

## Lecture 9: Spreading Phenomena

**Prof. Sotiris Nikoletseas**

*University of Patras*

*DDCDM M.Sc., Patras*  
2020 - 2021

# Introduction

- On February 21, 2003 a physician from southern China checked into the Metropole Hotel in Hong Kong. He previously treated patients suffering from a typical pneumonia that later was renamed as Severe Acute Respiratory Syndrome (**SARS**). Several days later he died from the same disease.
- Other guests of this hotel contracted the same disease and carried it with them to their homelands. Half of the 8,100 documented cases of SARS were traced back to the Metropole Hotel. The physician who brought the virus to Hong Kong became an example of a **super-spreader**.
- From the perspective of Network Science, **super-spreaders** are considered as **hubs**.

## Super-spreaders

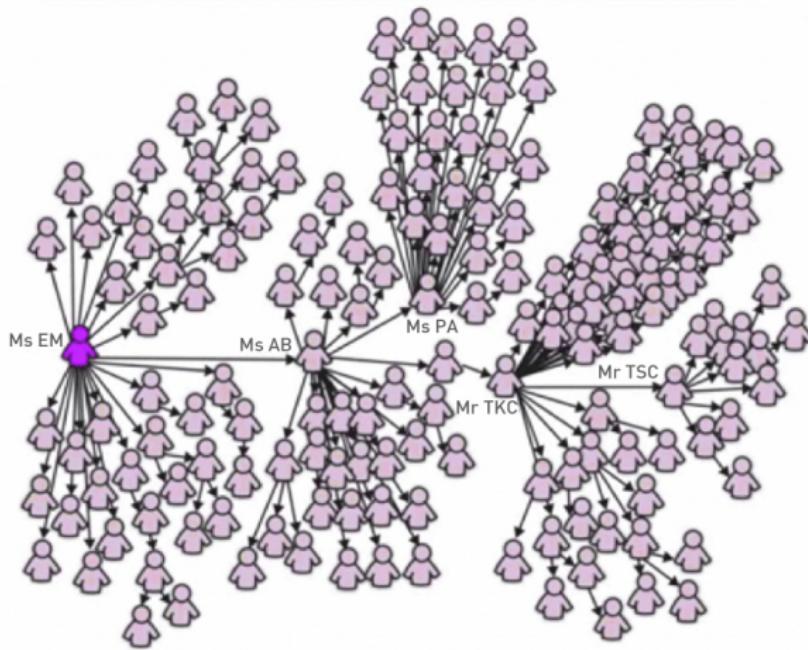


Figure: Five SARS patients spread out the disease to 144 individuals in Singapore

# The diversity of spreading phenomena

PHENOMENA	AGENT	NETWORK
Venereal Disease	Pathogens	Sexual Network
Rumor Spreading	Information, Memes	Communication Network
Diffusion of Innovations	Ideas, Knowledge	Communication Network
Computer Viruses	Malwares, Digital viruses	Internet
Mobile Phone Virus	Mobile Viruses	Social Network/Proximity Network
Bedbugs	Parasitic Insects	Hotel - Traveler Network
Malaria	Plasmodium	Mosquito - Human network

# Epidemic Modeling

- **How to model the spread of pathogens?**

Epidemiology has developed a framework that relies on two fundamental hypotheses:

- **Compartmentalization**

Each individual can be in one of three states or compartments: *Susceptible* (*S*), *Infectious* (*I*) and *Recovered* (*R*).

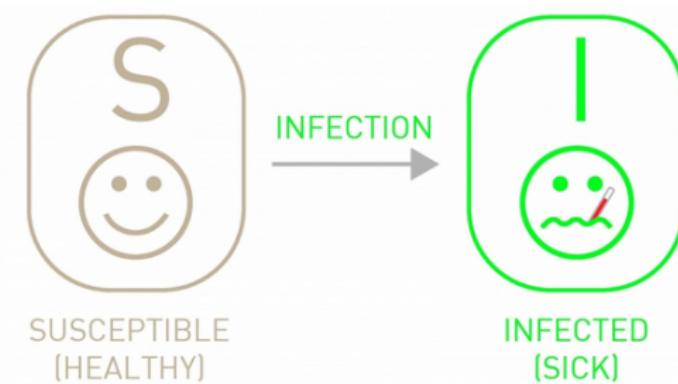
- **Homogeneous Mixing**

- Each individual has the same chance of coming into contact with an infected individual.
- There is no need of knowing the precise network topology on which a pathogen spreads.

- We will explore the dynamics of three most used epidemic models: **SI**, **SIS** and **SIR** models.

# Susceptible-Infected (SI) Model (I)

- $N$  individuals,  $S(t)$  number of susceptible individuals,  $I(t)$  number of infected individuals at time  $t$ . It is  $S(0) = N, I(0) = 0$ .
- We assume that each individual has  $\langle k \rangle$  contacts.
- We denote with  $\beta$  the likelihood the pathogen transmits from an infected to a susceptible individual at unit time.



## Susceptible-Infected (SI) Model (II)

- Assume  $I(0) = 1$ . **How many individuals will be infected at time  $t$  ?**
- An infected individual comes into contact with  $\langle k \rangle S(t)/N$  susceptible individuals in unit time where each one transmits the pathogen with rate  $\beta$ .
- The average number of new infections  $dI(t)$  during a timeframe  $dt$  is

$$dI(t) = \beta \langle k \rangle \frac{S(t)I(t)}{N} dt$$

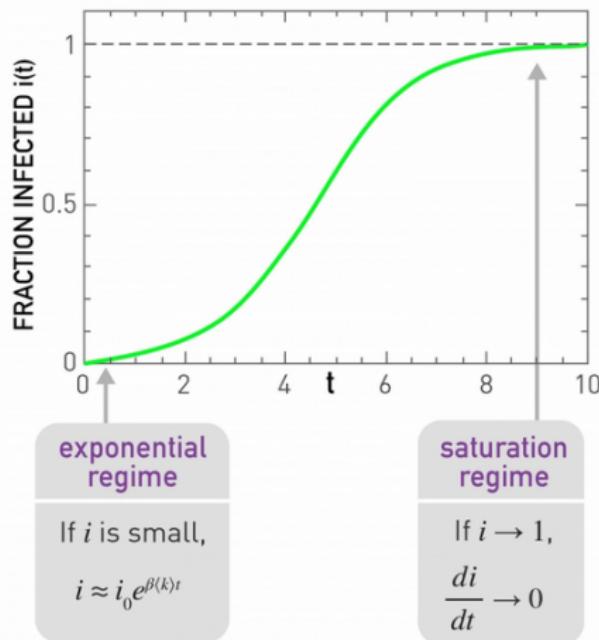
- Let  $s(t) = S(t)/N$  and  $i(t) = I(t)/N$ . We will denote these variables by  $s$  and  $i$  respectively. Re-writing the above equation:

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

- Let  $i_0 = i(t = 0)$ . The fraction of infected nodes increases in time as

$$i = \frac{i_0 e^{\beta \langle k \rangle t}}{1 - i_0 + i_0 e^{\beta \langle k \rangle t}}$$

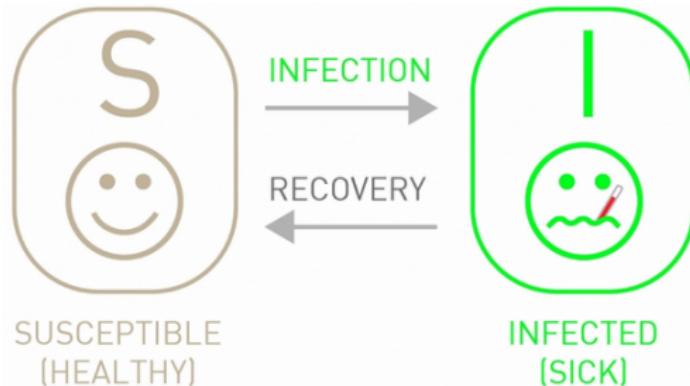
# Susceptible-Infected (SI) Model (III)



- at the beginning there are many susceptible people so spreading is exponential.
- with time, the spreading slows until all people are infected.

# Susceptible-Infected-Susceptible (SIS) Model (I)

- The infected individuals recover at a fixed rate  $\mu$ , becoming susceptible again.



- The equation describing the dynamics of this model is

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

where  $\mu$  is the recovery rate and the  $\mu i$  term captures the rate at which the population recovers from the disease.

# Susceptible-Infected-Susceptible (SIS) Model (II)

- The fraction of infected nodes increases in time as

$$i = \left(1 - \frac{\mu}{\beta \langle k \rangle}\right) \frac{C e^{(\beta \langle k \rangle - \mu)t}}{1 + C e^{(\beta \langle k \rangle - \mu)t}}$$

where the initial condition  $i_0 = i(t=0)$  gives  $C = \frac{i_0}{1 - i_0 - \frac{\mu}{\beta \langle k \rangle}}$ .

- At any moment only a finite fraction of the population is infected.
- The above equation predicts **two outcomes** in the **SIS** model:

- **Endemic State** ( $\mu < \beta \langle k \rangle$ )

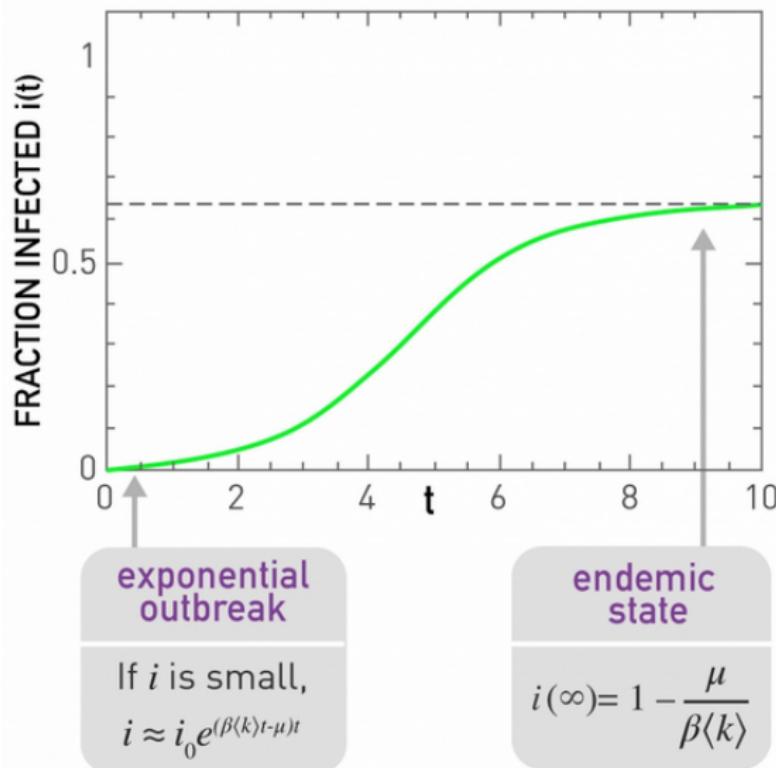
The number of newly infected individuals equals the number of individuals who recover from the disease. Hence the infected fraction of the population does not change with time ( $\frac{di}{dt} = 0$ ).

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

- **Disease-free State** ( $\mu > \beta \langle k \rangle$ )

The number of infected individuals **decreases** exponentially with time. Hence with time the pathogen disappears from the population.

# Susceptible-Infected-Susceptible (SIS) Model (III)



## Susceptible-Infected-Susceptible (SIS) Model (IV)

- The **SIS** model predicts that some pathogens will persist in the population while others die out shortly.
- To determine the outcome of an epidemic in this model, we define the **characteristic time**  $\tau$  of a pathogen as

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

that is the inverse of the speed of the pathogen spreading.

- We denote with  $R_0$  the **basic reproductive number** that represents the average number of susceptible individuals infected by an infected individual in a fully susceptible population.

$$R_0 = \frac{\beta \langle k \rangle}{\mu}$$

# Basic Reproductive Number $R_0$ (I)

$$\tau = \frac{1}{\mu(R_0 - 1)}$$

- If  $R_0 > 1$ , then  $\tau > 0$ . Each infected individual infects more than one healthy individual. The epidemic is in the **endemic state**.
- If  $R_0 < 1$ , then  $\tau < 0$ . Each infected individual infects less than one additional individual. The epidemic is in the **disease-free state**.
- The higher the  $R_0$  number of a pathogen, the faster is its spreading process in the population.

