Non-Poisson Contact Processes in Virus Spreading

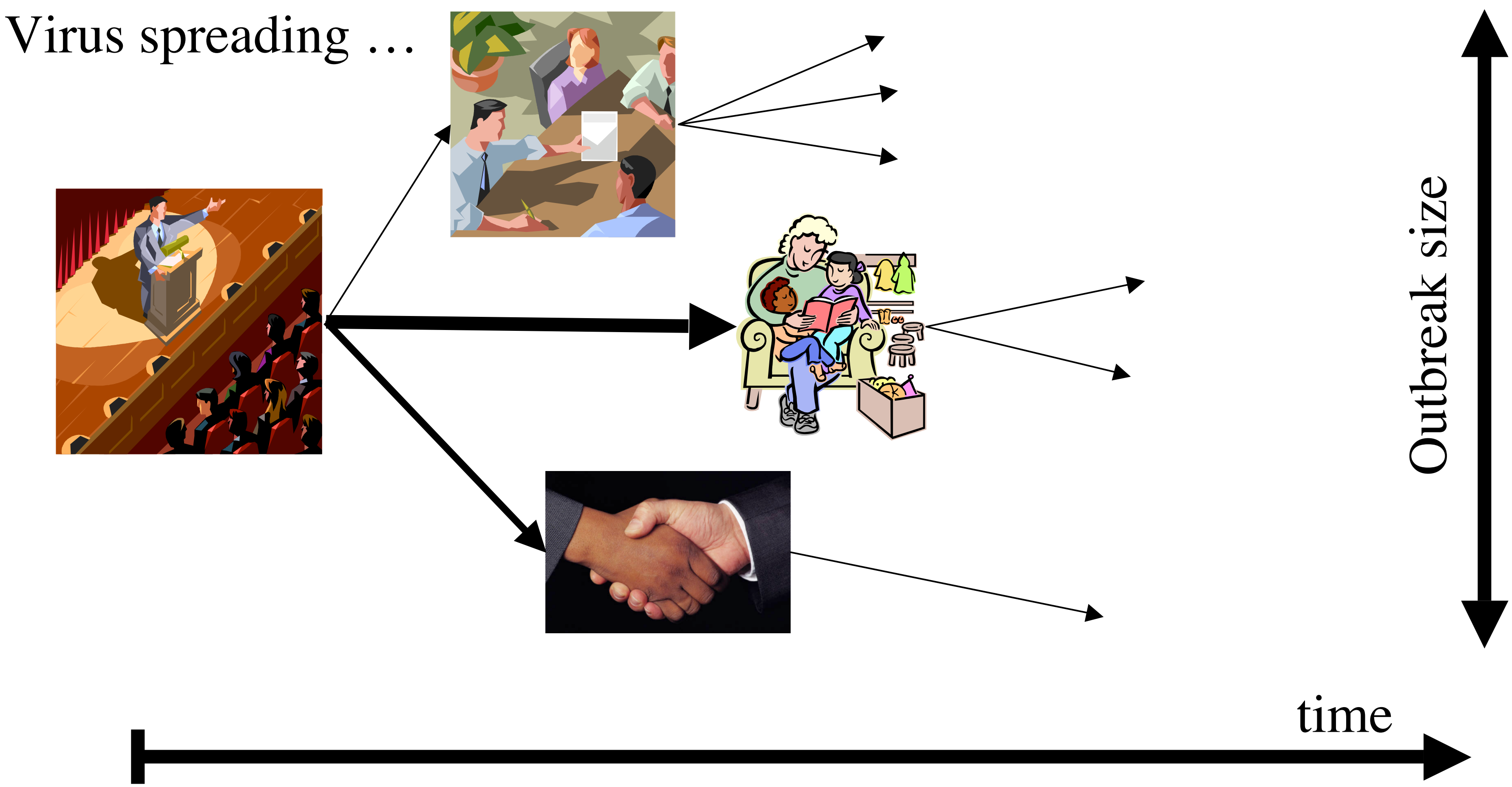
A. Vázquez and A.-L. Barabási
University of Notre Dame

www.nd.edu/~networks
www.nd.edu/~avazque1

Epidemic outbreaks

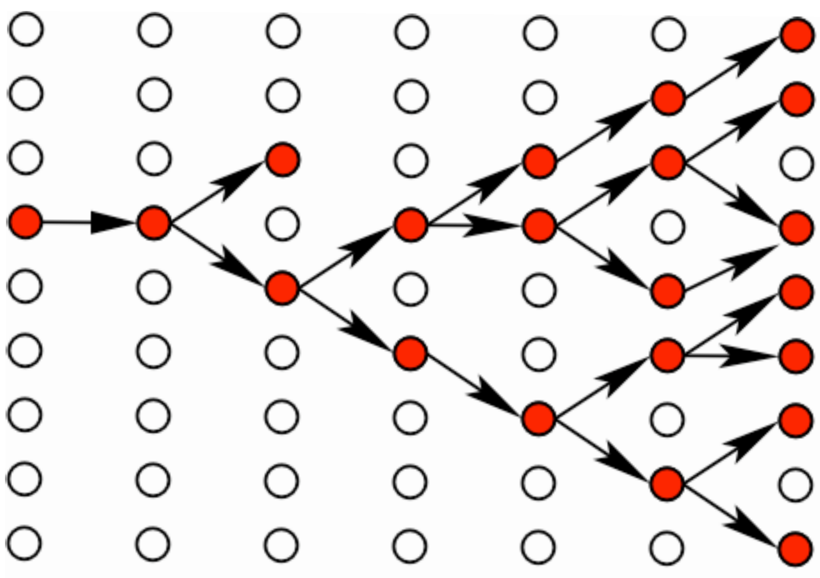
- Airborne viruses
  - SARS, Influenza
- Sexually transmitted diseases
  - HIV
- Computer viruses and worms
  - LoveLetter, Code-Red
- Rumors (“Infectious of the Mind”)
  - Chain Letters, Hoaxes

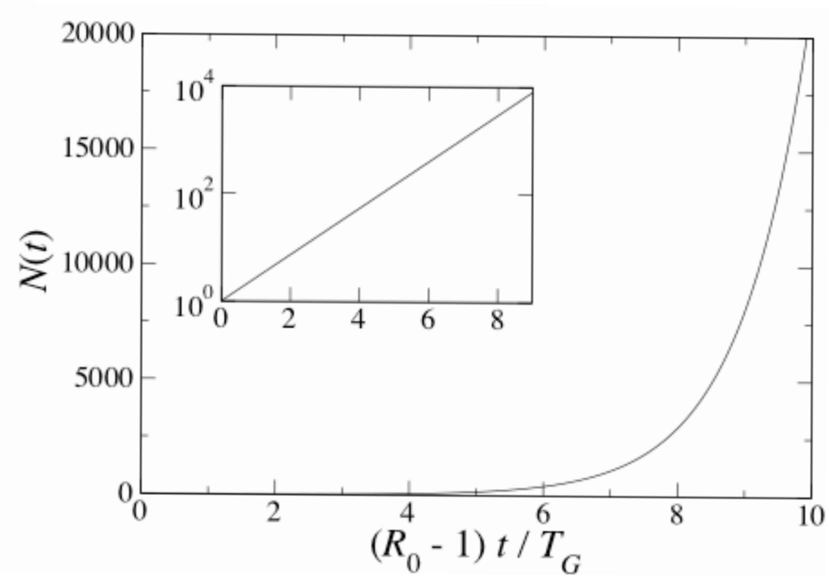
The New Jersey Outbreak



Basic epidemic model

○ Susceptible ● Infected (SI model)





R_0 Basic reproductive number

T_G Generation time

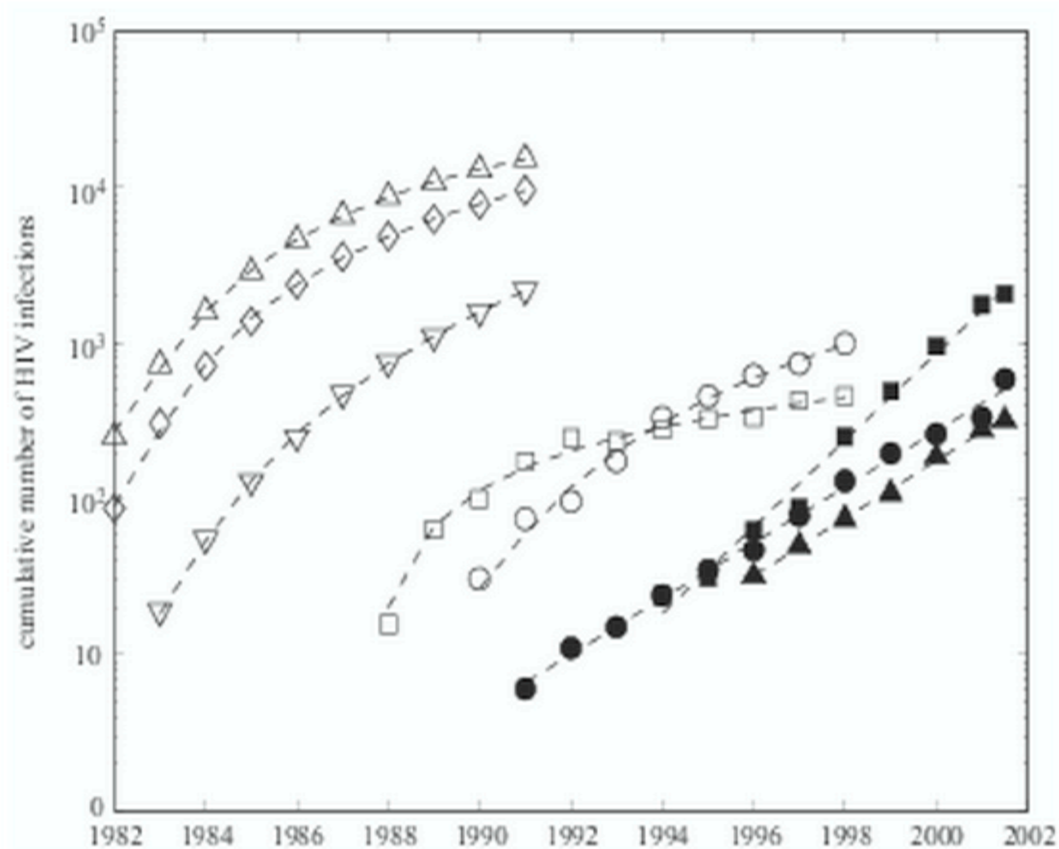
Main questions

- How many?
- How fast?
- How can we stop it?
- How can we avoid it?

