united nations educational, scientific and cultural organization

international atomic energy agency

the
abdus salam
international centre for theoretical physics

SMR/1423 - 19

SCHOOL ON
"STATISTICAL PHYSICS, PROBABILITY THEORY AND COMPUTATIONAL COMPLEXITY"

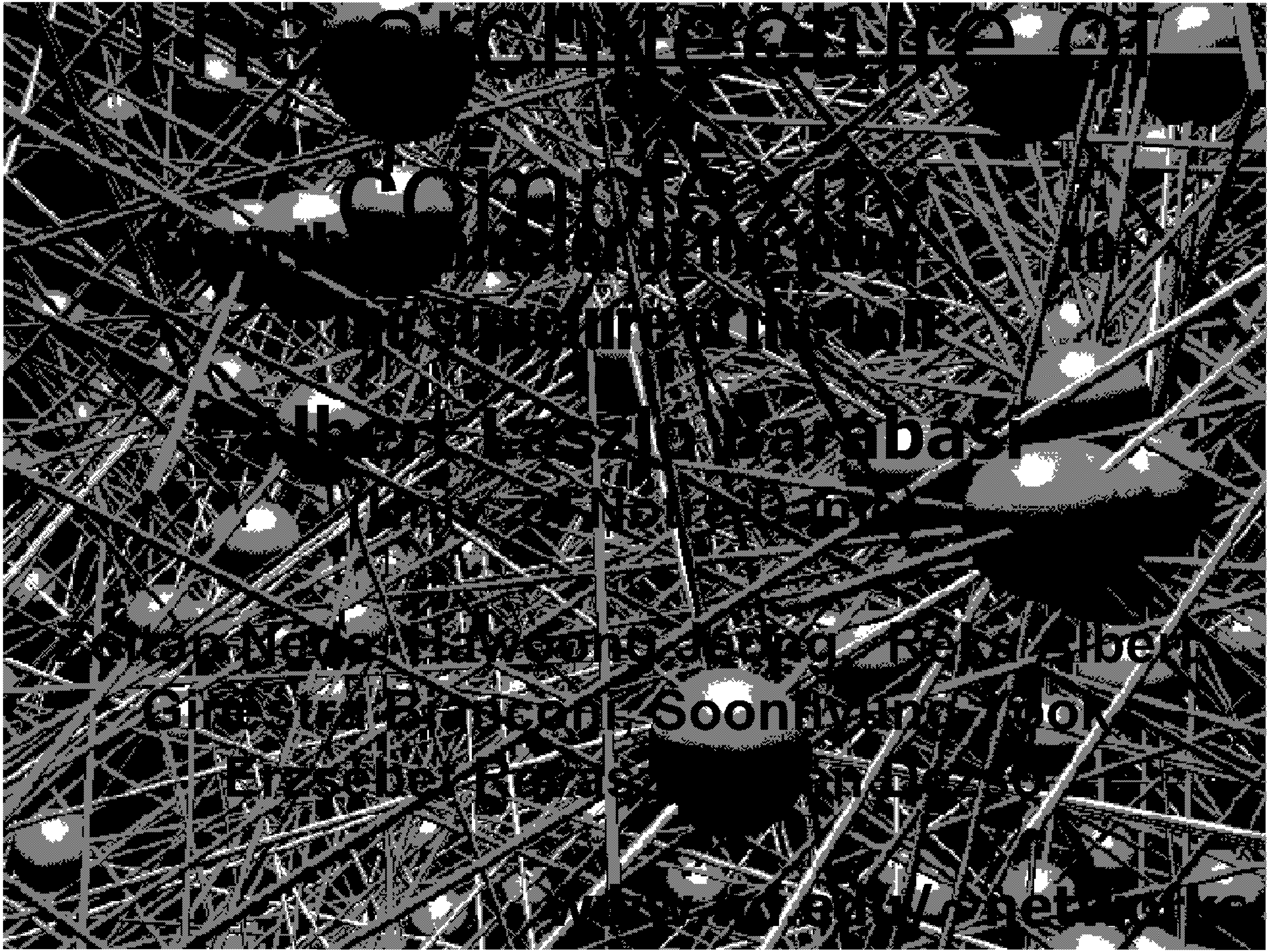
(26 August - 4 September 2002)

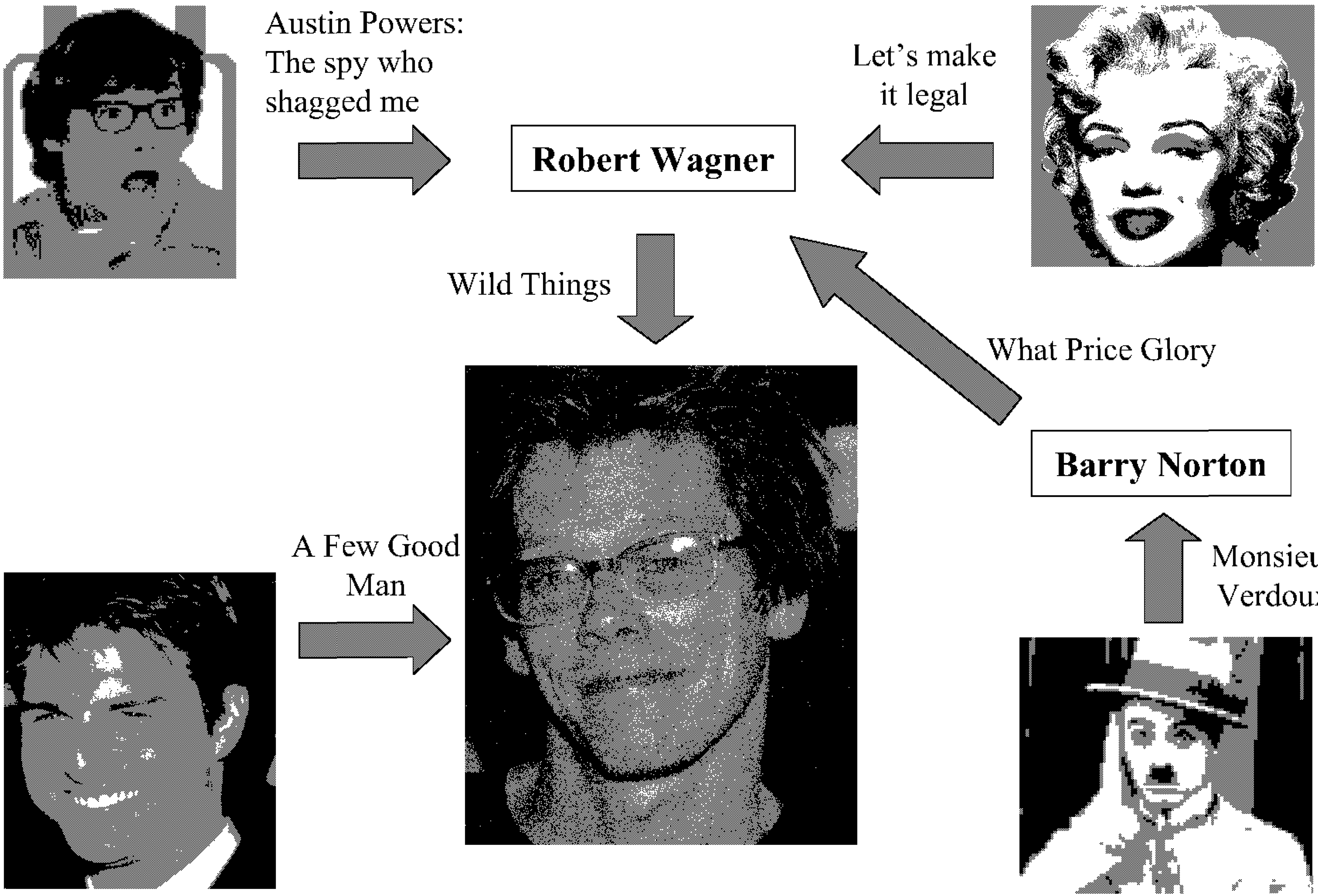
" Internet Graphs and their Generation "

presented by:

A.L. Barabasi
University of Notre Dame
U.S.A.

strada costiera, 11 - 34014 trieste italy - tel. +39 0402240111 fax +39 040224163 - sci_info@ictp.trieste.it - www.ictp.trieste.it


The architecture of
complexity
Albert-László Barabási
Notre Dame
Zoltán Néda, Hawoong Jeong, Réka Albert
Ginestra Bianconi, Soon-Hyung Yook
Erzsébet Ravasz


Austin Powers: The spy who shagged me
Robert Wagner
Let's make it legal
Wild Things
What Price Glory
Barry Norton
A Few Good Man
Monsieur Verdoux

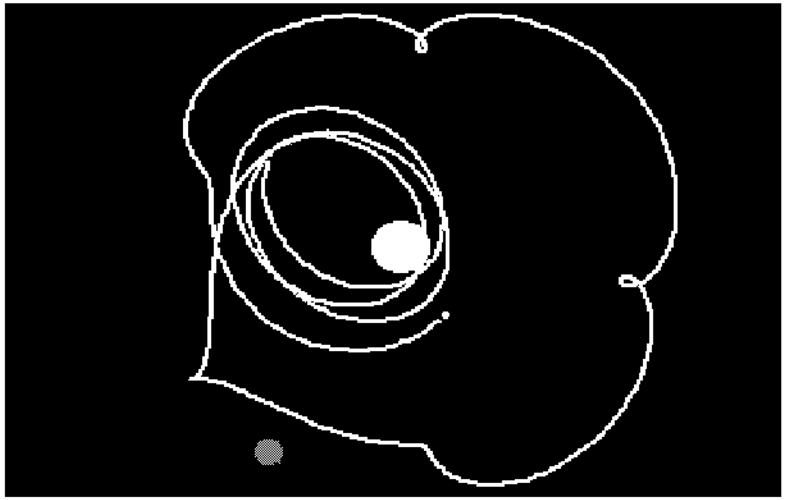
What is Complexity?

A popular paradigm: Simple systems display complex behavior

- non-linear systems
- chaos
- fractals

3 Body Problem

Earth( ) Jupiter ( ) Sun ( )



Main Entry: [1]**com·plex**
Function: *noun*
Etymology: Late Latin *complexus* totality, from Latin, embrace, from *complecti*
Date: 1643
1 : a whole made up of complicated or interrelated parts

Society

Nodes: individuals

Links: social relationship (family/work/friendship/etc.)

S. Milgram (1967)

John Guare

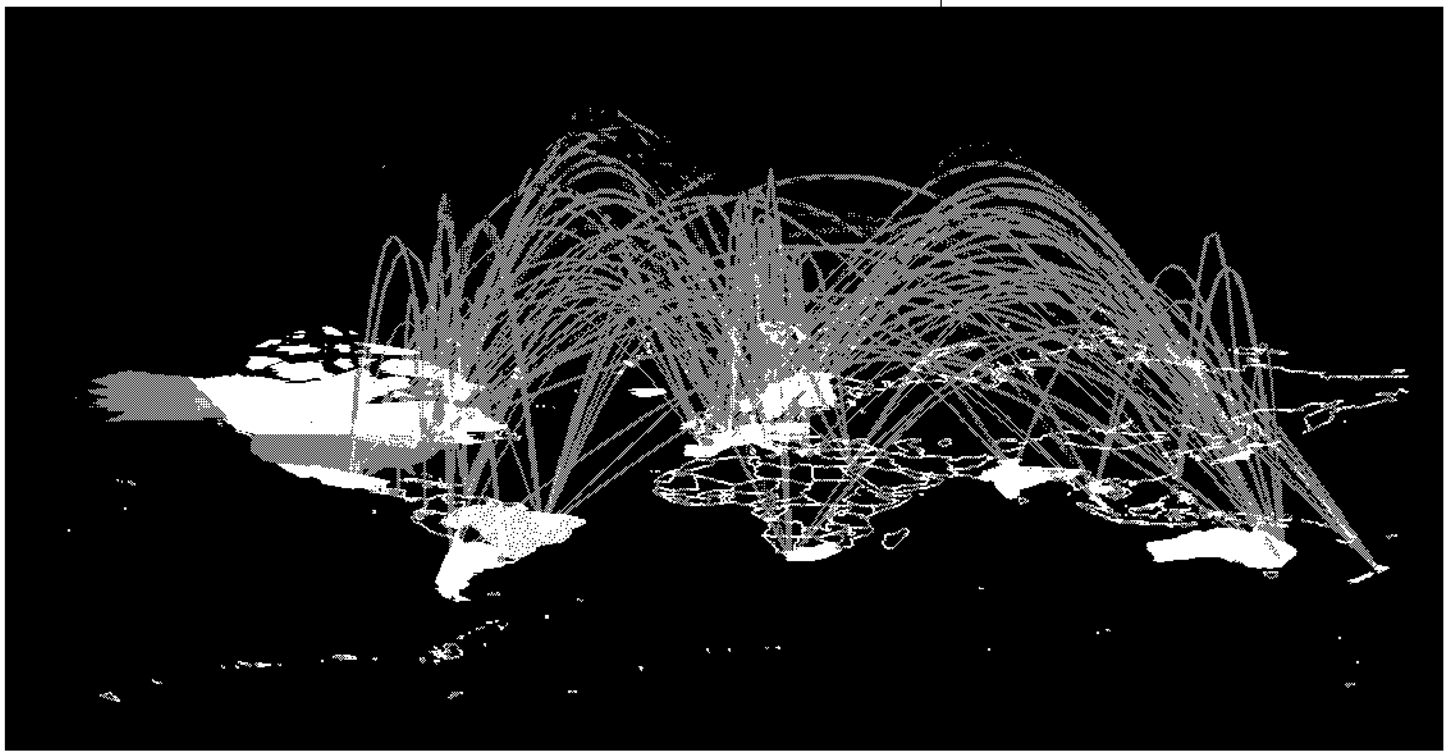
Six Degrees of Separation

Social networks: Many individuals with diverse social interactions between them.

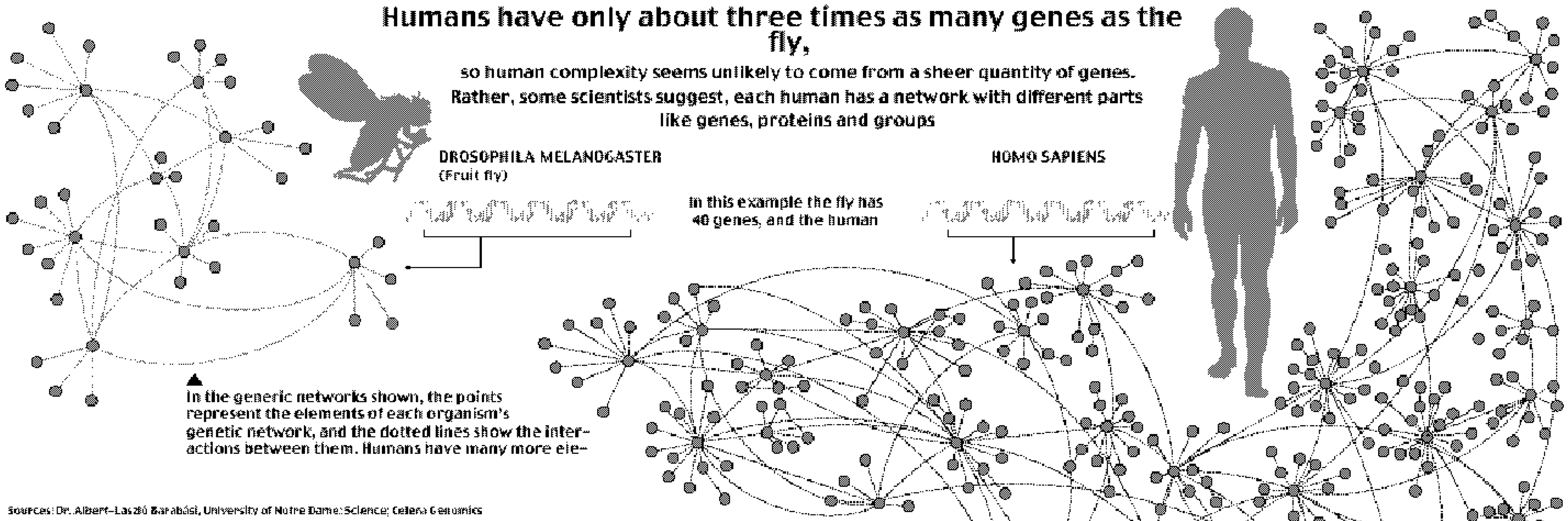
Communication networks

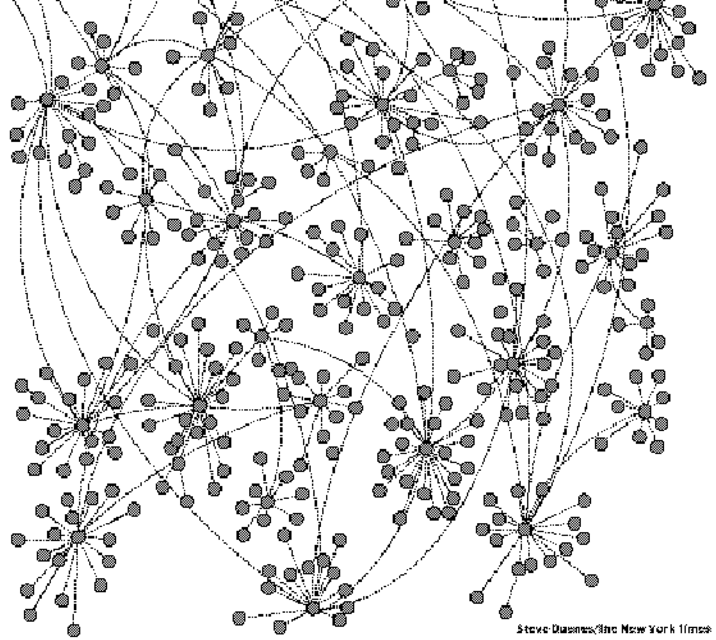
The Earth is developing an electronic nervous system, a network with diverse **<u>nodes</u>** and **<u>links</u>** are

nodes	links
-computers	-phone lines
-routers	-TV cables
-satellites	-EM waves



Communication networks: Many <u>non-identical</u> components with <u>diverse connections</u> between them.



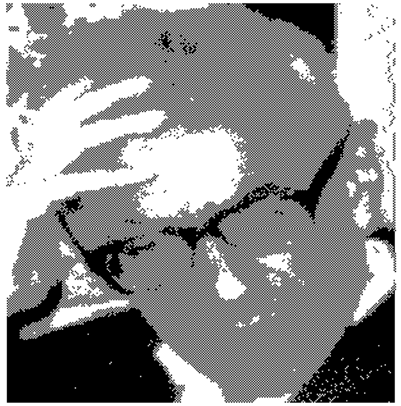


Complex systems

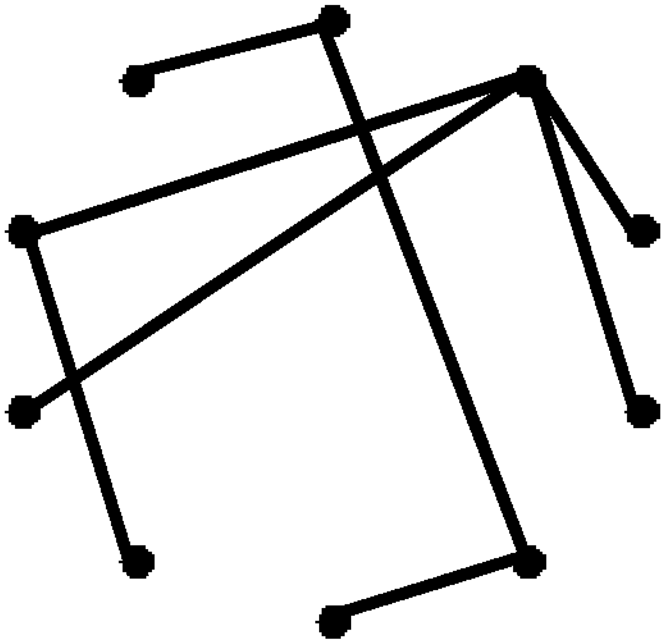
Made of many non-identical **elements** connected by diverse **interactions**.

NETWORK

Erdös-Rényi model
(1960)

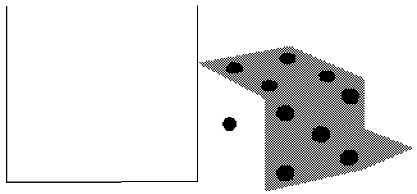


Pál Erdös
(1913-1996)



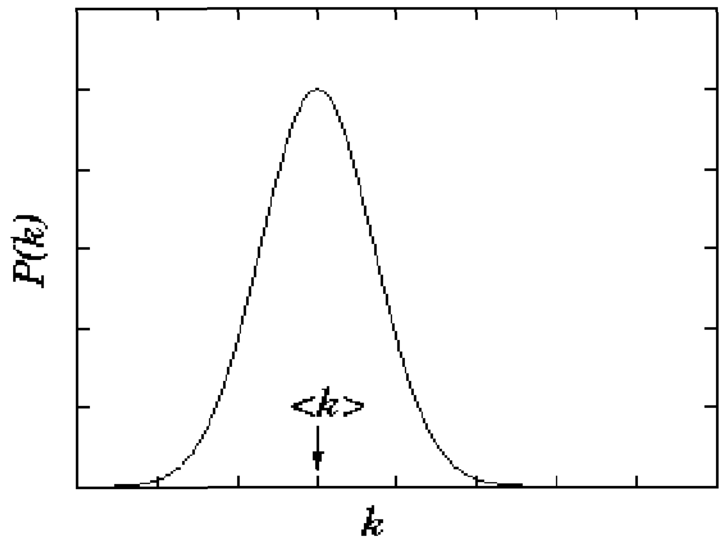
Connect with probability p

p=1/6
N=10 ⟨k⟩
~ 1.5



- **Democratic**
- **Random**

Poisson distribution



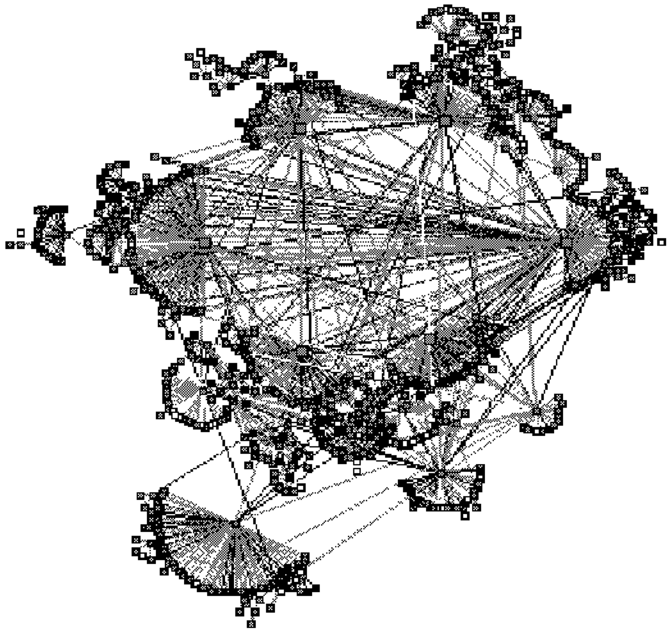
ARE COMPLEX NETWORKS REALLY RANDOM?

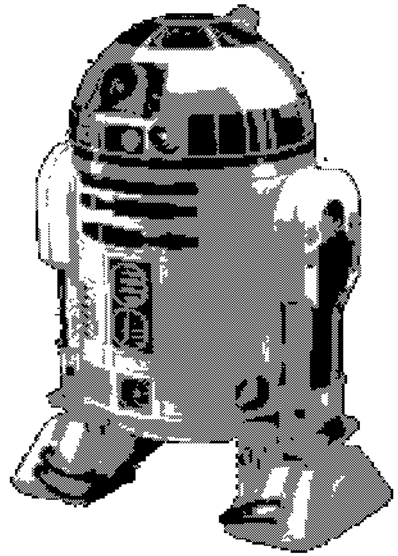
NO!

World Wide Web

Nodes: WWW documents
Links: URL links

800 million documents (S. Lawrence, 1999)

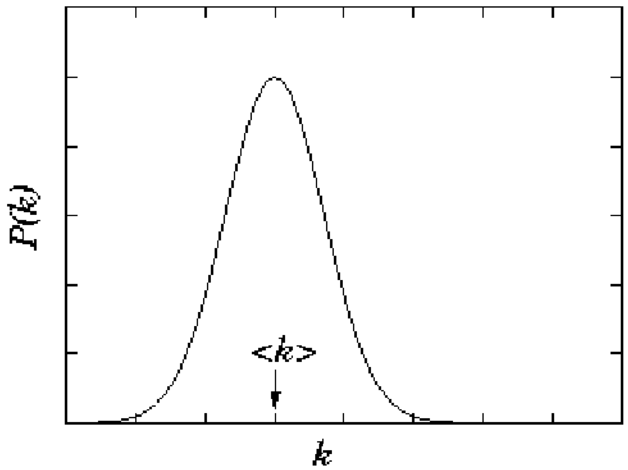




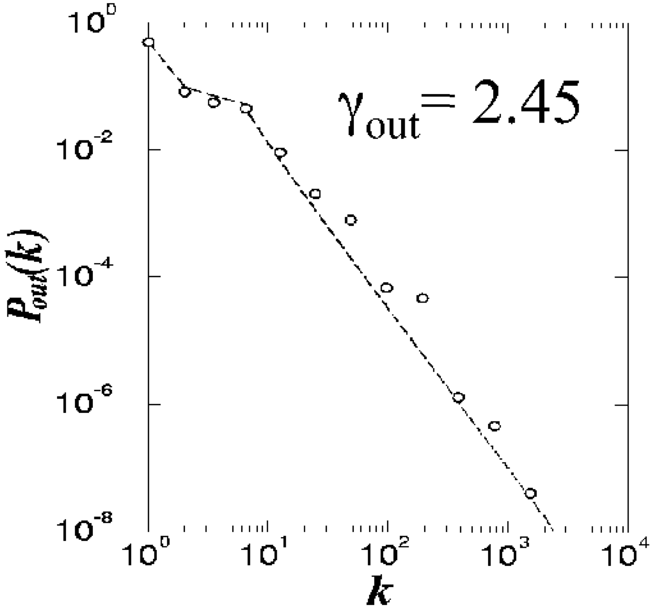
ROBOT: **collects all URL's found in a document and follows them recursively**

R. Albert, H. Jeong, A-L Barabasi, Nature, **401** 130 (1999)

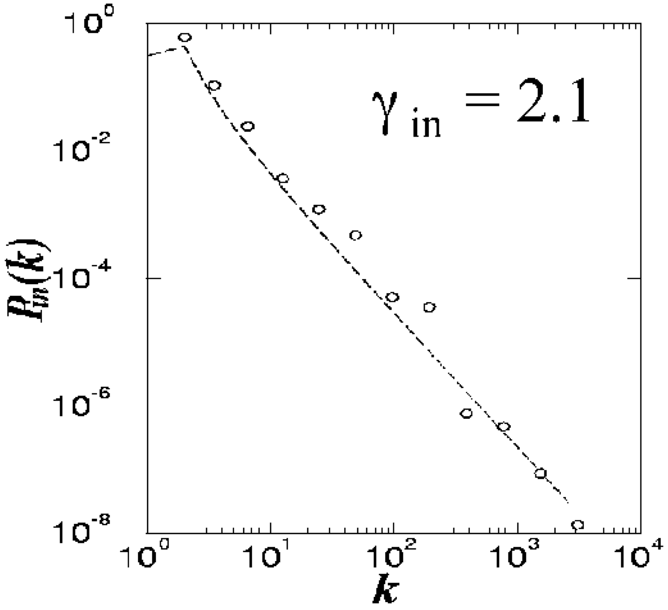
What did we expect?



