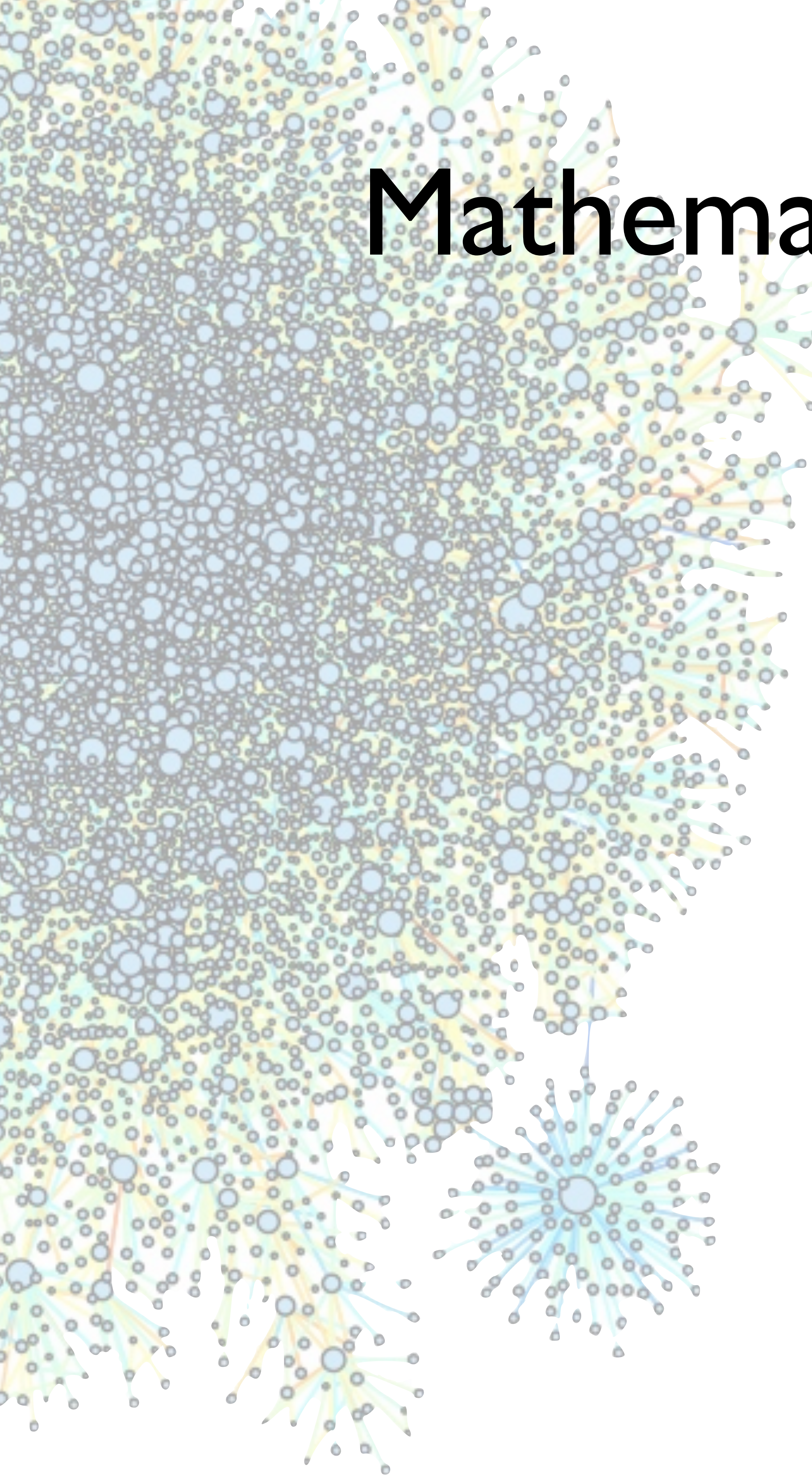


Mathematical Modeling of Human Communication

Giovanna Miritello

Esteban Moro

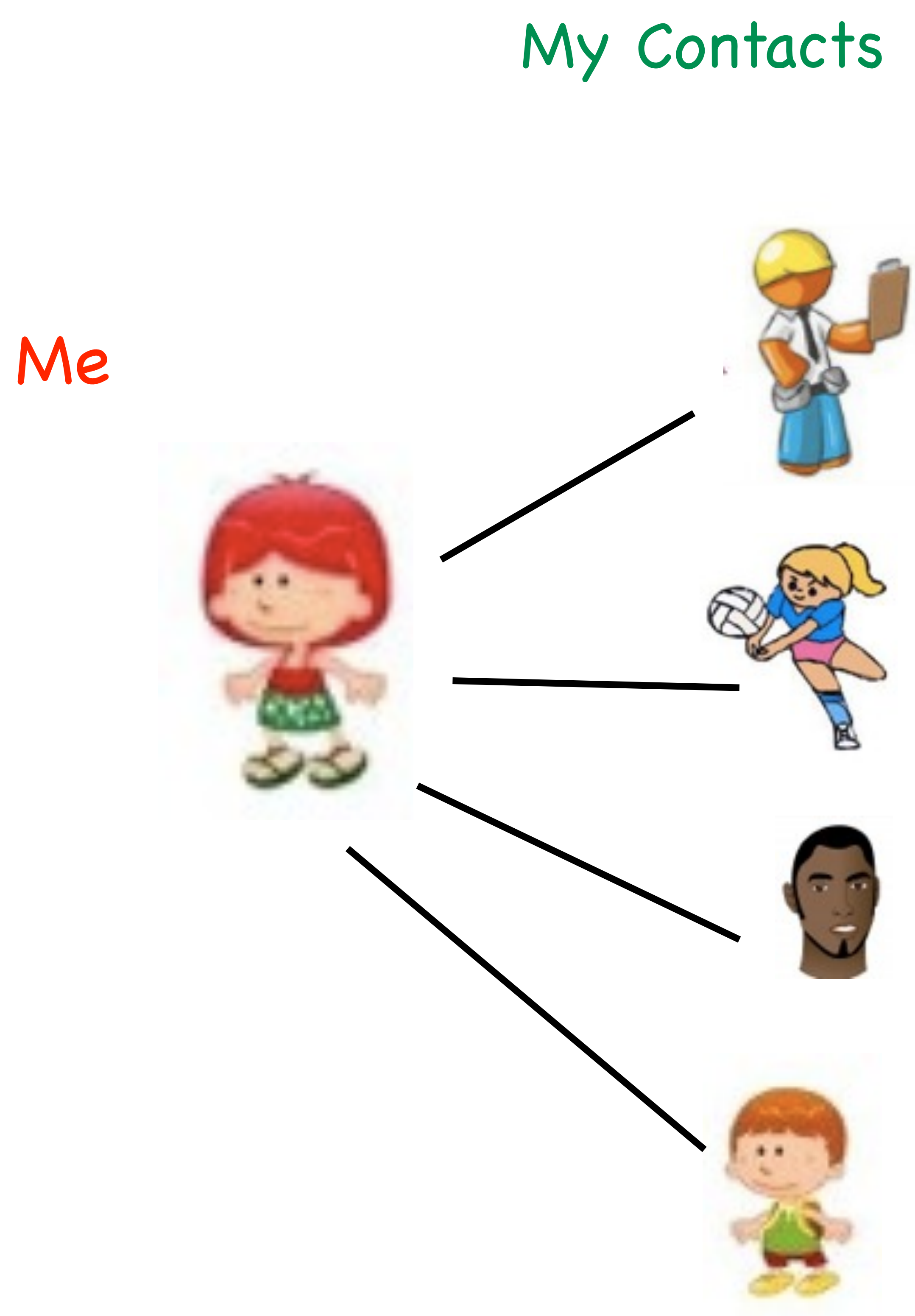
Rubén Lara

Universidad Carlos III de Madrid

Our work and relevance for Telefonica

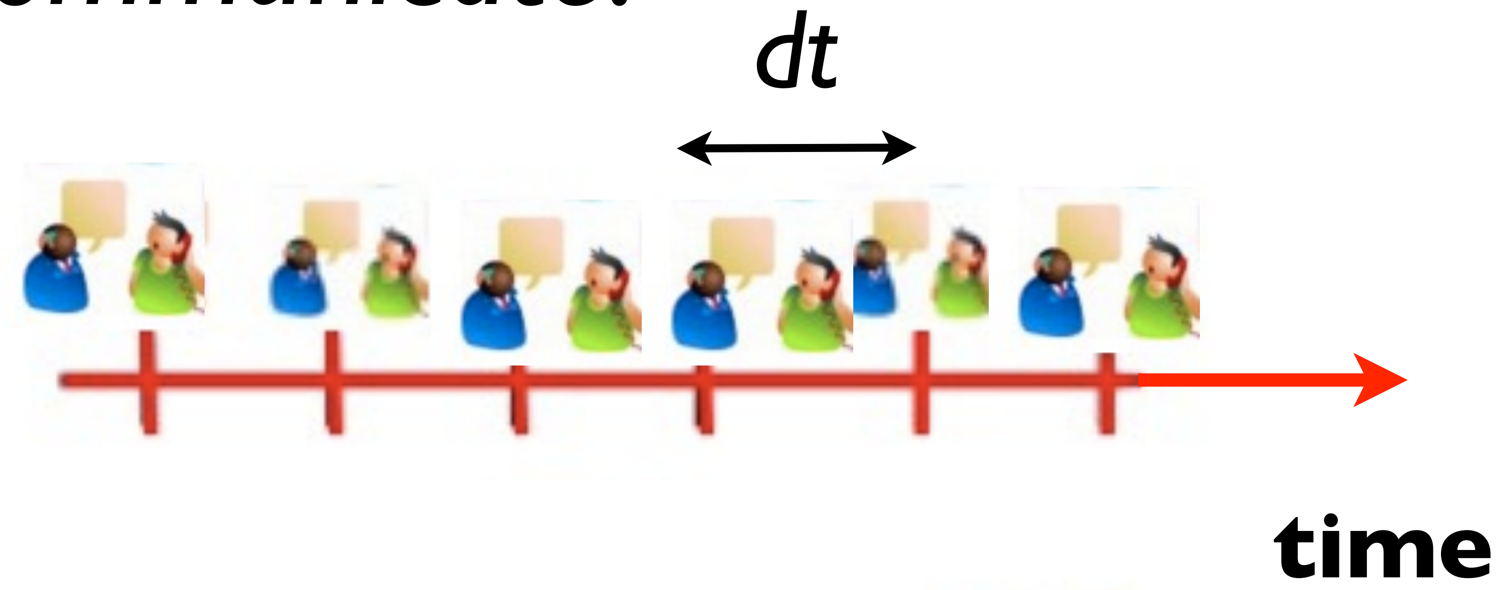
- *Characterize interactions between users*
- *Understand how people communicate*
- *Model (and predict) how information spreads between individuals*

How do people communicate?



