Dynamics on networks

János Kertész
Central European University

Pisa Summer School September 2019

"How can the universe start with a few types of elementary particles at the big bang, and end up with life, history, economics, and literature? The question is screaming out to be answered but it is seldom even asked. Why did the big bang not form a simple gas of particles or condense into one big crystal? We see complex phenomena around us so often that we take for granted without looking for further explanation. In fact, until recently very little scientific effort was devoted to understanding why nature is complex."

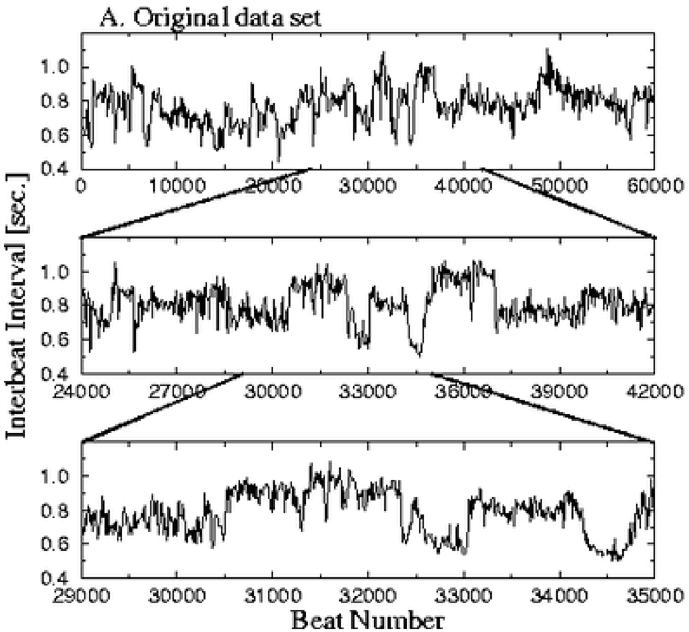
Per Bak, 1997

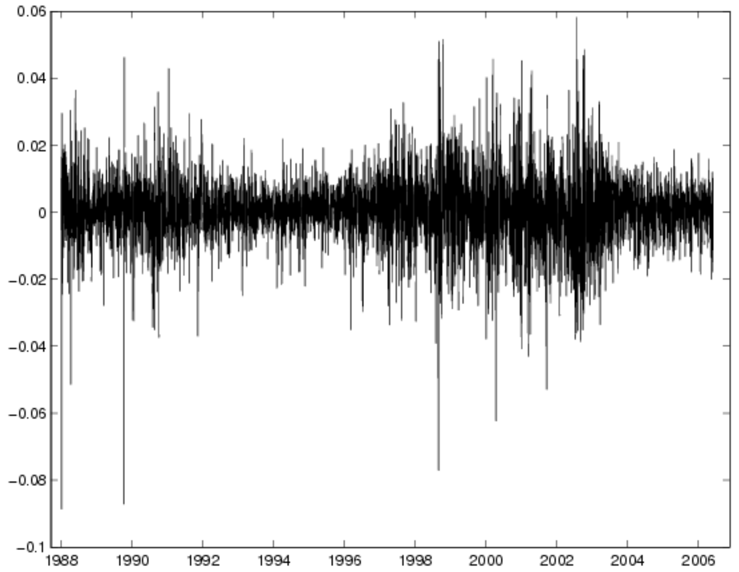
Complexity:
Inhomogeneity in space

Fractal objects:

Inhomogeneity in time:

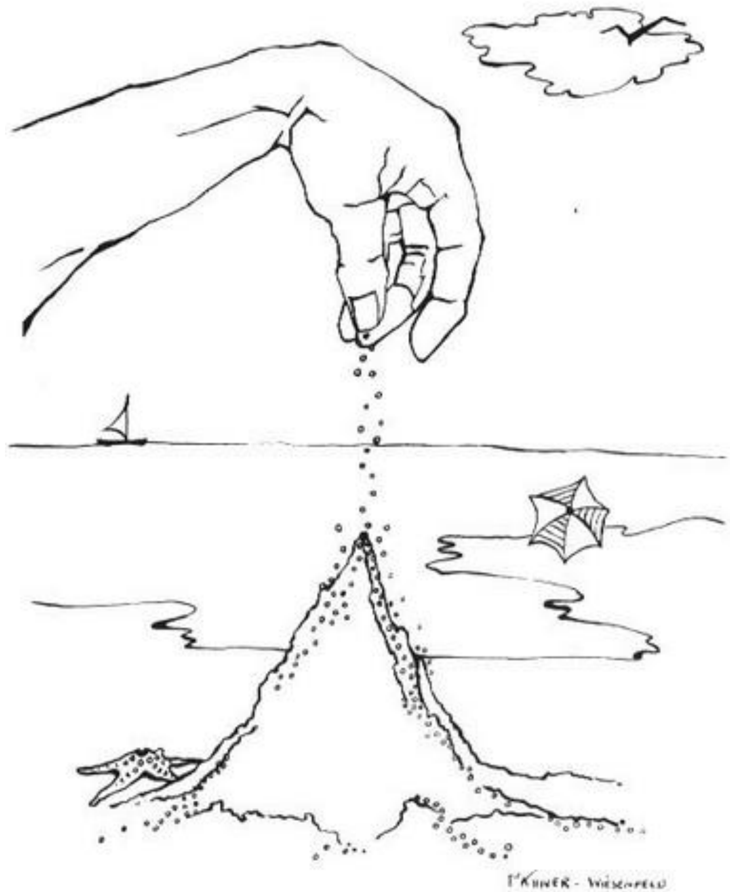
"Fractal" time series





Self-organized criticality

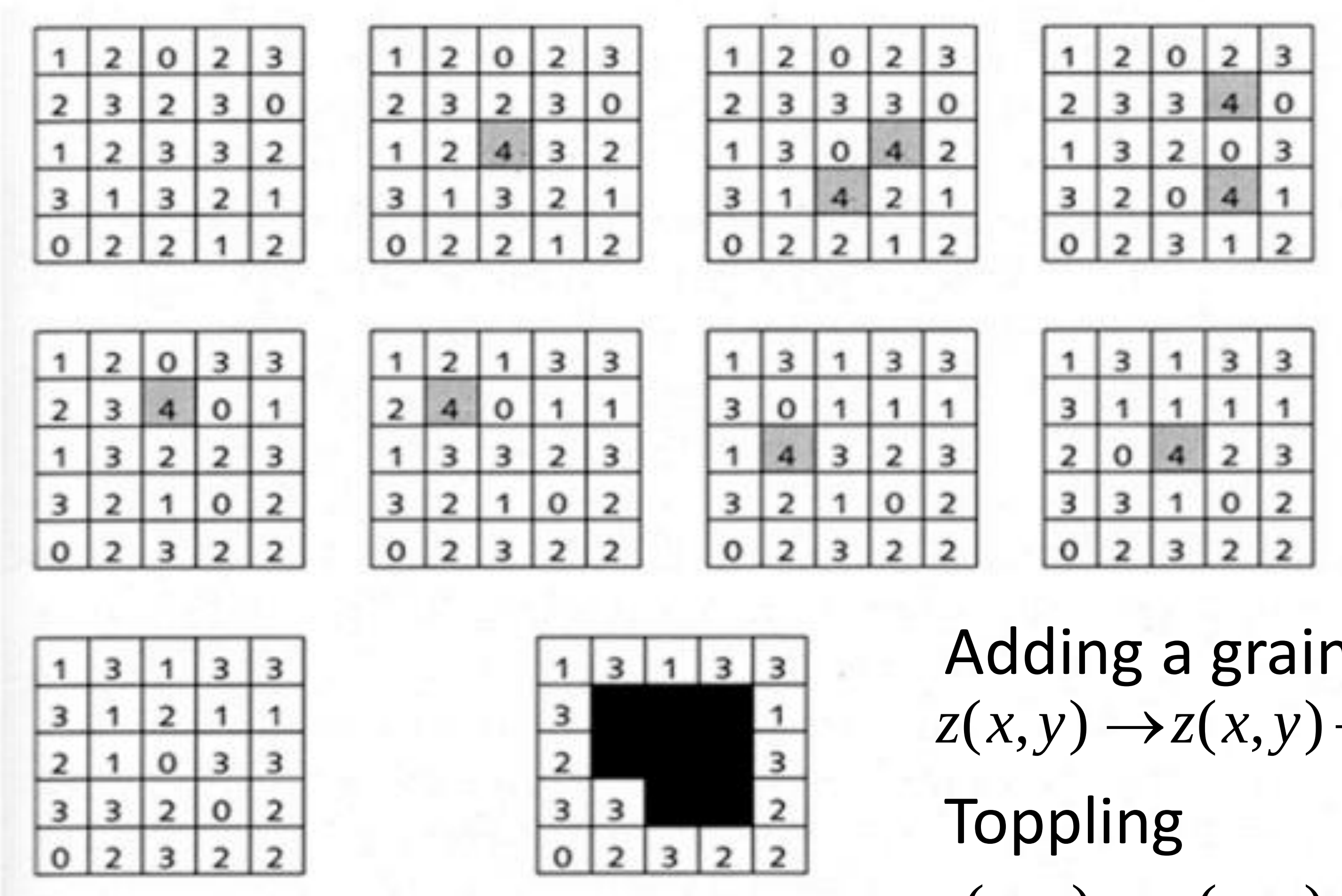
Sandpile model: emergence of scaling in driven systems



Slope is maintained constant with avalanches of inhomogeneous sizes and durations

Bak et al. PRL 1987

Cellular automaton model



Adding a grain

$$z(x,y) \rightarrow z(x,y)+1$$

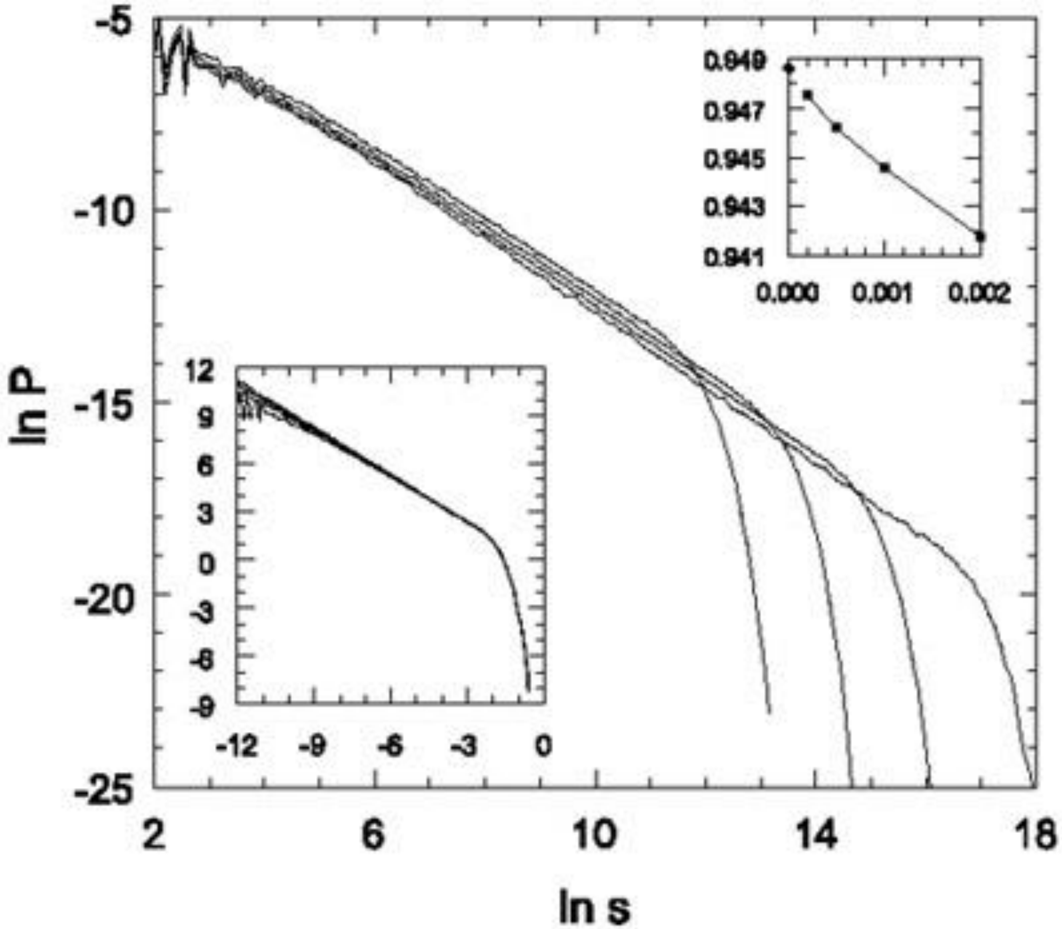
Toppling

$$z(x,y) \rightarrow z(x,y)-4$$

$$z(x\pm 1,y) \rightarrow z(x\pm 1,y)+1$$

$$z(x,y\pm 1) \rightarrow z(x,y\pm 1)+1$$

At the boundary grains get lost



s: size of avalanche
Similar power law
for duration

Non-trivial, scaling behavior also on complex networks (if threshold is dependent on k_i).

Goh et al. PRL (2003)

Interacting systems

In the above (network) models temporal heterogeneity results from interaction. There are units, which interact and result in complex behavior

The units are not always simple (but we hope that details are unimportant):

- Earthquakes (complex material processes) → GR law
- Solar flares (magneto-hydrodynamic processes) universal scaling distributions
- Neural activity (electrochemistry of cells) → scaling distributions
- Economy (including human agents) → price

Human communication

Interaction between people + human nature

Take a "simpler" view: Consider temporal pattern at single persons (nodes in the communication network) and calculate average behavior over many.

Simple model: Poissonian
(used until recently to design telephone crossbar capacities)
Advantage: A single parameter is enough
Disadvantage: Totally wrong

Poisson process

Events are separated by t_{ie} inter-event times with the independent distribution $P(t_{ie}) = e^{-\lambda t_{ie}}$
The expectation value of the number of events in an interval τ has a Poisson distribution:

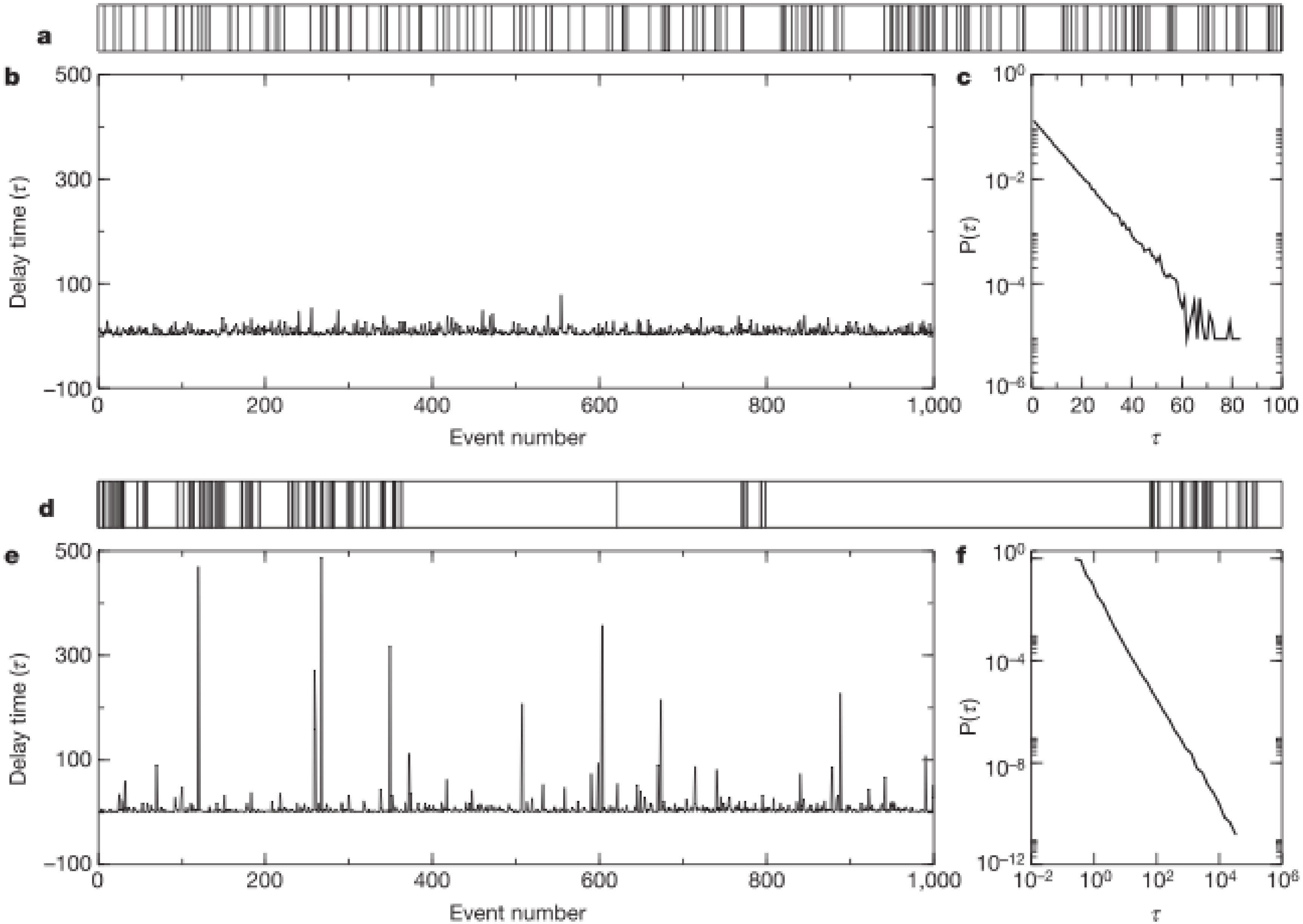
$$P(N_\tau) = e^{-\lambda\tau}\frac{(\lambda\tau)^{N_\tau}}{N_\tau!}$$ Homogeneous Poisson process

If λ depends on t

$$P(N_\tau) = e^{-\lambda_{ba}}\frac{(\lambda_{ba})^{N_\tau}}{N_\tau!} \quad \text{with } \lambda_{ba} = \int_a^b \lambda(t)dt$$

Non-homogeneous Poisson process

Poissonian vs bursty activity



A.-L. Barabási, Nature (2005)

Many bursty phenomena in human behavior and Nature

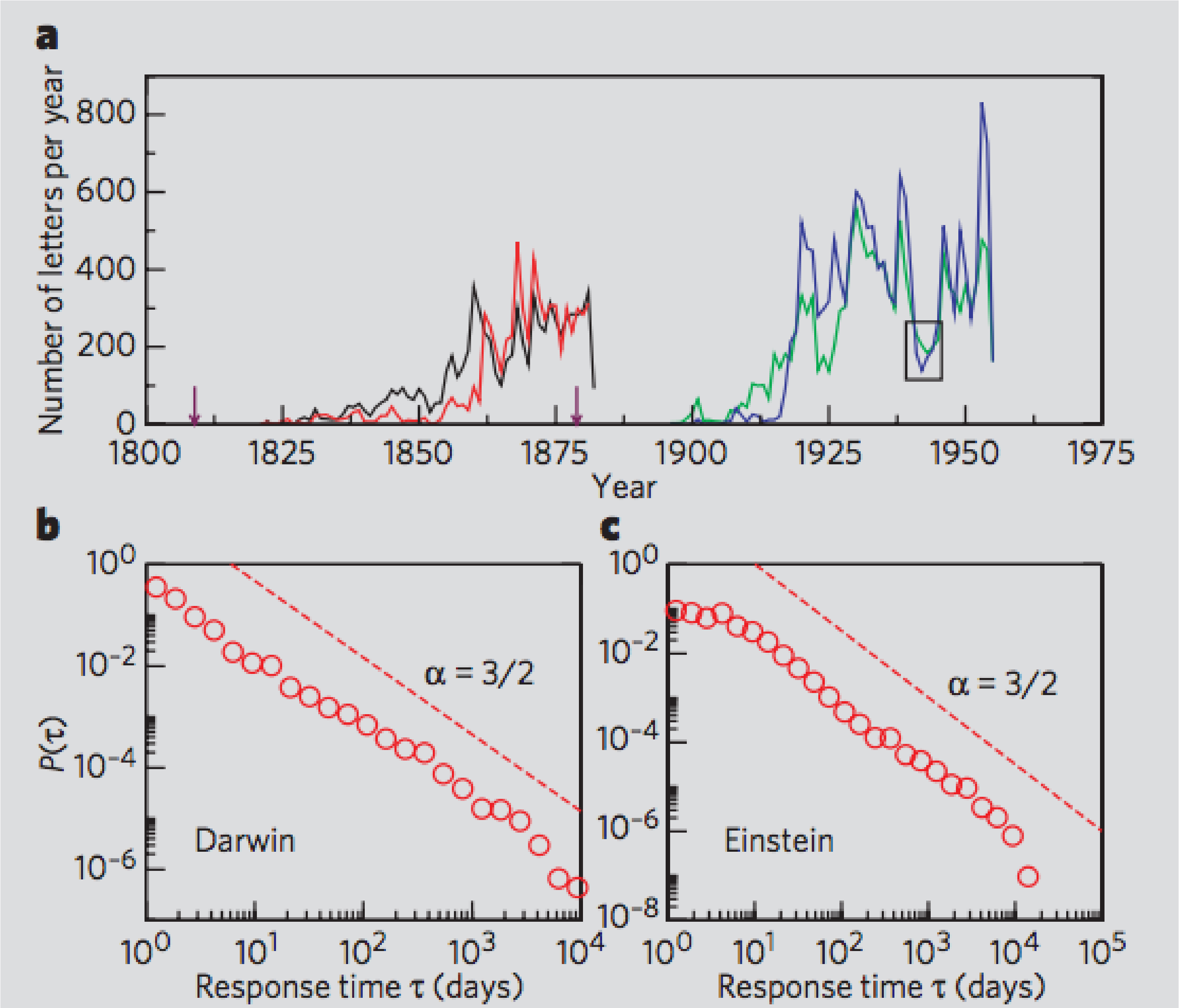
Examples for Poisson process include:

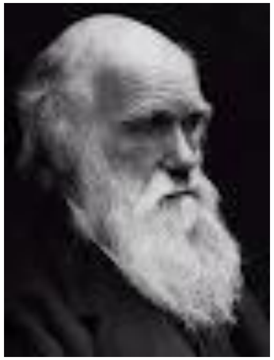
- radioactive decay
- low density road traffic
- light bulbs burning out

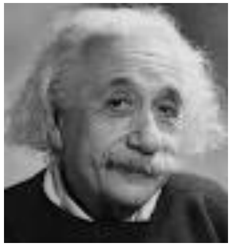
But MANY processes are not

- Human communication pattern
- Solar flares
- Earthquakes
- Price changes above threshold
- neuron firing
- etc.

