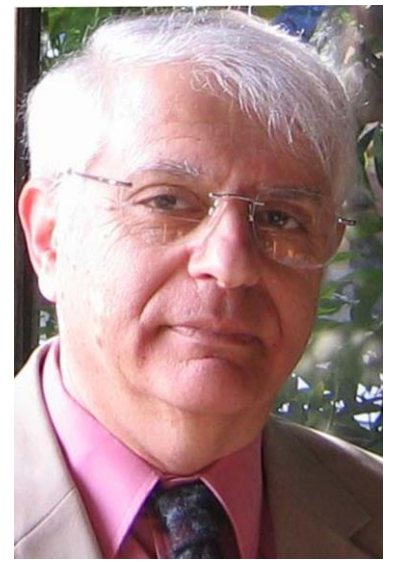


Dynamics of Networks

János Kertész
Central European University

Pisa Summer School September 2019

Mark Granovetter



The most pressing need for further development of network ideas is a move away from static analyses that observe a system at one point in time and to pursue instead systematic accounts of how such systems develop and change. Only by careful attention to this dynamic problem can social network analysis fulfill its promise as a powerful instrument in the analysis of social life.

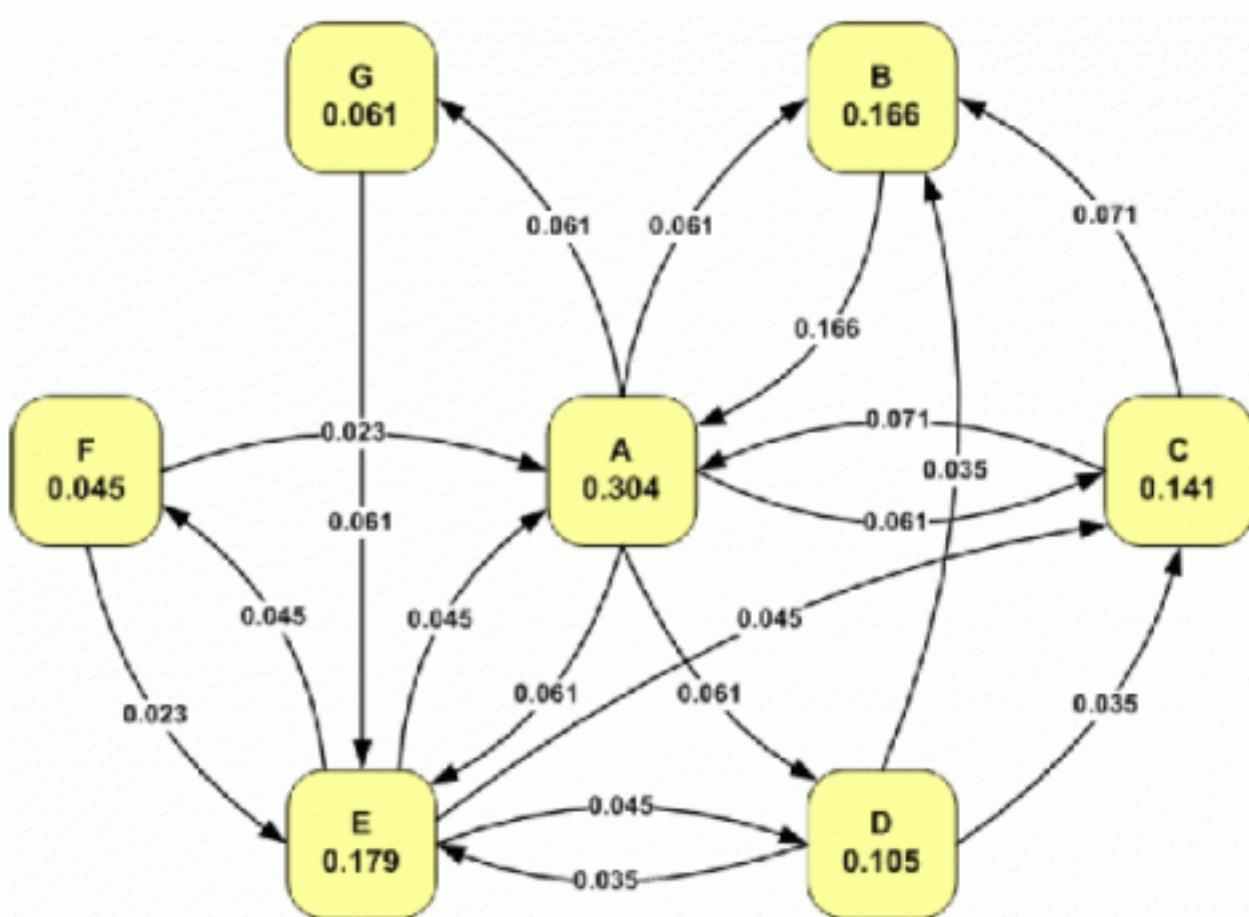
1983

Dynamics on and of Networks

- Dynamic processes on networks
 - Diffusion, random walk
 - Transport
 - Packet transfer according to protocol
 - Synchronization
 - Spreading
- Dynamics of networks
 - Network growth and development
 - Network shrinkage and collapse
 - Network restructuring, network adaptation
 - Temporal networks

Dynamics on Networks: Diffusion, random walk

Example: PageRank



PR is an iterative procedure to determine the importance of web pages based on random walk

Transport

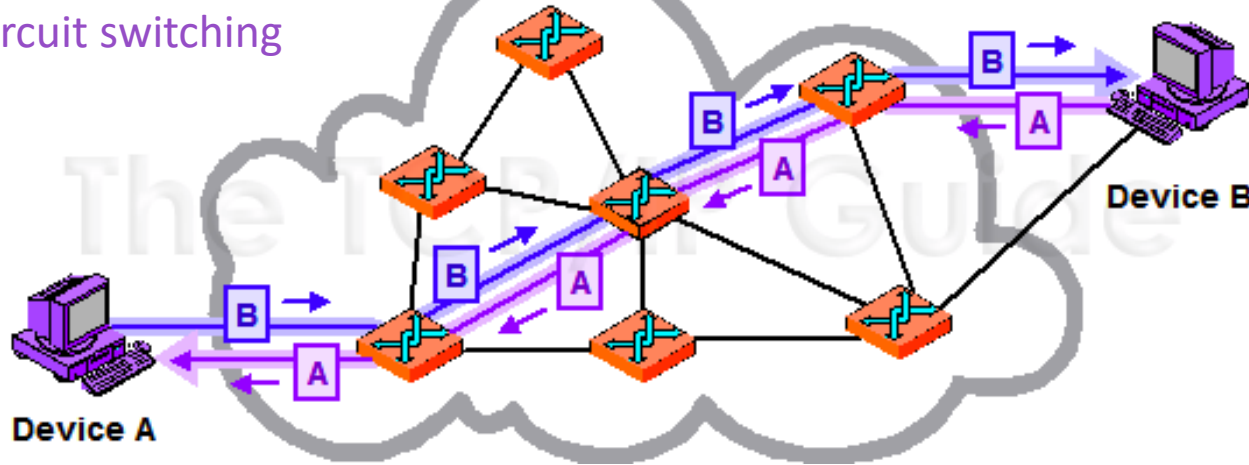
Truck Freight Flows, All Commodities

All truck types; highway freight density in tons



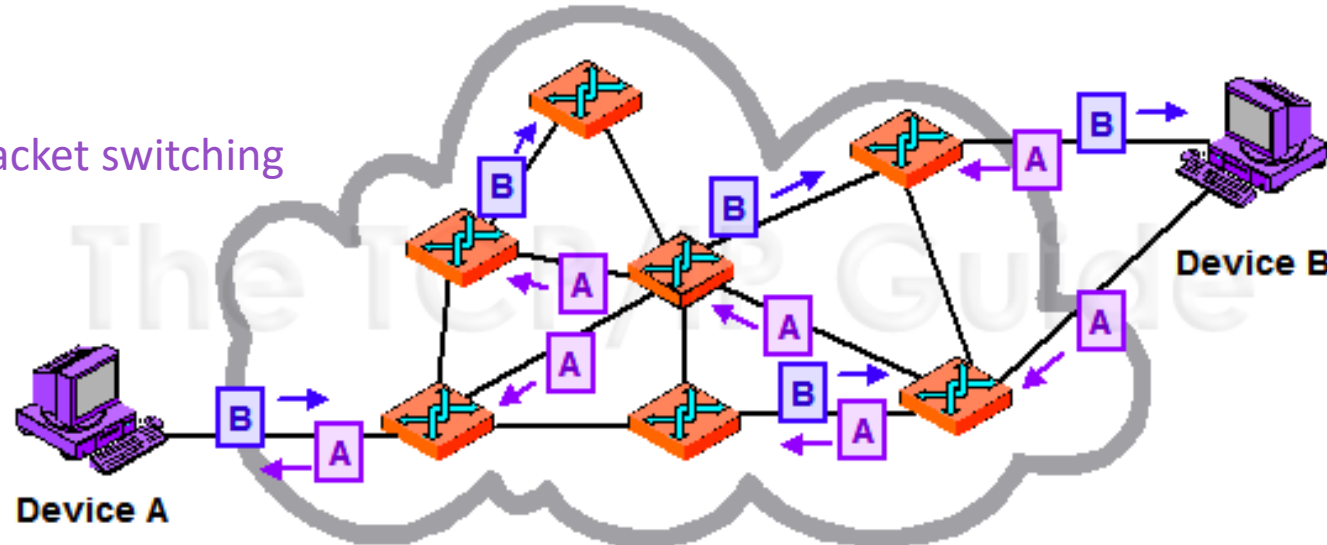
Packet transfer according to protocol

Circuit switching



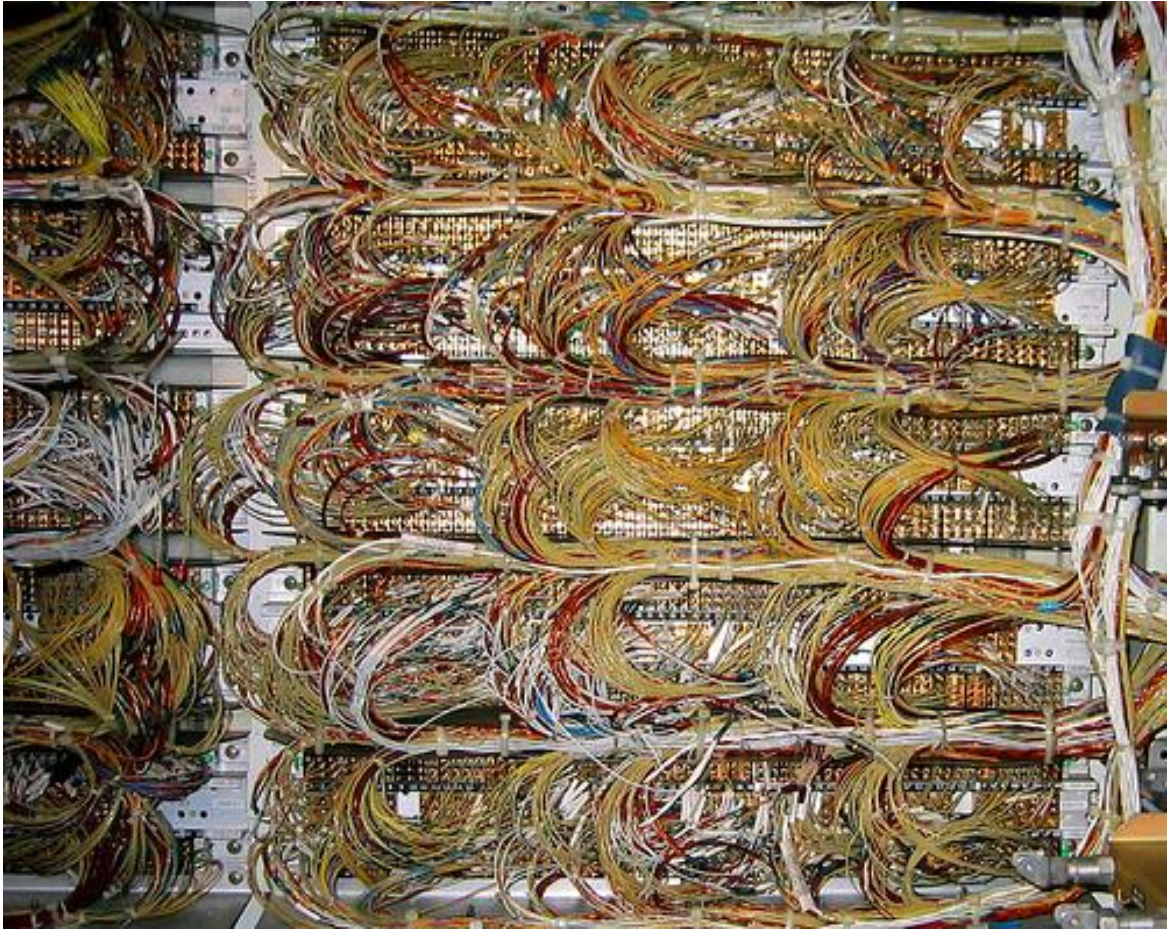
For communication a route has to be established and kept open throughout the exchange of information

Packet switching



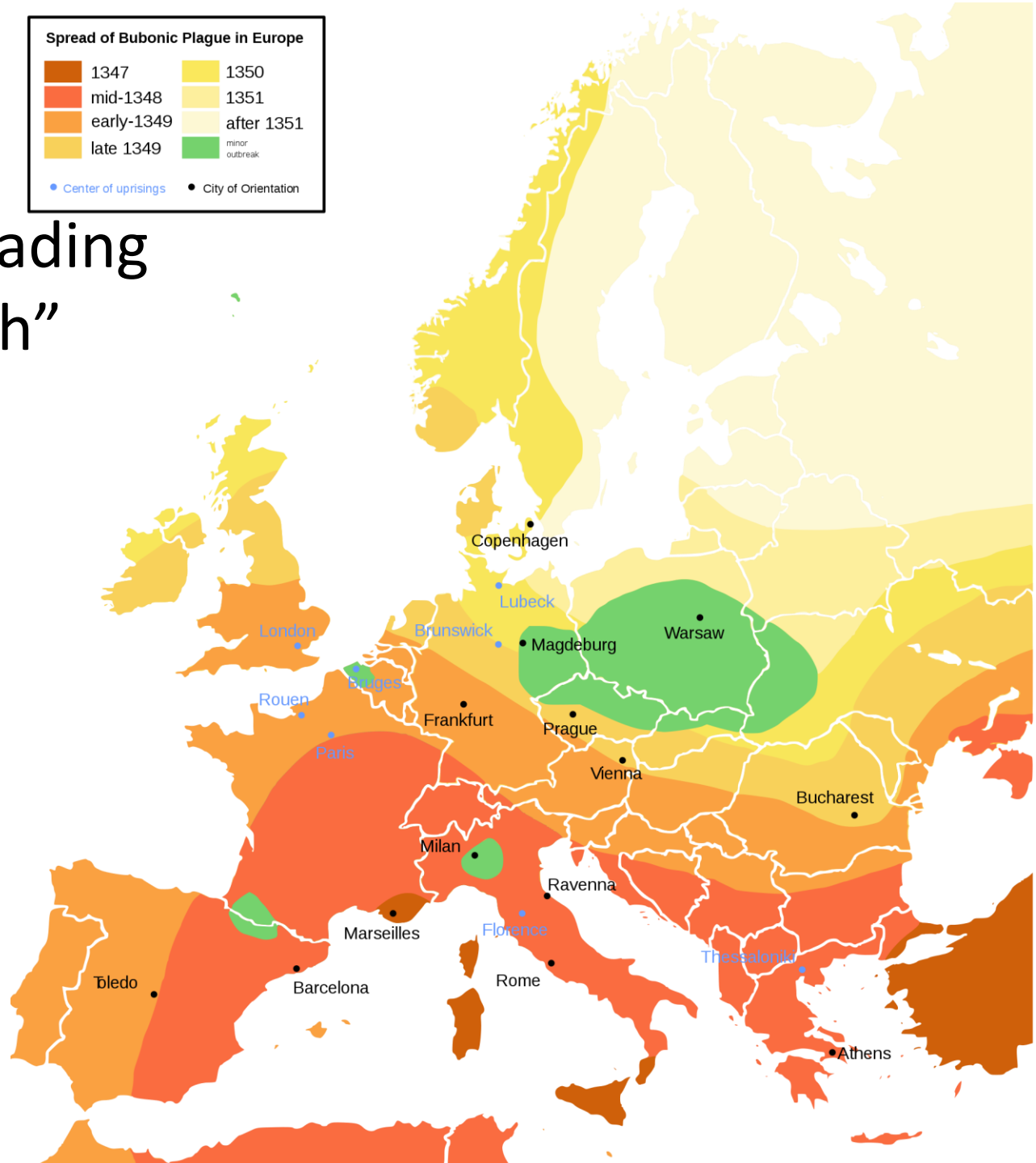
Information is chopped into pieces (packets), which travel on different routes and get reassembled finally

Complex electrical circuit

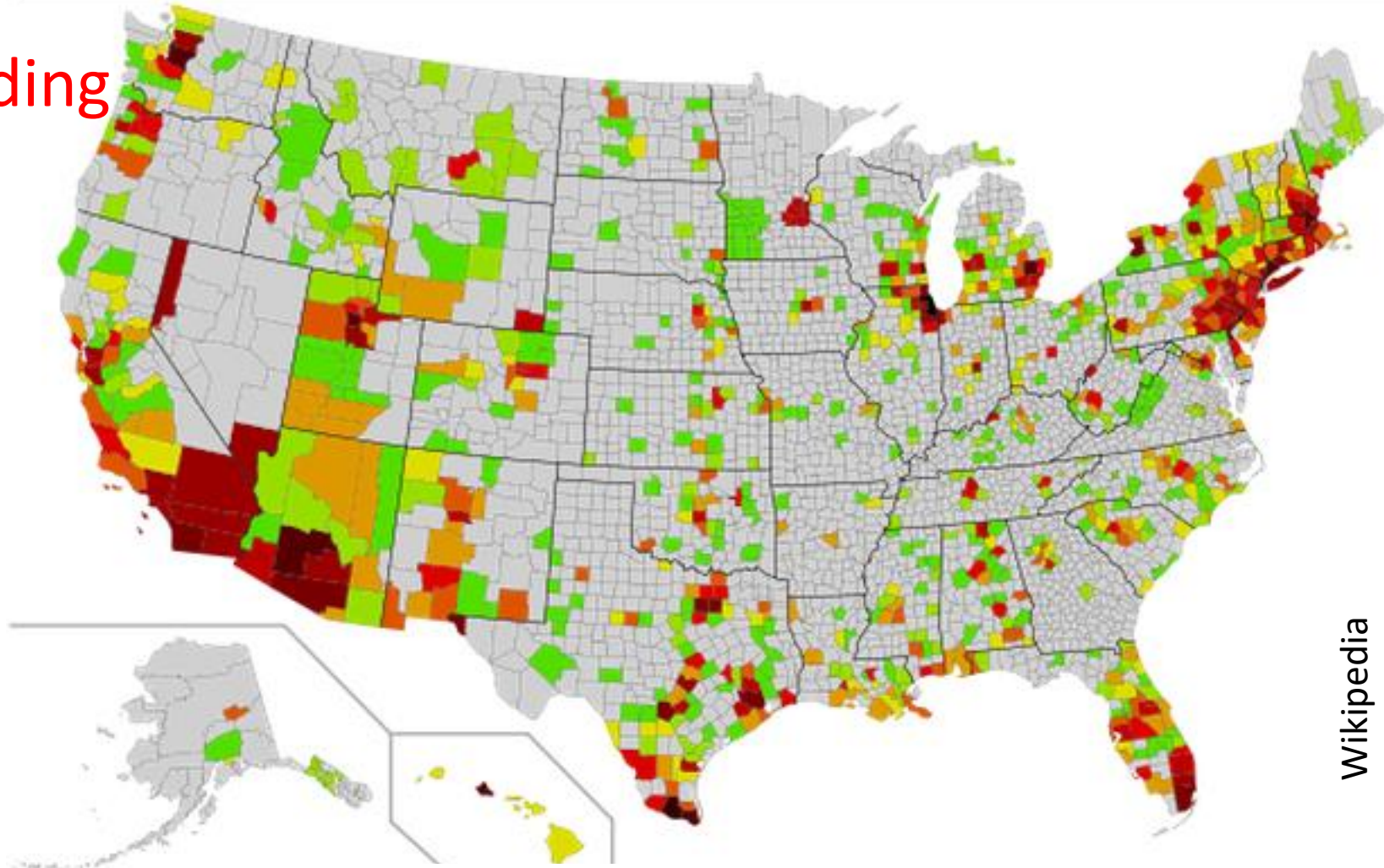


Spreading

Medieval spreading
of „Black Death”
(short range
interaction)



