Complex networks: an introduction

Alain Barrat

CPT, Marseille, France

ISI, Turin, Italy

http://www.cpt.univ-mrs.fr/~barrat
http://cxnets.googlepages.com

Plan of the lecture

I. INTRODUCTION
   I. Networks: definitions, statistical characterization
   II. Real world networks
II. DYNAMICAL PROCESSES
   I. Resilience, vulnerability
   II. Random walks
   III. Epidemic processes
   IV. (Social phenomena)
   V. Some perspectives

Random walks in complex networks

N nodes, W walkers

Node i => W_i walkers

$W = \sum_i W_i$

Diffusion rate out of i along (i,j): $d_{ij} = \frac{r}{k_i}$

=>Total escape rate out of i : r

Random walks in complex networks

Hypothesis= statistical equivalence of nodes with the same degree

=fundamental hyp. of **heterogeneous mean-field** approach

Degree block variables $$W_k = \frac{1}{N_k} \sum_{i|k_i=k} W_i$$

which evolve according to:

$$\partial_t W_k(t) = -rW_k(t) + k \sum_{k'} P(k'|k) \frac{r}{k'} W_{k'}(t)$$

Random walks in complex networks

$$\partial_t W_k(t) = -rW_k(t) + k\sum_{k'} P(k'|k)\frac{r}{k'}W_{k'}(t)$$

Uncorrelated random networks:

$$\partial_t W_k(t) = -rW_k(t) + \frac{k}{\langle k\rangle}\sum_{k'} P(k')rW_{k'}(t)$$

Stationarity =>

$$\boxed{W_k(t) = \frac{k}{\langle k\rangle}\frac{W}{N}}$$

Random walks in complex networks

An application: the PageRank algorithm

Definition:

$$\boxed{P_R(i) = \underbrace{\frac{q}{N}}_{\text{Random jumps}} + \underbrace{(1-q)\sum_j x_{ji}\frac{P_R(j)}{k_{out,j}}}_{\text{Random walk (on directed network)}}}$$

$P_R\,(i)$ = stationary probability of being on node i

PageRank: heterogeneous MF

$$P_R(i) = \frac{q}{N} + (1-q)\sum_j x_{ji}\frac{P_R(j)}{k_{out,j}}$$

Class of nodes of degree $\mathbf{k} \equiv (k_{in}, k_{out})$

$$\Rightarrow \boxed{P_R(\mathbf{k}) = \frac{1}{N_{\mathbf{k}}}\sum_{i\in\mathbf{k}} P_R(i)}$$

$$P_R(\mathbf{k}) = \frac{q}{N} + \frac{1-q}{N_{\mathbf{k}}}\sum_{i\in\mathbf{k}}\sum_{\mathbf{k}'}\frac{1}{k'_{out}}\sum_{j\in\mathbf{k}'} x_{ji}P_R(j)$$

PageRank: heterogeneous MF

$$\mathbf{k} \equiv (k_{in}, k_{out})$$

$$P_R(\mathbf{k}) = \frac{1}{N_{\mathbf{k}}} \sum_{i \in \mathbf{k}} P_R(i)$$

$$P_R(\mathbf{k}) = \frac{q}{N} + \frac{1-q}{N_{\mathbf{k}}} \sum_{i \in \mathbf{k}} \sum_{\mathbf{k}'} \frac{1}{k'_{out}} \sum_{j \in \mathbf{k}'} x_{ji} P_R(j)$$

Mean-Field approximation: $P_R(j) = P_R(\mathbf{k}')$

$$\Rightarrow \quad P_R(\mathbf{k}) = \frac{q}{N} + \frac{1-q}{N_{\mathbf{k}}} \sum_{\mathbf{k}'} \frac{P_R(\mathbf{k}')}{k'_{out}} \sum_{i \in \mathbf{k}} \sum_{j \in \mathbf{k}'} x_{ji}$$

$$\Rightarrow \quad P_R(\mathbf{k}) = \frac{q}{N} + \frac{1-q}{N_{\mathbf{k}}} \sum_{\mathbf{k}'} \frac{P_R(\mathbf{k}')}{k'_{out}} E_{\mathbf{k}' \to \mathbf{k}}$$

PageRank: heterogeneous MF

$$\mathbf{k} \equiv (k_{in}, k_{out})$$

$$P_R(\mathbf{k}) = \frac{q}{N} + \frac{1-q}{N_{\mathbf{k}}} \sum_{\mathbf{k}'} \frac{P_R(\mathbf{k}')}{k'_{out}} E_{\mathbf{k}' \to \mathbf{k}} \qquad P_R(\mathbf{k}) = \frac{1}{N_{\mathbf{k}}} \sum_{i \in \mathbf{k}} P_R(i)$$

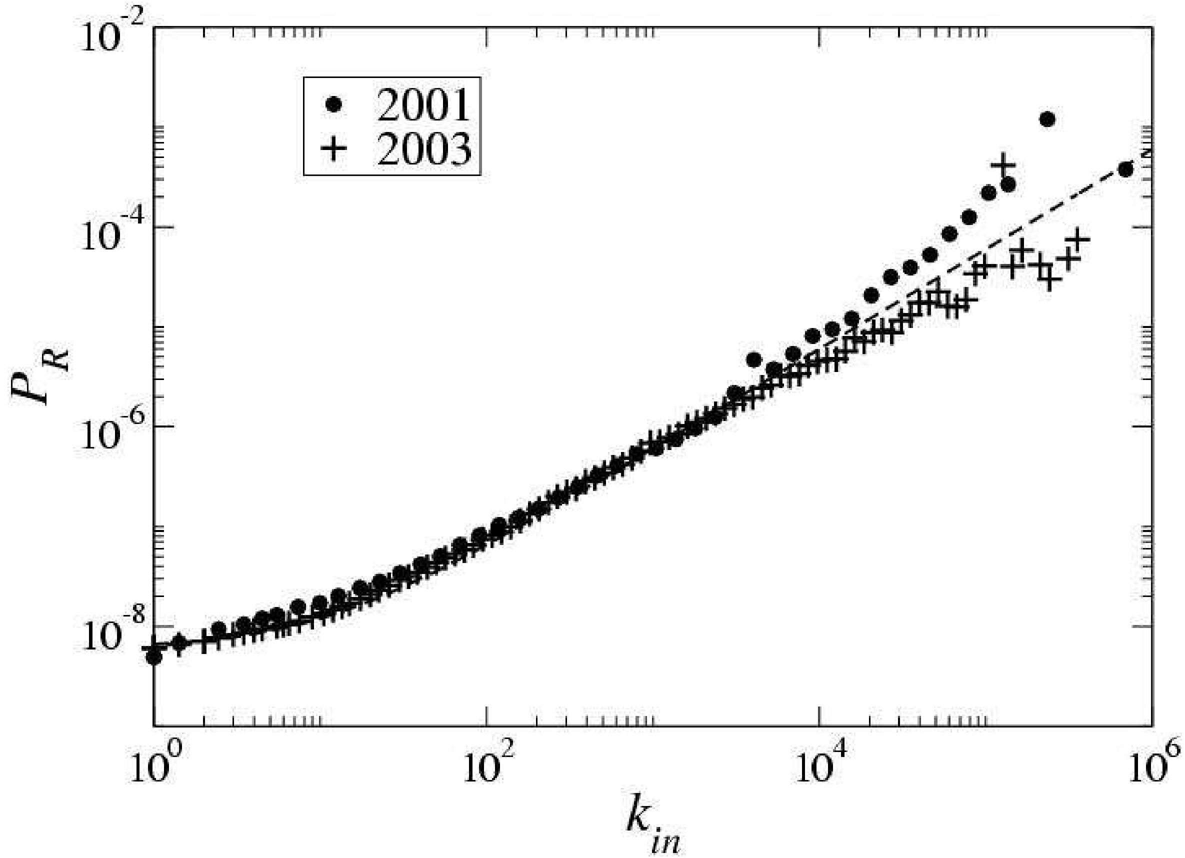
$$E_{\mathbf{k}' \to \mathbf{k}} = N k_{in} P(\mathbf{k}) P_{in}(\mathbf{k}'|\mathbf{k})$$

Uncorrelated networks: $P_{in}(\mathbf{k}'|\mathbf{k}) = \frac{k'_{out} P(\mathbf{k}')}{\langle k_{in} \rangle}$

$$P_R(\mathbf{k}) = \frac{q}{N} + (1-q) \frac{k_{in}}{\langle k_{in} \rangle} \sum_{\mathbf{k}'} P_R(\mathbf{k}') P(\mathbf{k}')$$

$$\boxed{P_R(\mathbf{k}) = \frac{q}{N} + \frac{1-q}{N} \frac{k_{in}}{\langle k_{in} \rangle}}$$

PageRank: heterogeneous MF



NB: MF result, important fluctuations are present

Random walks: mobility patterns

edges => weighted!

Ex: airport network: $w_{kk'} \sim w_0\,(k\,k')^{\theta}$, Traffic: $T_k \sim A\,k^{\theta+1}$

$$W_k = \frac{1}{N_k} \sum_{i|k_i=k} W_i$$

General evolution equation:

$$\partial_t W_k(t) = -r_k W_k(t) + k \sum_{k'} d_{k'k} P(k'|k) W_{k'}(t)$$

In uncorrelated networks:

$$\partial_t W_k(t) = -r_k W_k(t) + k \sum_{k'} d_{k'k} \frac{k' P(k')}{\langle k \rangle} W_{k'}(t)$$

Random walks: mobility patterns

$$\partial_t W_k(t) = -r_k W_k(t) + k \sum_{k'} d_{k'k} \frac{k' P(k')}{\langle k \rangle} W_{k'}(t)$$

First case: total escape independent from k:

$r_k = r$

a) Homogeneous diffusion $d_{k'k} = r/k'$

=> Recover previous case

b) Movements prop. to traffic intensities

$d_{k'k} = r\, w_0\, (kk')^{\theta} / T_{k'}$

$$\partial_t W_k(t) = -r W_k(t) + r k^{1+\theta} \frac{w_0}{A\langle k \rangle} \sum_{k'} P(k') W_{k'}(t)$$

Random walks: mobility patterns

$$\partial_t W_k(t) = -rW_k(t) + rk^{1+\theta}\frac{w_0}{A\langle k\rangle}\sum_{k'} P(k')W_{k'}(t)$$

Stationarity =>

$$\boxed{W_k(t) = \frac{k^{1+\theta}}{\langle k^{1+\theta}\rangle}\frac{W}{N}}$$

Random walks: mobility patterns

A different perspective:

We want:

W_i fixed

w_{ij} = number of travelers in a unit time, $w_{ij}=w_{ji}$

Each individual in subpopulation i has a diffusion rate $\sum_j w_{ij}/W_i$

$$\partial_t W_i = \sum_j W_j (w_{ji}/W_j) - W_i \sum_j (w_{ij}/W_i) = 0$$

Any population distribution is stationary

Random walks: mobility patterns

Or, in the degree block approximation:

$r_k = T_k / W_k$

$d_{k'k} = w_0\,(kk')^{\theta} / \boldsymbol{W_k},$

$$\partial_t W_k(t) = -r_k W_k(t) + k \sum_{k'} d_{k'k} \frac{k' P(k')}{\langle k \rangle} W_{k'}(t)$$

$$\partial_t W_k(t) = -T_k(t) + k^{1+\theta} w_0 \frac{\langle k^{1+\theta} \rangle}{\langle k \rangle} = 0$$

Plan of the lecture

I. INTRODUCTION
   I. Networks: definitions, statistical characterization
   II. Real world networks

II. DYNAMICAL PROCESSES
   I. Resilience, vulnerability
   II. Random walks
   III. Epidemic processes
   IV. (Social phenomena)
   V. Some perspectives

- Basic Epidemiology
- One population: network of contacts
  - ***effect of heterogeneous topology***
- Metapopulation models: transportation network
  - ***propagation pattern?***
  - ***epidemic forecasting?***
  - ***applications***

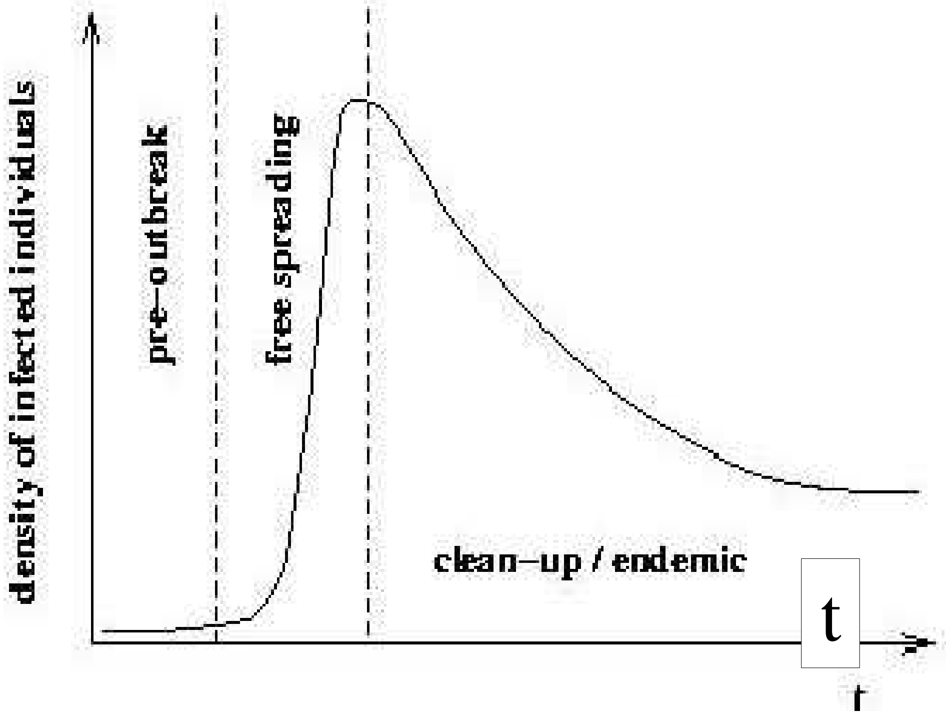
Epidemiology

Two levels:

•Microscopic: researchers try to disassemble and kill new viruses => quest for vaccines and medicines

•Macroscopic: statistical analysis and modeling of epidemiological data in order to find information and policies aimed at lowering epidemic outbreaks => macroscopic prophylaxis, vaccination campaigns...

