Complex networks: an introduction

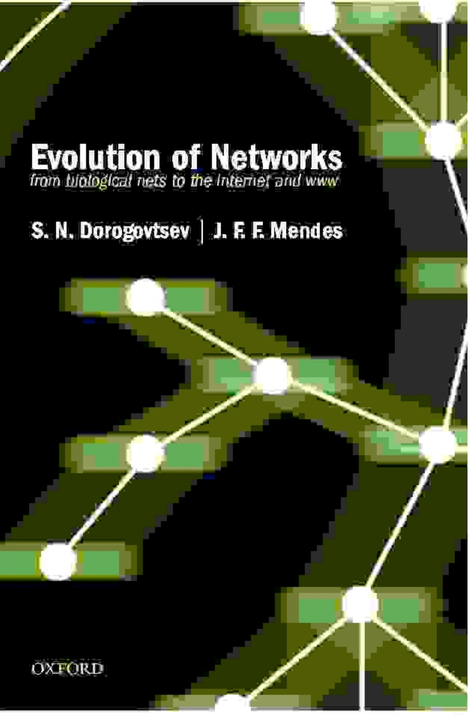
Alain Barrat

Centre de Physique Théorique, Marseille, France

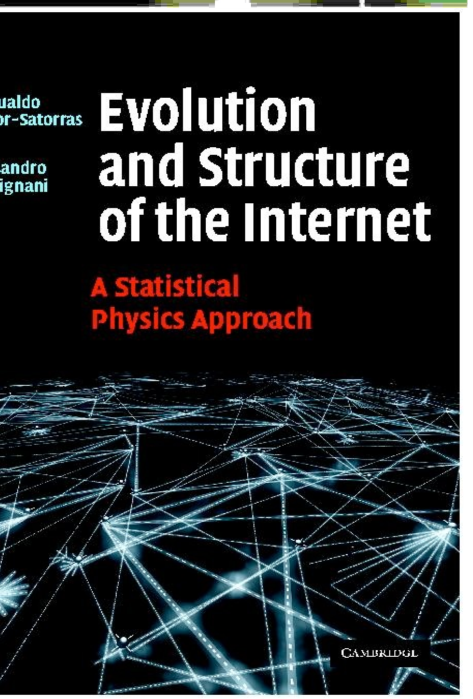
http://www.cpt.univ-mrs.fr/~barrat
http://www.sociopatterns.org
http://dataredux.weebly.com
alain.barrat@cpt.univ-mrs.fr

REVIEWS:

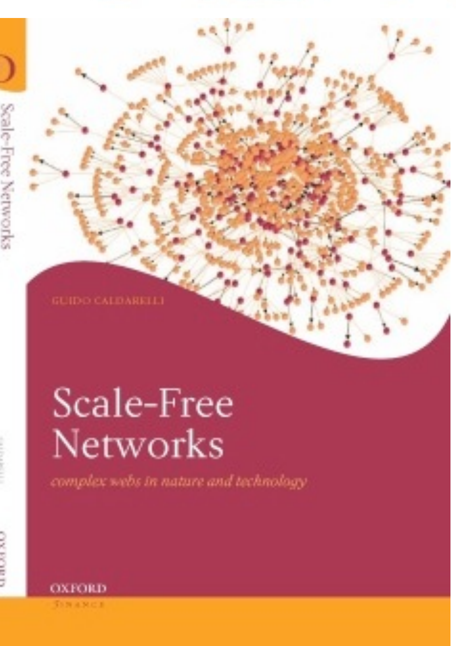
•**Statistical mechanics of complex networks**
R. Albert, A.-L. Barabasi, Reviews of Modern Physics 74, 47 (2002), cond-mat/0106096
•**The structure and function of complex networks**
M. E. J. Newman, SIAM Review 45, 167-256 (2003), cond-mat/0303516
•**Evolution of networks**
S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. 51, 1079 (2002) , cond-mat/0106144
•**Complex Networks: Structure and Dynamics**
S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Physics Reports 424 (2006) 175



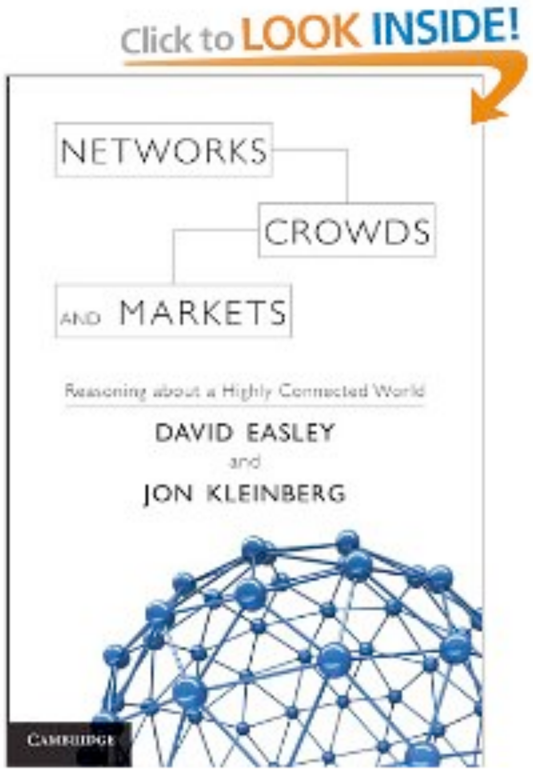
- ***Evolution of Networks: From Biological Nets to the Internet and WWW,*** S.N. Dorogovtsev and J.F.F. Mendes. Oxford University Press, Oxford, 2003.



- ***Evolution and Structure of the Internet: A Statistical Physics Approach***, R. Pastor-Satorras and A. Vespignani. Cambridge University Press, Cambridge, 2004.

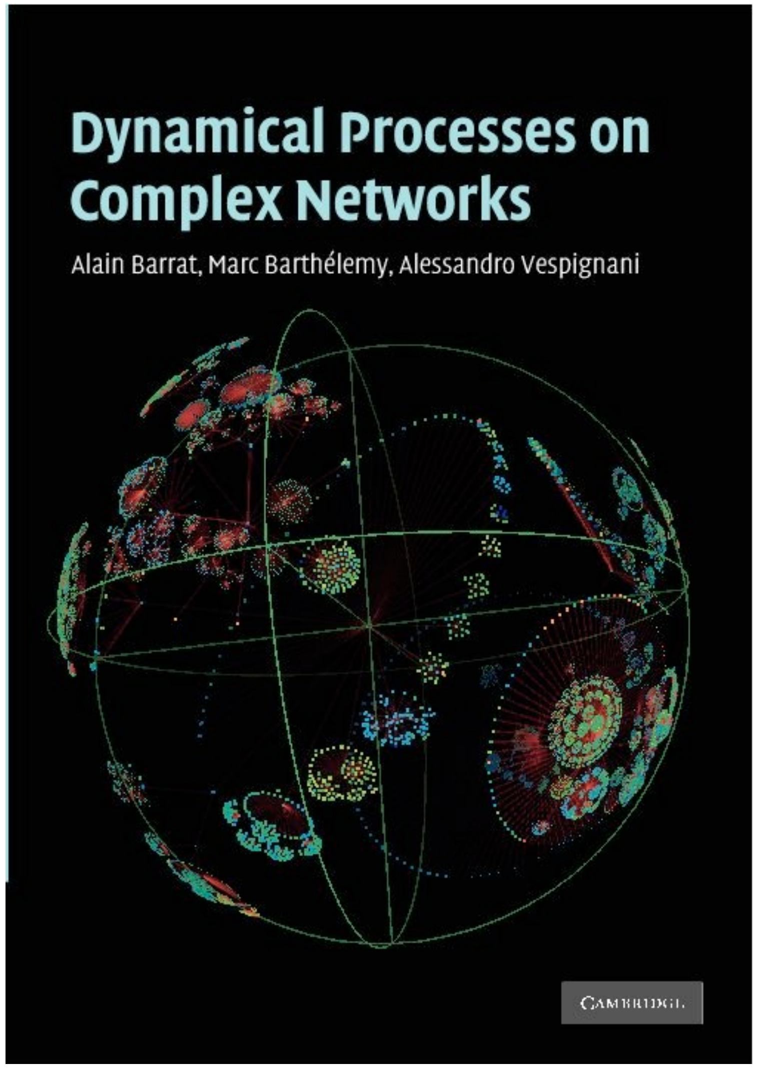


- ***Scale-free networks: Complex Webs in Nature and Technology,*** G. Caldarelli. Oxford University Press, Oxford, 2007



Networks, Crowds, and Markets: Reasoning About a Highly Connected World

D. Easley, J. Kleinberg

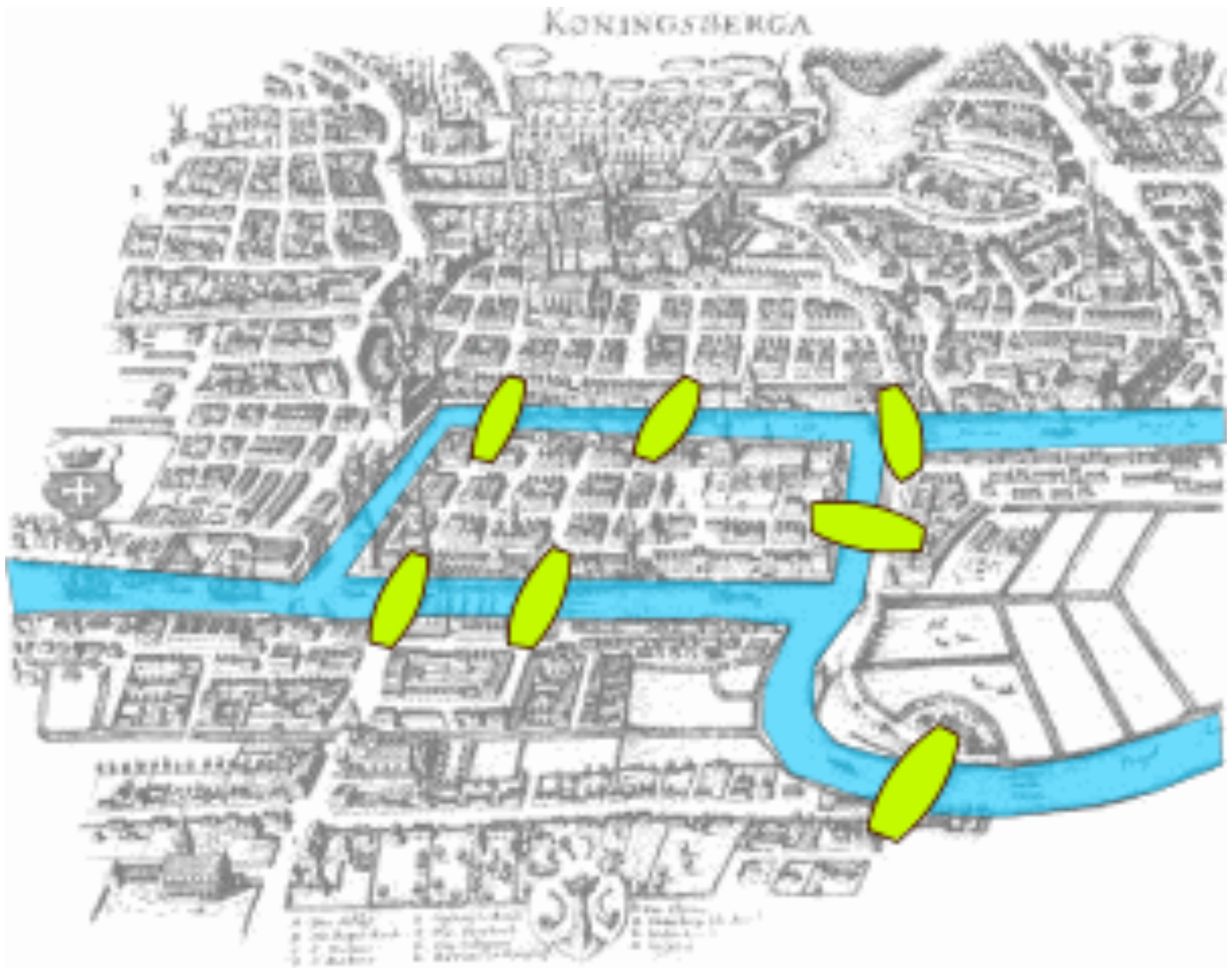


Outline of the lectures

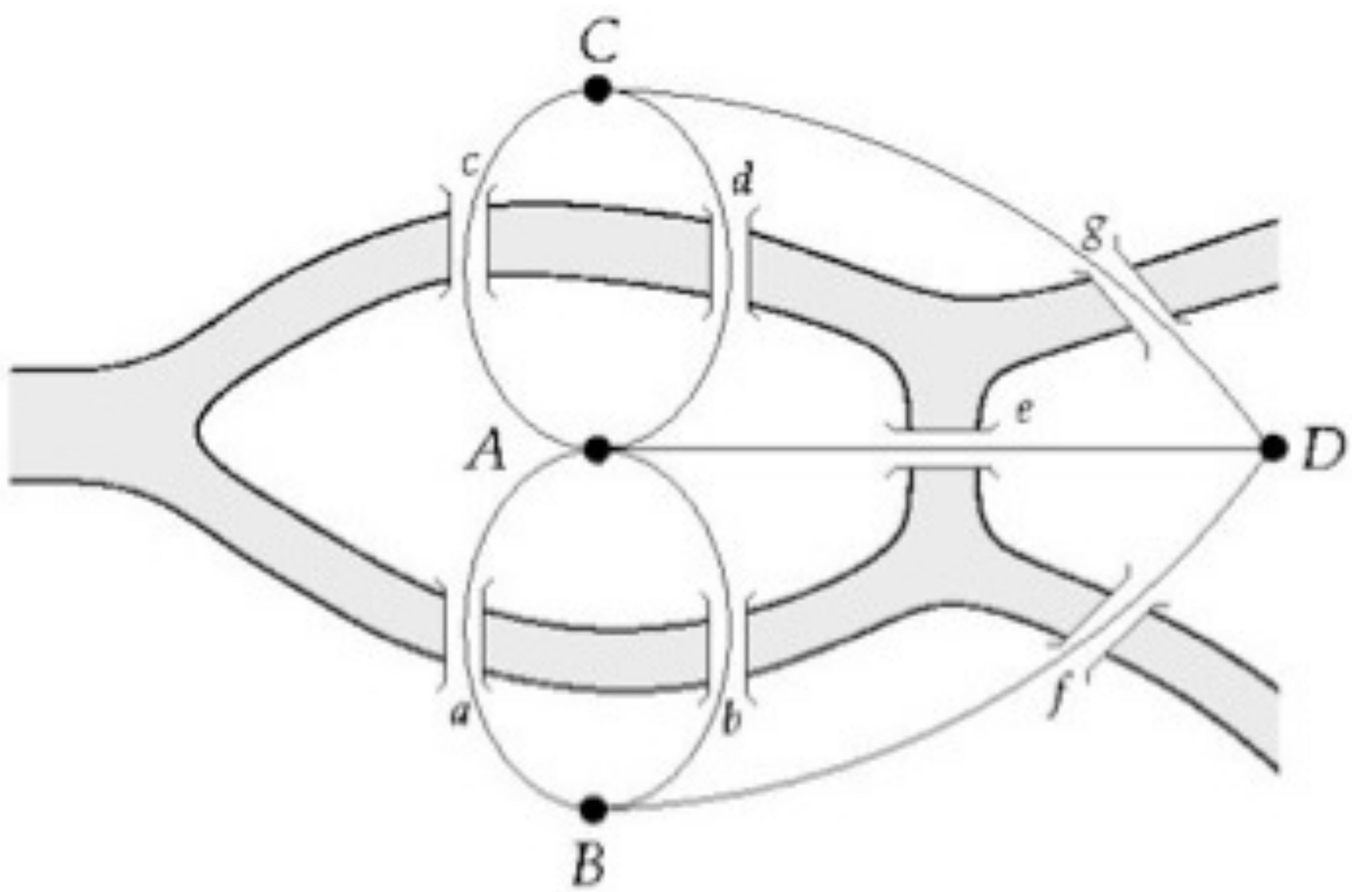
I. **Networks: definitions, statistical characterization, correlations, structures, hierarchies…**
II. Modeling frameworks
III. Resilience, vulnerability
IV. Temporal networks

The bridges of Koenigsberg

(now Kaliningrad, Russia)



L. Euler (18th century):
Can one walk across the seven bridges and never cross the same bridge twice?



land areas = nodes
bridges = connections

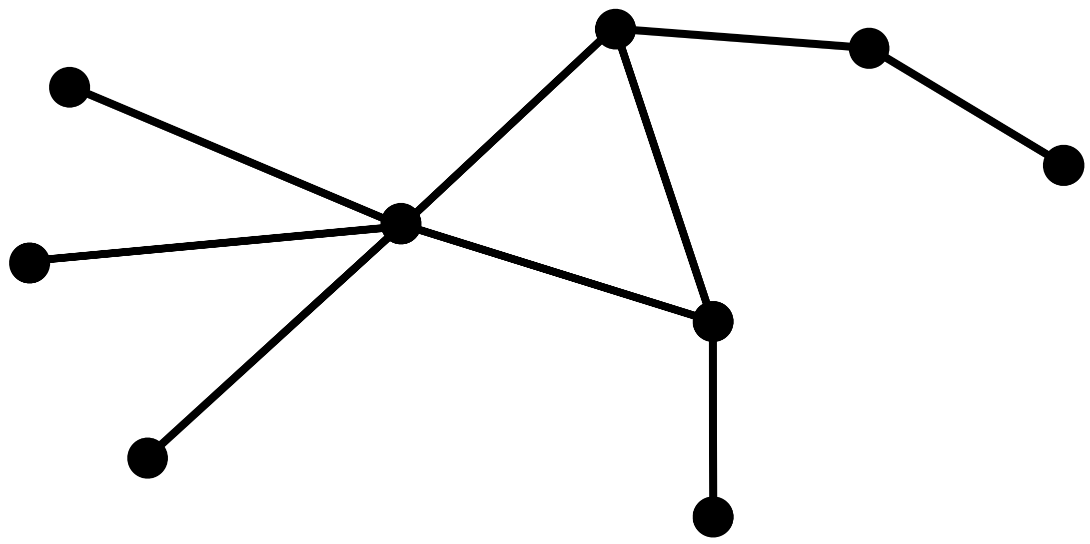
1735: **Leonhard Euler's theorem:**

(a) If a graph has nodes of odd degree, there is no path.

(b) If a graph is connected and has no odd degree nodes, it has at least one path.

What is a network

Network=set V of nodes joined by links (set E)

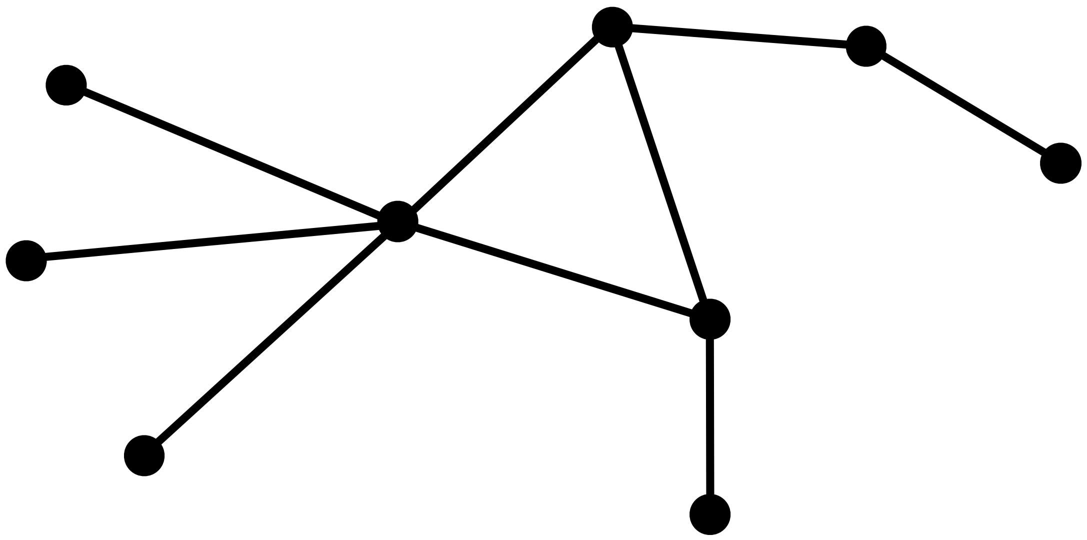


very abstract representation

very general

convenient to describe many different systems

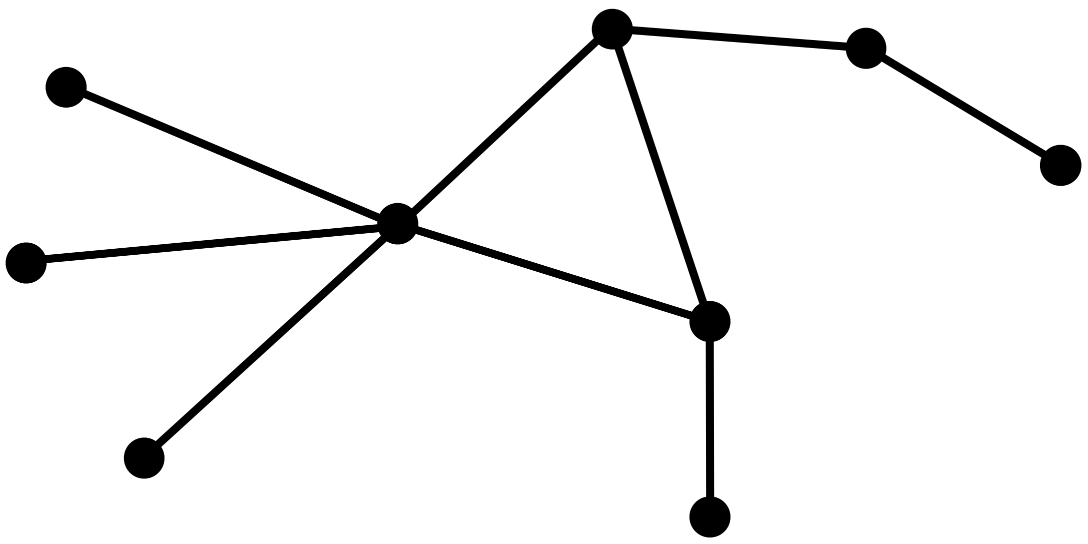
Graphs



graph theory

abstract tools for the description of graphs (degrees, paths, distances, cliques, etc…)

Networks



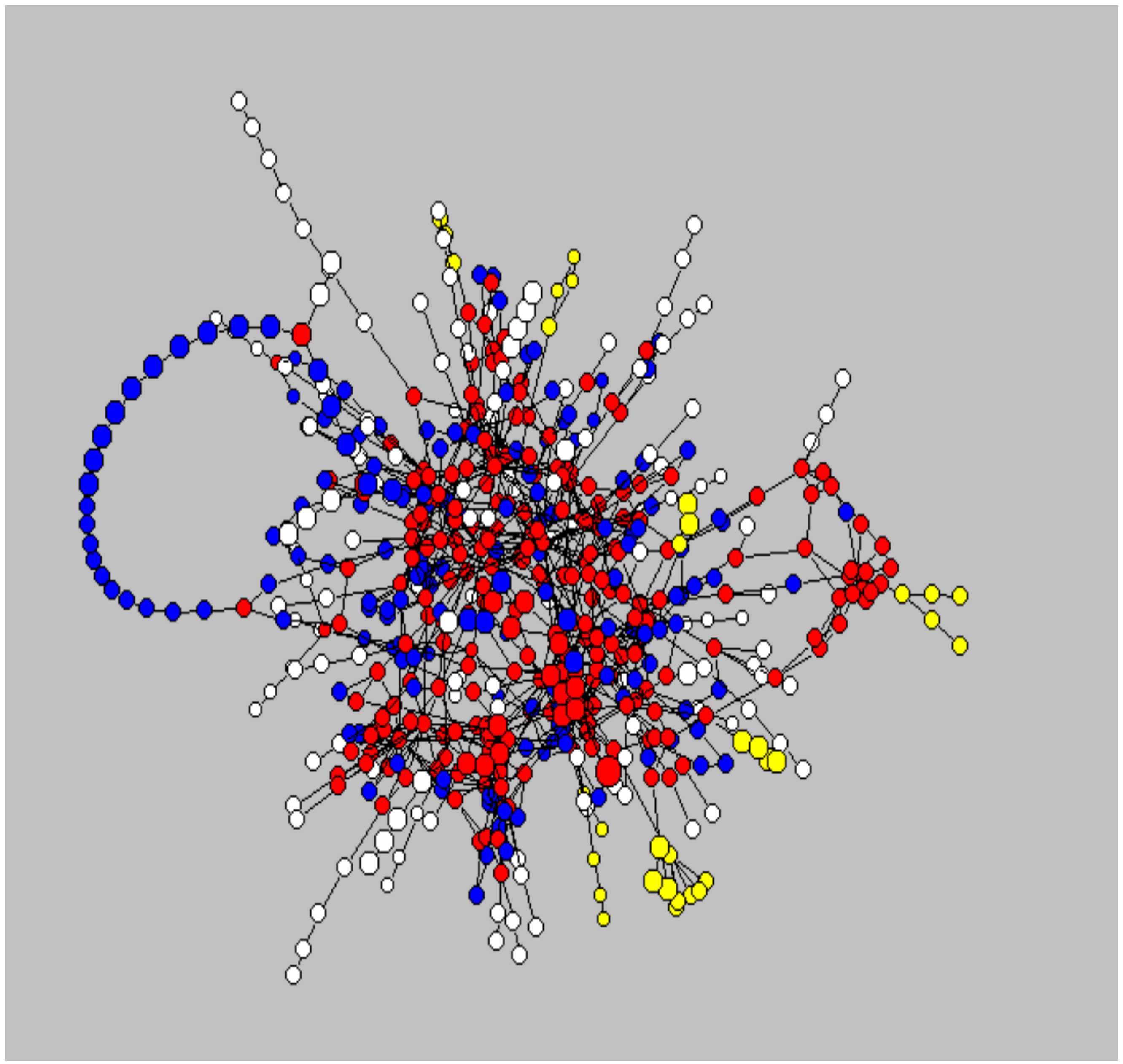
Nodes:
persons
computers
webpages
airports
molecules
....

Links:
social relationships
cables
hyperlinks
air-transportation
chemical reactions
....

Examples

Metabolic Network

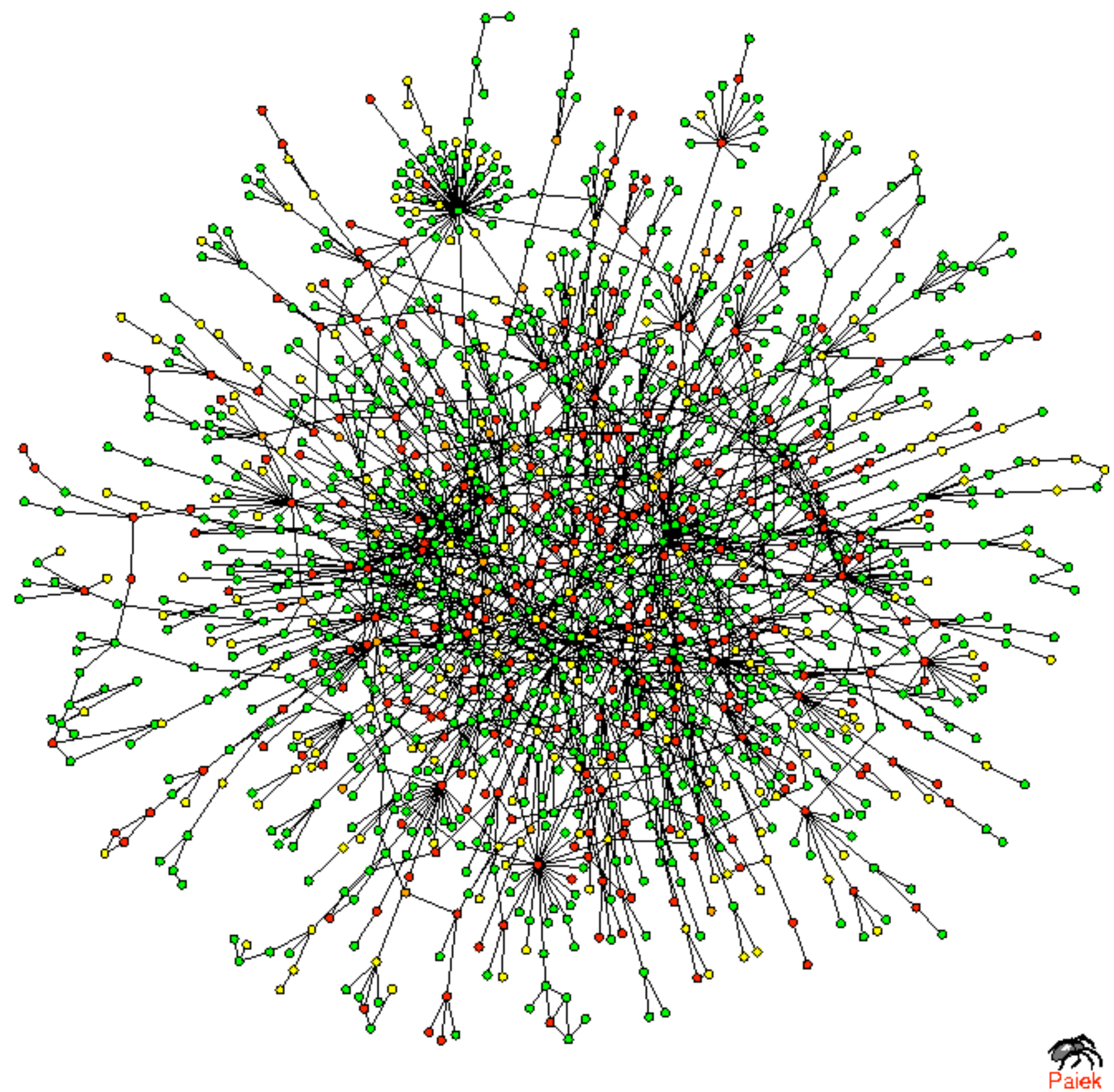
Nodes: metabolites
Links:chemical reactions