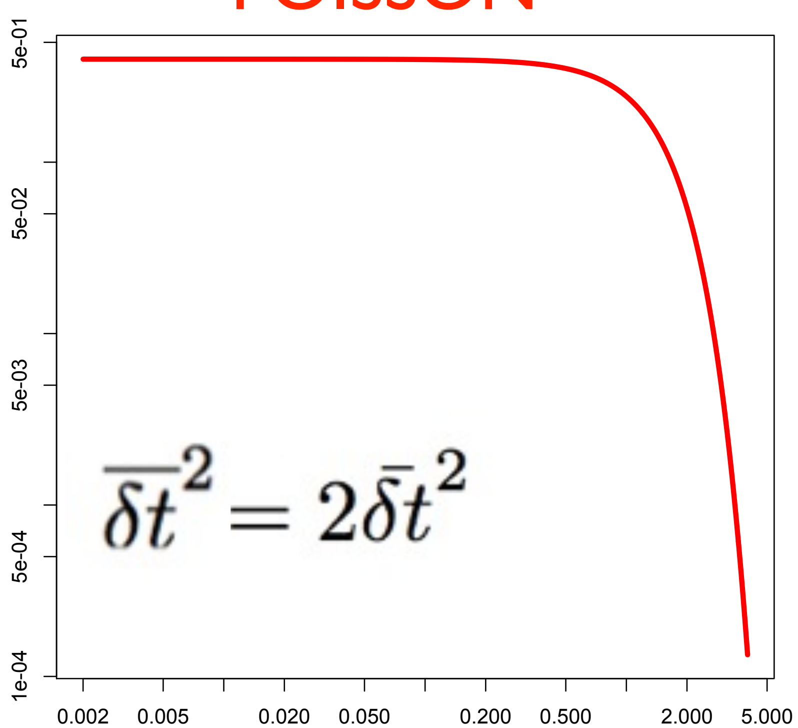
How do people communicate?

Homogeneous



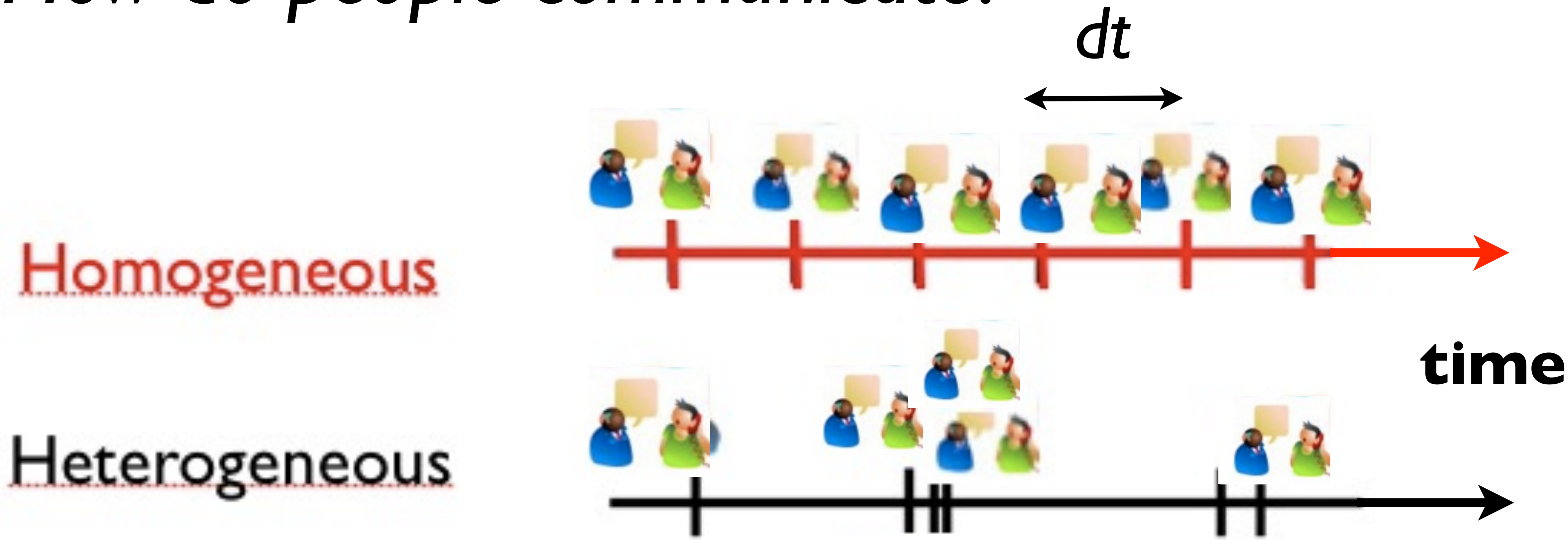
POISSON

P(dt)

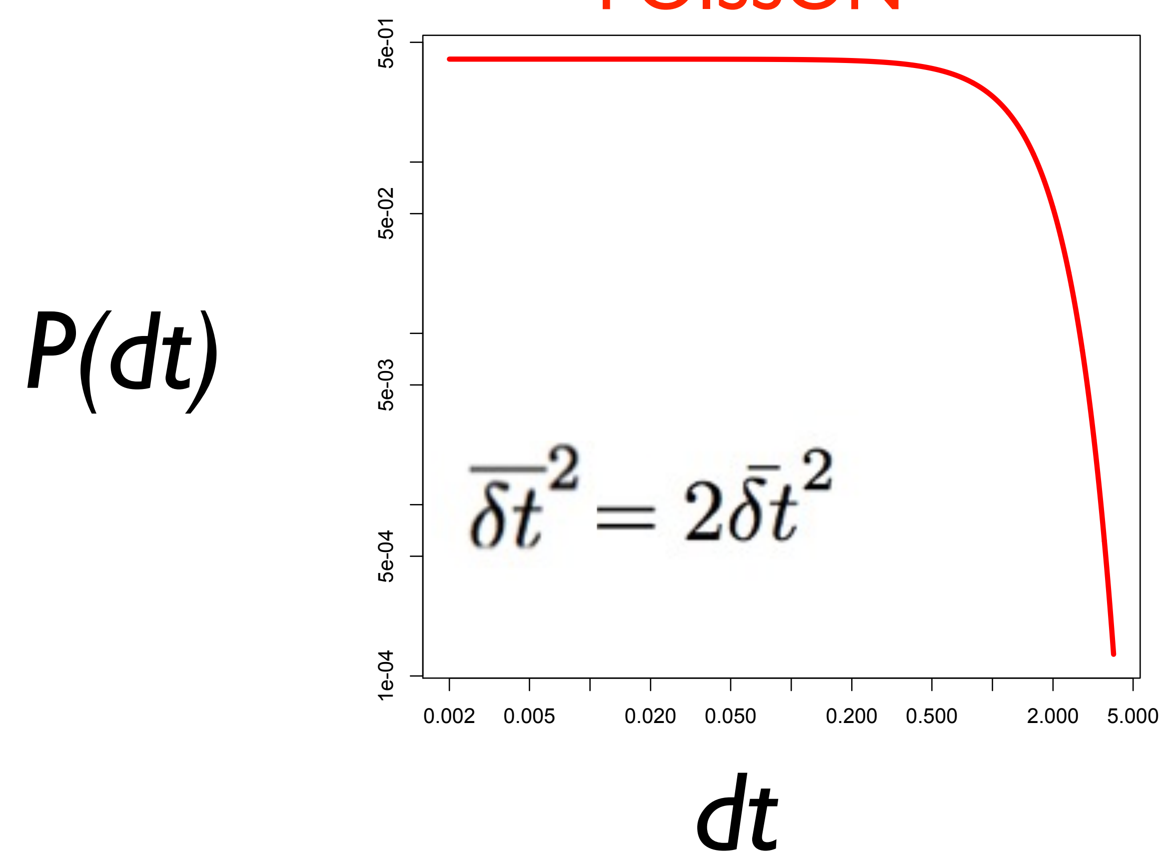


dt

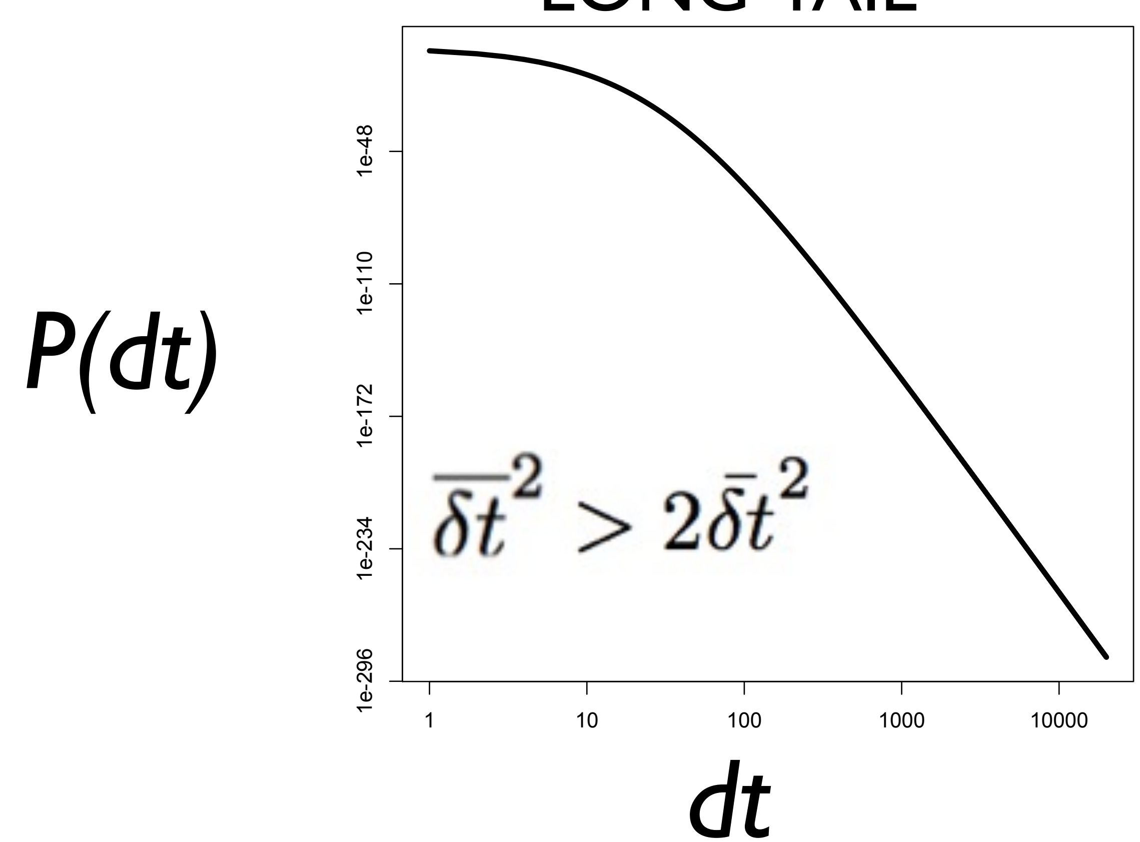
How do people communicate?