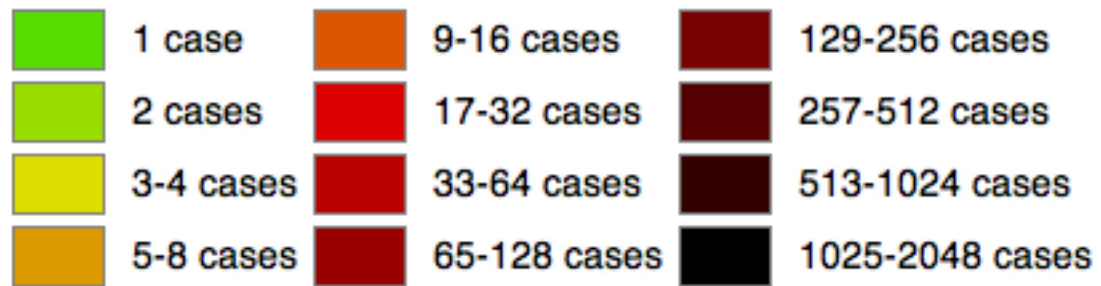
Spreading



Wikipedia

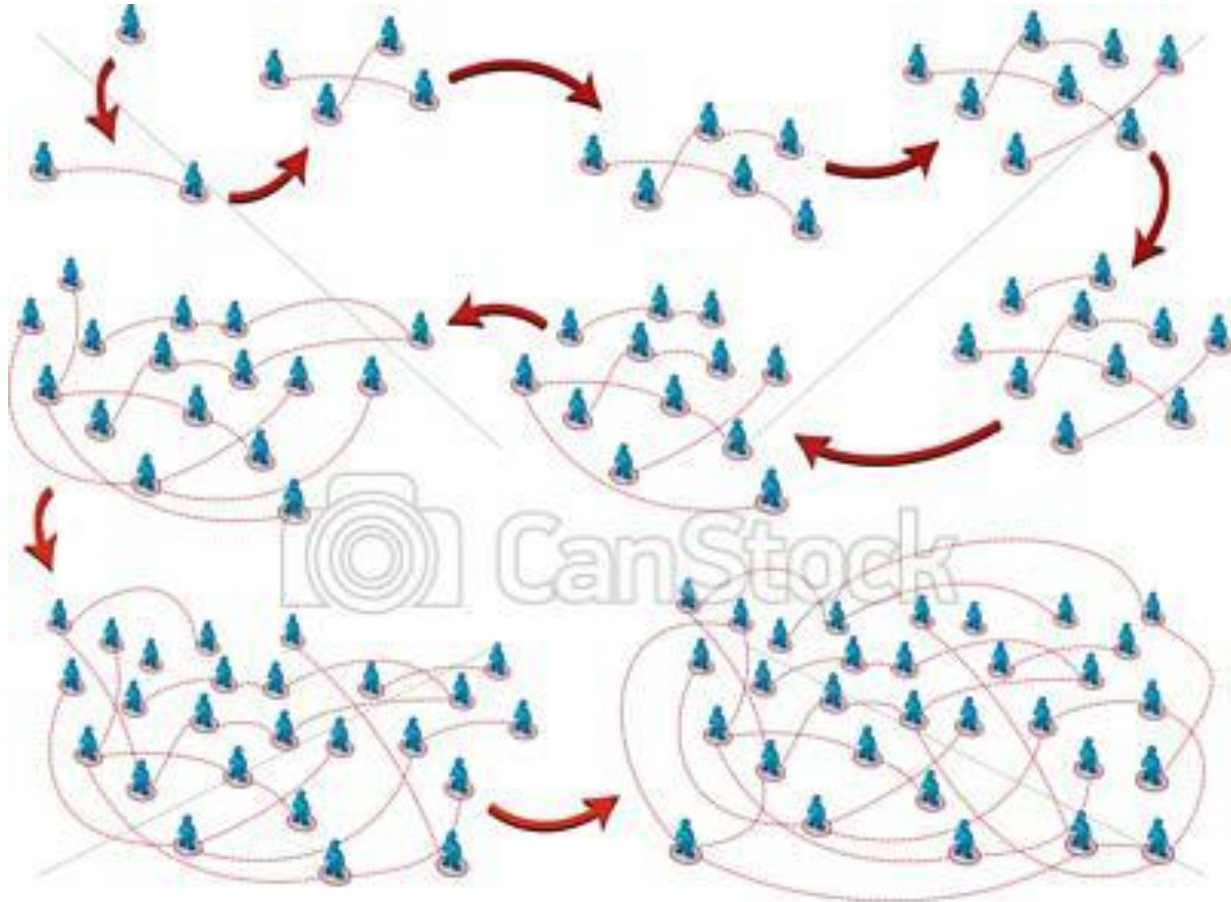
Swine flu June 2009
(long range
interaction)

Legend:



Dynamics of networks

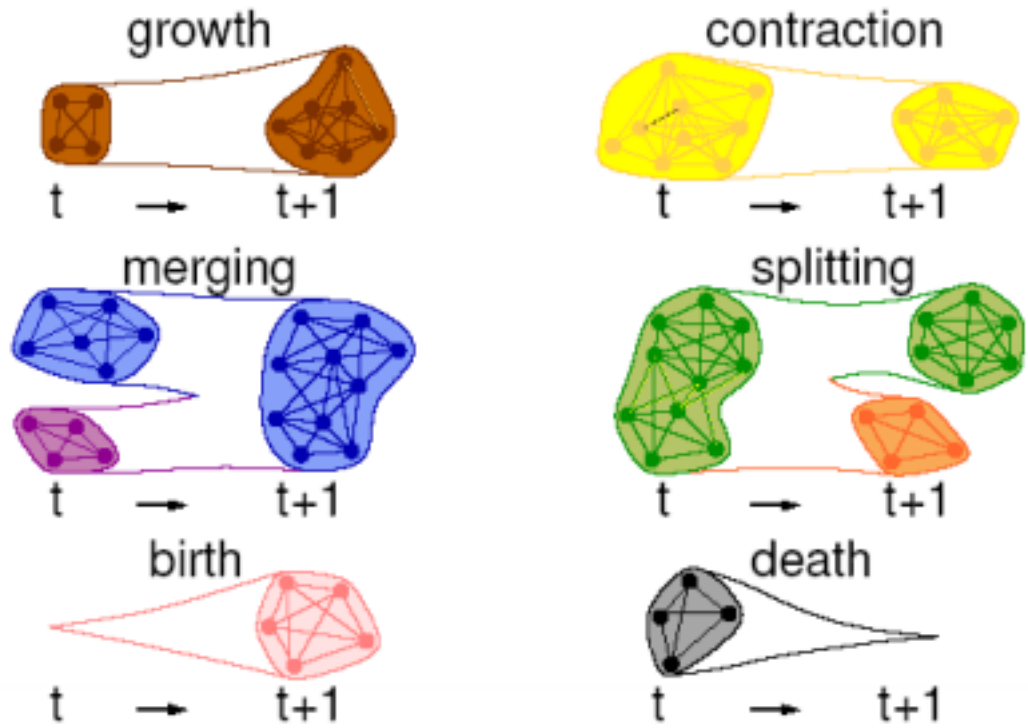
Network growth



© Can Stock Photo - csp8055353

See also network models, e.g., Barabási-Albert

Network restructuring



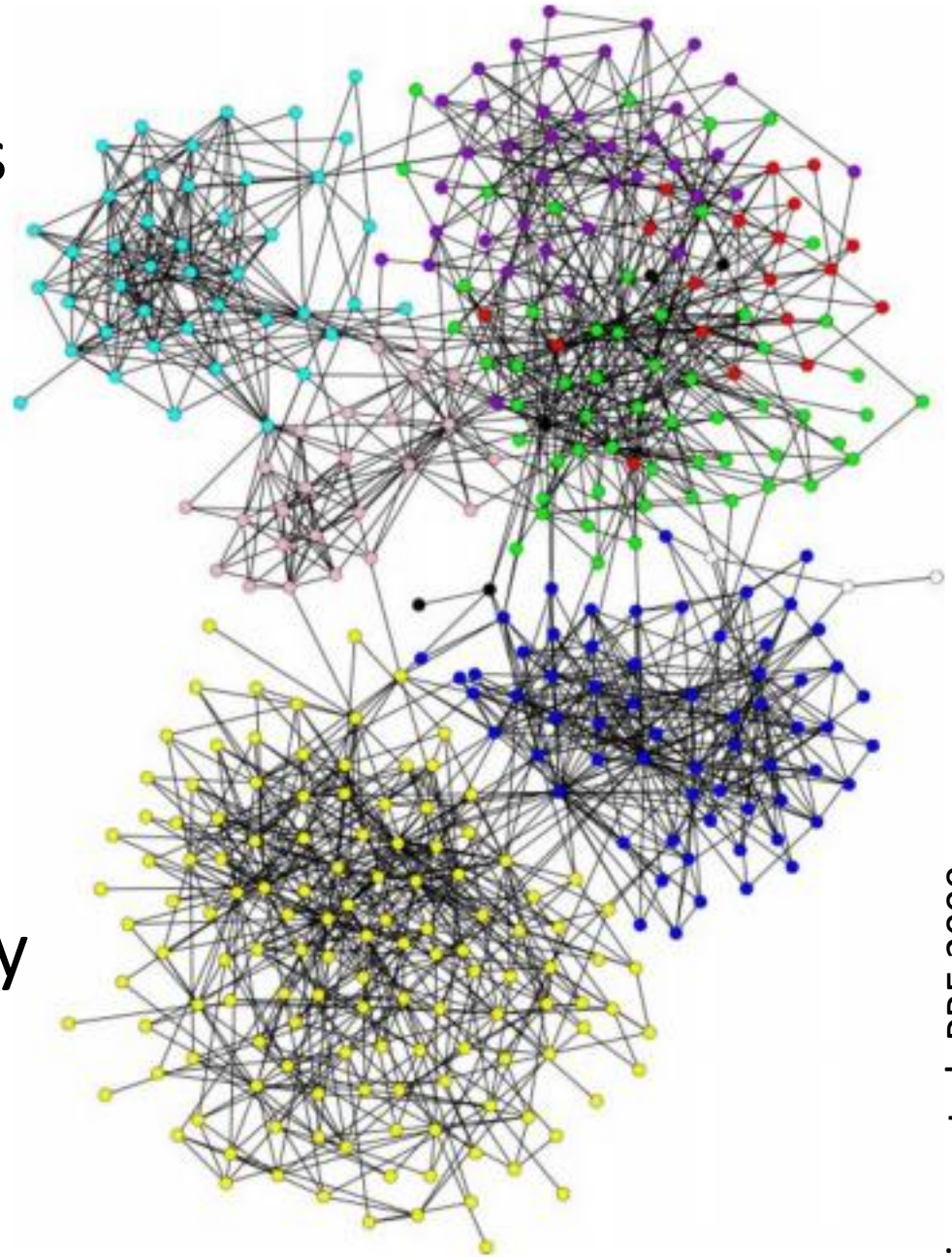
Group (community) evolution

Network adaptation

Network restructuring is coupled to an opinion dynamics mechanism

Nodes (people) look for more satisfactory connections.

The resulting community structure reflects the opinions



Time scales

In reality processes on the network and restructuring happen **simultaneously**.

Important: **Time scales**

If **time scales separate**, one can treat the dynamic degrees of freedom for the processes on the network separately from those of the network. Similar to the adiabatic approximation for solids.

E.g. road construction vs daily traffic

No separation of time scales

Reason:

- The characteristic times are similar (e.g., if the road is as frequently reconstructed as cars cross the static model of a network is meaningless.)
- There are no characteristic times (e.g., inter-event times are power-law distributed)

Even more so, if the network is **defined** by the events!

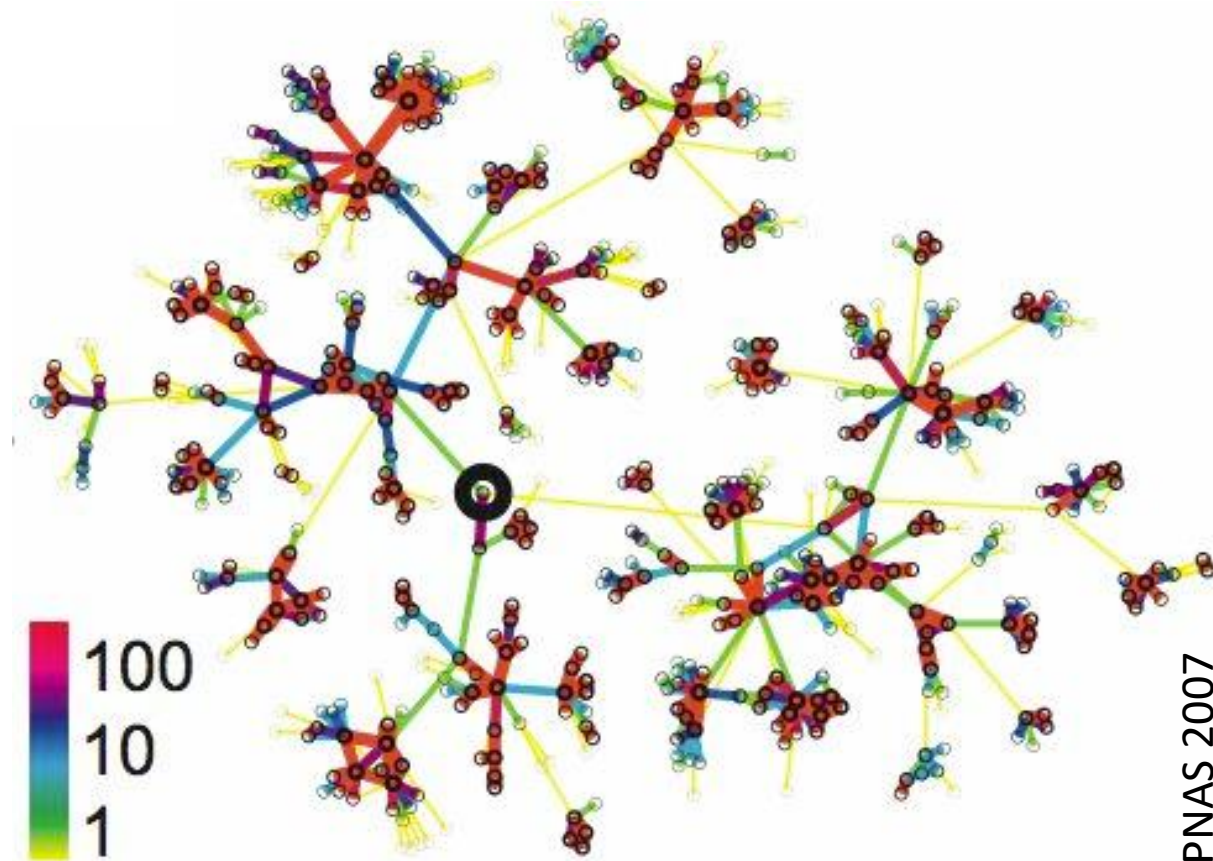
E.g.: communication

Temporal networks

Aggregate networks

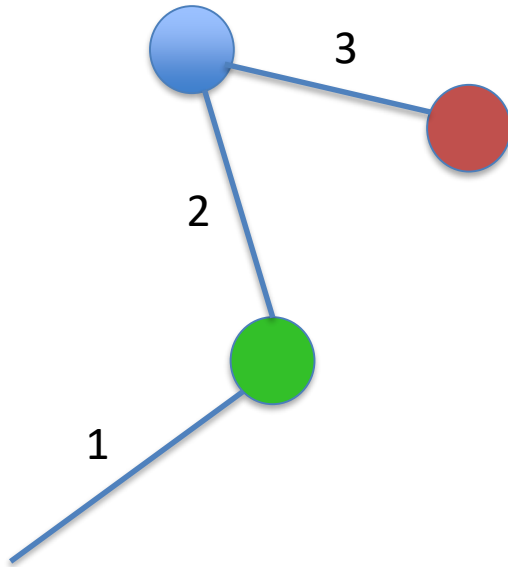
Consider all links over a period of time

Assuming that mobile phone calls represent social contacts, the aggregate network of call events is a proxy for the weighted human interaction network at societal level.



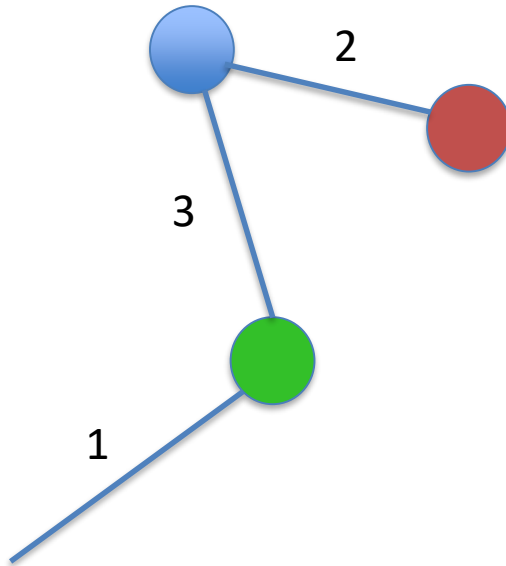
Spreading (of rumor, disease etc.)

Aggregation: information loss



Incoming information (1)
reaches everyone

Spreading (of rumor, disease etc.)