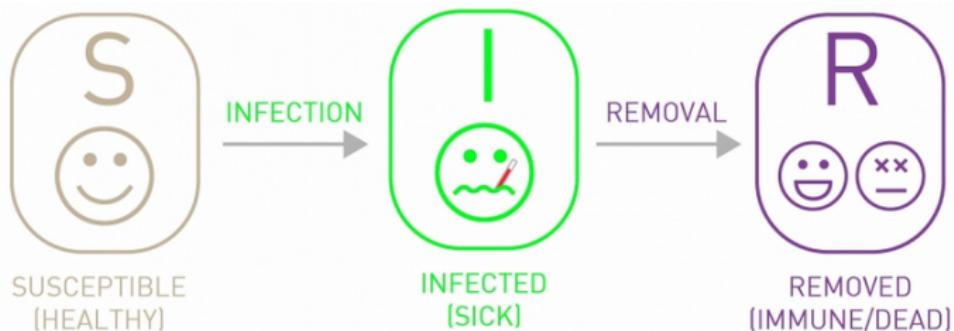
# Basic Reproductive Number $R_0$ (II)

DISEASE	TRANSMISSION	Ro
Measles	Airborne	12-18
Pertussis	Airborne droplet	12-17
Diphtheria	Saliva	6-7
Smallpox	Social contact	5-7
Polio	Fecal-oral route	5-7
Rubella	Airborne droplet	5-7
Mumps	Airborne droplet	4-7
HIV/AIDS	Sexual contact	2-5
SARS	Airborne droplet	2-5
Influenza (1918 strain)	Airborne droplet	2-3

Note:  $R_0$  (COVID 19)  $\simeq 2.9$ .

# Susceptible-Infected-Recovered (SIR) Model (I)

- For many pathogens, like most strains of influenza, individuals develop immunity after they recover from the infection.
- These individuals cannot be infected from the pathogen they recovered from, nor can they infect others.



- The fraction of infected individuals in the **SIR** model is given by the following equation:

$$di = (\beta \langle k \rangle i (1 - r - i) - \mu i) dt$$

where  $r$  is the fraction of the recovered individual that are removed from the susceptible population.

## Susceptible-Infected-Recovered (SIR) Model (II)

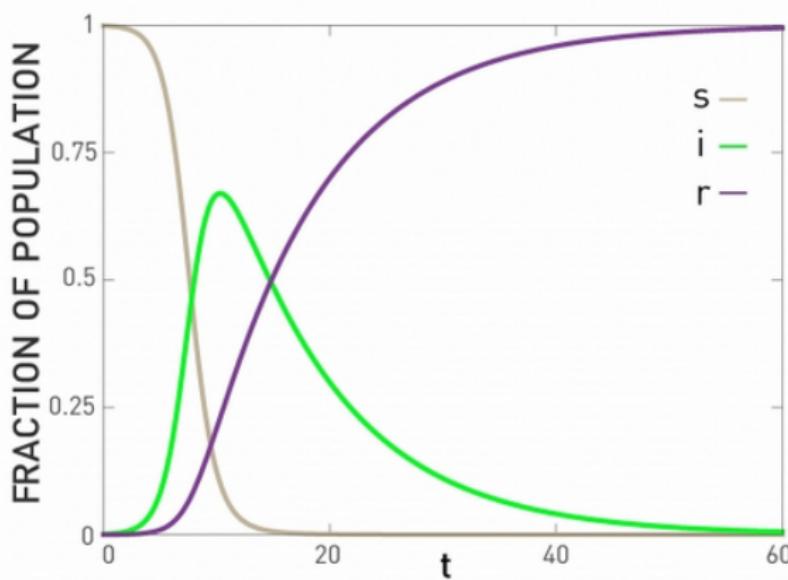
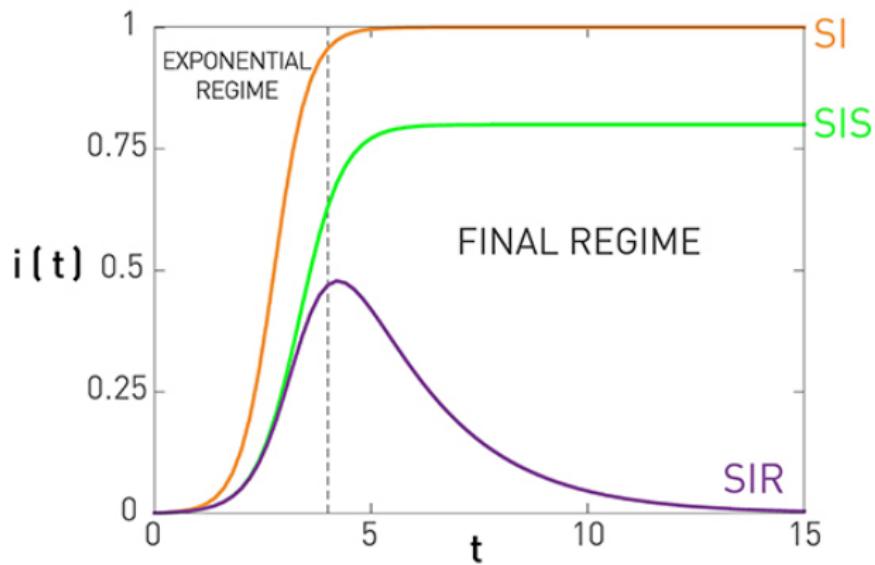


Figure: The time dependent behavior of  $s$ ,  $i$  and  $r$  in the SIR model.

# Comparison of the three epidemic models



**Figure:** The time dependent behavior of the fraction of infected individuals,  $i$ , in the SI, SIS and SIR models.

# Network Epidemics

- Epidemic modeling relies on the homogeneous mixing hypothesis and it also assumes that each individual has comparable number of contacts,  $\langle k \rangle$ .
- Both assumptions are false. Pathogens spread on complex contact networks as infected individuals transmit the pathogen only to the individuals they come into contact with.
- These contact networks are often scale-free, hence  $\langle k \rangle$  is not sufficient to characterize their topology. Therefore, **we need to explore the effects of network structure on epidemic spreading.**
- The **three basic epidemic models were extended** by *Romualdo Pastor-Satorras* and *Alessandro Vespignani* in 2001 to incorporate the topological characteristics of the underlying contact network a pathogen spreads on.

