Temporal (Dynamic) Networks

Lecture 20
19 November 2013
CSCI 5352, Network Analysis and Models

Prof. Aaron Clauset
University of Colorado, Boulder

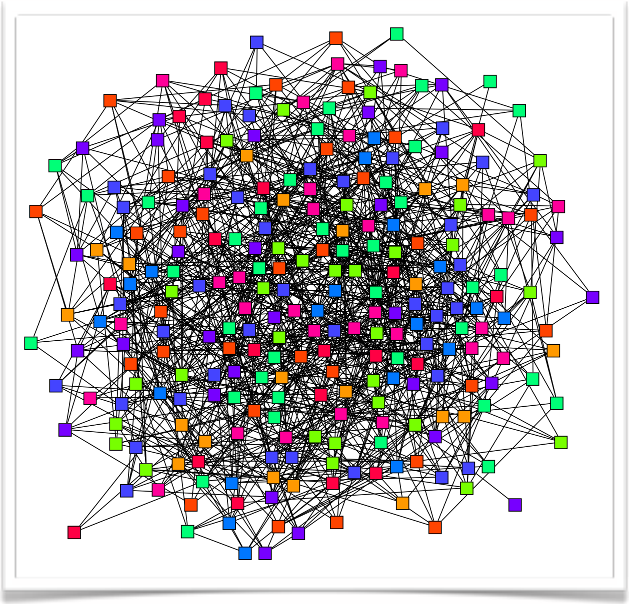
static network analysis

given network $G = (V, E)$

- centrality measures (degree-based, geometric, etc.)
- assortativity, transitivity, reciprocity
- distributions (degrees, distances, etc.)
- random walks on networks
- differences relative to configuration model
- community structure
- generative models
- etc.

static network analysis

G



empirical network

$f(G)$ →

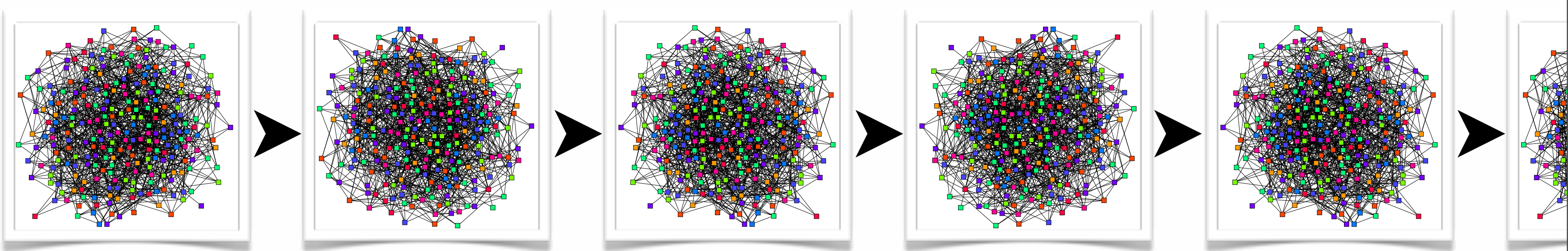
network
summary statistics

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}$$

temporal network analysis

idea 1:

empirical network sequence



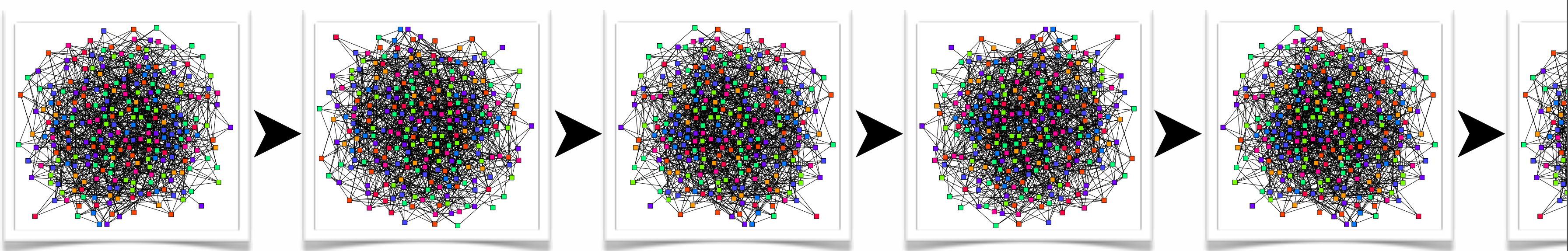
$G_{t=1}$ G_2 G_3 G_4 G_5

time-stamped interactions: $e = (i, j, t)$

temporal network analysis

idea 1:

empirical network sequence



$G_{t=1}$ G_2 G_3 G_4 G_5

time-stamped interactions: $e = (i, j, t)$

$$f(G_1) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_{t=1} \blacktriangleright \quad f(G_2) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_{2} \blacktriangleright \quad f(G_3) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_{3} \blacktriangleright \quad f(G_4) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_{4} \blacktriangleright \quad f(G_5) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_{5} \blacktriangleright$$

temporal network analysis

idea 1:

given network sequence $G_t = (V, E_t)$

- compute statistics for each "snapshot" in sequence
- makes time series of scalar or vector values

$$\vec{x} = x_1, x_2, x_3, \ldots, x_T$$

- apply standard time series analysis tools
  - autocorrelation (periodicities)
  - change-point detection, non-stationarity
  - covariance of features
  - etc.

temporal network analysis

idea 2:

edges have durations $e = (i, j, t_s, \Delta t)$

- durations of telephone calls
- time spent together
- etc.

temporal network analysis

idea 2:

edges have durations $e = (i, j, t_s, \Delta t)$

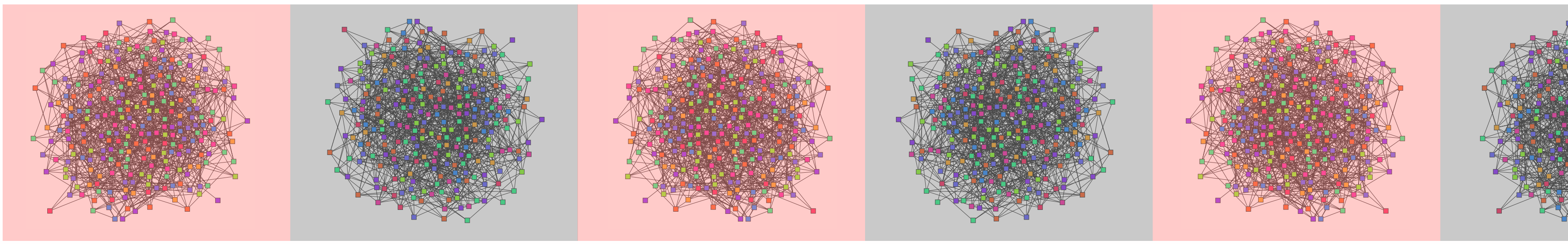
- durations of telephone calls
- time spent together
- etc.