POISSON

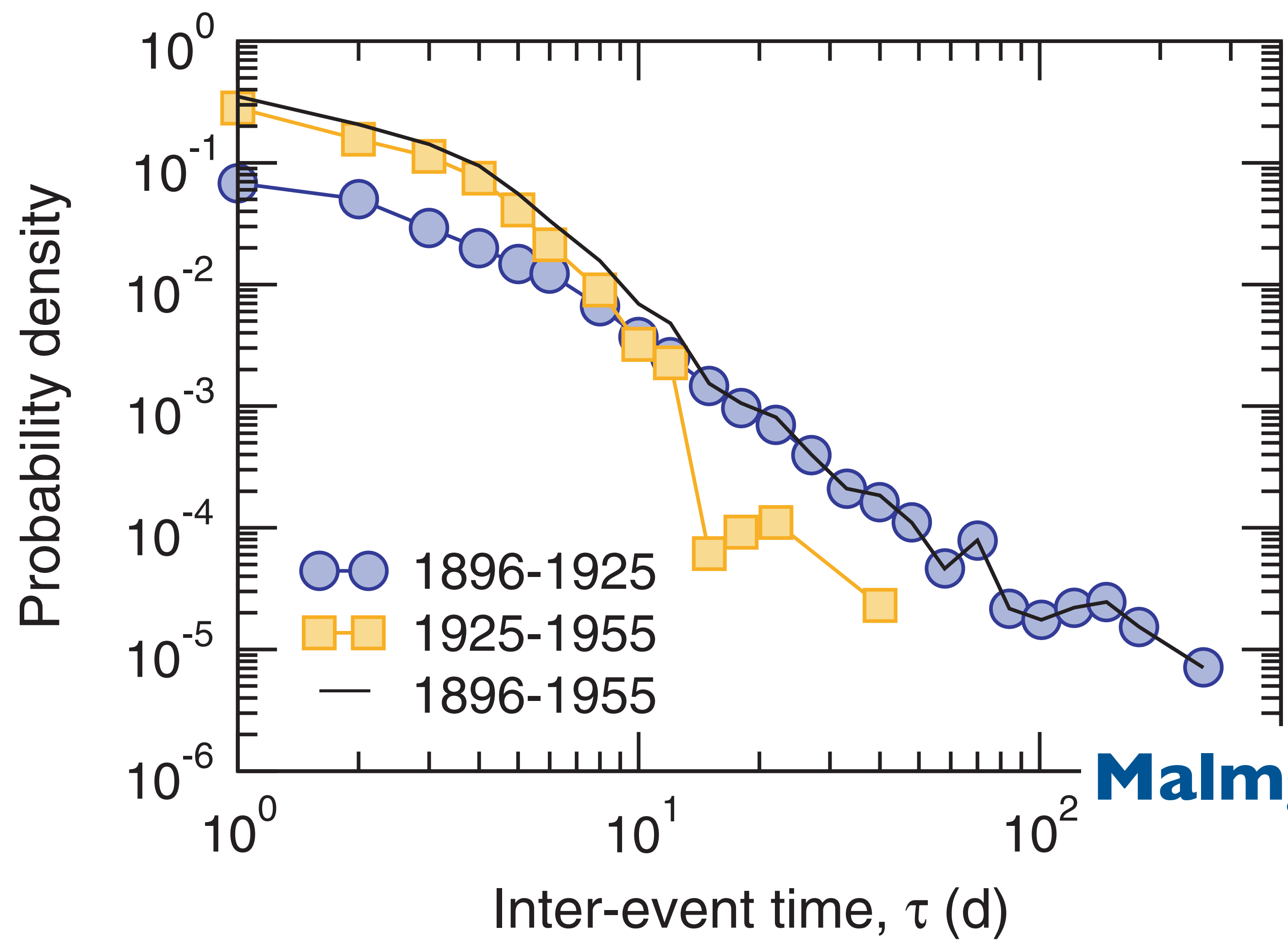


LONG-TAIL

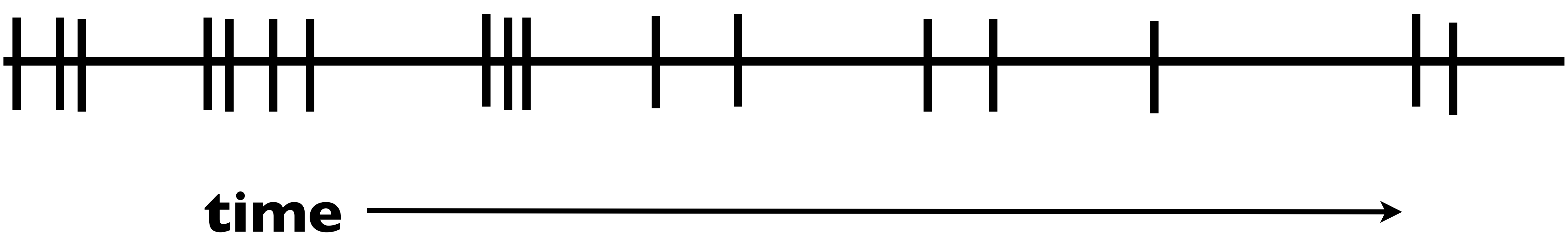


Human communication happens in cascades of events (bursts)

Albert Einstein's letter correspondence activity

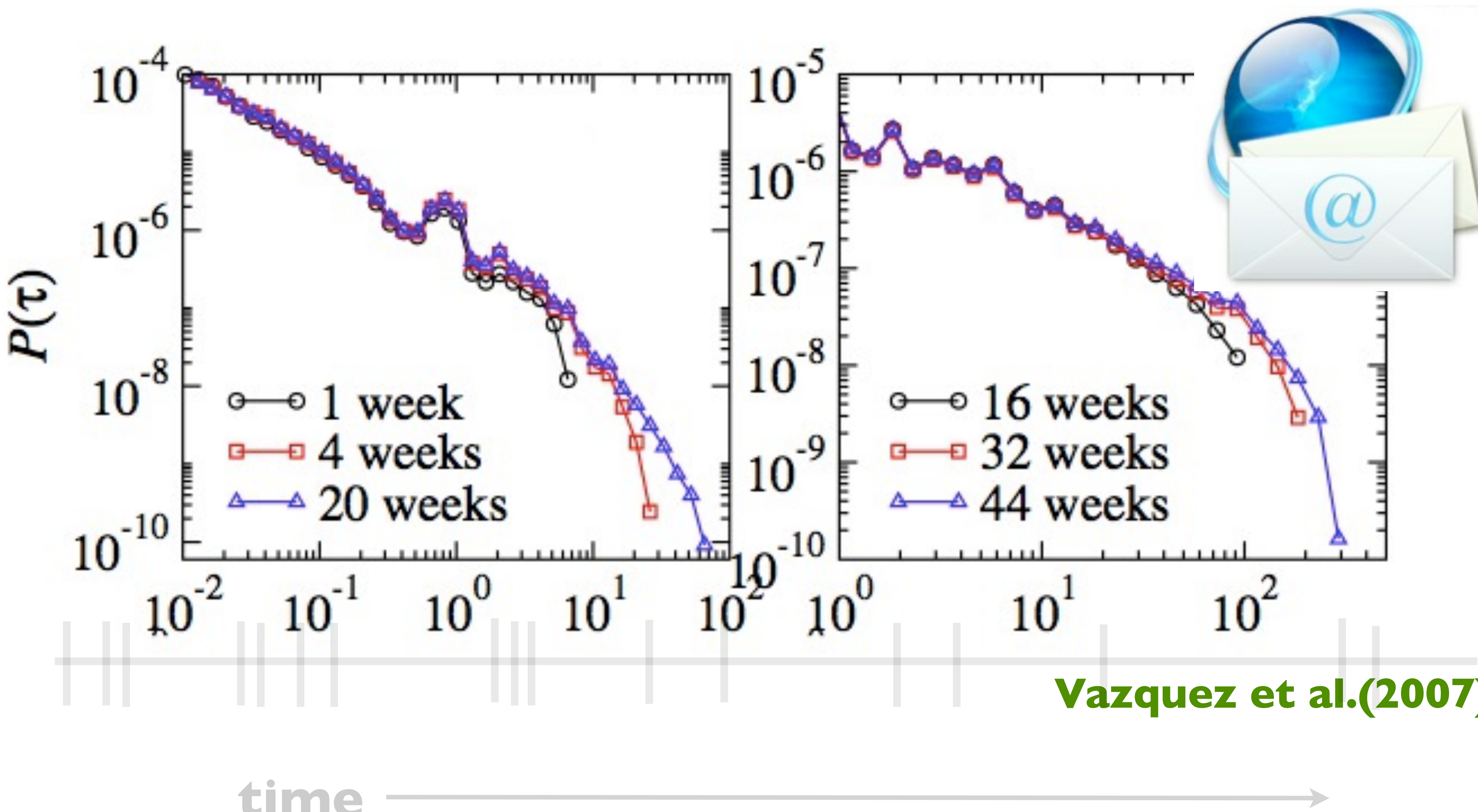


Malmgren et al.(2009)



