Freie Universität Berlin

Class 7:
Considering time in networks

Course: Computational Network Analysis

Prof. Dr. Claudia Müller-Birn

Institute of Computer Science, «Human-Centered Computing»

Mar 3, 2016

Recap

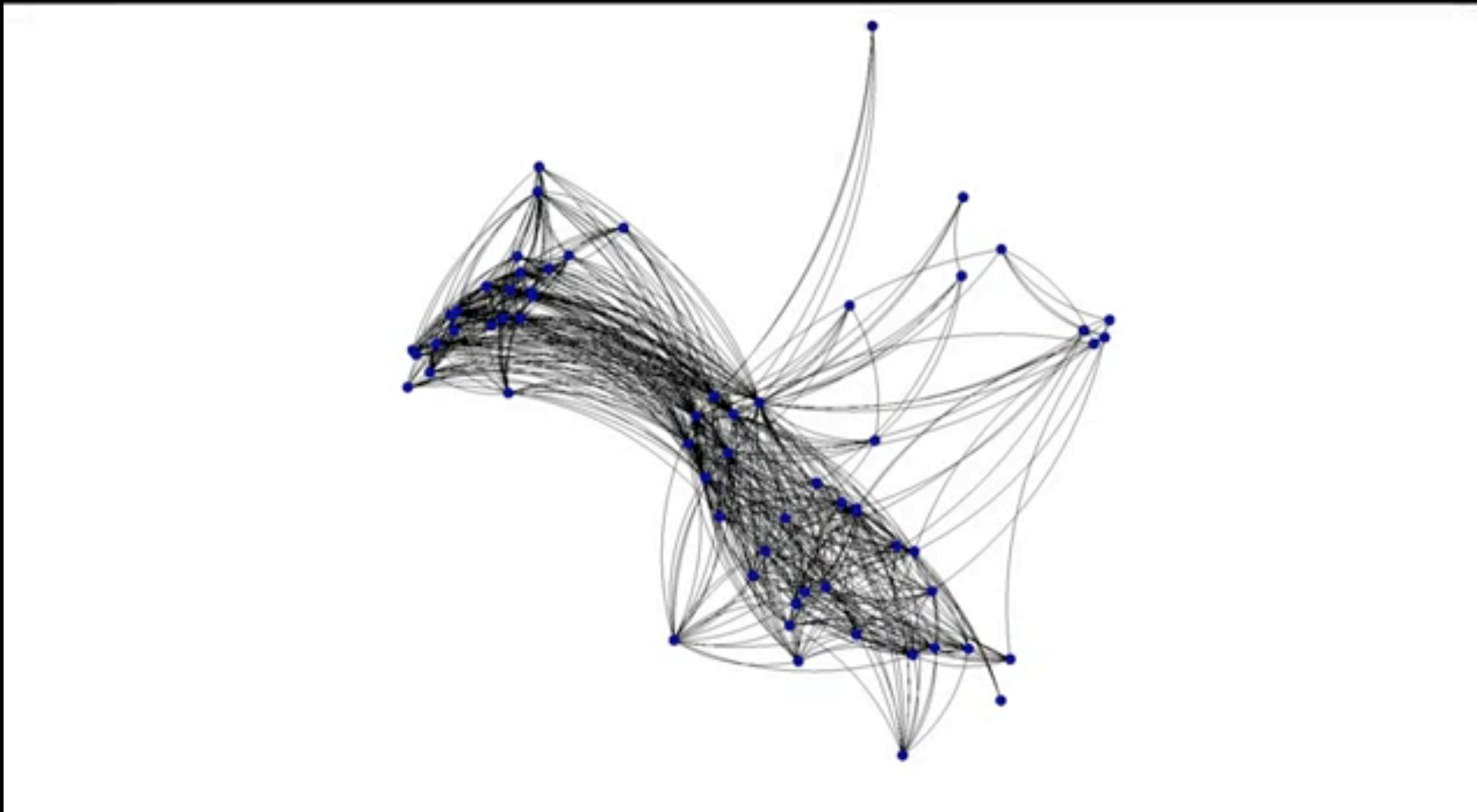
- You can describe the differences between degree, eigenvector, closeness, and betweenness.
- You know how to calculate all centrality measures on a vertex, normalized vertex and network level.
- You can explain for what questions you would use which measure.

Today’s outline

- Introduction
- Describing temporal networks
- Discrete approach(es)
- Continuous approach
- The difference

Motivation

- Most of the analysis we have discussed so far has been done on aggregated networks
- Time has not been kept into account
- **But why does this matter?**



https://www.sg.ethz.ch/research/dynamic-networks/

Motivation

- Most of the analysis we have discussed so far has been done on aggregated networks
- Time has not been kept into account
- But why does this matter?
- Examples for temporal relations
  - Timestamps, e.g. Facebook: friends added and removed over time
  - Duration, e.g. spending time close to other bees
  - Frequency, e.g. function calls
  - Time order, e.g. timetables in public transport systems

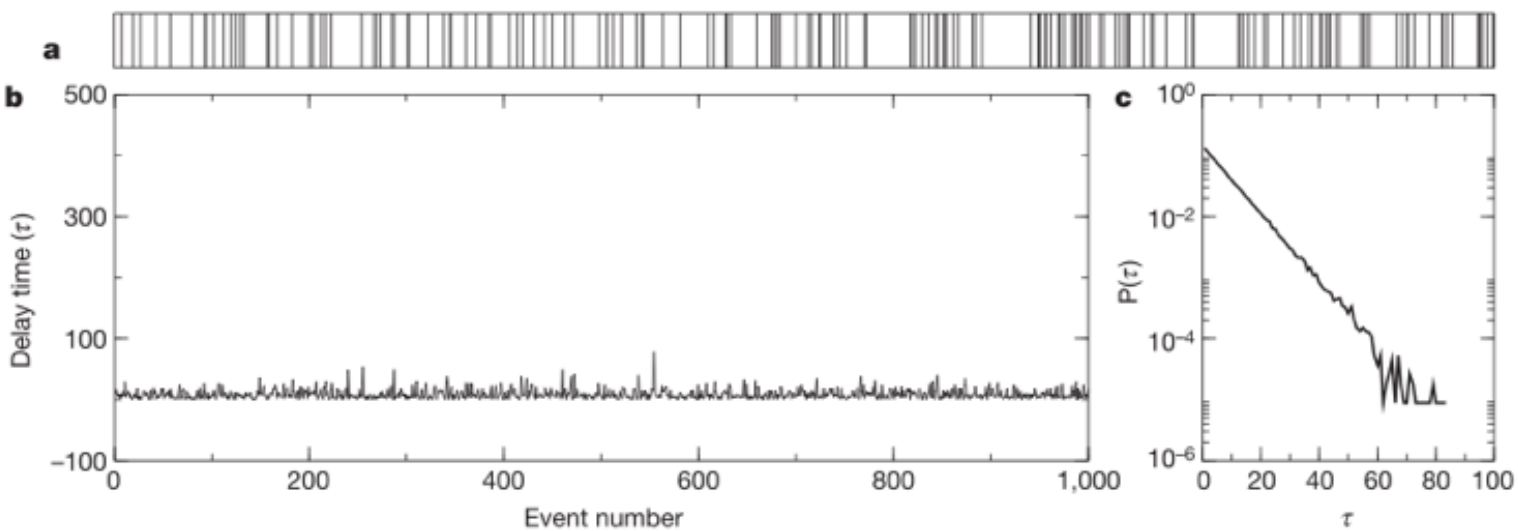
Defining a model

- It is not sufficient to collect network information at different moments and then to subsume them (Doreian, 1997)
- Events are not deterministic - they are stochastic
- For example, if two or three measuring points are selected randomly, then no or only unimportant events might occur
- The goal should be to define a model that is able to present all events in the data

Bursts in human dynamics

Barabasi, A.-L. 2005, 'The origin of bursts and heavy tails in human dynamics', Nature, 435, 7039, 207--211.

