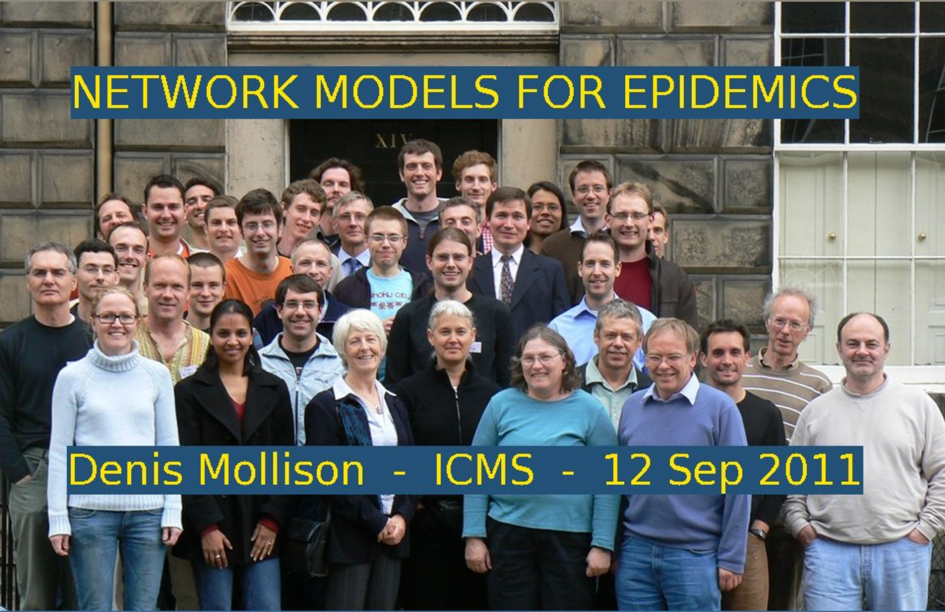
NETWORK MODELS FOR EPIDEMICS



Denis Mollison - ICMS - 12 Sep 2011

Deaths from infectious disease:

	11 Sep 2001	11 Sep 2011
AIDS	7900	4900
Diarrhoea	5500	6700
TB	4500	3700
Malaria	3100	1500

Deaths from infectious disease:

	11 Sep 2001	11 Sep 2008
AIDS	7900	4900
Diarrhoea	5500	6700
TB	4500	3700
Malaria	3100	1500

. . . all epidemiology, concerned as it is with the variation of disease from time to time or from place to place, must be considered mathematically, however many variables are implicated, if it is to be considered scientifically at all. To say that a disease depends upon certain factors is not to say much, until we can also form an estimate as to how largely each factor influences the whole result. And the mathematical method of treatment is really nothing but the application of careful reasoning to the problem at issue.

Ross (ca. 1911)

$R_0 = Mb^2cd$

where M is # mosquitoes/human

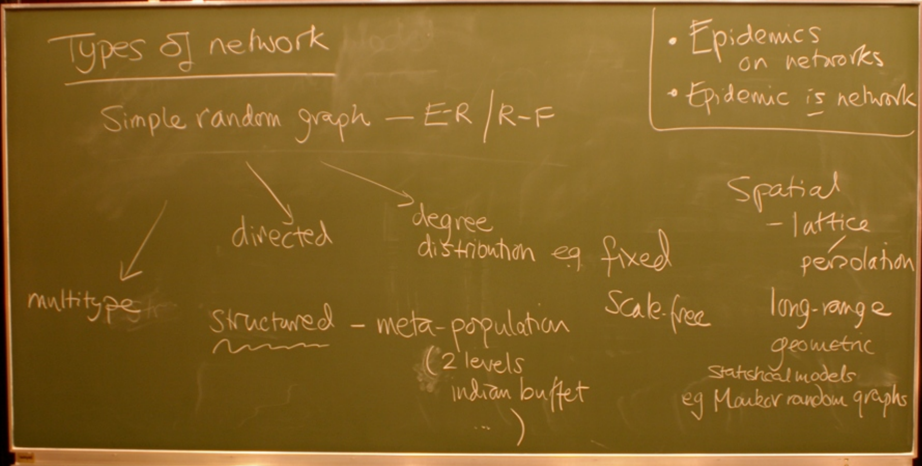
Ross (ca. 1911)

$R_0 = Mb^2cd$

where M is # mosquitoes/human,

whence

$R_0 < 1$ iff $M < M_c$.



Types of network

Simple random graph — E-R / R-F

multitype

directed

degree distribution eg fixed

Scale-free

structured — meta-population
(2 levels
indian buffet
...)