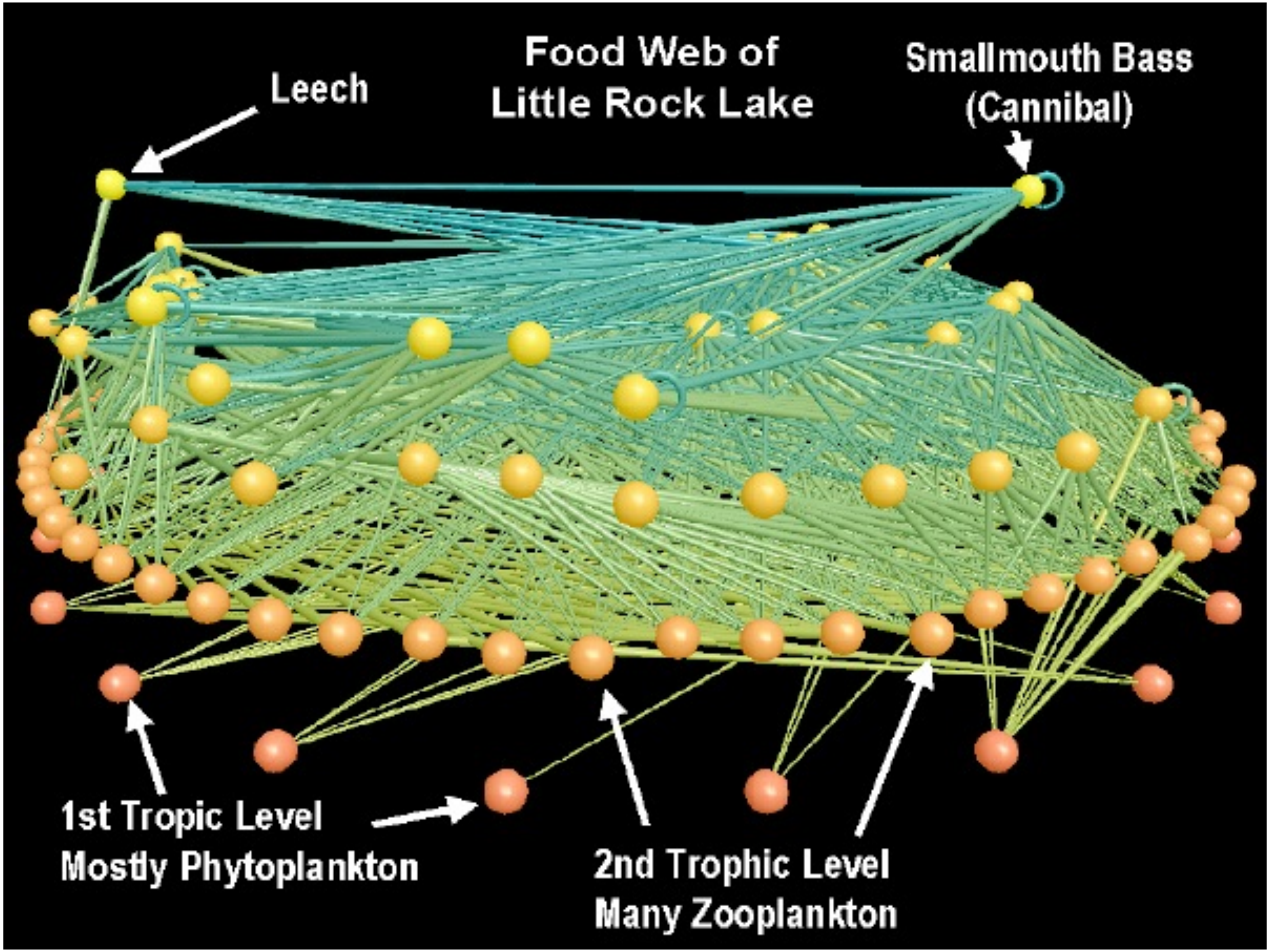
Protein Interactions

Nodes: proteins

Links: interactions

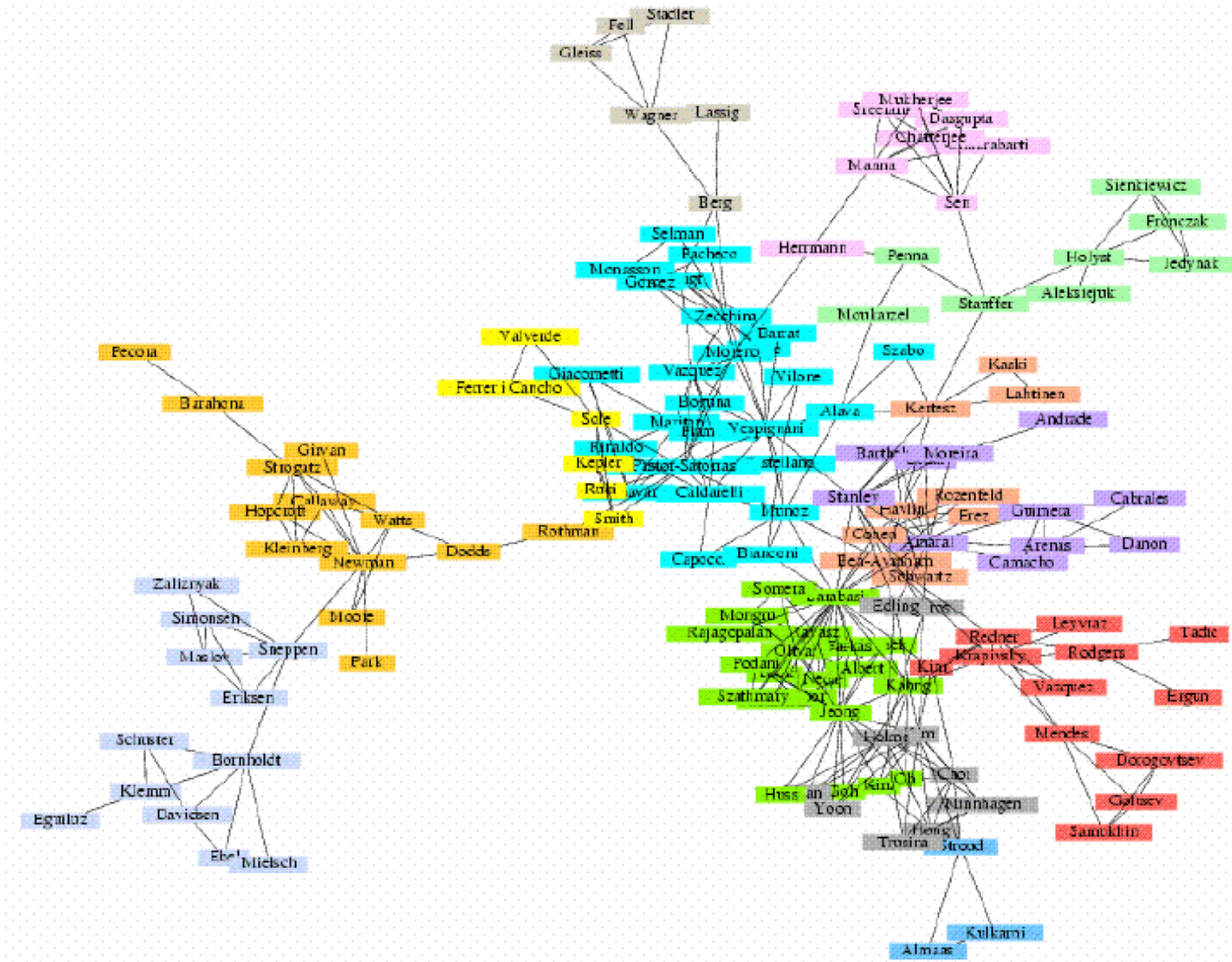


Food-webs



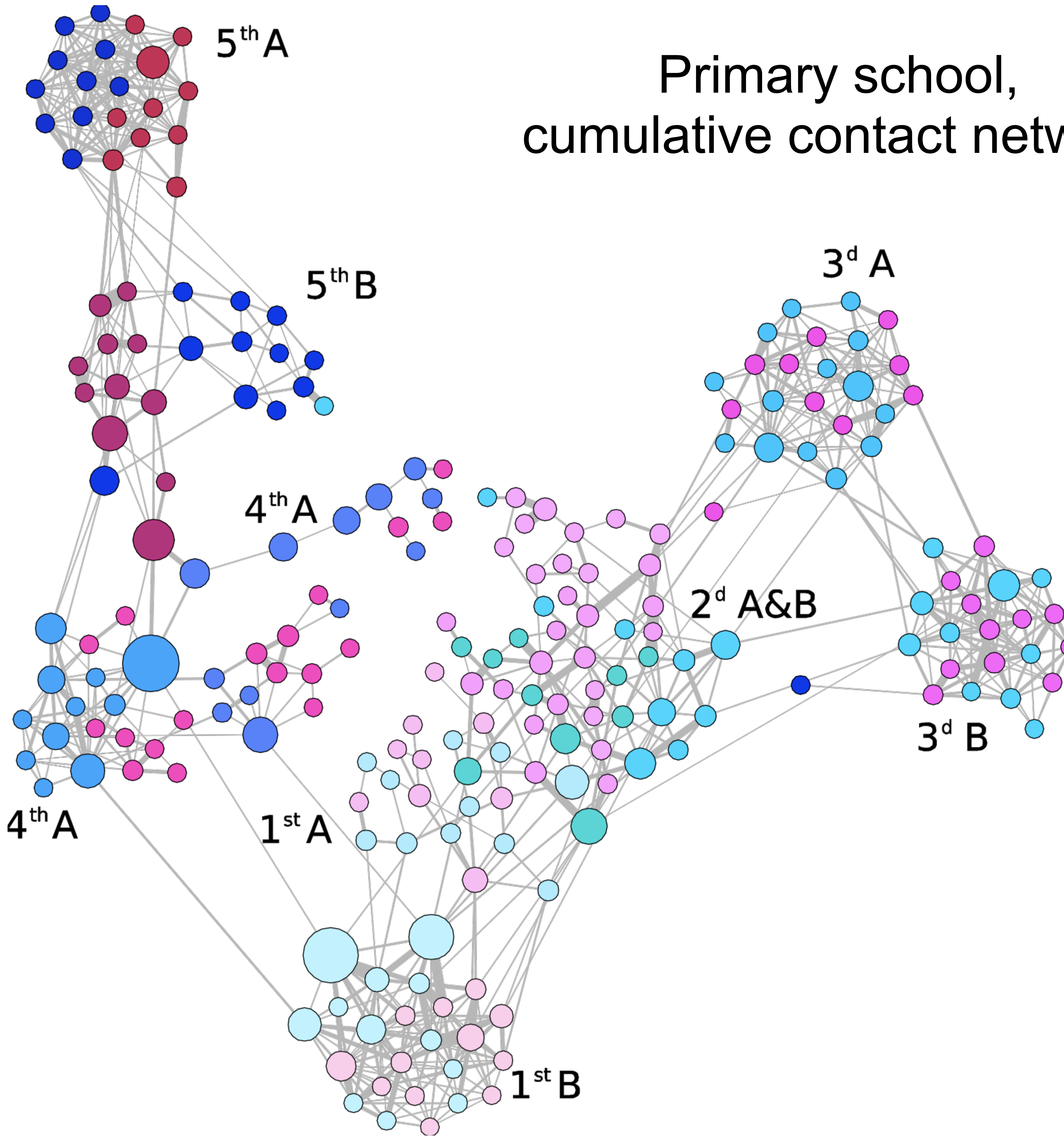
N. Martinez, R. Williams

Scientific collaboration network



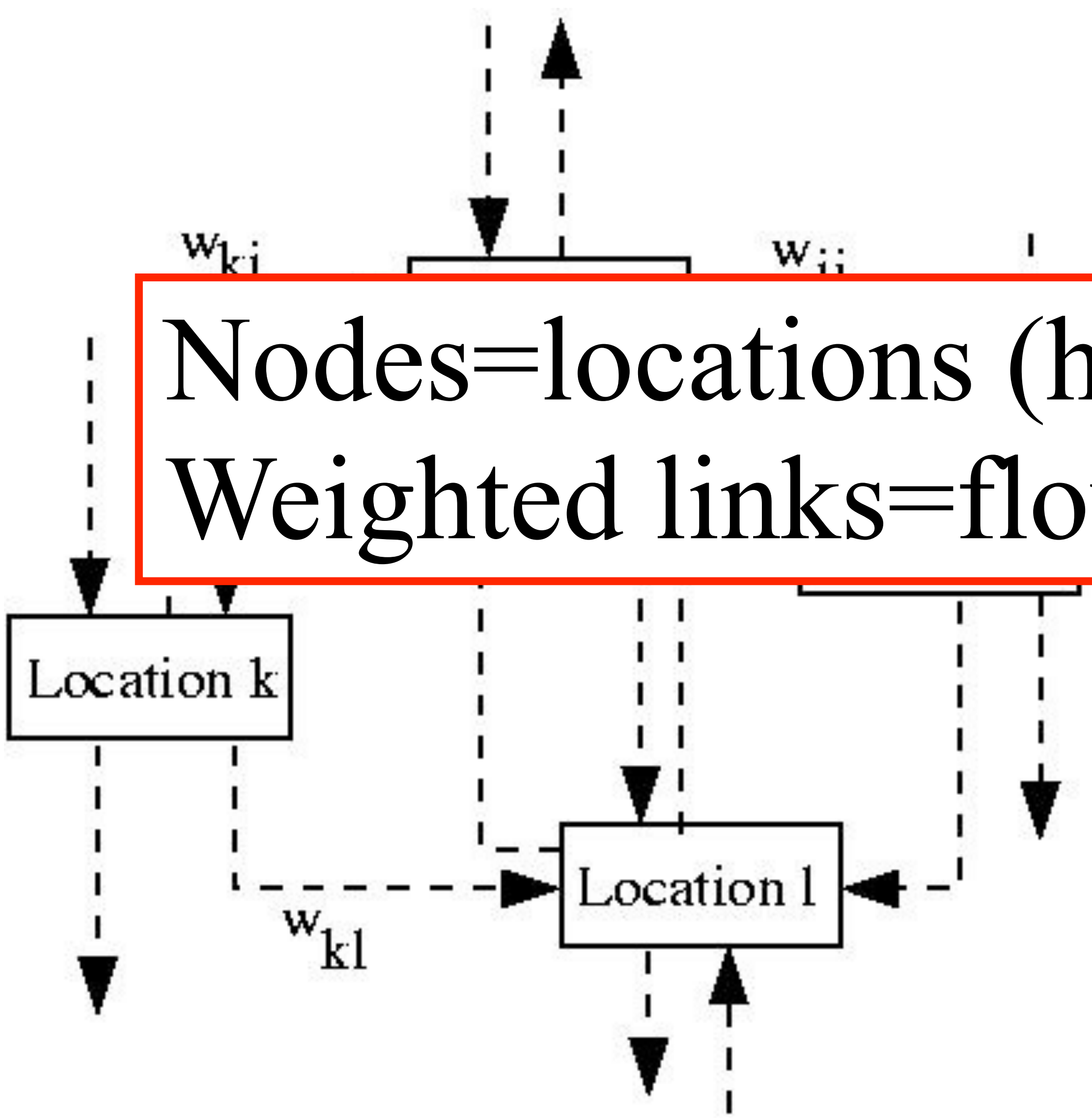
M. E. J. Newman and M. Girvan, *Physical Review E* **69**, 026113 (2004).
Image: MEJ Newman, http://www-personal.umich.edu/~mejn/networks/

Primary school, cumulative contact network



Online (virtual) social networks

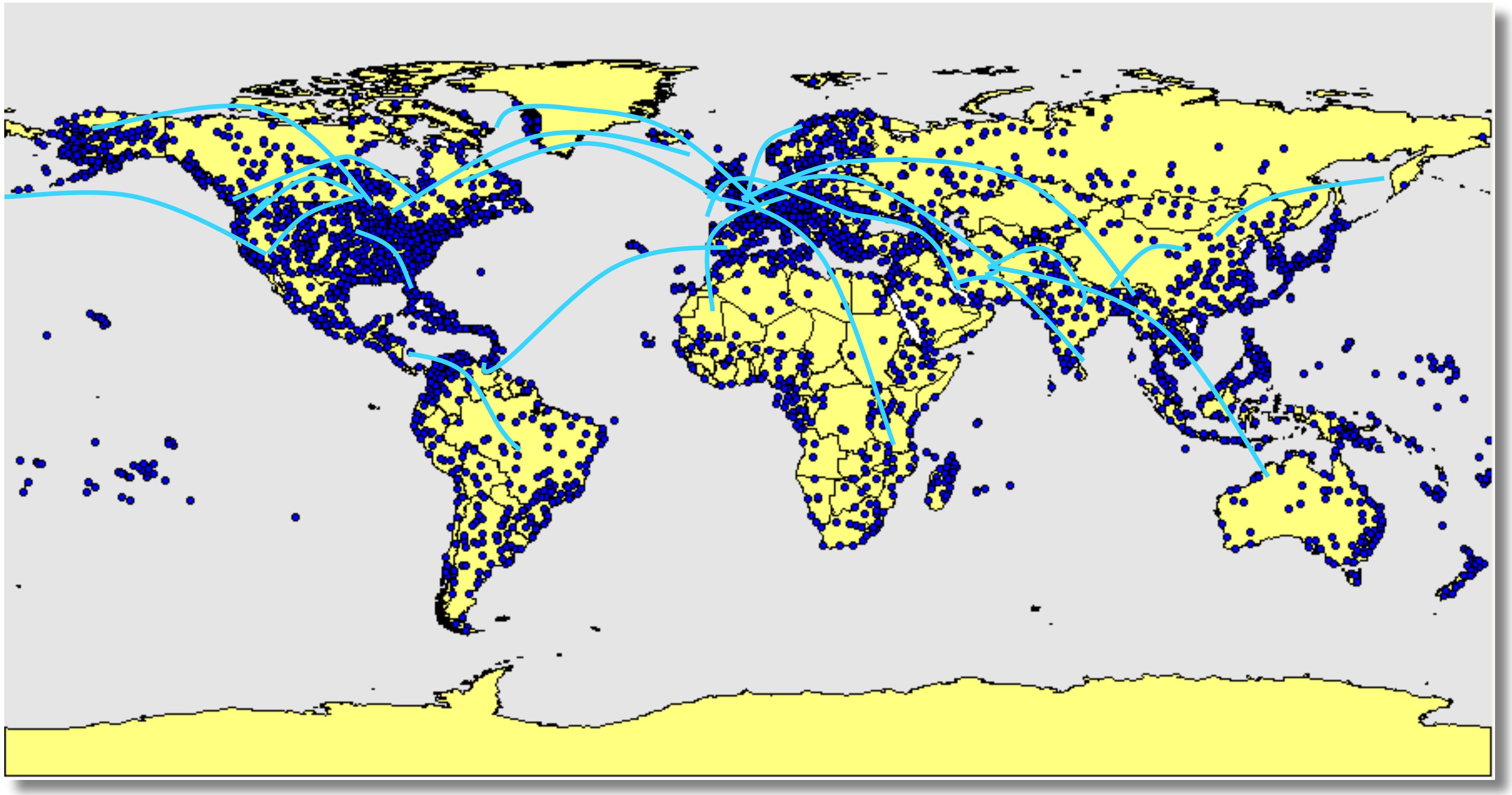
Transportation networks: Urban level



Nodes=locations (homes, shops, offices…)
Weighted links=flow of individuals

Person ID	Location ID	Location type	Arrival time	Departure time
116	4356	Home	00:00	07:00
116	21343	Social	19:30	21:00
116	4356	Home	21:00	07:00
324	12679	Home	00:00	07:00
324	431	School	08:00	14:00
324	12679	Home	14:30	19:00
⋮	⋮	⋮	⋮	⋮

World airport network



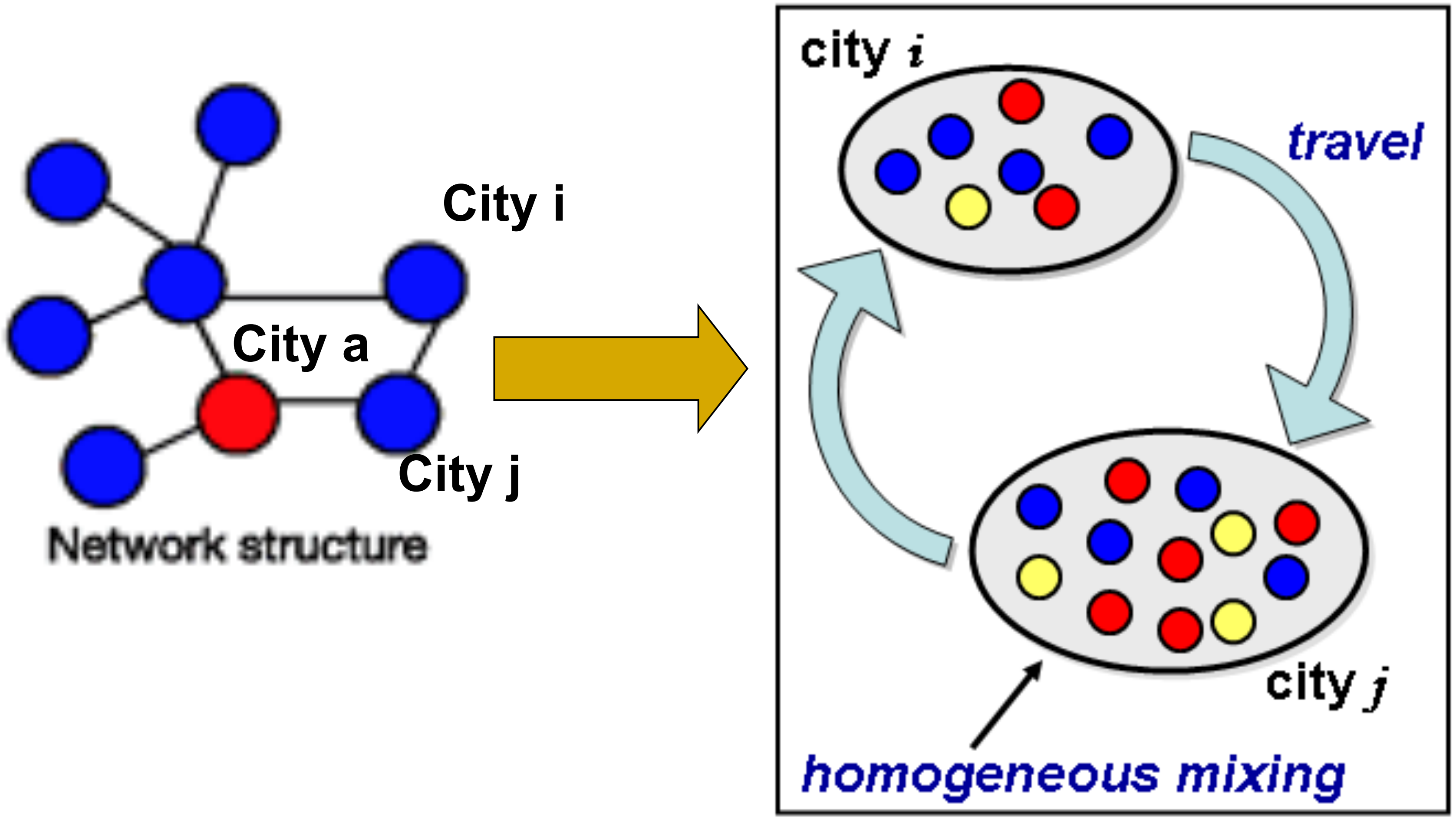
- Nodes = airports / geographical areas
- Edges = direct connections
- Weighted edges, weights = #seats / (time scale)

=> Metapopulation network

=> Used in large-scale models of epidemic propagation

Meta-population networks

Each node: internal structure
Links: transport/traffic

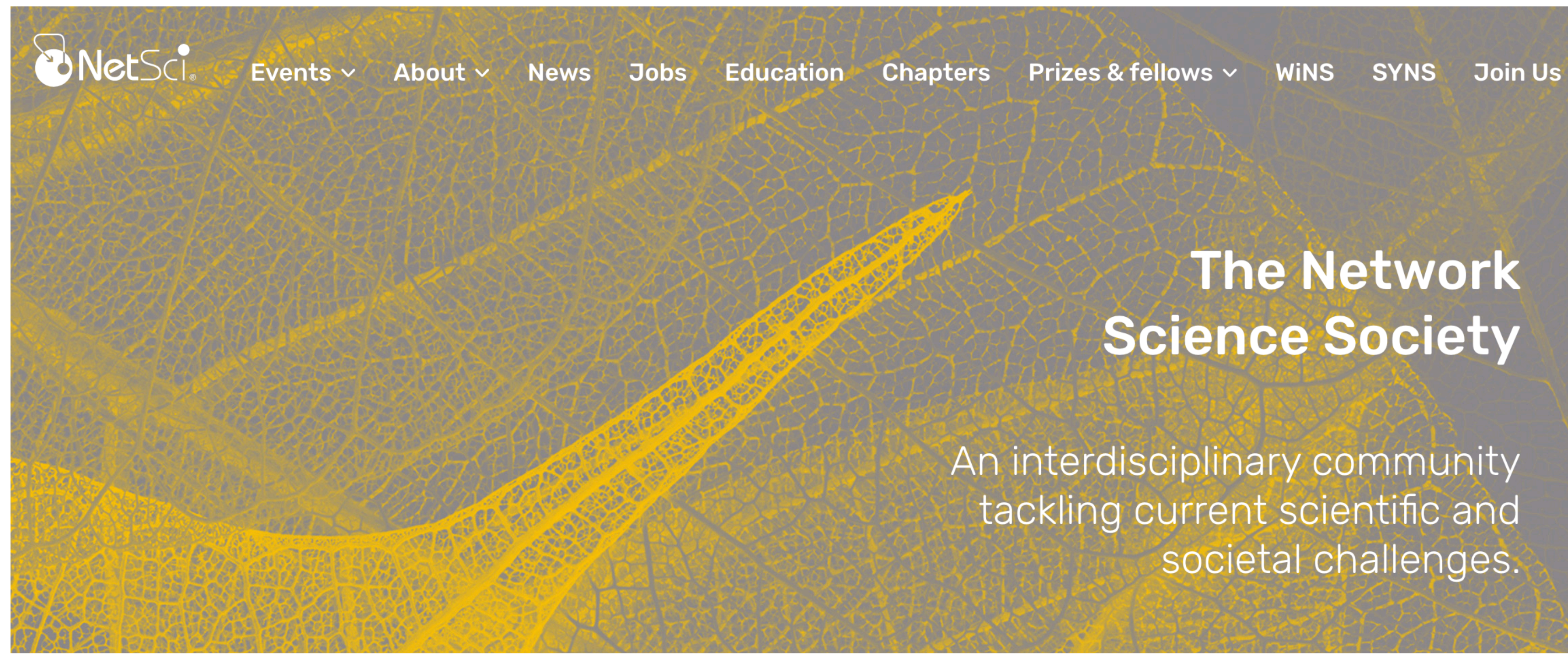


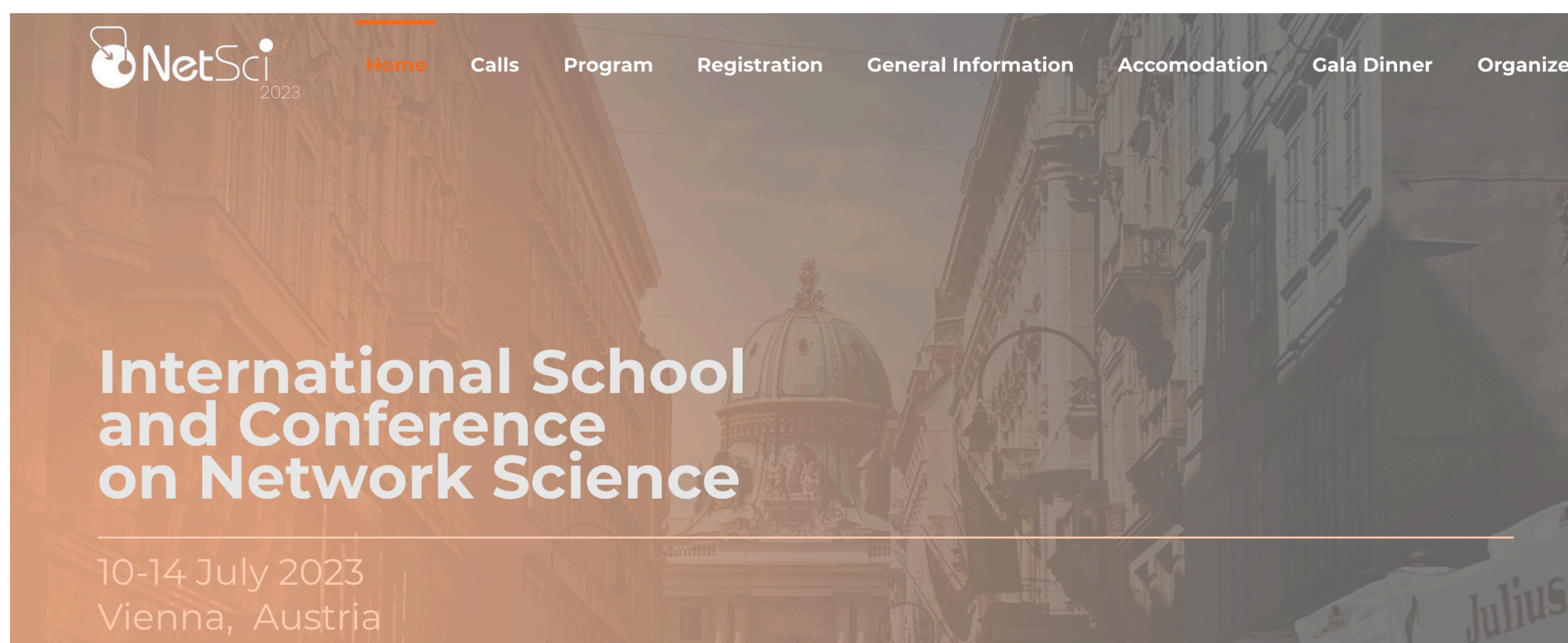
Interdisciplinary science

Science of complex networks ("Network science"):

- graph theory
- social sciences
- biology, neuroscience
- epidemiology
- physics
- computer science
- geography
- economy
- …

http://netscisociety.net





Interdisciplinary science

Science of complex networks:

- Empirics
- Characterization
- Modeling
- Dynamical processes
- Extensions (multiplexes, hyper graphs, temporal networks…)
- …

Data-driven
Tools both from graph theory and outside graph theory
Both general tools and domain specific tools

Networks characteristics

Networks: of very different origins

Do they have anything in common?
Possibility to find common properties?

the abstract character of the graph representation and graph theory allow to answer….

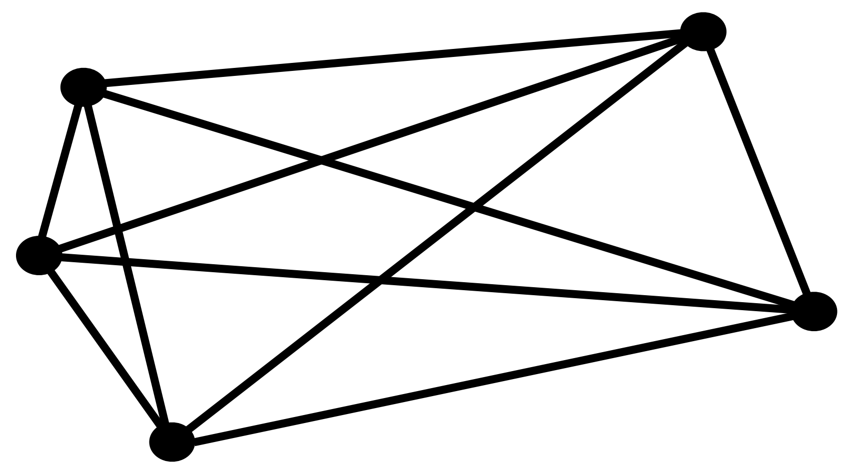
Graph theory: basics

$G=(V,E)$; $|V|=N$

Maximum number of edges

- Undirected: N(N-1)/2
- Directed: N(N-1)

Complete graph:



(all to all interaction/communication)

How to represent a graph

- List of nodes + list of edges

  i,j

- List of nodes + list of neighbors of each node (adjacency lists)

  1: 2,3,10,...

  2: 1,12,11

  3: 1,...

