Temporal Networks

aka time-varying networks, time-stamped graphs, dynamical networks...

Network Theory and Applications

ECS 253 / MAE 253

Spring 2016

Márton Pósfai (posfai@ucdavis.edu)

Sources

Reviews:

Physics Reports 519 (2012) 97–125

Contents lists available at SciVerse ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

Temporal networks

Petter Holme [a,b,c,*], Jari Saramäki [d]

Eur. Phys. J. B (2015) 88: 234
DOI: 10.1140/epjb/e2015-60657-4

THE EUROPEAN PHYSICAL JOURNAL B

Colloquium

Modern temporal network theory: a colloquium*

Petter Holme[a]

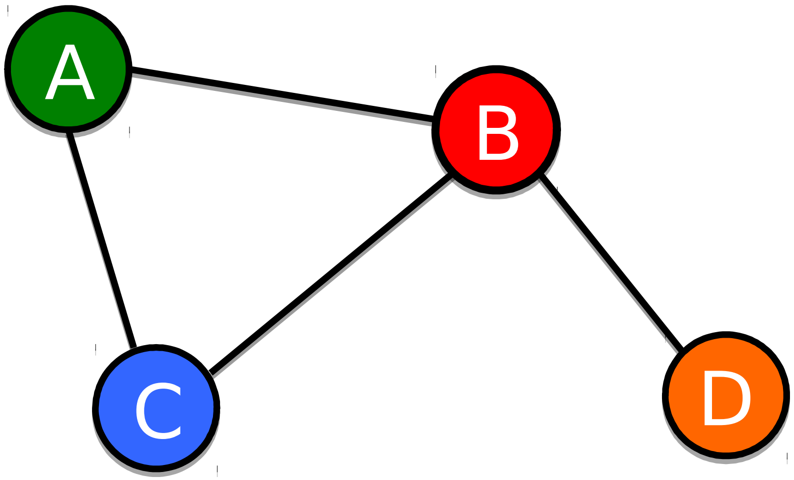
Department of Energy Science, Sungkyunkwan University, 440-746 Suwon, Korea

Courses:

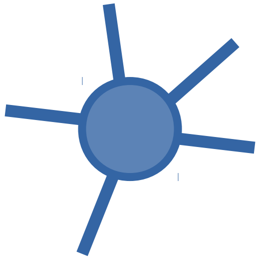
- ENS Lyon: Márton Karsai
- Northeastern University: Nicola Perra, Sean Cornelius, Roberta Sinatra

Temporal Networks

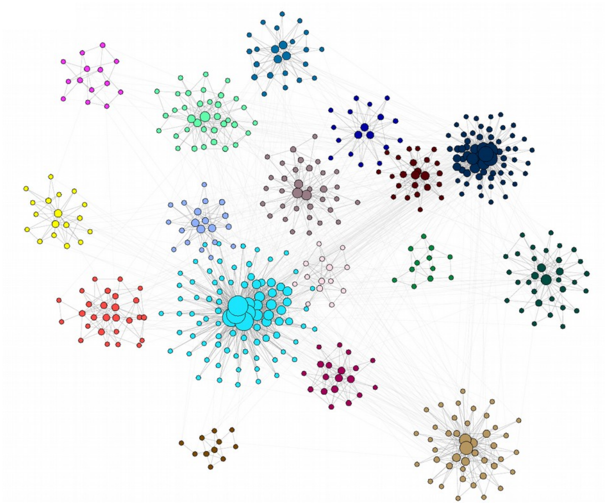
- So far: static network



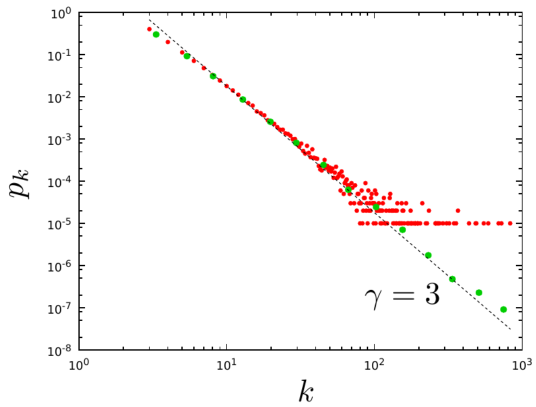
- Description:



Microscopic:
Node, link properties
(degree, centralities)



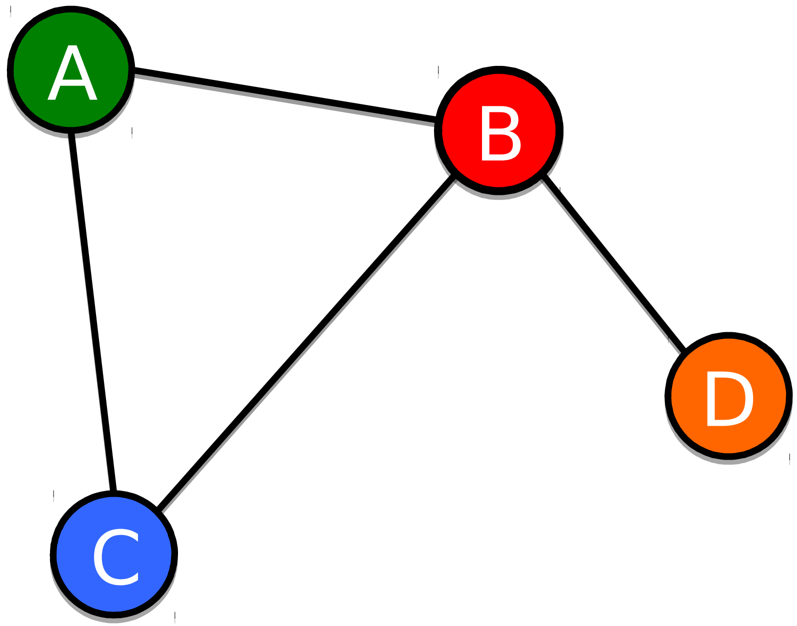
Mezoscopic:
Motives, communities



Macroscopic:
statistics

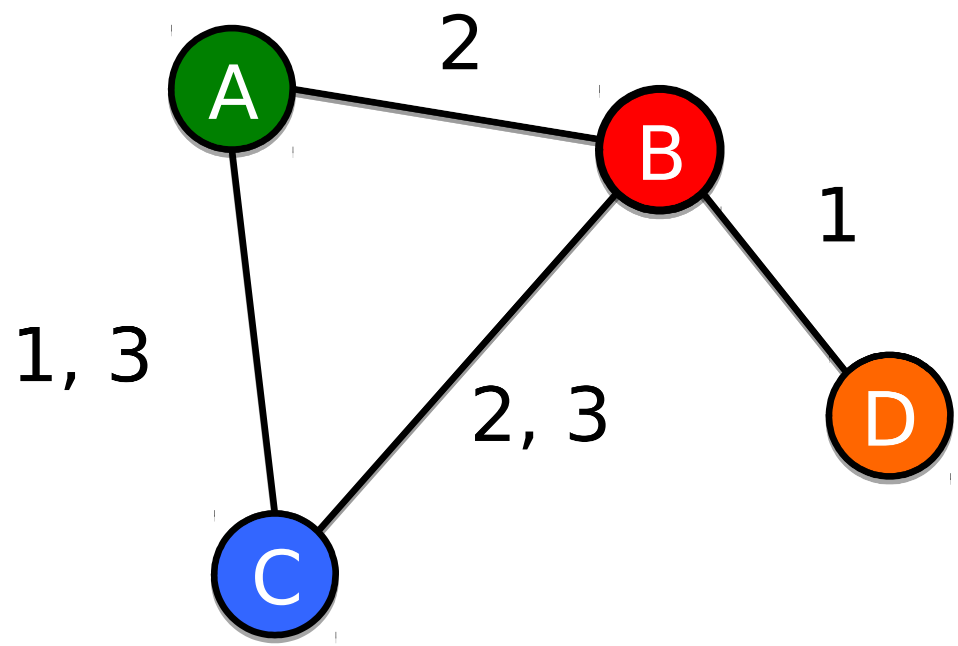
Temporal Networks

- Static network: Spreading process can reach all nodes starting from A.



Temporal Networks

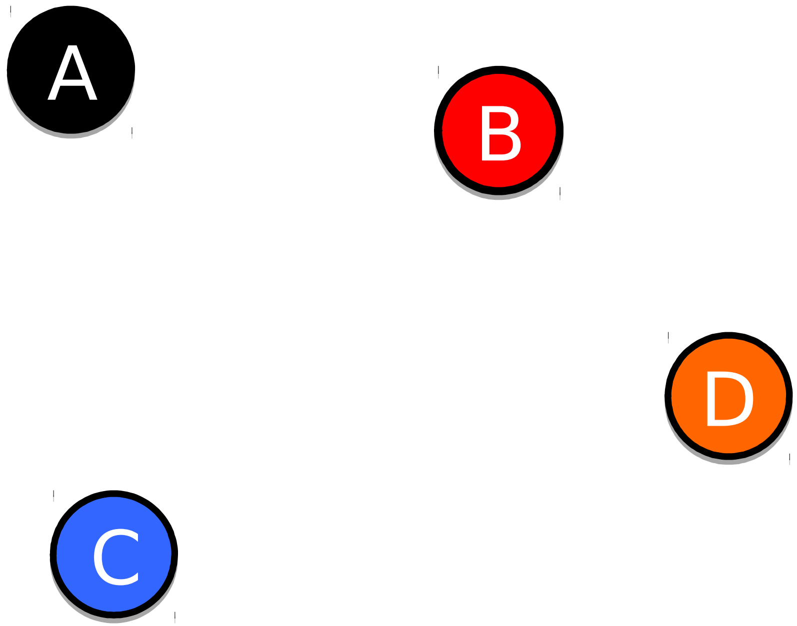
- Now: time of interaction



Temporal Networks

- Now: time of interaction

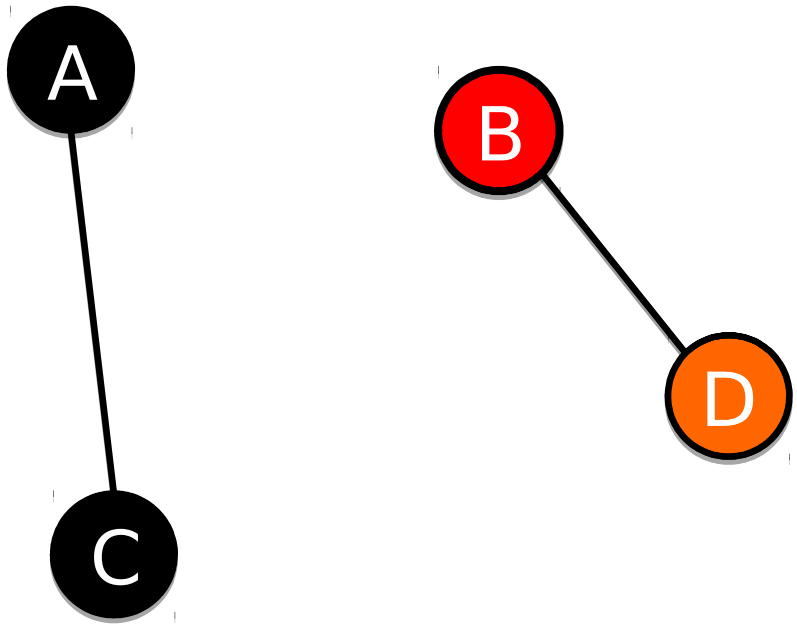
t=0



Temporal Networks

- Now: time of interaction

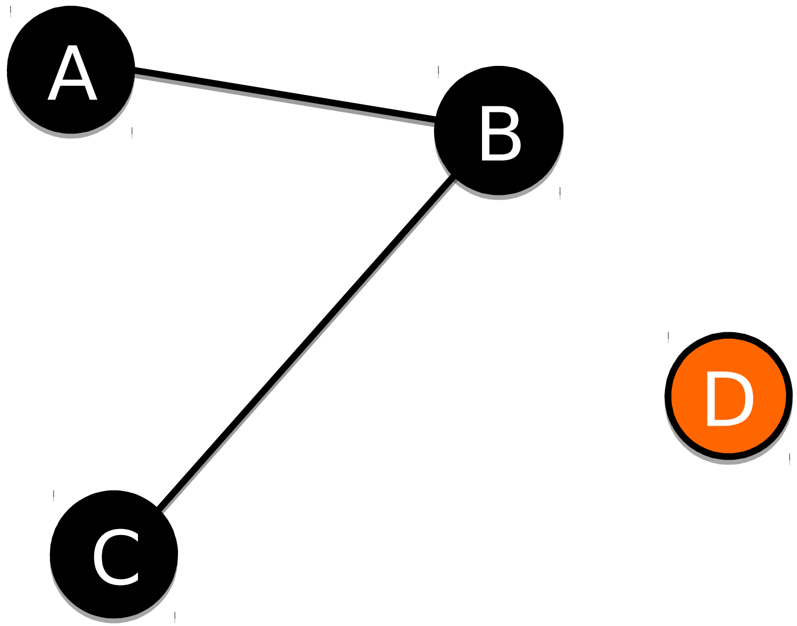
t=1



Temporal Networks

- Now: time of interaction

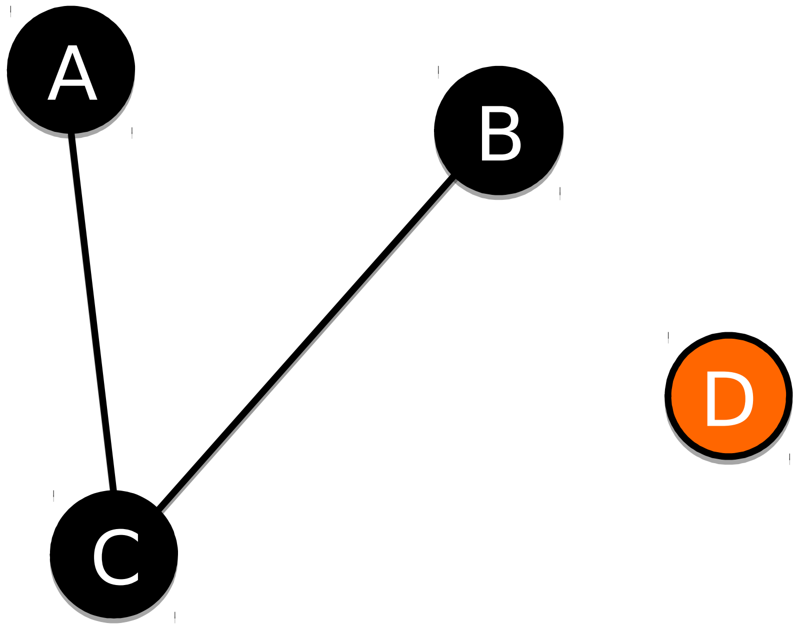
t=2



Temporal Networks

- Now: time of interaction

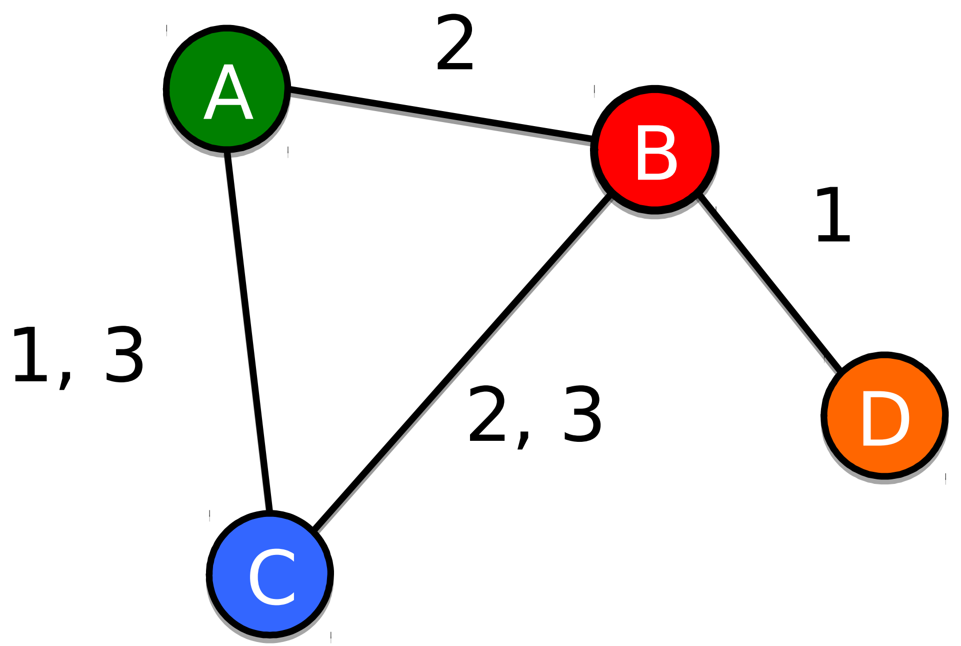
t=3



Temporal Networks

- Now: time of interaction

t=0



?

