SPREADING PROCESSES ON NETWORKS

Carlo PICCARDI

DEIB - Department of Electronics, Information and Bioengineering
Politecnico di Milano, Italy

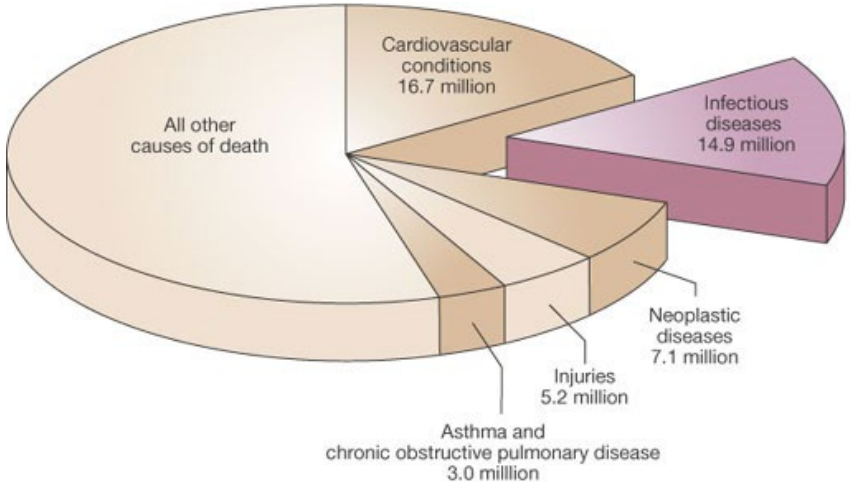
email carlo.piccardi@polimi.it
https://piccardi.faculty.polimi.it

Thanks to Renato Casagrandi (DEIB - Polimi) who prepared part of this material.

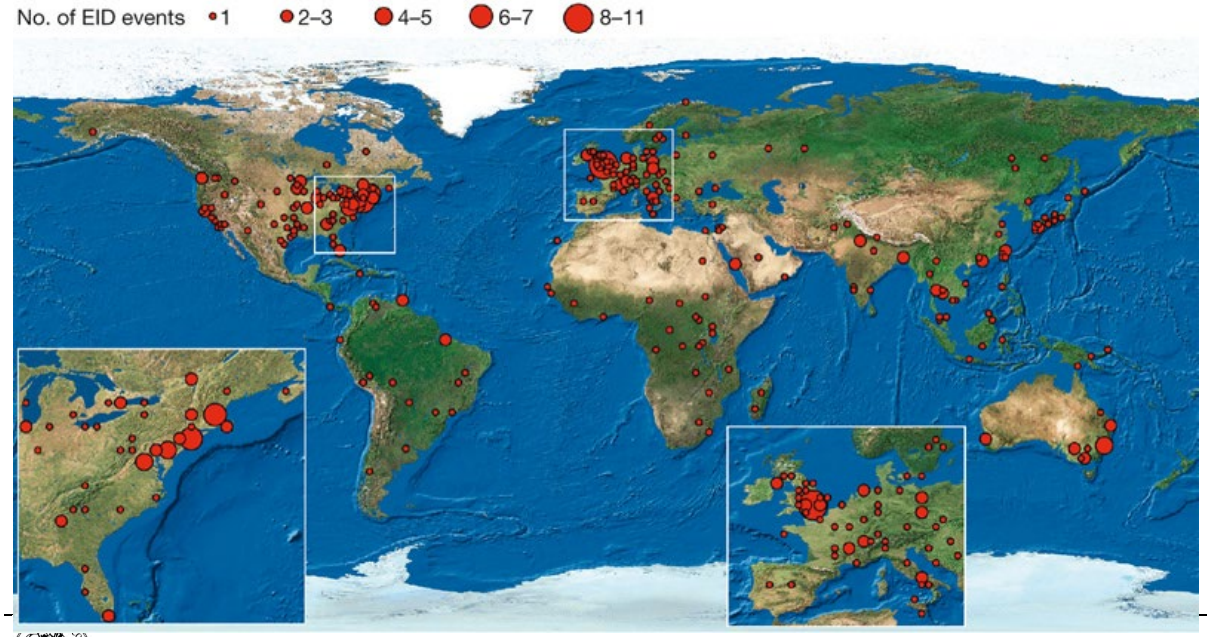
EPIDEMICS ON NETWORKS

The importance of infection diseases

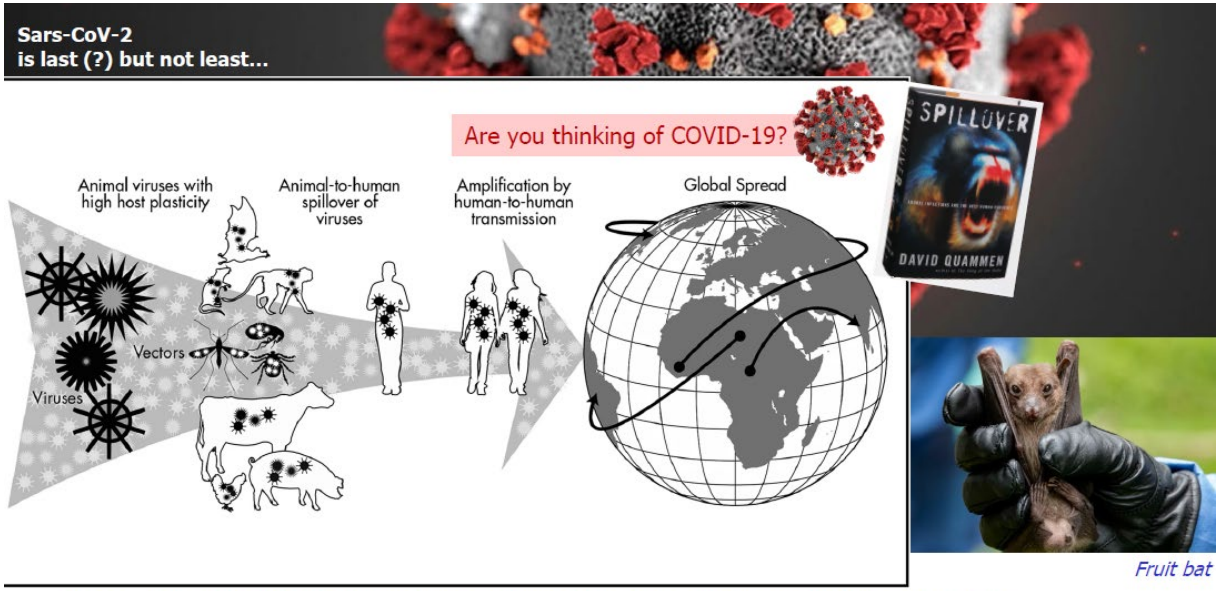
More than 25% of annual deaths worldwide are caused by infection diseases (Morens et al, Nature 2004)



Infectious diseases	Annual deaths (million)
Respiratory infections	3.96
HIV/AIDS	2.77
Diarrhoeal diseases	1.80
Tuberculosis	1.56
Vaccine-preventable childhood diseases	1.12
Malaria	1.27
STDs (other than HIV)	0.18
Meningitis	0.17
Hepatitis B and C	0.16
Tropical parasitic diseases	0.13
Dengue	0.02
Other infectious diseases	1.76