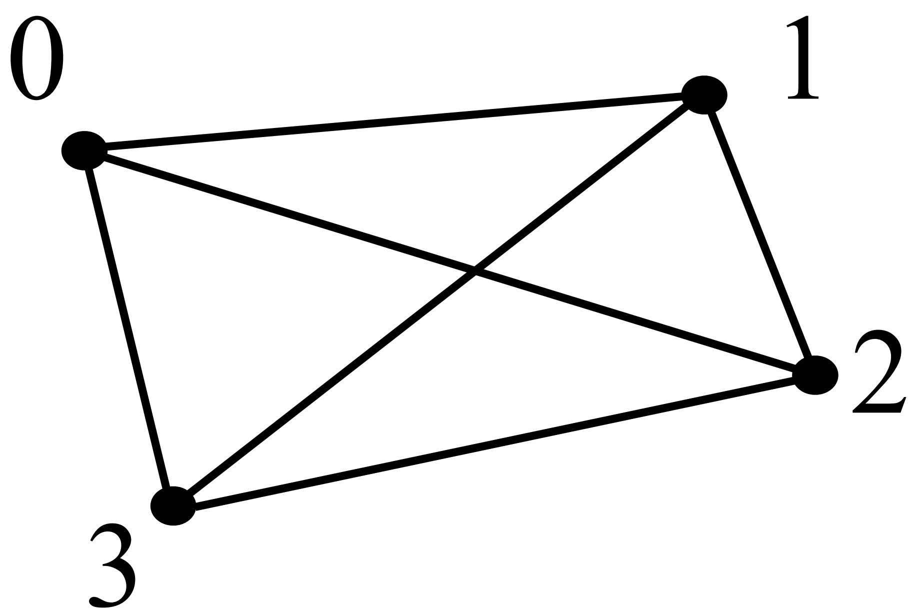
- Adjacency matrix

Adjacency matrix

N nodes i=1,…,N

$$a_{ij} = \begin{cases} 1 \text{ if } (i,j) \in E \\ 0 \text{ if } (i,j) \notin E \end{cases}$$

	0	1	2	3
0	0	1	1	1
1	1	0	1	1
2	1	1	0	1
3	1	1	1	0



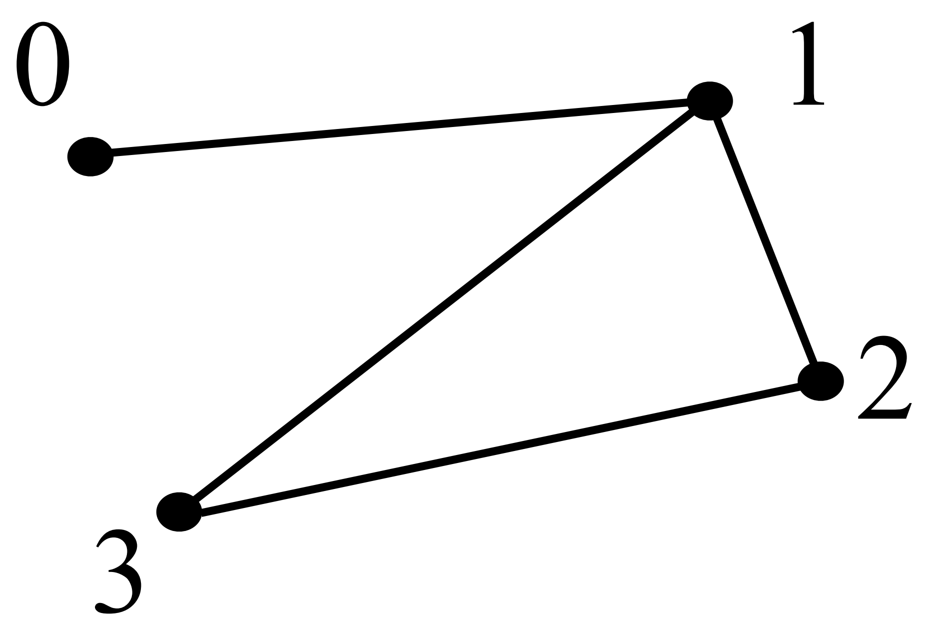
Adjacency matrix

N nodes i=1,….,N

$$a_{ij}= \begin{cases} 1 \text{ if } (i,j) \in E \\ 0 \text{ if } (i,j) \notin E \end{cases}$$

Symmetric
for undirected networks

	0	1	2	3
0	0	1	0	0
1	1	0	1	1
2	0	1	0	1
3	0	1	1	0



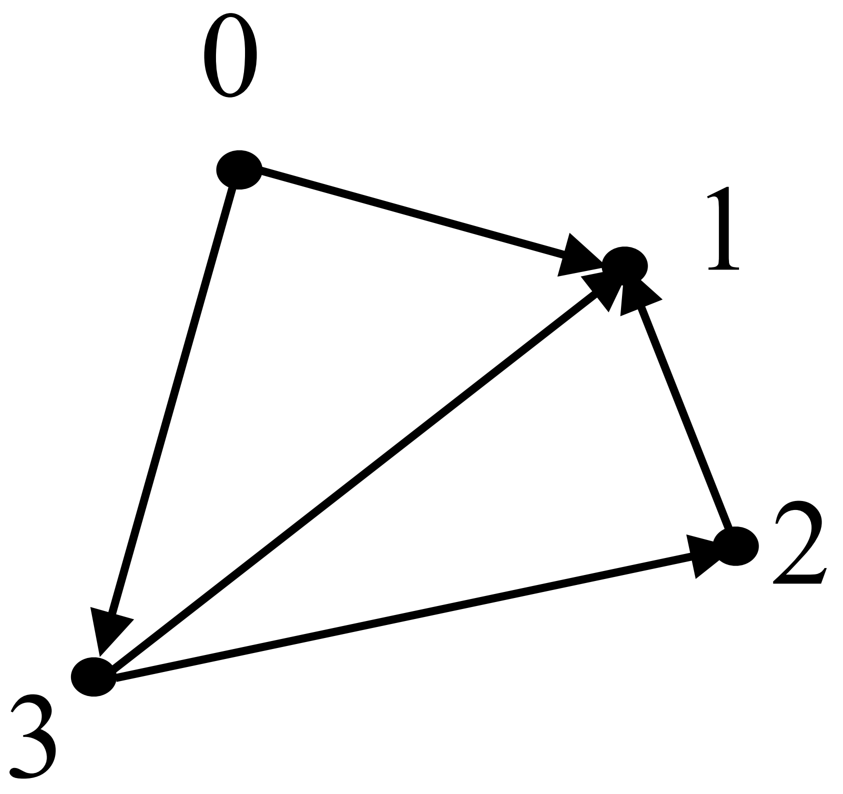
Adjacency matrix

N nodes i=1,….,N

$$a_{ij} = \begin{cases} 1 \text{ if } (i,j) \in E \\ 0 \text{ if } (i,j) \notin E \end{cases}$$

	0	1	2	3
0	0	1	0	1
1	0	0	0	0
2	0	1	0	0
3	0	1	1	0

Non symmetric
for directed networks



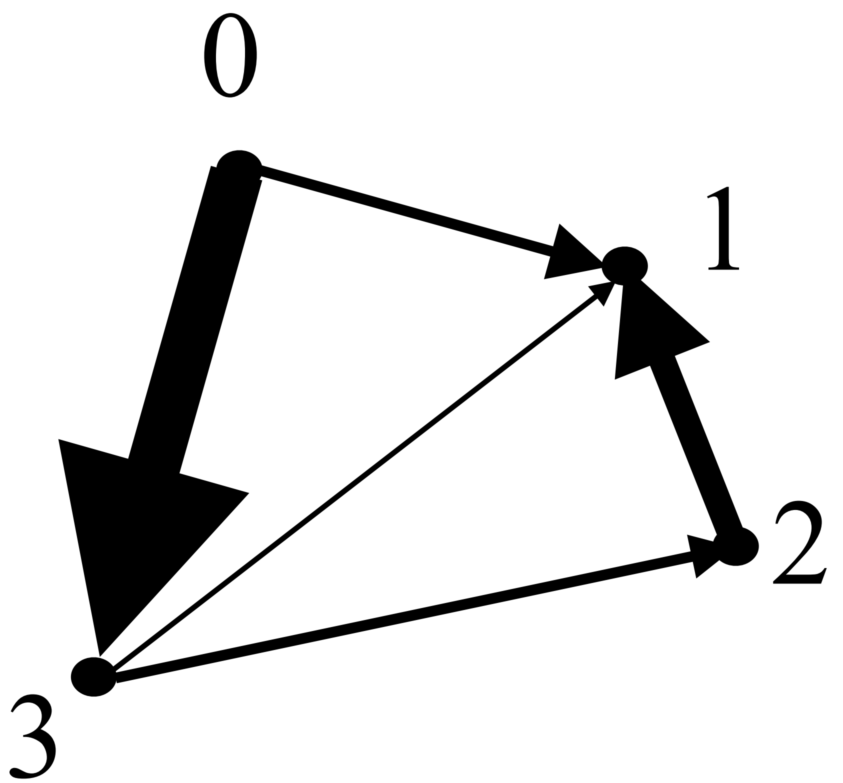
Matrix of weights

N nodes i=1,….,N

$$w_{ij} = \begin{cases} \neq 0 \text{ if } (i,j) \in E \\ 0 \text{ if } (i,j) \notin E \end{cases}$$

(Non symmetric for directed networks)

	0	1	2	3
0	0	2	0	10
1	0	0	0	0
2	0	5	0	0
3	0	1	2	0



Sparse graphs

Density of a graph D=|E|/(N(N-1)/2)

$$D = \frac{\text{Number of edges}}{\text{Maximal number of edges}}$$

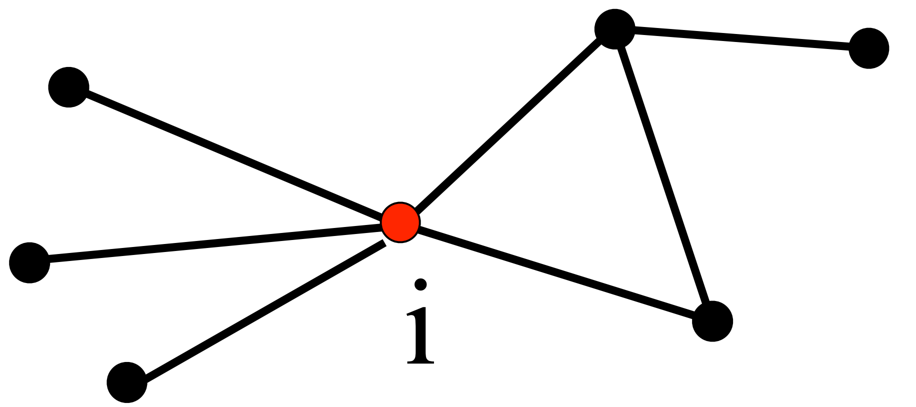
Sparse graph: D <<1 → Sparse adjacency matrix

→ Representation by lists of neighbours of each node (adjacency lists) better suited

Node characteristics: Degrees and strengths

Node characteristics

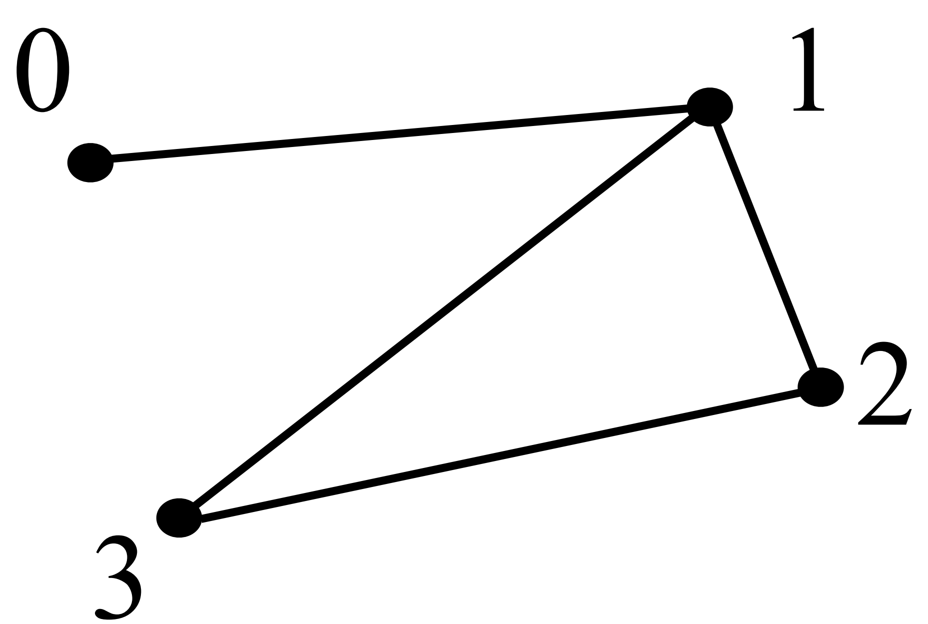
- Degree=number of neighbours=$\sum_j a_{ij}$



$k_i=5$

NB: in a sparse graph we expect $k_i << N$

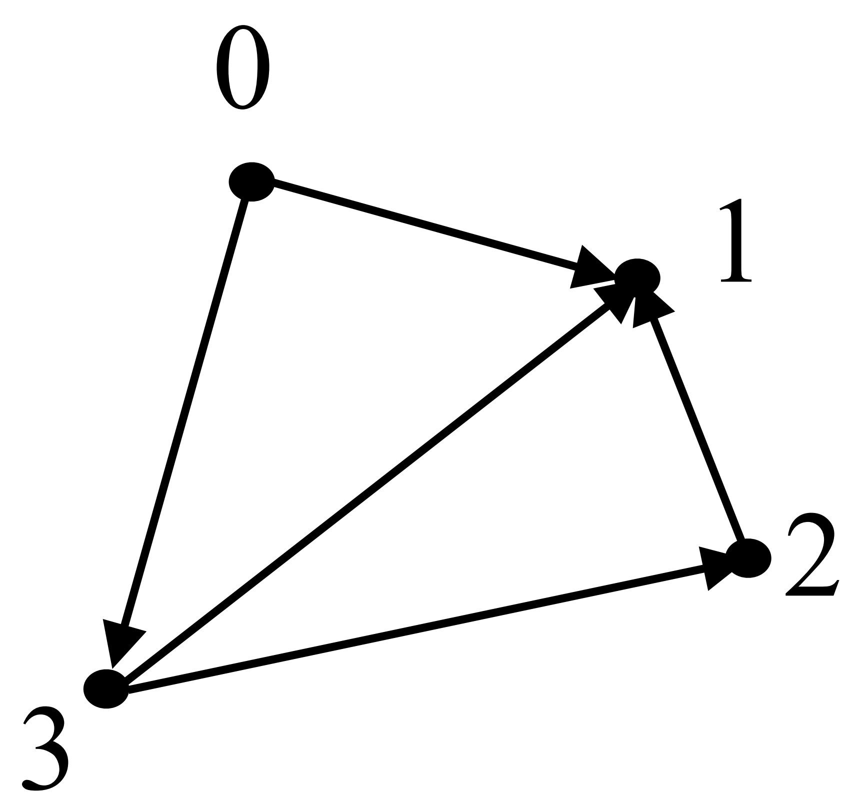
	0	1	2	3
0	0	1	0	0
1	1	0	1	1
i 2	0	1	0	1
3	0	1	1	0



Node characteristics

- Degree in directed graphs:
  - in-degree= number of in-neighbours= $\sum_j a_{ji}$
  - out-degree= number of out-neighbours= $\sum_j a_{ij}$

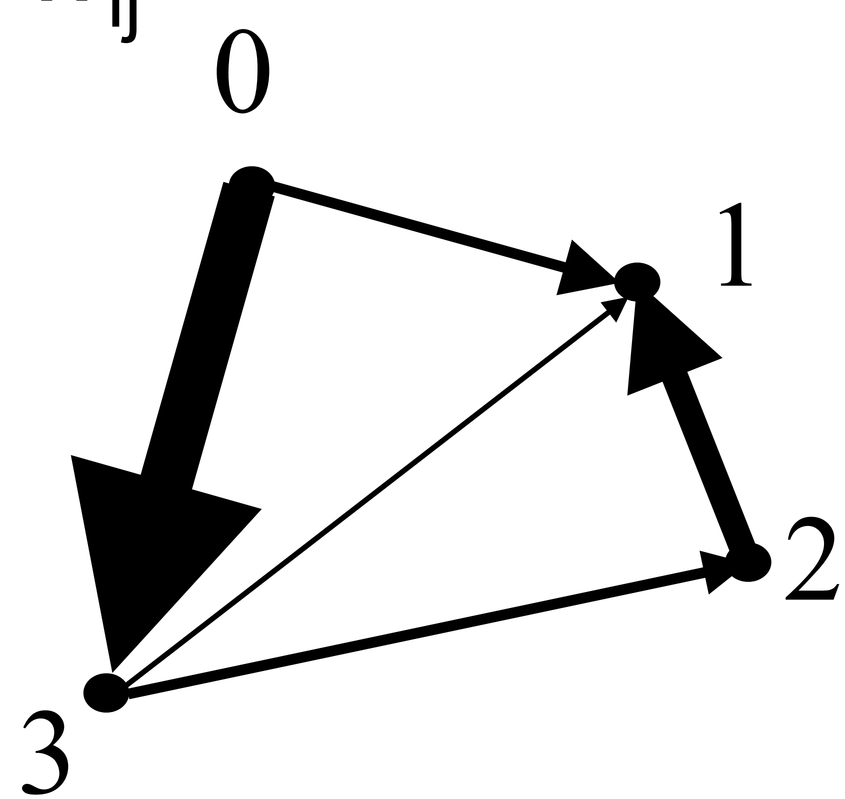
	0	1	2	3
0	0	1	0	1
1	0	0	0	0
2	0	1	0	0
3	0	1	1	0



Node characteristics

- Weighted graphs: Strength $s_i = \sum_j w_{ij}$
- Directed Weighted graphs:
  - –in-strength $s_i = \sum_j w_{ji}$
  - –out-strength $s_i = \sum_j w_{ij}$

	0	1	2	3
0	0	2	0	10
1	0	0	0	0
2	0	5	0	0
3	0	1	2	0



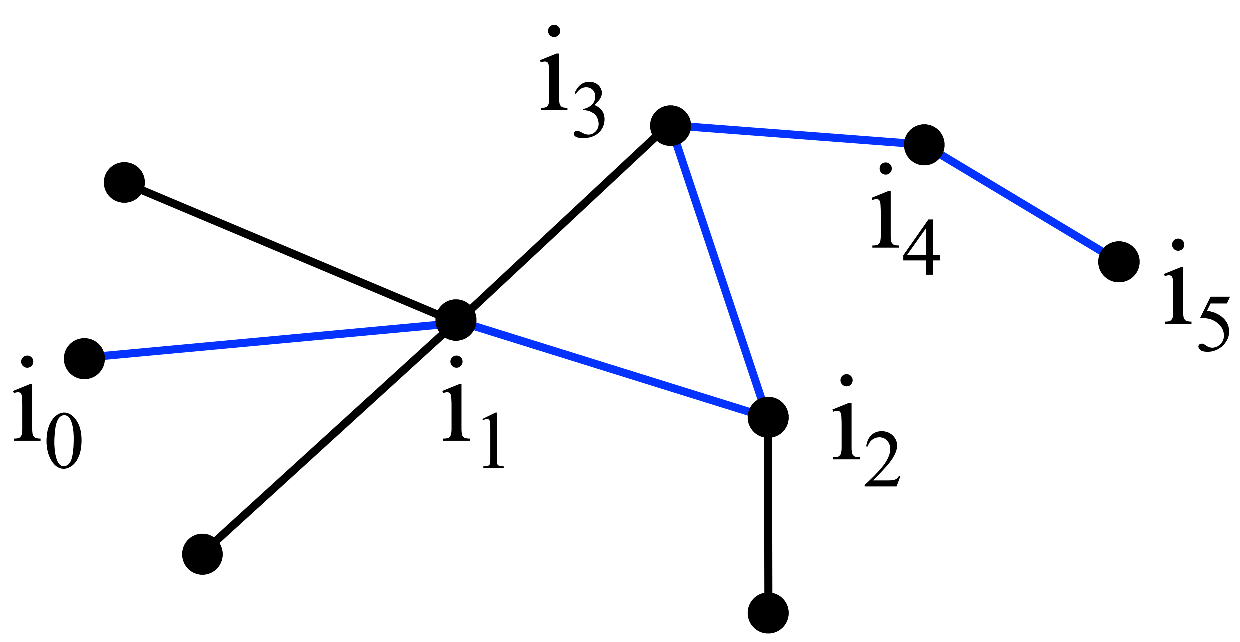
Paths, connectedness, small-world effect

Paths

$G=(V,E)$

Path of length n = ordered collection of

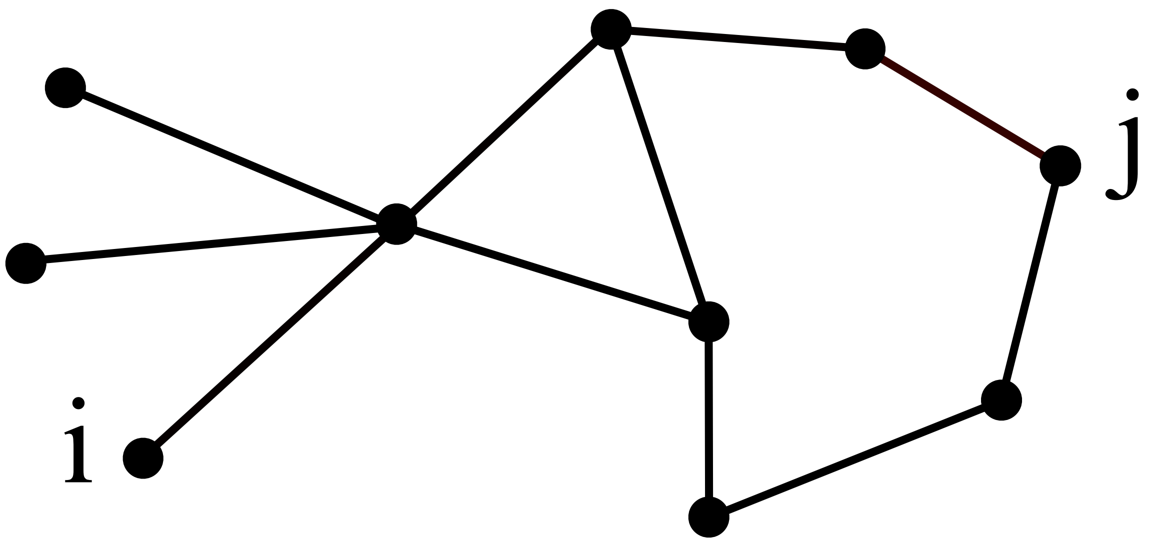
- n+1 vertices $i_0, i_1, \ldots, i_n \in V$
- n edges $(i_0,i_1), (i_1,i_2)\ldots,(i_{n-1},i_n) \in E$



Cycle/loop = closed path $(i_0=i_n)$

Tree=graph with no loops

Paths and adjacency matrix

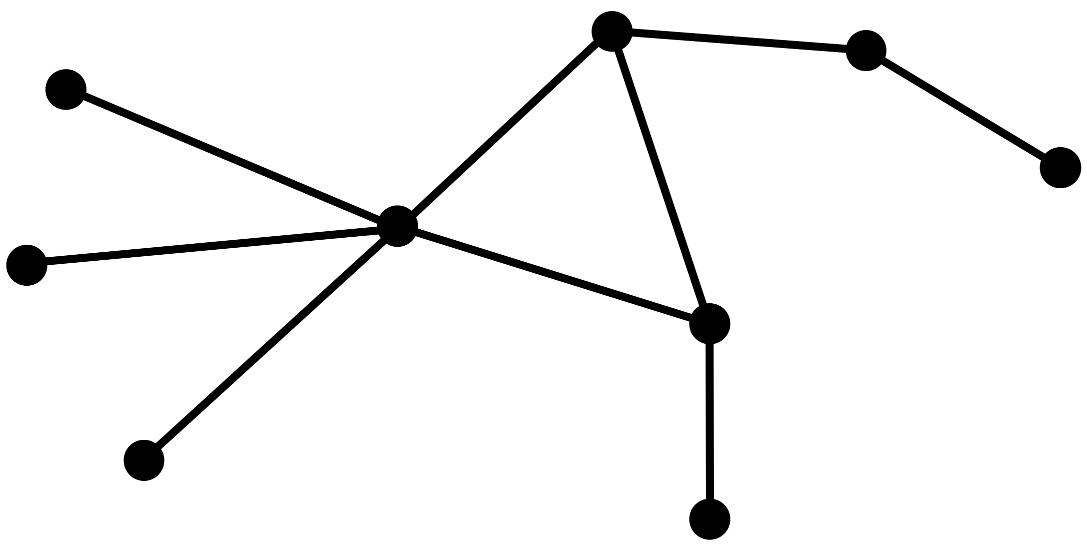


$(a^n)_{ij}$ =number of paths of length n between i and j

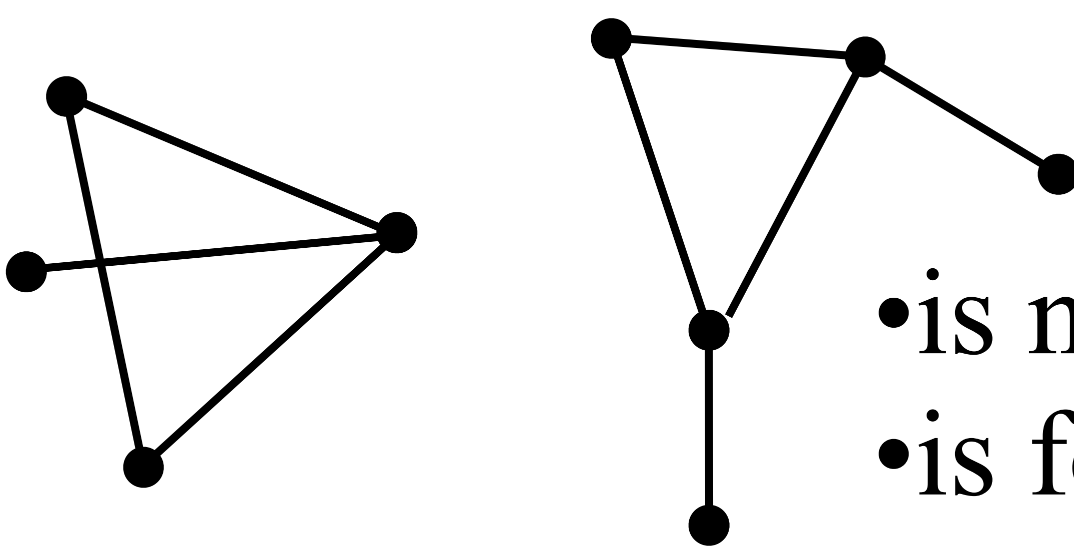
Example: $(a^2)_{ij} = \sum a_{ik}\, a_{kj}$

Paths and connectedness

G=(V,E) is connected if and only if there exists a path connecting any two nodes in G



is connected



- is not connected
- is formed by two components

Paths and connectedness

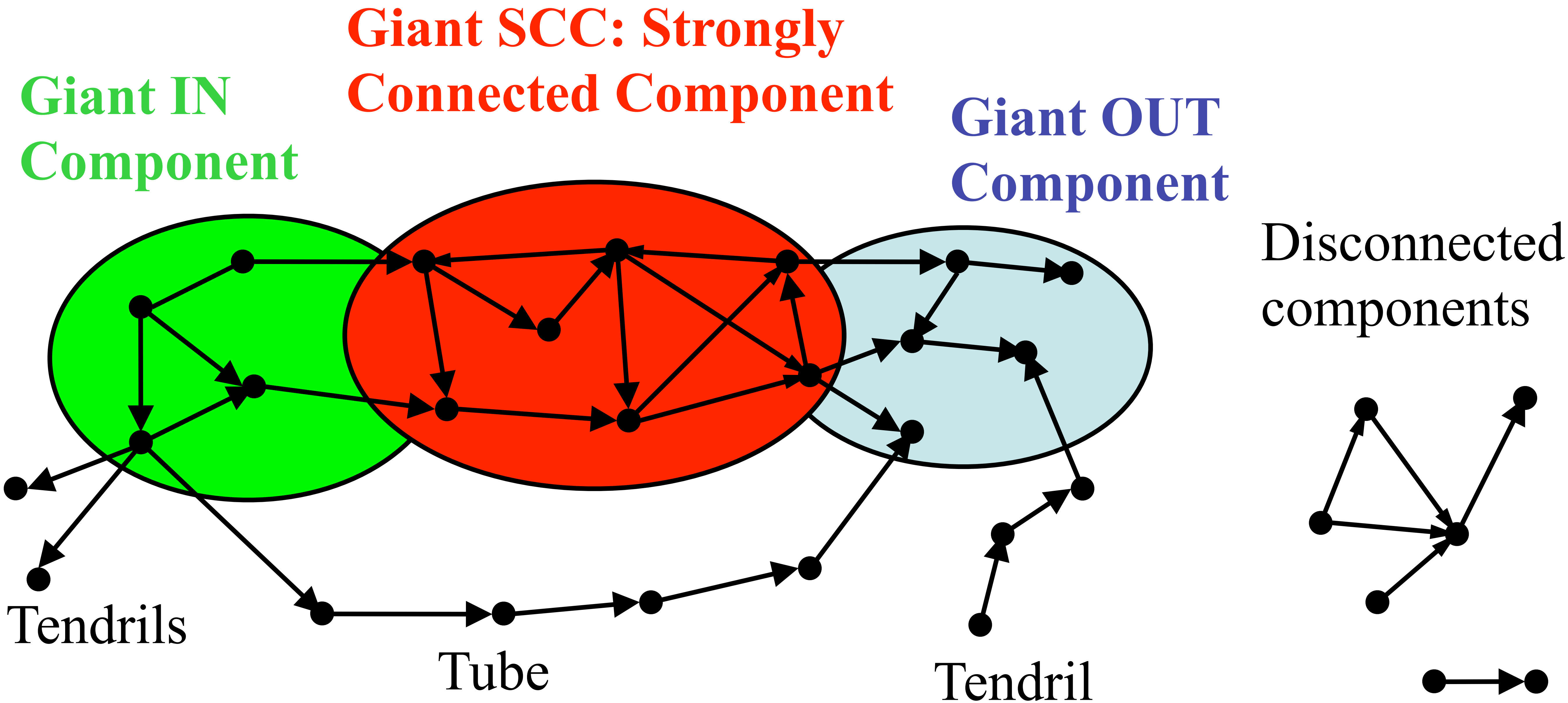
G=(V,E)=> distribution of components' sizes

Giant component= component whose size scales with the number of vertices N

Existence of a giant component ⟺ Macroscopic fraction of the graph is connected

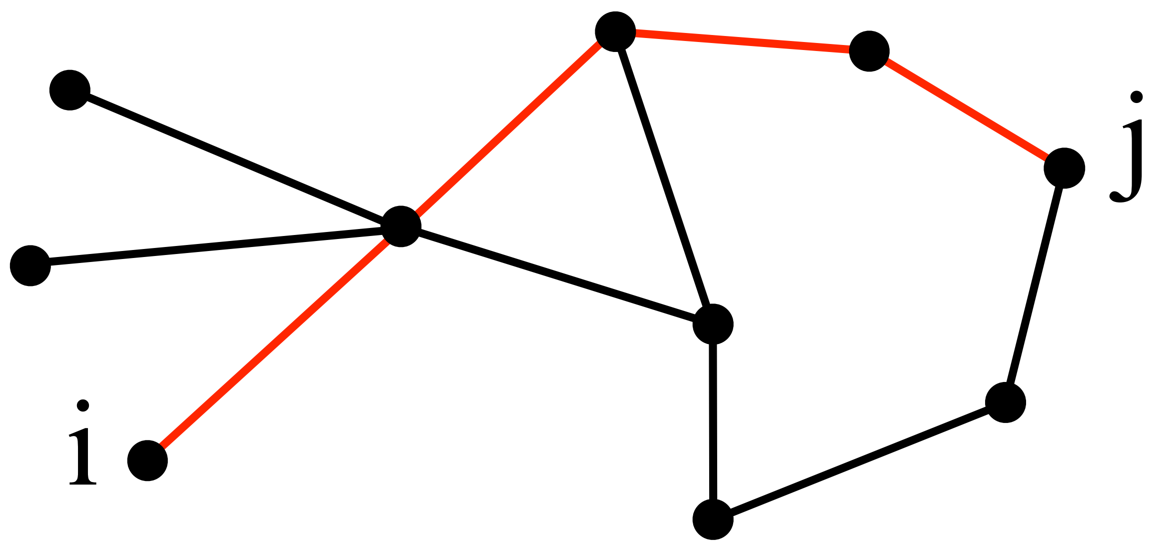
Paths and connectedness: directed graphs

Paths are ***directed***



Shortest paths

Shortest path between i and j: minimum number of traversed edges



distance $l(i,j)$=minimum number of edges traversed on a path between i and j

Diameter of the graph= $\max(l(i,j))$

Average shortest path= $\sum_{ij} l(i,j)/(N(N-1)/2)$

Complete graph: $l(i,j)=1$ for all i,j

"Small-world": "small" diameter

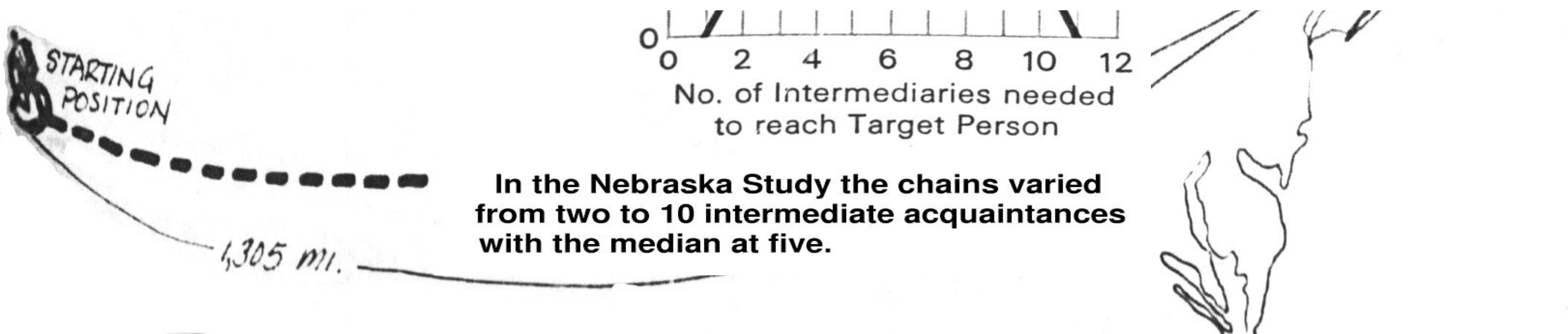
Social networks: Milgram's experiment

Milgram, Psych Today **2**, 60 (1967)

Dodds et al., Science **301**, 827 (2003)

"Six degrees of separation"