- Empirical evidence
- Models

HIV/AIDS

✓Szendroi & Czanyi 2004

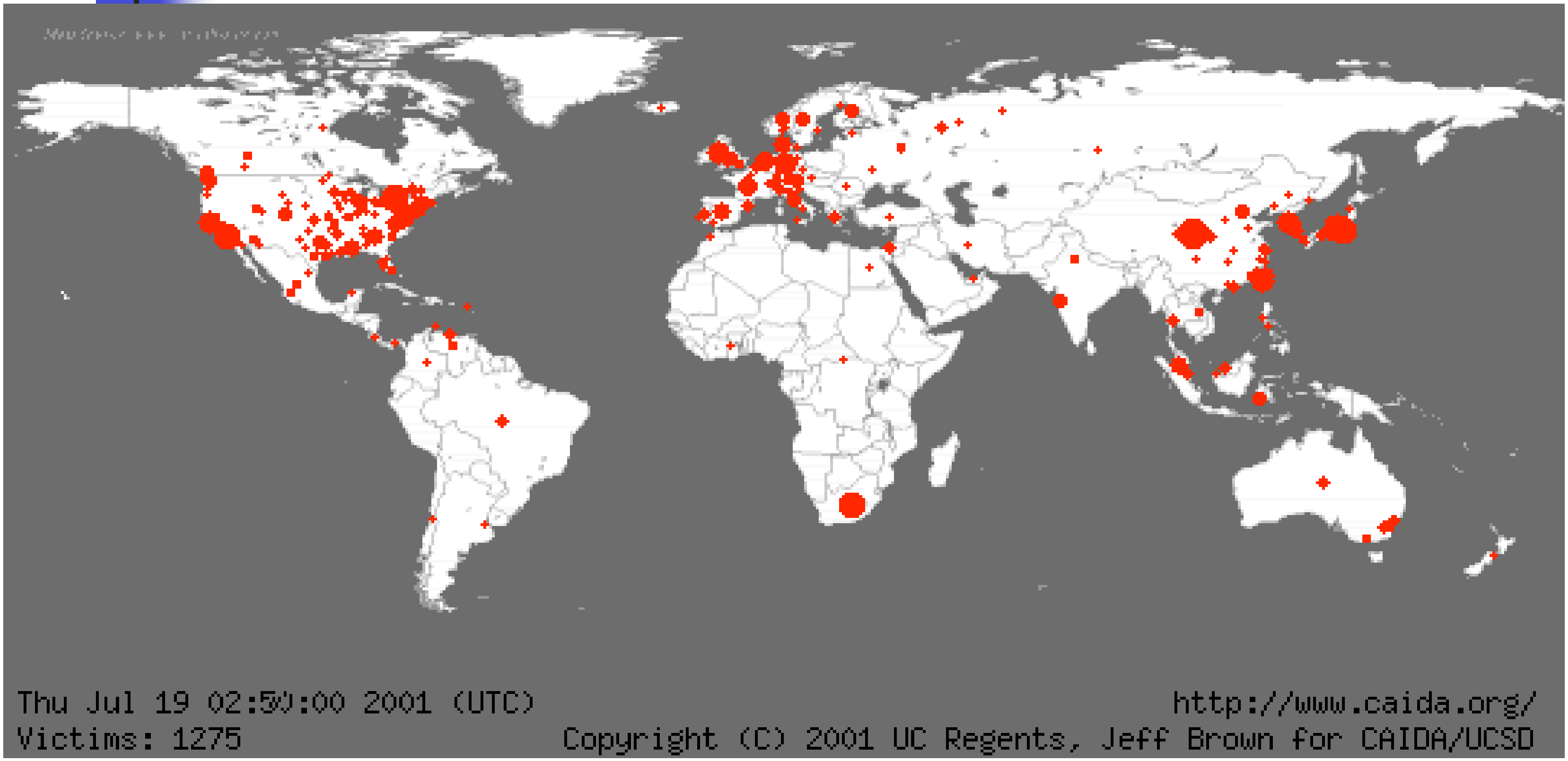


△ New York - HOM
▽ New York - HET
◇ San -Francisco - HOM
○ South Africa
□ Kenya

▲ Georgia
■ Latvia
● Lithuania

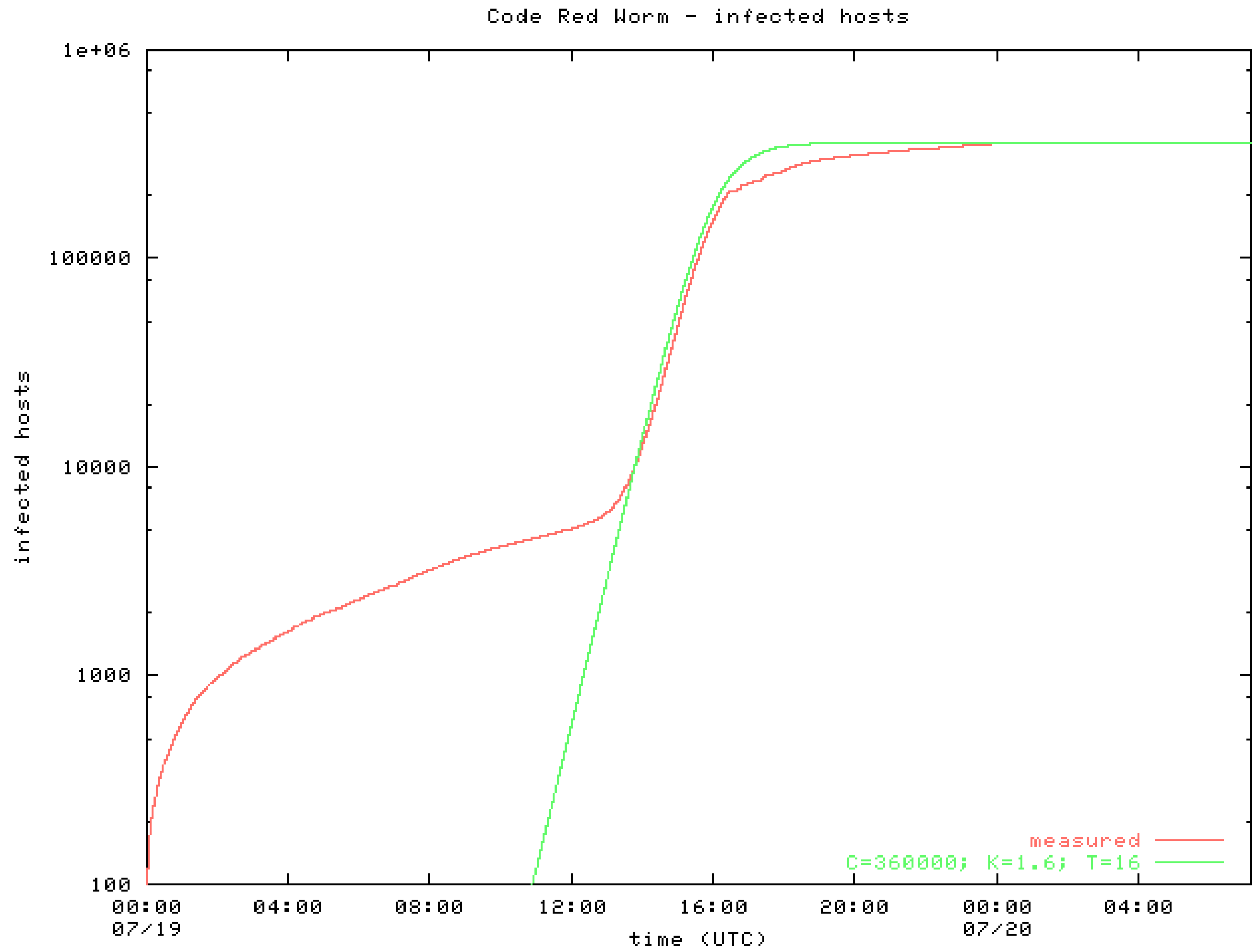
exponential=linear in a linear-log plot

Code-Red worm (CAIDA)