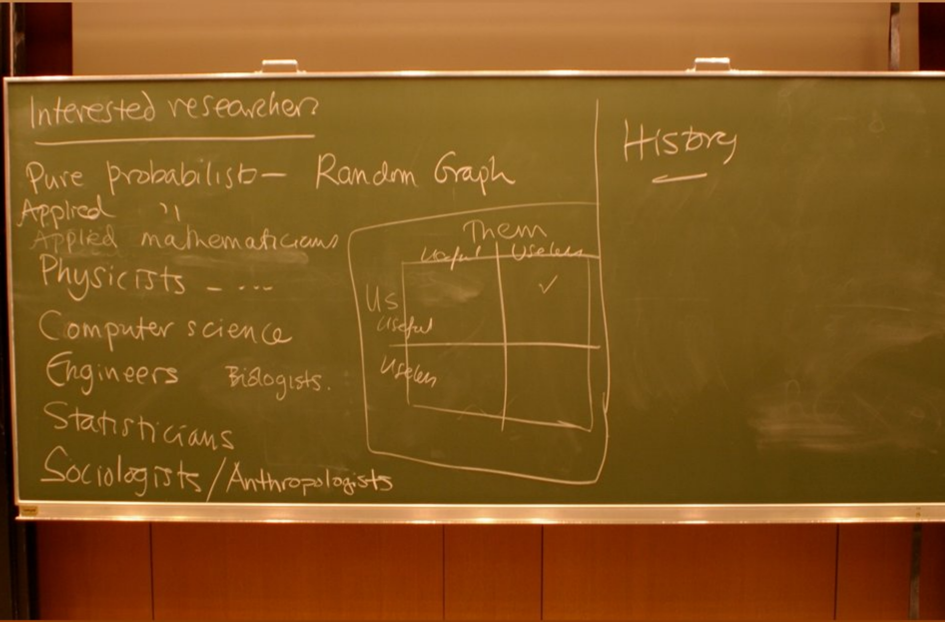
- Epidemics on networks
- Epidemic is network

Spatial
- lattice / percolation

long-range

geometric

statistical models
eg Markov random graphs



Interested researchers

Pure probabilists – Random Graph
Applied "
Applied mathematicians
Physicists ....
Computer science
Engineers   Biologists.
Statisticians
Sociologists / Anthropologists

Them

		Useful	Useless
Us	Useful		✓
	Useless		

History

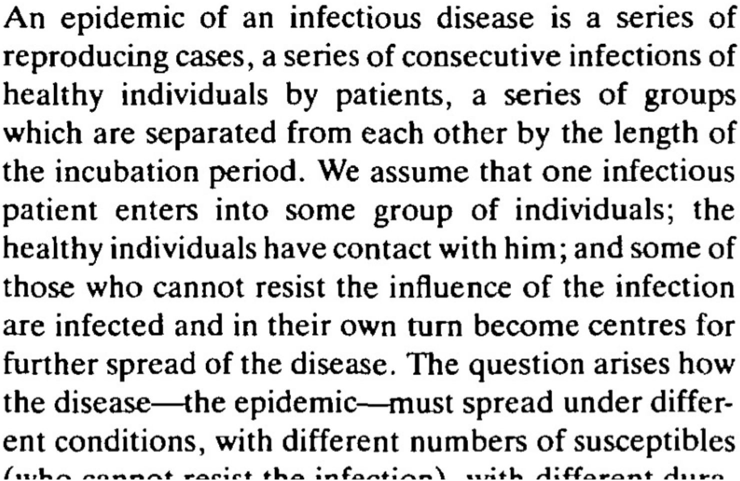
Basic network idea:

A pair of individuals are linked . . .

$$i \to j$$

. . . if i can (or does) infect j

An epidemic of an infectious disease is a series of reproducing cases, a series of consecutive infections of healthy individuals by patients, a series of groups which are separated from each other by the length of the incubation period. We assume that one infectious patient enters into some group of individuals; the healthy individuals have contact with him; and some of those who cannot resist the influence of the infection are infected and in their own turn become centres for further spread of the disease. The question arises how the disease—the epidemic—must spread under different conditions, with different numbers of susceptibles (who cannot resist the infection), with different dura



An epidemic of an infectious disease is a series of reproducing cases, a series of consecutive infections of healthy individuals by patients, a series of groups which are separated from each other by the length of the incubation period. We assume that one infectious patient enters into some group of individuals; the healthy individuals have contact with him; and some of those who cannot resist the influence of the infection are infected and in their own turn become centres for further spread of the disease. The question arises how the disease—the epidemic—must spread under different conditions, with different numbers of susceptibles (who cannot resist the infection), with different dura