SMALL-WORLD CHARACTER



Social networks as small-worlds: Milgram's experiment, revisited

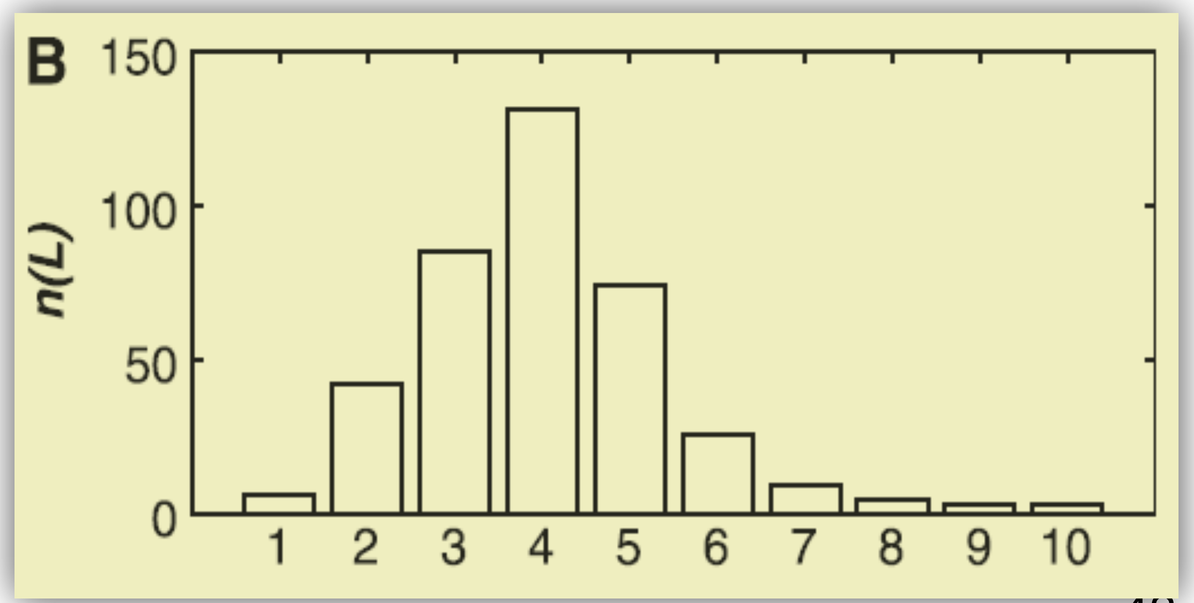
Dodds et al., Science **301**, 827 (2003)

email chains

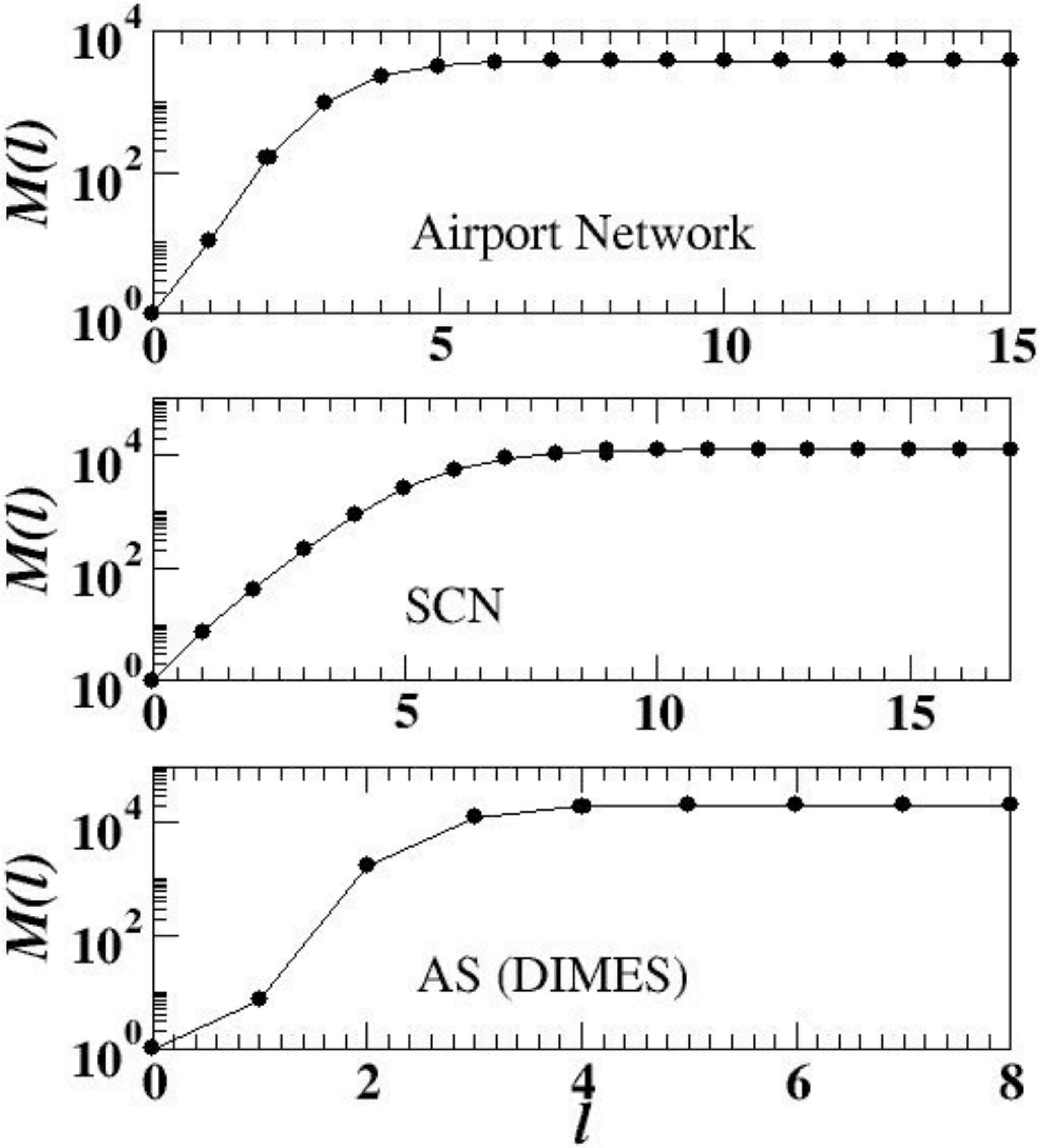
60000 start nodes

18 targets

384 completed chains



Small-world properties

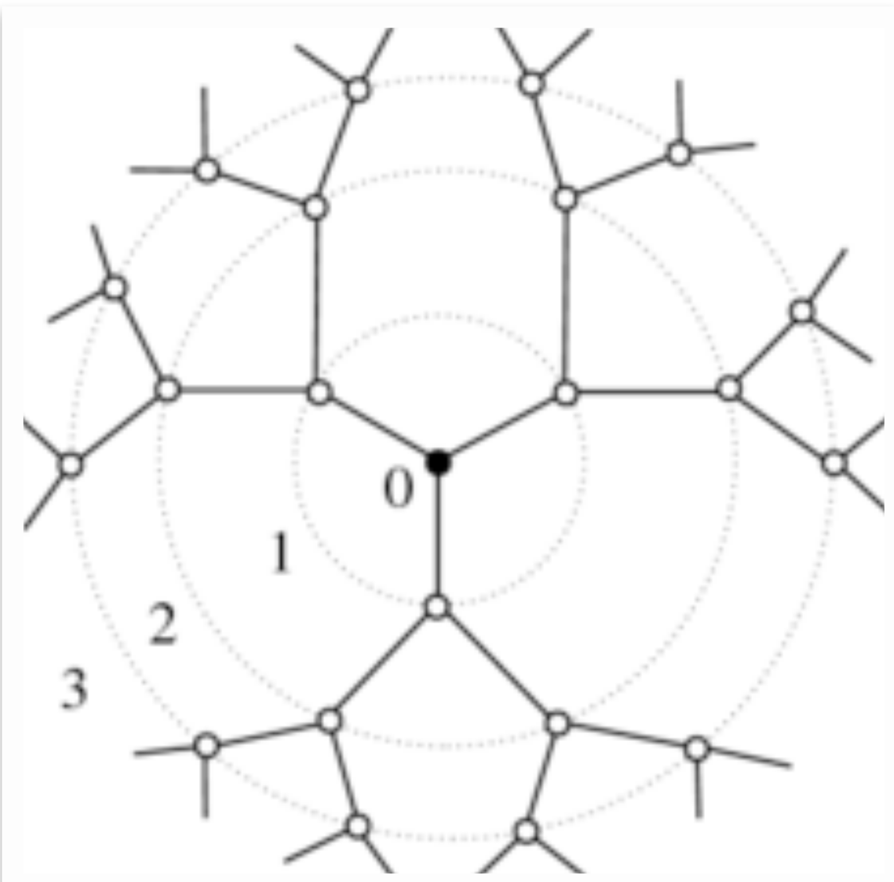


Average number of nodes within a chemical distance l

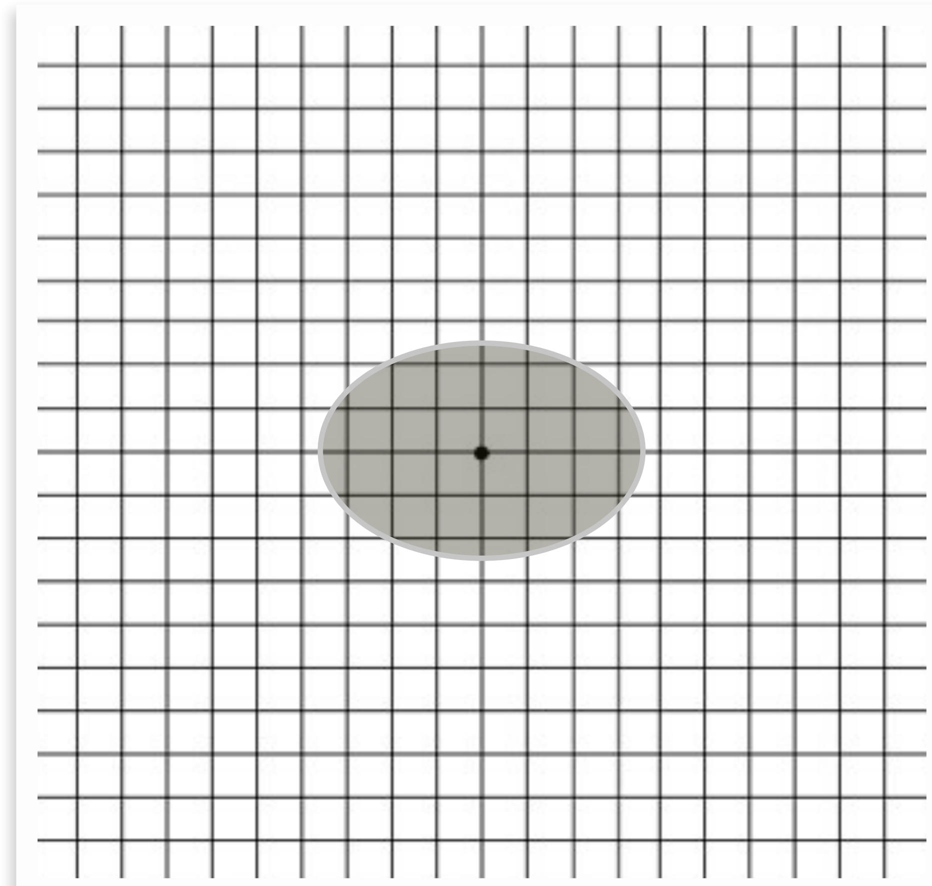
Scientific collaborations

Internet

The intuition behind the small-world effect



versus

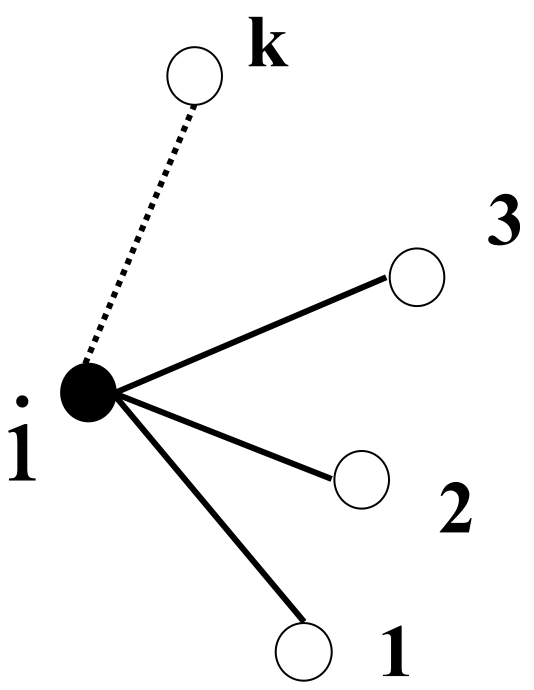


Tree:
number of reachable nodes grows very fast (exponentially) with the distance
=> distances logarithmic in size

(local) regular structure: slower growth of the number of reachable nodes (polynomial), because of path redundancy

Small-world yet clustered/locally cohesive

Structure of neighborhoods

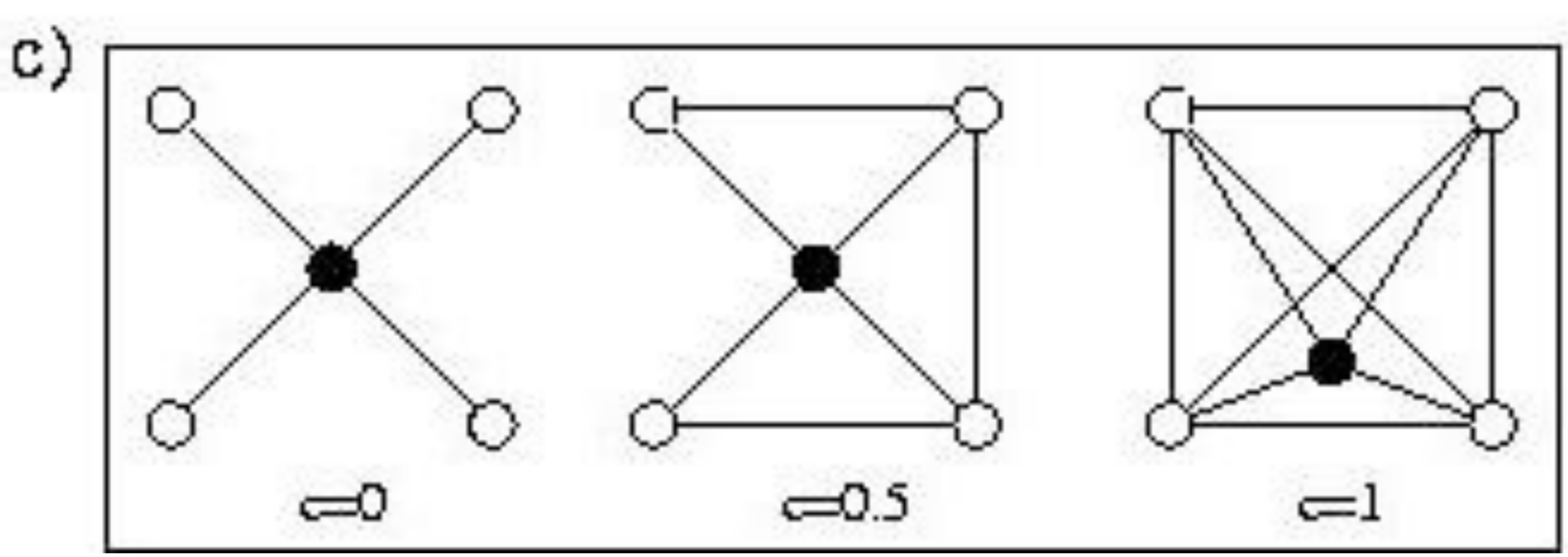


Clustering coefficient of a node

$$C(i) = \frac{\text{\# of links between 1,2,…n neighbors}}{k(k-1)/2}$$

$$C(i) = \frac{1}{k_i(k_i - 1)} \sum_{j \neq k} a_{ij} a_{jk} a_{ik}$$

Clustering in social networks: My friends will know each other with high probability!



Structure of neighborhoods

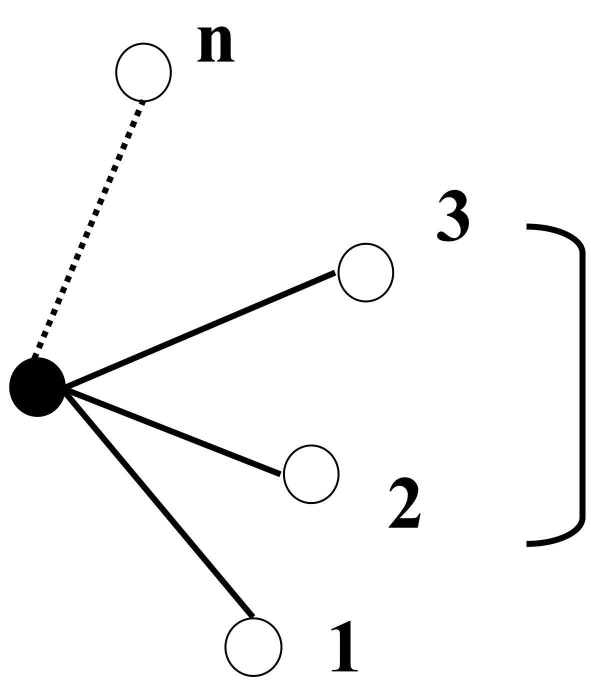
Average clustering coefficient of a graph

$C=\sum_i C(i)/N$

NB: slightly different definition from the fraction of transitive triples:

$$C' = \frac{3 \text{ x number of fully connected triples}}{\text{number of triples}}$$

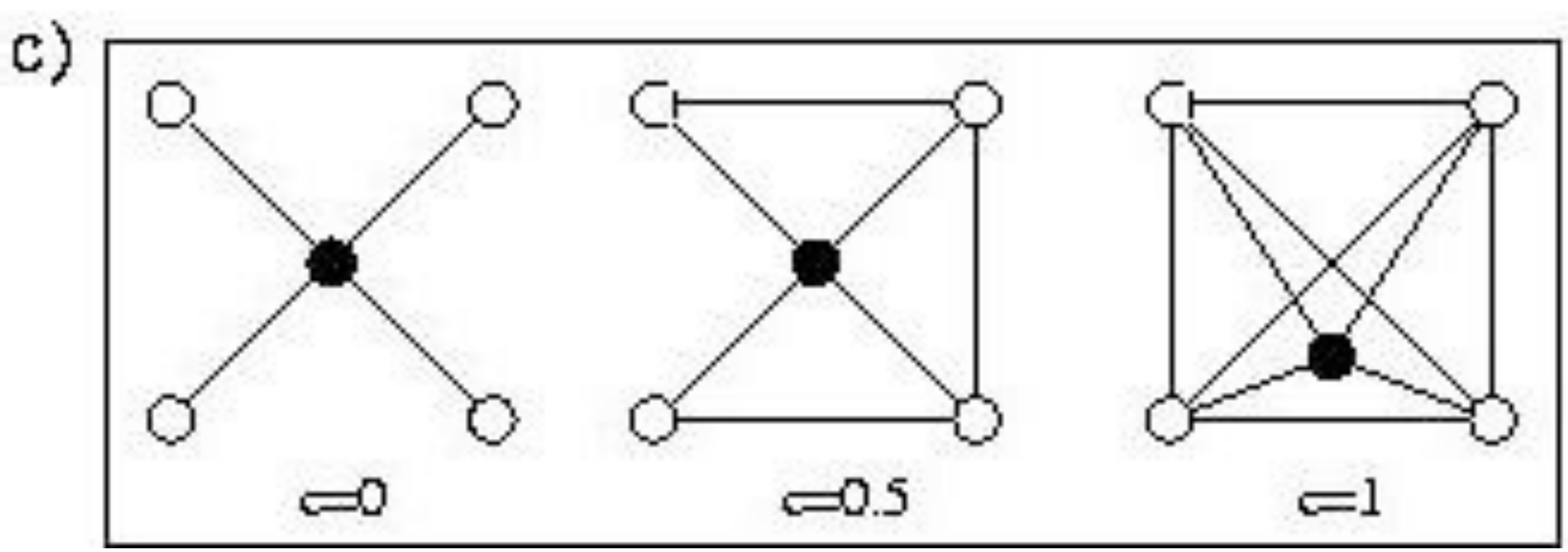
Clustering coefficient



Empirically: large clustering coefficients

Higher probability to be connected

Clustering in social networks: My friends will know each other with high probability!



Empirical networks:

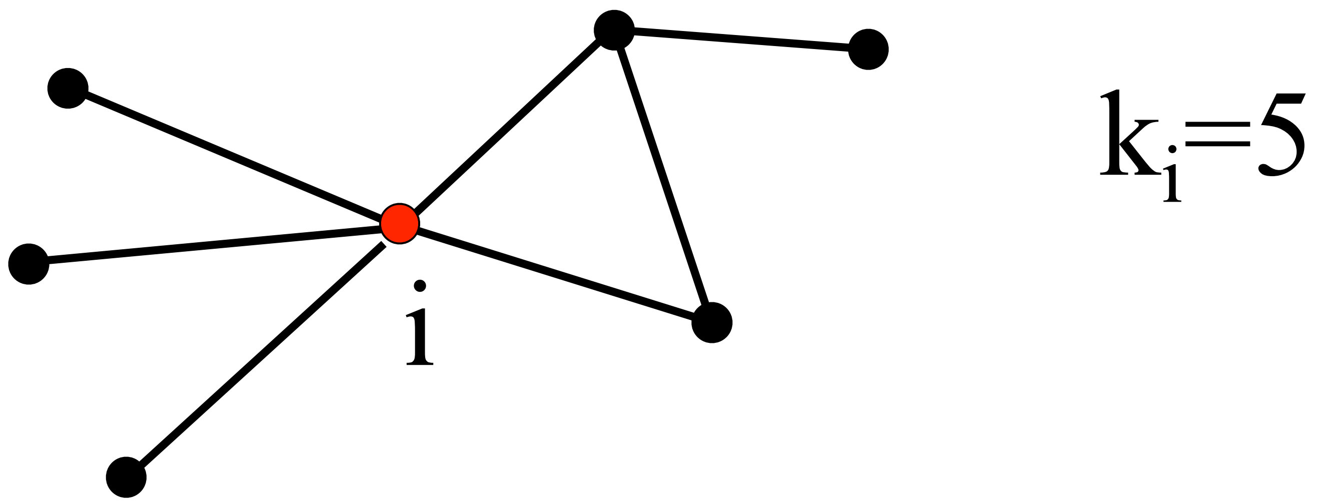
- small-world
- locally very "structured"

Ranking nodes

Centrality measures

How to quantify the importance of a node?

- Degree=number of neighbours=$\sum_j a_{ij}$



- Large degree nodes=”hubs”
- However: nodes with very large degree can be “peripheral”

Path-based centrality measures

- Closeness centrality

$$g_i = 1 / \sum_j l(i,j)$$

Quantifies the reachability of other nodes

Betweenness centrality

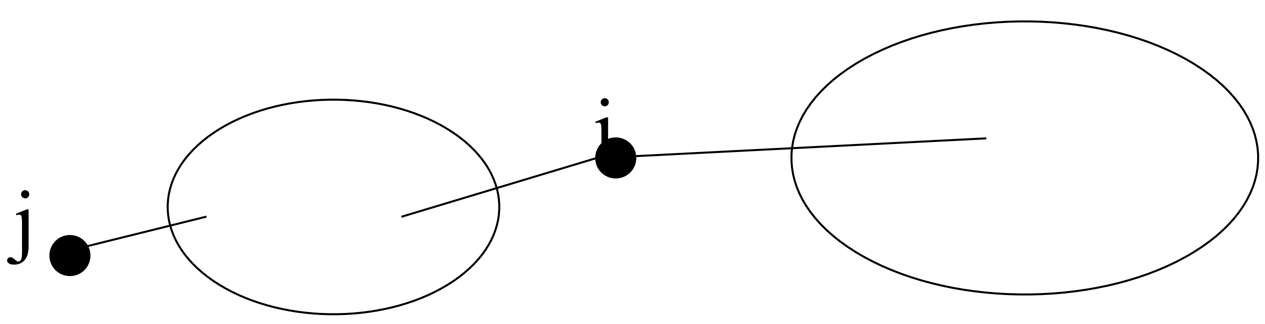
for each pair of nodes (l,m) in the graph, there are

σ^{lm} shortest paths between l and m

σ_i^{lm} shortest paths going through i

b_i is the sum of $\sigma_i^{lm} / \sigma^{lm}$ over all pairs (l,m)

path-based quantity



b_i is large (broker)

b_j is small

NB: generalization to *edge betweenness centrality*

Betweenness centrality

path-based quantity => bc(i) depends on all the nodes that are connected to i by at least one path

↓

non-local quantity

↓

"hard" to compute

"naive" algorithm: $O(N^3)$

Brandes algorithm: $O(N*E)$

Local approximations

Katz centrality

Rationale: a node is central if its neighbours are central

Centrality x_i of node *i*: depends on the centrality of its neighbours through

$$x_i = \alpha \sum_k a_{k,i} x_k + \beta$$

$$x = \beta(I - \alpha A)^{-1} = \beta \sum_{i=0}^{\infty} (\alpha A)^i$$

can be computed by matrix inversion or by iteration of the formula

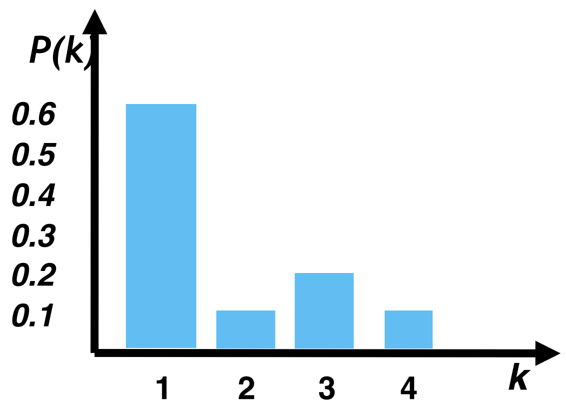
$\alpha < 1/\lambda_1$: damping factor
if too small: mainly short paths and close to degree

Statistical characterization of networks

Statistical characterization

Degree distribution

- List of degrees $k_1, k_2, \ldots, k_N$ ⟵ Not very useful!
- Histogram:
  N_k = number of nodes with degree k



- Distribution:
  $P(k)=N_k/N$ = probability that a randomly chosen node has degree k
- Cumulative distribution:
  $P_>(k)$ = probability that a randomly chosen node has degree at least k

Statistical characterization

Degree distribution

$P(k)=N_k/N=$probability that a randomly chosen node has degree k

Average$=\langle k \rangle = \sum_i k_i/N = \sum_k k\, P(k)=2|E|/N$

Sparse graphs: $\langle k \rangle \ll N$

Fluctuations: $\langle k^2 \rangle - \langle k \rangle^2$

$\langle k^2 \rangle = \sum_i k^2_i/N = \sum_k k^2\, P(k)$

$\langle k^n \rangle = \sum_k k^n\, P(k)$

Topological heterogeneity

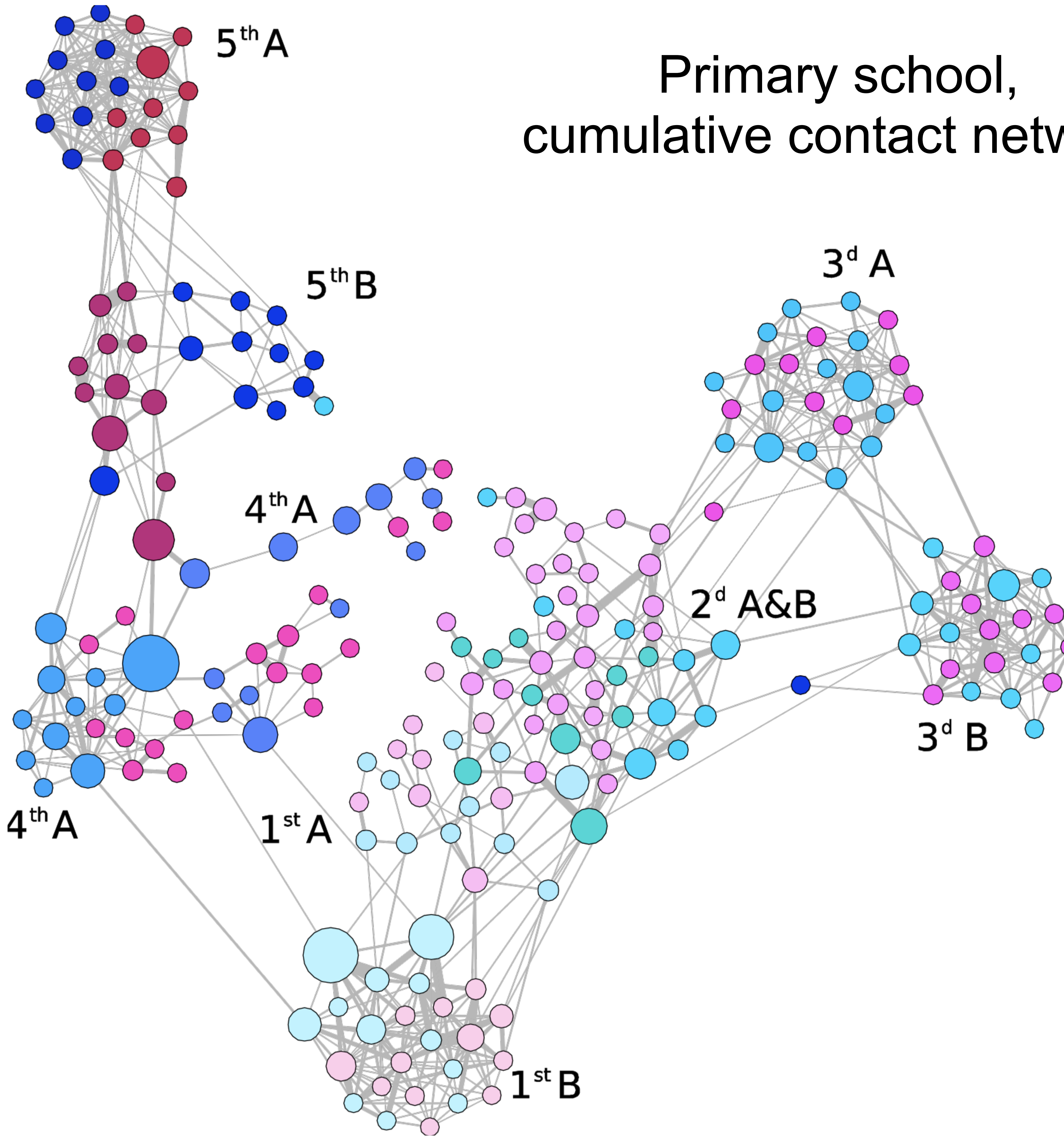
Statistical analysis of centrality measures:

$P(k)=N_k/N=$probability that a randomly chosen node has degree k

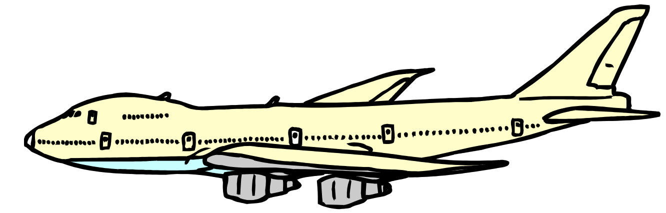
Two broad classes

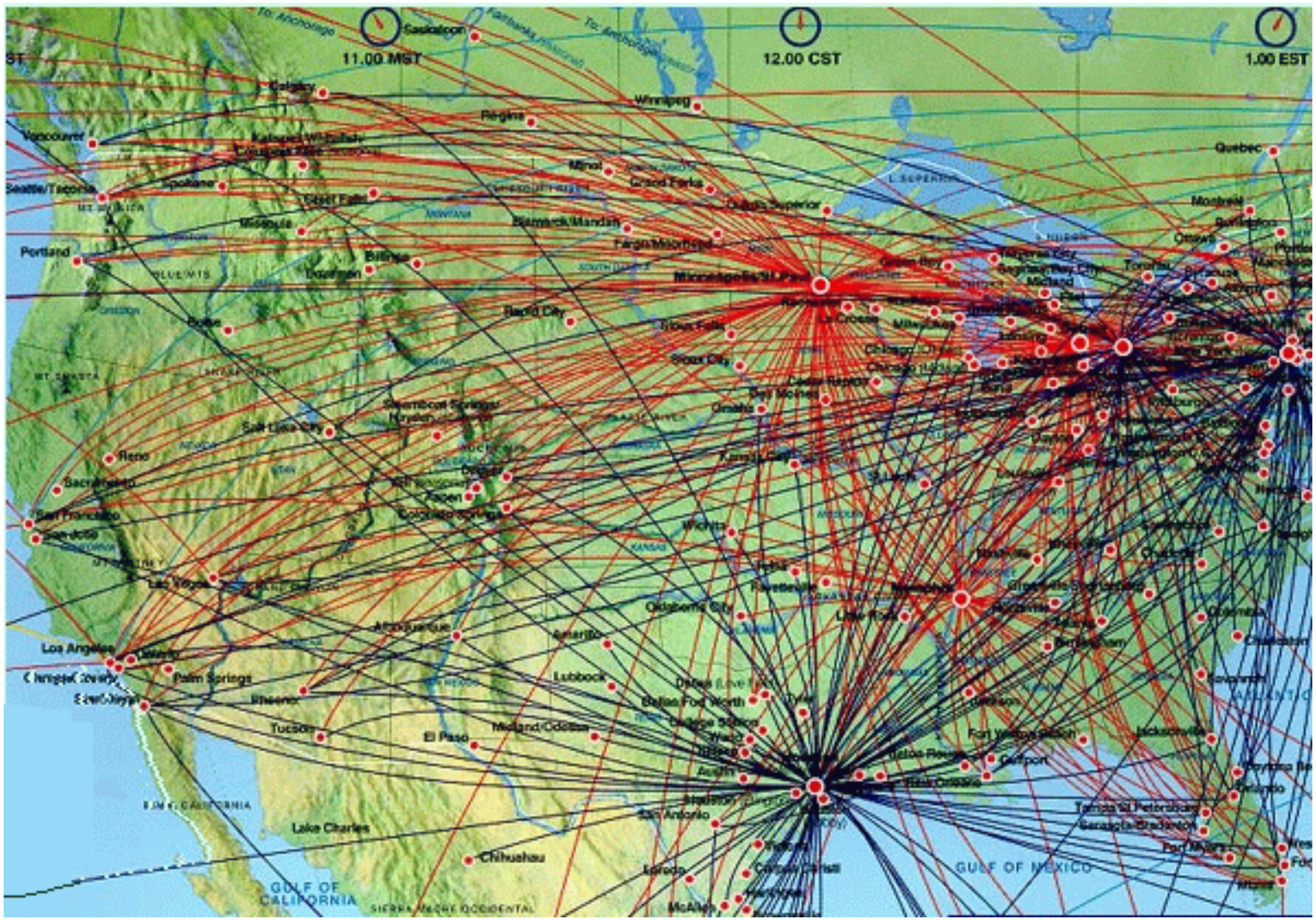
- homogeneous networks: light tails
- heterogeneous networks: skewed, heavy tails

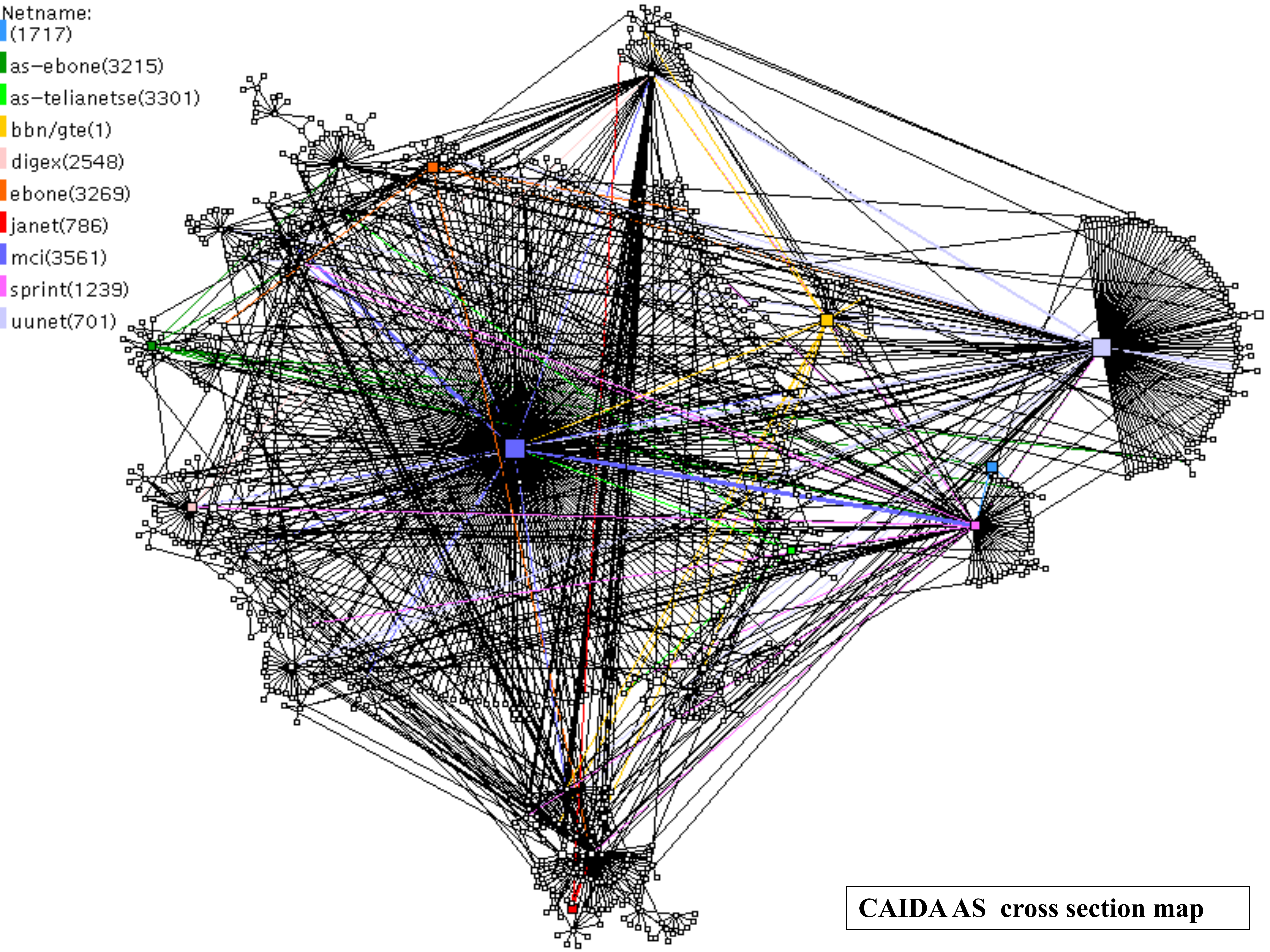
Primary school, cumulative contact network



Airplane route network



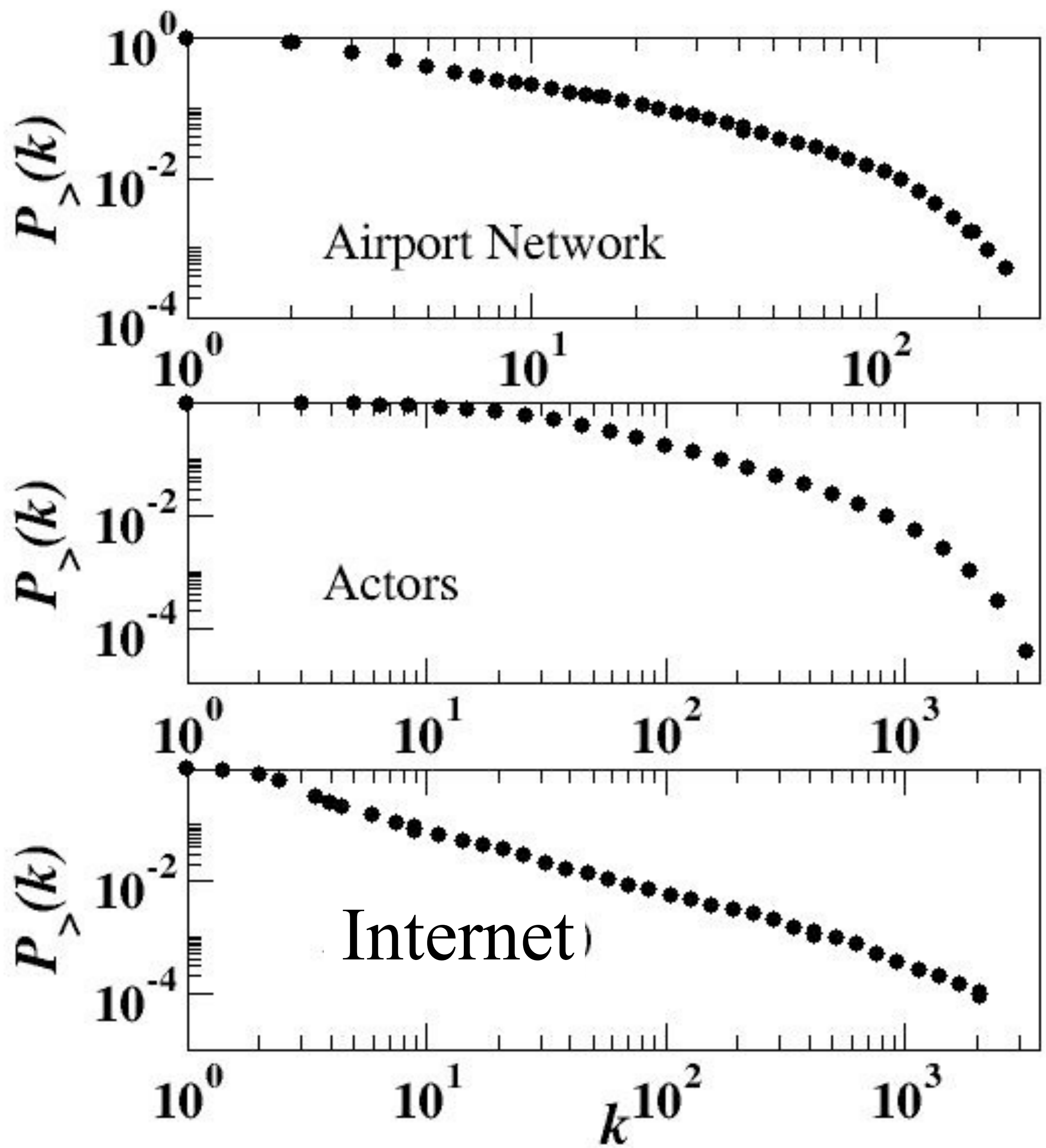




CAIDA AS cross section map

Topological heterogeneity

Statistical analysis of centrality measures



Broad degree distributions

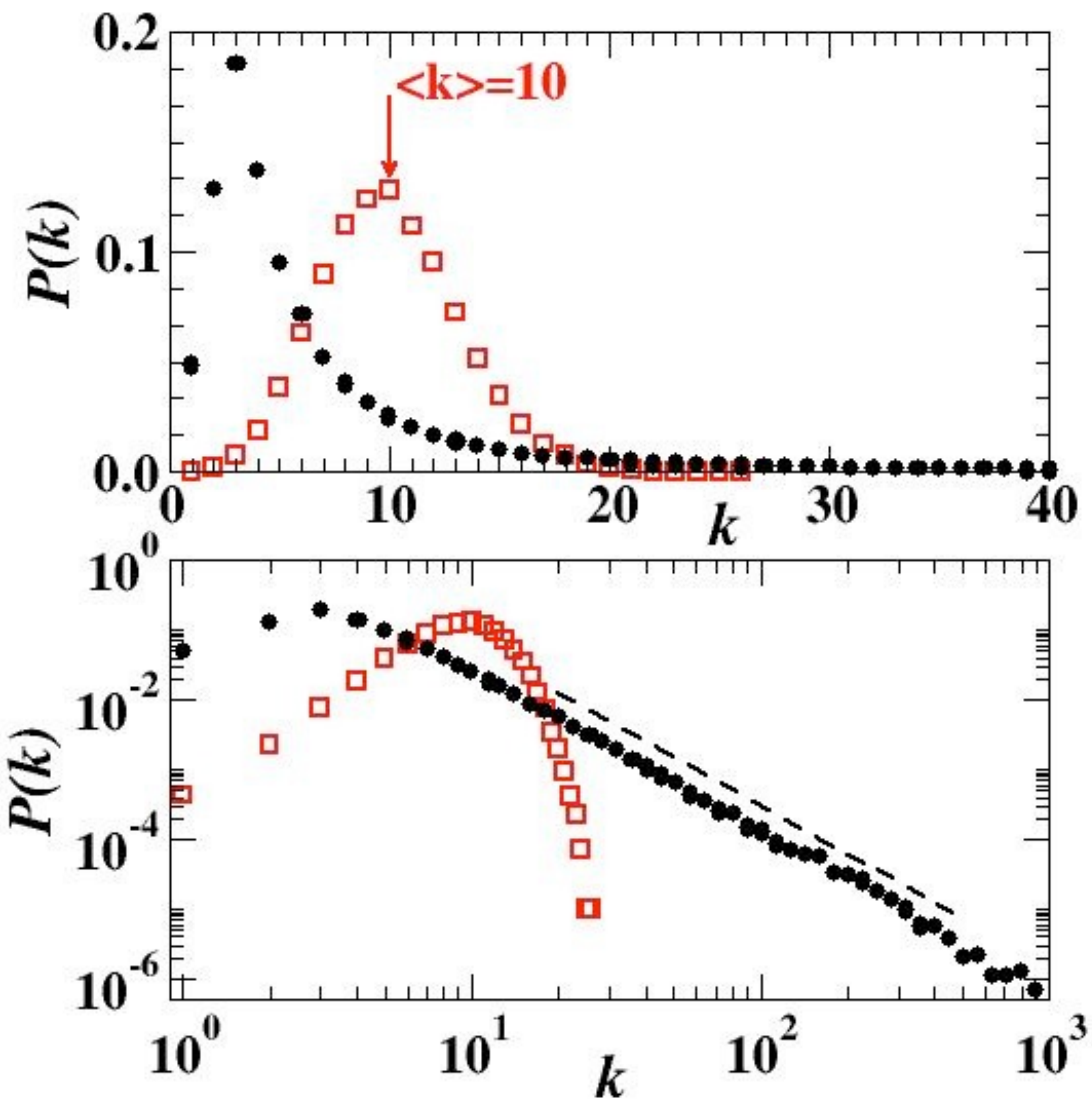
(often: power-law tails $P(k) \propto k^{-\gamma}$, typically $2 < \gamma < 3$)

No particular *characteristic* scale

Unbounded fluctuations

Topological heterogeneity

Statistical analysis of centrality measures:



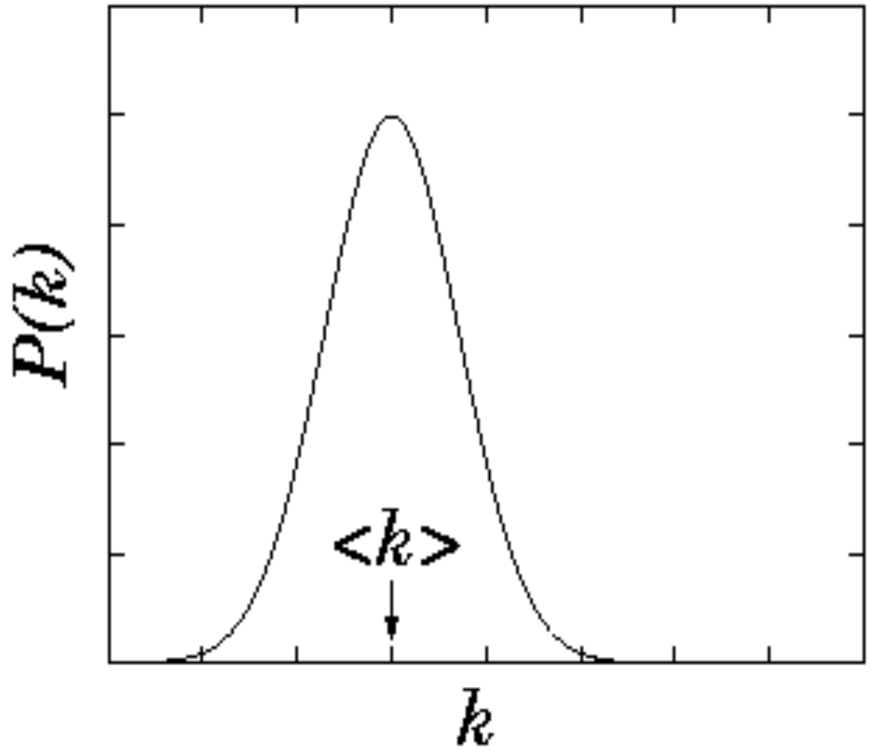
linear scale

Poisson
vs.
Power-law

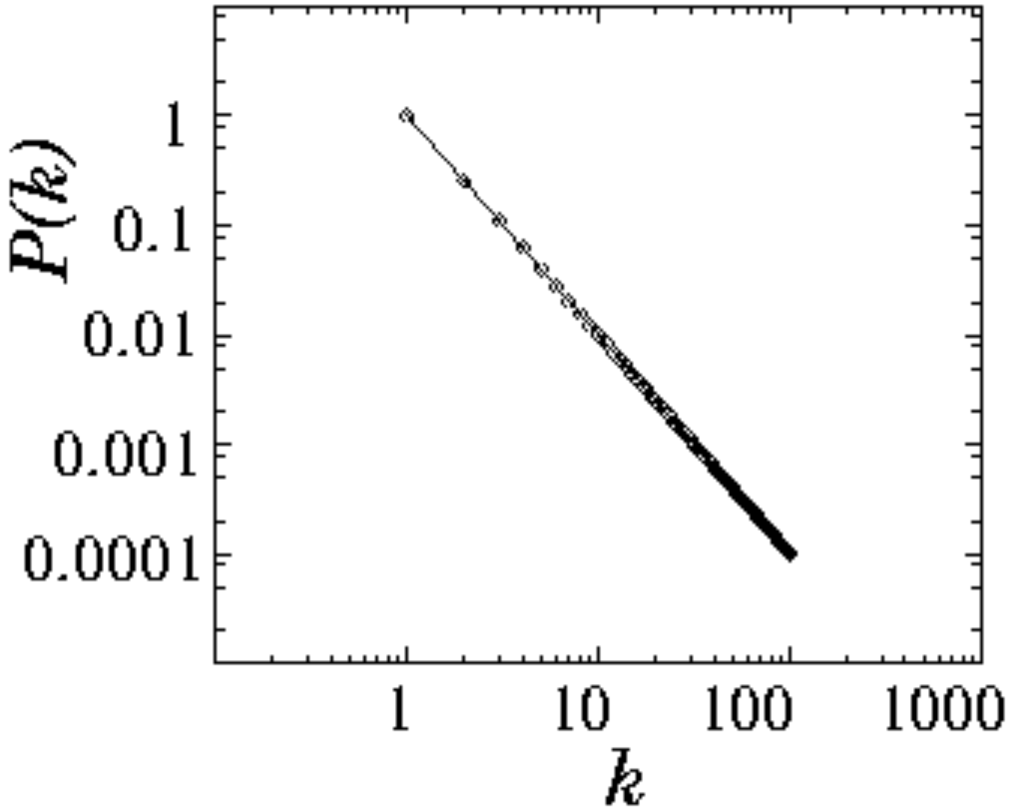
log-scale

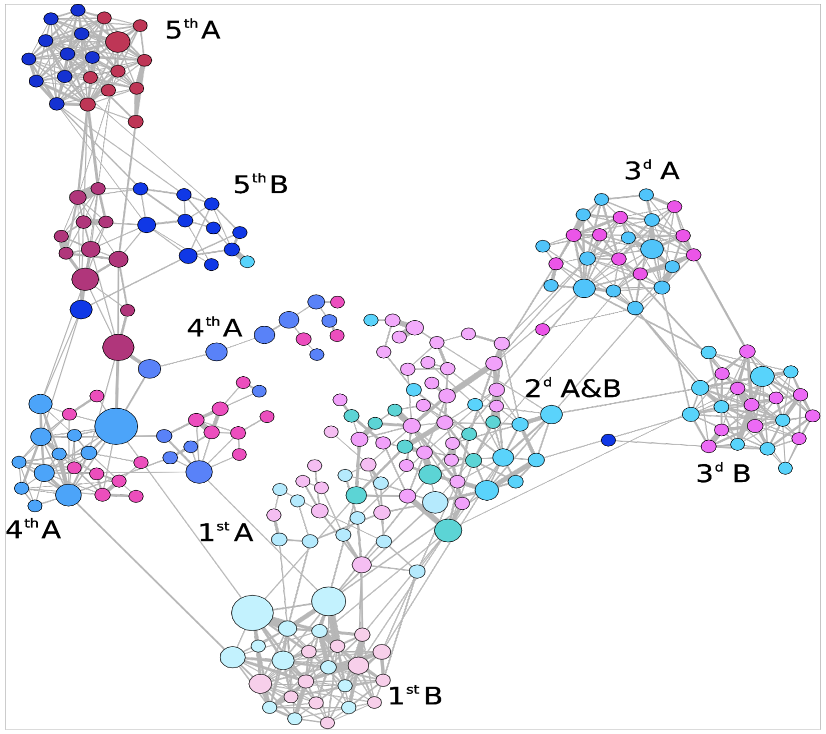
Exp. vs. Scale-Free

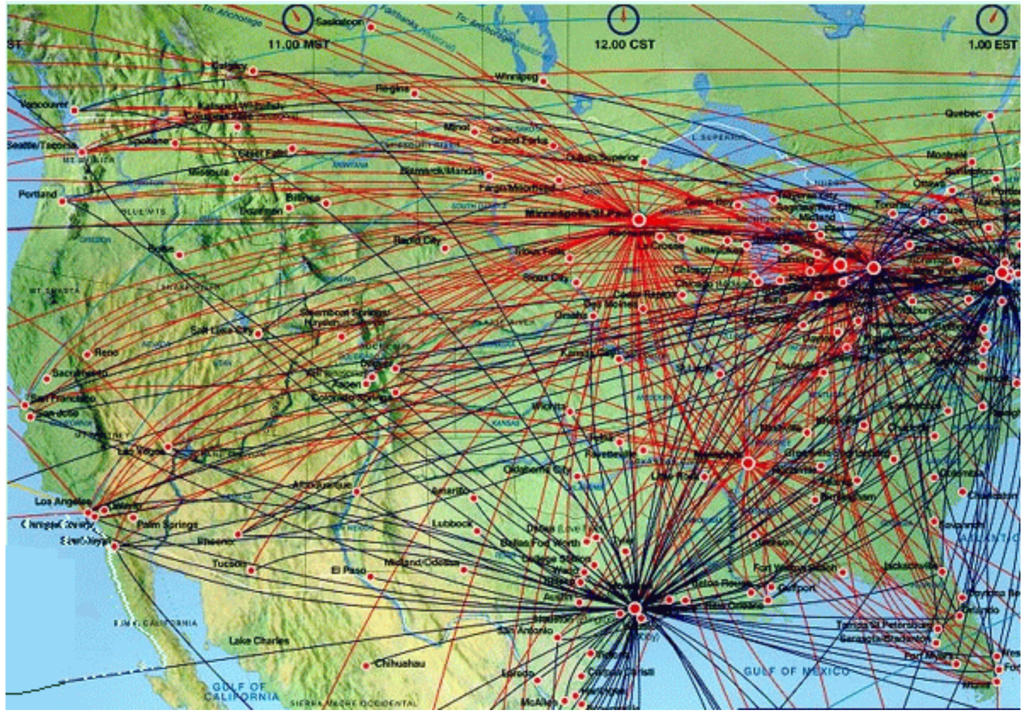
Poisson distribution



Power-law distribution







Homogeneous Network

Scale-free Network

Consequences

Power-law tails

$P(k) \propto k^{-\gamma}$

Average $= \langle k \rangle = \int k\, P(k)dk$

Fluctuations

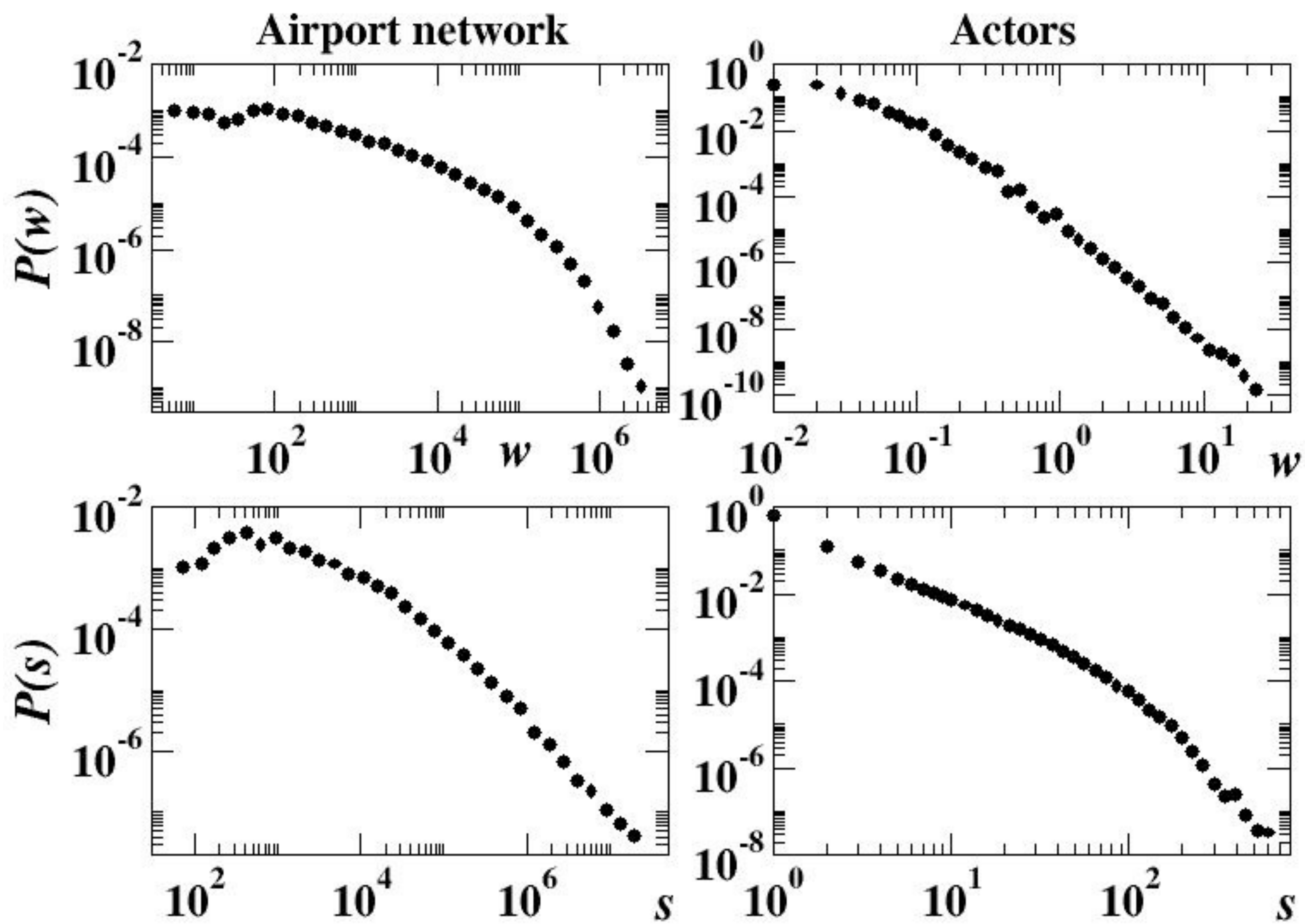
$\langle k^2 \rangle = \int k^2\, P(k)\, dk \propto k_c^{3-\gamma}$

k_c=cut-off due to finite-size

$N \to \infty$ => diverging degree fluctuations for $\gamma < 3$

Level of heterogeneity: $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$

Other heterogeneity levels



Weights

Strengths

Main things to (immediately) measure in a network

- Degree distribution
- Distances, average shortest path, diameter
- Clustering coefficient
- (Weights/strengths distributions)

Networks characteristics

Most often:

- Small diameter
- Large cohesiveness (clustering)
- Heterogeneities, broad distributions

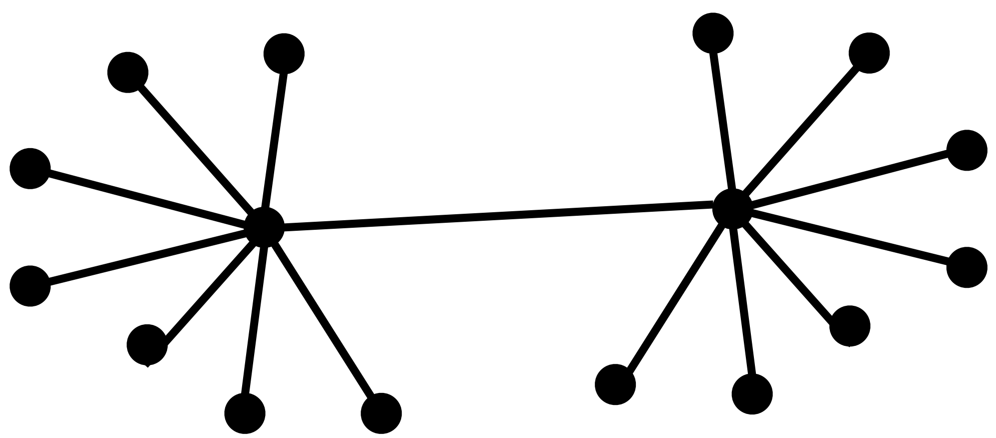
Of course, this is not the whole story...

