

# Digital Media and Social Networks

*Laurissa Tokarchuk*

[www.eecs.qmul.ac.uk/~laurissa](http://www.eecs.qmul.ac.uk/~laurissa)

[laurissa.tokarchuk@qmul.ac.uk](mailto:laurissa.tokarchuk@qmul.ac.uk)



## Lecture 7: Epidemic Spreading



\*some slides copyright Hamed Haddadi (QMUL) and Cecilia Mascolo (Cambridge)

## In this Lecture

- In this lecture we introduce the process of spreading epidemics in networks.
  - This has been studied widely in biology.
  - But it also has important parallels in information/idea diffusion in networks.



# Epidemics vs Cascade Spreading

- ❑ In cascade spreading nodes make *decisions* based on pay-off benefits of adopting one strategy or the other.
- ❑ In epidemic spreading
  - Lack of decision making.
  - Process of contagion is complex and unobservable
    - In some cases it involves (or can be modeled as) randomness.

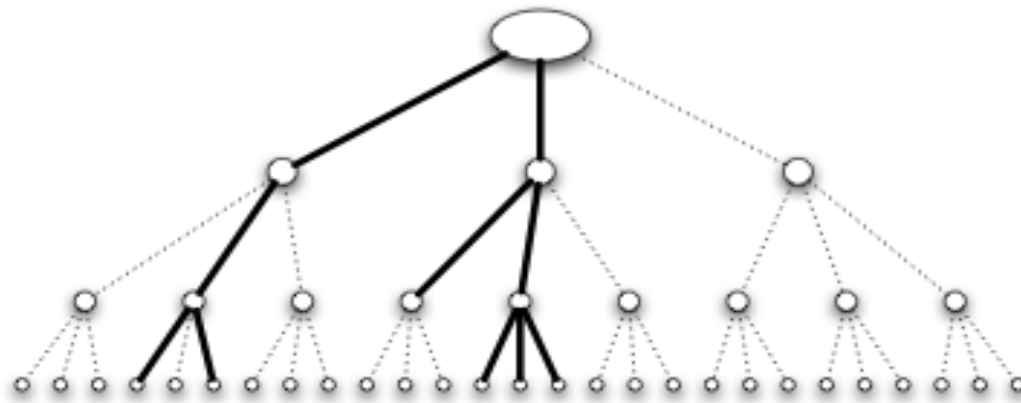


# Branching Process

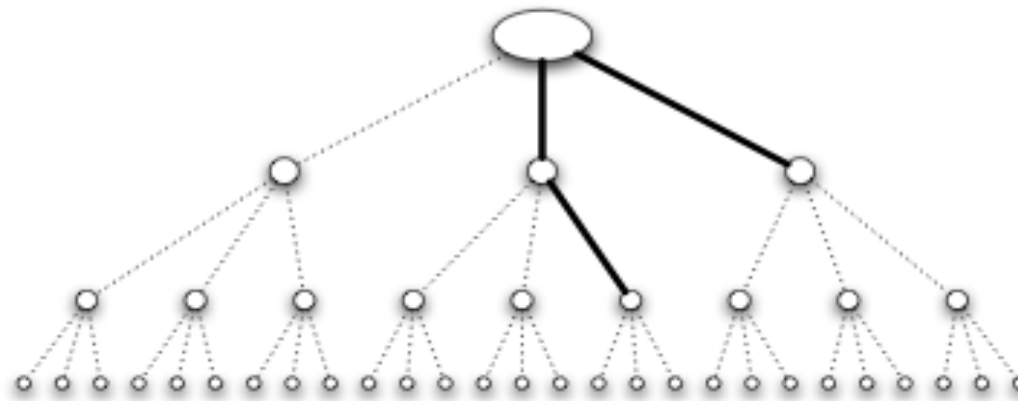
- Simple model.
- **First wave:** A person carrying a disease enters the population and transmits to all he meets with probability  $p$ . He meets  $k$  people: a portion of which will be infected.
- **Second wave:** each of the  $k$  people goes and meet  $k$  different people. So we have a second wave of  $k \times k = k^2$  people.
- **Subsequent waves:** same process.



## Example with $k=3$



High contagion probability:  
The disease spreads



Low contagion probability:  
The disease dies out



# Basic Reproductive Number

- Basic Reproductive Number  $R_0 = p * k$ 
  - $p$  = infection probability,  $k$  = branching number
  - It determines if the disease will spread or die out.
- In the branching process model, if  $R_0 < 1$  the disease will die out after a finite number of waves. If  $R_0 > 1$ , with probability  $> 0$ , the disease will persist by infecting at least one person in each wave.