?

?

Microscopic:
Node, link properties
(degree, centralities)

Mezoscopic:
Motives, communities

Macroscopic:
statistics

When are temporal nets useful?

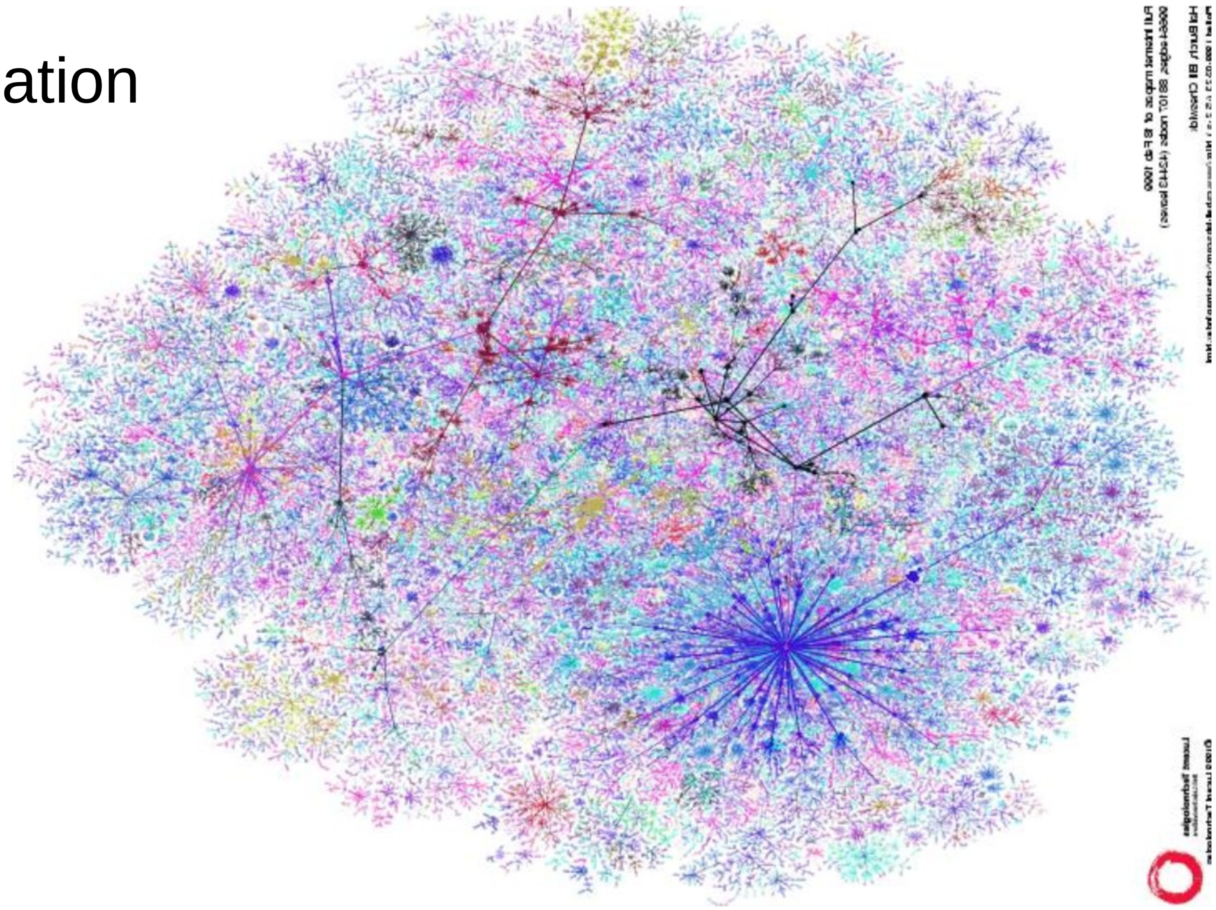
- Timescales:

τ_{D} : timescale of dynamics

τ_{N} : timescale of changes in network

- $\frac{\tau_{\mathrm{D}}}{\tau_{\mathrm{N}}} \gg 1$: static approximation

Example: Internet



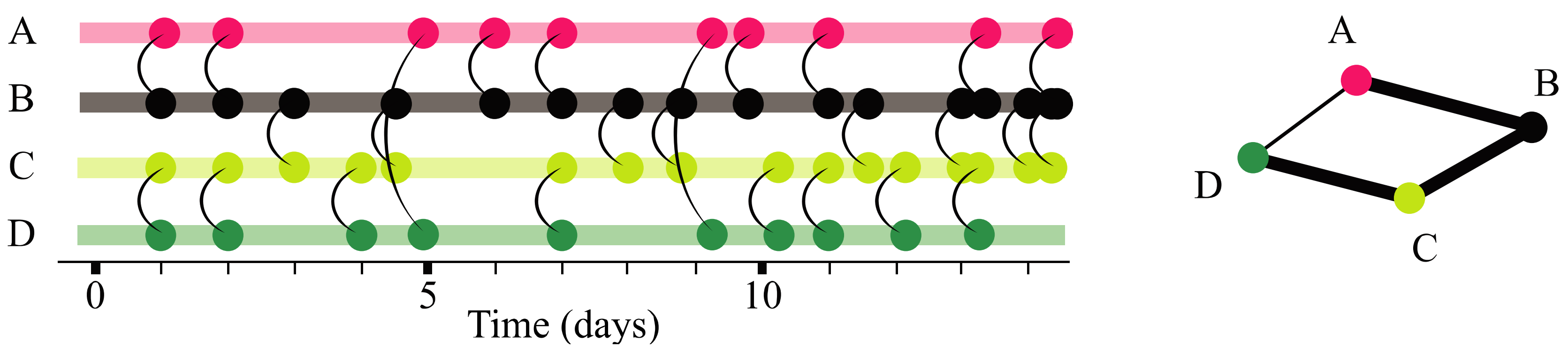
Branigan et al, *Internet Computing, IEEE* (2001)

When are temporal nets useful?

- Timescales:

τ_{D} : timescale of dynamics
τ_{N} : timescale of changes in network

- $\frac{\tau_{\mathrm{D}}}{\tau_{\mathrm{N}}} \ll 1$: annealed approximation



Example: PC or Mac

- Process slow enough that A meets all contacts
- Weight: how frequently they meet

P. Holme and J. Saramäki. "Temporal networks." Physics reports 519.3 (2012): 97-125.

When are temporal nets useful?

- Timescales:

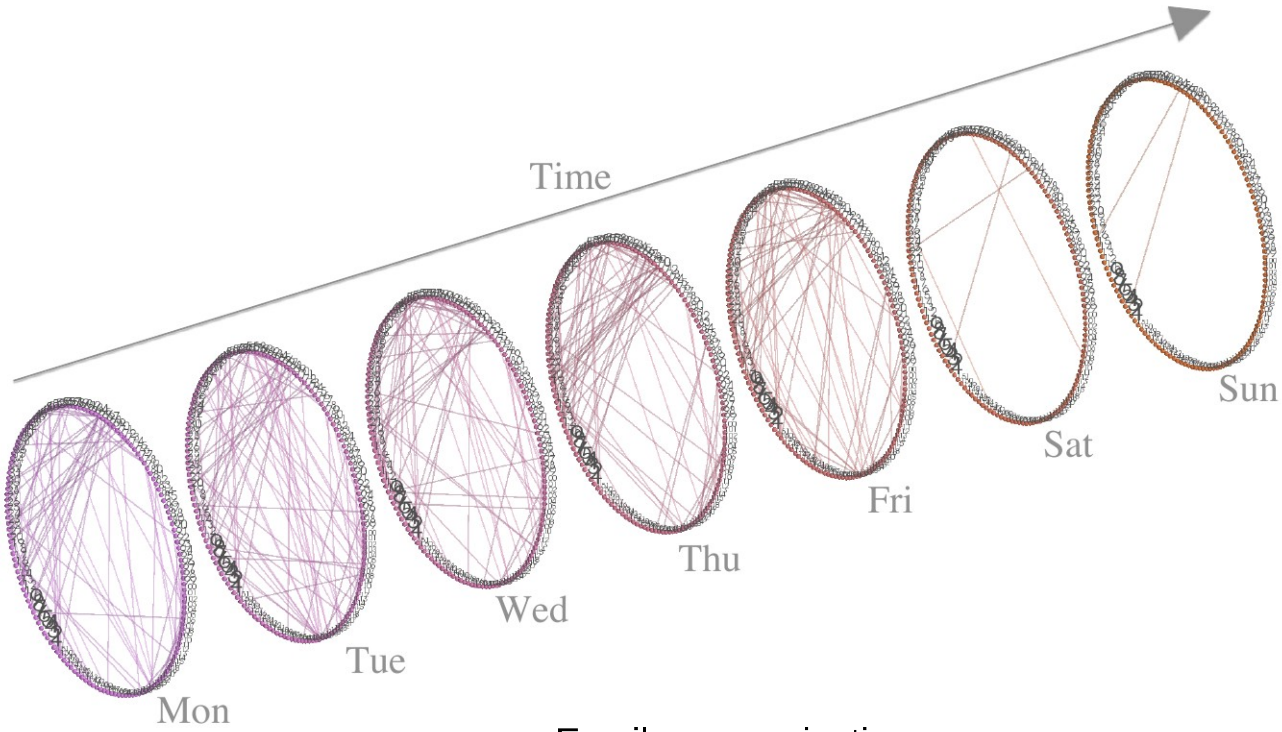
$\tau_{\rm D}$: timescale of dynamics
$\tau_{\rm N}$: timescale of changes in network

- $\dfrac{\tau_{\rm D}}{\tau_{\rm N}} \sim 1$:

TEMPORAL NETWORKS

Examples

- Communication: Email, phone call, face-to-face
- Proximity: same hospital room, meet at conference, animals hanging out
- Transportation: train, flights…
- Cell biology: protein-protein, gene regulation
- ...

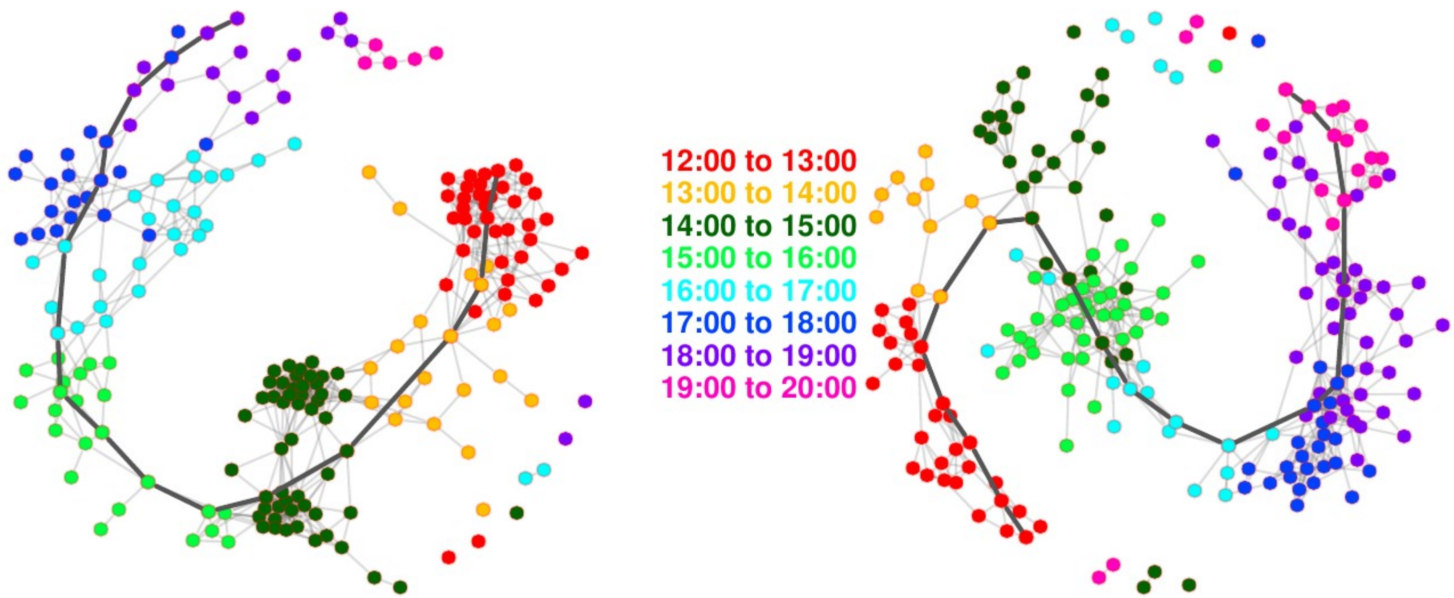


Email communication

J. Tang, John, Social Network Systems. ACM, 2010.

Examples

- Communication: Email, phone call, face-to-face
- Proximity: same hospital room, meet at conference, animals hanging out
- Transportation: train, flights…
- Cell biology: protein-protein, gene regulation
- ...

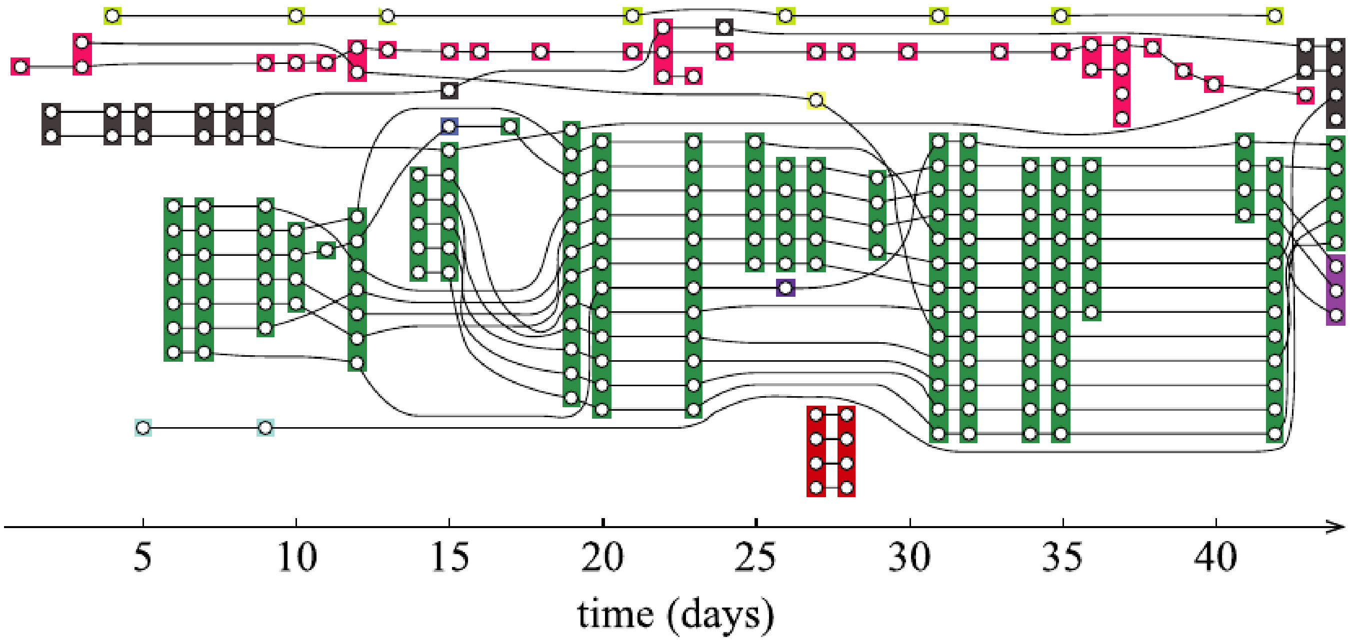


Visitors at exhibit.

