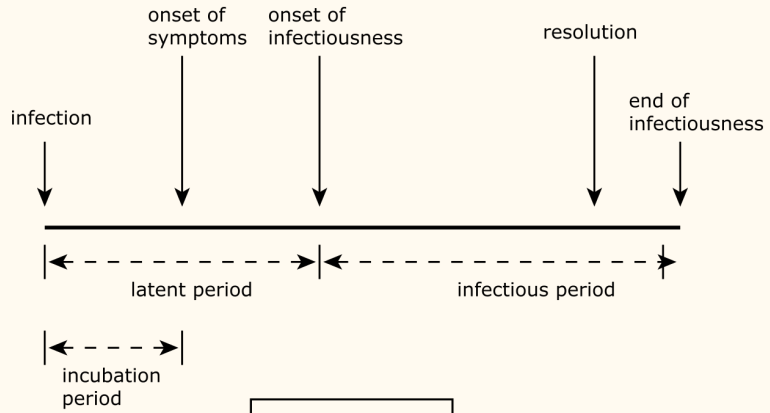
Real diseases – examples



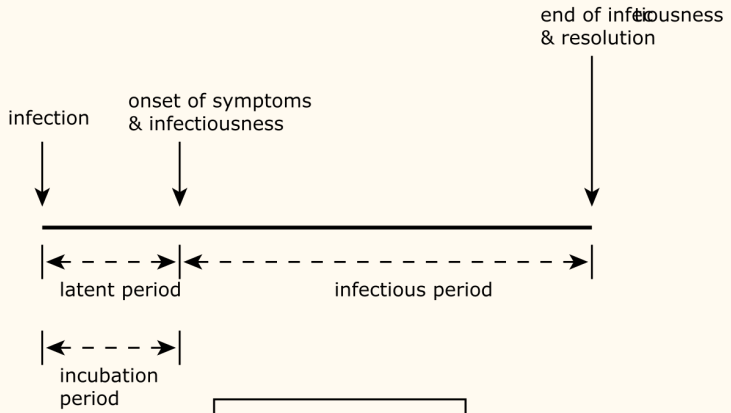
Measles (Type A)



SARS (Type B)



Ebola (Type B)



Zombie apocalypse

Real diseases – spread

Disease is spread by the exchange of a pathogen

- From infected to non-infected individuals

Different infection patterns

- How many other people does each person meet each day?
- How closely do they interact? For how long? In what way? For how long?

Infections defined by biology *and* environment

Real diseases – evolution

A person infected at the *end* of an epidemic doesn't get the same disease as a person infected at the *start*

- Pathogen is constantly mutating
- Lateral gene transfer from co-infecting pathogens
- Another reason to work to reduce transmission

Selection pressures often (but don't necessarily) introduce a particular dynamics

- More transmissible
- Less severe

$\mathcal{R}$ and all that [1]

$\mathcal{R}$, the case reproduction number

- Number of secondary cases per primary
- The exponent of an exponential growth process
- Especially $\mathcal{R}_0$, reproduction absent countermeasures

Typically averages over (unknown) distributions

- Details may be very significant
- For example may see "superspreaders" creating lots more infections

[1] Royal Society SET-C group. Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK, August 2020. URL https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf

The "wickedness" of covid-19

For "wild type" $\mathcal{R}_0 \approx 3$, not particularly infectious

- More infectious, less severe (maybe) variants emerge
- Some tendency towards vaccine escape
- Prior infection doesn't give clear-cut, long-term immunity

Substantial asymptomatic transmission

- Asymmetric costs (spreading *vs* dying, "long covid")
- Effective countermeasures are collective (and expensive)

Infection fatality rate is about 1%

- Too large to comfortably ignore
- …but too small to generate a consensus on seriousness

The goals of modelling

What are we trying to find out?

- Concrete: how will this *particular* outbreak behave, in this *particular* population?
- Abstract: how can diseases behave *in general*? Are there common mathematical structures?