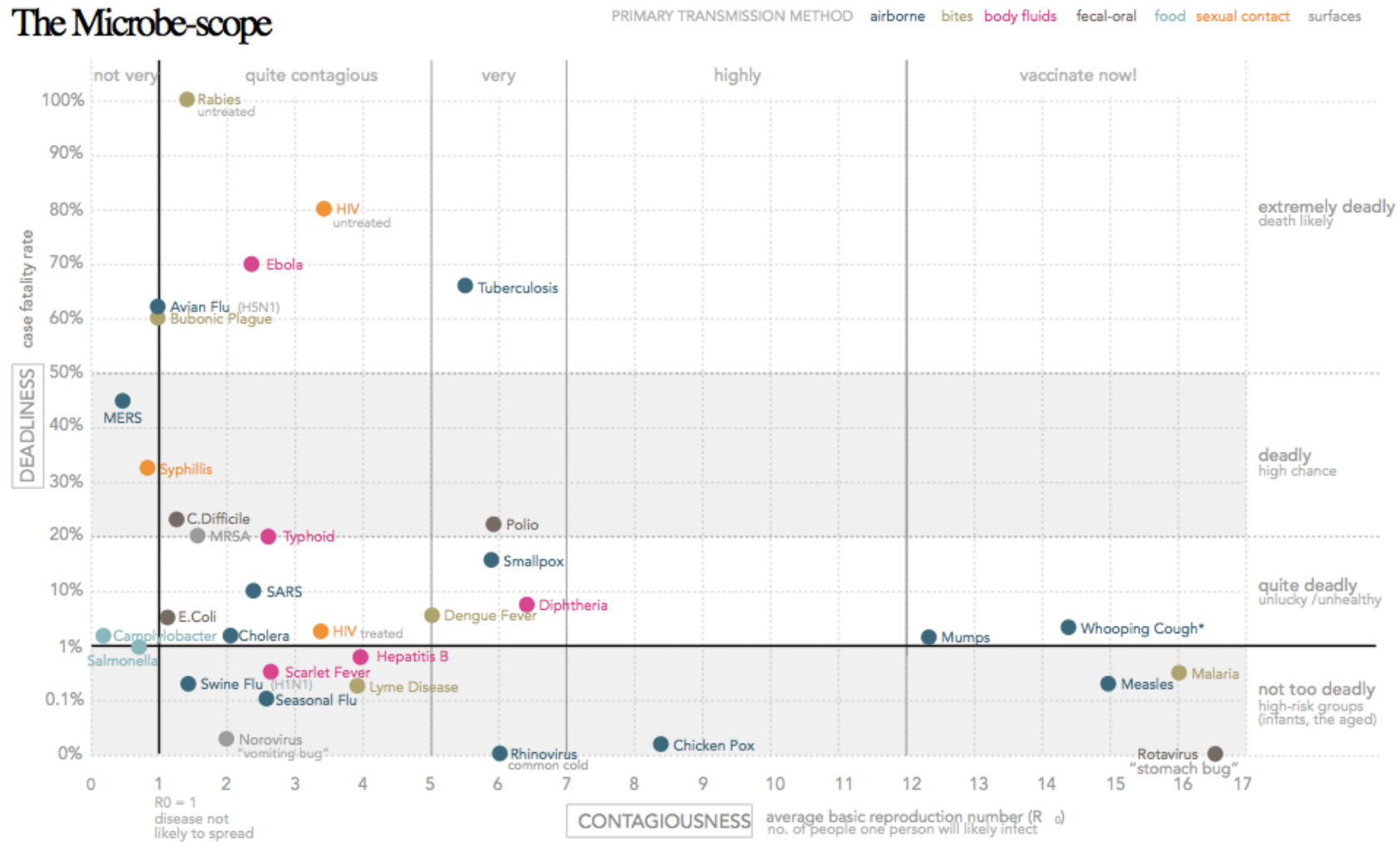
## Measures to limit the spreading

- When  $R_0$  is close 1, slightly changing  $p$  or  $k$  can result in epidemics dying out or happening.
  - Quarantining people/nodes reduces  $k$ .
  - Encouraging better sanitary practices reduces germs spreading [reducing  $p$ ].
- Limitations
  - No realistic contact networks: no triangles!
  - Nodes can infect only once.
  - No nodes recover.



# How contagious is Ebola?

## The Microbe-scope

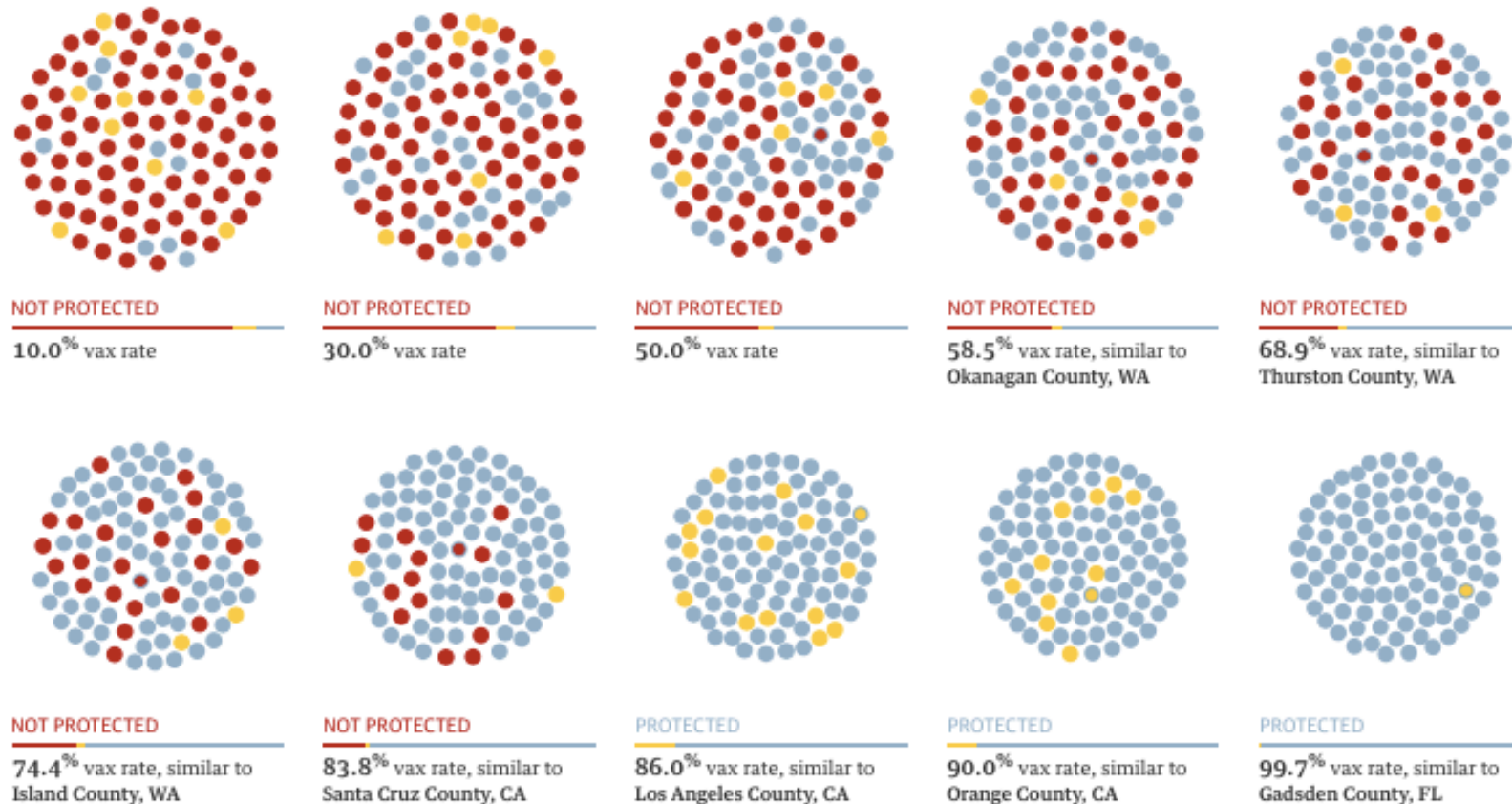


[http://infobeautiful2.s3.amazonaws.com/svg\\_microbescope.svg](http://infobeautiful2.s3.amazonaws.com/svg_microbescope.svg)



# Effect of vaccination on the herd

 vaccinated
  susceptible
  vaccinated but susceptible
  infected
  contact with an infected person

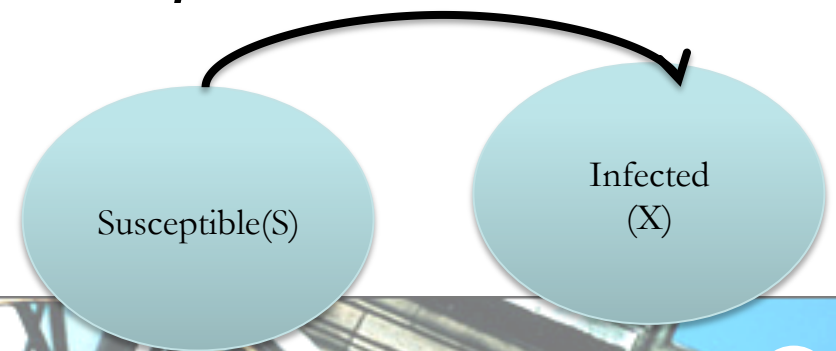


<http://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated>

# Formal Epidemics Models

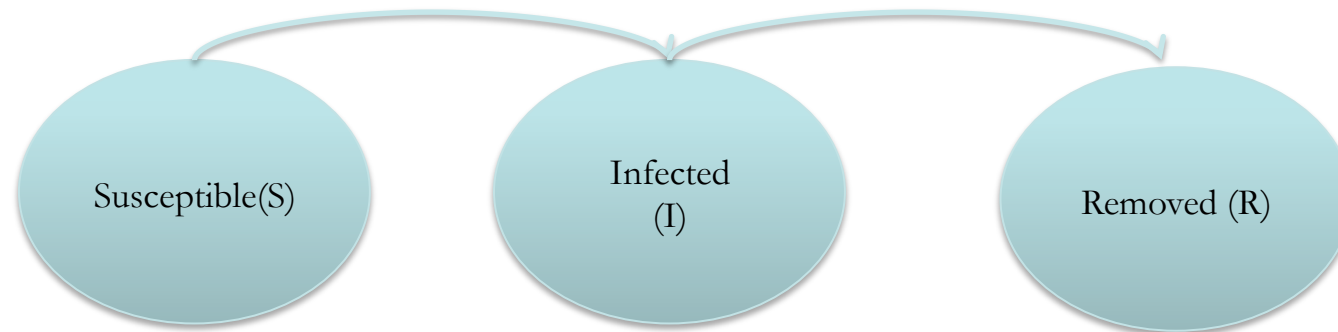
## The SI Model

- S: susceptible individuals.
- X: infected individuals, when infected they can infect others continuously.
- n: total population.
- $\langle k \rangle$  average contacts per individual
- $\beta = \lambda \langle k \rangle$  is the infection rate per individual ( $0 \leq \lambda \leq 1$ )
- Susceptible contacts per unit of time  $\beta S/n$ .
- Overall rate of infection  $X\beta S/n$ .



