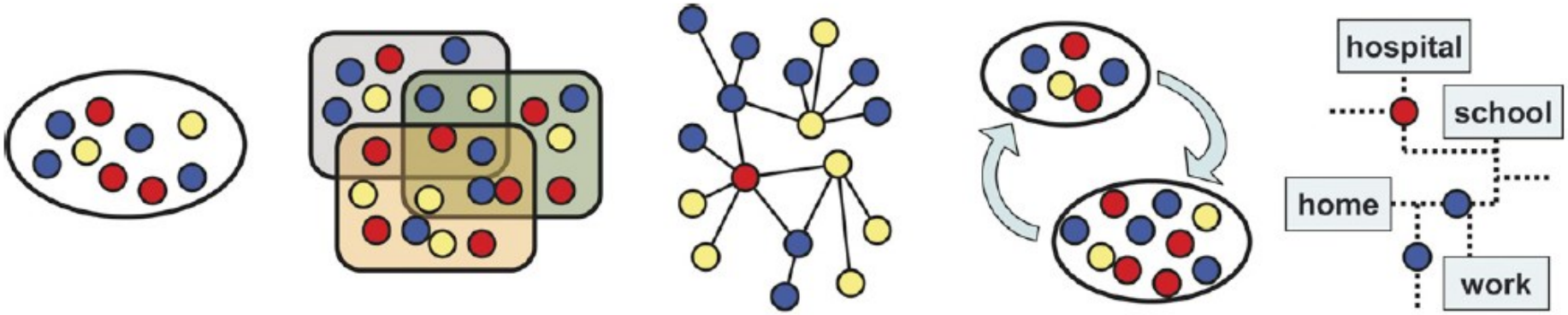
Systems Biology Across Scales: A Personal View XVI. Network Epidemiology

Sitabhra Sinha

IMSc Chennai

Multiple Scales of Epidemic Models



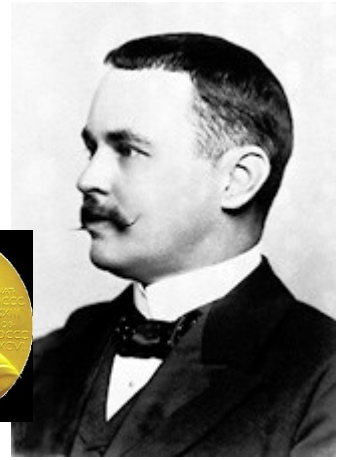
Homogeneous mixing	Social structure	Contact network models	Multi-scale models	Agent Based models
No consideration of spatial or social structure in a community	Individuals grouped according to demographic properties: age, gender,…	Agents affect those who are in direct contact with them	Spreading between communities in different spatial locations	Detailed consideration of spatial movements & interpersonal contact

Modeling how epidemics spread

E.g., can be based on Lotka-Volterra predator-prey dynamics

Growth rate of prey: $dX/dt = a X - b XY$
Growth rate of predator: $dY/dt = c XY - dY$

1902



Ronald Ross
(1857-1932)

The earliest mathematical model of epidemic spreading
Malaria: Ronald Ross (1909),
modified by George MacDonald (1950-3)

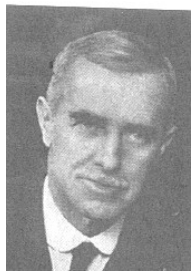
For a vector-borne disease
Infected humans X: $dX/dt = ab\,Y\,(N - X) - g\,X$
Infected mosquitos Y: $dY/dt = ac\,X\,(M - Y) - dY$

a: biting rate per human per mosquito
b,c: mosquito to human & human to mosquito transmission probability per bite
g: recovery rate of humans
N: total number of humans
d: death rate of mosquitos
M: total number of mosquitos

A Simple Model of Epidemic Dynamics

Modeling spreading by direct contact (SIR model)
Kermack-McKendrick (1927)

W O Kermack
(1898-1970)

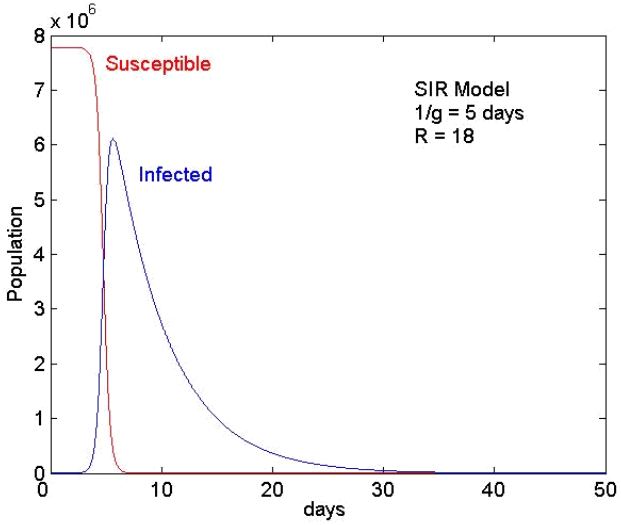


A G McKendrick
(1876-1943)

Under assumption of homogeneous mixing,
i.e., anyone is equally likely to infect anyone else:

Susceptible population : $dS/dt = -\beta\ S\ I$

Infected population : $dI/dt = \beta\ SI - \gamma\ I$



$\Rightarrow$ Epidemic : $dI/dt > 0$, i.e., $S(t=0) > \gamma / \beta$

β : rate of infection spreading

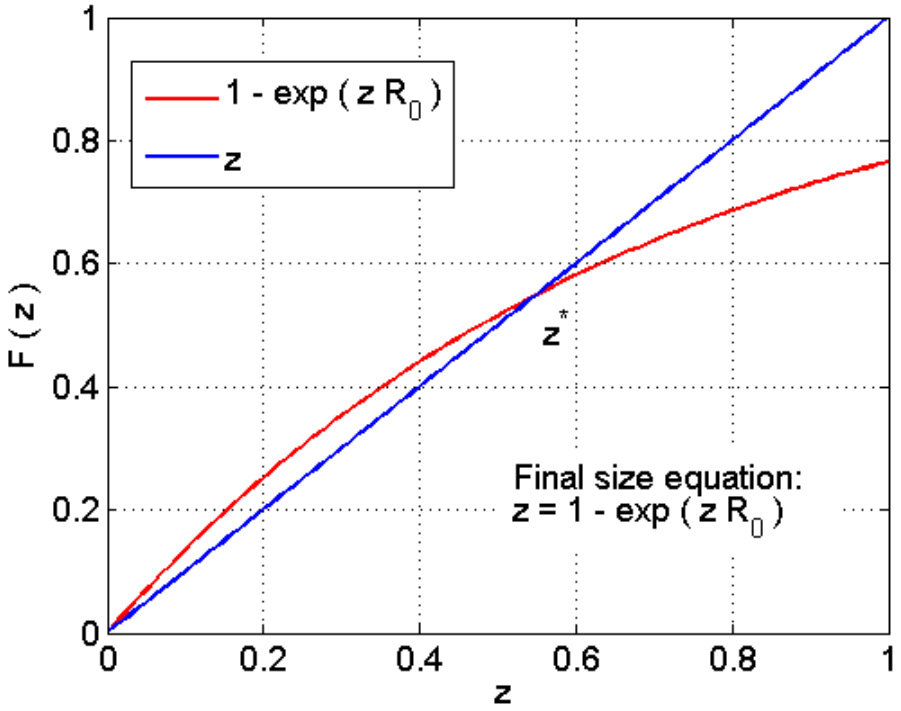
γ : recovery rate (= [avg infectious period, τ]$^{-}$

$S\ (t = 0) = N$, the total population

An important parameter: *effective reproduction number* R(t), the average number of new infections per infected person

$\Rightarrow$Epidemic if *basic reproduction number* $R_0 = R\ (t=0) = N\ \beta\ \tau > 1$

Utility of R_0 : Estimating the final size



An individual that is susceptible at time t_0 and experiences a force of infection $\lambda(t)$ for $t > t_0$ will escape from being infected with probability $\mathcal{F}(t)$ defined by

$$\left.\begin{array}{l} \frac{d\mathcal{F}}{dt} = -\lambda\mathcal{F} \\ \mathcal{F}(t_0) = 1 \end{array}\right\} \Rightarrow \mathcal{F}(t) = e^{-\int_{t_0}^{t} \lambda(\tau)\, d\tau}.$$

For large populations we shall have $s(\infty) = \mathcal{F}(\infty)$, i.e. the fraction that remains susceptible equals the probability to remain susceptible. Let us call $\int_{t_0}^{\infty} \lambda(\tau)\, d\tau$ the total cumulative force of infection. The fraction $z = 1 - s(\infty)$ that falls victim to the infection generates a total cumulative force of infection equal to

$$\frac{1}{N} pc\Delta T z N = R_0 z$$

(we have to divide by N since contacts are with probability $1/N$ with the susceptible individual that we consider). Hence

$$s(\infty) = \mathcal{F}(\infty) = e^{-R_0(1-s(\infty))},$$

Diekmann & Heesterbeek, 2000

The final size of the epidemic z is the proportion of population which will experience infection by the end of the epidemic.

Given by the implicit equation: $z = 1 - \exp(z R_0)$

Can be solved graphically

For $R_0 \approx 1.45$, we obtain $z \approx 55\%$

Utility of R_0 : Minimum immunization coverage required to stop epidemic

Let us recall the SIR model equation: $dI/dt = \beta SI - \gamma I$

$\Rightarrow$ To stop epidemic we need to make $dI/dt < 0$, i.e., $S(t=0) < \gamma / \beta$

where

β : rate of infection spreading

γ: recovery rate (= [avg infectious period, τ]$^{-1}$)

Let total population be N

Thus, proportion of the population that is susceptible, $s = S(t=0)/N$ needs to be made smaller than $1/(N\beta\tau) = 1/R_0$ (because $R_0 = N\beta\tau$)

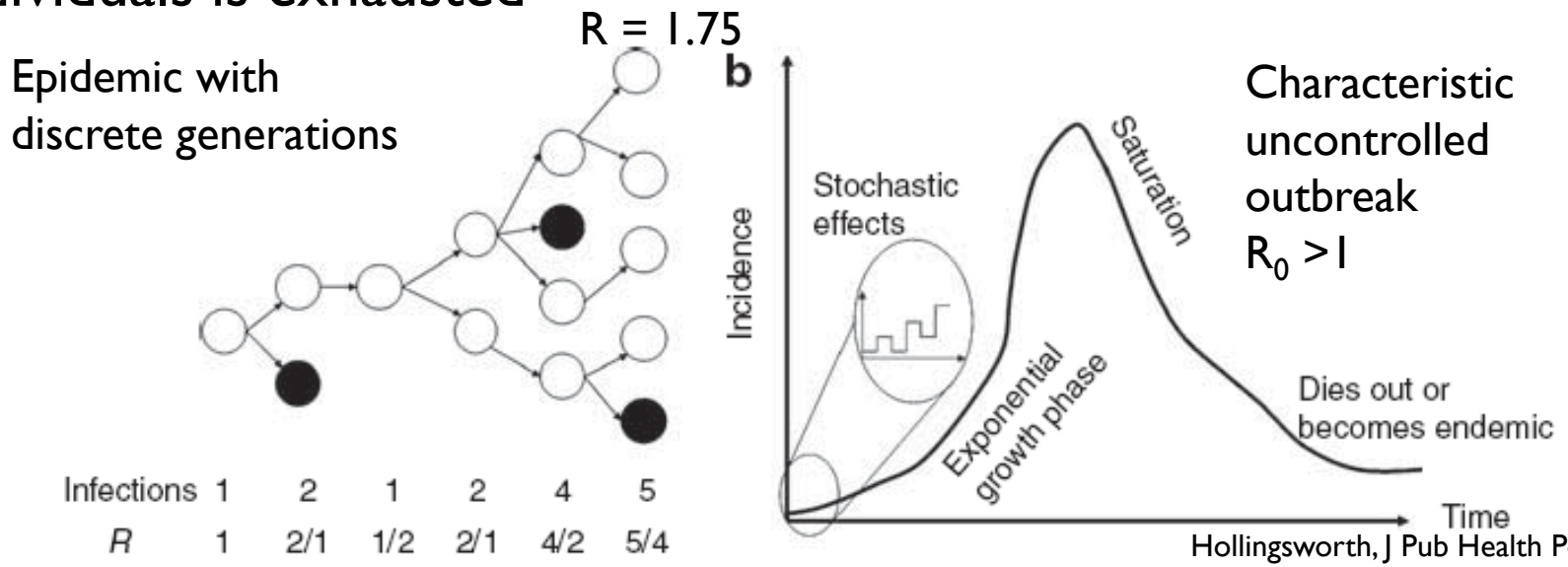
$\Rightarrow$The fraction of population that needs to be immunized to stop the epidemic (assuming homogeneous mixing) is $p > 1 - (1/R_0)$