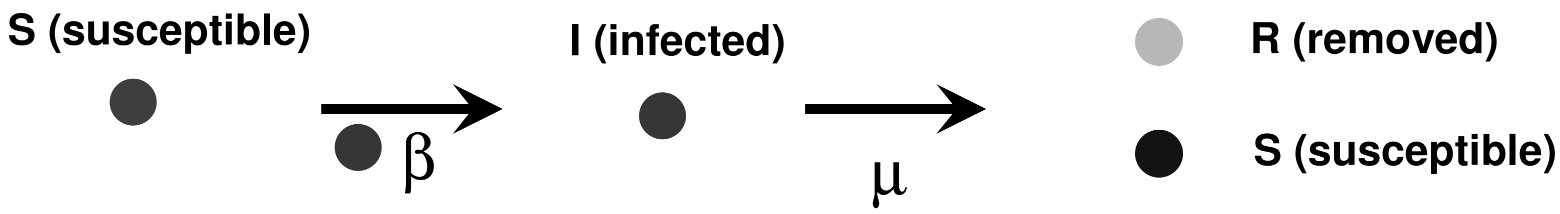
Stages of an epidemic outbreak



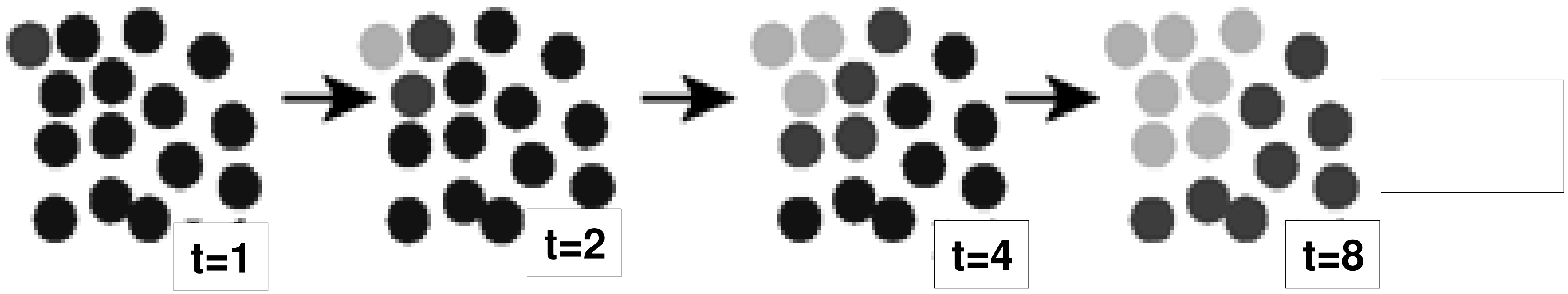
Infected individuals => prevalence/incidence

Standard epidemic modeling

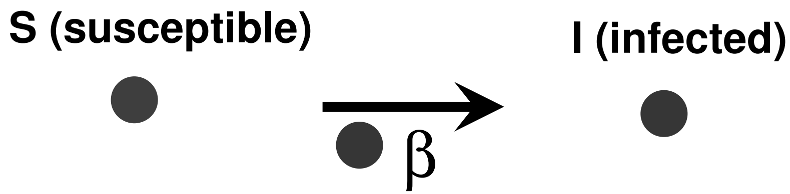
<u>Compartments:</u> S, I, R...



Homogeneous mixing assumption (Mean-field)



The SI model



N individuals

$I(t)$=number of infectious, $S(t)=N-I(t)$ number of susceptible

$i(t)=I(t)/N$, $s(t)=S(t)/N$

Individual with k contacts, among which n infectious:

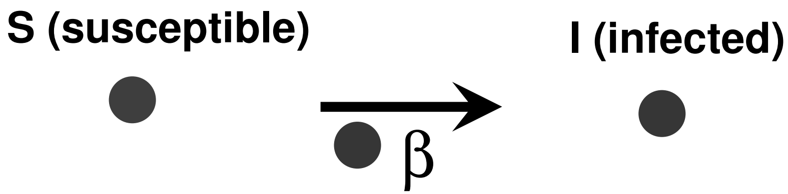
in the **homogeneous mixing approximation**, the probability to get the infection in each time interval dt is:

$$1 - (1-\beta dt)^n \approx \beta \, k \, i \, dt \quad (\beta \, dt << 1)$$

If k is the same for all individuals (homogeneous network):

$$\frac{di}{dt} = \beta \langle k \rangle i(1-i)$$

The SI model



N individuals

$I(t)$=number of infectious, $S(t)=N-I(t)$ number of susceptible

$i(t)=I(t)/N$, $s(t)=S(t)/N$

$$\frac{di}{dt} = \beta \langle k \rangle i(1-i)$$

$$i(t) = \frac{i_0 \exp(t/\tau)}{1 + i_0(\exp(t/\tau) - 1)} \qquad \tau = 1/(\beta\langle k \rangle)$$

The SIS model

N individuals

$I(t)$=number of infectious, $S(t)=N-I(t)$ number of susceptible

$i(t)=I(t)/N$, $s(t)=S(t)/N$

Homogeneous mixing

$$\frac{di}{dt} = \beta\langle k\rangle i(1-i) - \mu i$$

Competition of two time scales

The SIR model

N individuals

I(t)=number of infectious, *S(t)* number of susceptible, *R(t)* recovered

i(t)=I(t)/N , s(t)=S(t)/N, r(t)=R(t)/N=1-i(t)-s(t)

Homogeneous mixing:

$$\frac{ds}{dt} = -\beta\langle k\rangle i(t)s(t)$$

$$\frac{di}{dt} = \beta\langle k\rangle i(t)s(t) - \mu i(t)$$

$$\frac{dr}{dt} = \mu i(t)$$

SIS and SIR models: linear approximation

Short times, $i(t) << 1$ (and $r(t) << 1$ for the SIR)

$$\frac{di}{dt} \approx (\beta\langle k\rangle - \mu)i(t)$$

Exponential evolution $exp(t/\tau)$, with

$$\boxed{1/\tau = \beta\langle k\rangle - \mu}$$

If $\beta <k> > \mu$: exponential growth

If $\beta <k> < \mu$: extinction

$$\boxed{\text{Epidemic threshold condition: } \beta <k> = \mu}$$

Long time limit, SIS model

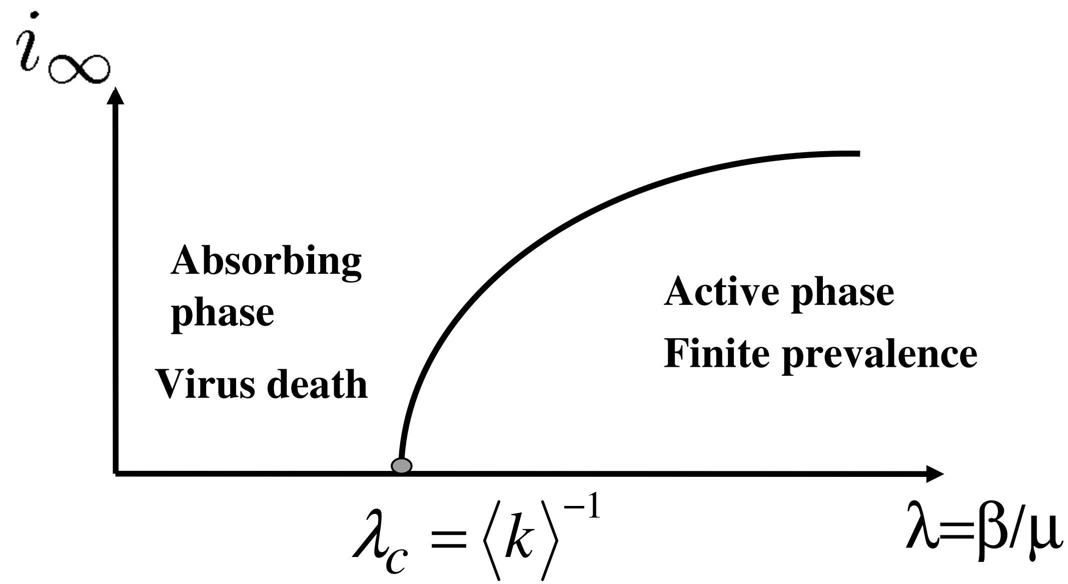
Stationary state: *di/dt = 0* $$\mu i_\infty = \beta\langle k\rangle i_\infty(1-i_\infty)$$

If β <*k*> < μ : $i_\infty = 0$

$\boxed{\text{Epidemic threshold condition: } \beta <k> = \mu}$

If β <*k*> > μ : $i_\infty = 1 - \mu/(\beta\langle k\rangle)$

Phase diagram:



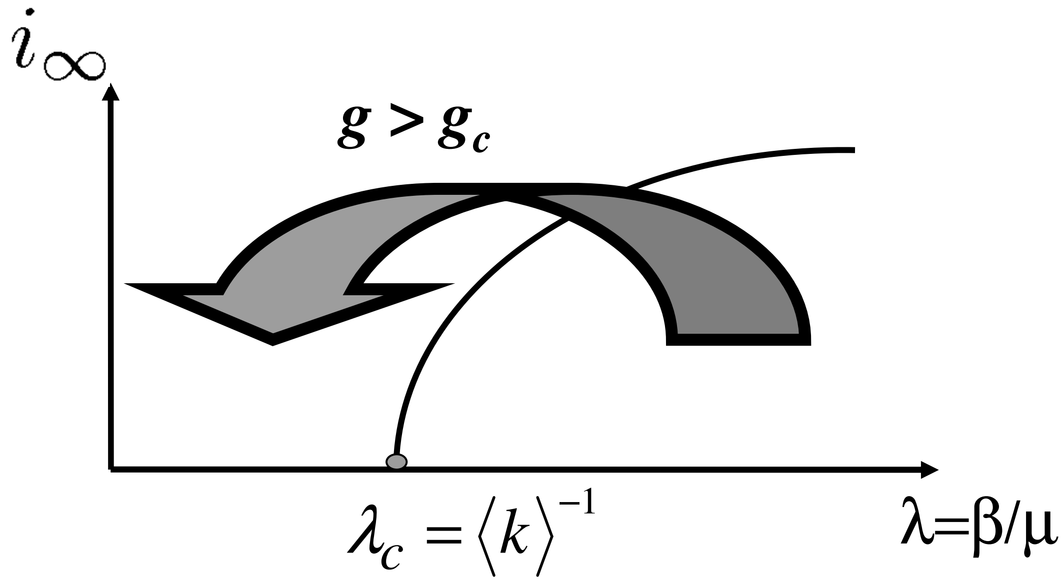
Immunization

Fraction g of immunized (vaccined) individuals:

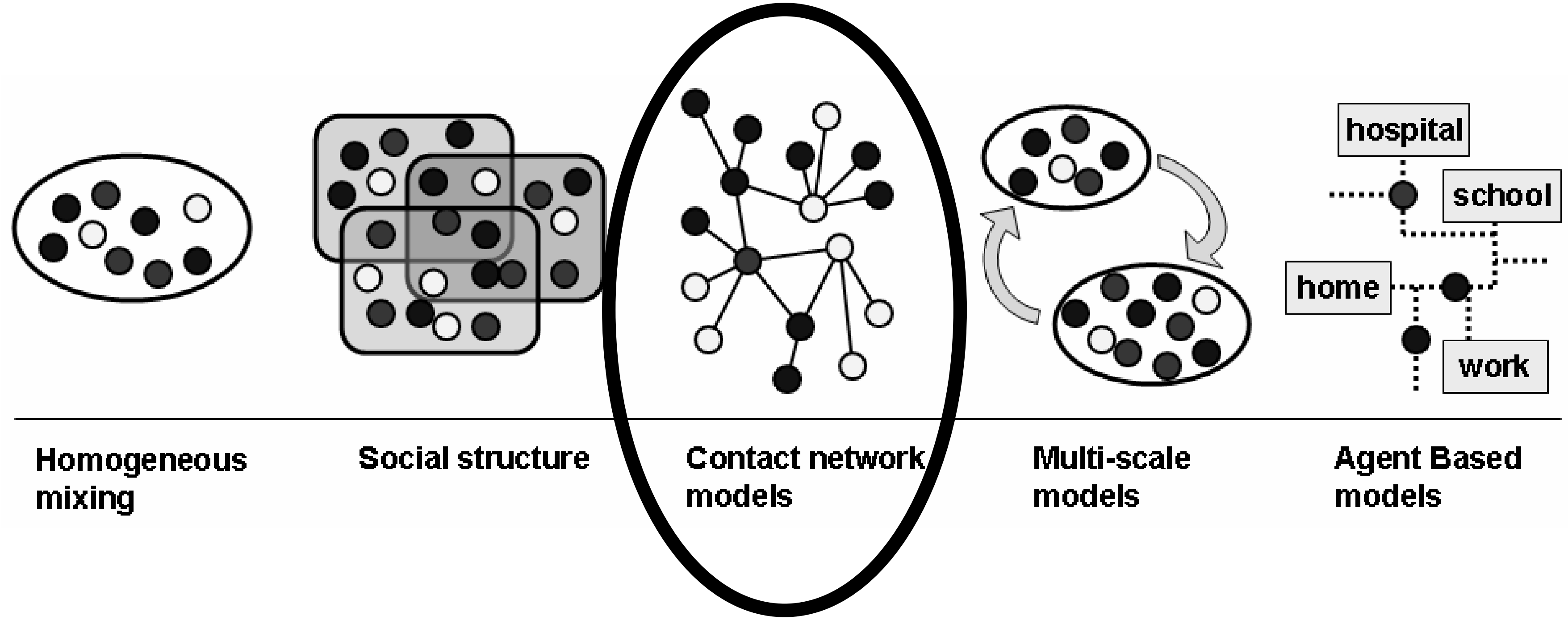
$$\beta \to \beta\,(1-g)$$

=>critical immunization threshold

$$\boldsymbol{g_c = 1 - \mu/(\beta\langle k\rangle)}$$



Wide spectrum of complications and complex features to include…



Simple → **Realistic**

Ability to explain trends at a population level

Model realism looses in transparency.
Validation is harder.