Correspondence







Oliveira and Barabási, Nature (2005)

Barabási model

Priority list

Model:

a) Pick the task with highest priority with probability *p* or randomly one with 1- *p* and execute it
b) generate a new task with a random priority

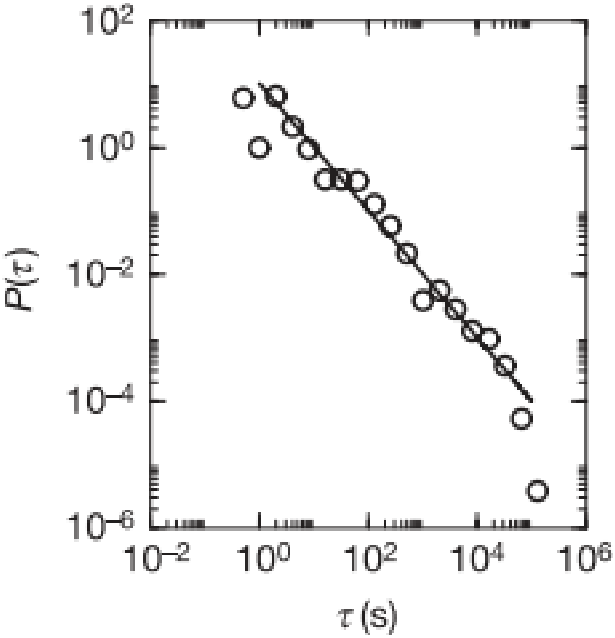
waiting ≠ inter-event time

Two types of tasks…

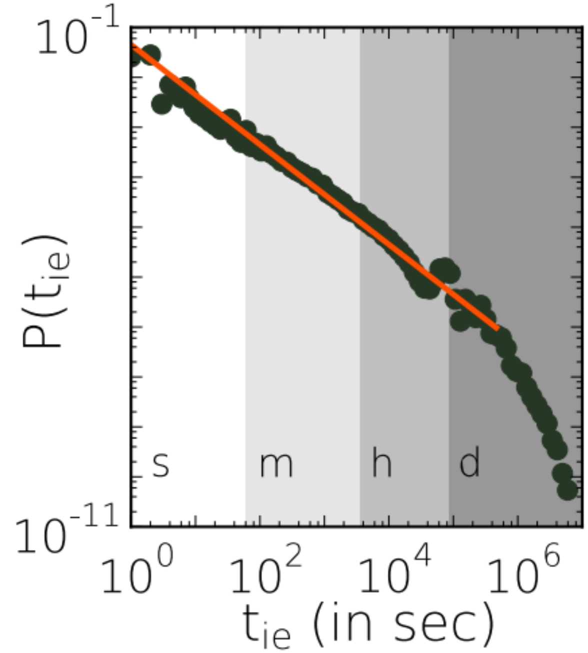
task	priority
1	x_1
2	x_2
3	x_3
4	x_4
5	x_5
6	x_6
…	…
L	x_L

A.-L. Barabási, Nature (2005)

Info-communication data



Email Barabási 2005



Karsai et al. 2012

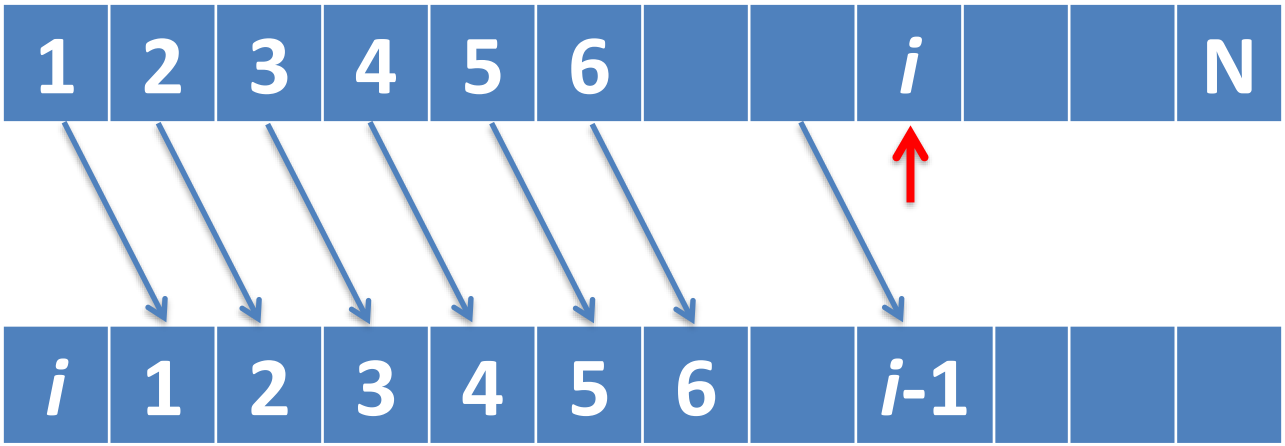
Power law valid only within a day

α close to 1

Measured are the inter-event times, not the waiting times.

Priority arranged list model

Bursty behavior consists of excited (active) and normal periods. I.e., there is some persistence.



a) Choose task *i* with probability $w_i \sim i^{-\sigma}$
b) Put task *i* to position 1
c) Shift all tasks 1→ *i-1* by one to the right

There are 2 kinds of tasks: A and B → inter-event times

Markovian property

ABAABBAAAAB…
AABAABBAAAB…
BAABAABAAAB…

$$\underline{P}_A(t+1) = \underline{P}_A(t)A$$

Master equation

$$A = \begin{pmatrix} w_1 & \sum_{i=2}^{N} w_i & 0 & \cdots & 0 \\ w_2 & w_1 & \sum_{i=3}^{N} w_i & 0 & \vdots \\ w_3 & 0 & w_1 + w_2 & \sum_{i=4}^{N} w_i & 0 \\ \vdots & \vdots & 0 & \ddots & w_N \\ w_N & 0 & \cdots & 0 & \sum_{i=1}^{N-1} w_i \end{pmatrix}$$

Karsai et al 2012

Solution for the inter-event time

There is only one A task

$q_n(t)$ is the prob. that A is at n at time t for the first time.

$$q_n(0) = (1 - w_1)\delta_{n,2}$$

$$q_n(t+1) = \begin{cases} P_1 q_2(t) & \text{if } n = 2, \\ P_{n-1} q_n(t) + (1 - P_{n-1}) q_{n-1}(t) & \text{if } n > 2 \end{cases}$$

$$P_n = \sum_{k=1}^{n} w_i$$

Vajna et al. 2012

$Q(t) = \sum_{n=2}^{\infty} q_n(t)$ is the prob. not to recur until t and the prob. of first recurrence at t is

$$Q(t) - Q(t+1)$$

Introducing discrete Laplace transformation

$\Gamma(\lambda,t)=\sum_{k=2}^{\infty} q_k(t)e^{-k\lambda}$ and using the form of w_i we get

$$\Gamma(\lambda,t+1)=\Gamma(\lambda,t)-\sum_{k=2}^{\infty} k^{-\sigma+1}q_k(t)e^{-k\lambda}+\sum_{k=2}^{\infty}(k+1)^{-\sigma+1}q_k(t)e^{-(k+1)\lambda}$$

Applying $(-\frac{\partial}{\partial\lambda})^{\sigma-1}$ to both sides

$$\left(-\frac{\partial}{\partial\lambda}\right)^{\sigma-1}[\Gamma(\lambda,t+1)-\Gamma(\lambda,t)]=\left(e^{-\lambda}-1\right)\Gamma(\lambda,t)$$

$$\left(-\frac{\partial}{\partial\lambda}\right)^{\sigma-1}\frac{\partial\Gamma(\lambda,t)}{\partial t}=(e^{-\lambda}-1)\Gamma(\lambda,t)$$

which has the scaling solution for small λ:

$$\Gamma(\lambda,t)=t^{\frac{1}{\sigma}-1}\phi\left(\lambda t^{\frac{1}{\sigma}}\right)$$

leading to

$$P(t_{ie})\sim t^{-(2-\frac{1}{\sigma})}$$

Renewal processes

Stochastic process of instantaneous events separated by random IID inter-event times τ, distributed according to $P_{ie}(\tau)$.

Generalization of the Poisson process, where $P_{ie}(\tau)=\exp(-a\tau)/a$

Autocorrelation function

$$\mathscr{A}(t) = \frac{\mathbb{E}[X(0)X(t)] - \langle \mathbb{E}[X(t)]\rangle_t^2}{\langle \mathbb{E}[X(t)]\rangle_t - \langle \mathbb{E}[X(t)]\rangle_t^2}.$$

where $X(t)$ is the indicator of the event

Scaling law for renewal processes

For a renewal process with power law tailed inter-event time $P_{ie}(t) \sim t^{-\beta}$ we have $\mathscr{A}(t) \sim t^{-\alpha}$ with a scaling law:

$$\alpha + \beta = 2 \qquad (*)$$

The Laplace transform of $A(t)$ can be expressed by that of $P_{ie}(t)$

$$g(\lambda) = \sum_{t=0}^{\infty} e^{-\lambda t} A(t) = \sum_{t=0}^{\infty} e^{-\lambda t} P(t = \{\tau_0, \tau_1, \tau_2, \ldots, \tau_m, \ldots\}) =$$

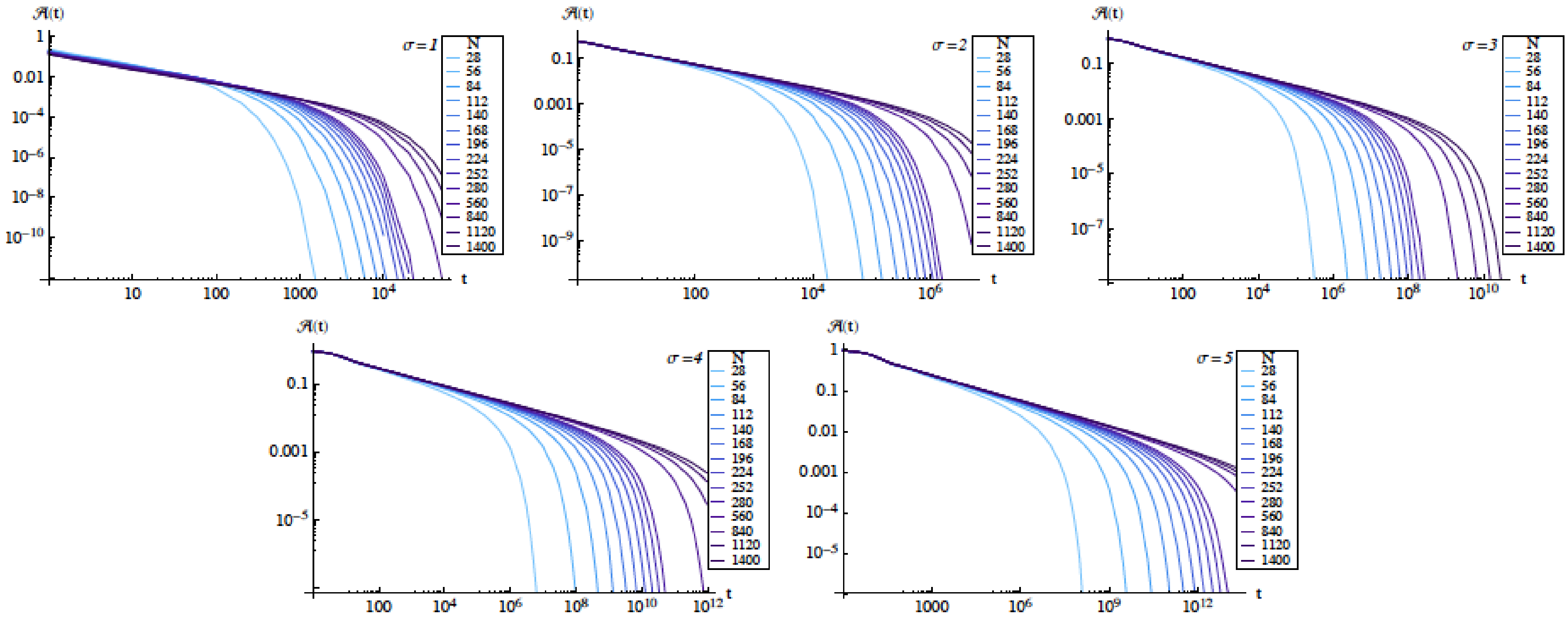
$$= \sum_{m=0}^{\infty} \mathrm{E}(e^{-\lambda \tau_m}) = \sum_{m=0}^{\infty} \mathrm{E}(e^{-\lambda(\tau + \tau' + \tau'' + \ldots + \tau^{(m)})}) = \sum_{m=0}^{\infty} \left[\mathrm{E}(e^{-\lambda \tau})\right]^m =$$

$$= \left(1 - \mathrm{E}(e^{-\lambda \tau})\right)^{-1}$$

From which (*) follows via Tauber theorems

Sz. Vajn et al. 2012

Empirical results for the priority list model



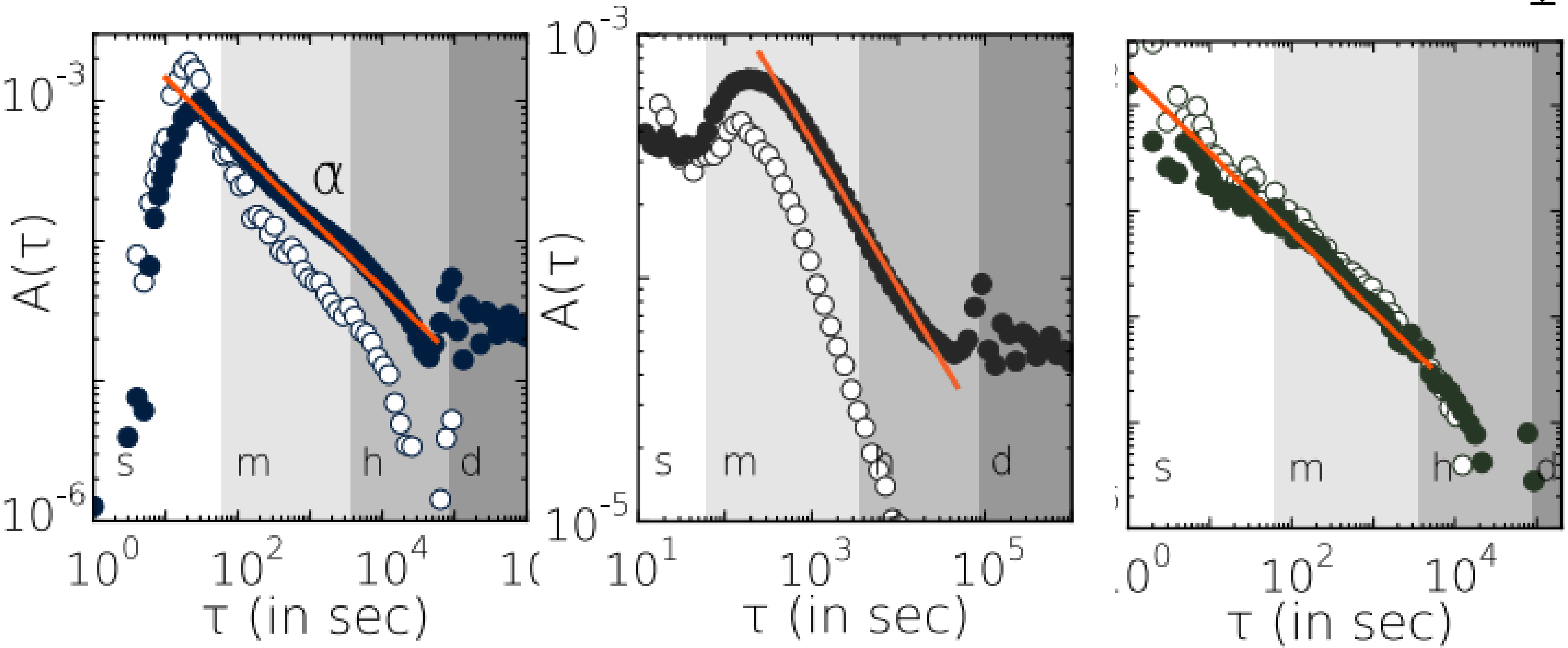
α=2-β= 1/σ perfectly verified

Empirical autocorrelation function

$$\mathscr{A}(t) = \frac{\mathbb{E}[X(0)X(t)] - \langle \mathbb{E}[X(t)] \rangle_t^2}{\langle \mathbb{E}[X(t)] \rangle_t - \langle \mathbb{E}[X(t)] \rangle_t^2}.$$

A(*t*)-s also decay as power law!

Karsai et al. Sci. Rep. 2012



phone calls

SMS

Email

Renewal processes?

	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\alpha+\beta}$
Phone calls	0.5	0.7	1.2
SMS	0.6	0.7	1.3
Emails	0.7	1.0	1.7

$\alpha + \beta < 2$, i.e. the process is NOT a renewal one, there is dependence between the events.

Renewal processes?

	α	β	α'	$\alpha'+\beta$
Phone calls	0.5	0.7	1.1	1.8
SMS	0.6	0.7	1.2	1.8
Emails	0.7	1.0	0.8	1.8

$\alpha + \beta < 2$, i.e. the process is NOT a renewal one, there is dependence between the events.

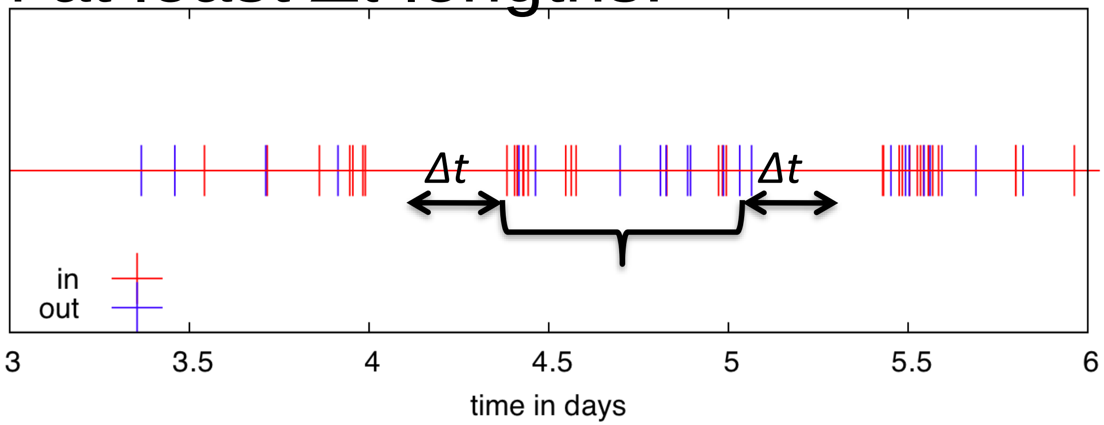
If shouffled, $\alpha'+\beta$ ~ 1.8, much closer to the scaling law holds

Measuring dependence

Bursty behavior means that there are high activity periods separated by low activity ones

We define a „bursty period" relative to a window Δt:

A bursty period (or train of bursts) is a sequence of events separated from the rest by empty periods of at least Δt lengths.

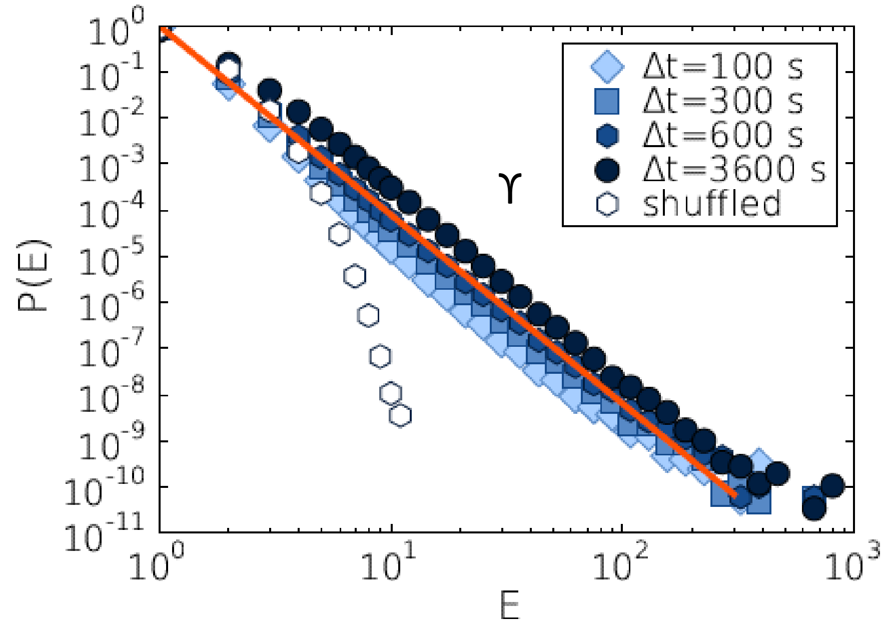


Calculate the distribution P(*E*) of the number *E* of events within the trains

The $P(E)$ distribution measures dependence

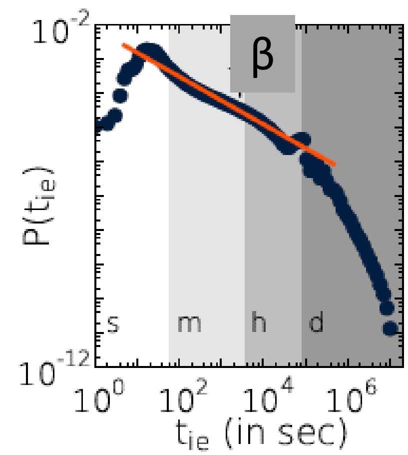
For any independent inter-event time distribution $P(E)$ decays exponentially:

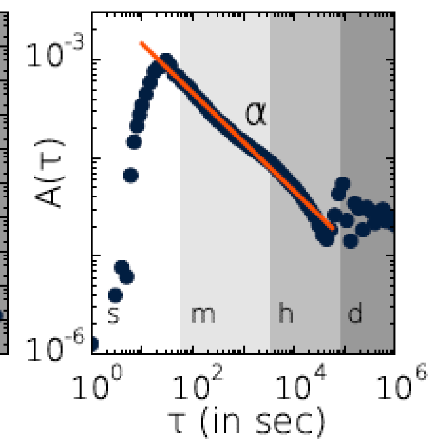
$$P(E=n)\sim\left(\int_0^{\Delta t}P(\tau)\,d\tau\right)^n = e^{-an}$$



Empirical results show the presence of intrisic correlations. We find: $P(E)\sim E^{-\gamma}$

	α	**β**	**γ**
Phone calls	0.5	0.7	4.1
SMS	0.6	0.7	3.9
Emails	0.7	1.0	2.5





Relation to memory

Power law $P(E)$ shows: the process is non-Markovian:

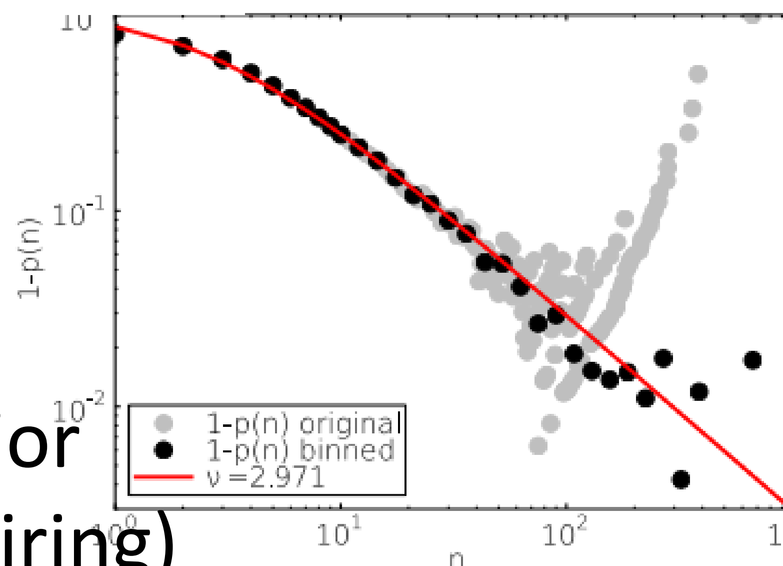
Memory

$p(n)$ is the prob. that a bursty train, which has already the length n will get longer.

$$p(n) = \frac{\sum_{E=n+1}^{\infty} P(E)}{\sum_{E=n}^{\infty} P(E)}$$

Assuming perfect power law behavior for $P(E)$ leads

$$p(n) = \left(\frac{n}{n+1}\right)^{\gamma - 1}$$



Persistence in human behavior

(Also: earthquakes, neuron firing)

M. Karsai et al. Sci. Rep. 2012

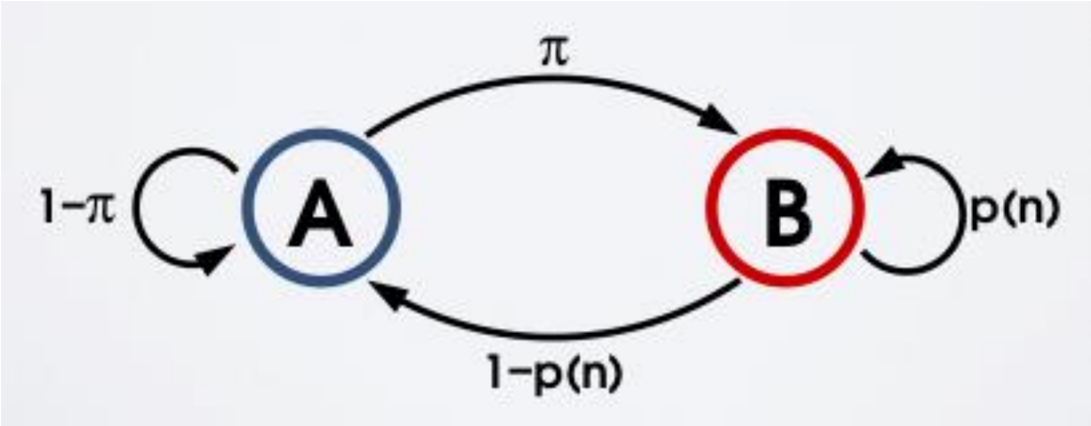
Modeling

Can we construct a model for all features?