- Current models of human activity are based on Poisson processes, and assume that in a dt time interval an individual (agent) engages in a specific action with probability qdt, where q is the overall frequency of the monitored activity
- This model predicts that the time interval between two consecutive actions by the same individual, called the waiting or inter-event time, follows an exponential distribution

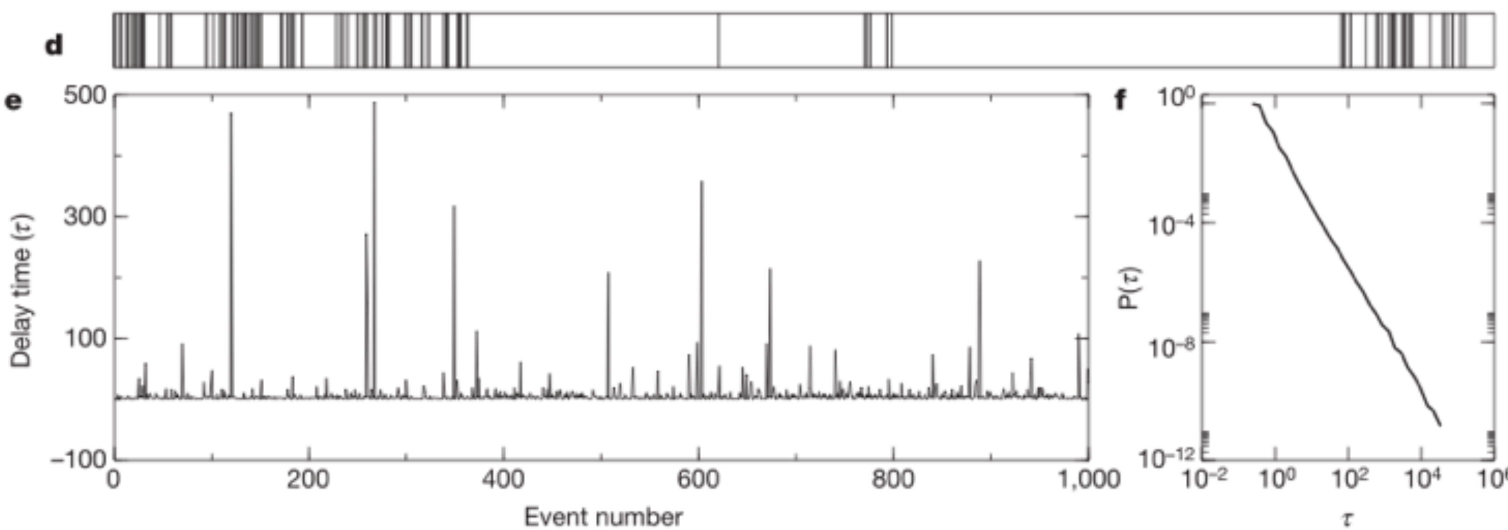


a) Succession of events predicted by a Poisson process. The horizontal axis denotes time, each vertical line corresponding to an individual event; b) Plot shows the delay times t for 1,000 consecutive events, the size of each vertical line corresponding to the gaps seen in a); c) Probability of finding exactly n events within a fixed time interval

Bursts in human dynamics

Barabasi, A.-L. 2005, 'The origin of bursts and heavy tails in human dynamics', Nature, 435, 7039, 207--211.

- An increasing number of recent measurements indicate that the waiting or inter-event times are being better approximated by a heavy tailed or Pareto distribution
- The slowly decaying, heavy-tailed processes allow for very long periods of inactivity that separate bursts of intensive activity



a) The succession of events for a heavy-tailed distribution; b) The waiting time t of 1,000 consecutive events, where the mean event time was chosen to coincide with the mean event time of the Poisson process; c) Delay time distribution

Freie Universität Berlin

Describing temporal networks

Basic approaches two consider time

Discrete approach

Cross-sectional analysis of graphs where the main focus lies on the changes of network stages

Continuous approach

Each single interaction with a start and end is considered.

Moody, J., McFarland, D., & Bender-deMoll, S. (2005). Dynamic Network Visualization. American Journal of Sociology, 110(4), 1206–1241.

Discrete approach

Definitions

- A temporal graph G_t refers to a discrete description of graph *G* that consists of a sequence of interactions

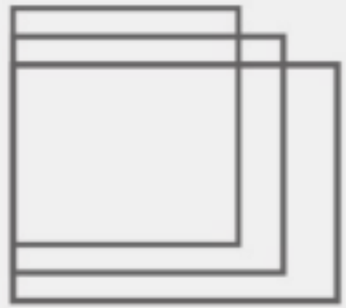
- Each element g_t of a sequence corresponds to all interactions in t

- The interval t is ranging from(t,t+1) and all interactions within this time interval are aggregated to the end of the interval t+1 in g_t

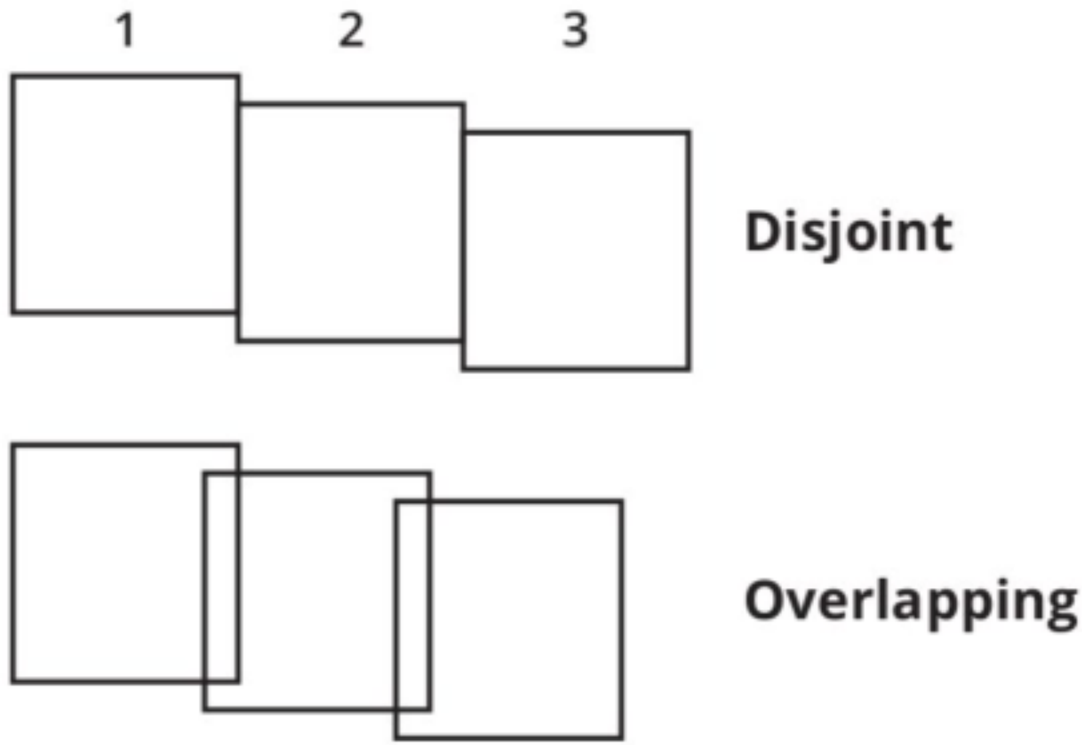
- Approaches to discretize the interactions of a graph G to g_i with i=1,…,t
  - the cumulative approach
  - the time window approach