We find:

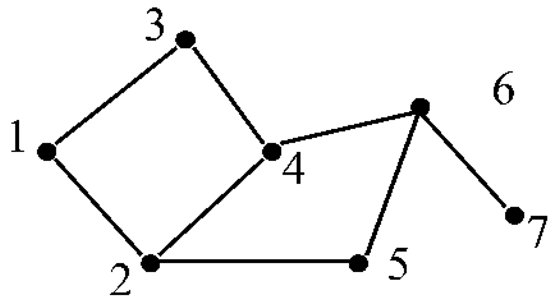


$P_{out}(k) \sim k^{-\gamma out}$



$P_{in}(k) \sim k^{-\gamma in}$

19 degrees of separation



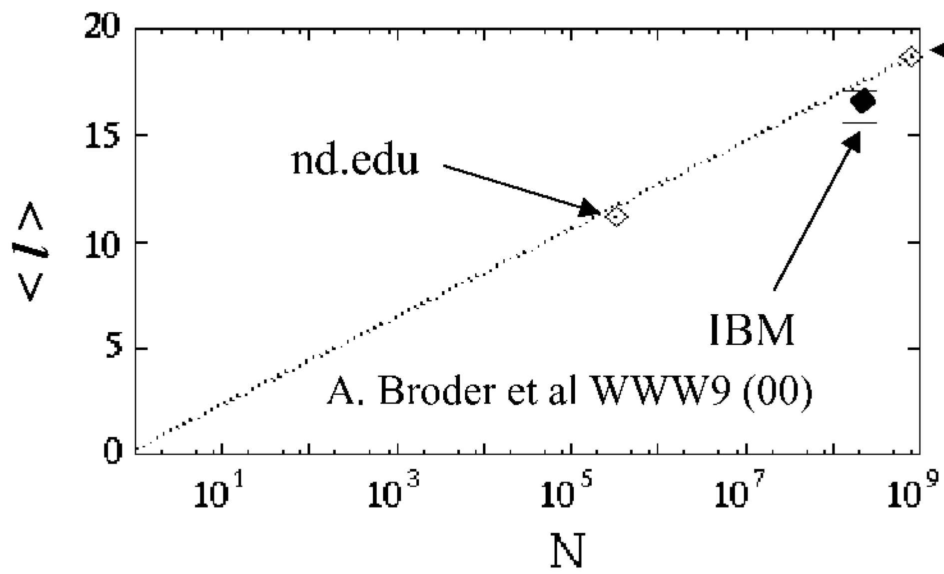
$l_{15} = 2$ [1→2→5]

$l_{17} = 4$ [1→3→4→6 → 7]

… $< l > = ??$

- **Finite size scaling:** create a network with N nodes with $P_{in}(k)$ and $P_{out}(k)$

$$< l > = 0.35 + 2.06 \log(N)$$

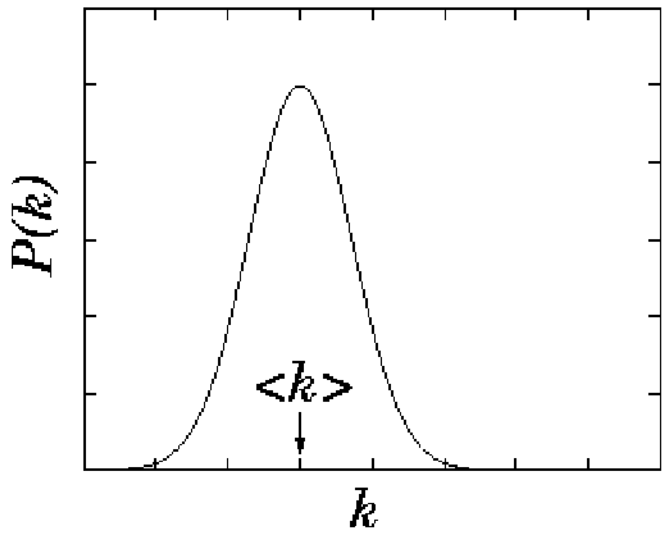


19 degrees of separation R. Albert et al Nature (99)

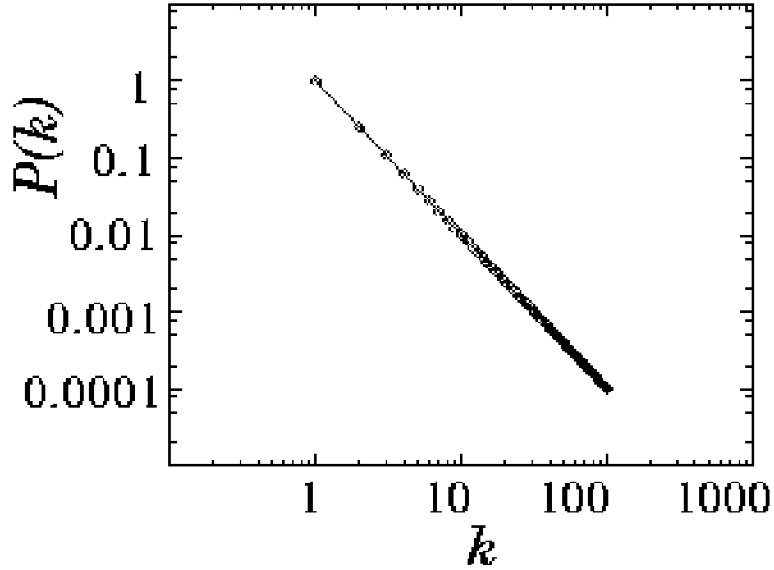
based on 800 million webpages [S. Lawrence et al Nature (99)]

What does it mean?

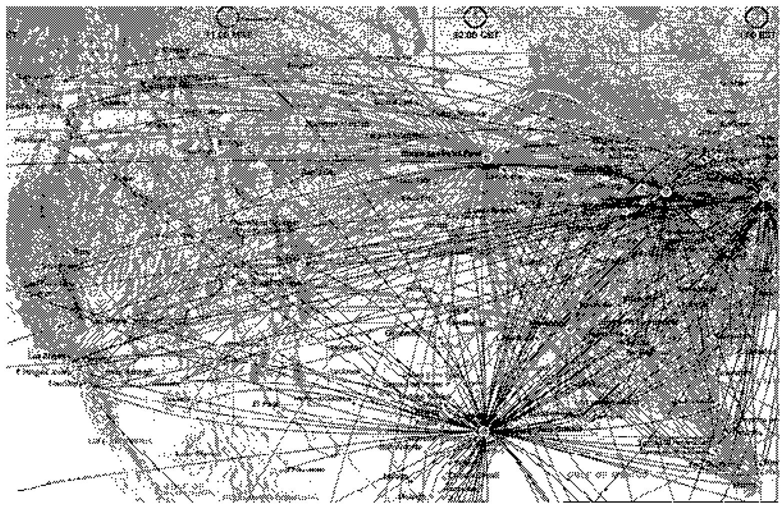
Poisson distribution



Power-law distribution



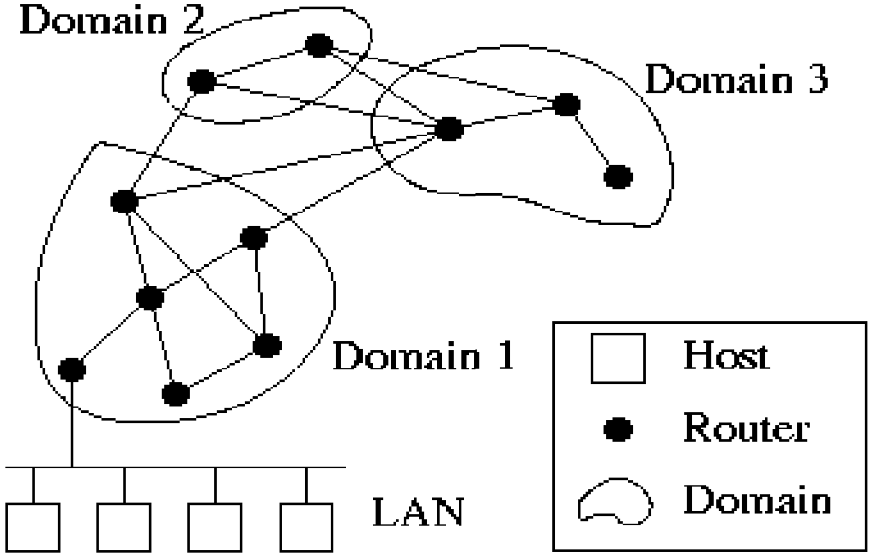
Exponential Network

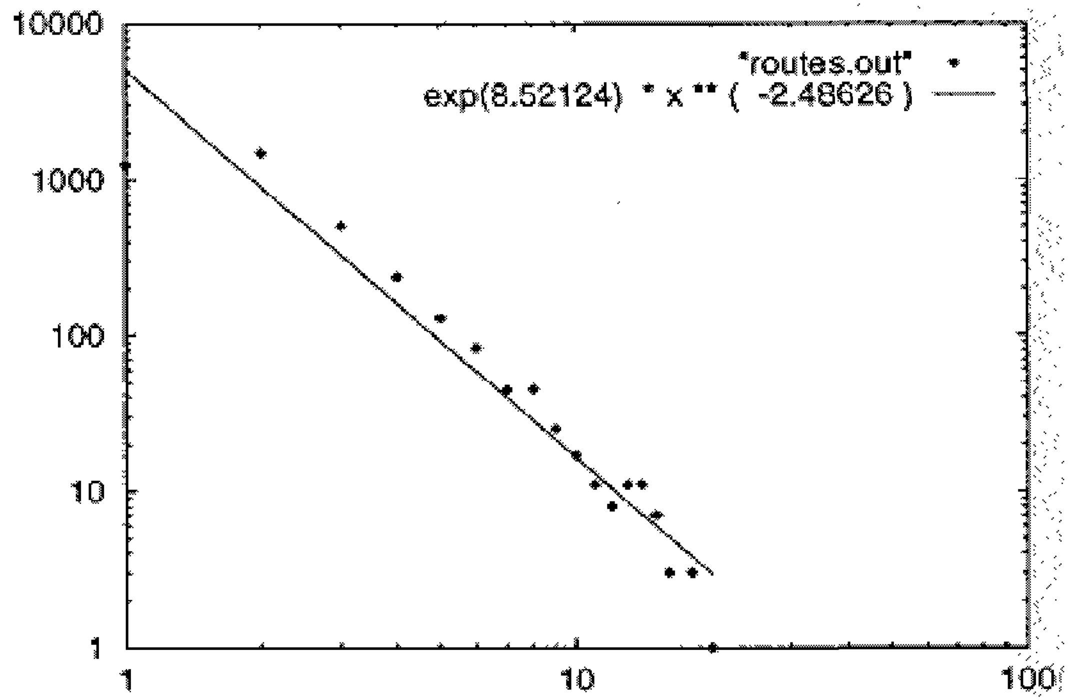


Scale-free Network

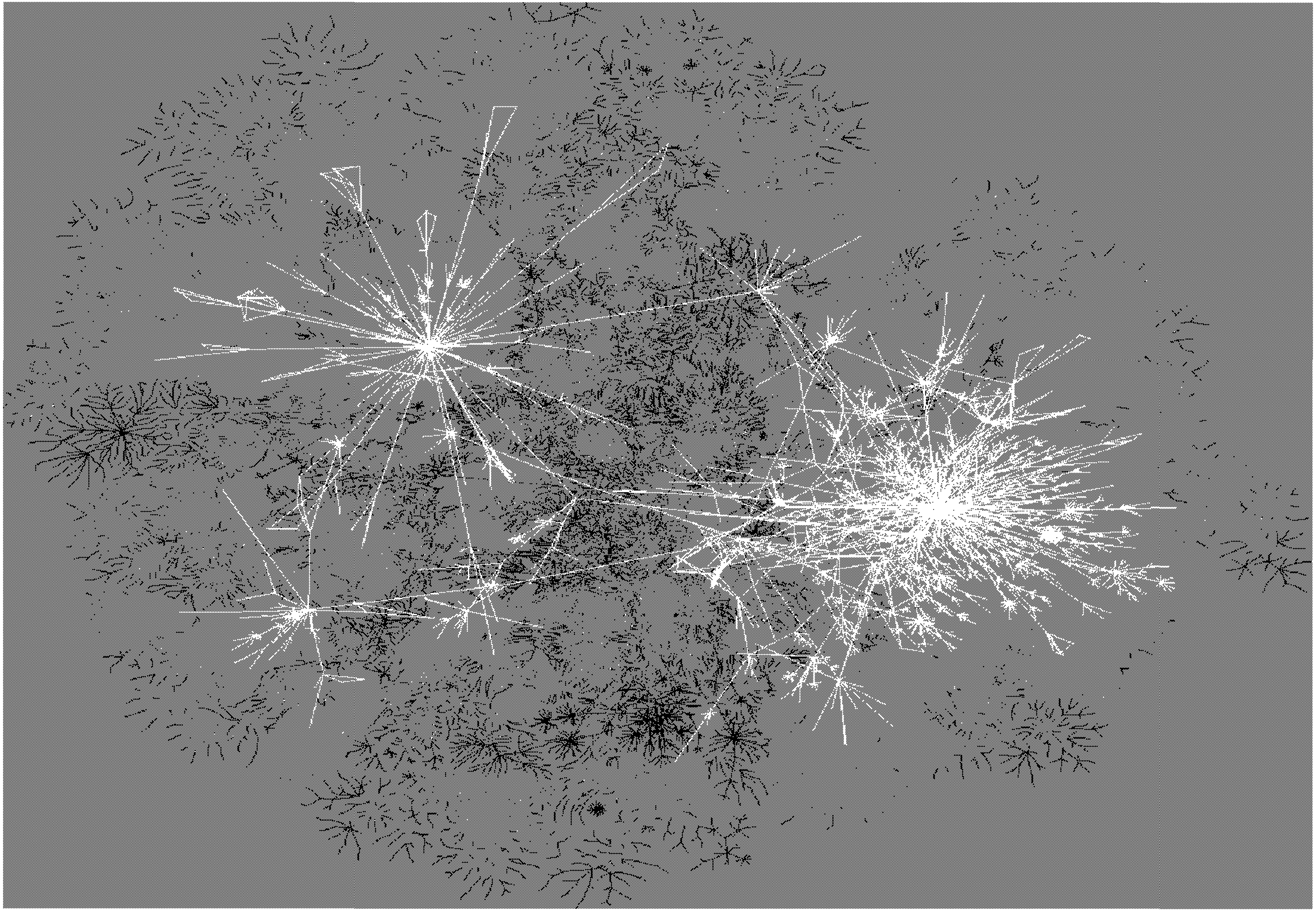
INTERNET BACKBONE

Nodes: computers, routers
Links: physical lines





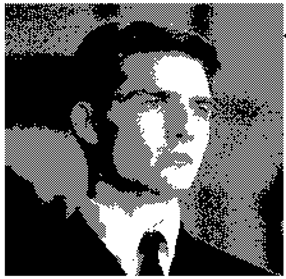
(Faloutsos, Faloutsos and Faloutsos, 1999)



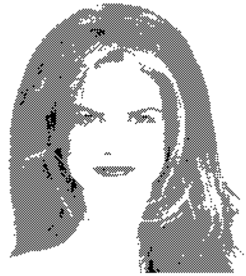
ACTOR CONNECTIVITIES

Nodes: actors **Links**: cast jointly

IMDb Internet Movie Database



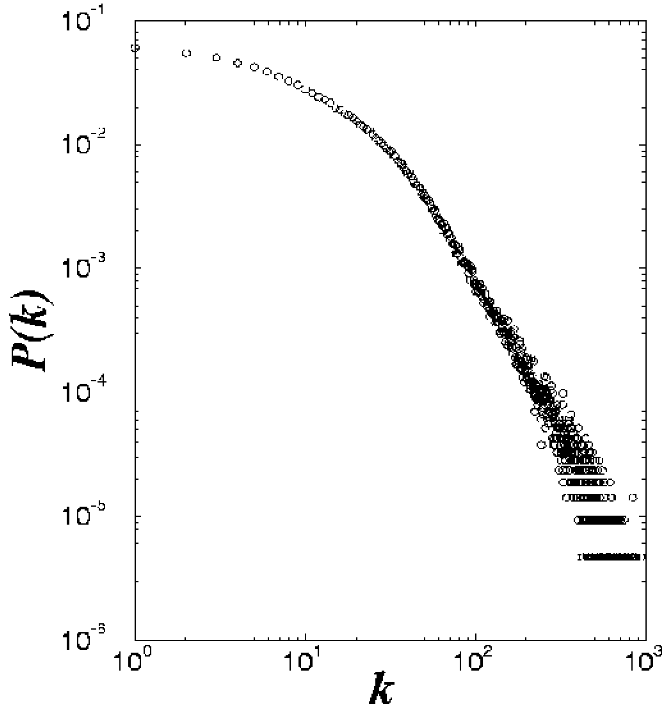
Days of Thunder (1990)
Far and Away (1992)
Eyes Wide Shut (1999)



N = 212,250 actors
$\langle k \rangle = 28.78$

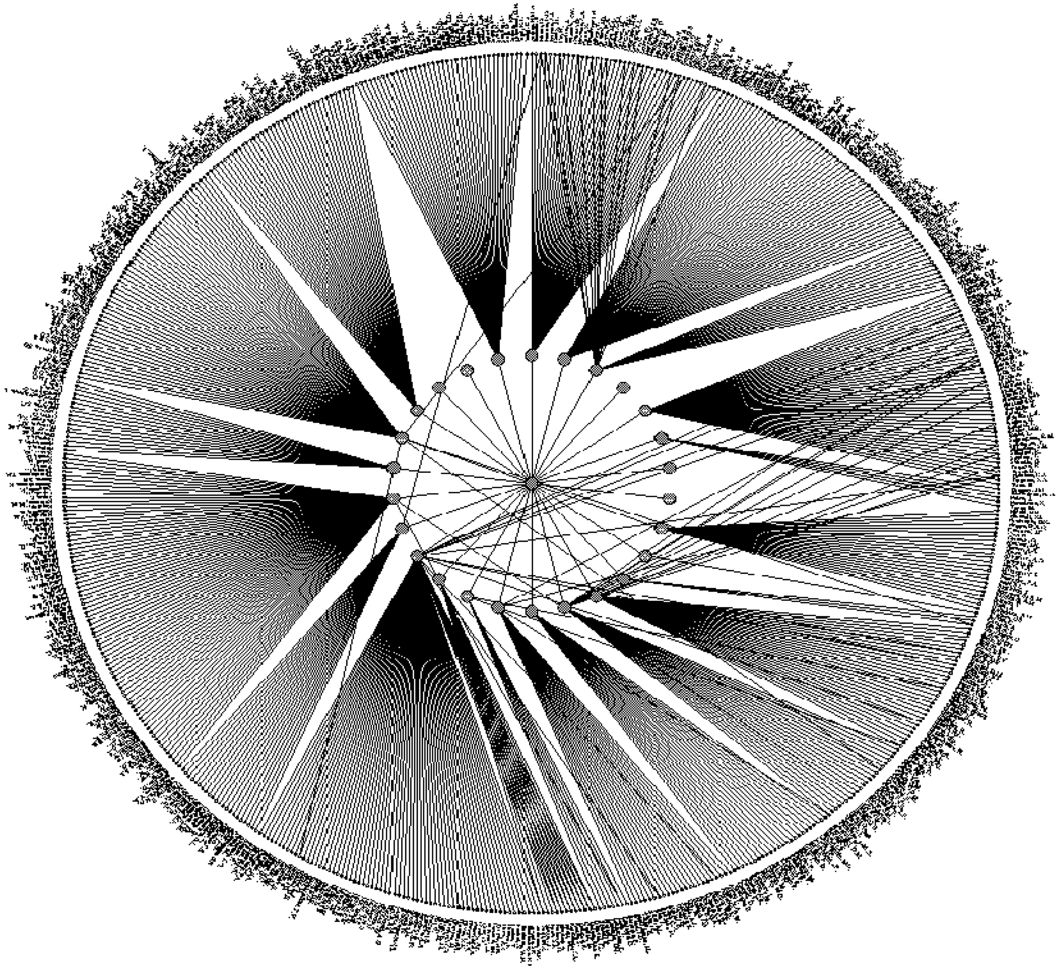
$P(k) \sim k^{-\gamma}$

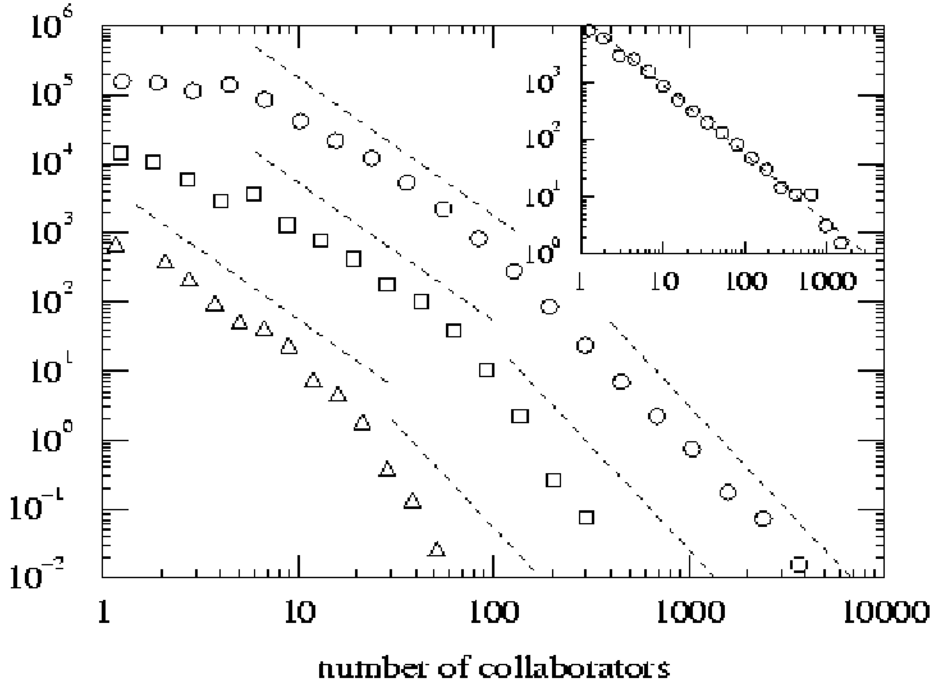
$\gamma = 2.3$



SCIENCE COAUTHORSHIP

Nodes: scientist (authors)
Links: write paper together



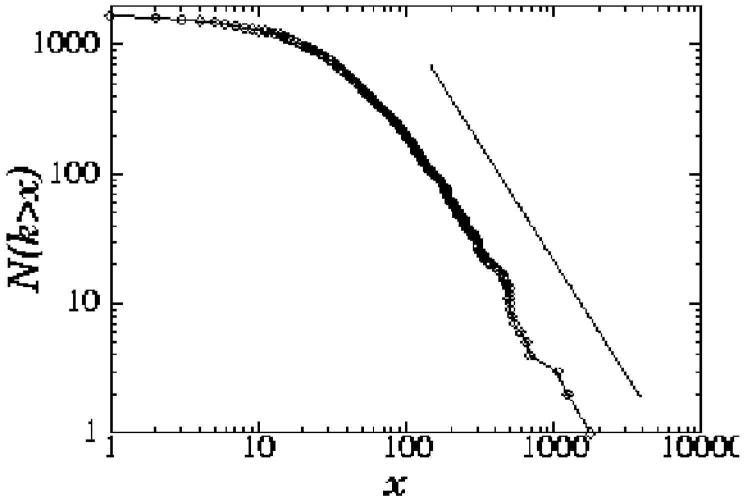


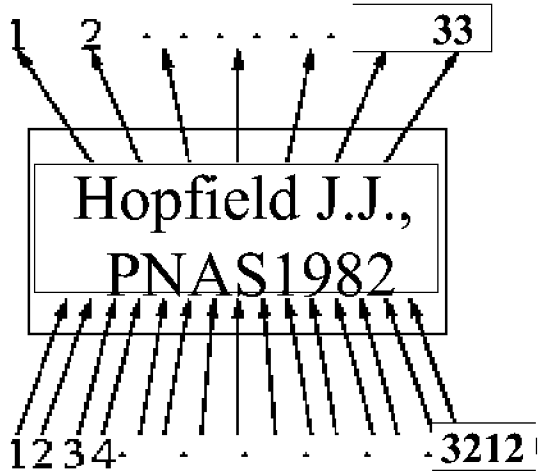
(Newman, 2000, H. Jeong et al 2001)