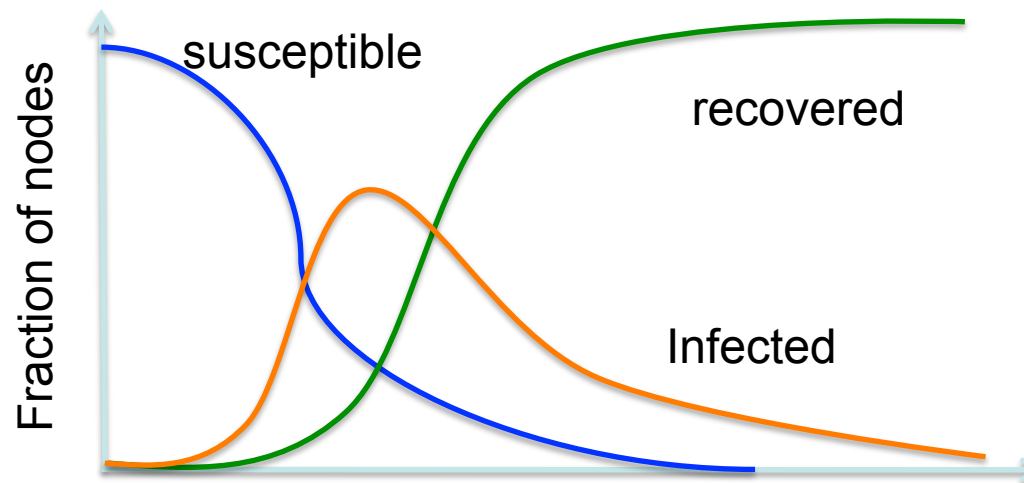
# SIR Model

- Infected nodes recover at a rate  $\gamma$ .
- A node stays infected for  $\tau$  time.
- Branching process is SIR with  $\tau=1$ .

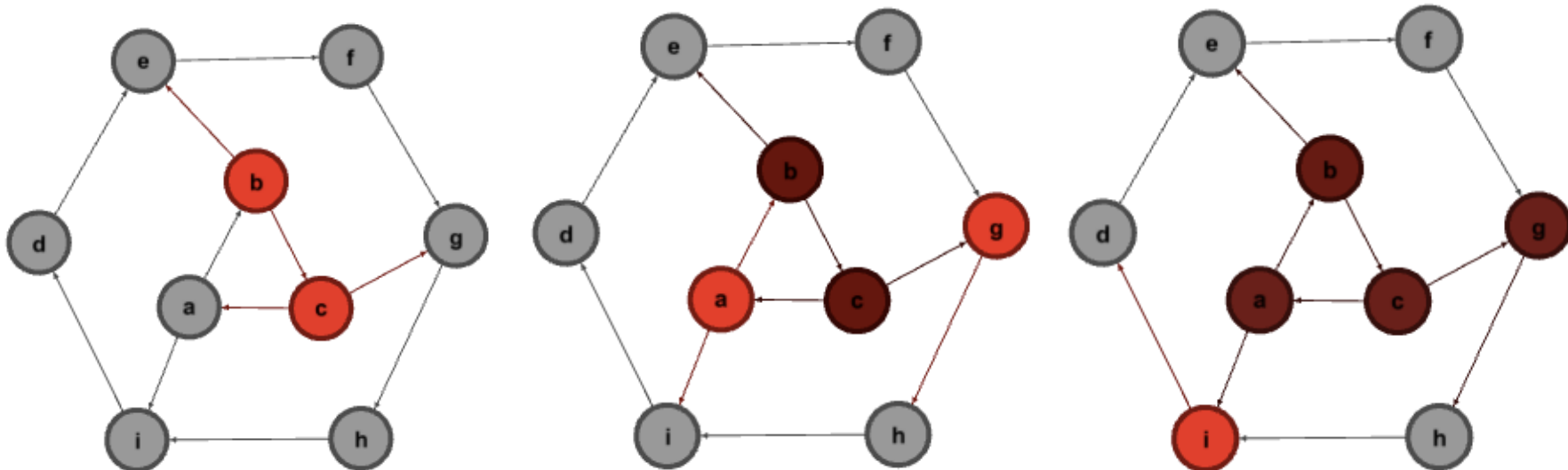


# Example

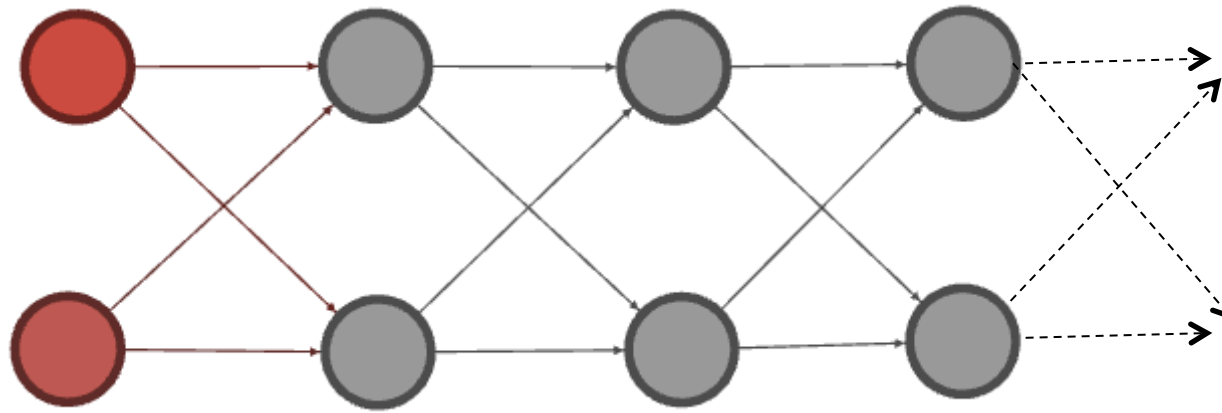
- Numerical examples of solution:
- $\beta=1$ ,  $\gamma=0.4$ ,  $s(\text{at start})=0.99$ ,  $x(\text{at start})=0.01$ ,  $r(\text{at start})=0$