discretize time and reduce to idea 1

temporal network analysis

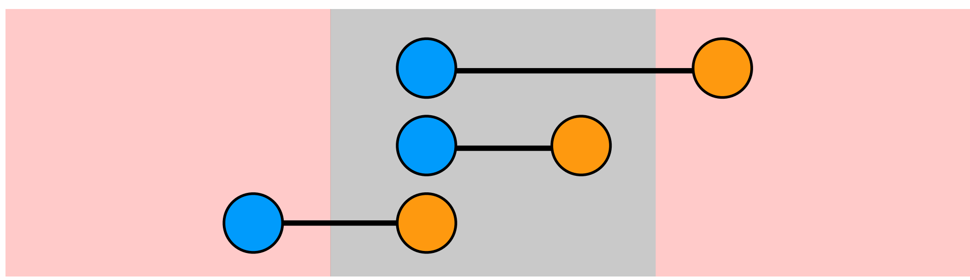
idea 2:

edges have durations $e = (i, j, t_s, \Delta t)$



$G_{t=1}$ G_2 G_3 G_4 G_5

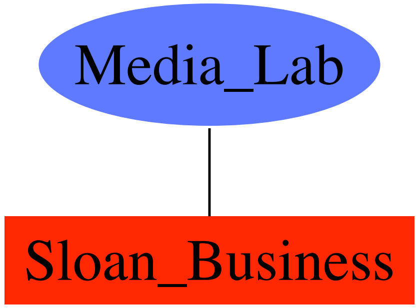
edge in G_t if



$t-1$ t $t+1$

dynamic proximity network

- MIT Reality Mining Project
- 100 mobile phones, 2 groups
- scan area with bluetooth
- every 5 minutes
  for 12 months
  (~100,000 minutes of data)
- record proximate devices (range: 5m)
- convert to dynamic proximity network
  (assume phone = person

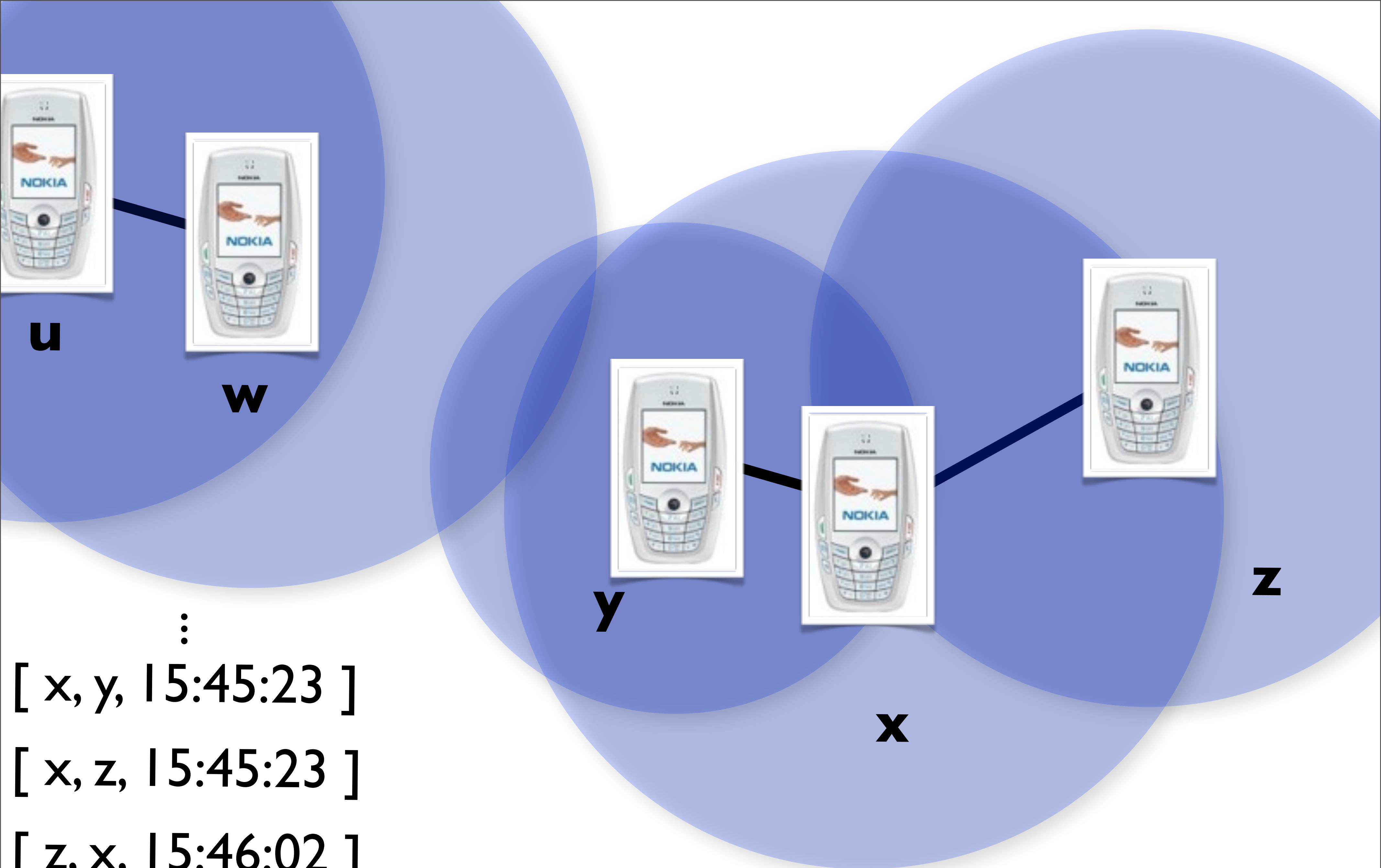


paper reference:

Persistence and periodicity in a dynamic proximity network*

Aaron Clauset[‡,*] and Nathan Eagle[†]

http://realitycommons.media.mit.edu



⋮

[x, y, 15:45:23]

[x, z, 15:45:23]

[z, x, 15:46:02]

[u, w, 15:46:12]

