Dynamic networks

(aka temporal or evolving networks)

Network Analysis and Modeling, CSCI 5352
Prof. Aaron Clauset

Lecture 13
2017

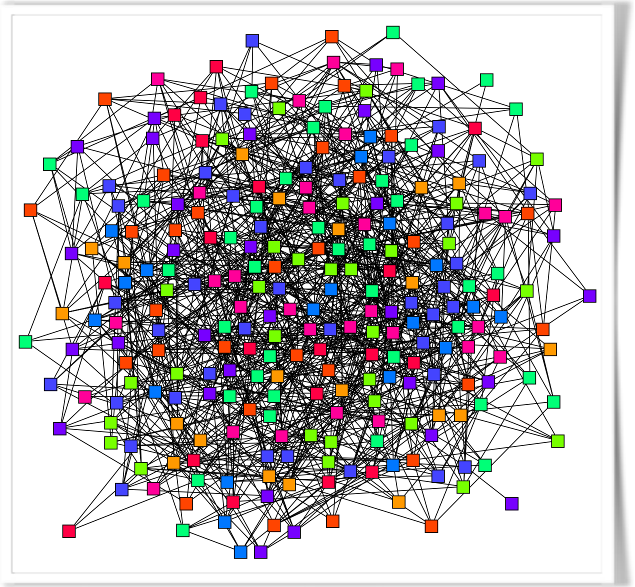
static network analysis

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

$f(G)$ →

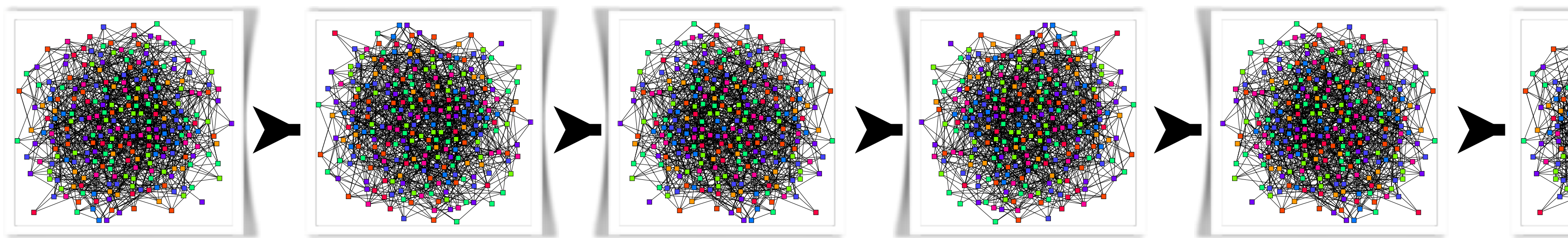
network
summary statistics

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

temporal network analysis

idea 1:

empirical network sequence



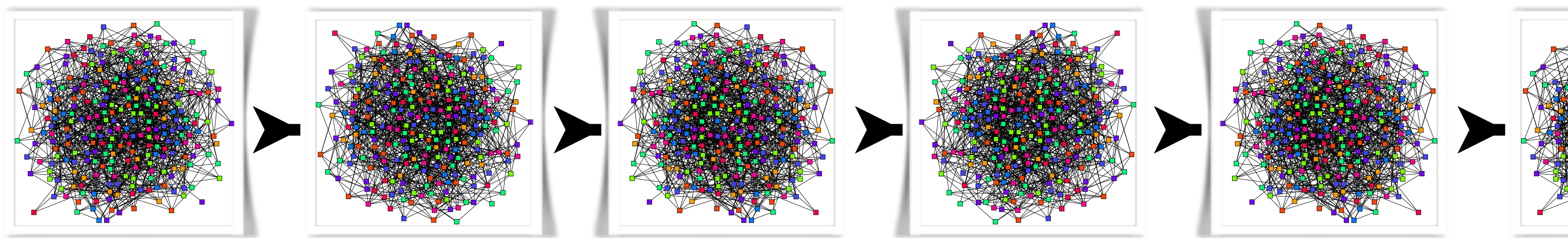
$G_{t=1}$ G_2 G_3 G_4 G_5

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

temporal network analysis

idea 1:

empirical network sequence



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

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

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

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

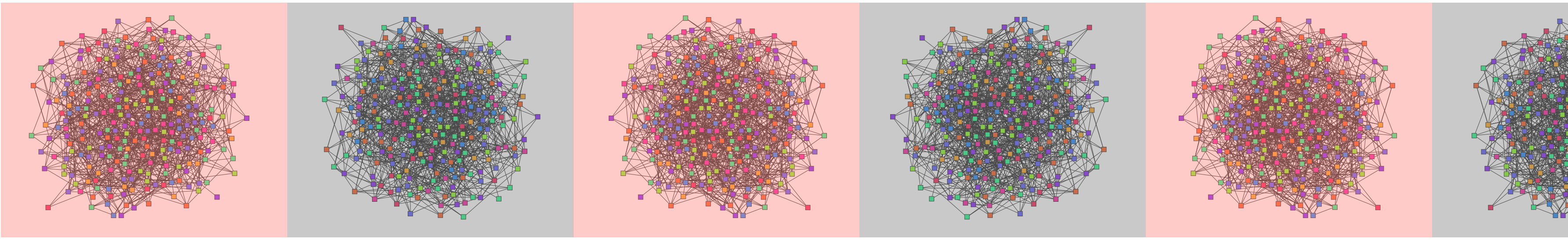
- durations of telephone calls
- time spent together
- etc.

discretize time and reduce to idea 1

temporal network analysis

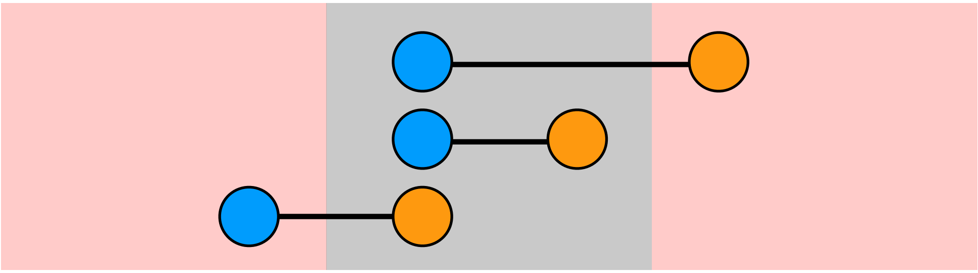
idea 2:

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



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

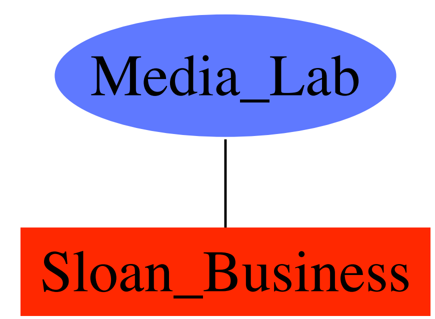
edge in G_t if



$t-1$ t $t+1$

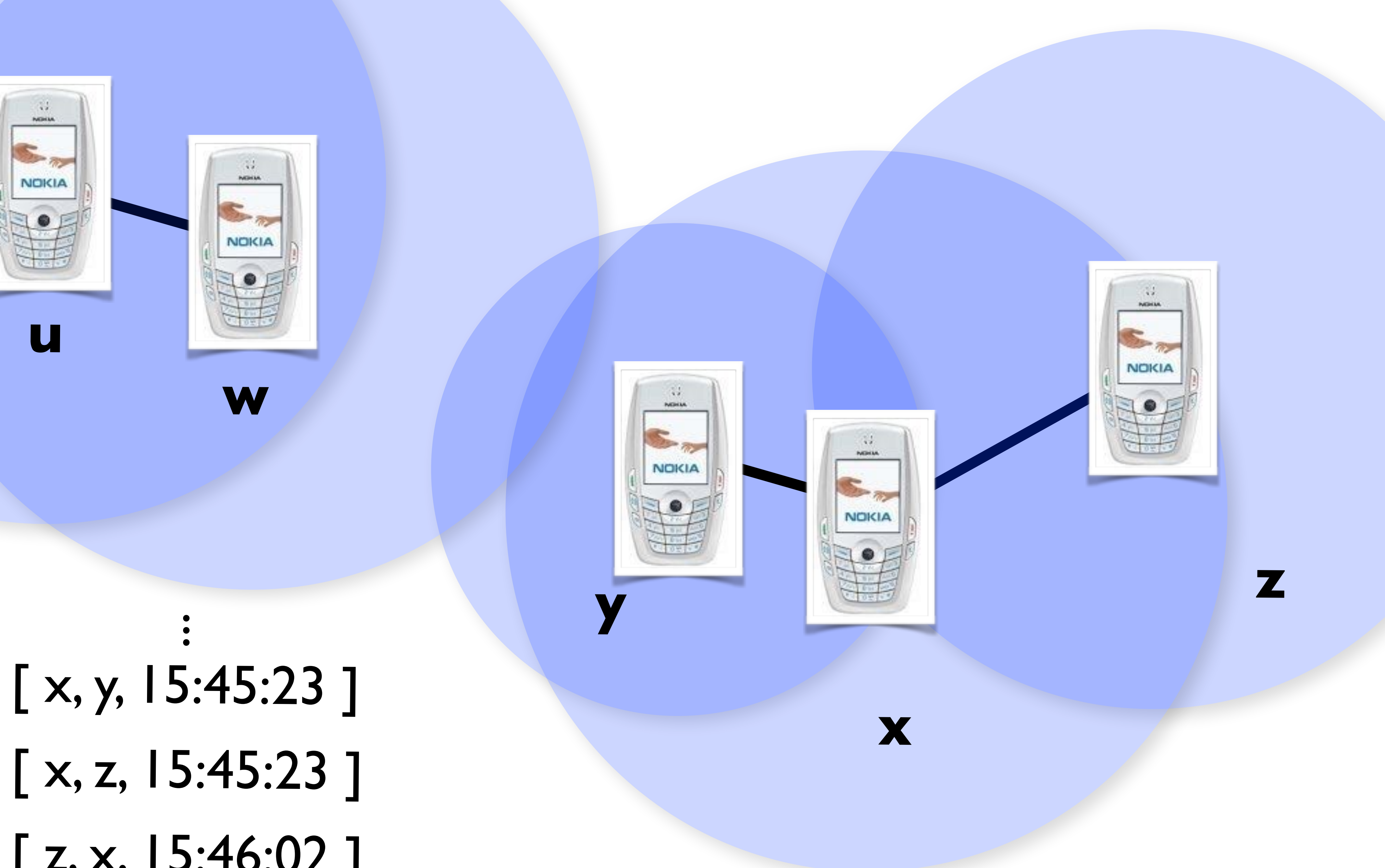
dynamic proximity network

- 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



Clauset and Eagle, "Persistence and periodicity in a dynamic proximity network."
DIMACS Workshop on Computational Methods for Dynamic Interaction Networks (2007)
arxiv:1211.7343