# SIR epidemic ( $t_1 = 1$ , b and c initially infectious)



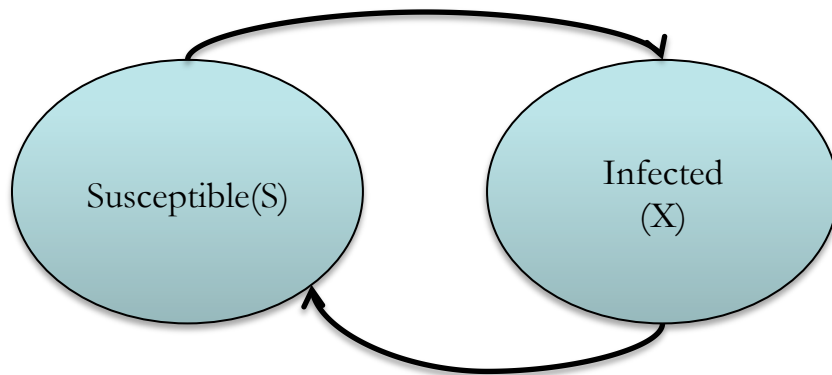
## Network structure and $R \geq 1$



- $p = 2/3, k = 2 \Rightarrow R_0 = 4/3$ 
  - 1/3 probability that the disease will fail to transmit on any edge
  - $(1/3)^4$  that it will fail on all 4 edges
- ***Different networks structures affect contagion.***

# SIS Model

- If  $\beta > \gamma$  growth curve like in SI but never reaching all population infected. The fraction of infected  $\rightarrow 0$  as  $\beta$  approaches  $\gamma$ .
- If  $\beta < \gamma$  the infection will die out exponentially.
- SIS has the same  $R_0$  as SIR.





# Epidemic Threshold

- When would the epidemic develop and when would it die out?
- It depends on the relationship of  $\beta$  and  $\gamma$ :
  - Basic Reproductive Number  $R_0 = \beta/\gamma$
  - If the infection rate [per unit of time] is higher than the removal rate the infection will survive otherwise it will die out.
  - In SI,  $\gamma=0$  so the epidemics always happen.





# Limitations of SIR

- ❑ Contagion probability is uniform and “on-off”
- ❑ Extensions
  - Probability  $q$  of recovering in each step.
  - Infected state divided into intermediate states (early, middle and final infection times) with varying probability during each.
  - **We have assumed homogenous mixing** : assumes all nodes encounter each others with same probability: we could assume different probability per encounter.



# Relaxing Assumptions

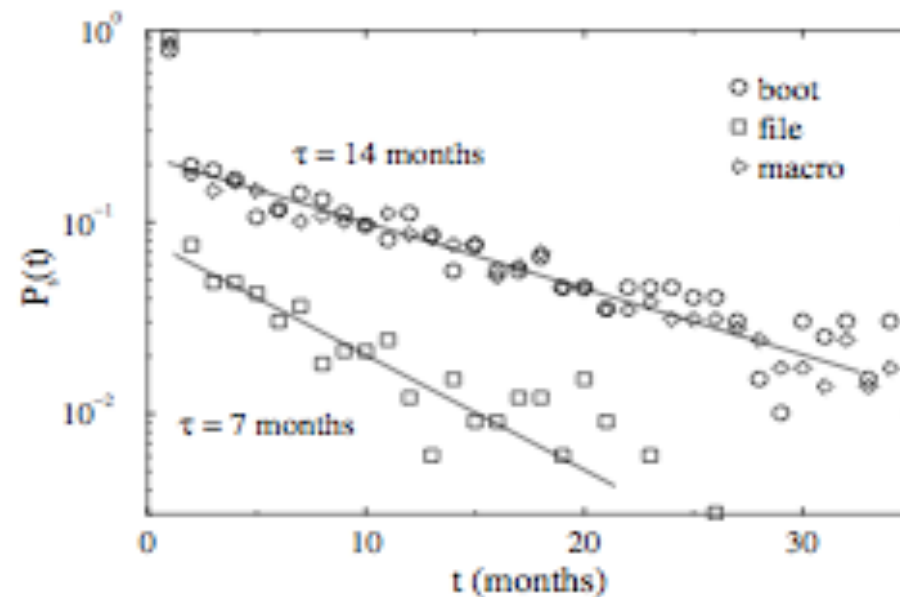
- ❑ Homogeneous Mixing: a node connects to the same average number of other nodes as any other.
- ❑ Most real networks are not Erdos-Renyi random networks (for which the homogeneous mixing assumption holds).
- ❑ Most networks have heterogeneous degree distributions.
  - Scale free networks!



# Would the model apply to SF?

- Pastor-Satorras and Vespignani [2001] have considered the life of computer viruses over time on the Internet:

Surviving probability of virus



**Virus survived on average 6-9/14 months depending on type**

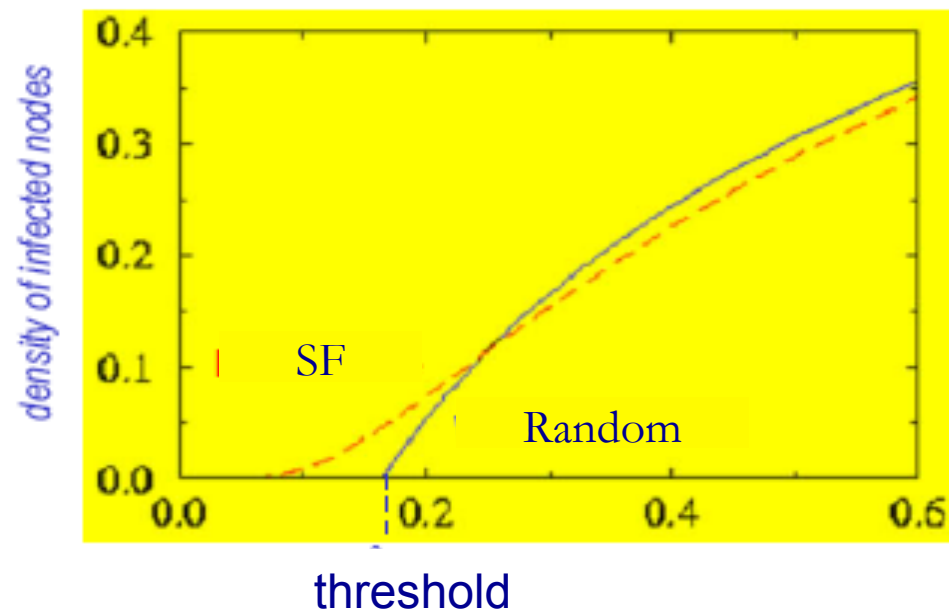


## How to justify this survival time?

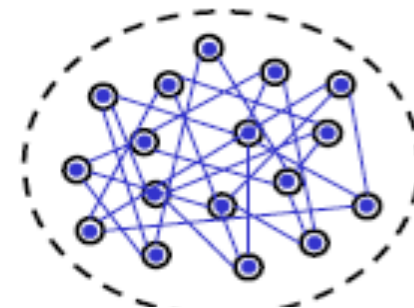
- ❑ Times too long considering antivirus software available within days/weeks!
- ❑ The virus survival time is considerably high with respect to the results of epidemic models of spreading/recovering:
  - Something wrong with the epidemic threshold!
- ❑ Experiment: SIS over a generated Scale Free network (exponent -3).



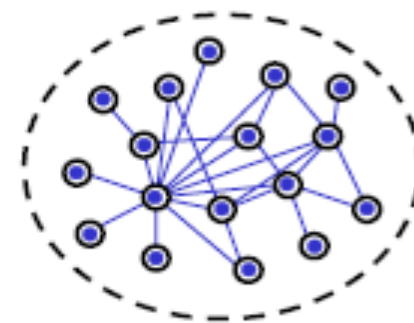
# No Epidemic Threshold for SF!