Isella, Lorenzo, et al. Journal of theoretical biology 271.1 (2011): 166-180.

Examples

- Communication: Email, phone call, face-to-face
- Proximity: same hospital room, meet at conference, animals hanging out
- Transportation: train, flights…
- Cell biology: protein-protein, gene regulation
- ...



Temporal network of zebras
C. Tantipathananandh, et.al. (2007)

Plan

1) Mathematical representation
2) Path- based measures of temporal
3) Temporal heterogeneity
4) Processes and null models
5) Motifs

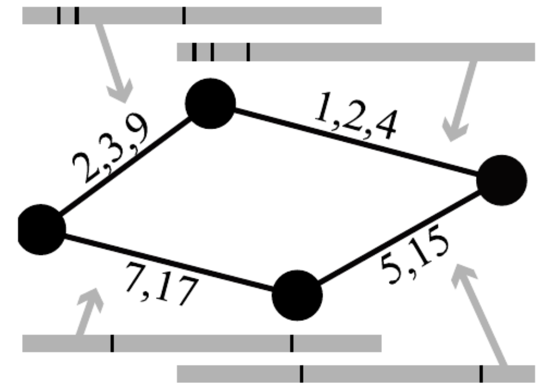
Mathematical Description

Mathematical representation

- **Temporal graph:** $G_t = (V, E, T_e)$

Set of vertices. Set of edges. $T_e = \{t_1, t_2, ..., t_n\}$

Set of times when edge *e* is active.



1. Contact sequence

$$E \subset T \times V \times V$$

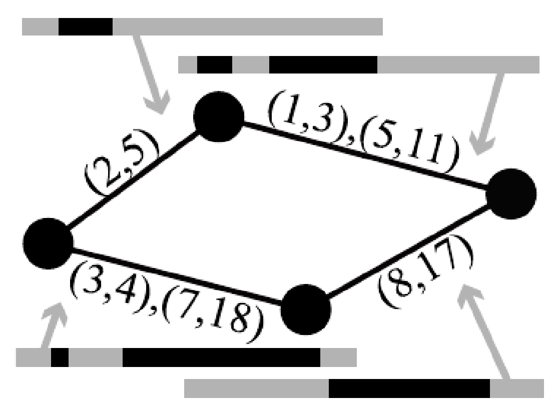
2. Adjacency matrix sequence

$$A_{ij}(t)$$

- **Interval graph:** $G_t^d = (V, E, T_e^d)$

$$T_e^d = \{(t_1, t_1'), ..., (t_n, t_n')\}$$

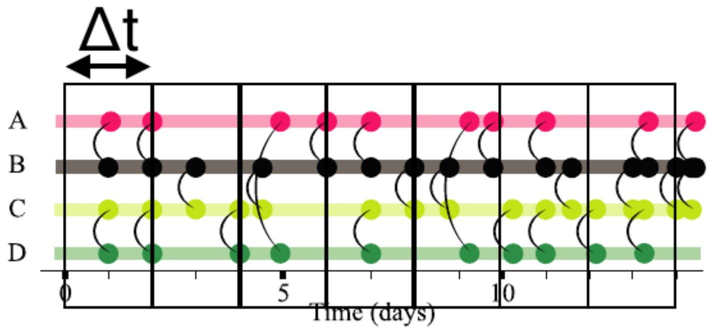
Set of intervals when edge *e* is active.



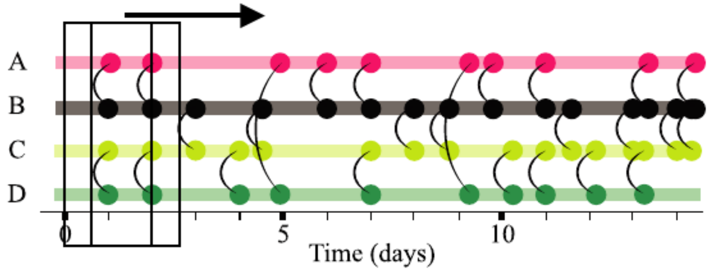
P. Holme and J. Saramäki. "Temporal networks." Physics reports 519.3 (2012): 97-125.

Aggregating in time windows

- Sequence of snapshots



Consecutive windows



Sliding windows

- Lossy method
- Sometimes data is not available
- Convenient: Static measures on snapshots → Time series of measures

- Problem: snapshots depend on window size?
- How to choose?

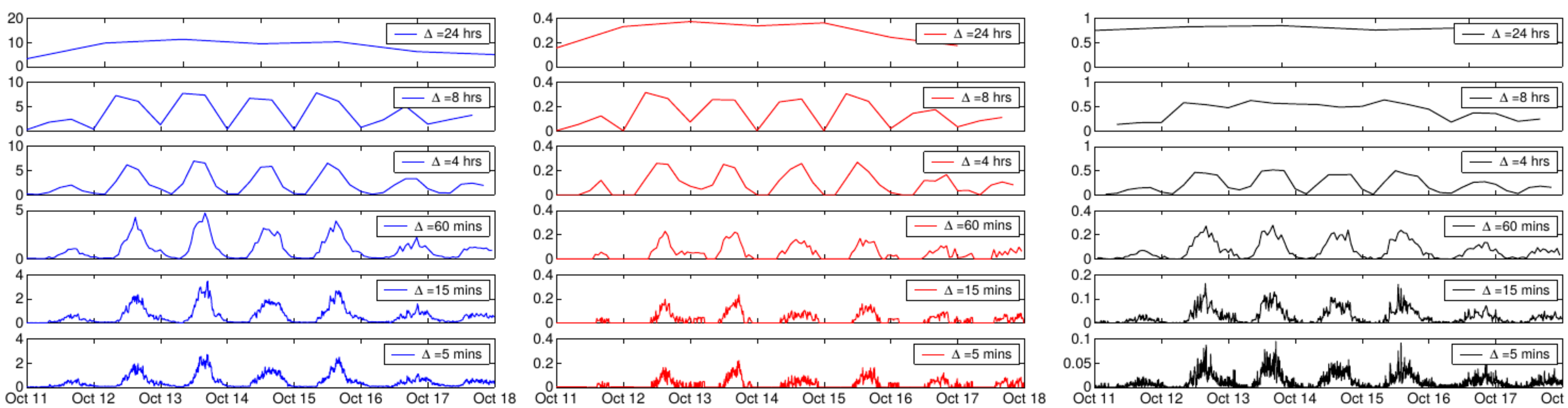
Window size?

- MIT reality mining project: high resolution proximity data
- Snapshots: $A^{(1)}, A^{(2)}, \ldots, A^{(T)}$ Time window: Δ
- Adjacency correlation:

$$\gamma_j = \frac{\sum_{i \in N(j)} A_{i,j}^{(x)} \; A_{i,j}^{(y)}}{\sqrt{\left(\sum_{i \in N(j)} A_{i,j}^{(x)}\right) \left(\sum_{i \in N(j)} A_{i,j}^{(y)}\right)}}$$

$N(j)$: set of nodes that are connected to *j* at *x* of *y*

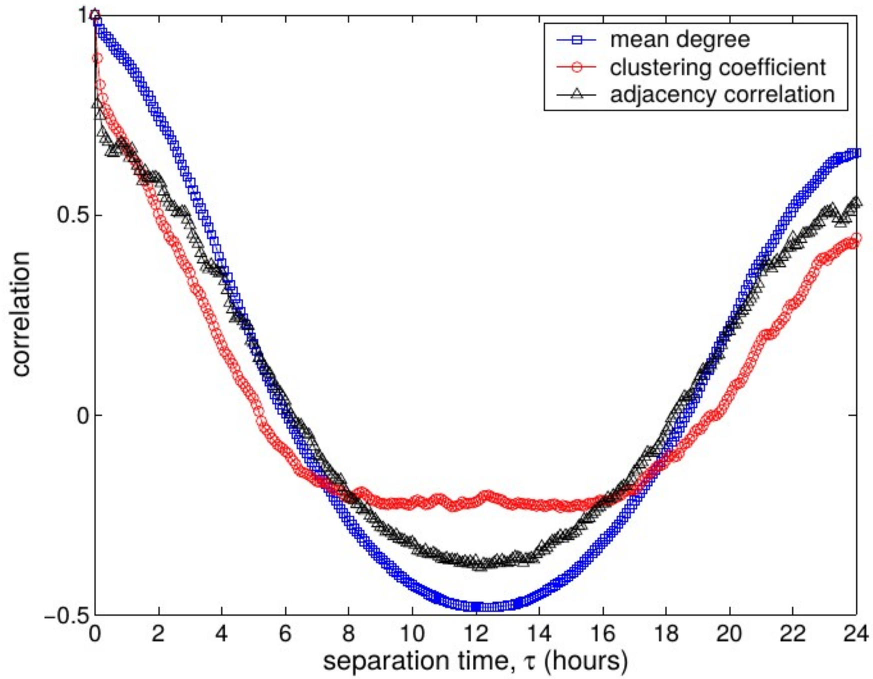
- 0 uncorrelated, 1 if the same



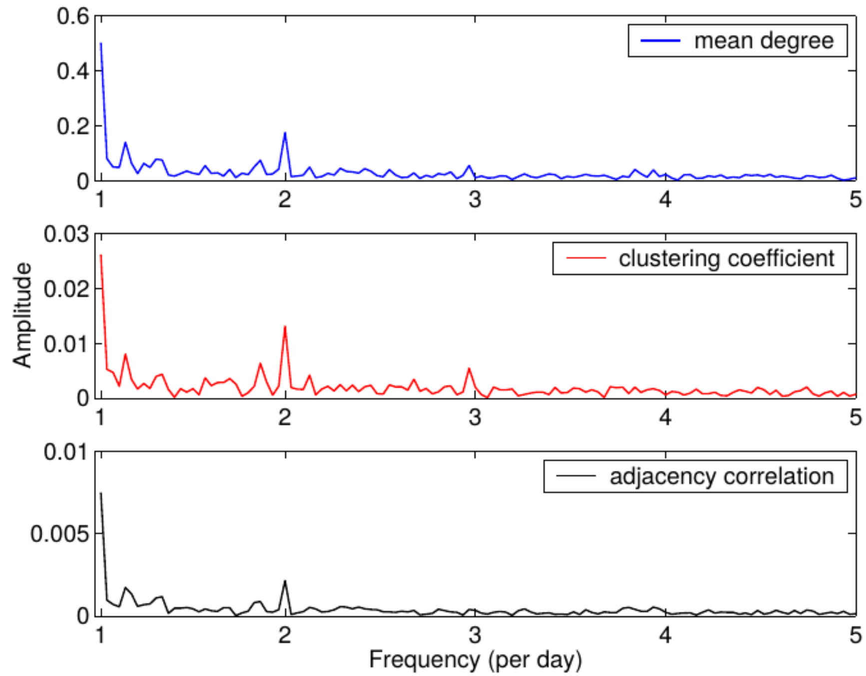
Average degree Clustering coef. 1-Adjacency corr.

Clauset and Eagle (2012)

Window size?



Time series autocorrelation
for Δ=5 min



Fourrier spectrum

Driven by periodic patterns → Sampling rate should be twice the highest frequency

$\Delta = 4$ hours

Clauset and Eagle (2012)

Path-based measures

Path-based measures

Time-respecting paths:

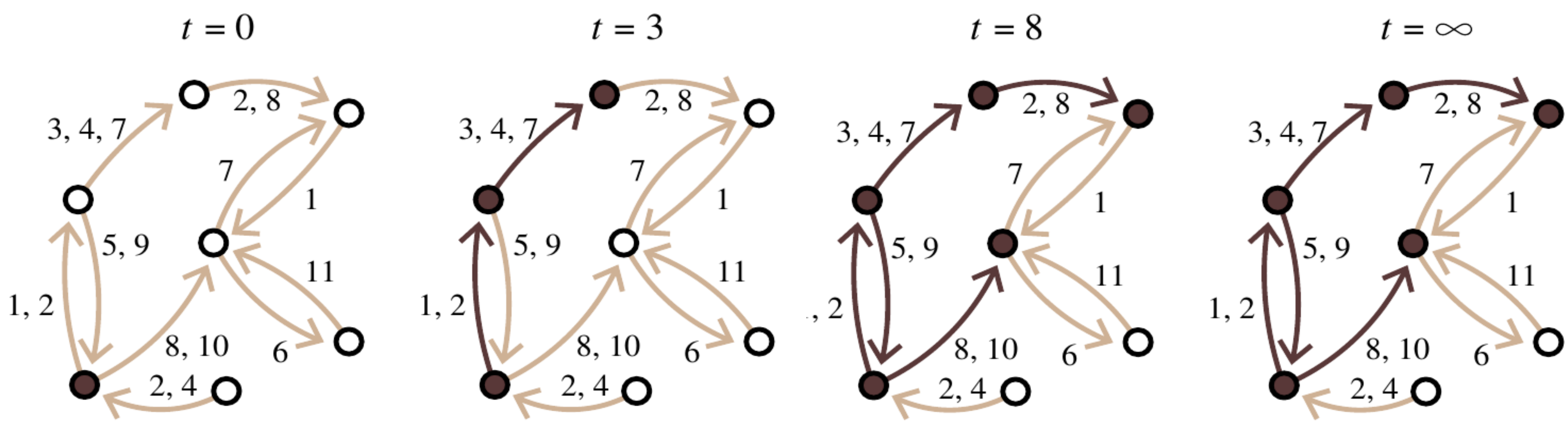
- Takes into account the temporal order and timing of contacts.

$$\{(i,k,t_1),(k,l,t_2),\ldots,(p,j,t_n)\} \qquad t_1<t_2<\ldots<t_n$$

- Optional: maximum wait time

Properties:

- **No reciprocity:** path $i \to j$ does not imply path $j \to i$.
- **No transitivity:** path $i \to j$ and path $j \to k$ does not imply path $i \to j \to k$.
- **Time dependence:** path $i \to j$ that starts at t does not imply paths at $t' > t$.



P. Holme, Physical Review E 71.4 (2005)

Path-based measures

Observation window $[t_0, t_1]$

Influence set of node *i*

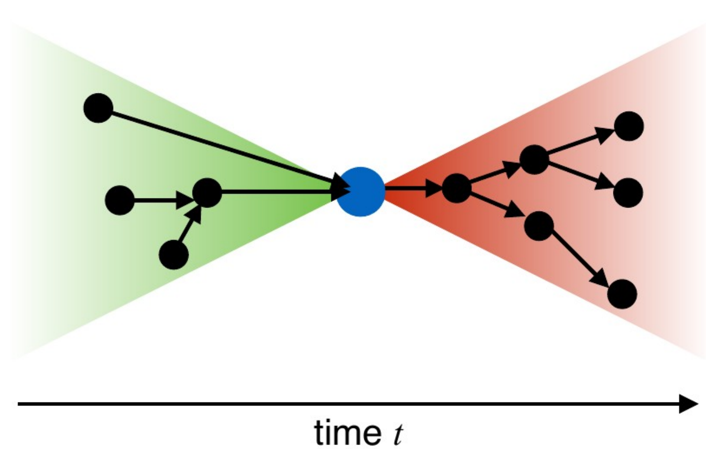
- Nodes that can be reached from node *i* within the observation window.
- Reachability ratio *f*: fraction of nodes that can be reached

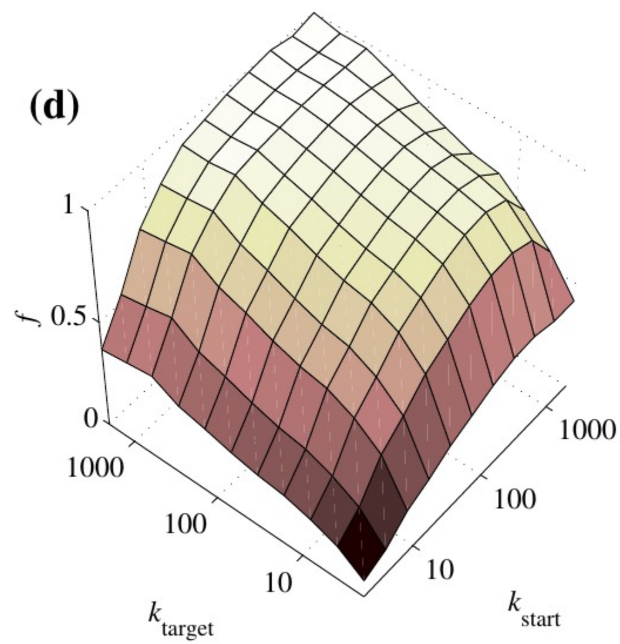
Source set of node *i*

- Nodes from node *i* is reached within the observation window.