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

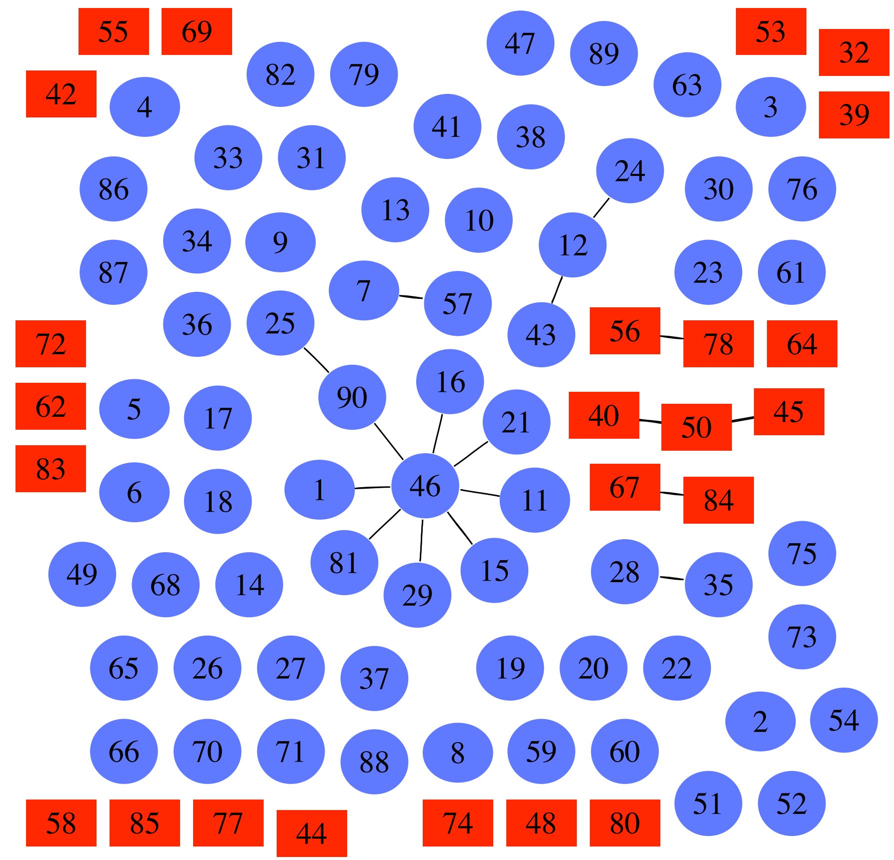
- proximities are time-stamped (i, j, t)
- we want to infer durations $(i, j, t_s, \Delta t)$
- proximities are noisy [some edges unobserved]
- high-resolution temporal sampling [every 5 mins]