⋮

proximity inference rule

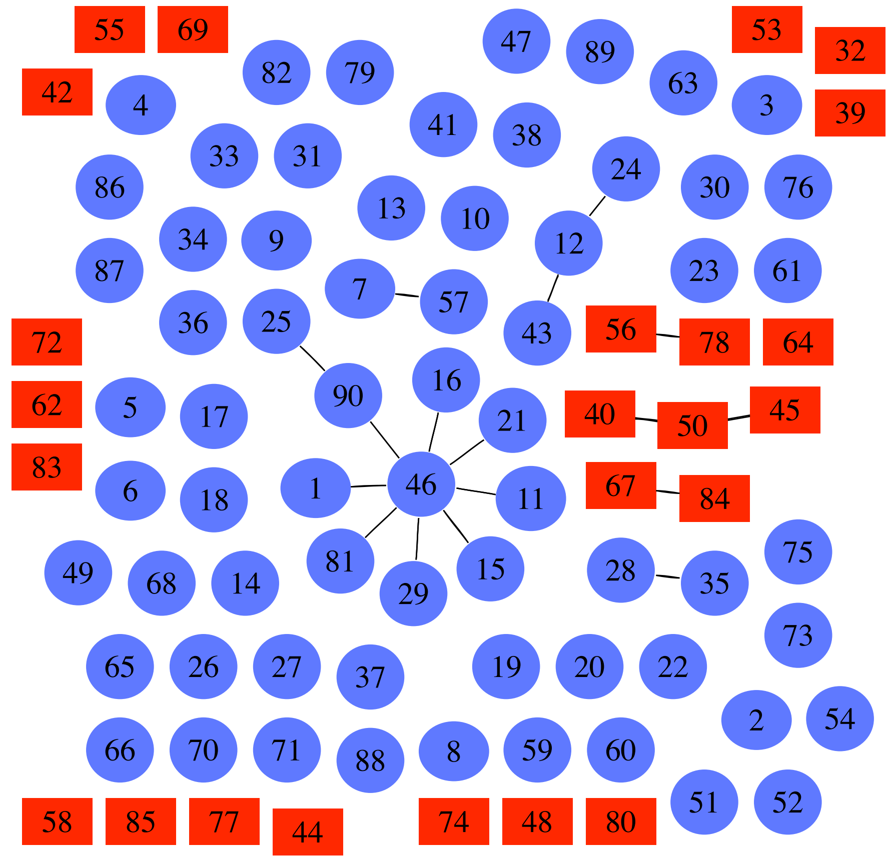
- proximities are time-stamped (i, j, t)
- we want to infer durations $(i, j, t_s, \Delta t)$
- proximities are noisy [some edges unobserved]
- high-resolution temporal sampling [every 5 mins]

- rule:
    - define tolerance τ; if gap less than τ, assume continuous proximity