Compartmented models

Traditional epidemic modelling uses the framework of a *compartmented model* of a disease

- A number of "compartments" that hold some fraction of the population
- Can also think of a compartment as the state of each individual within the population (we'll come back to this)
- Rules on how these fractions change over time

Continuous SIR

The model

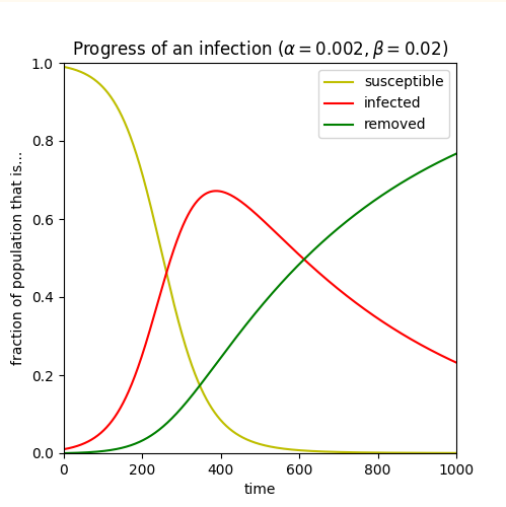
- **S**usceptible individuals can catch the infection from **I**nfected individuals
- …who then are **R**emoved from the dynamics by recovery (or death)

Epidemic dynamics

- Susceptibles infected per contact with probability β
- Infecteds removed with probability α
- Gives rise to $\mathcal{R}_0 = \frac{\beta}{\alpha}$

$$\frac{dS}{dt} = -\beta SI \qquad \frac{dI}{dt} = \beta SI - \alpha I \qquad \frac{dR}{dt} = \alpha I$$

Solution



Different disease structures [2]

- SIR – complete immunity post-infection
- SIS – infection confers no immunity
- SEIR – exposed individuals are infectious before symptoms
- MSEIR – initial immunity passed from mother to child
- SEIRS – immunity wears off with time
- …

[2] H. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, December 2000. URL doi://10.1137/S0036144500371907

Structure of this talk

Network science

Networks (or graphs)

- Model objects and relationships in an abstract mathematical form
- Use as a substrate for processes that affect the states of objects and their relationships over time

Social networks

- Individuals and their social contacts
- May be real or synthetic

Network science for epidemic modelling

Use a network as the substrate for the epidemic [3]

- Only adjacent nodes can interact
- Compartment = label on node
- Infection passes over **SI** edges

Pros and cons

- ✘ Doesn't scale as well as the differential equations (we're modelling explicit individuals)
- ✓ Can build contact structures and systems of equations we can't solve (but can simulate)

[3] M. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66, July 2002. URL doi://10.1103/PhysRevE.66.016128

Basic treatment – networks

Start from a simple model of a population

- As a random process, collected from contact data, …
- Actually a lot we don't know about how people interact

Add fine structure

- Structured contact patterns
- More- and less-well-connected sub-populations
- Adaptive behaviour to change features over time and/or in response to the disease

Basic treatment – processes

Assign a state vector to each node

- For epidemics, this might be the node's compartment

Process defines changes to state vectors

- A function of current states of the node and its immediate neighbours
- Generally stochastic, applied with some probability

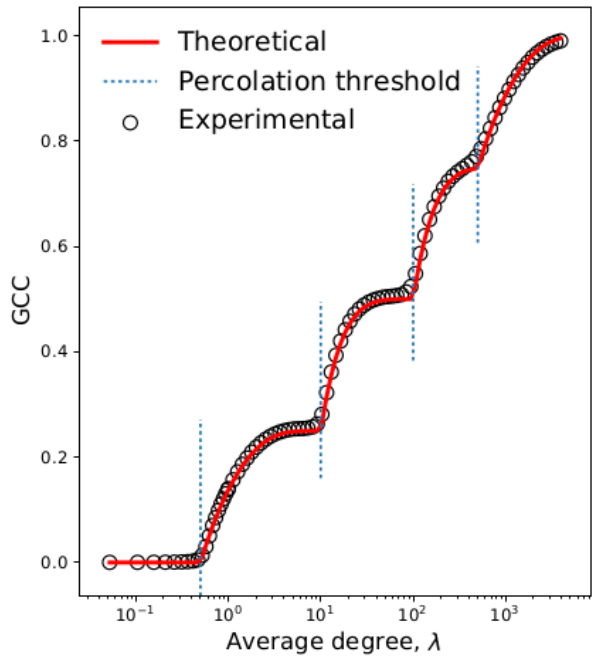
Seed the network with initial state vectors