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

single day of proximities

single day of proximities

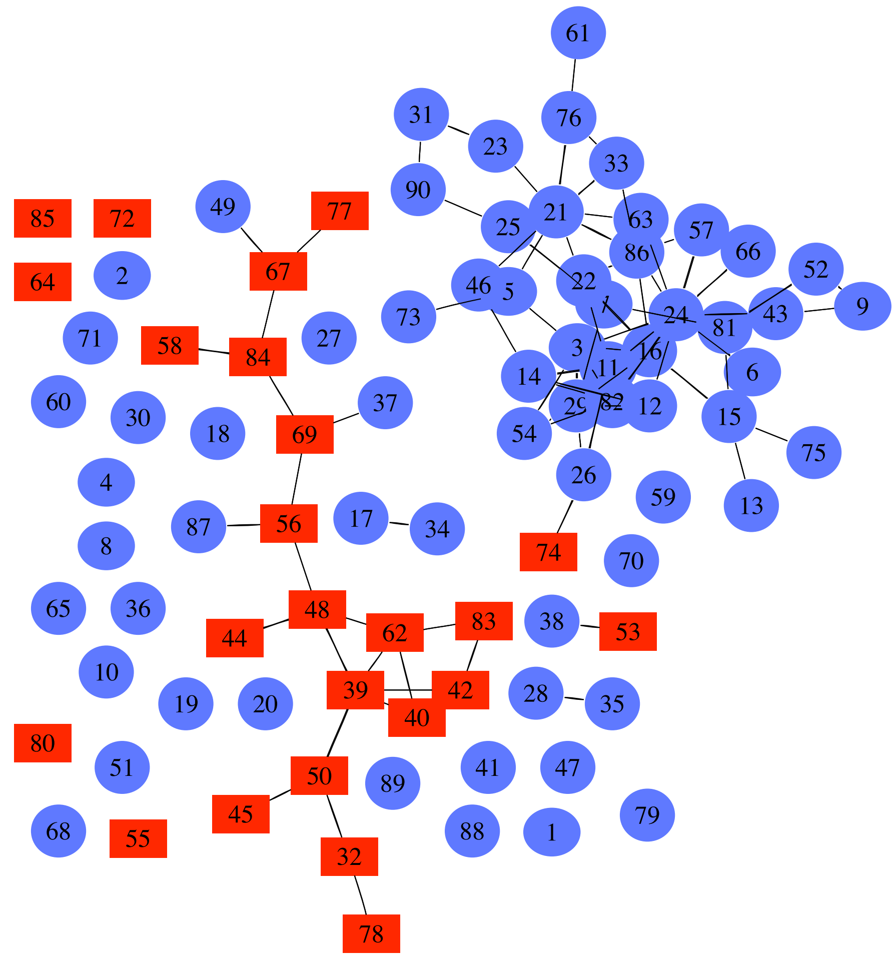
Tuesday, 19 Oct 2004



very few connections

single day of proximities

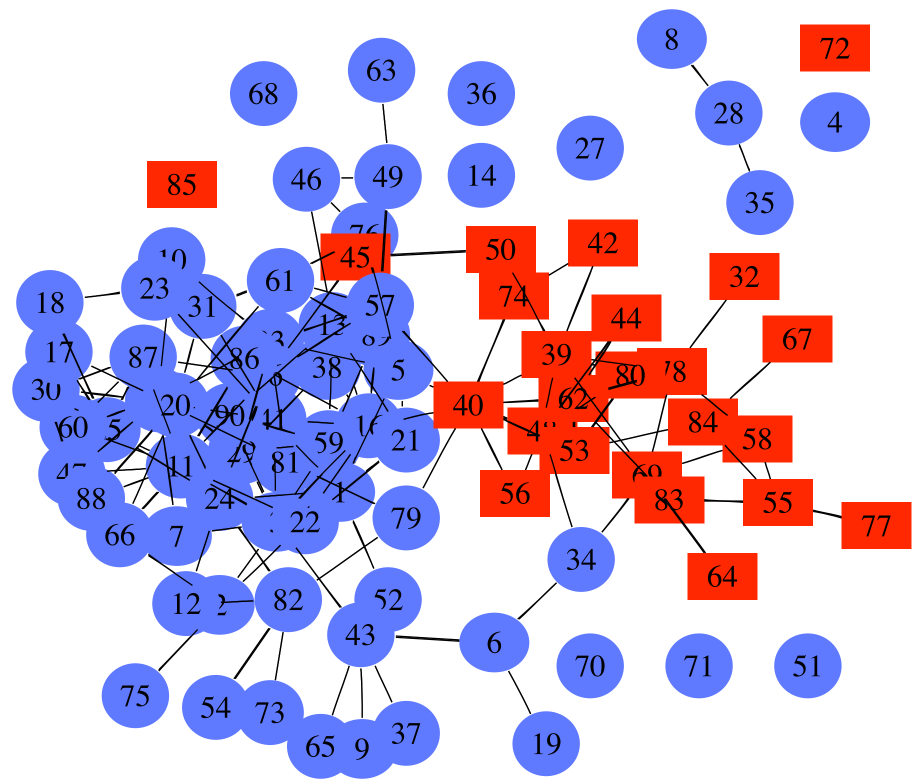
Tuesday, 19 Oct 2004



more connections

single day of proximities

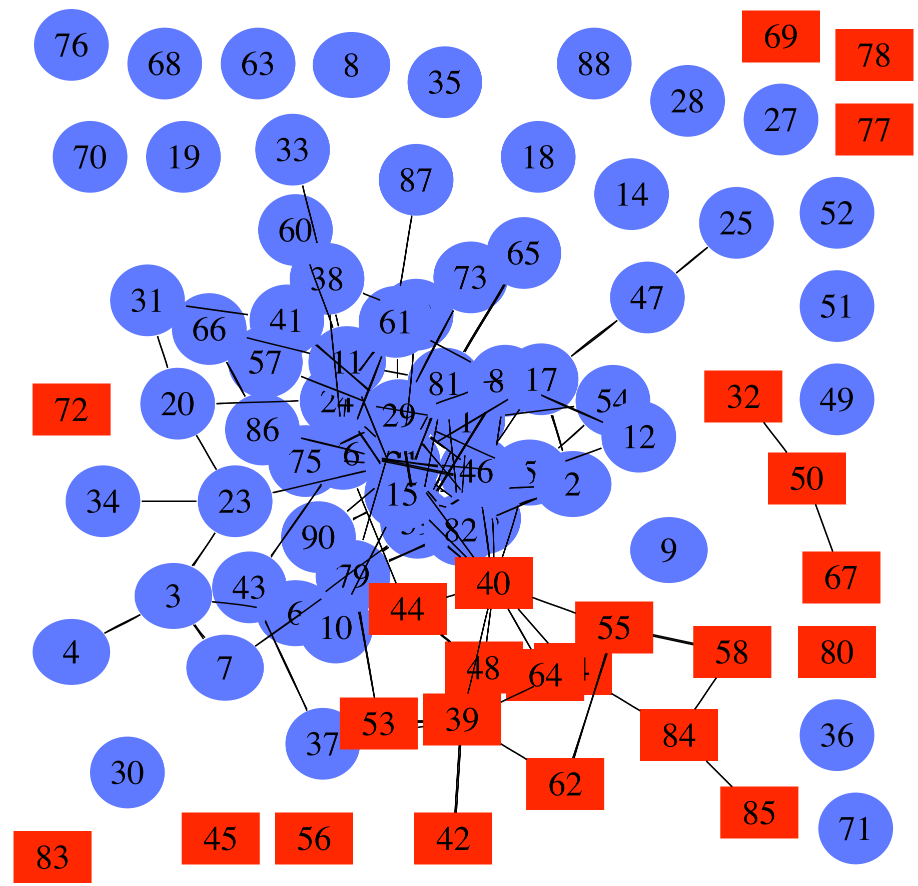
Tuesday, 19 Oct 2004



peak connections,
two communities

single day of proximities

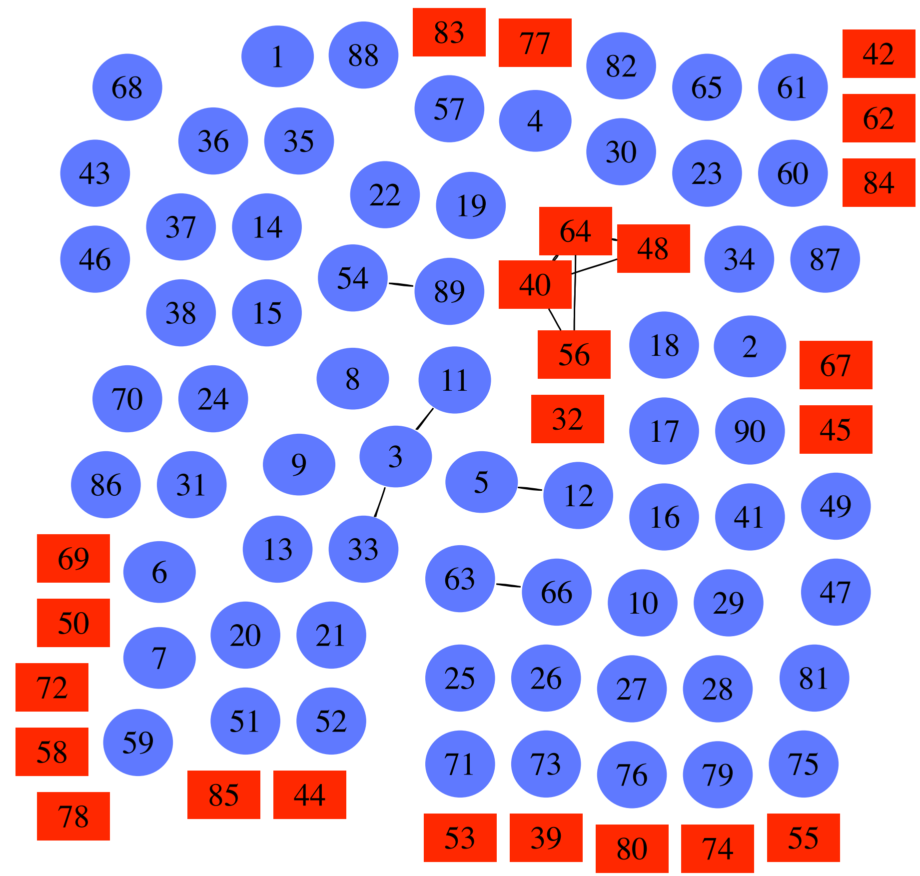
Tuesday, 19 Oct 2004



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

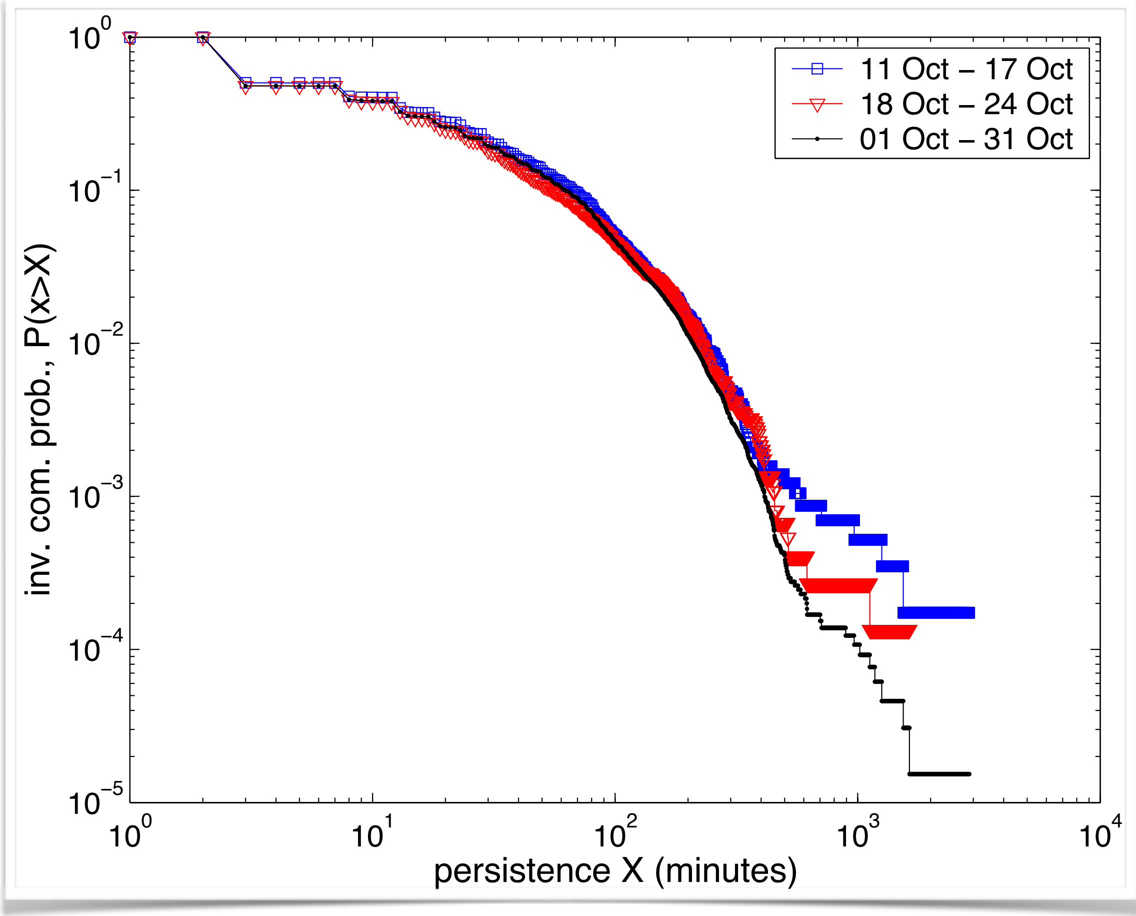
how long do edges last?