single day of proximities

single day of proximities

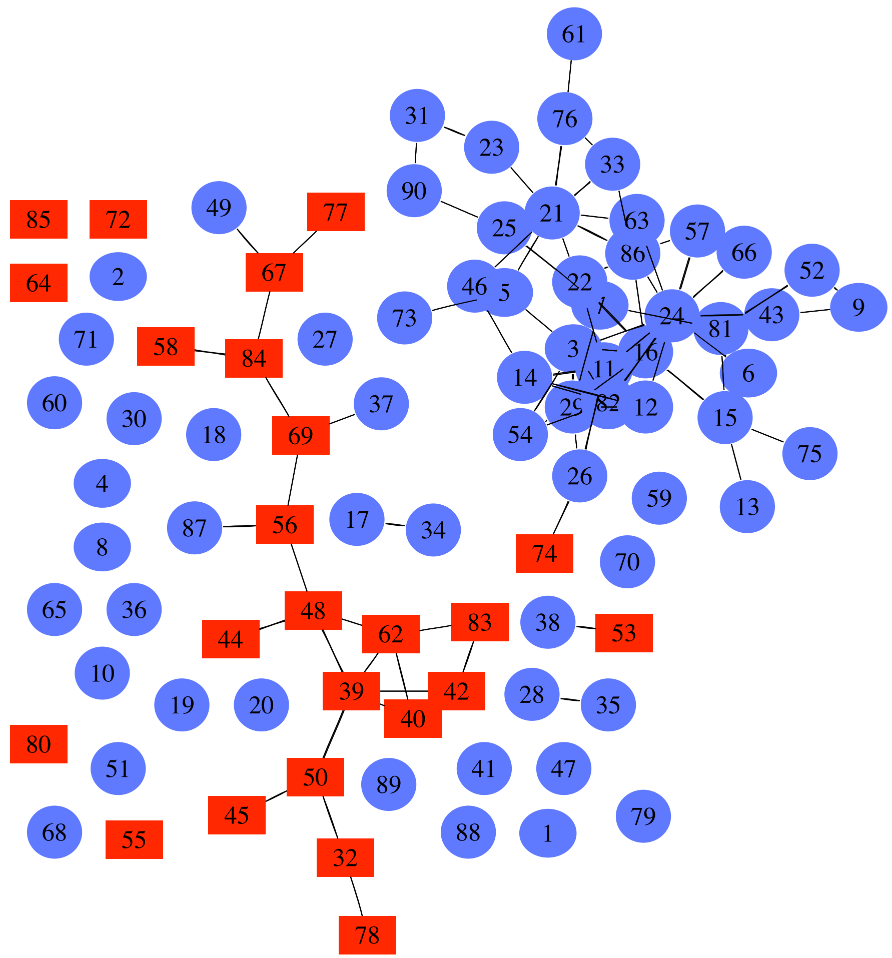
Tuesday, 19 Oct 2004



very few connections

single day of proximities

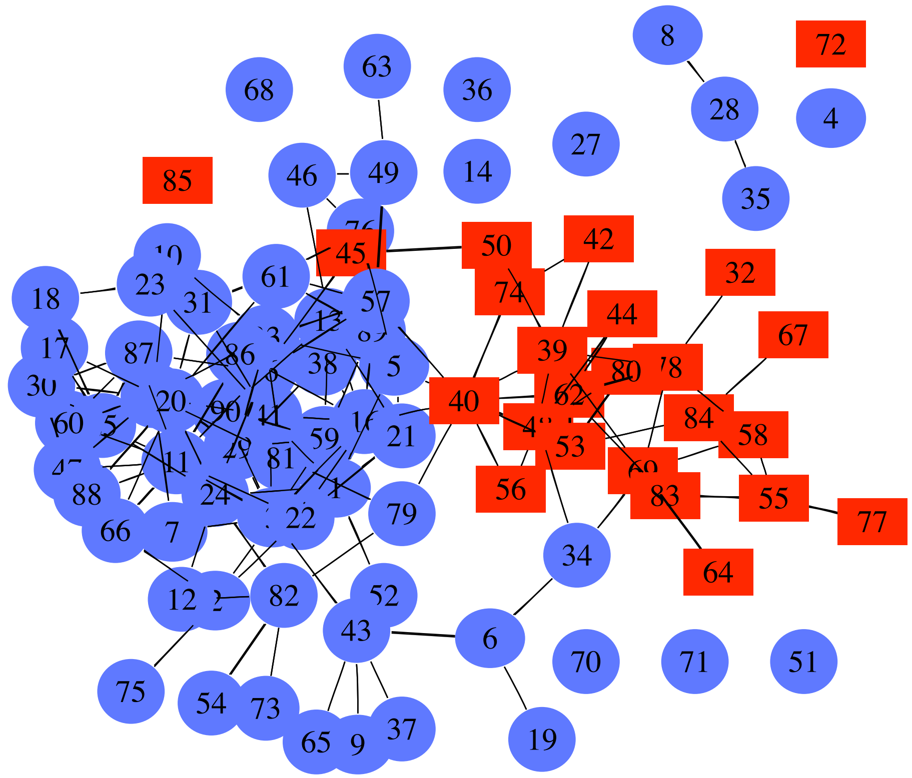
Tuesday, 19 Oct 2004



more connections

single day of proximities

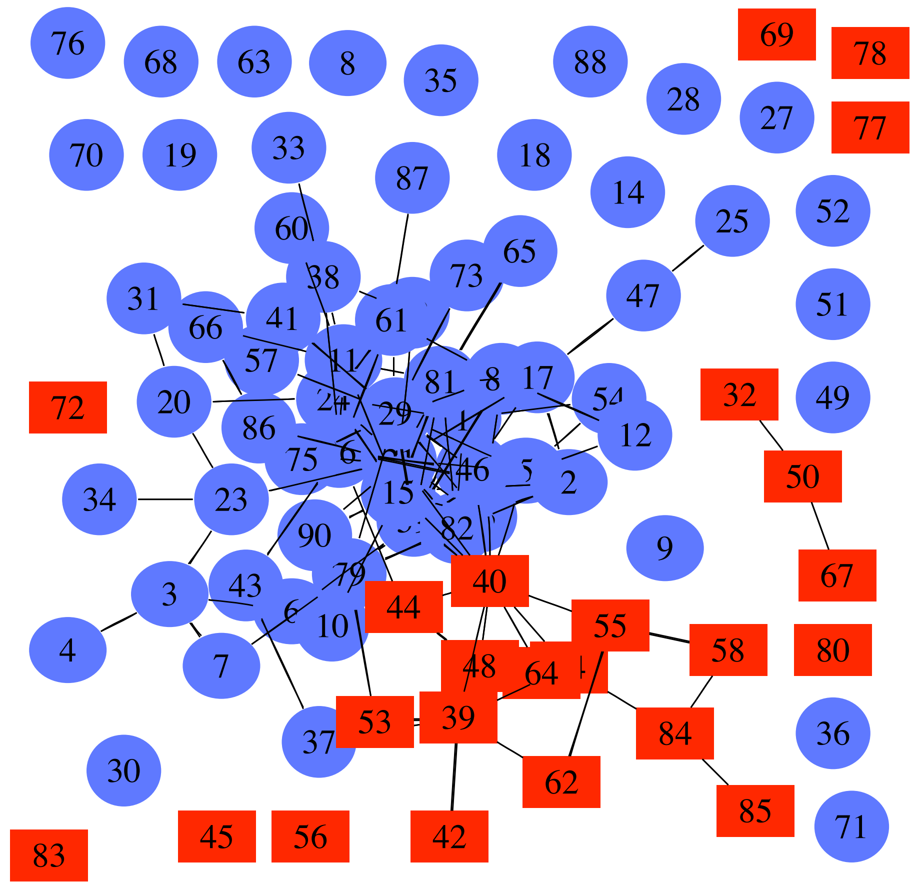
Tuesday, 19 Oct 2004



peak connections,
two communities

single day of proximities

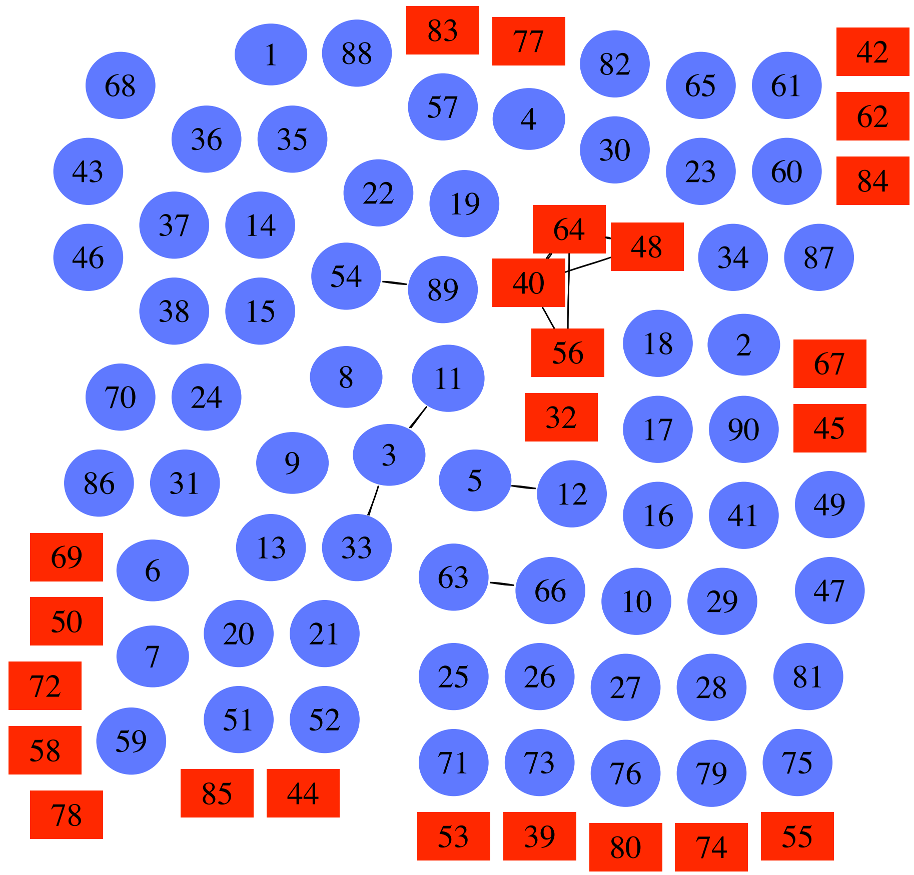
Tuesday, 19 Oct 2004



fewer connections

single day of proximities

Tuesday, 19 Oct 2004



very few connections

timing is everything?