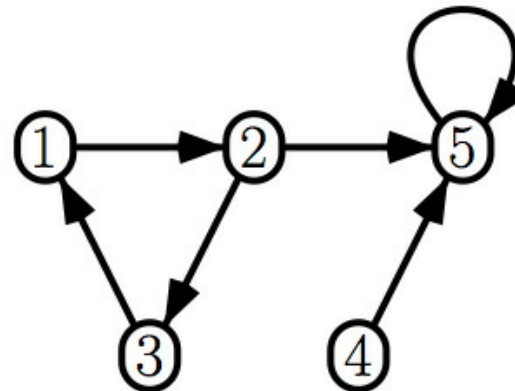
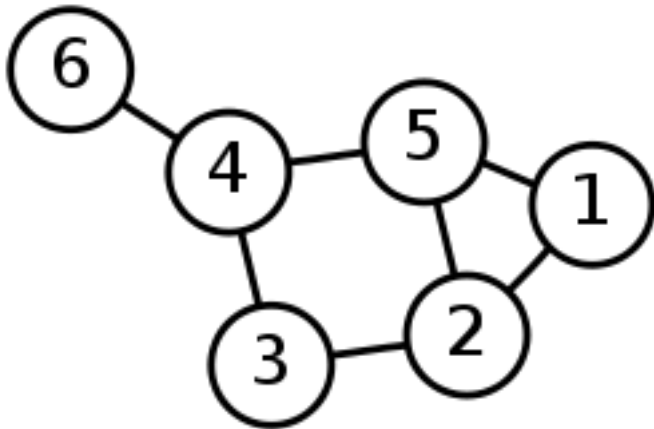


Incoming information (1)
does not reach 

The sequence of calls is crucial for the process

Network definition

Networks (graphs) are defined as $G = \{V, E\}$ where V is the set of nodes (vertices) and E is the set of – possibly directed – links (edges). Given the number N of nodes, the network is uniquely defined by the $N \times N$ adjacency matrix A_{ij} indicating that there is a link from i to j : $A_{ij} = 1$ or $A_{ij} = 0$ otherwise for non-weighted networks.



Temporal network definition

A **temporal network** (contact sequence) is defined as $\mathcal{T} = \{V, S\}$ where V is the set of nodes and S is the set of – possibly directed – event sequences assigned to pairs of nodes. For $s_{ij} \in S$

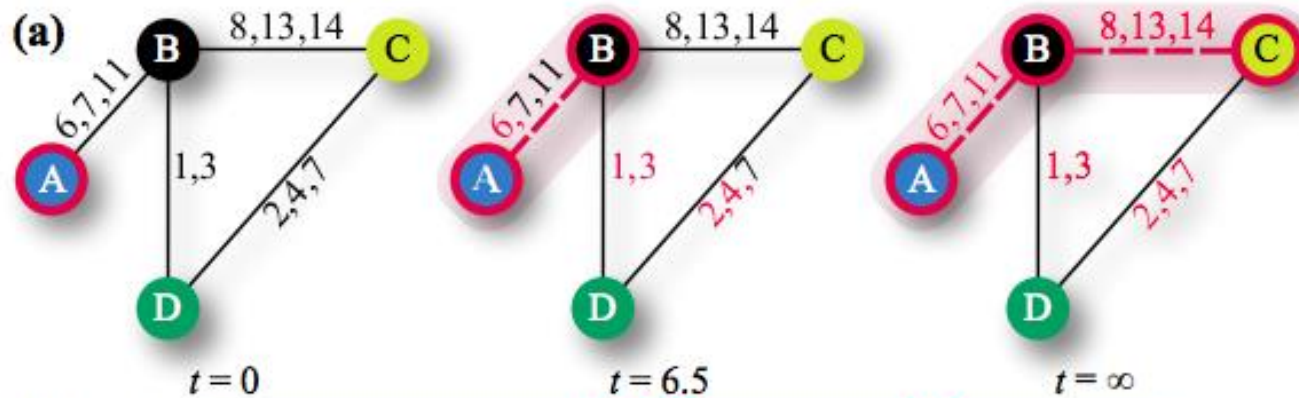
$$s_{ij} = \left\{ t_{ij}^{(1)}, \tau_{ij}^{(1)}; t_{ij}^{(2)}, \tau_{ij}^{(2)}; \dots; t_{ij}^{(n)}, \tau_{ij}^{(n)}; \dots \right\}$$

where t_{ij} -s are the beginnings and τ_{ij} -s the durations of events $i \rightarrow j$ within a time window

$$A(i, j, t) = \begin{cases} 1 & \text{if } i \rightarrow j \text{ connected at } t \\ 0 & \text{otherwise} \end{cases} \quad \begin{matrix} \tau_{ij}=0 \text{ can often be assumed} \\ \text{adjacency index} \end{matrix}$$

continuous or discrete

Temporal network visualization



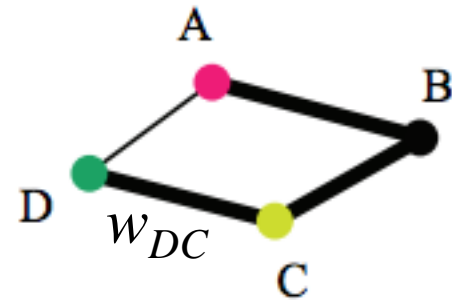
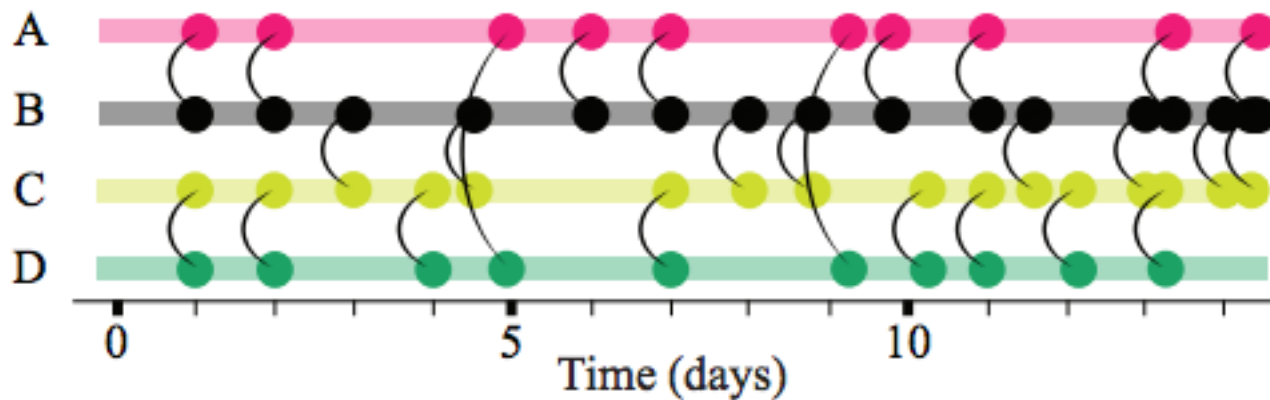
Holme, Saramaki : Phys. Rep. 519, 97-125 (2012)

Figures are taken from that review if not indicated otherwise

When are temporal networks important?

Always, if sequence of events is important (spreading) or temporal inhomogeneities matter (jamming).

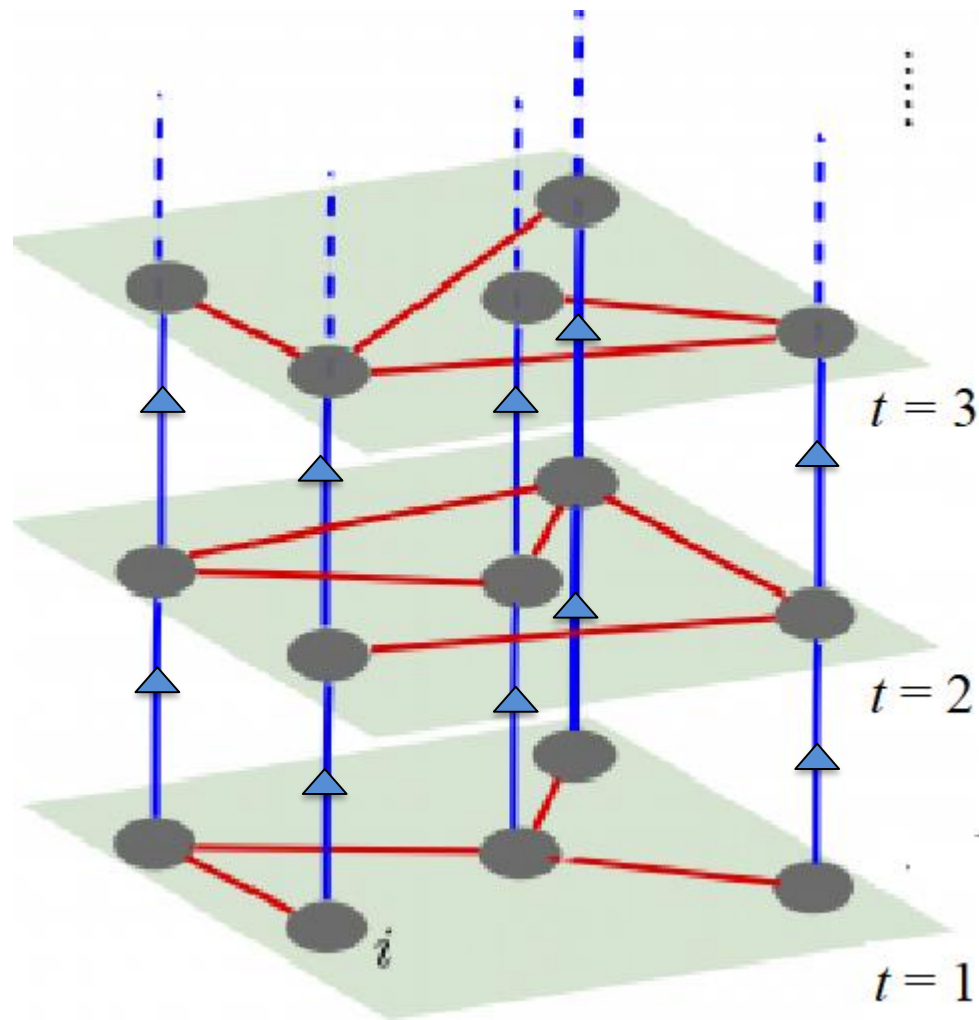
From each temporal network a (weighted) static network can be constructed by aggregation.



$$w_{ij} = \int_{t_{\min}}^{t_{\max}} A(i, j, t) dt: \quad w_{ij} = \# \text{ or total duration of events}$$

This can be used to model dynamic phenomena if processes are simple (Poissonian).

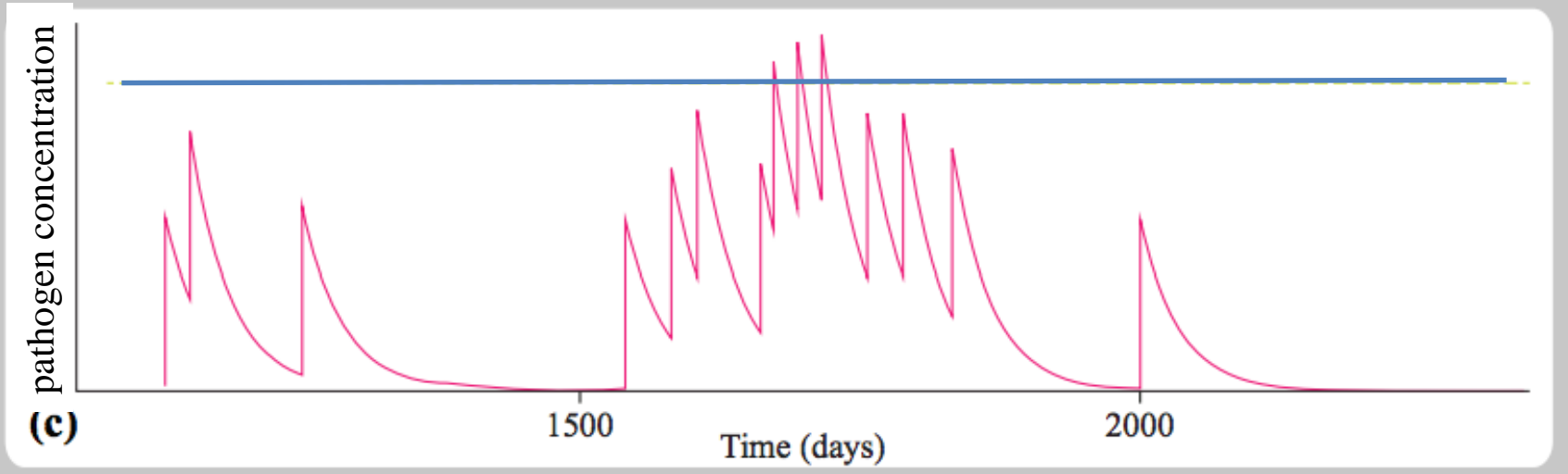
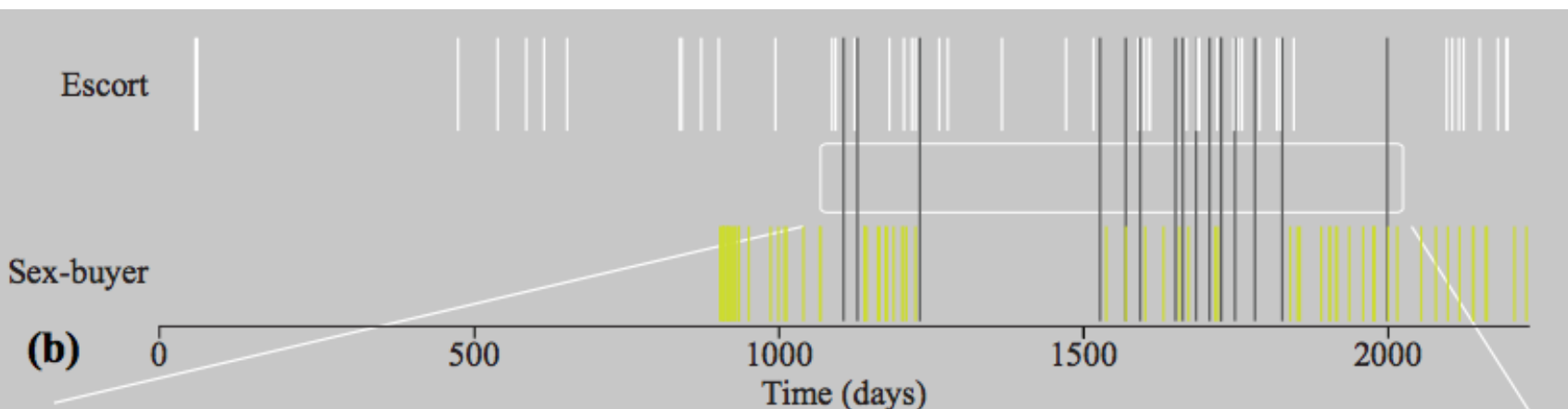
Relation to multiplex networks: discrete time



Blue lines are strictly directed

Consequences of strong temporal inhomogeneities

Temporal behavior is often non-Poissonian, bursty. This may have different reasons from seasonalities to external stimuli and to intrinsic burstiness.



Examples of temporal networks

- Communication networks
- Physical proximity
- Gene regulatory networks
- Parallel and distributed computing
- Neural networks
- etc.

Examples of temporal networks

- Communication networks
- Physical proximity
- Gene regulatory networks
- Parallel and distributed computing
- Neural networks
- etc.

Temporal communication networks

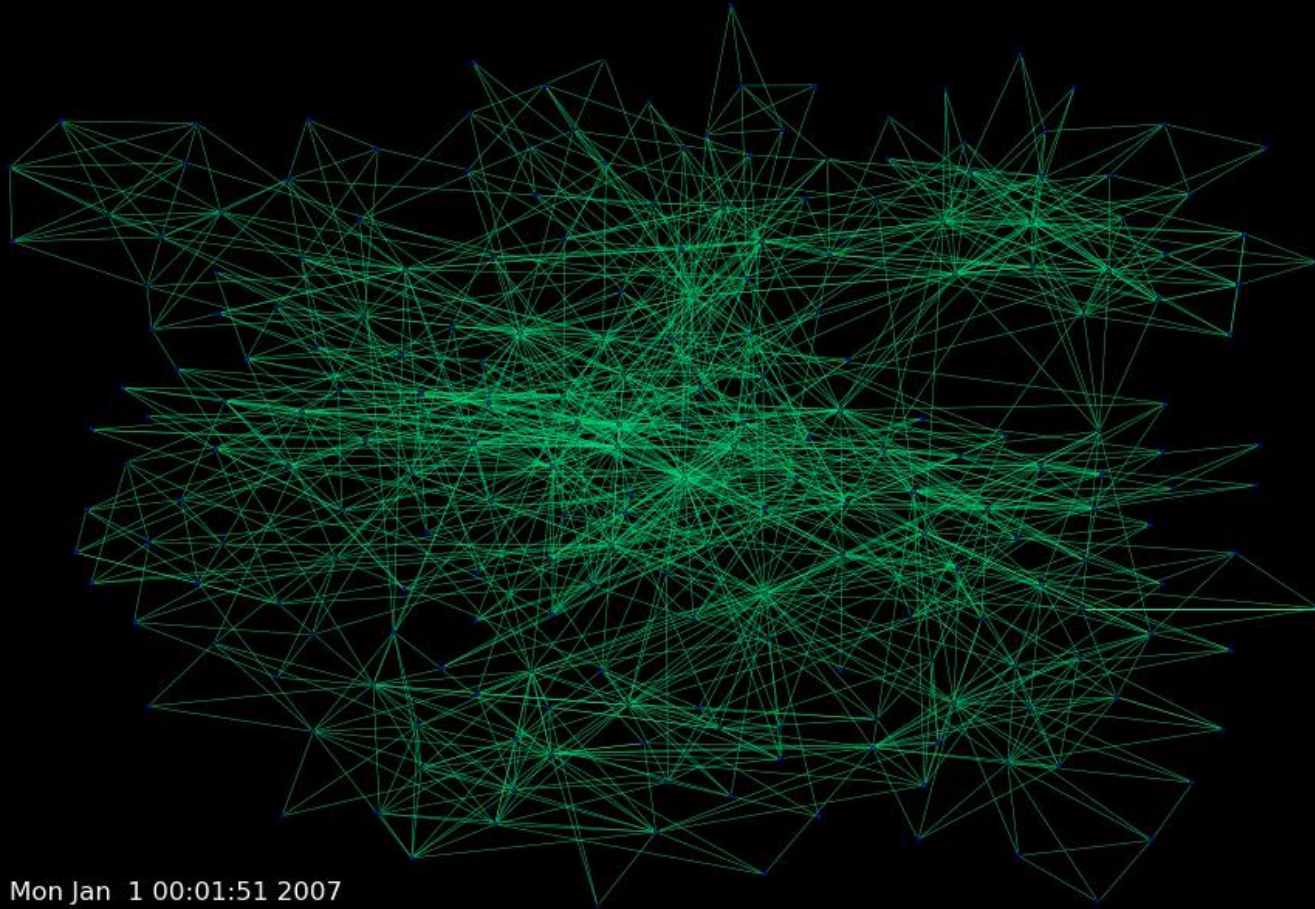
- One to one
 - face to face
 - phone
 - SMS
 - email
 - chat
- One to many
 - lecture
 - multi address SMS
 - multi address email
 - twit, blog