# Degree Block Approximation

- Individuals with more links are more likely to be in contact with an infected individual. Thus, they are more likely to be infected.
- The degree of each node must be considered as an implicit variable in the mathematical formalism of the model.
- This is achieved by **degree block approximation**. It assumes that nodes (individuals) with the same degree,  $k$ , behave similarly.
- Therefore, the fraction of infected nodes  $I_k$  with degree  $k$  among all  $N_k$  degree- $k$  nodes in the network is  $i_k = I_k/N_k$ .
- The total fraction of infected nodes,  $i$ , is the sum of all infected degree- $k$  nodes:

$$i = \frac{\sum_k I_k}{N} = \frac{\sum_k p_k \cdot N \cdot i_k}{N} = \sum_k p_k i_k$$

# Network Epidemics - SI model (I)

- In the SI model,  $\tau^{SI}$  is the characteristic time of the spread of the pathogen

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

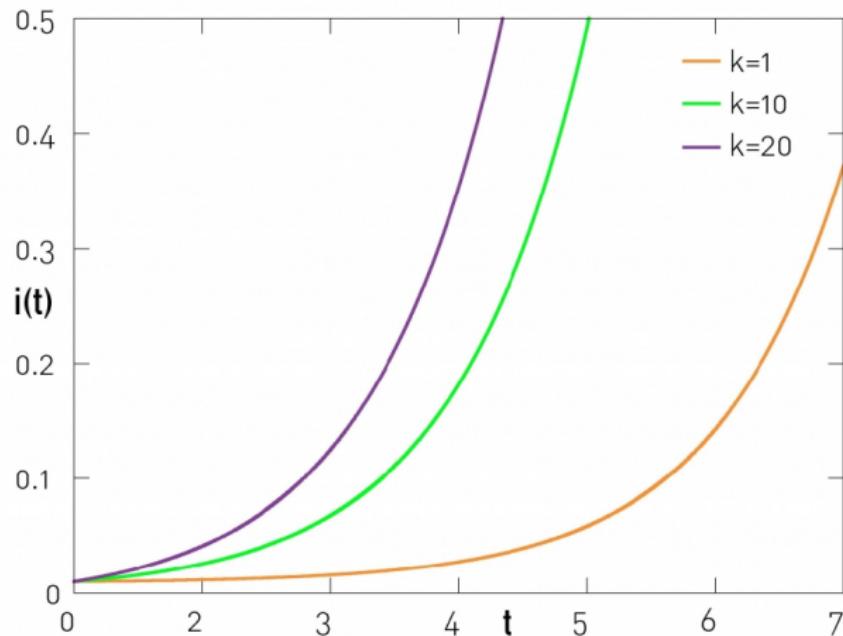
- Then, the fraction of infected degree- $k$  nodes is

$$i_k = i_0 \left( 1 + \frac{k(\langle k \rangle - 1)}{\langle k^2 \rangle - \langle k \rangle} \left( e^{t/\tau^{SI}} - 1 \right) \right)$$

- Finally, the total fraction of infected nodes is calculated by integrating over all the  $k_{max}$  degrees

$$i = \int_0^{k_{max}} i_k p_k dk$$

## Network Epidemics - SI model (II)



**Figure:** Fraction of infected nodes in the SI Model. At any time the fraction of high degree nodes that are infected is higher than the fraction of low degree nodes.

# Network Epidemics - SI model (III)

- The characteristic time

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

depends not only  $\langle k \rangle$  but also on the network's heterogeneity  $\langle k^2 \rangle$ .

- **Random Network** For a random network  $\langle k^2 \rangle = \langle k \rangle(\langle k \rangle + 1)$ .  
Hence,  $\tau^{SI}$  is finite and greater than zero

$$\tau_{ER}^{SI} = \frac{1}{\beta \langle k \rangle}$$

The same behavior is predicted for **scale-free networks** with  $\gamma \geq 3$ .

- **Scale-free Networks with  $\gamma \leq 3$**  As  $N \rightarrow \infty$ ,  $\langle k^2 \rangle \rightarrow \infty$ .  
Hence,  $\tau^{SI} \rightarrow 0$ , meaning that the spread of a pathogen is **instantaneous**.  
Hubs play a significant role in the spreading of a pathogen. Once they are infected, they become **super-spreaders**.
- Generally, in **inhomogenous networks** where  $\langle k^2 \rangle > \langle k \rangle(\langle k \rangle + 1)$ ,  $\tau^{SI}$  is reduced and spreading is fast.

# Network Epidemics - SIS model

- The characteristic time for the spread of a pathogen is estimated as

$$\tau^{SIS} = \frac{\langle k \rangle}{\beta \langle k^2 \rangle - \mu \langle k \rangle}$$

- We observe that  $i_k$  decays exponentially when the recovery rate  $\mu$  has a sufficiently large value. However, this condition also depends on the heterogeneity of the network through  $\langle k^2 \rangle$ .
- Therefore, in order to predict if a pathogen persists in the population, a metric called **spreading rate** is defined as  $\lambda = \frac{\beta}{\mu}$ .
- The spreading rate  $\lambda$  depends only on the biological characteristics of the pathogen,  $\mu$  and  $\beta$ .

# Epidemic Threshold - SIS model (I)

- A pathogen persists in the population only if its characteristic time  $\tau$  is non-negative.
- **Random Network**

$$\tau_{ER}^{SIS} = \frac{1}{\beta(\langle k \rangle + 1) - \mu} > 0$$

This implies that

$$\lambda > \frac{1}{\langle k \rangle + 1}$$

Hence, the pathogen spreads in a random network only if its spreading rate is above the **epidemic threshold**

$$\lambda_c = \frac{1}{\langle k \rangle + 1}$$

As  $\lambda_c$  is always nonzero, a pathogen will disappear from the population when  $\lambda < \lambda_c$ .

# Epidemic Threshold - SIS model (II)

- **Scale-free Network**

$$\tau^{SIS} = \frac{\langle k \rangle}{\beta \langle k^2 \rangle - \mu \langle k \rangle} > 0$$