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

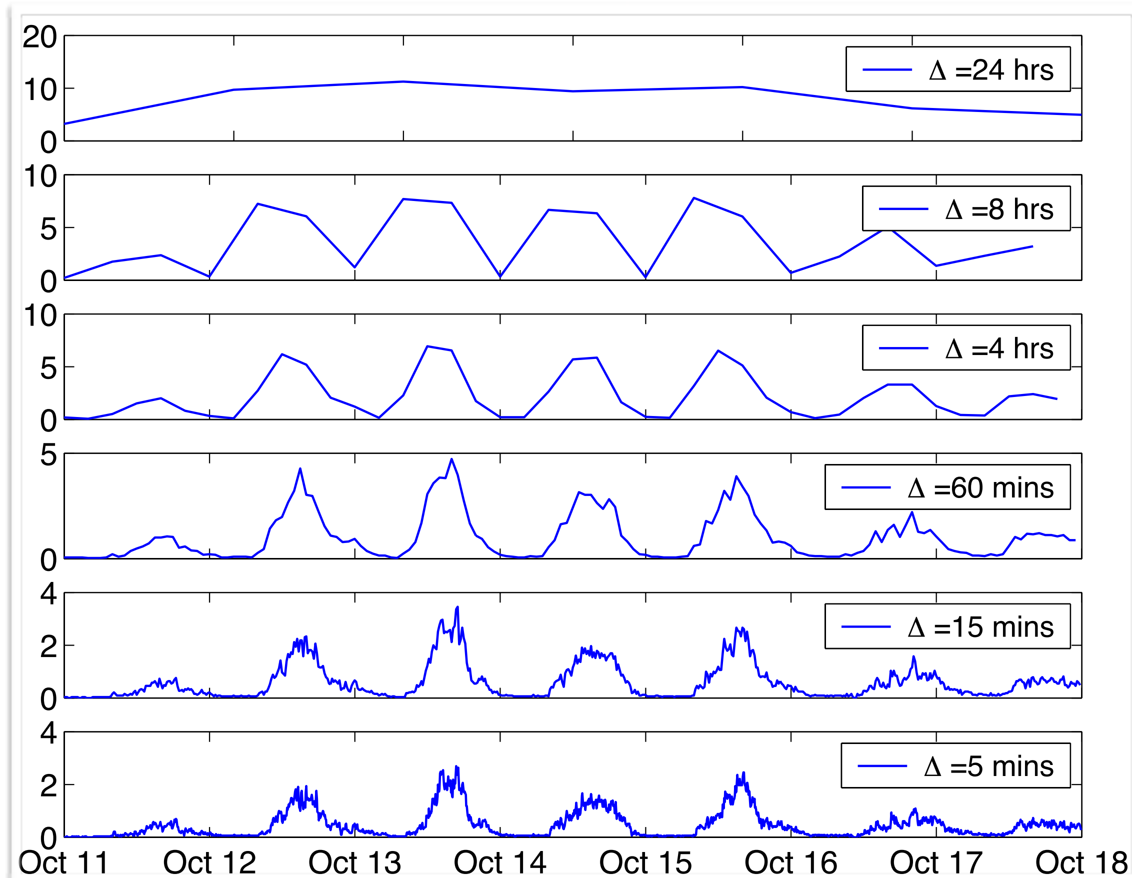
network dynamics

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



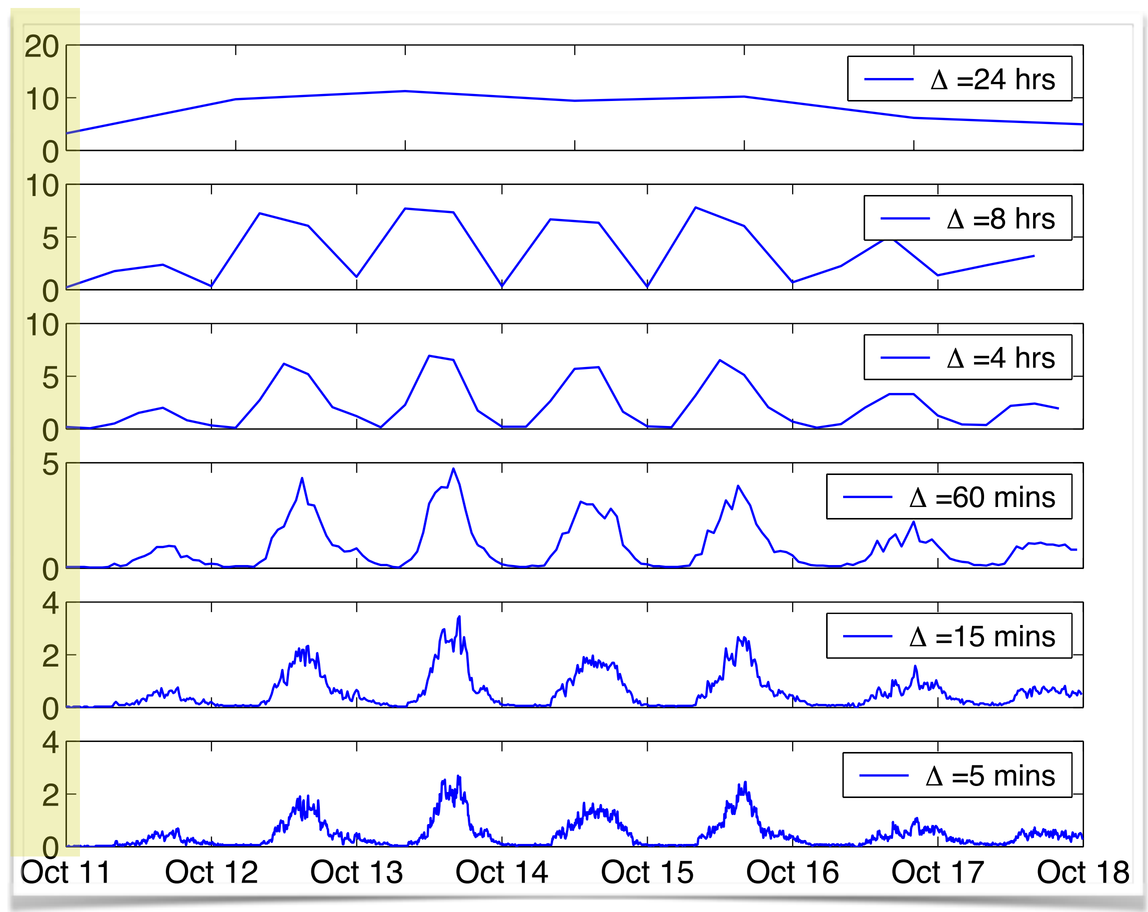
network dynamics

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 $\langle \gamma \rangle$

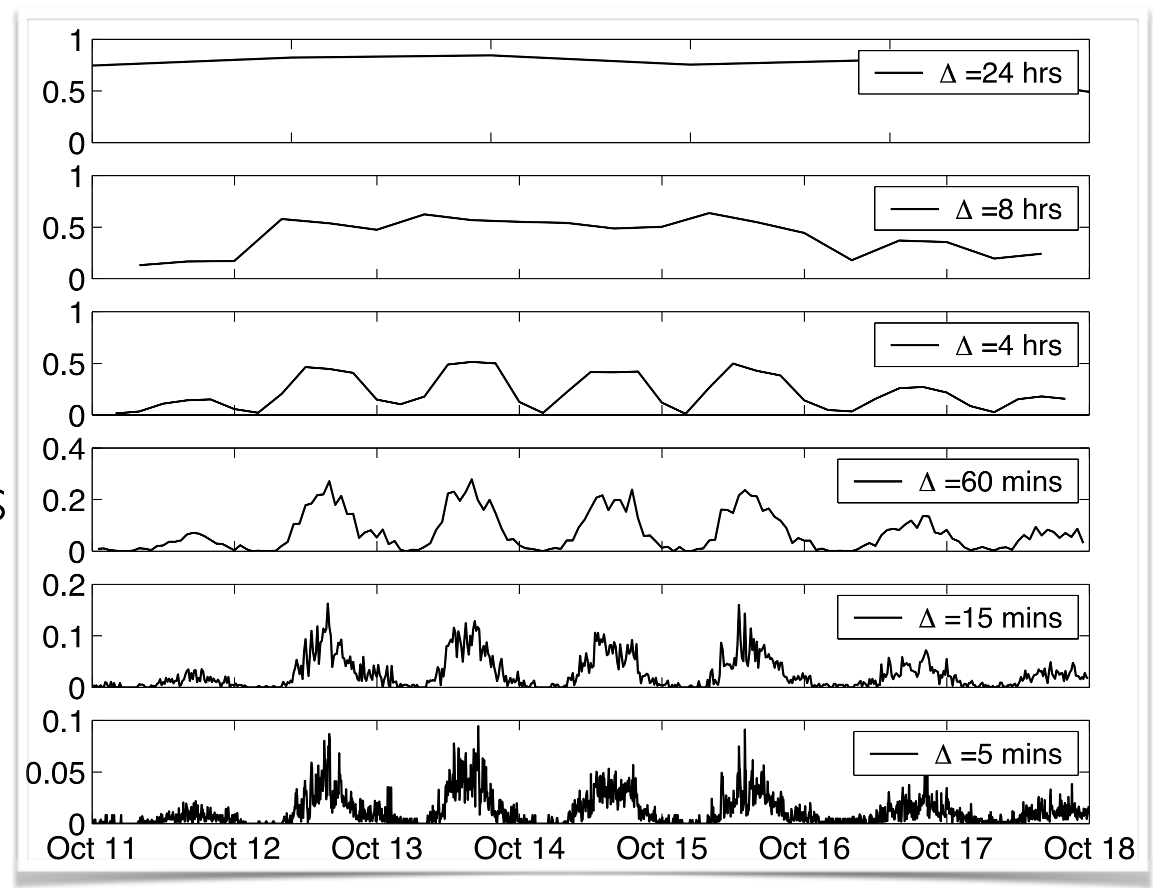
network dynamics

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**

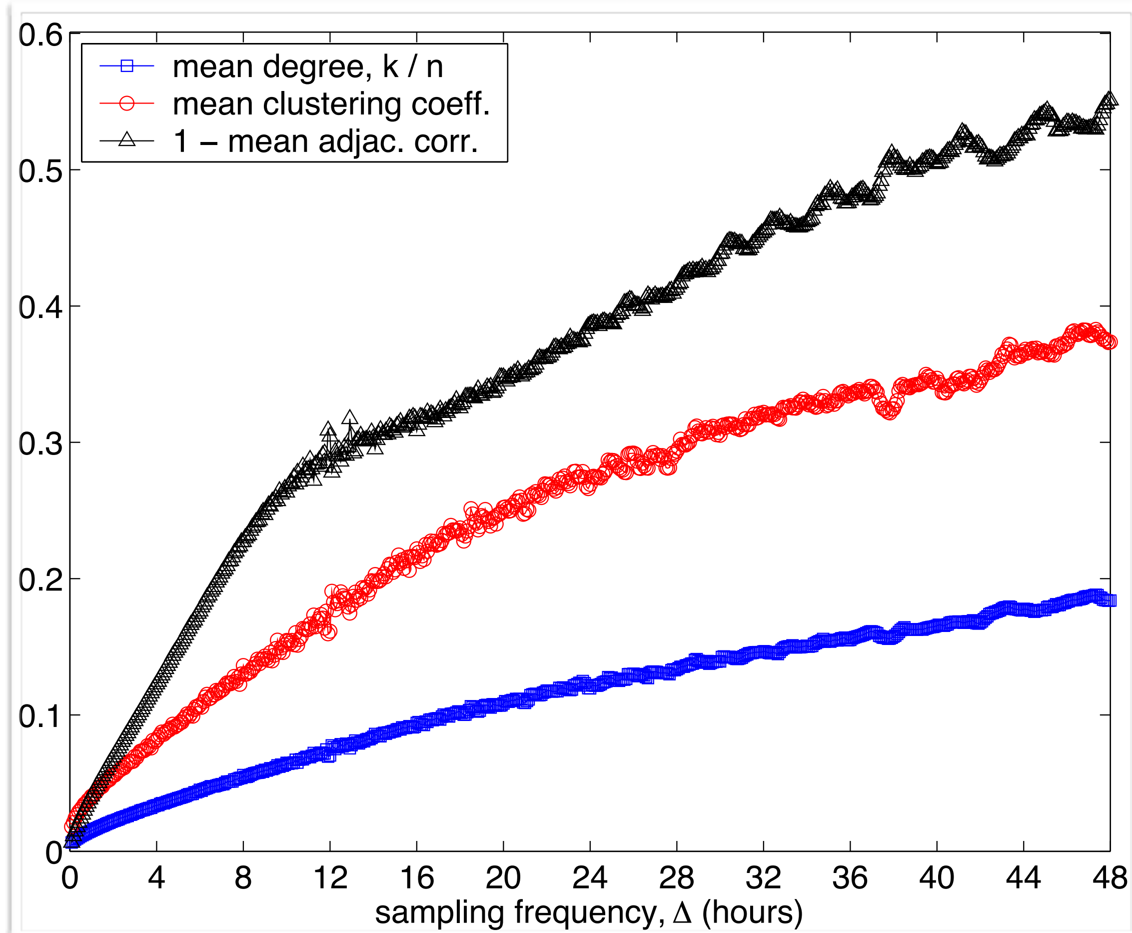
network dynamics

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

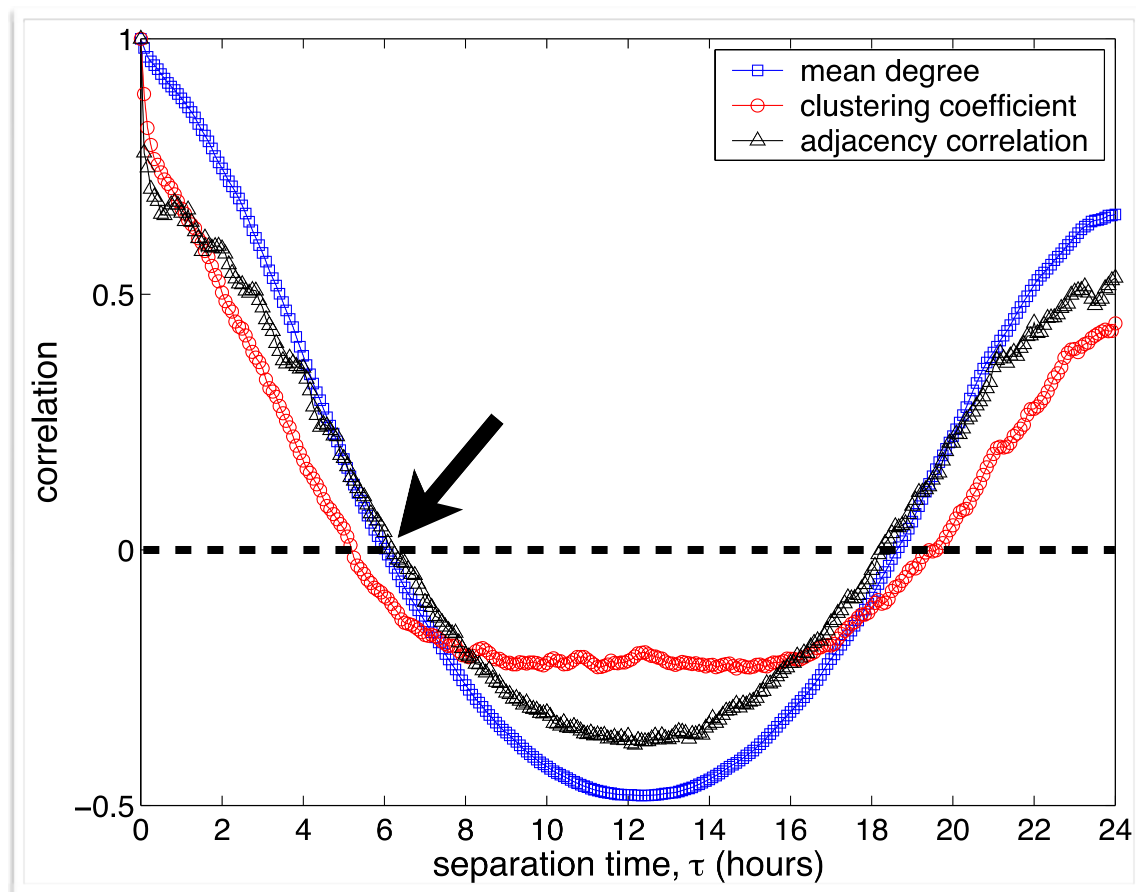
compute **autocorrelation function on network measures**

network dynamics

how to choose aggregation time?