- For SIR, mainly susceptible with a few infected

How to do analysis

The "gold standard" is an analytic model with numerical validation

- Find an analytic description for what happens under different infection parameters
- Run process on random networks with different topologies
- Lots of repetitions to squeeze out variance
- (Hopefully) sample points land on solutions to the equations [4]



[4] P. Mann, V. A. Smith, J. Mitchell, and S. Dobson. Random graphs with arbitrary clustering and their applications. *Physical Review E*, 103(1), January 2021a. URL https://doi.org/10.1103/PhysRevE.103.012309

Discrete-event simulation

A simulation consists of a large series of *events*

- An infected person infected a susceptible person
- An exposed person developed symptoms
- An infected person recovered

Events selected using Gillespie's algorithm [5]

- $P(\tau, e)\, d\tau$ the probability that an event e occurs in the next interval $(t + \tau, t + \tau + d\tau)$
- Draw a pair (τ, e) from this distribution

[5] D. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25): 2340—2361, 1977

Tooling

There wasn't any standard tooling, so we built some

A flexible way to express networks and processes

- `epydemic`, a simulation framework using `networkx`
- Reference epidemic (and other) processes
- Support for the main mathematical techniques, such as generating functions

A way to perform repeated, repeatable, experiments

- `epyc`, a computational experiment manager
- Experiment submission, parallelism, remote evaluation

Example code

```
import numpy
import pandas
from epyc import ClusterLab, HDF5LabNotebook, RepeatedExperiment
from epydemic import ERNetwork, SIR, StochasticDynamics

# notebook for results and lab with connection to compute cluster
nb = HDF5LabNotebook('test.h5', description='My experiments in networking')
lab = ClusterLab(profile='hogun', notebook=nb)

# set up the experimental parameters
lab[ERNetwork.N] = 10000
lab[ERNetwork.KMEAN] = 40
lab[SIR.P_INFECTED] = 0.001
lab[SIR.P_REMOVE] = 0.002
lab[SIR.P_INFECT] = numpy.linspace(0.00001, 0.0002, num=50)

# construct the experiment: a process and a class of networks
m = SIR()
g = ERNetwork()
e = StochasticDynamics(m, g)

# repeat runs across the parameter space
lab.runExperiment(RepeatedExperiment(e, 100))

# retrieve for analysis
df = nb.current().dataframe(only_successful=True)
```

Structure of this talk

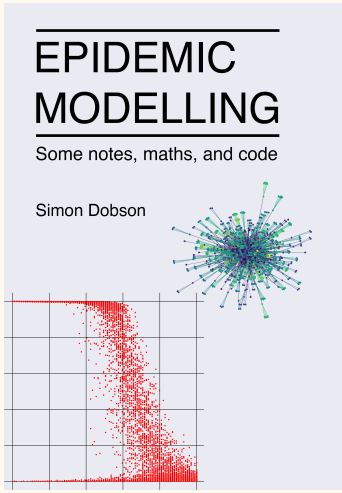
Explorations

We've been experimenting with different network structures

- Especially interested in "clustered" networks: friends-of-friends and larger cycles
- Fine structure affects how processes evolve

Make the science more accessible [6]