SCIENCE CITATION INDEX

Nodes: papers
Links: citations

1736 PRL papers (1988)





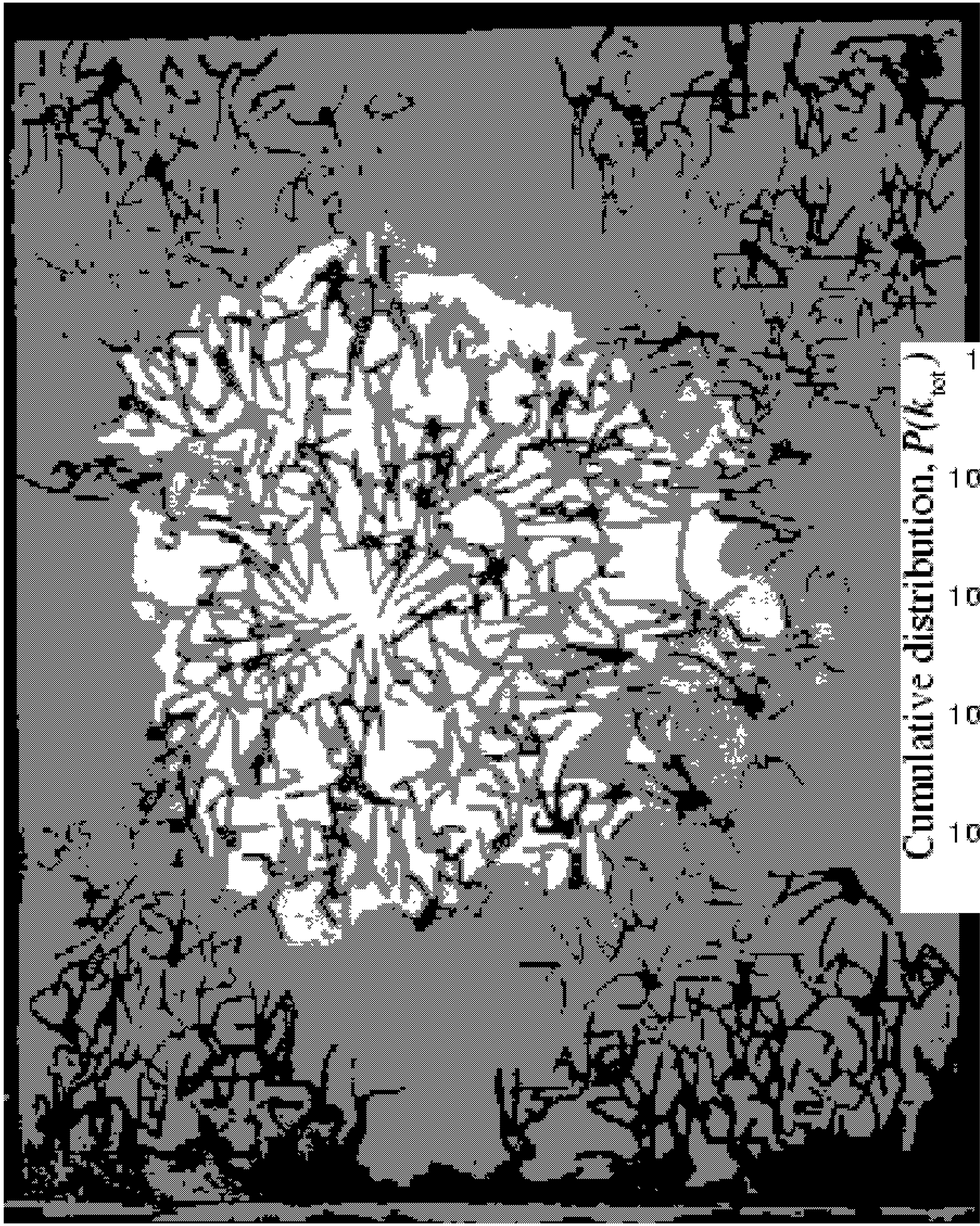
$$P(k) \sim k^{-\gamma}$$

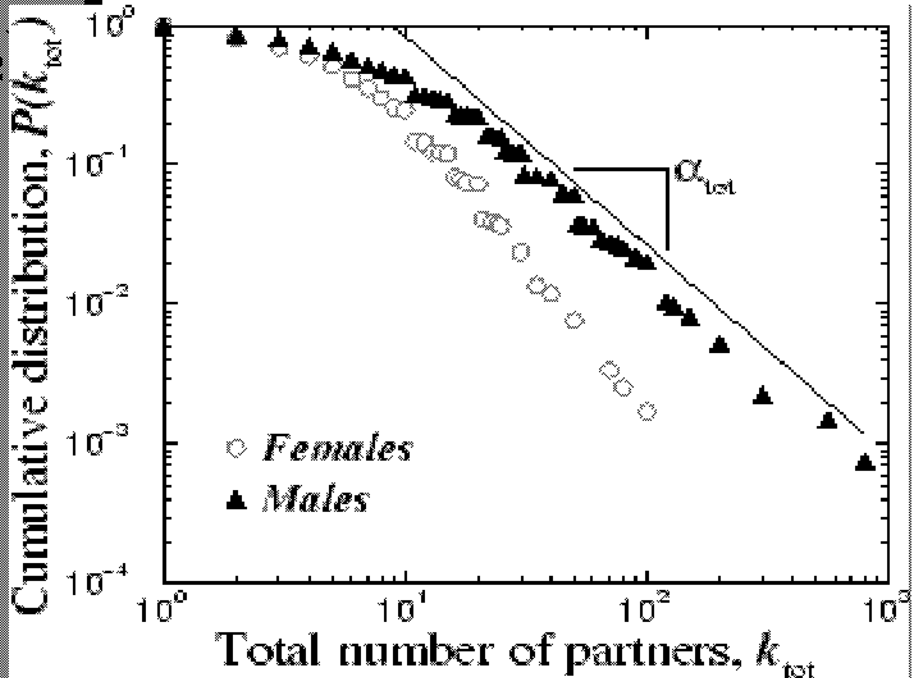
$$(\gamma = 3)$$

(S. Redner, 1998)

Swedish sex-web

Nodes: people (Females; Males)
Links: sexual relationships





4781 Swedes; 18-74;
59% response rate.

Liljeros et al. Nature 2001

Many real world networks have the same internal structure:

Scale-free networks

Why?

SCALE-FREE NETWORKS

(1) The number of nodes (N) is NOT fixed.

Networks continuously expand by the addition of new nodes

Examples:
WWW : addition of new documents
Citation : publication of new papers

(2) The attachment is NOT uniform.

A node is linked with higher probability to a node that already has a large number of links.

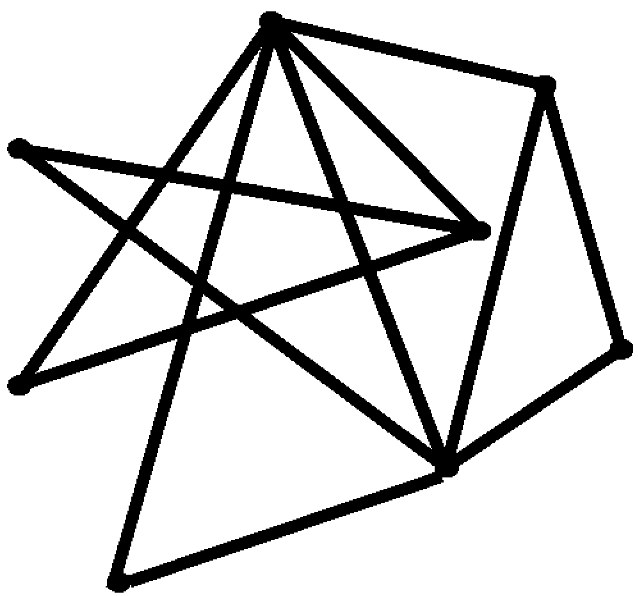
Examples :
WWW : new documents link to well known sites (CNN, YAHOO, NewYork Times, etc)
Citation : well cited papers are more likely to be cited again

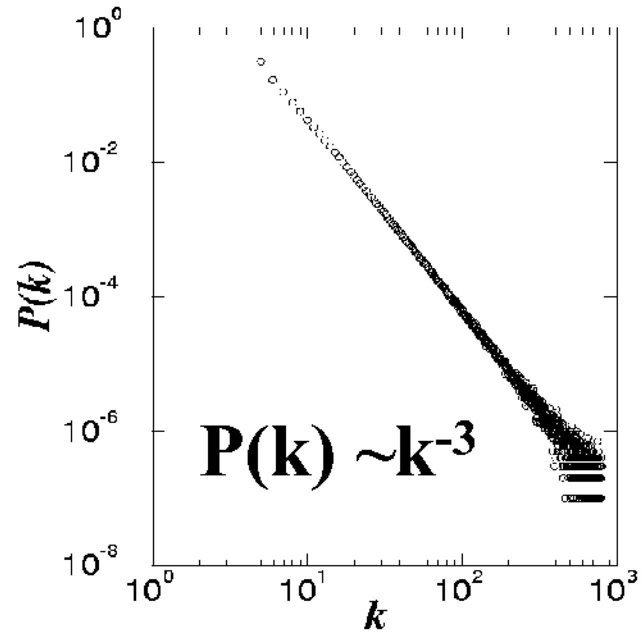
Scale-free model

(1) GROWTH : At every timestep we add a new node with *m* edges (connected to the nodes already present in the system).

(2) PREFERENTIAL ATTACHMENT :
The probability _ that a new node will be connected to node *i* depends on the connectivity k_i of that node

$$\Pi(k_i) = \frac{k_i}{\Sigma_j k_j}$$



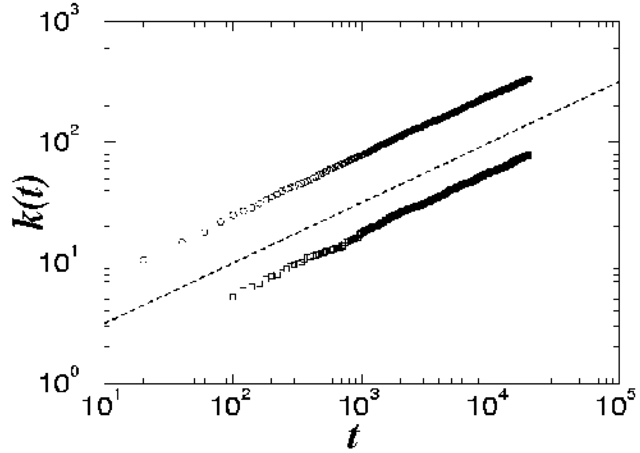


A.-L.Barabási, R. Albert, Science **286,** 509 (1999)

Mean Field Theory

$$\frac{\partial k_i}{\partial t} \propto \Pi(k_i) = A\frac{k_i}{\sum_j k_j} = \frac{k_i}{2t}$$, with initial condition $k_i(t_i) = m$

$$\boxed{k_i(t) = m\sqrt{\frac{t}{t_i}}}$$



$$P(k_i(t) < k) = P_t(t_i > \frac{m^2 t}{k^2}) = 1 - P_t(t_i \le \frac{m^2 t}{k^2}) = 1 - \frac{m^2 t}{k^2(m_0 + t)}$$

$$\therefore P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^2 t}{m_o + t}\frac{1}{k^3} \sim k^{-3}$$

$$\boxed{\_ = 3}$$

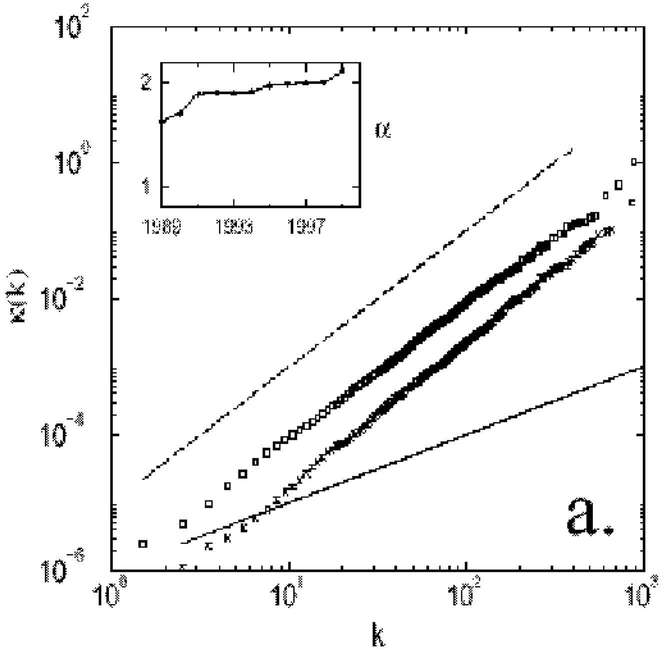
A.-L.Barabási, R. Albert and H. Jeong, Physica A **272,** 173 (1999)

Preferential Attachment

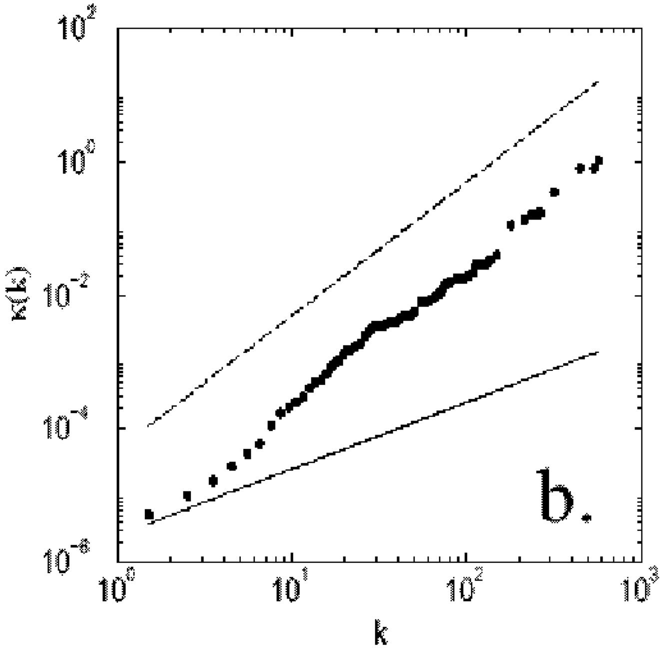
$$\frac{\partial k_i}{\partial t} \propto \Pi(k_i) \sim \frac{\Delta k_i}{\Delta t}$$

For given $\Delta t, \therefore \Delta k \propto \Pi(k)$

k vs. Δk : increase in the No. of links in a unit time



Citation network

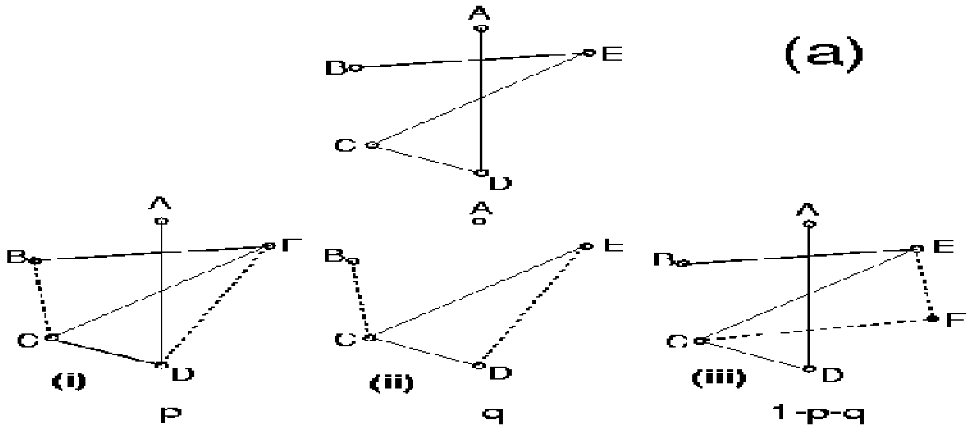


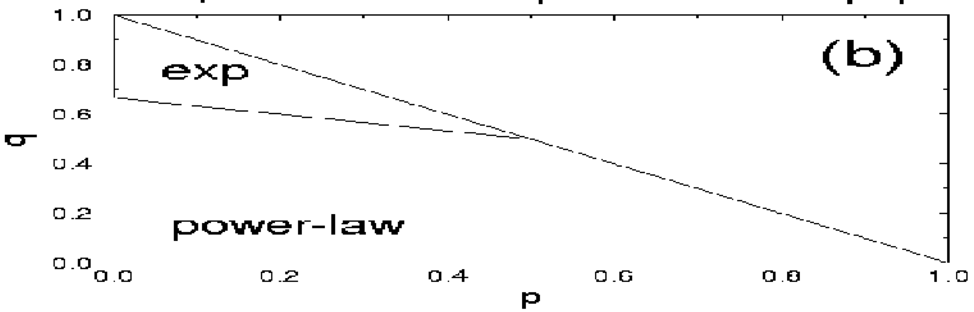
Internet

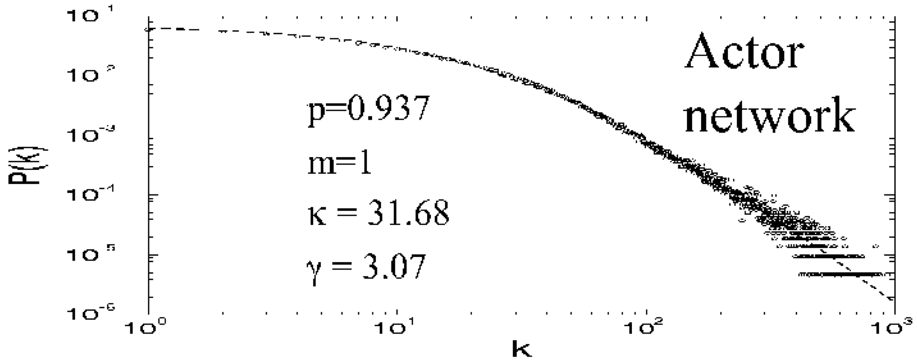
(Jeong, Neda, A.-L. B, cond-mat/0104131)

Universality?

WWW (in)	Internet	Actor	Citation index	Sex Web	Cellular network	Phone call network	linguistics
$\gamma = 2.1$	$\gamma = 2.5$	$\gamma = 2.3$	$\gamma = 3$	$\gamma = 3.5$	$\gamma = 2.1$	$\gamma = 2.1$	$\gamma = 2.8$







Extended Model

- prob. p : internal links
- prob. q : link deletion
- prob. 1-p-q : add node

$$P(k) \sim (k+\kappa(p,q,m))^{-\gamma(p,q,m)}$$

$$\gamma \in [1,\infty)$$

- Predict the network topology from microscopic processes with parameters (p,q,m)
- Scaling but no universality

Other Models

• **Non-linear preferential attachment :** $\Pi(k) \sim k^{\alpha} \rightarrow P(k) \sim$ **no scaling for** $\alpha \neq 1$

⇒ $\alpha<1$ **: stretch-exponential**

⇒ $\alpha>1$ **: no-scaling (**$\alpha>2$ **: "gelation")**

(Krapivsky et al (2000).)

• **Initial attractiveness :** $\Pi(k) \sim A+k^{\alpha}$

$\rightarrow P(k) \sim k^{-\gamma}$ **where** $\gamma=2 + A/m$

(Dorogovtsev et al (2000).)

• **Aging : each node has a lifetime → node cannot get links after retirement. (actor)**

⇒ $P(k)$ **: power-law with exponential cutoff**

(Amaral et al (2000).)

Other Models (continued)

• Saturation : each node has maximum link number.

→ node cannot get links after finite # of links

⇒P(*k*) : power-law with exponential cutoff

(Amaral et al (2000).)

•Walking on the network:

Each new node :
→ connects to a randomly selected node
→ with prob. p to nodes the selected node points to
(repeated recursively with the new links)

For p=1: $P(k) \sim k^{-2}$

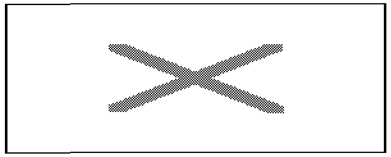
(Vazquez, 2000)

Can Latecomers Make It? *Fitness Model*

SF model: $k(t) \sim t^{-}$ → (first mover advantage)

Real systems: nodes compete for links -- ***fitness***

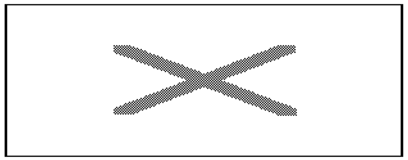
Fitness Model: fitness (η)

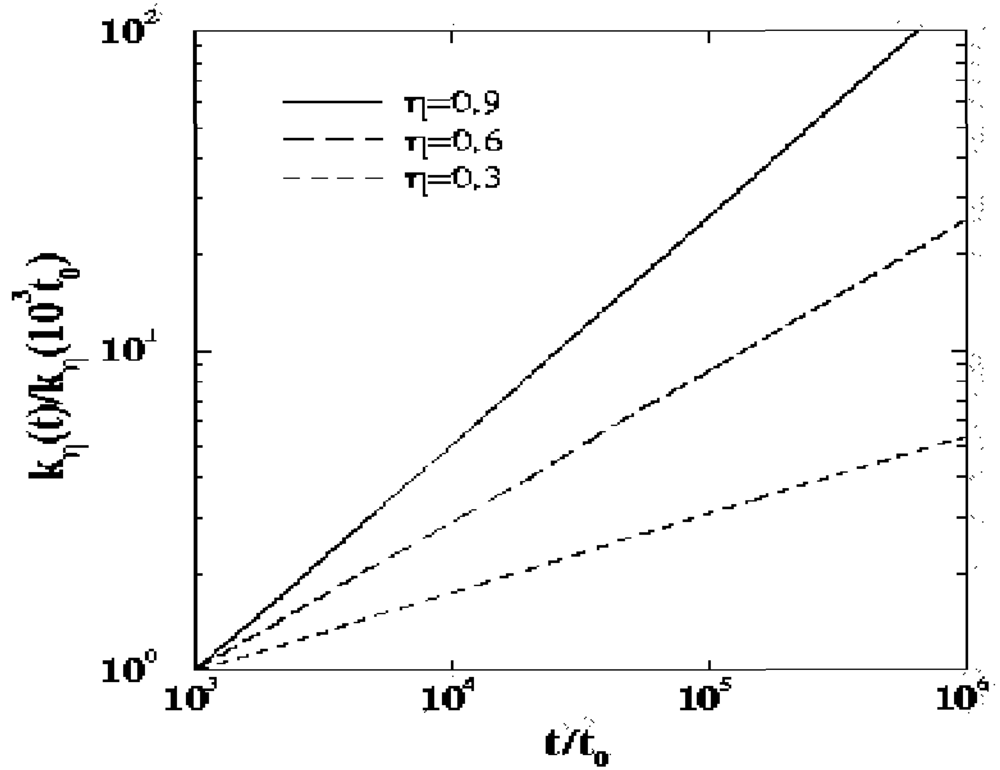


$$\boldsymbol{k(\eta,t) \sim t^{\beta(\eta)}}$$

where

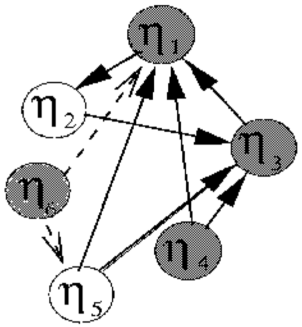
$\beta(\eta) = \eta/C$

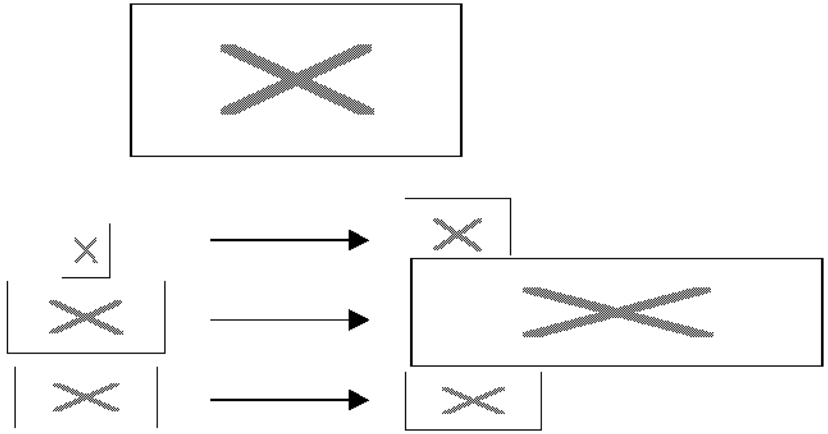




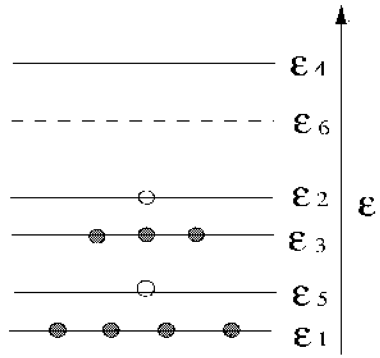
Bose-Einstein Condensation in Evolving Networks

Network

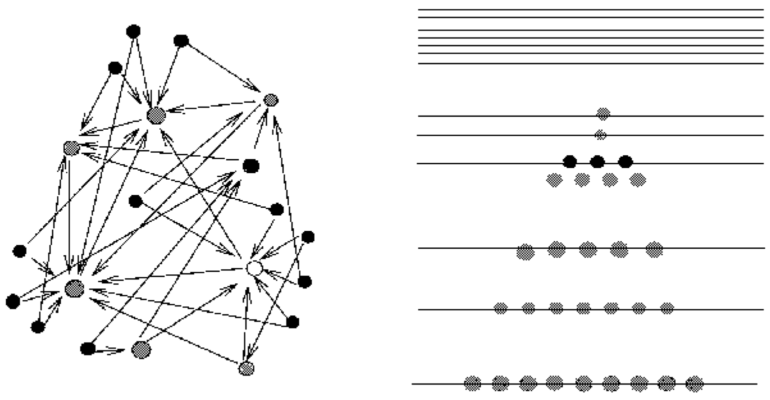




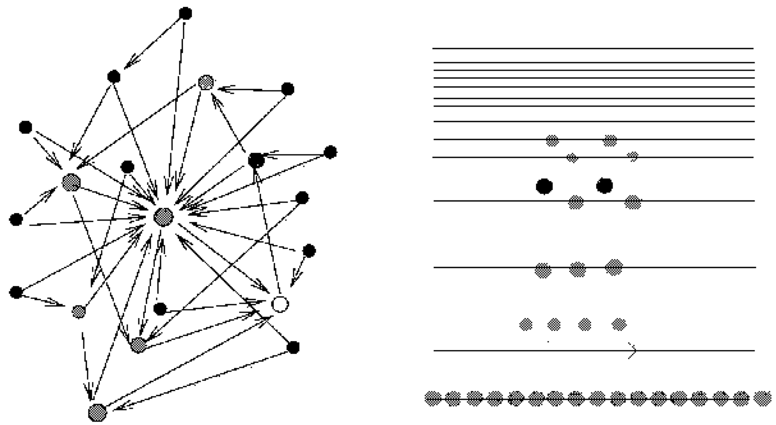
Bose gas



Fit-gets-rich

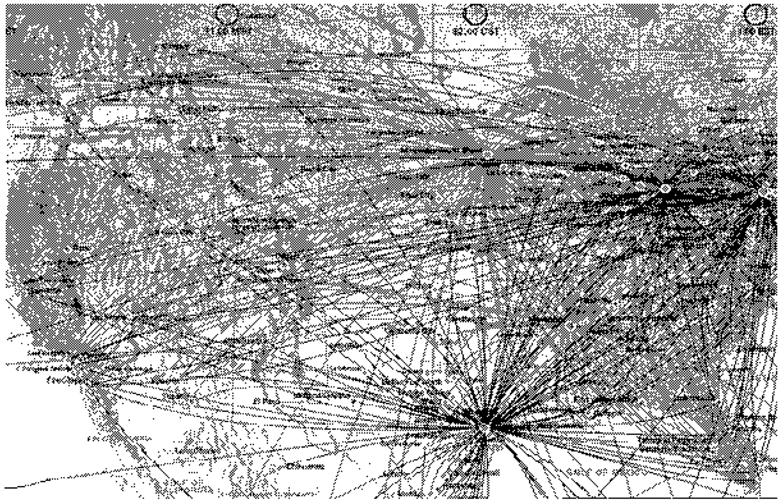


Bose-Einstein condensation



G. Bianconi and A.-L. Barabási, Physical Review Letters 2001; cond-mat/0011029

What is the topology of cellular networks?