Infections proliferate in SF networks  
independently of their spreading rates!



Random Network



Scale Free Network



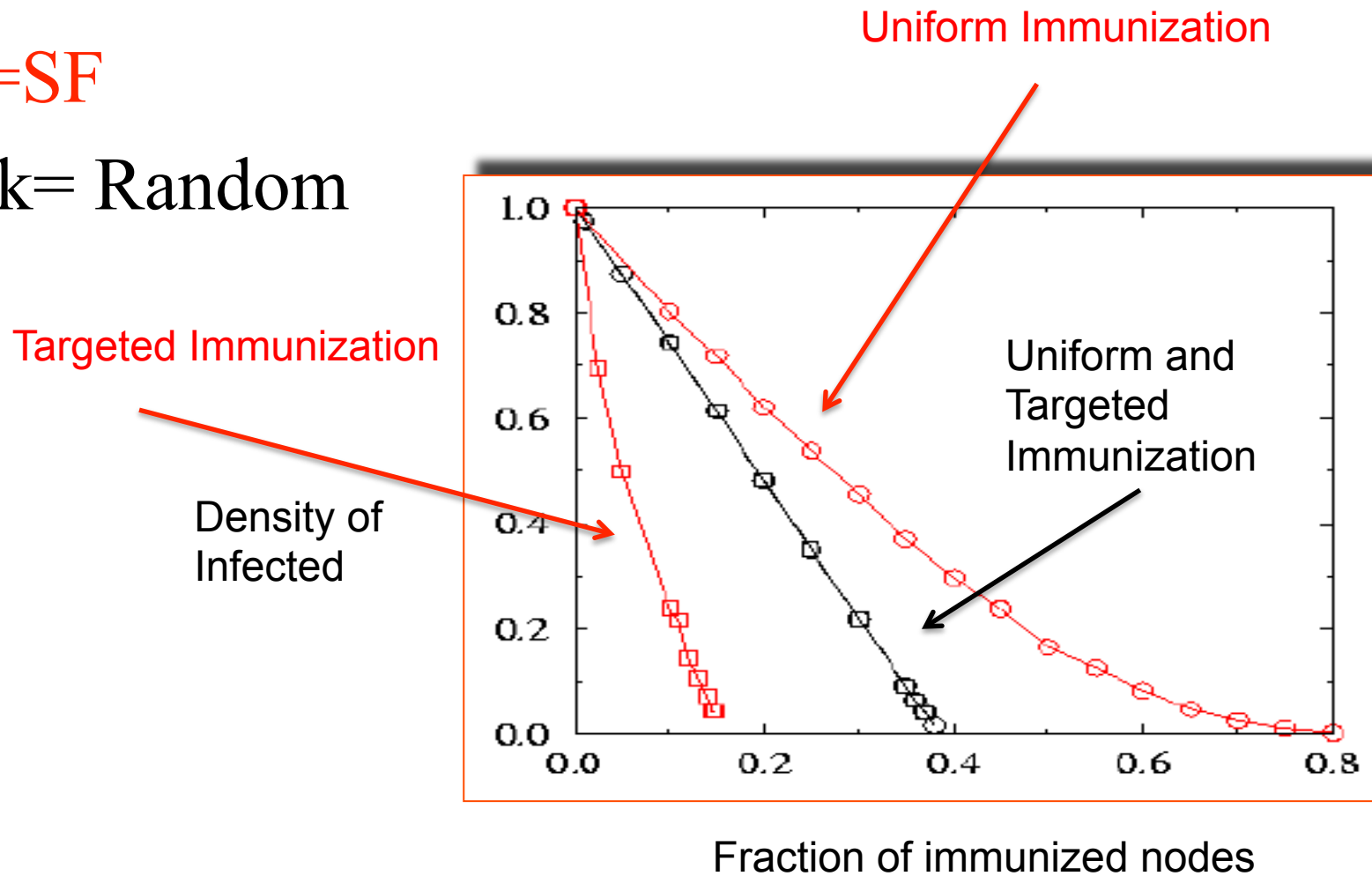
## Following result on Immunization

- Random network can be immunized with some sort of uniform immunization process [oblivious of the characteristics of nodes].
- **This does not work in SF networks** no matter how many nodes are immunized [unless it is all of them].
- Targeted immunization needs to be applied
  - Keeping into account degree!



# Immunization on SF Networks

- Red=SF
- Black= Random



# Local Immunization

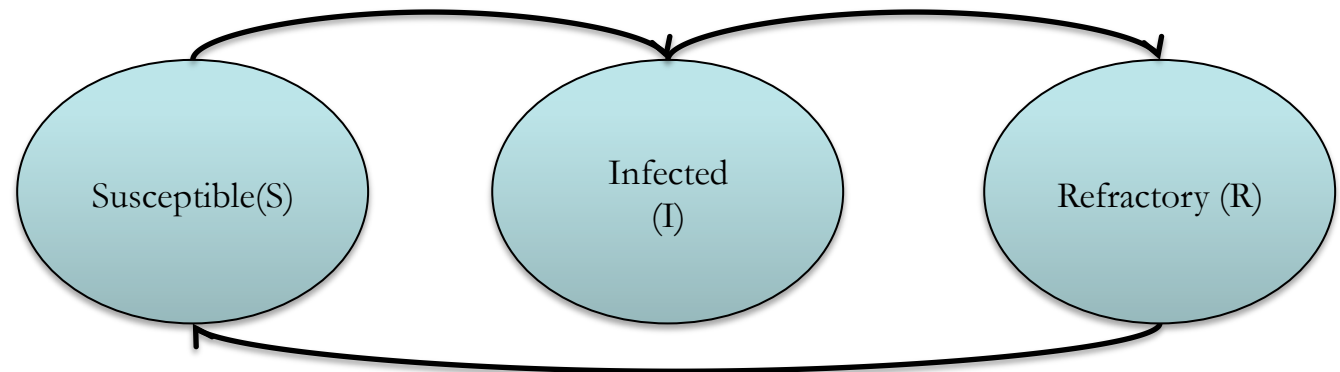
- ❑ Global knowledge on the network structure is rarely (or never) available
- ❑ Local immunization strategy:
  - Select  $g$  nodes at random
  - Ask to each of them to pass over the vaccine to one of their neighbors
  - As a result, a node with degree  $k$  is immunized with a probability  $kP(k)$  (hubs are immunized with higher probability!)





# SIRS Model

- SIR but after some time an R node can become susceptible again.
- A number of epidemics spread in this manner (remaining latent for a while and having bursts).