Reachability ratio

- Fraction of node pairs *(i,j)* such that path $i \rightarrow j$ exists.





P. Holme, Physical Review E 71.4 (2005)

Path-based measures

Temporal path length – Duration

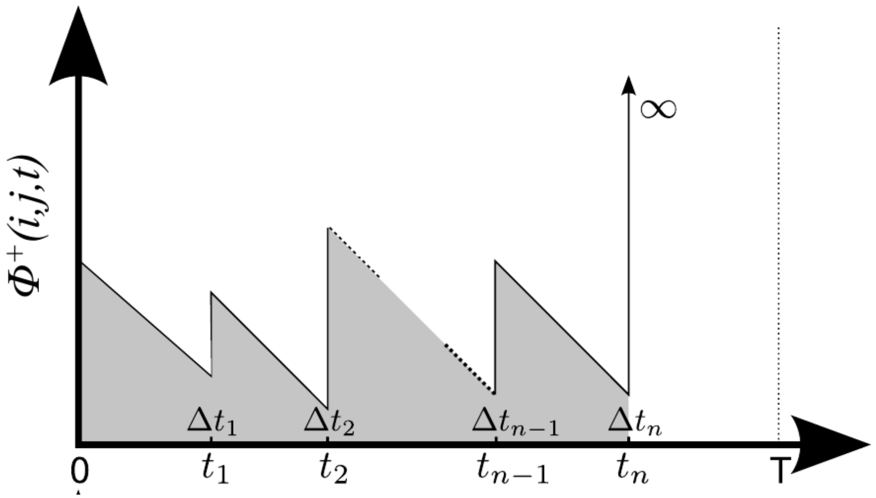
- Duration = $t_n - t_0$

Temporal distance – Latency

- $\Phi^{+}{}_{i,t}(j)$ the shortest (fastest) path duration from *i* to *j* starting at *t.*

Information latency

- $\lambda_{i,t}(j)$ the age of the information from *j* to *i* at *t*



- End of the observation window: paths become rare.
- Solution: periodic boundary, throw away end

R. K. Pan, and J. Saramäki.Physical Review E 84.1 (2011)

Path-based measures

Strongly connected component

- All node pairs are connected in both directions within *T*.

Weakly connected component

- All node pairs are connected in at least one direction within *T*.

Temporal betweenness centrality

- Static:

$$b(i) = \frac{\sum_{i \neq j \neq k} v_{jk}(i)}{\sum_{j \neq k} v_{jk}}$$

Temporal:

$$b(i,T) = \frac{\sum_{i \neq j \neq k} v_{jk}(i,T)}{\sum_{j \neq k} v_{jk}(T)}$$

Etc.

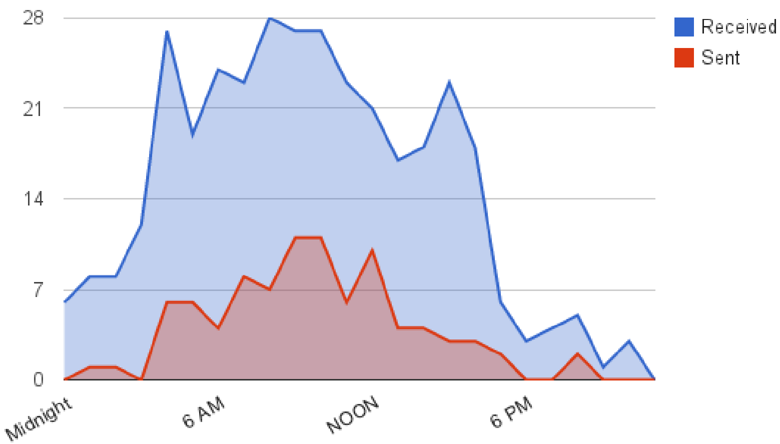
R. K. Pan, and J. Saramäki.Physical Review E 84.1 (2011)

Temporal heterogeneity

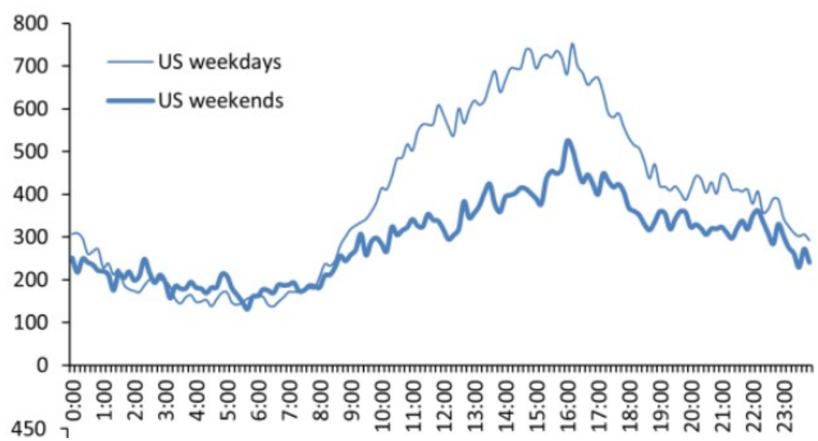
Source 1: Periodic patterns

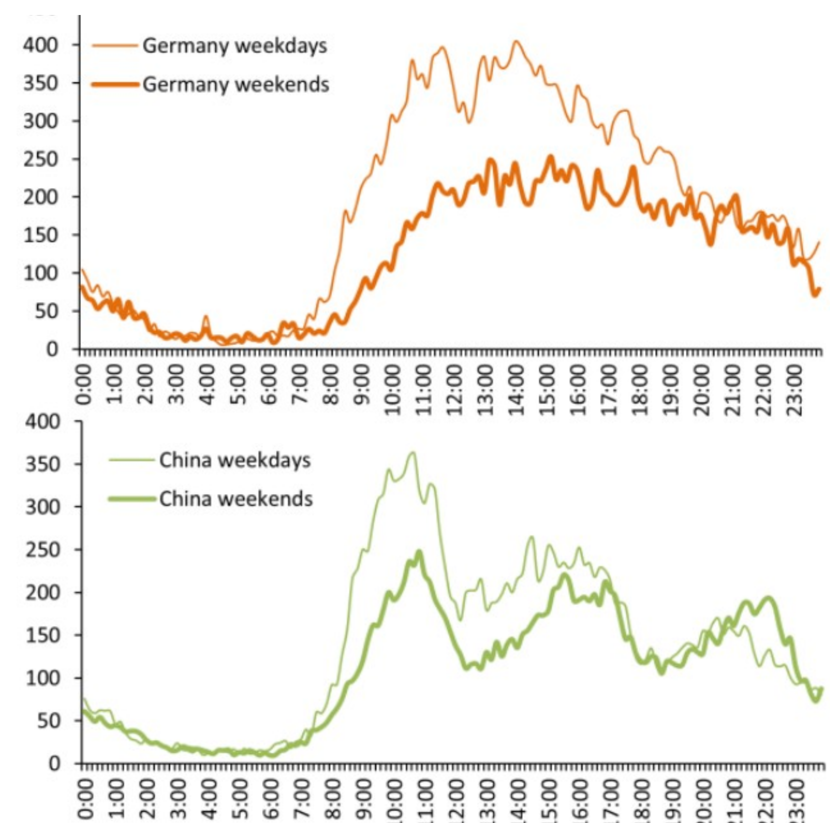
For example, circadian rythm

- My email usage



- Scientists work schedule

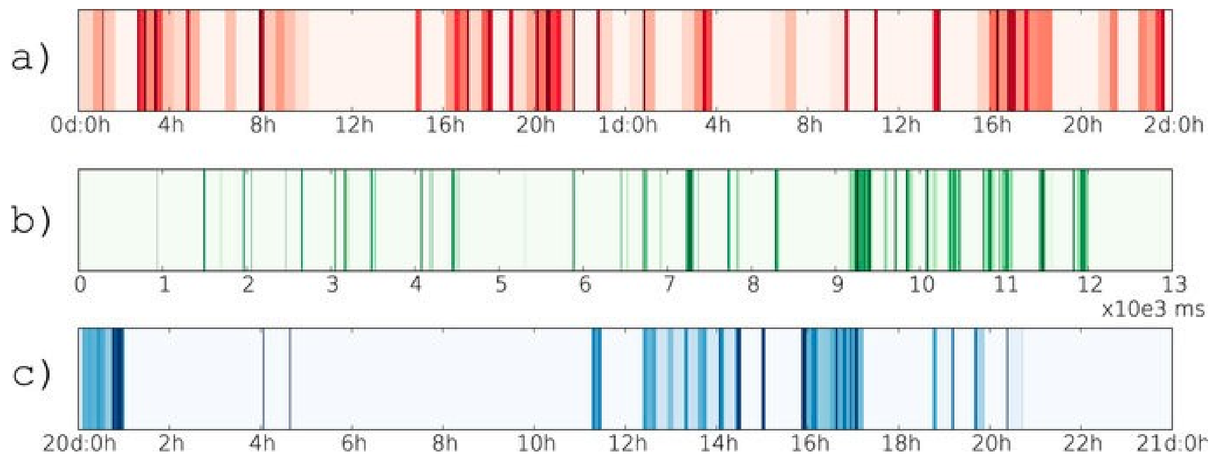




Wang et al, *Journal of infometrics* (2012)

Source 2: Burstiness

- Humans and many natural phenomena show heterogeneous temporal behavior on the individual level.
- Switching between periods of low activity and high activity bursts.
- Sign of correlated temporal behavior



Earthquakes in Japan

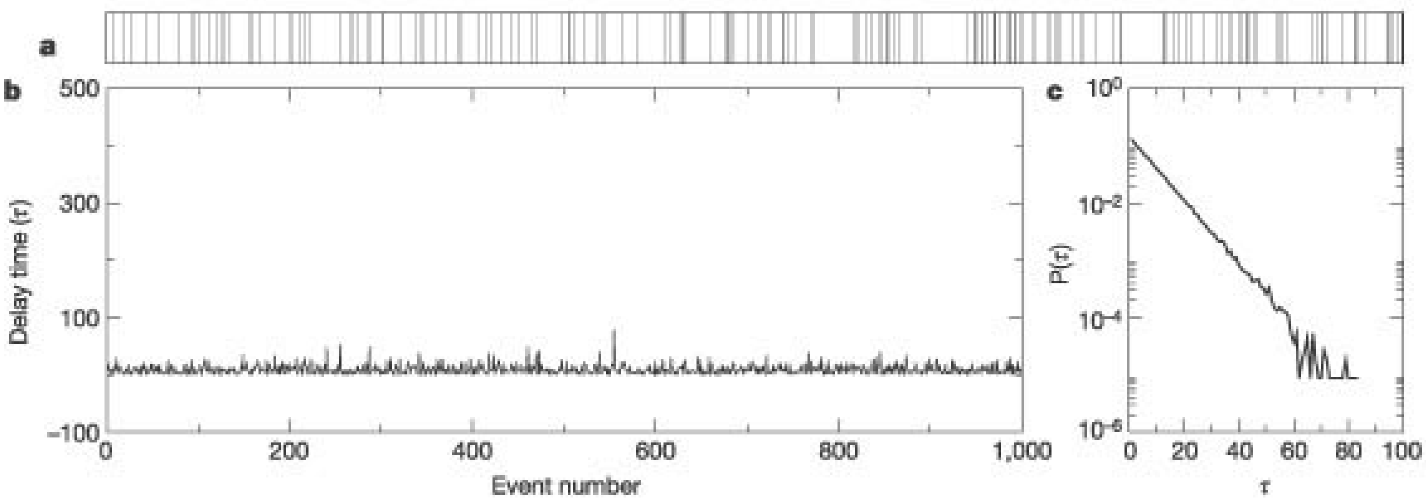
Neuron firing

Phone calls

Karsai et al, *Scientific Reports* (2011)

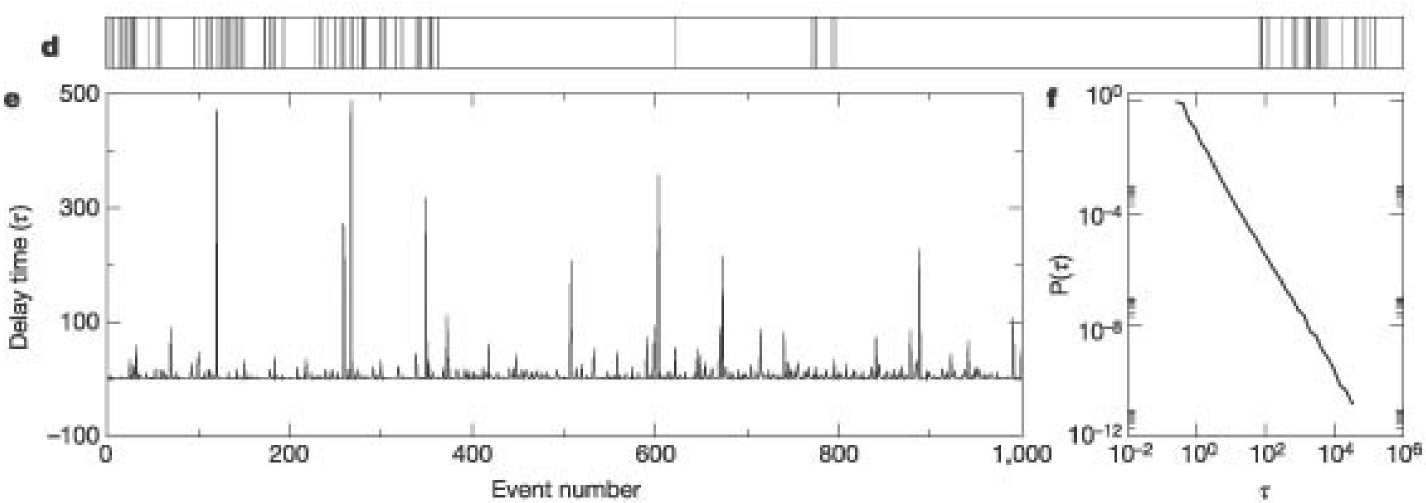
Source 2: Burstiness

- Sign of correlated temporal behavior
- Reference: Poisson process, events uncorrelated



$$P(\tau) \sim e^{-q\tau}$$

Poisson process



$$P(\tau) \sim \tau^{-\gamma}$$

Bursty signal

Barabási, *Nature* (2005)