For $R_0 \approx 1.45$, $p_{min} \approx 31\%$

Basic reproduction number R_0

Mean number of new infections caused by a single infectious individual in a <u>wholly susceptible</u> population (as in the beginning of an epidemic): If each infected person on average infects more than one other individual, $R_0 > 1 \Rightarrow$ Epidemic

Initially the epidemic may die out due to stochastic fluctuations, but once established it grows <u>exponentially</u> until the pool of susceptible individuals is exhausted



Hollingsworth, J Pub Health Policy, 2009

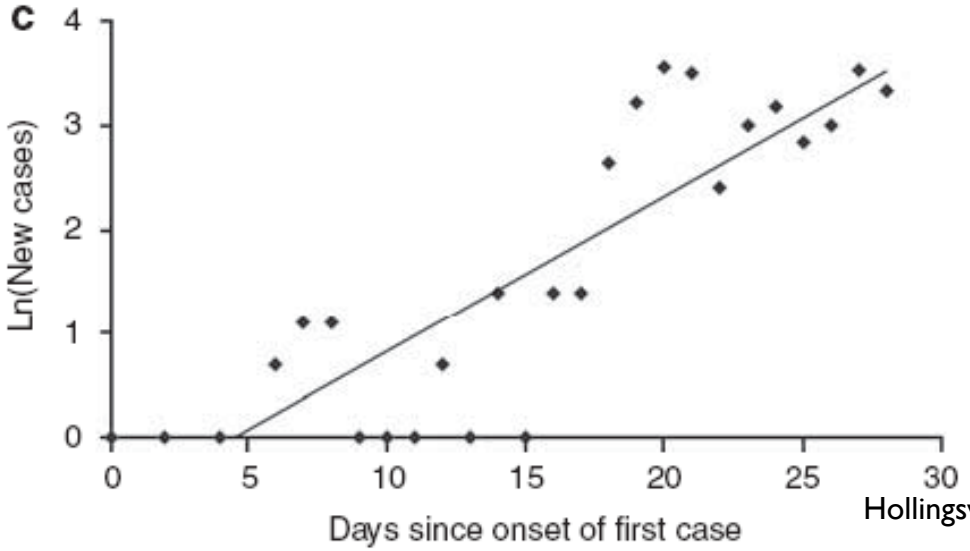
Estimating basic reproduction number R_0

A frequently used approach is

• to fit an exponential function to the incidence data to obtain the exponential growth rate λ, and,

• to use the approximate relation $R_0 \approx \exp(\lambda \tau_g)$ where τ_g is the observed mean generation interval of the epidemic (defined as "the sum of the average latent and the average infectious period" [Anderson & May, 1991])

For small $\lambda \tau_g$ we can further approximate $R_0 \approx 1 + \lambda \tau_g$



Hollingsworth, J Pub Health Policy, 2009

Stochastic SIR dynamics

Transmission of infection and recovery are essentially stochastic processes

The deterministic SIR model does not account for fluctuations, particularly important at the beginning of an epidemic when I is small

The stochastic infection dynamics:

Infection: $S + I \xrightarrow{\alpha} 2I$ Recovery: $I \xrightarrow{\beta} 0$

α: infection rate

β: recovery rate

Quantity of interest: probability of finding *S* susceptibles and *I* infected in a population of size *N* at time *t*, $p(S, I, t)$

The master equation for the evolution of this probability :

$$\partial_t p(S, I; t) = \frac{\alpha}{N}(S+1)(I-1)\, p(S+1, I-1; \quad + \beta(I+1)\, p(S, I+1; t)$$

$$- \left(\frac{\alpha}{N} SI\, p - \beta I\right) p(S, I; t)$$

With initial condition

$$p(S, I; t = t_0) = \delta_{I,I_0}\delta_{S,N-I_0}$$

In the limit of large *N*, this approaches the deterministic SIR model results because the fluctuations decay as $1/\sqrt{N}$

Going beyond the assumptions of homogeneous mixing

Homogeneous mixing: any agent is equally likely to infect any other agent

In reality,
(A) social relations and (B) physical or geographical proximity,
make some people more likely to be infected than others

Important to consider either or both of these effects in any realistic model of epidemic spreading

(A): role of contact structure (spreading through a network)
(B): role of space (diffusion of an epidemic front)

Epidemic spreading on WS networks

Watts & Strogatz: SIR dynamics on small-world network

Epidemics spread much faster on WS networks than an equivalent regular network – and far more difficult to control by partial removal of nodes than on a random network !

The shortcut links make the transport process fast

Unlike a random network, where every node is more or less equivalent, so that removing a certain fraction of susceptible agents (making $R < 1$) ensures the epidemic dies, here the few nodes that are terminals of shortcut links are principally responsible for rapid transit of infection

For efficient control of epidemic on WS networks, necessary to identify the "shortcuts" and preferentially control those
In the context of STD or AIDS: "Control the truck drivers"

No threshold for epidemics in *scale-free networks*

Networks of sexual relations have been claimed to be scale-free !
A few highly promiscuous individuals act as "hub" nodes
May play a crucial role in spreading sexually transmitted diseases !

If the contact structure of a disease is network with inhomogeneous degree distribution, the condition for occurrence of an epidemic is:

$$R_0 = \beta N / \gamma \quad > \quad \langle k \rangle / \langle k^2 \rangle$$

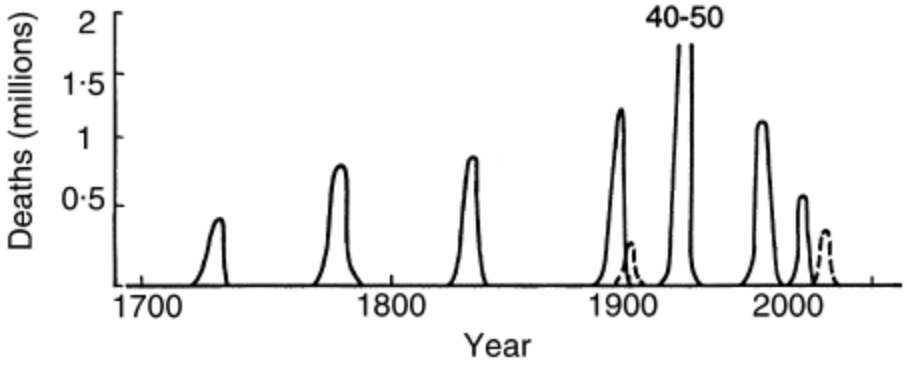
Initial popn of susceptibles, S(t = 0) = N, the total population

For a scale-free network having degree exponent $2 < \alpha \leq 3$, $\langle k^2 \rangle \rightarrow \infty$
$\Rightarrow$There is no epidemic threshold !