En'ko 1889

In each generation,

Pr(escape infection)

$$= (1 - i)^{pN} \qquad \text{(Enko 1889)}$$

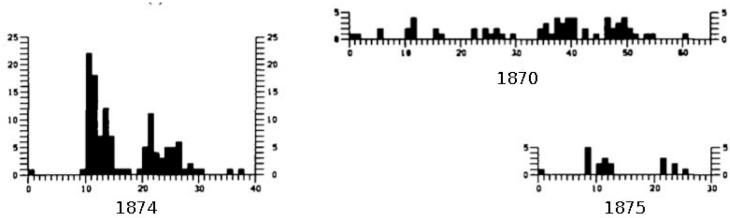
$$= (1 - p)^{iN} \qquad \text{(Reed-Frost)}$$

where $p = \mathrm{P(contact)}$, $i =$ proportion infected

$\longrightarrow$

$$I_{n+1} \sim \mathrm{Binomial}(S_n, (1 - (1 - p)^{I_n}))$$

Data analysis:



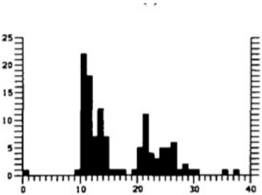
1874

1870

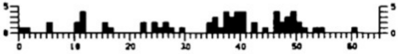
1875

Three measles outbreaks in the Educational College for the Daughters of the Nobility

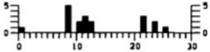
En'ko 1889



1874



1870



1875

1874: $1 \rightarrow 70 \rightarrow 45 \rightarrow 2$

Best fitting simulation: $1 \rightarrow 79 \rightarrow 53$

when $N = 400, S_0 = 133, S_T = 0, A = 360$

Reed-Frost

The chain-binomial, with

$I_{n+1} \sim \text{Binomial}(S_n, (1 - (1 - p)^{I_n}))$

is generally credited to Reed & Frost, ca. 1928, but ..

Reed-Frost

The chain-binomial, with

$I_{n+1} \sim \text{Binomial}(S_n, (1-(1-p)^{I_n}))$

is generally credited to Reed & Frost, ca. 1928, but ..

.. was not widely known until ca. 1950.

Main applications in small populations (households) (see *e.g.* Becker 1989)

Reasons R-F went out of fashion

- difficulties of calculation

- attractions not appreciated ..

- .. or held against it

- rise of continuous-time models

  (techniques including DEs, PGFs, and

  branching-process approximations)

  with emphasis on ‘mass-action’

Mass-action models

Deterministic:

Hamer 1906	discrete-time
Ross 1908	continuous-time
Kermack & McKendrick 1927	DEs for SIR

Stochastic:

McKendrick 1926

Bartlett 1949

Example (value of “unnecessary” detail):

Simple birth and death process

(1) $r_{n,n+1} = an, \quad r_{n,n-1} = bn$

(2) Independent individuals, each with birth rate a and death rate b.

P(extinction) μ_n when initial pop. $= n$?

Example (value of “unnecessary” detail):

Simple birth and death process

(1) $r_{n,n+1} = an, \quad r_{n,n-1} = bn$

(2) Independent individuals, each with birth rate a and death rate b.

P(extinction) μ_n when initial pop. $= n$

$$= \mu_1^n$$

Spatial processes

Network models first used explicitly for spatial case, because individual-based models more obviously needed:

Broadbent & Hammersley 1957 motivated **percolation theory** with the example of “spread of disease in an orchard’

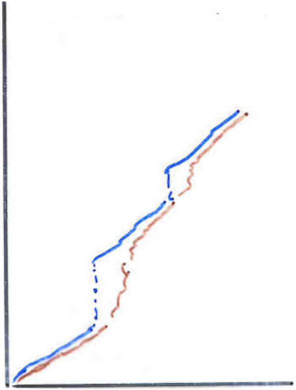
Specific spatial epidemic models

Morgan & Welsh 1965

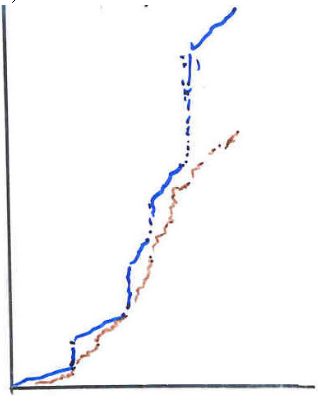
Mollison 1972:

Velocity of 1-D stochastic epidemics

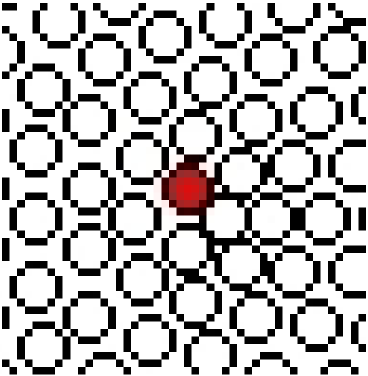
- theory
- simulations
- comparison with DE models



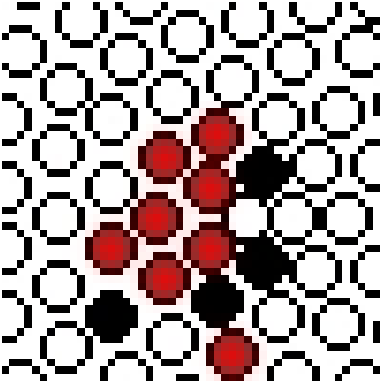
$v(s) \sim s^{-4}$



$v(s) \sim s^{-3}$



Time 0



Time 1

Comparing different local structures

(Kuulasmaa & Zachary 1984)

A a subset of neighbours of i,

$q(A) = P[i$ doesn't infect any of $A]$.