- Many to many
 - meeting
 - conference call

IT related communication data are precious:
Large in number and accurate





Mon Jan 1 00:01:51 2007

Examples of temporal networks

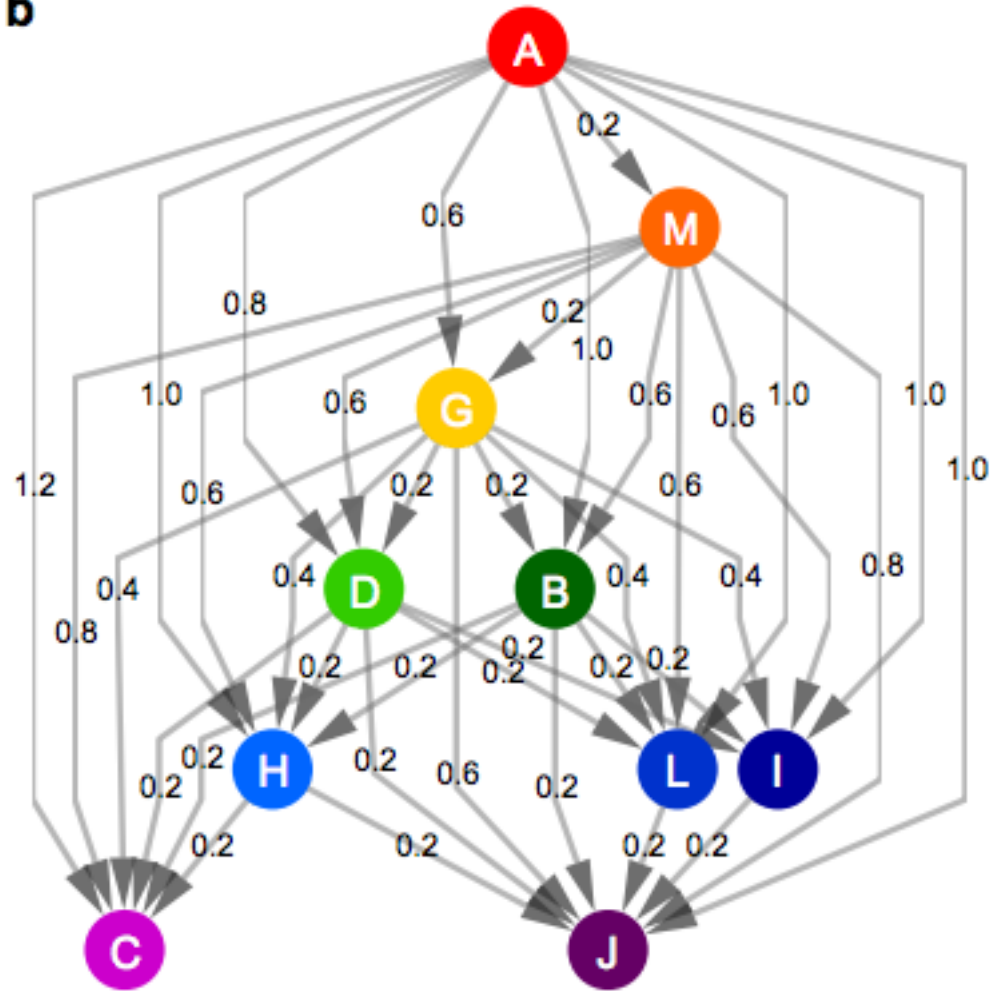
- Communication networks
- Physical proximity
- Gene regulatory networks
- Parallel and distributed computing
- Neural networks
- etc.

Physical proximity

Human or animal
proximity
Important, e.g., for
spread of airborne
pathogens or mobile
phone viruses
transmitted via
bluetooth



b

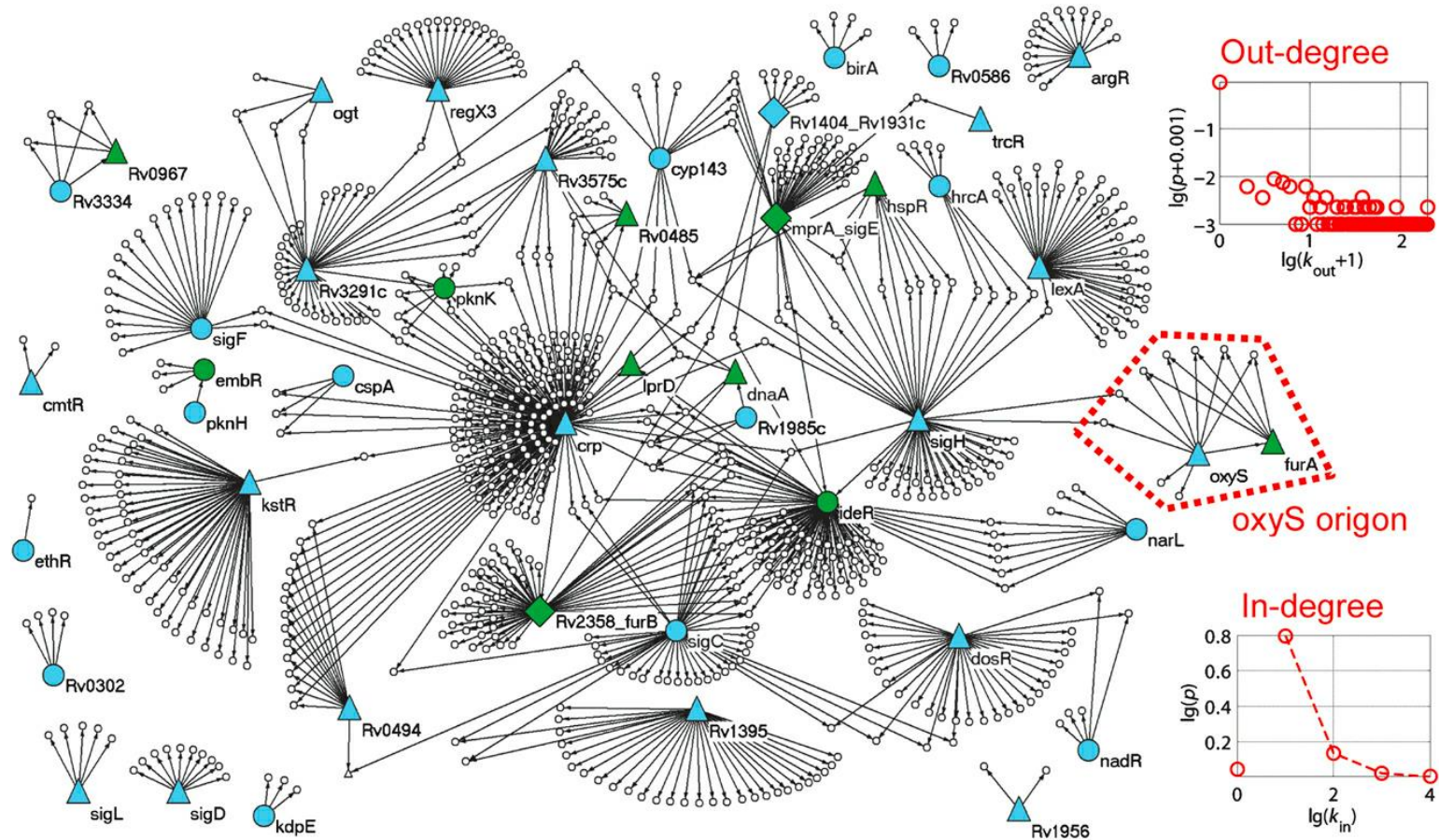


Data: MIT Reality mining (Bluetooth), Barrat group (RFID), OtaSizzle (tower, WiFi), Copenhagen Network Study (CDR, Wi-Fi), traffic (GPS)

Examples of temporal networks

- Communication networks
- Physical proximity
- Gene regulatory networks
- Parallel and distributed computing
- Neural networks
- etc.

Gene regulatory networks



Aggregate NW, in reality: Sequence of chemical reactions. Order pivotal!

Examples of temporal networks

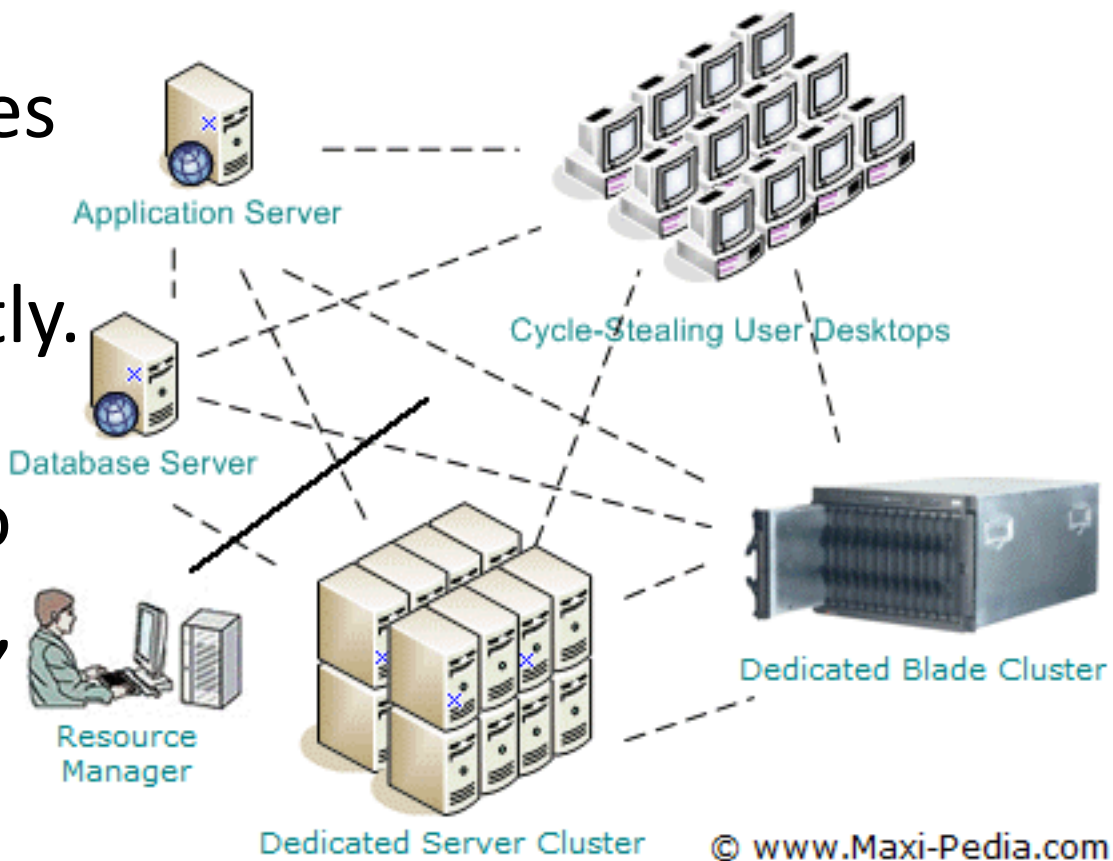
- Communication networks
- Physical proximity
- Gene regulatory networks
- Parallel and distributed computing
- Neural networks
- etc.

Parallel and distributed computing

DC: Put all resources together to solve a single task efficiently.

Problems similar to parallel computing, where many processors work simultaneously.

Data transfer: Processes use results of other units – timing is crucial.

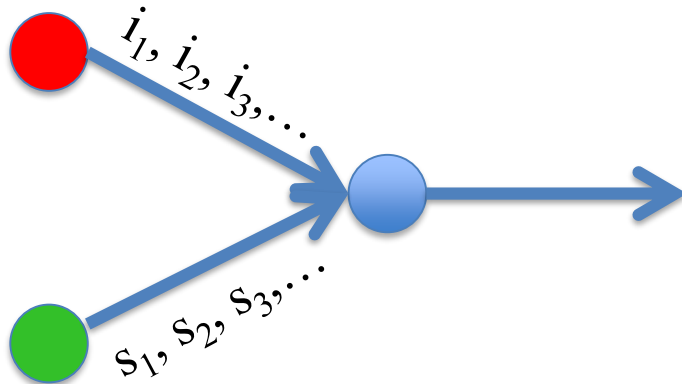


Examples of temporal networks

- Communication networks
- Physical proximity
- Gene regulatory networks
- Parallel and distributed computing
- Neural networks
- etc.

Neural networks

Neurons get stimulating or inhibitory impulses from other ones



Output heavily depends on the sequence of the inputs:

$s_1, i_1, s_2, i_2, s_3, i_3, s_4, \dots$ is totally different from
 $s_1, s_2, s_3, s_4, \dots, i_1, i_2, i_3, \dots$

Characterizing networks

Aggregated networks can be considered as static ones:
An arsenal of concepts and measures exist:

- path, distance, diameter
- degree
- centrality measures
- correlations (e.g., assortativity)
- components
- minimum spanning tree
- motifs
- communities

Characterizing temporal networks

Similarities with directed networks – due to the arrow of time.

Difference: sequential order matters

Need for generalization of concepts

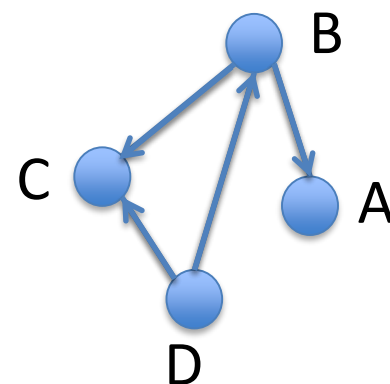
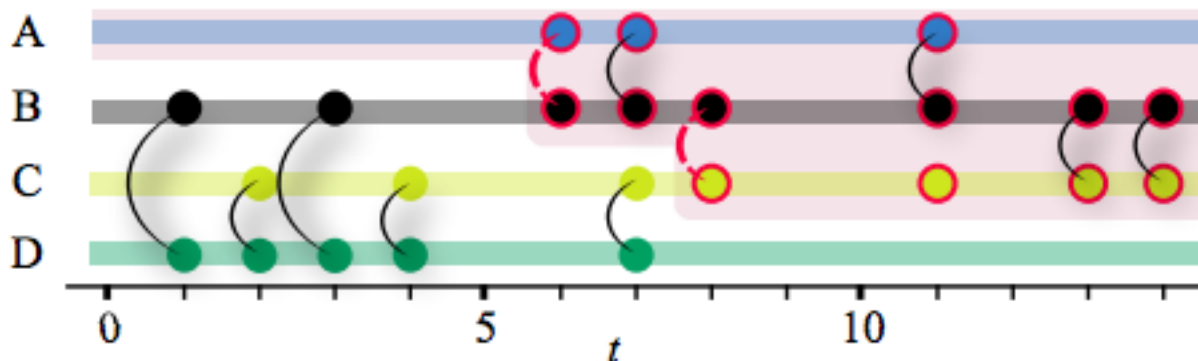
- path, distance, diameter
- centrality measures
- components
- motifs

Paths vs reachability

A **path** in a graph consists of a series of subsequent edges without visiting a node more than once.

$$\mathcal{P}(1, n) = \left\{ e_{12}, e_{23}, e_{34}, \dots, e_{n-1, n} \mid e_{ij} \in E \right\}$$

A path from i to j on the aggregate graph does not mean that j is reachable from i .



There is a path DA, which is symmetric for undirected graphs. A can be reached from D but not D from A.

Like for directed graphs

Time respecting path (journey)

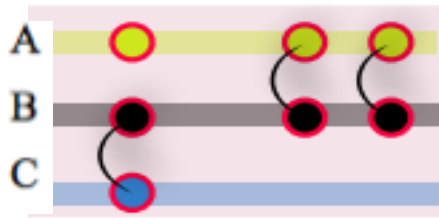
Temporal networks should be studied with respect to a **time window** $t_{ij} \in (t_{\min}, t_{\max})$.

$$\mathcal{J}_{1 \rightarrow n} = \{t_{12}, t_{23}, t_{34}, \dots, t_{n-1,n} \mid t_{12} < \dots < t_{n-1,n}\},$$

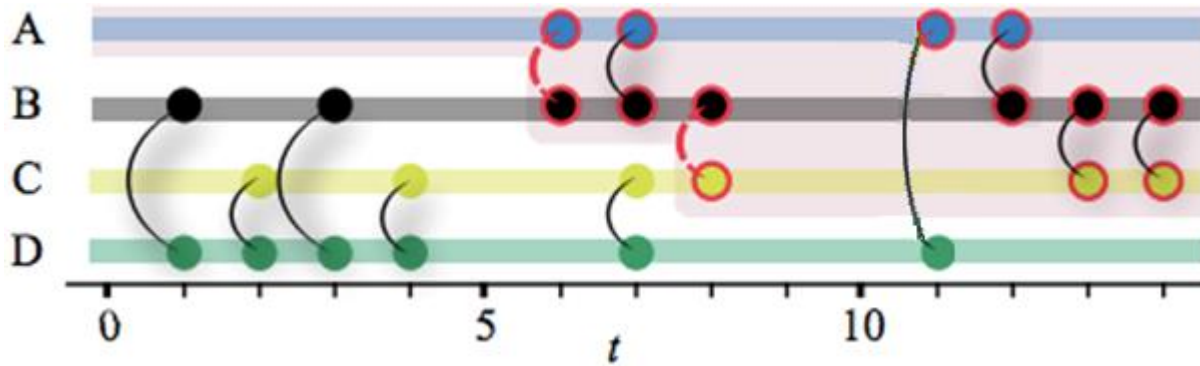
where t_{ij} -s are event times and the nodes $\{1, 2, \dots, n\}$ form a path in the aggregate network.

Time respecting paths define the **set of influence** of node i within this window: $\mathcal{F}_i(t) = \{\forall j \mid j \in V, \exists \mathcal{J}_{i \rightarrow j}\}$ such that all times $> t$ in $\mathcal{J}_{i \rightarrow j}$ -s are within the window.