- With available and re-usable code
- With explanations

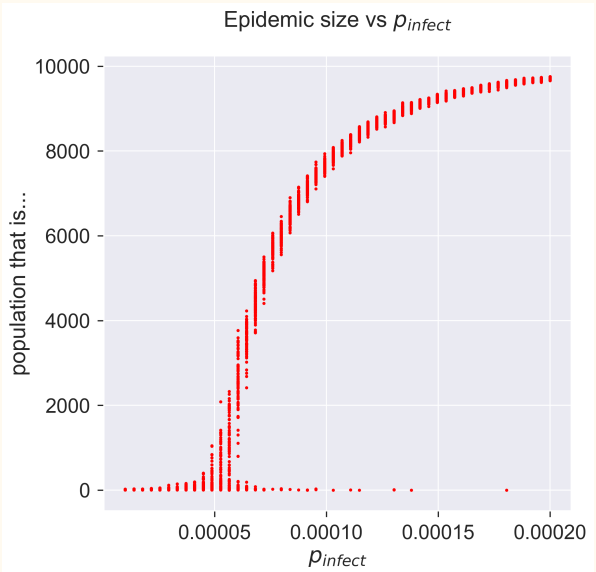


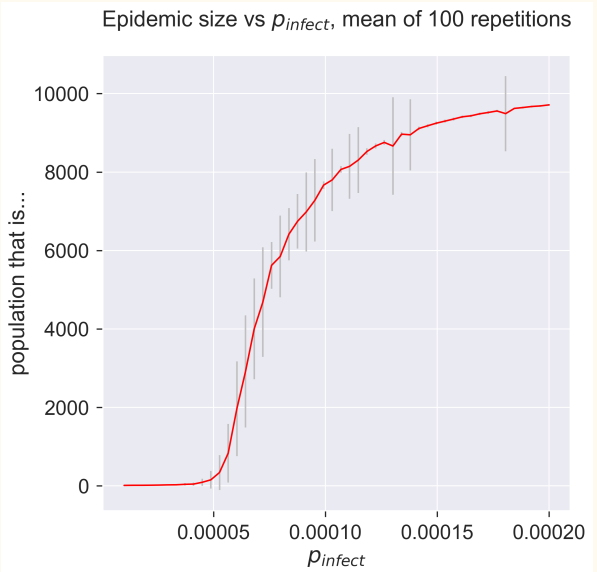
[6] S. Dobson. *Epidemic modelling – Some notes, maths, and code*. Independent Publishing Network, 2020. ISBN 978-183853-565-0. URL https://simoninireland.github.io/introduction-to-epidemics/

The epidemic threshold

Erdős-Rényi (ER) networks

- For N nodes build the complete network K_N
- For each edge, retain ("occupy") it with probability p_{infect}
- Leads to p_k normally distributed around $\langle k \rangle = p_{infect} N$