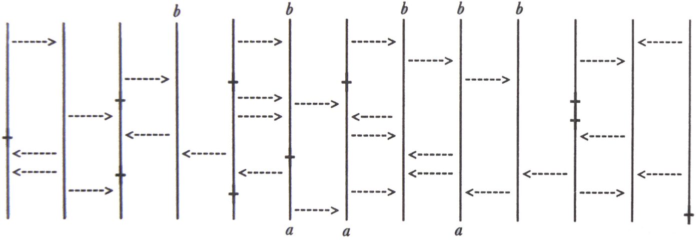
If $q_1(A) \leq q_2(A)$ for all A,

then "$1 \geq 2$"

[*E.g.* indep. contacts $\geq$ correlated contacts]

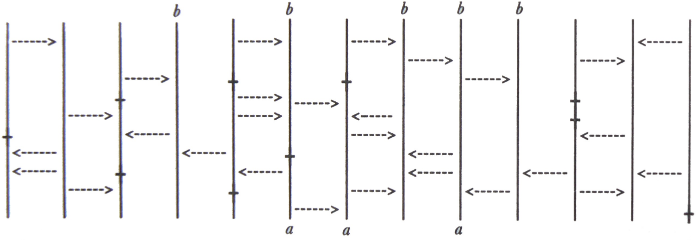
Harris 1974: the "contact" process (= SIS)

A coupling argument . . .



Harris 1974: the "contact" process (= SIS)

A coupling argument ...



...shows that this is monotone with initial set

Cox & Durrett (1988): the "contact process" has an asymptotic velocity

Liggett 1985 *Interacting Particle Systems*

Liggett 1985 *Interacting Particle Systems*

Meanwhile, in another part of the wood, more probabilists were at work …

Simple random graphs

Erdos and Renyi,

followed by Bollobas and others

were finding out more about a very simple model than En'ko could have hoped for.

Simple random graphs

Erdos and Renyi,

followed by Bollobas and others

were finding out more about a very simple model than En'ko could have hoped for (or wanted?).

(En'ko / Reed-Frost revisited)

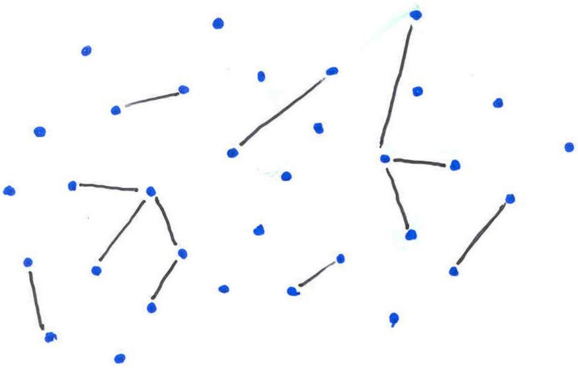
In each generation,

Pr(escape infection) $\quad = (1-p)^{iN}$

$\longrightarrow$

$I_{n+1} \sim \text{Binomial}(S_n, (1-(1-p)^{I_n}))$

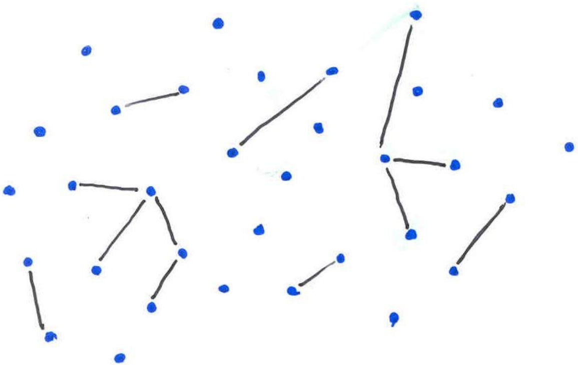
Simple random graph



N individuals, each pair linked with probability p

- Why are these links undirected?
- Why are they independent?