recall highly periodic dynamics

compute **autocorrelation function on network measures**



network dynamics

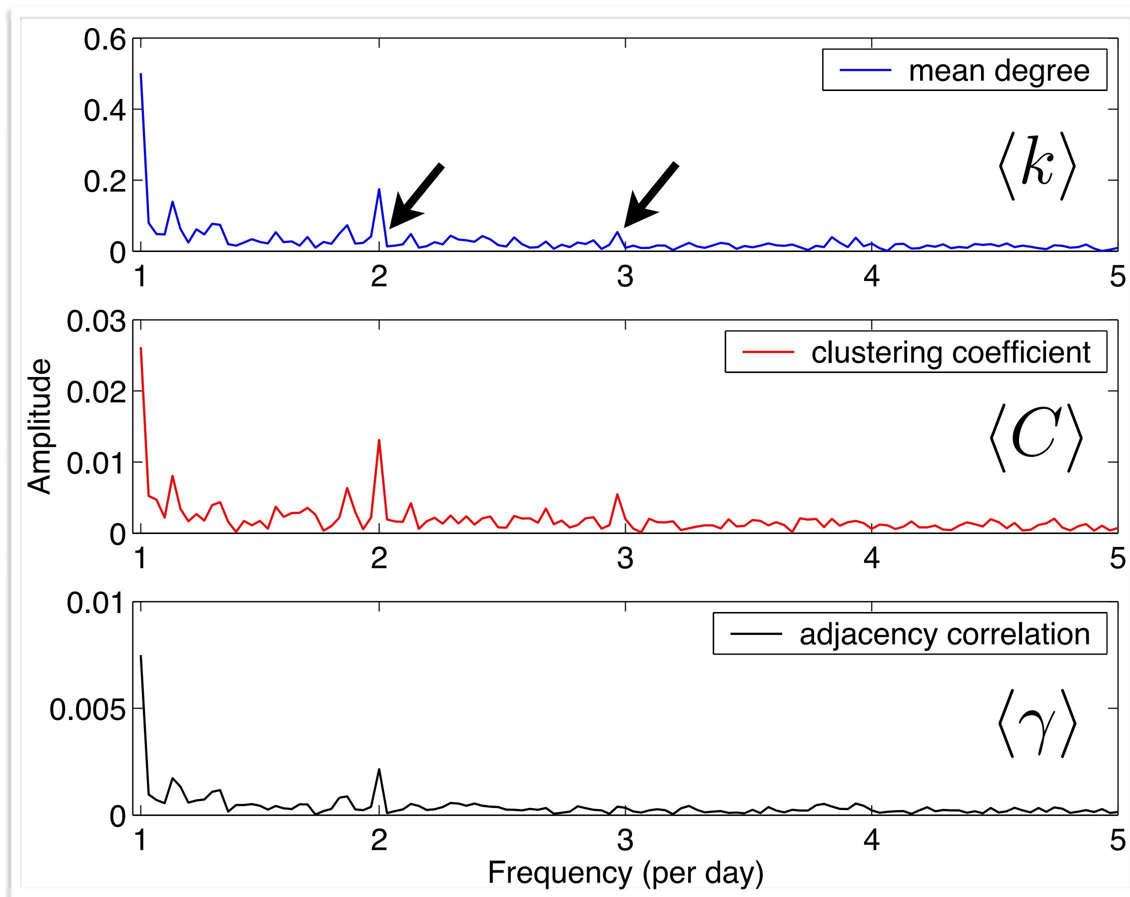
how to choose aggregation time?

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.

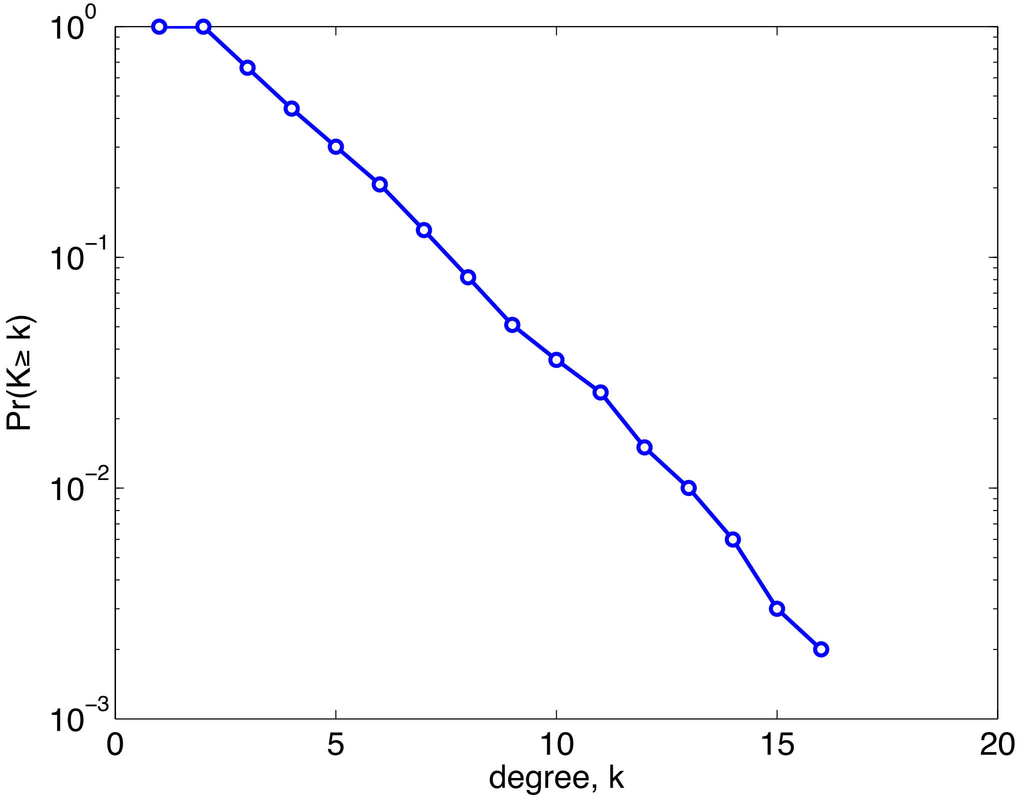
"densification laws"

- small-world phenomenon: diameter = $O(\log n)$
- how does diameter change in evolving networks?
- consider simple randomly-grown network:
  - at each time t, add vertex with degree c
  - attach each new edge via uniform attachment mechanism $\Pr(k_i \to k_i + 1) \propto \text{const.}$