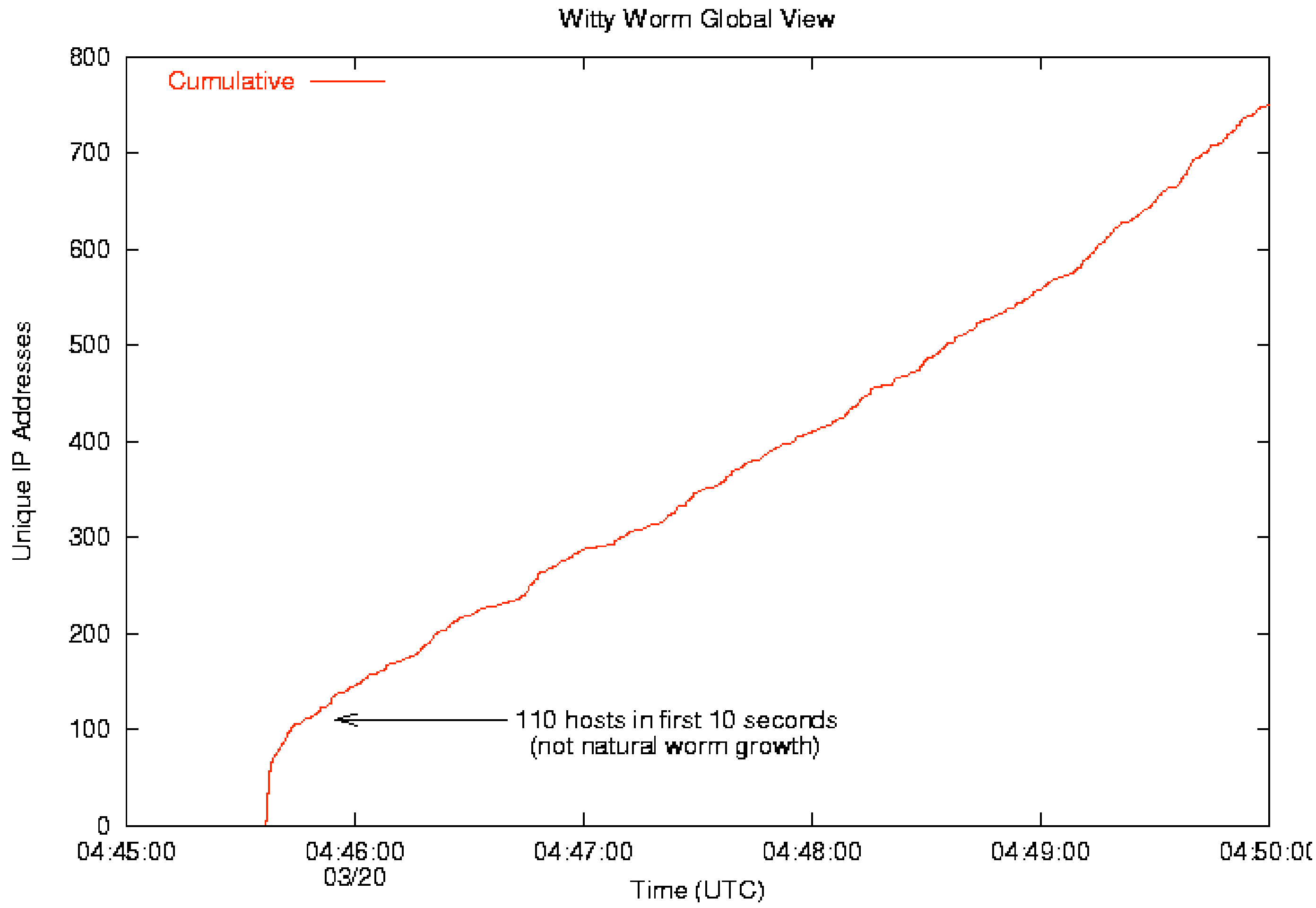


Estimated damage $2.6 billion

Code-red worm (CAIDA)

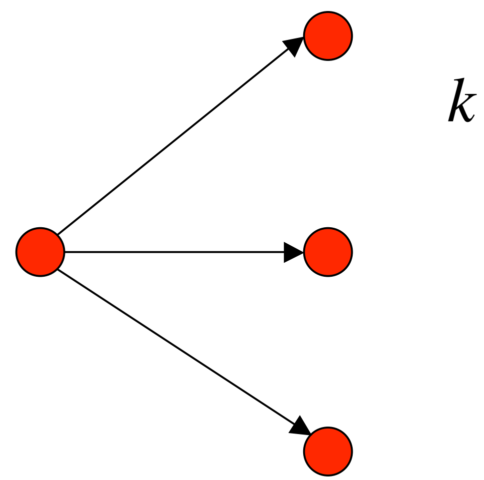


Witty worm (CAIDA)



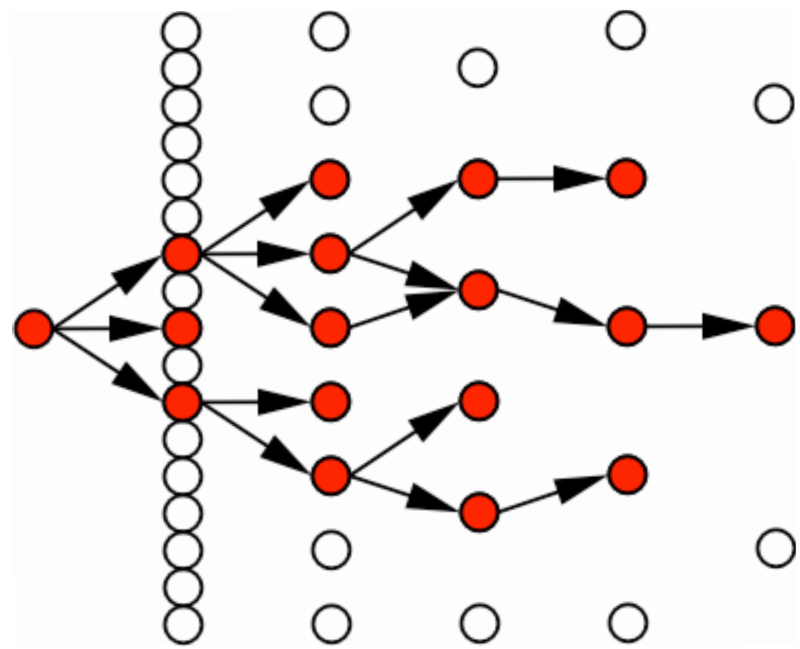
Contact heterogeneity

- $P(k) \sim k^{-\gamma}$
  - Sexual contacts
    - ✓ Liljeros *et al* 2001
  - Email contacts
    - ✓ Ebel *et al* 2002
    - ✓ Eckman *et al* 2004
  - Urban contacts
    - ✓ Eubank *et al* 2004



Contact heterogeneity

✓May & Anderson 1988



✓May & Anderson 1988
✓Barthelemy, Barrat, Vespignani 2004 ➔ exponential

Time between contacts

?

Temporal activity patterns

- Airborne viruses
  - Visitation of public places
- Sexually transmitted diseases
  - Sexual activity patterns
- Computer viruses
  - Email, Login sessions
- Rumors
  - Email, SMS, Phone

Poisson process

- Basic assumption:

  Contacts take place at constant rate $\lambda = 1/T_G$

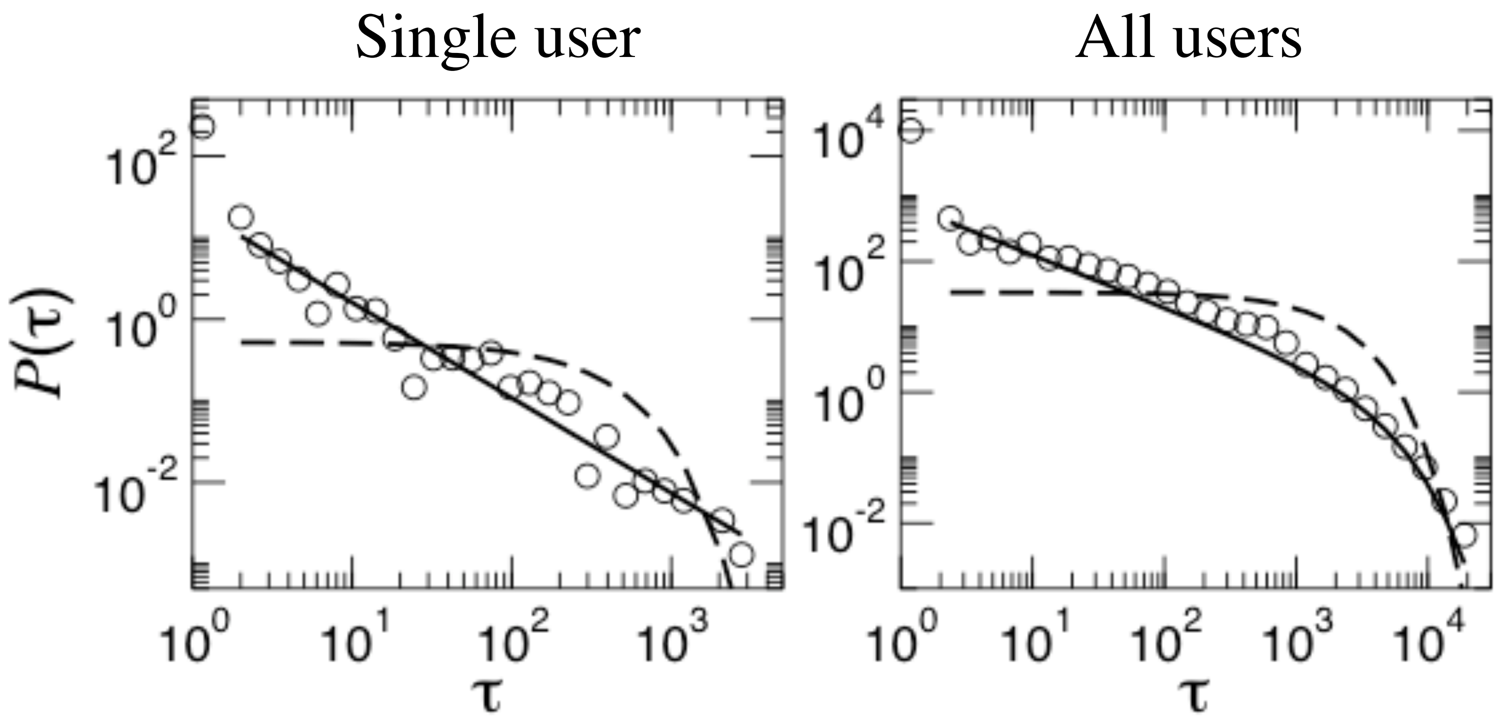
- Time interval distribution

$$P(\tau) = \lambda e^{-\lambda\tau}$$

$$P_0(\tau_0) = \lambda e^{-\lambda\tau_0}$$

$\boxed{\text{Empirical data?}}$

Library data / airborne viruses

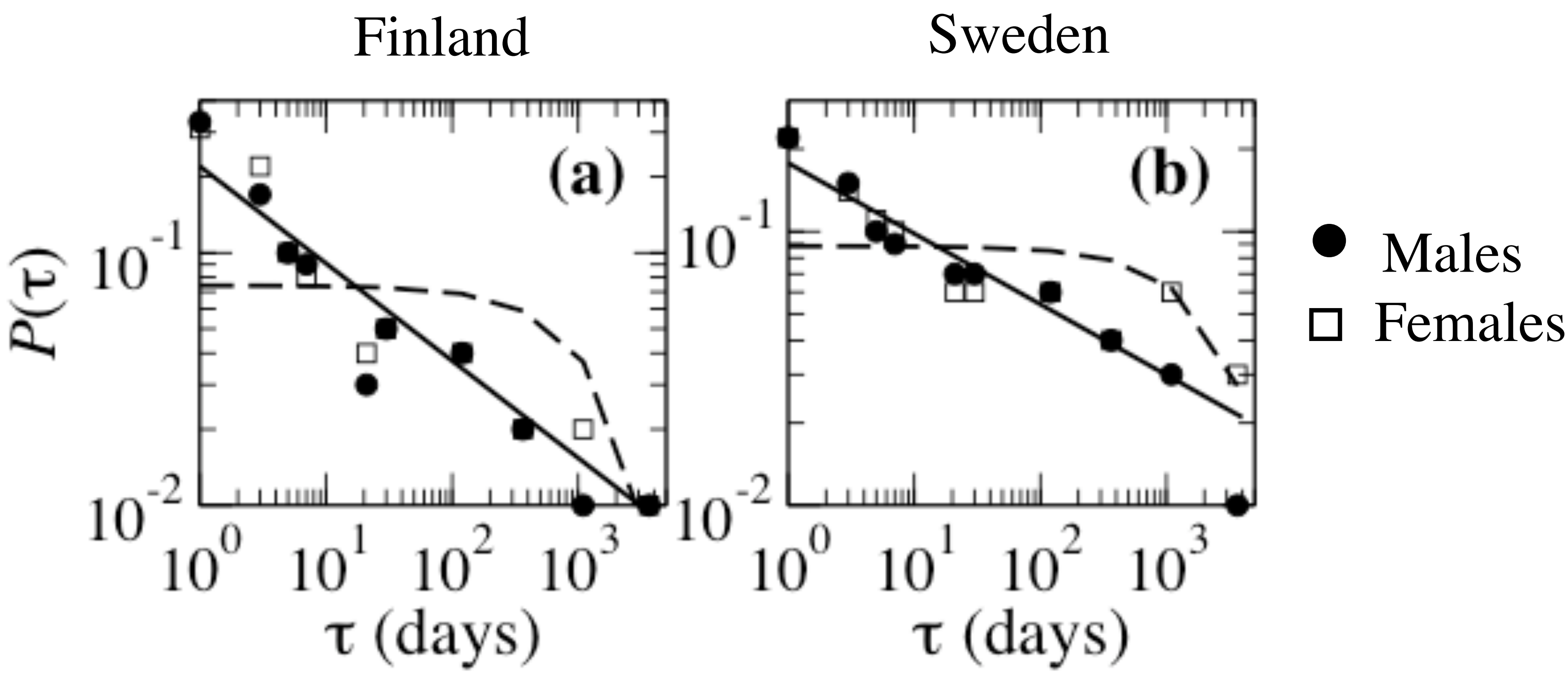


τ: time between two consecutive loans

Deszo *et al*, unpublished

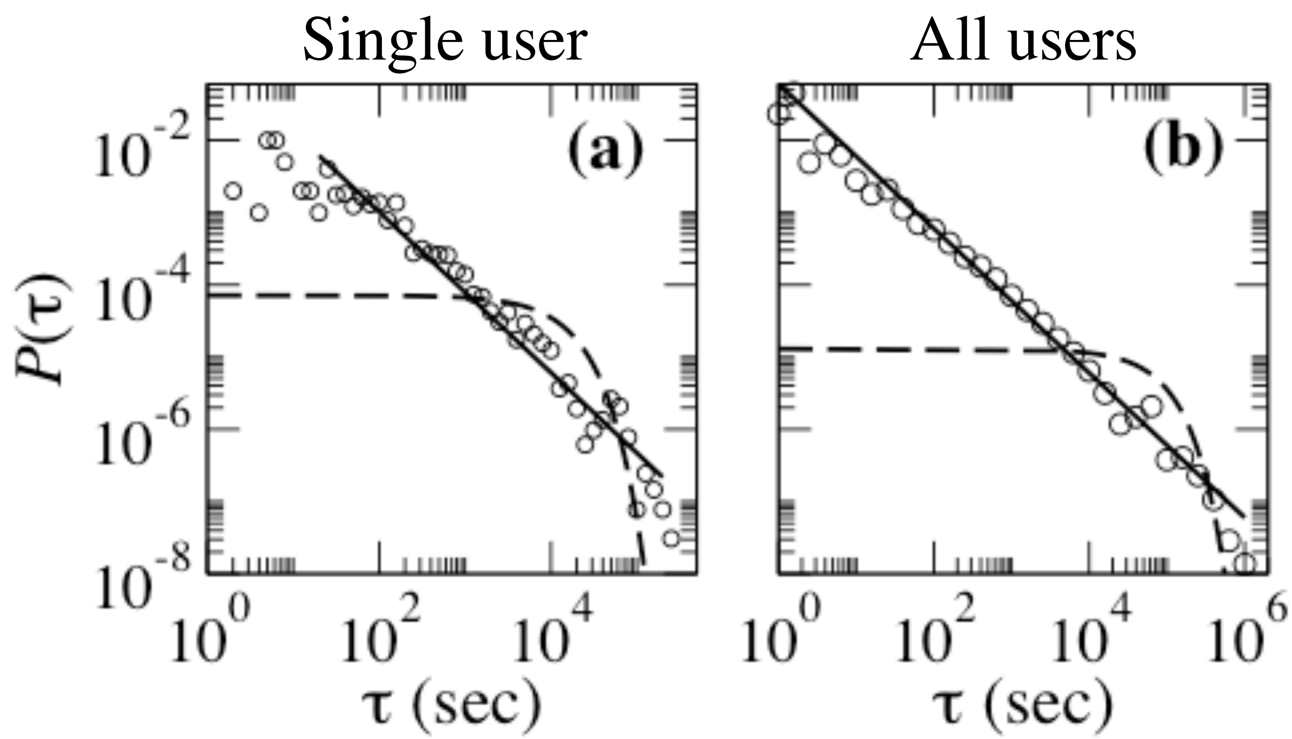
Power law=linear in a log-log plot

Sexual activity / STD



τ: time since the last sexual intercourse

Emails / Computer viruses



τ: time between two consecutive emails sent by a user

Poisson vs heavy tailed