Two state model:

A normal state - it performs independent events with relatively long inter-event times

B excited state - it performs correlated bursty events with relatively short inter-event times



Inter-event times are generated by reinforcement processes

- the longer an entity waits after an event, the larger the probability it will wait longer (see Stehlé, et al. PRE (2010)).
- Here the inter-event time of an event depends on the actual state
- Different reinforcement functions for state A and B

$$f_{A,B}(t_{ie}) = \left(\frac{t_{ie}}{t_{ie}+1}\right)^{\mu_{A,B}}$$

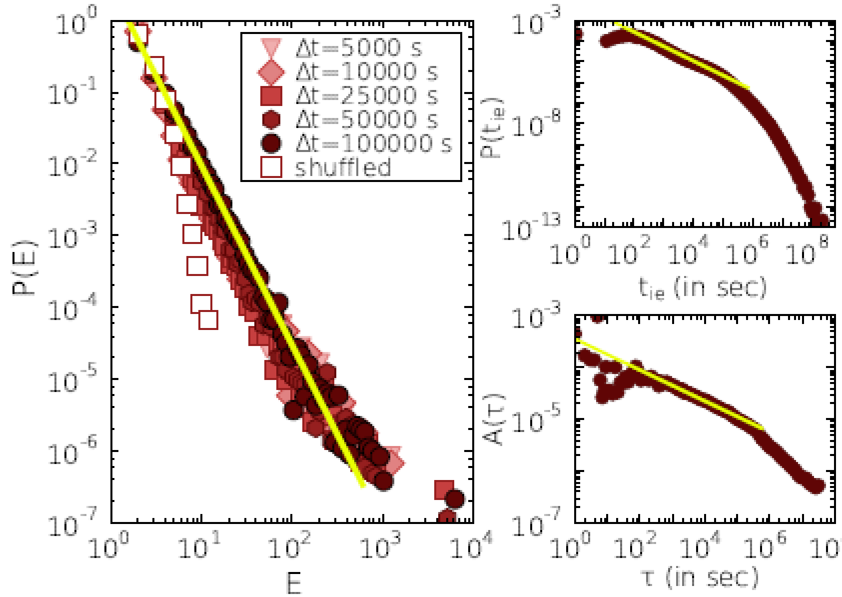
If $\mu_A << \mu_B$ long bursty trains develop. $\beta=\mu+1$

Other systems

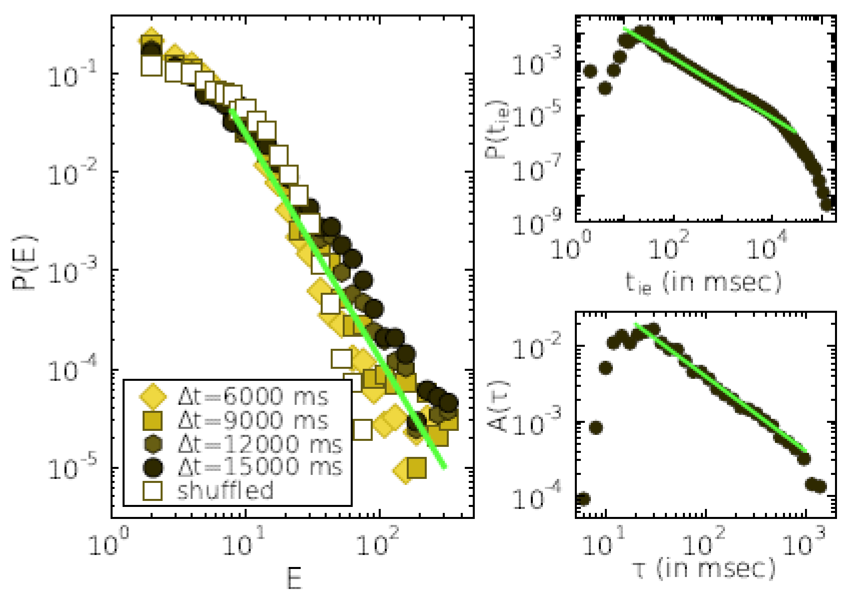
Very different systems show similar behavior (weak universality, exponents are different)

Is there a common mechanism behind?

Threshold phenomena



(a)Japanese earthquake sequence



(b)Neuron firing sequence

Measuring burstiness

How to measure burstiness?

$$B \equiv \frac{\sigma_\tau - m_\tau}{\sigma_\tau + m_\tau}$$

with m_τ being the mean and σ_τ the variance of the empirical inter-event distribution.

B = 0 for Poisson distribution

B = -1 for delta distribution

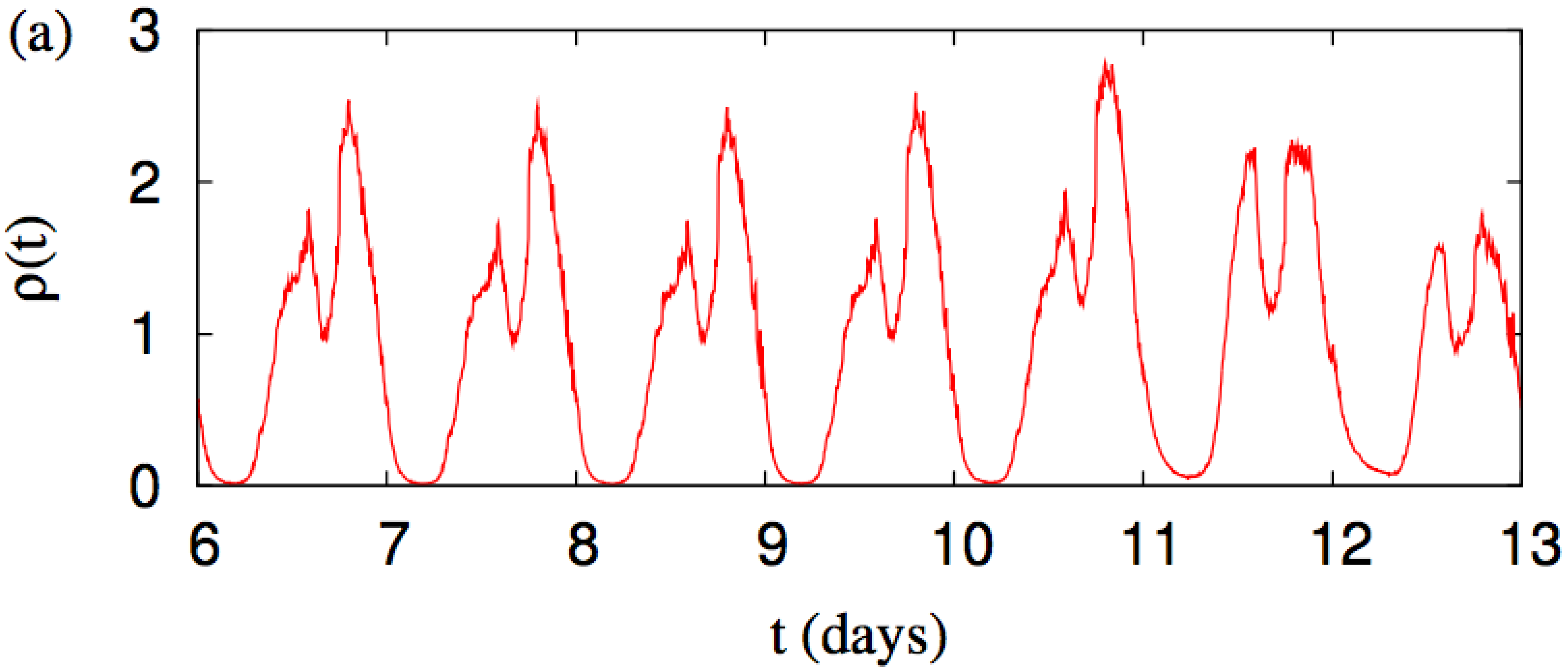
B = 1 if the second moment diverges

The larger B > 0 the more bursty

Goh and Barabási, 2011

Circadian pattern

Activity pattern for mobile phone data



Jo et al. NJP 2012

How is this related to the observed burstiness?

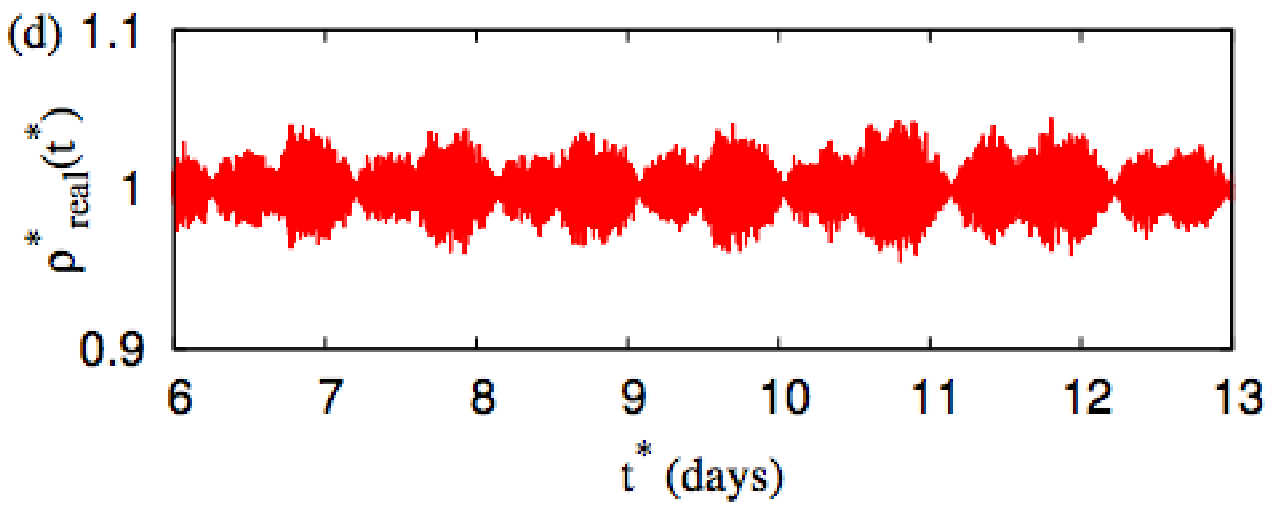
Is there "intrinsic" burstiness? Deseasoning

Rescaling time

$$dt^* = \frac{c(t)}{C_T}\,dt = \rho(t)dt$$

where *c*(*t*) is the event density at *t*, C_T is the average density over period *T*

$$\rho^*(t^*)dt^* = \rho(t)dt$$

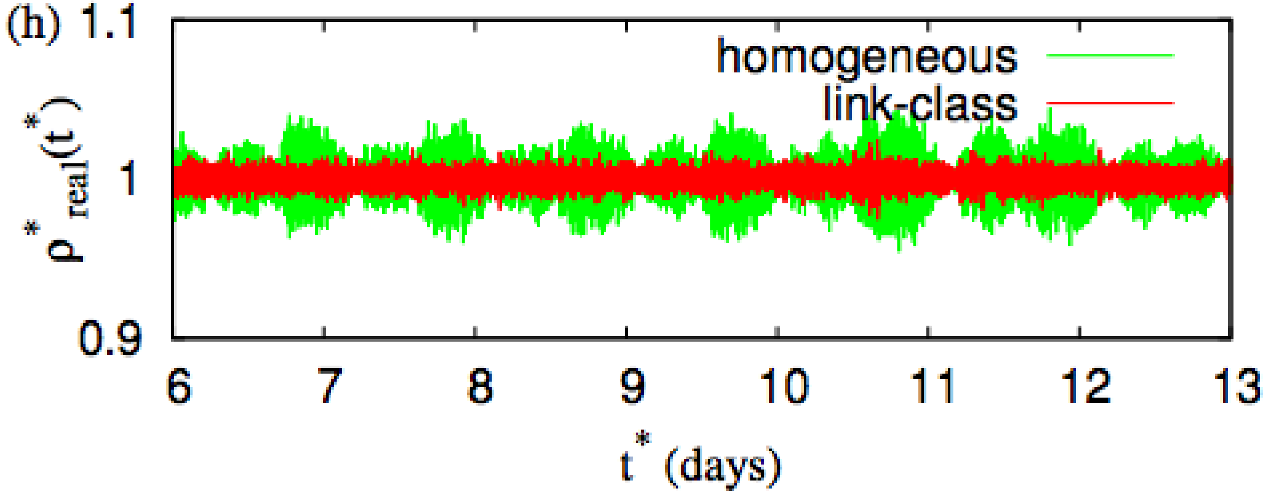


There is still circadian pattern observed!

Activity classes

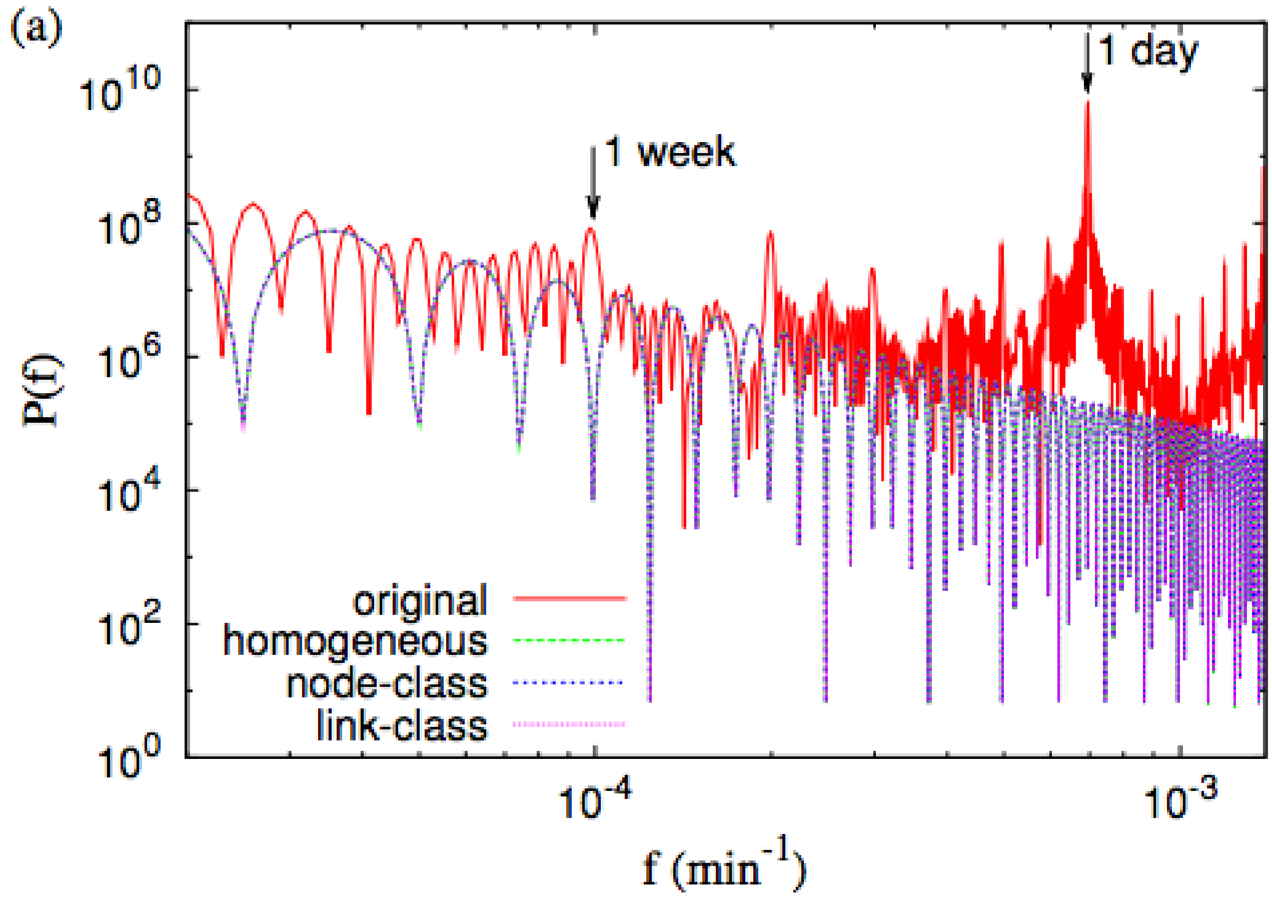
When averaging over the whole sample, we mix a very inhomogeneous sample

Introduce classes of activities and do the rescaling on them.

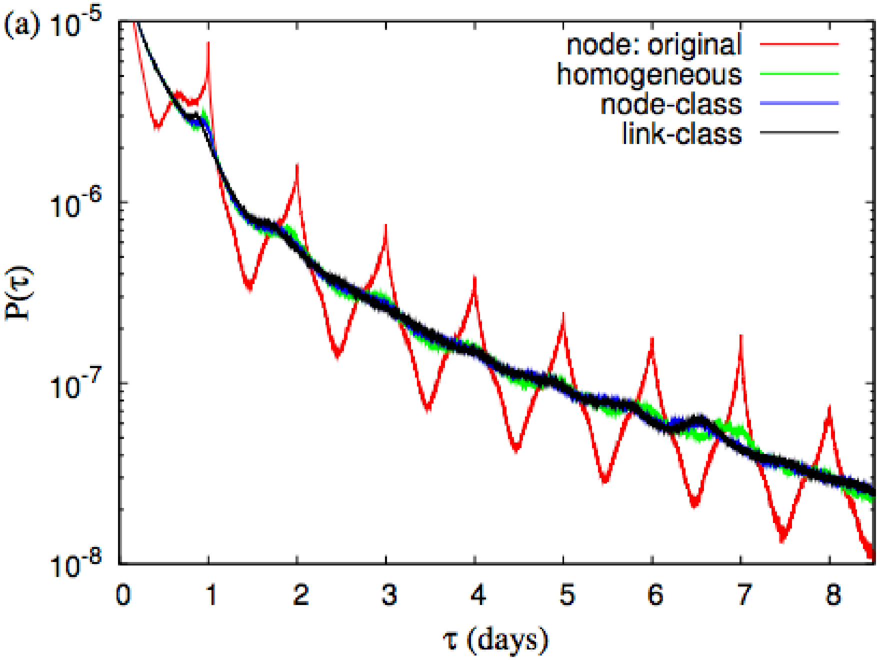


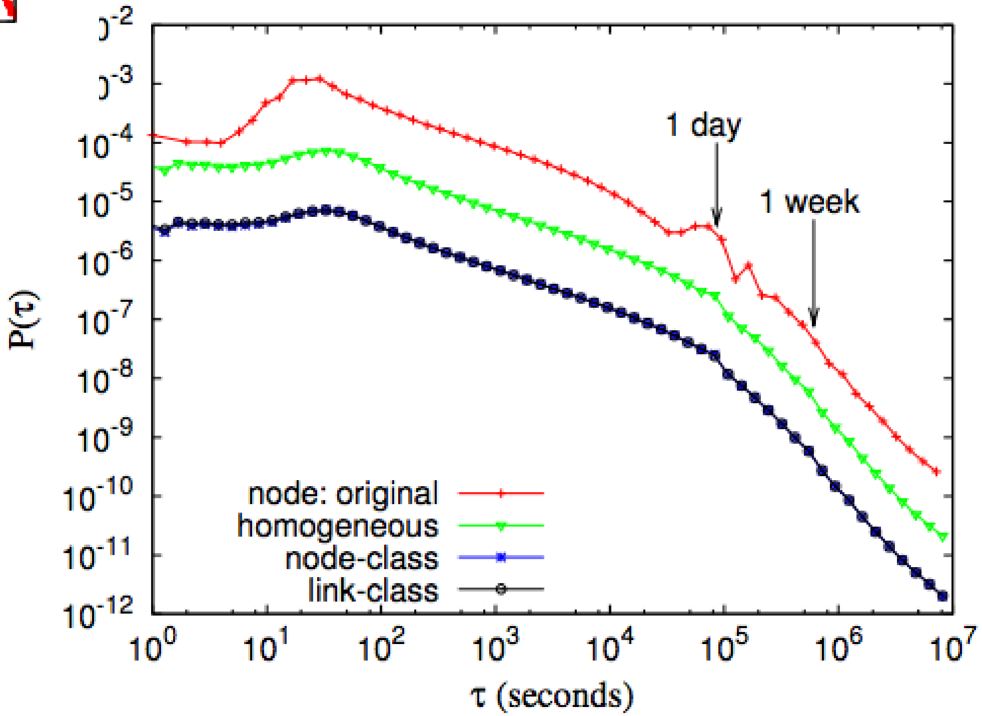
Pattern is almost entirely removed!

Power spectrum

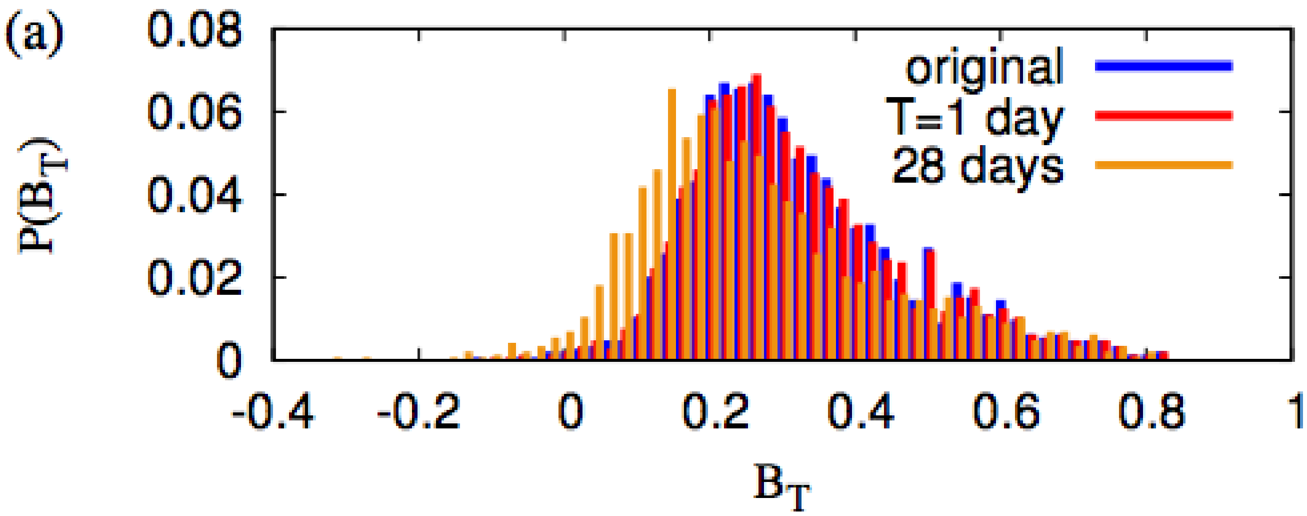


Rescaled inter-event times



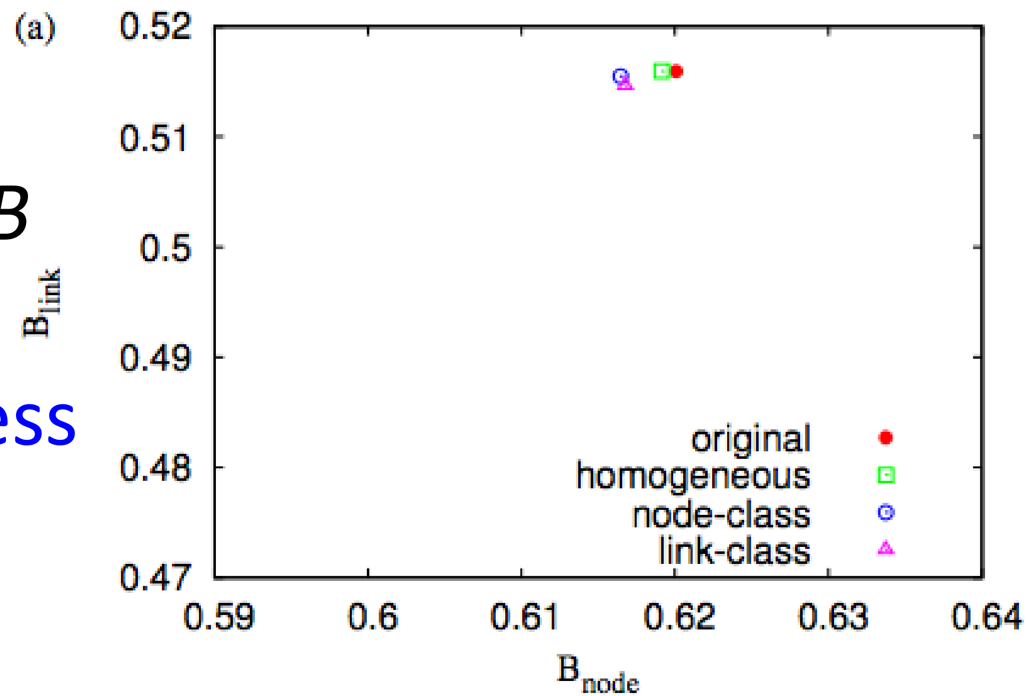


Burstiness



Distribution of *B* for an activity class of users

De-seasoning:
-little effect on the total *B*
-scaling improves
There is intrinsic burstiness related to human task execution



Burstiness and conflicts in Wikpedia edits

WP is a collaborative, free, WEB-based, multilingual encyclopedia written by volunteers from all around the world.

Fully recorded: Every single edit, discussion, interaction – the full history of a (special) society.

Yasseri et al. 2012

What is the mechanism of arriving at a consensus in a collaborative environment?

Mostly constructive activity, sometimes conflicts, edit wars.