335 new diseases emerged from 1940 to 2004 (Jones et al, Nature 2008)



Fruit bat

Johnson, C.K. et al. (2015) Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Scientific Reports* **5**: 14830

3.3 By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases.

"Classical" modelling approach

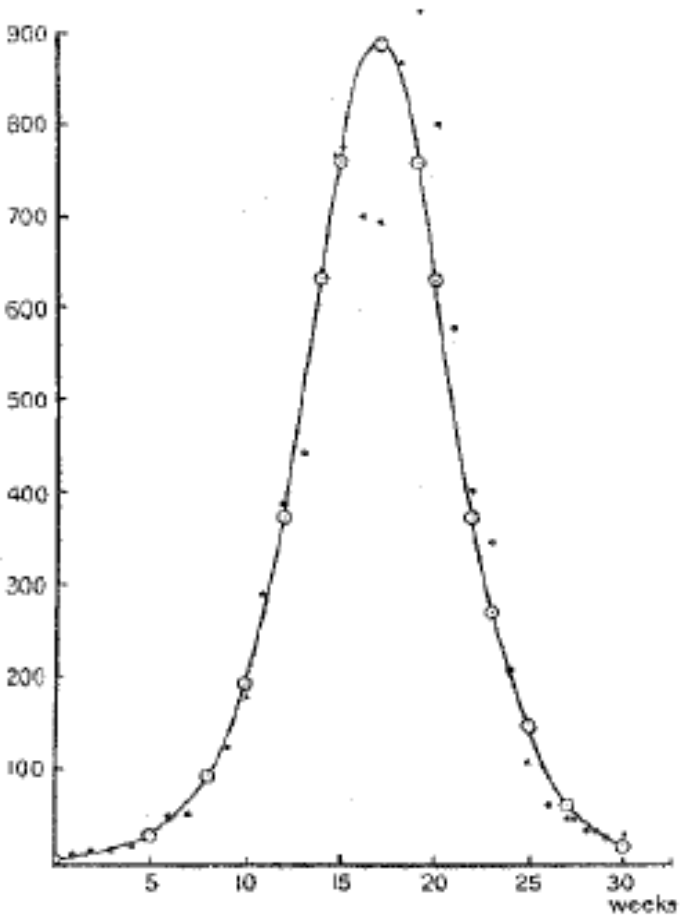
(Kermack and McKendrick, 1927)

Basic assumption:
Homogeneous mixing among individuals

$y(t)$ = *"the total number who are ill"* at time t (infectious, thus infectives)

$x(t)$ = *"denotes the number of individuals still unaffected"* (susceptibles)

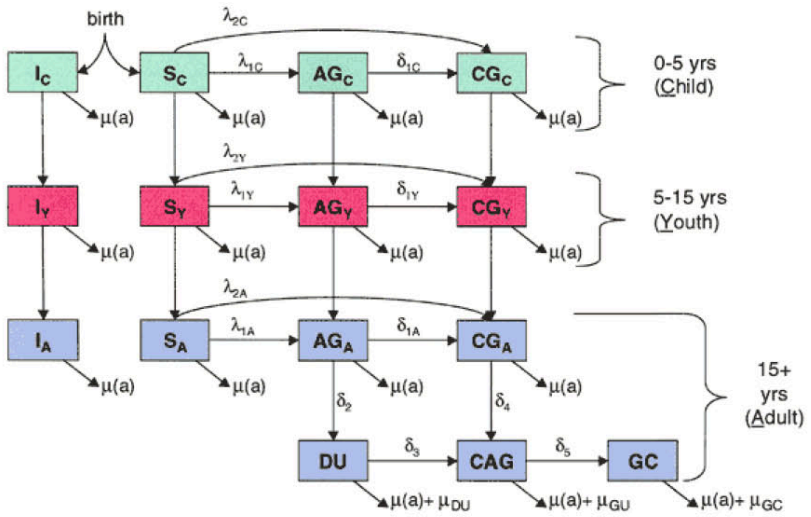
$z(t)$ = *"the number who have been removed by recovery and death"* (recovered)



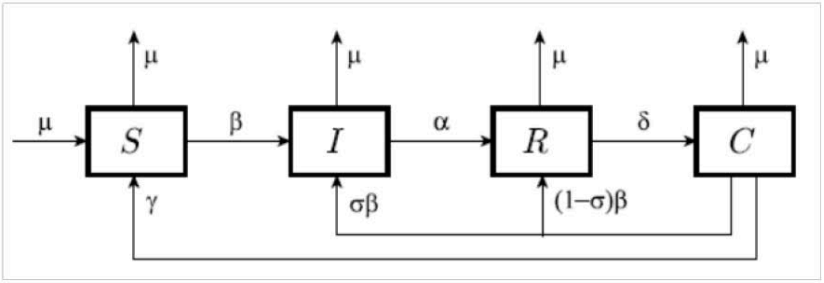
$$\left.\begin{aligned} \frac{dx}{dt} &= -\kappa xy \\ \frac{dy}{dt} &= \kappa xy - ly \\ \frac{dz}{dt} &= ly \end{aligned}\right\}$$

of infection.) This follows since the chance of an infection is proportional to the number of infected on the one hand, and to the number not yet infected on the other.