- how long do edges last?
- how does structure vary over time?
- how stable is a local neighborhood?
- how does discrete time impact measures?

edge persistence

how long do edges last?

measure durations $\Pr(\Delta t)$

edge persistence

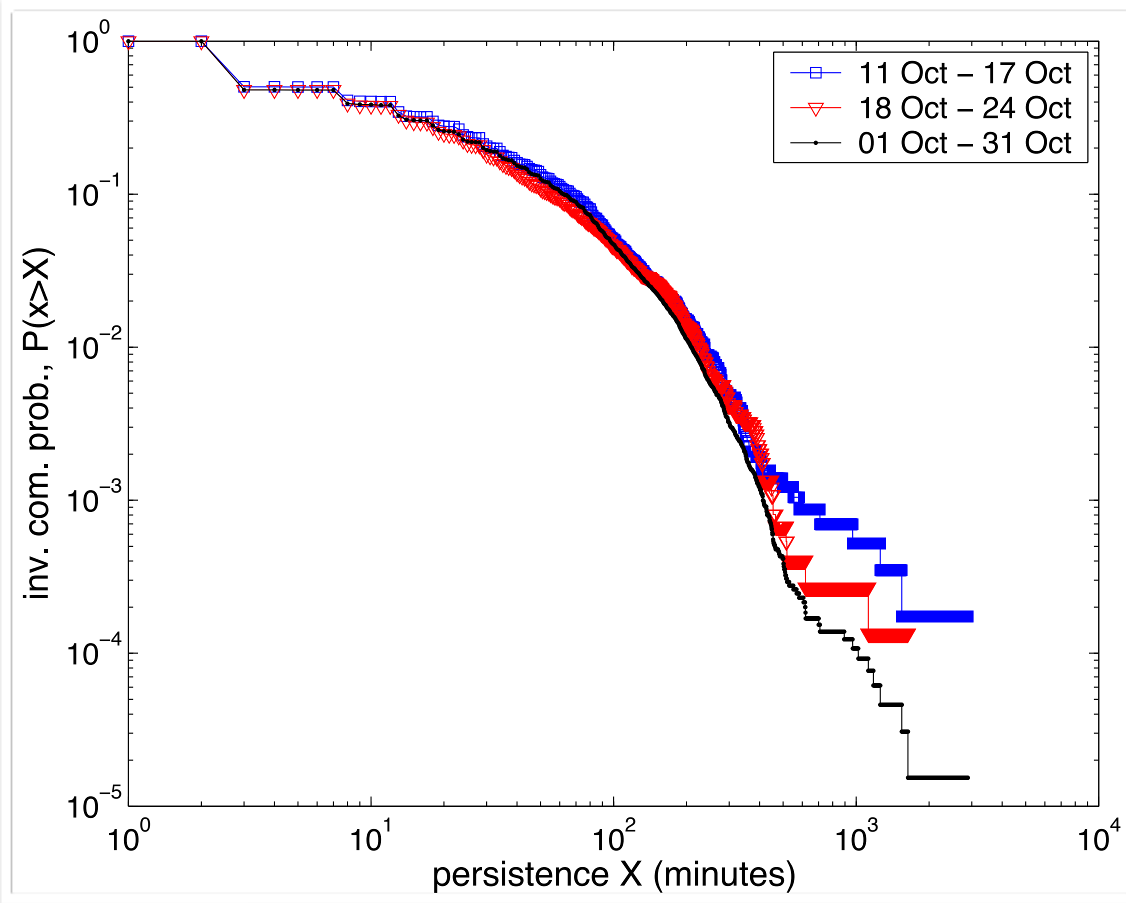
how long do edges last?

measure durations $\Pr(\Delta t)$

- month of October
- broad distribution

$\langle \Delta t \rangle = 22.8$ minutes

- changes at many time scales
- consistent up to $\Delta t < 400$ minutes



network dynamics

how does structure vary over time?

vary aggregation window for snapshots

compute **mean degree** over time

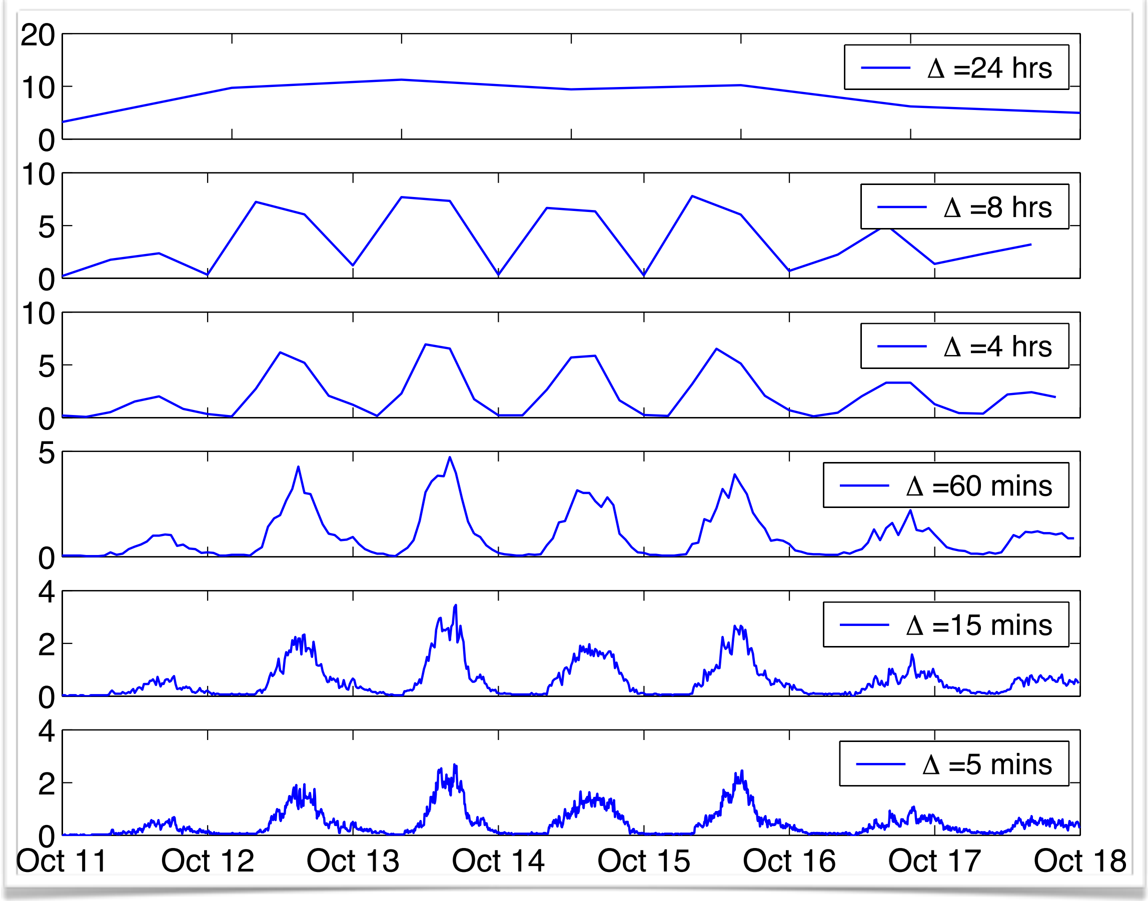
network dynamics

how does structure vary over time?