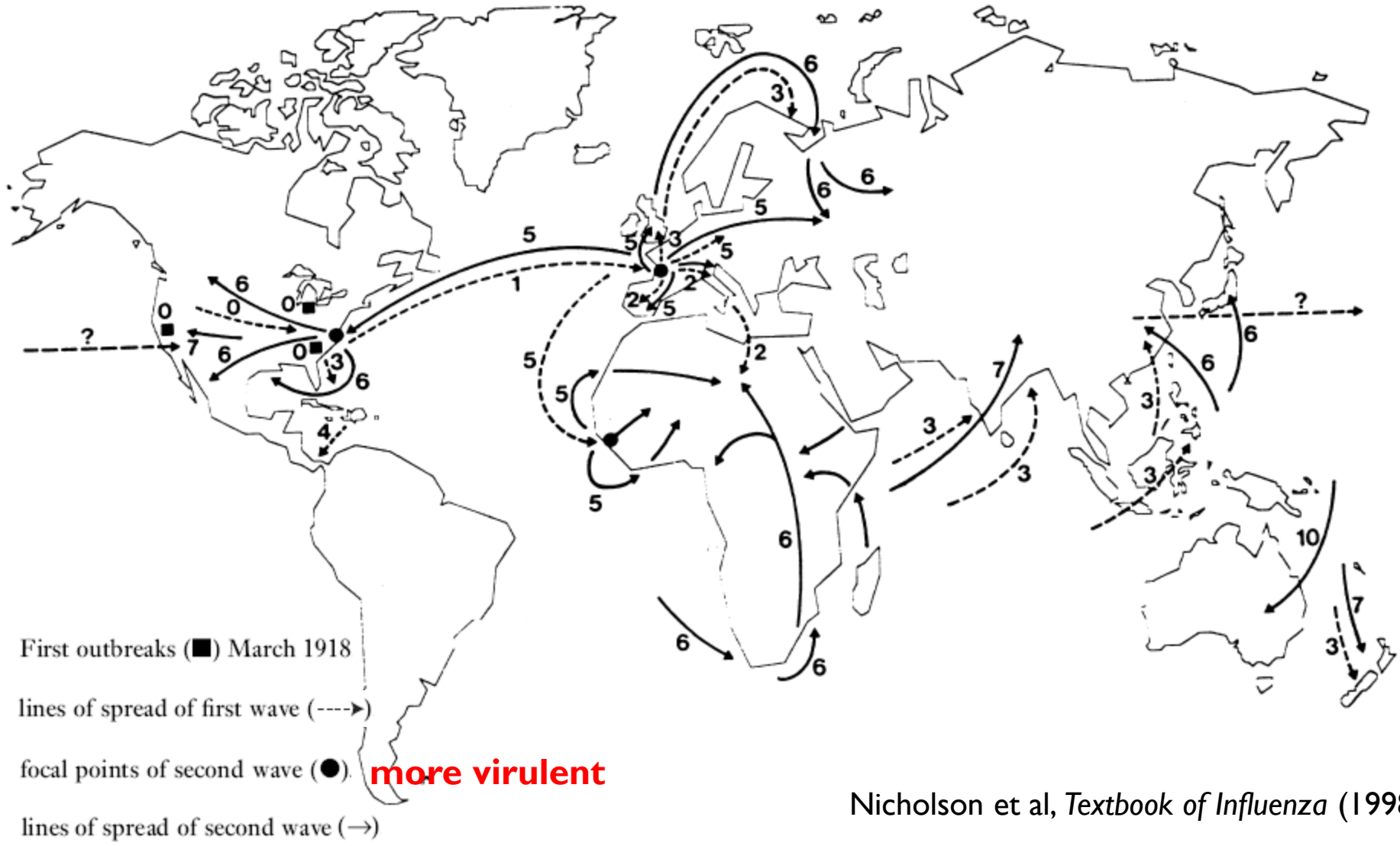
Even diseases with extremely low transmission probabilities are likely to cause a major outbreak involving a signficant fraction of population