- easy to simulate numerically

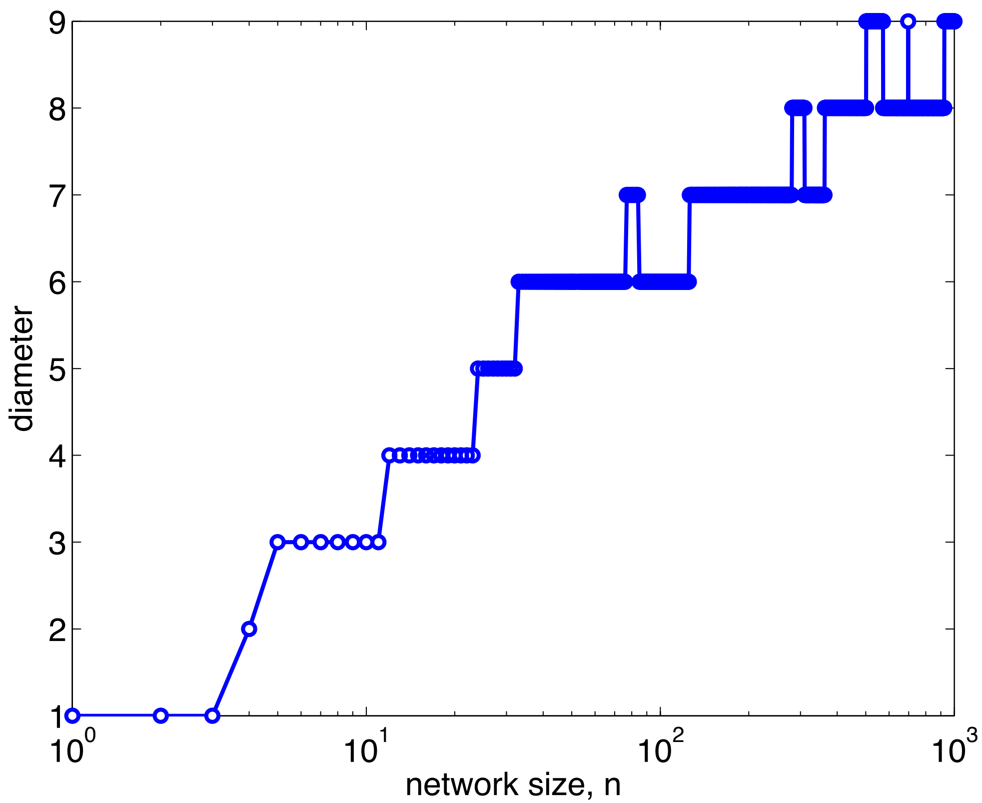
randomly grown networks

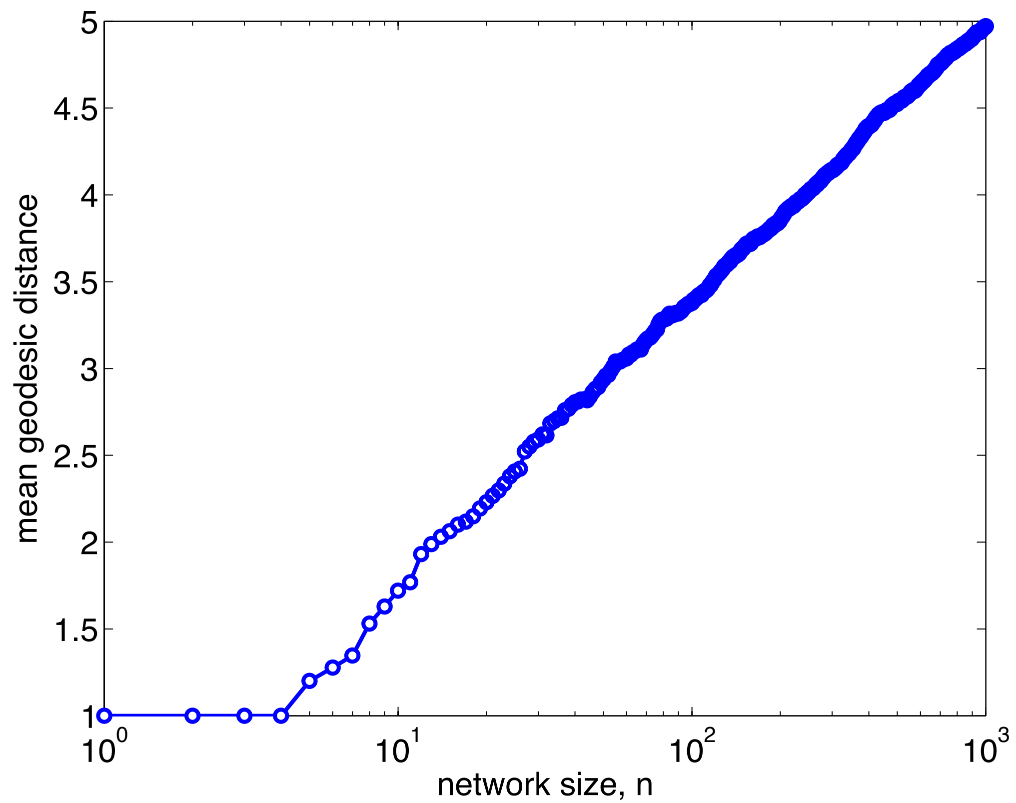
- choose $c = 2$ and $n = 10^3$



randomly grown networks

- choose $c = 2$ and $n = 10^3$ = increasing diameter $O(\log n)$





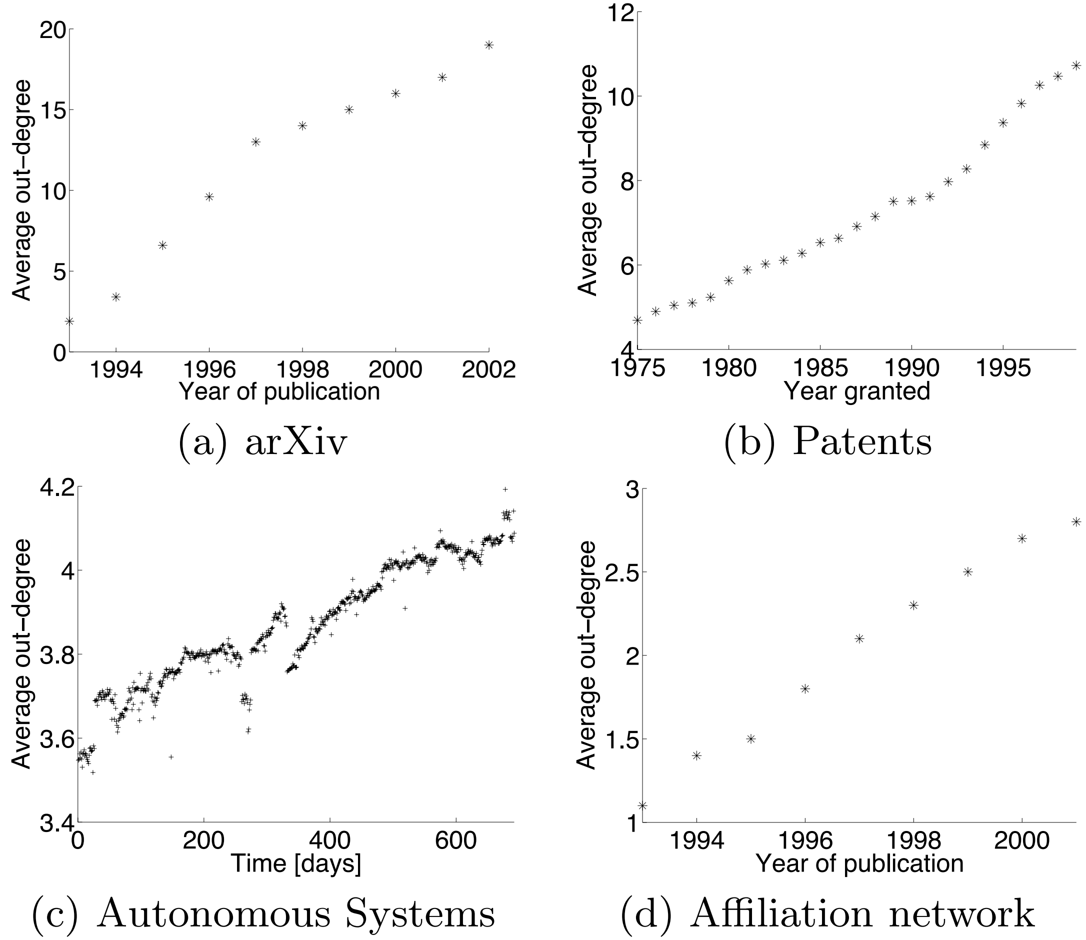
"densification laws"

- Leskovec, Kleinberg, Faloutsos (2005)
- examined 4 networks:
  - citation network (from arxiv.org),
  - US patents citation networks (from NBER)
  - Autonomous Systems (BGP) graph
  - author-paper bipartite network (from arxiv.org)

Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

"densification laws"

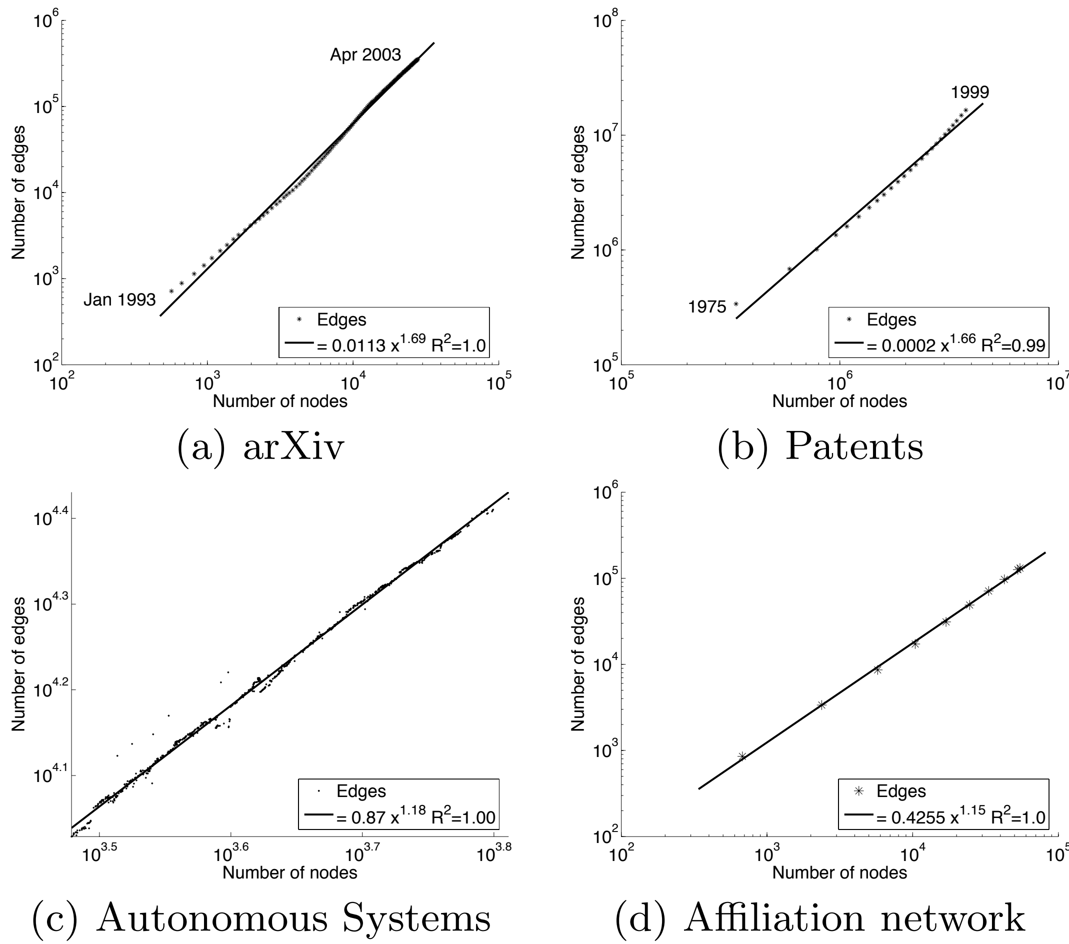
- mean degree over time



Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

"densification laws"

- mean degree (again) over time



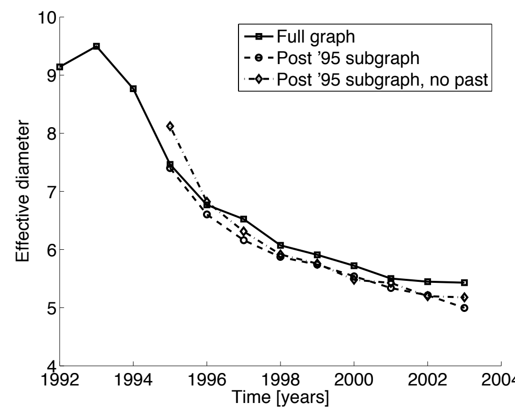
(a) arXiv (b) Patents

(c) Autonomous Systems (d) Affiliation network

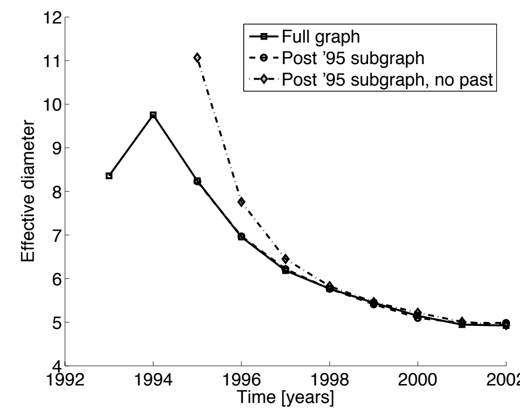
Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

"densification laws"

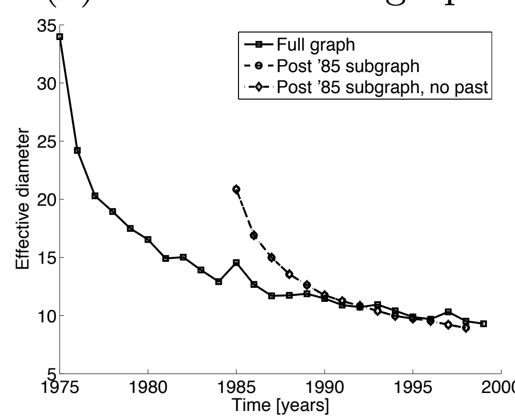
- "effective" diameter over time (90% of node-pairs within this distance)



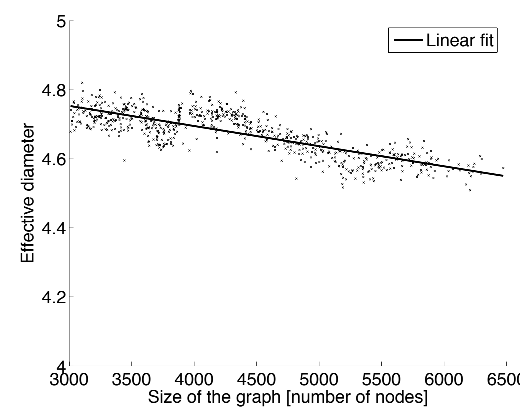
(a) arXiv citation graph



(b) Affiliation network



(c) Patents



(d) AS

Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

"densification laws"

- (effective) diameter is clearly shrinking in these networks
- mean degree is seems to be increasing, but slowly
- key questions:
  - is the mean degree really increasing? [need statistics]
  - can an increasing mean degree cause shrinking diameter in a growing network? [need a model]
  - how much does it need to increase? [more model]
  - what else could be going on with these networks? [need intuition, more data]

Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

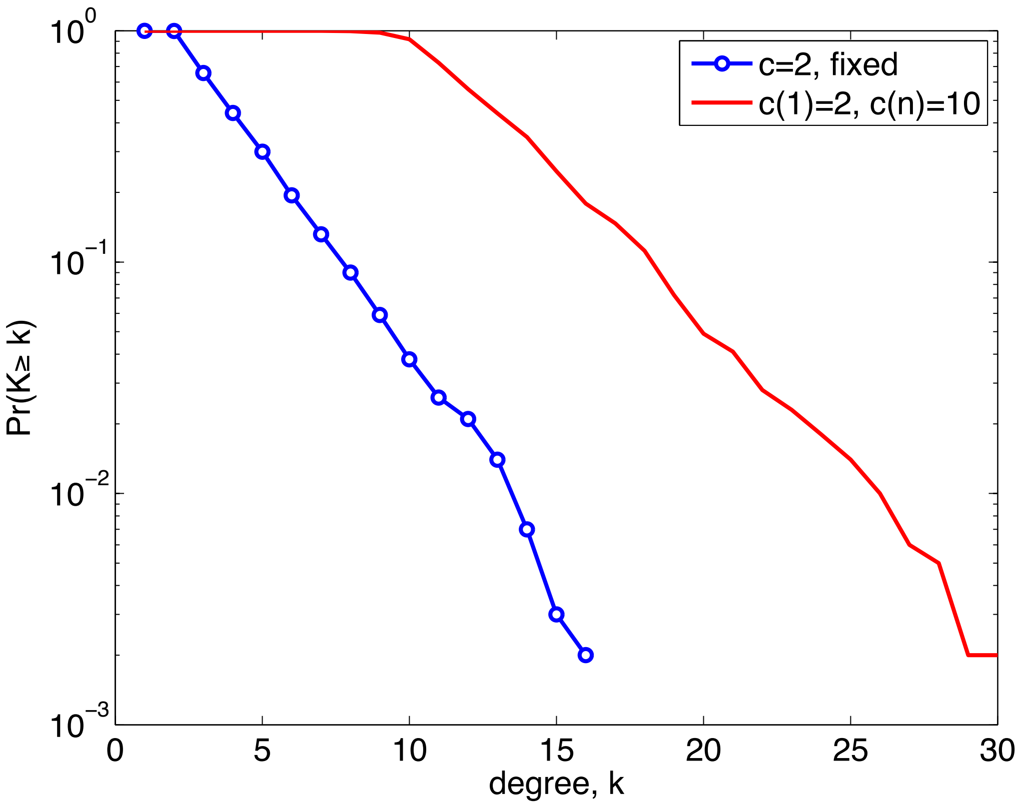
randomly grown networks (redux)