Approaches to discretize the interactions

Cumulative approach



Time window approach



Börner, Katy, and David E. Polley. Visual insights: A practical guide to making sense of data. MIT Press (2014)

Cumulative approach

- Aggregation of all interactions in a graph G_t over a period of time
- It is defined by the Cartesian product of each element of the graph sequence

$$\mathcal{G}_t = g_1 \oplus g_2 \oplus \ldots \oplus g_t = \bigoplus_{i=1}^{t} g_i = \mathcal{G}_{t-1} \oplus g_t$$

where expression on the right-hand side is the cumulative summary at time t which is the sum of the cumulative summary at time t - 1 and the network activity at time step t (Cortes et al., 2003)

Please note: In graph theory the Cartesian product of two graphs G and H is the graph denoted by
G×H whose vertex set is the (ordinary) Cartesian product V(G)×V(H) and such that two vertices (u,v) and (u′,v′) are adjacent in G×H if and only if u is adjacent to u′ and v is adjacent to v′.

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

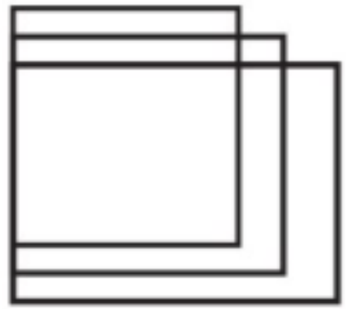
Cumulative approach

- The graph has an infinite memory of all active edges in the past and grows up to its final size which is equal to the static network's size at this time
- Even though, changes in the graph structure are visible, the sequence and existing interdependencies in network processes are not identifiable

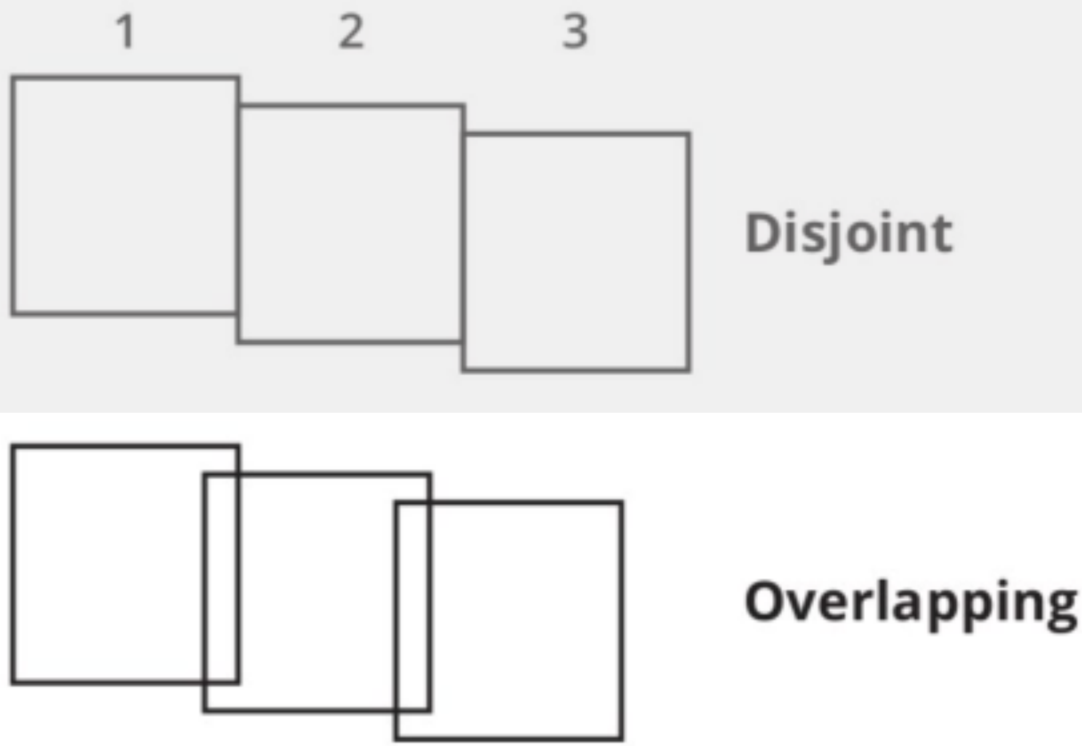
Moody, J., McFarland, D., & Bender-deMoll, S. (2005). Dynamic Network Visualization. American Journal of Sociology, 110(4), 1206–1241.

Approaches to discretize the interactions

Cumulative approach



Time window approach



Börner, Katy, and David E. Polley. Visual insights: A practical guide to making sense of data. MIT Press (2014)

Time window approach

- Analysis of a small period of time of a network or sub-network

- Main objective is to identify the impact of previous events on following events
  - A sequence of events can be evaluated based on time of occurrence and rhythm
  - Events can be specific attributes of vertices or edges as well as can be positional changes

- The goal is not to define a time interval that is preferably small but a time interval that contains the amount of edge information necessary to describe existing interactions
  - The size of the time frame depends on the number of occurred events
  - Only events in a specific time frame are aggregated to the network

- This simple approach detaches network data completely from each other

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

Example

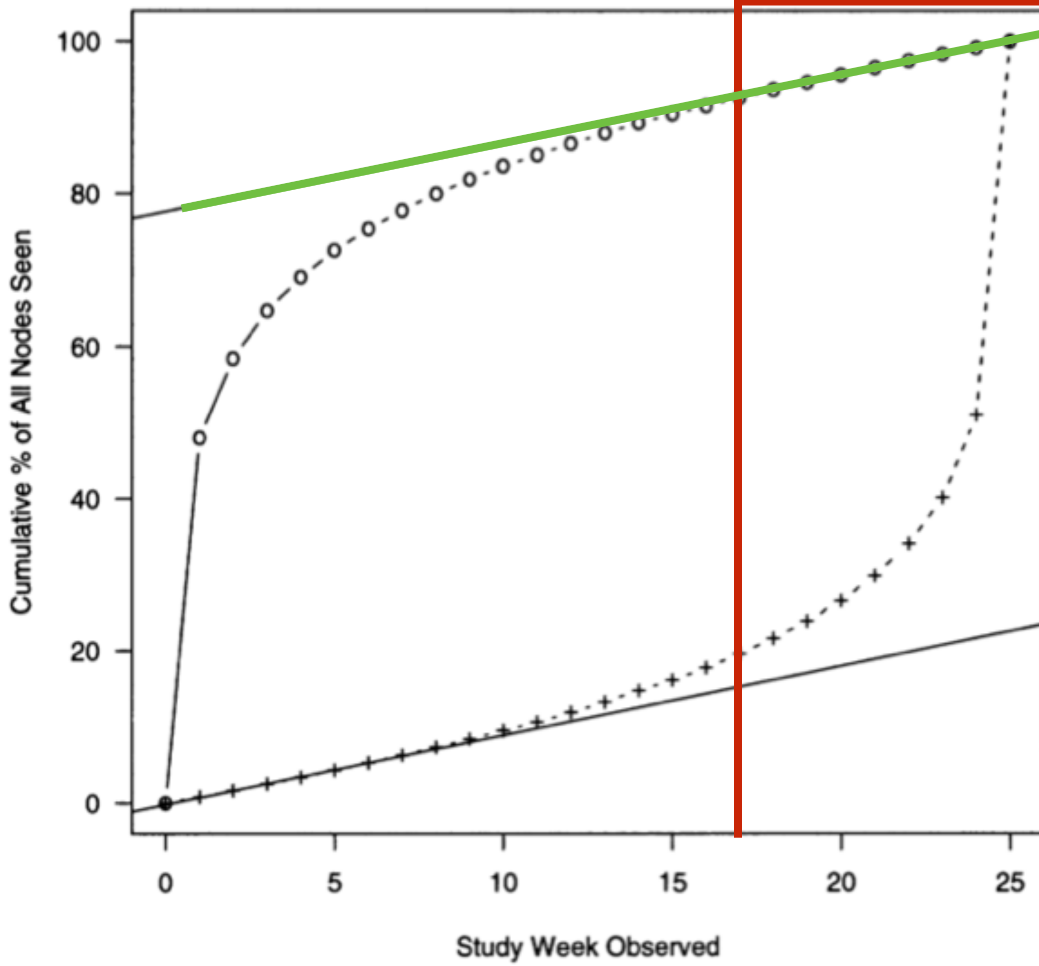
- Calls carried on a large telecommunications network
- Large number of nodes (hundreds of millions user accounts) and edges (several billion edges in an average week)
- Sparse connectivity
- Dynamics have stable macro effects but substantial variation in micro effects
- Data represent 25 successive weeks of activity
- Describing node activity:
  - *Node addition*: number of new nodes that we see in week i that we haven't yet seen through week i-1.
  - *Node attrition*: number of nodes that we see for the last time in week i