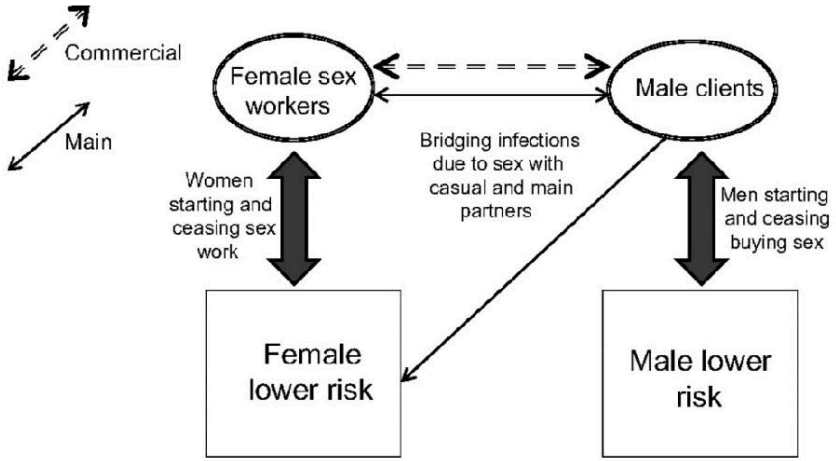
Generalized compartmental models



Rupnow et al (2000), *Emerg Infec Dis* **6**: 228



Casagrandi et al (2006), *Math Biosc* **200**: 152



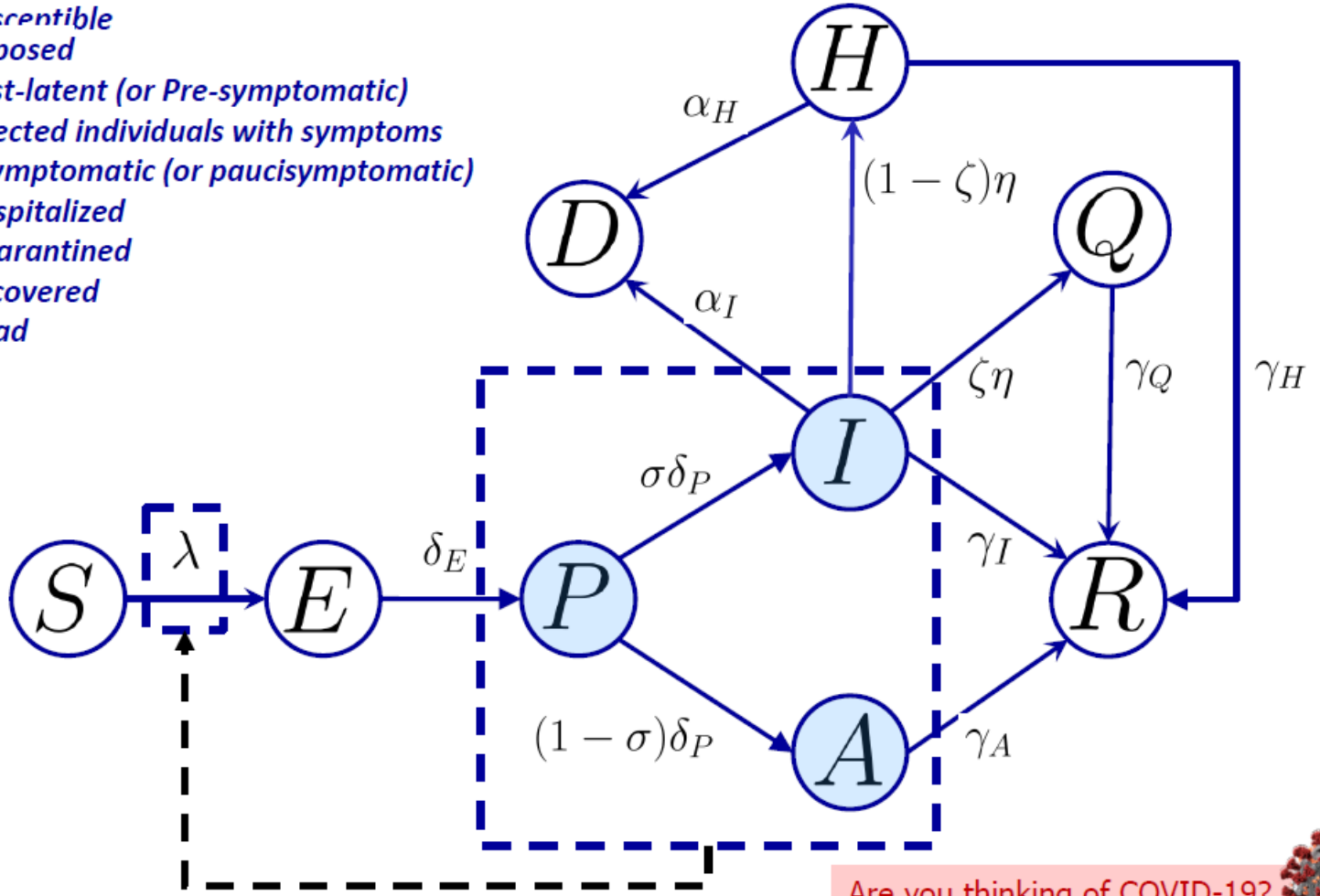
Pickles et al (2010), *Sex Transm Infect* **86**: i33

$$\frac{dS_1(t)}{dt} = \lambda_1 - \mu S_1(t) - B_1(t)S_1(t) + C_{21}S_2(t) - C_{12}S_1(t),$$
$$\frac{dI_1(t)}{dt} = B_1(t)S_1(t) - (\mu + \sigma + D_{12})I_1(t) + D_{21}I_2(t),$$
$$\frac{dS_2(t)}{dt} = \lambda_2 - \mu S_2(t) - B_2(t)S_2(t) + C_{12}S_1(t) - C_{21}S_2(t),$$
$$\frac{dI_2(t)}{dt} = B_2(t)S_2(t) - (\mu + \sigma + D_{21})I_2(t) + D_{12}I_1(t),$$
$$\frac{dA(t)}{dt} = \sigma(I_1(t) + I_2(t)) - (\mu + \gamma)A(t),$$

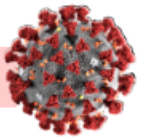
Greenhalgh et al (2001), *IMA* **18**: 225

The core SEPIA model for COVID-19

Susceptible
Exposed
Post-latent (or Pre-symptomatic)
Infected individuals with symptoms
Asymptomatic (or paucisymptomatic)
Hospitalized
Quarantined
Recovered
Dead



Are you thinking of COVID-19?



Gatto, Bertuzzo, Mari, Miccoli, Carraro, Casagrandi and Rinaldo (2020), *PNAS* **117:** 10484

The need for network approaches: the small-world effect

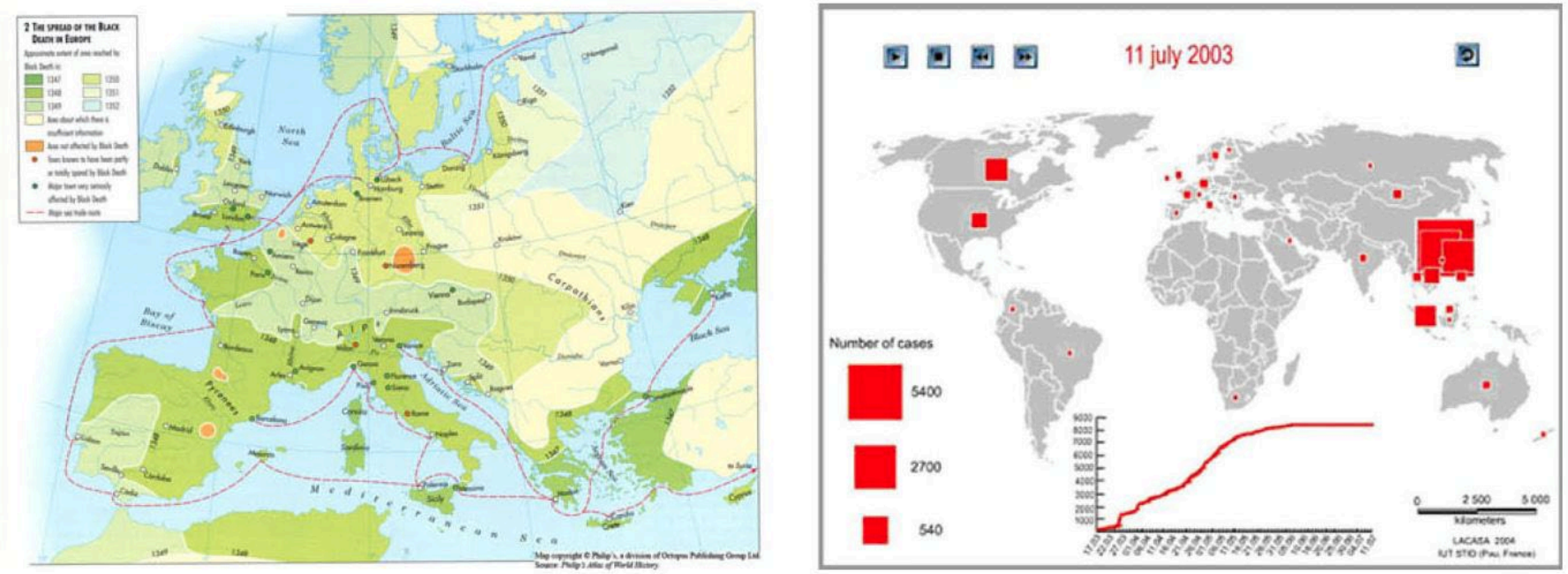
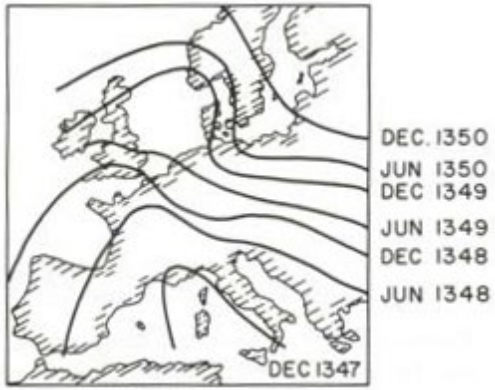