vary aggregation window for snapshots

compute **mean degree** over time

- one week of October
- highly periodic
- aggregation time matters



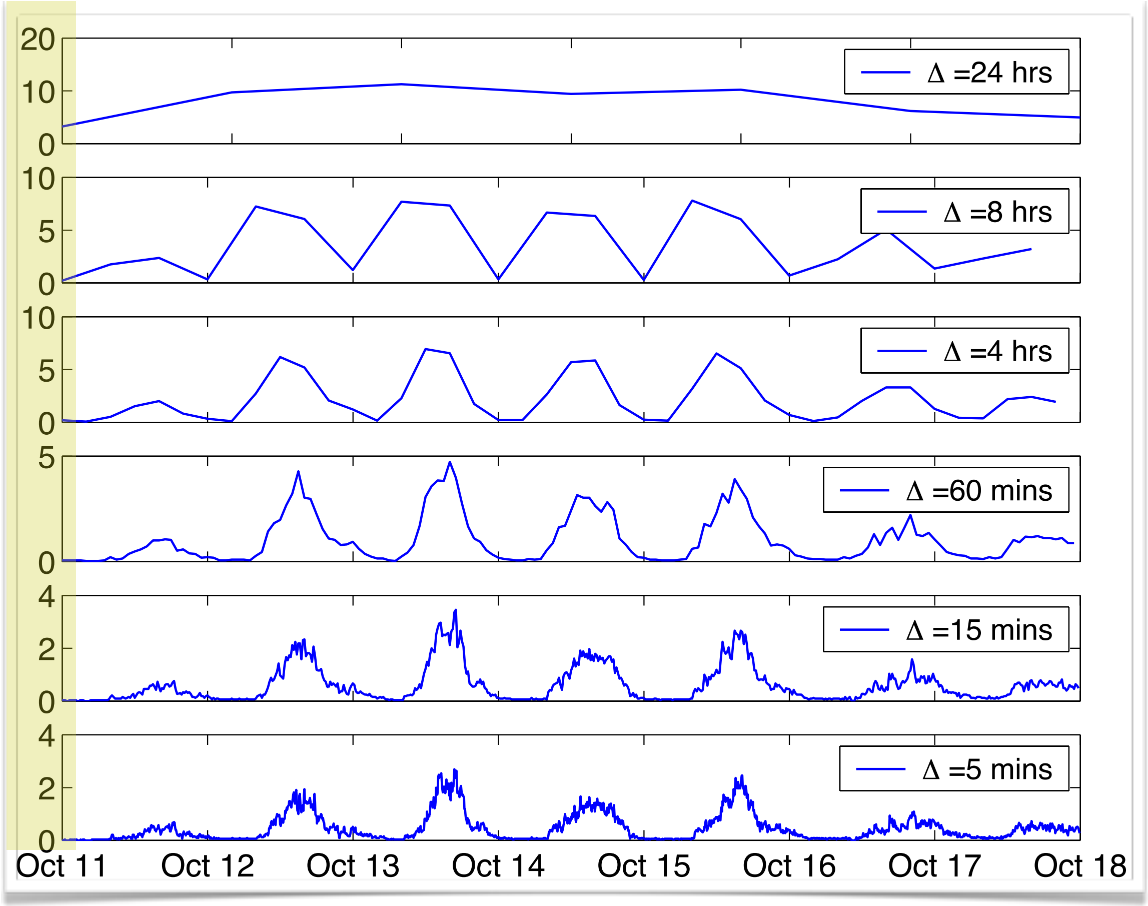
network dynamics

how does structure vary over time?

vary aggregation window for snapshots

compute **mean degree** over time

- one week of October
- highly periodic
- aggregation time matters



network dynamics

how stable are local neighborhoods?

vary aggregation window for snapshots

compute **adjacency correlation** over time

$$\gamma_j = \frac{\sum_{i \in N(j)} A_{ij}^{(x)} A_{ij}^{(y)}}{\sqrt{\sum_{i \in N(j)} A_{ij}^{(x)} \sum_{i \in N(j)} A_{ij}^{(y)}}}$$