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

Node addition and attrition through time

Steady state is reached after about 18 weeks and lines are fitted to the remaining last and first 7 weeks,respectively.

The slopes correspond to addition and attritrion rates of just under1%.



Upper curve: shows the cumulative percentage of unique nodes in seen each study week that had not been seen before

Lower curve: shows the cumulative percentage of nodes seen for the last time in each study week

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

Basic parameters

- ***Size*** *of the time window*
  - How can the optimal time window be chosen?
  - One possible approach is to check the density over time; if it is somehow stable, the optimal time window should have this size
- ***Type*** *of time window*
  - Which time window type does offer better results, the overlapping or non-overlapping window? How can the size of the overlap be defined?
- ***Edge weight*** (absolute vs. normalized weight)
  - How should such a weight be defined?
  - For example, in a commit contribution network, one possible approach is the quotient of the number of contributions of a single author and the sum of all contributions of all authors in this period of time

Defining the size of the time interval

Approach 1

- The size of a time interval is based on the duration of an activity in a network
- For example, in an email communication network, the duration of an activity is determined by the mean time between the first and the last mail that has been sent with the same subject

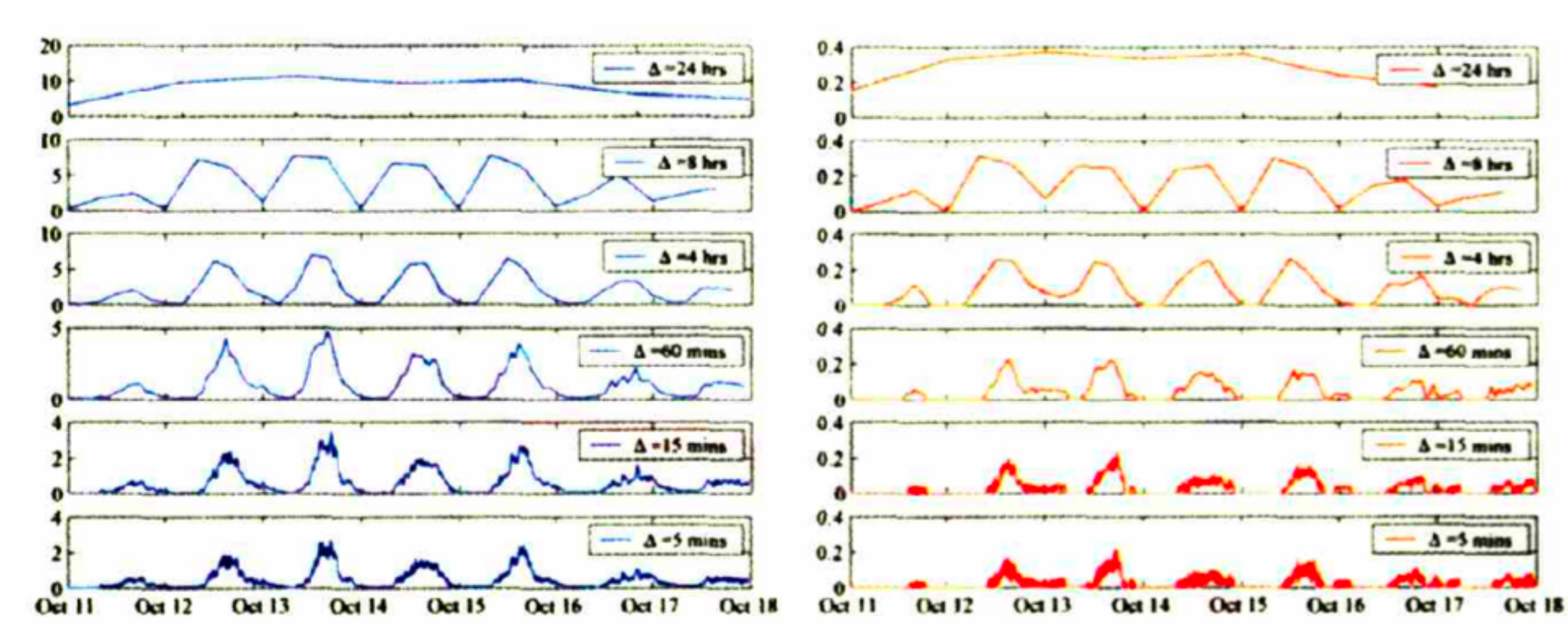
Approach 2

- The size of a time interval is chosen by comparing the properties of graphs on different time intervals
- The goal is to determine a size that exhibit as much information possible (or the other way around that looses as less information as possible) and that is as stable as possible

Kossinets, G., & Watts, D. J. (2006). Empirical Analysis of an Evolving Social Network. Science, 311(88).