Complex networks

Viruses propagate on networks:

- Social (contact) networks
- Technological networks:
  - Internet, Web, P2P, e-mail...

...which are **complex, heterogeneous networks**

Epidemic spreading on heterogeneous networks

Number of contacts (degree) can vary a lot

huge fluctuations ($\langle k^2 \rangle \gg \langle k \rangle$)

Heterogeneous mean-field: density of

- Susceptible in the class of degree k, $s_k=S_k/N_k$
- Infectious in the class of degree k, $i_k=I_k/N_k$
- Recovered in the class of degree k, $r_k=R_k/N_k$

$$s(t)=\sum P(k)\, s_k,\quad i(t)= \sum P(k)\, i_k,\quad r(t)= \sum P(k)\, r_k$$

Epidemic spreading on heterogeneous networks

Relative density of infected nodes with given degree *k*: i_k

SI model:

$$\boxed{\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k}$$

Θ_k=Proba that any given link points to an infected node

For the SI model:

$$\boxed{\Theta_k = \sum_{k'} \frac{k' - 1}{k'} P(k'|k) i_{k'}}$$

$P(k'|k)$ = the probability that a link originated in a node with connectivity *k* points to a node with connectivity *k'*

SI model on heterogeneous networks

$$\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k \qquad \Theta_k = \sum_{k'} \frac{k'-1}{k'} P(k'|k) i_{k'}$$

In uncorrelated networks: $$\Theta_k = \Theta = \sum_{k'} \frac{k'-1}{\langle k \rangle} P(k') i_{k'}$$

Short times, $i_k(t) << 1$

$$\boxed{\frac{d\Theta}{dt} = \beta \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \Theta}$$

Exponential growth

$$\boxed{\tau = \frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)}}$$

SIS model on heterogeneous networks

$$\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k - \mu i_k \qquad \Theta_k = \sum_{k'} P(k'|k) i_{k'}$$

In uncorrelated networks: $\Theta_k = \Theta = \sum_{k'} \frac{k'}{\langle k \rangle} P(k') i_{k'}$

Short times, $i_k(t) << 1$

$$\boxed{\frac{d\Theta}{dt} = \left(\beta \frac{\langle k^2 \rangle}{\langle k \rangle} - \mu\right) \Theta}$$

Epidemic threshold condition $\boxed{\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}}$

Long time limit, SIS model on het. networks

$$\frac{di_k}{dt} = 0 \Rightarrow i_k(\infty) = \frac{\beta k \Theta_k(\infty)}{\beta k \Theta_k(\infty) + \mu}$$

$$\Theta_k = \sum_{k'} P(k'|k) i_{k'}$$

Sum over *k*

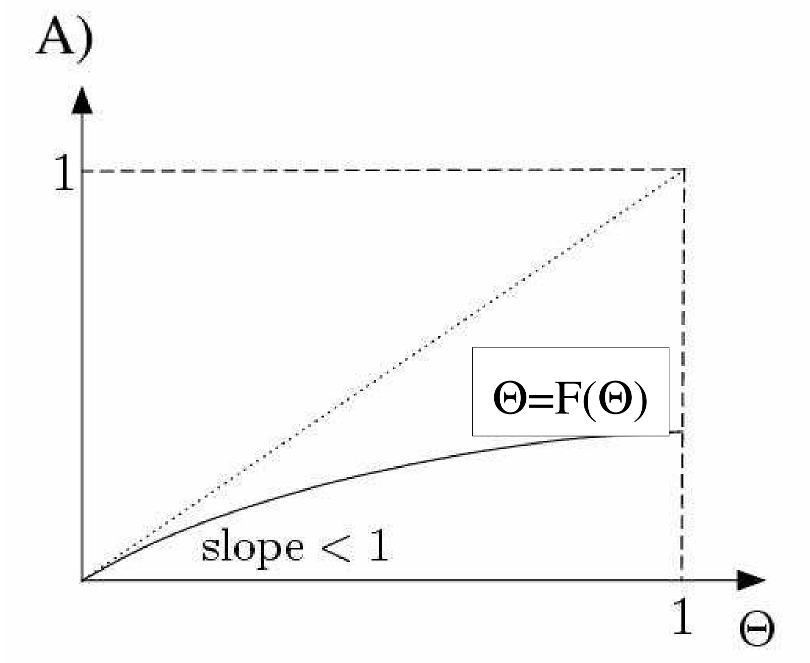
In uncorrelated networks: $\Theta_k = \Theta = \sum_{k'} \frac{k'}{\langle k \rangle} P(k') i_{k'}$

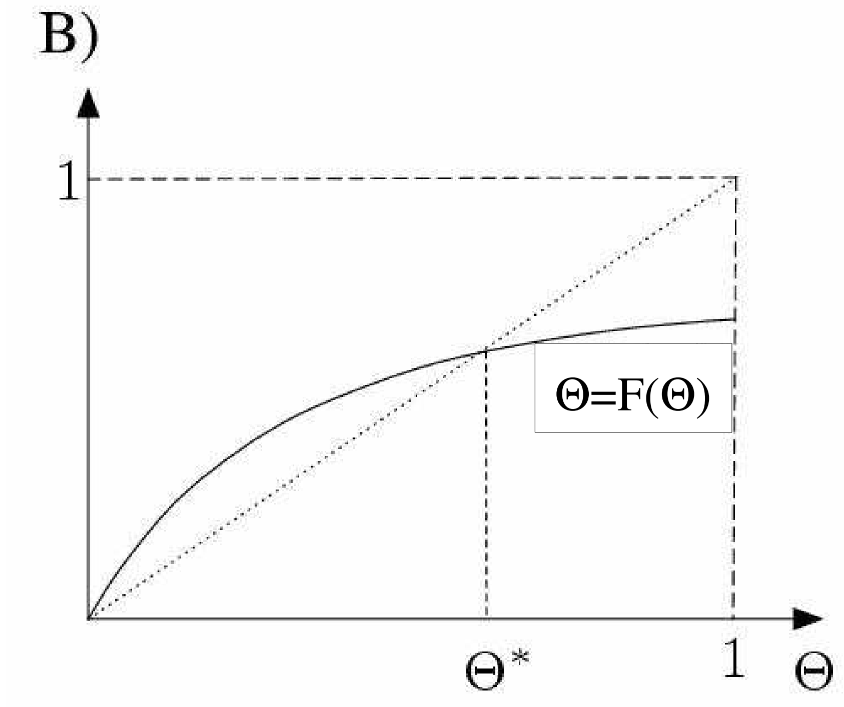
$$\Theta(\infty) = \frac{1}{\langle k \rangle} \sum_k \frac{\beta k^2 P(k) \Theta(\infty)}{\beta k \Theta(\infty) + \mu}$$

Self-consistent equation of the form Θ=F(Θ)

with F(0)=0, F' > 0, F'' < 0

Graphical solution





Epidemic threshold:

existence of a non-zero solution for Θ $\quad$ F'(0) > 1 :

$$\Longleftrightarrow \quad \sum_k \frac{\beta k^2 P(k)}{\mu \langle k \rangle} > 1 \quad \Longleftrightarrow \quad \frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} > 1$$

Epidemic threshold in uncorrelated networks

$$\boxed{\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}}$$

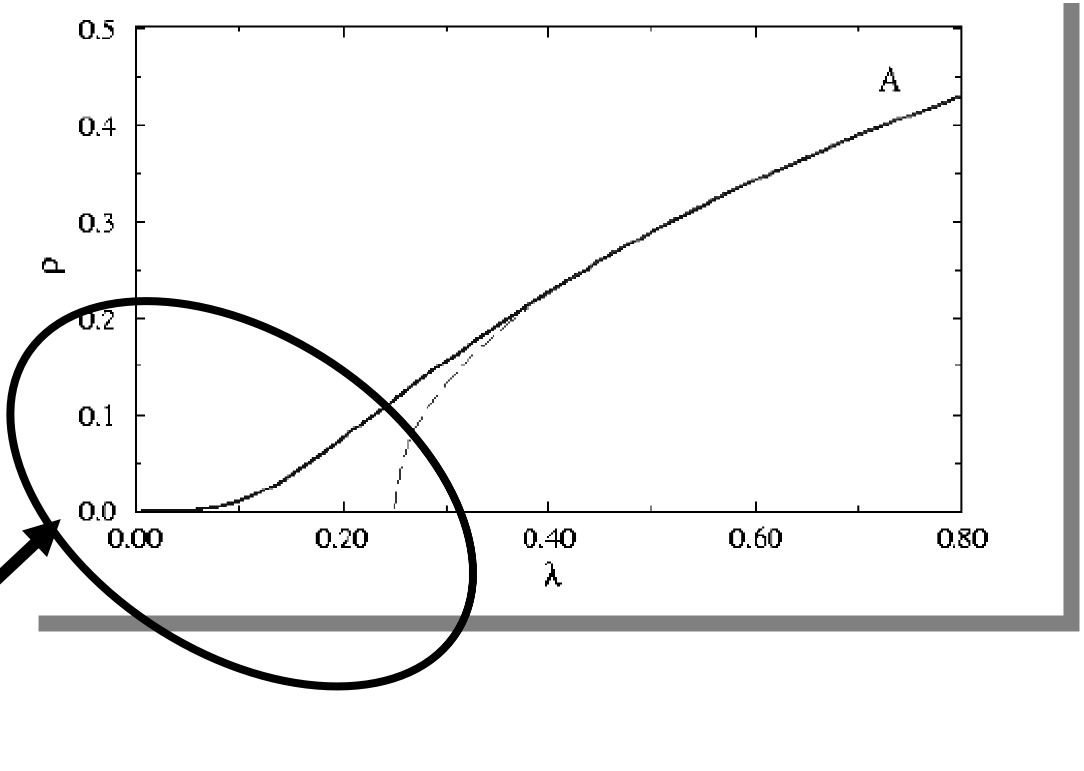
Heterogeneous, infinite network:

$$\langle k^2 \rangle \to \infty$$

Condition always satisfied

Finite prevalence for any spreading parameters β, μ

Epidemic phase diagram in heterogeneous networks



- Wide range of spreading rate with low prevalence
- Lack of healthy phase = standard immunization cannot drive the system below threshold!!!

Immunization strategies

'Usual', uniform immunization:

Fraction g of randomly chosen immunized (vaccined) individuals:

$$\beta \rightarrow \beta\,(1-g)$$

=> inefficient

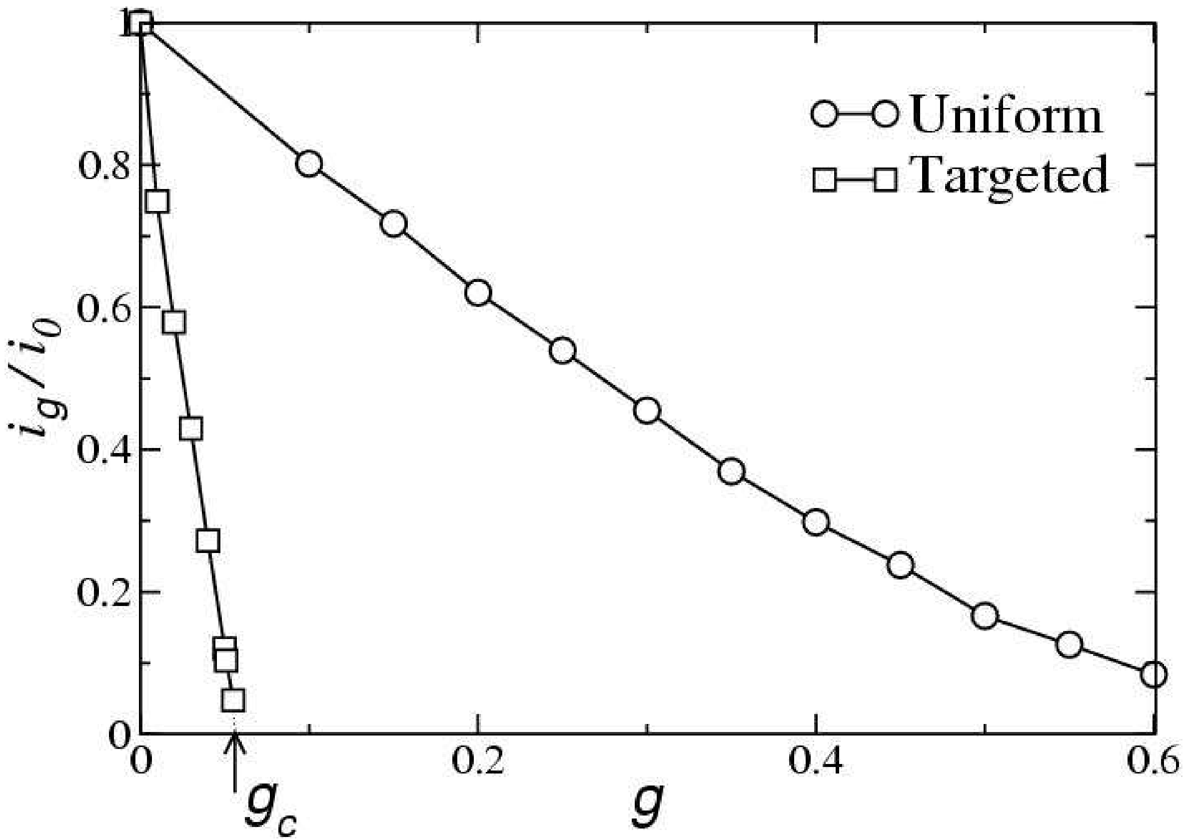
Targeted immunization:

Vaccination of the most connected individuals

=> efficient

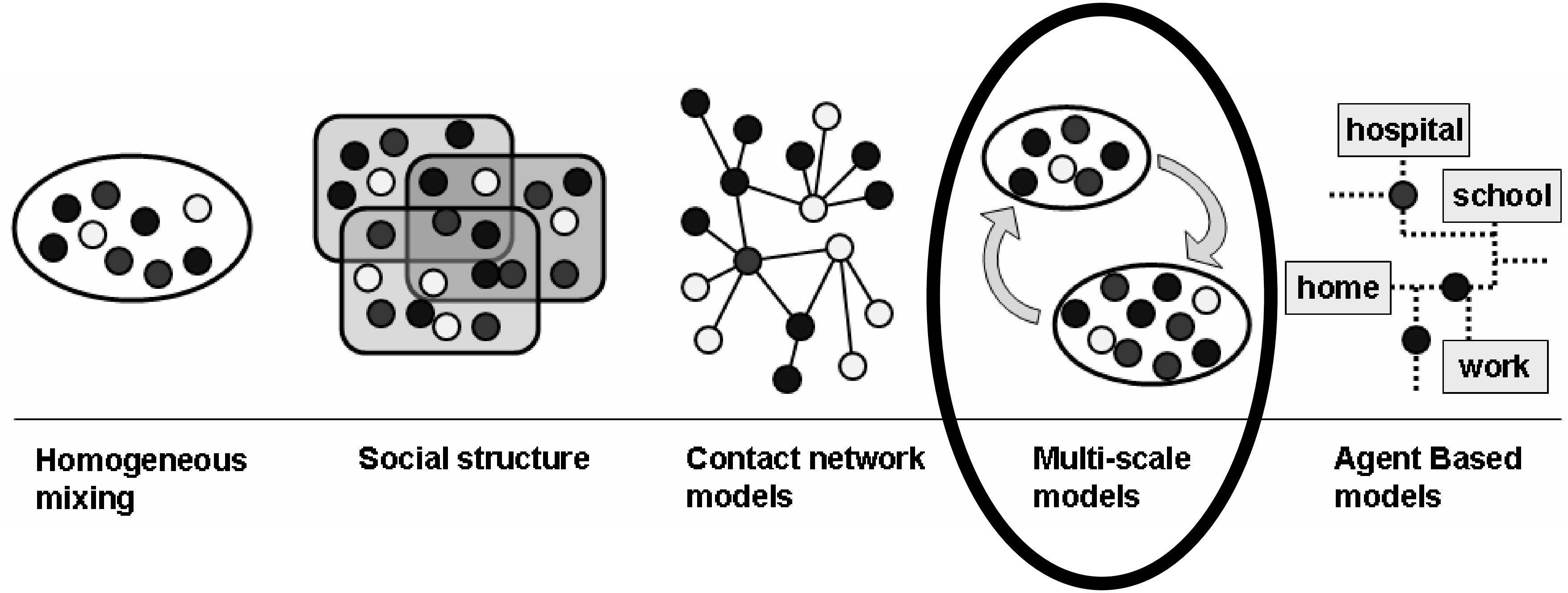
(cf resilience vs fragility to attacks)

Immunization



NB: when network's topology unknown: acquaintance immunization

Wide spectrum of complications and complex features to include…



Simple → **Realistic**

Ability to explain trends at a population level

Model realism looses in transparency.
Validation is harder.

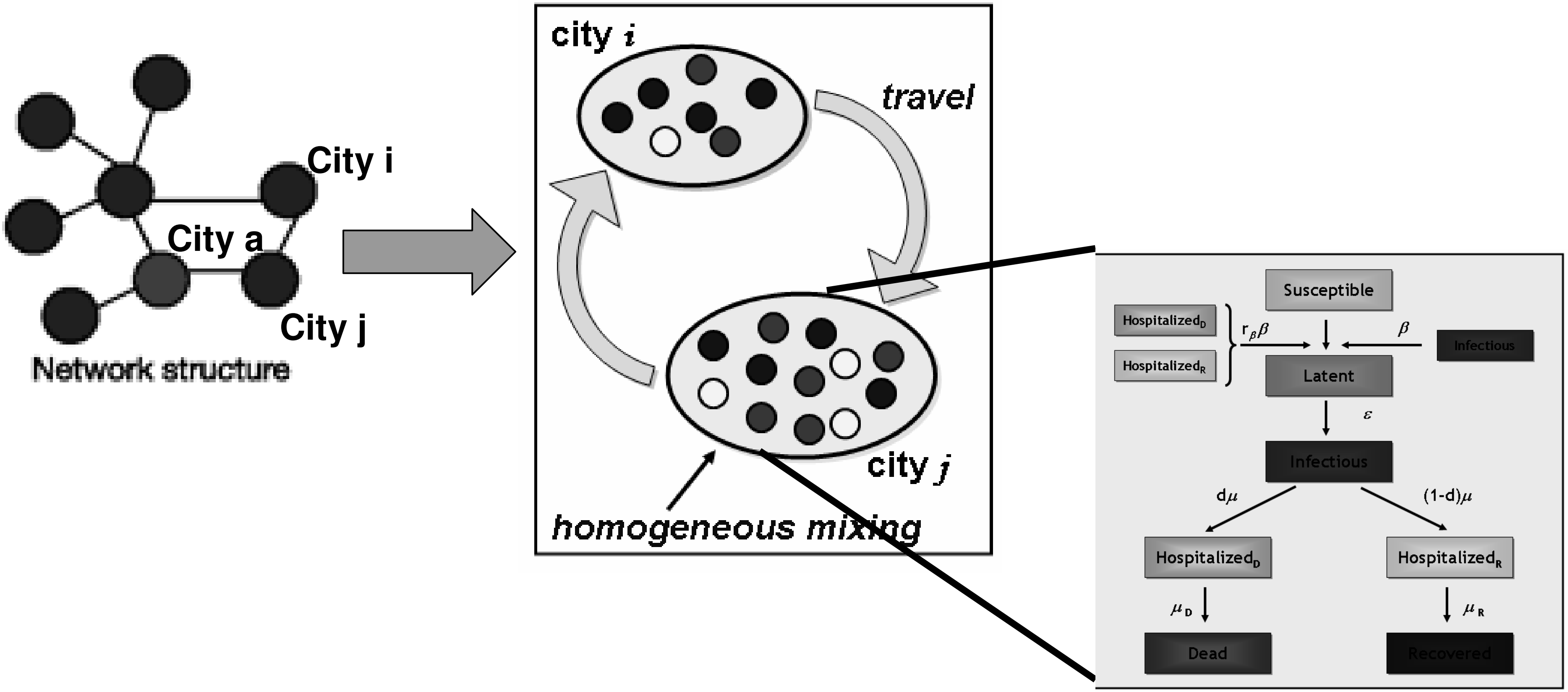
General framework:
Bosonic reaction-diffusion processes

- Previous cases: (at most) *one* particle/individual *per site*
- In general: reaction-diffusion processes on networks => no restriction on the number of particles per site

"Particles"

- diffusing along edges
- reacting in the nodes

Example: Meta-population models



Intra-population infection dynamics by stochastic compartmental modeling

Baroyan *et al.* (1969)
Ravchev, Longini (1985)

Bosonic RD processes on complex networks

- A_α , α=1,…,S types of particles
- Diffusion coefficients D_α
- Reactions (*r=1,…R*):

$$\sum_{\alpha=1}^{S} q_\alpha^r A_\alpha \xrightarrow{\lambda_r} \sum_{\alpha=1}^{S} (q_\alpha^r + p_\alpha^r) A_\alpha$$

Bosonic RD processes on complex networks

Heterogeneous mean-field formalism:

Densities $\rho_{\alpha,k} \equiv \frac{n_{\alpha,k}}{N_k} = \sum_{i \in k} \frac{n_{\alpha,i}}{N_k}$

$$\rho_\alpha(t) = \sum_k \rho_{\alpha,k}(t) P(k)$$

$$\partial_t \rho_{\alpha,k}(t) = \mathcal{D}_\alpha + \mathcal{R}_\alpha$$

Diffusion

Reaction

Bosonic RD processes
on complex networks

Diffusion term:

For site i: $$-D_\alpha n_{\alpha,i}(t) + D_\alpha \sum \frac{x_{ji}}{k_j} n_{\alpha,j}(t)$$

$$\mathcal{D}_\alpha = -D_\alpha \rho_{\alpha,k}(t) + D_\alpha k \sum_{k'} \frac{P(k'|k)}{k'} \rho_{\alpha,k'}(t)$$

Bosonic RD processes on complex networks

$$\sum_{\alpha=1}^{S} q_\alpha^r A_\alpha \xrightarrow{\lambda_r} \sum_{\alpha=1}^{S} (q_\alpha^r + p_\alpha^r) A_\alpha$$

Reaction term: $\mathcal{R}_\alpha = \sum_r p_\alpha^r \lambda_r \prod_r (\rho_{\beta,k})^{q_\beta^r}$

See Baronchelli et al, Phys. Rev. E 78, 016111 (2008)
for various examples of further computations

A concrete example:

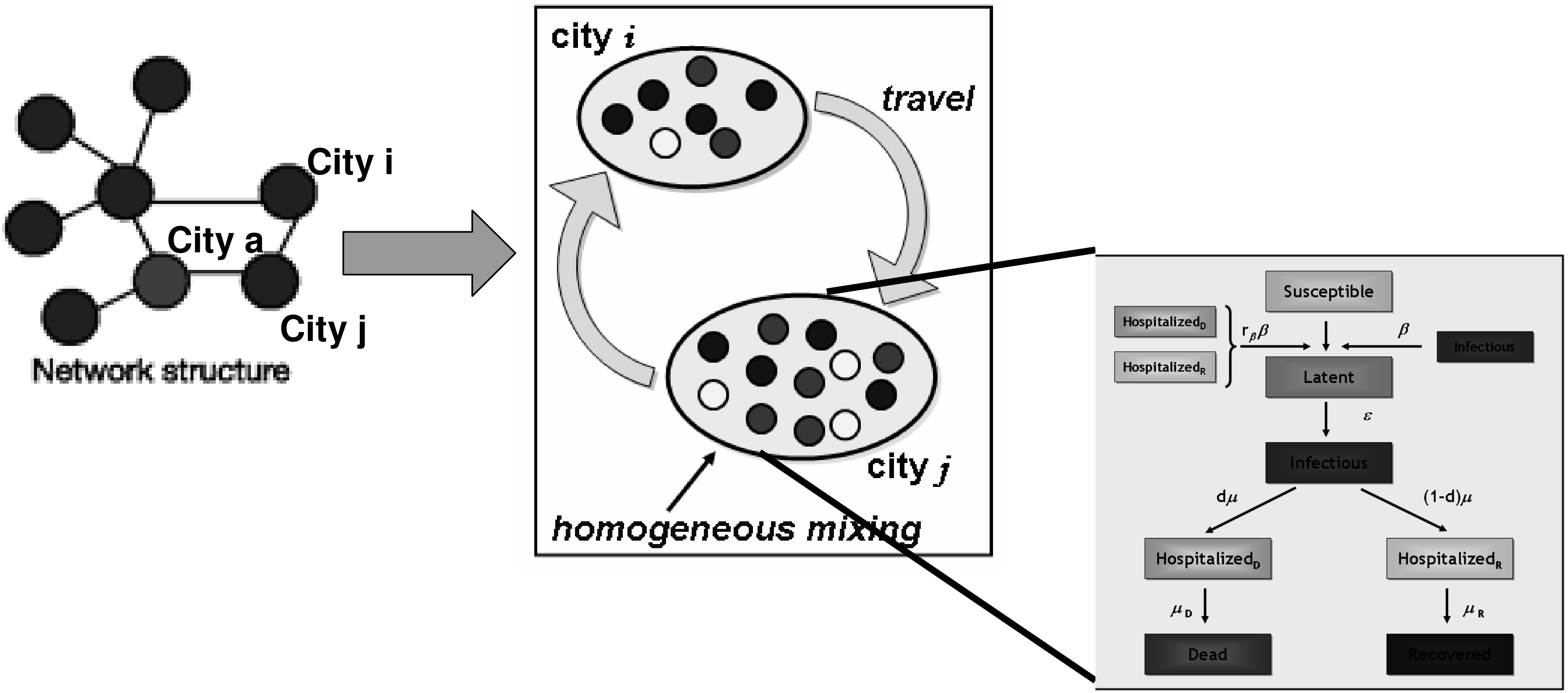
epidemic meta-population models

Unprecedented amount of data…..