- how to get a shrinking diameter?
- idea: increase degree over time in our simple model
- simple randomly-grown network:
  - at each time t, add vertex with degree $c(t) \propto t$
  - attach each new edge via uniform attachment mechanism $\Pr(k_i \to k_i + 1) \propto \text{const.}$

- easy to simulate numerically

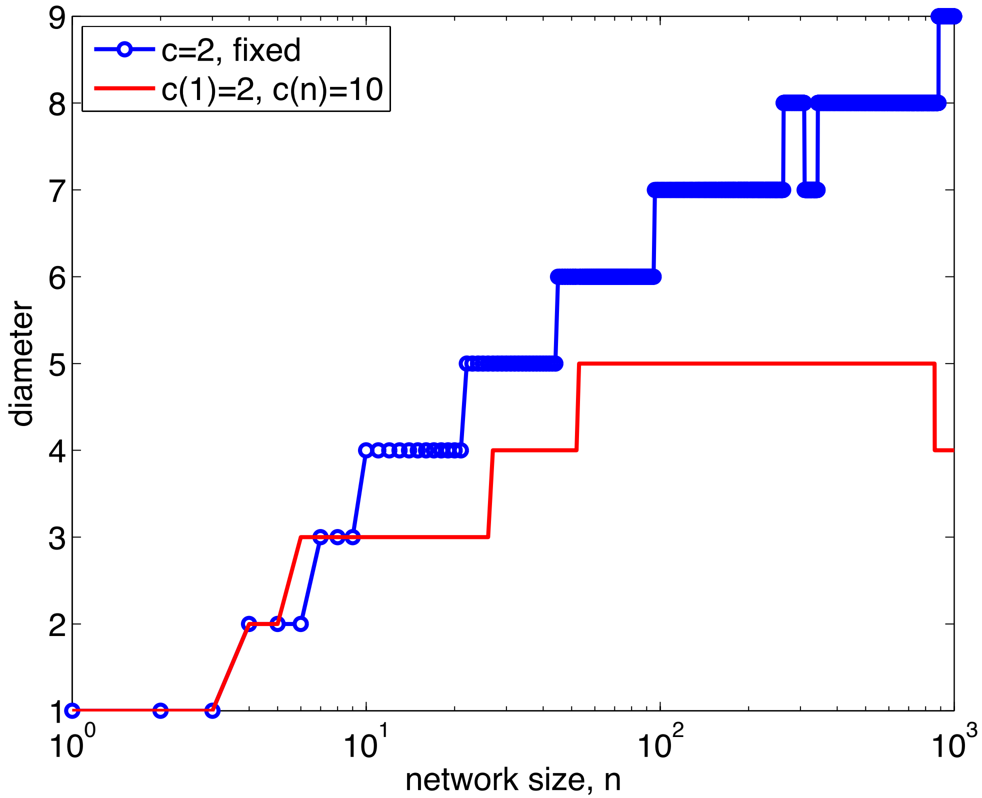
randomly grown networks (redux)

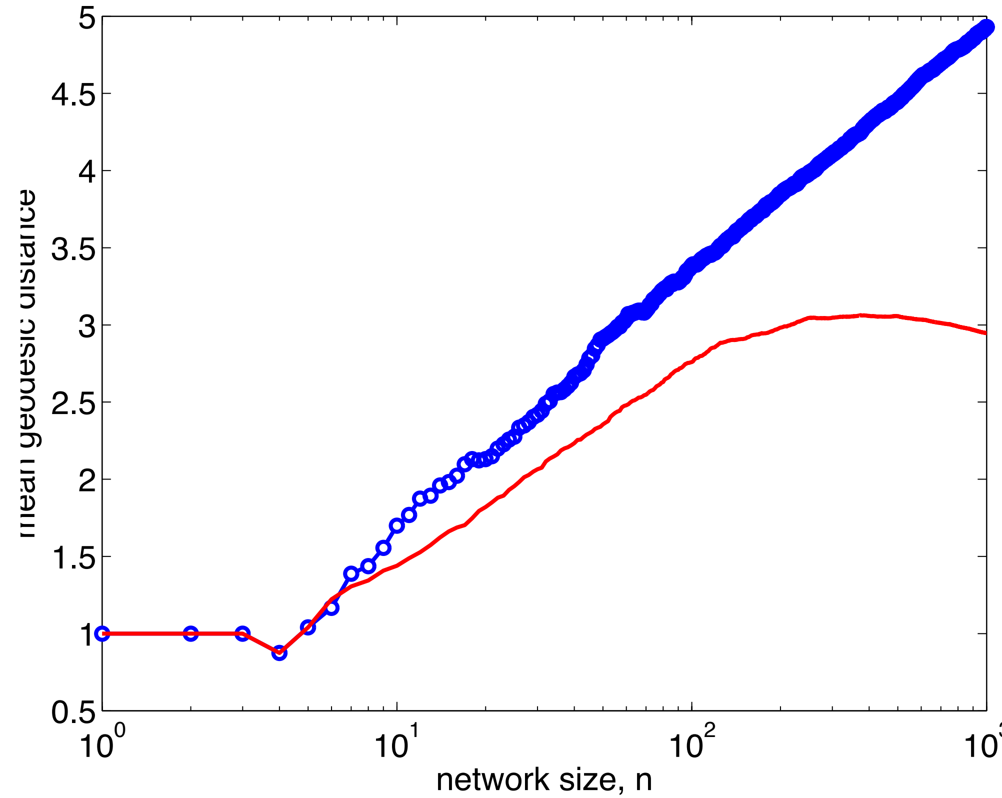
- choose $c(1) = 2$, growing linearly to $c(n) = 10$ for $n = 10^3$



randomly grown networks (redux)

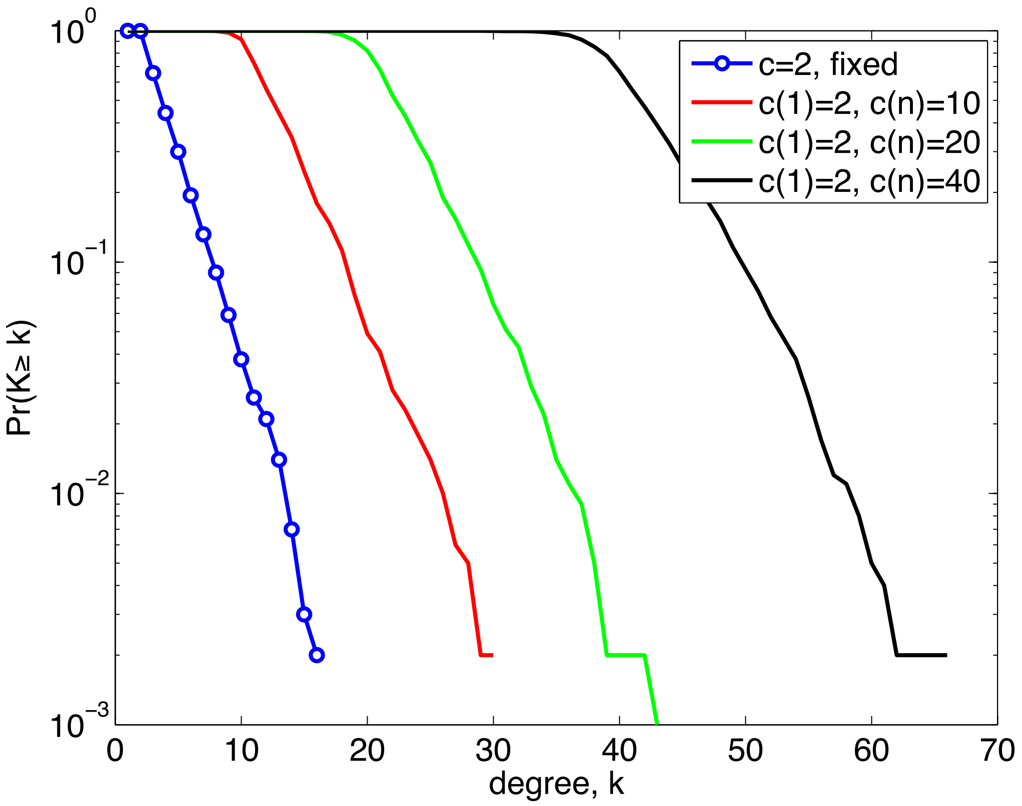
- choose $c(1) = 2$, growing linearly to $c(n) = 10$ for $n = 10^3$





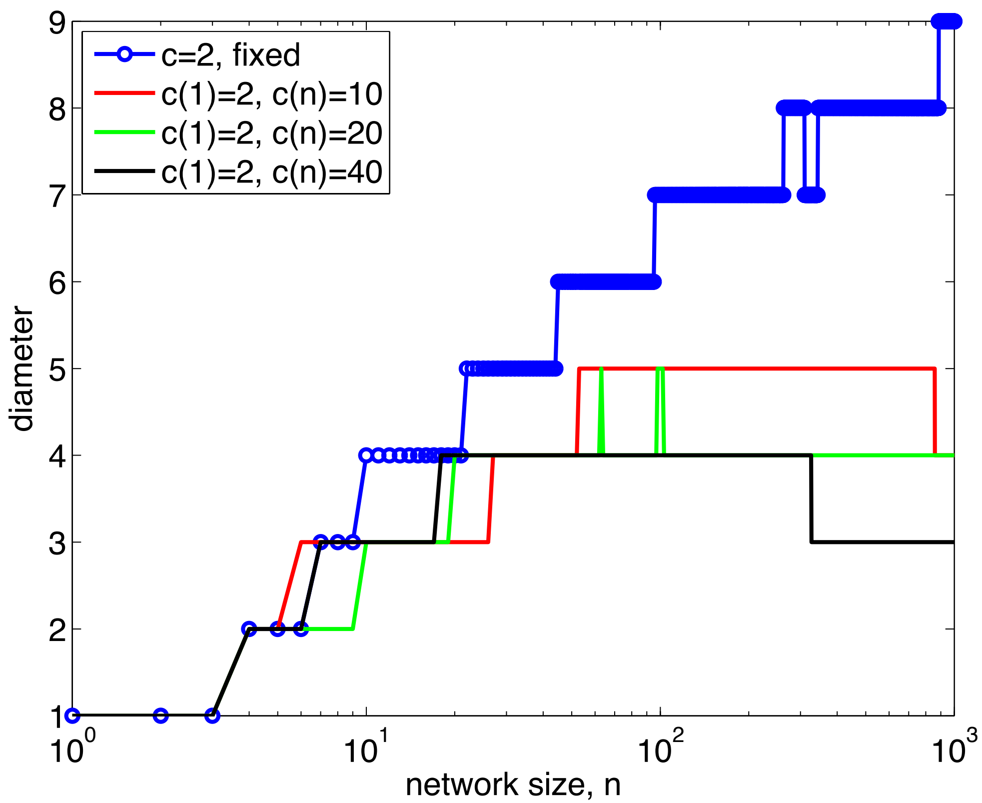
randomly grown networks (redux)