Identification of conflicts

English WP 2011 most controversial articles

$$M = E \times \sum_{\substack{(i,j)\in\text{reverters}\\ \backslash\text{max pair}}} \min\left[N_i^d, N_j^r\right]$$

E: total # of reverting editors
(larger army → worse war)

$N_i^d(N_j^r)$: # of reverts of reverte**d**(**r**)
(more mature editors → worse war

\ max pair: Avoid conflict btw 2 persons

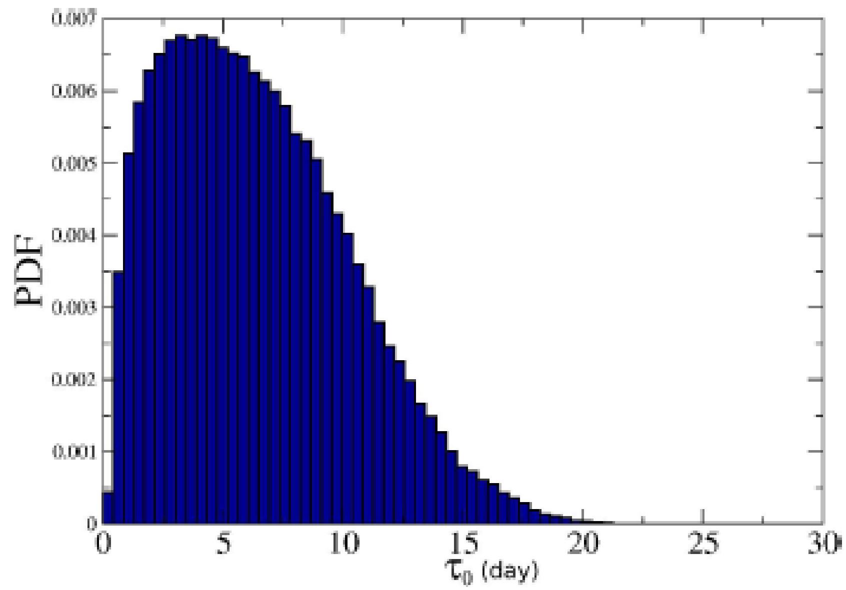
George W. Bush
Anarchism
Muhammad
Circumcision
Race and intelligence
Global warming
United States
Barack Obama
Jesus
Christianity
Michael Jackson

Edit frequency

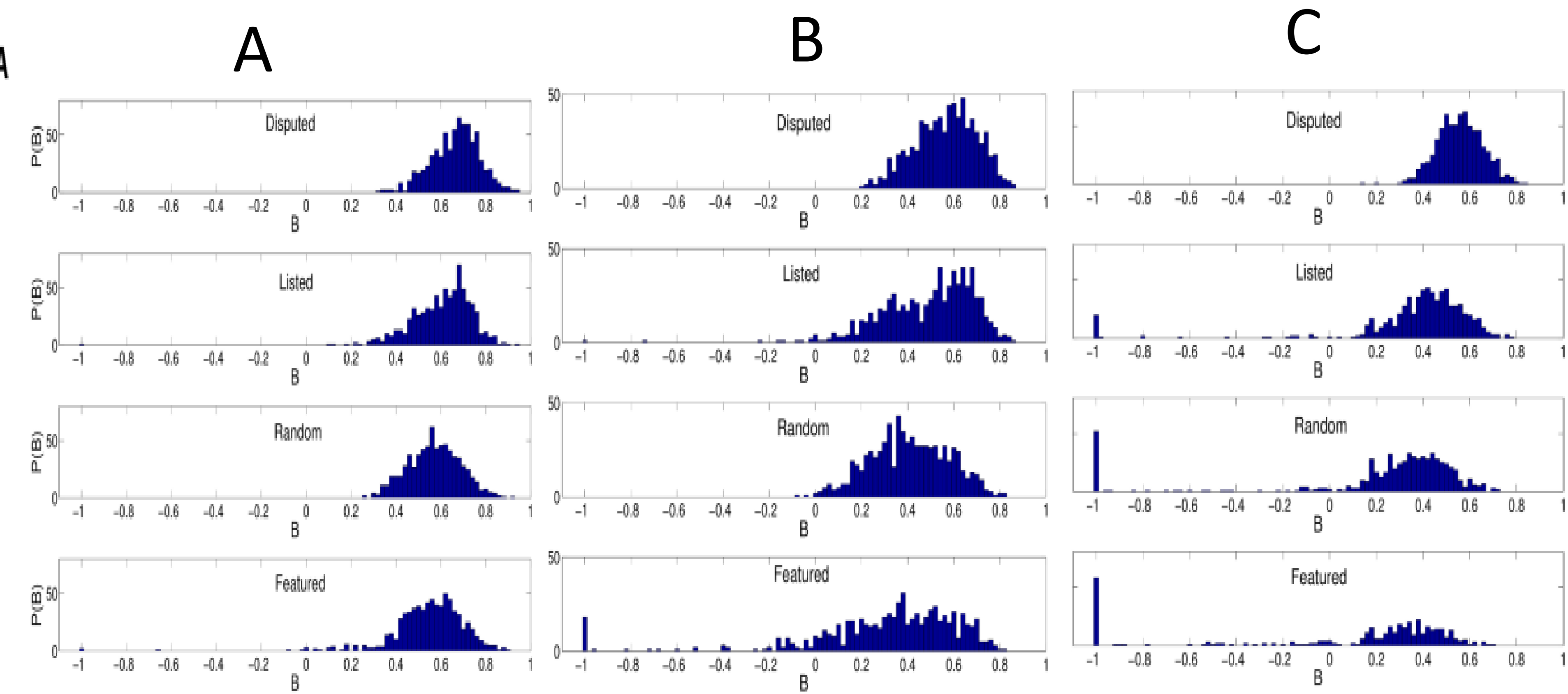
Edit patterns:

t := timestamp assigned to each edit
τ := time-interval between two successive edits on an article
τ_0 := $\langle \tau \rangle$ over entire life of the article

Average edit interval

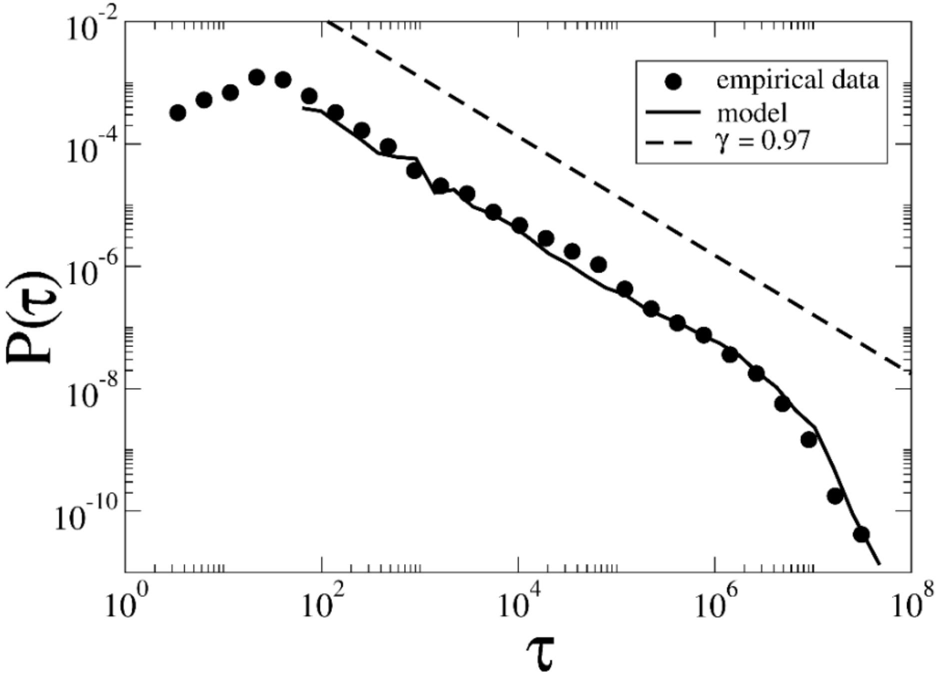


Burstiness

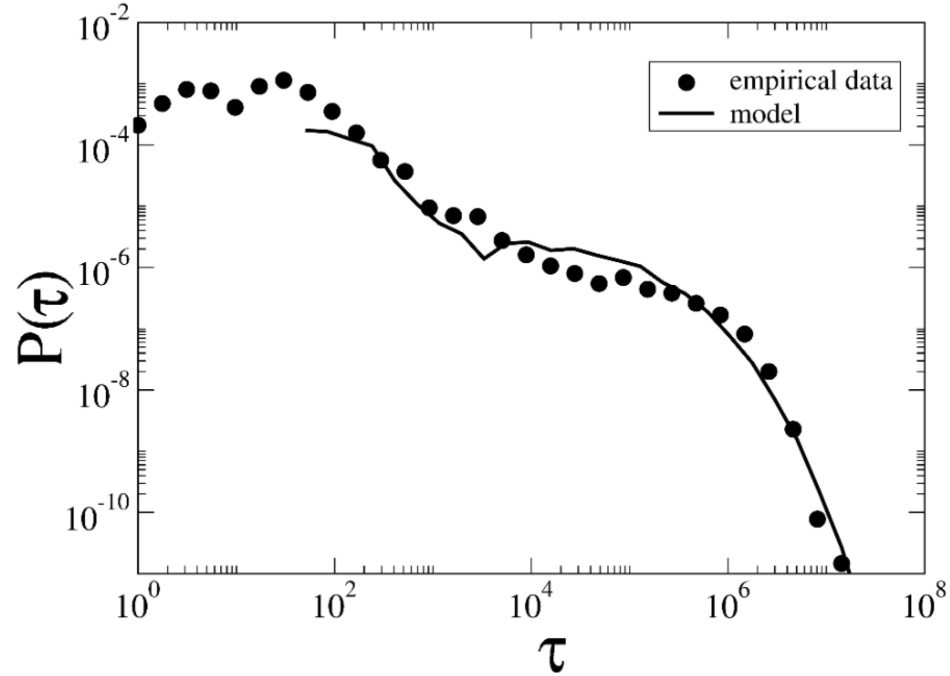


All edits Reverts Mutual reverts

Inter-event time distribution



Conflict articles



Normal articles

DYNAMICS OF SPREADING IN A COMMUNICATION NETWORK

Spreading phenomena in networks

- epidemics (bio- and computer)
- social contagion (rumors, information, opinion, innovation)

Difference in the transmission:

Epidemics (in simple cases, like influensa) – binary

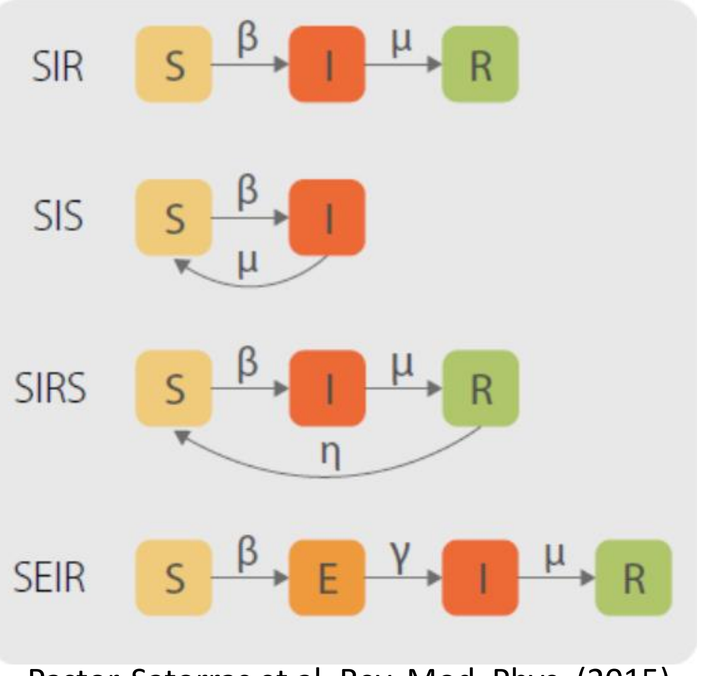
Social contagion: complex (multiple nodes participate)

Epidemic spreading theory

Epidemic spreading among individuals

Different states – compartments:

- **S**usceptible
- **I**nfected
- **R**ecovered (immune)
- **E**xposed (infected but not yet infecting)



Pastor-Satorras et al. Rev. Mod. Phys. (2015)

Resulting in different models in the spirit of reaction-diffusion processes, e.g., $S + I \rightarrow 2I$ (SI model).

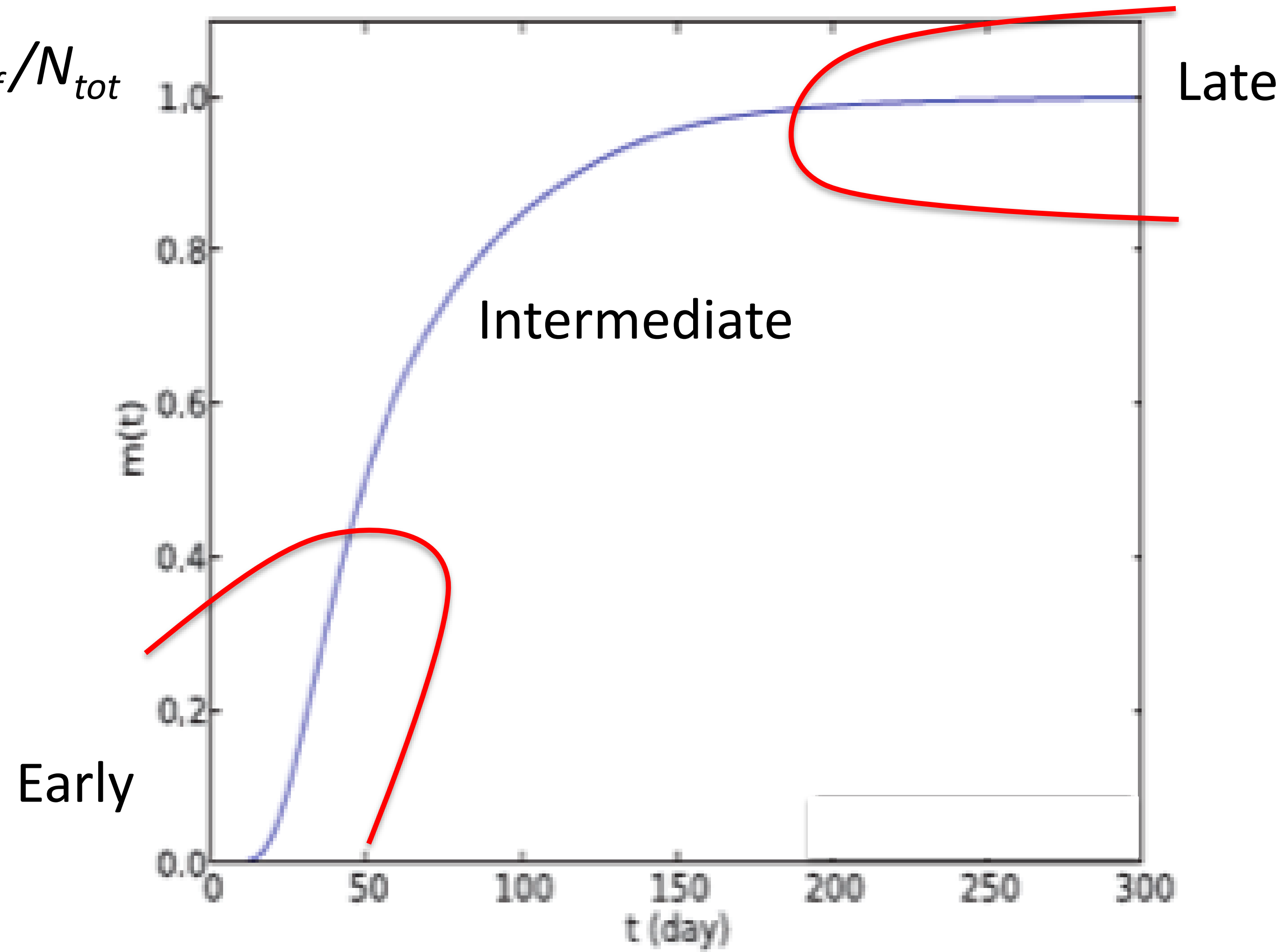
β, μ, η, γ are rates by which the reactions happen. In the simplest case "homogeneous or perfect mixing" is assumed: Everybody can meet everybody with the probability proportional to the concentrations (mean field approximation).

In simple cases solvable, epidemic threshold, relation to percolation.

DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK

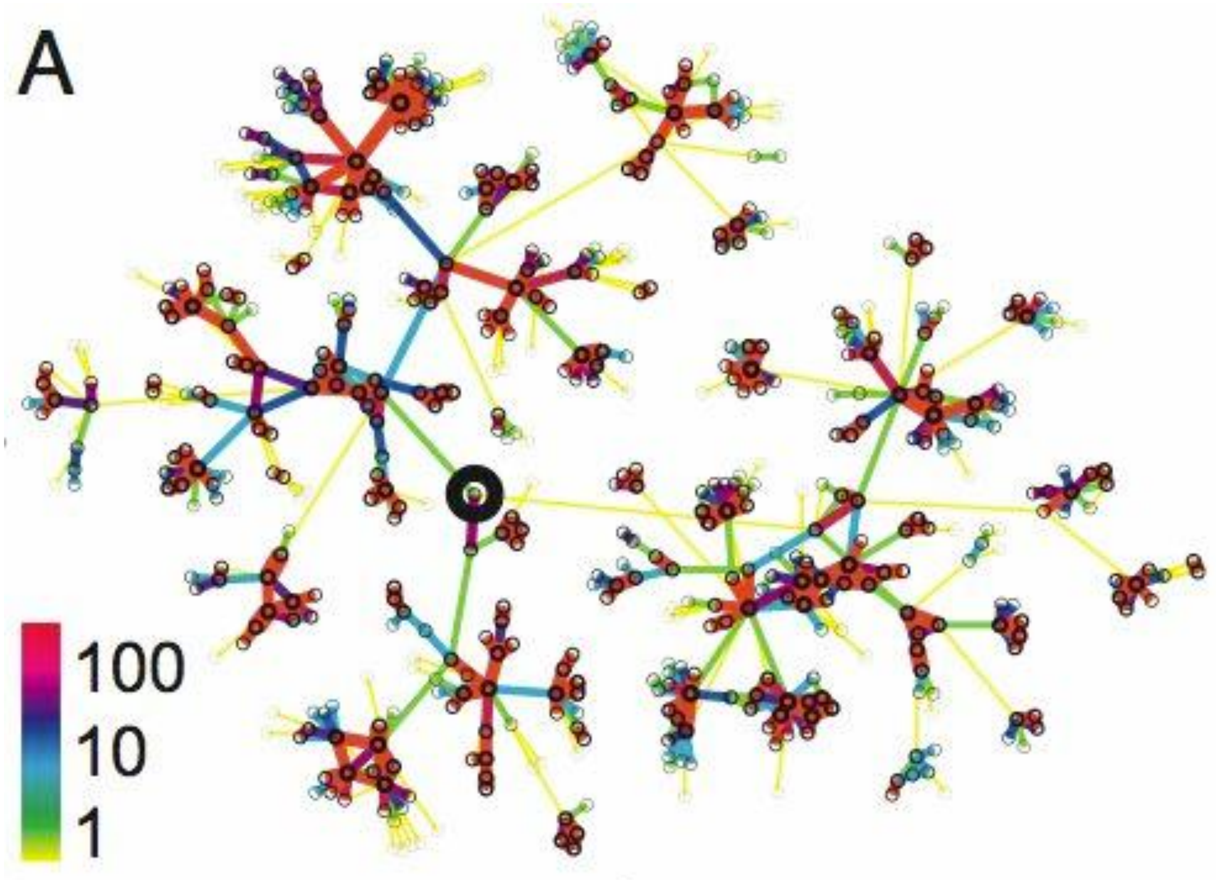
Spreading curve for SI (simplest model)

$m(t)=N_{inf}/N_{tot}$



Important: **speed of spreading**

Aggregate network



Granovetterian structure: Strength of week ties

DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK

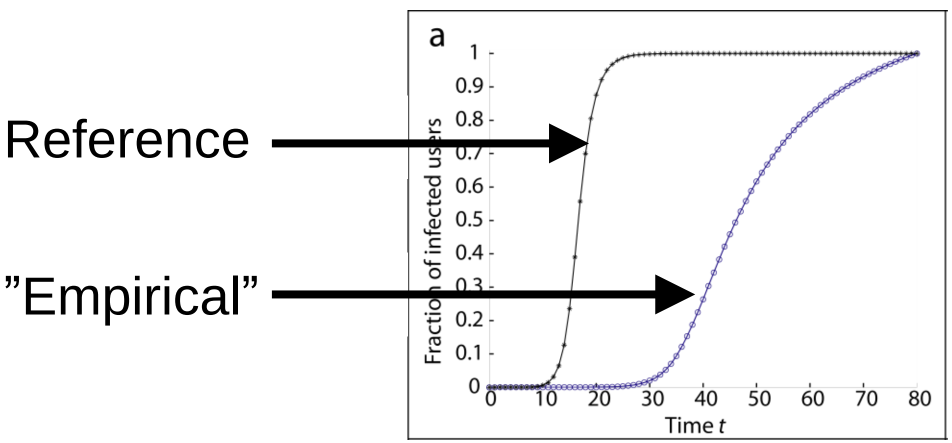
Consequence of the Granovetterian structure:

Strongly wired communities slow down spreading.

Simulation: SI model with hopping rates p_{ij}

(1) Empirical: $p_{ij} \propto w_{ij}$

(2) Reference: $p_{ij} \propto \langle w \rangle$



DYNAMICS OF SPREADING IN A TEMPORAL NETWORK

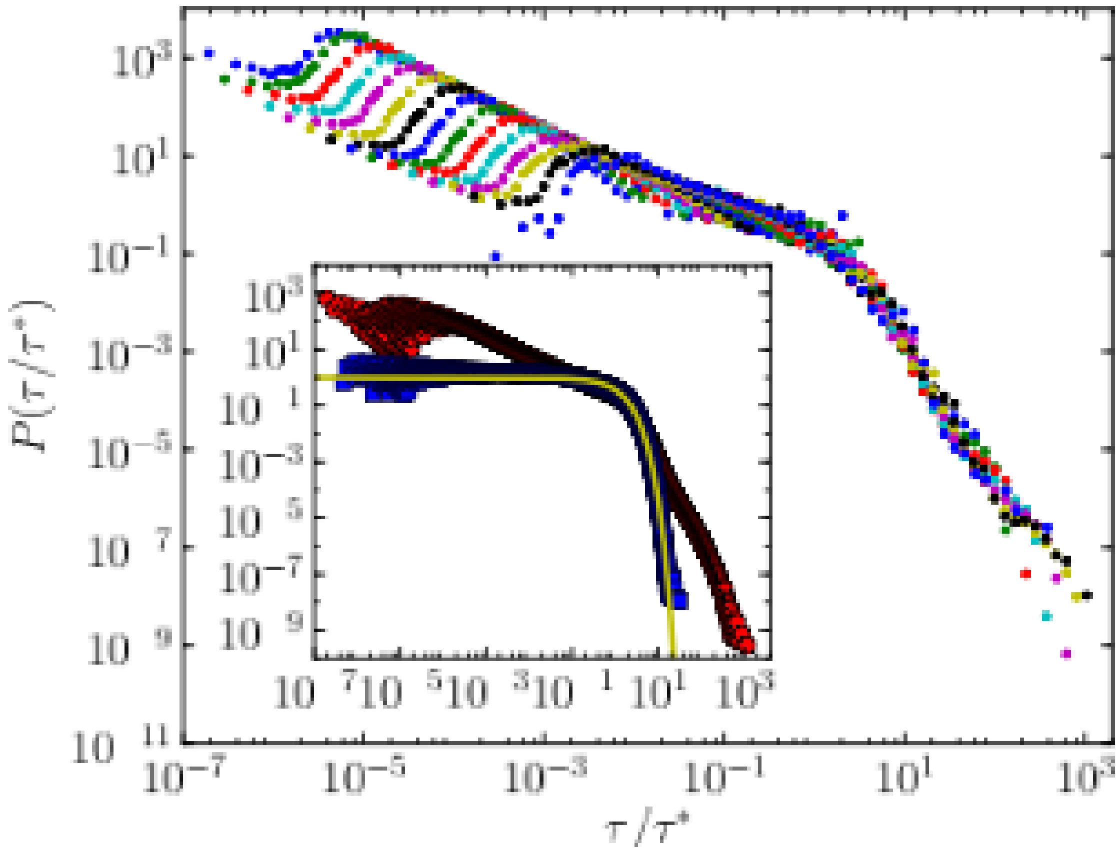
The process is in reality non-Poissonian! Inhomogeneities not only in the topology but also in the temporal behavior (remember the movie!)

Characterizing inhomogeneities
306 million mobile call records of 4.9 million individuals during 4 months with 1s resolution

- Burstiness (fat tailed inter-event time distribution)
- Circadian, weekly pattern
- Triggered activity, temporal motifs

M.Karsai et al. Phys. Rev. E83, 025102 (2011)

DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK



Scaled inter-event time distr.