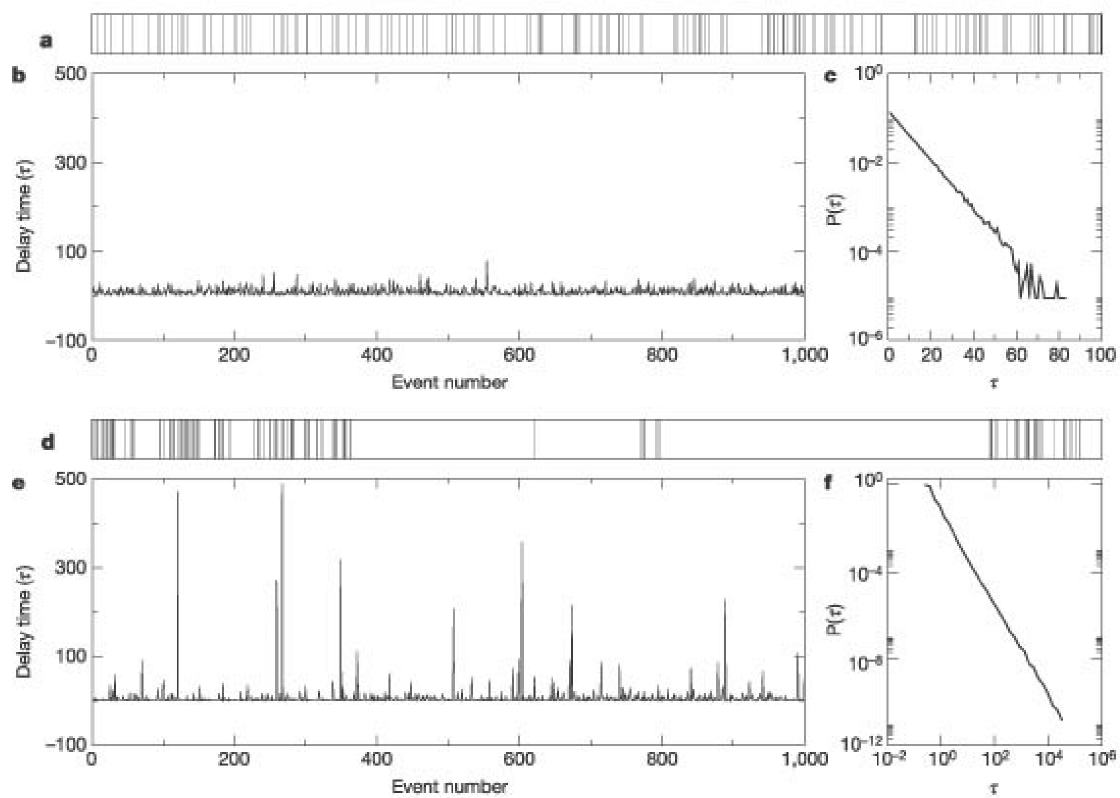
Source 2: Burstiness

- Measure of burstiness: $$B \equiv \frac{(\sigma_\tau/m_\tau - 1)}{(\sigma\tau/m_\tau + 1)} = \frac{(\sigma_\tau - m_\tau)}{(\sigma_\tau + m_\tau)}$$

m_τ - average inter-event time

σ_τ - STD of inter-event time

B=-1	B=0	B=1
Max. regular	Poisson	Max bursty



Barabási, *Nature* (2005)

Possible explanation for burstiness

- Executing tasks based on prioirity
- *L* types of tasks, one each (e.g. work, family, movie watching,...)
- Each task *i* has a priority x_i draw uniformly from [0,1]
- Each timestep one task is executed, probability of choosing *i*:

$$\Pi(i) = \frac{x_i^\gamma}{\sum_{j=1}^{L} x_j^\gamma}$$

- And a new task is added of that type

$\gamma = 0$

Random

$$P(\tau) \sim e^{-\tau}$$

$\gamma = \infty$

Deterministic

$$P(\tau) \sim \tau^{-1}$$

Barabási, *Nature* (2005)

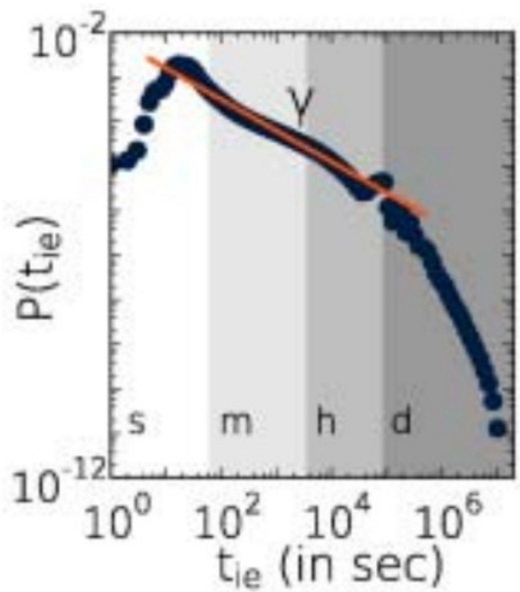
Is inter-event time power-law?

$$P(\tau) \sim e^{-q\tau}$$

Poisson process

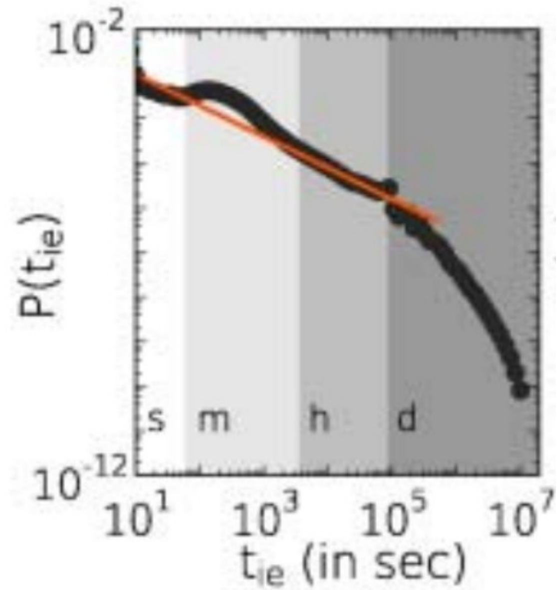
$$P(\tau) \sim \tau^{-\gamma}$$

Bursty signal



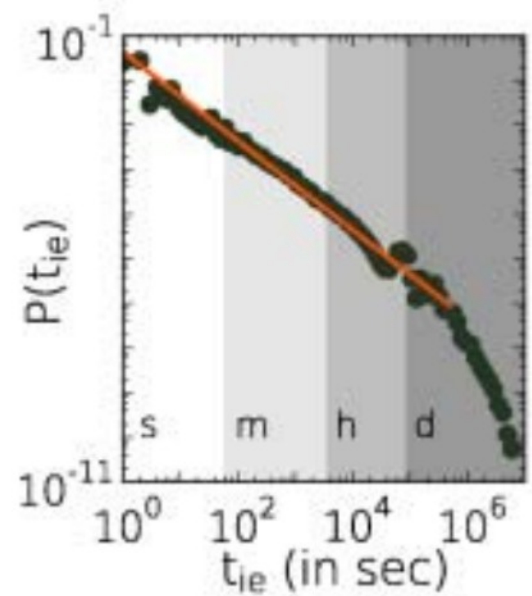
Phone calls

$\gamma \approx 0.7$



Text message

$\gamma \approx 0.7$



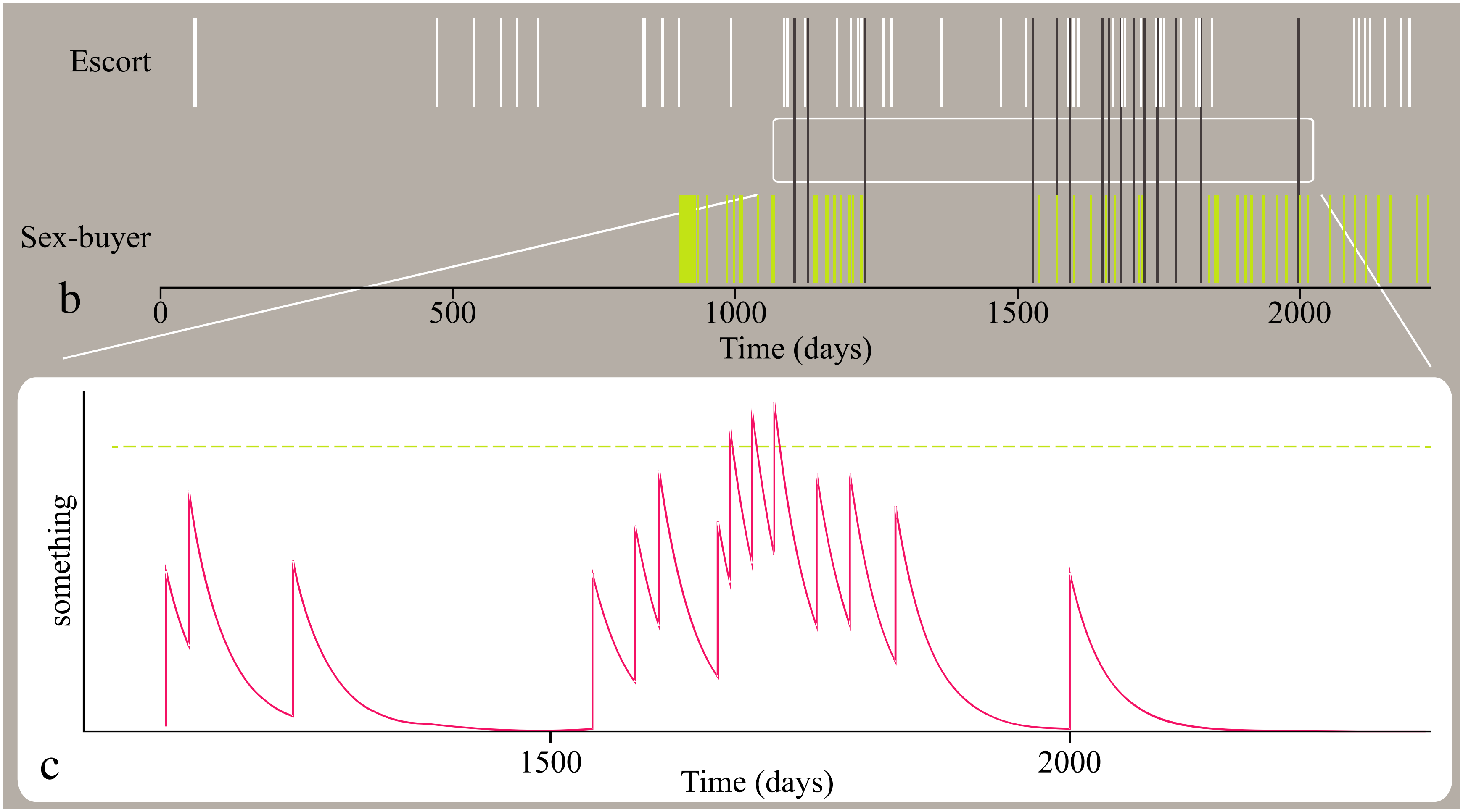
Email

$\gamma \approx 1.0$

- Is this a power-law? Definitely not exponential.

Karsai et al, *Scientific Reports* (2012)

Does burstiness matter?



L. EC Rocha, F. Liljeros, and P. Holme. PNAS 107.13 (2010)

Processes and null models

Structure and dynamics

- So far: various measures to characterize network
- How does structure affect processes on the network?
- Possibility:

  1) Generate model networks with given parameters.

  2) Run dynamics model → measure outcome.

  3) Scan parameters.
- Models from scratch: many parameters to set → few models, no dominant

- Instead:

  1) Take empirical network.

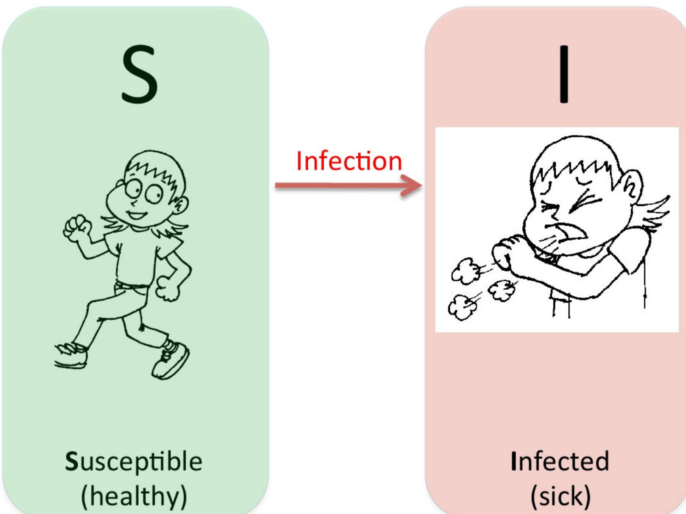
  2) Remove correlations by randomization.

  3) Run dynamics model → measure outcome.

L. EC Rocha, F. Liljeros, and P. Holme. PNAS 107.13 (2010)

Dynamics

- Model: SI model



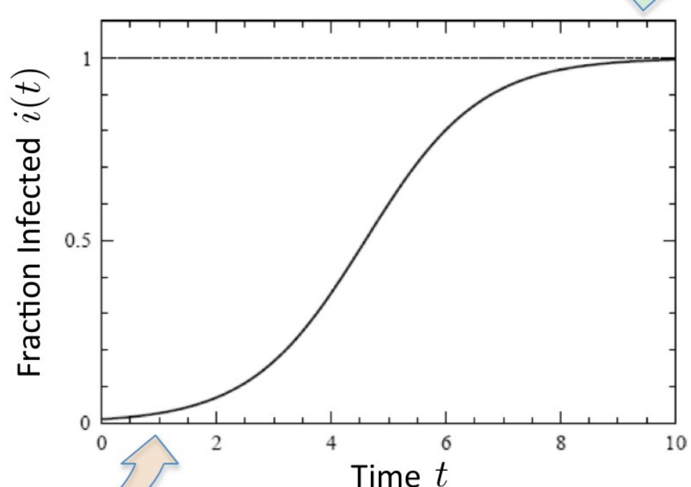
- Infection rate: β

If $i(t)$ is small,

$$\frac{di(t)}{dt} \approx \beta\langle k\rangle$$

$$i(t) \approx i_0 e^{\beta\langle k\rangle t}$$

exponential outbreak



As $i(t) \rightarrow 1$

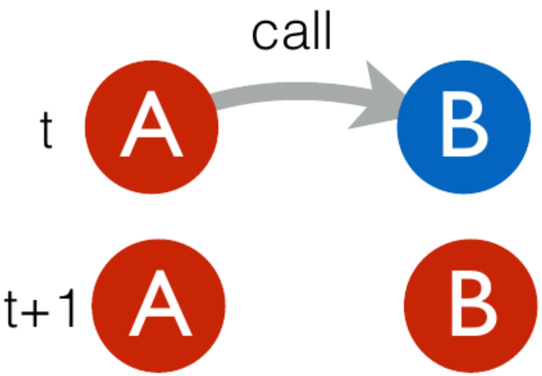
$$\frac{di(t)}{dt} \rightarrow 0$$

saturation

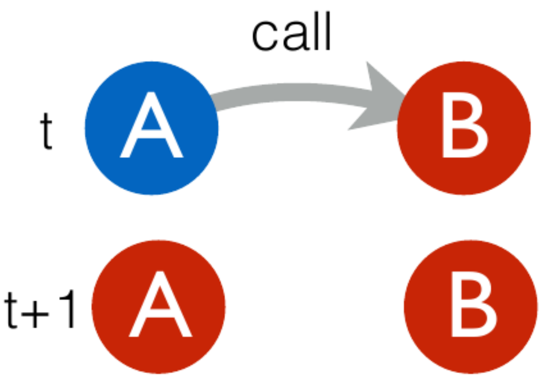
If connected

Dynamics

- Model: SI model on a temporal network
- Simplest model of information spreading
- Infection only spreads along active contacts
- Infection can spread both ways
- Infection rate $\beta = 1$
- Single seed at *t=0*