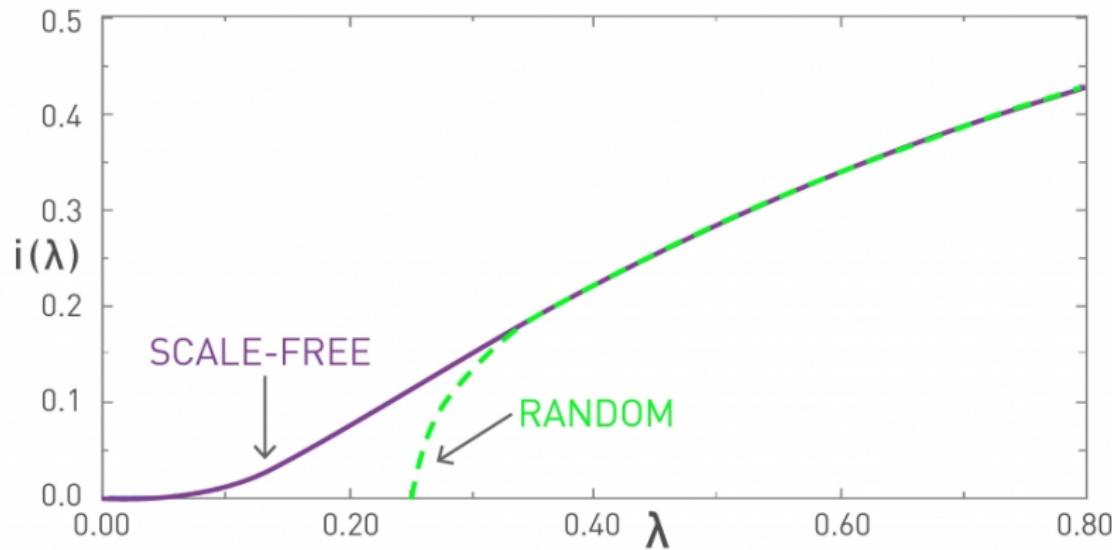
This implies that

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

For large networks ( $N \rightarrow \infty$ ),  $\langle k^2 \rangle \rightarrow \infty$ . Hence,  $\lambda_c \rightarrow 0$  (**vanishing epidemic threshold**).

- This means that even pathogens with small spreading rate  $\lambda$  can persist in the population. This is a consequence of hubs' ability to transmit a pathogen to a large number of other nodes, once they get infected.

## Epidemic Threshold - SIS model (III)



**Figure:** The fraction of infected individuals  $i(\lambda) = i(t \rightarrow \infty)$  in the endemic state of the SIS model.

# Network Epidemics - SIR model

- The characteristic time for the spread of a pathogen is estimated as

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

- Hence, the epidemic threshold is

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

# Contact Networks - Diseases (I)

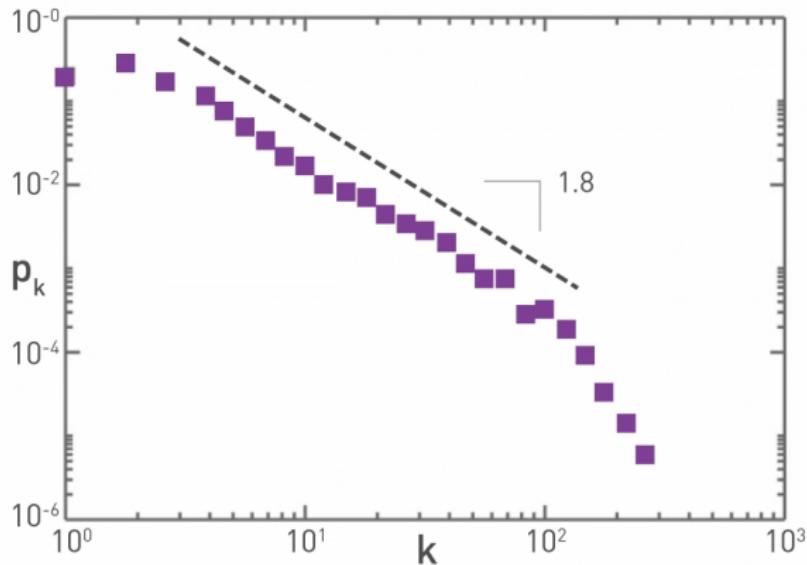
- **Sexually Transmitted Diseases**

- The HIV virus spreads through the contact network that captures who had sexual relationship with whom.
- A survey of the sexual habits of the Swedish population revealed the **scale-free nature of the sexual network**.
- Hence, the sexual network enhances the spreading of HIV virus as it lowers both  $\tau$  and  $\lambda_c$ .

- **Airborne Diseases**

- Airborne diseases, like influenza, SARS and H1N1, spread on the contact network that captures the set of individuals a person comes into physical proximity.
- The structure of this network is explored at two levels: **the global travel network** and **local contact patterns**.
- The **air transportation network** plays a significant role in modeling the spread of pathogens worldwide.
- Also, many airborne pathogens spread through the **location network** that its scale-free nature enhances the spreading process.

## Contact Networks - Diseases (II)



**Figure:** The degree distribution of the air transportation network is well approximated by a power-law.

# Contact Networks - Digital Viruses

- **Computer Viruses**

- Many computer viruses spread as email attachments through the **email network** which is scale-free.
- Other viruses exploit various communication protocols spreading through the scale-free **internet network**.

- **Mobile Viruses**

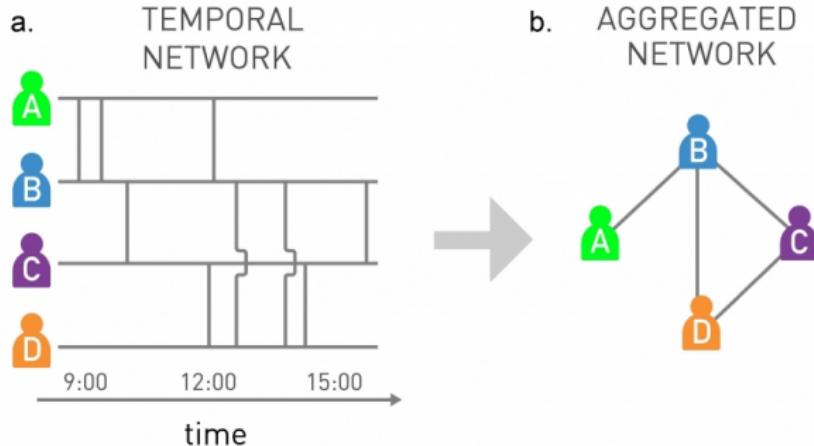
- Bluetooth and MMS technologies are two ways for the spreading of mobile phone viruses.
- Both **location network** and **mobile communications network**, respectively, are scale-free with high  $\langle k^2 \rangle$ .

# Beyond the degree distribution

- The degree distribution is not enough to capture all the characteristics of real networks.
- Real complex networks have a number of properties: **contact burstiness, communities, weighted or temporal networks.**
- Such characteristics should be taken into account when predicting the spreading of a pathogen in the underlying network.

# Temporal Networks

- Most interactions between nodes in social networks are infrequent and they have a **finite duration**.