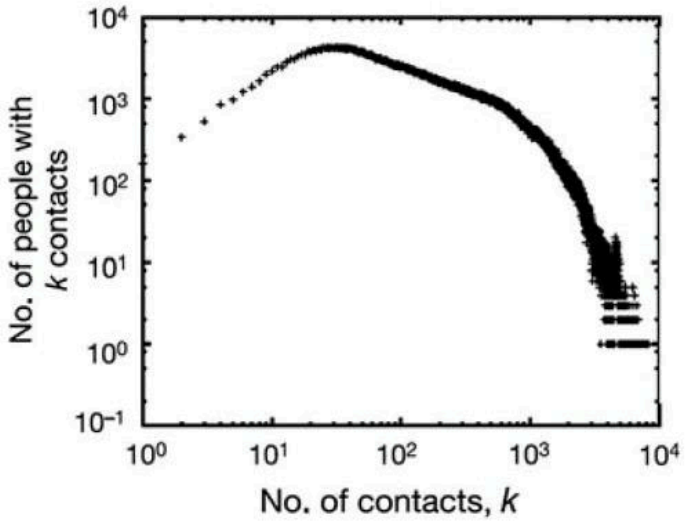
Fig. 3 *Left* the geographical spreading of the Black Death in the fourteenth century is described by a classical diffusive model [13, p. 105]; *right* in the case of the 2003 SARS epidemic, the epidemic spreads over large distances rapidly, due to air transport [3]

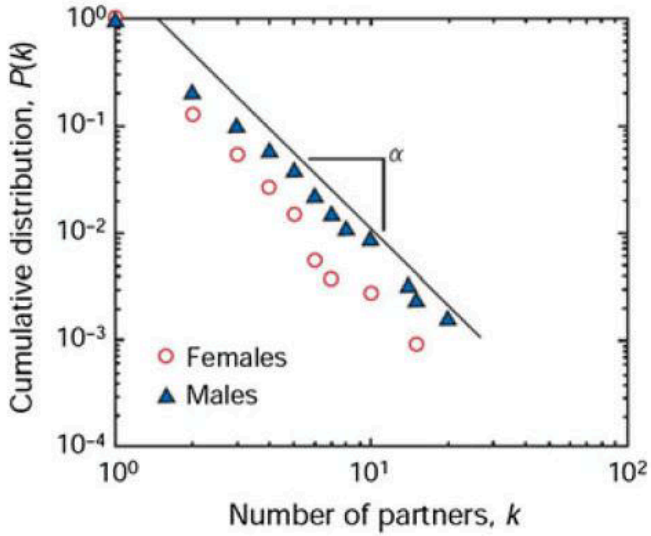


Covid-19 cases on March 29, 2020 [from mapbox.cn]

The need for network approaches: heterogeneous networks

Fig. 2 The number of proximity contacts in the residents of a big city (*left*, [6]) or the number of sexual partners in the previous 12 months declared by the individuals of a community (*right*, [8]) vary greatly, suggesting a strong heterogeneity in the connectivity of the individuals themselves





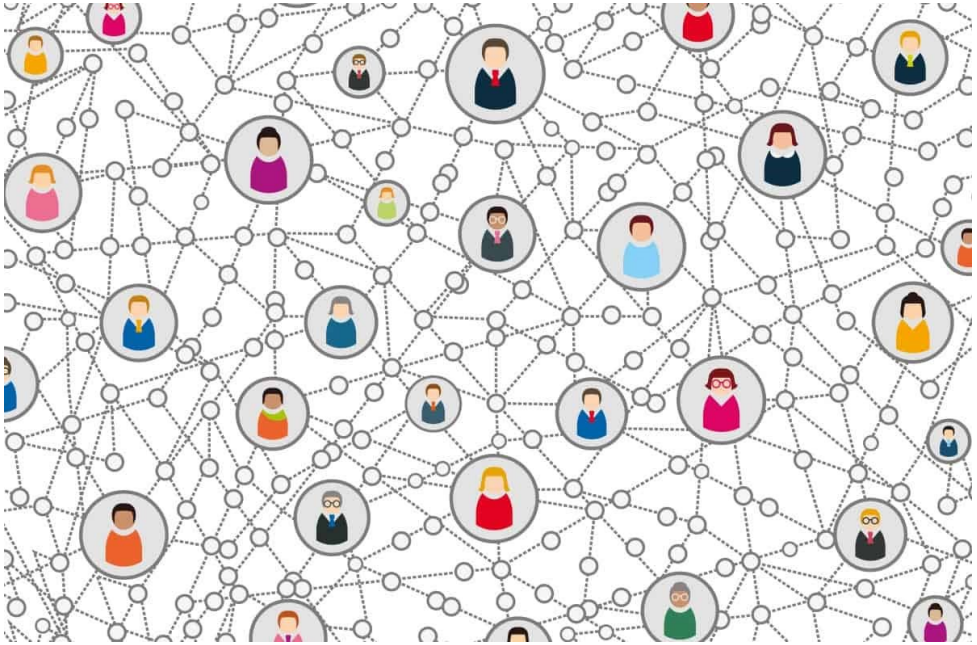
CONTAGION AND EPIDEMICS ON NETWORKS

Probabilistic cellular automata are used to model the spread of infectious diseases over the network.