The role of long-distance transportation networks

Influenza pandemics 1700-2000



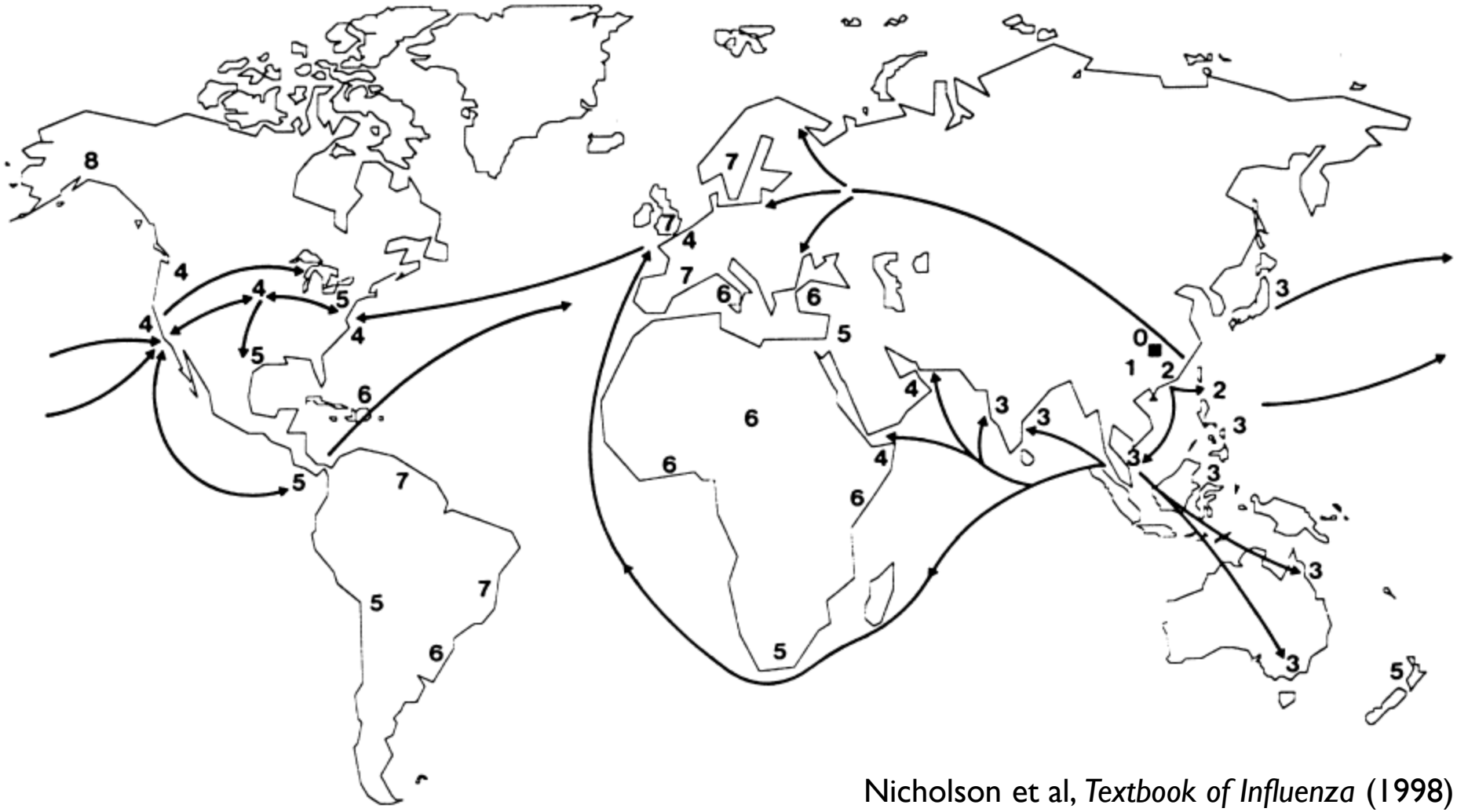
The influenza pandemic of 1918-1920



more virulent

Nicholson et al, *Textbook of Influenza* (1998)

The influenza pandemic of 1957-1958

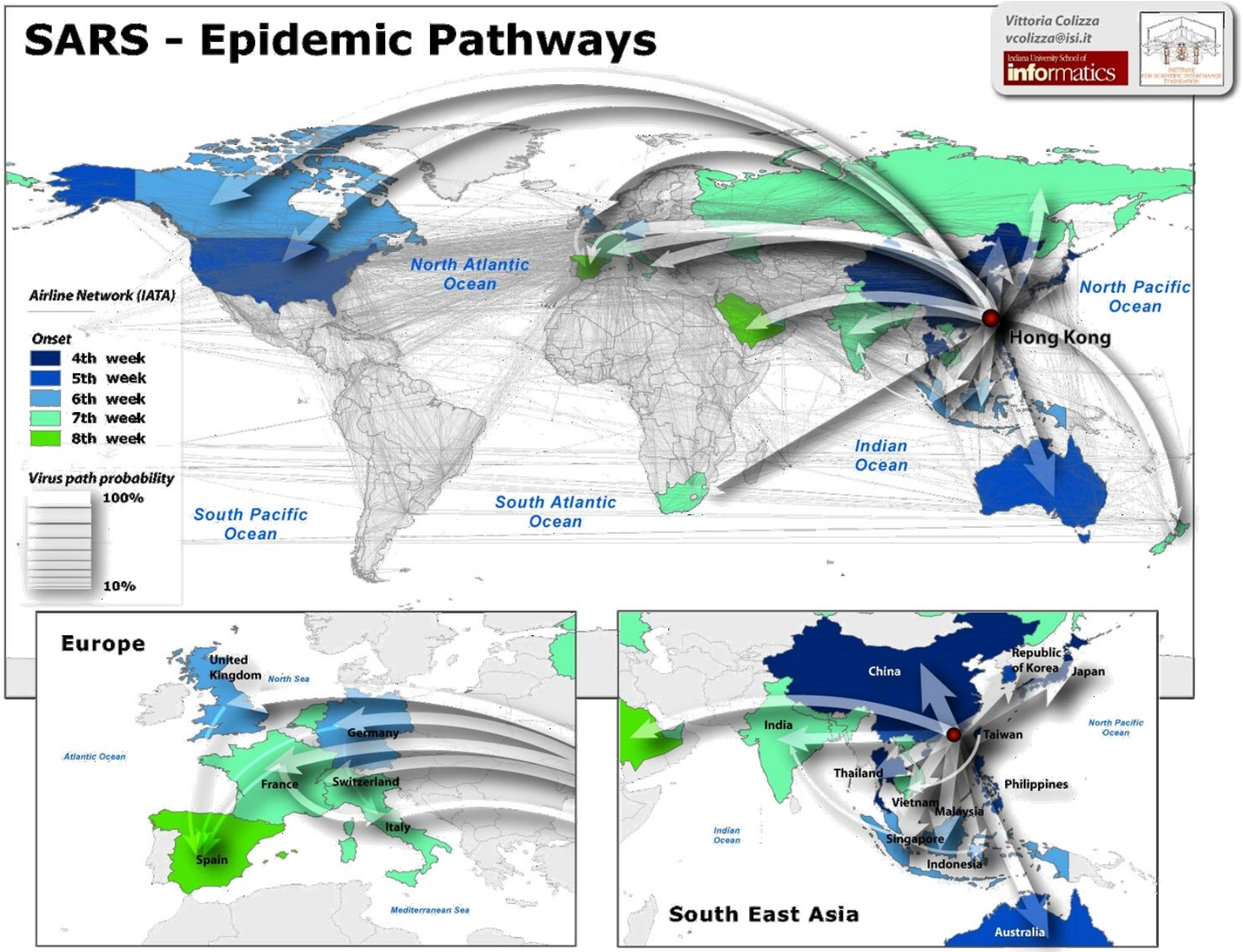


Nicholson et al, *Textbook of Influenza* (1998)

Point of origin (■) February 1957 lines of spread of pandemic (→)

number of months after February 1957 (0) when epidemic infection was recorded (number accompanies corresponding arrow)

Worldwide spread of SARS through international airline network



Epidemic spreading via Networks

Can be described by SIR model with non-local diffusion operator

Infected population in j-th community: $\partial I_j / \partial t = \beta S_j I_j - \gamma I_j + \Omega_j (\{ I_j \})$