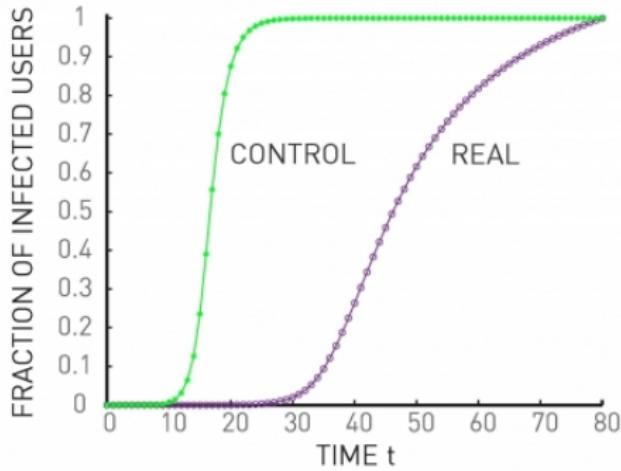
**Figure: Temporal Network:** There is a temporal path from A to D. Hence, a pathogen cannot spread from D to A. **Aggregated Network:** The pathogen can reach all individuals independent of its starting point; This fallacy is the result of disregarding the temporal dimension.

# Bursty Contact Patterns

- The epidemic models assume that the timing of interactions between two individuals are **random** following a **Poisson** distribution.
- However, in most social networks, e.g., email and mobile communications networks, the inter-event times between consecutive contacts follow a **power law** distribution (periods of frequent interactions but also very long gaps).
- Bursty contact patterns alter the dynamics of the spreading process of the pathogen by **increasing the characteristic time  $\tau$** . Thus, the spread is slower and decay takes more time.

# Communities

- Nodes in the same community interact repeatedly leading to tie strengths between them. When an individual gets infected, it will pass the pathogen to individuals that spend more time with high probability.



**Figure:** In a **control network** where all link weights are equal the pathogen spreads more quickly than in a **real network**. The reduced speed observed in the real network indicates that the pathogen is trapped within communities (yet the virus spreads very fast and widely within each community).

# Immunization

- Immunization strategies are guided by an important prediction of the traditional epidemic models: If a pathogen's spreading rate  $\lambda$  is reduced under its critical threshold  $\lambda_c$ , the virus naturally dies out.
- However, in scale-free networks with the vanishing epidemic threshold, immunization strategies can not move  $\lambda$  under  $\lambda_c$ .
- Therefore, immunization strategies must consider the underlying network topology a pathogen spreads to effectively counter the impact of the vanishing epidemic threshold.

# Random Immunization

- A randomly selected  $g$  fraction of individuals are immunized.
- Assuming that the pathogen follows the SIS model, the effective degree of a susceptible node changes from  $\langle k \rangle$  to  $\langle k \rangle(1 - g)$ .
- Consequently, the spreading rate  $\lambda$  decreases to  $\lambda' = \lambda(1 - g)$
- **Random Networks** The fraction of the immunized individuals that at least is needed to push  $\lambda$  under  $\lambda_c$  is

$$g_c = 1 - \frac{\mu}{\beta} \frac{1}{\langle k \rangle + 1}$$

- **Heterogenous Networks** For a pathogen spreading on a network with high  $\langle k^2 \rangle$ , the immunization threshold is

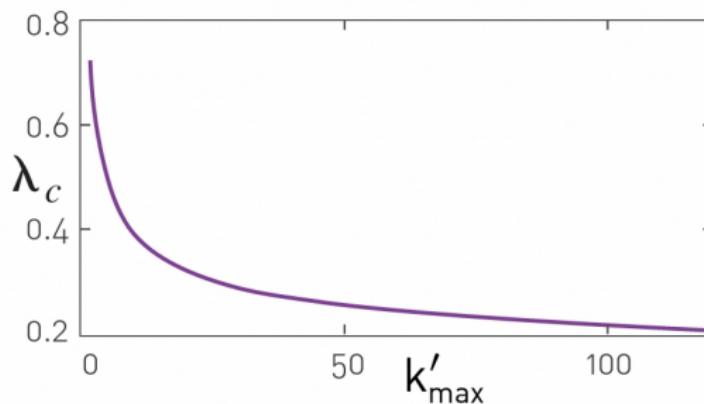
$$g_c = 1 - \frac{\mu}{\beta} \frac{\langle k \rangle}{\langle k^2 \rangle}$$

For a scale-free network with  $\gamma < 3$ ,  $g_c \rightarrow 1$ . In other words, we need to immunize virtually all nodes to stop the epidemic.

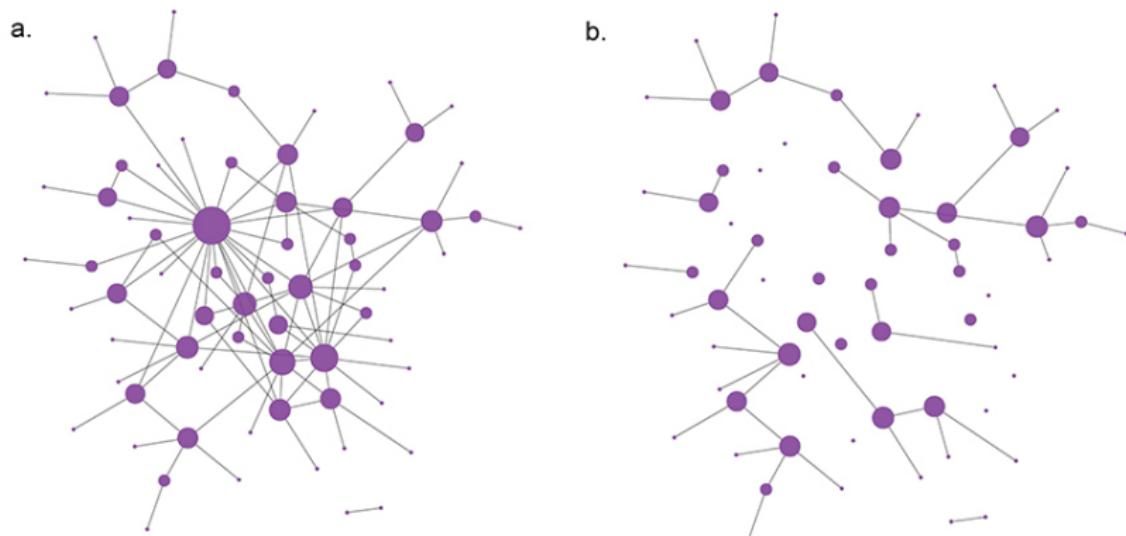
# Vaccination Strategies in Scale-Free Networks (I)

- A way to eradicate the transmission of a pathogen in scale-free networks is to increase the epidemic threshold.
- This can be achieved by reducing the heterogeneity  $\langle k^2 \rangle$  of the underlying network, i.e., by **immunizing the hubs whose degree exceeds a degree threshold  $k'_{max}$** .
- If we immunize all nodes with degrees  $k > k'_{max}$ , the  $\lambda_c$  changes to

$$\lambda'_c \approx \frac{\gamma - 2}{3 - \gamma} \frac{k_{min}^{2-\gamma}}{(k'_{max})^{\gamma-3}}$$