Example for approach 2

mean degree k **mean clustering coefficient C**



- Calculated over a function of time for delta={1440,480,240,60,15,5}(minutes) during the week of 11 October through 17 October for the core 66 subjects
- Note: as d grows, undersampling clearly washes out higher frequency fluctuations

(Eagles, 2005)

Basic parameters

- ***Size*** *of the time window*
  - How can the optimal time window be chosen?
  - One possible approach is to check the density over time; if it is somehow stable, the optimal time window should have this size

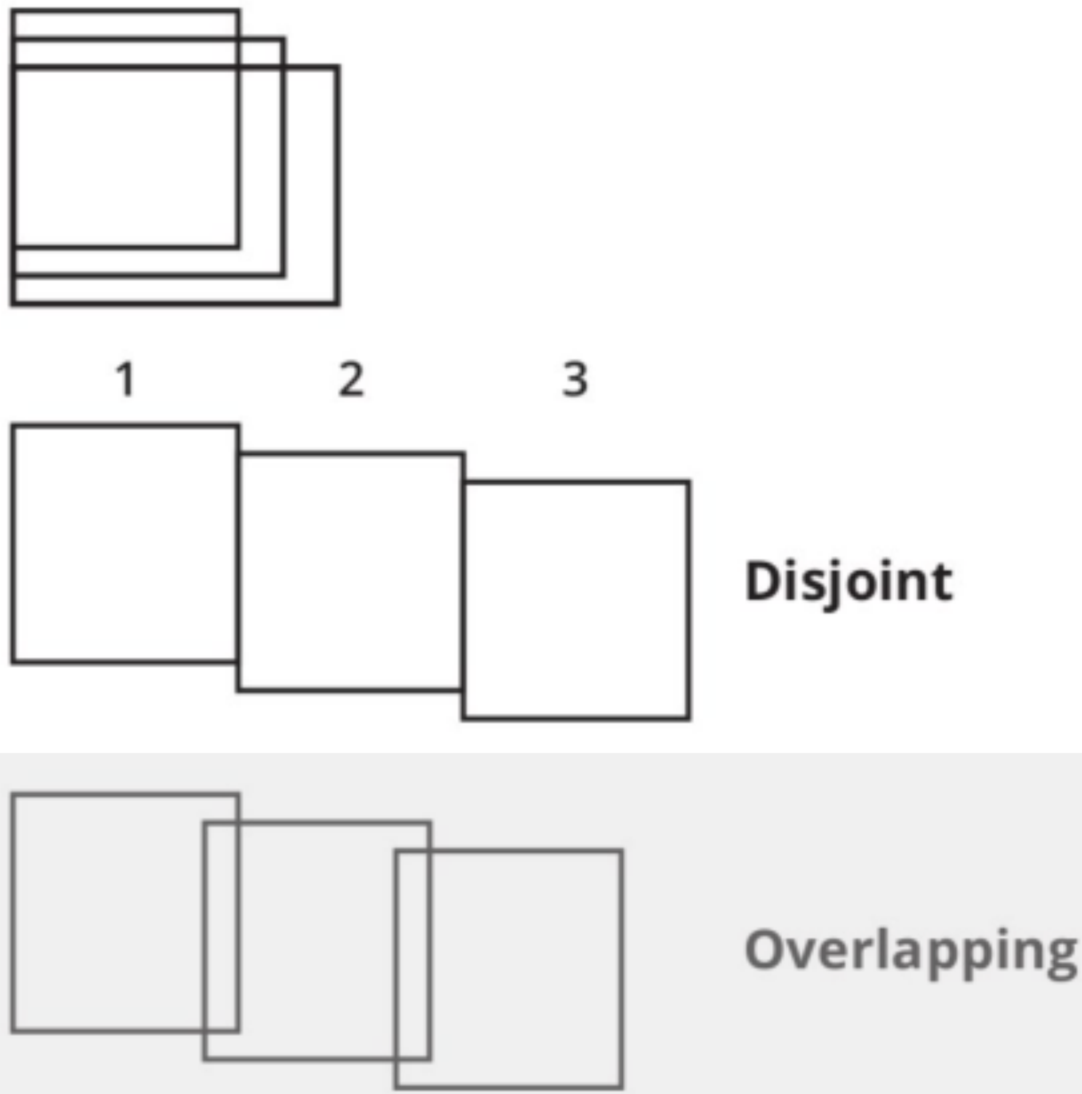
- ***Type*** *of time window*
  - Which time window type does offer better results, the overlapping or non-overlapping window? How can the size of the overlap be defined?

- ***Edge weight*** (absolute vs. normalized weight)
  - How should such a weight be defined?
  - For example, in a commit contribution network, one possible approach is the quotient of the number of contributions of a single author and the sum of all contributions of all authors in this period of time

Approaches to discretize the interactions

Cumulative approach

Time window approach



Börner, Katy, and David E. Polley. Visual insights: A practical guide to making sense of data. MIT Press (2014)

Defining the overlap of time windows

- The time-span of the sliding window can be called the “relevancy horizon” τ
- The relevancy horizon τ defines which past events are considered within the time window
- Example for defining the size of τ
  - In an email network it can be based on a number of people who have exchanged at least one email within the defined horizon
  - The instantaneous strength of an edge in an email network can be determined by the geometric average of the bilateral exchange within a time window of width τ

$$w_{ij}(t,\tau) = \sqrt{m_{ij}m_{ji}}/\tau$$

where m_{ij} and m_{ji} are the number of emails a person i and j have exchanged within the period(t-τ,t].

Defining the overlap of time windows (cont.)

- Often it is desirable to blend network activity in a way that discounts the past in favor of recent behavior
- It is opposed to the former approach which simply uses an moving average (past sequences are equally weighted)

Hill, S., Agarwal, D. K., Bell, R., & Volinsky, C. (2006). Building an Effective Representation for Dynamic Networks. Journal of Computational and Graphical Statistics, 15(3), 584–608. http://doi.org/doi: 10.1198/106186006X139162.

Defining the overlap of time windows (cont.)

- Parameter θ determines the relative priority given to recent data
- Exponentially decreasing weights over time (called exponential smoothing) are considered with

$$w_i = \theta^{t-i}(1-\theta)$$

where

$$0 \leq \theta \leq 1$$

- Therefore, each sequence of a graph is defined as a function of its preceding sequences (recurrence relation)

$$G_t = \theta G_{t-1} \otimes (1-\theta) g_t$$

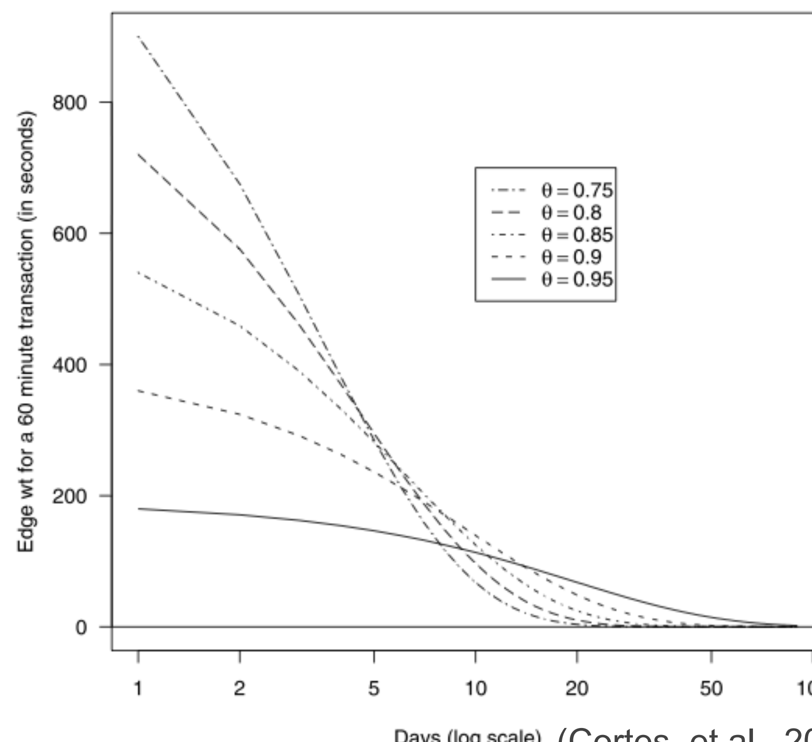
- This weight function provides a smooth dynamic evolution of G_t

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

Example

- What does the parameter θ mean?
  - The parameter θ determines the relative priority given to recent data
  - As θ near 0 puts almost all weight on the most recent time period and therefore, as θ approaches 1, we are blending in more historical data

- Another interpretation
  - θ determines the decay of the weight of a given transaction in the aggregated edge
  - The figure shows the weight over time given to one hour transaction for different values of θ



(Cortes. et al., 2003)

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

Global thresholding

- Problem: the smoothed graph described so far decays, all edges exponentially over time, but never deletes them
- To avoid maintaining a lot of edges with arbitrarily small weights (performance problems), a global threshold parameter ε can be defined that prunes all edges with weights less than ε
- Trade-off between deleting edges if they do not represent an important relationship for an entity (and yet consume storage space) and deleting information that may be relevant to an entity and can be useful in analysis

Cortes, Corinna, Daryl Pregibon, and Chris Volinsky. "Computational methods for dynamic graphs." Journal of Computational and Graphical Statistics (2012).