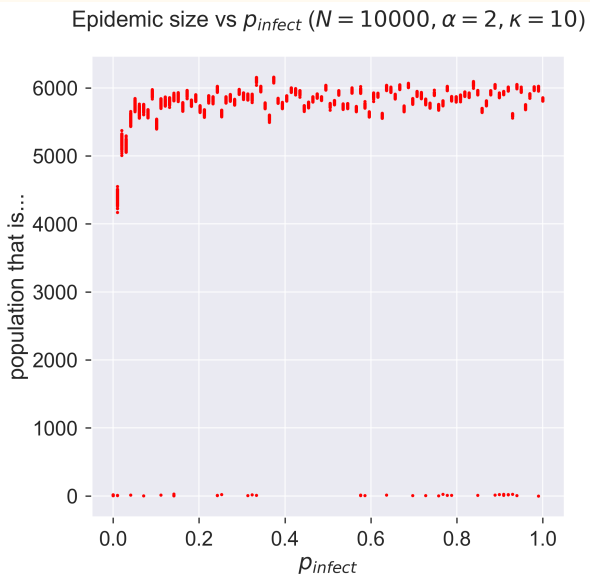


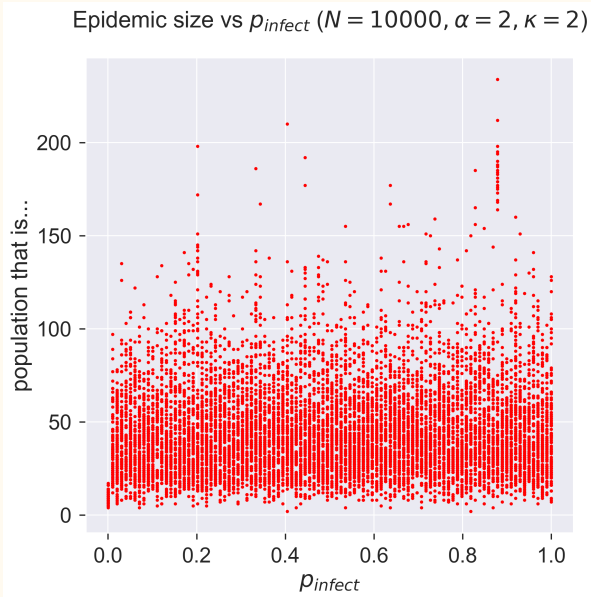


Not all networks behave like this

Too "even" to be a good model of human contacts

- Powerlaw with cutoff, $p_k \propto k^{-\alpha}\, e^{K/\kappa}$



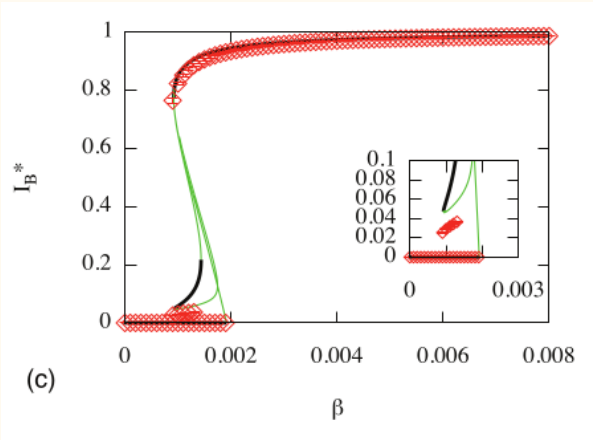


- Relatively insensitive to p_{infect}, but sensitive to α and κ

Adaptive networks

Things can become even more complicated when the network responds to the disease [7]

- For example quarantine
- Social contacts with infected people are reduced



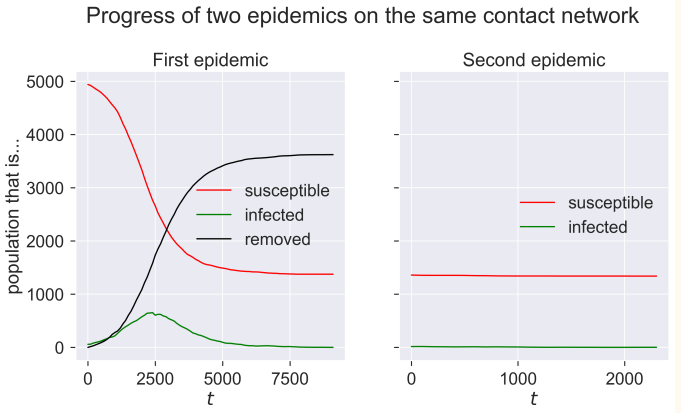
- Rewiring can balance (and even reverse) infection

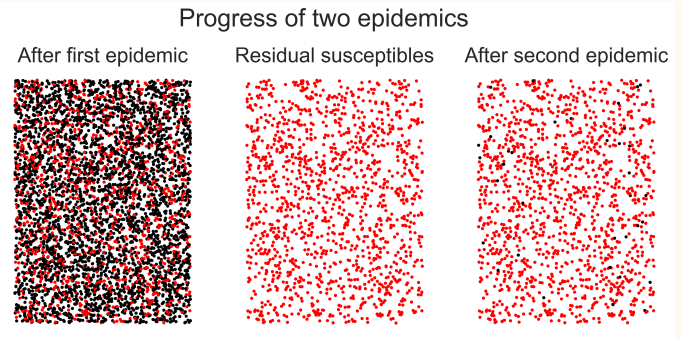
[7] S. Shai and S. Dobson. Coupled adaptive complex networks. *Physical Review E*, 87(4), April 2013. URL https://dx.doi.org//10.1103/PhysRevE.87.042812

Herd immunity

Sufficient immune/recovered individuals to stop an epidemic propagating

- Infecteds never adjacent to enough susceptibles
- First epidemic changes the effective topology
- "Effective" $\langle k \rangle$ drops from 20 to 5.5





Why pursuing herd immunity is a bad idea

Herd immunity was advocated by some as a COVID-19 strategy[8]

Ignores some rather inconvenient facts

- A 1% death rate = 700K UK deaths, about one year's excess
- At a rate that would collapse health services
- Immunity looks not to be permanent – which changes how herd immunity behaves (is it appears at all)
- Long COVID not accounted for in the costs

[8] See the "Great Barrington Declaration", https://gbdeclaration.org

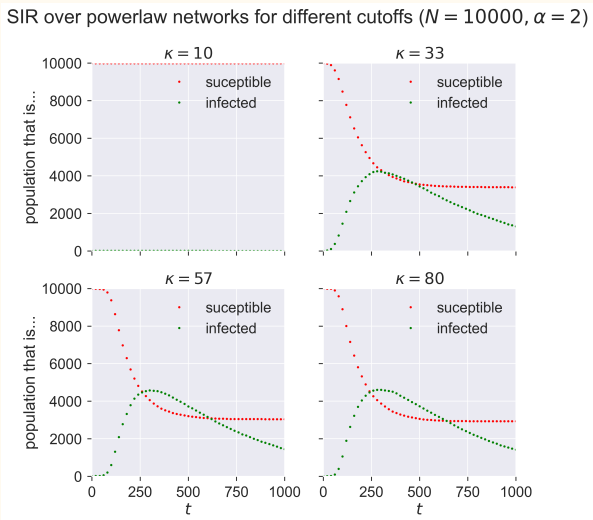
Vaccination

"Herd immunity without the bad bits"