- Transportation infrastructures
- Behavioral Networks
- Census data
- Commuting/traveling patterns

  - Different scales:
    - International
    - Intra-nation (county/city/municipality)
    - Intra-city (workplace/daily commuters/individuals behavior)

Meta-population models



Intra-population infection dynamics by stochastic compartmental modeling

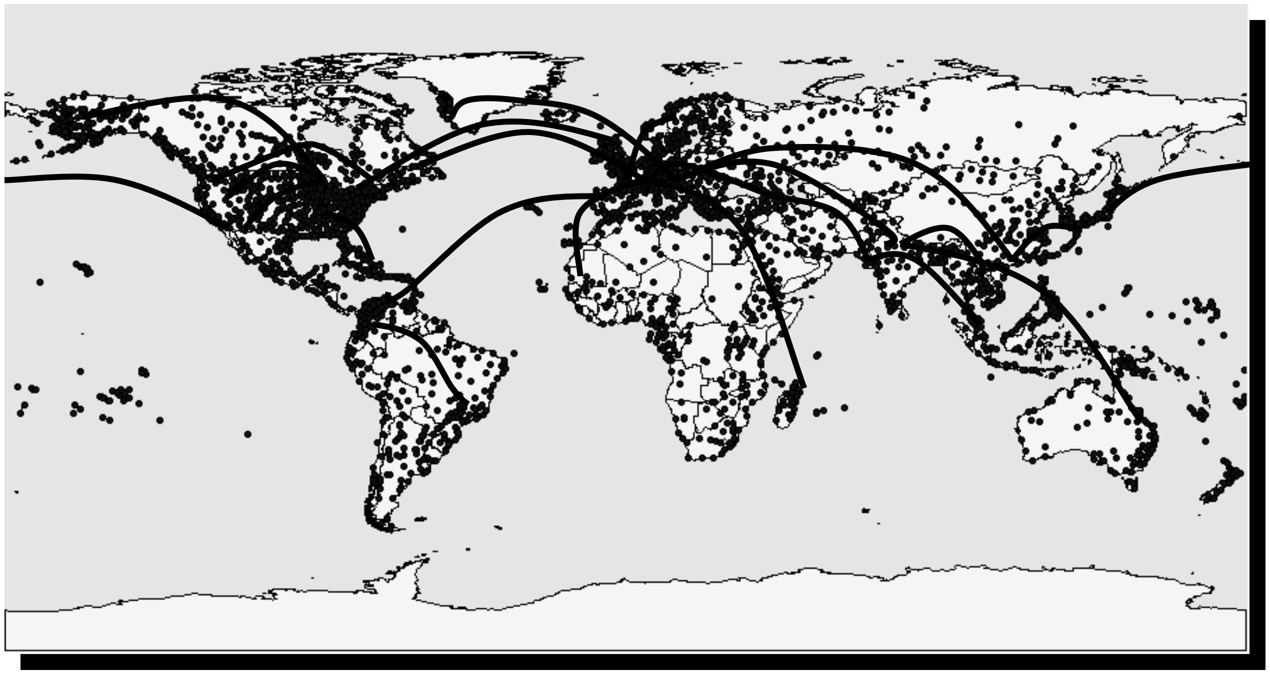
Baroyan *et al.* (1969)
Ravchev, Longini (1985)

Modeling of global epidemics propagation

multi-level description :

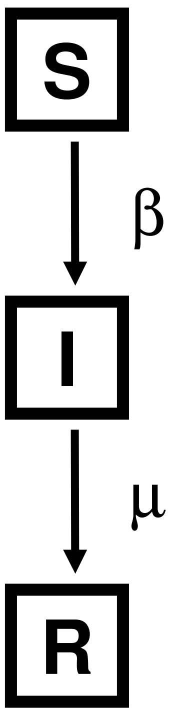
§ intra-city **epidemics**

§ inter-city **travel**



Baroyan *et al.* (1969)
Ravchev, Longini (1985)

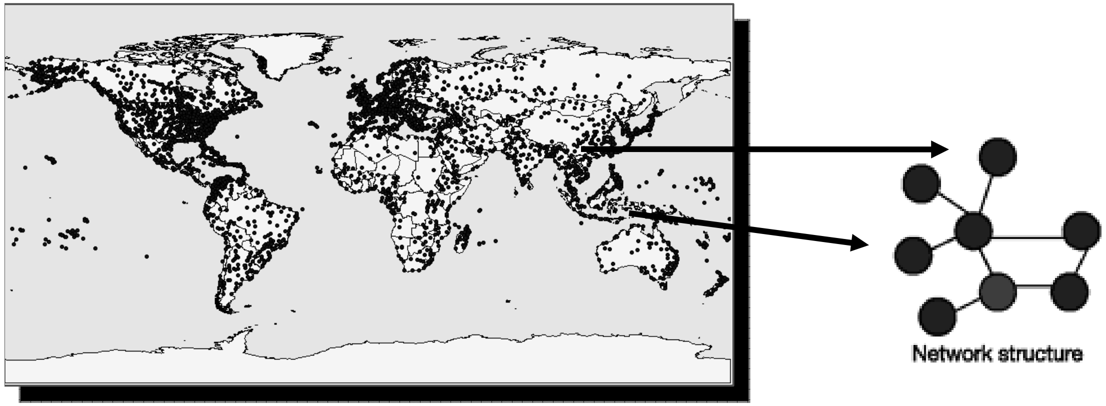
Inside a city



- **Homogeneous assumption**

- β rate of transmission
- μ^{-1} average infectious period

Global spread of epidemics on the airport infrastructure



Urban areas
+
Air traffic flows

World-wide airport network

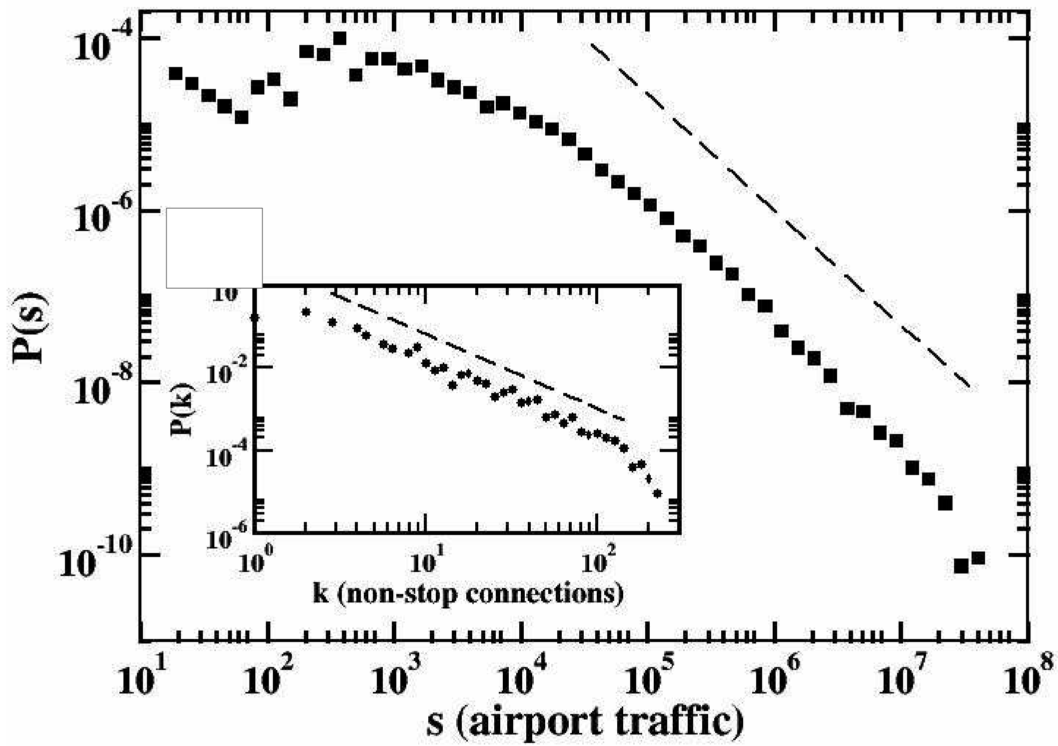
- **complete IATA database**
  - ***V = <u>3100</u>* airports**
  - ***E = <u>17182</u>* weighted edges**
  - **w_{ij} #seats / (different time scales)**

> 99% of total traffic

- **N_j urban area population (UN census, …)**

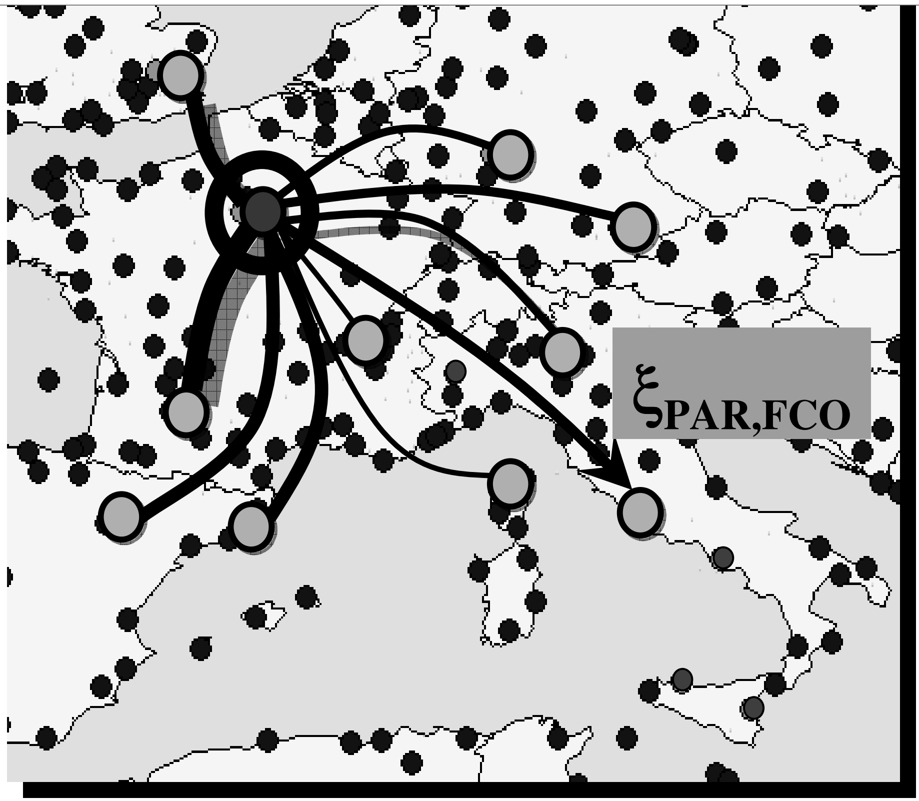
Statistical distributions…

- Skewed
- Heterogeneity and high variability
- Very large fluctuations (variance>>average)



Barrat, Barthélemy, Pastor-Satorras, Vespignani. *PNAS* (2004)
Colizza, Barrat, Barthélemy, Vespignani. *PNAS* (2006)

Stochastic model: travel term

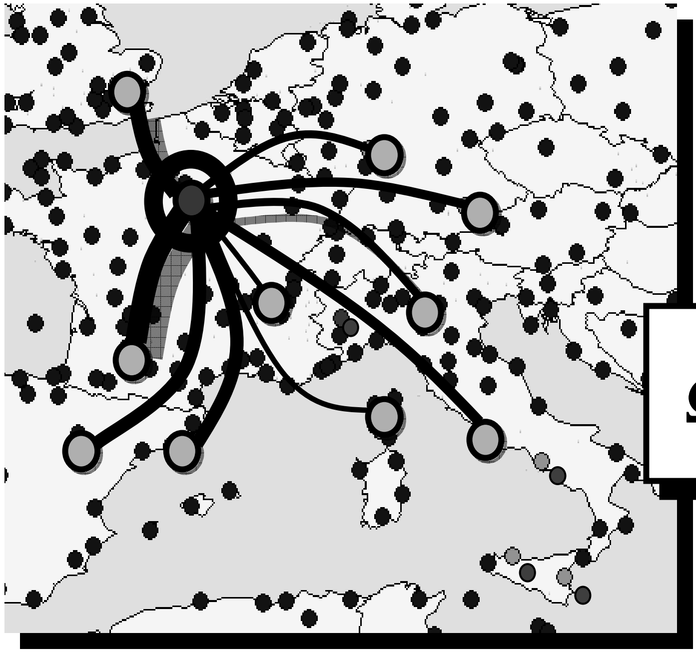


Travel probability from *PAR* to *FCO*:

$$p_{PAR,FCO} = \frac{w_{PAR,FCO}}{N_{PAR}} \Delta t$$

$\xi_{PAR,FCO}$ # passengers from *PAR* to *FCO*: *Stochastic variable*, multinomial distr.