Human communication happens in cascades of events (bursts)

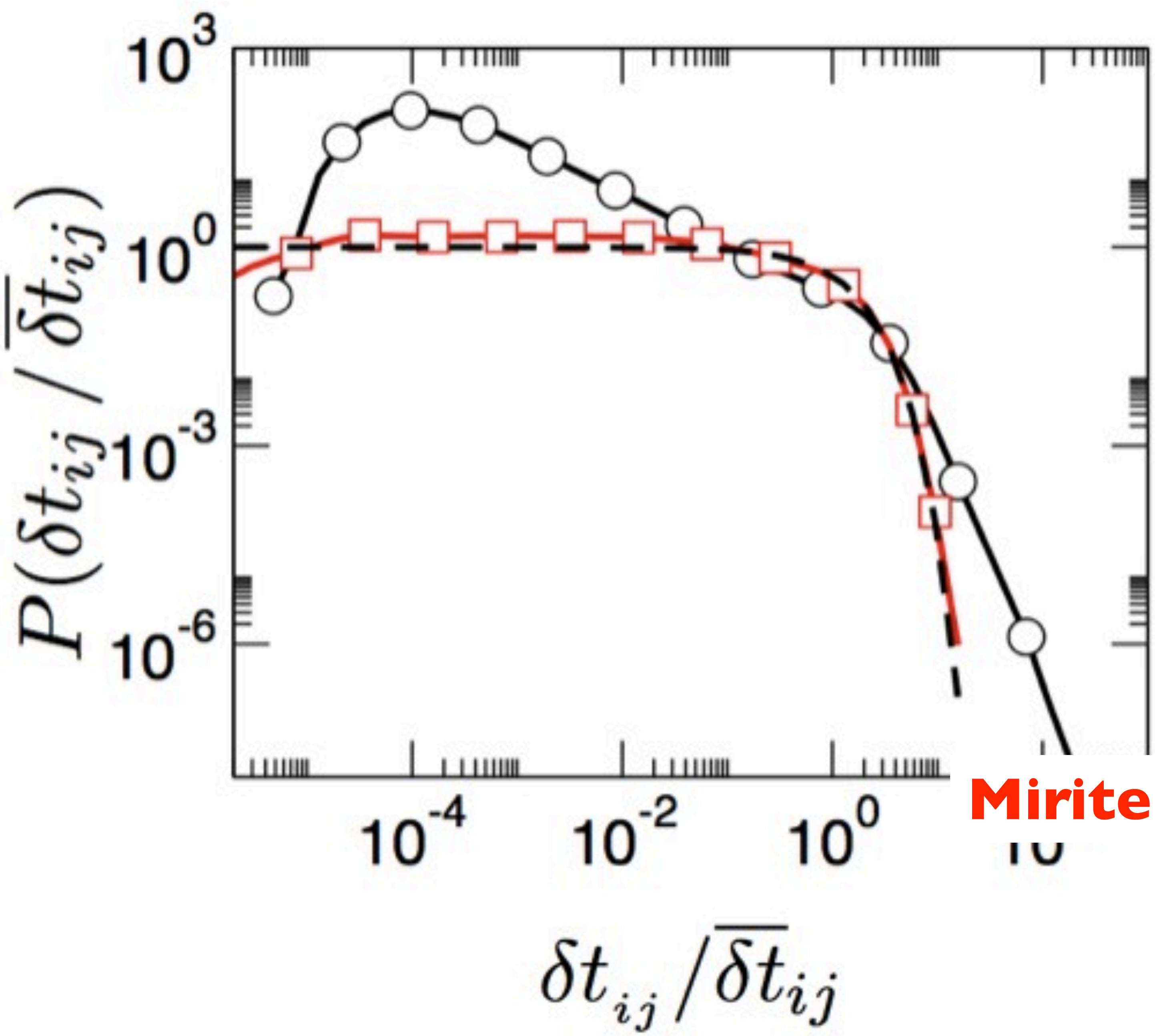
E-mails communication



Vazquez et al.(2007)

Human communication happens in cascades of events (bursts)

Mobile phone communication



Miritello, Moro, Lara (2010)

Telefónica

Universidad Carlos III de Madrid

What is the impact of real bursty behavior in information spreading?

Universidad Carlos III de Madrid

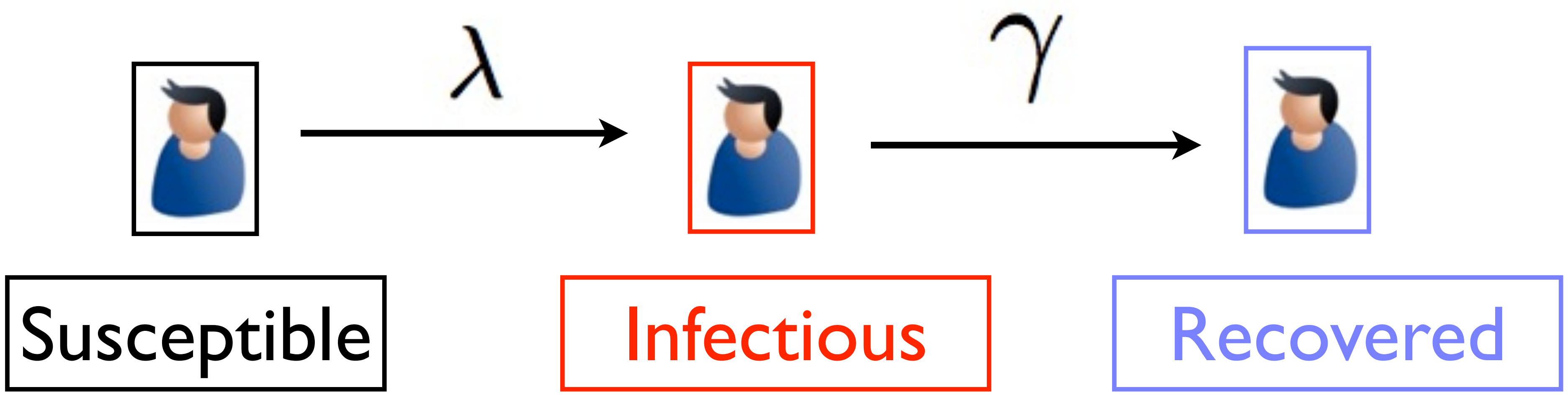
What is the impact of real bursty behavior in information spreading?

Information diffusion model from epidemiology

SIR MODEL

Daley and Kendall , Nature, vol. 204, no. 4963 (1964)

Murray, Math. Biol, vol.17, Springer, NY (2002)



- Fully mixed population
- Each user has the same probability of infection
- Homogeneous time

$$\frac{ds}{dt} = -\lambda i s$$

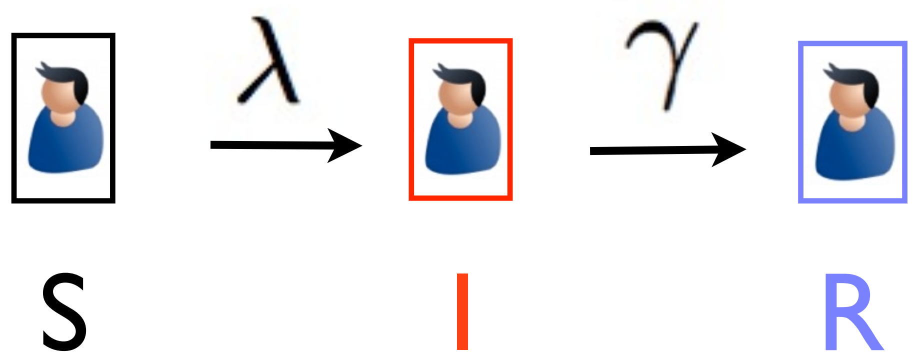
$$\frac{di}{dt} = \lambda i s - \gamma i$$