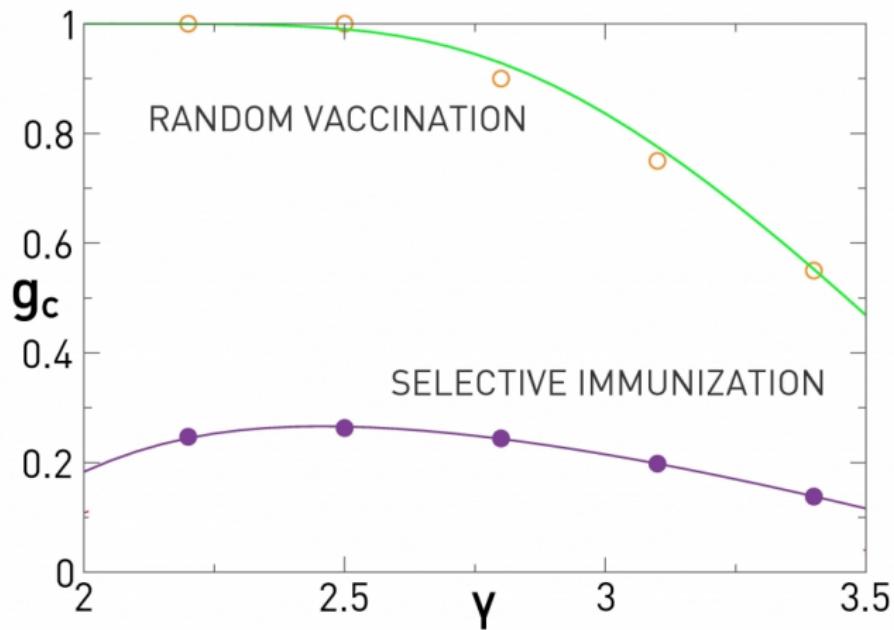


# Vaccination Strategies in Scale-Free Networks (II)



**Figure:** By immunizing the hubs, we are fragmenting the contact network, making more difficult for the pathogen to reach the nodes in other components.

# Random vs Selective Immunization



**Figure:** The critical immunization threshold  $g_c$  in function of the degree exponent  $\gamma$  of the contact network on which the pathogen spreads following the SIS model.

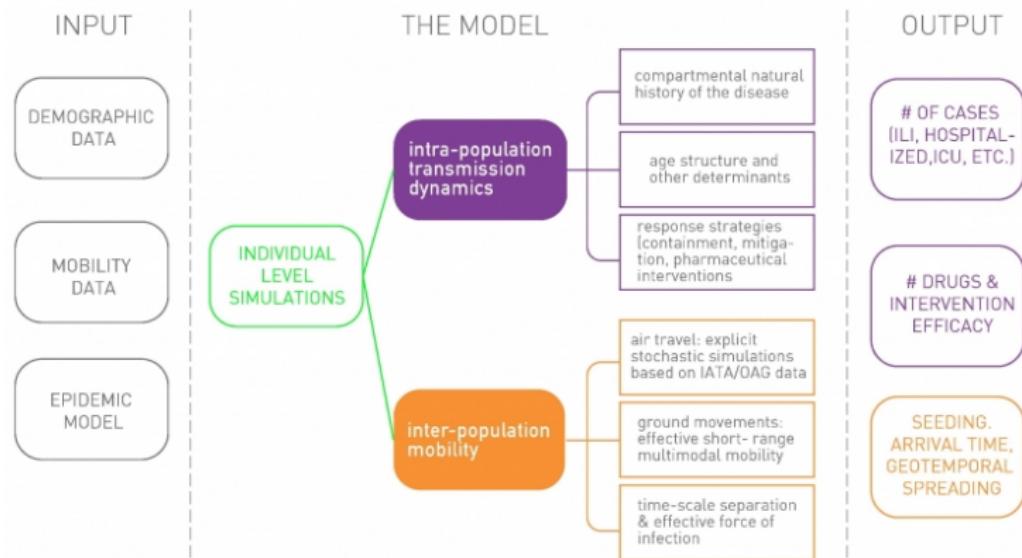
# How to halt an epidemic?

- Some of the most common interventions safety officials should rely on include:
  - **Transmission-Reducing Interventions**, such as face masks, gloves, hand washing and condoms
  - **Contact-Reducing Interventions**, such as quarantine patients and closing down frequently visited public spaces
  - **Vaccinations**, Nevertheless, there have not been developed effective vaccines for all known pathogens.
  - **Epidemic prediction tools** are required for the effectiveness of these interventions.

# Epidemic Prediction

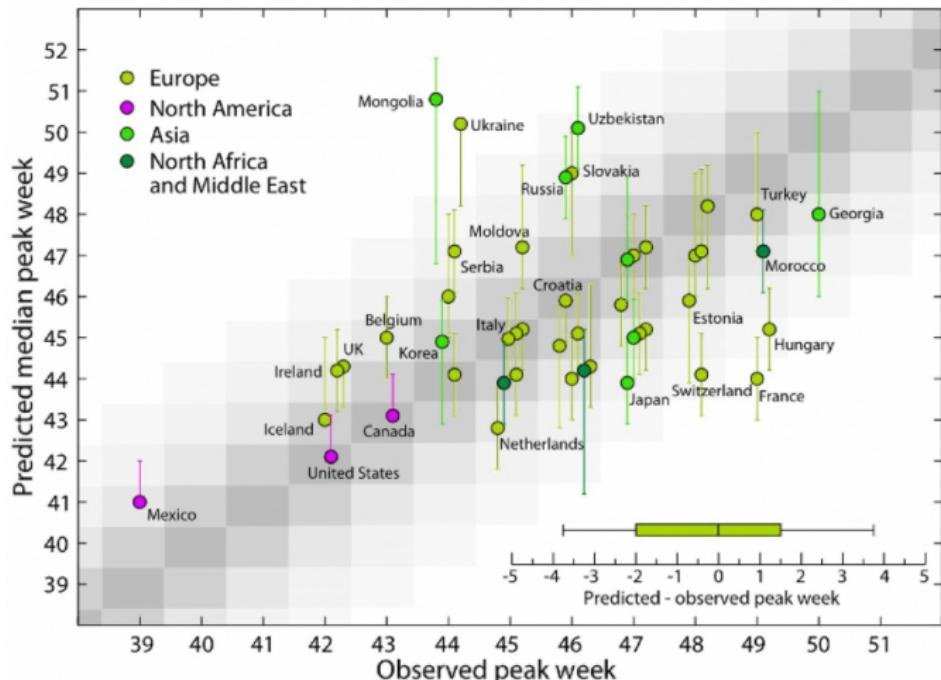
- **Real-Time Forecast**

- The first successful real time pandemic forecast based on network science relied on the **Global Epidemic and Mobility (GLEAM)** computational model.