- Aim for the herd immunity threshold, generally about 60% of the population
- …without anyone actually being infected

Epidemic proceeds at different rates depending on topology

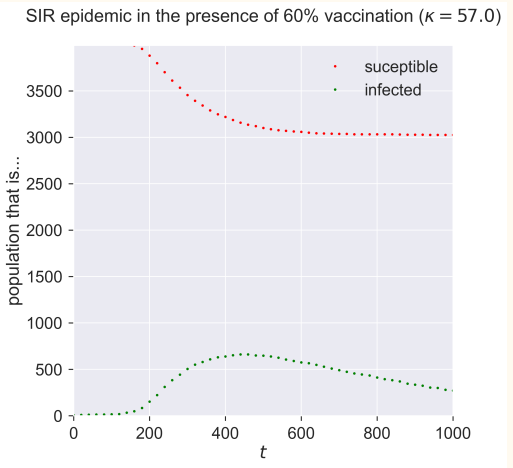
- "Enough" contacts removed to stabilise the size of outbreak



Vaccination strategies

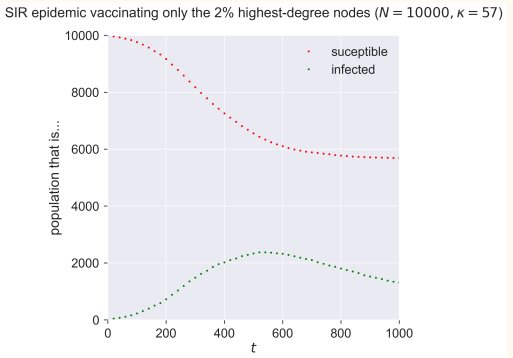
Randomly vaccinate

- Massive reduction in epidemic size
- …but need to get ¿60% of the population
- Only catching high-degree nodes at random



Target 2% highest-degree nodes

- Immunise the most likely super-spreaders
- Can also use social knowledge



Physical distancing

What does a physically-distanced contact network look like?

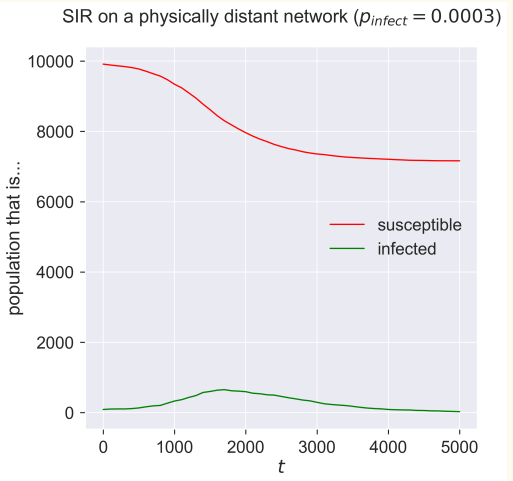
- Good question: *lots* of assumptions, especially about compliance

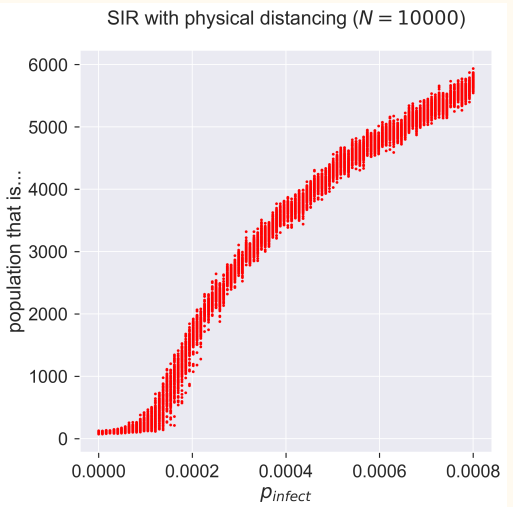
Changes the epidemic threshold compared to an ER network