- FINITE STATE SET: node (=individual) i is in state $s^i \in \Sigma = \{1,2,\ldots,\sigma\}$ at time t

  e.g.:
  Σ=*{Susceptible, Infected, Recovered}*
  in epidemics

- LOCAL RULES (=CONTAGION MECHANISM):
  the next state s^i_{t+1} depends (according to probabilistic rules) on s^i_t and on the state s^j_t of the neighbours



from physicsworld.org

Example: the SIS process

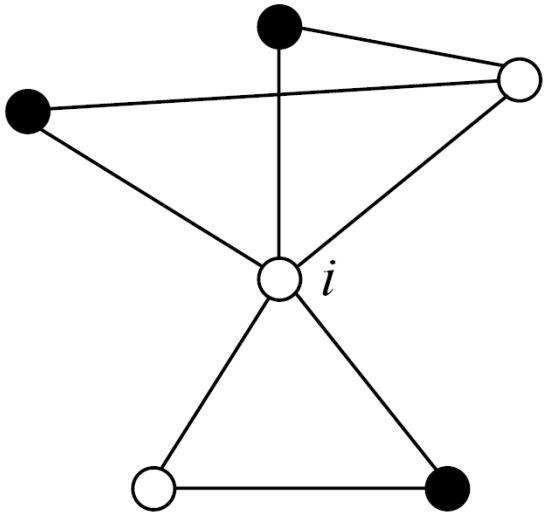
At time t, each node is

- susceptible (S) (= it is healthy but can potentially be infected), or
- infected (I) (= it is infected and capable of transmitting the infection)

LOCAL RULES:

At each time step (of length Δ):

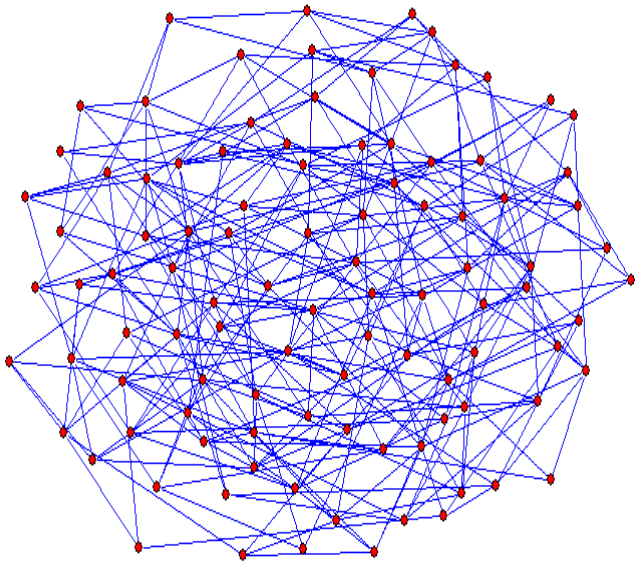
- infection: a node i in state S becomes I with probability $\beta I_i \Delta$, i.e. proportional to the number I_i of infected neighbors
- recovering: a node in state I returns S with probability $\gamma\Delta$

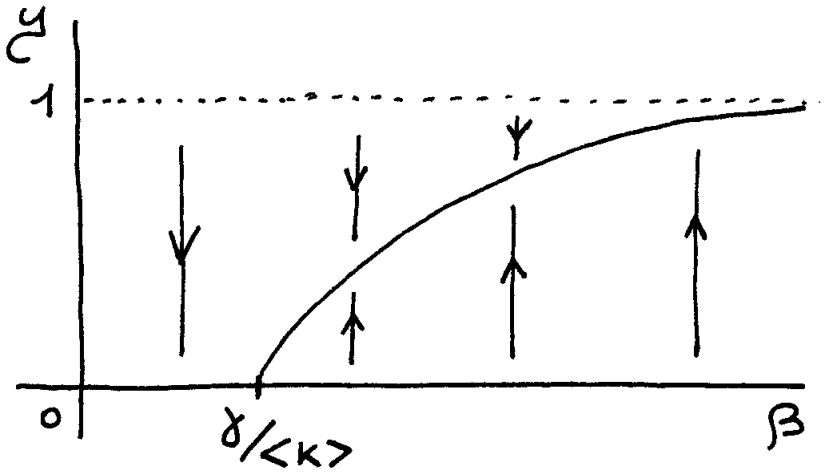


What is the global behaviour of the epidemics?

In a homogeneous (or almost homogeneous) network:

- if $\beta \leq \frac{\gamma}{< k >}$ $\Longrightarrow$ the fraction $y(t)$ of infected tends to 0 (=the epidemic dies out)
- if $\beta > \frac{\gamma}{< k >}$ $\Longrightarrow$ the fraction y of infected increases with the transmission rate β





This result is consistent with the classical epidemiology (Kermack and McKendrick, 1927):

No epidemic can survive if the transmission rate is below the epidemic threshold.

Some technical details...

Let $Y(t) \in [0, N]$ *be the number of I and* $y(t) = Y(t)/N \in [0,1]$ *their density (prevalence).*

$$Y(t+1) = Y(t) - \gamma \Delta Y(t) + \beta \Delta \Theta(t)(N - Y(t))$$

where $\Theta(t)$ *is the estimate of the average number of I among the neighbors of any S.*

Assuming $\Theta = < k > y(t)$ *(average n. of neighbors × prob. that a neighbor is I) we obtain (for* $\Delta \to 0$*) the classical SIS model:*

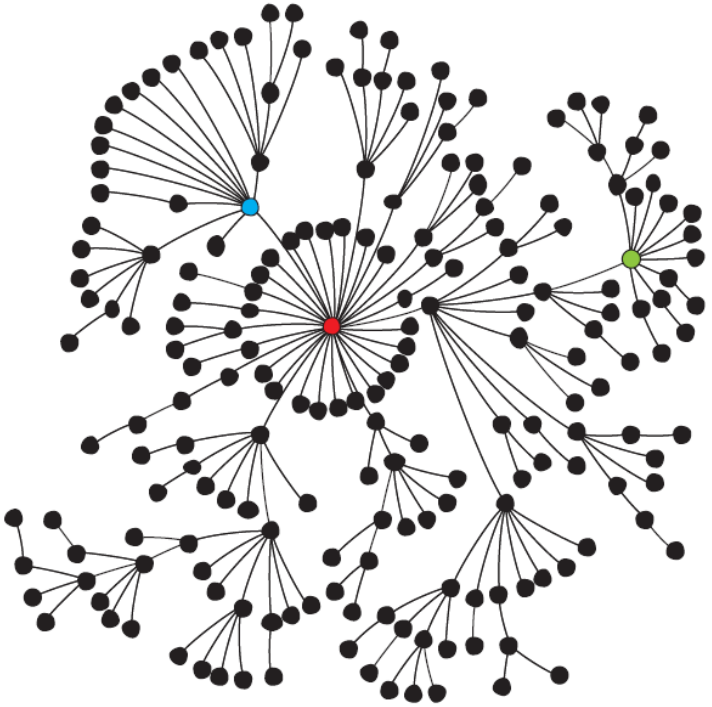
$$\dot{y}(t) = -\gamma y(t) + \beta < k > y(t)(1 - y(t))$$

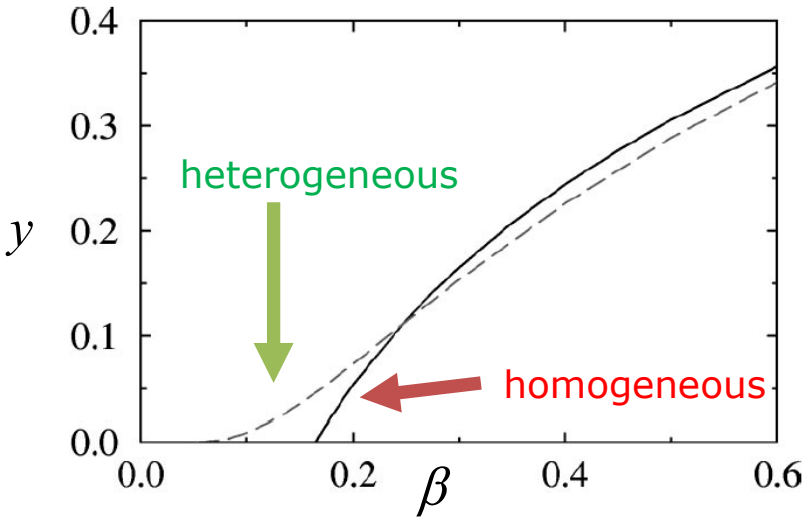
The non-trivial (> 0)*, asymptotically stable equilibrium* $y = 1 - \gamma / (\beta < k >)$ *exists iff* $\beta > \gamma / < k >$.