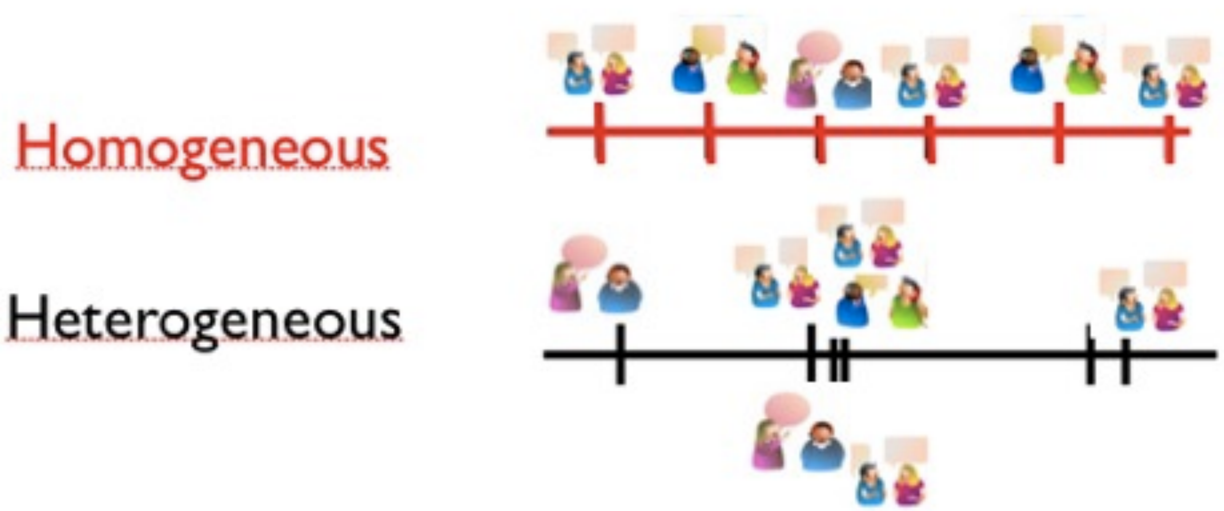
$$\frac{dr}{dt} = \gamma i$$

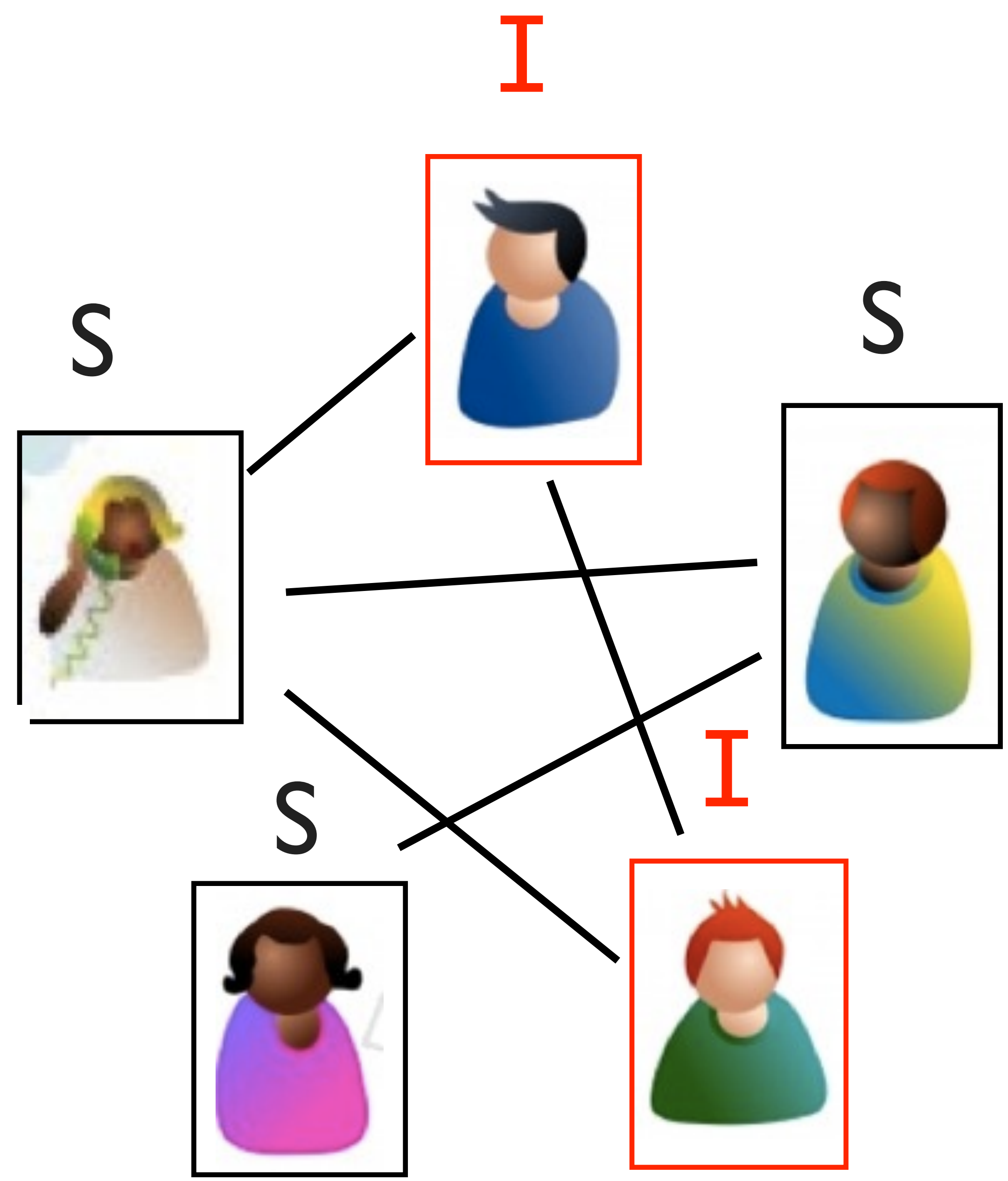
$$s + i + r = 1$$

What is the impact of real bursty behavior in information spreading?

SIR MODEL







SIR Model on Real Data from Mobile Phone Calls

Country X

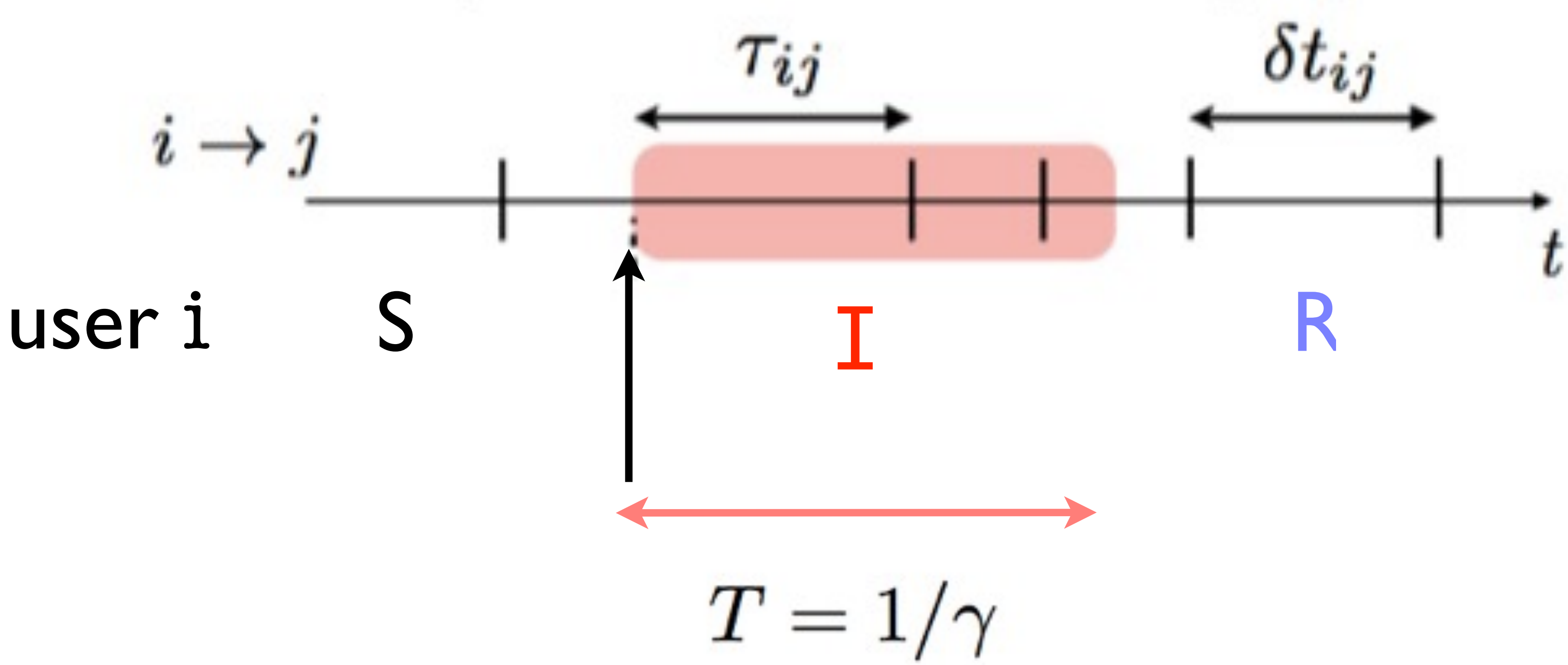
- CDR
- 11 months of data
- 23M people
- 2300M (call) links

07514422955	07872909981	15	000003	C
07802715977	07783276078	15	347	C
07872620521	07872618243	18	000036	C
07740086040	07821842408	24	000069	C
07543053528	07595184583	26	000061	C
07740469727	07766415523	34	20	C
07516626552	07522398677	35	000033	C
07907193548	07856616326	35	000004	C
07907378133	07873905835	41	000021	C
07783276078	07543466075	42	000175	C
07729856086	07769974265	47	000017	C
07894668339	07711598505	48	000009	C
07521620630	07597377490	53	000209	C
07889800622	07776301842	53	145	C
07725326776	07595030724	55	292	C