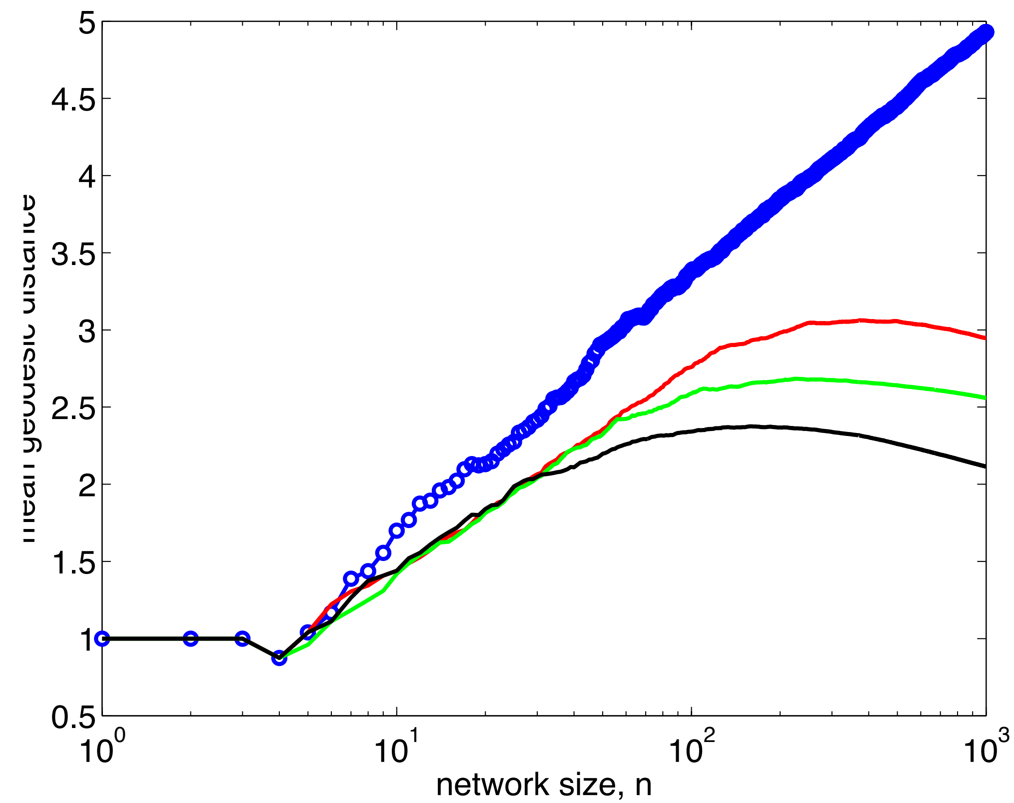
- choose $c(1) = 2$, growing linearly, for $n = 10^3$



randomly grown networks (redux)

- choose $c(1) = 2$, growing linearly, for $n = 10^3$





randomly grown networks (redux)

- very simple model!
- increasing mean degree can shrink the diameter
- but not *enough* shrinkage under this model
- how should we improve the model?
  [what’s "wrong" with our model?]

"densification laws"

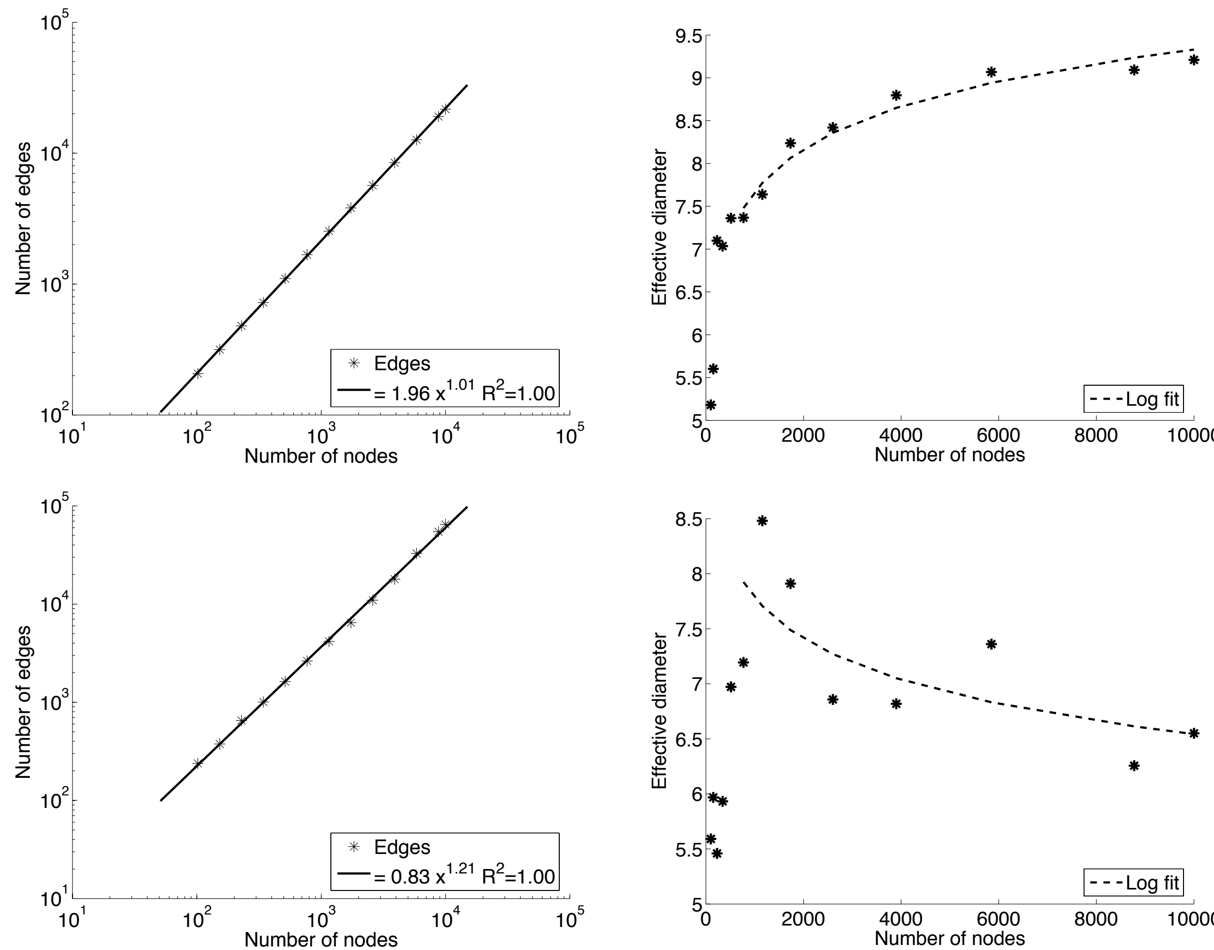
- Leskovec, Kleinberg, Faloutsos propose
  - community guided attachment
    [links form preferentially within a vertex's community]
  - "forest fire model" : a kind of recursive vertex-copy model
    1. each new node u chooses uniformly random existing node v
    2. links to each of v neighbors* with probability p
    3. repeat step 2 for each of linked neighbor

Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

* FF model is presented here as undirected. in LKF2005, FF model is directed, and in-links are selected r times less often than out-links in the burning

"densification laws"

- simulated results. FF model can produce both growing and shrinking diameters
- yields heavy-tailed degree distributions
- also makes lots of triangles



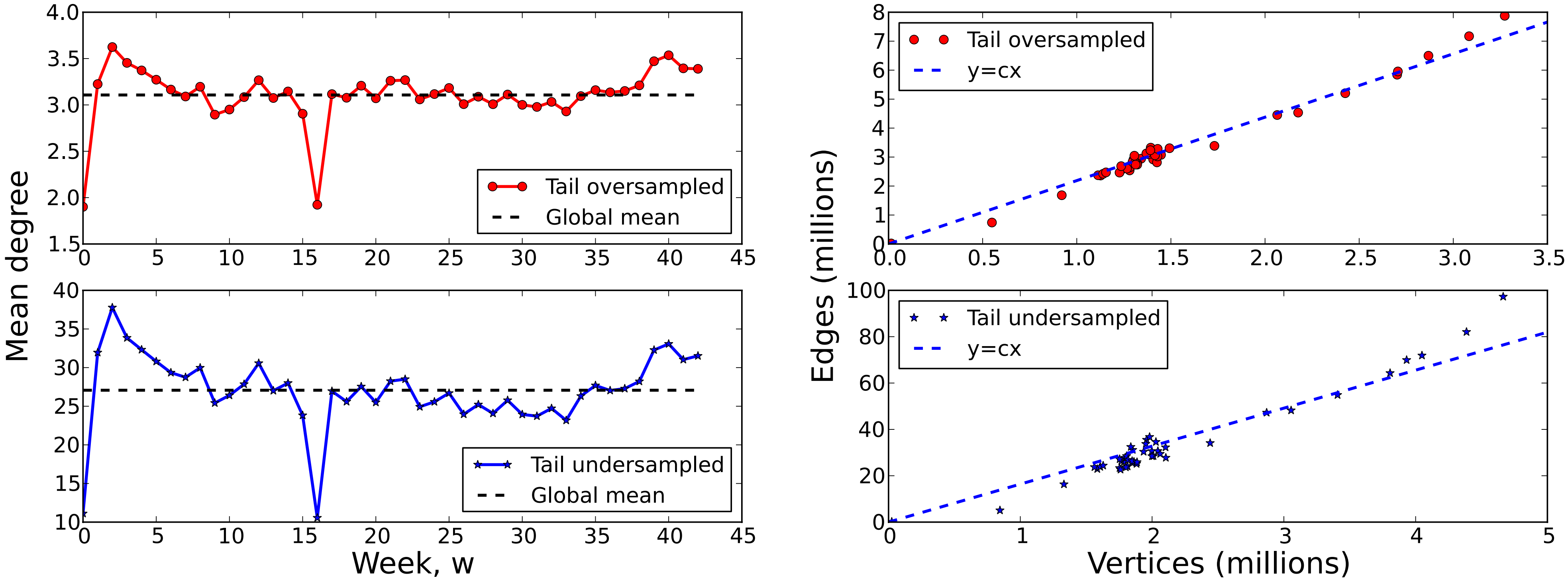
Leskovec, Kleinberg, Faloutsos, "Graphs over Time: Densification Laws, Shrinking Diameters, and Possible Explanations." KDD (2005)

are these results universal?

- densification "law"
- suggestion that this pattern occurs in all time-varying (growing?) networks

are these results universal?

- but it’s not universal
- case in point: human friendship network (from Halo data)
- mean degree constant, so no densification behavior



are these results universal?

- but it's not universal
- case in point: human friendship network (from Halo data)
- distances stable, so no shrinking diameter (despite ongoing turnover in vertices, edges)
- how can this be?
    - *idea: non-trivial edge maintenance costs mean degrees must remain bounded, no densification, no shrinking diameter*