We assume homogeneous mixing within a community

The network of interactions describing Ω could be
(depending on the scale of description one is interested in)

- The world-wide air-transportation network (the communities being the population of the urban agglomeration being served by an airport)

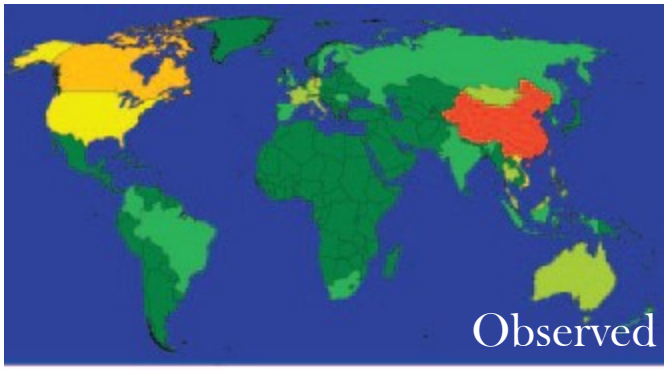
- The rail-road-air network of a country

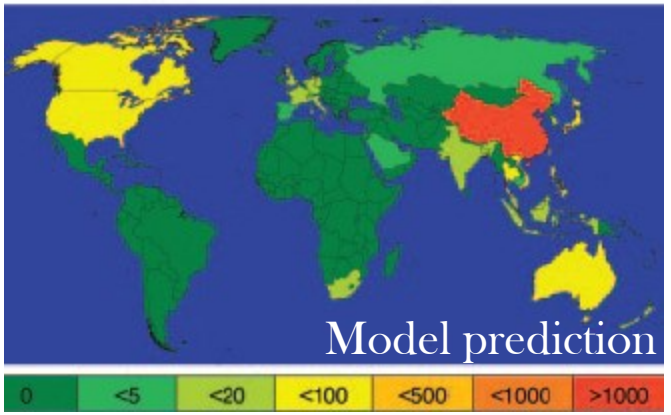
- The local transport within a district

The network may be strongly heterogeneous in terms of connection topology (e.g., the number of links for a given node may differ for a "hub" and a "leaf") , distribution of transport density along each connection and size of the different communities

Using the Global Aviation Network to forecast SARS spreading

Hufnagel et al, *PNAS* 101 (2004) 15124

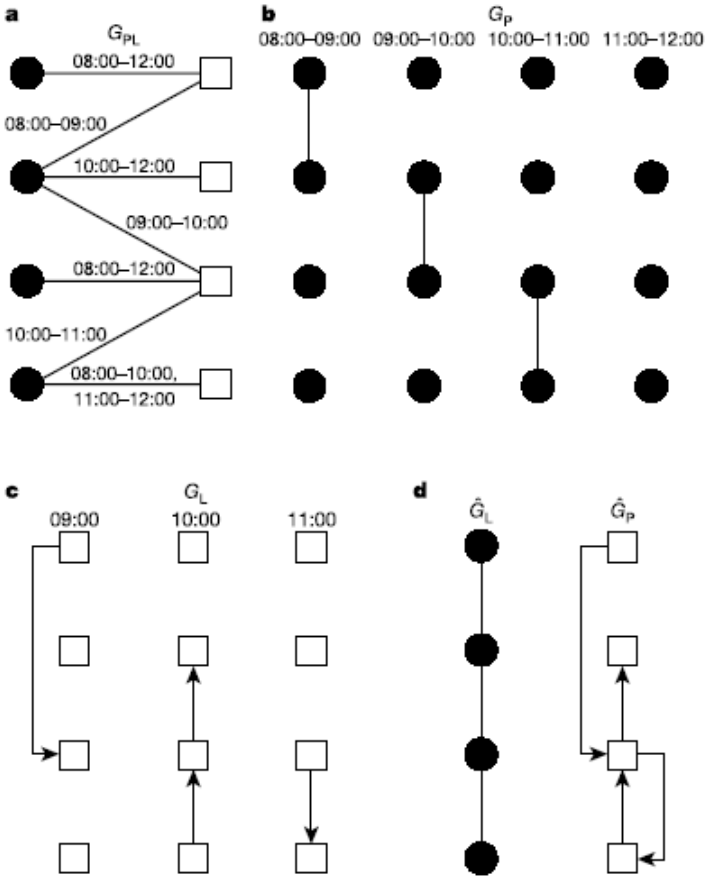




Model = local stochastic SIR infection dynamics + global transport along international civil aviation network, links weighted by passengers

The close match between predicted & observed spreading of SARS
⇒ model can be used to forecast how disease will spread if the initial infection begins elsewhere

On a smaller scale, we can fruitfully combine network approach with spatial or geographical information about location of agents/infection "hotspots" to understand the evolution of epidemics

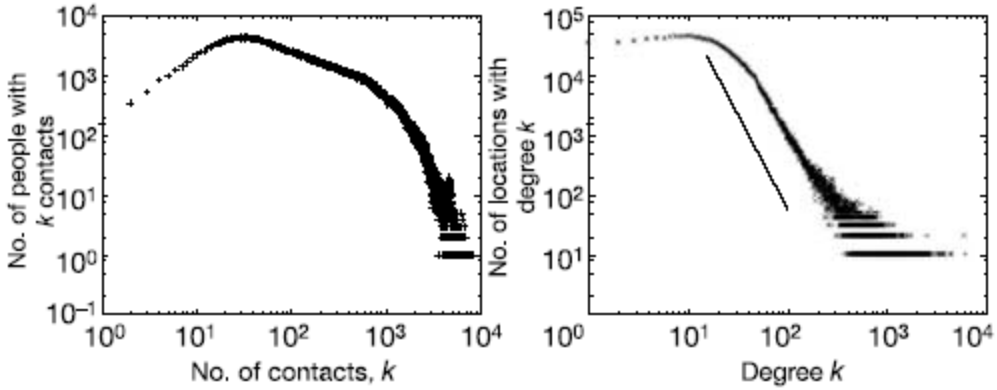


Eubank et al, *Nature* 429 (2004) 180

Eubank et al: Agent-based model (EPISIMS) to generate bipartite network of individuals and locations

The corresponding individual contact network is not scale-free degree distributed, but the location network is !