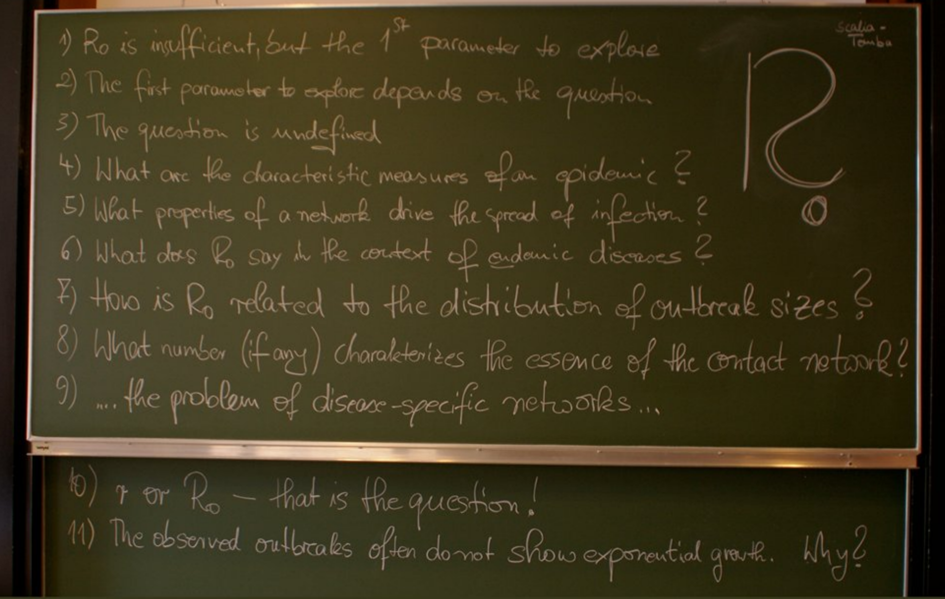
Simple random graph



Here $R_0 \equiv Np$ is < 1

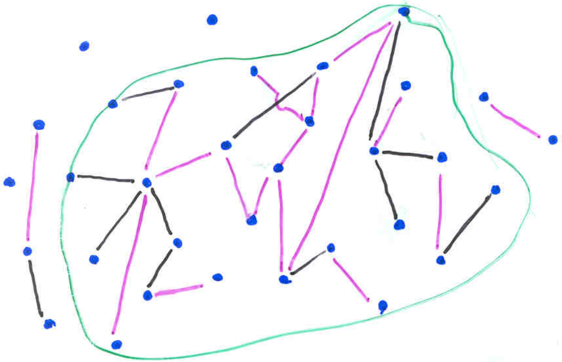
R_0

The ***basic reproductive ratio*** of an epidemic is the mean number of new infections made by an infected individual in a mostly susceptible population



Scalia-Tomba

1) R_0 is insufficient, but the 1st parameter to explore
2) The first parameter to explore depends on the question
3) The question is undefined
4) What are the characteristic measures of an epidemic?
5) What properties of a network drive the spread of infection?
6) What does R_0 say in the context of endemic diseases?
7) How is R_0 related to the distribution of outbreak sizes?
8) What number (if any) charakterizes the essence of the contact network?
9) ... the problem of disease-specific networks...
10) r or R_0 — that is the question!
11) The observed outbreaks often do not show exponential growth. Why?



Here $R_0 \equiv Np$ is > 1

Results for simple random graph:

Giant component exists iff $R_0 > 1$.

Results for simple random graph:

Giant component exists iff $R_0 > 1$.

Diameter of giant, $T \sim \log N$.

Results for simple random graph:

Giant component exists iff $R_0 > 1$.

Diameter of giant, $T \sim \log N$.

Final size (and probability of a large outbreak) are both given by the largest solution of

$$z = 1 - \exp(-R_0 z)$$

Deterministic mass-action equivalent,
a differential equation model ('SIR'):

$$\begin{aligned}\dot{S} &= -cSI \\ \dot{I} &= cSI - dI \\ \dot{R} &= dI\end{aligned}$$

Results for ‘SIR’:

Large outbreak *always* occurs if $R_0 \equiv c/d > 1$,
duration $T \sim \log N$,
and the final size z is given by

Results for ‘SIR’:

Large outbreak *always* occurs if $R_0 \equiv c/d > 1$,
duration $T \sim \log N$,
and the final size z is given by

$$z = 1 - \exp(-R_0 z)$$

Epidemiologists are interested in more than just the Simple Random Graph

Structural choices for network models

- Directed or undirected?
- Degree – fixed? Poisson? power-law?
- Large-scale structure
  (mean-field to spatial)

Undirected links??

$A_{ij} =$ "i infects j"

In R-F, A_{ij}s are all i.i.d. w.p. p

– this requires:

(a) infectious period T_i constant

(b) $P[i \rightarrow j] = P[j \rightarrow i]$

Then (c) in any realisation we are interested in only one of A_{ij} and A_{ji}, so we can represent them by a single (undirectional) link.