Poisson

$$P(\tau) = \lambda e^{-\lambda\tau}$$

$$P(\tau) = \frac{A}{\tau^{\alpha}}$$

Heavy tailed

long delays

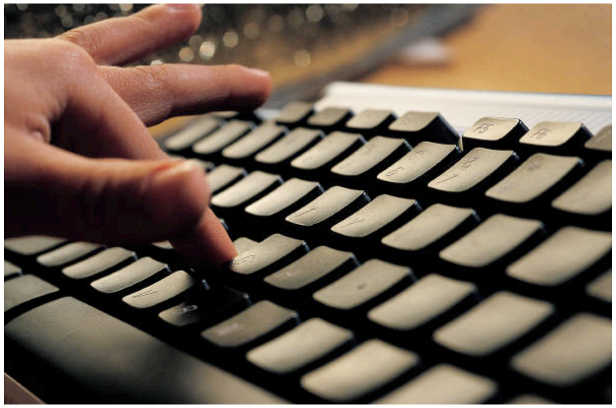
many contacts
in a short time

Epidemic growth

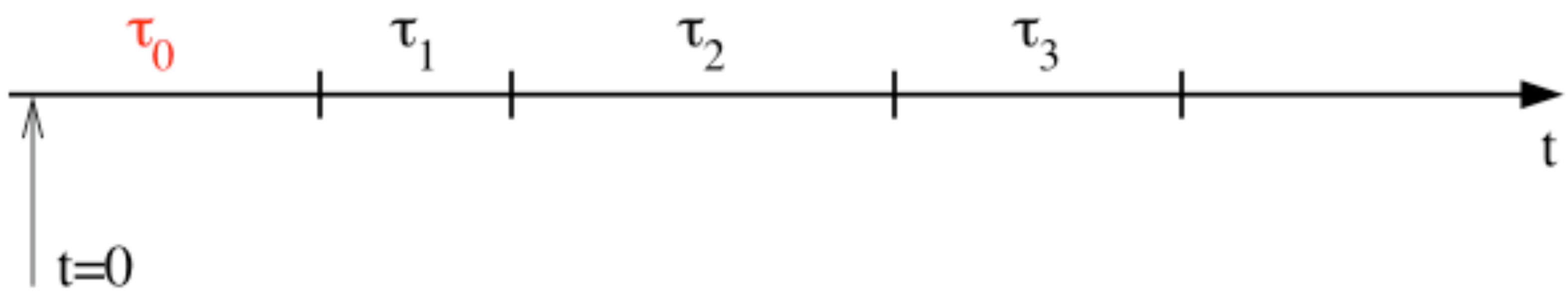
Non-Poisson
contact processes

Spreading via Emails

Infected
Email user



Single Email user

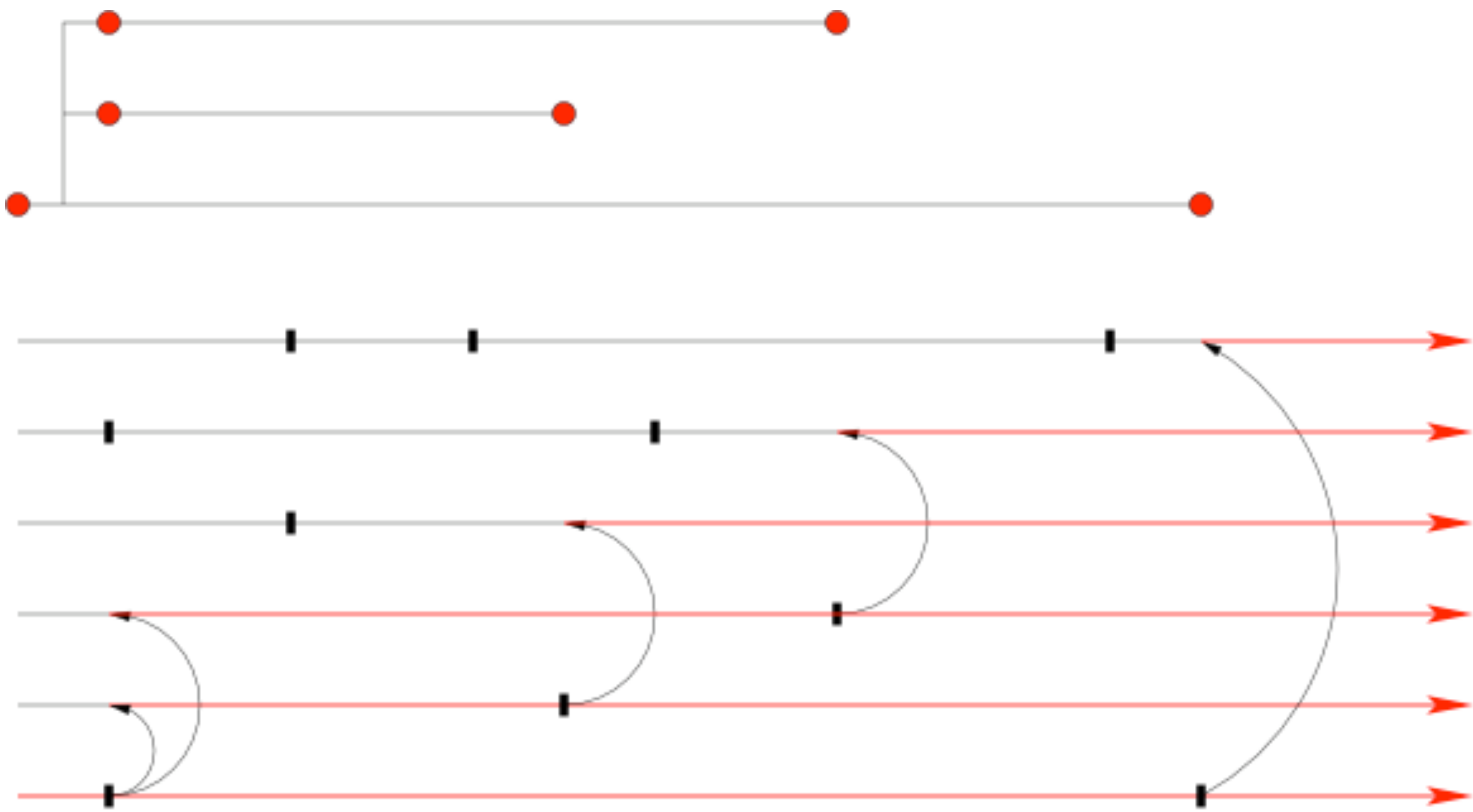


Renewal process

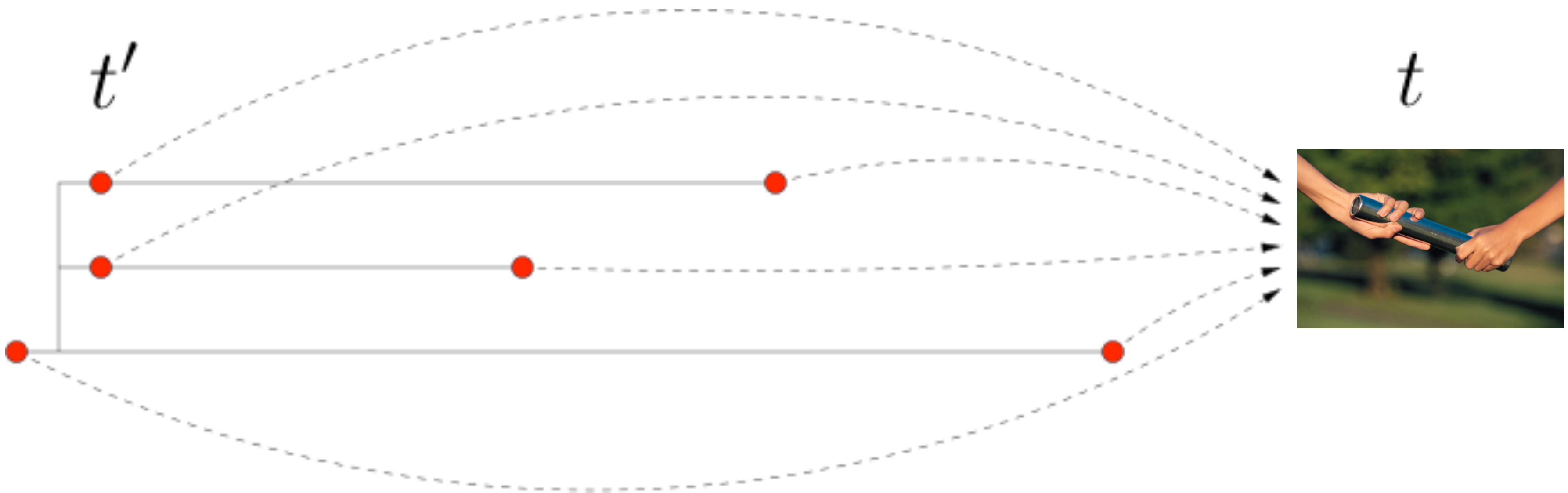
$\tau_1, \tau_2, \tau_3, \ldots$: inter-contact times $\quad P(\tau)$

τ_0: initial delay $\quad P_0(\tau_0) = \frac{1}{\langle \tau \rangle} \int_{\tau_0}^{\infty} d\tau P(\tau)$