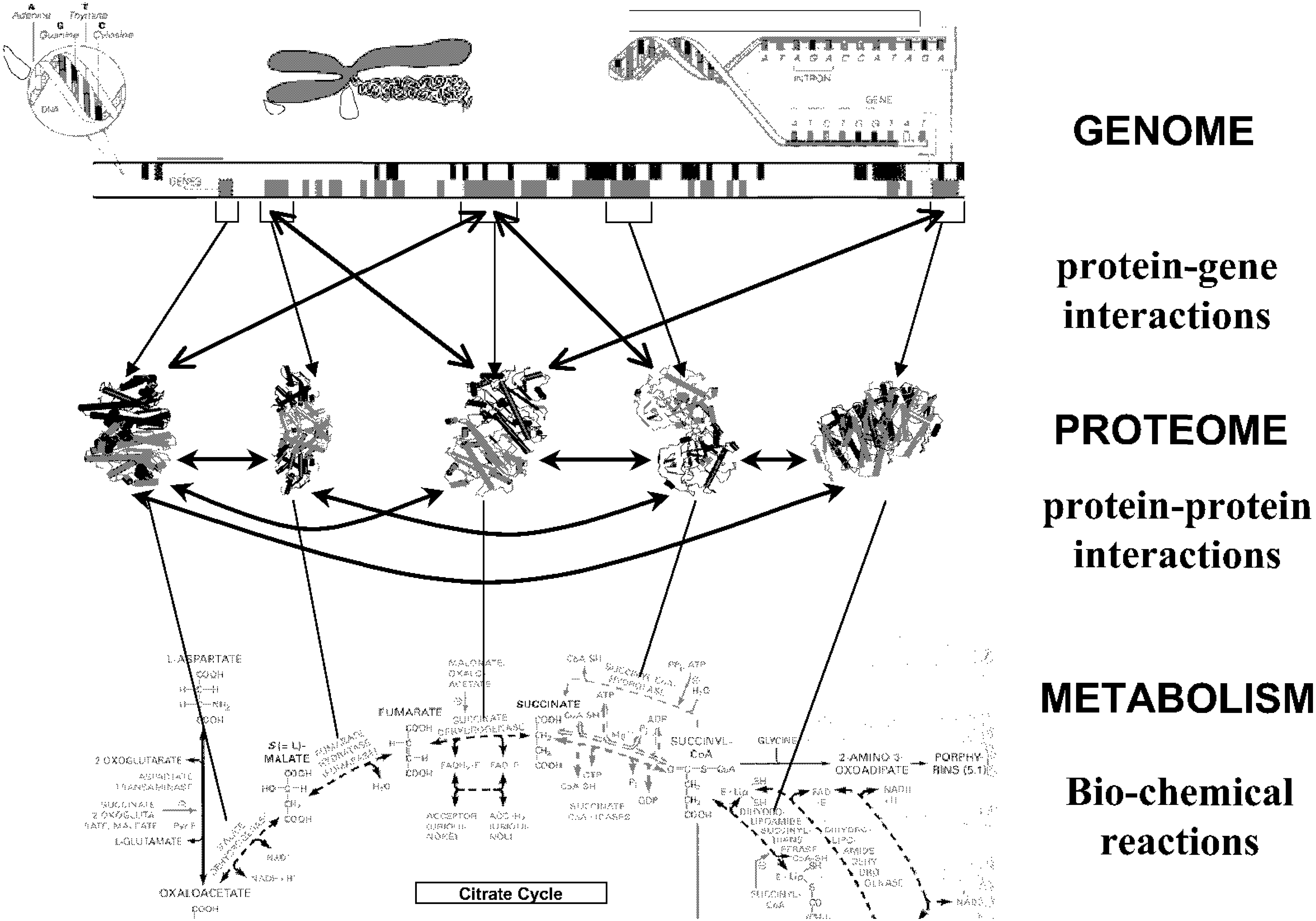


Argument 1:
Cellular networks are scale-free!

Reason: They formed one node at a time…

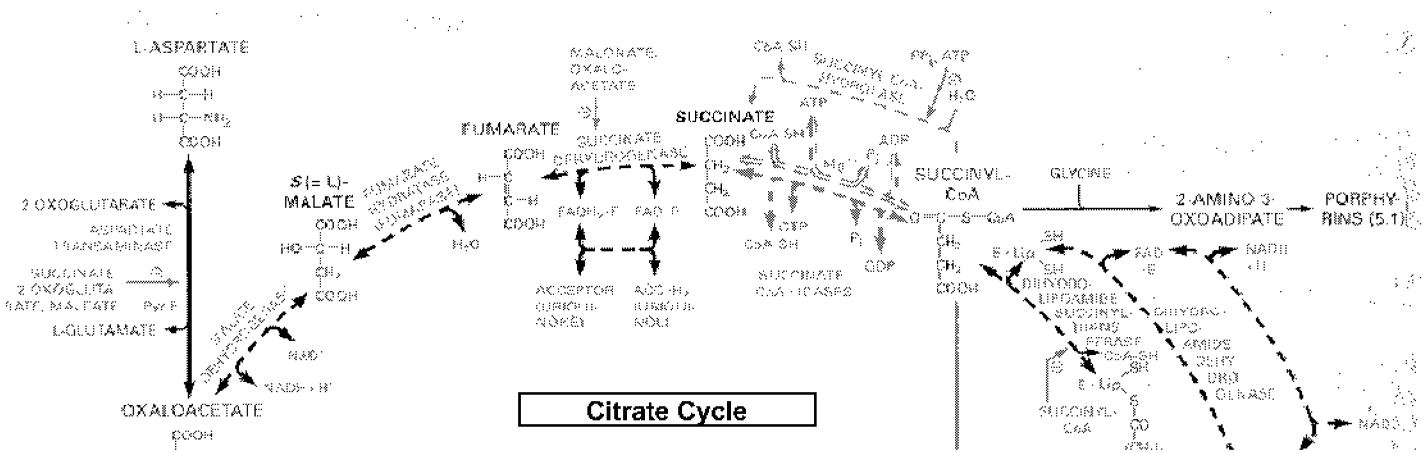
Argument 2:
Cellular networks are exponential!

Reason: They have been streamlined by evolution...

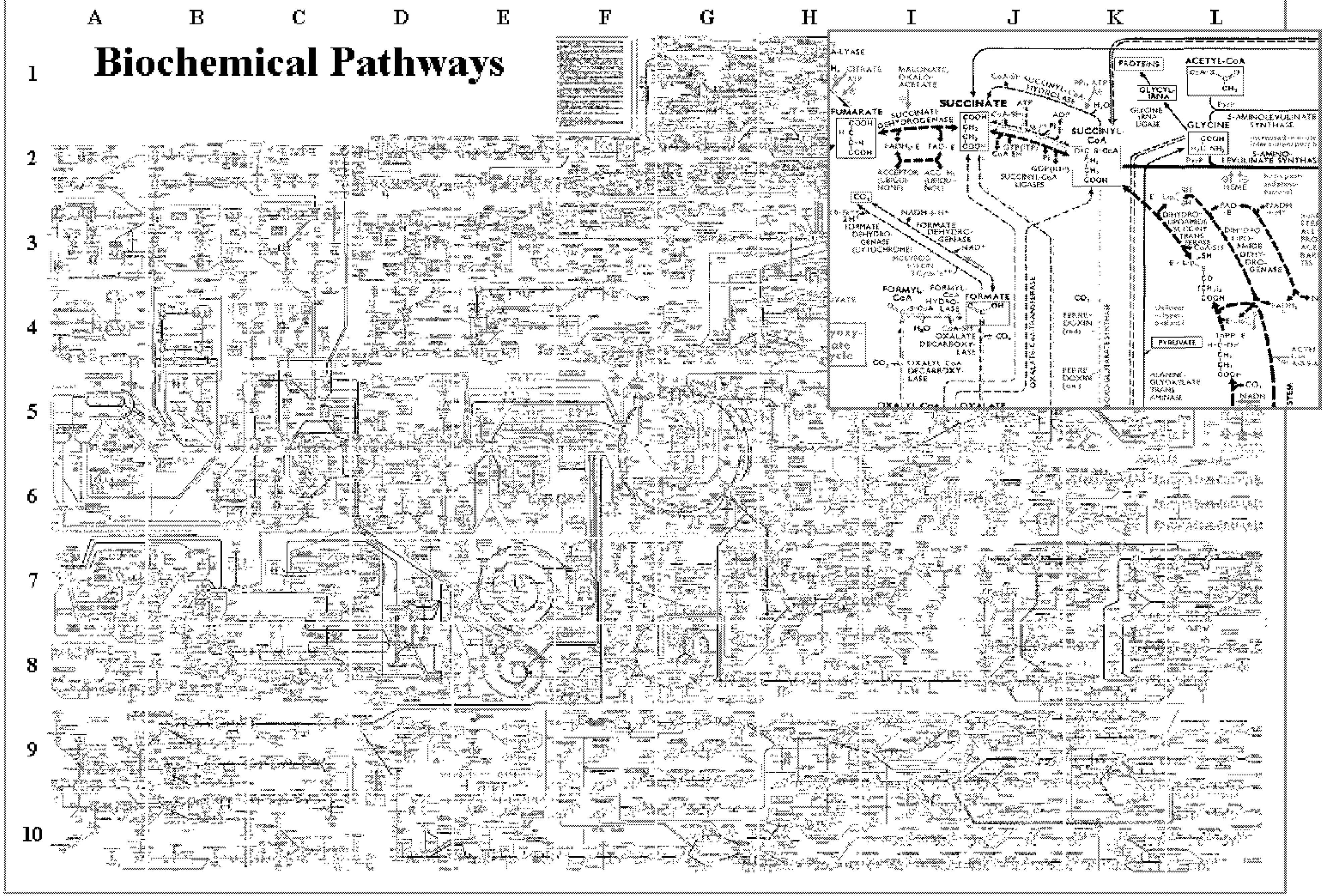

GENOME
protein-gene interactions
PROTEOME
protein-protein interactions
METABOLISM
Bio-chemical reactions
INTRON
GENE
L-ASPARTATE
FUMARATE
SUCCINATE
SUCCINYL-CoA
GLYCINE
2-AMINO-3-OXOADIPATE
PORPHYRINS (5.1)
OXALOACETATE
L-GLUTAMATE
Citrate Cycle

METABOLISM

Bio-chemical reactions



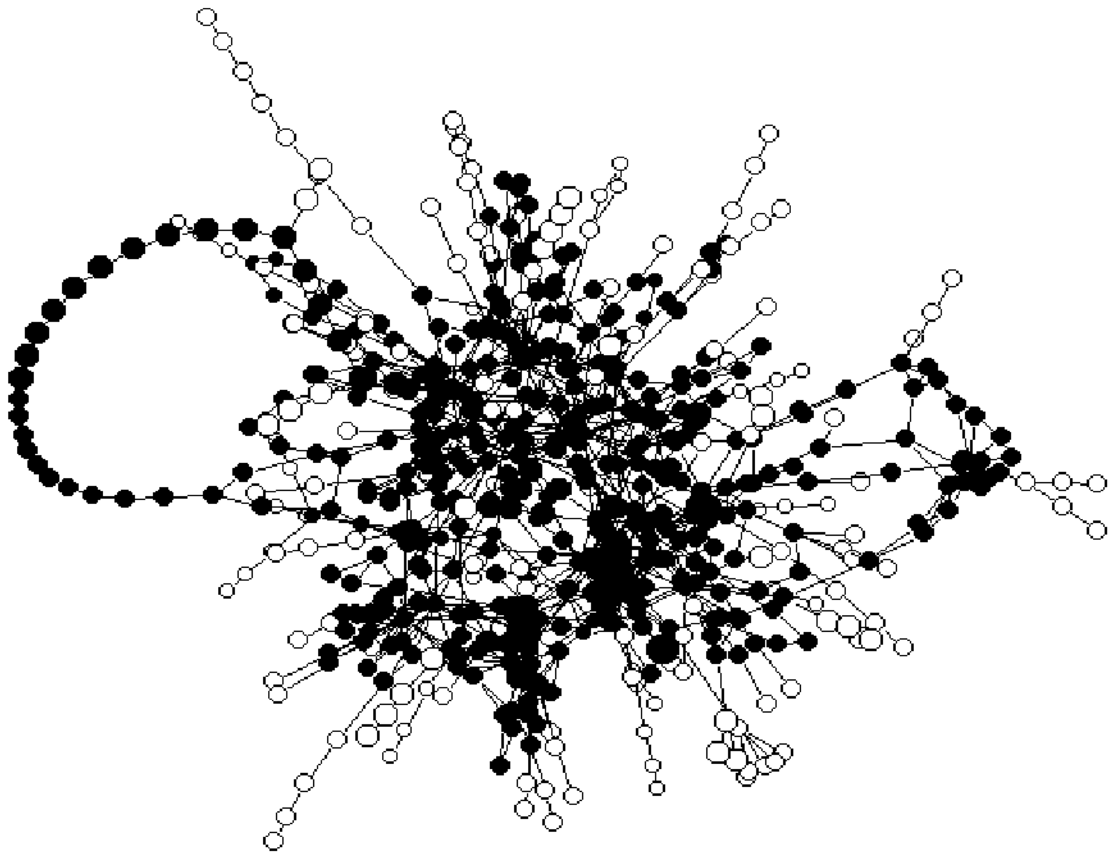
Biochemical Pathways



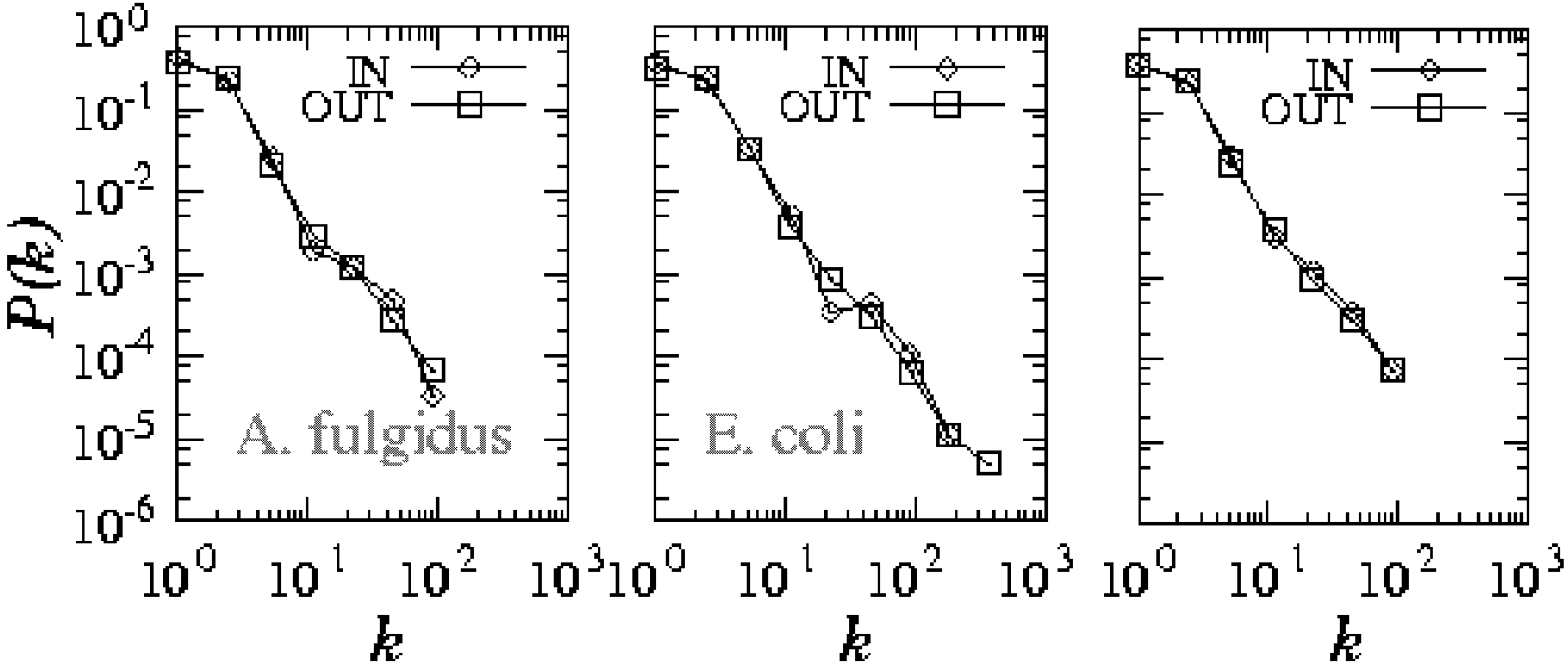
Metabolic Network

Nodes: chemicals (substrates)

Links: bio-chemical reactions



Metabolic network

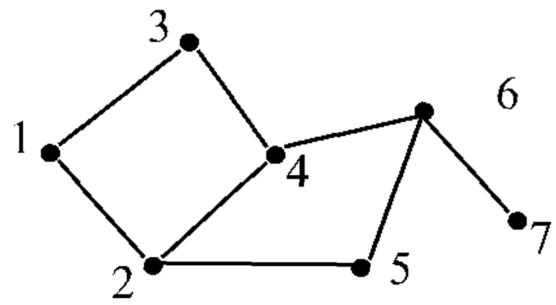


Archaea Bacteria Eukaryotes

Organisms from all three domains of life are **scale-free** networks!

H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, Nature, **407** 651 (2000)

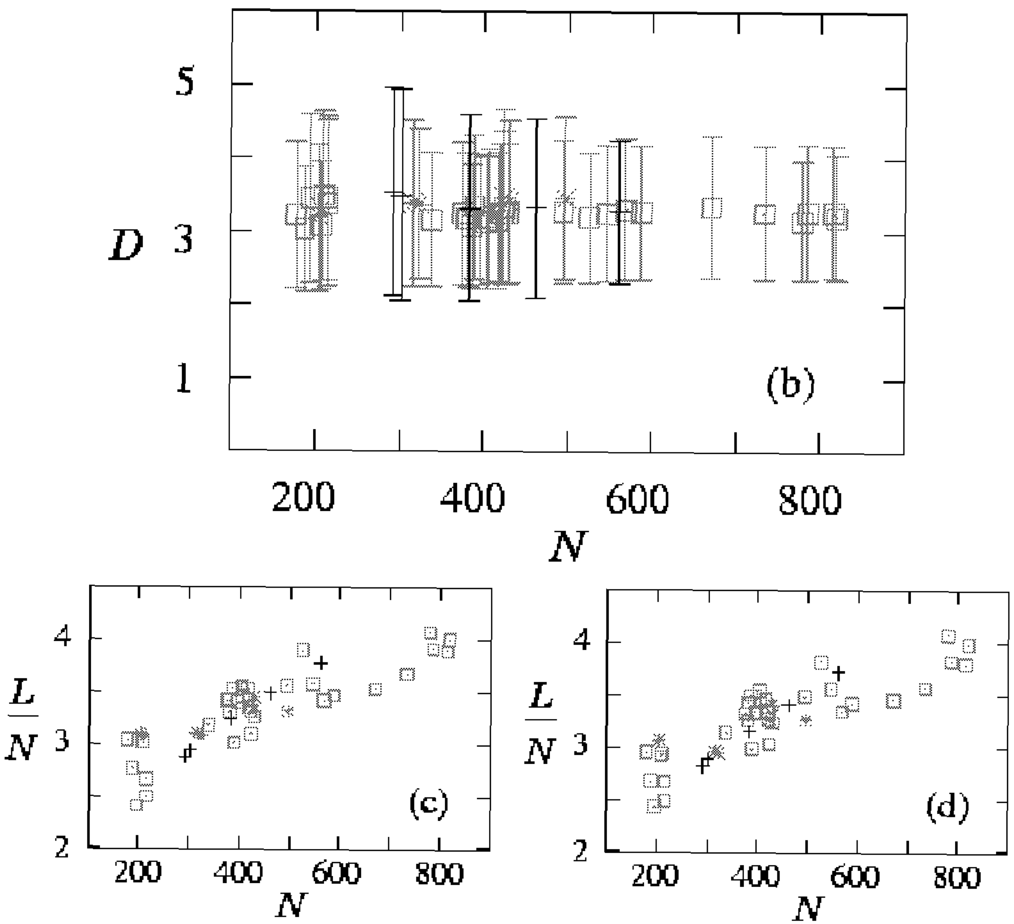
Node-node distance in metabolic networks



$D_{15} = 2$ [1→2→5]

$D_{17} = 4$ [1→3→4 → 6→7]

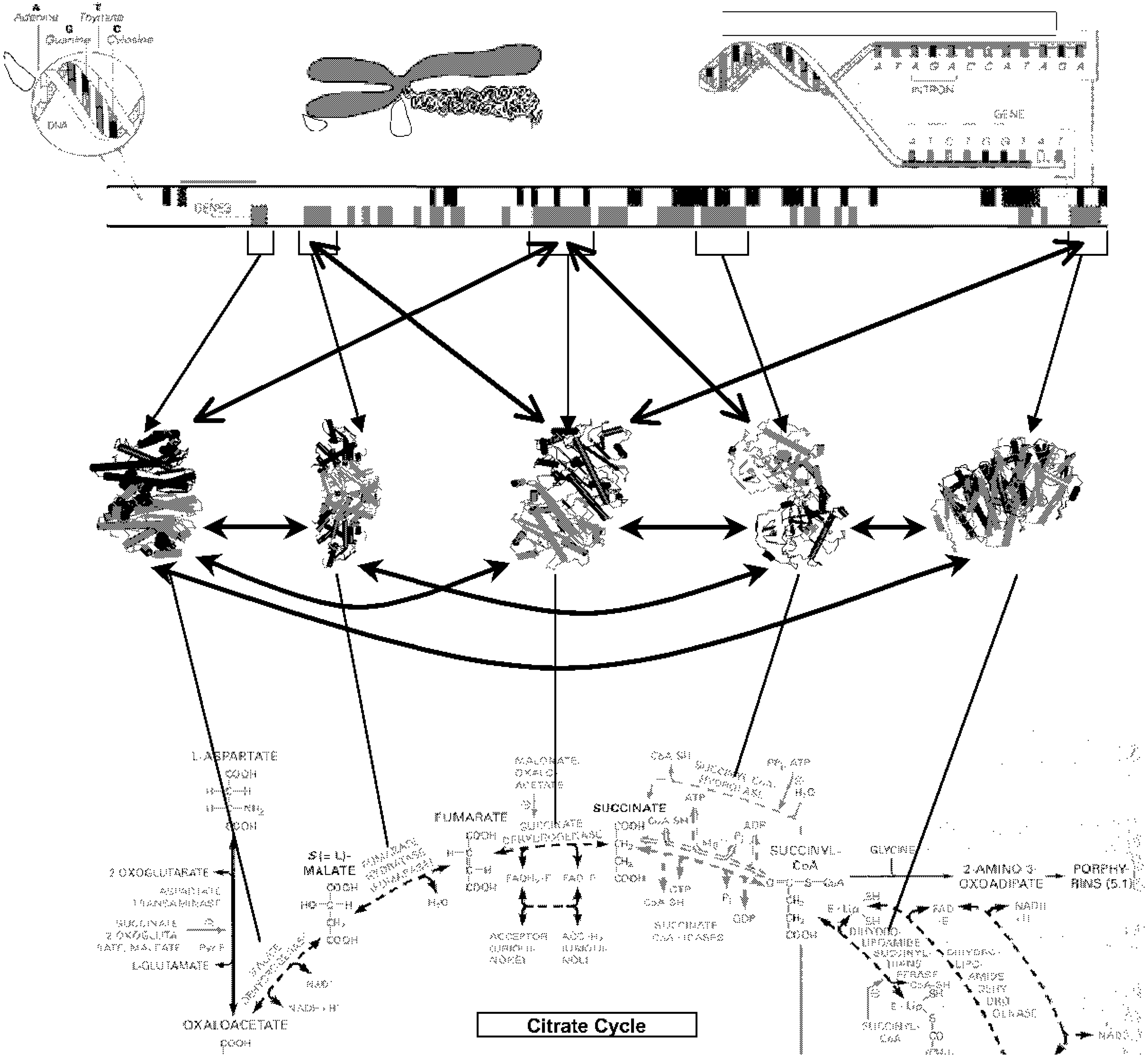
… ***D* = ??**



Scale-free networks:

D~log(N)

Larger organisms are expected to have a larger diameter!