Contacts or potential contacts?

In R-F we can either think of all others as potential contacts, each an actual contact with probability p; or of a Binomial (asymptotically Poisson) degree distribution prescribing "realised" contacts.

One generalisation is to take the latter approach with arbitrary degree distribution.

$R_0 = E[D] = \sum d\pi_d$?

Effective value, $R_0' = \sum d\pi_d' - 1$

where $\pi_d' = d\pi_d / \sum d\pi_d$

whence $R_0' = E[D^2]/E[D] - 1$

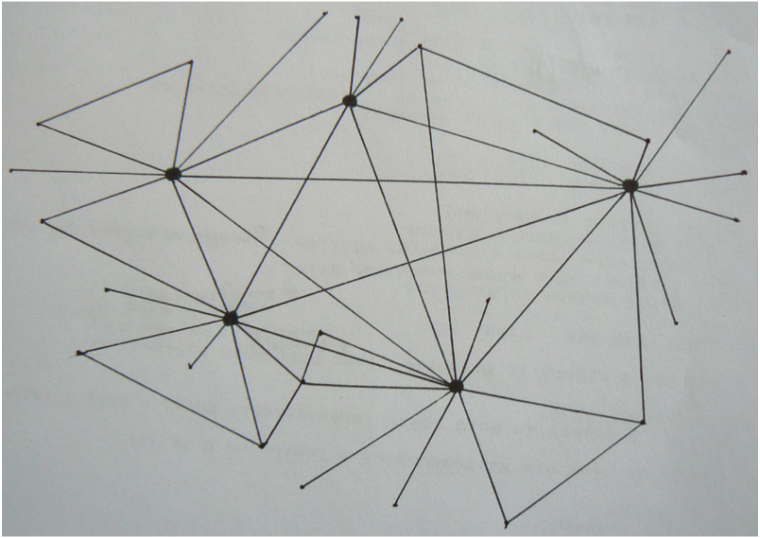
($= E[D]$ for Poisson degree distribution).

Note

Contrast traditional (?) epidemic models where numbers of incoming and outgoing links are not correlated, so we don't get this "size-biased" effect.

An extreme example exhibiting size bias is

'Scale-free' models:

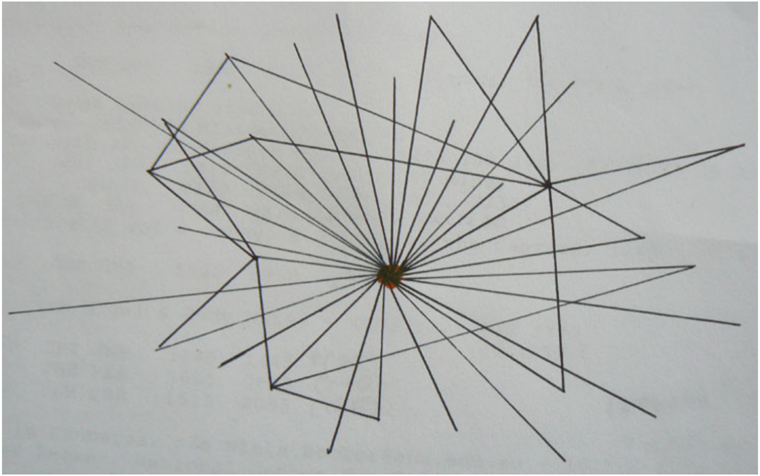


$T = 3$

Sector times and distances may vary according to seasonal weather, flying conditions, aircraft type and route variations. Routes shown are indicative only.

- Location of The Qantas Clubs or associated lounges
- ——— Qantas routes
- -·-·- Jetstar routes
- - - - QantasLink routes
- ······ Routes operated by other airlines for Qantas

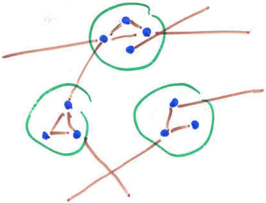
- ✵ International Gateway Port
- ● National Capital



$T = 2$

Metapopulation models *(BMS-T 1997)*

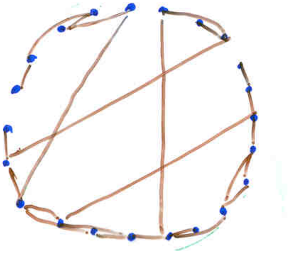
Consider a population with local and global contacts



where the geography can be either mean-field

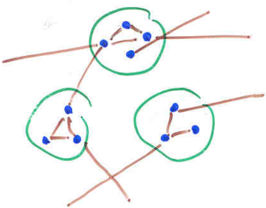
. . .

. . . or spatial



(‘great circle’ or ‘small world’ model)

Consider first the process including only global contacts, with reproductive ratio $R_0 = Nq$.