for two adjacency matrices $A^{(x)}$, $A^{(y)}$

measures similarity among neighbors observed in either network

average overlap = mean value $\langle \gamma \rangle$

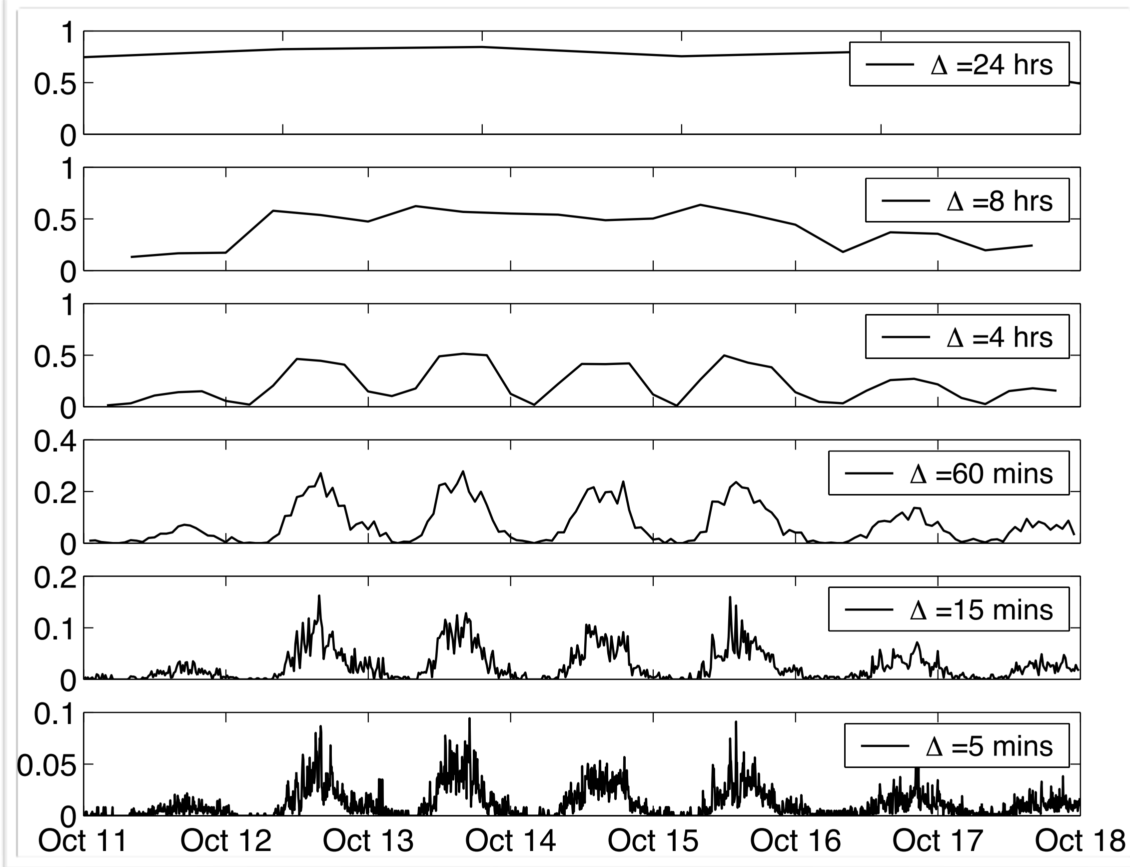
network dynamics

how stable are local neighborhoods?

vary aggregation window for snapshots

compute **adjacency correlation** over time

- one week of October
- highly consistent neighborhoods
- daily / weekly periodicity
- aggregation time matters



how does discrete time impact measures?

vary aggregation window for snapshots

compute **summary statistics**

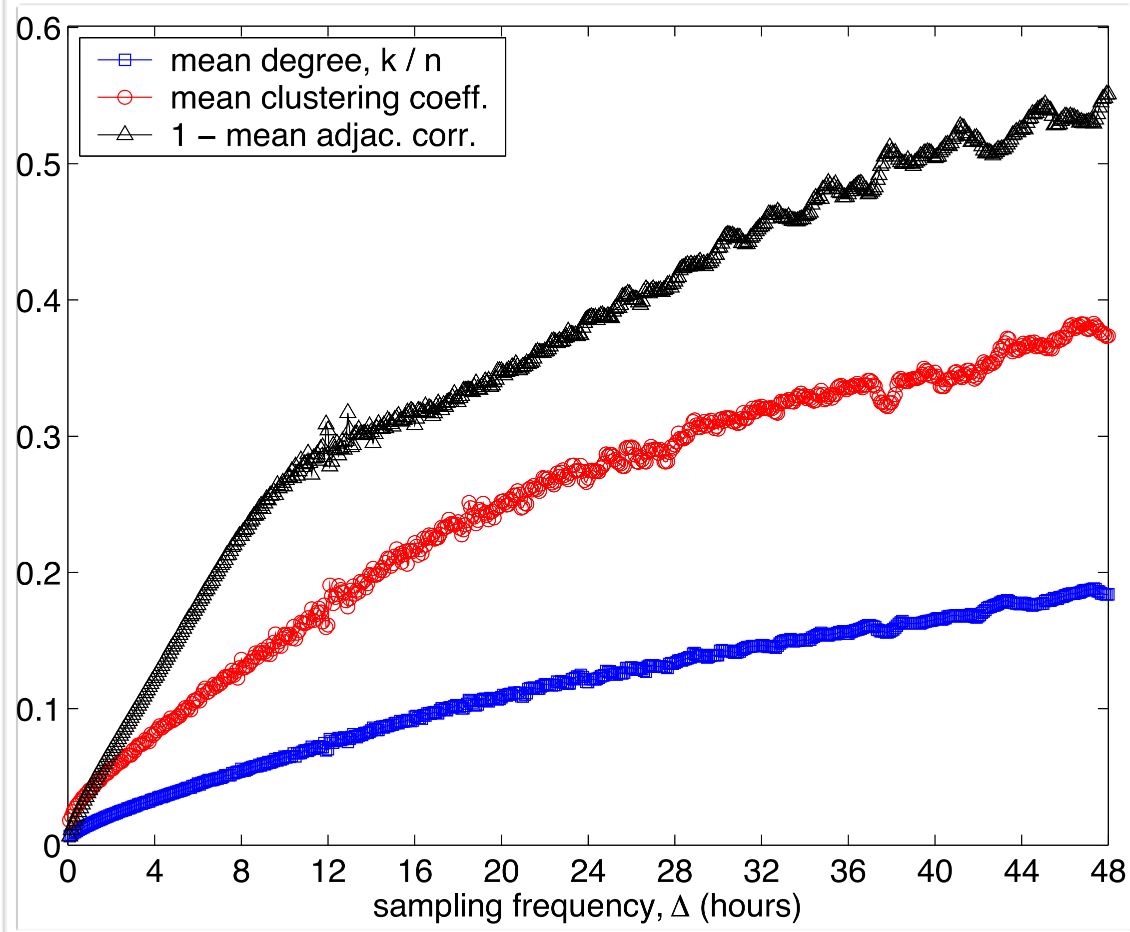
network dynamics

how does discrete time impact measures?

vary aggregation window for snapshots

compute **summary statistics**

- all statistics depend on aggregation duration
- choose a time scale = choose a statistical value



network dynamics

how to choose aggregation time?