GENOME

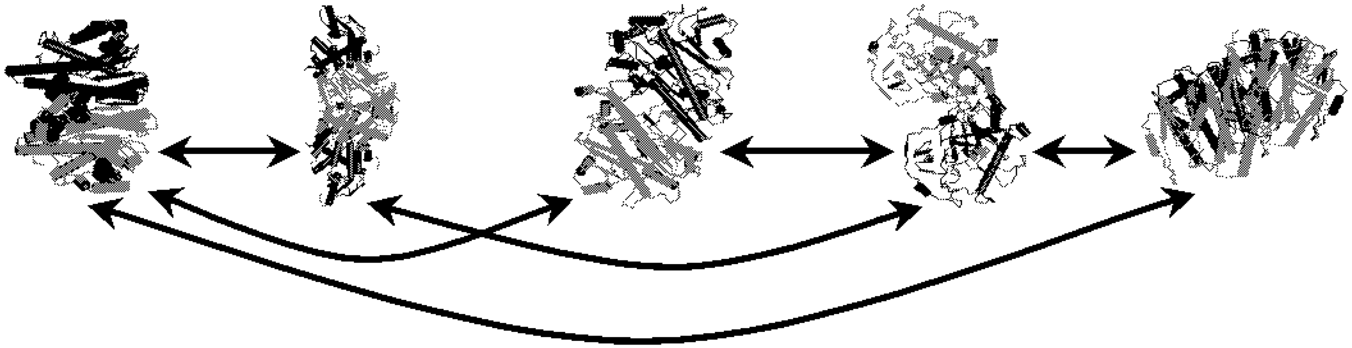
protein-gene interactions

PROTEOME

protein-protein interactions

METABOLISM

Bio-chemical reactions



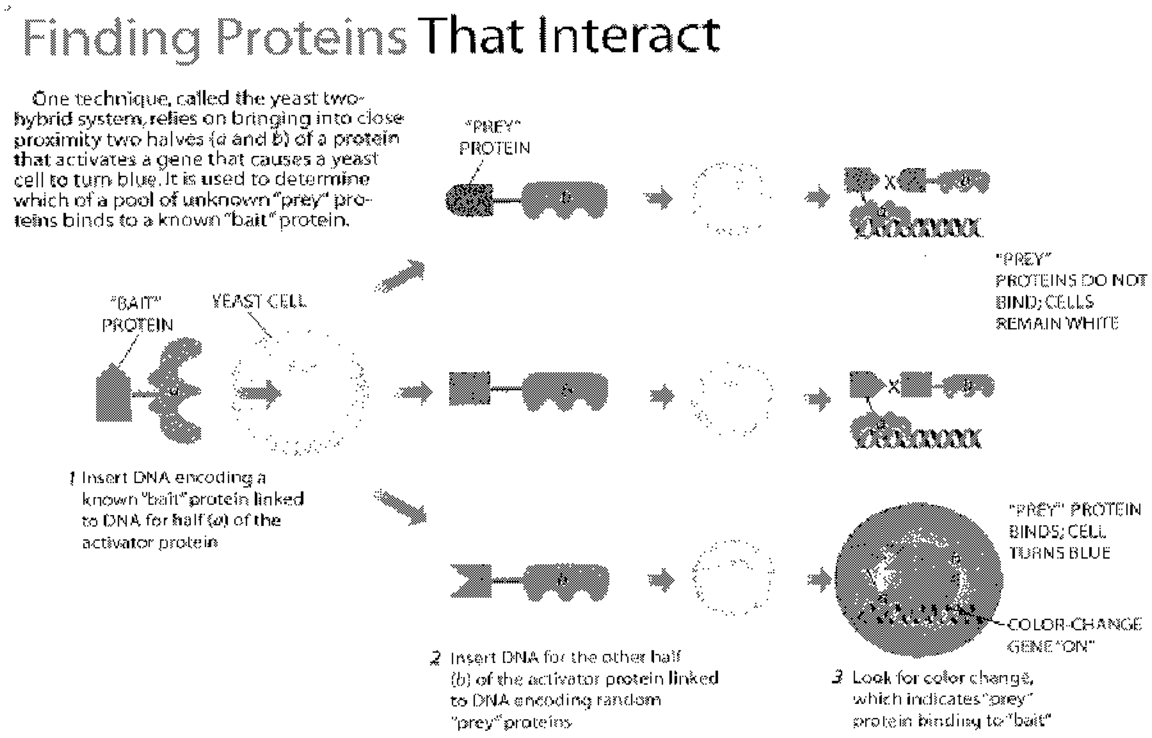
PROTEOME

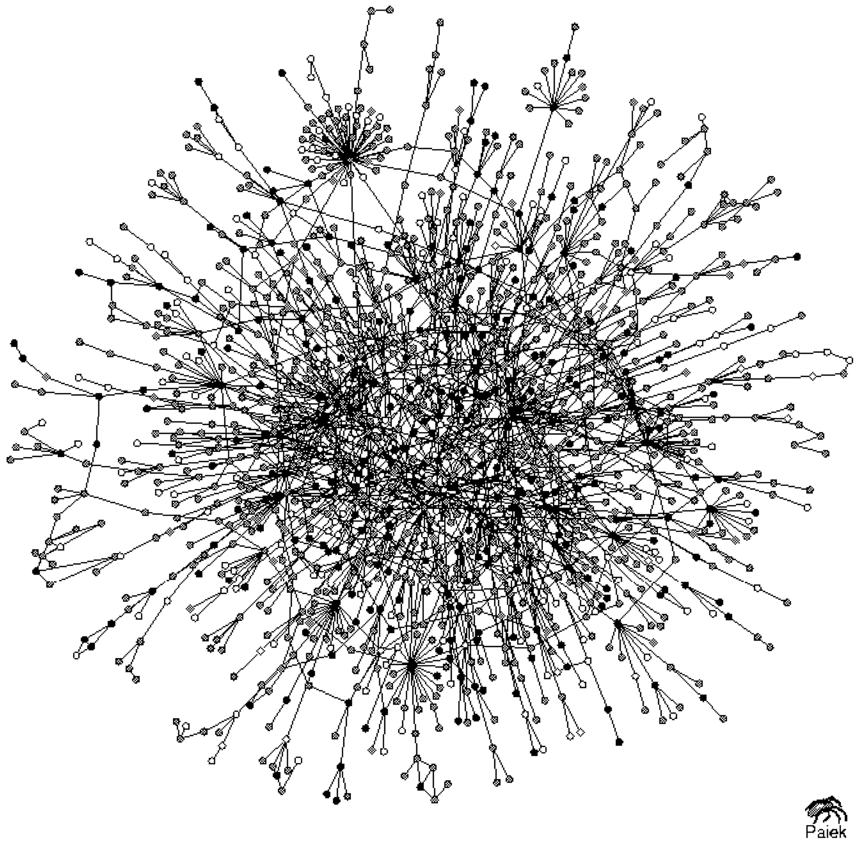
protein-protein interactions

Yeast protein network

Nodes: proteins

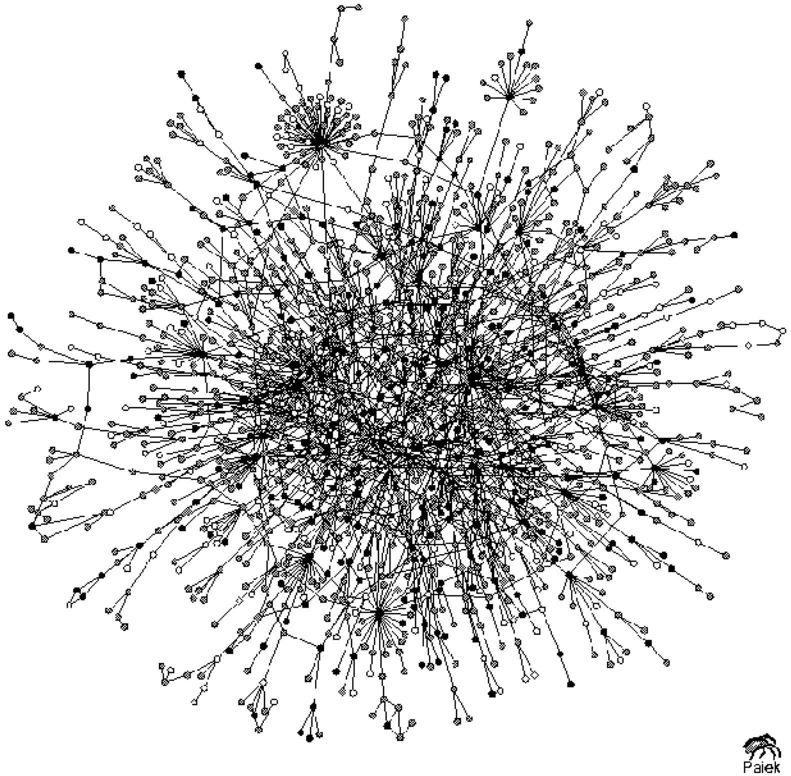
Links: physical interactions (binding)

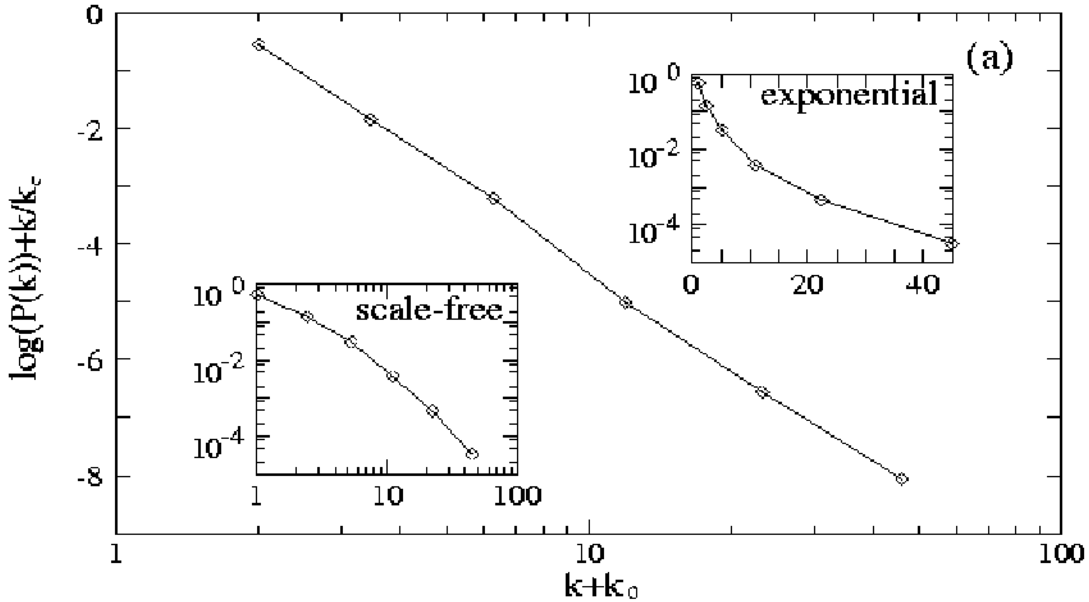




P. Uetz, *et al.* *Nature* **403**, 623-7 (2000).

Topology of the protein network

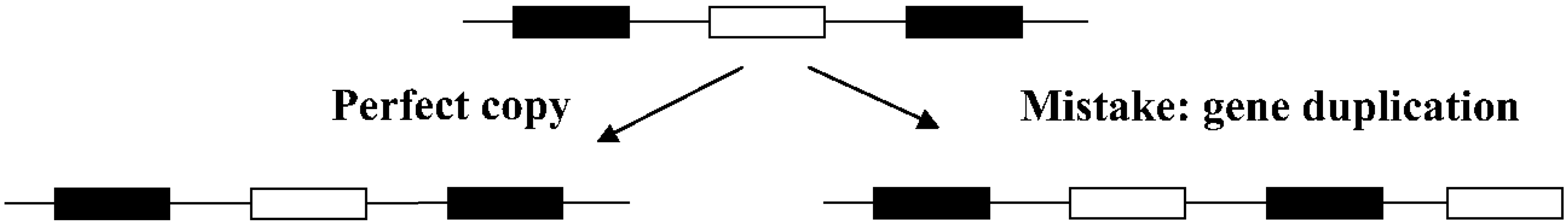


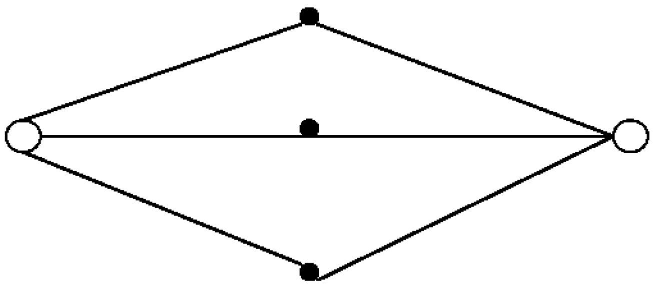


$$P(k) \sim (k+k_0)^{-\gamma} \exp(-\frac{k+k_0}{k_\tau})$$

H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Nature 411, 41-42 (2001)

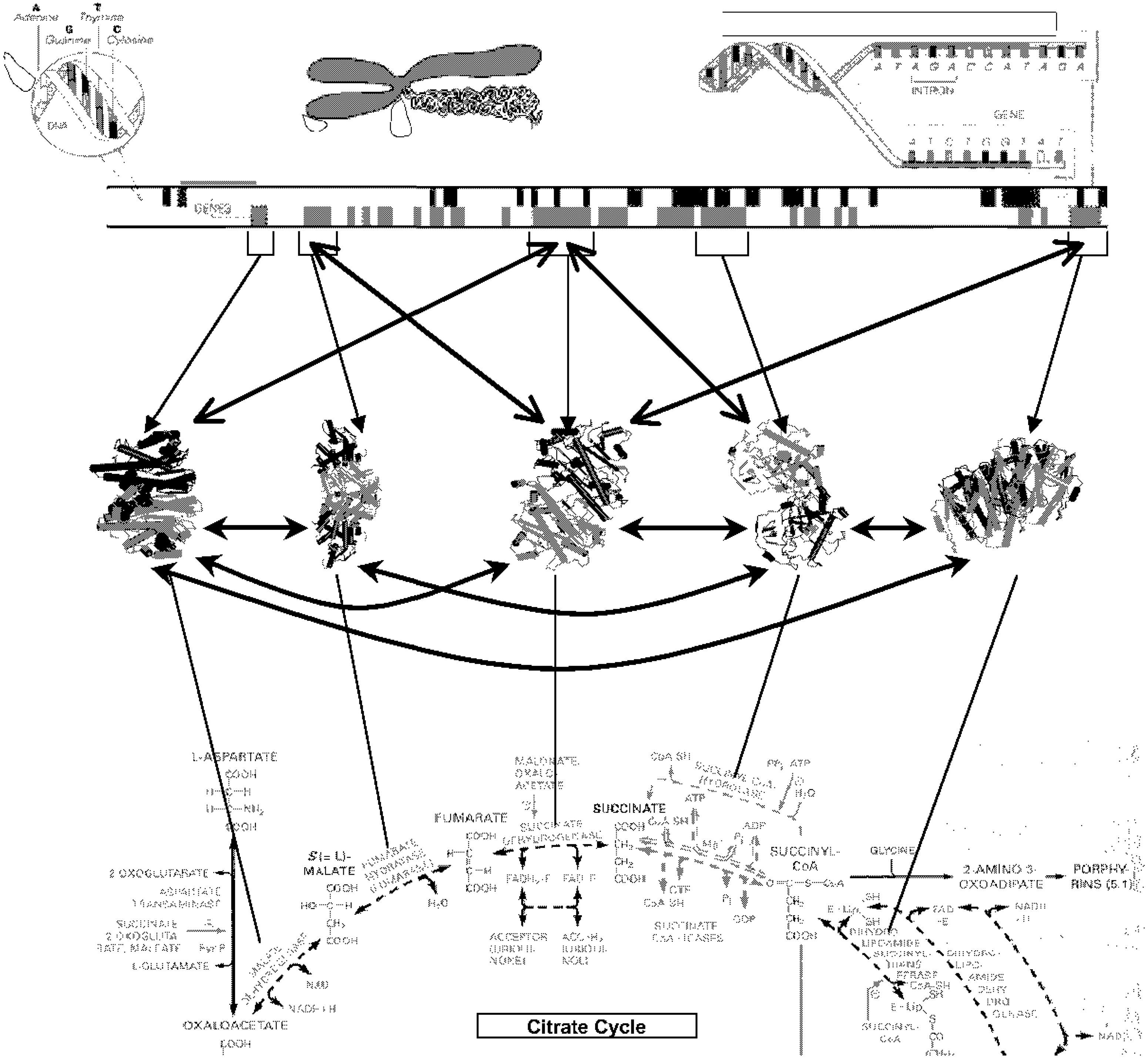
Origin of scaling in protein interaction networks





Proteins with more interactions are more likely to get a new link:
_(k)~k
(preferential attachment).

Vazquez, Flammini, Maritan, Vespignani cond-mat/0108043
Sole, Pastor-Satorras, Smith, Kepler, Adv. Compl. Syst. 2001



GENOME

protein-gene interactions

PROTEOME

protein-protein interactions

METABOLISM

Bio-chemical reactions

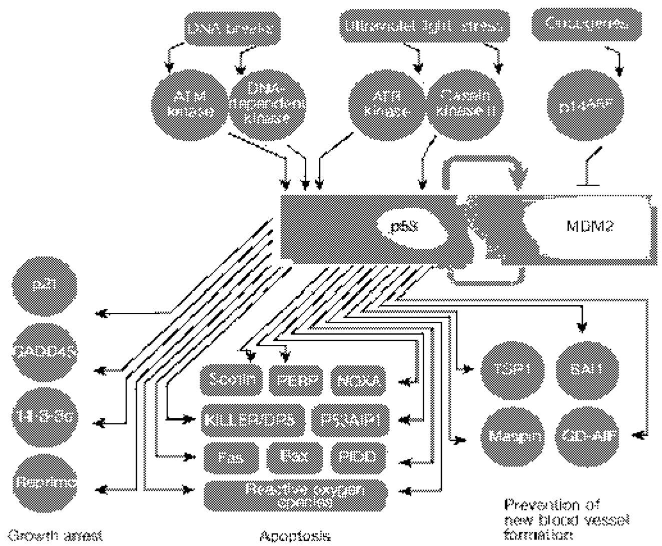
Nature **408** 307 (2000)

news and views feature

Surfing the p53 network

Bert Vogelstein, David Lane and Arnold J. Levine

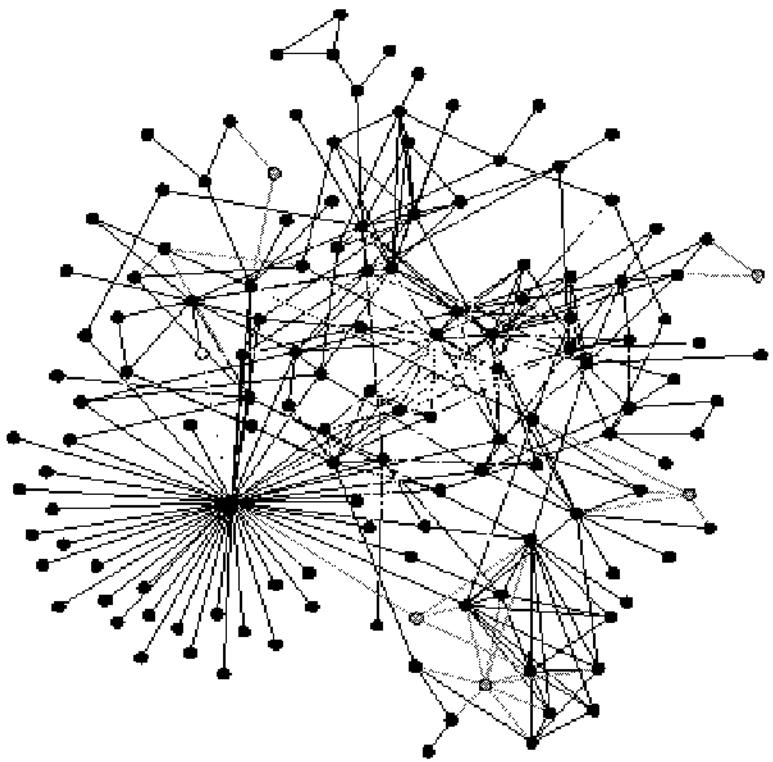
The p53 tumour-suppressor gene integrates numerous signals that control cell life and death. As when a highly connected node in the Internet breaks down, the disruption of p53 has severe consequences.

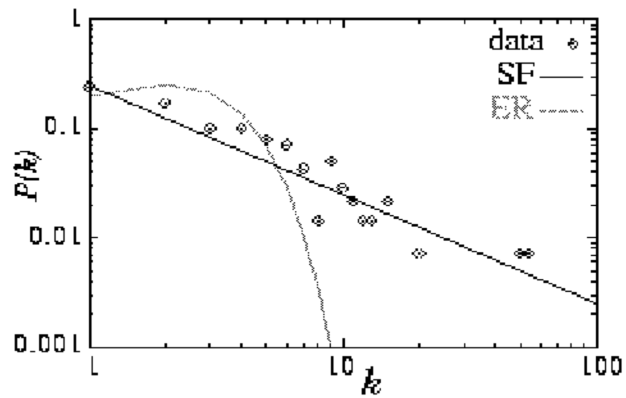


…

"One way to understand the p53 network is to compare it to the Internet. The cell, like the Internet, appears to be a '**scale-free network**'."