Relative to this 'global-only' process, local contacts have an amplifying effect.

Hence the overall reproductive ratio is

$$R_T = R_0 \mu$$

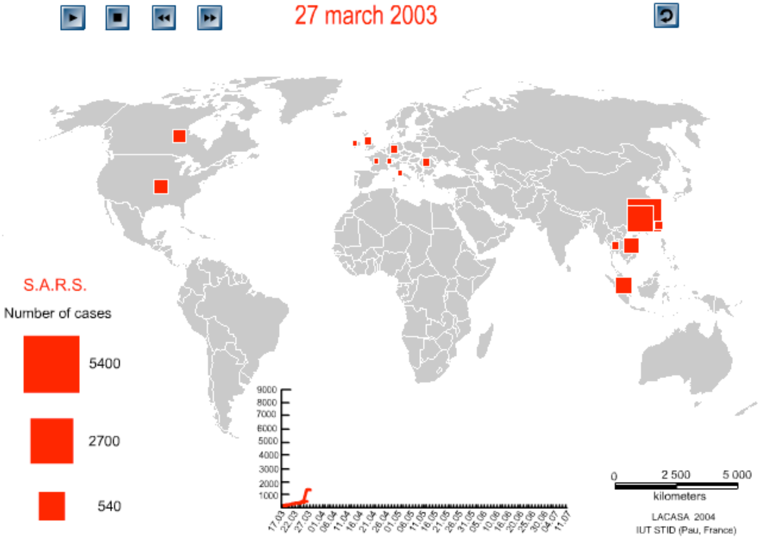
where μ is the mean size of a local outbreak.

Hence the overall reproductive ratio is

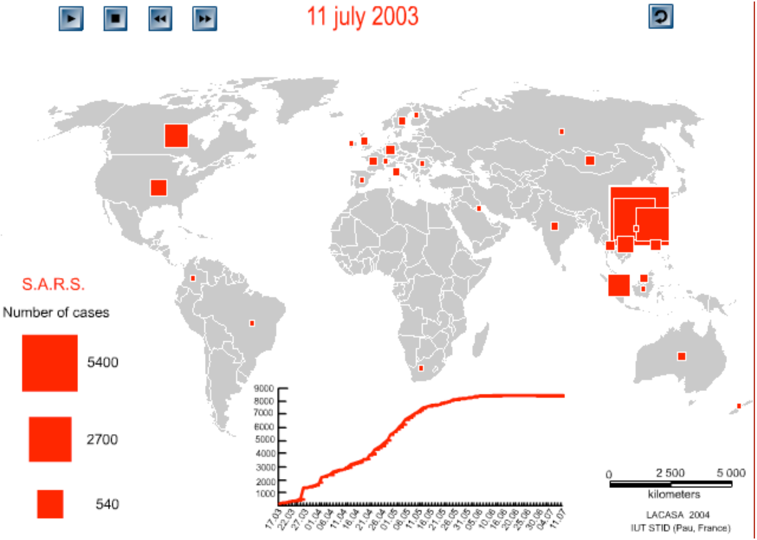
$$R_T = R_0 \mu$$

where μ is the mean size of a local outbreak.

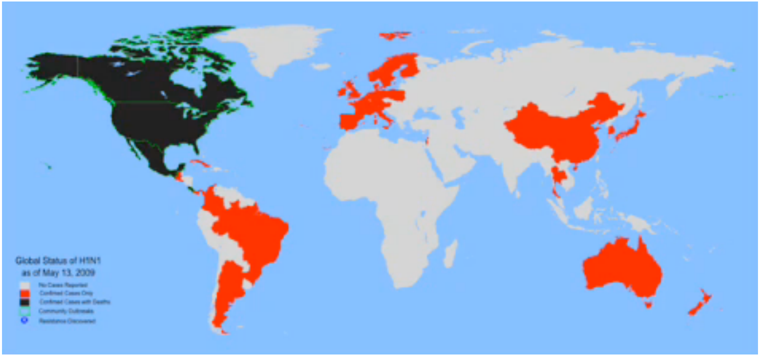
A key question for control is whether you can get local outbreaks below threshold (compare SARS and swine flu?)