Stochastic model: travel term



Transport operator:

$$\mathbf{\Omega_{PAR}\,(\{X\}) = \sum_l \left(\xi_{l,PAR}\,(\{X_l\}) - \xi_{PAR,l}\,(\{X_{PAR}\})\right)}$$

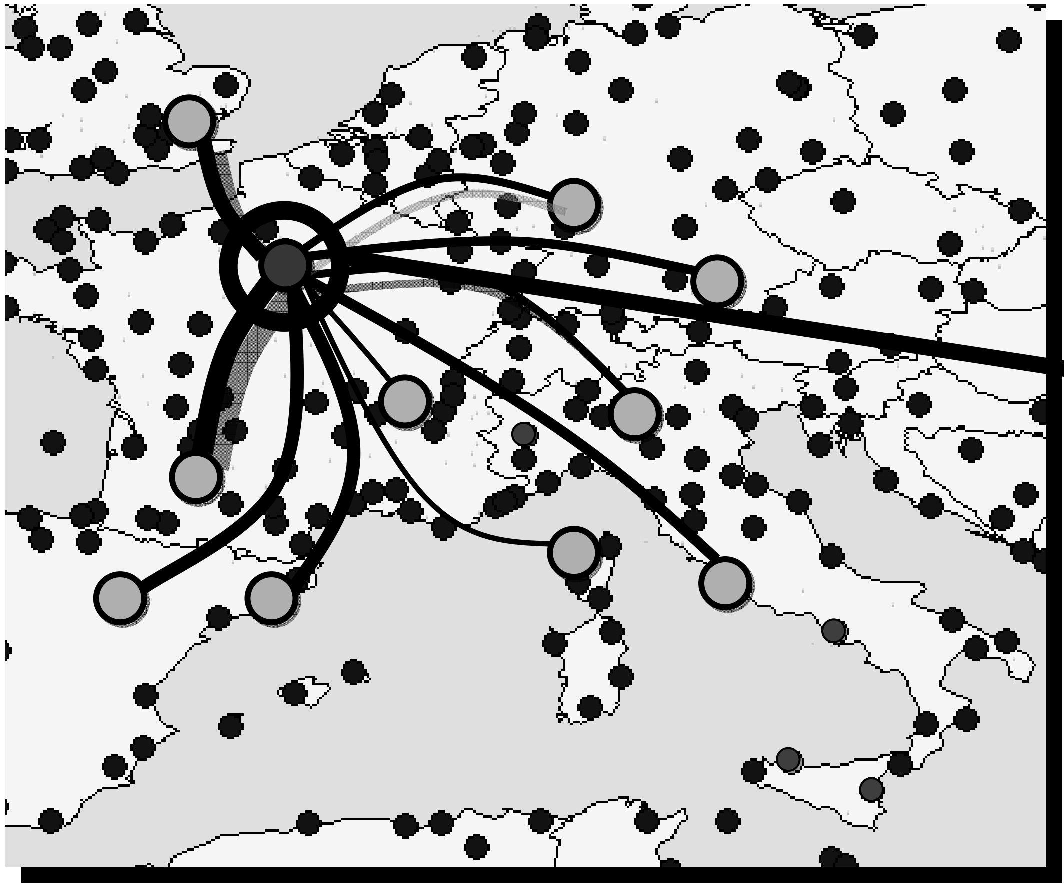
ingoing **outgoing**

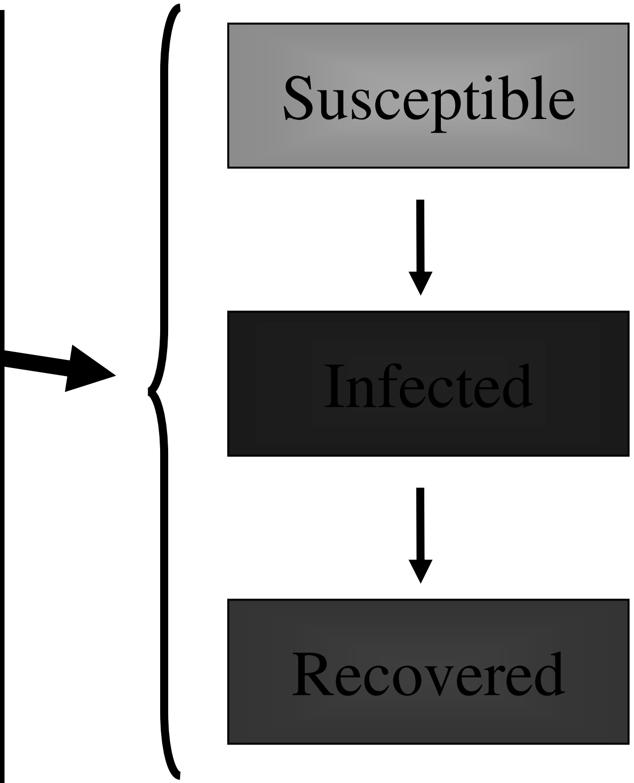
§ other source of **noise**: $w_{jl}^{noise} = w_{jl}\,[\alpha+\eta(1-\alpha)]$ α=70%

§ **two-legs** travel: $\Omega_j(\{X\}) = \Omega_j^{(1)}(\{X\}) + \Omega_j^{(2)}(\{X\})$

Stochastic large-scale model

compartmental model + **air transportation data**



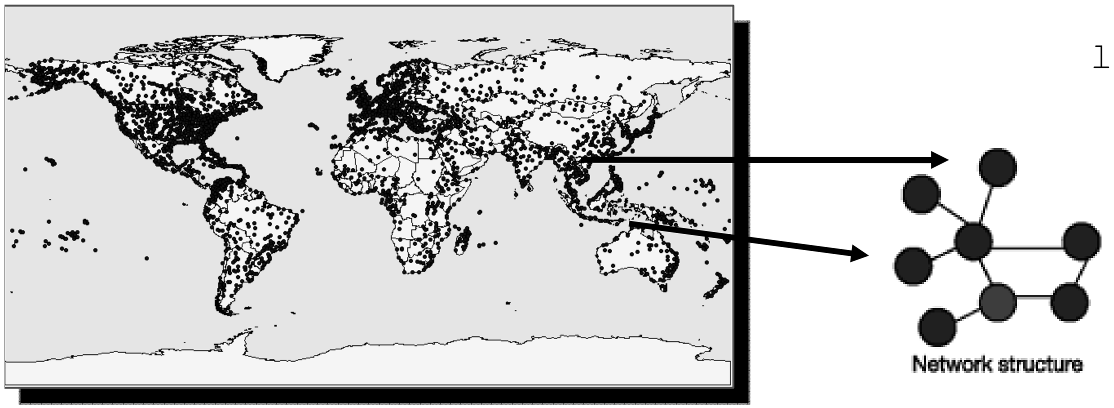


Large-scale model

Stochastic evolution equations describing the disease evolution at a mean-field level in each city

…

Coupled through transport operators



- **complete IATA database**
  - ***V = <u>3100</u>* airports**
  - ***E = <u>17182</u>* weighted edges**
  - **w_{ij} #seats / (different time scales)**
- **N_j urban area population (UN census, …)**

> 99% of total traffic

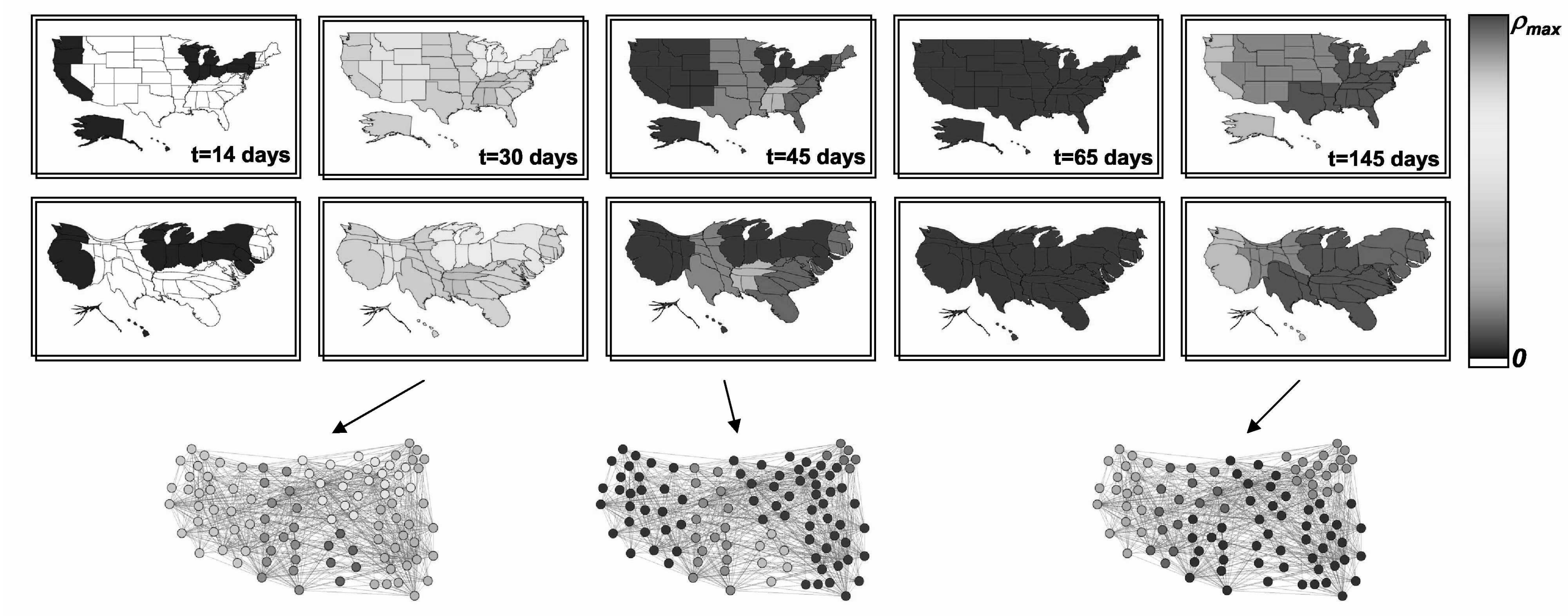
Directions…..

- **Applications…**
  - Historical data
  - Scenarios forecast
  - complicated, realistic disease models

- Basic **theoretical** questions…
  - simple SIS, SIR models
  - features determining **propagation pattern**?
  - issue of **predictability**, epidemic forecasting?

Propagation pattern

Epidemics starting in Hong Kong (SIR model)



Colizza, Barrat, Barthélemy, Vespignani, *PNAS* 103, 2015 (2006); *Bull. Math. Bio.* (2006)

Heterogeneity

- maps heterogeneity epidemic spread

- appropriate measure ?
- role of specific structural properties: topology, traffic, population ?
- comparison with null hypothesis

Heterogeneity: quantitative measure

$$i_j(t) = \frac{I_j(t)}{N_j}$$ prevalence in city *j* at time *t*

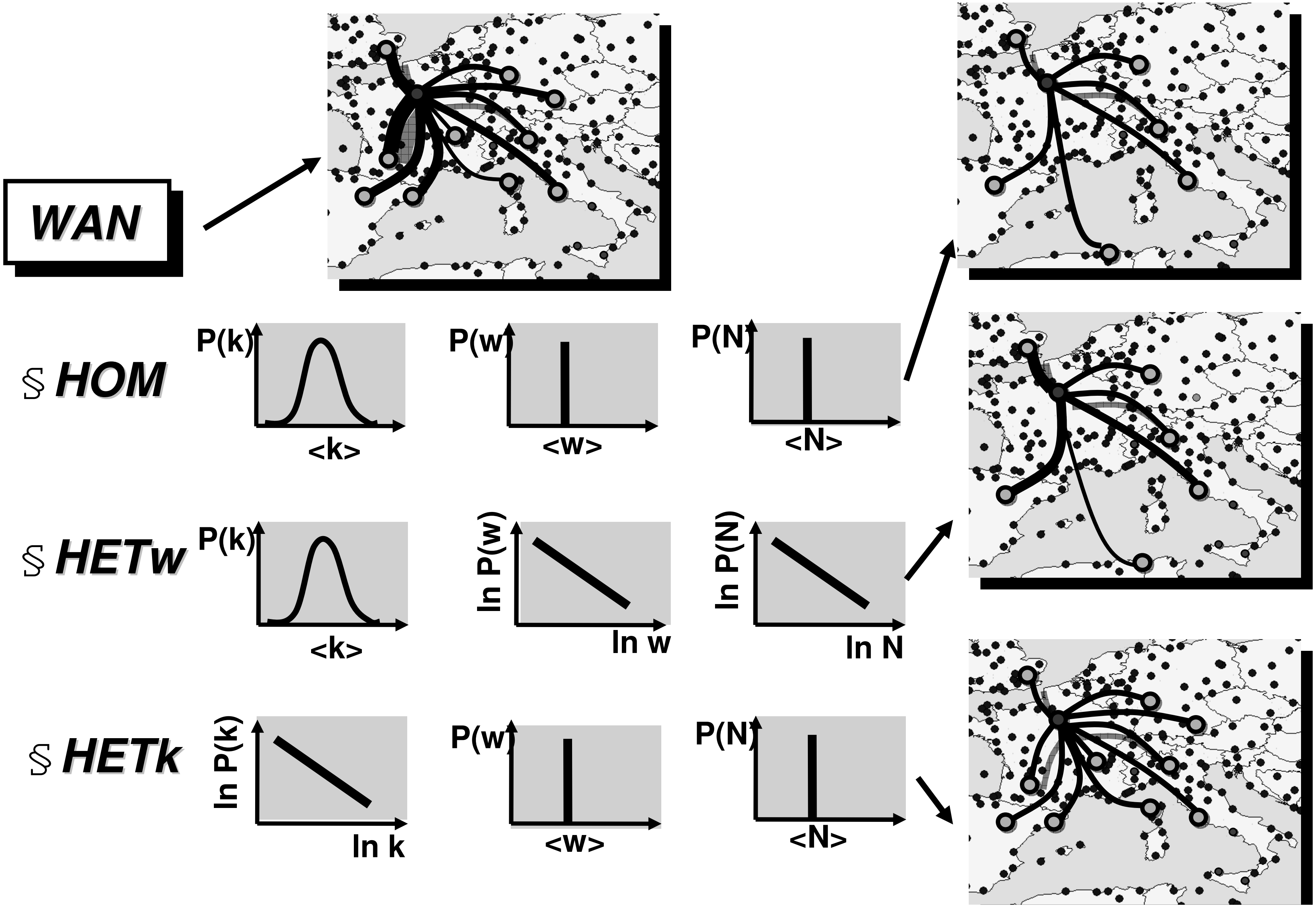
$$\rho_j(t) = \frac{i_j(t)}{\sum_l i_l(t)}$$ normalized prevalence

Entropy:

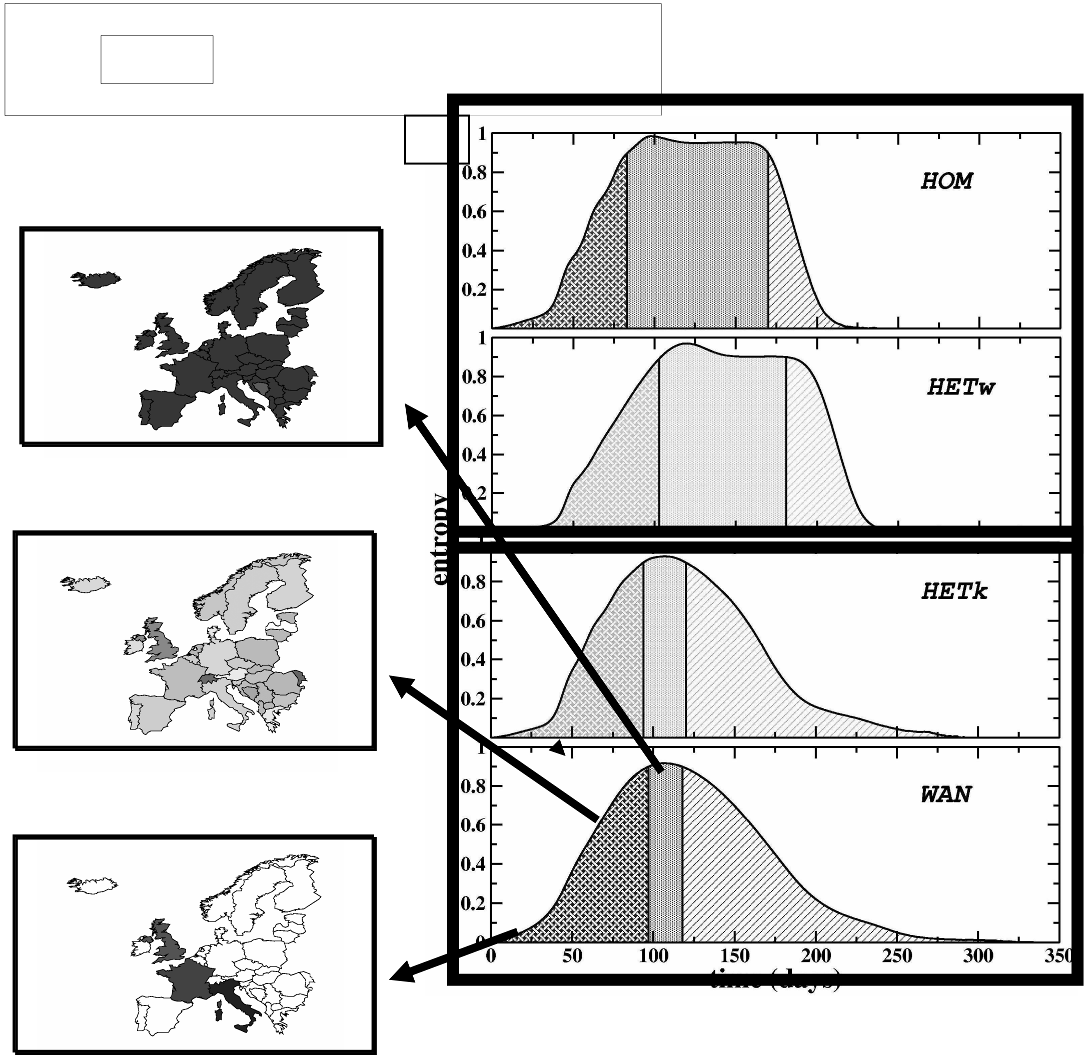
$$\boxed{H(t) = -\frac{1}{\ln V}\sum_j \rho_j \ln \rho_j}$$

H* ∈ *[0,1]
H=0 most het.
H=1 most hom.

...and: compare with null hypothesis!



Results: Heterogeneity



- global properties
- average over initial *seed*
- central zone: *H>0.9*
- ***HETk*** $\cong$ ***WAN***
  importance of ***P(k)***

Prediction and predictability

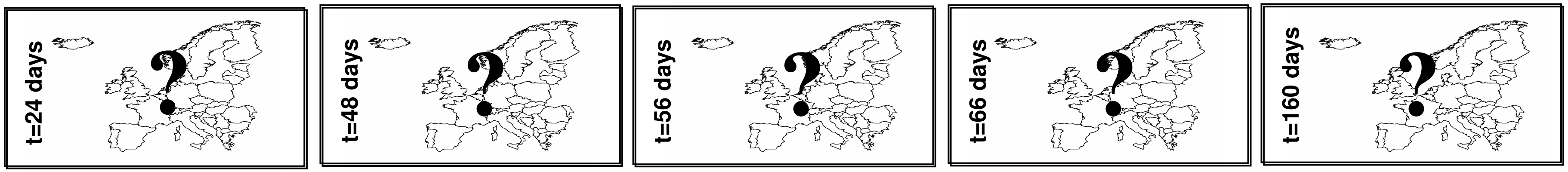
- Do we have consistent scenario with respect to different stochastic realizations?
- What are the network/disease features determining the predictability of epidemic outbreaks
- Is it possible to have **epidemic forecasts?**

Colizza, Barrat, Barthélemy, Vespignani, *PNAS* 103, 2015 (2006); *Bull. Math. Bio.* (2006)

Predictability

One outbreak realization:

Another outbreak realization ?



§ epidemic forecast
§ containment strategies

Quantitative characterization of epidemic predictability

Statistical similarity of two outbreaks (I and II) with the same initial conditions subject to different noise realizations

Observable: infected probability distribution

$$\vec{\pi}(t), \qquad \pi_j(t) = \frac{I_j(t)}{\sum_l I_l(t)}$$

$$sim(\vec{\pi}^I, \vec{\pi}^{II}) = \sum_j \sqrt{\pi_j^I \pi_j^{II}}$$

Quantitative characterization of epidemic predictability

NB: The normalized distribution similarity is the same in the case of different total prevalence

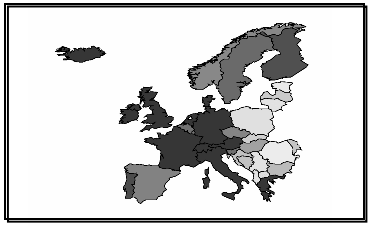
$$\vec{i}(t) = (i, 1-i), \quad i = \sum_j I_j / \sum_j N_j$$

Overlap function:

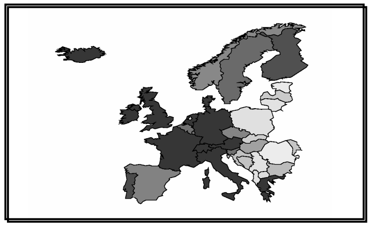
$$\boxed{\Theta(t) = sim(\vec{i}^{\,I}, \vec{i}^{\,II}) \times sim(\vec{\pi}^{I}, \vec{\pi}^{II})}$$

Quantitative characterization of epidemic predictability

$$\mathbf{\Theta(t) \in [0,1]}$$

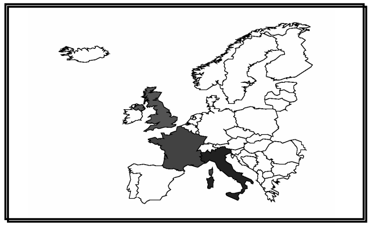


time t



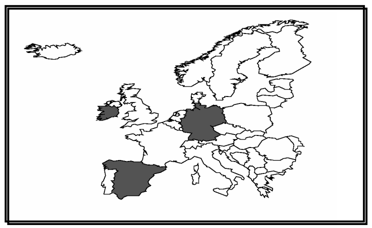
time t

$\mathbf{\Theta(t)=1 \rightarrow}$ **2 identical outbreaks**



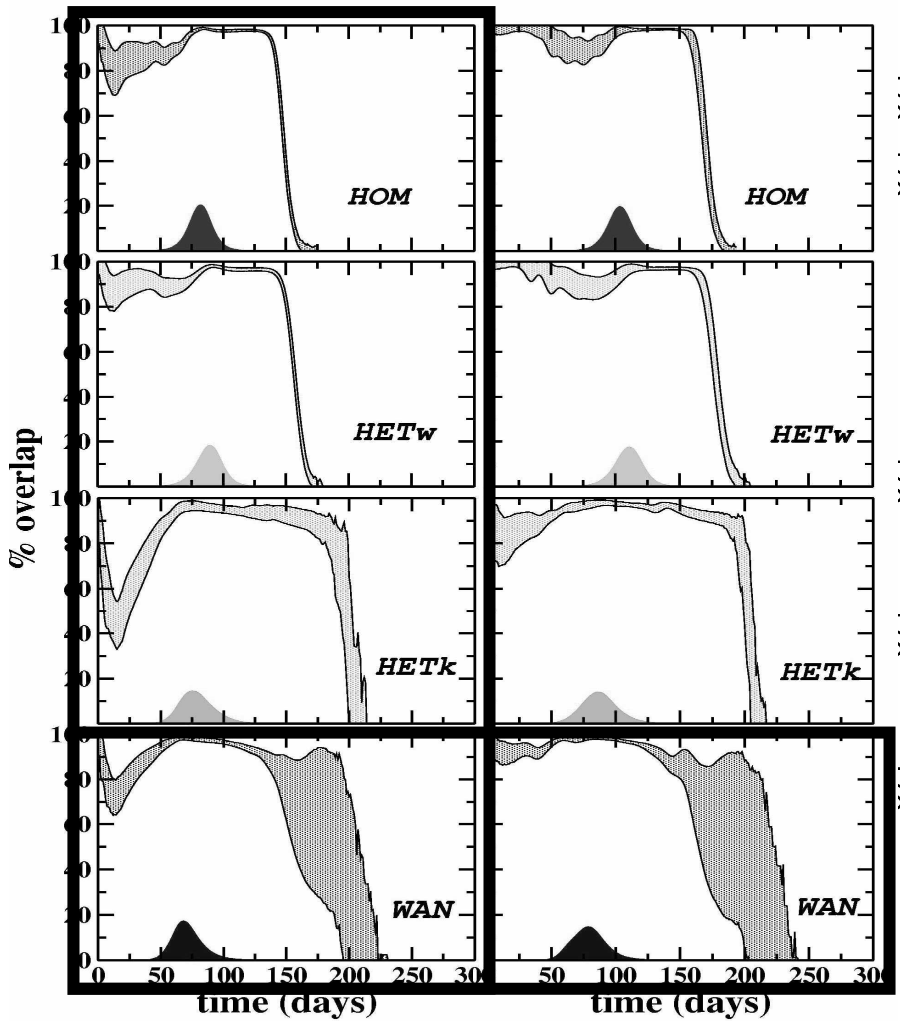
time t

$\mathbf{\Theta(t)=0 \rightarrow}$ **2 distinct outbreaks**



time t

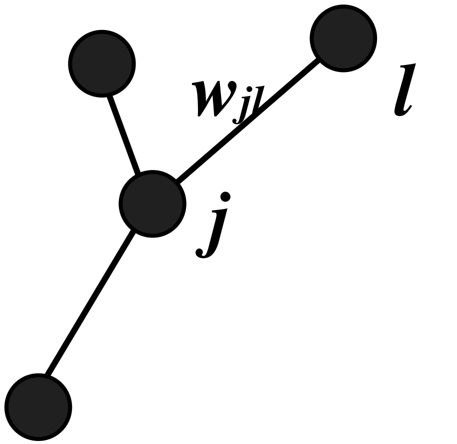
Results: predictability



- **left:** *seed* = airport hubs
- **right:** *seed* = poorly connected airports

- ***HOM & HETw*** high overlap
- ***HETk*** low overlap
- ***WAN*** increased overlap !!

Results: predictability



HOM: $k_i \approx \langle k \rangle \to$ few channels

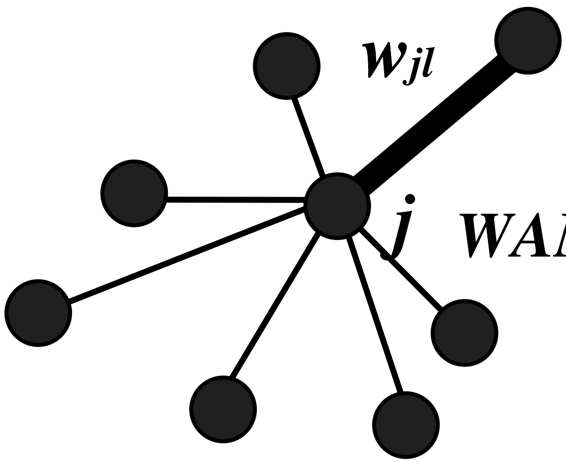
$\to$ high overlap

+ degree heterog.

HETk: *broad P(k)* $\to$ lots of channels!

$\to$ low overlap

+ weight heterog.