Similarly, the **source set** is defined as the set of nodes from which i can be reached by t within the window $\mathcal{P}_i(t) = \{\forall j \mid j \in V, \exists \mathcal{J}_{j \rightarrow i}\}$



Journeys are **non-transitive**: $A \rightarrow B$ and $B \rightarrow C$ does not imply $A \rightarrow C$.



$$\mathcal{F}_B(10) = \{A, C\}$$

$$\mathcal{P}_C(5) = \{B, D\}$$

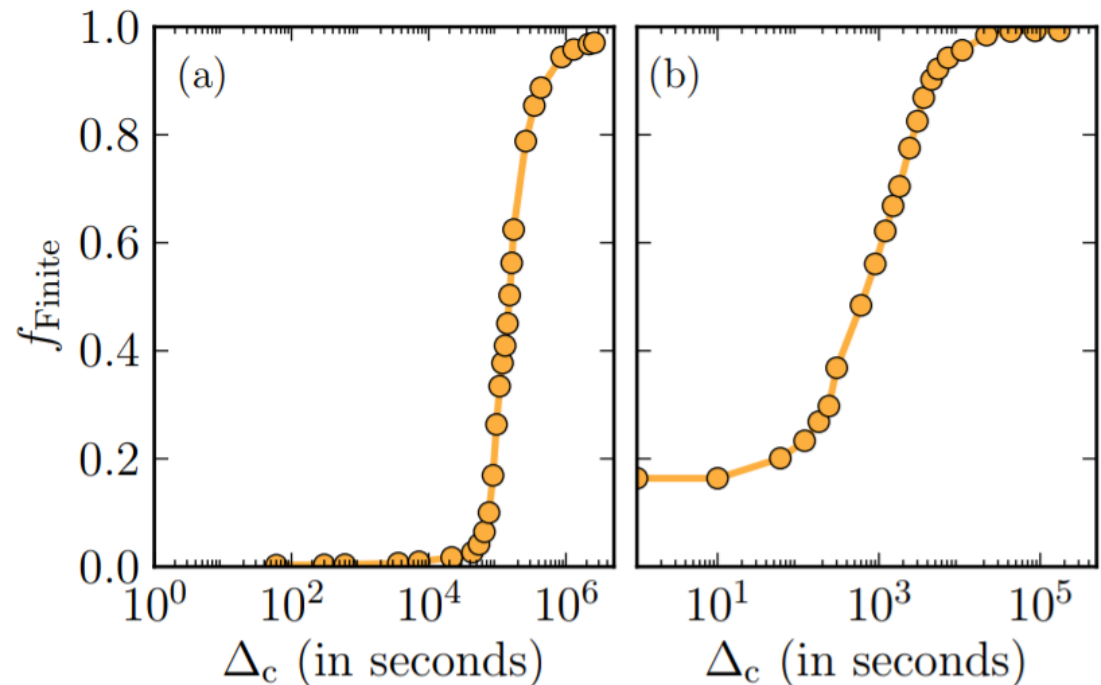
Journeys with max. waiting times

$$\mathcal{J}_{1 \rightarrow n}^{\Delta_c} = \{t_{12}, t_{23}, t_{34}, \dots, t_{n-1,n} | t_{12} < \dots < t_{n-1,n}; t_{i,i+1} - t_{i-1,i} < \Delta_c\}$$

Similarly, sets of influence and source set can be defined with respect to Δ_c . Reachability ratio:

$$f_{\text{Finite}}(\Delta_c) = \frac{1}{N} \sum_{k=1}^N \mathcal{F}_k^{\Delta_c}(t_{\min})$$

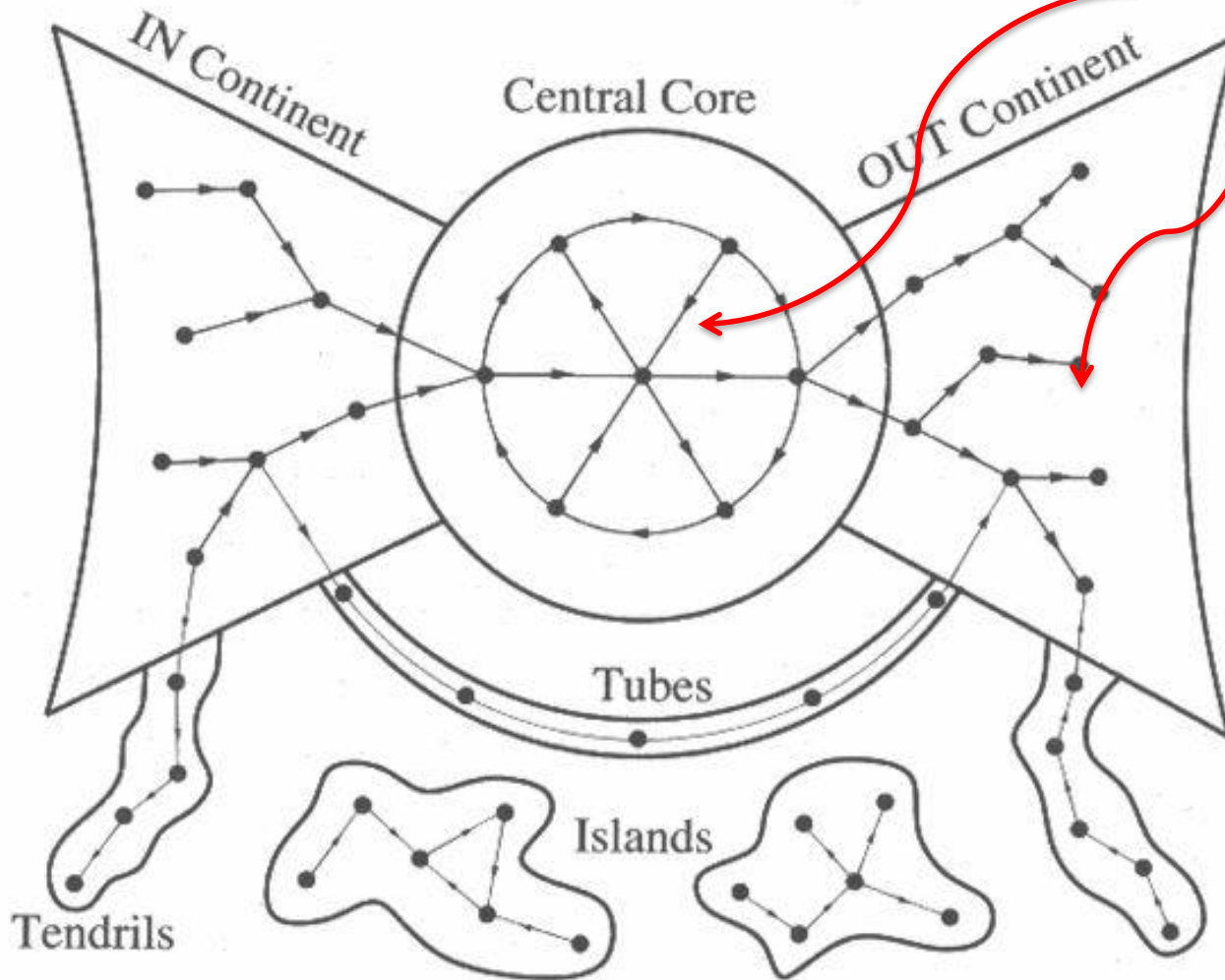
- a) Mobile call data
char. time: 1-2d
- b) Air traffic
char. time: 30 min
(~transfer time)



Connectivity and components

For directed networks:

Strongly connected
weakly conn.
components



Analogously
for temporal
graphs

**WINDOW-
DEPENDENCE!**

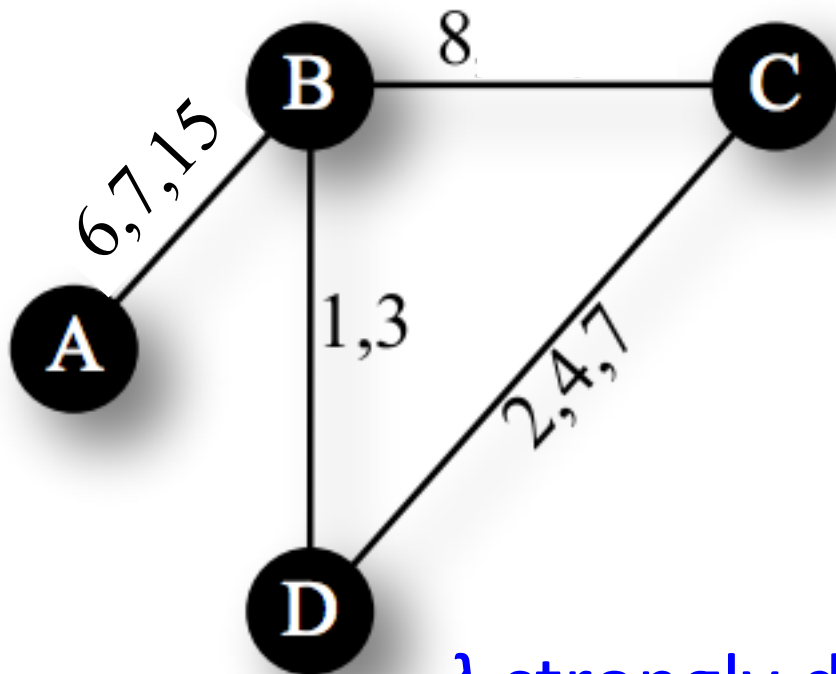
Shortest paths, fastest journeys

Length ℓ of a path is the number of edges in it.

Distance $d(i,j)$ is the length of the shortest path.

Duration $\delta(1,n)$ of a journey is the time $t_{n,n-1} - t_{1,2}$

Latency $\lambda(i,j)$ is the duration of the fastest journey.



$$\ell(C,D,B,A)=3$$

$$d(C,B,A)=2$$

$$\delta(C,B,A)=15-8=7$$

$$\lambda(C,D,B,A)=3-2+6-3=4$$

λ strongly depends on the time window

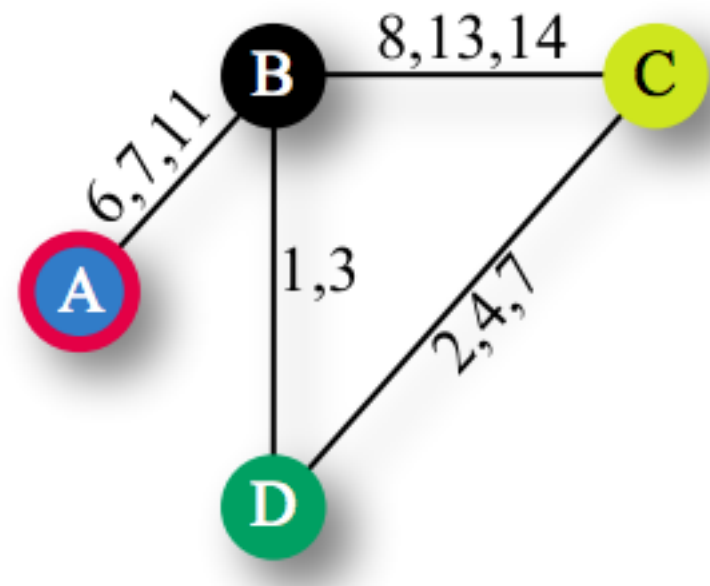
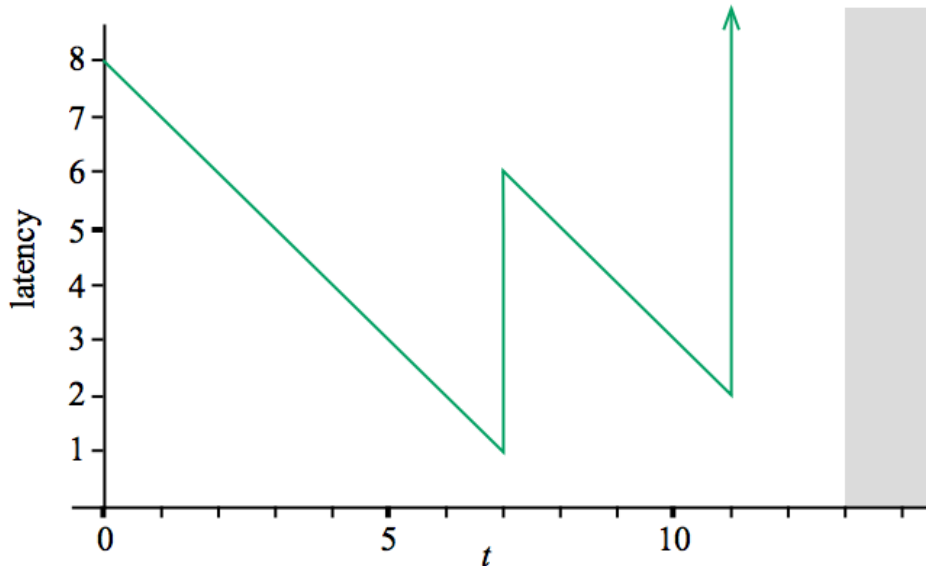
Mean shortest path, average latency

$$\bar{d} = 1/|E| \sum_{(i,j)} d(i,j)$$

Defined for a connected component (E is its edge set)

Generalization to average latency is non-trivial.

1. Mean shortest path tells about spatial reachability, latency is about time
2. There are strong variations even for the average over a single link.



Temporal boundary cond.

Centrality measures I.

Detect importance of elements

Closeness centrality in graphs: inverse average distance from i

$$C_C(i) = \frac{N-1}{\sum_{i \neq j} d(i, j)}$$

Temporal analogue

$$C_C(i, t) = \frac{N-1}{\sum_{i \neq j} \lambda_t(i, j)}$$

where $\lambda_t(i, j)$ is the latency from $i \rightarrow j$ at time t

Centrality measures II.

Betweenness centrality in graphs: proportional to the number of shortest paths through element

$$C_B(i) = \frac{\sum_{i \neq j \neq k} v_i(j, k)}{\sum_{i \neq j \neq k} v(j, k)}$$

where $v_i(j, k)$ is the number of shortest paths through i and $v(j, k) = \sum_i v_i(j, k)$

Temporal analogue

Possibilities:

- a) Shortest paths ratio conditioned by reachability
- b) Fastest path ratio

Temporal BC-s!

Motifs

Static motifs

Main task of studying (static) complex network is to understand the relation between topology and function.

Centrality measures try to identify most important elements.

What are the most important **groups of elements**?

Motif: set topologically equivalent (isomorphic) subgraphs

Cardinality of a motif shows its relevance with respect to a (random) null model.

Relevance of static motifs

If the cardinality of a motif is significantly high, it is expected that the represented subgraphs are relevant for some kind of function.

If it is small, the related function is irrelevant

Null model: Configuration model, no degree-degree correlations. The studied NW is a single sample, the null model is an ensemble leading to distributions in properties.

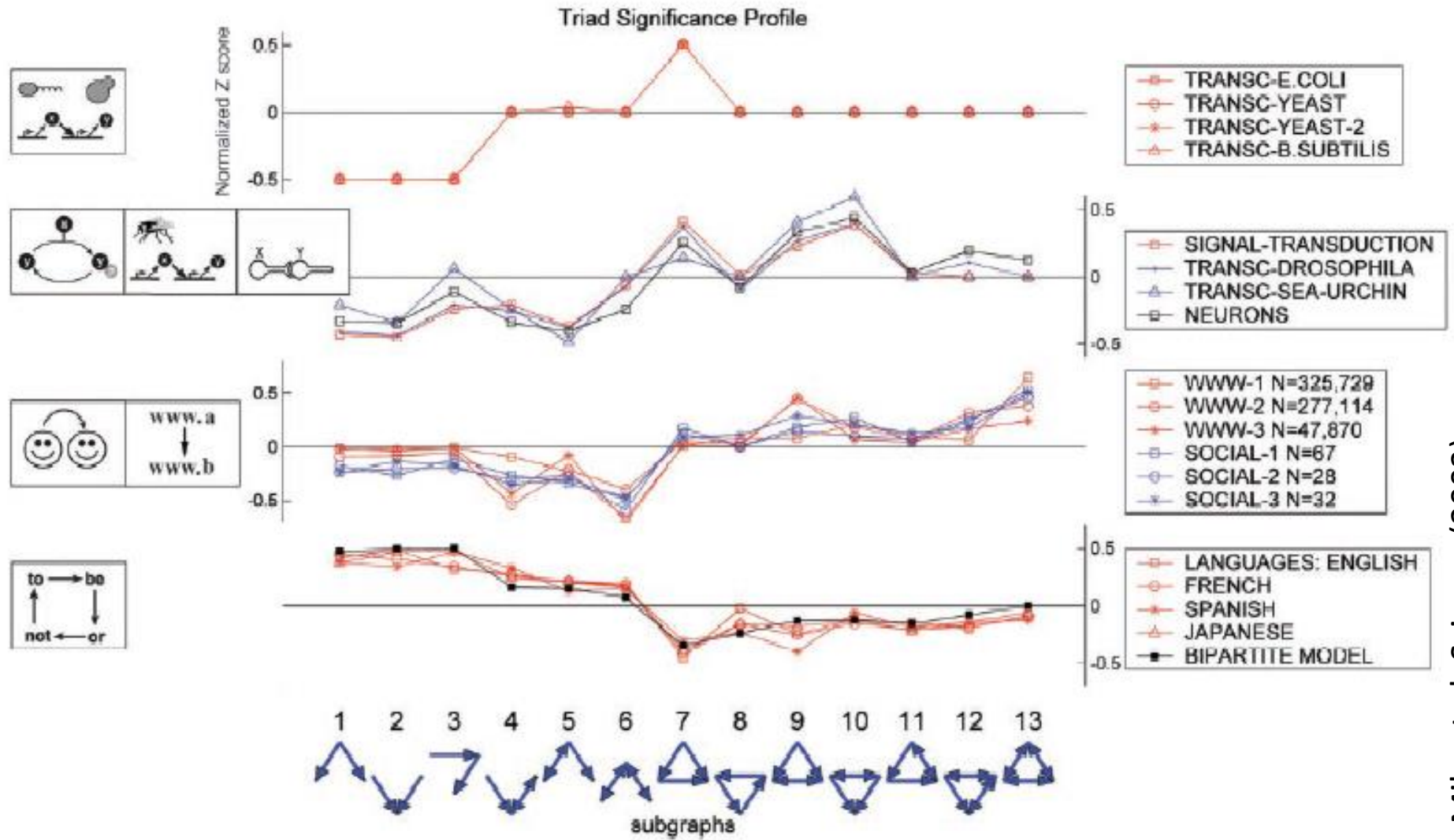
Measure: z-score
$$z_m = \frac{N_m(emp) - N_m(rnd)}{\sigma_m}$$

$N_m(emp)$ cardinality of motif m in the empirical NW

$N_m(rnd)$ average cardinality of motif m

σ_m is its standard deviation in the null model

Example for static motifs



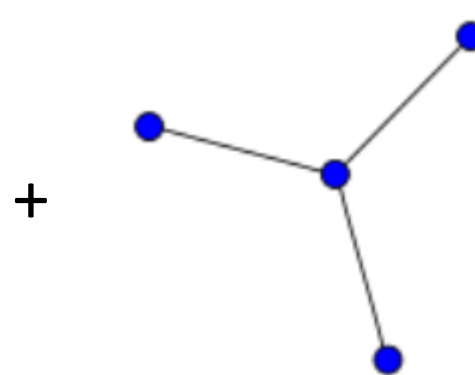
Induced subgraphs

Let's take a situation, where a star subgraph exists in the static graph under consideration:



This would contribute to the following motifs:

$$3 \times \text{---} + 3 \times \text{---}$$



Only the last one is "real", the others cause counting and interpretation problems

Only induced subgraphs should be considered!