Correlations

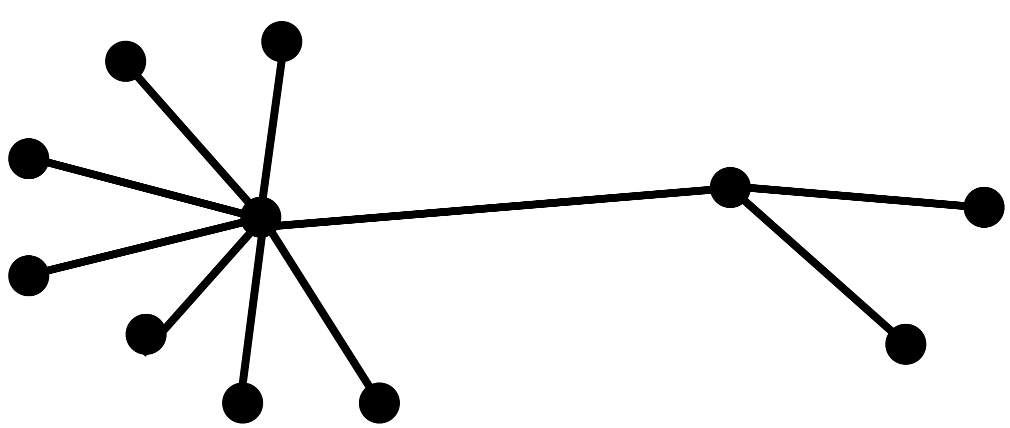
Statistical characterization

Multipoint degree correlations

P(k): not enough to characterize a network



Large degree nodes tend to connect to large degree nodes
Ex: social networks



Large degree nodes tend to connect to small degree nodes
Ex: technological networks

Statistical characterization

Multipoint degree correlations

Measure of correlations:

P(k',k'',….$k^{(n)}$|k): conditional probability that a node of degree k is connected to nodes of degree k', k'',…

Simplest case:

P(k'|k): conditional probability that a node of degree k is connected to a node of degree k'

often inconvenient (statistical fluctuations)

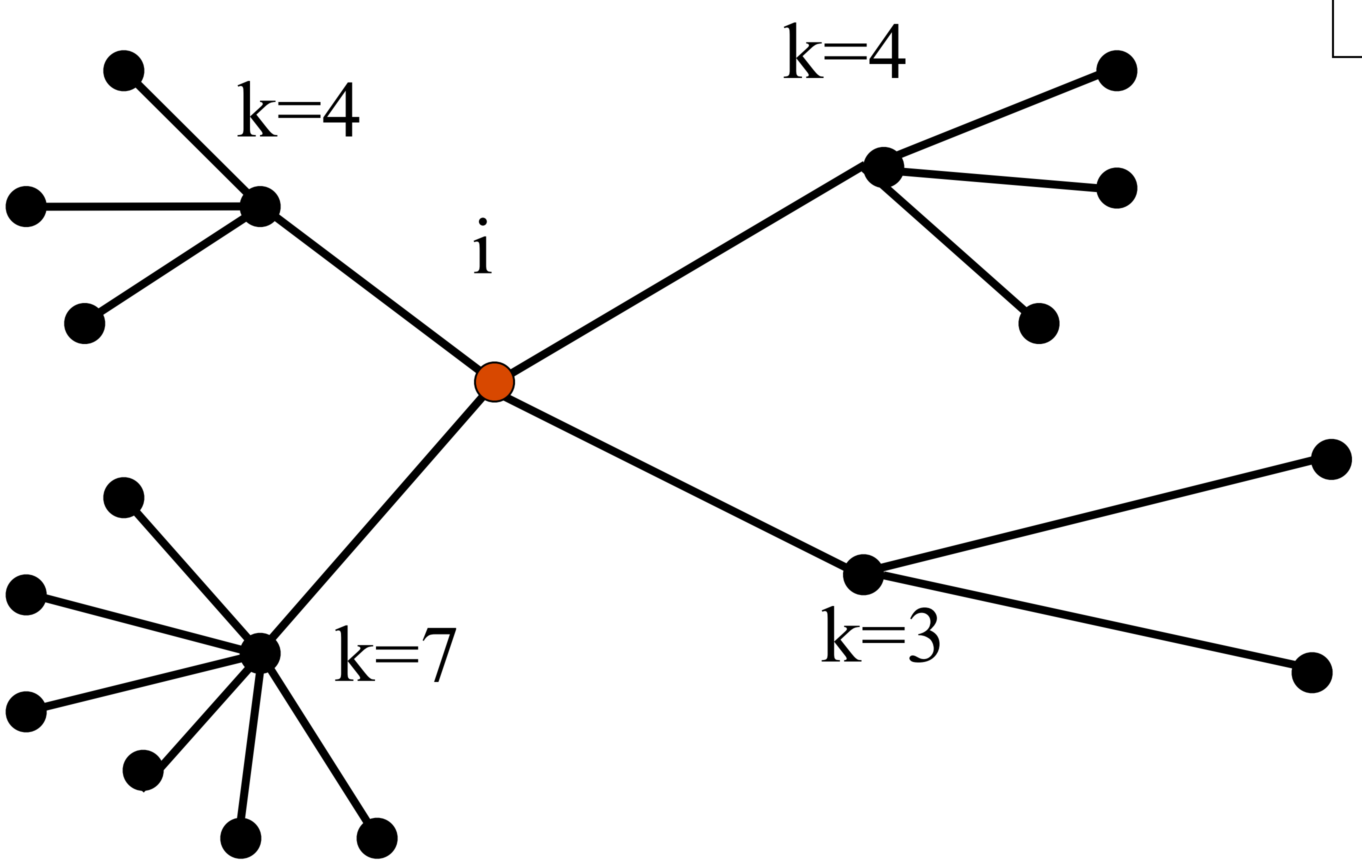
Statistical characterization

Multipoint degree correlations

Practical measure of correlations:

average degree of nearest neighbors

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in V(i)} k_j$$



k_i=4

$k_{nn,i}$=(3+4+4+7)/4=4.5

Statistical characterization

average degree of nearest neighbors

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in V(i)} k_j$$

Correlation spectrum:

putting together nodes which have the same degree

$$k_{nn}(k) = \frac{1}{N_k} \sum_{i/k_i=k} k_{nn,i}$$

class of degree k

$$k_{nn}(k) = \sum_{k'} k' P(k'|k)$$

Statistical characterization

case of *random uncorrelated* networks

P(k'|k)

- independent of k
- proba that an edge points to a node of degree k'

$$\frac{\text{number of edges from nodes of degree k'}}{\text{number of edges from nodes of any degree}} = \frac{k' N_{k'}}{\sum_{k''} k'' N_{k''}}$$

$$P^{unc}(k'|k) = k' P(k') / \langle k \rangle$$

proportional to k' itself

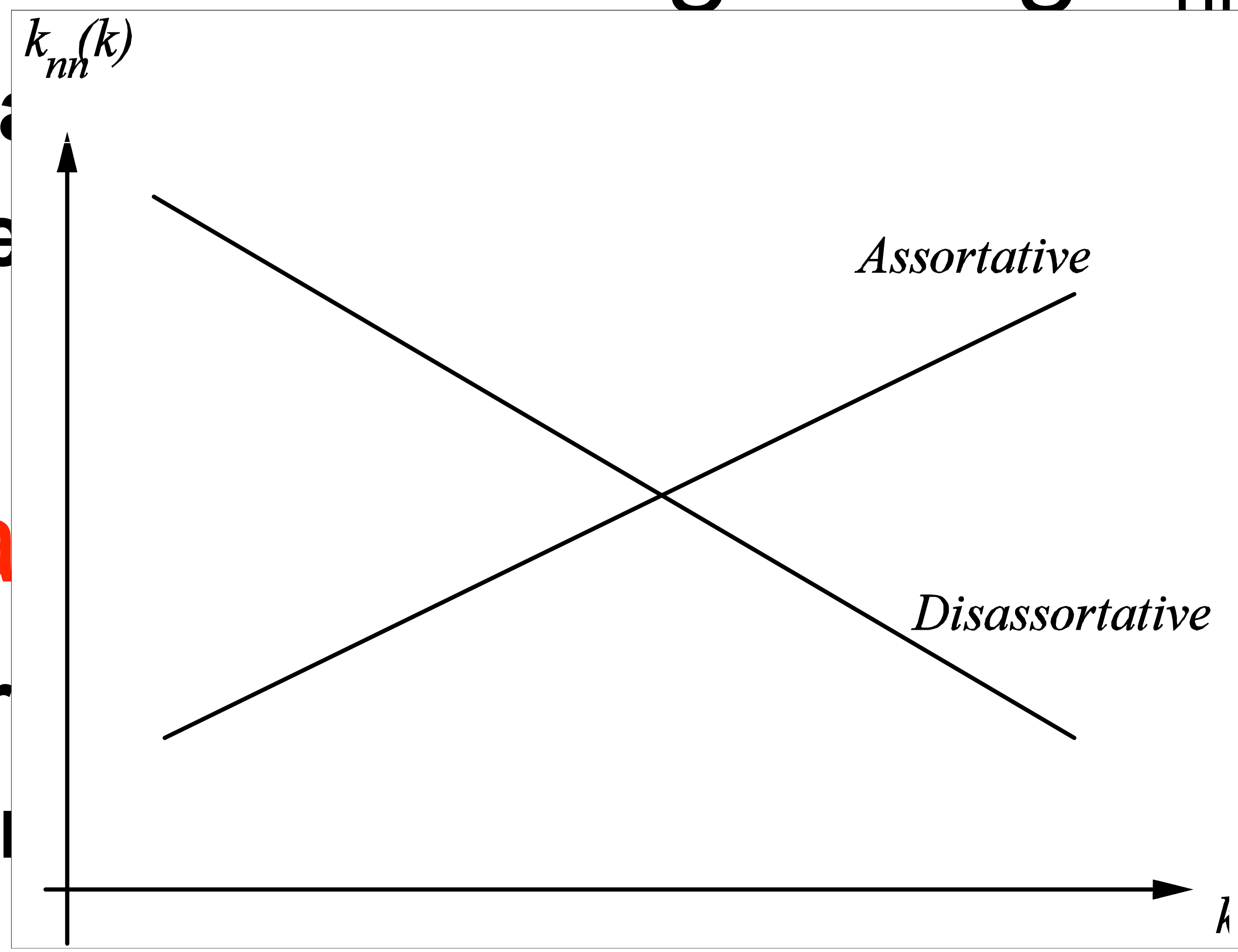
$$k_{nn}^{unc}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

Typical correlations

- Assortative behaviour: growing $k_{nn}(k)$

Example: soci

Large sites are



- Disassorta ing $k_{nn}(k)$

Example: inter

Large sites co archical structure

Correlations: Clustering spectrum

- P(k',k''|k): cumbersome, difficult to estimate from data
- Average clustering coefficient C=average over nodes with very different characteristics

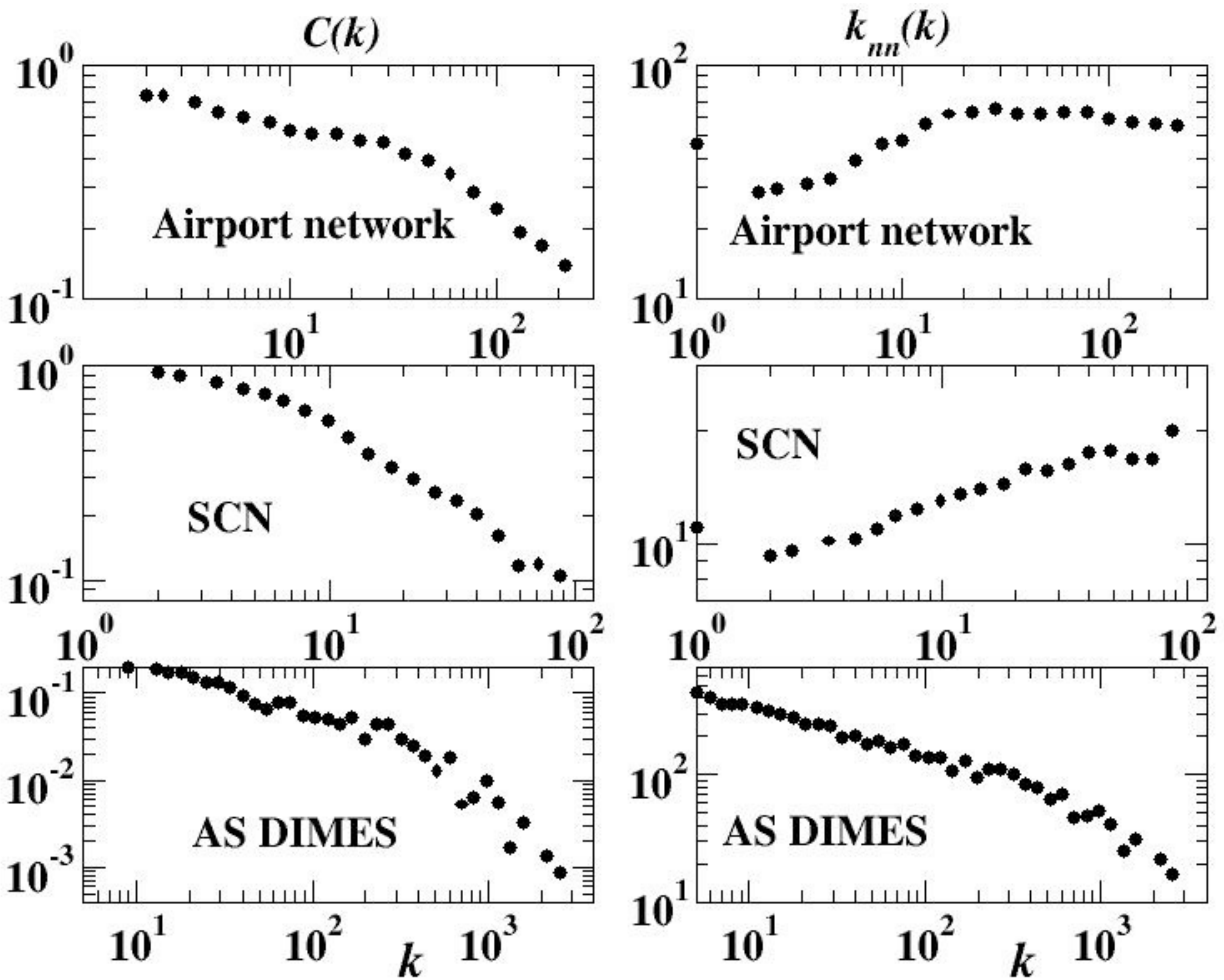
Clustering spectrum:

putting together nodes which have the same degree

$$C(k) = \frac{1}{N_k} \sum_{i/k_i=k} C(i)$$

class of degree k

Empirical clustering and correlations



non-trivial structures

No special scale

Weights-topology correlations

Weighted graphs:

- Strength $s_i = \sum_j w_{ij}$
- Degree $k_i = \sum_j a_{ij}$

$\langle s(k)\rangle$=average strength of nodes of degree k

- no correlations, or random weights:

$$\langle s(k)\rangle \sim \langle w\rangle\, k$$

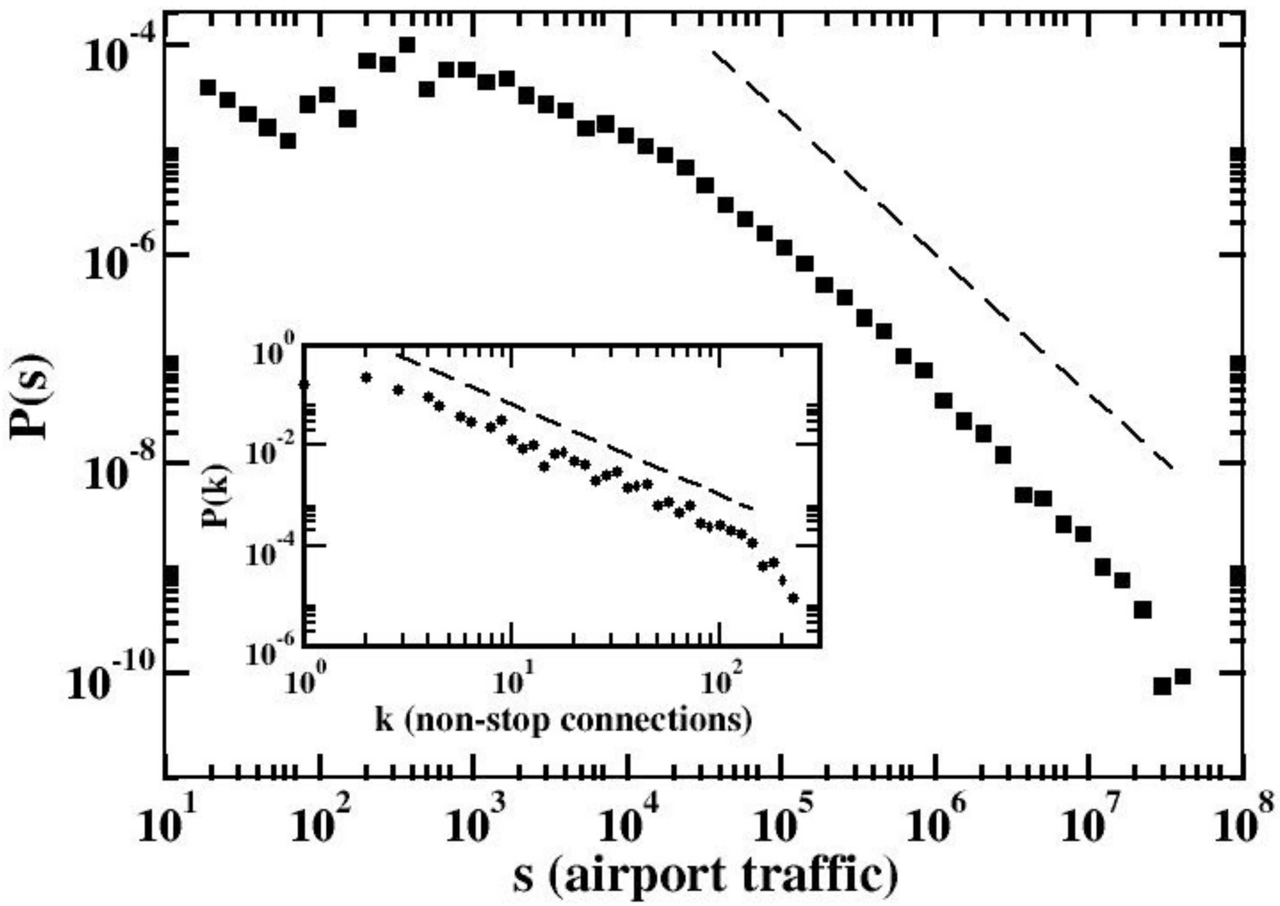
- existence of correlations:

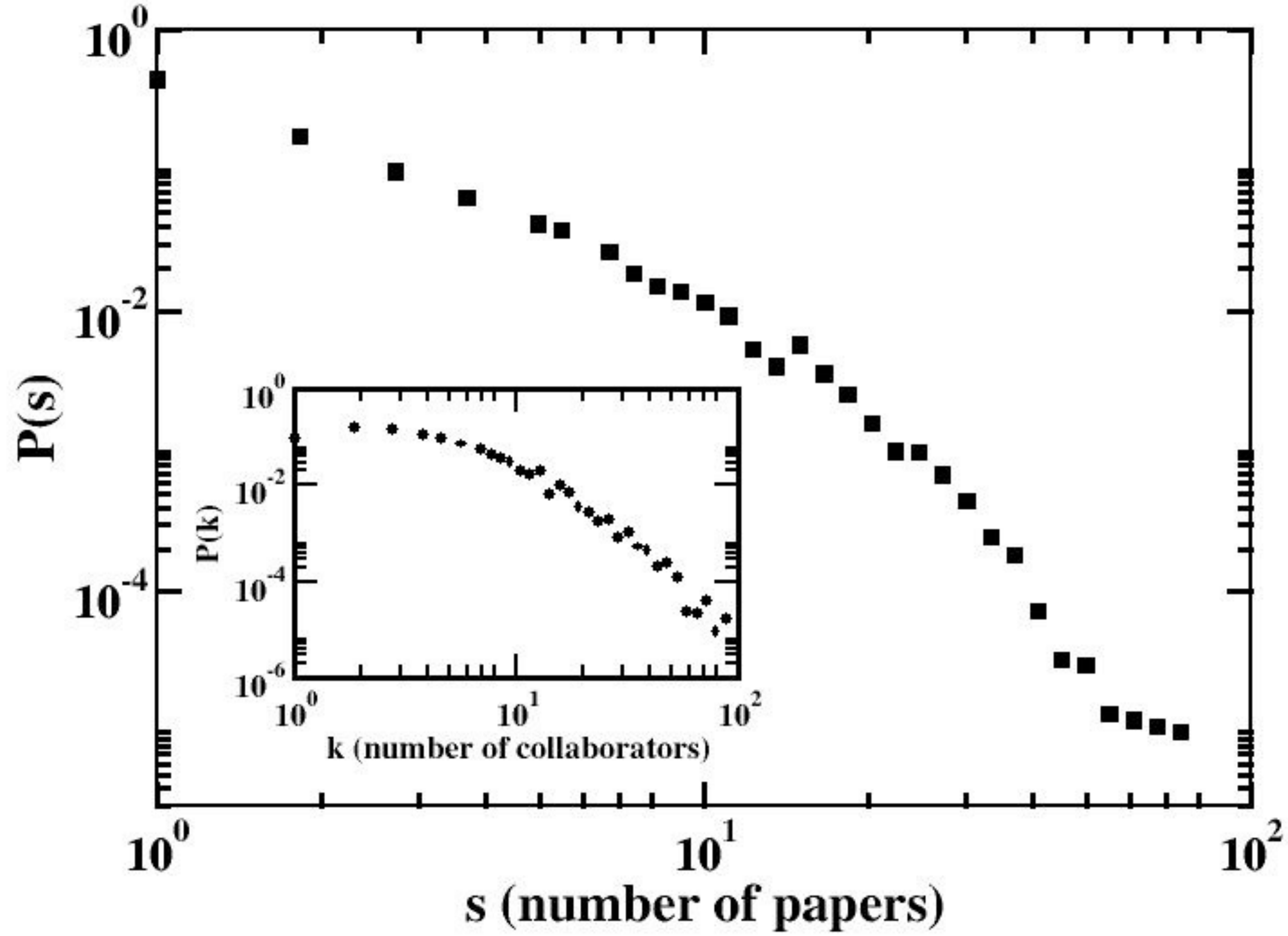
$$\langle s(k)\rangle \sim k^{\alpha}$$

Examples:

- Scientific collaborations: cond-mat archive; N=12722 authors, 39967 links
- Airports' network: data by IATA; N=3863 connected airports, 18807 links
- Human contact networks

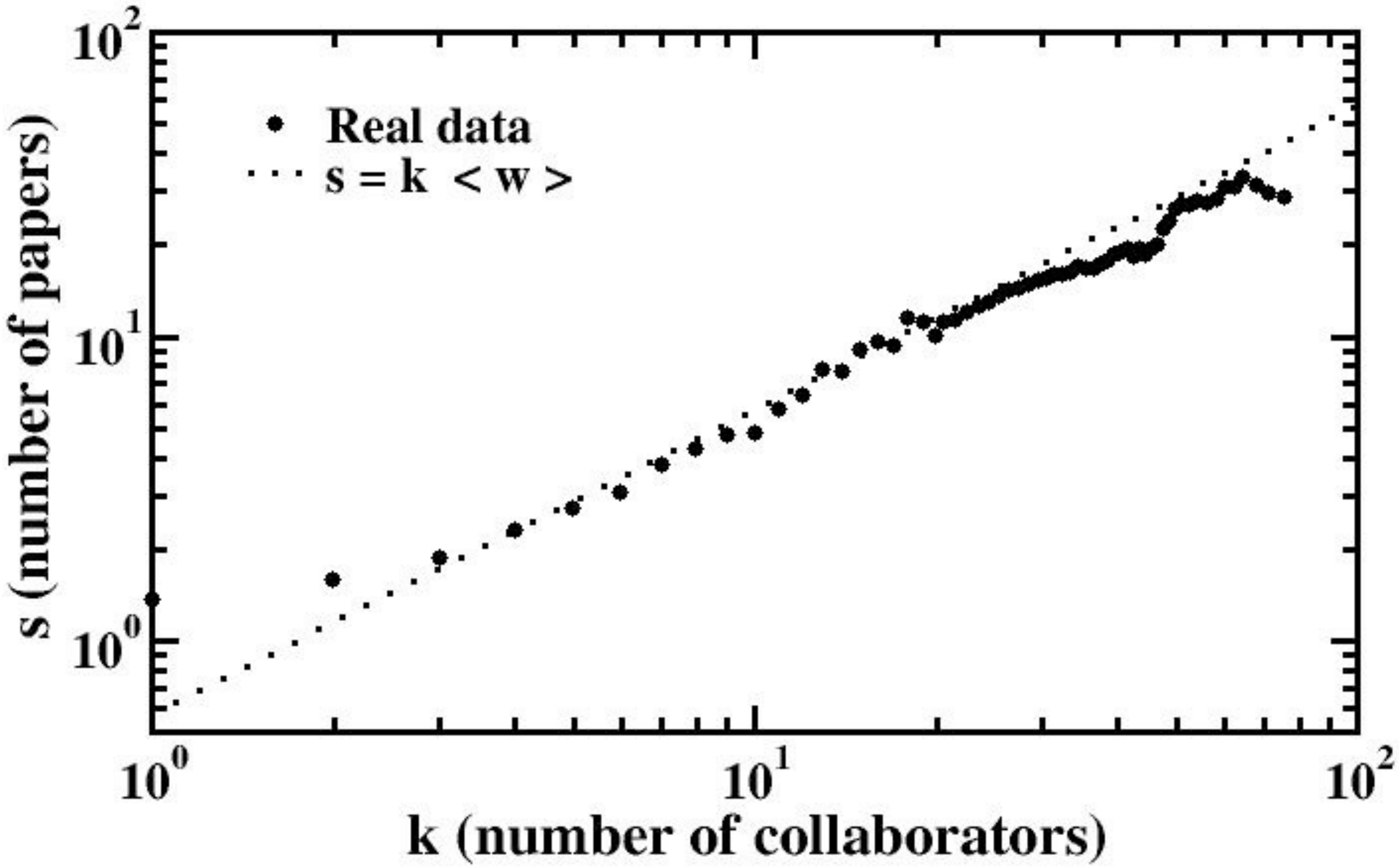
Strengths and degrees





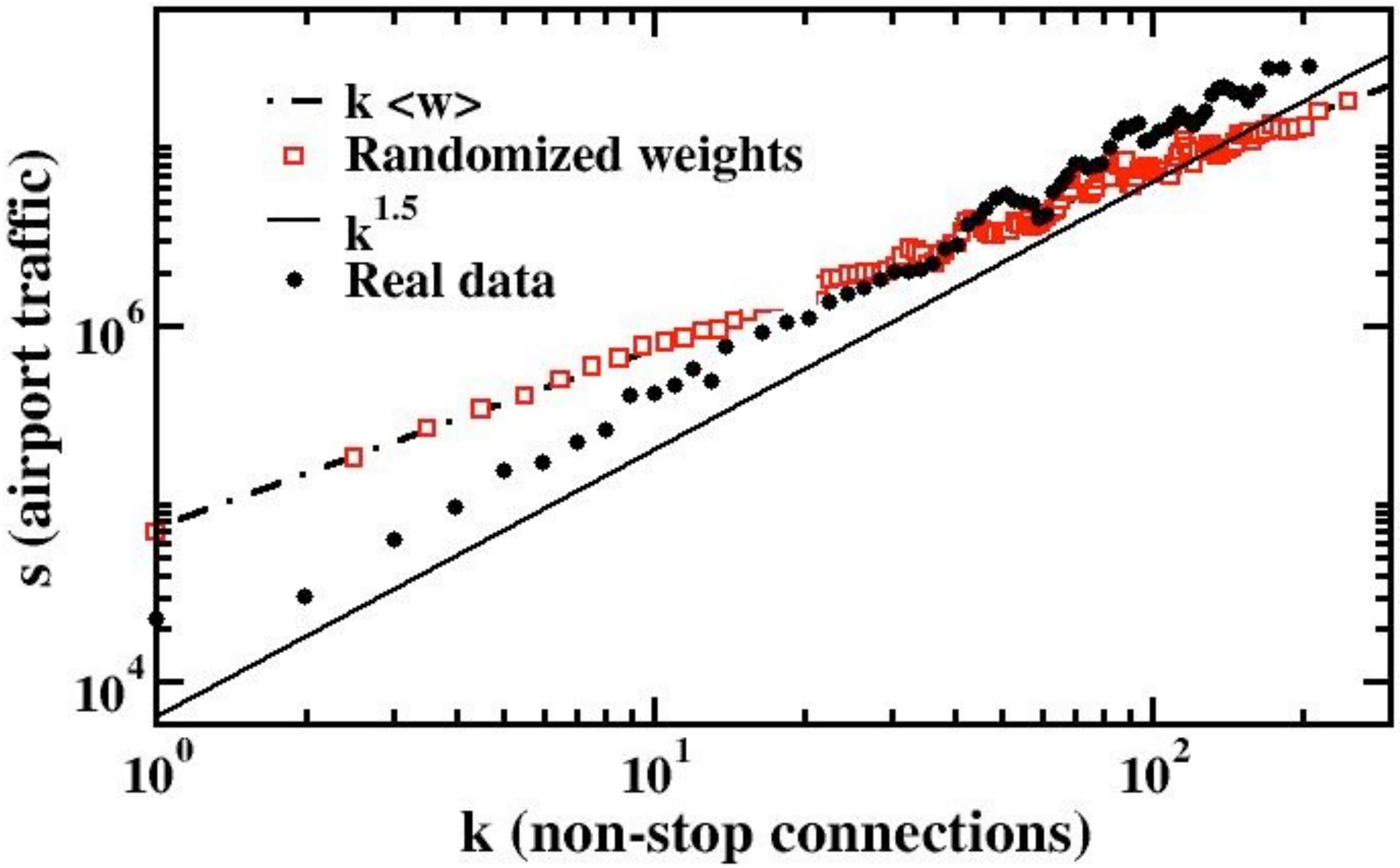
Strengths vs degrees

Collaboration network



Strengths vs degrees

Air transportation network



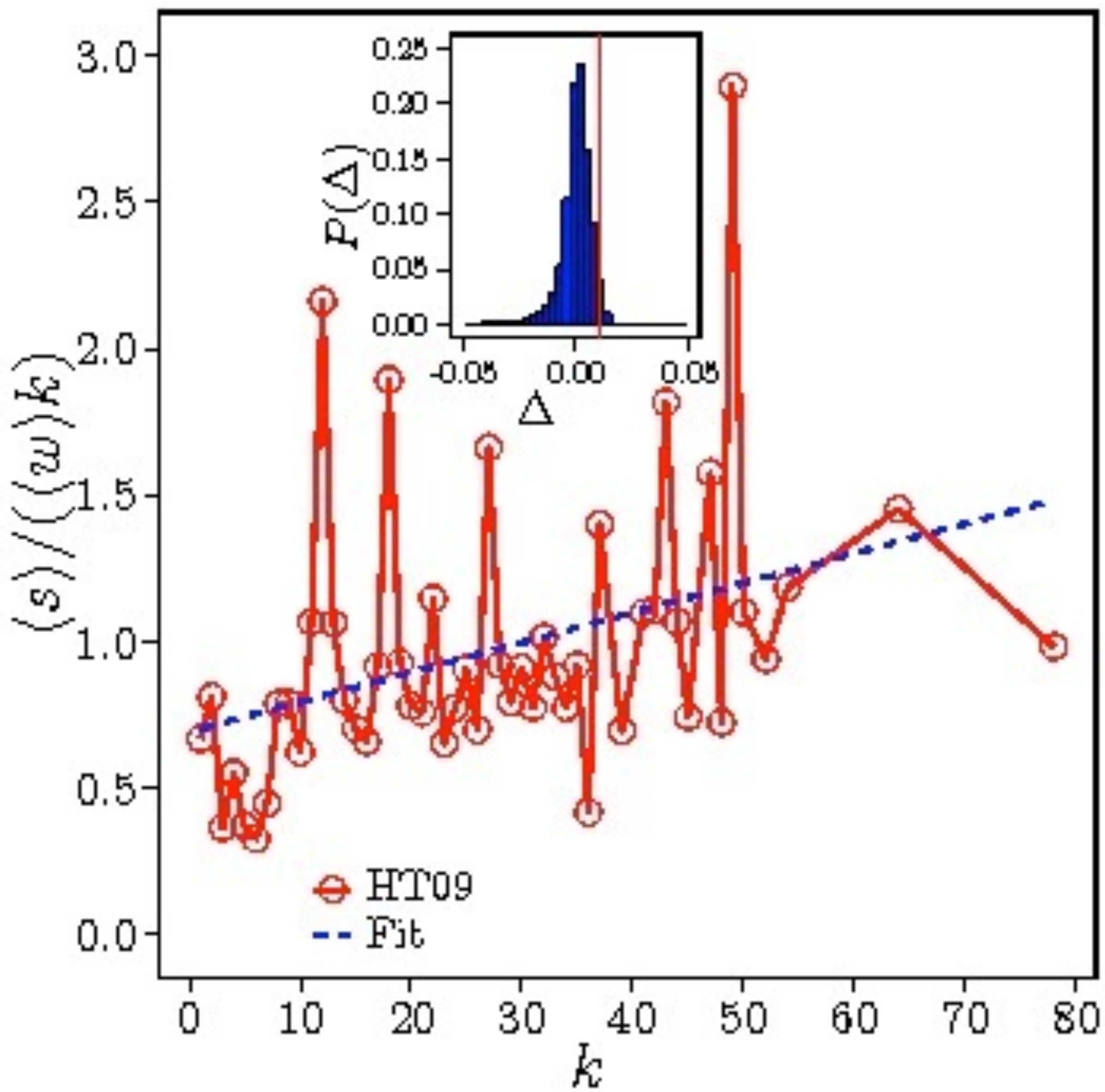
S(k) proportional to k^{β}, $\beta=1.5$

Randomized weights: $\beta=1$

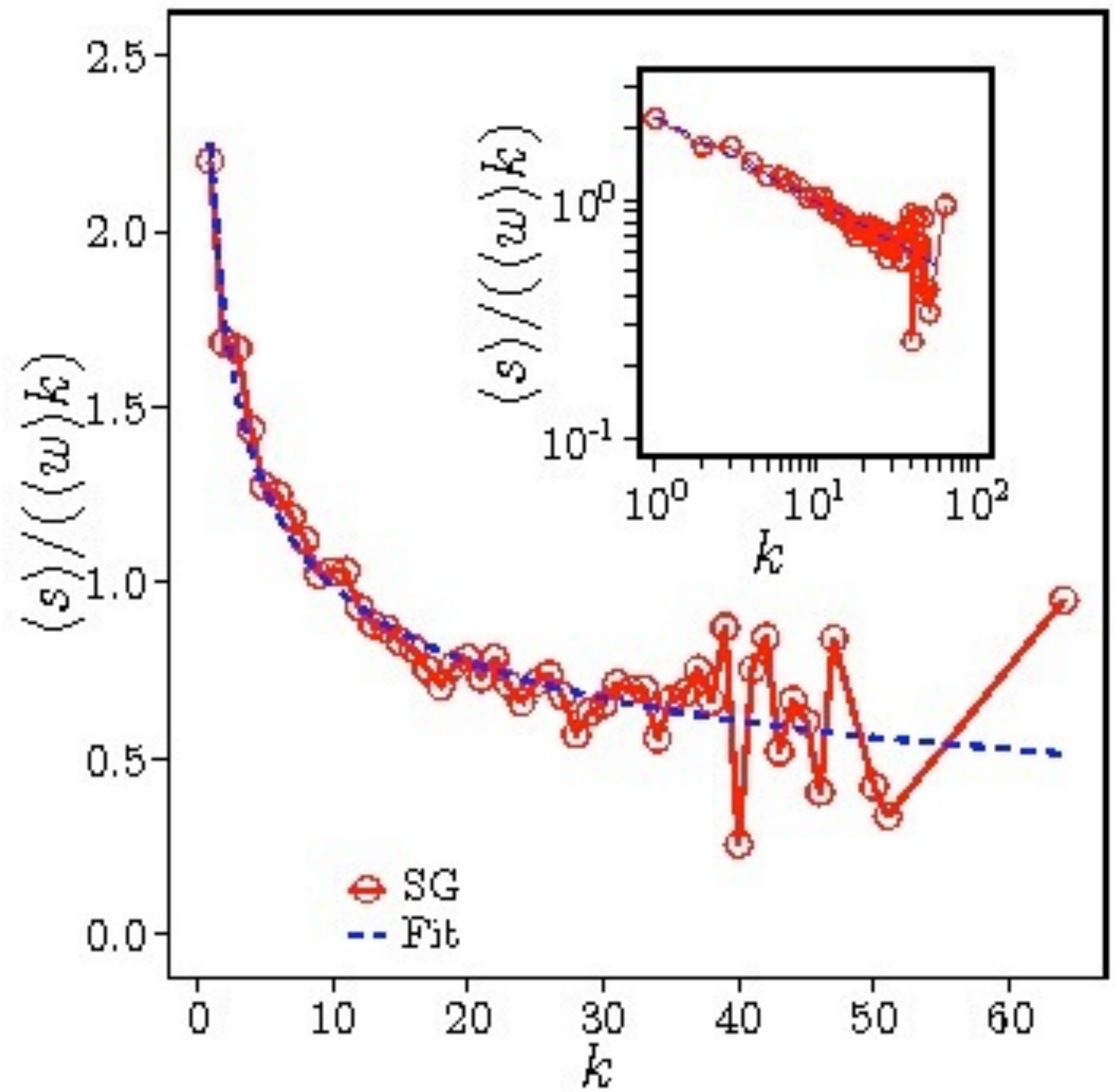
Strong correlations between topology and dynamics