$\beta_c = \gamma / < k >$ *is the epidemic threshold.*

In a heterogeneous (e.g., scale-free) network (Pastor-Satorras and Vespignani, 2001):

- the epidemic threshold is $\beta_c = \gamma < k > / < k^2 >$, then it may tend to 0 for large networks ($N \to \infty$)
- $\Longrightarrow$ $y(t)$ never vanishes, whatever the value of the transmission rate β
- the nodes with larger degree are rare but have a large probability of being infected





The epidemic is able to survive with arbitrarily small transmission rate β (but with vanishing prevalence y).

Some technical details...

How can we model the epidemic dynamics when the network is strongly inhomogeneous?

We must model y *separately for each ensemble of nodes having the same degree* k*:*

$$\dot{y}_k(t) = -\gamma y_k(t) + \beta\Theta_k(t)(1 - y_k(t)) \quad , \quad k = k_{\min}, \ldots, k_{\max}$$

$\Theta_k(t)$ =*(n. of neighbors* $\times$ *prob. that a neighbor is I)*$= k\tilde{y}(t) = k(\Sigma_h hP(h)y_h(t))/ < k >.$

At the equilibrium ($\dot{y}_k = 0$*) we obtain*

$$y_k = \frac{\beta k\tilde{y}}{\gamma + \beta k\tilde{y}} = \frac{1}{1 + \gamma/(\beta k\tilde{y})}$$

Thus the prevalence y_k *grows with* k *and tends to 1 as* $k \to \infty$ *(=nodes with a very large number of connections are rare but, most likely, they are infected).*

The (global) prevalence is given by

$$y(t) = \Sigma_k P(k)y_k(t)$$

Immunization of homogeneous/heterogeneous networks

immunization = vaccination policy aimed at eradicating the epidemic ($y \to 0$)

g = fraction of vaccinated nodes

If nodes to be vaccinates are selected at random:

g = prob. that a randomly chosen susceptible is vaccinated

$$\beta\big(1-y(t)\big) \Longrightarrow \beta(1-g)\big(1-y(t)\big)$$

which is equivalent to replace $\beta \Longrightarrow \beta(1-g)$.

What is the vaccination fraction g_c above which the epidemic is eradicated?

Homogeneous networks (random vaccination)

no epidemic below the threshold $\beta < \beta_c = \frac{\gamma}{<k>} \Longrightarrow \beta(1-g) < \frac{\gamma}{<k>}$

$$g > g_c = 1 - \frac{\gamma}{\beta < k >}$$

If $g > g_c$ nodes are vaccinated at random, the epidemic is eradicated ($y \to 0$).

Heterogeneous networks (random vaccination)

no epidemic below the threshold $\beta < \beta_c = \frac{\gamma<k>}{<k^2>} \Longrightarrow \beta(1-g) < \frac{\gamma<k>}{<k^2>}$

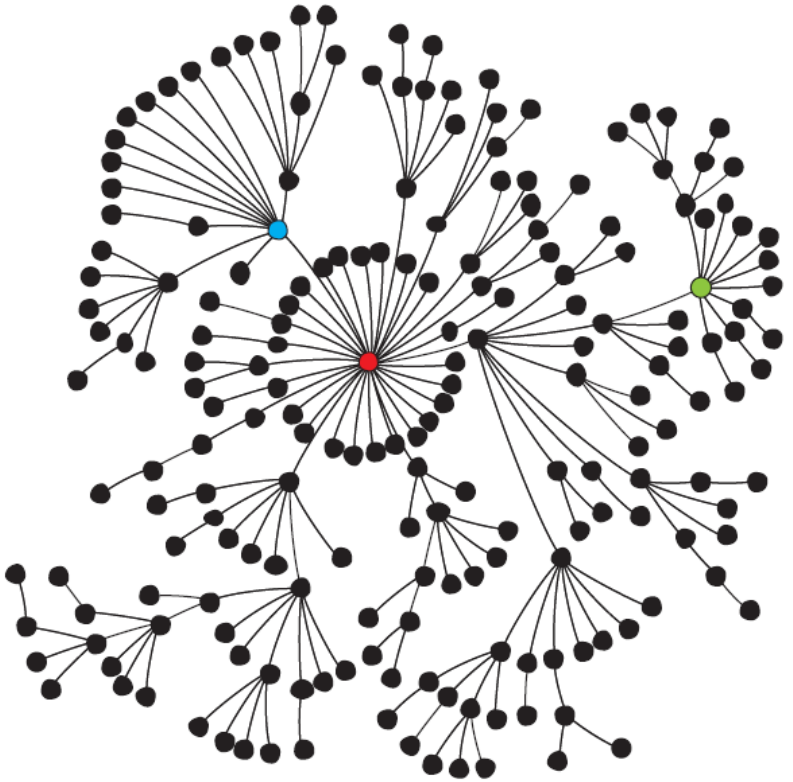
$$g > g_c = 1 - \frac{\gamma < k >}{\beta < k^2 >}$$

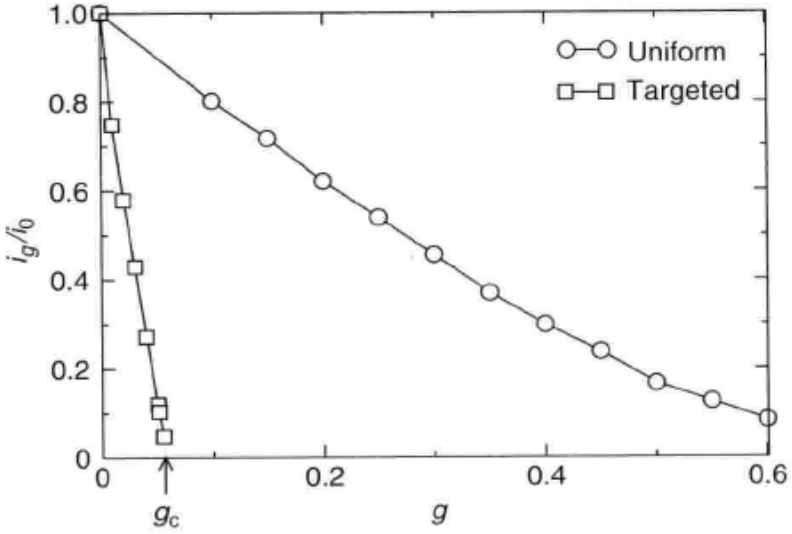
If $< k^2 > \to \infty$ (scale-free nets with large N) then $g_c \to 1$: need to vaccinate all population.