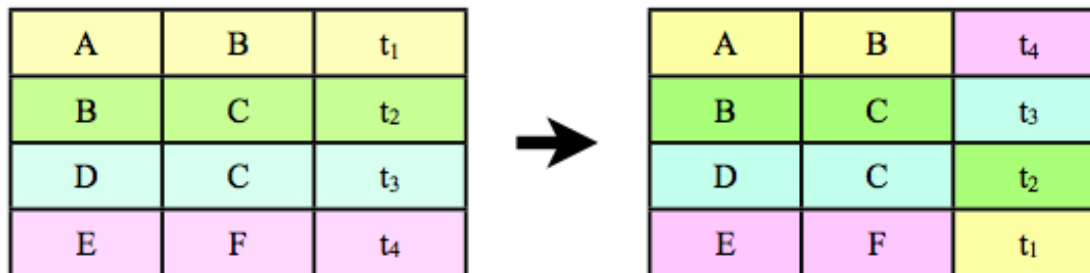
Motifs: Temporal aspects

- Time dependence of static motifs
- Daily mobility patterns
- Trigger statistics (causality)
- Temporal motifs
- Analysis of role of tagged nodes in temporal networks

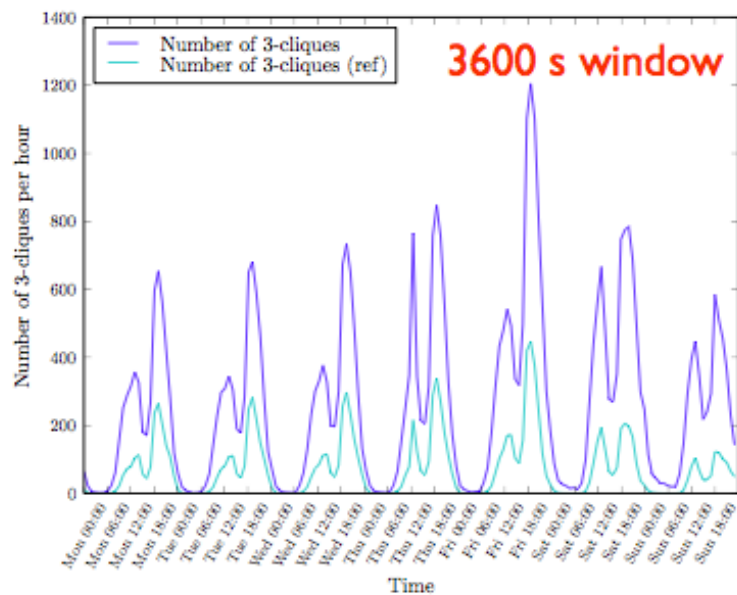
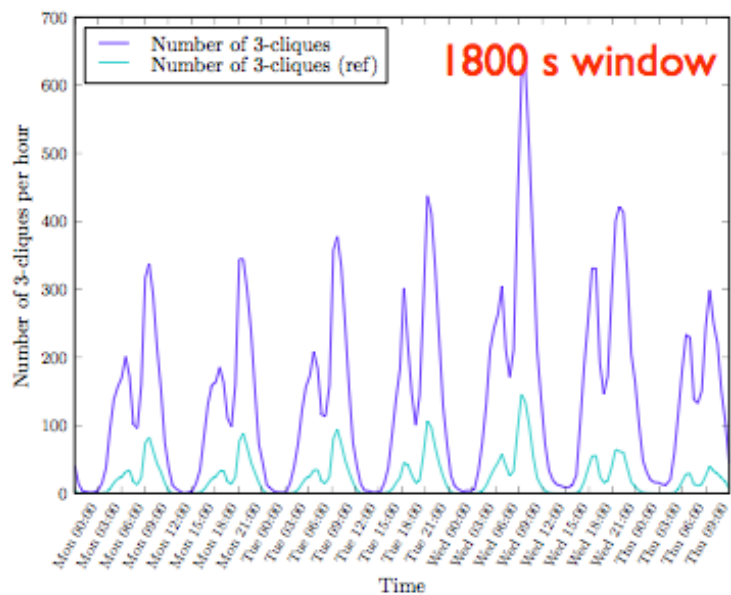
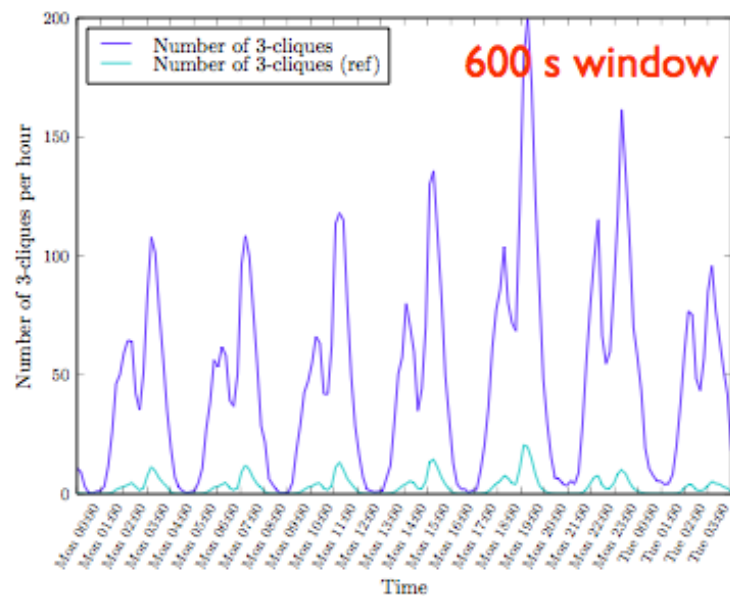
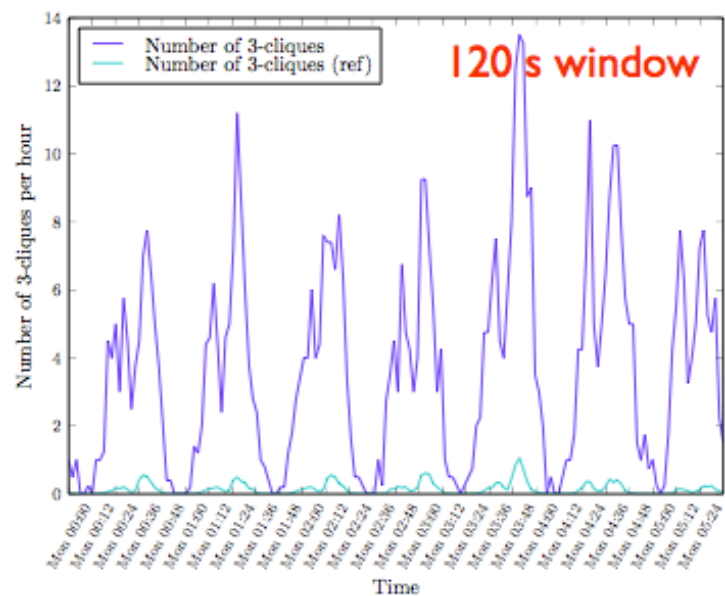
Activity counts on static motifs

Data: Mobile phone time records

- Target:
detect topological objects, where each edge occurred within a short time window
- Sliding-window counts over the whole data
- “Shuffled-times-reference”: take original event data, reshuffle all event times

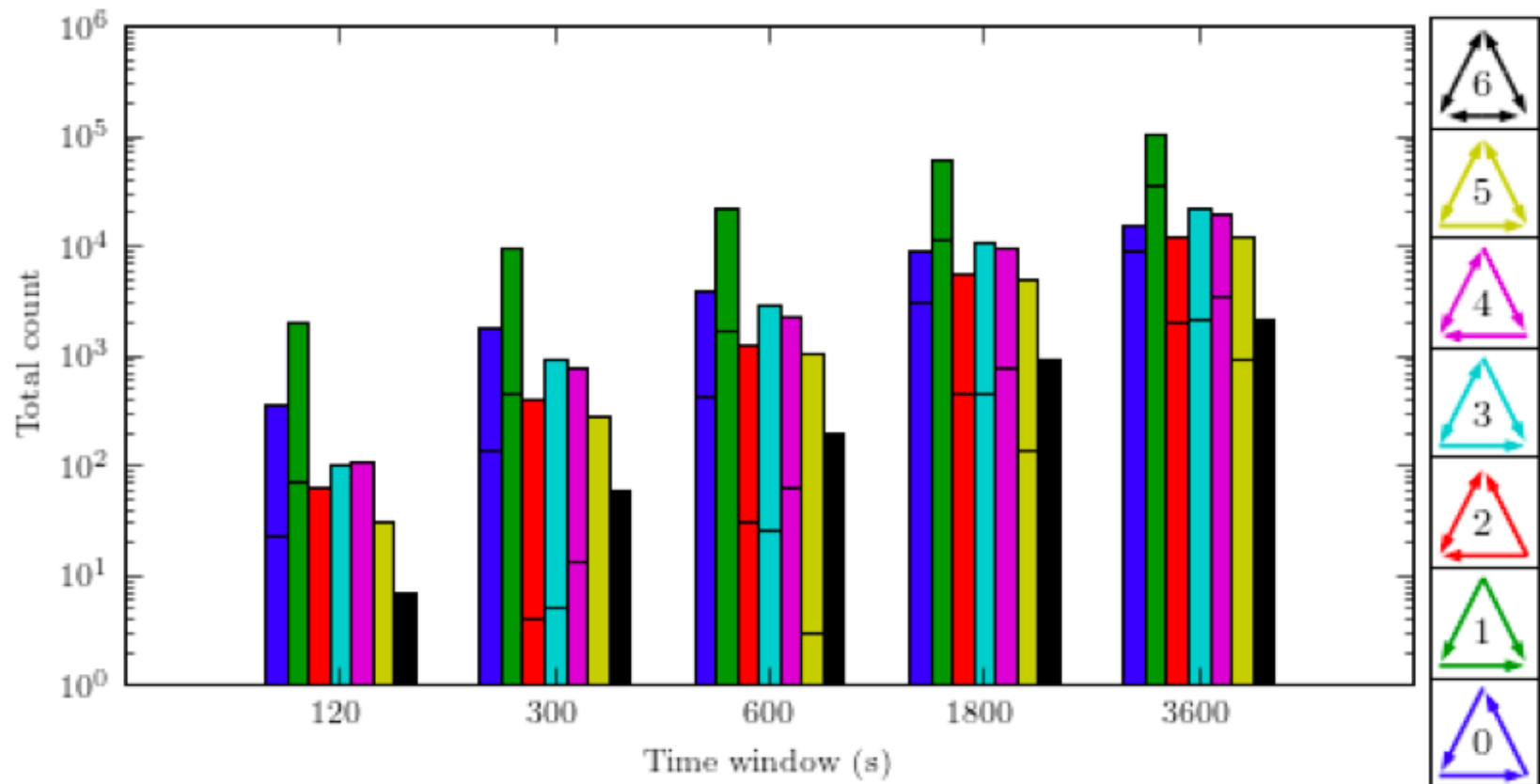


Activity counts on triangles



Activity counts on directed triangles

- Shown below: total number of directed triangles for time windows
- 7 possible cases
- Horizontal line: time-reshuffled reference



Evolution of motifs

Data from Chinese and European mobile phone services

Time stamped, who calls whom (hashed)

Problem: Which link is representing real social tie?

(And not commercial or technical calls)

Statistically validated network

How are static motifs present in the aggregate form in time? What is the characteristic time scale?







Change of the participation of nodes in the largest connected component

		Original				Bonferroni			
		Daily	Weekly	Monthly	All	Daily	Weekly	Monthly	All
Ch	Mean (%)	29.33	68.62	81.47	85.69	0.41	48.96	72.50	79.12
	SD (%)	8.53	2.66	2.55	—	0.26	4.82	2.47	
EU	Mean (%)	11.45	75.85	93.89	98.85	0.018	34.53	81.35	96.79
	SD (%)	3.69	1.77	0.44	—	0.008	4.79	2.38	—

Strong effect, underlining the importance of filtering
 Giant component exists in the original but not in the
 filtered nw. As time windows grows giant comp. emerges.







Relevance of morifs

Examples of evolution of overrepresented motifs (Ch)

	μ	σ	μ_{ref}	σ_{ref}	
110 	0.4 (+)	0.04	3.88e-6	1.7e-5	daily
238 	0.07 (+)	0.01	1.03e-6	1.07e-5	
<hr/>					
110 	1.55 (+)	0.12	2.89e-5	6.32e-6	weekly
238 	0.61 (+)	0.07	3.06e-7	3.36e-7	
<hr/>					
110 	2.16 (+)	0.07	7.41e-5	6.86e-6	monthly
238 	1.39 (+)	0.06	1.12e-6	6.83e-7	

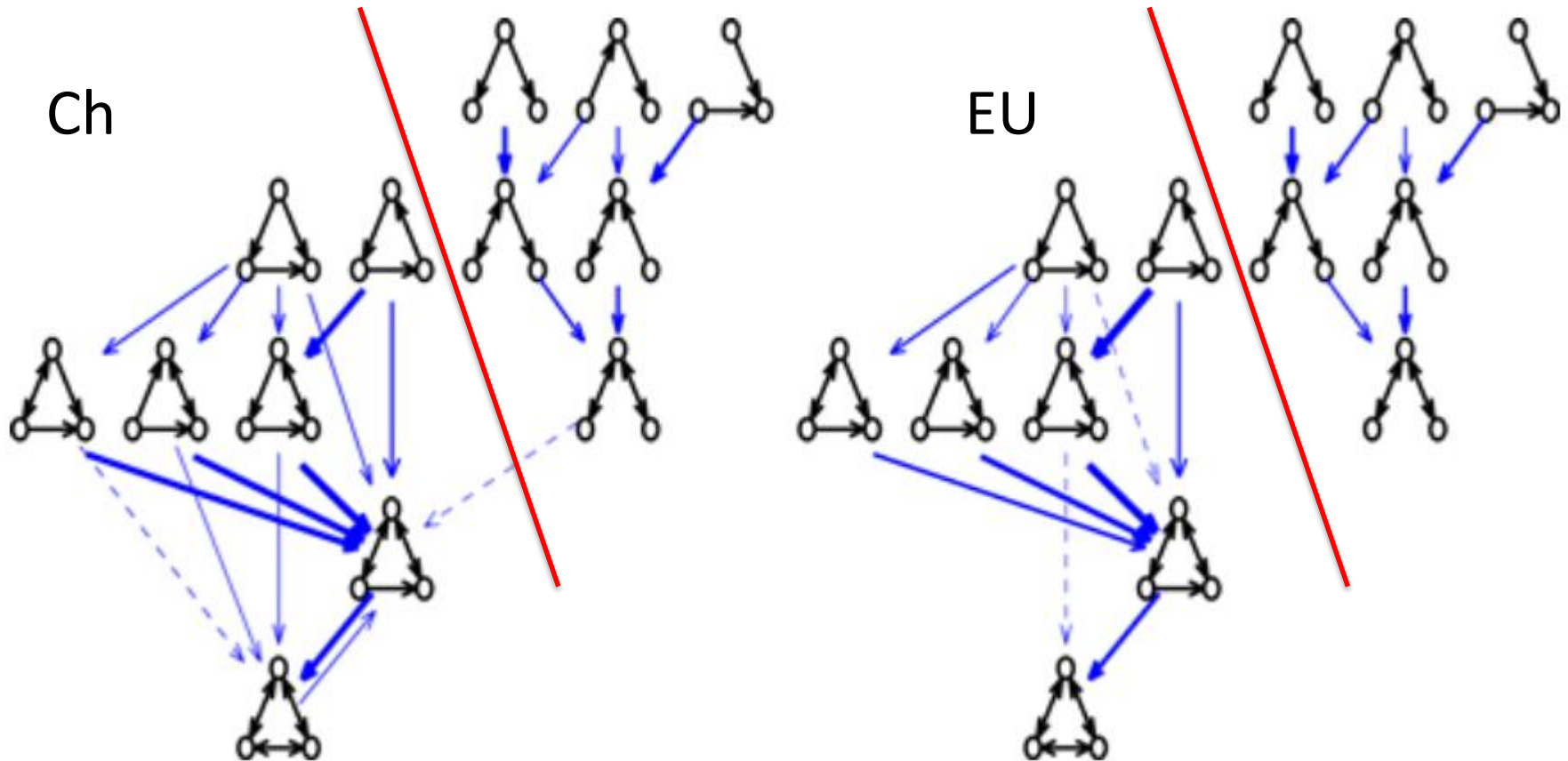
ref: shuffled network, keeping in/out degrees and bidirectional links

Examples of evolution of underrepresented motifs (Ch)

	μ	σ	μ_{ref}	σ_{ref}	
74 	12.58	0.34	13.28	0.38	daily
78 	3.45	0.22	3.84	0.25	
<hr/>					
74 	19.6 (-)	0.41	20.85	0.44	weekly
78 	10.29 (-)	0.68	12.43	0.87	
<hr/>					
74 	22.47 (-)	0.16	23.55	0.15	monthly
78 	16.35 (-)	0.34	20.3	0.45	

Correlations and evolution of motifs:

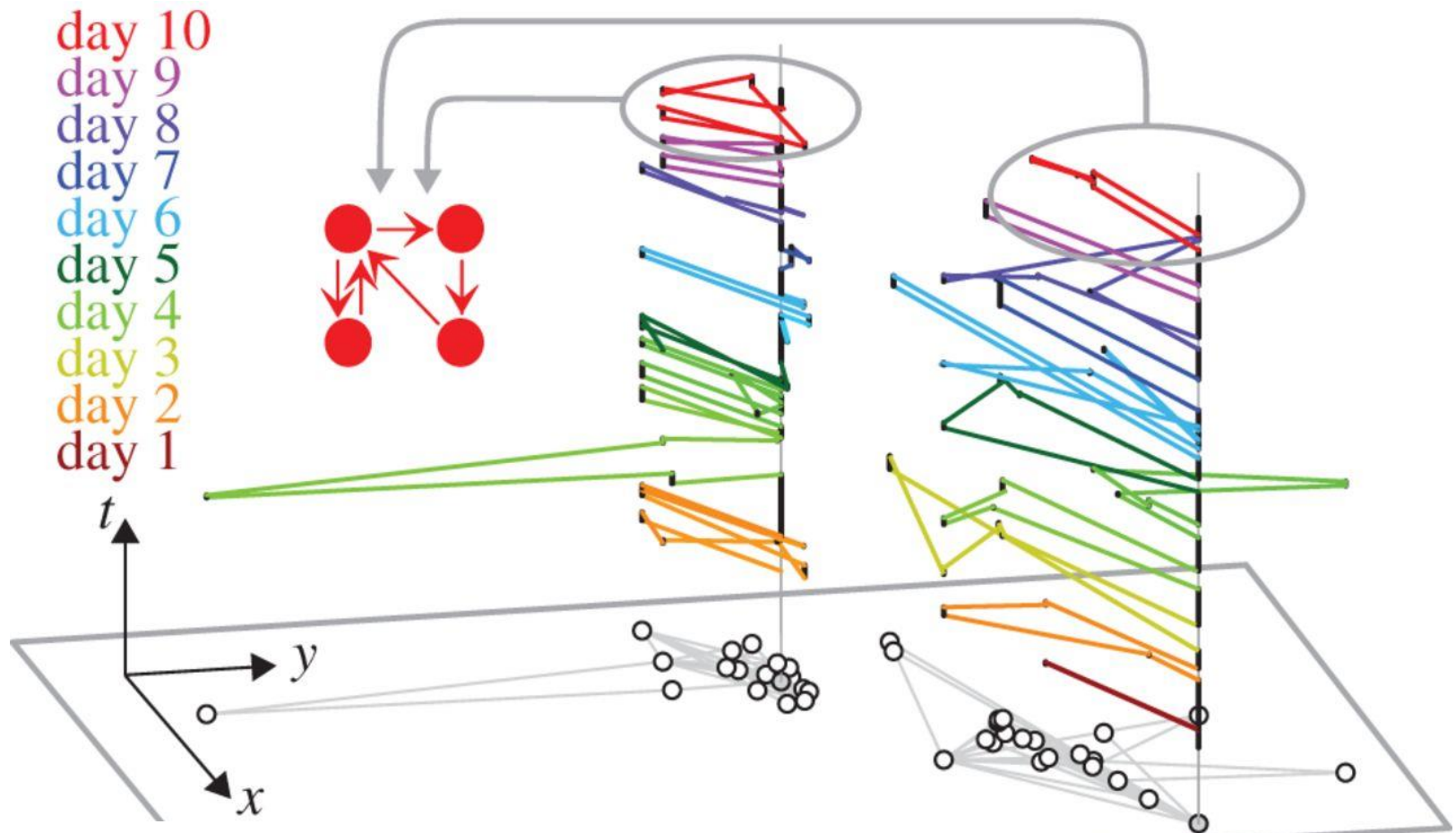
arrows indicate conditional probabilities from day Monday \rightarrow Monday + Tuesday



Closed triangles form on an intraday scale!