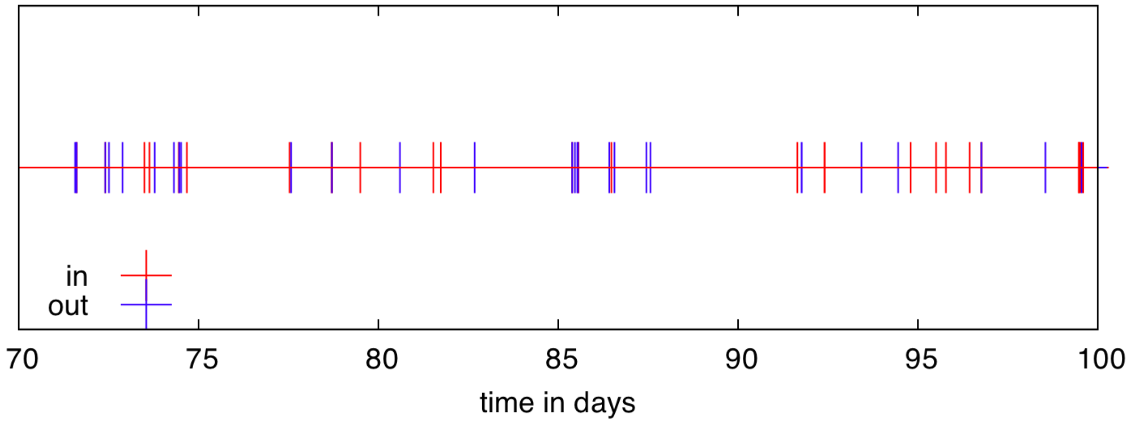
Binned according to weights (here: number of calls)

Calls are non-Poissonian

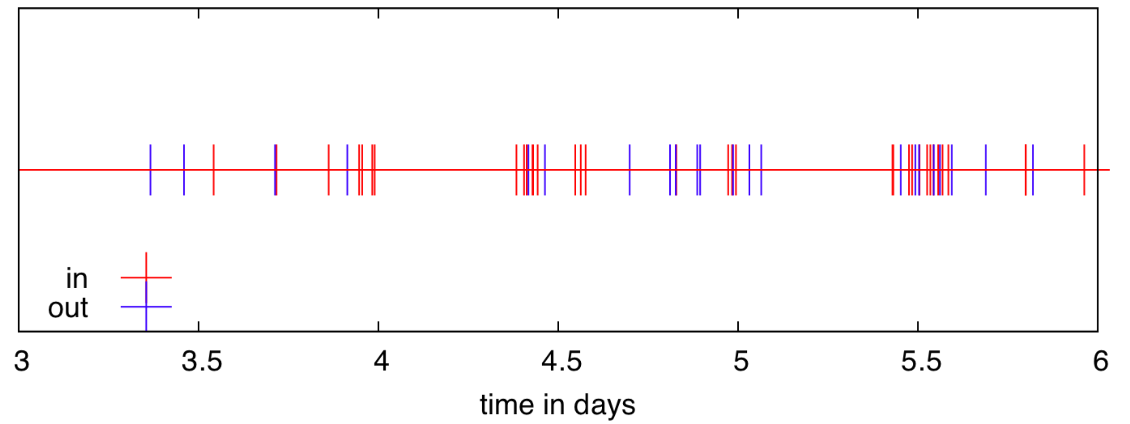
Inset: time shuffled

DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK

Bursty call patterns for individual users



Average user



Busy user

Note the different scales

DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK

Correlations influence spreading speed

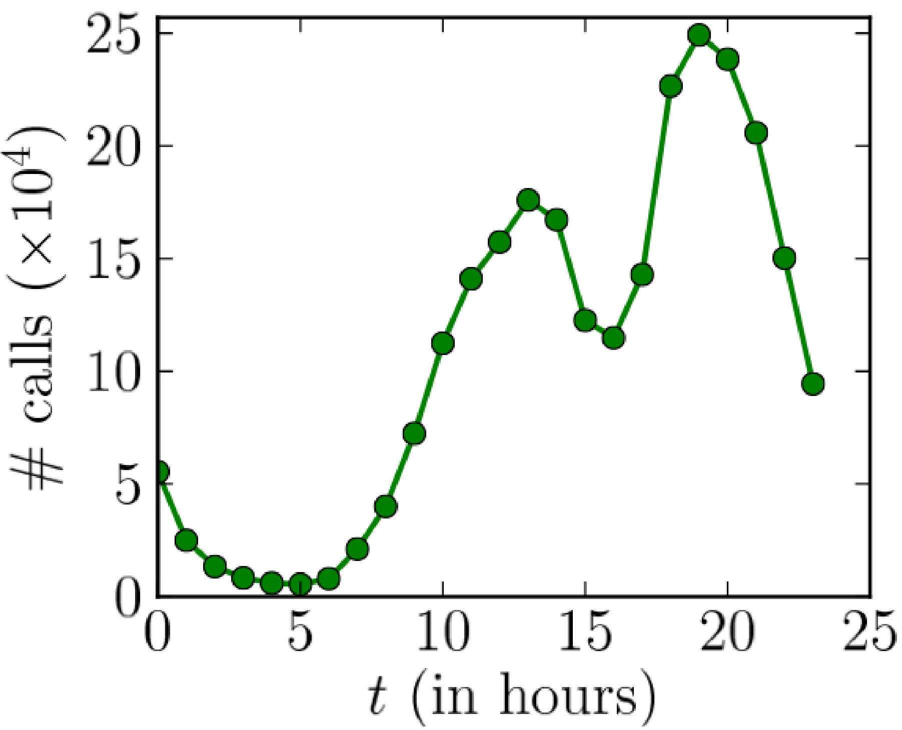
-Topology (community structure)

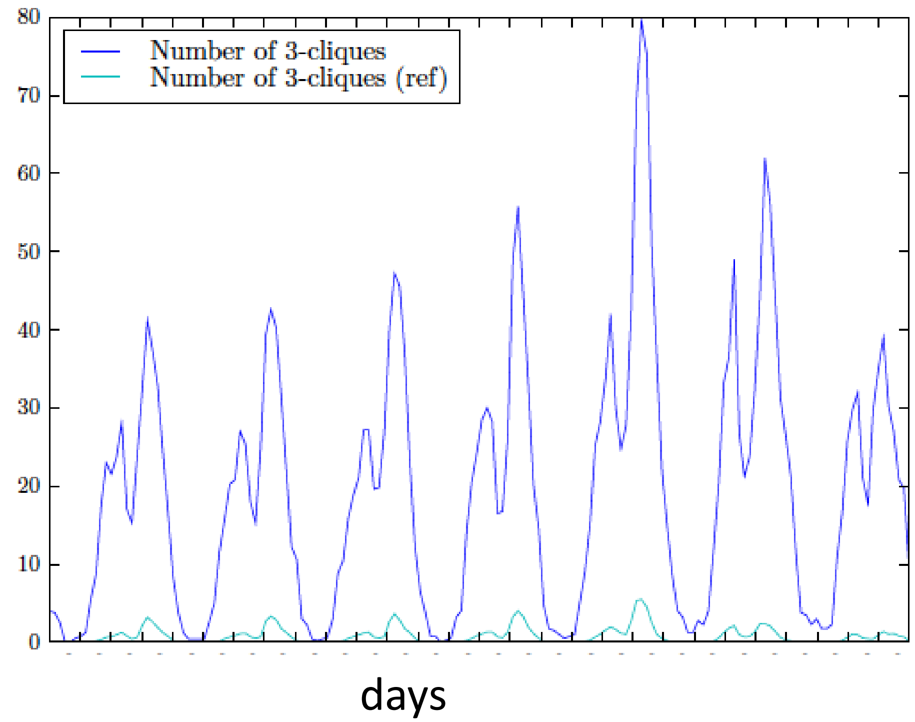
- Weight-topology (Granovetterian structure)

- Daily, weekly patterns

- Bursty dynamics

- Link-link dynamic correlations





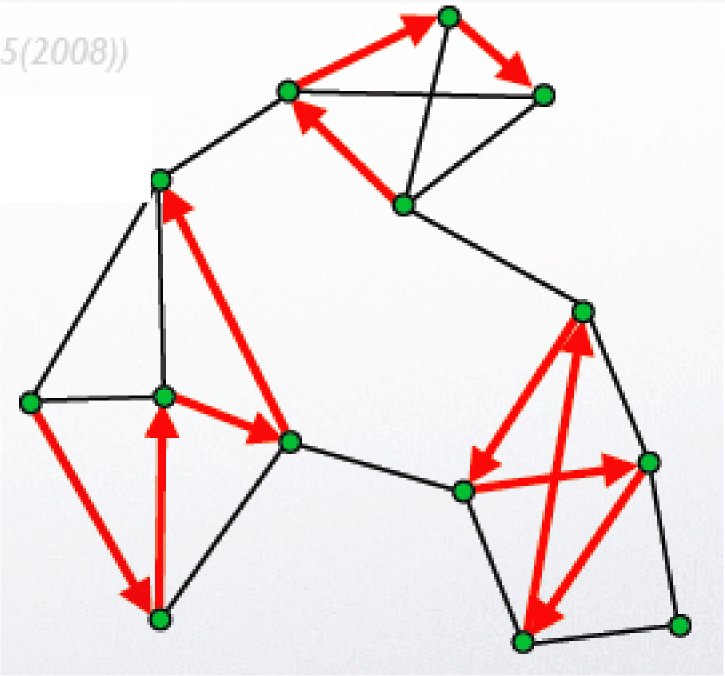
Can be eliminated by inhom. scale transformation

DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK

- Link-link dynamic correlations

Triggered calls, cascades, etc.

Temporal motifs



Experiment: ”Infect” a random node and assume that ”infection” is transmitted by each call (SI).

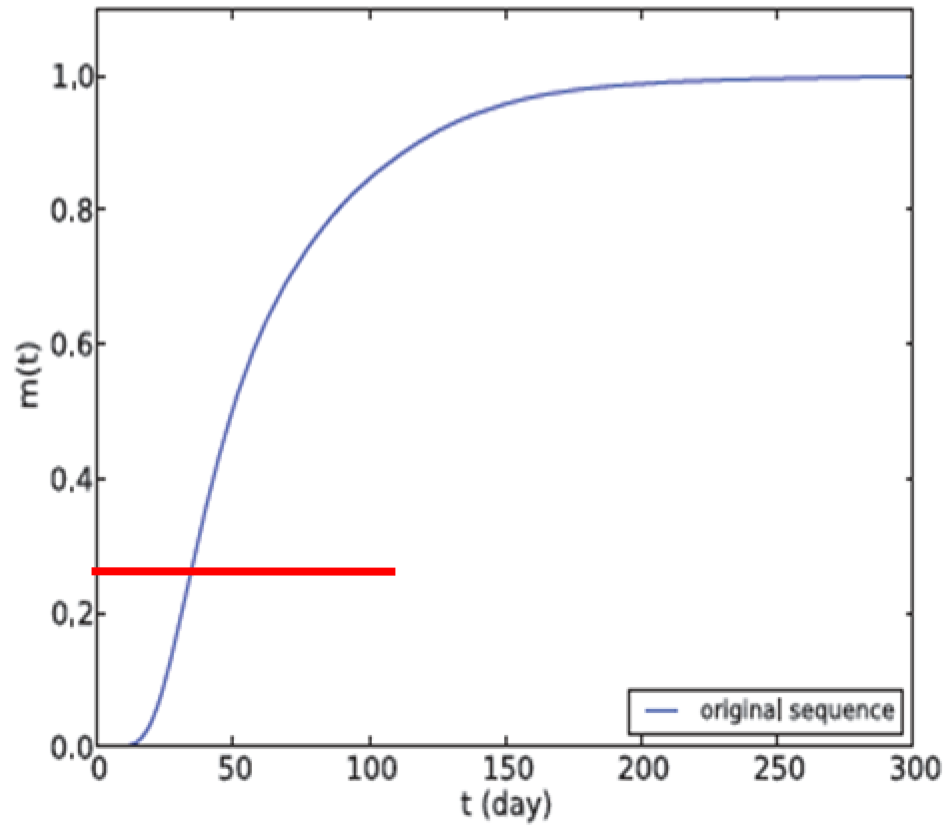
How to identify the effect of the different correlations on spreading?

Introduce different null models by appropriate shuffling of the data.

Correlations: CS: community structure
WT: Weight-topology
BD: Bursty dynamics
LL: Link-link correlations

Original network

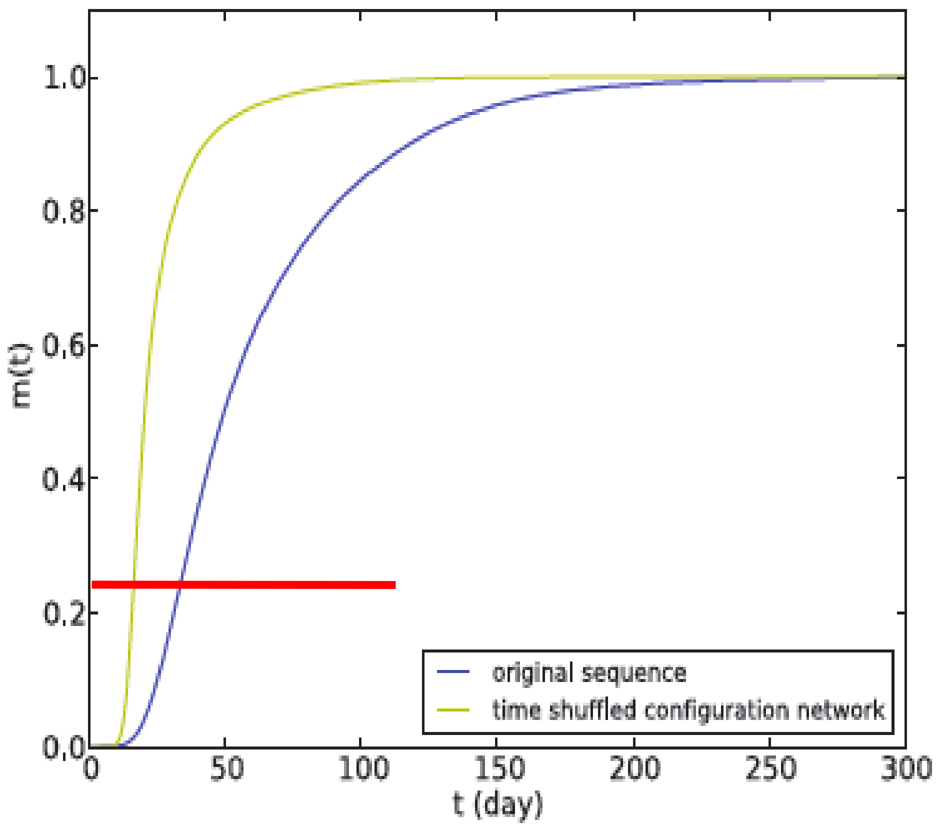
	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33.7



Time shuffled configuration network

- Using configuration model to **destroy community structure**, but keep N, |E| and the network connected
- Shuffle the event times to **destroy bursty dynamics**
- No correlation takes place in the system

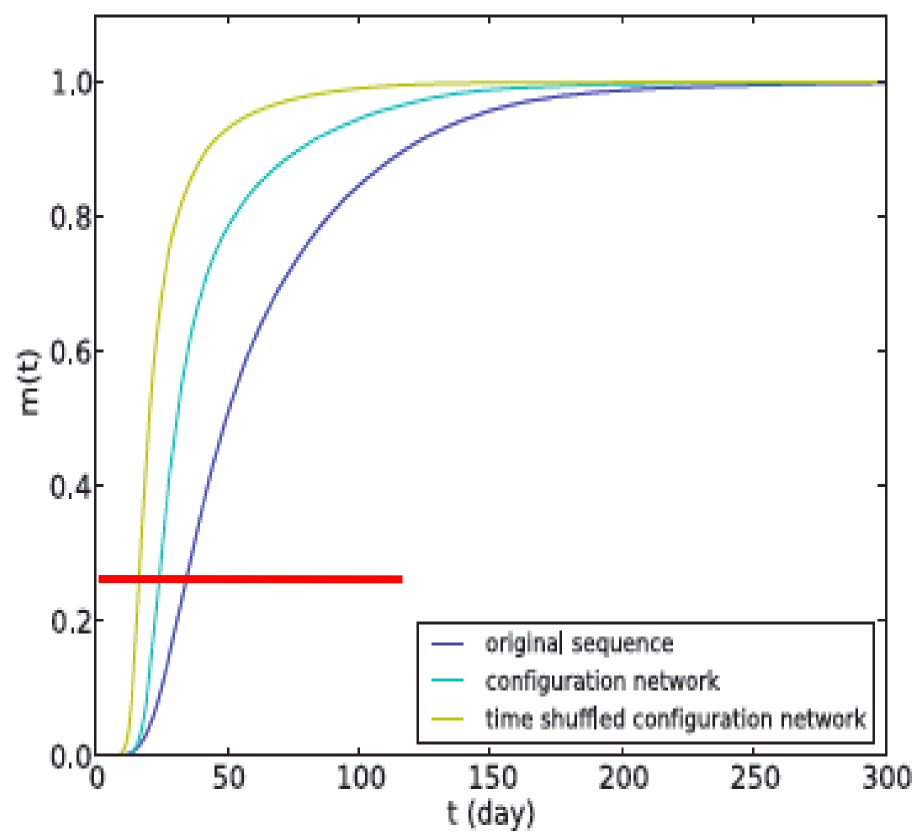
	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33.7
TimeConf	X	X	X	X	16.4



Configuration network

- Using the same configuration method to **destroy community structure**
- Only **bursty dynamical behavior** is kept
- The infection speed is slowed down by bursty dynamics

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33.7
TimeConf	X	X	X	X	16.4
Config.	X	✓	X	X	23.8



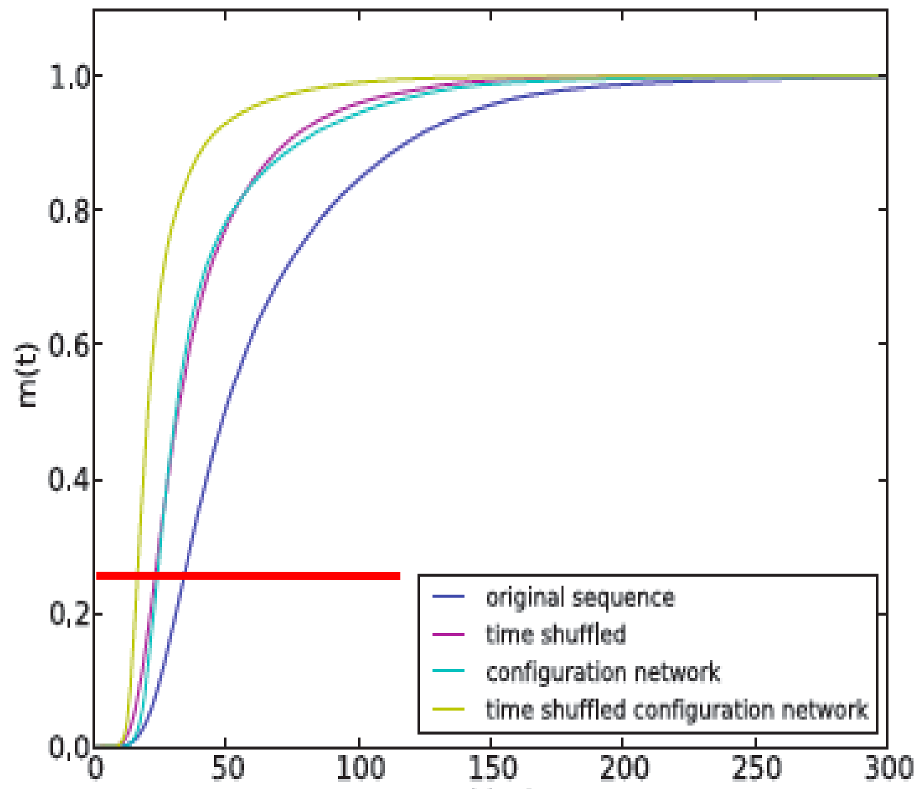
Time shuffled event sequence

- Shuffle the event times but keep **community structure** and **weight-topology** correlations unchanged

- Bursty dynamics and link-link correlations are switched off

- Bursty event clustering is slowing down the dynamics

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33.7
TimeConf	✗	✗	✗	✗	16.4
Config.	✗	✓	✗	✗	23.8
Time	✓	✗	✗	✓	22.9



Time shuffling

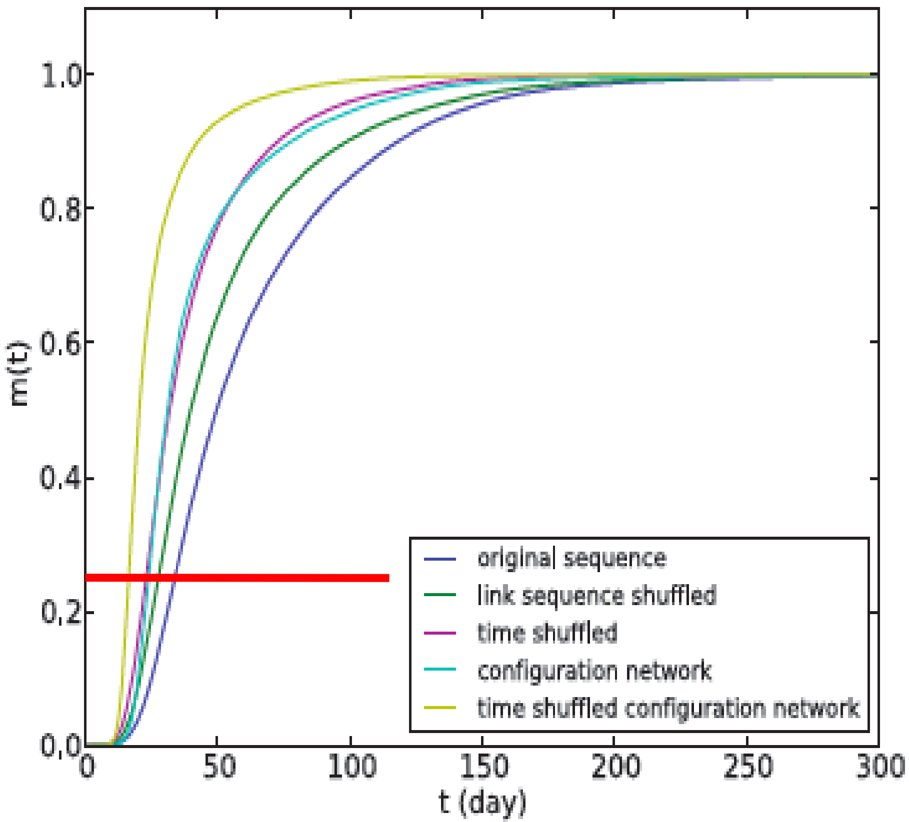
Link1	Link2	Link3...	LinkN
t_{11}	t_{21}	t_{31}...	t_{N1}
t_{12}	t_{22}	t_{32}...	t_{N2}
.	.	.	.
.	.	t_{3n_3}...	.
t_{1n_1}	.		.
	t_{2n_2}		.
			t_{Nn_N}

Destroyes burstiness (and link-link correlations) but keeps weight and daily pattern

Link sequence shuffled event sequence

- Shuffle link call sequences between randomly chosen links
- **Link-link** and **weight-topology** correlations are switched off
- Weight-topology correlations also slow down the dynamics

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33.7
TimeConf	X	X	X	X	16.4
Config.	X	✓	X	X	23.8
Time	✓	X	X	✓	22.9
Link	X	✓	X	✓	27.5

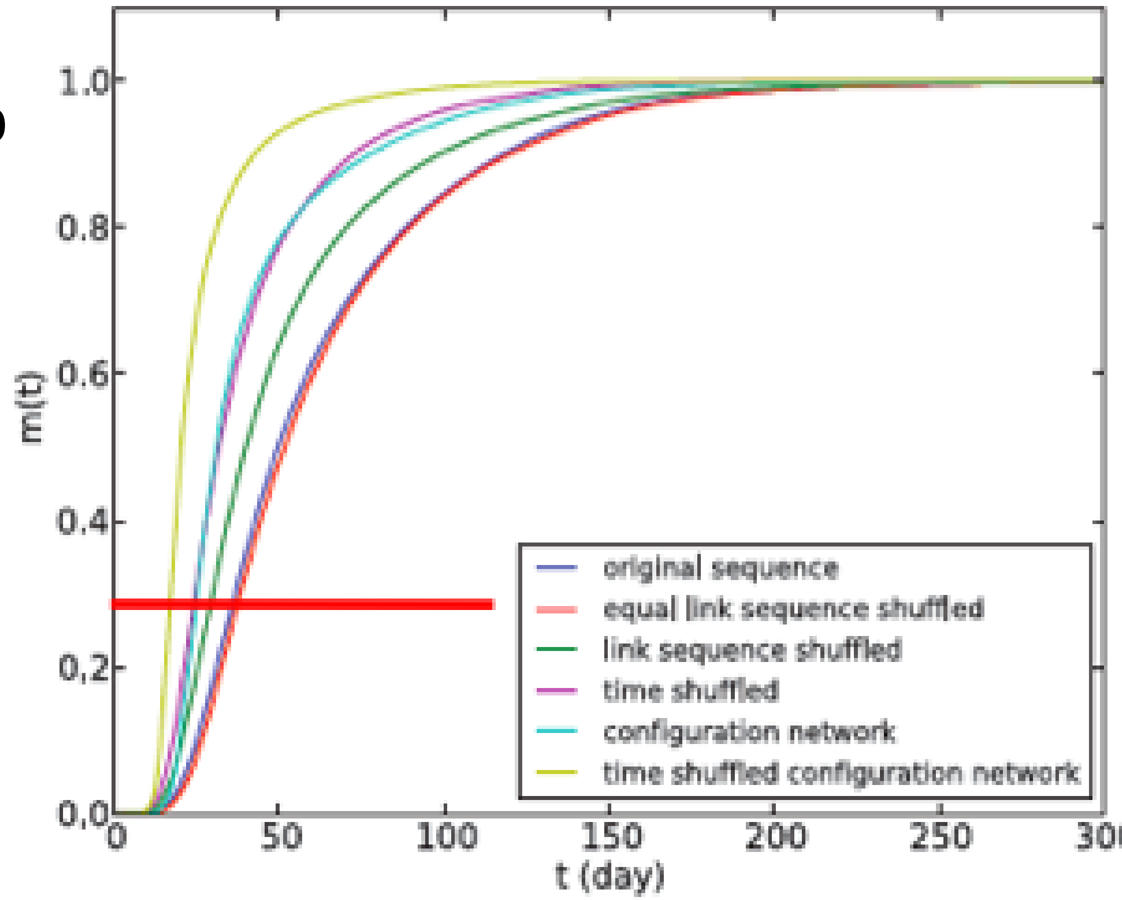


DYNAMICS OF SPREADING IN MOBILE COMMUNICATION NETWORK

Results:

Strong slowing down due to
- topology (communities)
- link-topology correlations
- burstiness

Minor effect:
- circadian etc. patterns
- temporal motifs



Small but slow world

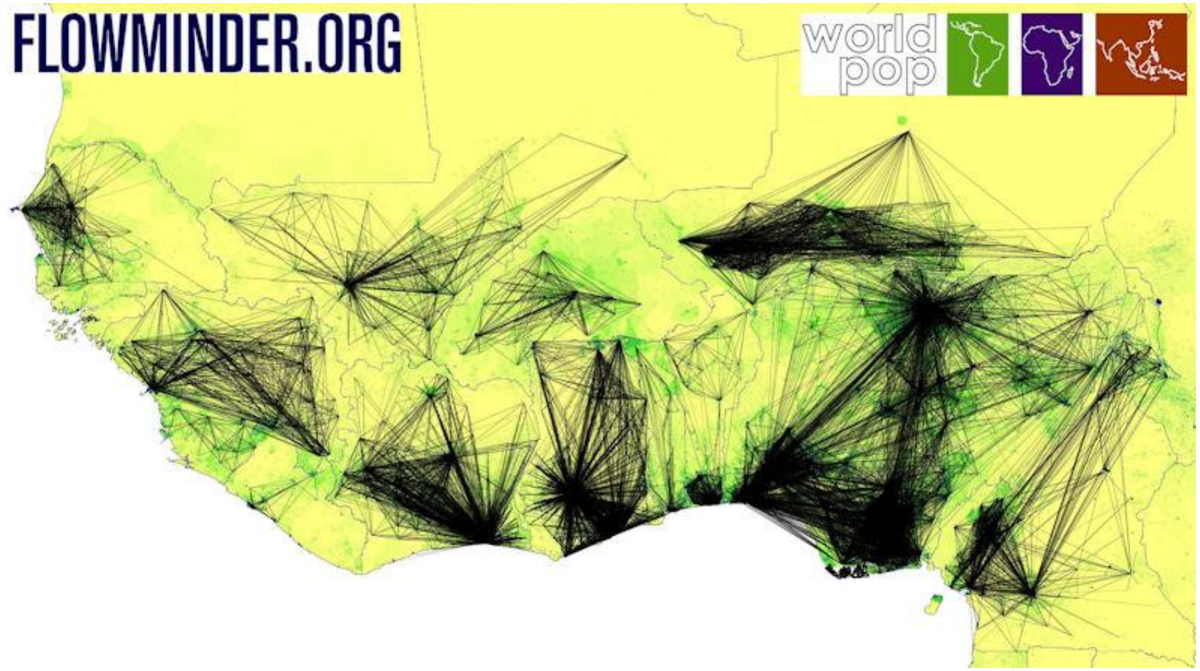
Effect of burstiness

- Empirically: Slowing down
- Analytical model (Infinite complete graph, Cayley tree): speeding up!
- Clean numerical models (ER, BA): Mostly speeding up, but:

- Model calculations for pure power law inter-event time distributions
- CORRELATIONS (in addition to power law inter-event times)
- NON-STATIONARITY!