Heterogeneous networks (*targeted* vaccination)

The weakness of highly heterogeneous networks (low resilience to targeted attacks) becomes a defensive strategy against spreading processes.

Vaccinate a fraction g of nodes starting from those with highest degree.





In Barabási-Albert networks, immunization is obtained if

$$g > g_c \approx \exp\left(\frac{-2\gamma}{\beta k_{min}}\right)$$

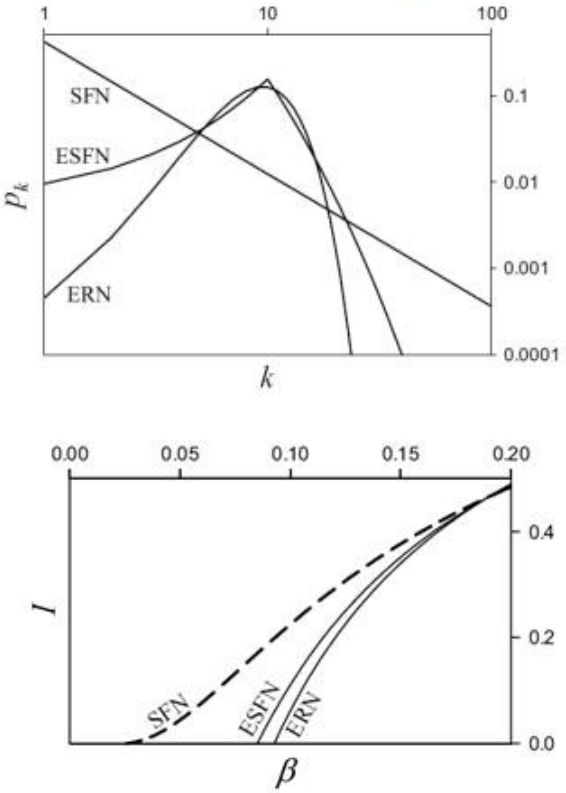
not depending on network heterogeneity.

Problem: how to identify the network (then the hubs) in a real-world population?

More complex models (just a glance…)

- Non-elementary contagion mechanisms (saturation, mean-field dependence, …)

- Birth/death/infection processes



Piccardi and Casagrandi (2009), Springer

- Epidemics on adaptive networks

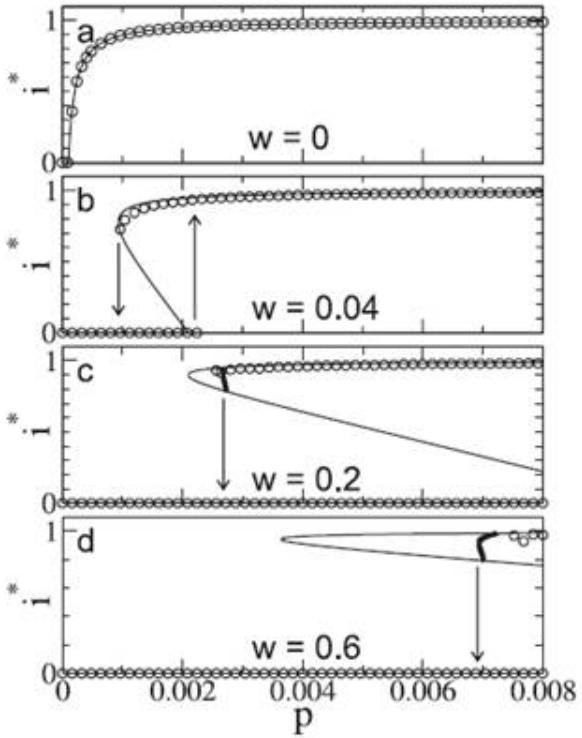


FIG. 3. Bifurcation diagram of the density of the infected i^* as a function of the infection probability p for different values of the rewiring rate w. In each diagram i^* has been computed analytically from Eqs. (2)–(4) (thin lines). Along the stable branches these results have been confirmed by the explicit

Gross et al (2006) *PRL* **96**:208701

Models of Influence Propagation ("Social Contagion")

They model the diffusion in a network-structured population of

- information (news, rumor, opinion, …)
- behavior (adoption of a product/service, food/smoking habits, …)

The main mechanism of propagation is "social contagion" (word-of-mouth).

Assumptions:

- two states Σ={Inactive , Active}
- progressive contagion {once turned to Active, a node remains such forever}

Problems:

- analysis: given the initial active set S_0 ("seed set"), find the final active set S_∞
- optimization: find the initial active set S_0, subject to the budget $|S_0| = m \ll N$, so to maximize the size $|S_\infty|$ of the final active set

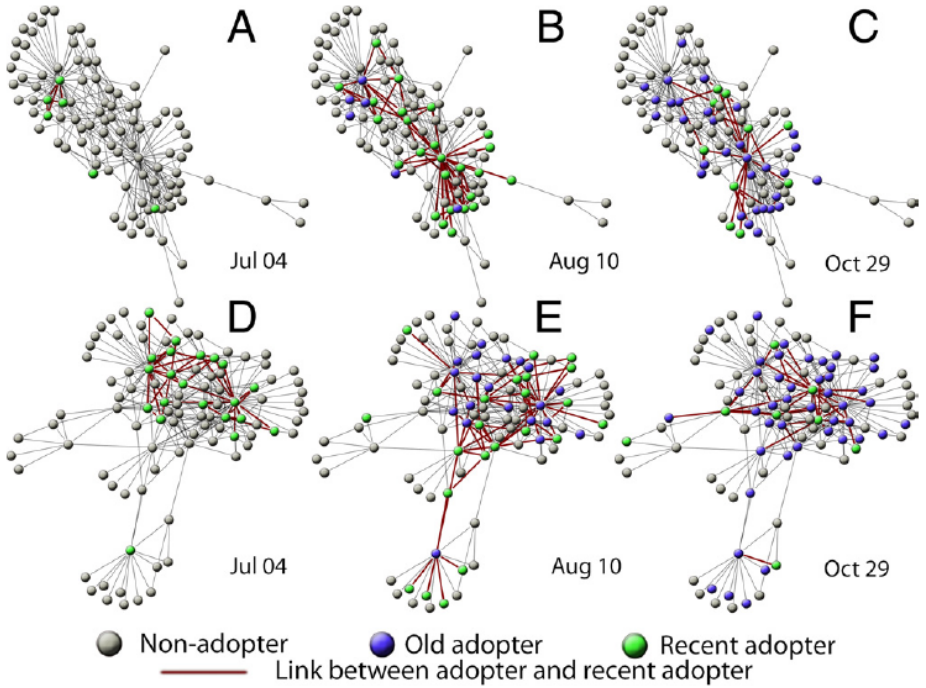


Fig. 1. Diffusion of Yahoo! Go over time. (*A–C* and *D–F*) Two subgraphs of the Yahoo! IM network colored by adoption states on July 4 (the Go launch date), August 10, and October 29, 2007. For animations of the diffusion of Yahoo! Go over time see Movies S1 and S2.