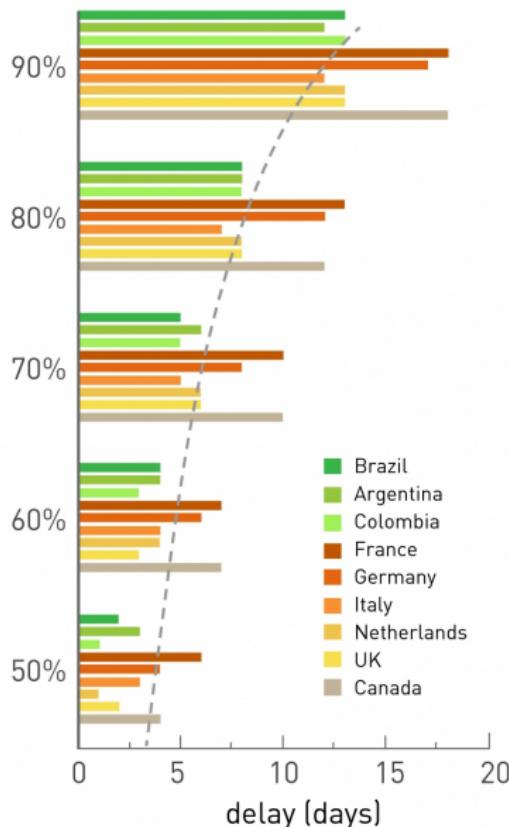
**Figure:** The flowchart of the GLEAM computational model for predicting the real-time spread of pathogens.

# Peak Time of an Epidemic



**Figure:** The observed and predicted peak time estimated by GLEAM for the H1N1 virus in several countries. Peak time is the week when most individuals are infected.

# Travel Reduction



*Figure: The impact of travel reduction on the arrival time of the H1N1 virus from Mexico to various countries, compared with the reference scenario of no travel reduction. The percentages show the degree of travel reduction implemented around the world.*

# Effective Distance

- It is obvious that geographic locations that are nearby to an infected region are more likely to get infected.
- Today, with airline travel, geographic distance has lost its relevance for epidemic phenomena.
- **Effective distance** derived by the mobility network is proposed instead to view the spread of an epidemic and it is defined as

$$d_{ij} = (1 - \ln p_{ij}) \geq 0$$

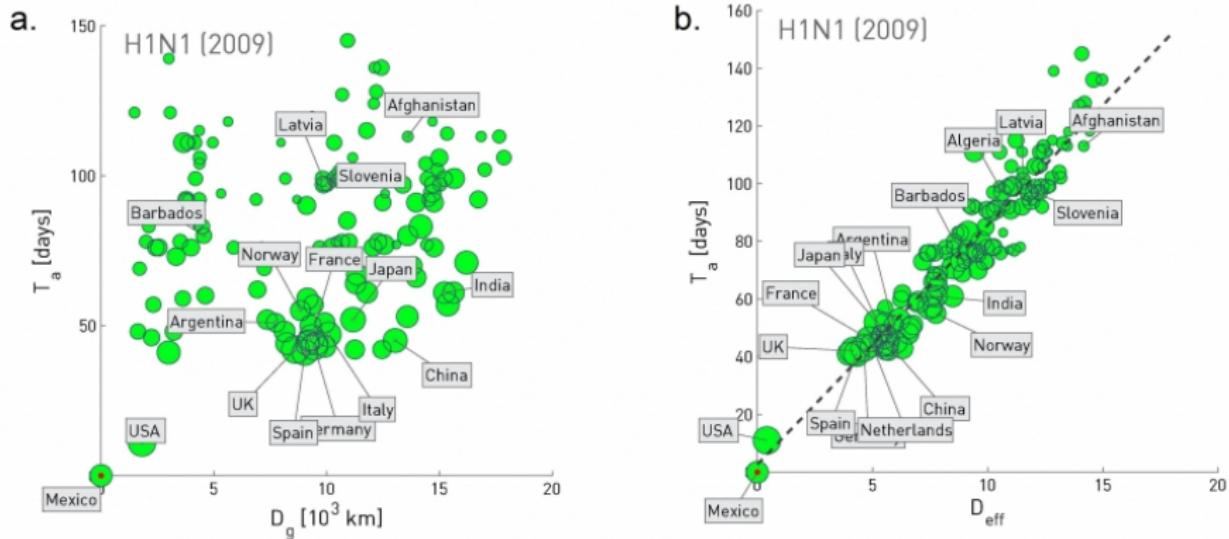
where  $p_{ij}$  represents the fraction of individuals that travel from node  $i$  to  $j$  and  $d_{ij} \neq d_{ji}$

- The **arrival time**  $T_\alpha$  of a pathogen depends on the effective distance  $d_{eff}$  as

$$T_\alpha = \frac{d_{eff}(P)}{V_{eff}(\beta, R_0, \gamma, \epsilon)}$$

where  $V_{eff}$  is the **effective speed** of the pathogen.

# Arrival Time (I)



**Figure:** The arrival time of H1N1 appears to be random if plotted in function of the physical distance, but it correlates strongly with the effective distance

## Arrival Time (II)

- The arrival time of the pathogen to a location is difficult to be measured at the beginning of an epidemic outbreak as the epidemiological parameters of the pathogen are not known.
- Yet, the **relative arrival time** from a node  $i$  to nodes  $j$  and  $l$  is independent from the epidemiological parameters

$$\frac{T_\alpha(j/i)}{T_\alpha(l/i)} = \frac{d_{eff}(j/i)}{d_{eff}(l/i)}$$

- Therefore, the mobility patterns which are unique and model-independent are sufficient to predict the spread of an epidemic.
- Recent advances in data collection and network epidemics have offered the capability to predict the real-time spread of a pathogen.

# Acknowledgement

An initial version of this lecture was nicely prepared by Konstantinos Tsinganos, an excellent student of the 2019-2020 class of the YDA postgraduate program.