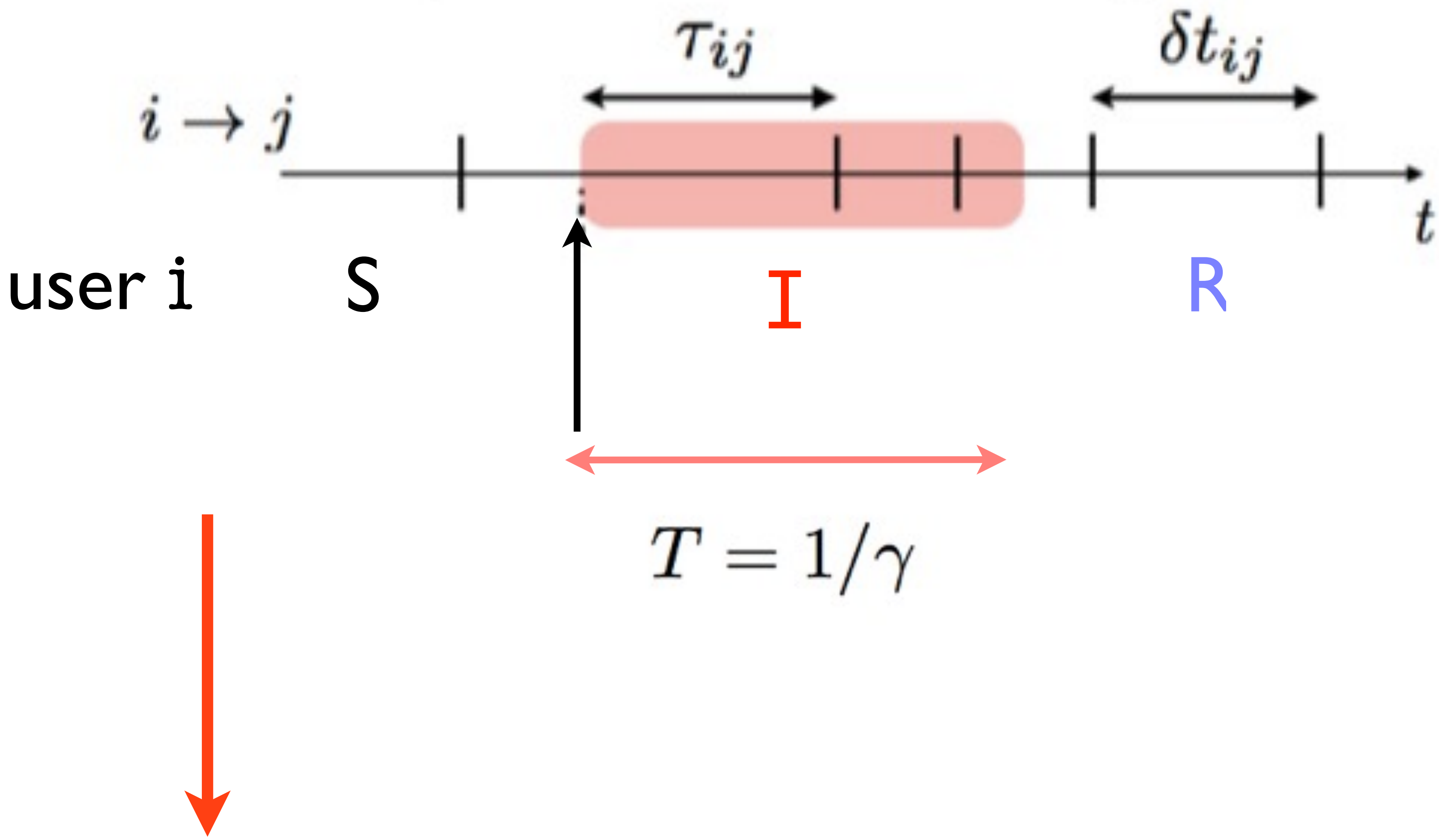
Telefonica

Universidad Carlos III de Madrid

Modeling Human Communication

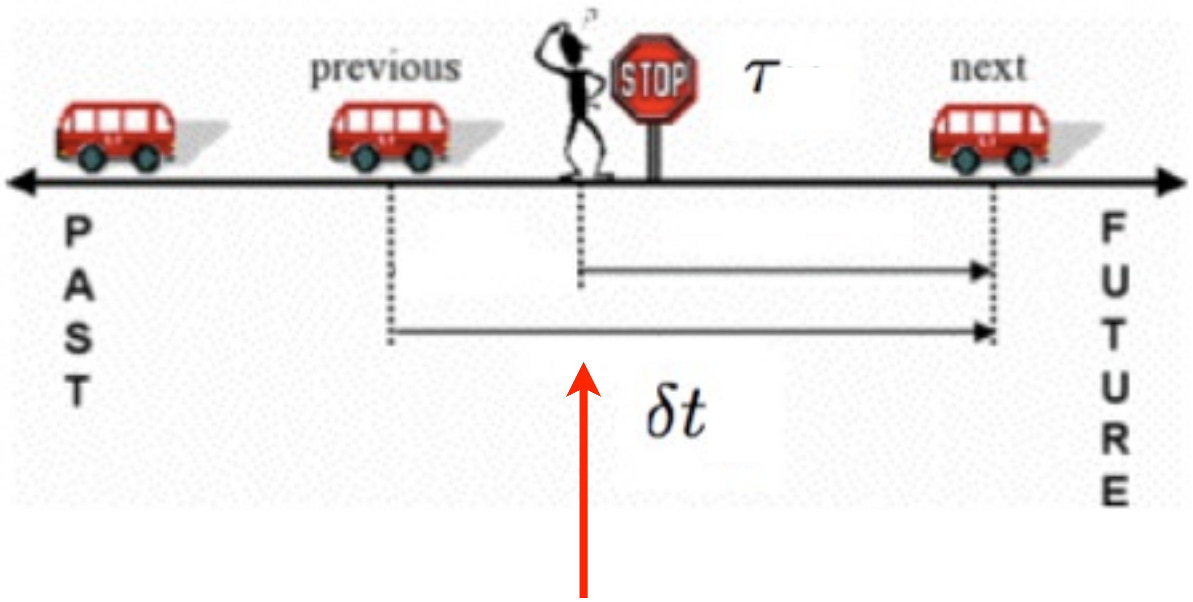


Modeling Human Communication



Bus Waiting Time Paradox: how long do you wait for the bus?

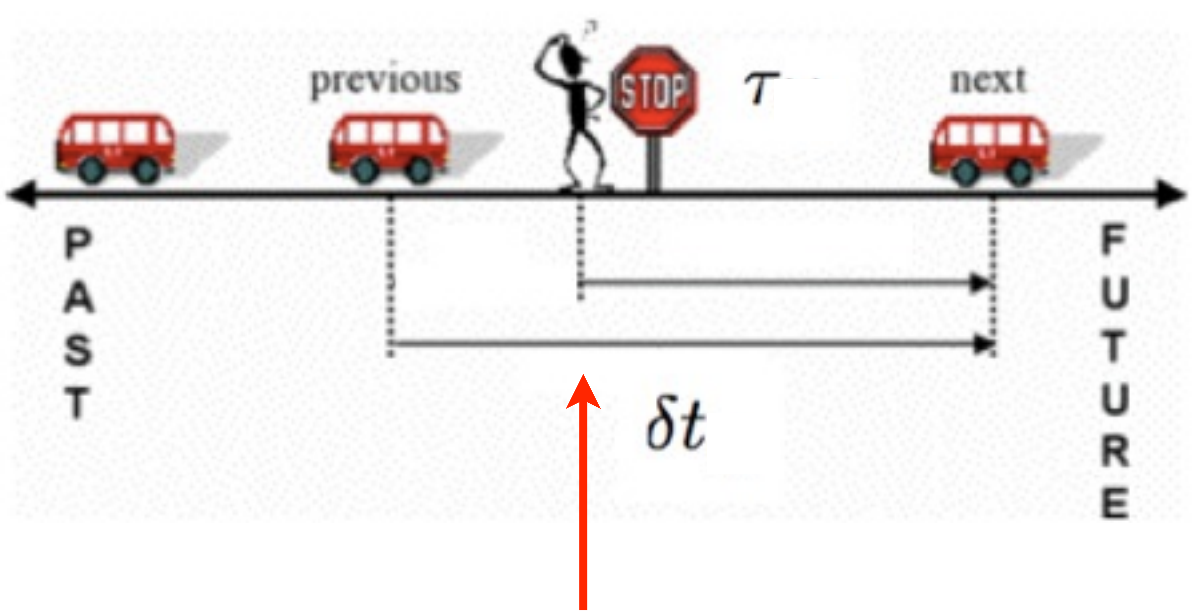
Bus Waiting Time Paradox: how long do you wait for the bus?



- you arrive at random t
- buses arrive according to $Q(\delta \mathrm{t})$
- the mean interval between buses is $\bar{\delta t}$

$$\bar{\tau}\ ?$$

Bus Waiting Time Paradox: how long do you wait for the bus?



- you arrive at random t
- buses arrive according to $Q(\delta t)$
- the mean interval between buses is $\bar{\delta t}$

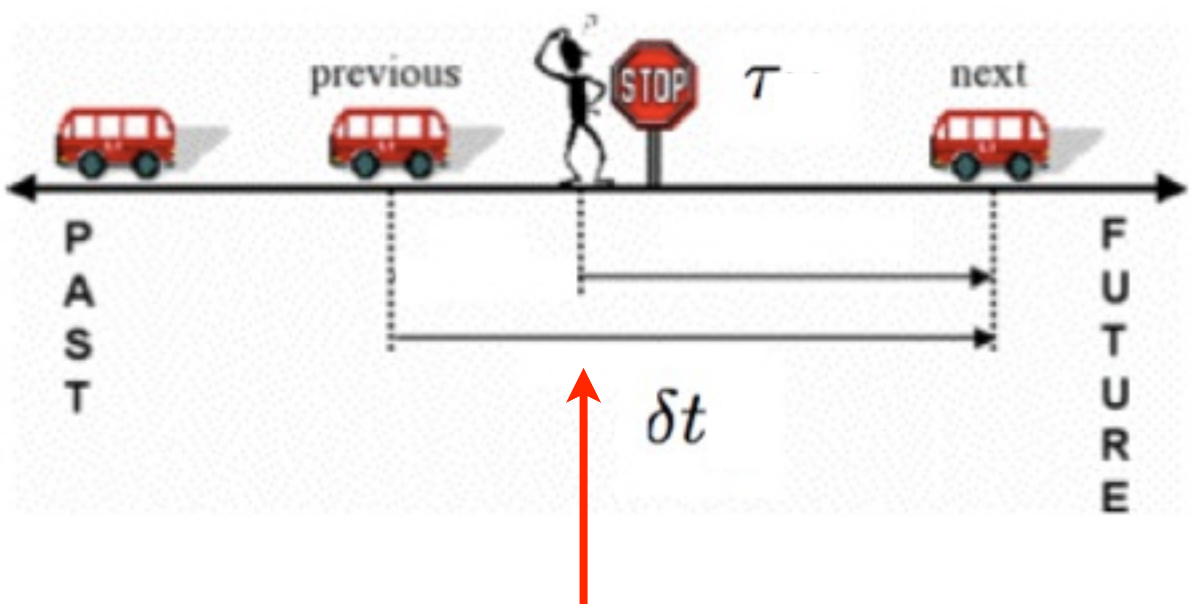
$$\bar{\tau}\ ?$$

Probability to arrive at an interval δt

$$\frac{\delta t Q(\delta t)}{\bar{\delta t}}$$

Bus Waiting Time Paradox: how long do you wait for the bus?