Basic approaches to consider time

Discrete approach

Cross-sectional analysis of graphs where the main focus lies on the changes of network stages

Continuous approach

Each single interaction with a start and end is considered.

(Moody, 2005)

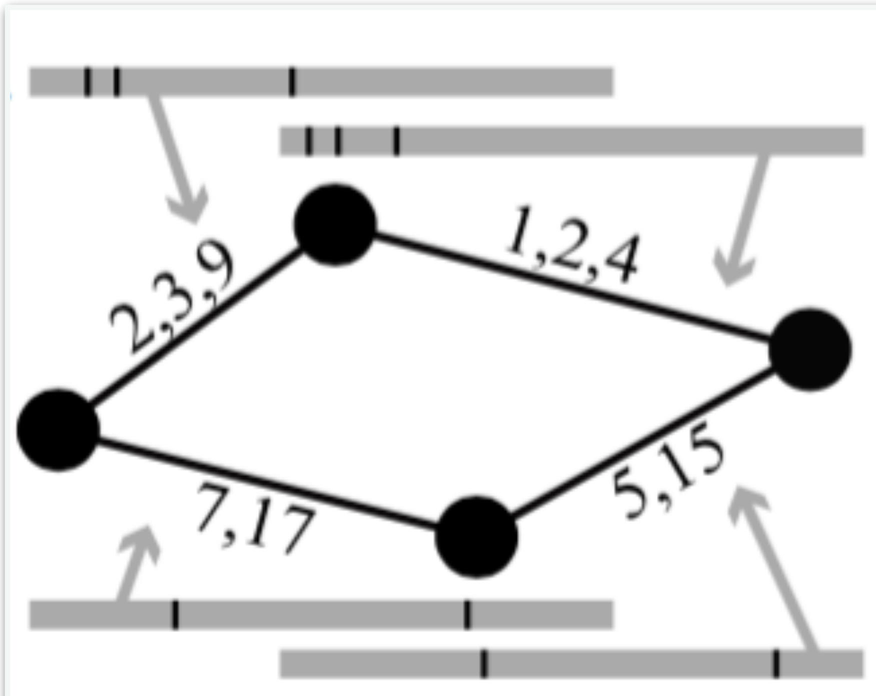
Continuous approach

Contact sequence

- There is a set of N vertices V interacting with each other at certain times
- Durations of the interactions are negligible
- A contact sequence is a set of C contacts, represented as triples (i, j, t) where i, j $\in$ V and t denotes time
- Typical systems suitable to be represented
  - Communication data (sets of e-mails, phone calls, text messages, etc.)
  - Physical proximity data where the duration of the contact is less important (e.g. sexual networks)

P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Example



P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Interval graphs

- Edges are not active over a set of times but rather over a set of intervals

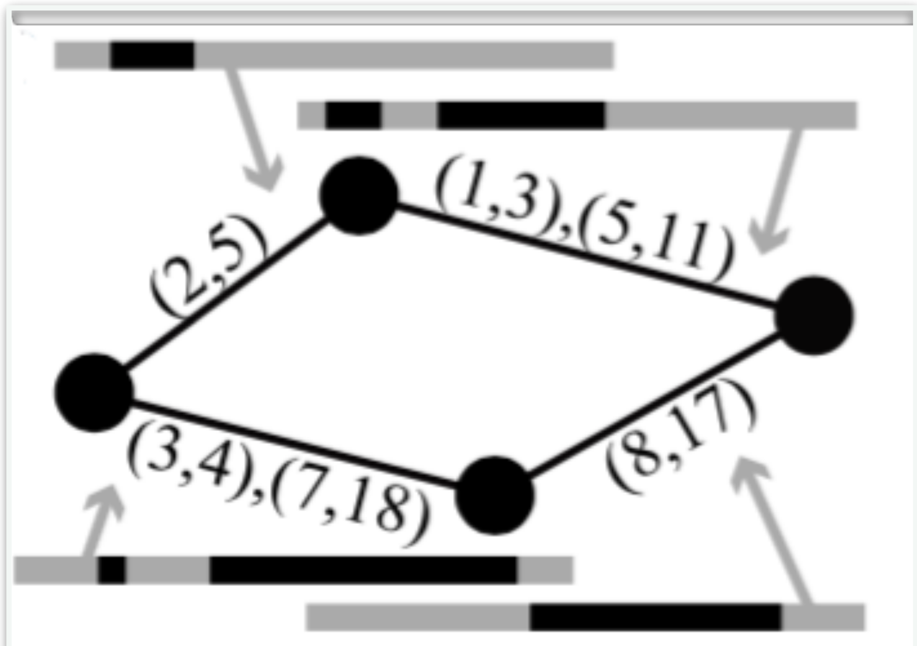
- An interval $T_e = \{(t_1, t_1'), \ldots, (t_n, t_n')\}$, where the parentheses indicate the periods of activity, where t—is the beginning and t_n' the end of the interval

- The static graph with an edge between i and j if and only if there is a contact between i and j is called the (time) aggregated graph

- Examples
  - Proximity networks, where a contact can represent that two individuals have been close to each other for some extent of time
  - Seasonal food webs, where a time interval represents that one species is the main food source of another at some time of the year

P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Example



P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Adjacency index

$$a(i, j, t) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected at time } t \\ 0 & \text{otherwise.} \end{cases}$$

- A triple of a contact sequence never occurs twice
- Thus, we can order the contacts uniquely (first by the time stamps, then by their smallest vertex index and finally by their largest vertex index)

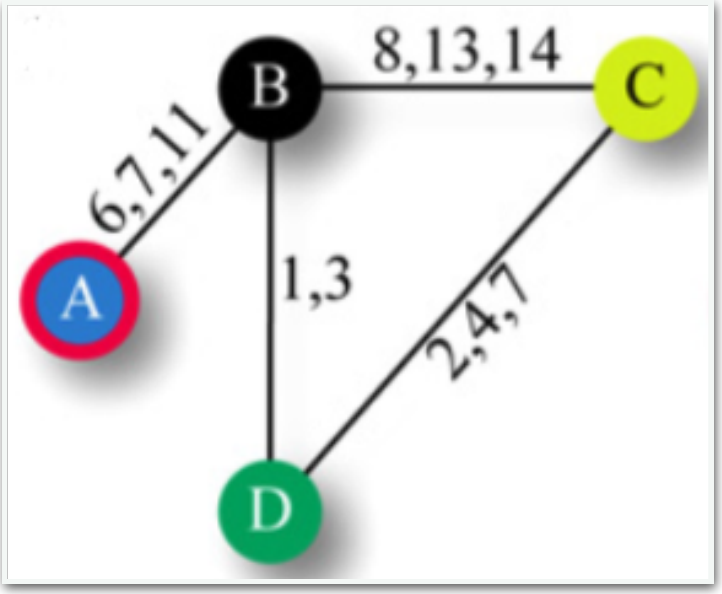
- For interval graphs, we assume that there are no empty or overlapping intervals
- Consider two intervals (t_i, t_i'), $(t_j, t_j') \in T_e$; then the following three statements need to be true
  - $t_i < t_i'$
  - $t_j < t_j'$
  - $t_i < t_j$ if and only if $t_i' < t_j$