Spreading dynamics

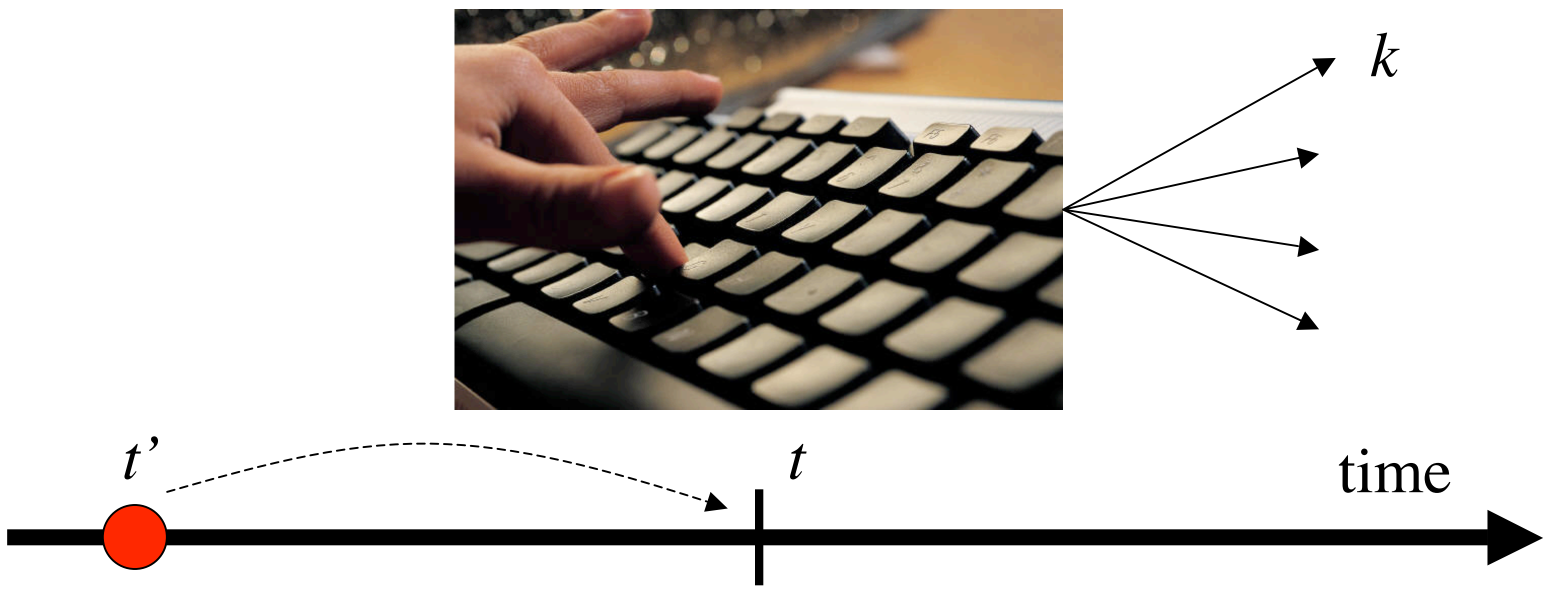


Spreading dynamics



$$n(t) = \int_0^t dt' n(t') \beta(t', t)$$

Spreading dynamics



$$\beta(t', t) = \langle k \rangle C(t', t)$$

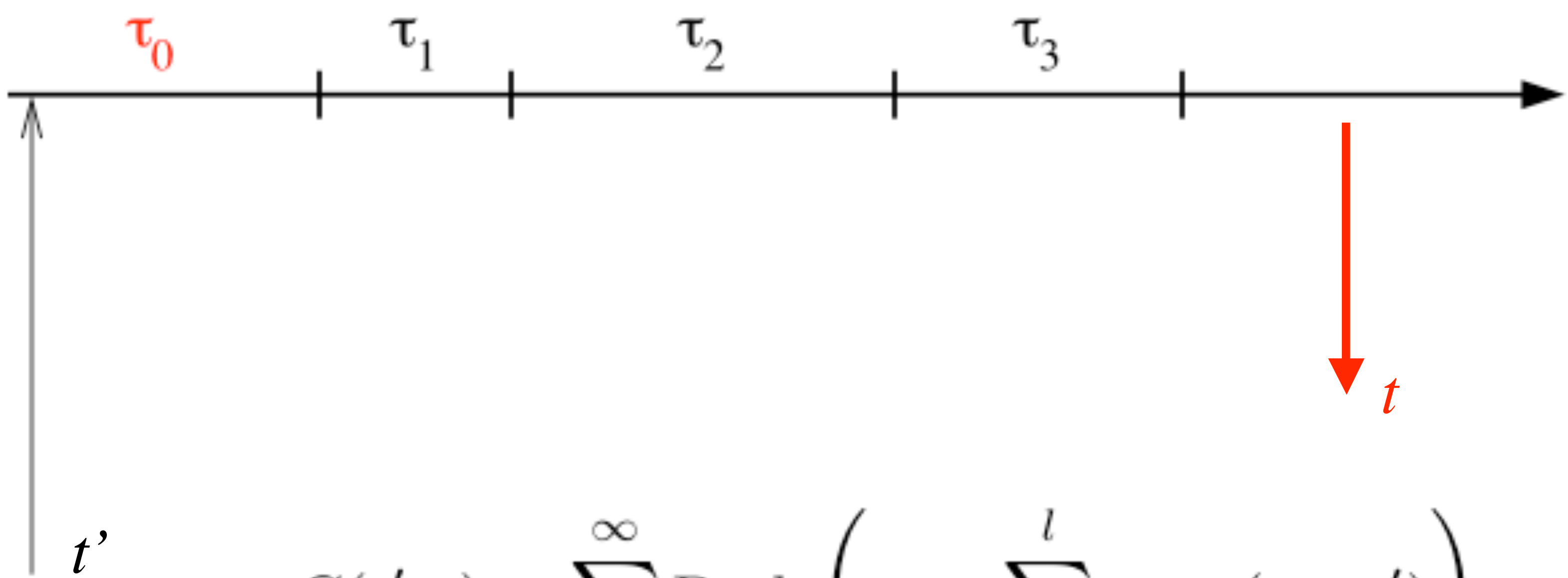
Poisson process

$$C(t', t) = \lambda \,, \quad t' \geq 0$$

$$\Downarrow$$

$$n(t) = \langle k \rangle \lambda e^{\langle k \rangle \lambda t}$$

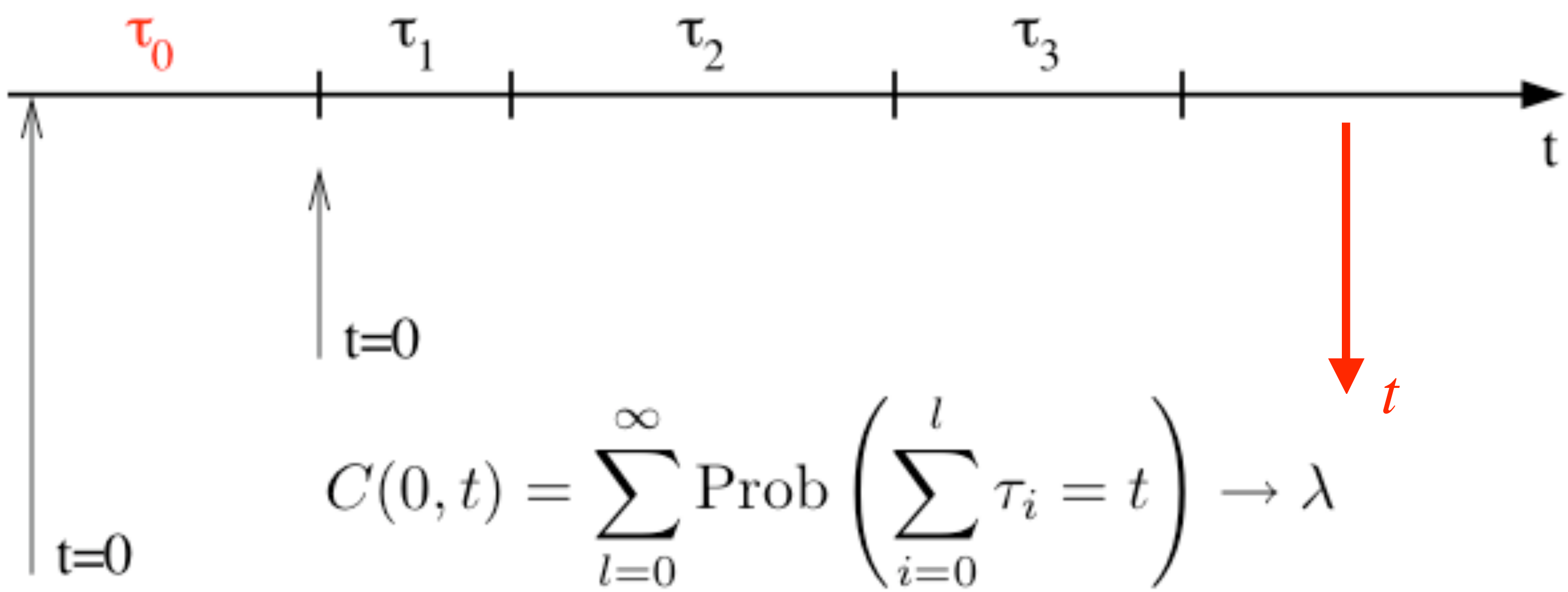
Secondary infected users



$$C(t',t) = \sum_{l=0}^{\infty} \mathrm{Prob}\left(\tau_0 + \sum_{i=0}^{l} \tau_i = (t-t')\right) = \lambda$$

$$\lambda = \frac{1}{\langle \tau \rangle}$$

Patient zero

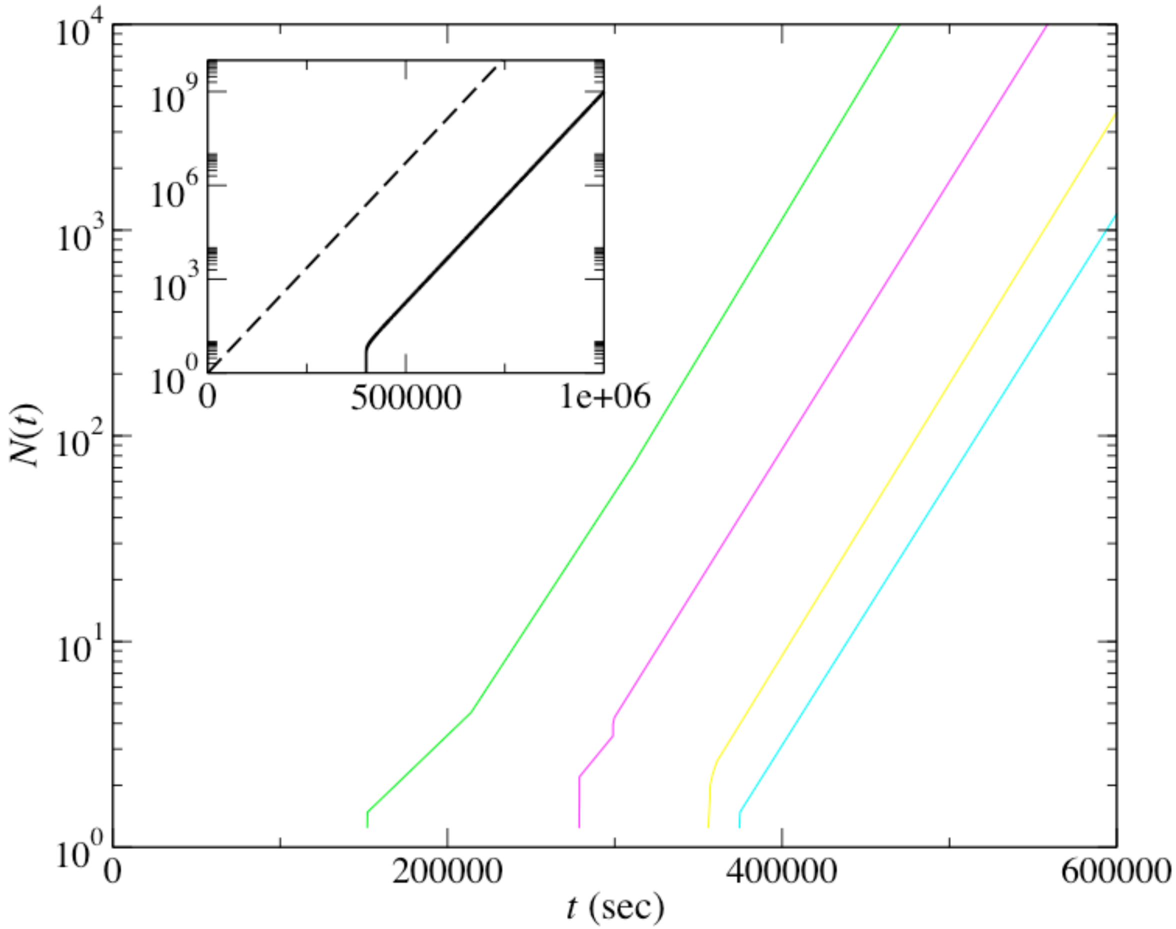


$$C(0,t) = \sum_{l=0}^{\infty} \text{Prob}\left(\sum_{i=0}^{l} \tau_i = t\right) \rightarrow \lambda$$