WAN: *broad P(k),P(w)* lots of channels, but…

$\to$ emergence of **preferred channels**

$\to$ **increased overlap !!!**

Taking advantage of complexity…

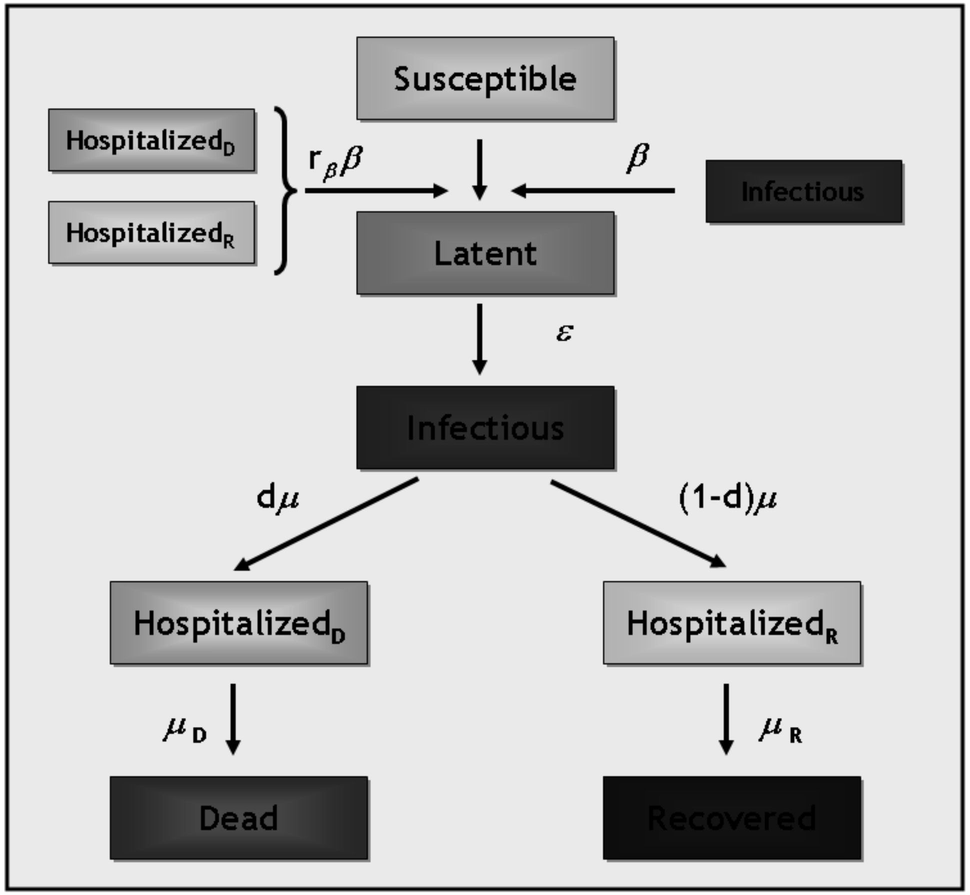
- Two competing effects
  - Paths degeneracy (connectivity heterogeneity)
  - Traffic selection (heterogeneous accumulation of traffic on specific paths)

- Definition of ***epidemic pathways*** as a backbone of dominant connections for spreading

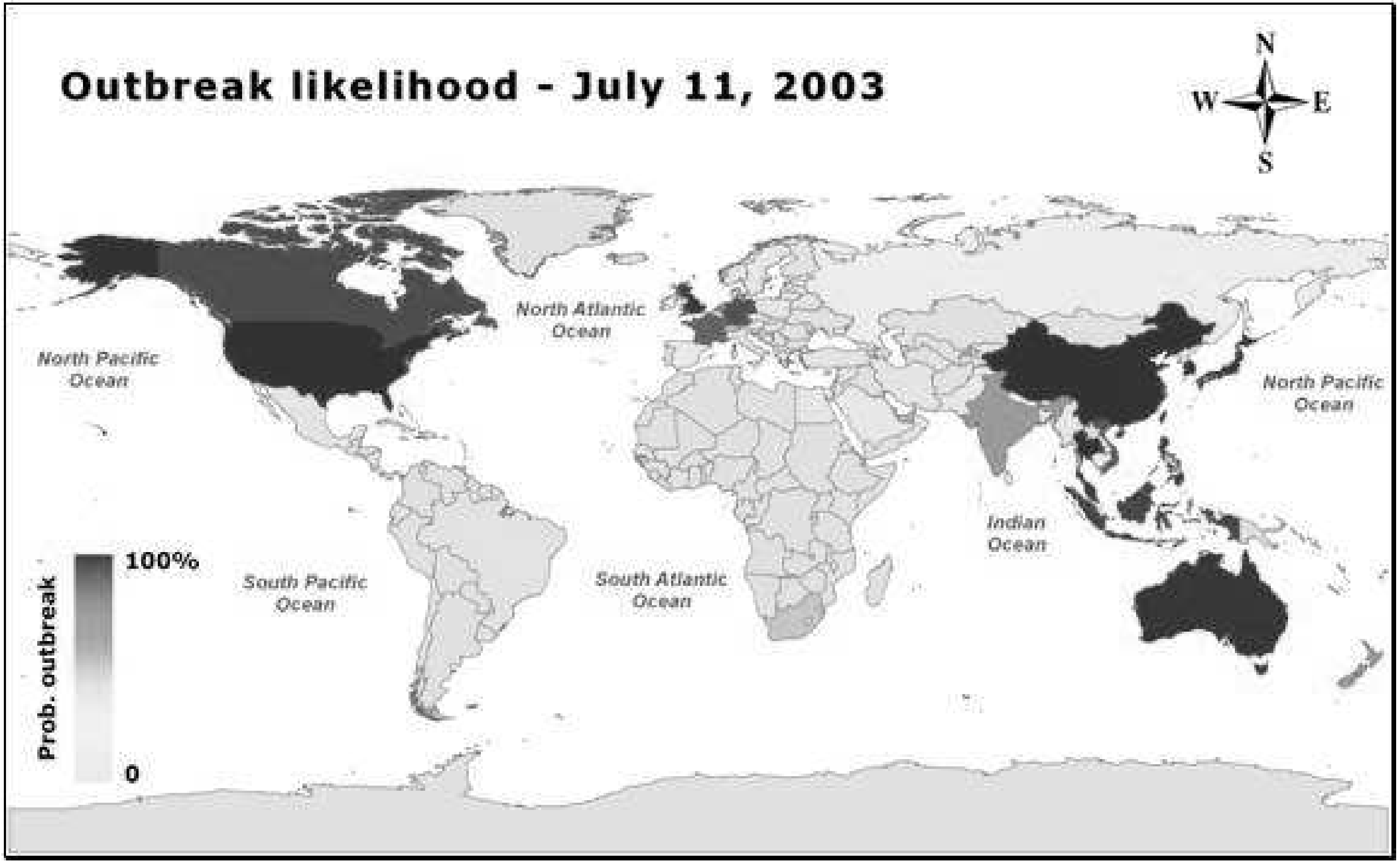
Applications

- Historical data: SARS
  - ***more involved model***
  - ***validation vs real data***
- Pandemic forecast
  - ***effect of travel limitations***
  - ***scenario evaluation***

Historical data :
The SARS case…

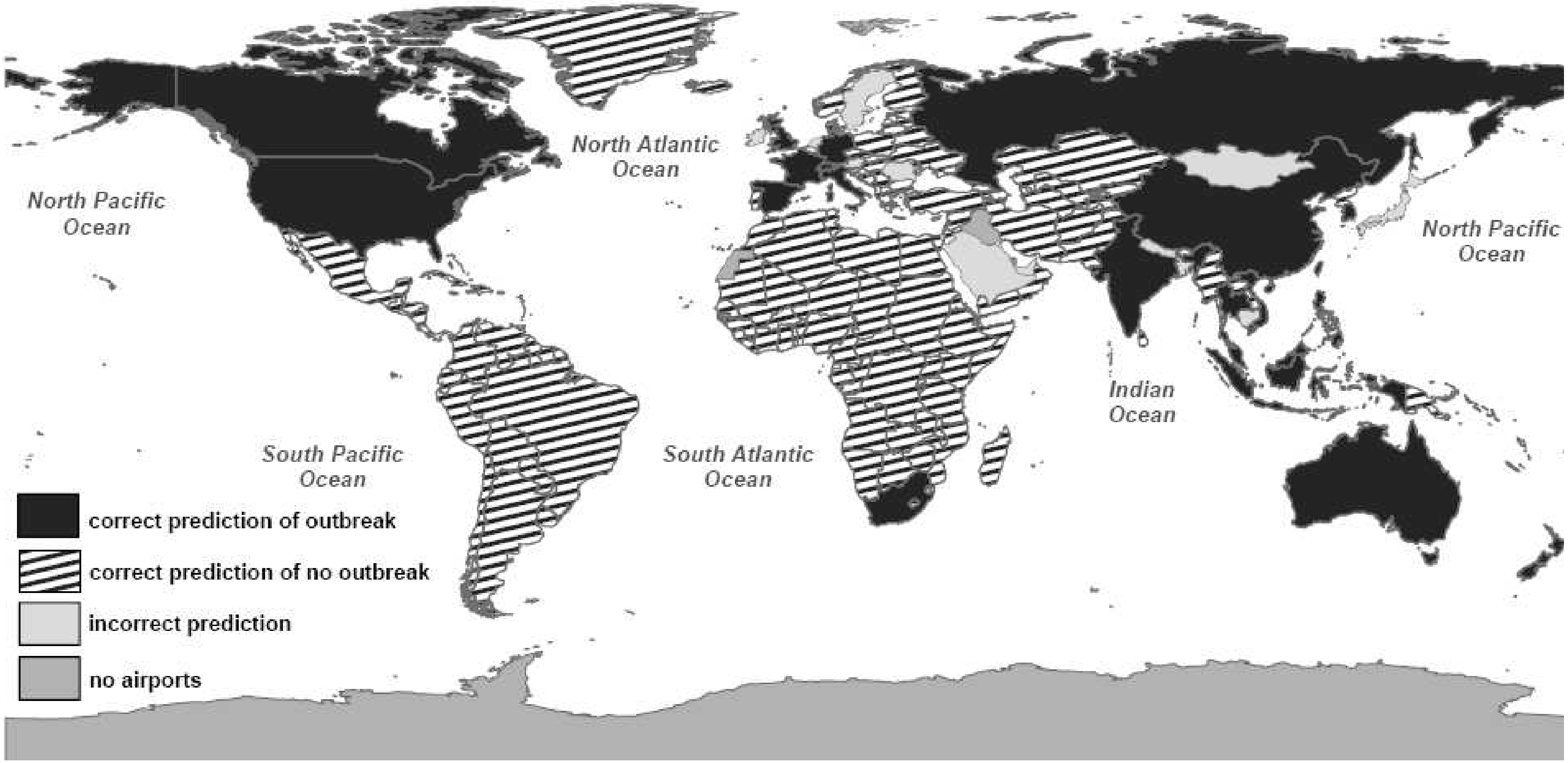


Predictions…

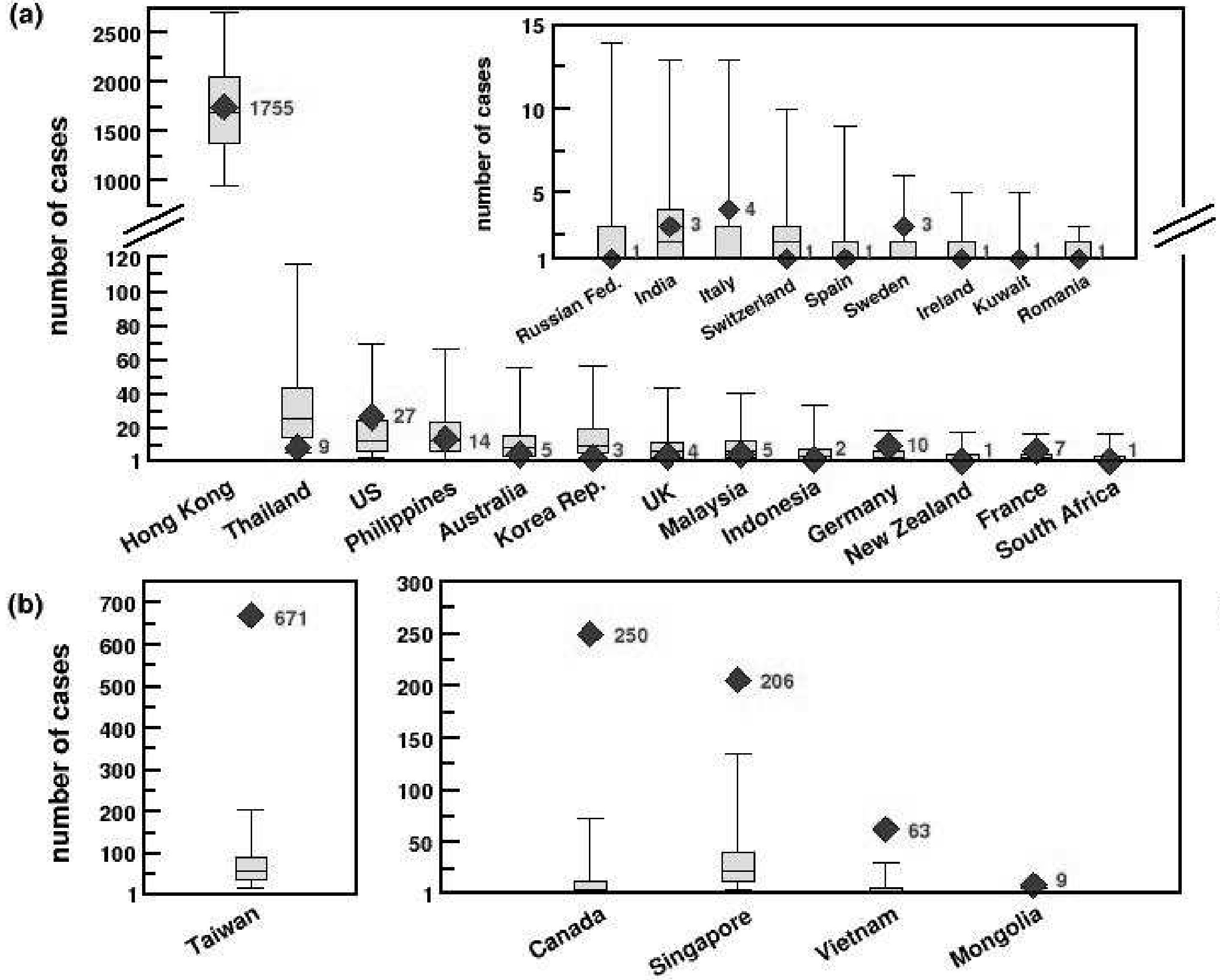


Predictions…

SARS - July 11, 2003



Quantitatively speaking



NB: populations of millions of individuals!!

Very good results because…