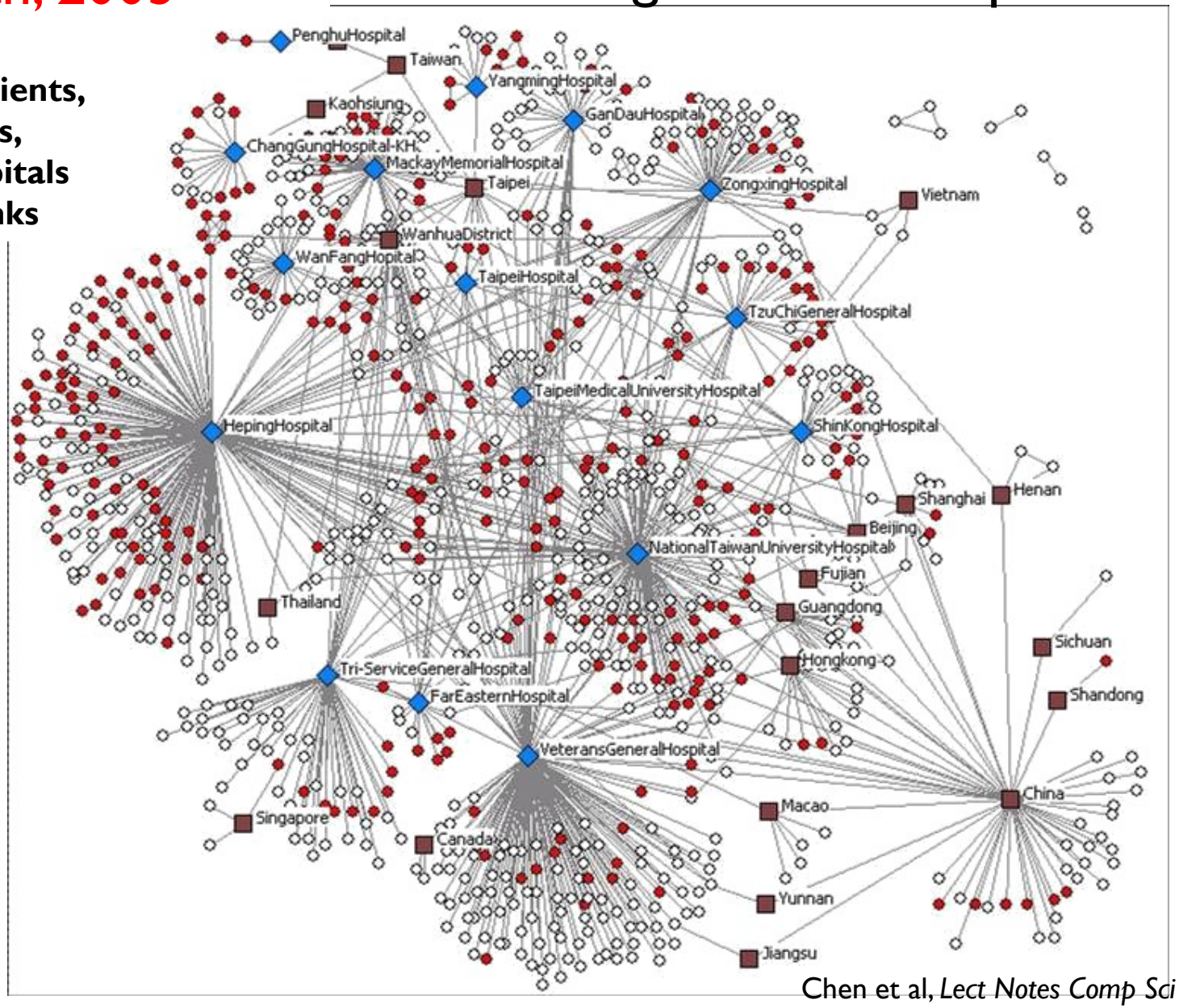
⇒ Importance of locations ("hotspots") in spreading infection