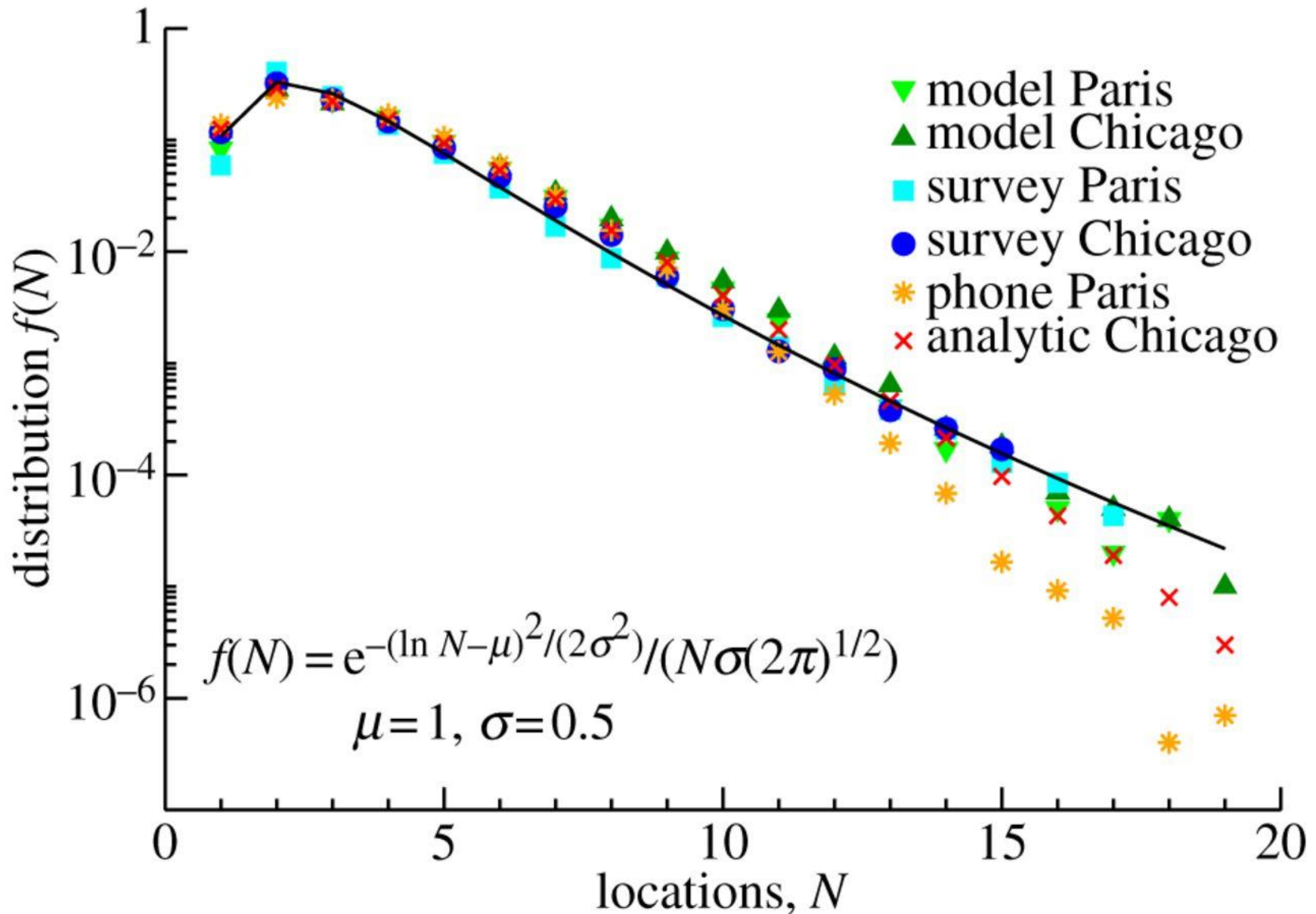
Mobility patterns

Data: Mobile call records with tower position, surveys (Paris, Chicago)



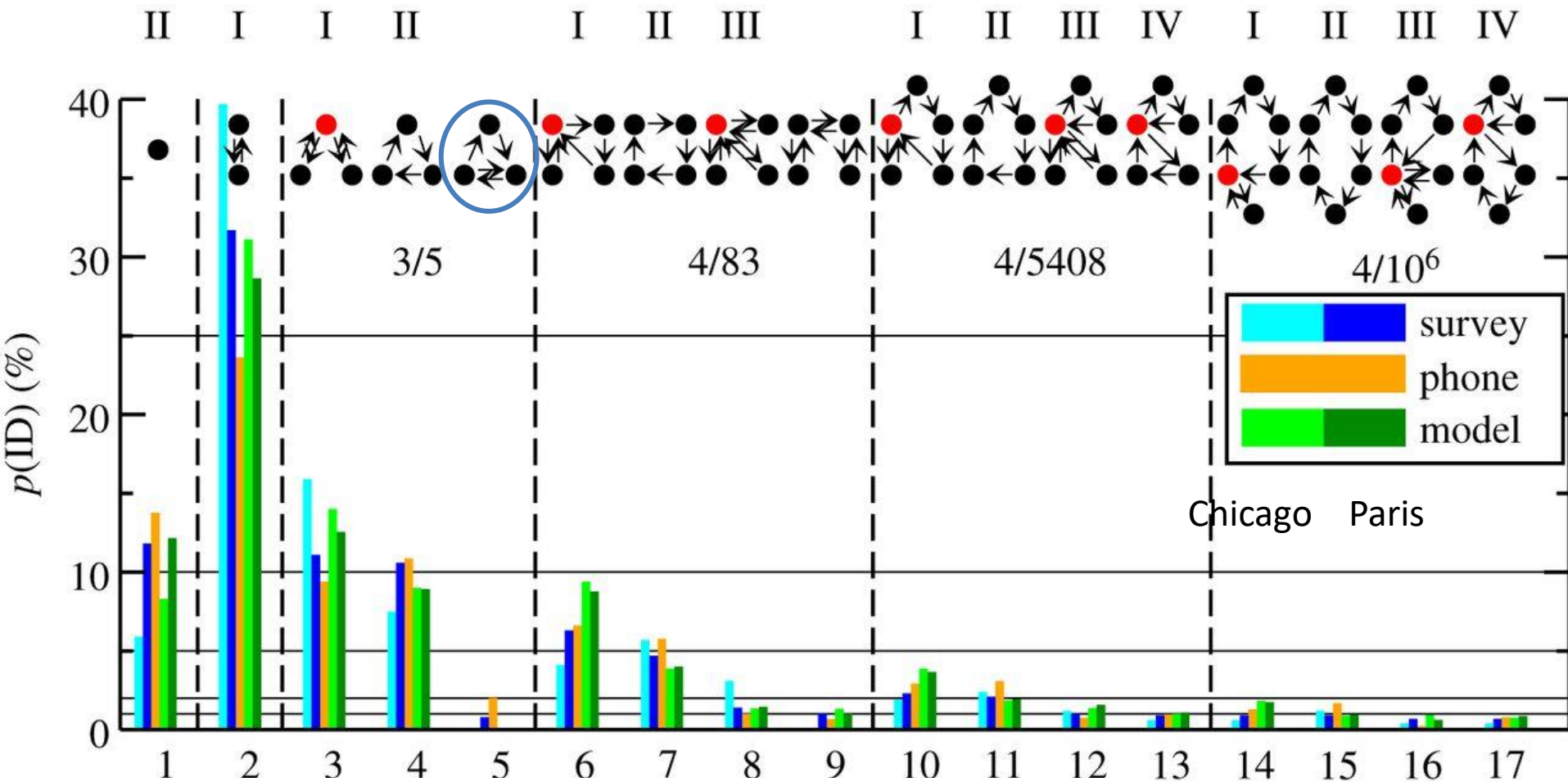
Spatio-temporal resolution

Distribution of the number of distinct visited sites



Mobility motifs – these 16 motifs explain 90%

Groups: # nodes. Numbers: N_0/N_f with N_0 the # of observed and N_f the total number of Eulerian paths.

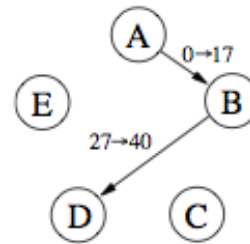


Mostly Eulerian cycles – home, sweet home

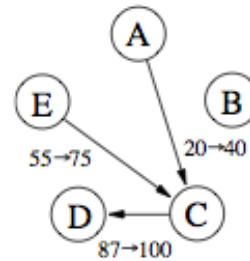
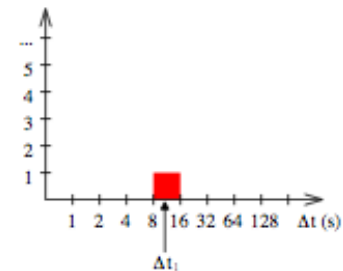
Action triggers (order important)

Data: Mobile call records

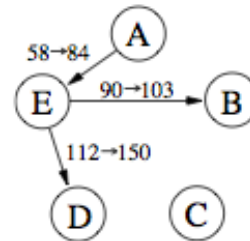
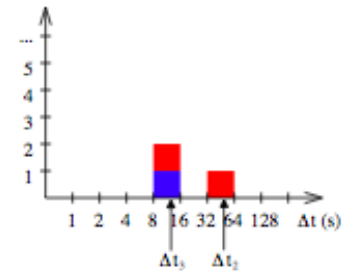
- Motivation: detect “causal” chains of A calling B, who then calls A or C
- Construction:
 - i) take an outgoing event ($t=t_2$),
 - ii) take earlier incoming event(s) ($t=t_1$),
 - iii) increase event counter at $\Delta t=t_2-t_1$
- Do this for all outgoing events



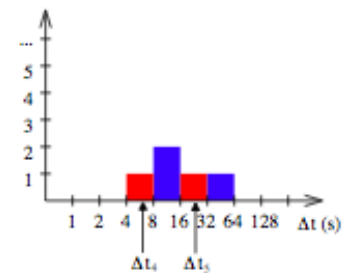
$$\Delta t_1 = 27 - 17 = 10 \text{ s}$$



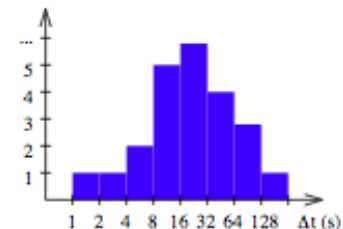
$$\Delta t_2 = 87 - 40 = 47 \text{ s}$$
$$\Delta t_3 = 87 - 75 = 12 \text{ s}$$



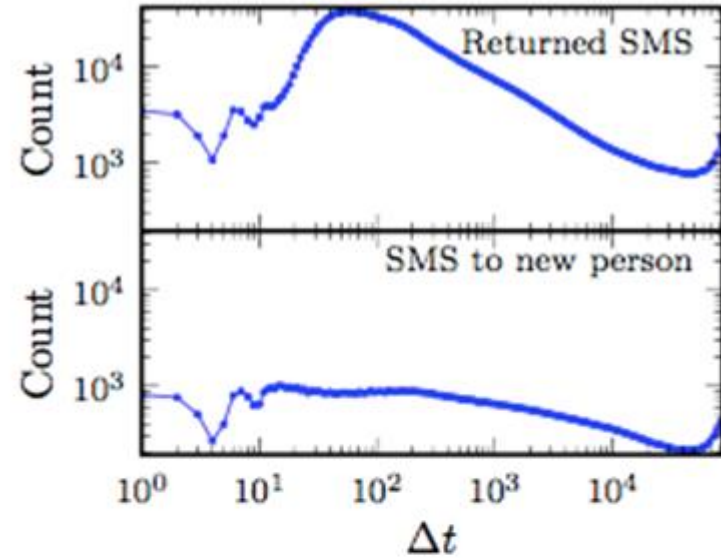
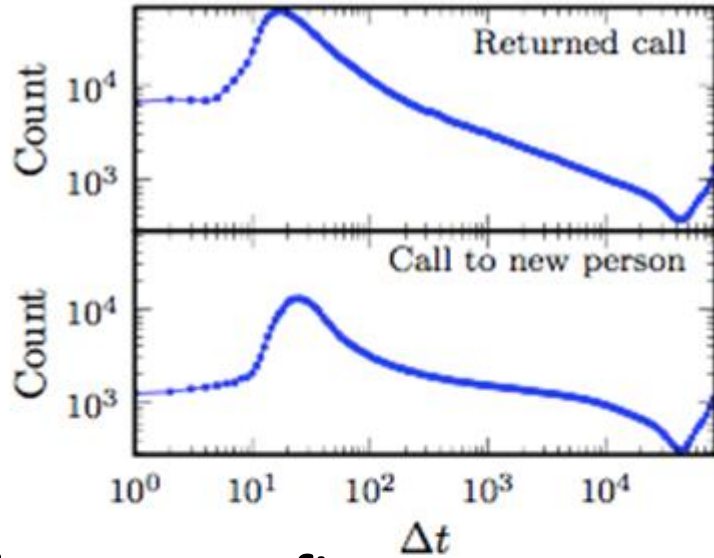
$$\Delta t_4 = 90 - 84 = 6 \text{ s}$$
$$\Delta t_5 = 112 - 84 = 28 \text{ s}$$



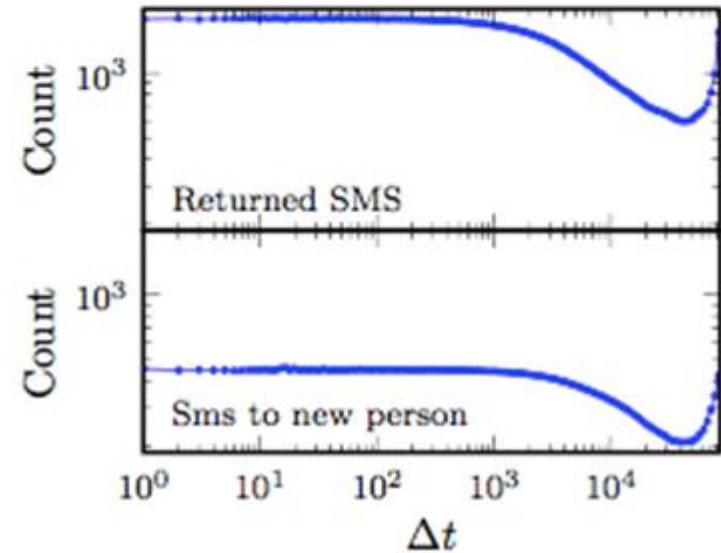
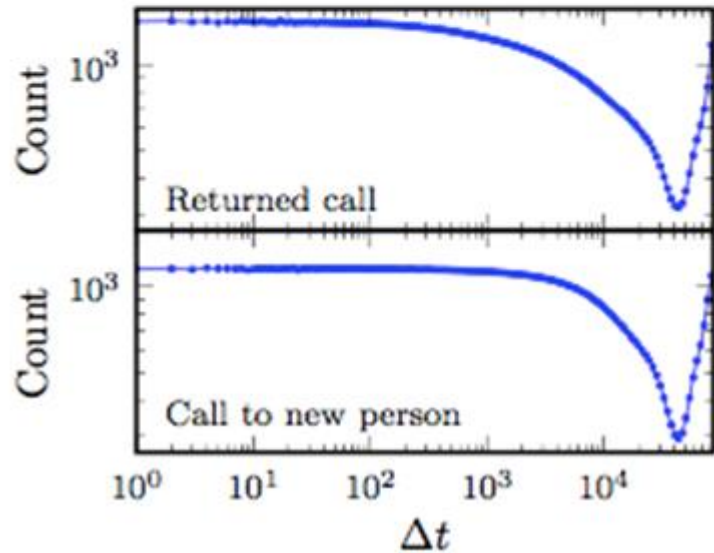
⋮



Action triggers: characteristic reaction time



Ref: average first response



Temporal motifs (formalism)

s_{ij} and s_{jk} are Δt -adjacent events if their time difference is not longer than Δt .

s_{ij} and s_{nm} are Δt -connected events if there is a sequence of Δt -adjacent events connecting i and m .

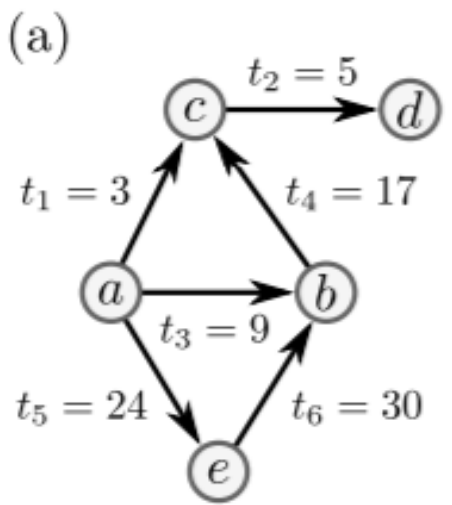
(There is no ordering requested, m is not necessarily reachable from i .)

A temporal subgraph with respect to Δt is set of events, which are mutually Δt -connected.

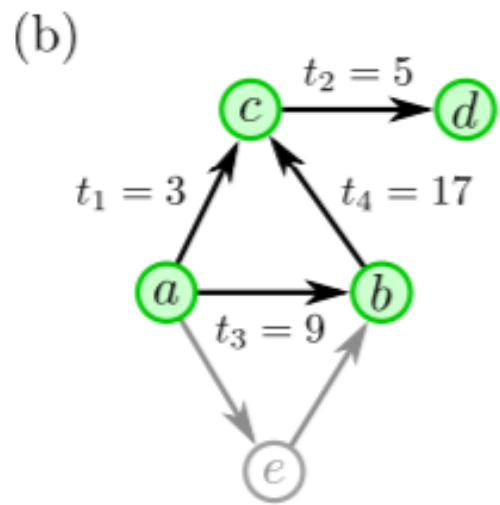
A temporal subgraph is valid if no event is skipped at any node to construct it. It is "consecutive".

A temporal motif is a set of isomorphic valid temporal subgraphs, where isomorphism is defined with respect to the order of events.