P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Measures of temporal–topological structure

Time-respecting paths and reachability

- Paths must be constrained to sequences of link activations that follow one another in time
- In a temporal graph, paths are usually defined as sequences of contacts with non-decreasing times that connect sets of vertices
- This is called ‘’time-respecting’’ path



P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Characteristics of a time-respecting path

- Time-respecting paths define which vertices can be reached from which other vertices within some observation window $t \in [t_0 , T]$
- It might be the case that i is reachable by time-respecting paths from j, but j cannot be reached from i
- The existence of time-respecting paths from i to j and j to k does not imply that there is a path from i to k (paths are not transitive)

- **Set of influence** of i is a set of vertices that can be reached by time-respecting paths from vertex i
- **Reachability ratio** is the average fraction of vertices in the sets of influence of all vertices
- **Source set** of i is a set of vertices that can reach i through time-respecting paths within the observation window

P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Connectivity and components

- **Strongly connected components**
  - Two vertices i and j of a temporal network are strongly connected if there is a directed, time-respecting path connecting i to j and vice versa
- **Weakly connected components**
  - If there are undirected time-respecting paths from i to j and j to i, i.e. the directions of the contacts are not taken into account.
- **Transitive connectivity**
  - A subgraph is transitively connected if time respecting paths from i to j and j to k implies a time respecting path from i to k

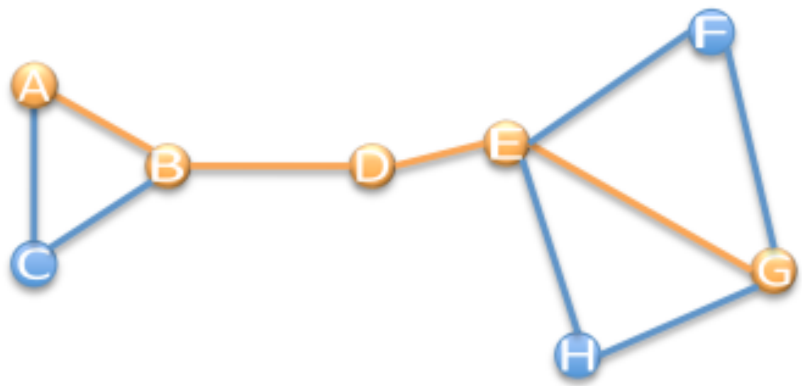
P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Distances, latencies, and fastest paths

- A time-respecting path is associated with a **duration**, measured as the time difference between the last and first contacts on the path
- The fastest time-respecting path(s) between two nodes, i.e. the shortest time within which i can reach j is called their **latency** (or **temporal distance**)
- These concepts can be used to define the overall **velocity** of the temporal network, i.e. measuring how quickly vertices can on average transmit something to each other along the contact sequences
- **Diameter** should be a number as small as possible such that, increasing it would not make you find more pairs of vertices connected by a time-respecting path that is shorter than the diameter

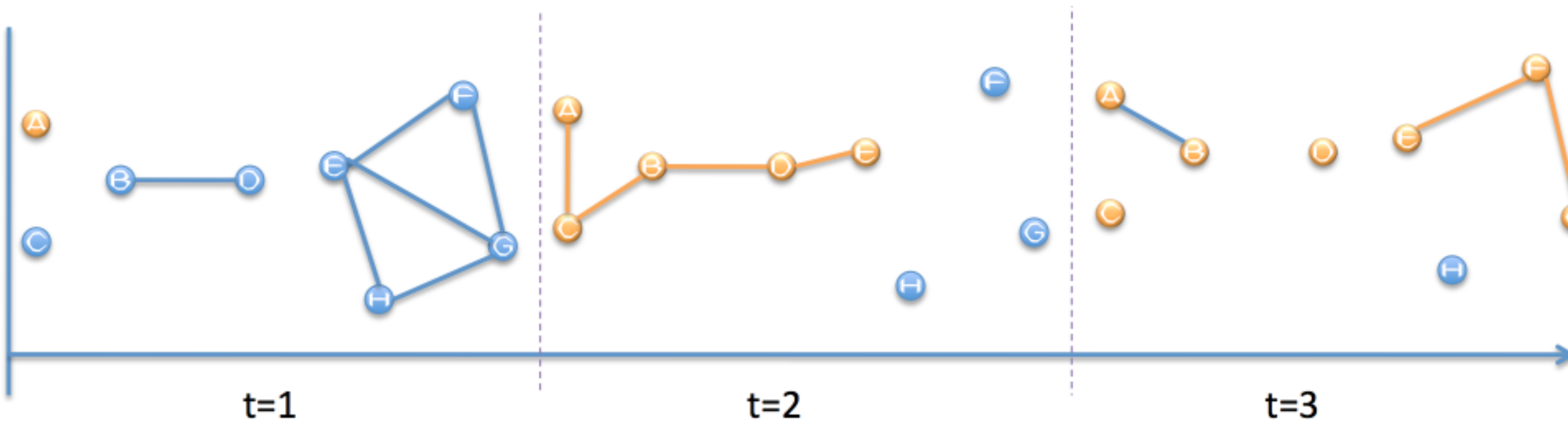
P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Example: Static network