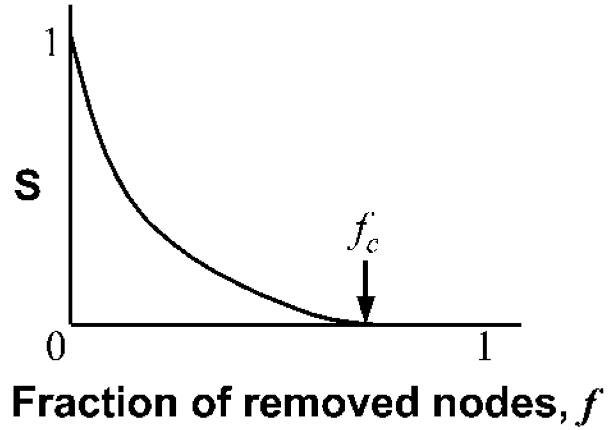
p53 network (mammals)

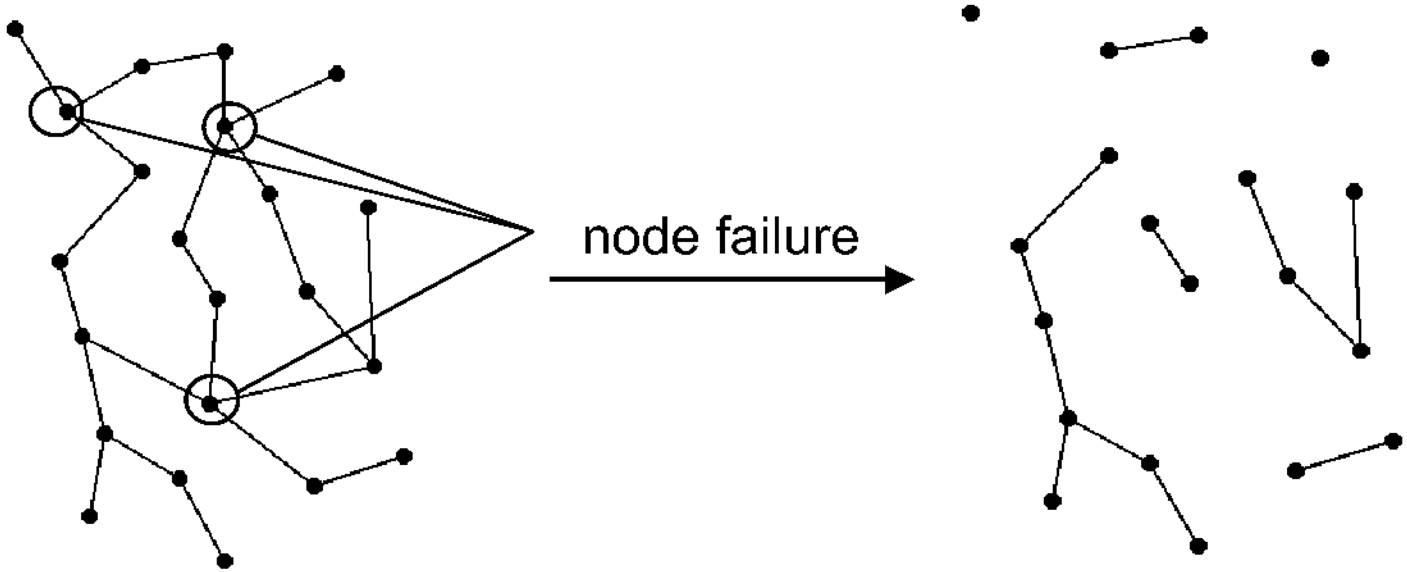




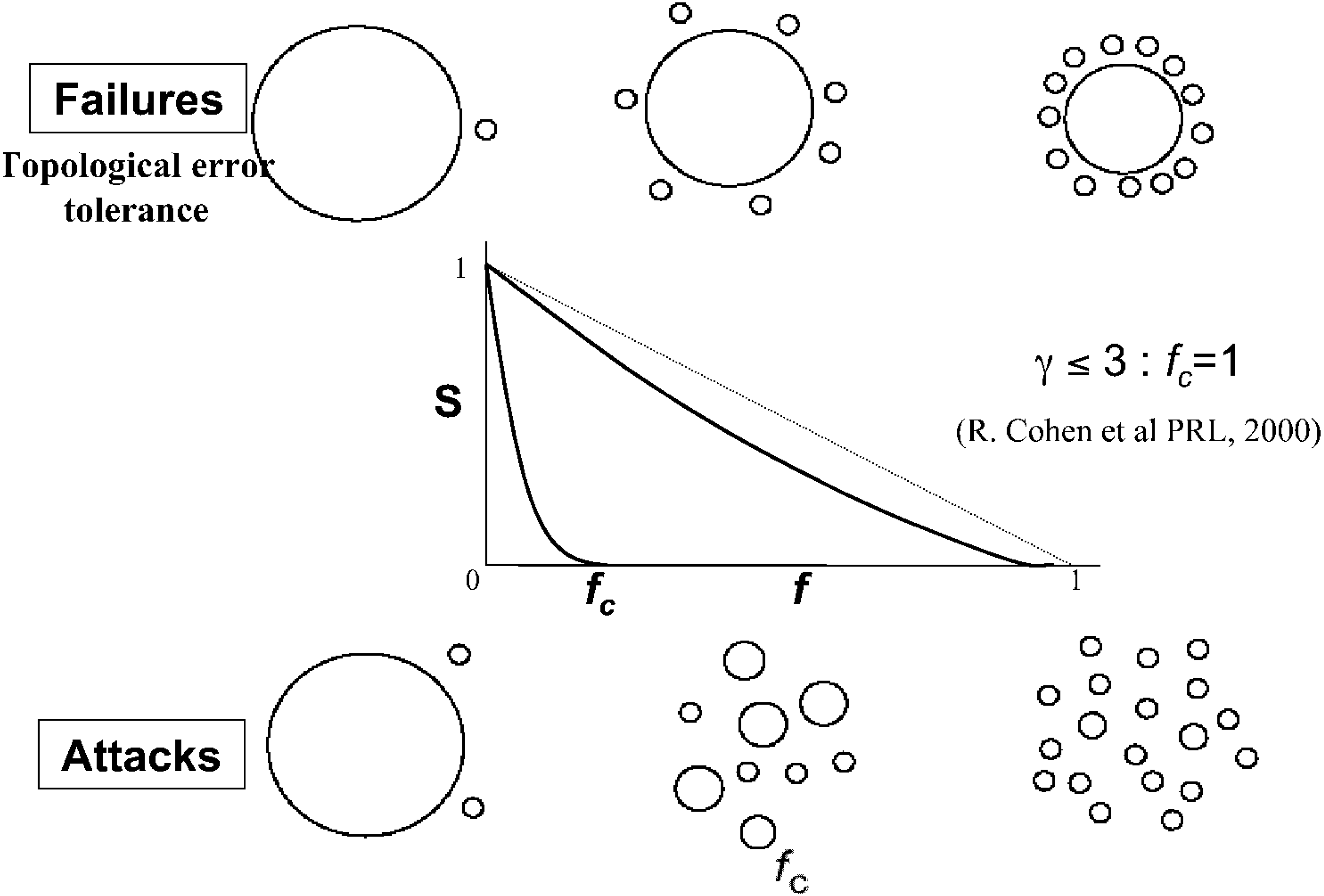
Robustness

Complex systems maintain their basic functions
even under errors and failures
(cell → mutations; Internet → router breakdowns)

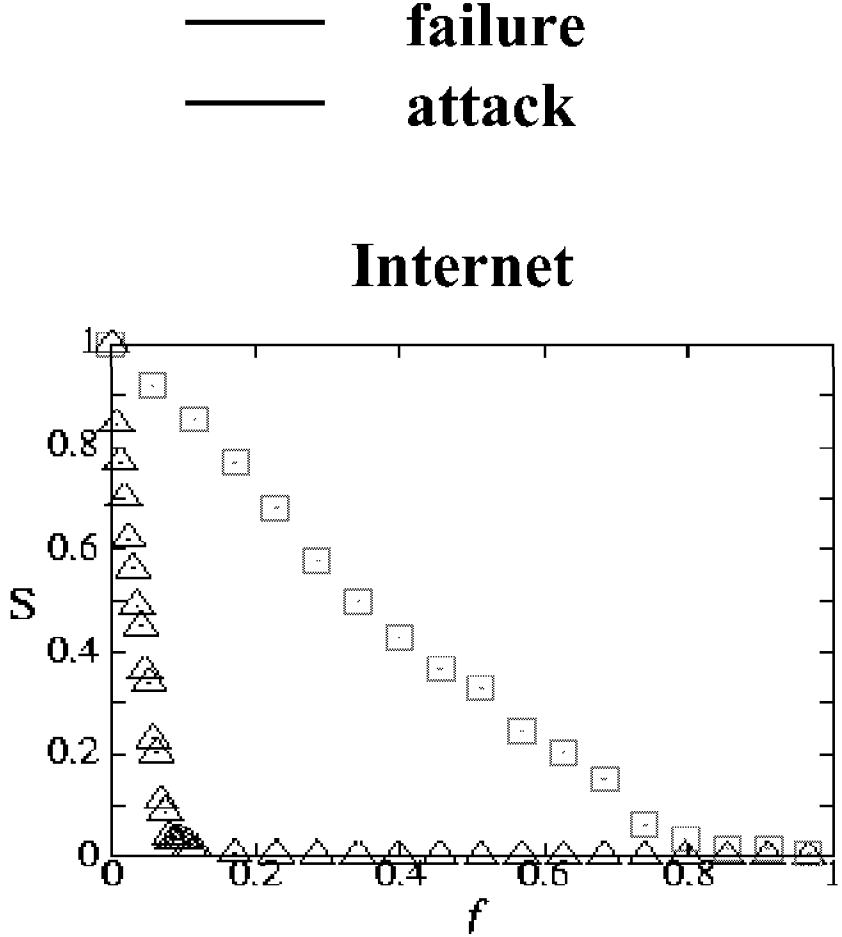




Robustness of scale-free networks

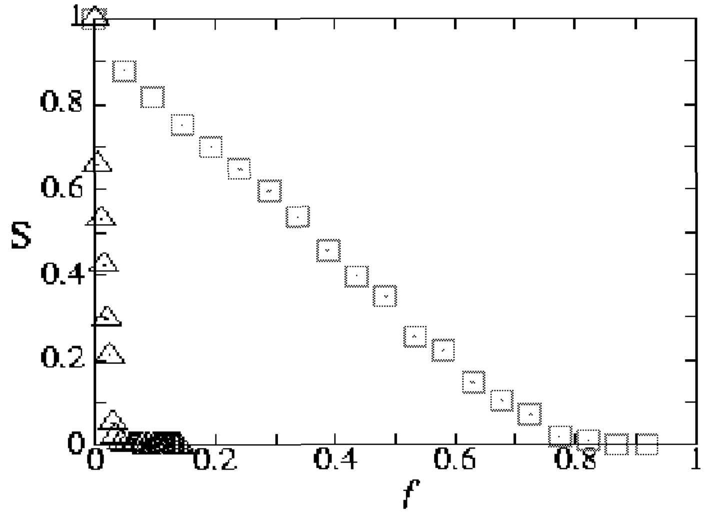


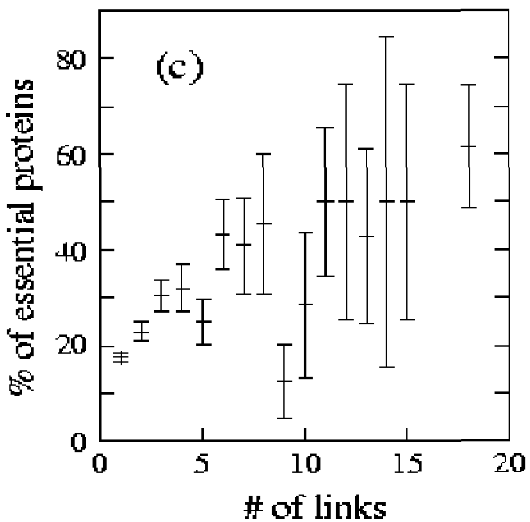
Achilles' Heel of complex networks



R. Albert, H. Jeong, A.L. Barabasi, Nature **406** 378 (2000)

Yeast protein network
- lethality and topological position -





Highly connected proteins are more **essential (lethal)**...

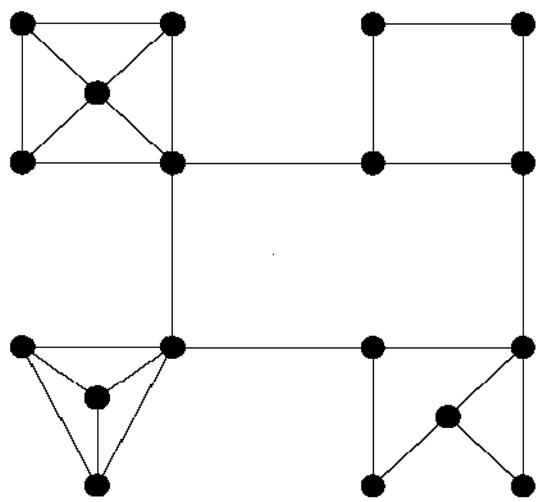
H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Nature 411, 41-42 (2001)

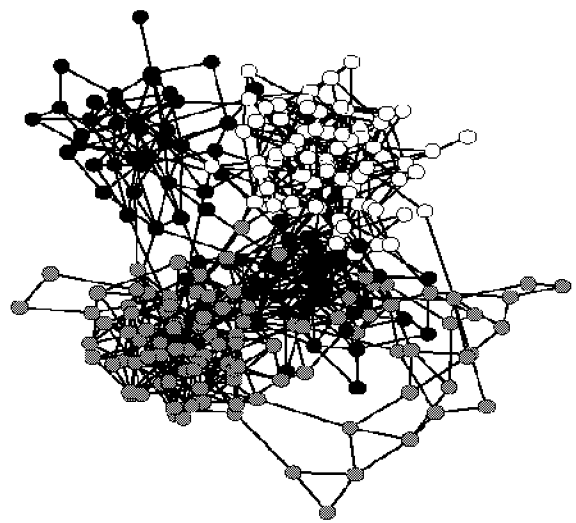
Modularity in Cellular Networks

- Hypothesis:
  Biological function are carried by discrete functional modules.

❖Hartwell, L.-H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999).

- Traditional view of modularity:

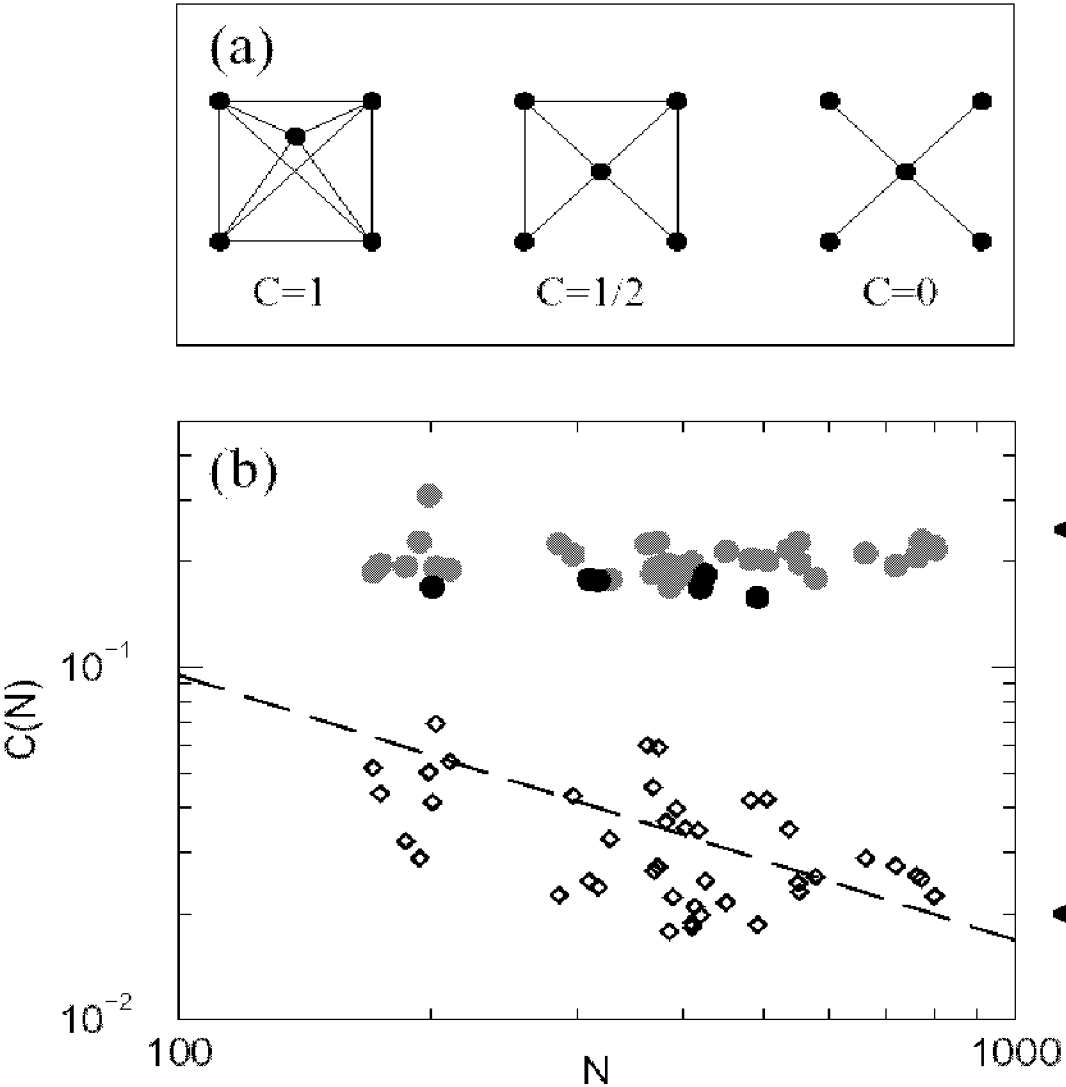




- Question: Is modularity a myth, or a structural property of biological networks?
  (are biological networks fundamentally modular?)

Ravasz, Somera, Mongru, Oltvai, A-L. B, ***Science*** **297,** 1551 (2002).

Modularity in the metabolism



Clustering Coefficient:

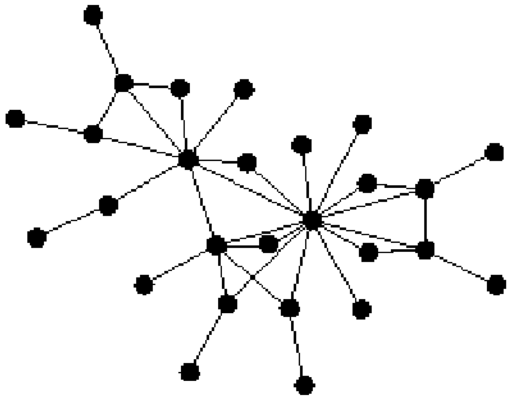
$$C(k) = \frac{\#\text{ links between k neighbors}}{k(k-1)/2}$$

←Metabolic network (43 organisms)

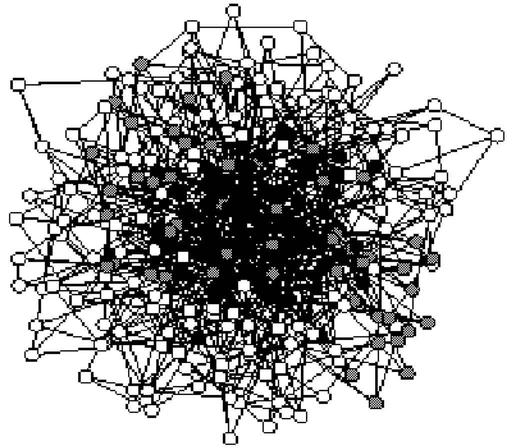
← Scale-free model

Modular vs. Scale-free Topology

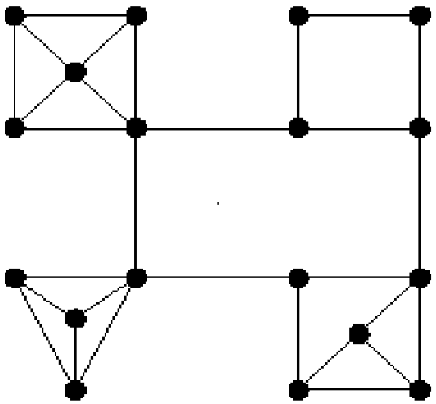
(a)



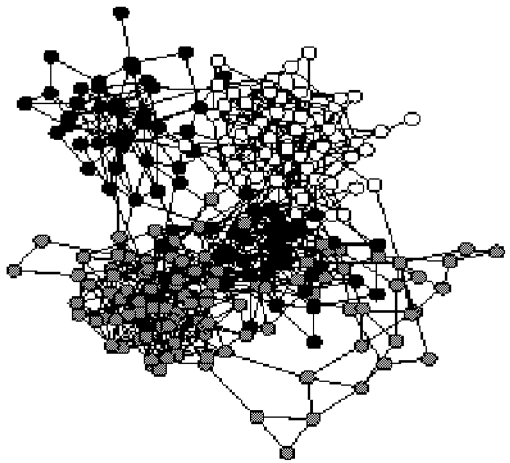
Scale-free



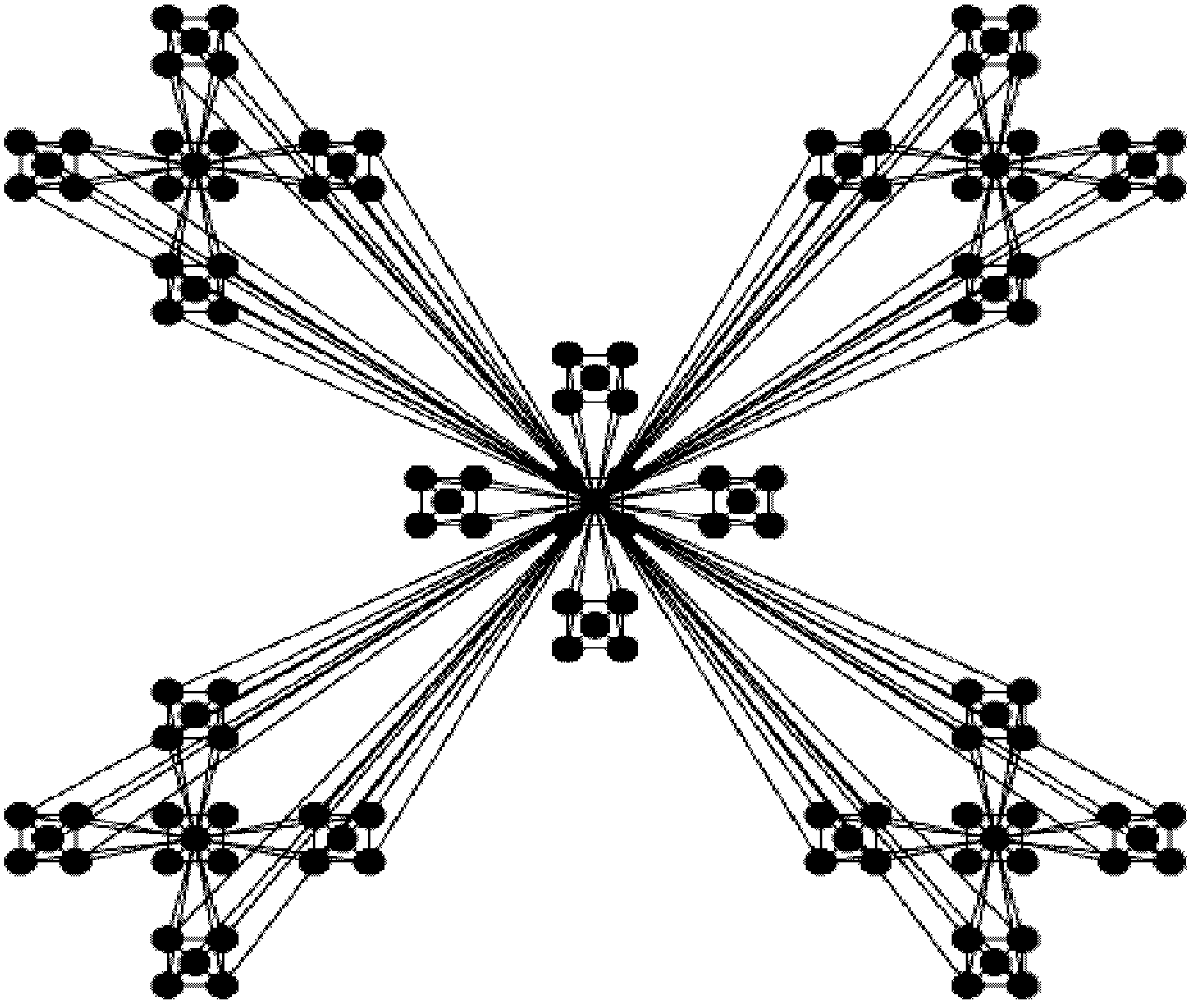
(b)