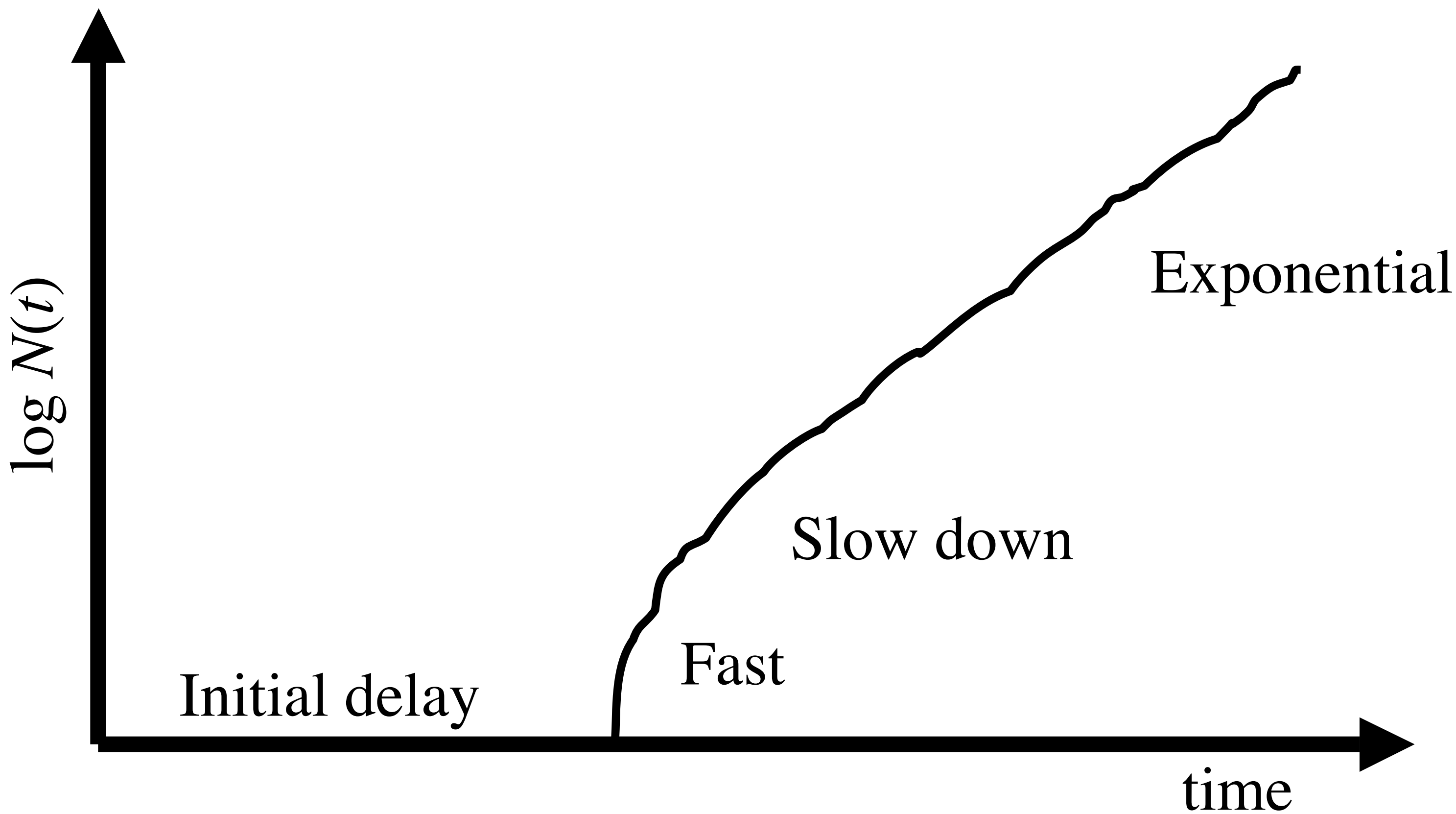
$$C(0,t) = \sum_{l=0}^{\infty} \text{Prob}\left(\tau_0 + \sum_{i=0}^{l} \tau_i = t\right) = \lambda$$

$$N(t) = 1 + \langle k \rangle \int_0^t dt' C(0,t') e^{\langle k \rangle \lambda (t-t')}$$