- you arrive at random t
- buses arrive according to $Q(\delta t)$
- the mean interval between buses is $\bar{\delta t}$



$$\bar{\tau}\ ?$$

given the interval δt

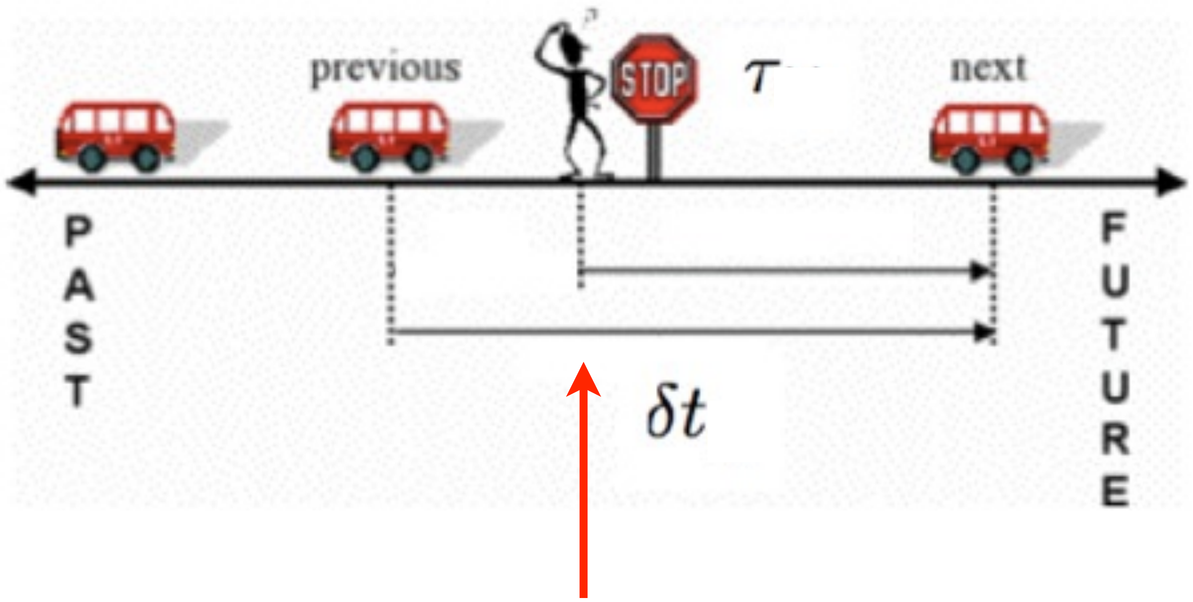
τ is equally distributed

Probability to arrive at an interval δt_{ij}

$$\frac{\delta t Q(\delta t)}{\overline{\delta t}} \quad \frac{1}{\delta t}$$

Bus Waiting Time Paradox: how long do you wait for the bus?

- you arrive at random t
- buses arrive according to $Q(\delta t)$
- the mean interval between buses is $\bar{\delta t}$



$$\bar{\tau}\,?$$

given the interval δt_{ij}

τ_{ij} is equally distributed

Probability to arrive at an interval δt_{ij}

Finally we average over all the $\delta t \geq \tau$

$$P(\tau) = \int_{\tau}^{\infty} d\delta t \, \frac{\delta t Q(\delta t)}{\bar{\delta t}} \, \frac{1}{\delta t}$$

Bus Waiting Time Paradox: how long do you wait for the bus?

$$P(\tau) = \int_\tau^\infty d\delta t \; \frac{\delta t Q(\delta t)}{\overline{\delta t}} \; \frac{1}{\delta t} \; = \frac{1}{\overline{\delta t}} \int_\tau^\infty d\delta t Q(\delta t)$$

$$\bar{\tau} \, ?$$

$$\bar{\tau} = \int_0^\infty d\tau \; \tau \; P(\tau) \; = \frac{\overline{\delta t}}{2} \left(1 + \frac{\sigma_{\delta t}^2}{\overline{\delta t}^2} \right)$$

$$\bar{\tau} = \frac{1}{2} \frac{\overline{\delta t^2}}{\overline{\delta t}}$$

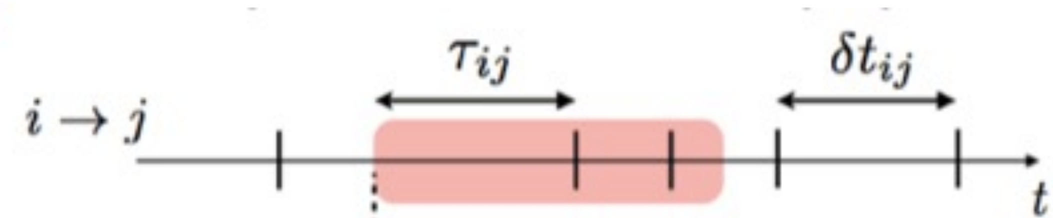
Bus Waiting Time Paradox: how long do you wait for the bus?

$$P(\tau) = \int_\tau^\infty d\delta t \, \frac{\delta t Q(\delta t)}{\overline{\delta t}} \, \frac{1}{\delta t} = \frac{1}{\overline{\delta t}} \int_\tau^\infty d\delta t Q(\delta t)$$

$$\bar{\tau} \, ?$$

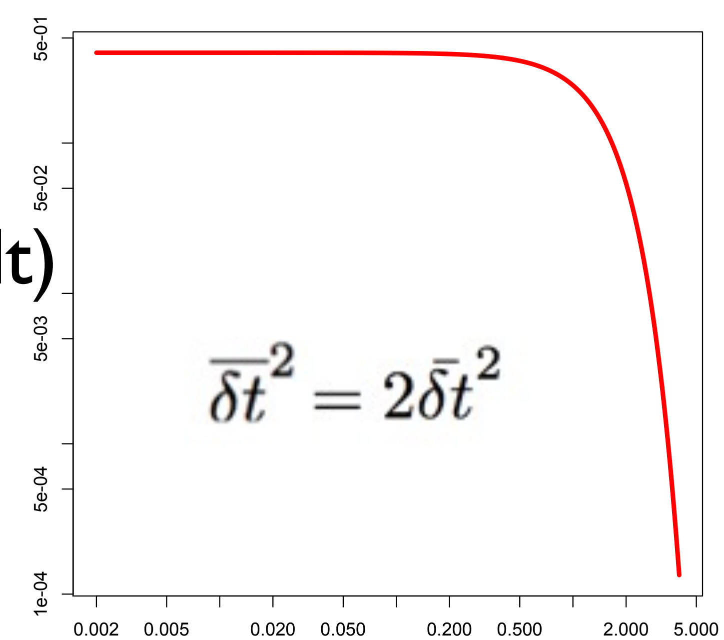
$$\bar{\tau} = \int_0^\infty d\tau \, \tau \, P(\tau) = \frac{\overline{\delta t}}{2} \left(1 + \frac{\sigma^2_{\delta t}}{\overline{\delta t}^2}\right)$$

$$\bar{\tau} = \frac{1}{2} \frac{\overline{\delta t^2}}{\overline{\delta t}}$$



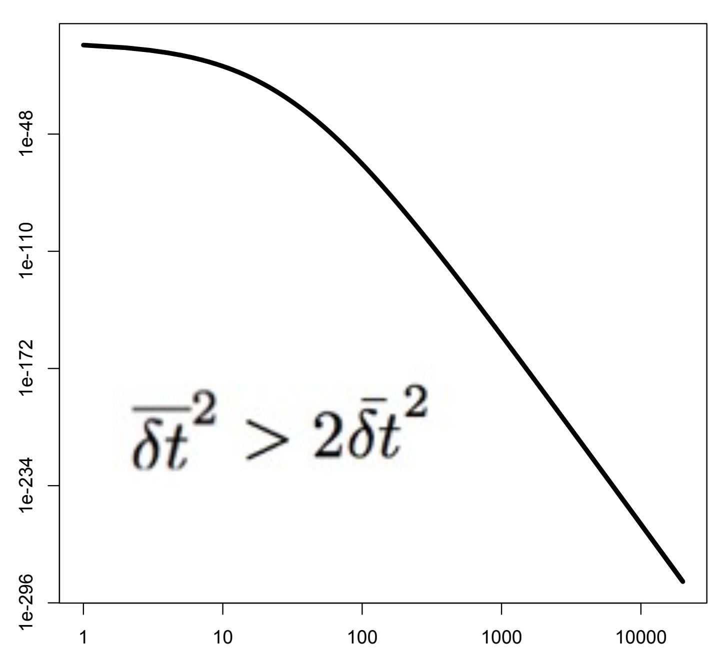
POISSON

Q(dt)



dt

LONG-TAIL



dt

The long-tail of

$Q(\delta t)$

makes the waiting time longer than the Poisson case

Telefonica

Universidad Carlos III de Madrid

Implications in information spreading

SIR Model on Real Data from Mobile Phone Calls

Country X

- CDR
- 11 months of data
- 23M people
- 2300M (call) links

07514422955	07872909981	15	000003	C
07802715977	07783276078	15	347	C
07872620521	07872618243	18	000036	C
07740086040	07821842408	24	000069	C
07543053528	07595184583	26	000061	C
07740469727	07766415523	34	20	C
07516626552	07522398677	35	000033	C
07907193548	07856616326	35	000004	C
07907378133	07873905835	41	000021	C
07783276078	07543466075	42	000175	C
07729856086	07769974265	47	000017	C
07894668339	07711598505	48	000009	C
07521620630	07597377490	53	000209	C
07889800622	07776301842	53	145	C
07725326776	07595030724	55	292	C

Telefónica

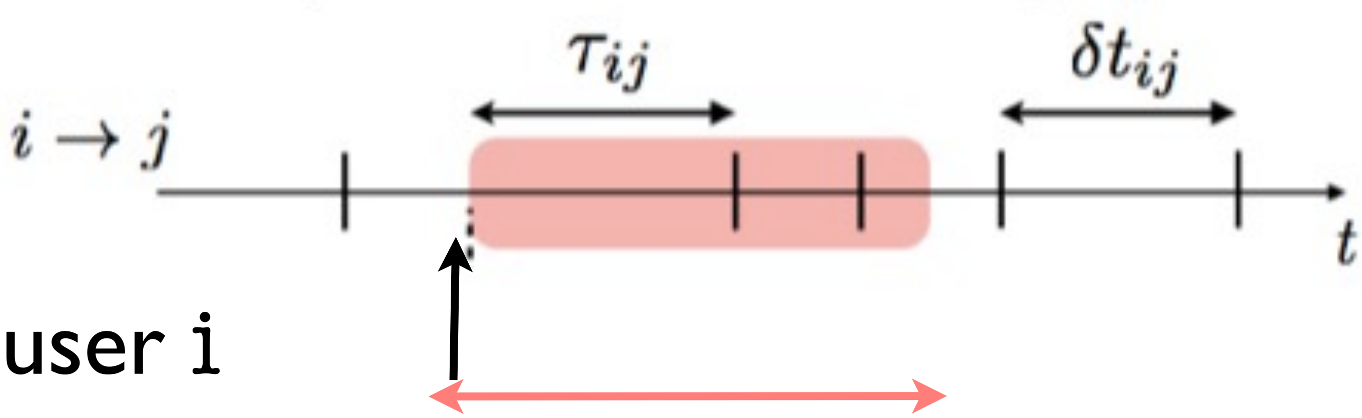
Universidad Carlos III de Madrid

Implications in information spreading

probability that the information
will be transmitted from i to j

Case $\lambda = 1$

M.E.Newman, Phys. Rev. E 66, 16128 (2002)