Karsai et al. Sci. Rep. 2012
Horvath and JK: NJP 2014
Jo et al. PRX 2014

Spreading: Spatiotemporal process



Mobility pattern in West Africa as mapped out by cell phones

Successful efforts to simulate epidemic spreading real time → prediction

Mobility and demographic data should be included.

Vespignani group (Northeastern+ISI Torino)

Social contagion

Similarities and Differences

	Network	Transmission	External influence
Physics systems	Lattice or amorphous	Contact	External field
Biological epidemics	Social	Contact	None
Social contagion	Social	Social pressure	Media

Complex contagion process

(D. J. Daley, D. G. Kendall, Epidemics and rumours. Nature 204, 1118 (1964))

Cascading Phenomena

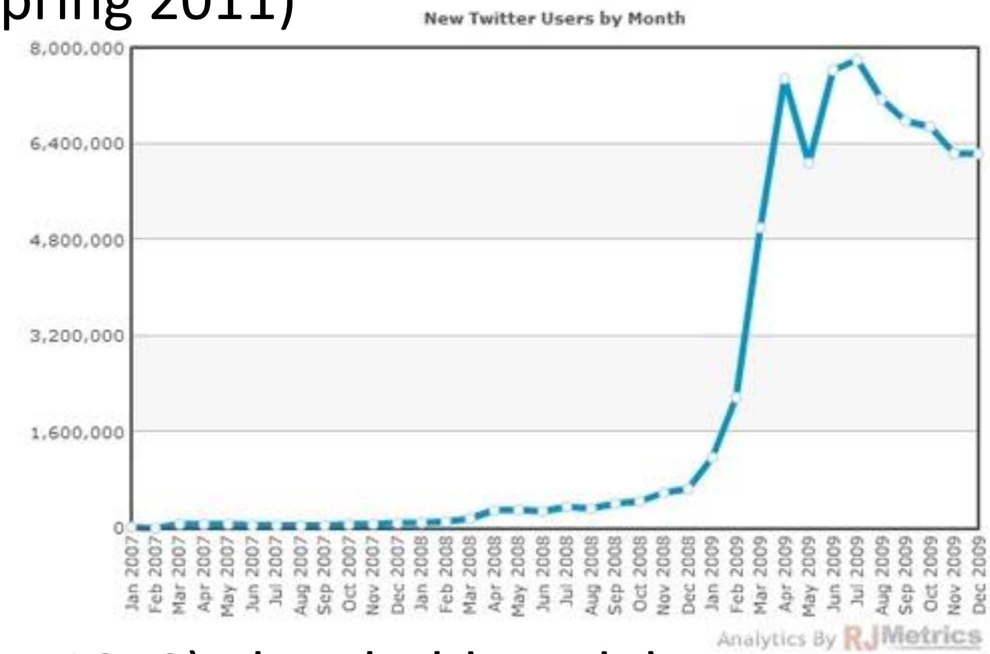
Complex social contagion can be surprisingly fast. A triggering perturbation may release rapid spreading.

Examples:

Rumor (e.g., false Hungarian nuclear breakdown 2002)

Political movements (Arab spring 2011)

Innovation: Twitter



Granovetter (Am. J. Sociology 1978) Threshold models

D. Watts (PNAS 2002) Mathematical form

Threshold Model

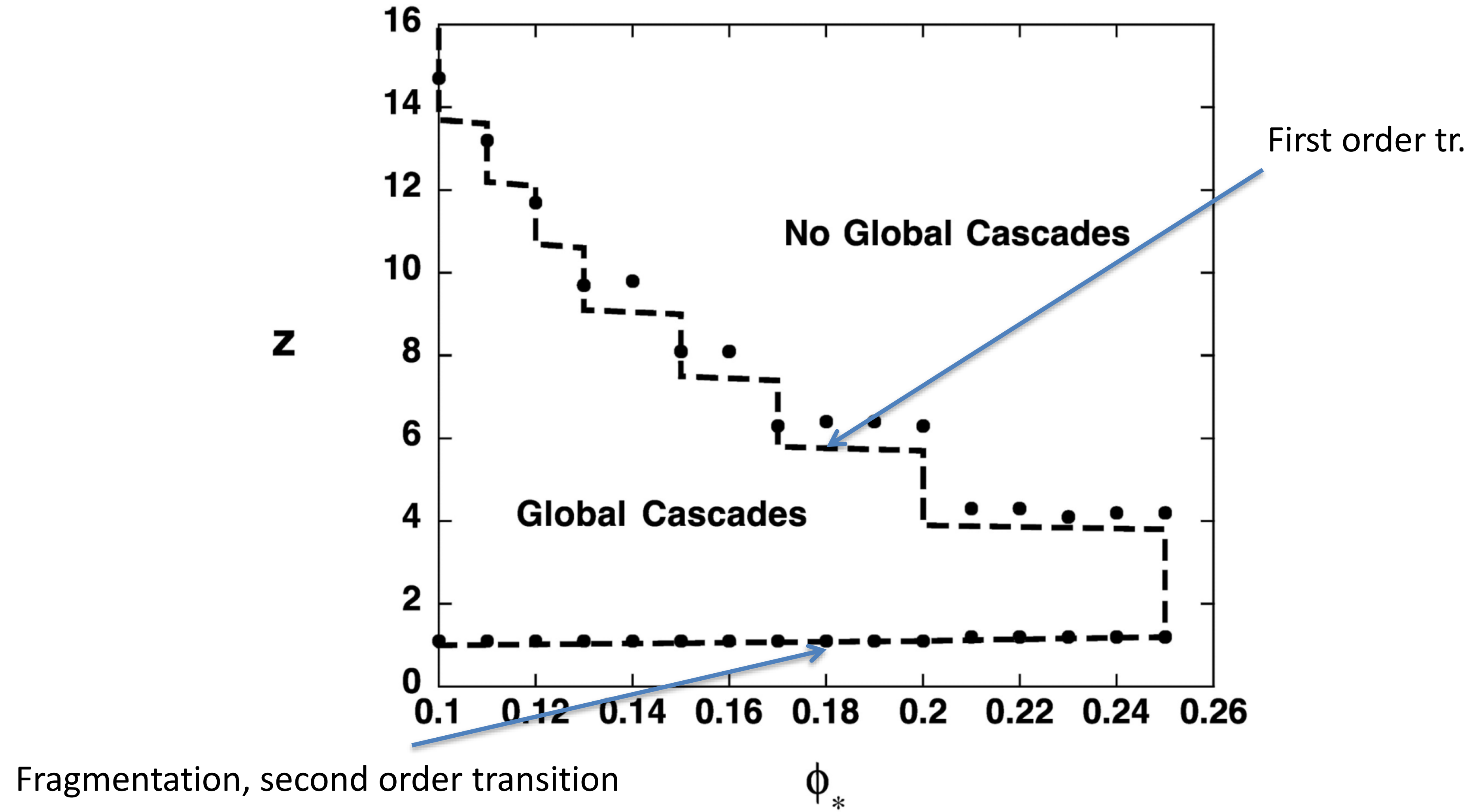
Random network with degree distribution p_k and average degree $\langle k \rangle = z$. Every node *i* has a threshold ϕ_i indicating the critical ratio of adopting neighbors needed to make the node adopting.

There are **vulnerable** nodes, which get infected if they have one adopting neighbor: $\phi \leq 1/k$.

The others are **stable**.

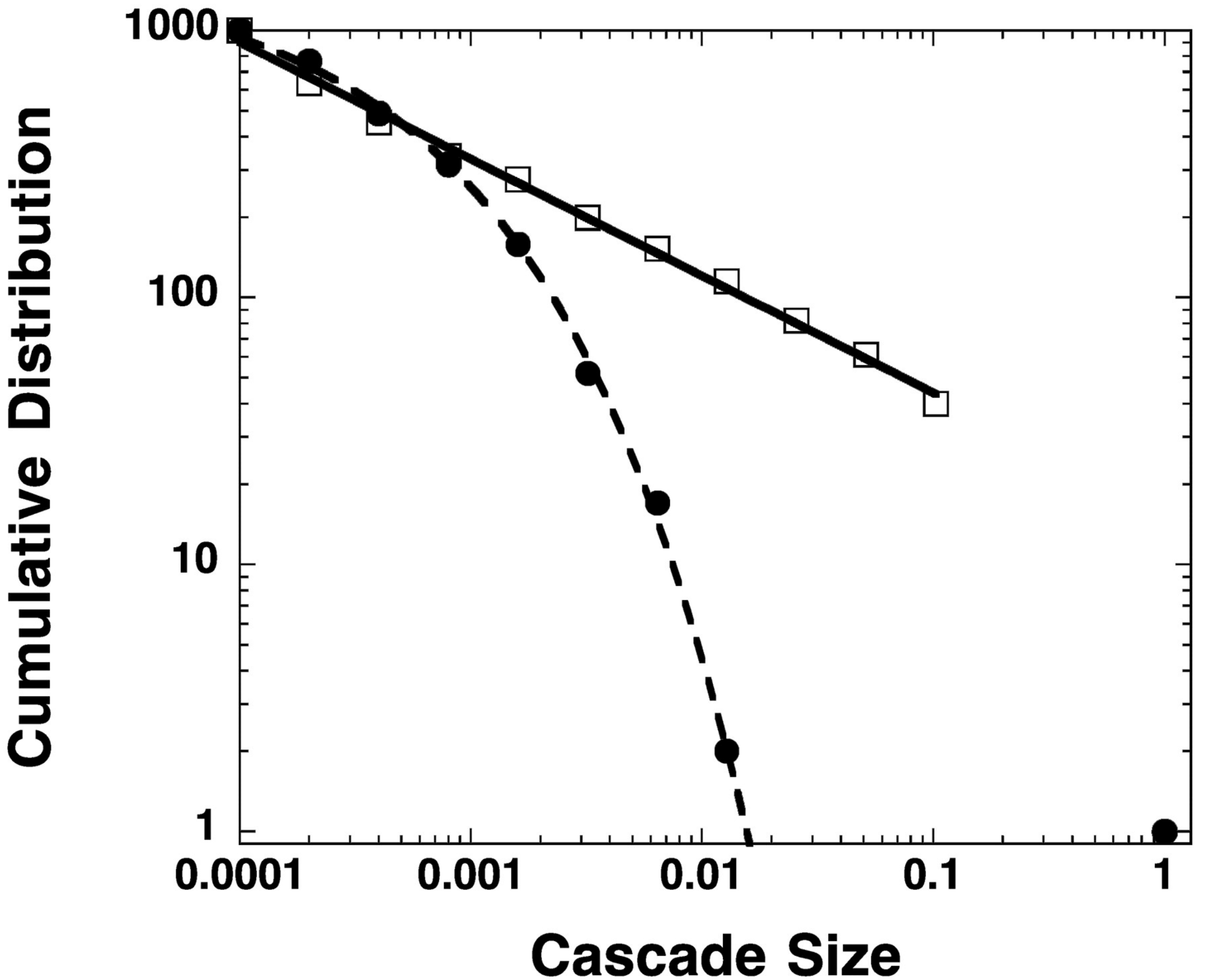
The phase diagram can be calculated.

Cascade windows for the threshold model. (ER graph)



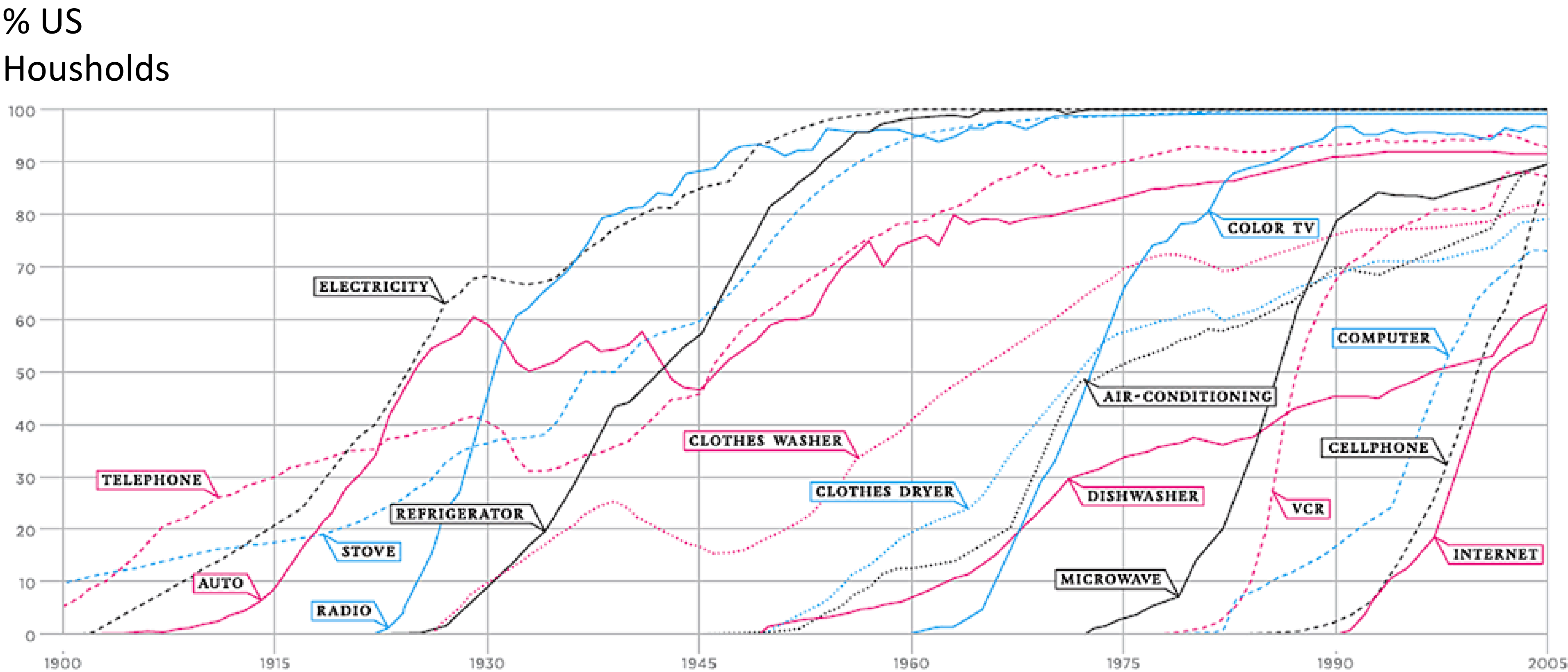
Watts D J PNAS 2002;99:5766-5771

Cumulative distributions of cascade sizes at the lower and upper critical points, for n = 1,000 and z = 1.05 (open squares) and z = 6.14 (solid circles), respectively.



Watts D J PNAS 2002;99:5766-5771

PNAS



Adoption speed can be very different for different innovations

http://www.nytimes.com/imagepages/2008/02/10/opinion/10op.graphic.ready.html

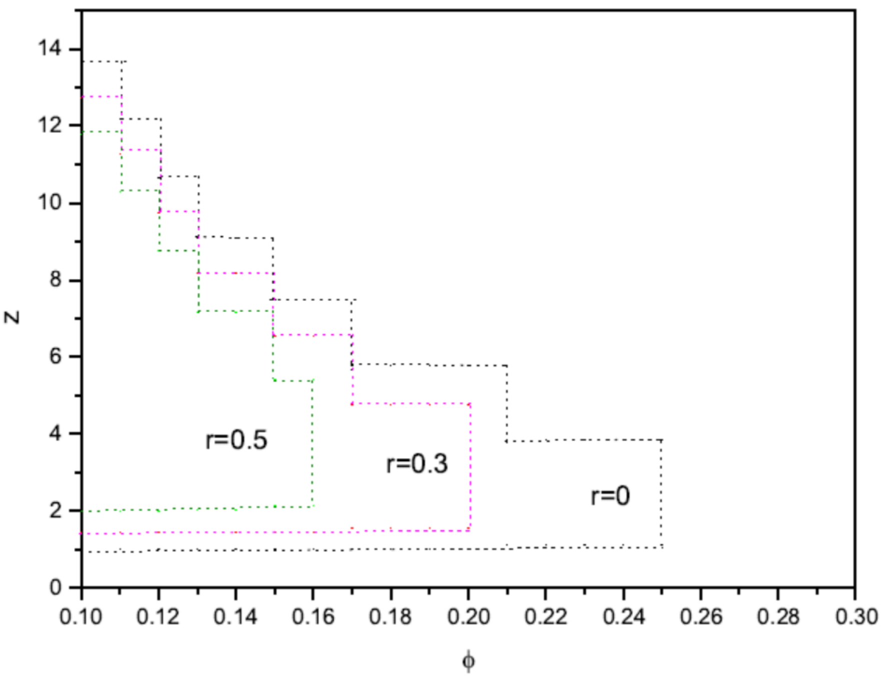
Generalized Watts Model

In the Watts model the criteria for a dynamic process (cascades) is traced back to a static problem, the existence of the percolating vulnerable cluster.

Incomplete picture:
1. There are more than one spontaneous innovators
 due to external information(Korniss et al. 2013) (Still static.)
2. Some nodes are blocked. Some people are reluctant to adopt (have a satisfactory service, have some reasons on principle etc.) (still static)
3. There are spontaneous innovators appearing
External information flows continuously (intrinsically dynamic)

Blocked Nodes

Nodes are blocked with probability *r* (quenched disorder).
Blocked nodes make it more difficult to fulfil the threshold criterion.
The problem can be solved similarly to the original Watts case, with the generating function method.
The result is a three-dimensional phase diagram:

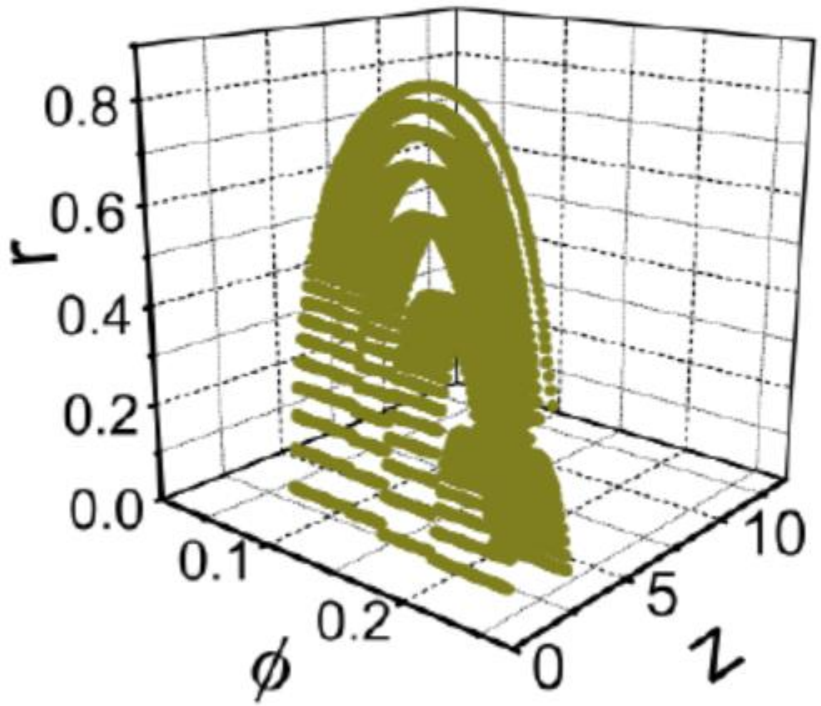


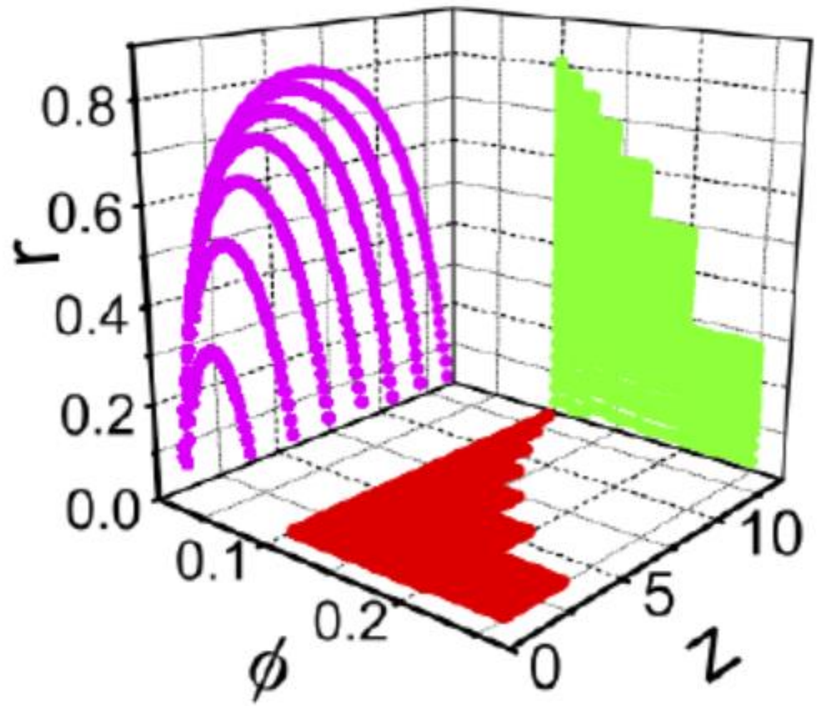
ER graph with average degree *z*, uniform threshold ϕ and blocking probability *r*.