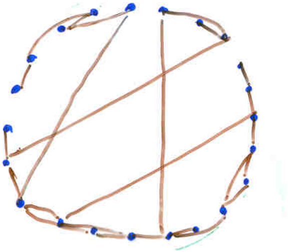
http://www.cybergeo.eu/index12803.html



http://www.cybergeo.eu/index12803.html


Global Status of H1N1
as of May 13, 2009
No Cases Reported
Confirmed Cases Only
Confirmed Cases with Deaths
Community Outbreaks
Resistance Discovered

Small worlds



Threshold: $R_T = R_0\mu > 1$ (as for metapopulation model)

T reduces from $\sim N$ to $\sim \log N$ as the number of global links increases

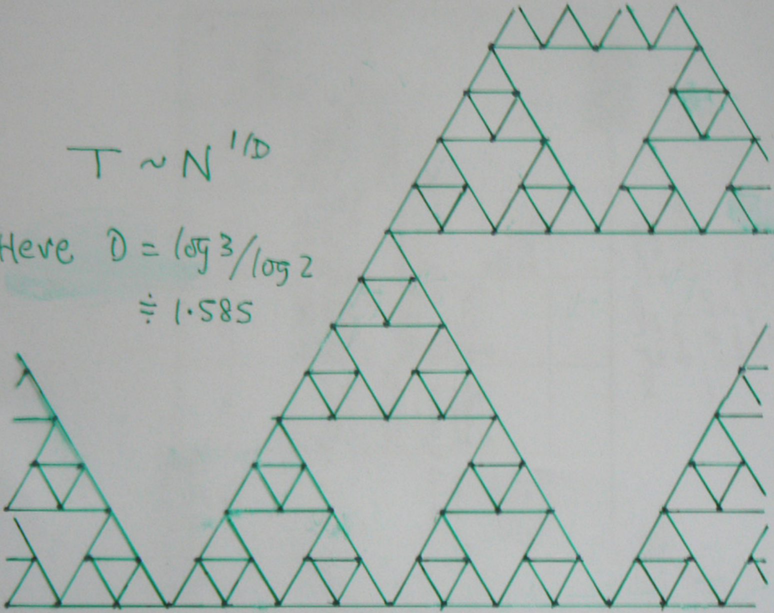
‘Small world’ phenomenon:

The proportion of global links required to collapse the spatial model to one close to homogeneous mixing, reducing T to $\sim \log N$, is surprisingly small.

$T \sim N^{1/D}$

Here $D = \log 3 / \log 2$

$\doteqdot 1.585$



Advantages of network models

+ “links” $i \rightarrow j$ captures idea of infectious contact

+ clarity (potentially)

− but not the only approach

Have tried to include examples of some nice techniques, whether network-based or not.

One last example . . .

Sellke construction for R-F

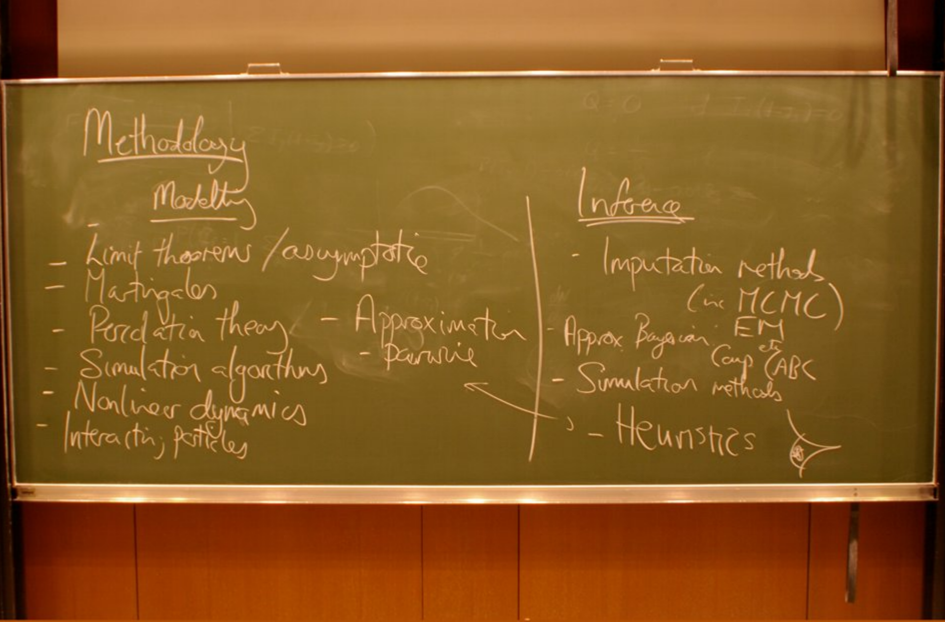
Note that $P[\text{escape } n \text{ attacks}] = (1-p)^n$

Choose X_i i.i.d. Uniform$[0,1]$

Start with initial set of infected;

when cumulative total $= I$,

i becomes infected iff $X_i > (1-p)^I$.



Methodology

Modelling

- Limit theorems / asymptotics
- Martingales
- Percolation theory
- Simulation algorithms
- Nonlinear dynamics
- Interacting particles

- Approximation
  - pairwise

Inference

- Imputation methods (inc MCMC)
- Approx Bayesian Comp (ABC
- EM etc
- Simulation methods
- Heuristics