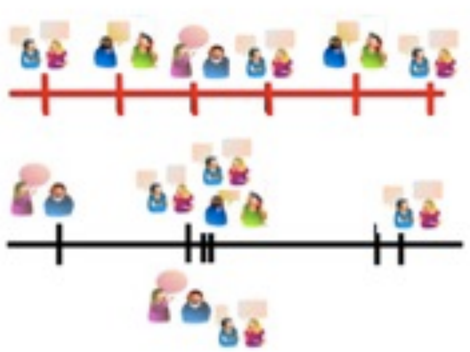


user i

$T = 1/\gamma$

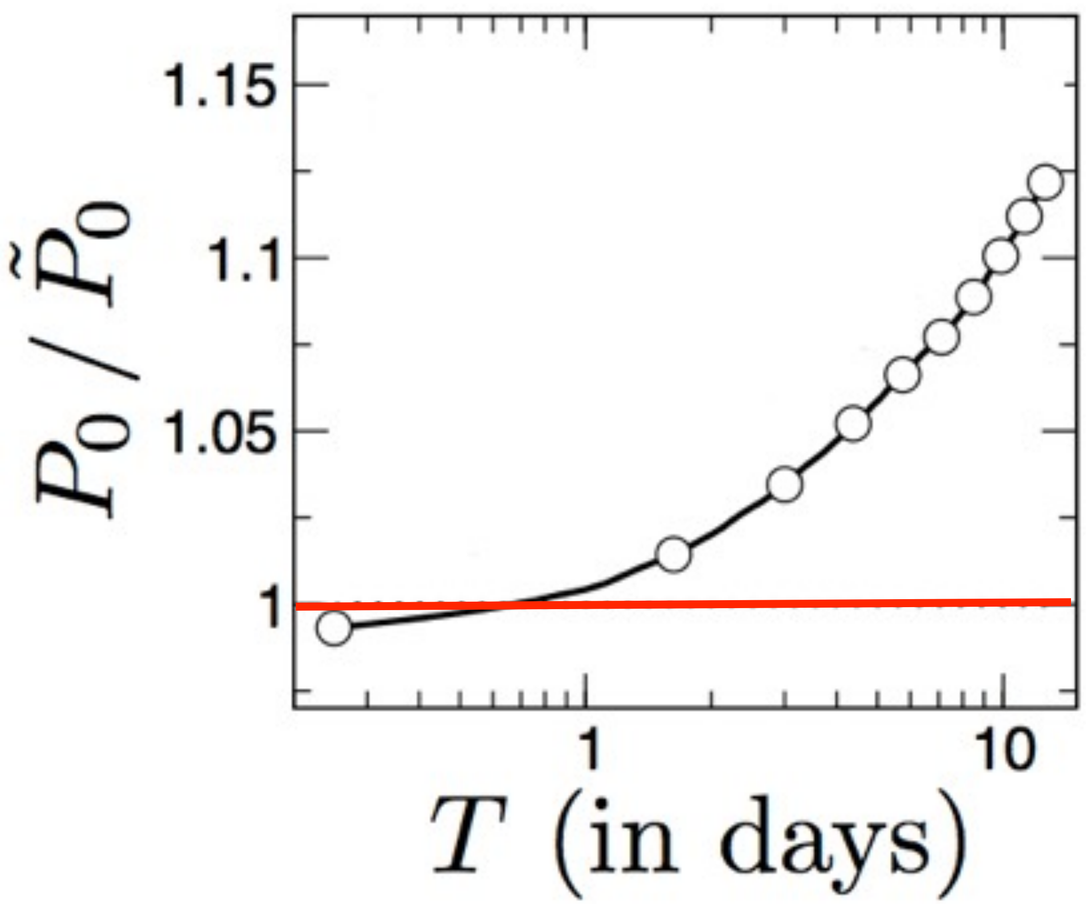
$$\mathcal{T}_{ij} = 1 - P^0_{ij}$$

$$P^0_{ij} = \int_T^\infty P(\tau_{ij})d\tau_{ij}$$



$$P^0_{ij} > \widetilde{P}^0_{ij}$$

$$\mathcal{T}_{ij} < \widetilde{\mathcal{T}}_{ij}$$

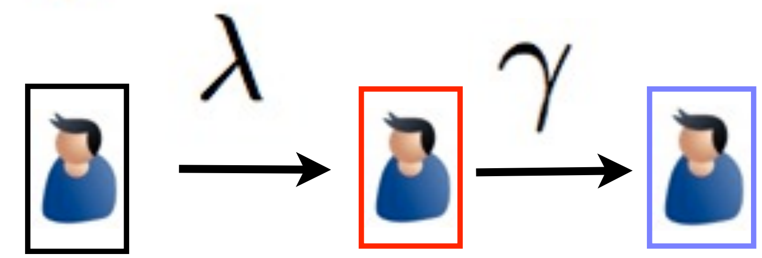


Telefonica
Universidad Carlos III de Madrid

Implications in information spreading

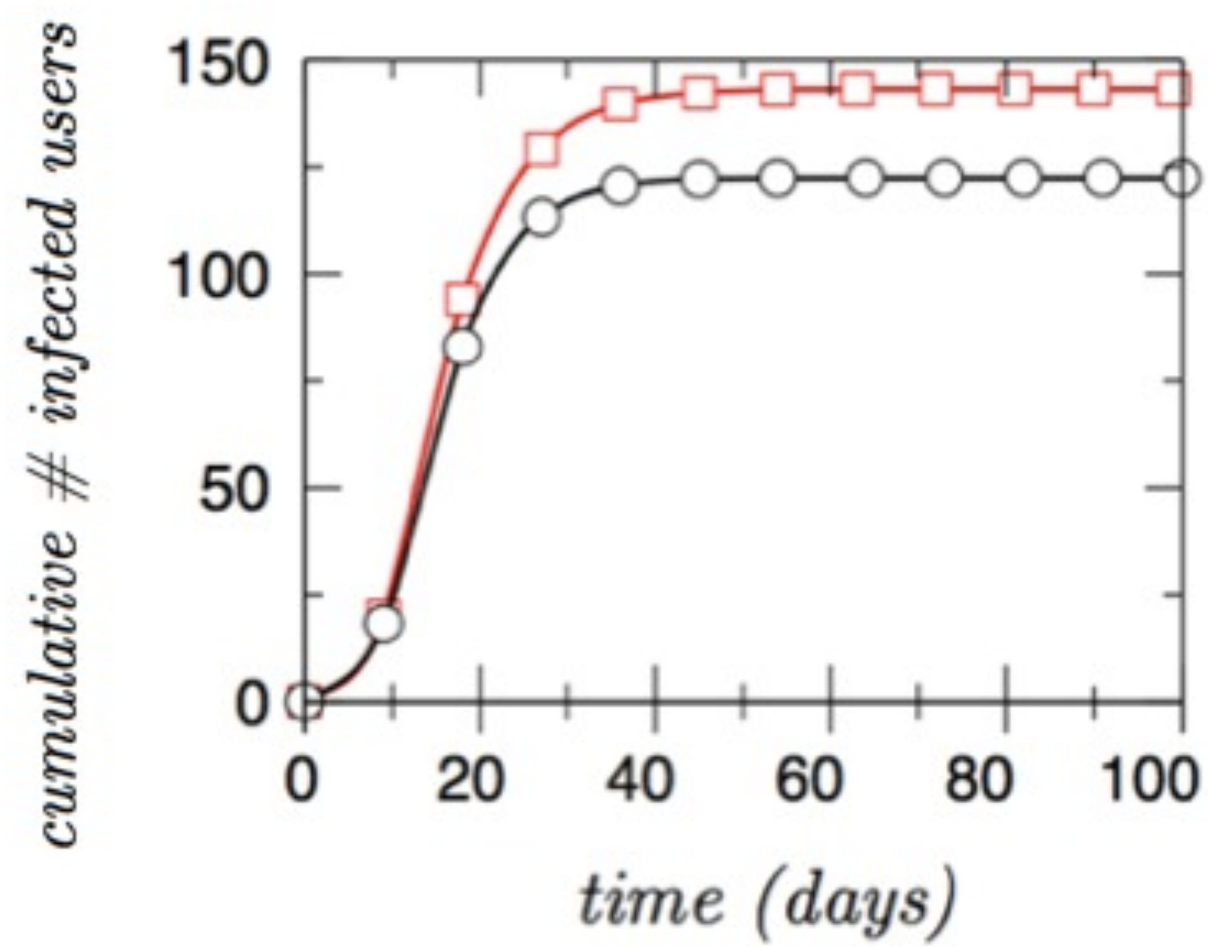
$$P^0_{ij} > \widetilde{P}^0_{ij} \quad \longrightarrow \quad T_{ij} < \widetilde{T}_{ij}$$

SIR MODEL on Real Data



Assumptions

- in any call infection happens with probability $\lambda = 1$
- an infected node recovers after a fixed $T = 1/\gamma$



The bursty real dynamics makes the reach narrower!

Summary

- The way humans communicate is very heterogeneous - bursty behavior - (hurry up and wait)

Barabasi, Bursts: The Hidden Pattern Behind Everything We Do, Dutton Books (2010)

- The observed behavior can be modelled and understood by means of mathematical tools

Daley and Kendall , Nature, vol. 204, no. 4963 (1964)
Murray, Math. Biol, vol.17, Springer, NY (2002)

M.E.Newman, Phys. Rev. E 66, 16128 (2002)

- Heterogeneity affects dynamical real processes e.g. information spreading

Miritello, Moro & Lara, PRE Rapid Comm. (2011)
http://arxiv.org/abs/1011.5367