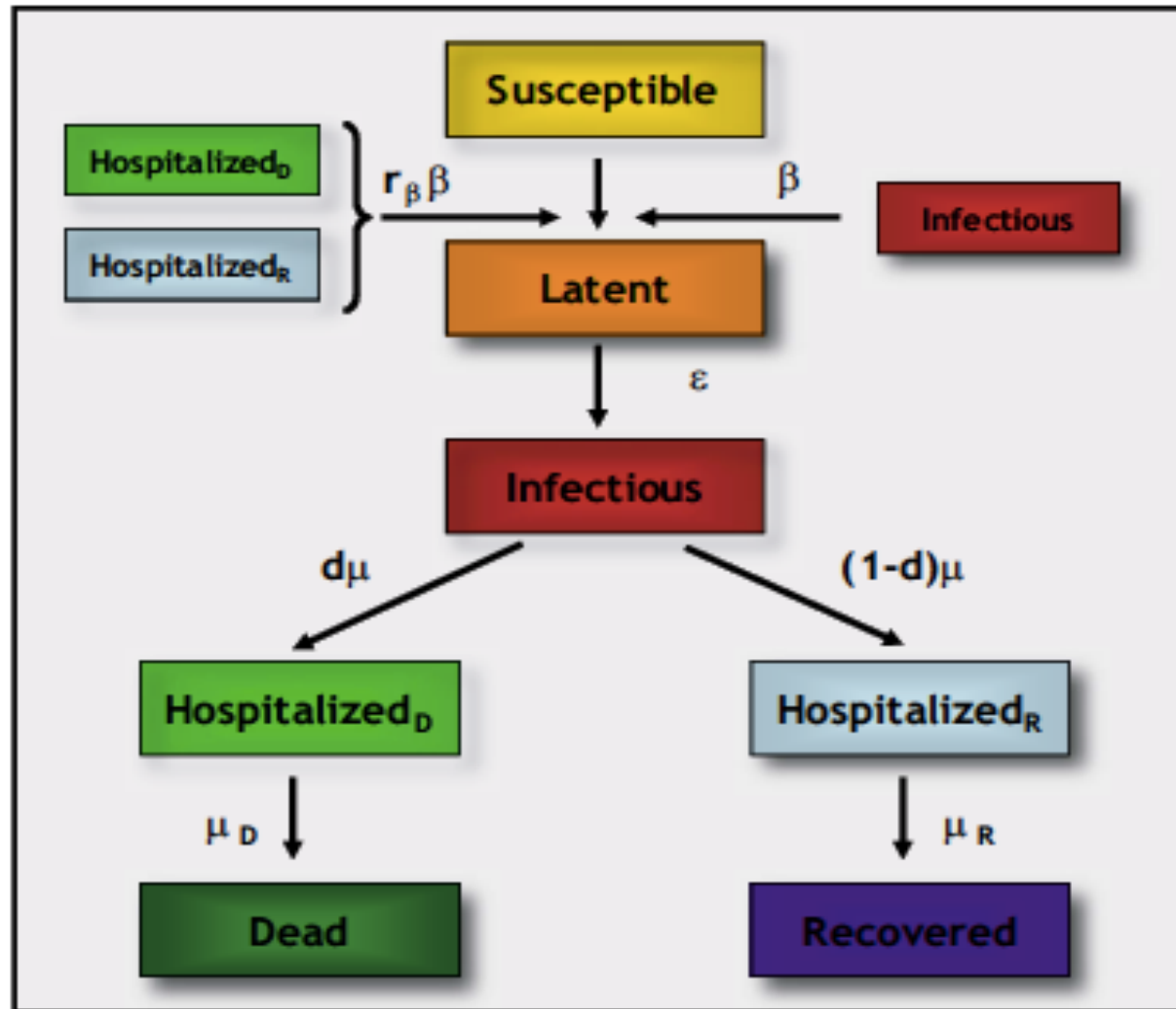


# SARS Prediction

- ❑ SARS: severe acute respiratory syndrome
- ❑ Model based on global stochastic meta-population model
  - SIR like model with more parameters and homogenous mixing
  - Travel data and census data
- ❑ WHO data about SARS spreading to evaluate the model
- ❑ Outbreak starts in Hong Kong



# Epidemics Flow



# Adding Travel and Geography

- A model per city
- Probabilities of an individual of moving from one city to the next follows proportions of traffic observed in the air travel data