3D Phase Diagram

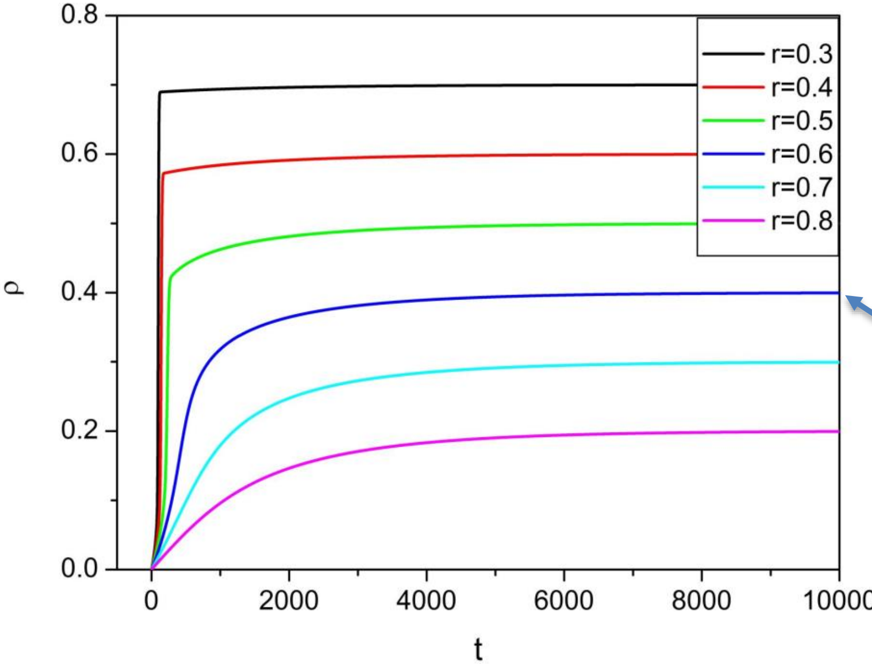
For Erdős-Rényi graph p_k is Poisson, parametrized by z.
Assuming uniform ϕ with $k_c = \lfloor 1/\phi \rfloor$

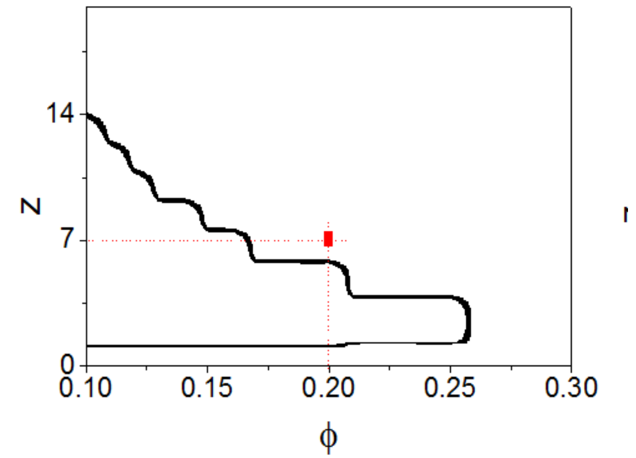
$$(1-r)e^{-z}\sum_{k=2}^{k_c}\frac{z^k}{(k-2)!} - z = 0$$





Spontaneous Adopters + Blocked Nodes

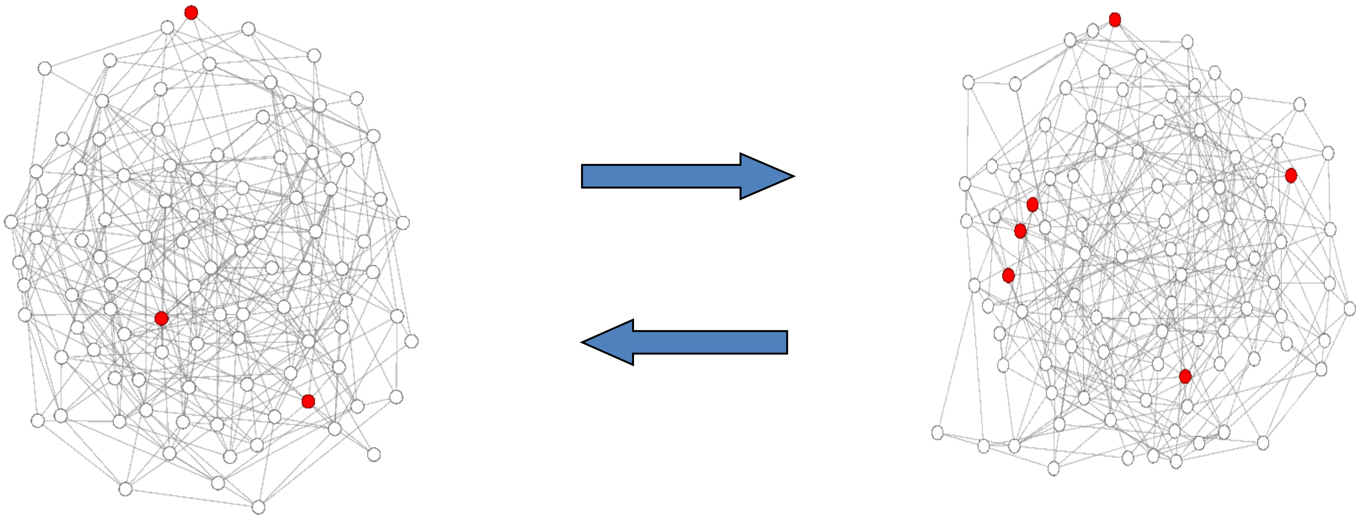




$1 - r$

ER, $z = 7,\ \phi = 0.2,\qquad p = 5 \times 10^{-4}$

Spontaneous vs. Induced Adoption

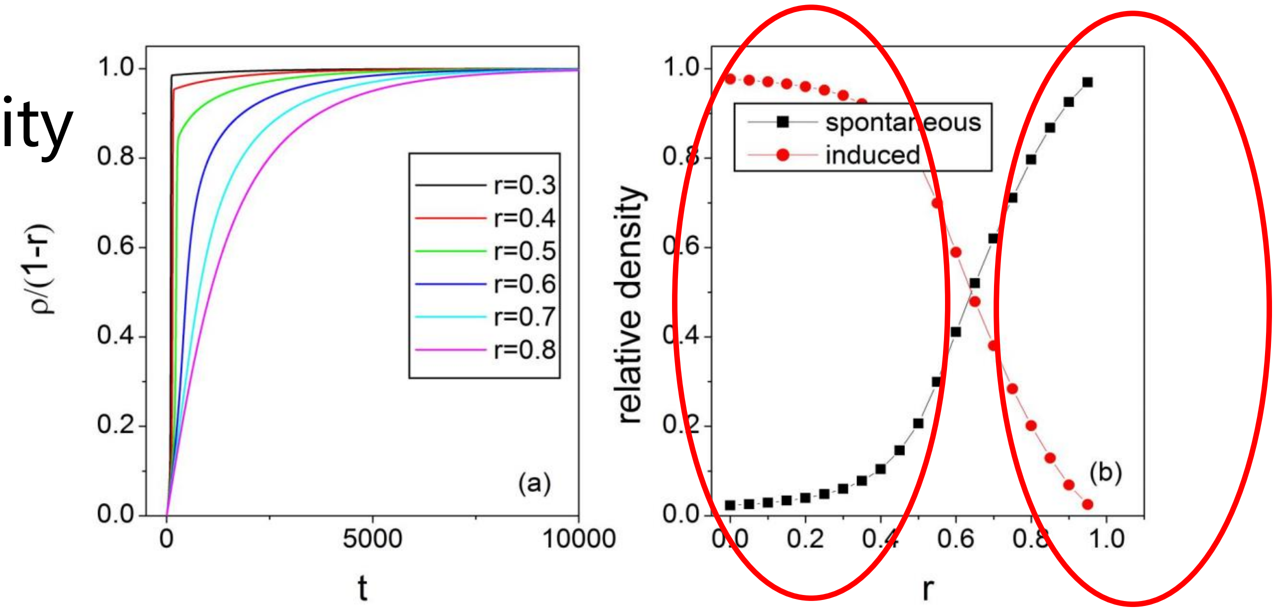


spontaneous induced

Evolution of Adopter Density

Different mechnisms?

Normalized adopter density



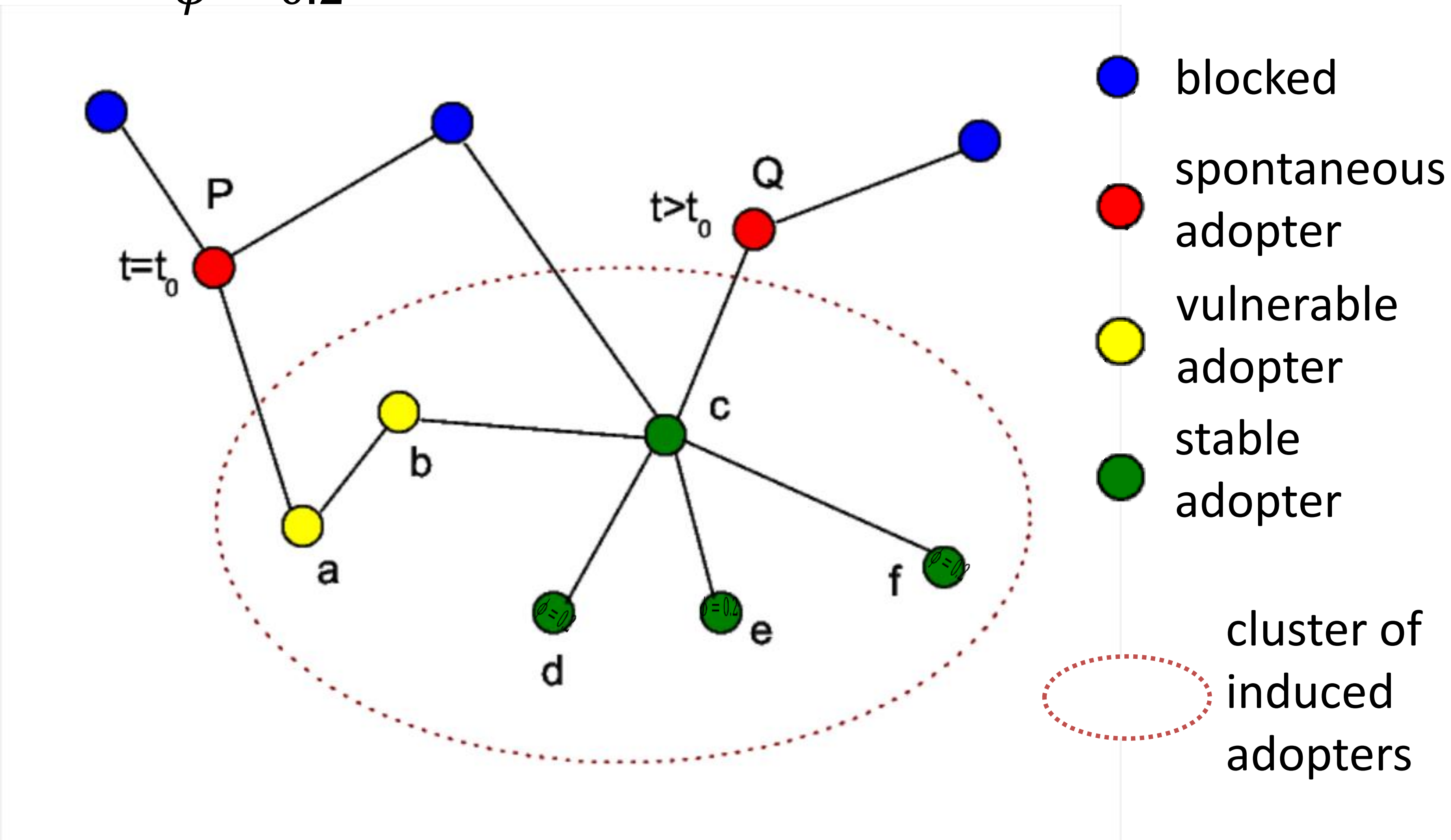
ER , $z = 7,\ \phi = 0.2, \qquad p = 5 \times 10^{-4}$

$r^* = 1 - 1/z = 0.86$ is the percolation threshold

Is there an $r_\times < r^*$ where the kinetics changes?

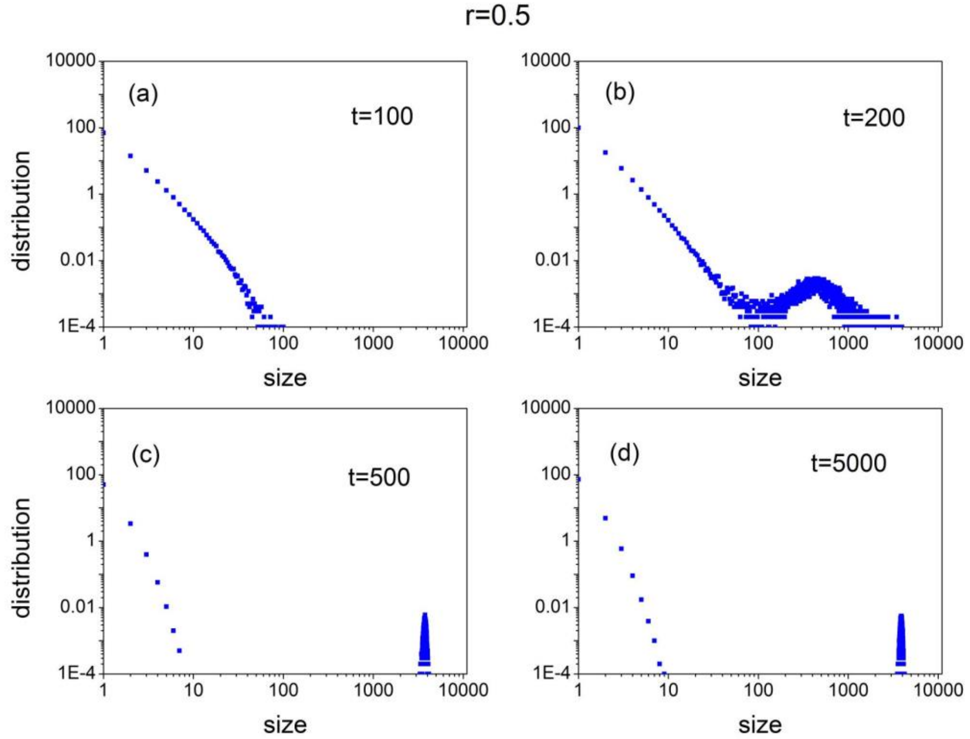
Node types

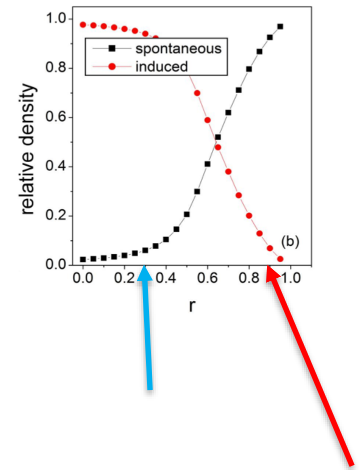
$\phi = 0.2$



Distribution of Induced Clusters ($r < r_{\times}$)

scenario I

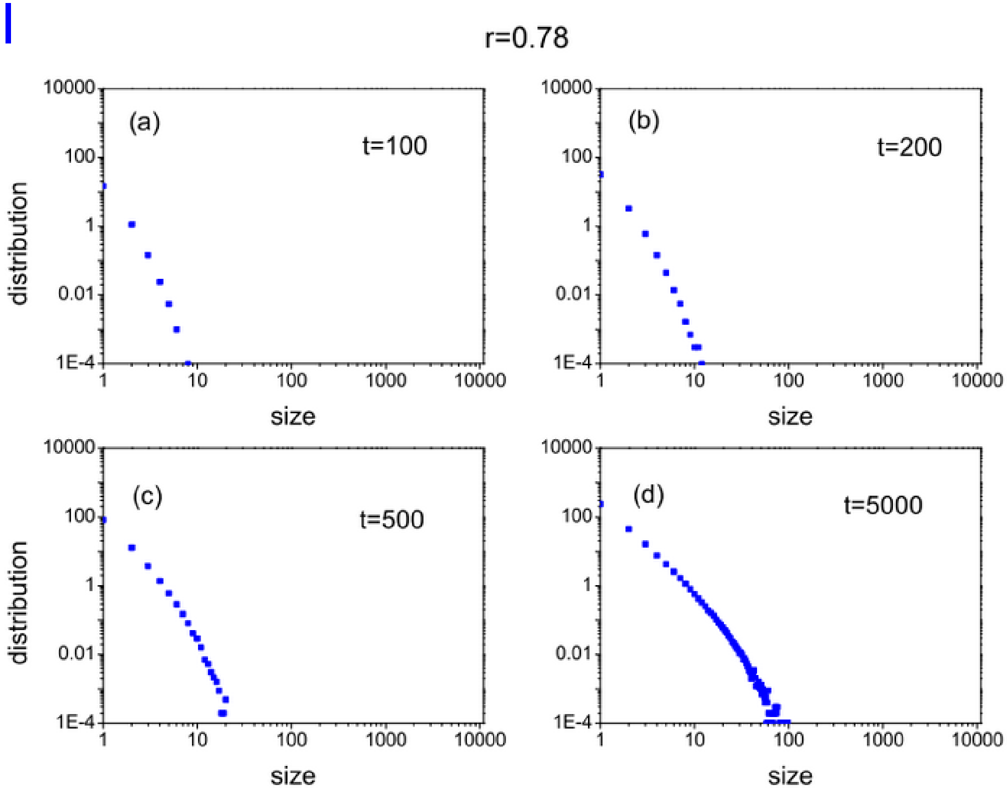


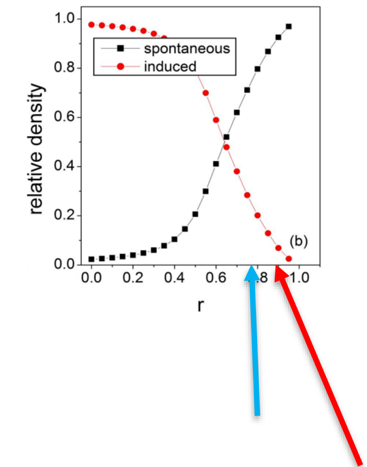


$$r = 0.5, \qquad z = 7, \qquad \phi = 0.2, \qquad p = 5 \times 10^{-4}$$

Distribution of Induced Clusters ($r_\times < r < r^*$)

scenario II



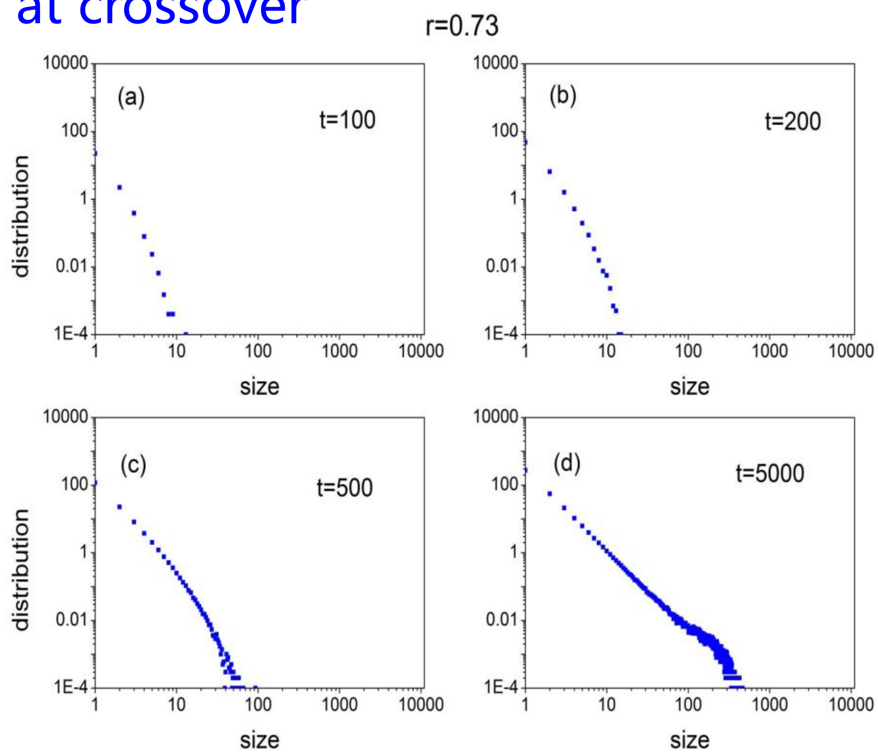


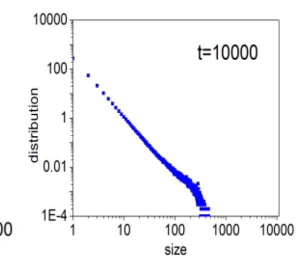
$r = 0.78, \qquad z = 7, \qquad \phi = 0.2, \qquad p = 5 \times 10^{-4}$

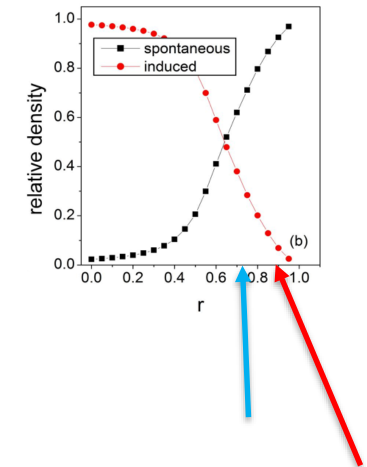
$r^* = 0.86$

Distribution of Induced Clusters ($r \sim r_{\times}$)

scenario at crossover

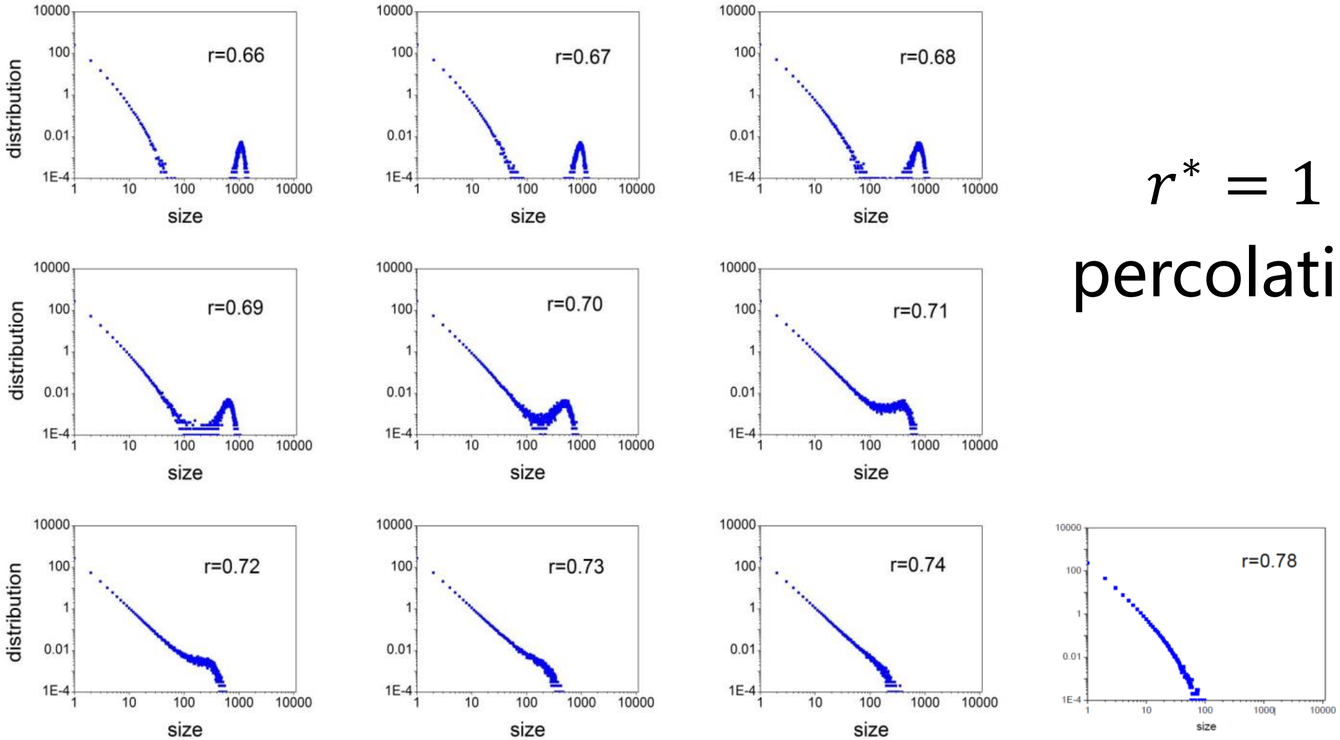






$$r = 0.73, \qquad z = 7, \qquad \phi = 0.2, \qquad p = 5 \times 10^{-4}$$

Asymptotic Distribution of Induced Clusters



$r^* = 1 - 1/z = 0.86$

percolation threshold

$z = 7, \qquad \phi = 0.2, \qquad p = 5 \times 10^{-4}, \qquad t = 5000$

Ruan, et al. PRL 2016

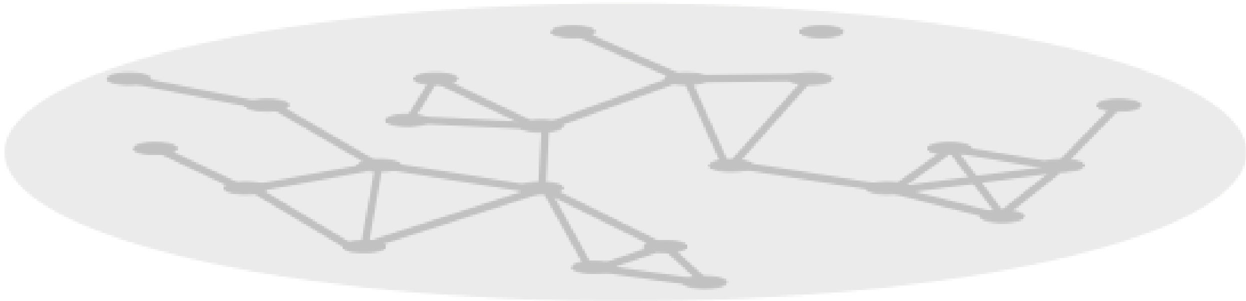
Instead of Anecdotes: Big Data

Information about:
- Basic service network
- Adoption of additional services
- Data about location (IP)