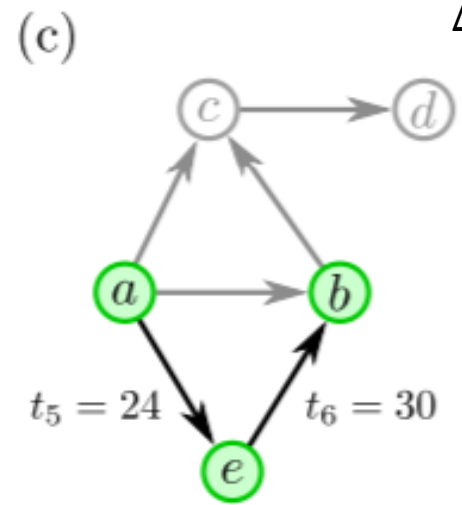
$\Delta t = 10$



$E = \{e_1, \dots, e_6\}$



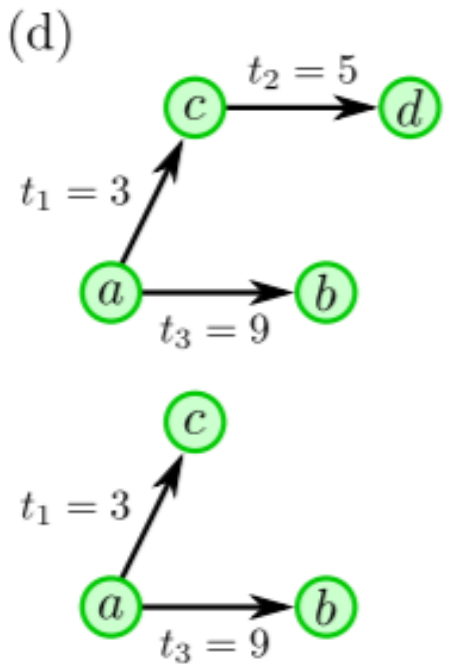
$E^*_{\max} = \{e_1, e_2, e_3, e_4\}$



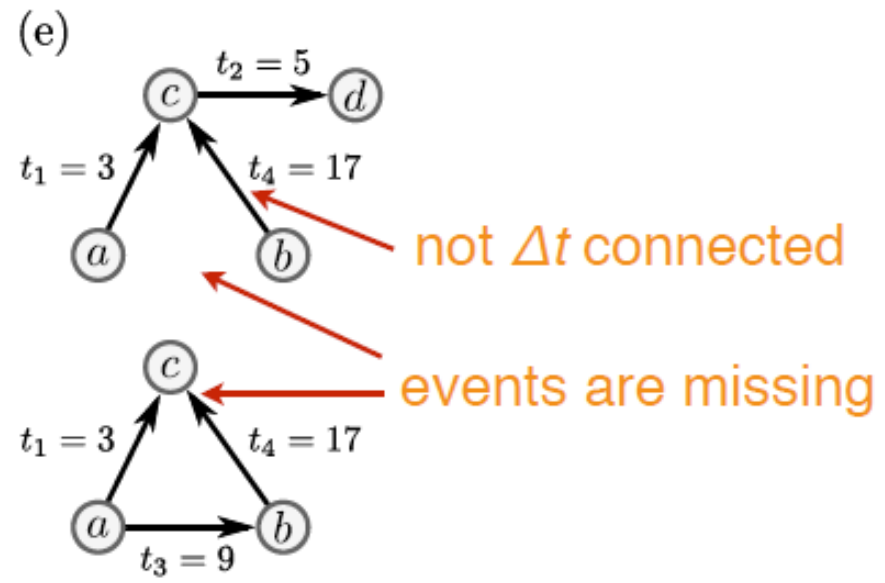
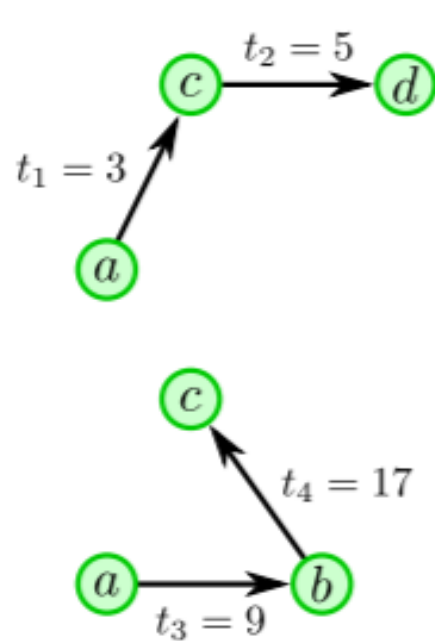
$E^*_{\max} = \{e_5, e_6\}$

a) Temporal graph (no durations)

b), c) Maximal subgraphs

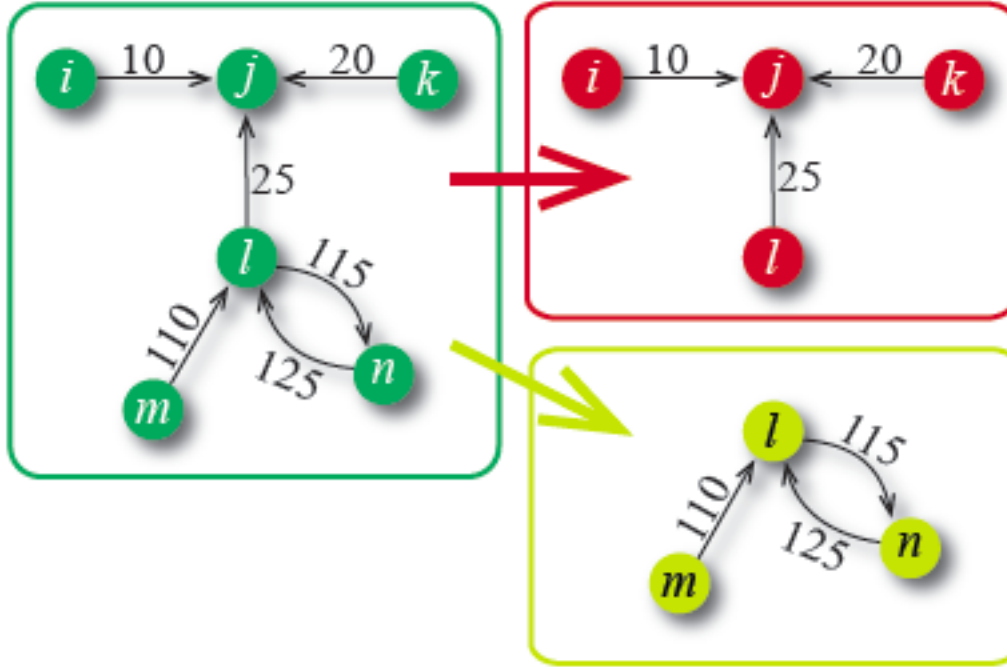


d) Valid subgraphs



Non-valid subgraphs in (b)

Maximal temporal motifs



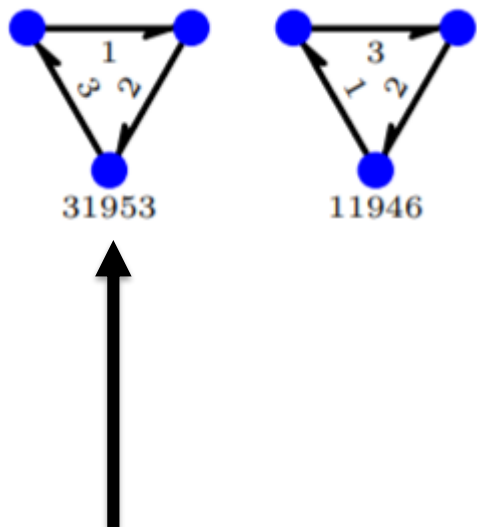
$$\Delta t = 15$$

Two maximal temporal subgraphs (max. set of pairwise Δt -conn events)

Temporal motifs based on maximal subgraphs are **maximal temporal motifs**.

Importance of a temporal motif is measured by its cardinality.

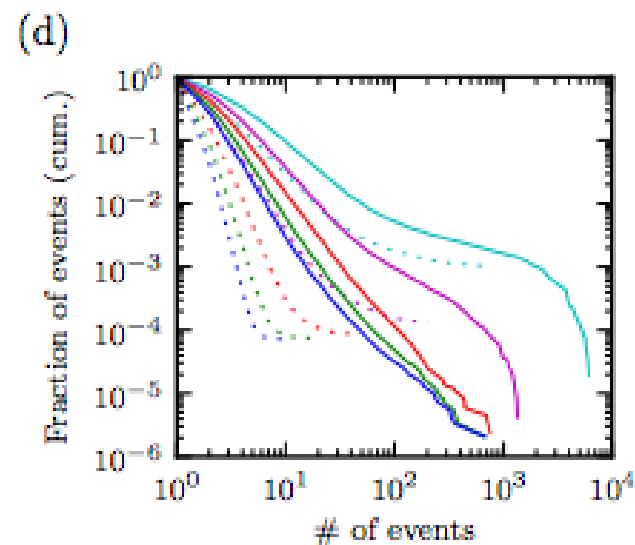
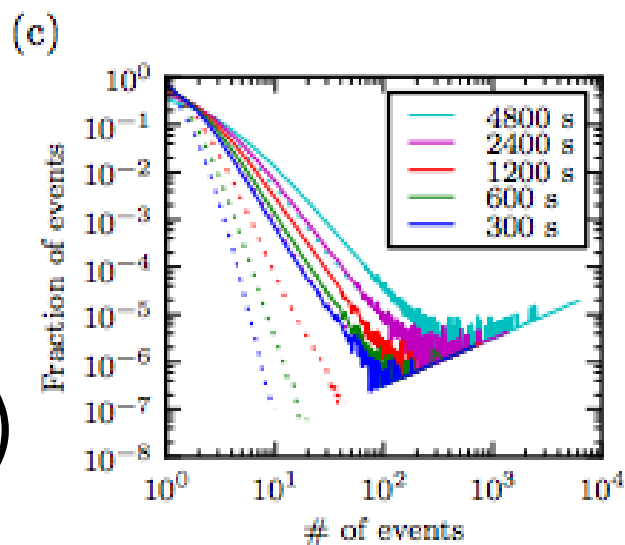
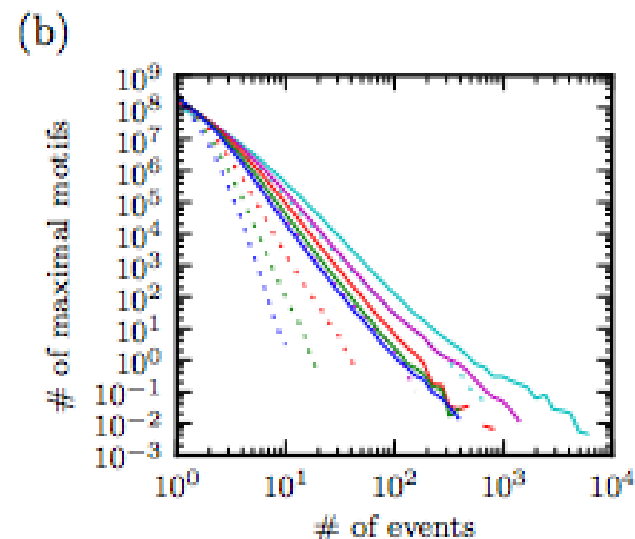
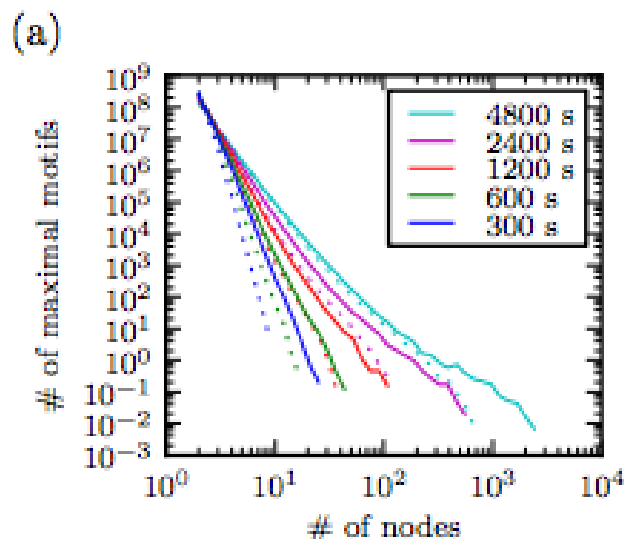
Results on maximal temporal motifs (different Δt -s)



“Causal”

Mobile
phone data
($\Delta t = 10$ min)

Reference systems?



Null models

The comparison of the empirical data with a statistics on a null model tells whether the properties of the null model give a good null hypothesis.

(E.g., strong deviations from the configuration model suggest that topological correlations are important for static models.)

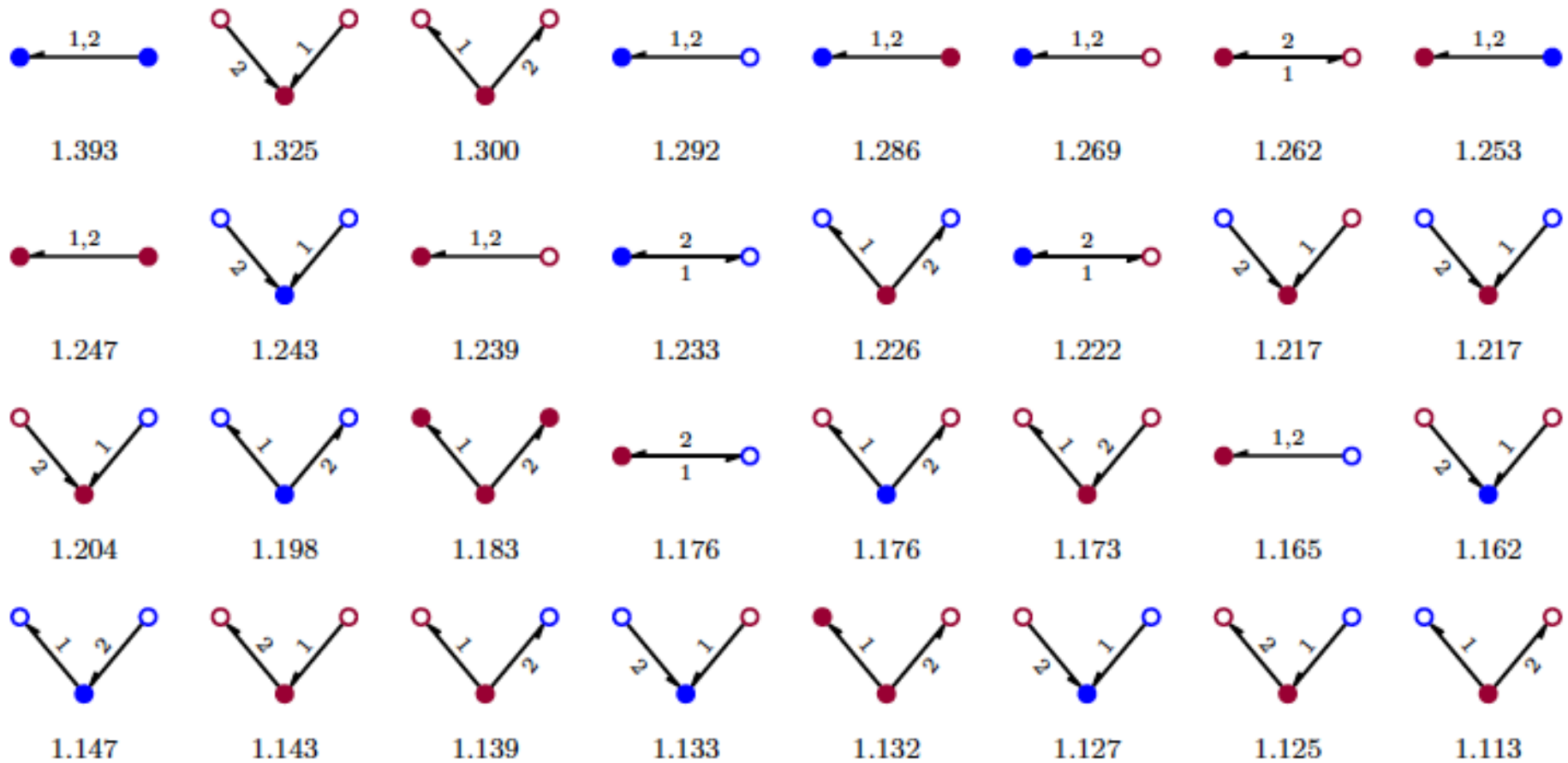
Simple time shuffling leads to relevance of too many motifs.

Better: Check relevance of temporal aspects for node properties (gender, age, type of user): Colored temporal networks

Simple randomizing the types of nodes does not give a good null model for their role, since weight may play a role.

A proper null model can be constructed if the weight distribution of the aggregate network is taken into account when randomizing the colors. The null model is created by counting the motifs assuming dependence only on edge weight but not on node type.

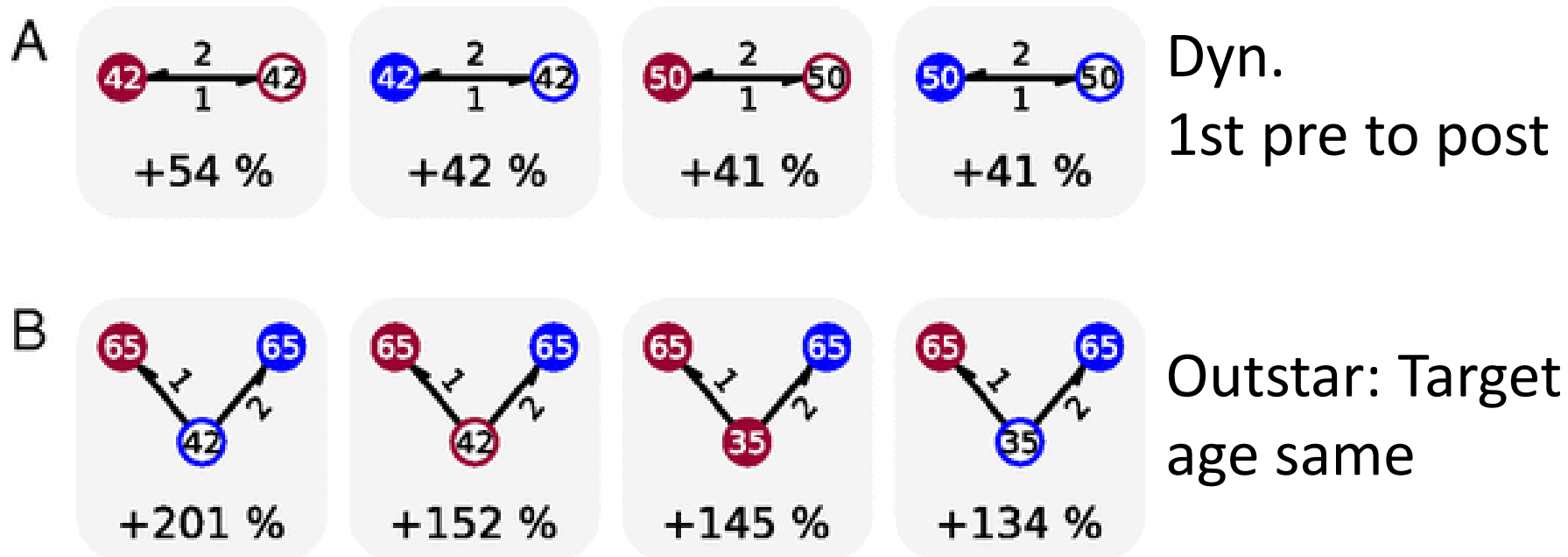
Results on motifs as compared to null model



Most frequent motifs



Example of temporal effects



 Female, 42 ± 2 years old, prepaid user

 Male, 50 ± 2 years old, postpaid user

Observations:

- Clear indication of temporal homophily. Very strong for prepaid – socioeconomic background
- Outstars with same category of target are overrepresented
- Chains and stars overrepresented for females
- Local edge density correlates with temporal overrepresented motifs (temporal Granovetter effect)

Summary

Temporal networks are important for dynamic processes on complex networks if links are defined by the events and events happen inhomogeneously in time and/or the sequence of events is crucial

Temporal networks are defined with respect to a time window of observation.

Many concepts can be generalized:
path, distance, connectivity, motifs etc.

Motifs: static in evolution, mobility, temporal

Burstiness has a decelerating effect on spreading (!)

Broad field of applications