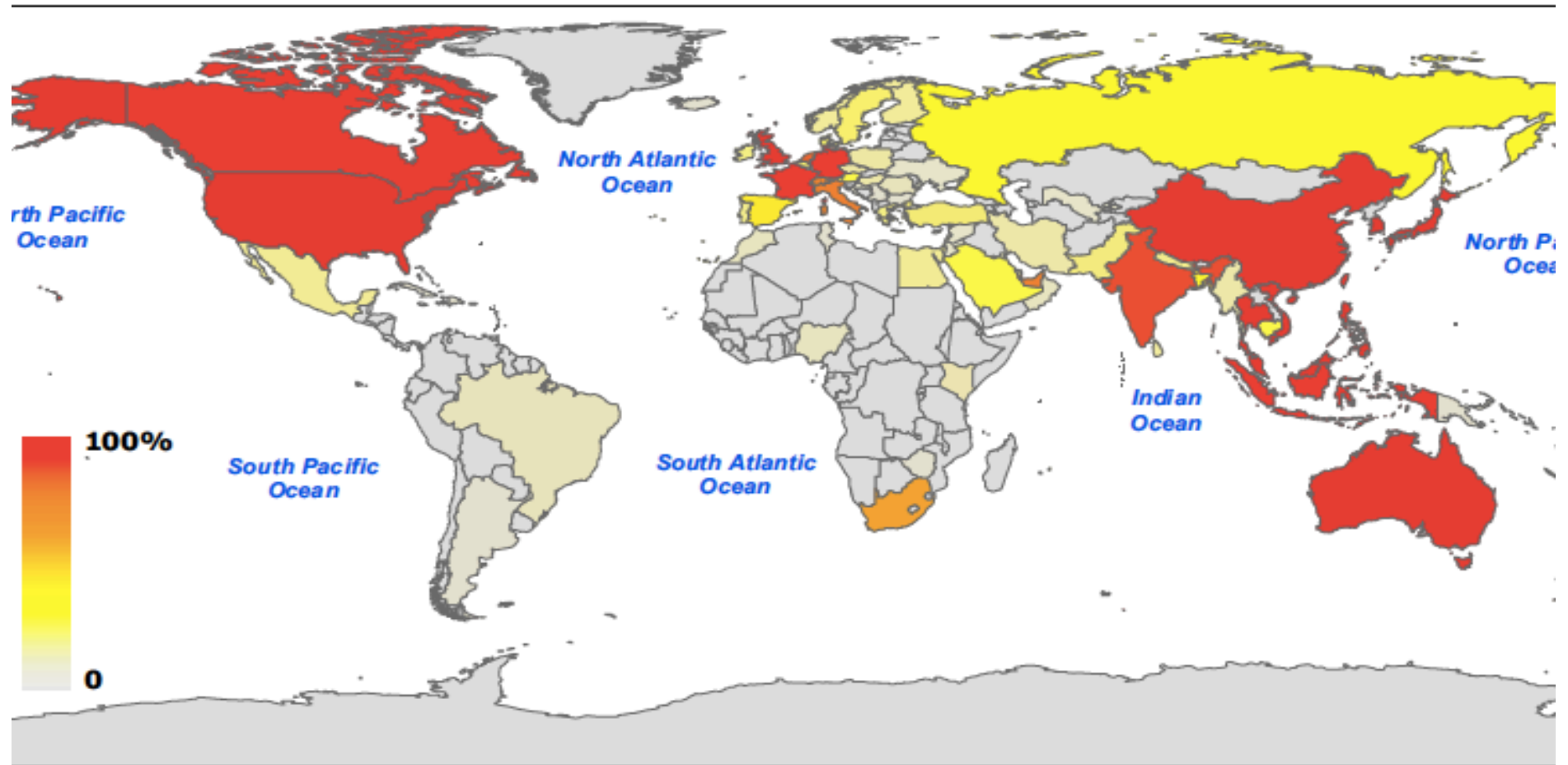
# The Parameters of the Model

## □ Parameters used:

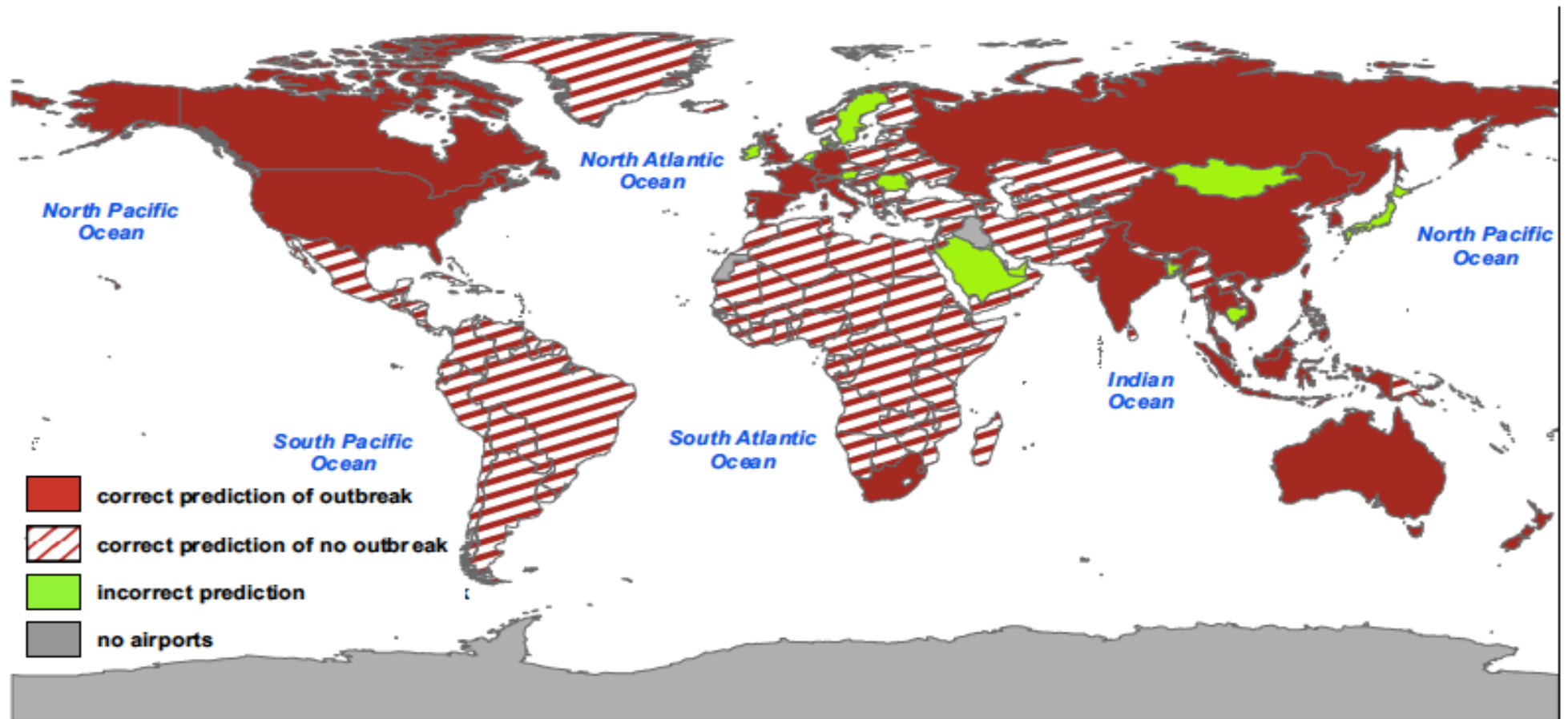
Parameter	Description	Baseline value
$T_0$	Initial offset from 21 February (days)	3*
$\beta$	Rate of transmission	0.57*
$L(t = 0)$	Number of initial latent individuals	10*
$s_f(t)$	Scaling factor for the rate of transmission	21 February + $T_0$ -20 March
		1.00
		21 March – 9 April
$r_\beta$	Relative infectiousness of patients at the hospital	0.37
		10 April – 11 July
		0.06
$\varepsilon^{-1}$	Average latency period (days)	0.2
$\mu^l(t)$	Average period from onset of symptoms to admission (days)	4.6
		21 February + $T_0$ -25 March
		4.84
$\mu_R^{-1}$	Average period from admission to recovery (days)	25 March – 1 April
		3.83
		2 April – 11 July
$\mu_D^{-1}$	Average period from admission to death (days)	3.67
$d$	Case fatality rate	23.5
		35.9
		0.2



# Predicted Outbreak Likelihood



# Comparison with Data





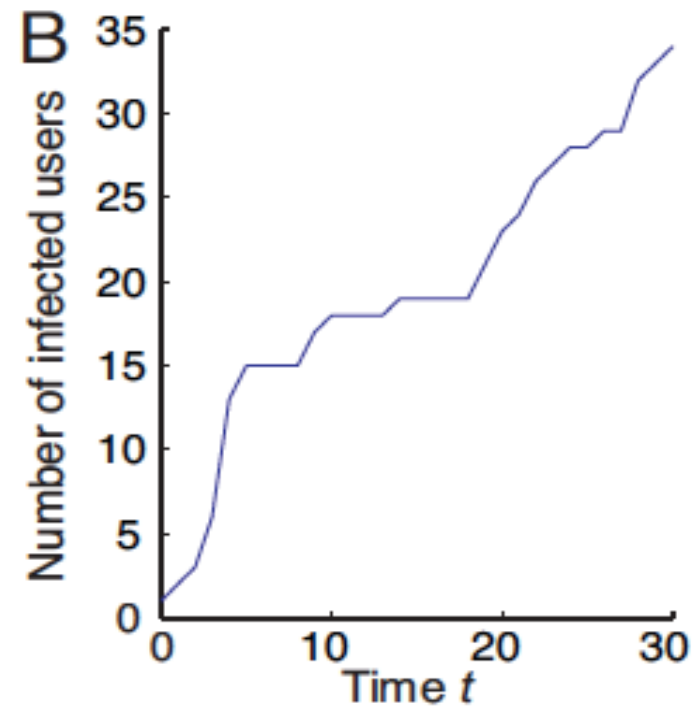
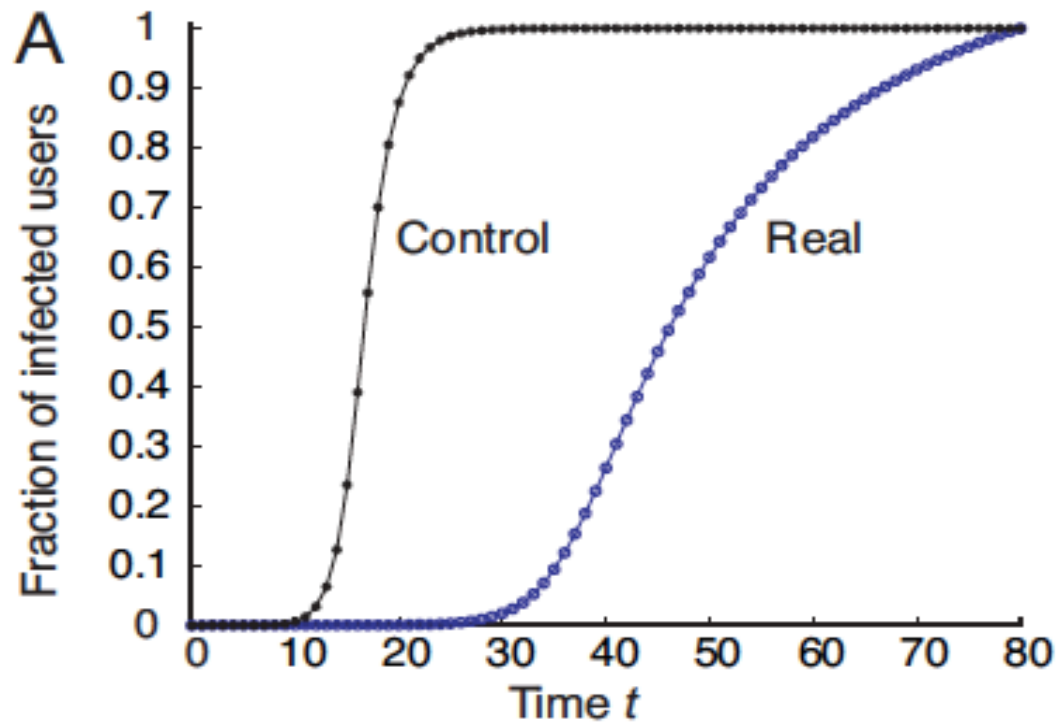
## SI and Weak Ties (Onnela 2007)