Human contact networks

Conference



Museum



k=number of distinct persons contacted
s=total time spent in contact
Random weights: s ~ <w>k

Structures at various scales

Subgraphs

Subgraphs

A subgraph of G=(V,E) is a graph G'=(V',E')
such that
V' ⊆ V and E' ⊆ E
i.e., V' and E' are subsets of nodes and edges of G

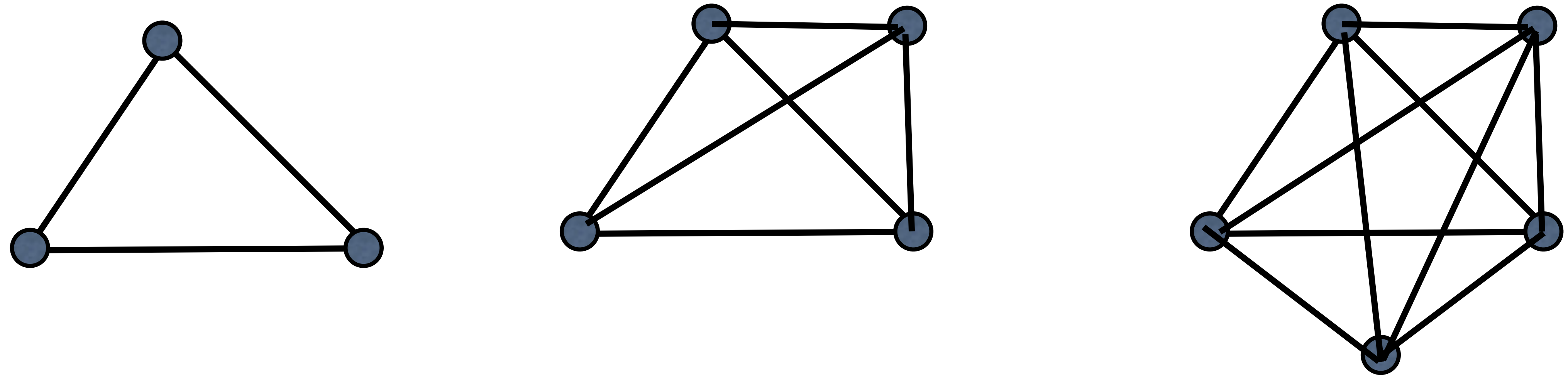
Special case: subgraph *induced* by a set of nodes=
-this set of nodes
-and all links of G between these nodes

Particular subgraphs=connected components

Cliques

A clique is a set C of nodes of G=(V,E) such that for all i,j ∊ C, (i,j)∊ E

Examples:



Motifs

REPORTS

Network Motifs: Simple Building Blocks of Complex Networks

R. MILO, S. SHEN-ORR, S. ITZKOVITZ, N. KASHTAN, [...], AND U. ALON +1 authors Authors Info & Affiliations

SCIENCE • 25 Oct 2002 • Vol 298, Issue 5594 • pp. 824-827 • DOI: 10.1126/science.298.5594.824

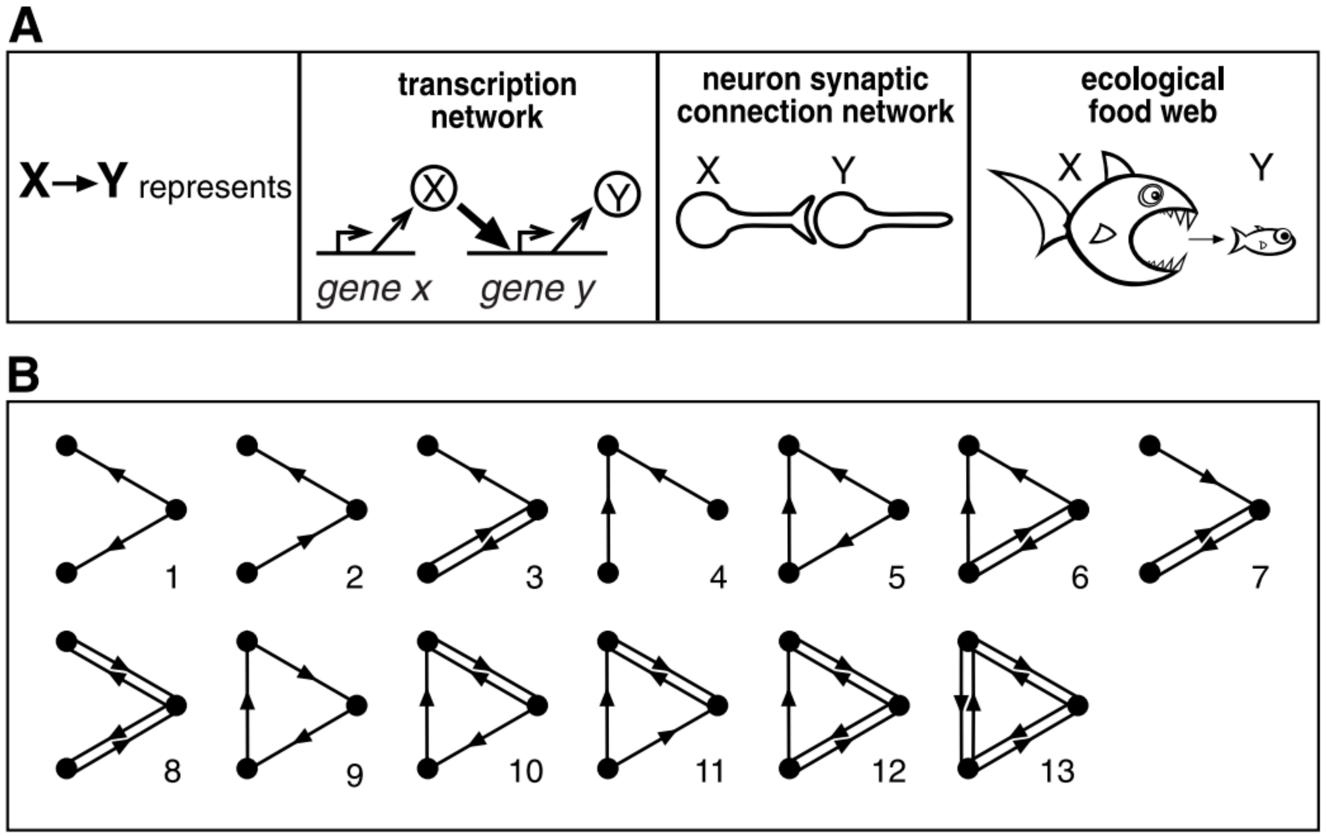


Fig. 1. (**A**) Examples of interactions represented by directed edges between nodes in some of the networks used for the present study. These networks go from the scale of biomolecules (transcription factor protein X binds regulatory DNA regions of a gene to regulate the production rate of protein Y), through cells (neuron X is synaptically connected to neuron Y), to organisms (X feeds on Y). (**B**) All 13 types of three-node connected subgraphs.

Motifs: occur more often than expected "by chance"

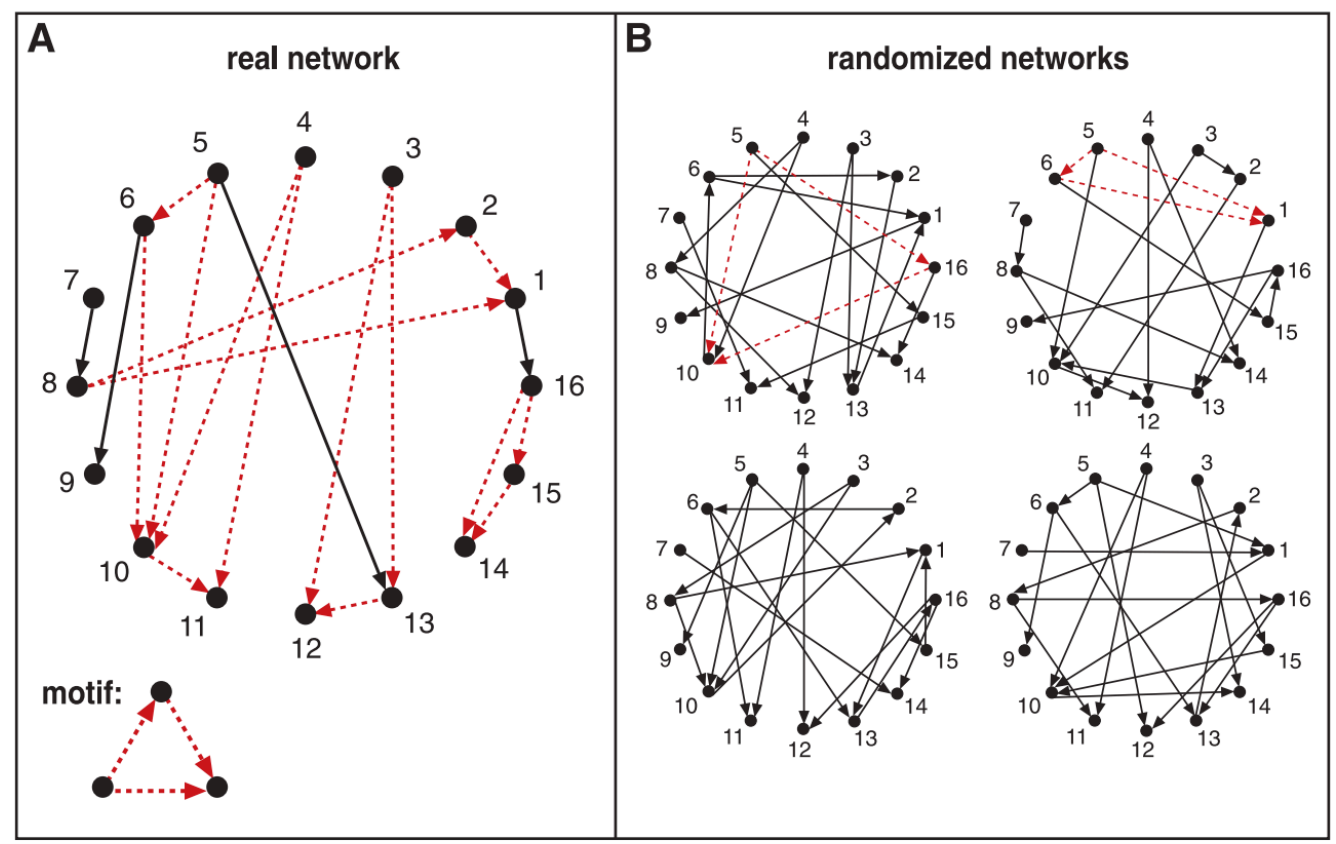
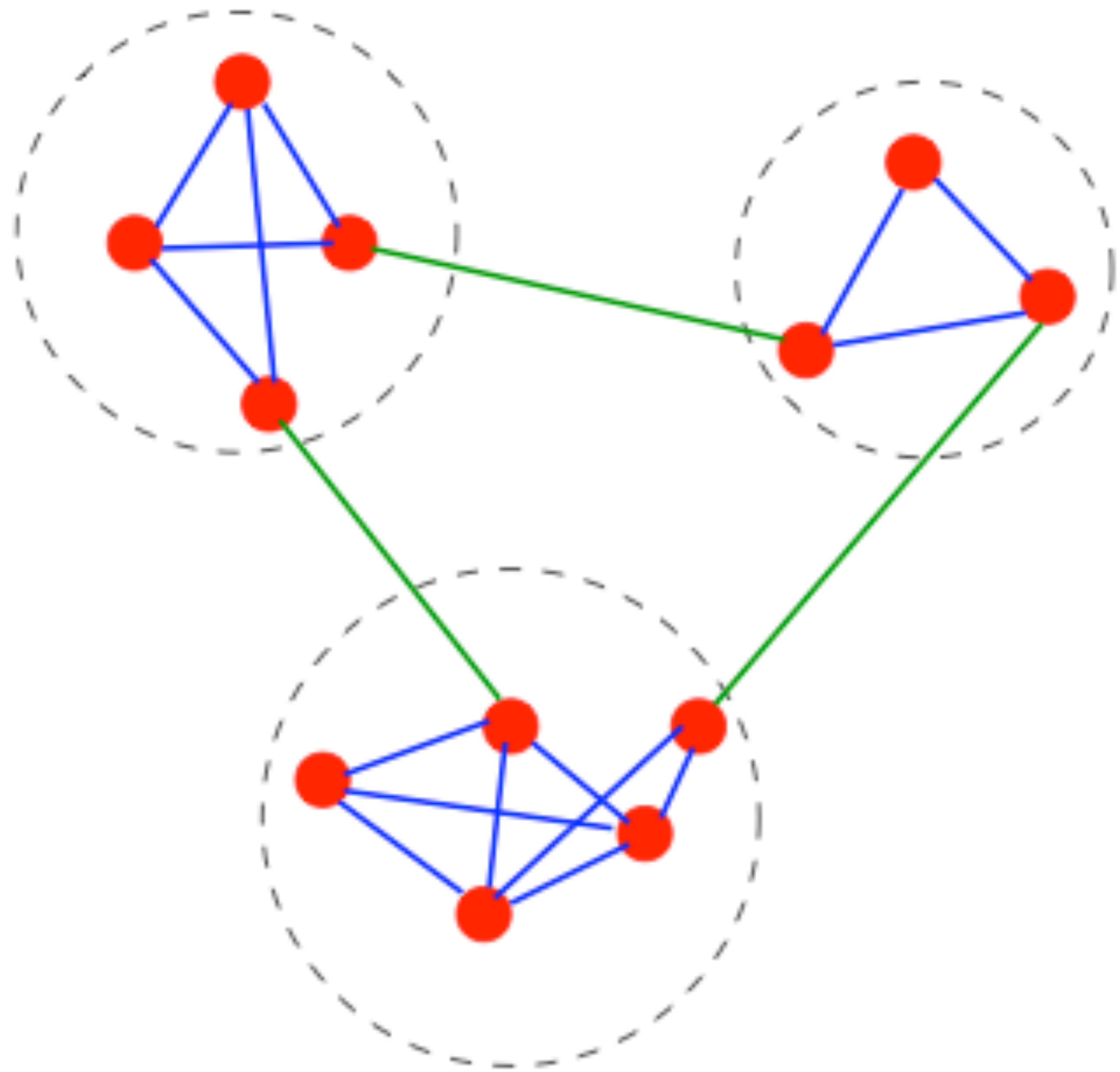


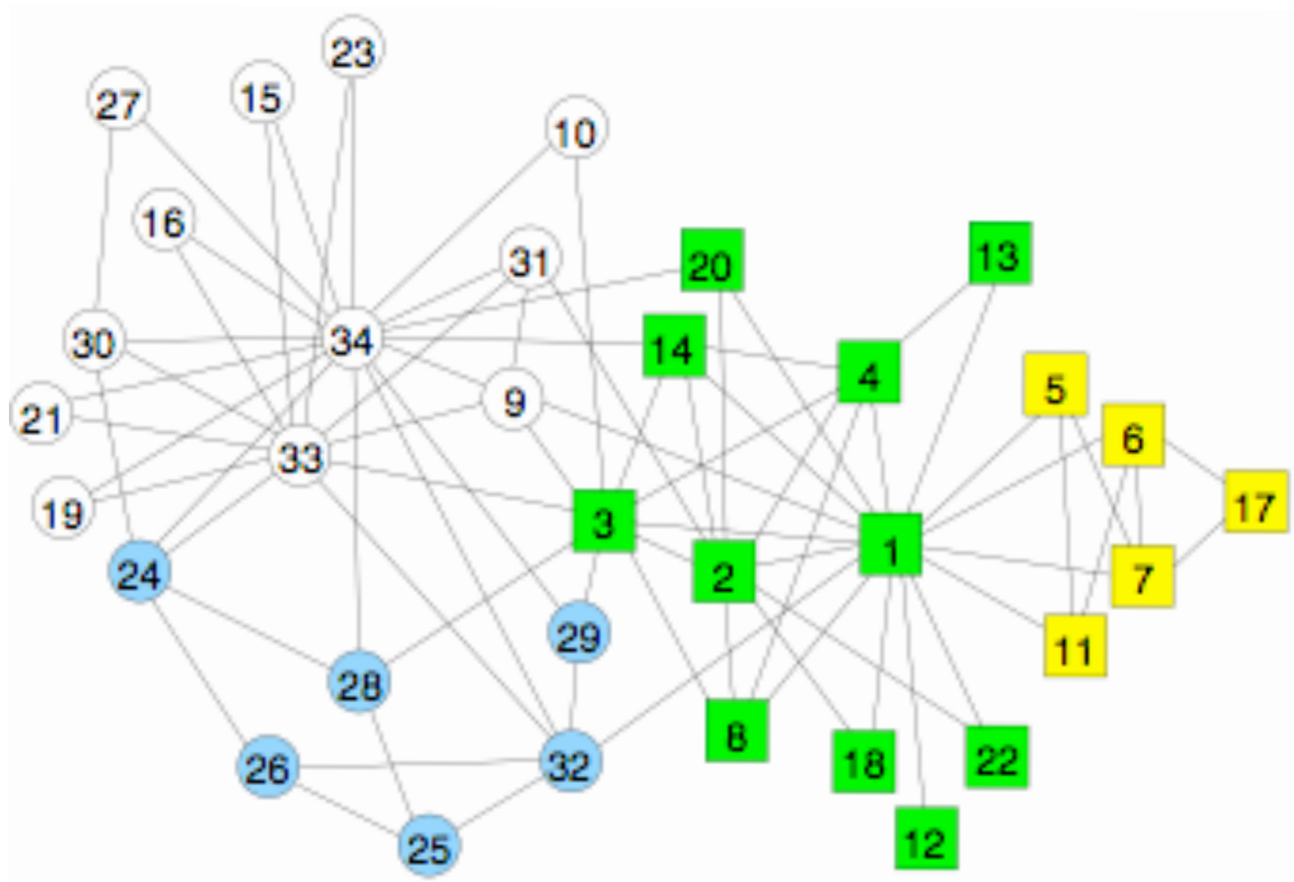
Fig. 2. Schematic view of network motif detection. Network motifs are patterns that recur much more frequently (**A**) in the real network than (**B**) in an ensemble of randomized networks. Each node in the randomized networks has the same number of incoming and outgoing edges as does the corresponding node in the real network. Red dashed lines indicate edges that participate in the feedforward loop motif, which occurs five times in the real network.

Communities and community detection

Communities: examples

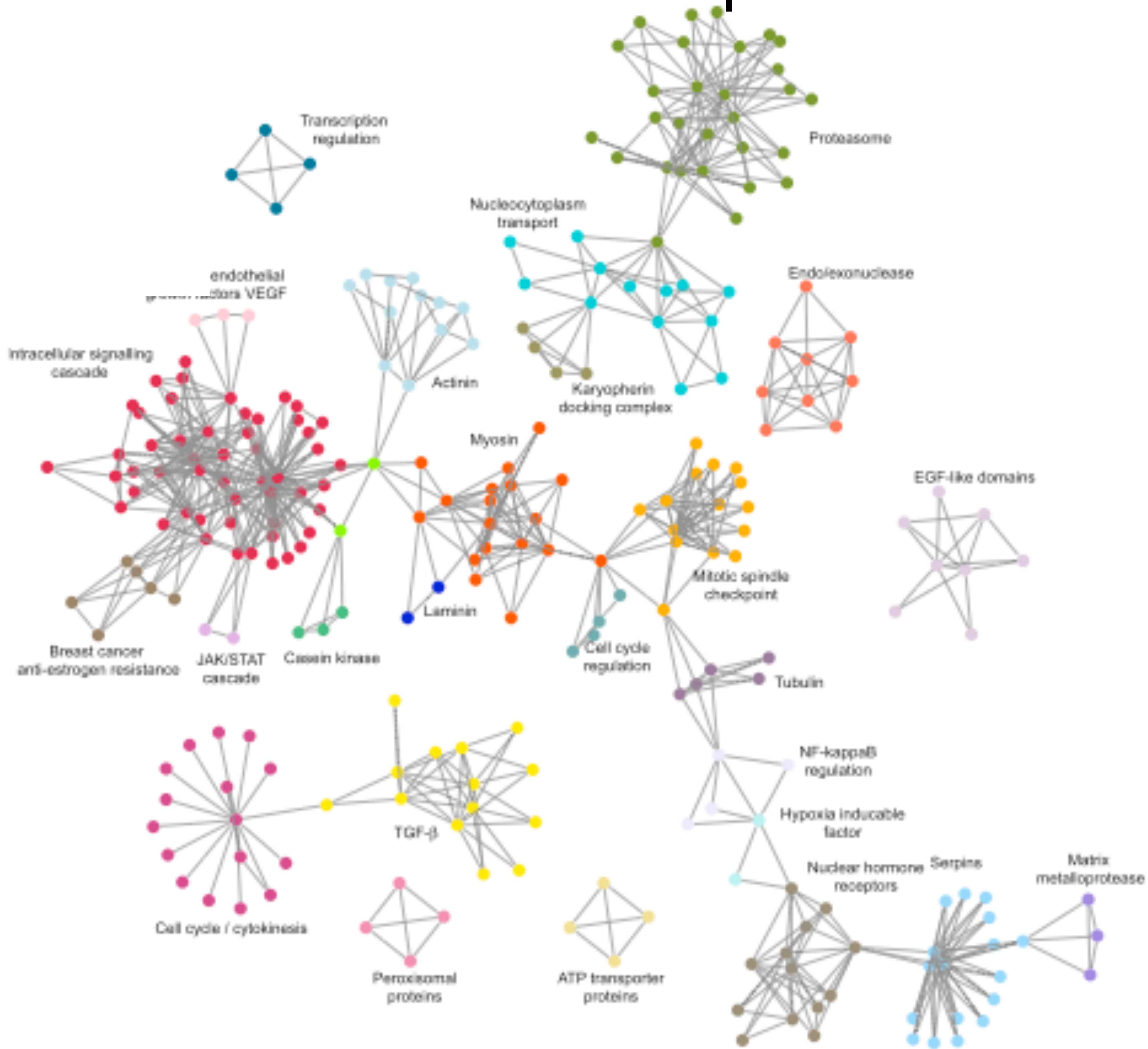


Communities: examples



Social network
(Zachary's karate club)

Communities: examples



Protein-protein interaction network

Communities: (loose) definition

Group of nodes
that are more tightly linked together
than with the rest of the graph

Why are communities interesting?

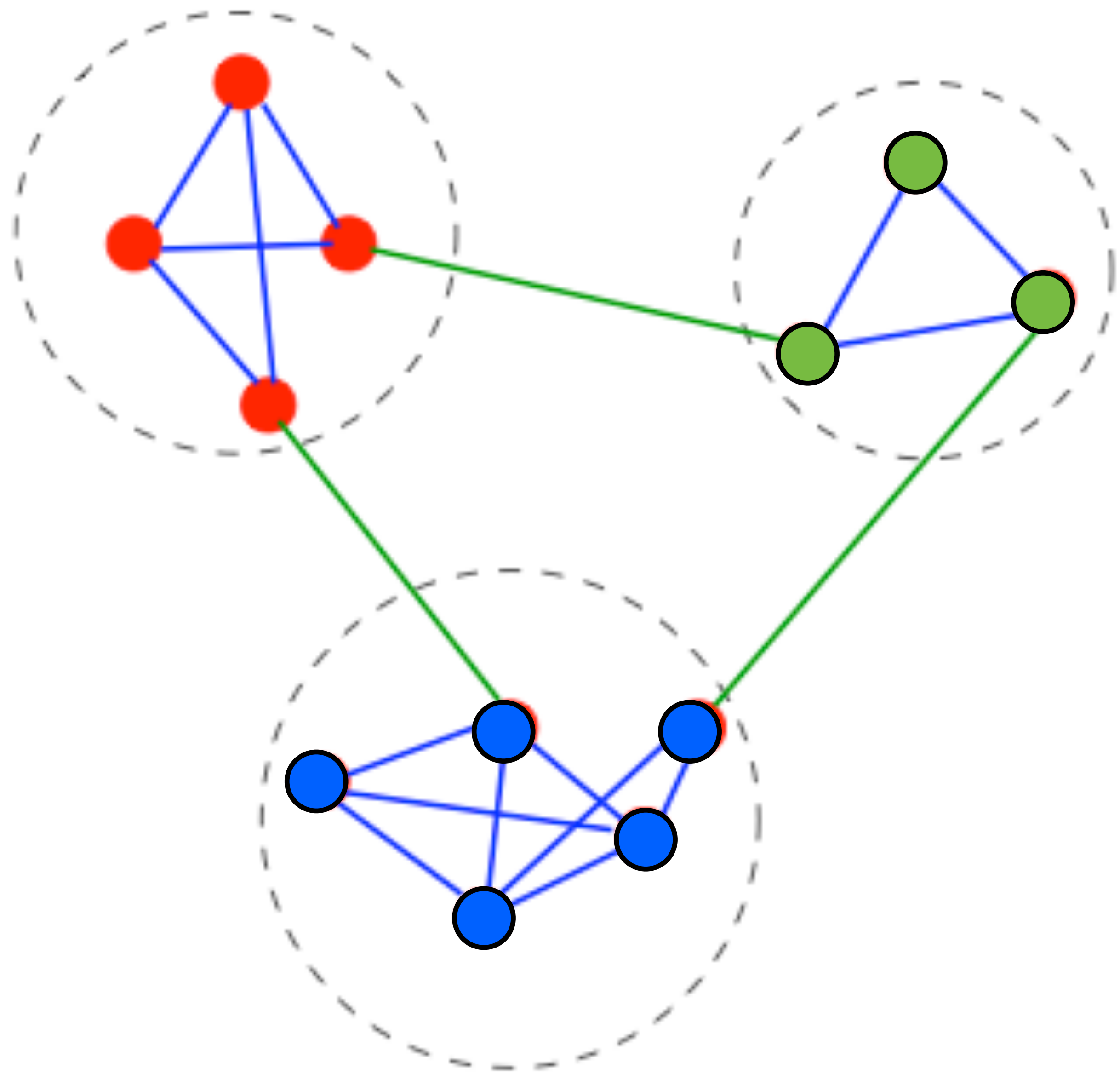
Discover groups in social networks

Discover common interests

Understand opinion formation mechanisms

Communities: consequences on dynamical processes

opinion formation/consensus
disease/information spreading
synchronisation
....



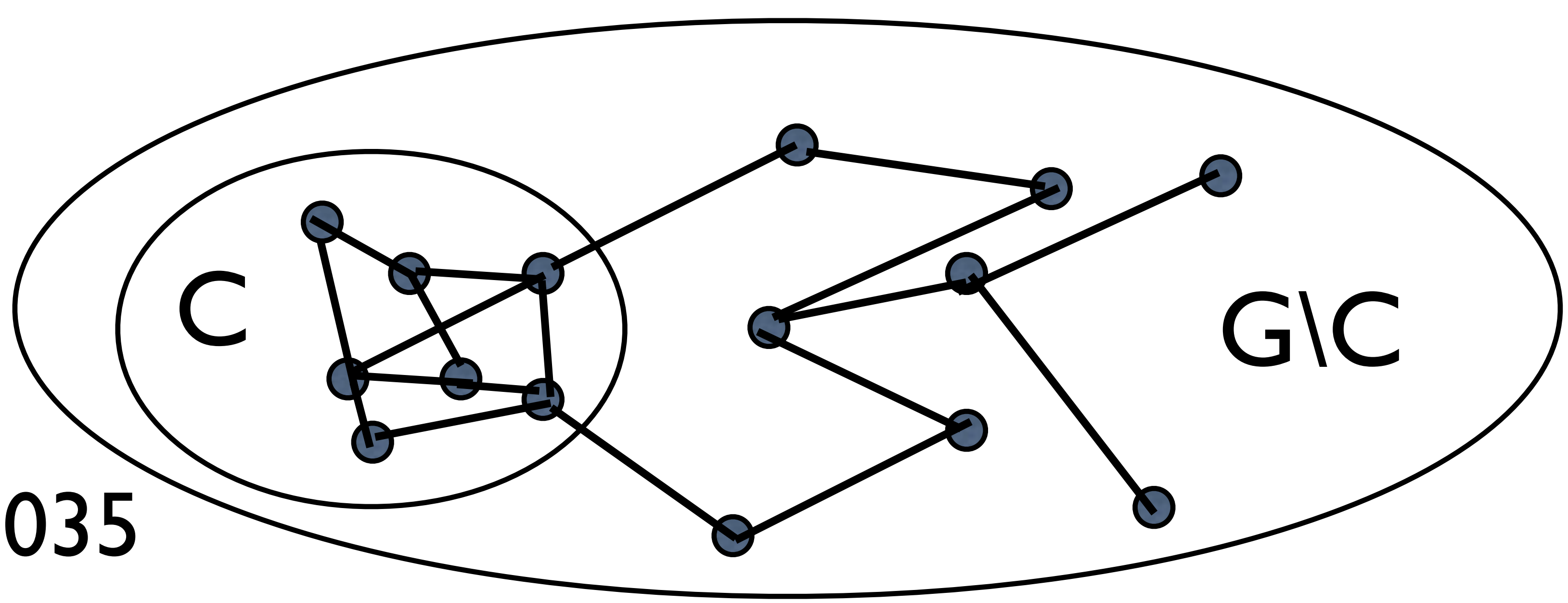
Communities: (loose) definition

subgraph C of G, n nodes, e links
internal density $d = e/(n(n-1)/2)$

For C to be a community, we expect:

- d (much) larger than density of G
- d (much) larger than the density of links towards G\C, given by $d' = e'/(n(N-n))$, where e'=number of links between nodes of C and nodes of G\C

$n=7$, $e=10$
$N-n=8$, $e'=2$
$d=0.24$, $d'=0.035$



Communities: detection problem

Group of nodes
that are more tightly linked together
than with the rest of the graph

- How to (systematically) detect such groups?
- How to partition a graph into communities?
- How to check if it makes sense?