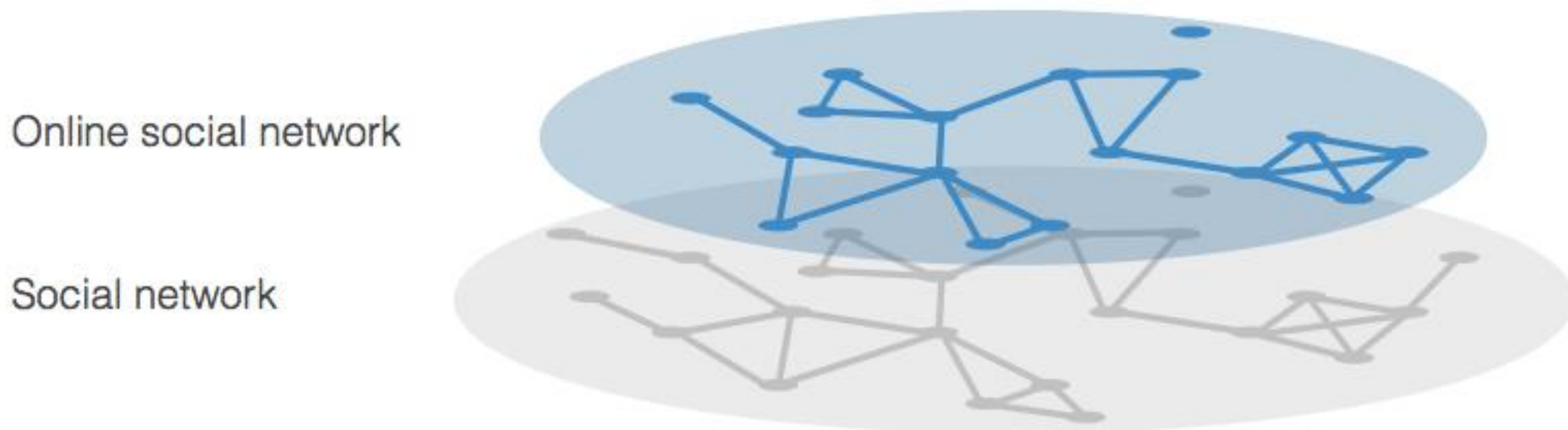
- 700+ million users world-wide
  - September 2003 - March 2011 (2738 days)
  - Registration dates
  - Location & self reported demographic data
  - Spamming accounts are removed
- Link creation dynamics
  - Time stamped link addition events
  - Only confirmed links
- Free and Payed services
  - 6 free and 9 payed services
  - Time of adoption
  - Usage activity sequences
- Country networks
  - For calculations we selected users in single countries
  - For selected users we considered all first neighbors
  - Look for the behaviour of country users only

Social network layer

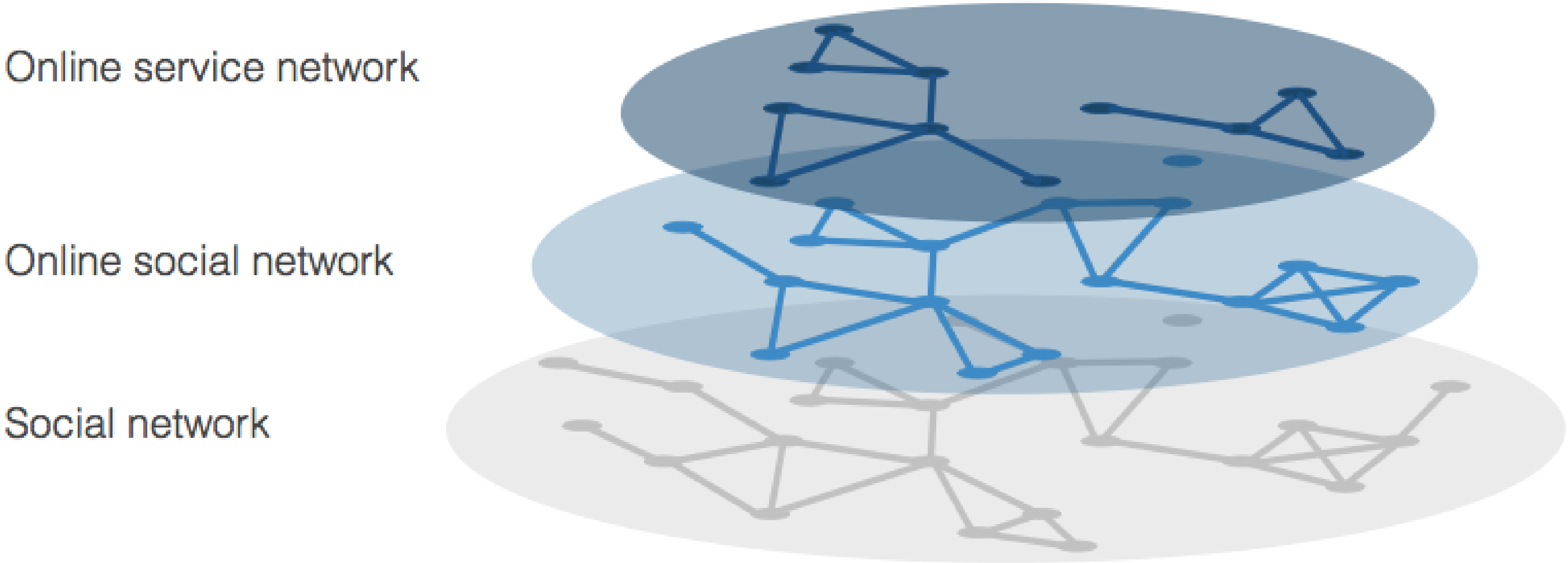
Social network



Online social network layer



Online service network layer



unknown

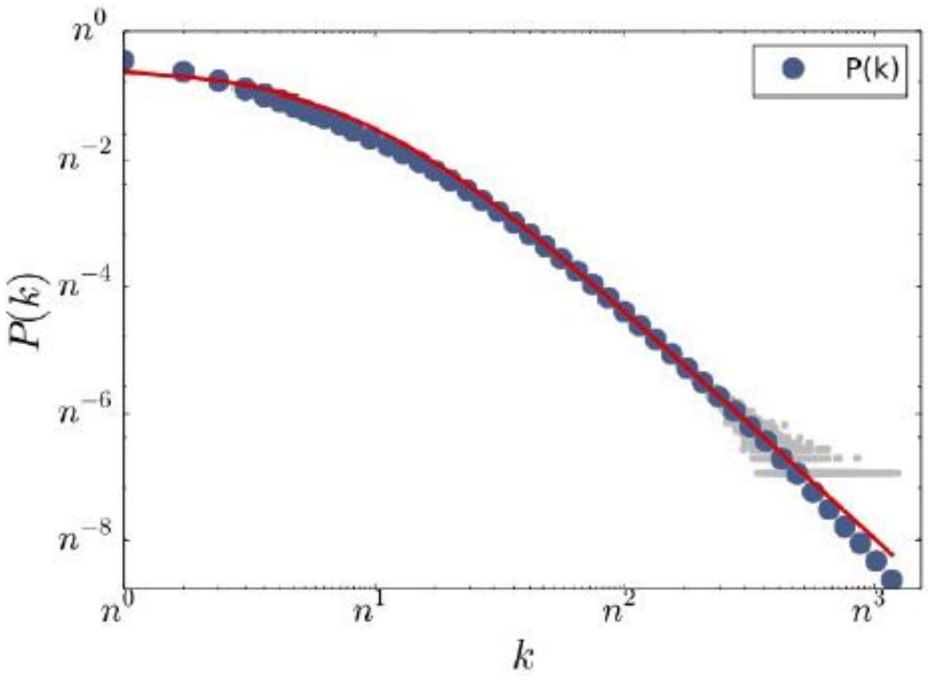
Earlier work: M. Karsai, G. Iniguez, K. Kaski and J.K:
Journal of the *Royal* Society *Interface* 11 (101), 20140694, 2014.

Empirical Results

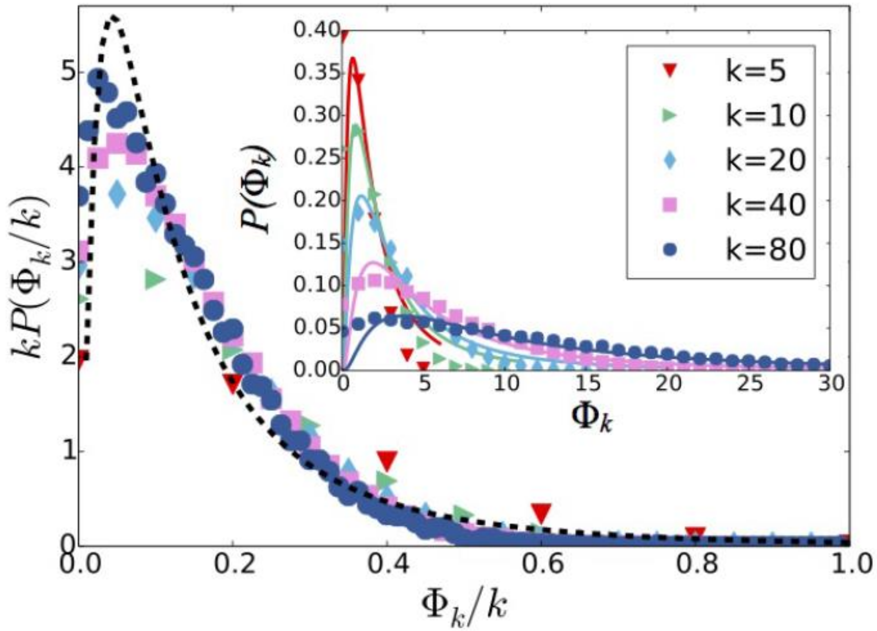
Here we know the underlying network: 520 M nodes of the Voice over Internet service.

r=0.95. The network is NOT ER, broad degree distribution.

Shifted power law with γ=3.8 (z=8.56).

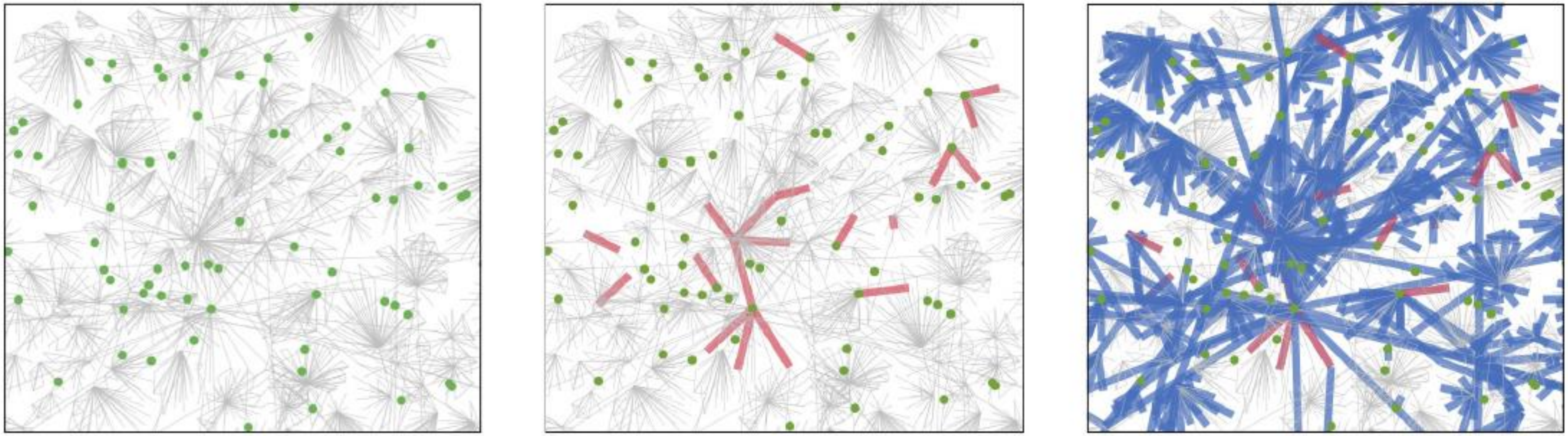


Empirical Results



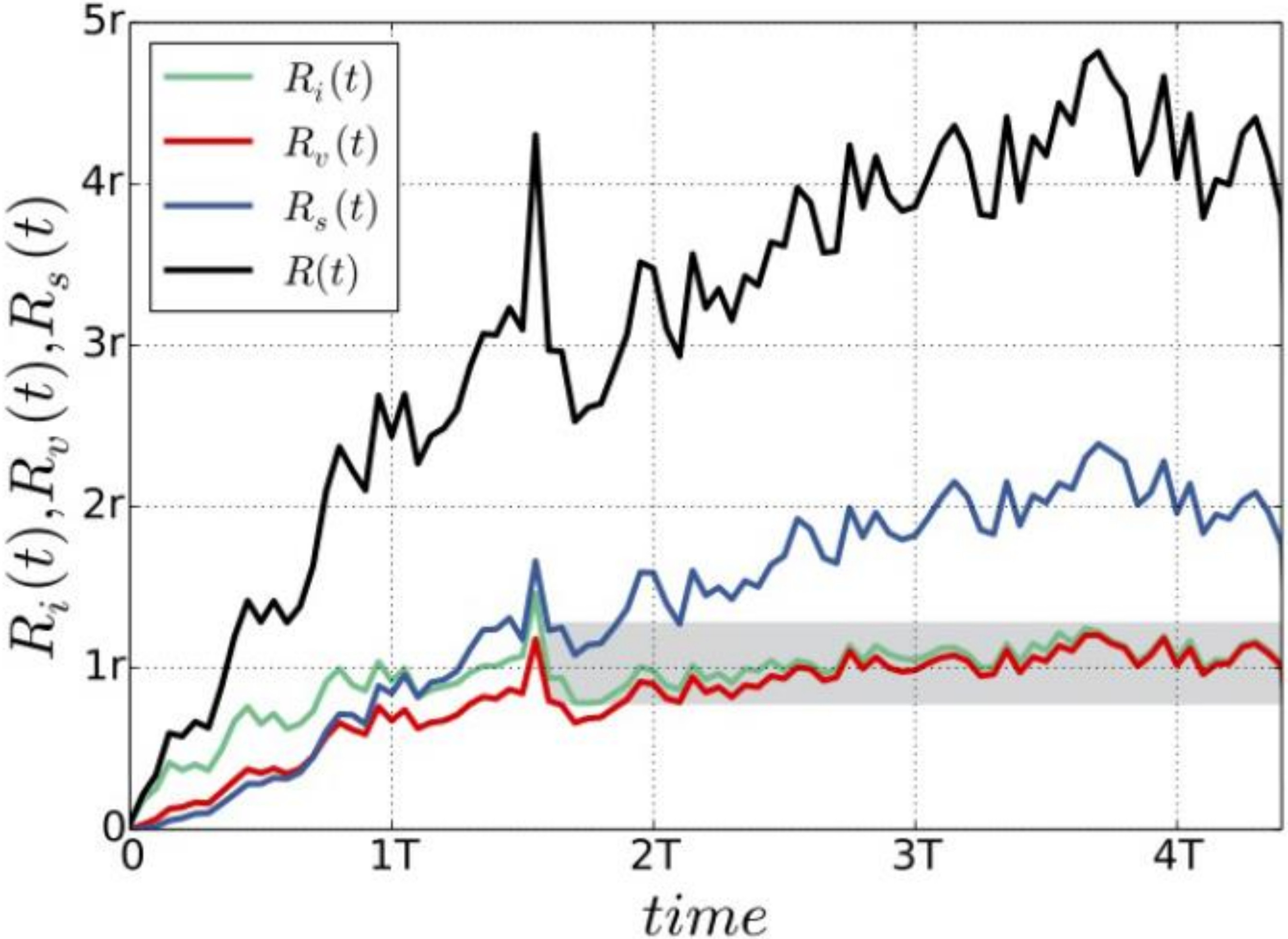
Empirical threshold distribution: log-normal $\langle \varphi \rangle = 0.19$

Empirical Results



Initiators vulnerable clusters adopters

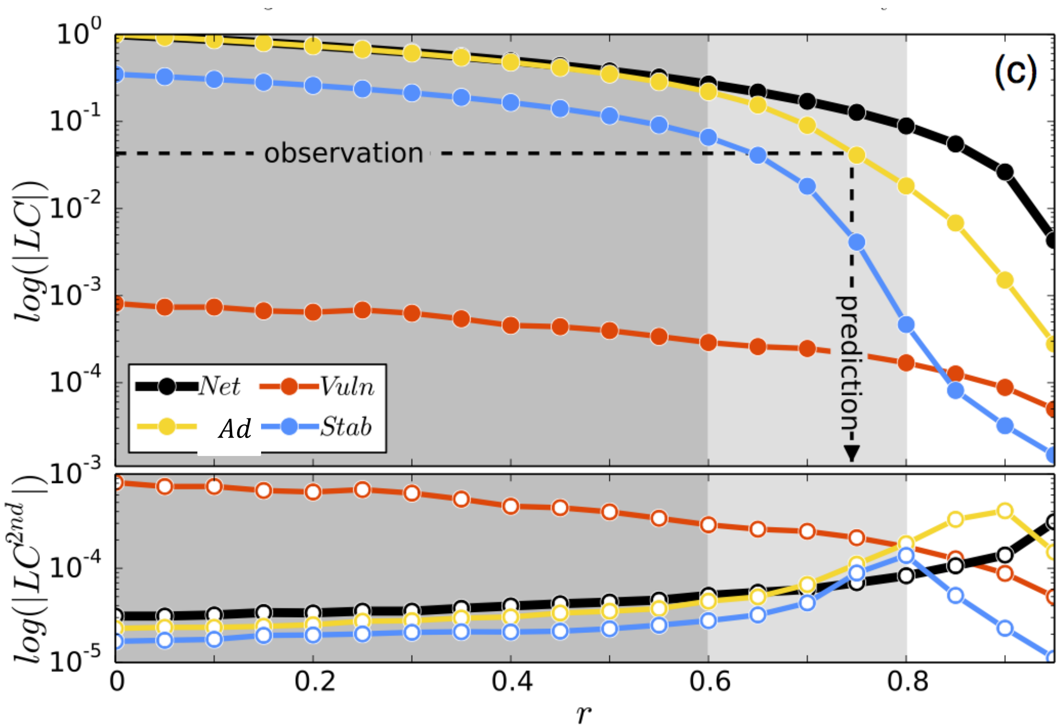
Empirical Results



Linear growth

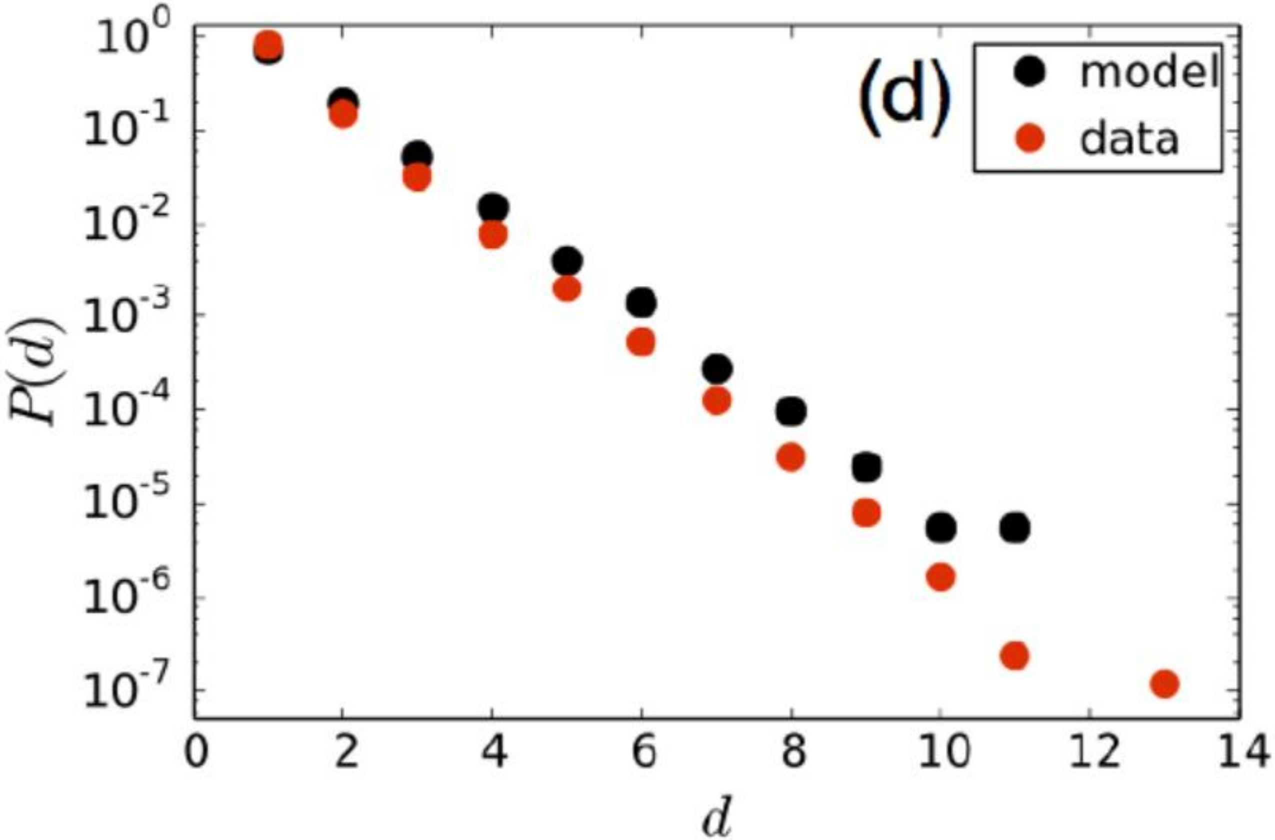
Rates

Empirical Results – Comparison with Model



Model calculation with empirical threshold and degree distributions and evolution time. The density r_{emp} is determined from the plot: $r_{\mathrm{emp}} = 0.745$.

Empirical Results – Comparison with Model



Distribution of depth of vulnerable trees

REFERENCES

- J.-P. Onnela et al. *PNAS* **104**, 7332-7336 (2007), *New J. Phys.* **9**, 179 (2007)
- M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, J. Saramäki: Small But Slow World: How Network Topology and Burstiness Slow Down Spreading, *Phys. Rev. E* **83**, 025102 (2011)
- Márton Karsai, Kimmo Kaski, Albert-László Barabási, János Kertész: Universal features of correlated bursty behavior, *Scientific Reports* **2**, Article number: 397 (2012)
- Márton Karsai, Kimmo Kaski, János Kertész: Correlated dynamics in egocentric communication networks: *PLoS ONE* **7**(7) e40612 (2012)
- H.-H. Jo, M Karsai, J Kertész, K Kaski: Circadian pattern and burstiness in mobile phone communication, *New Journal of Physics* **14**, 013055 (2012)
- Hang-Hyun Jo, Juan I. Perotti, Kimmo Kaski, János Kertész: Enhanced Spreading Dynamics by Non-Poissonian Processes, *Phys. Rev. X,* **4**, 011041 (2014)
- Dávid X. Horváth, János Kertész: Spreading dynamics on networks: the role of burstiness, topology and non-stationarity, *New Journal of Physics* arXiv:1404.2468

Reviews, Textbooks

- P. Holme J. Saramäki: Temporal Networks, Phys. Rep. 519, 97-125 (2012)
- M. Karsai et al. Bursty Human Dynamics, Springer 2018
- A.-L. Barabási: Network Science (Cambridge, 2017)
- R. Pastor-Satorras et al. Epidemic spreading in complex networks, Rev. Mod. Phys. 85, 925 (2015)

Summary

- Cascade model can be extended to describe rich kinetics of spreading by inclusion of blocked nodes and spontaneous innovators.
- Blocked nodes problem can be solved by generating function method. 3D phase diagram.
- The general rate equation catches important features of the model. Fast and slow regimes.
- Simulations show that there is a percolation transition of induced clusters in the background.
- ICT based data help in understanding the laws of innovation spreading. Two levels of Skype data: Free and payed services
- The spreading of payed service is relatively slow due to the large number of „blocked" individuals.