- What is the role of tie strength in information diffusion/epidemic spreading?
- Let us remember the findings in Lecture 2...



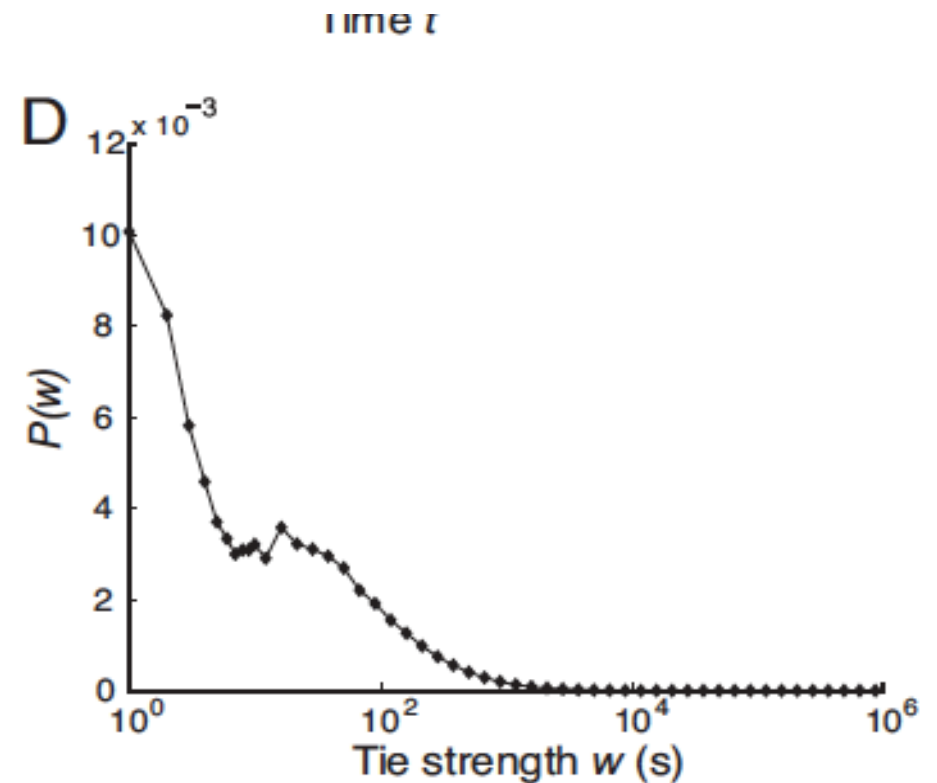
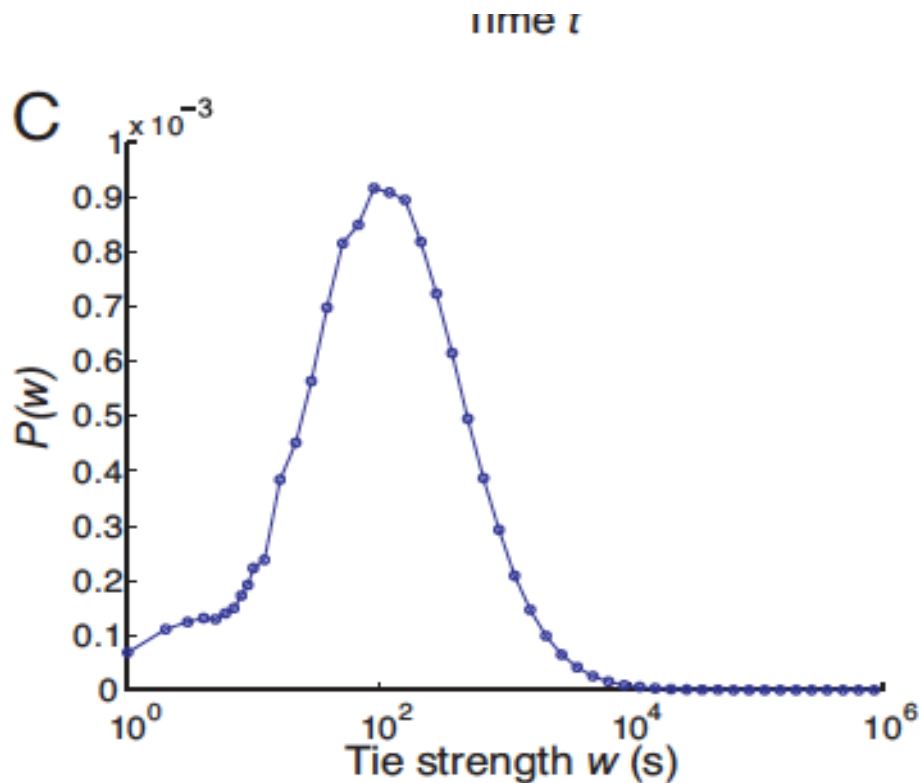


# SI over Communication Network



# SI and tie strength

Medium ties are the most used:  
Weak ties: not enough time to communicate  
Strong ties: stuck inside a community



## Summary

- ❑ Epidemics are very complex processes.
- ❑ Existing models have been increasingly capable of capturing their essence.
- ❑ However there are still a number of open issues related to the modelling of real disease spreading or information dissemination.
- ❑ Mixing geographic factors with epidemics model is very relevant and important.
- ❑ Tie strength add another perspective.



# References

- ❑ Chapter 21
- ❑ Pastor-Satorras, R. and Vespignani, A. Epidemic Spreading in Scale-Free Networks. Phys. Rev. Lett.(86), n.14. Pages = 3200--3203. 2001.
- ❑ Pastor-Satorras, R. and Vespignani, A. Immunization of Complex Networks. Physical Review E 65. 2002.
- ❑ Marcelo Kuperman and Guillermo Abramson. Small world effect in an epidemiological model. Physical Review Letters, 86(13): 2909–2912, March 2001.
- ❑ V. Colizza, A. Barrat, M. Barthélemy and A. Vespignani. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. BMC Medicine 2007, 5:34