- Needs a higher infectivity to take off

Slower take-off

- Not like a powerlaw network
- "Bursts" if the infection gets into a bubble





When multiple variants emerge

As the pathogen evolves, we see different variants with different behaviours

- Often more transmissible but less severe
- Coupled of increased immunity, leads to epidemic dying out

Environment controls selection pressures

- In systems with only short-range connections, highly contagious variants are often contained by previous infections
- Whereas in systems with long-range connections, the most contagious variant almost always spreads globally [9]

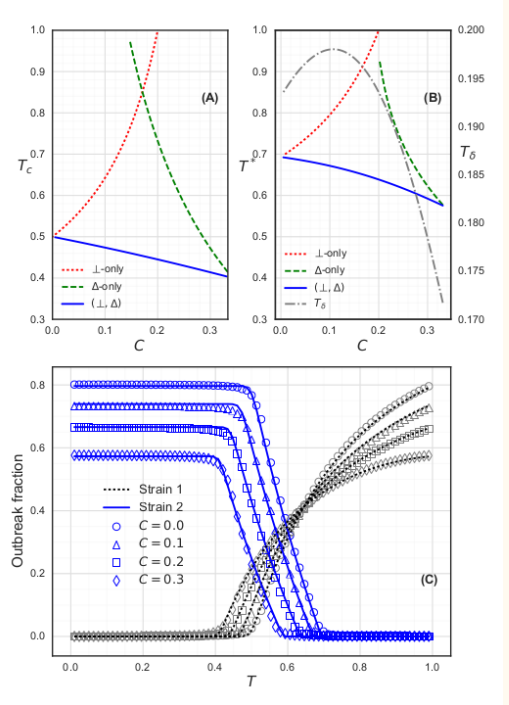
[9] M. Boots and A. Sasaki. 'Small worlds' and the evolution of virulence: Infection occurs locally and at a distance. *Proceedings of the Royal Society B*, 266(1432):1933–1938, October 1999

Co-infection dynamics

What happens when variants co-exist?

- May co-operate: previous infection with one makes you more sensitive to the next
- Or may compete: having one reduces the risk of re-infection [10]

Lots more work to do to understand this



[10] P. Mann, V. A. Smith, J. Mitchell, and S. Dobson. Two-pathogen model with competition on clustered networks. *Physical Review E*, 103(6), June 2021b. URL https://doi.org/10.1103/PhysRevE.103.062308

Structure of this talk

Research directions

Multiple variants

- What happens when disease evolve?
- More detailed co-infection dynamics

We're now very interested in network fine structure

- How do processes behave in detail?
- Can they be "steered" by disrupting small local features?
- New analytical techniques, based on graph signal processing
- Improved tooling, new software and algorithms

Three things to take away

1. Epidemic spreading still isn't fully understood – there's lots of exciting work still to do, mathematically and computationally
2. Interactions between network and process can be very subtle, and may have significant effects
3. We can explore the space of public policy decisions as "citizen scientists", and also counter misinformation

References


M. Boots and A. Sasaki. 'Small worlds' and the evolution of virulence: Infection occurs locally and at a distance. *Proceedings of the Royal Society B*, 266(1432):1933–1938, October 1999.


S. Dobson. *Epidemic modelling – Some notes, maths, and code*. Independent Publishing Network, 2020. ISBN 978-183853-565-0. URL https://simoninireland.github.io/introduction-to-epidemics/.


D. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25): 2340—2361, 1977.


H. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, December 2000. URL doi://10.1137/S0036144500371907.


P. Mann, V. A. Smith, J. Mitchell, and S. Dobson. Random graphs with arbitrary clustering and their applications. *Physical Review E*, 103(1), January 2021a. URL https://doi.org/10.1103/PhysRevE.103.012309.


P. Mann, V. A. Smith, J. Mitchell, and S. Dobson. Two-pathogen model with competition on clustered networks. *Physical Review E*, 103(6), June 2021b. URL https://doi.org/10.1103/PhysRevE.103.062308.


M. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66, July 2002. URL doi://10.1103/PhysRevE.66.016128.


Royal Society SET-C group. Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK, August 2020. URL https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf.


S. Shai and S. Dobson. Coupled adaptive complex networks. *Physical Review E*, 87(4), April 2013. URL https://dx.doi.org//10.1103/PhysRevE.87.042812.