- Mobile call data as underlying network

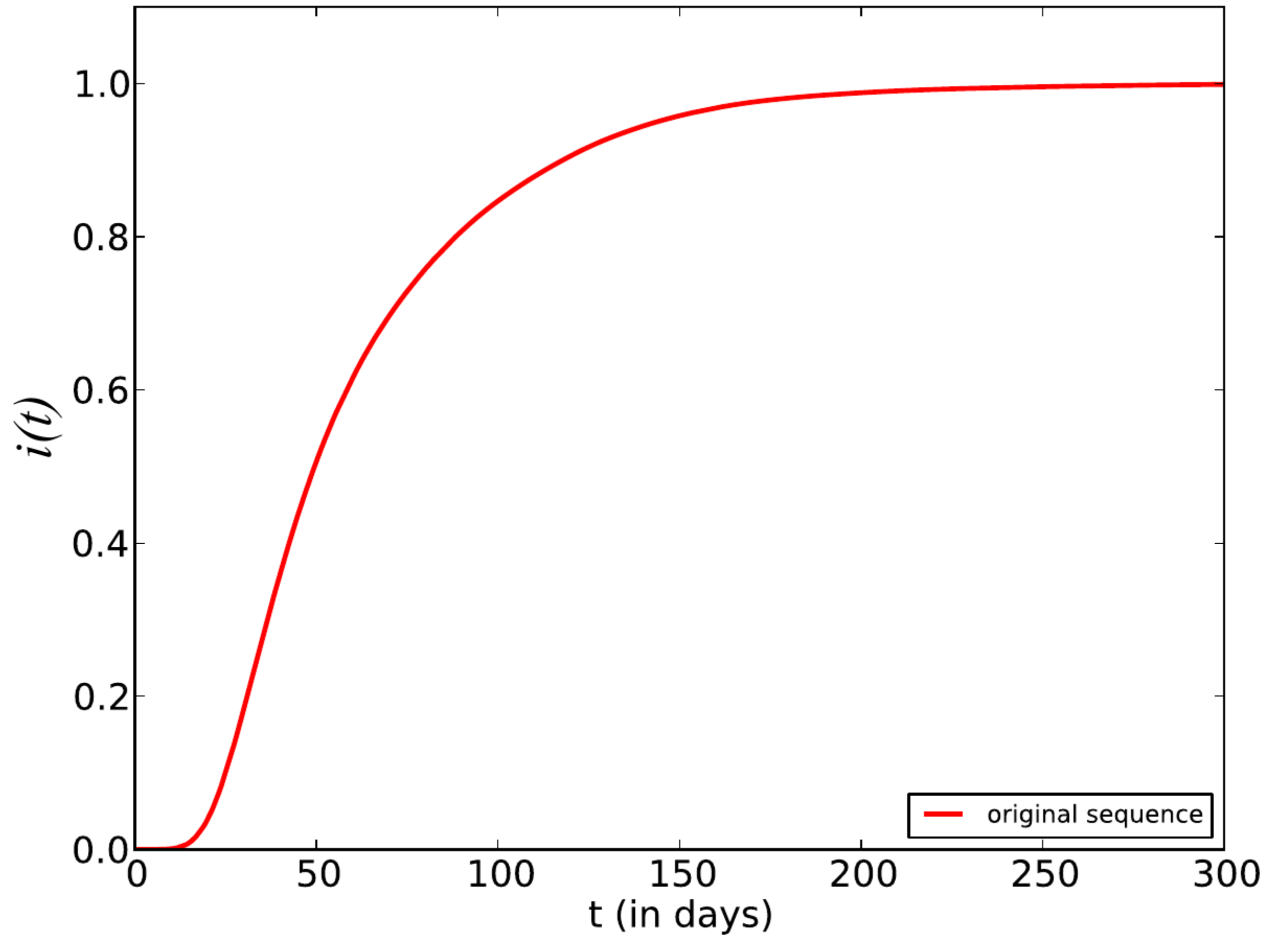


Slide adopted from Márton Karsai

Original temporal net

- We run SI model and measure $i(t) = \frac{I(t)}{N}$

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7



- Now what? Is this because burstiness, community structure?

Slide adopted from Márton Karsai

Properties of the original

- **Bursty dynamics (BD)**
  Heterogeneous inter-event time distribution

- **Community structure (CS)**
  Densely connected subgroups
  (Any other structure beyond degree distribution)

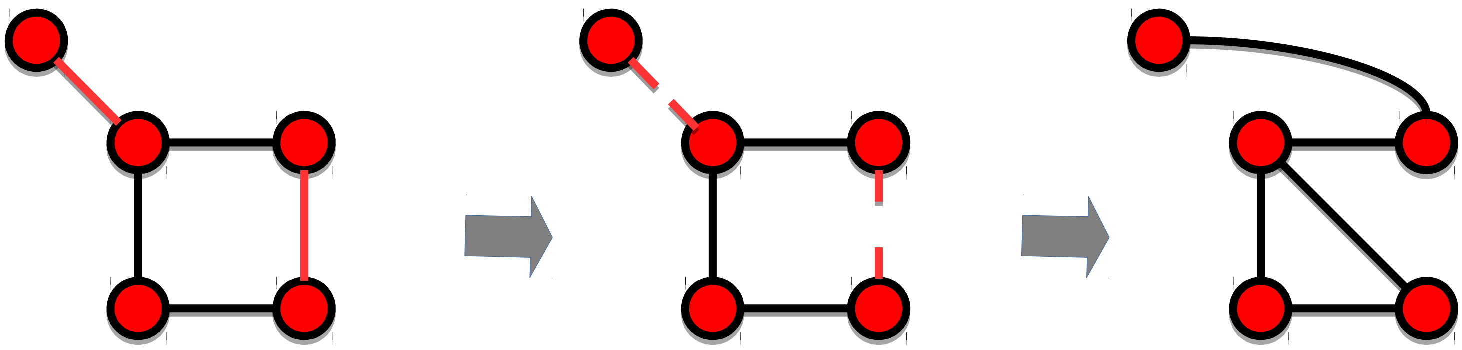
	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7

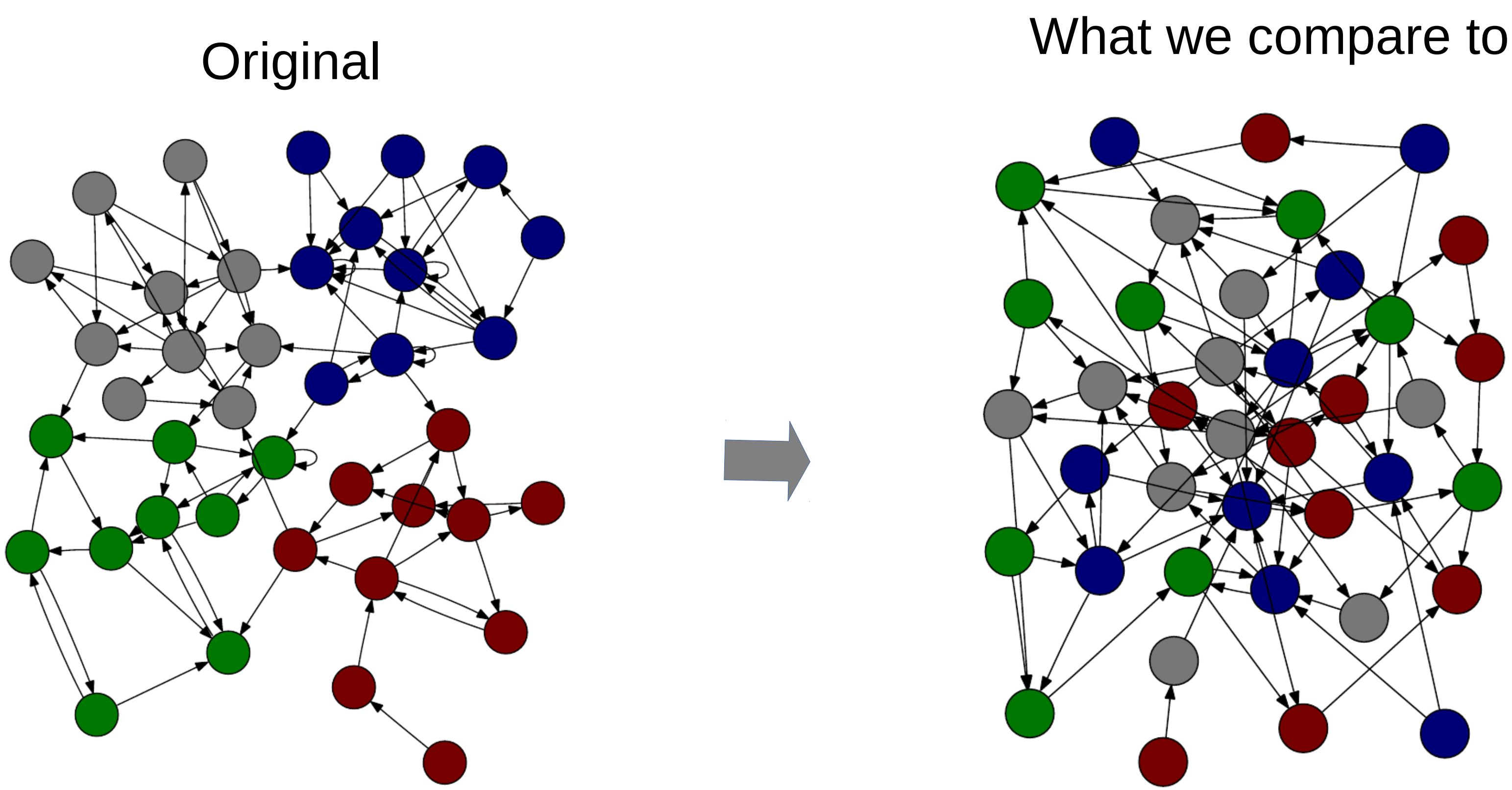
- **Link-link correlations (LL)**
  Causality between consecutive calls

- **Weight-topology correlation (WT)**
  Strong ties within local communities, weak ties connect different communities
  Weight = total call time
  s = strength of a node = some of adjacent link weights
  Onnela, J-P., et al. "Structure and tie strengths in mobile communication networks." PNAS 104.18 (2007).

Recap from community detection

- Randomization to remove community structure of a static network



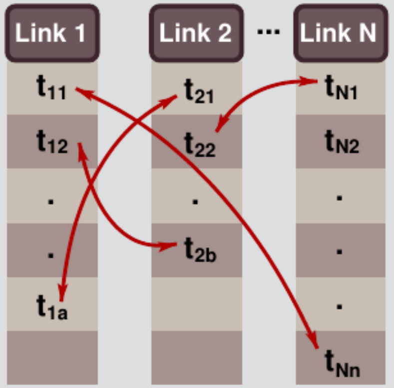


Randomization 1: temp. config. model

- Degree preserved randomization to remove **CS** and **WT**
- Shuffle event times to remove **BD** and **LL**

Shuffling

- Shuffle the event times of calls and destroy temporal heterogeneities
- keep P(w), P(k), P(s), w-top correlations
- destroy $P(t_{ie})$, link-link correlations



Time shuffling

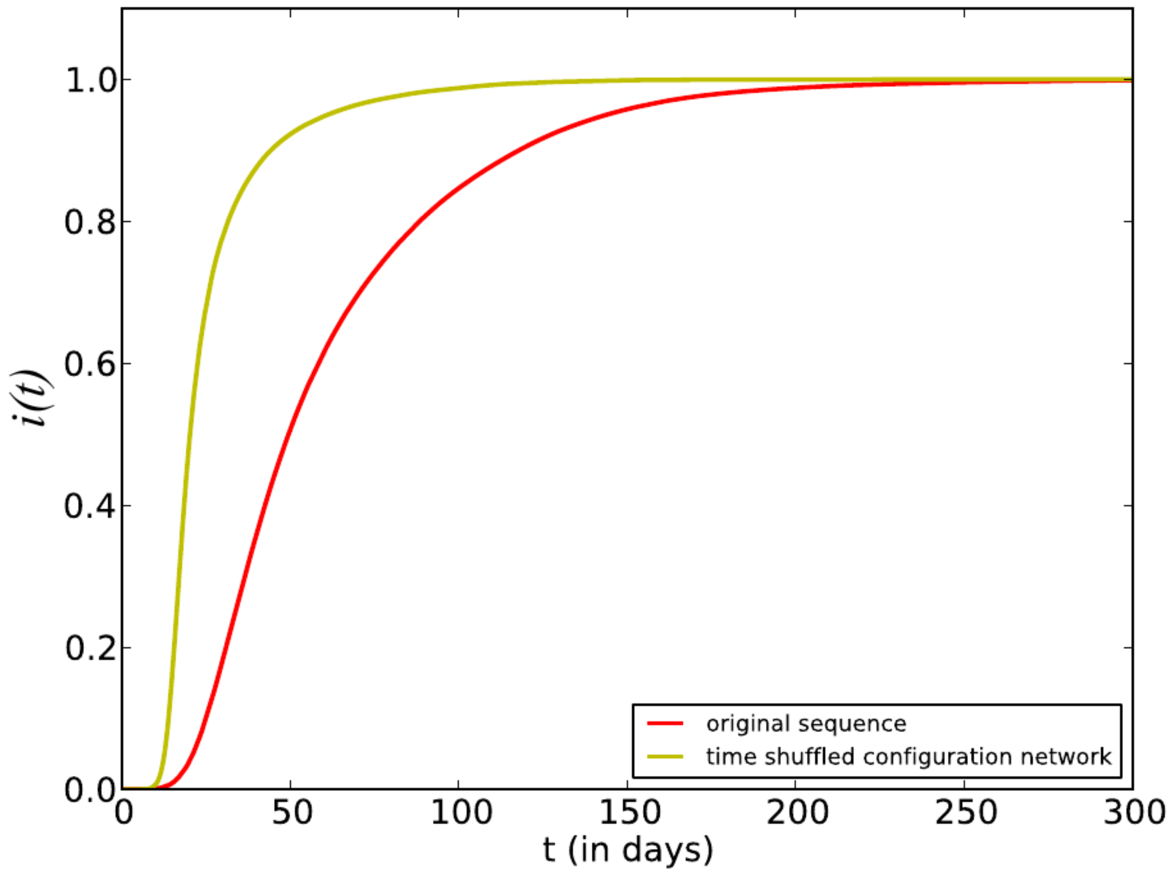
Slide adopted from Márton Karsai

Randomization 1: temp. config. model

- Degree preserved randomization to remove **CS** and **WT**
- Shuffle event times to remove **BD** and **LL**
- **No correlation left.**

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7
TimeConf	✕	✕	✕	✕	16,4

→ **Correlations slow the spread of information.**

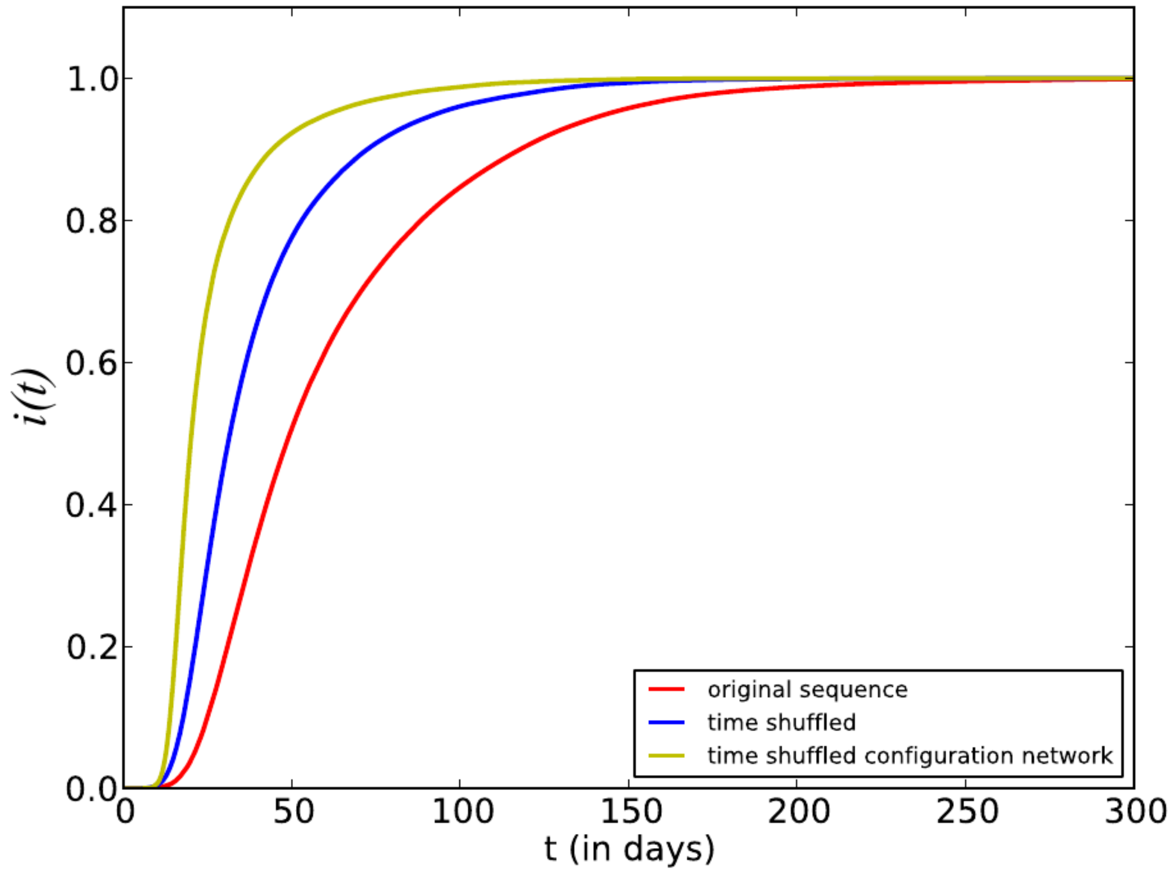


Slide adopted from Márton Karsai

Randomization 2: time shuffled network

- Shuffle event times to remove **BD** and **LL**
- **CS** and **WT** remain.