Spread of SARS in Taiwan, 2003

Contagion propagation in populations through contact in specific locations

961 patients,
22 areas,
14 hospitals
1313 links



Area

Hospital

Confirmed SARS patient

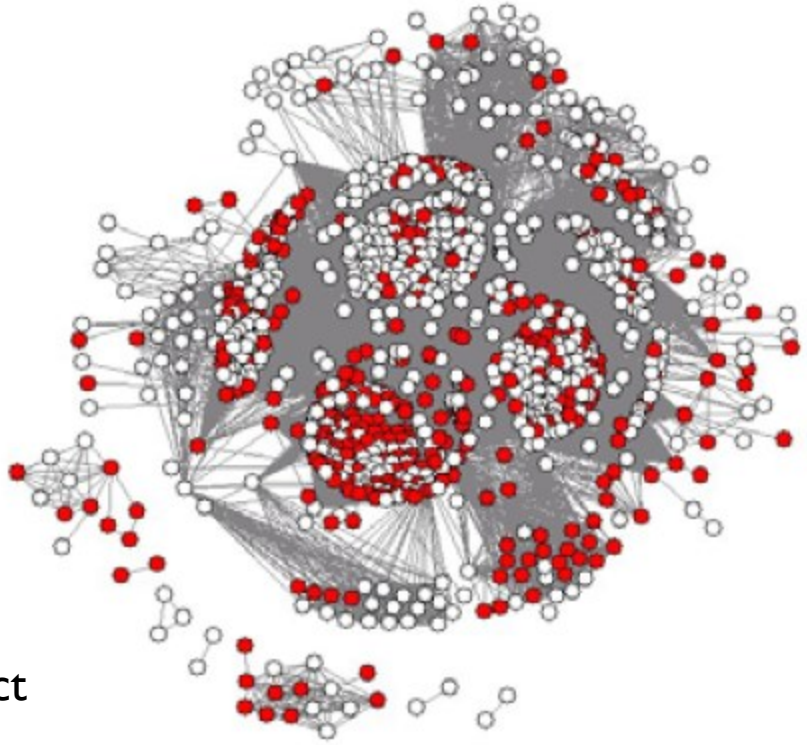
Suspected SARS patient

Chen et al, *Lect Notes Comp Sci* 4506 (2009) 23

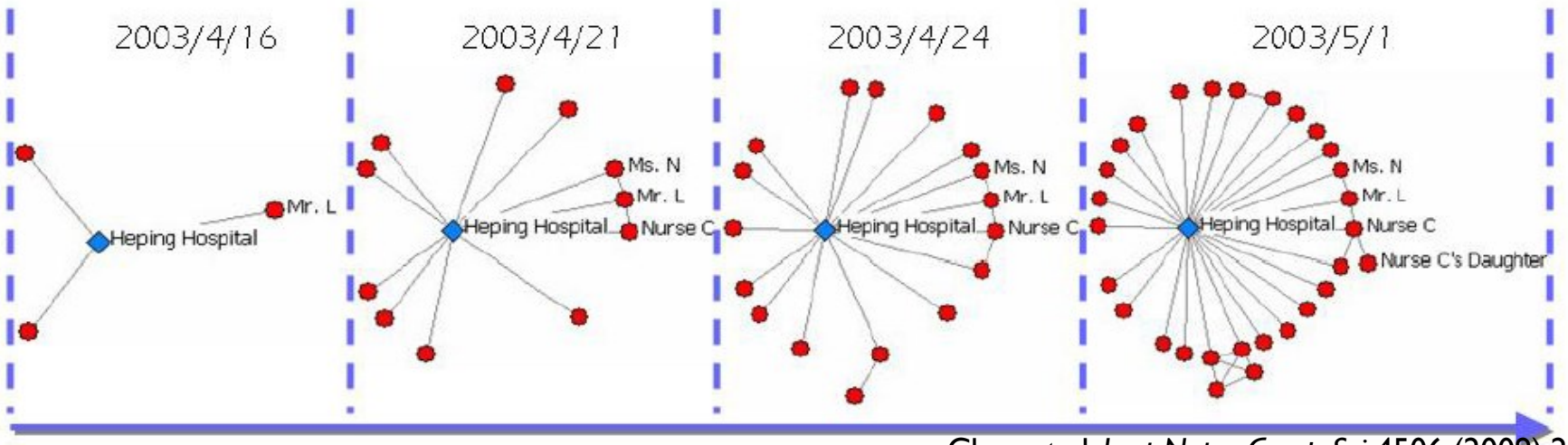
Contact network analysis put in a spatial context

Straightforward contact network analysis often looks only at the set of inter-personal contacts between infected agents

However, bi-partite graphs – comprising both agent nodes and location nodes – may explain better the evolution of disease along the contact network



Infection jumps from one to other at specific "high-risk" locations (*hub* location nodes)



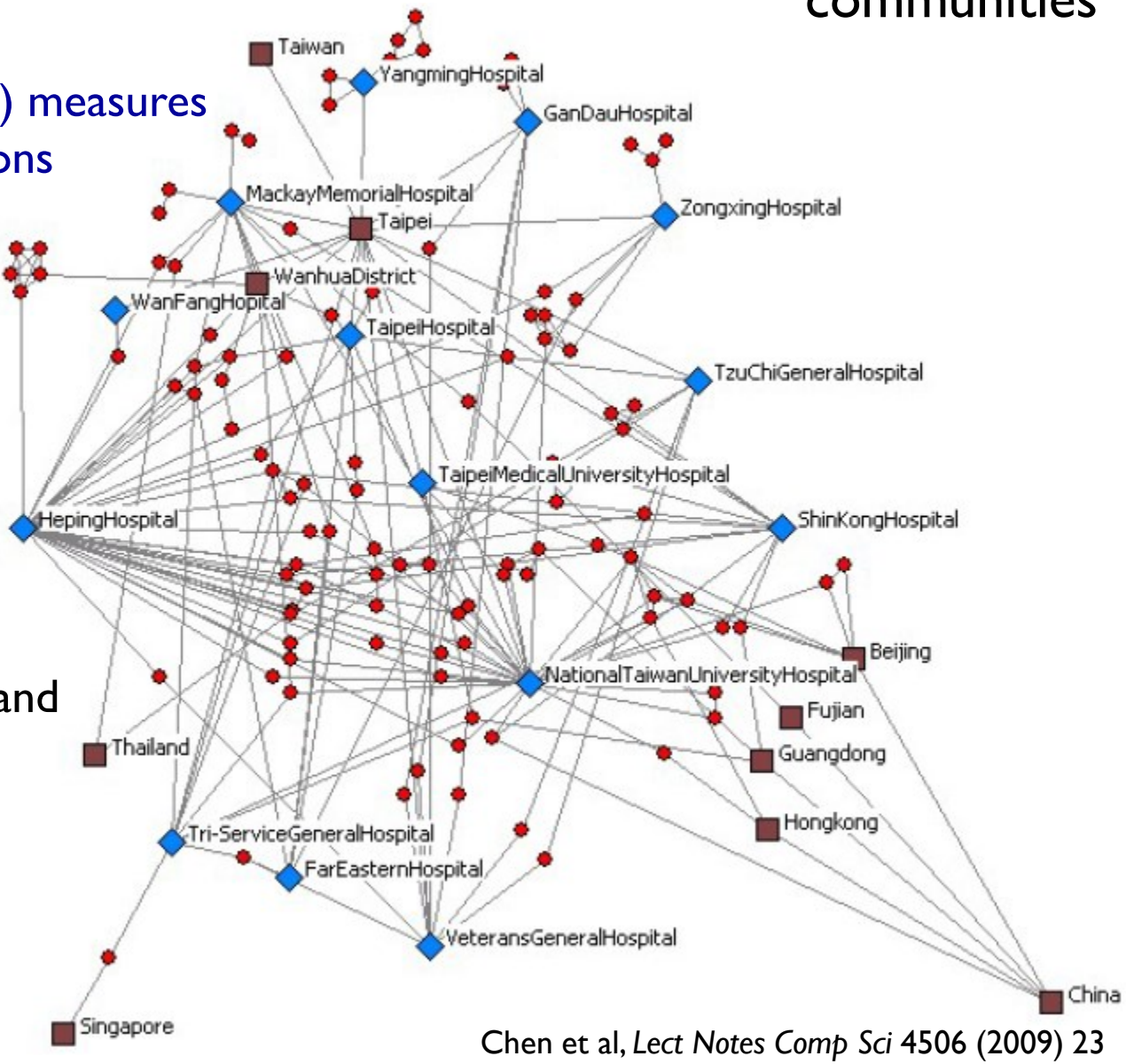
Chen et al, *Lect Notes Comp Sci* 4506 (2009) 23

Incorporating geographical / spatial information in network modeling helps reveal nodes acting as bridges for disease transfer between communities

Betweenness centrality (BC) measures can reveal individuals/locations acting as key bridges among households, hospitals and different regions in the contact network

Node BC and Edge BC provide complementary information on key agents and connections helping in spreading the disease

Can provide targets for control measures



Chen et al, *Lect Notes Comp Sci* 4506 (2009) 23