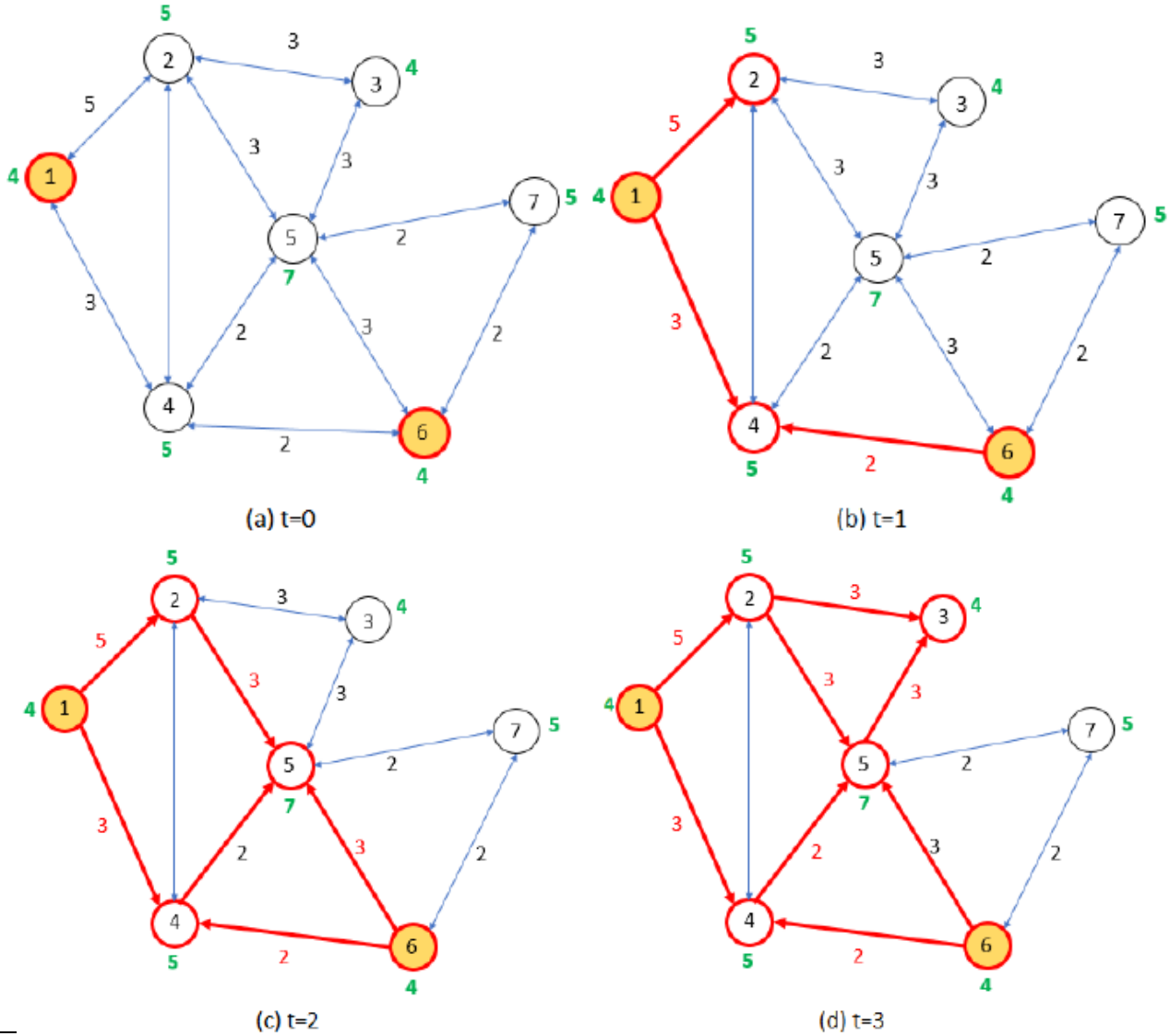
Linear Threshold (LT) model

A node becomes Active when the (weighted) count of Active neighbors exceeds a certain threshold.

- Each node i has a threshold $\theta_i > 0$ and is influenced by the incoming neighbors $N_i = \{j | a_{ji} = 1\}$ with a weight $w_{ji} > 0$.
- If A_t is the set of Active nodes at time t, then node i becomes active at time $t+1$ if

$$\sum_{j \in N_i \cap A_t} w_{ji} \geq \theta_i$$

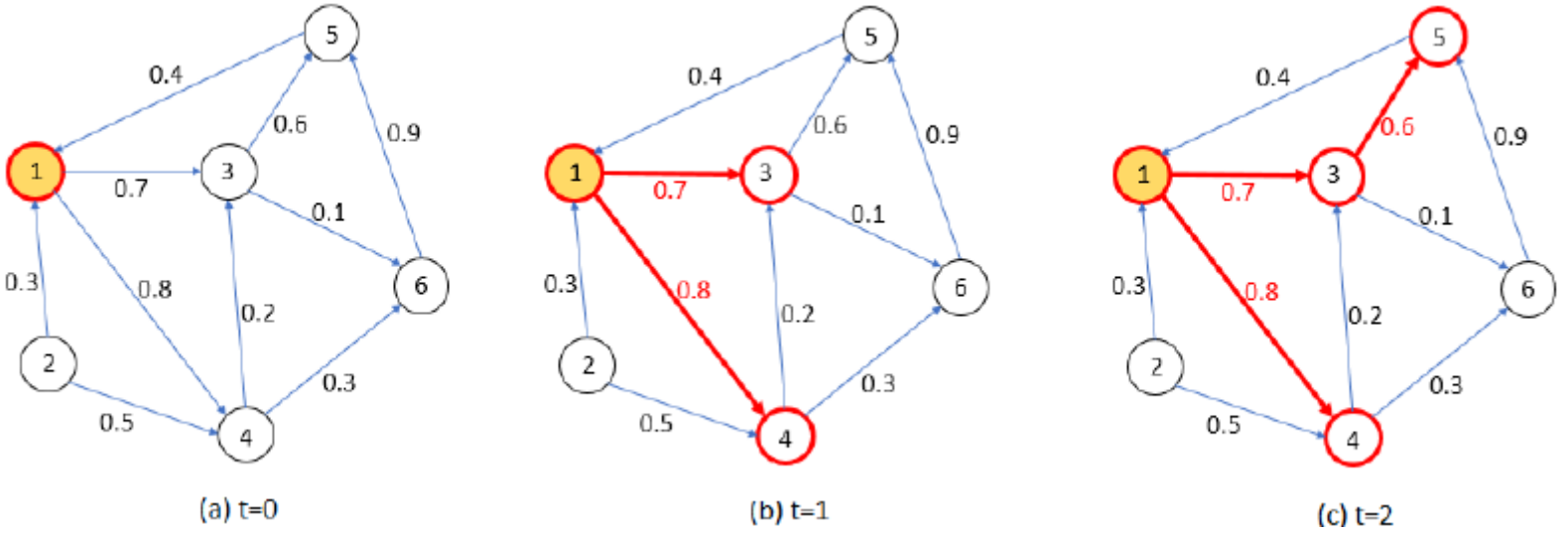
Notice: the model is deterministic.



Independent Cascade (IC) model

A node which becomes Active has exactly one chance to activate each Inactive neighbor.

- Each node i becoming active at time t is given a single chance to activate each of its currently inactive outgoing neighbors $N_i = \{j | a_{ij} = 1\}$.
- The activation attempt succeeds with probability p_{ij}, and j will become Active at time $t + 1$.



Notice: the model is stochastic.