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7
TimeConf	×	×	×	×	16,4
Time	✓	×	×	✓	22,9



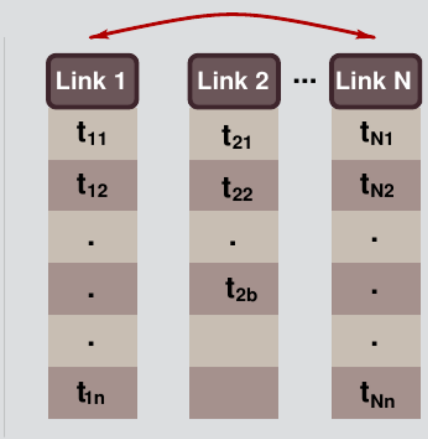
Slide adopted from Márton Karsai

Randomization 2: time shuffled network

- Shuffle event sequences to remove **WT** and **LL**

Shuffling

- Change complete call sequences of individuals regardless of their edge weight
- keep P(w), P(k), $P(t_{ie})$
- destroy P(s), link-link correlations, w-top correlations



Eq.-w. link-seq. shuffling

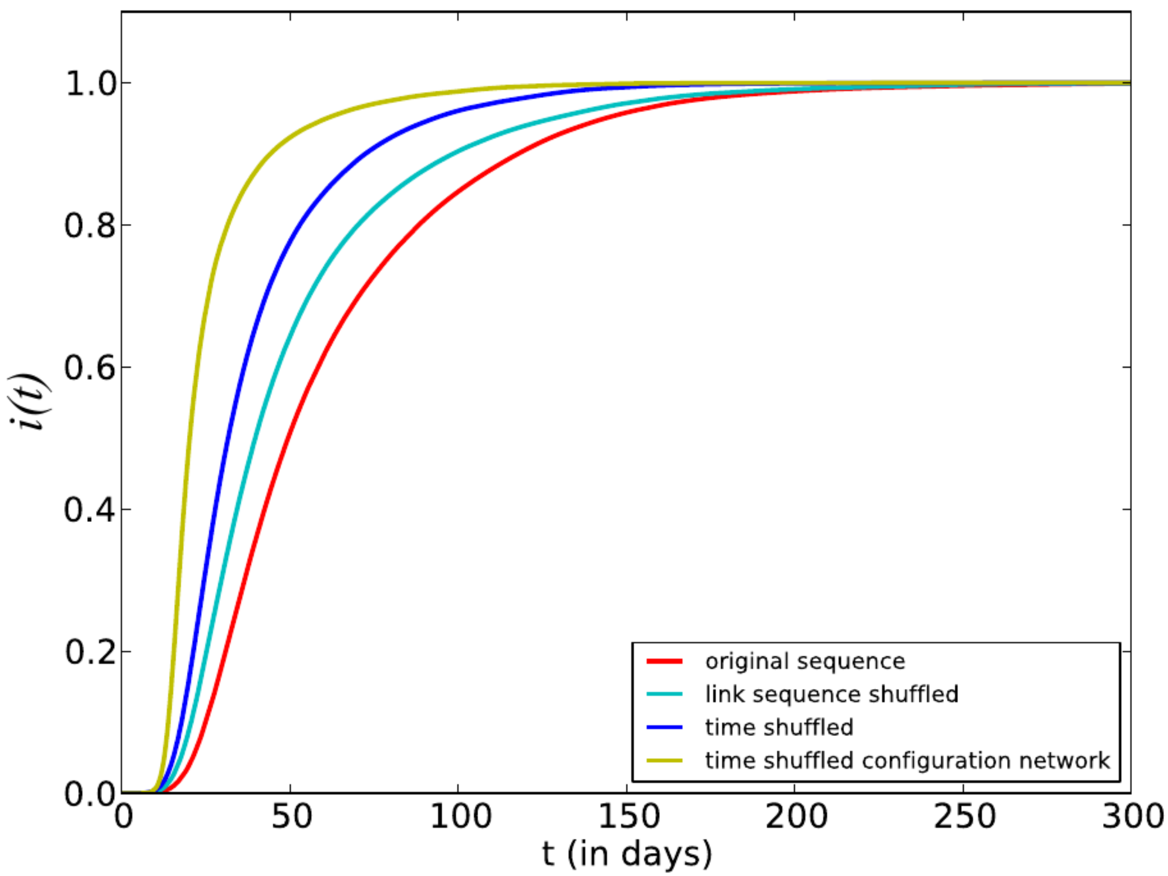
- **CS** and **BD** remain.

Slide adopted from Márton Karsai

Randomization 3: time seq. shuffled

- Shuffle event sequences to remove **WT** and **LL**
- **CS** and **BD** remain.

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7
TimeConf	×	×	×	×	16,4
Time	✓	×	×	✓	22,9
Link	×	✓	×	✓	27,5

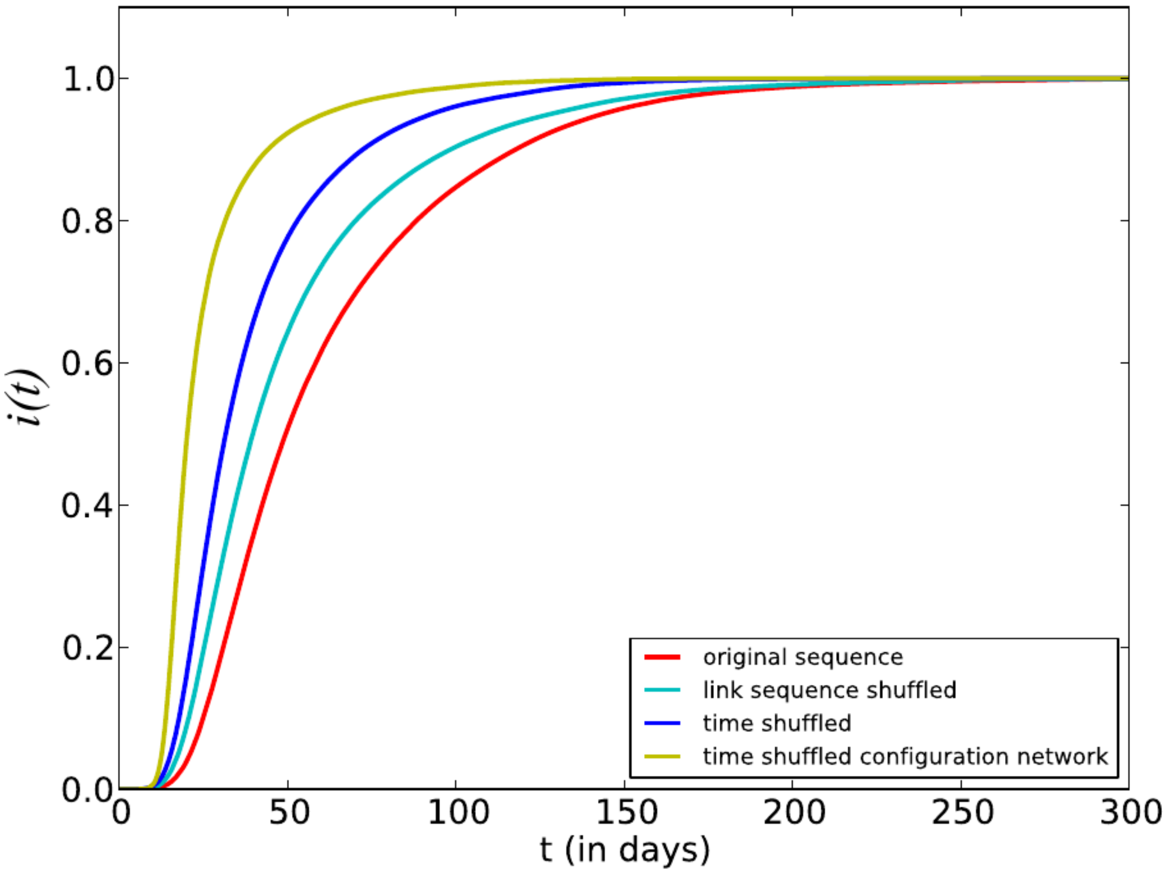


Slide adopted from Márton Karsai

Randomization 3: time seq. shuffled

- Shuffle event sequences to remove **WT** and **LL**
- **CS** and **BD** remain.

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7
TimeConf	×	×	×	×	16,4
Time	✓	×	×	✓	22,9
Link	×	✓	×	✓	27,5

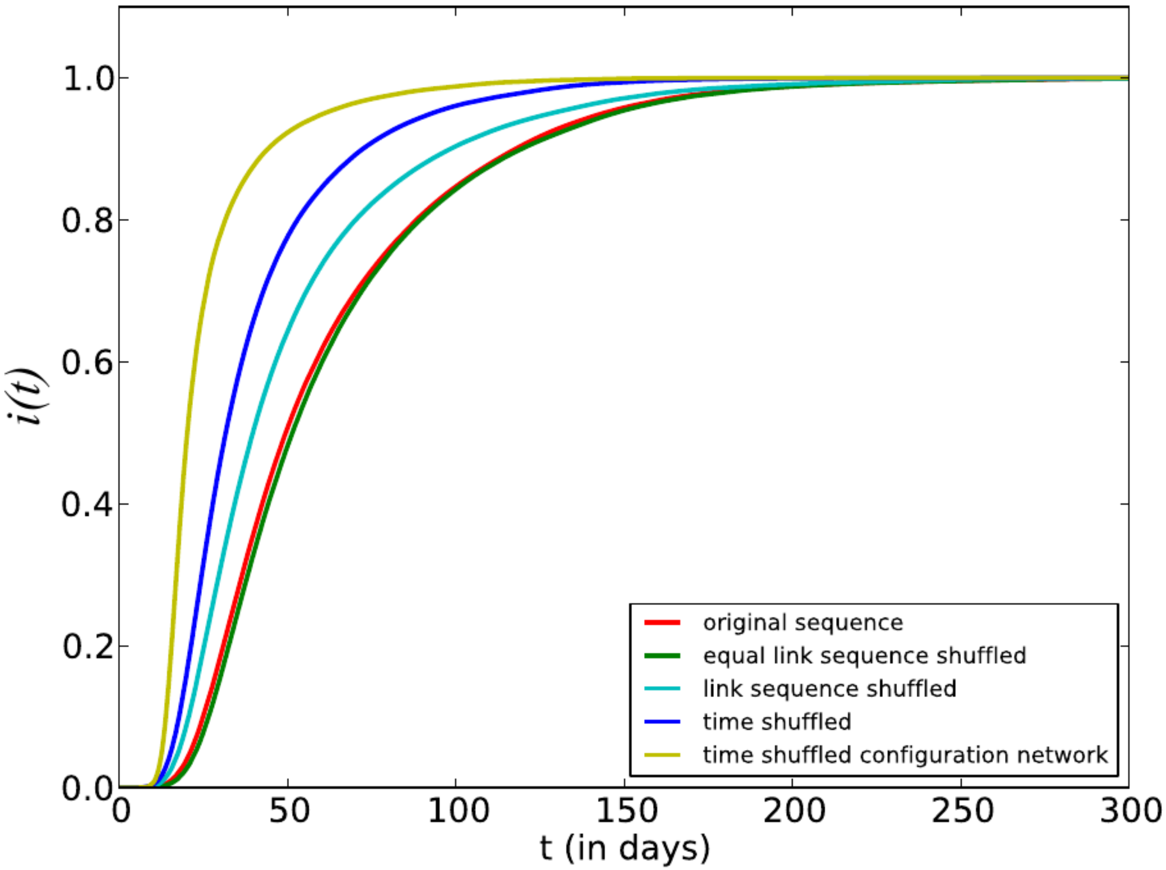


Slide adopted from Márton Karsai

Rand. 4: equal link time seq. shuffled

- Shuffle event sequences if they have the same weight to remove **LL**
- **CS**, **WT** and **BD** remain.

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7
TimeConf	×	×	×	×	16,4
Time	✓	×	×	✓	22,9
Link	×	✓	×	✓	27,5
Equal link	✓	✓	×	✓	35,3



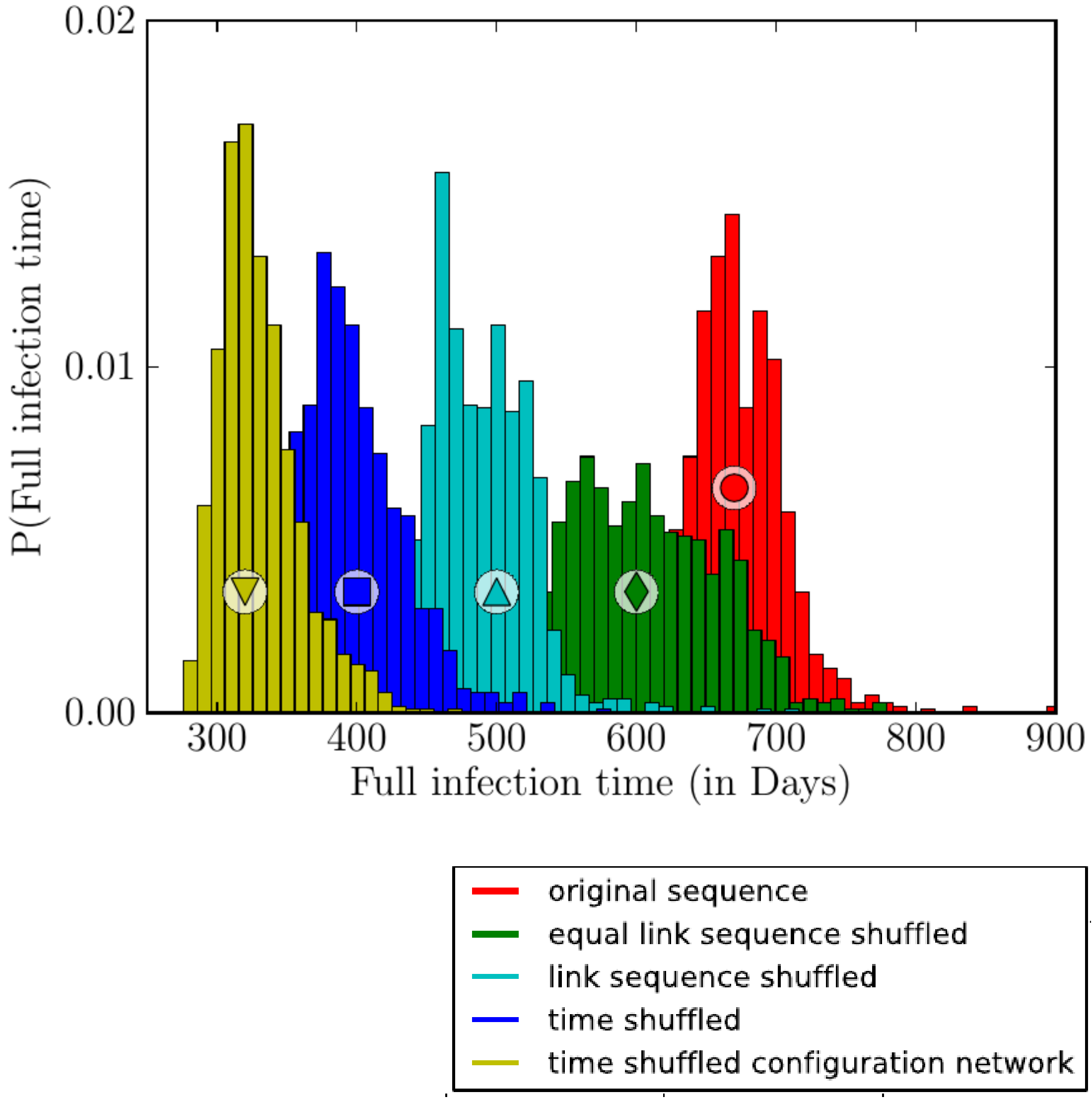
→ **Multi-link processes slightly accelerate the spread.**