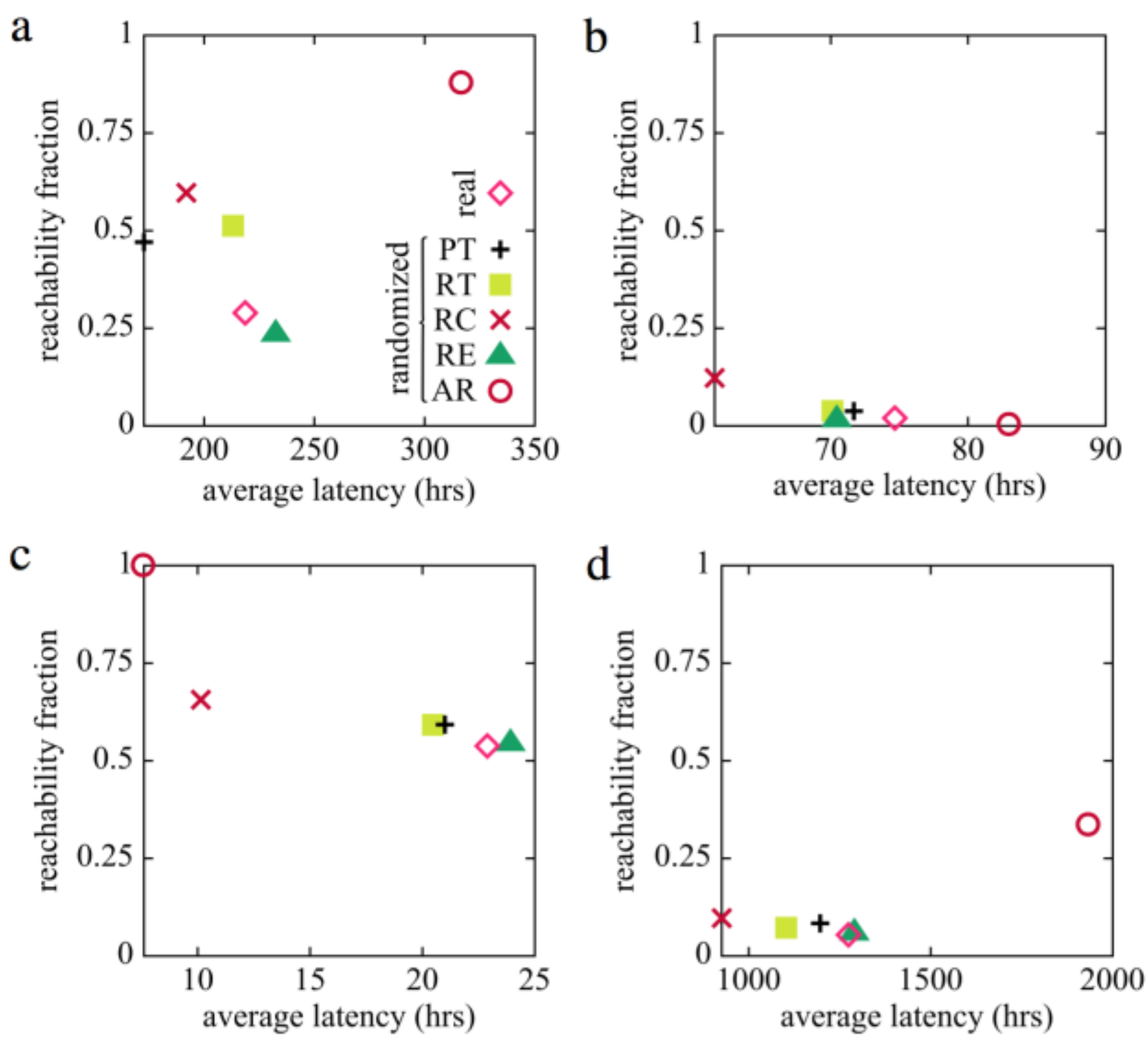
- Shortest path(A,G)=[A,B,D,E,G]
- Shortest path length(A,G)=4 hops

Example: Temporal network



- Shortest path (A,G)=[A,C,B,D,E,F,G]
- Shortest path length (A,G) = 6 hops
- Time = 3 seconds

Average latency and reachability ratio of some empirical contact sequences



P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Each panel corresponds to a dataset: (a) is from contacts of an Internet community; (b)–(d) comes from e-mail exchange

Closeness Centrality

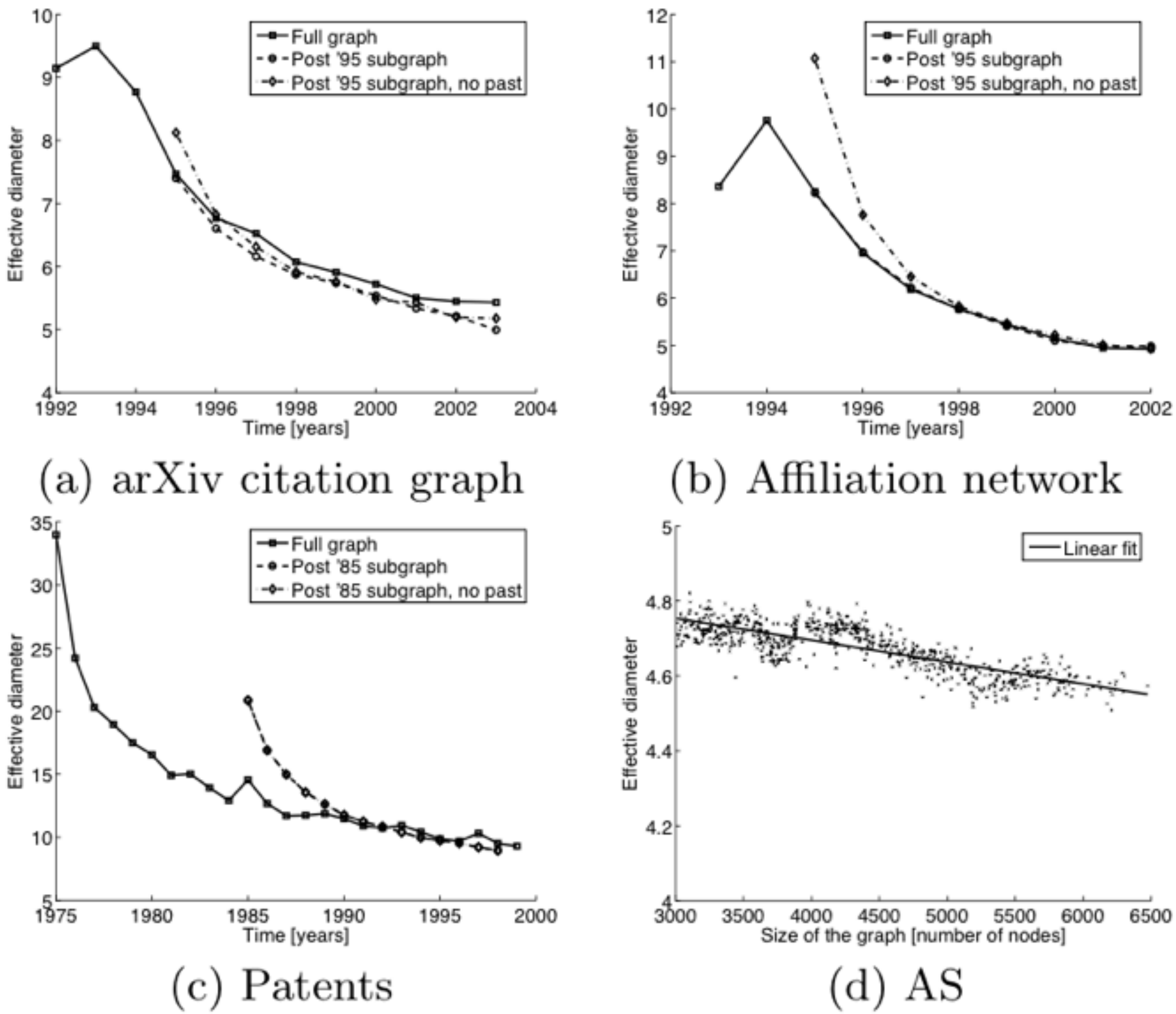
- How quickly a vertex may on average reach other vertices
- It is based on pairwise latencies averaged over the observation window
- If there are no paths at all between i and j, the average latency is infinite for this pair and the reciprocal latency is defined as zero
- Thus, the temporal closeness centrality as

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

where
$\lambda_{i,t}$ (j) is the latency between i and j and
1/(λi,t (j)) is defined as zero if there are no time-respecting paths from j to i arriving at time t or earlier

P. Holme, Network reachability of real-world contact sequences, Phys. Rev. E 71 (2005) 046119.

Shrinking diameter over time



Here the effective diameter d is used; meaning that 90 percent of the node pairs that are connected by a shortest path with a length of at most d (authors used their own function).

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2005). Graphs over time: densification laws, shrinking diameters and possible explanations (pp. 177–187). Presented at the KDD '05: Proceedings of the eleventh ACM SIGKDD

Discrete vs. continuous approach

- The two approaches should not be considered as totally different
- The separation in the lecture is more caused by terminology
- However, the origins of articles are quite different and the nomenclature is very diverse
- *Be aware!*

Questions?