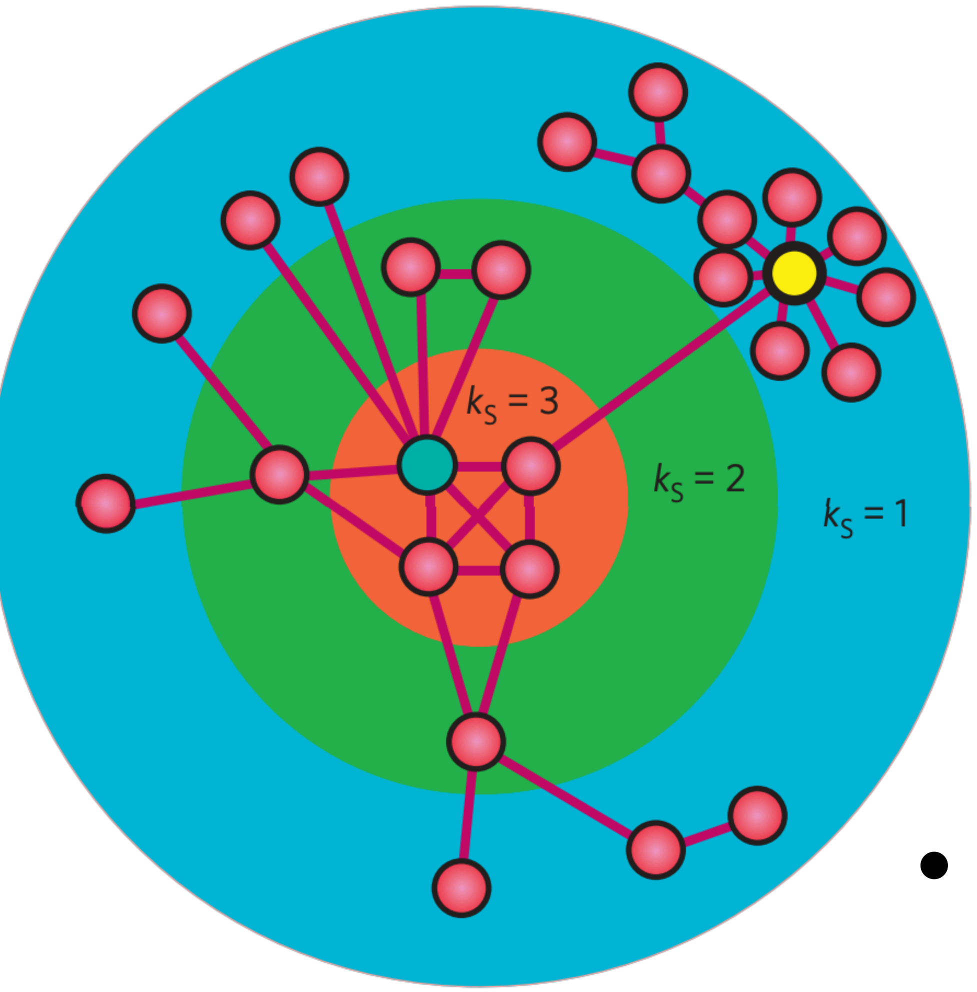
Many algorithms available, most often open

http://www.cfinder.org/
http://www.oslom.org/
http://www.tp.umu.se/~rosvall/code.html
https://skewed.de/tiago

For a review
S. Fortunato, Phys. Rep. **486**, 75-174, 2010
(http://sites.google.com/site/santofortunato/)

Hierarchies

k-core decomposition



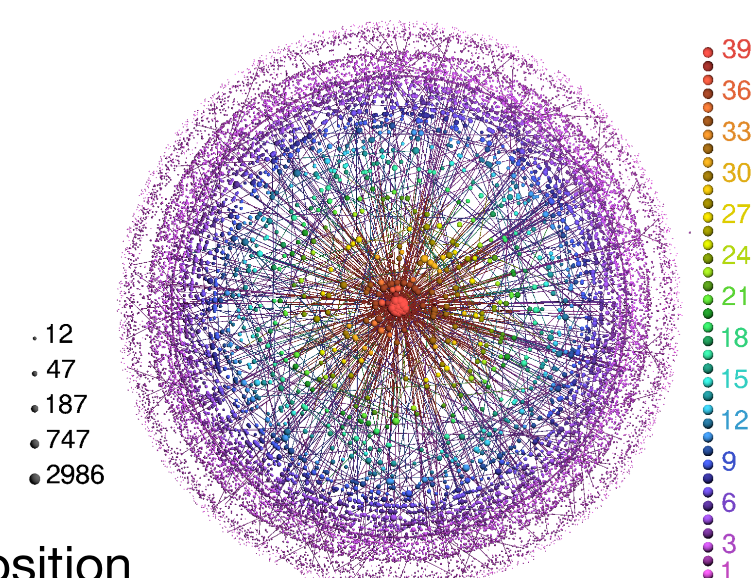
(picture from Kitsak et al., Nat Phys 2010)

graph G=(V,E)

k-core of graph G: maximal subgraph such that for all vertices *in this subgraph* have degree at least k

- vertex i has shell index k iff it belongs to the k-core but not to the (k+1)-core
- k-shell: ensemble of all nodes of shell index k

http://lanet-vi.fi.uba.ar/



Large scale networks fingerprinting and visualization using the k-core decomposition
I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, A. Vespignani, Neural Information Processing Systems (2005)

nature physics

Published: 29 August 2010

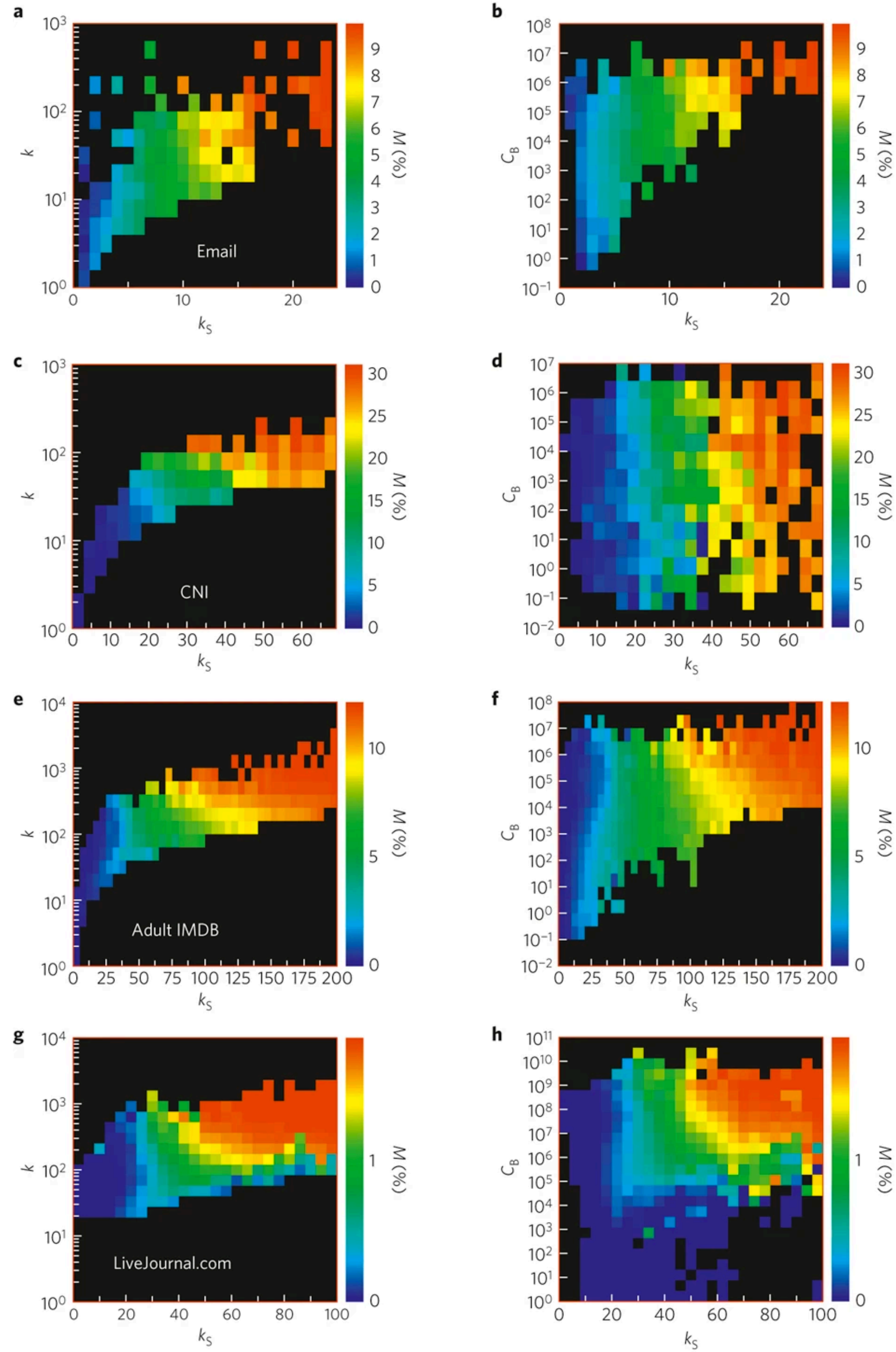
Identification of influential spreaders in complex networks

Maksim Kitsak, Lazaros K. Gallos, Shlomo Havlin, Fredrik Liljeros, Lev Muchnik, H. Eugene Stanley & Hernán A. Makse

Nature Physics **6**, 888–893(2010) | Cite this article

Size of an outbreak as a function of the seed's properties

=> largely determined by coreness



k-core analysis

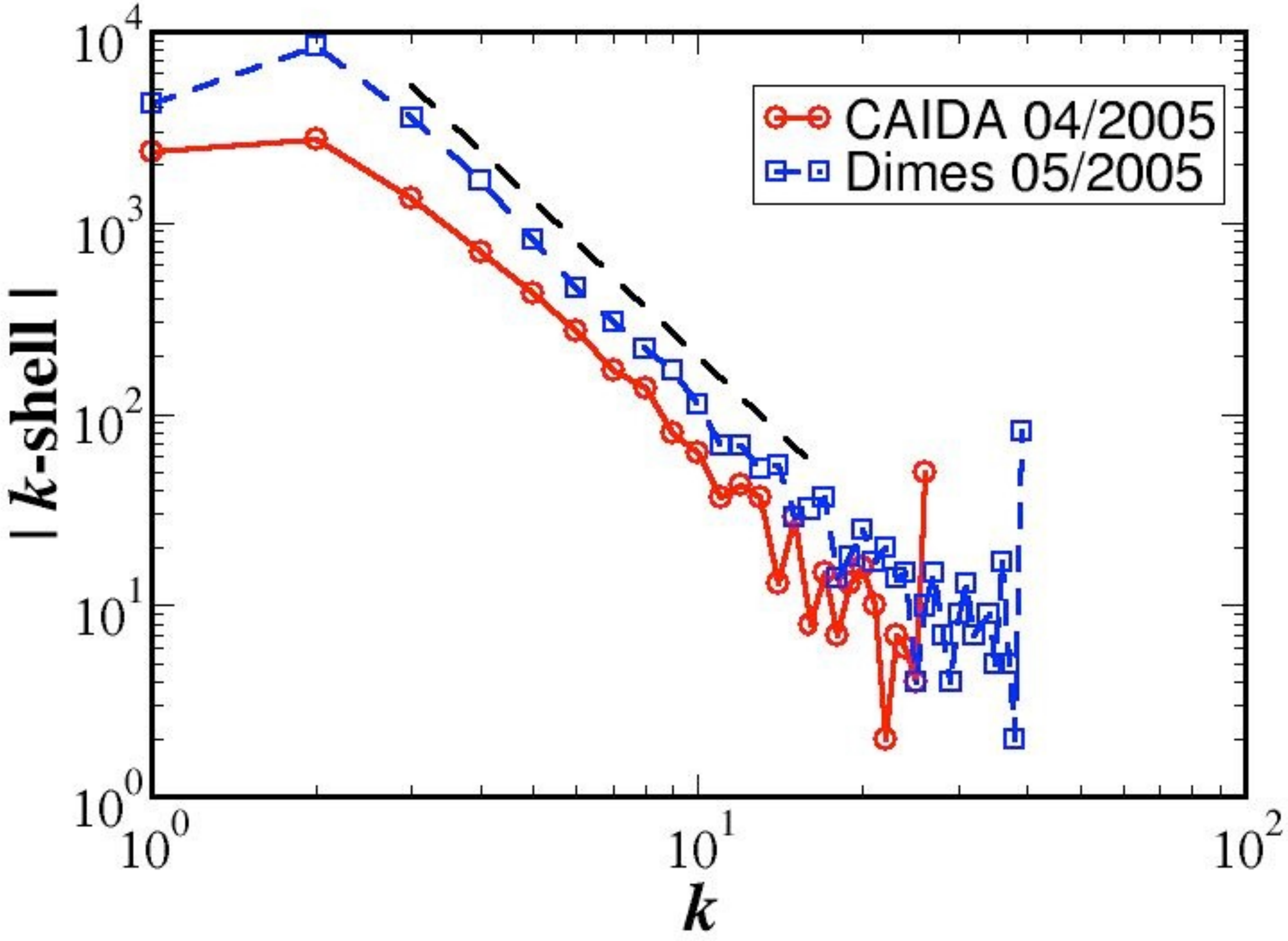
- **maximal** shell index
- **size** of shells versus shell index
- **structure** of the successive cores:

  degree distribution, correlations, clustering...

Remarks

- Additional investigation tool
- Easy to construct
- Focuses on more and more central parts of the network
- Visualization tool

Analysis of AS Internet maps



PHYSICAL REVIEW E
covering statistical, nonlinear, biological, and soft matter physics

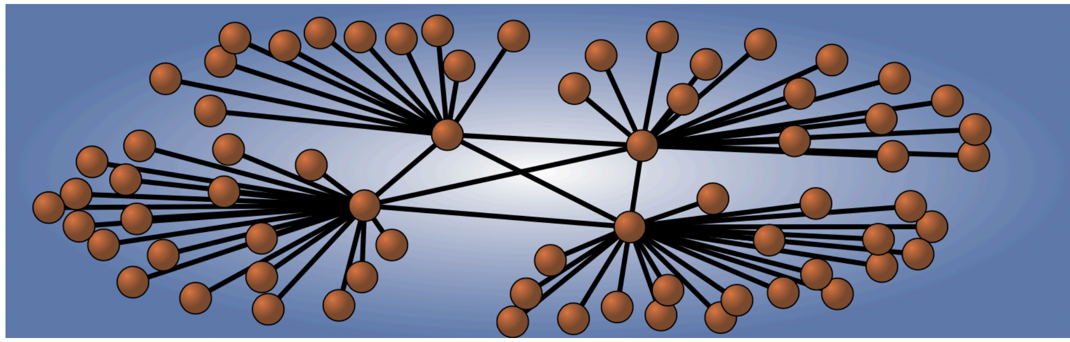
s-core network decomposition: A generalization of k-core analysis to weighted networks

Marius Eidsaa and Eivind Almaas

Phys. Rev. E **88**, 062819 – Published 30 December 2013

Rich-club

Are "rich" nodes (large degree) more inter connected (than chance), i.e., forming a "***rich club***"?



Journals & Magazines > IEEE Communications Letters > Volume: 8 Issue: 3 ? (2004)

The rich-club phenomenon in the Internet topology

Publisher: IEEE **Cite This** **PDF**

Shi Zhou ; R.J. Mondragon **All Authors**

Published: 15 January 2006

Detecting rich-club ordering in complex networks

V. Colizza, A. Flammini, M. A. Serrano & A. Vespignani

Nature Physics **2**, 110–115 (2006) | Cite this article

$S_{>k}$ = set of nodes with degree $> k$;
$N_{>k} \equiv |S_{>k}|$; $E_{>k} \equiv$ # edges between nodes of $S_{>k}$

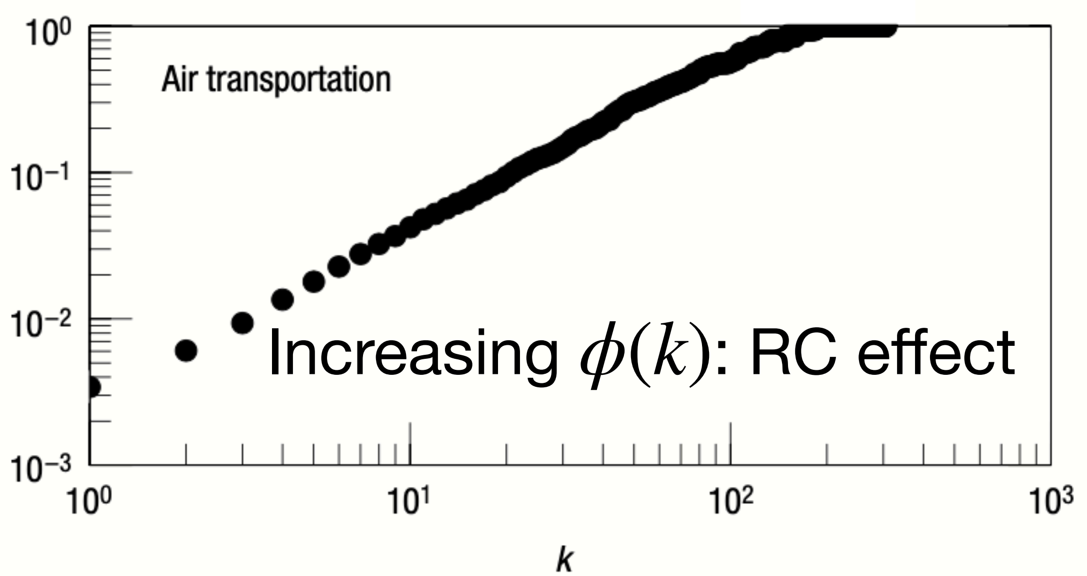
Rich Club coefficient:

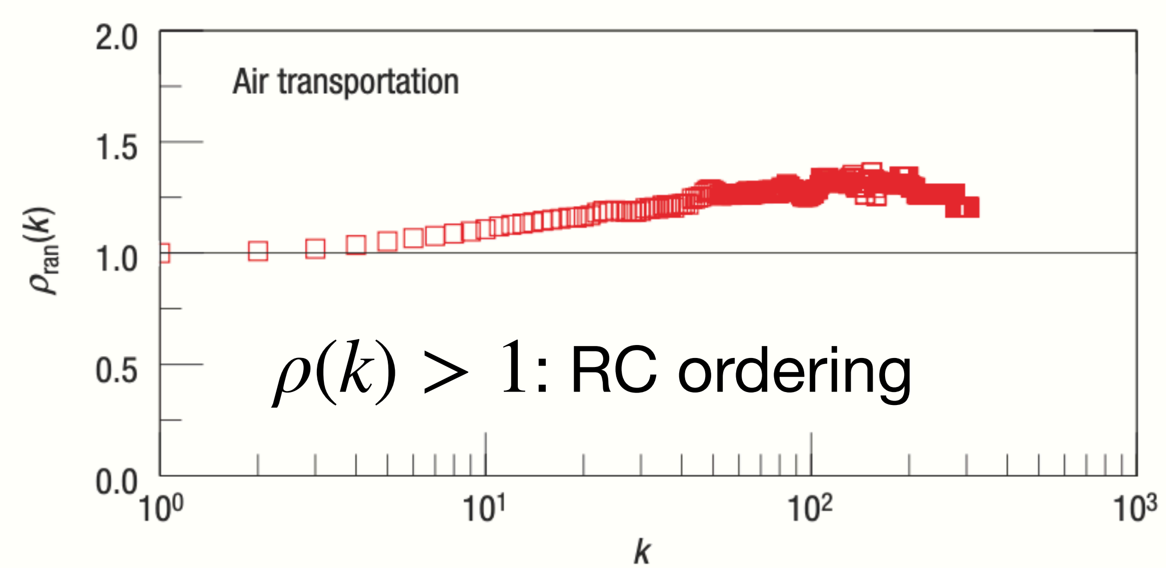
$$\phi(k) \equiv \frac{2E_{>k}}{N_{>k}(N_{>k}-1)}$$

comparison with null model →

Rich Club ordering:

$$\rho(k) \equiv \frac{\phi(k)}{\phi_{rand}(k)}$$





Zhou & Mondragon 2004, Colizza et al. 2006

Core-periphery

ELSEVIER

Social Networks 21 (1999) 375–395

SOCIAL NETWORKS

www.elsevier.com/locate/socnet

Models of core/periphery structures

Stephen P. Borgatti [a,*], Martin G. Everett [b,1]

[a] Department of Organization Studies, Carroll School of Management, Boston College, Chestnut Hill, MA 02467, USA

[b] School of Computing and Mathematical Sciences, University of Greenwich, 30 Park Row, London SE10 9LS, UK

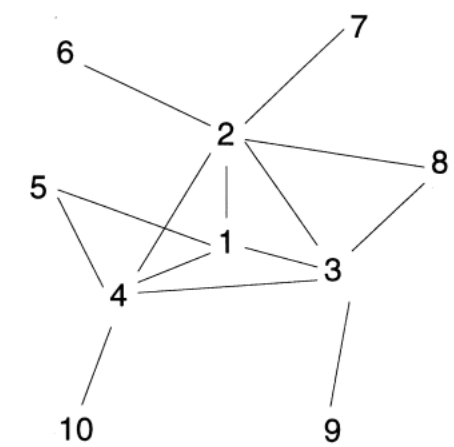


Fig. 1. A network with a core/periphery structure.

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	0	0	0	0	0
2	1		1	1	0	1	1	1	0	0
3	1	1		1	0	0	0	1	1	0
4	1	1	1		1	0	0	0	0	1
5	1	0	0	1		0	0	0	0	0
6	0	1	0	0	0		0	0	0	0
7	0	1	0	0	0	0		0	0	0
8	0	1	1	0	0	0	0		0	0
9	0	0	1	0	0	0	0	0		0
10	0	0	0	1	0	0	0	0	0	

The matrix has been blocked to emphasize the pattern, which is that core nodes are adjacent to other core nodes, core nodes are adjacent to some periphery nodes, and periphery nodes do not connect with other periphery nodes. In blockmodeling terms, the core/core region is a 1-block, the core/periphery regions are (imperfect) 1-blocks, and the periphery/periphery region is a 0-block. We claim that this pattern is characteristic of core/periphery structures and is in fact a defining property. [2]

Core-periphery

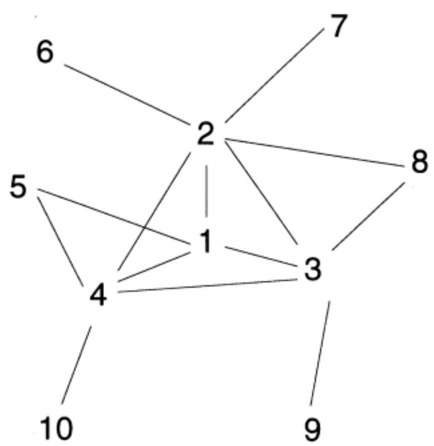


Fig. 1. A network with a core/periphery structure.

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	0	0	0	0	0
2	1		1	1	0	1	1	1	0	0
3	1	1		1	0	0	0	1	1	0
4	1	1	1		1	0	0	0	0	1
5	1	0	0	1		0	0	0	0	0
6	0	1	0	0	0		0	0	0	0
7	0	1	0	0	0	0		0	0	0
8	0	1	1	0	0	0	0		0	0
9	0	0	1	0	0	0	0	0		0
10	0	0	0	1	0	0	0	0	0	

Table 2
Idealized core/periphery structure

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	1	1	1	1	1
2	1		1	1	1	1	1	1	1	1
3	1	1		1	1	1	1	1	1	1
4	1	1	1		1	1	1	1	1	1
5	1	1	1	1		0	0	0	0	0
6	1	1	1	1	0		0	0	0	0
7	1	1	1	1	0	0		0	0	0
8	1	1	1	1	0	0	0		0	0
9	1	1	1	1	0	0	0	0		0
10	1	1	1	1	0	0	0	0	0	

Automatic detection??

Core-periphery

Open Access | Published: 19 March 2013

Profiling core-periphery network structure by random walkers

Fabio Della Rossa, Fabio Dercole & Carlo Piccardi

Scientific Reports **3**, Article number: 1467 (2013) | Cite this article

Let w_{ij} be the weight of the edge $i \to j$ in a (possibly) directed, strongly connected[10,14] network with nodes $N = \{1, 2, \ldots, n\}$. At each (discrete) time step, a random walker which is in node i jumps to j with probability $m_{ij} = w_{ij} \Big/ \sum_h w_{ih}$. Let $\pi_i > 0$ be the asymptotic probability of visiting node i, i.e., the fraction of time steps spent on i. Given a subnetwork S (defined by the node subset $S \subset N$ with all the edges of the original network linking pairs of nodes in S), the *persistence probability* α_S denotes the probability that a random walker which is currently in any of the nodes of S remains in S at the next time step. It is thus a measure of cohesiveness and, indeed, it proved to be an effective tool for finding and testing the community structure of networks[15]. The value of α_S can be made explicit (see Methods) as

Core-periphery profile. In a network with ideal core-periphery structure[11], peripheral nodes (p-nodes) are allowed to link to core nodes only, namely no connectivity exists among p-nodes. This implies that $\alpha_S = 0$ for any subnetwork S composed of p-nodes only, since a random walker is constrained to immediately escape from the set of p-nodes. This suggests a strategy to identify the periphery: find the largest subnetwork with zero persistence probability. In most real-world networks, however, the structure is not ideal although the core-periphery structure is evident: a weak (but not null) connectivity exists among the peripheral nodes. This calls for the generalized definition of α-*periphery*, which denotes the largest subnetwork S with $\alpha_S \le \alpha$: a random walker which is in any of the nodes of the α-periphery, will escape, at the next step, with probability $1 - \alpha$.

Core-periphery

Open Access | Published: 19 March 2013

Profiling core-periphery network structure by random walkers

Fabio Della Rossa, Fabio Dercole & Carlo Piccardi

Scientific Reports **3**, Article number: 1467 (2013) | Cite this article

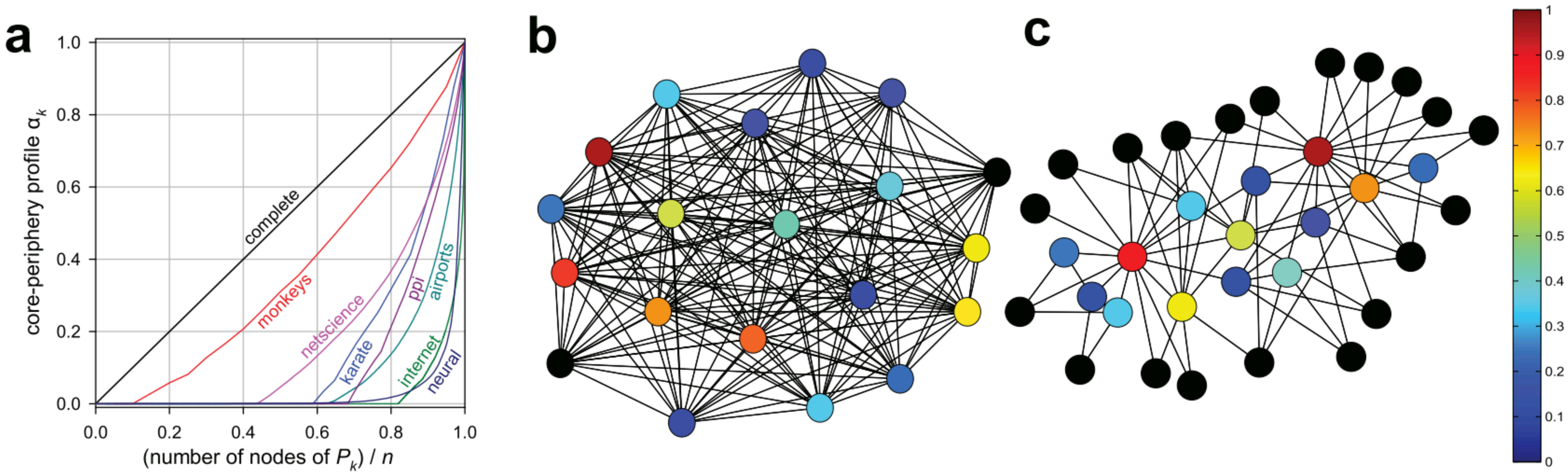


Figure 2 | Core-periphery analysis of real-world networks. (a). The core-periphery profiles of the networks describing: the social interactions within a troop of *monkeys*, $n = 20$ (graph in panel (b)); the friendship among the members of a *karate* club, $n = 34$ (graph in panel (c)); the coauthorships among scientists working on networks (*netscience*), $n = 379$; the protein-protein interaction (*ppi*) network of *Saccharomyces cerevisiae*, $n = 1458$; the international *airports* network, $n = 2868$; the *Internet* in 2006, at the level of autonomous systems, $n = 11745$; and the *neural* network of the worm *Caenorhabditis elegans*, $n = 239$. In graphs (b) and (c), nodes are coloured according to their coreness: p-nodes ($\alpha_k = 0$) are in black.

Comparing networks

Motivation

- Avalanche of data sets: classification?

- Compare structures in different contexts
    - transportation networks in different countries
    - biological networks for different cells
    - brain networks in different conditions, patients vs. baseline, etc
    - social networks in different contexts
    - online vs. real-life interaction networks
    - foodwebs in different ecological conditions
    - …

- Evolving networks: compare different periods

- Model validation

- …

Many different methods

scientific reports

Explore content | About the journal | Publish with us

nature > scientific reports > articles > article

Article | Open Access | Published: 26 November 2019

Comparing methods for comparing networks

Mattia Tantardini, Francesca Ieva, Lucia Tajoli & Carlo Piccardi

Scientific Reports **9**, Article number: 17557 (2019) | Cite this article

The Journal of Open Source Software

Product | Solutions | Open Source | Pricing | Search | Sign in | Sign up

netsiphd / netrd Public

Notifications | Fork 40 | Star 141

Code | Issues 14 | Pull requests 1 | Actions | Projects | Security | Insights

master | 1 branch | 6 tags | Go to file | Code

djpasseyjr Two small fixes and black file auto-formatting. (#3... c1c85aa on Mar 29 820 commits

.github/workflows Update CI coverage and tests (#329) 3 months ago

About

A library for network {reconstruction, distances, dynamics}

netrd.readthedocs.io/en/latest/

netrd: A library for network reconstruction and graph distances

Stefan McCabe[1], Leo Torres[1], Timothy LaRock[1], Syed Arefinul Haque[1], Chia-Hung Yang[1], Harrison Hartle[1], and Brennan Klein[1, 2]

1 Network Science Institute, Northeastern University, Boston, MA, USA **2** Laboratory for the Modeling of Biological and Socio-Technical Systems, Northeastern University, Boston, USA

DOI: 10.21105/joss.02990

Many different methods

- Distances between adjacency matrices, e.g. DeltaCon
- Compare statistics, e.g. NetSimile

arXiv > cs > arXiv:1209.2684

Sea

Help

Computer Science > Social and Information Networks

[Submitted on 12 Sep 2012]

NetSimile: A Scalable Approach to Size-Independent Network Similarity

Michele Berlingerio, Danai Koutra, Tina Eliassi-Rad, Christos Faloutsos

- choose list of features for each node (degree, clustering, etc)
- compute statistics for each feature (mean, median, sd, skewness…)
- => signature vector for each graph
- compare signature vectors of different graph (using Canberra distance)

Many different methods

- Distances between adjacency matrices, e.g. DeltaCon

- Compare statistics, e.g. NetSimile

- Use paths

Open Access | Published: 09 January 2017

Quantification of network structural dissimilarities

Tiago A. Schieber, Laura Carpi, Albert Díaz-Guilera, Panos M. Pardalos, Cristina Masoller & Martín G. Ravetti

Nature Communications **8**, Article number: 13928 (2017) | Cite this article

  - build distributions of nodes at distance d for each node, compare them across nodes => Network Node Dispersion NND
  - compute Katz centrality of each node

  - compare the NNDs and the centrality distributions of different networks

Many different methods

- Distances between adjacency matrices, e.g. DeltaCon
- Compare statistics, e.g. NetSimile
- Use paths

Home > Applied Network Science > Article

Research | Open Access | Published: 16 July 2019

An information-theoretic, all-scales approach to comparing networks

James P. Bagrow & Erik M. Bollt

Applied Network Science **4**, Article number: 45 (2019) | Cite this article

Network portraits

Network portraits were introduced in (Bagrow et al. 2008) as a way to visualize and encode many structural properties of a given network. Specifically, the network portrait B is the array with (ℓ,k) elements

$$B_{\ell,k} \equiv \text{the number of nodes who have } k \text{ nodes at distance } \ell \qquad (2)$$

Comparing networks by comparing portraits

Given that a graph G admits a unique B-matrix makes these portraits a valuable tool for network comparison. Instead of directly comparing graphs G and G', we may compute their portraits B and B', respectively, and then compare these matrices. We

Networks and complexity

Complex networks

Complex is not just "complicated"

Cars, airplanes…=> complicated, not complex

Complex (no unique definition):

- many interacting units
- no centralized authority, self-organized
- complicated at all scales
- evolving structures
- emerging properties (heavy-tails, hierarchies…)

Examples: Internet, WWW, Social nets, etc…

Main features of complex networks

- **Many interacting units**
- **Self-organization**
- **Small-world**
- **(Scale-free) heterogeneity**
- **Dynamical evolution**

Standard graph theory

Random graphs

- **Static**
- **Ad-hoc topology**

Example: Internet topology generators
Modeling of the Internet structure with ad-hoc algorithms tailored on the properties we consider more relevant