Slide adopted from Márton Karsai

Long time behavior

- Distribution of complete infection time
- Evidence of effect of correlations in the late time stage.
- Multi-link correlations have contrary effect compared to early stage
- **WT** and **BD** are the main factors in slowing down

	WT	BD	LL	CS	25%m
Original	✓	✓	✓	✓	33,7
TimeConf	×	×	×	×	16,4
Time	✓	×	×	✓	22,9
Link	×	✓	×	✓	27,5
Equal link	✓	✓	×	✓	35,3



Slide adopted from Márton Karsai

Summary

- Timescale of dynamics and changes in network structure comparable

  → Temporal networks

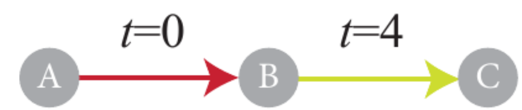
- Time respecting paths profound effect on spreading
- Temporal inhomogeneities: circadian rhythm and burstiness
- Measures more involved, computationally more difficult

Temporal motifs

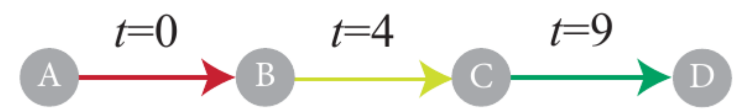
Definitions

Δt=5

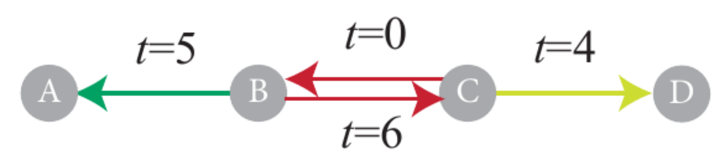
- **Δt adjacent** are two events if they share at least one node and are performed in *Δt*



- **Δt connected** are two events if there exists a sequence of events $e_i = e_{k0} e_{k1} e_{k2} \ldots e_{kn} = e_j$ such that all pairs of consecutive events are *Δt* adjacent



- **Connected temporal subgraph** consists of set of events, which are all *Δt* connected



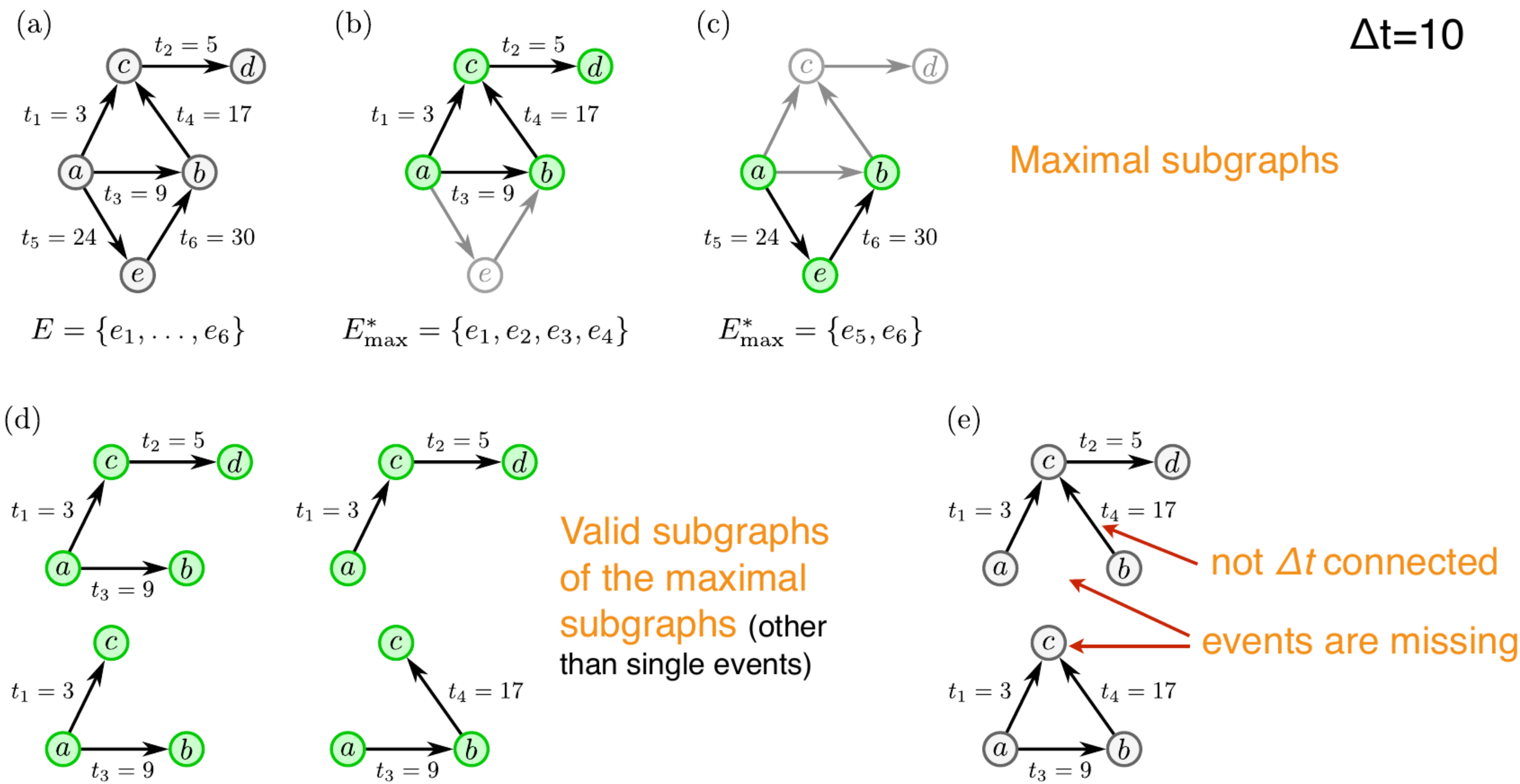
- **Valid temporal subgraph** are connected temporal subgraphs where all *Δt* connected events of each node are consecutive

- **Maximal temporal subgraph** for an event e_i is a unique maximal subgraph E^*_{max} that contains e_i and in which all event pairs are *Δt* connected

Slide adopted from Márton Karsai

Detection

Mezoscopic correlated and casual temporal structures with topological and temporal order isomorphism



Slide adopted from Márton Karsai

Algorithm

Mezoscopic correlated and casual temporal structures with topological and temporal order isomorphism

- To detect them we need to group events into equivalent classes where timing not but direction and ordering matters

1. Find all maximum connected subgraphs E^{*}_{max}
   - start from an event e_i
   - iterate forward and backward to find all Δt adjacent events
   - repeat it for all new events

2. Find all valid subgraphs E^{*}
   (this can be reduced to find all induced subgraphs of a static graph)

3. Identify the motifs for all E^{*} subgraphs
   (map to directed coloured graphs and find isomorphic structures with equivalent ordering, e.g. using the bliss algorithm (Junttila and Kaski (2007)))

Slide adopted from Márton Karsai

What to compare to?

Candidates null models

1. Time-shuffled reference: randomly redistribute event times between events
   - Destroys all temporal correlations and casual correlations

a	b	t_1	→	a	b	t_3
a	c	t_2		a	c	t_4
e	f	t_3		e	f	t_1
f	a	t_4		f	a	t_2
	…				…	

2. Time-reversed reference: read the event sequence in a reversed order
   - Destroys all casual correlations but keeps all temporal correlations

a	b	t_1	→	f	a	t_1
a	c	t_2		e	f	t_2
e	f	t_3		a	c	t_3
f	a	t_4		a	b	t_4
	…				…	

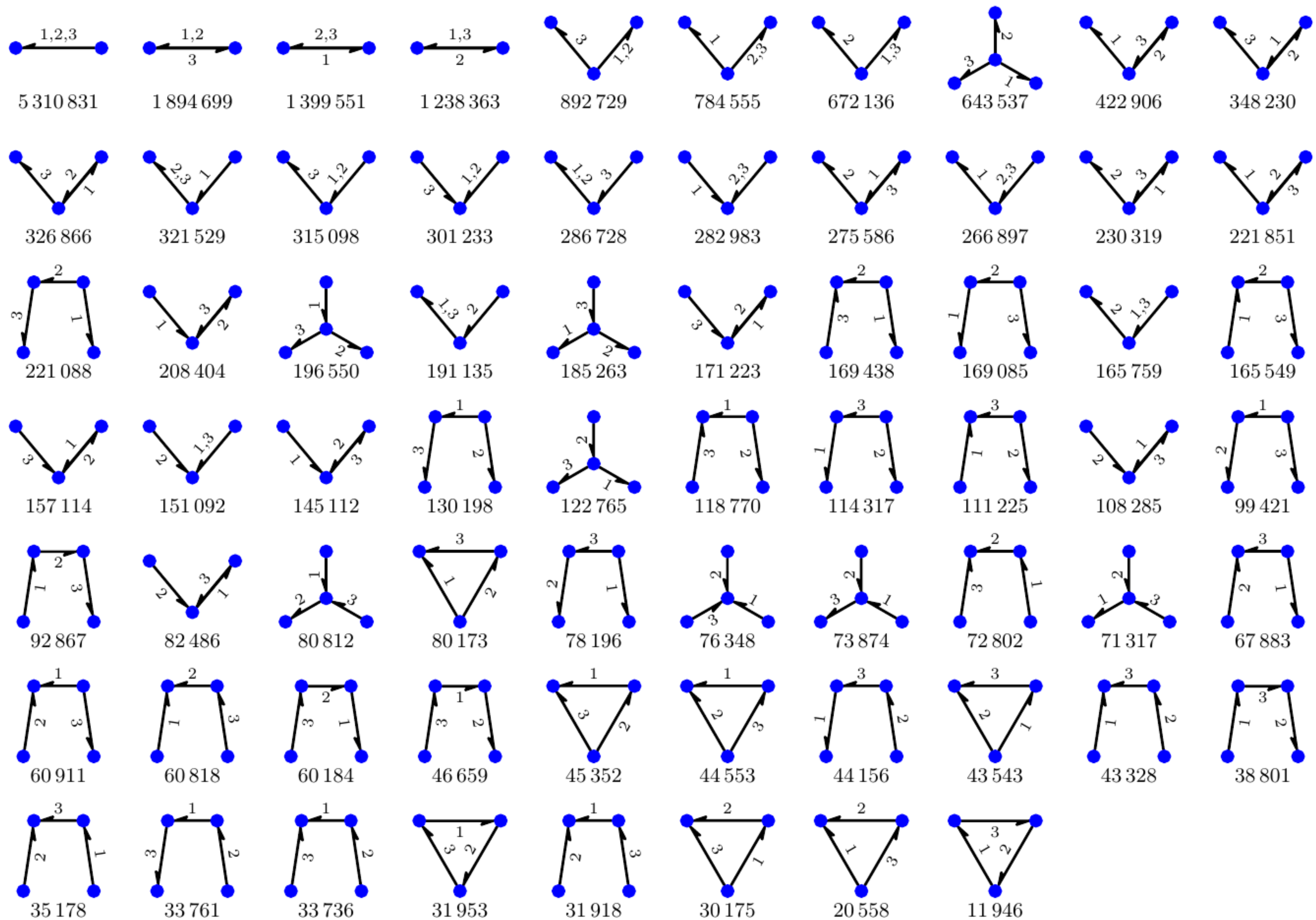
3. Self reference: compare different periods of the sequence to each other
   - Highlights seasonal dependencies

a	b	t_1	↔	d	w	t_{91}
a	c	t_2		y	a	t_{92}
e	f	t_3		a	b	t_{93}
f	a	t_4		q	s	t_{94}
	…				…	

Slide adopted from Márton Karsai

Phone call network

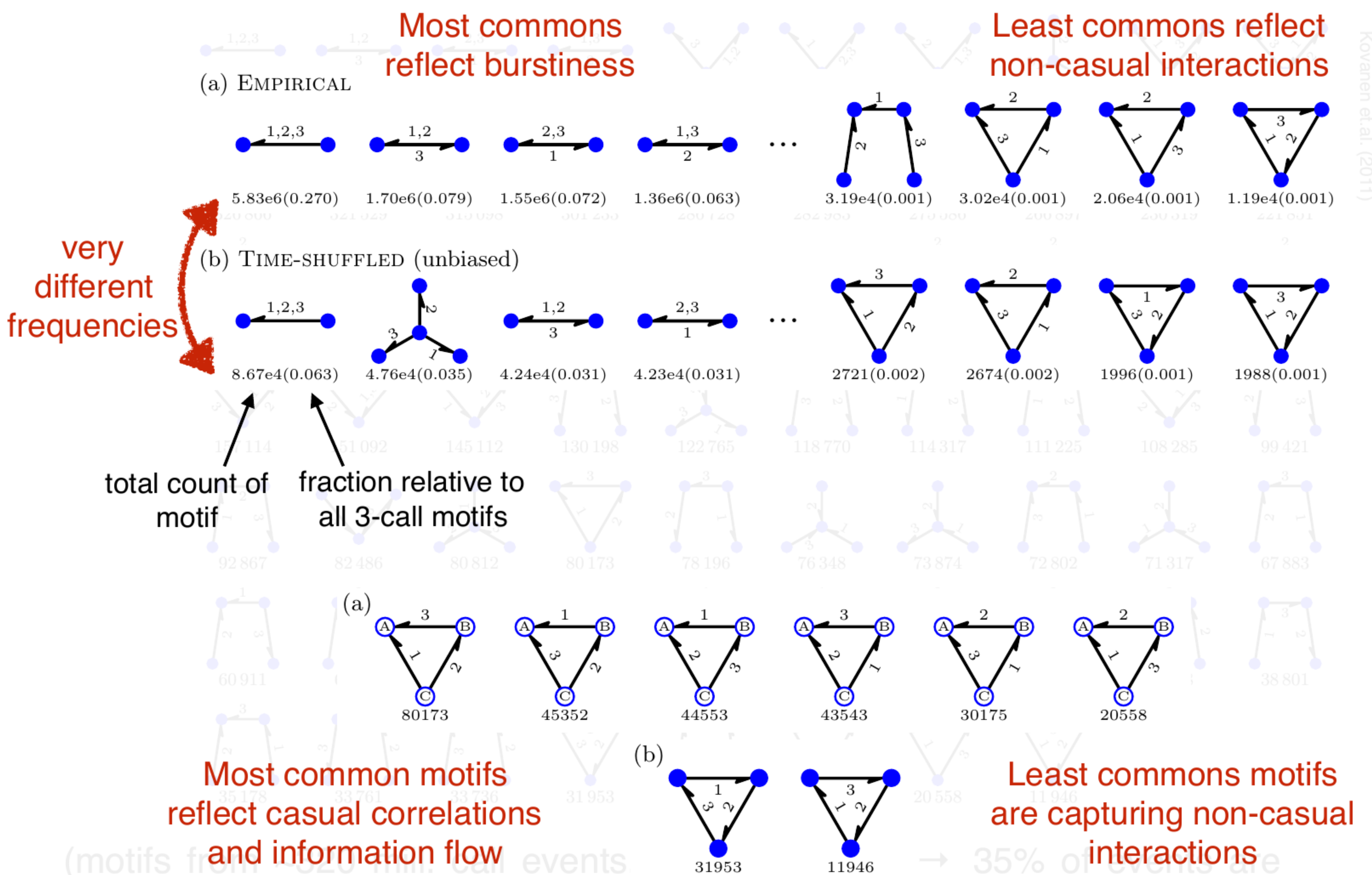
All 3-call motifs



Slide adopted from Márton Karsai

Phone call network

All 3-call motifs



Slide adopted from Márton Karsai