recall highly periodic dynamics

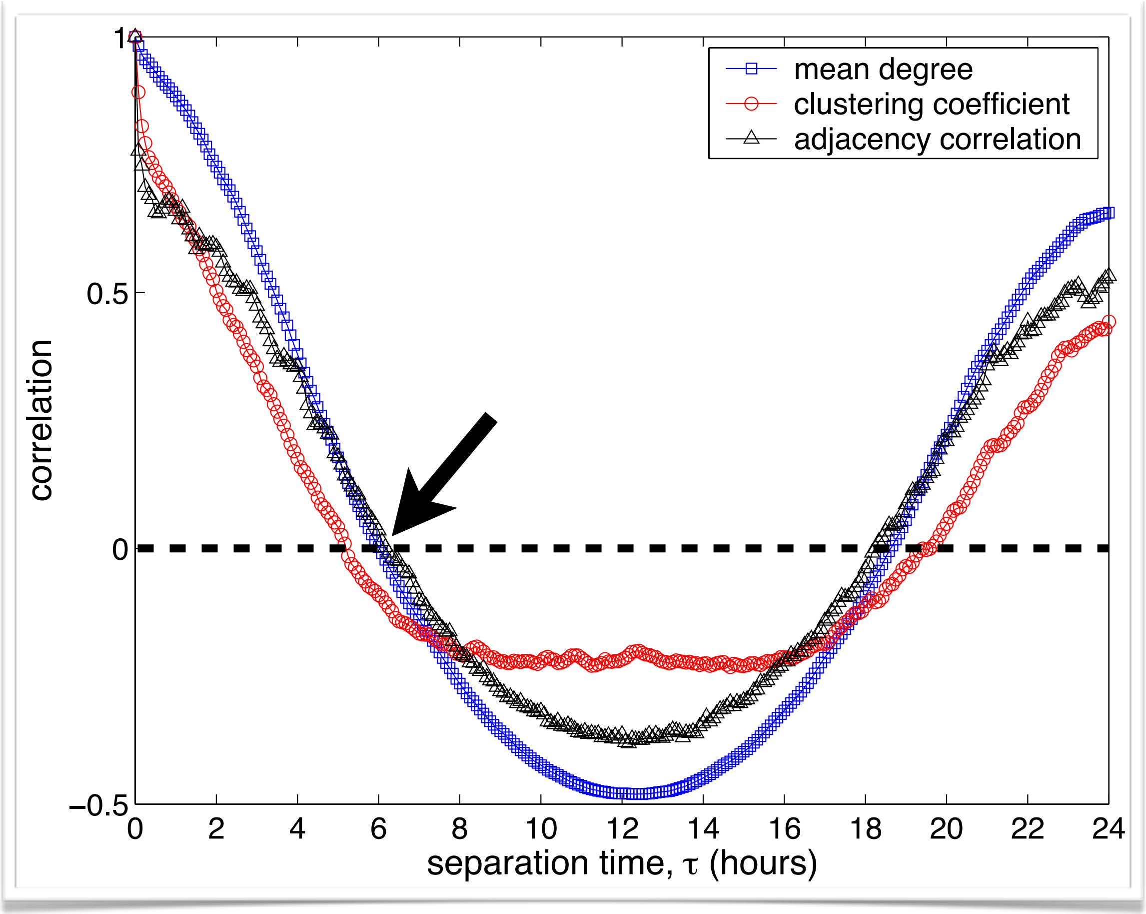
compute **autocorrelation function on network measures**

network dynamics

how to choose aggregation time?

recall highly periodic dynamics

compute **autocorrelation function on network measures**



network dynamics

how to choose aggregation time?

recall highly periodic dynamics

use frequency spectrum to choose sampling rate

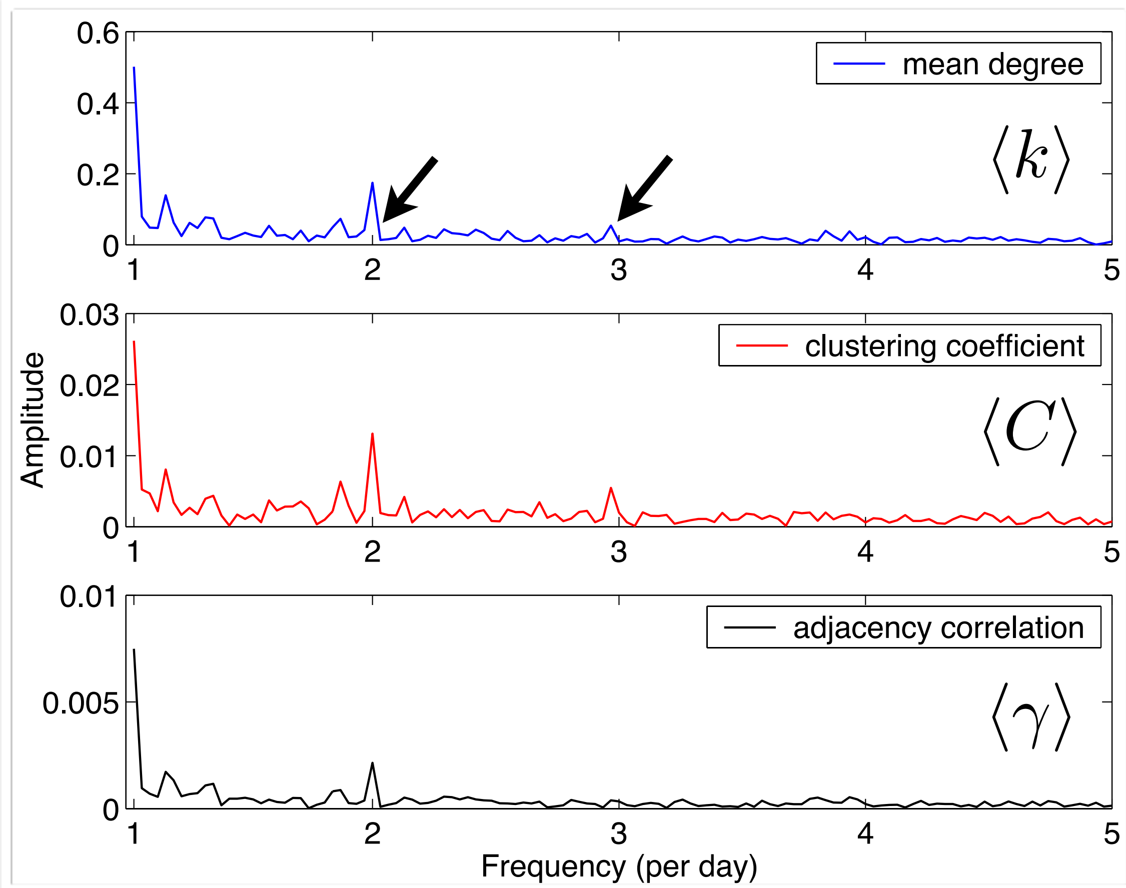
- periodicity at 1,2,3 samples per day
- Nyquist rate

$$\Delta_{\rm nat} \simeq 4 \text{ hours}$$

degree $\langle k\rangle_{\rm nat} = 2.24$

triangles $\langle C\rangle_{\rm nat} = 0.084$

adj. corr. $\langle \gamma\rangle_{\rm nat} = 0.88$



other ideas

- temporal "reachability" and continuous-time methods
- different parts evolving at different rates
- generative models?
- densification dynamics?
- temporal anomalies
- etc.