t=0 ➔ Average outbreak size



General picture



Power law distribution

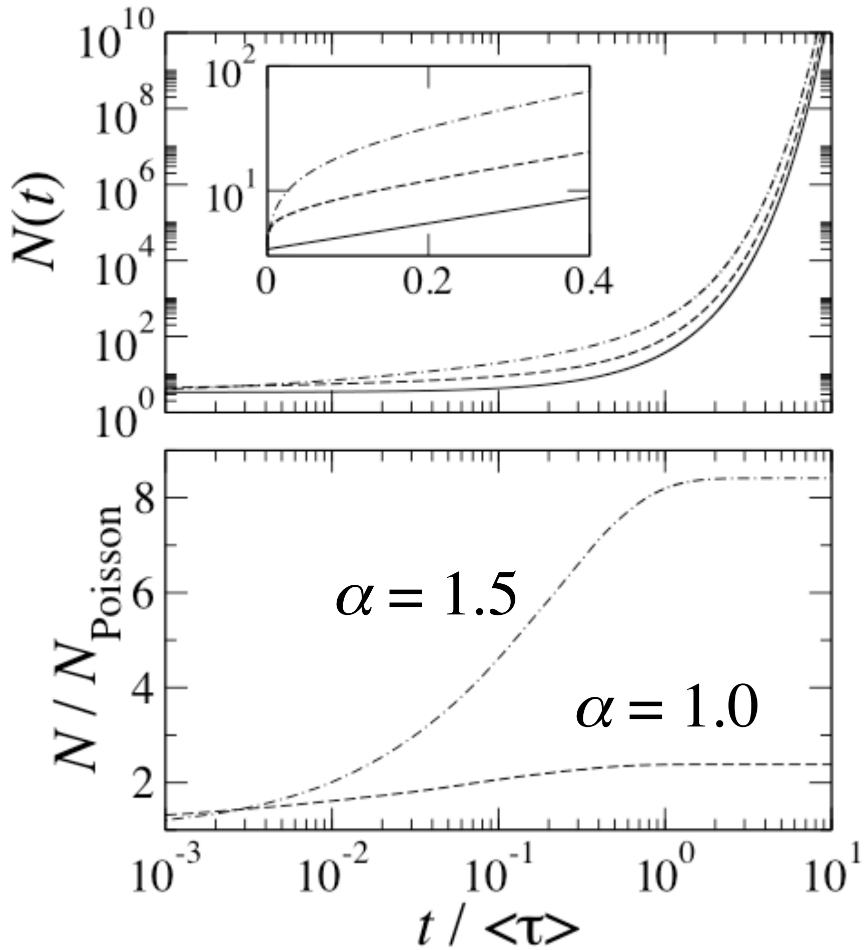
• Power law intercontact distribution

$$P(\tau) = \frac{A}{\tau^{\alpha}}$$

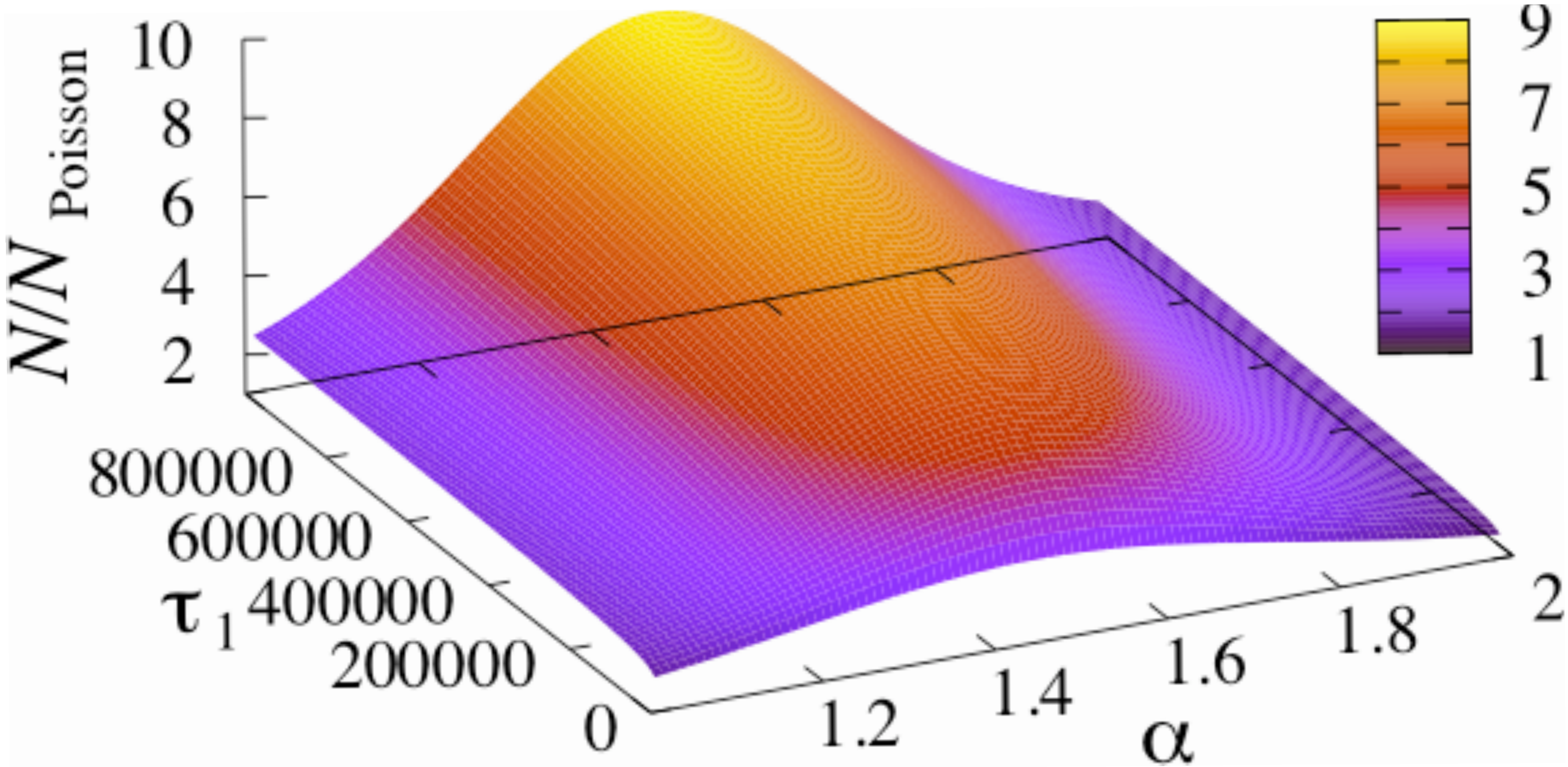
$$\tau_0 \leq \tau \leq \tau_1$$

$$\tau_0 = 1$$

$$\tau_1 = 10^6$$



Final outbreak size



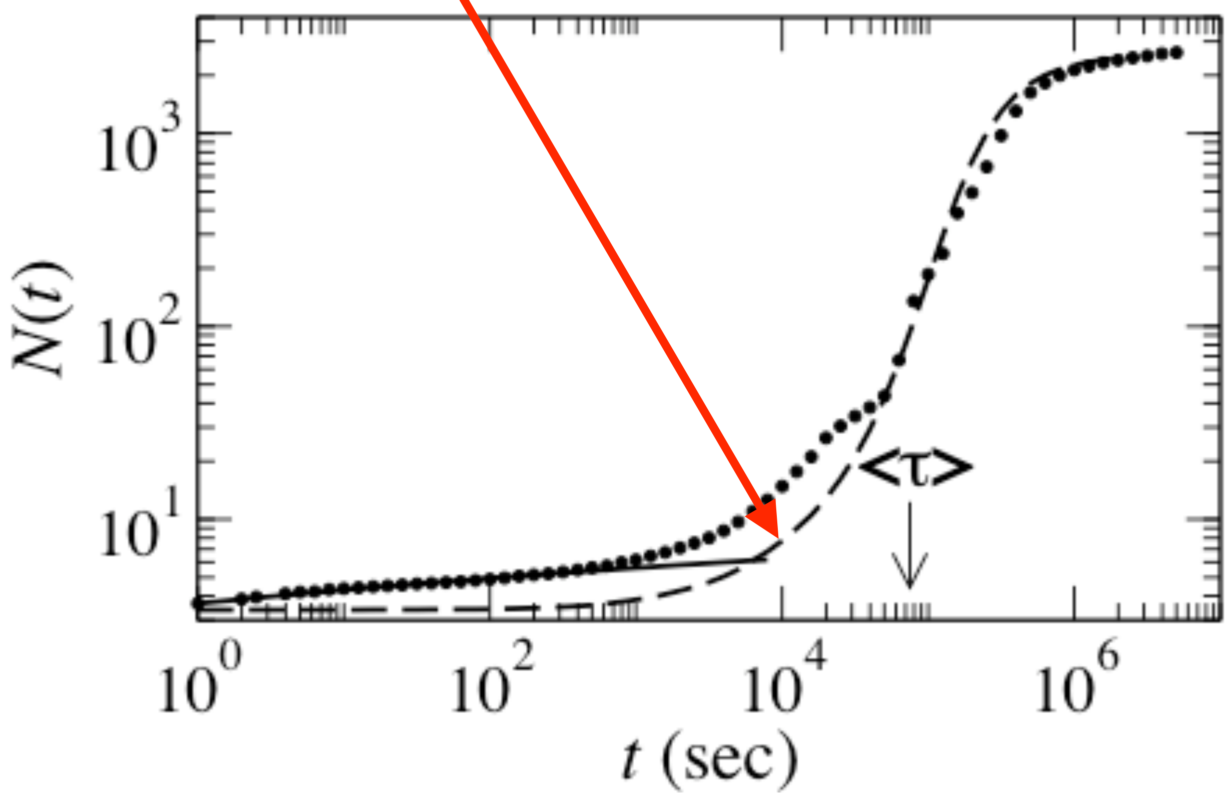
Real Email history

Contact heterogeneity is not the main factor

3,188 users
3 month time interval

● pass a virus
to a user, and
follow its spread

- - - Poisson timing,
same contacts as in
the real data



Immunization

Infected individuals are are removed at rate μ

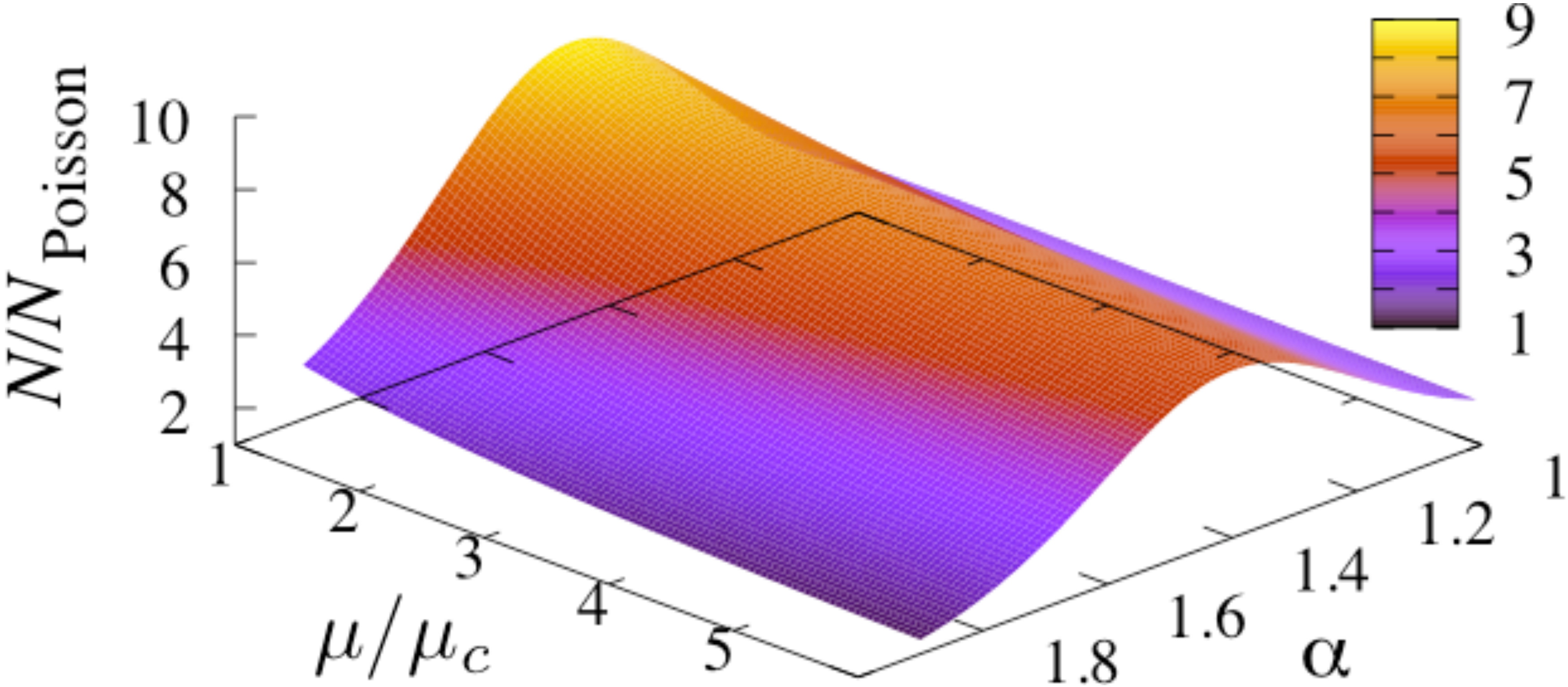
$$\beta(t',t) \quad \longrightarrow \quad \beta(t',t)e^{-\mu(t-t')}$$

$$n(t) \quad \longrightarrow \quad n(t)e^{-\mu t}$$

Final outbreak size

$$N(\infty) = \int_0^\infty dt\; n(t)\big|_{\mu=0}\, e^{-\mu t}$$

Immunization



$\mu_c = \lambda\langle k\rangle$

Conclusions

- Empirical evidence:
  - In many contact processes the inter-contact time distribution is subexponential
- Consequences:
  - Long delay for the first infectious contact
  - <u>Fast→Subexponential</u>→Exponential→Saturation
  - Larger outbreak size
- Outlook:
  - Empirical measurements
  - Epidemic growth models

Outlook: Computer worms

- Email viruses
  - Timing is the main factor
- IP address-scanning Worms
  - Timing may be the main factor: login sessions
- Self-broadcasting Email worms
  - The contact heterogeneity may be more relevant than in the case of Email viruses

Outlook: HIV/AIDS

- Contact heterogeneity is also determinant:
  - ✓ Vazquez 2005 (unpublished)

Outlook: Airborne viruses

- Eubank … next week