Modular



Hierarchical Networks

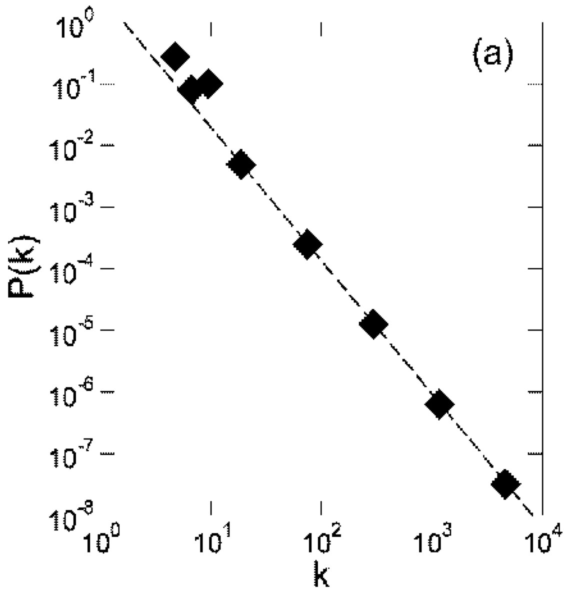


Properties of hierarchical networks

1. Scale-free

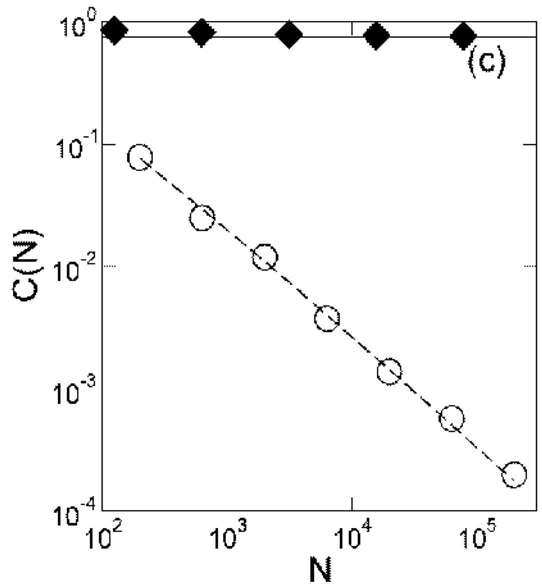
$$\gamma = 1 + \frac{\ln 5}{\ln 4}$$

$$= 2.161$$



2. Clustering coefficient independent of N

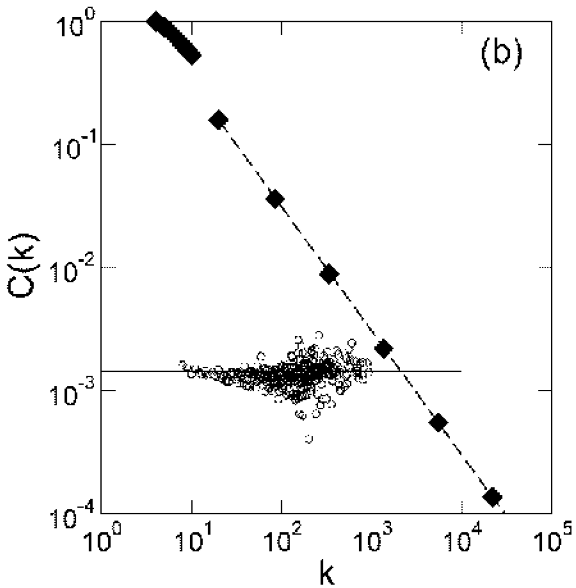
$$C(N) = \text{const.}$$

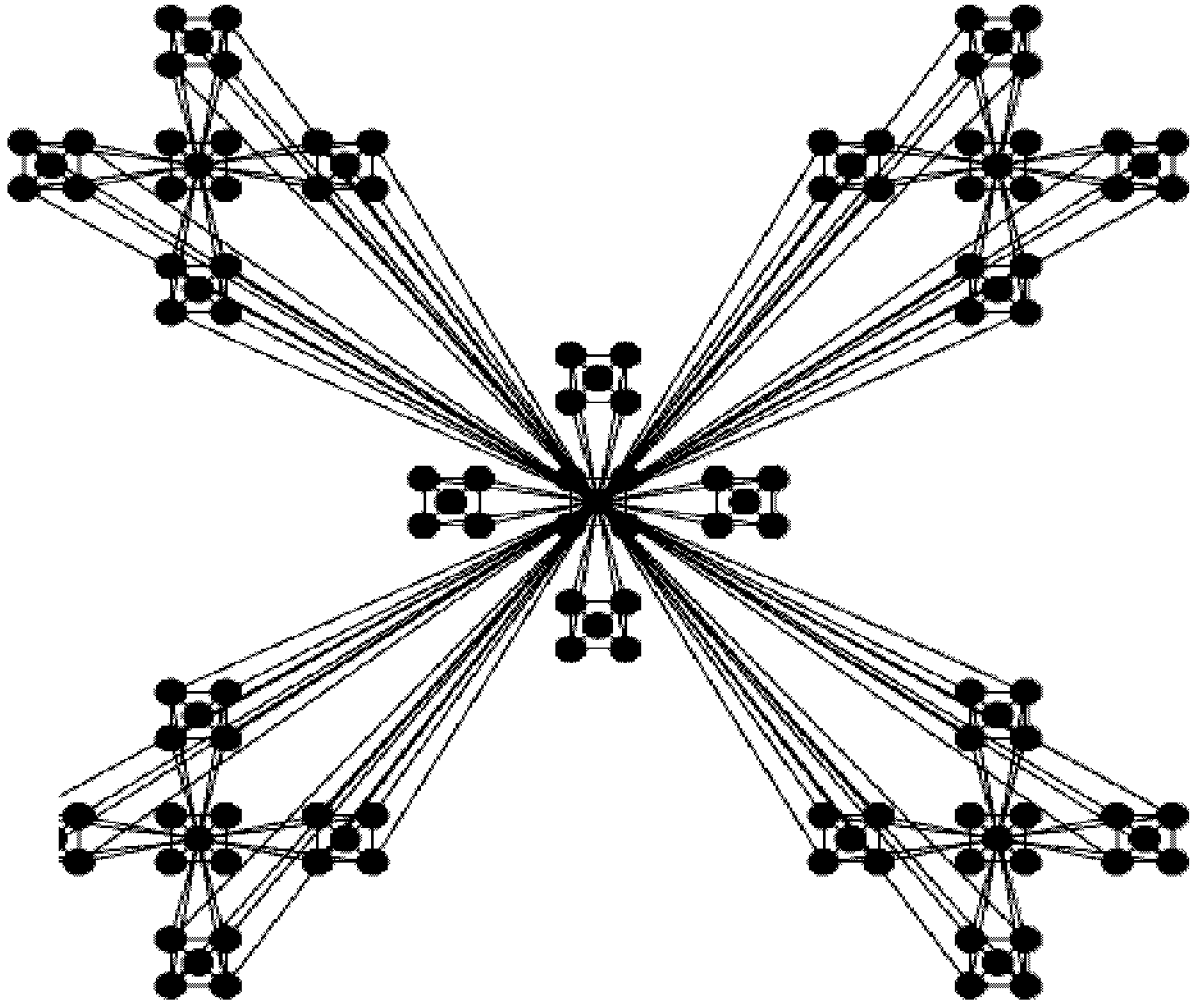


Hierarchical Networks

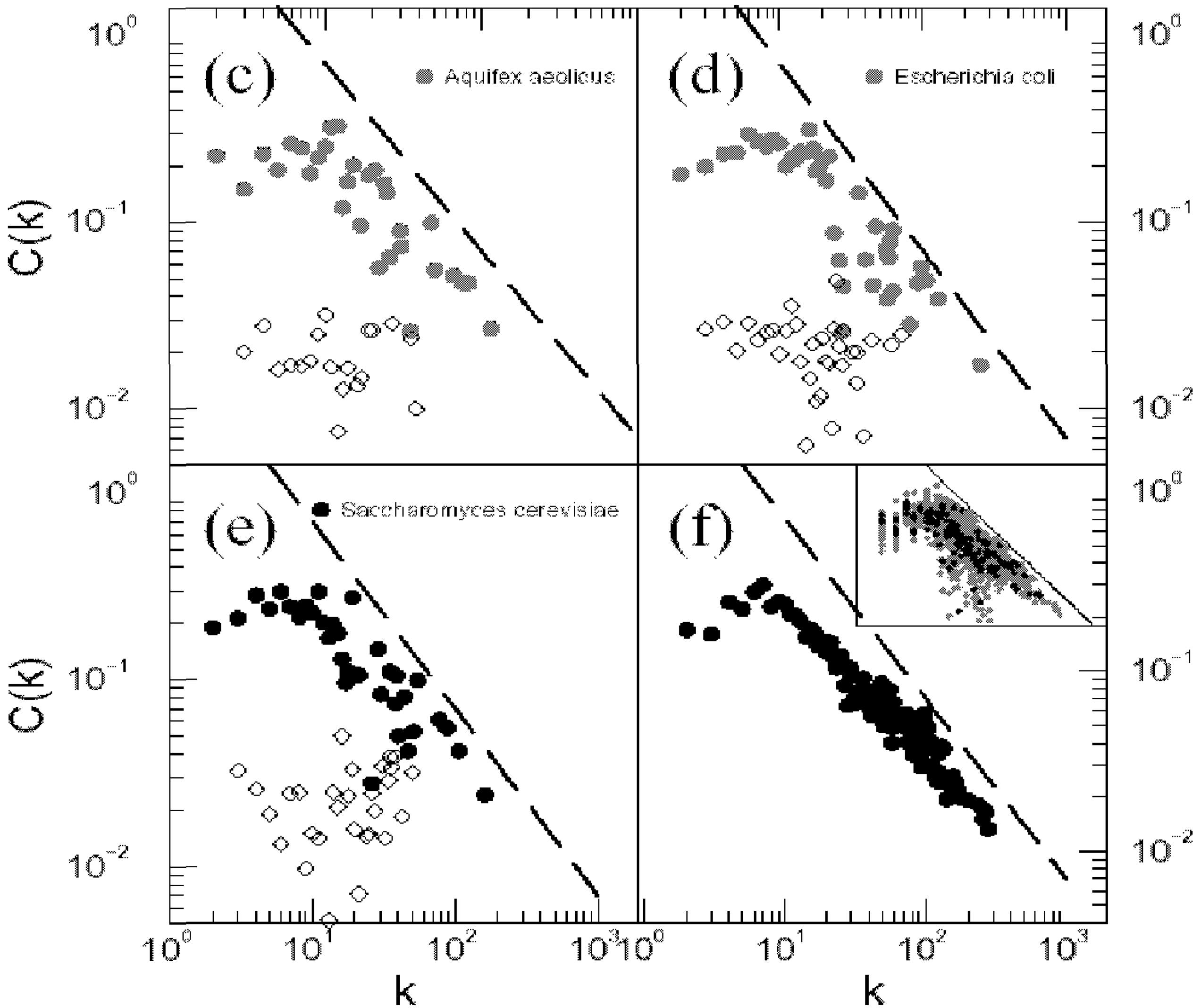
3. Clustering coefficient depends on *k*

$$C(k) \sim k^{-1}$$





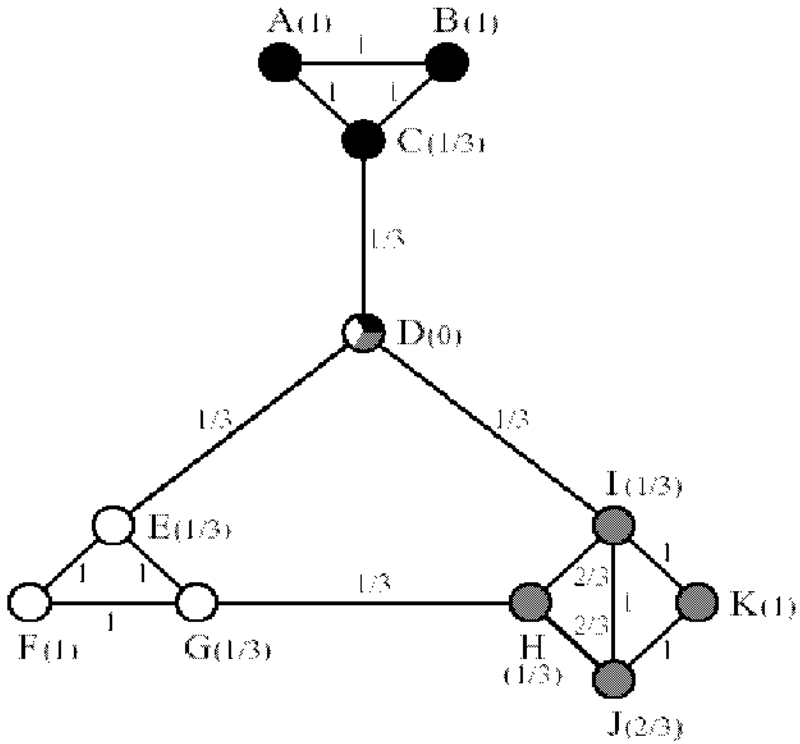
Scaling of the clustering coefficient C(k)

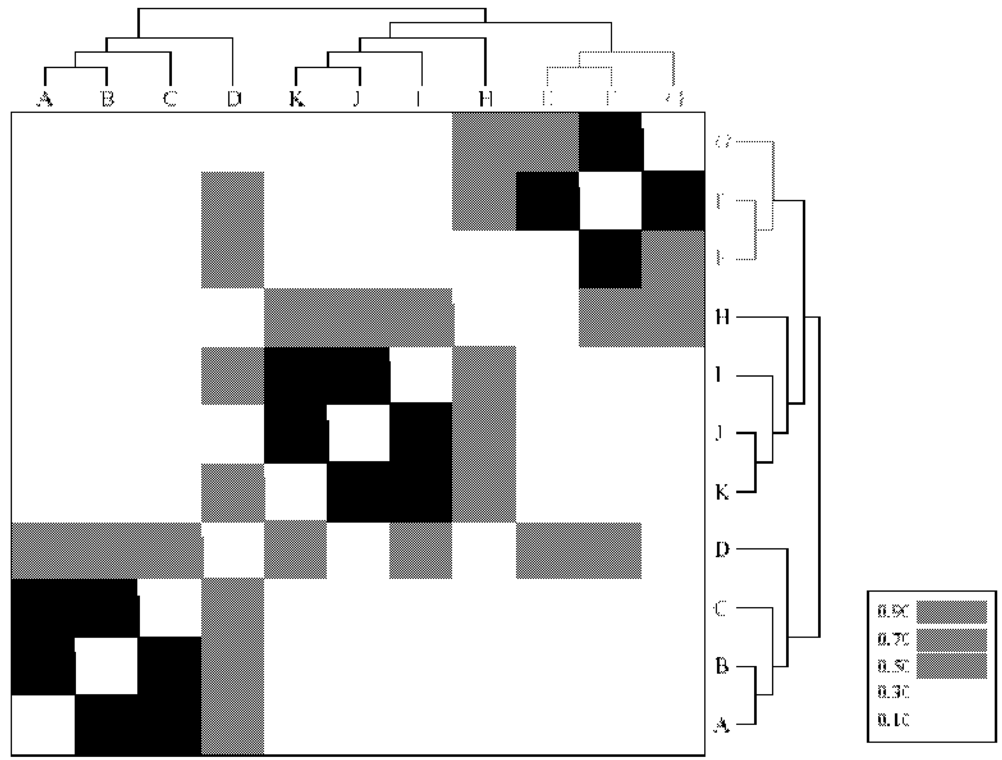


The metabolism forms a hierachical network.

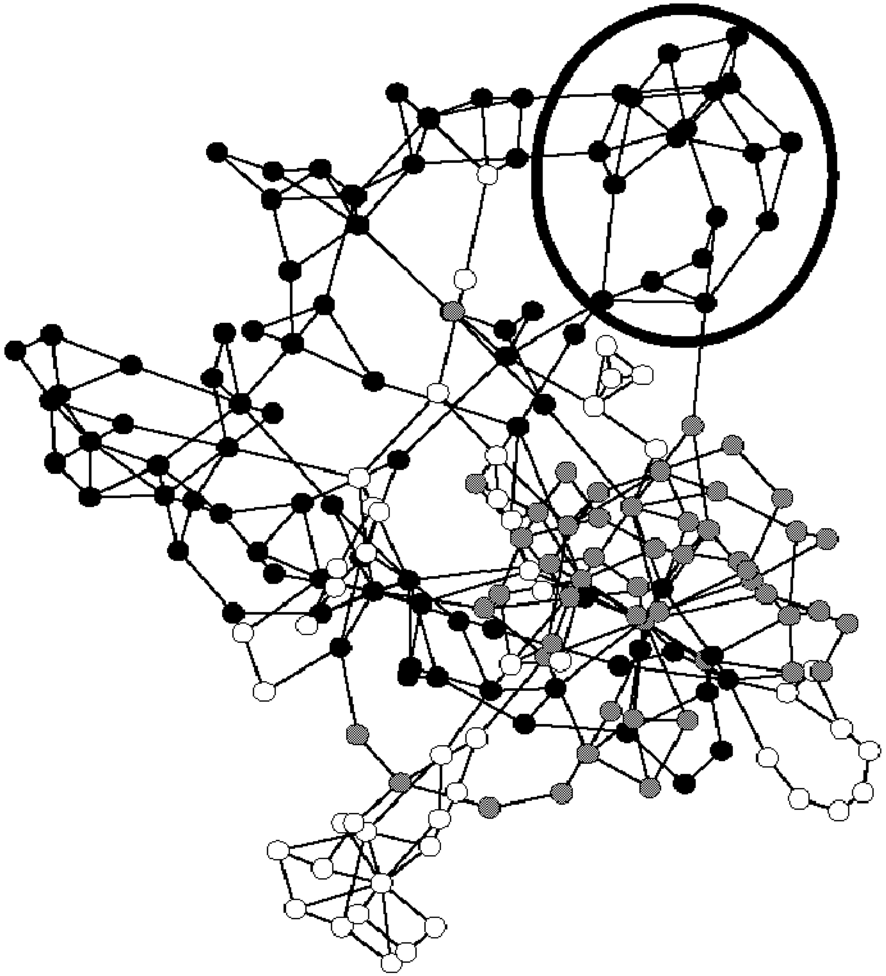
Ravasz, Somera, Mongru, Oltvai, A-L. B, ***Science*** **297,** 1551 (2002).

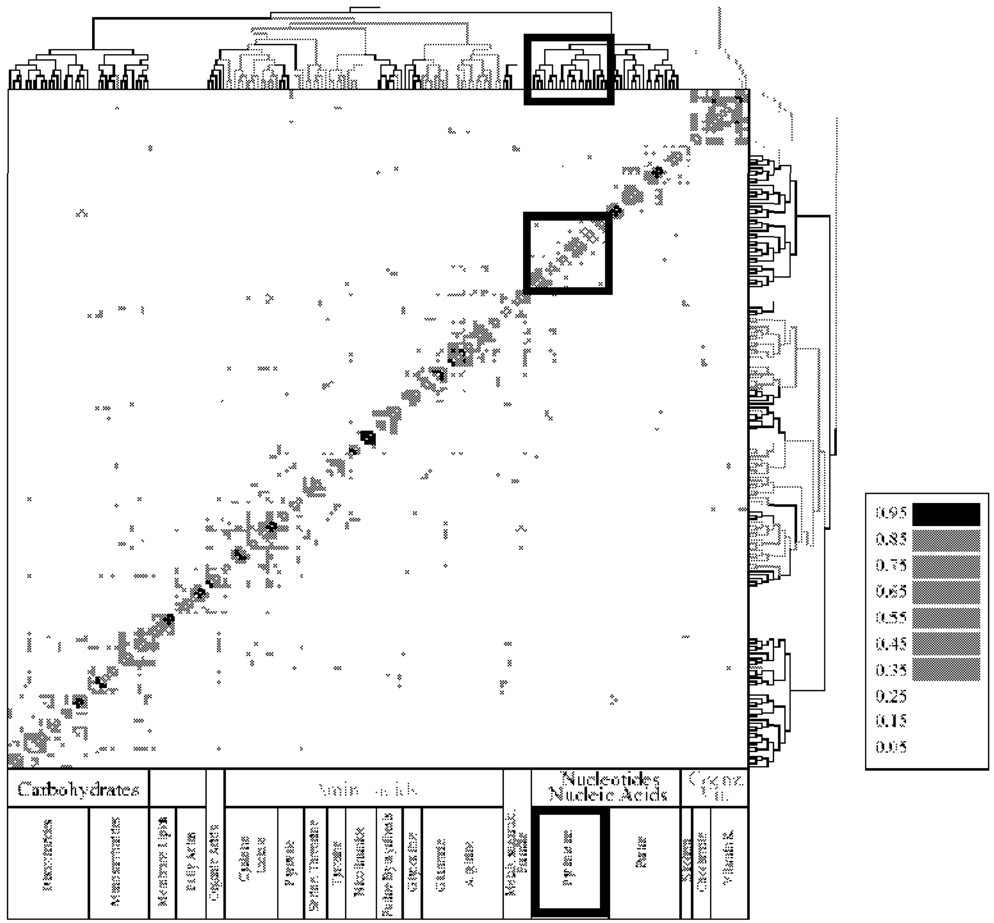
Can we identify the modules?



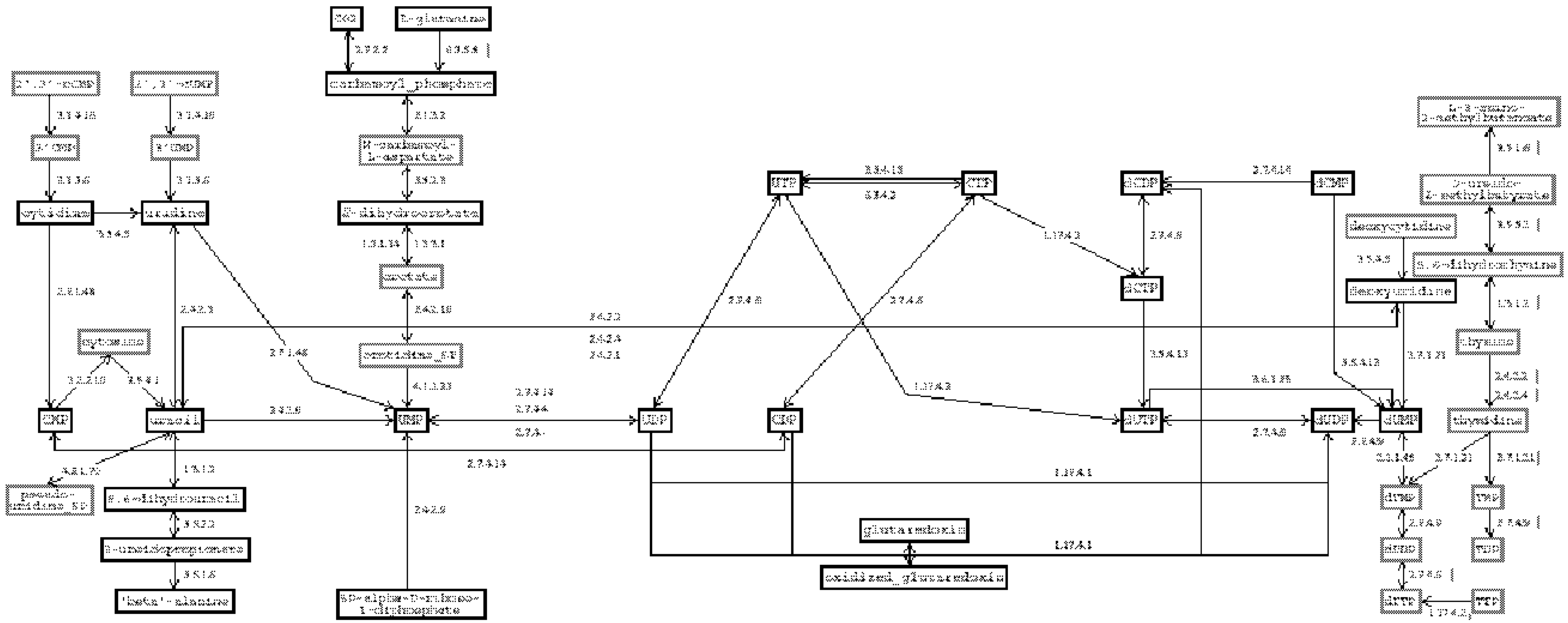


Modules in the *E. coli* metabolism





The structure of pyrimidine metabolism

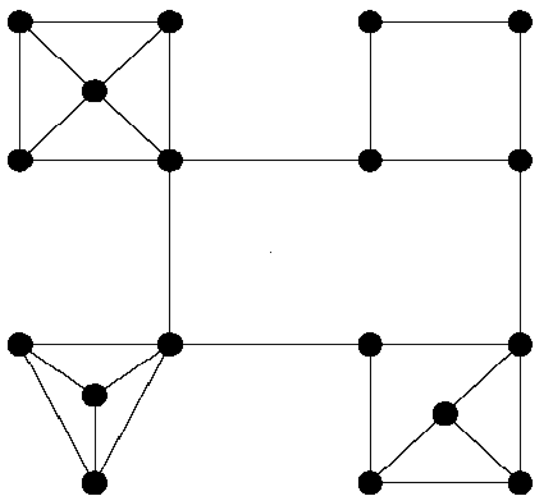


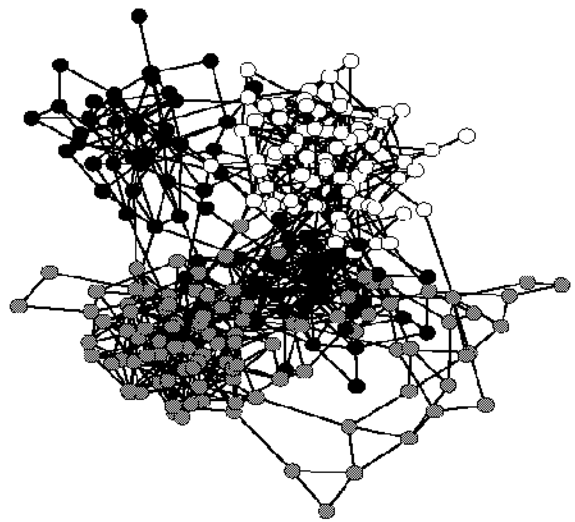
Modularity

- High C → real networks are fragmented into group or modules

  - ❖ Society: Granovetter, M. S. (1973) ; Girvan, M., & Newman, M.E.J. (2001); Watts, D. J., Dodds, P. S., & Newman, M. E. J. (2002).
  - ❖ WWW: Flake, G. W., Lawrence, S., & Giles. C. L. (2000).
  - ❖ Biology: Hartwell, L.-H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999).
  - ❖ Internet: Vasquez, Pastor-Satorras, Vespignani(2001).

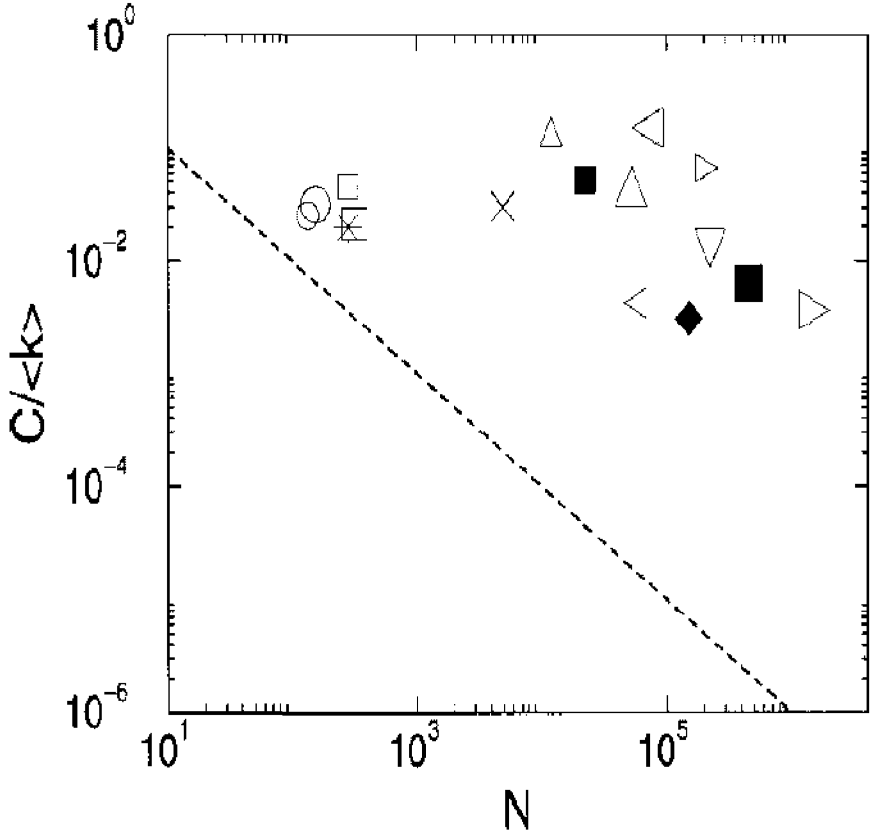
- Traditional view of modularity:



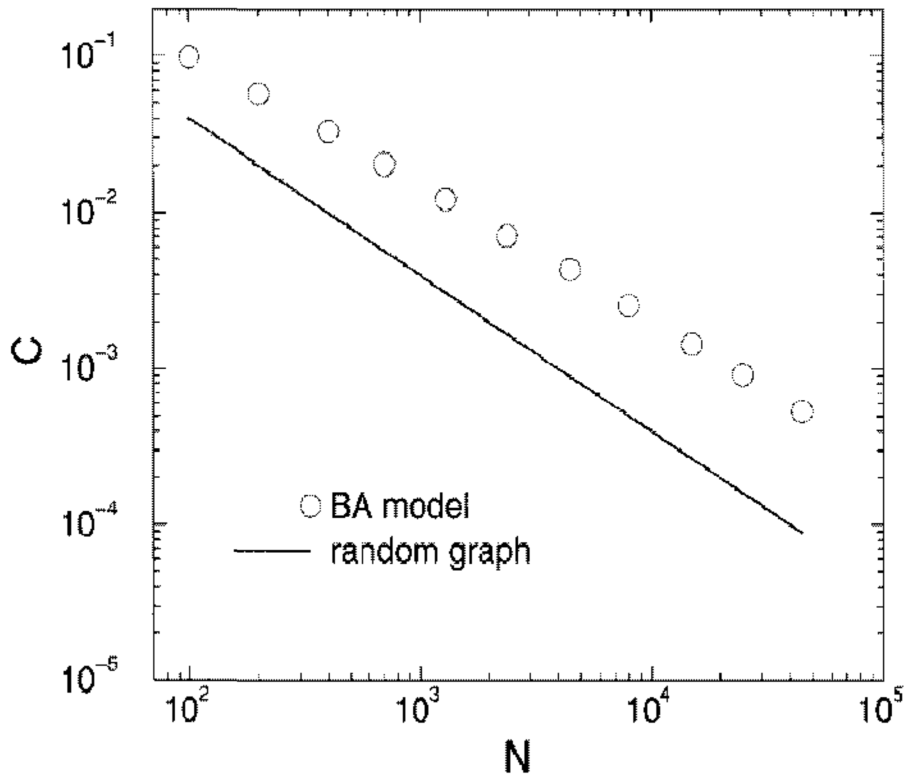


Clustering in non-biological networks

C is independent of N

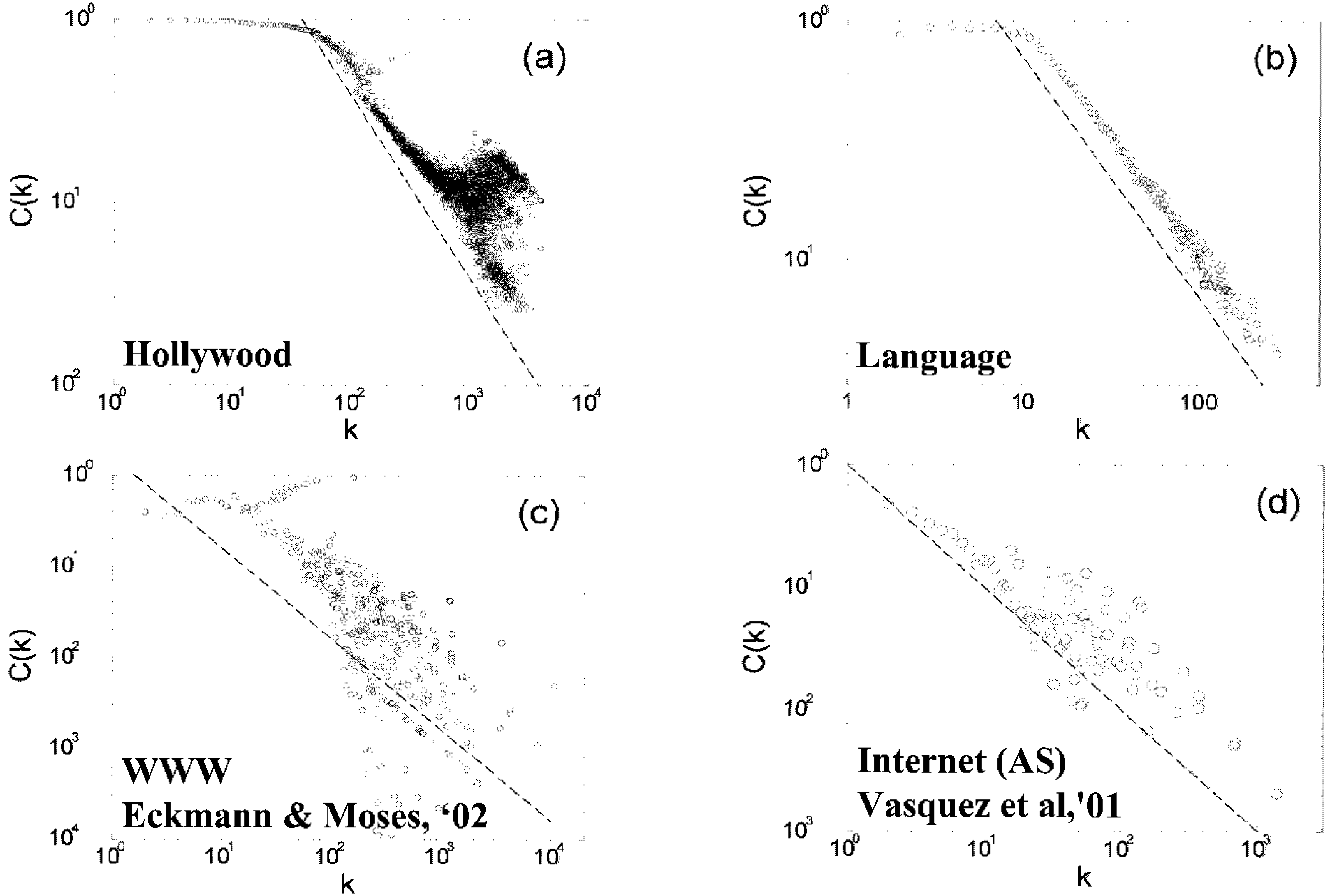


C decreases with N



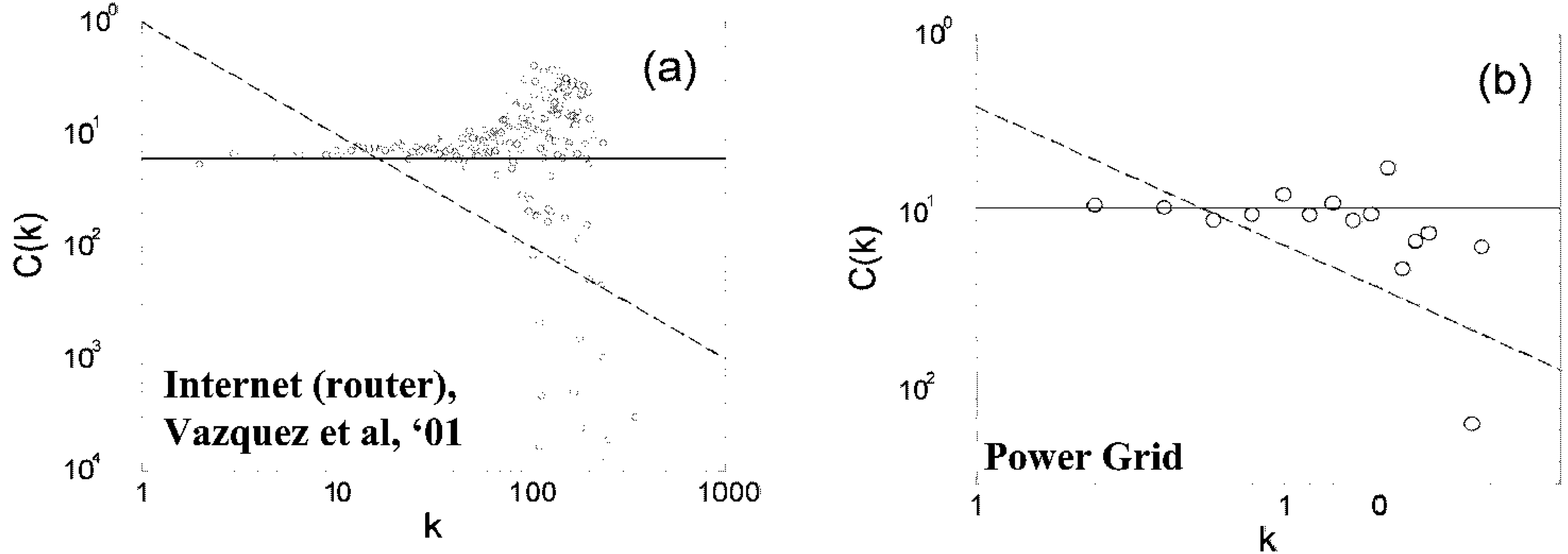
$C_i = 2n_i / k_i(k_i - 1)$

Real Networks



E. Ravasz and A.-L. B, **cond-mat/0206130**

Exceptions:
Geographically Organized Networks:

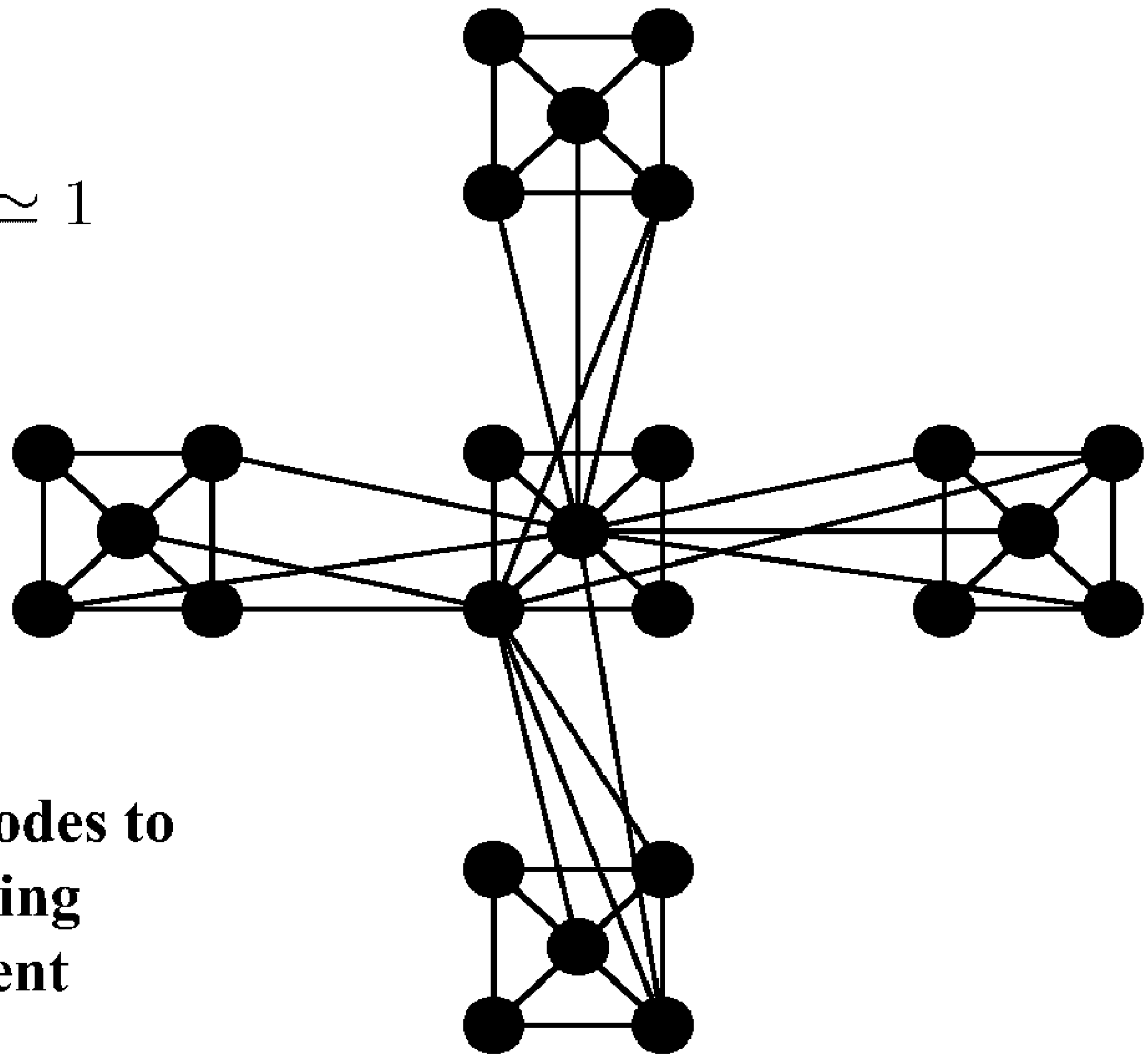


<u>Common feature:</u>

economic pressures towards shorter links

Is the hierarchical exponent _ universal?

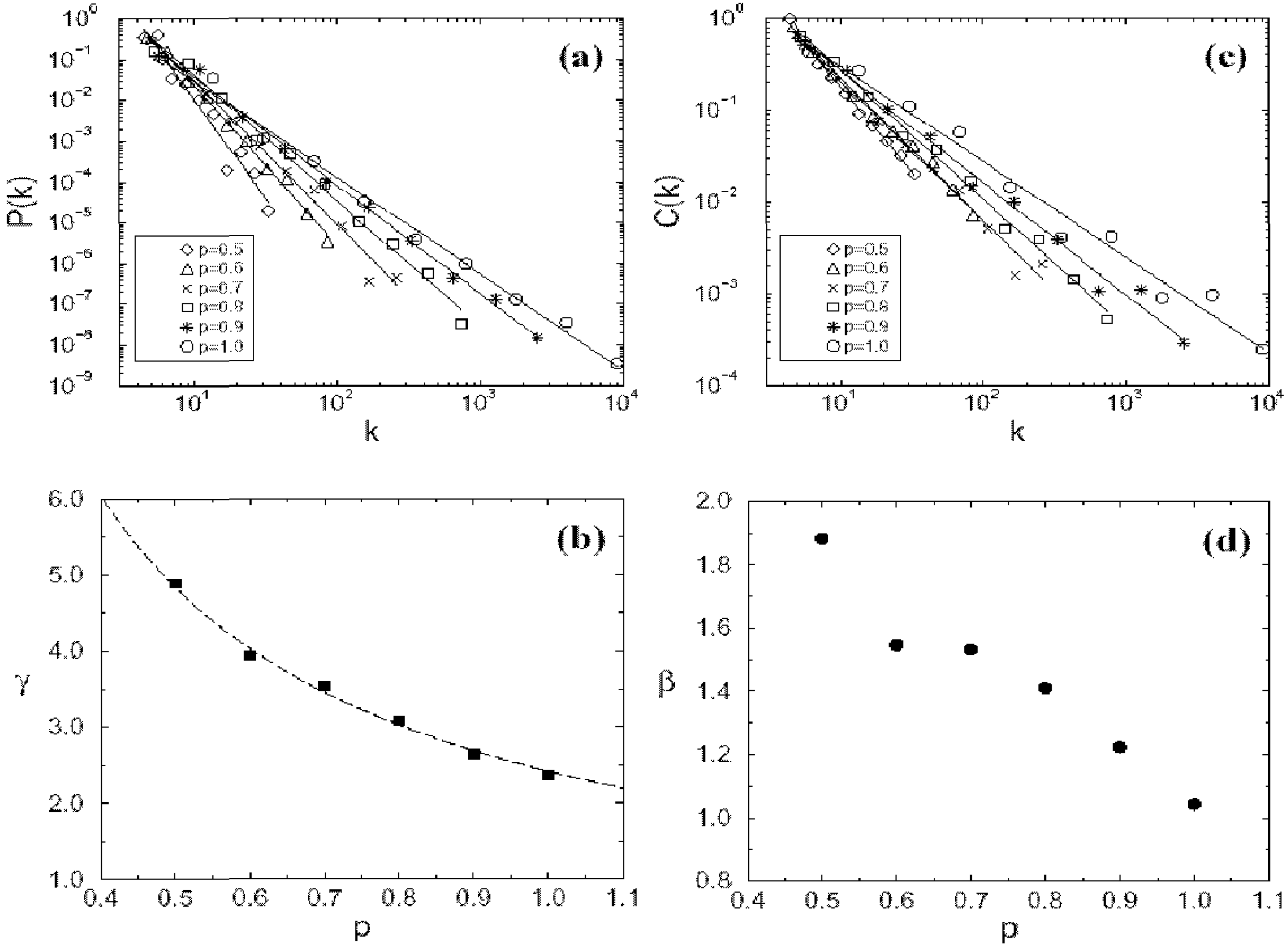
- $C(k) \sim k^{-\beta}$
- For most systems: $\beta \simeq 1$



Connect a *p* fraction of nodes to the central module using preferential attachment

Stochastic Hierarchical Model

$P(k) \sim k^{-\gamma}$ $C(k) \sim k^{-\beta}$



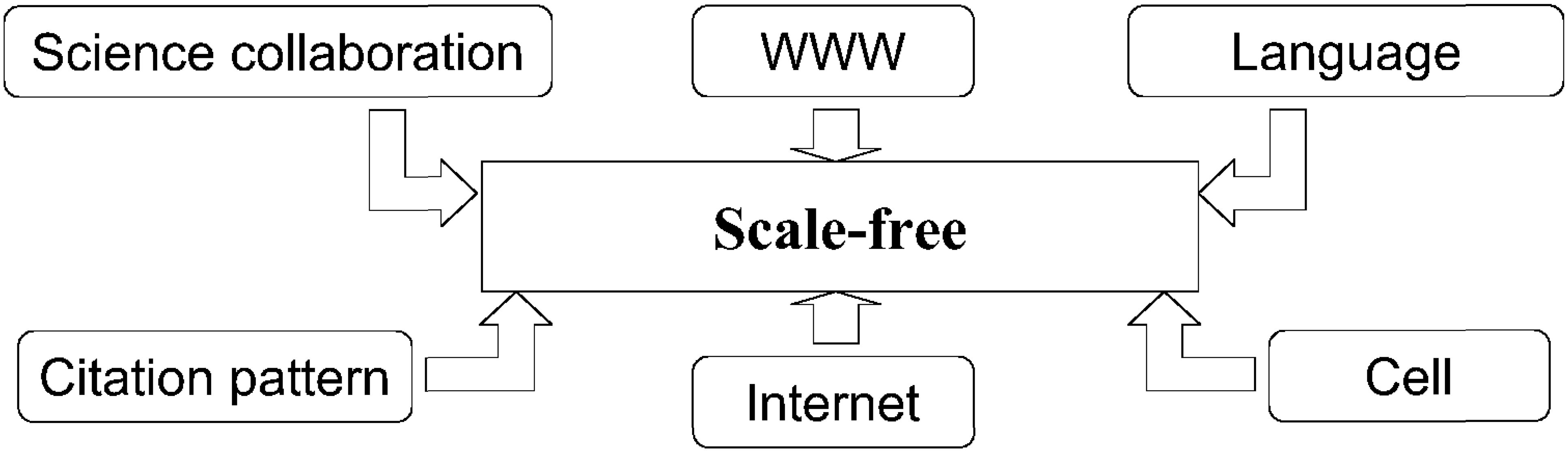
Is hierarchy present in network models?

NO:

-Scale-free model (alb& Albert,1999)
-Erdos-Renyi model (1959)
-Watts-Strogatz (1998)

YES:

Dorogovtsev, Goltsev, Mendes, 2001 (determ.)
-Klemm and Eguiluz, 2002
-Vasquez, Pastor-Satorras,Vespignani (2001)*
$\Rightarrow$ Bianconi & alb (fitnesss model) (2001)



Hierarchical Networks

Scale-free | **Modular** | **Hierarchical**

$P(k) \sim k^{-}$ | $C(N) = const.$ | $C(k) \sim k^{-}$

Traditional modeling: Network as a static graph

Given a network with N nodes and L links

⇓

Create a graph with statistically identical topology

RESULT: model the static network topology

PROBLEM: **Real networks are dynamical systems!**

Evolving networks

OBJECTIVE: capture the network dynamics

METHOD :
- identify the processes that contribute to the network topology
- develop dynamical models that capture these processes

⇓

BONUS: **get the topology correctly**.

Bonus: Why Kevin Bacon?

Measure the average distance between Kevin Bacon and all other actors.

Kevin Bacon

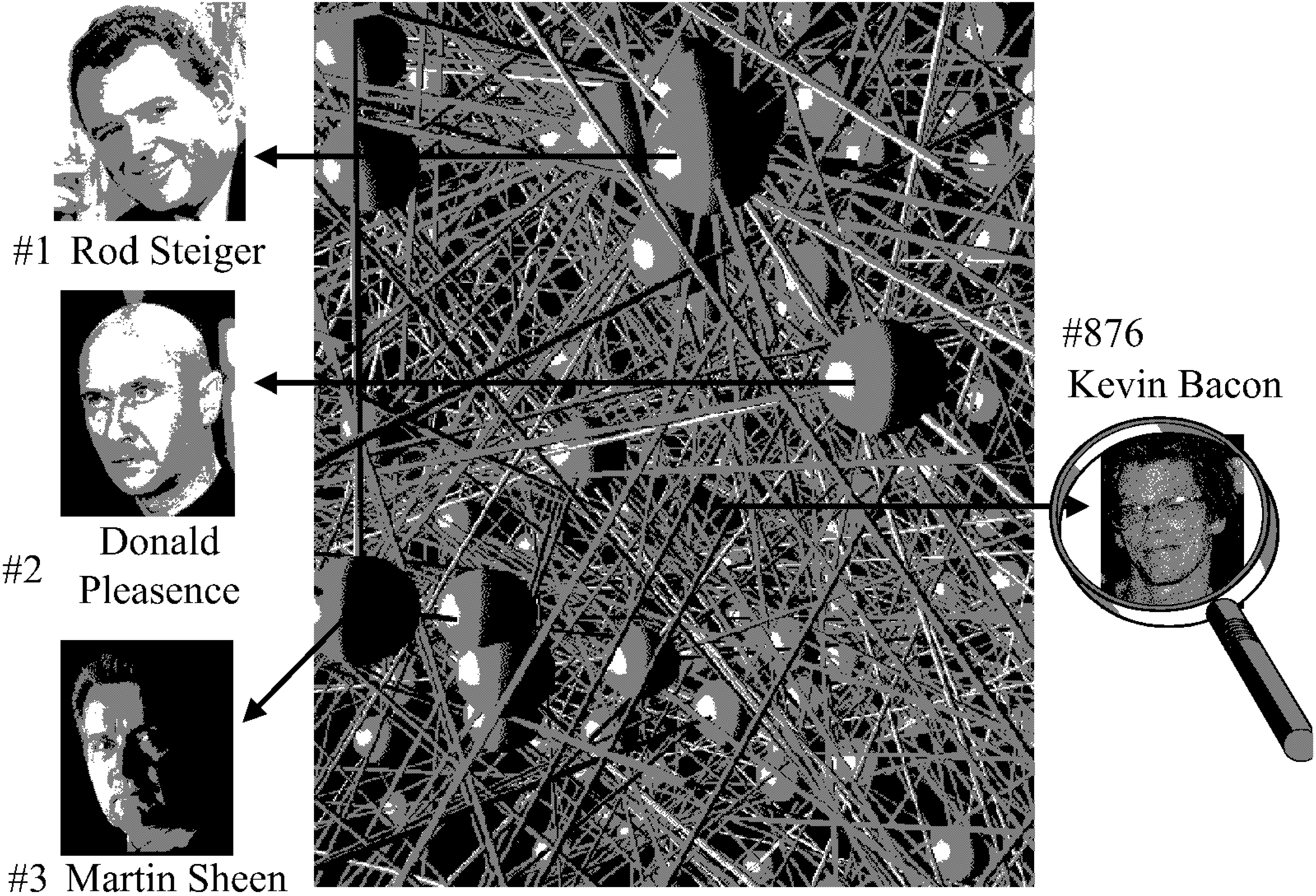
No. of movies : 46   No. of actors : 1811

Average separation: 2.79

Is Kevin Bacon the most connected actor?

NO!

Rank	Name	Average distance	# of movies	# of links
1	Rod Steiger	2.537527	112	2562
2	Donald Pleasence	2.542376	180	2874
3	Martin Sheen	2.551210	136	3501
4	Christopher Lee	2.552497	201	2993
5	Robert Mitchum	2.557181	136	2905
6	Charlton Heston	2.566284	104	2552
7	Eddie Albert	2.567036	112	3333
8	Robert Vaughn	2.570193	126	2761
9	Donald Sutherland	2.577880	107	2865
10	John Gielgud	2.578980	122	2942
11	Anthony Quinn	2.579750	146	2978
12	James Earl Jones	2.584440	112	3787
…				
876	**Kevin Bacon**	**2.786981**	**46**	**1811**
…				


#1 Rod Steiger
#2 Donald Pleasence
#3 Martin Sheen
#876 Kevin Bacon

http://www.nd.edu/~networks

References

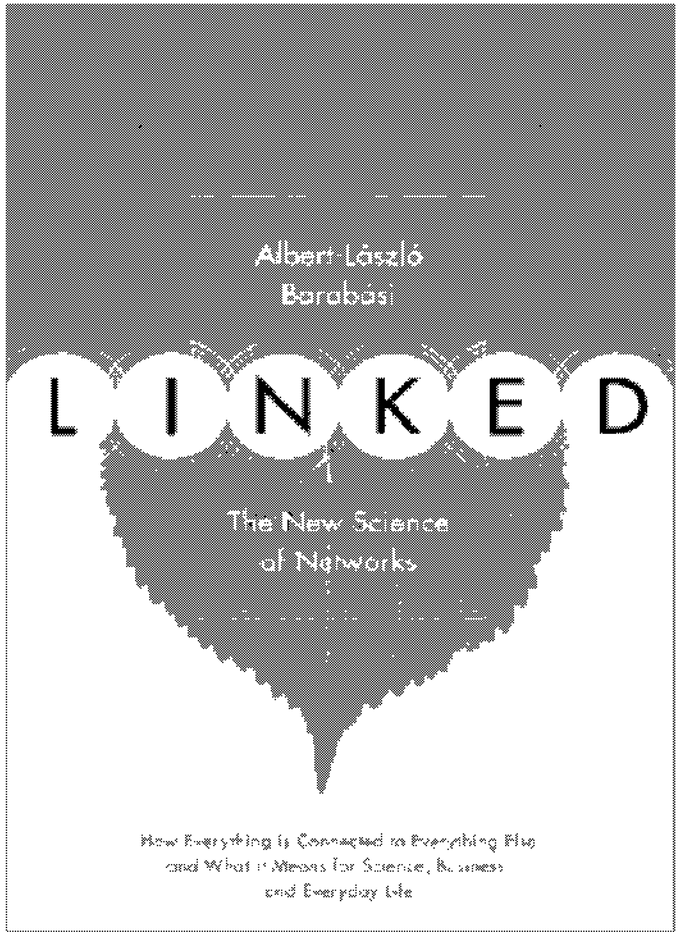
- R. Albert, H. Jeong, A.L. Barabasi, Nature **401** 130 (1999).
- R. Albert, A.L. Barabasi, Science **286** 509 (1999).
- A.L. Barabási, R. Albert and H. Jeong, Physica A **272,** 173 (1999)
- R. Albert, H. Jeong, A.L. Barabasi, Nature **406** 378 (2000).
- H.Jeong, B.Tombor, R.Albert, Z.N.Oltvai, A.L.Barabasi, Nature **407** 651 (2000).
- H. Jeong, S.P. Mason, A.L. Barabasi, Z.N. Oltvai, Nature (in press).

URL: http://www.nd.edu/~networks

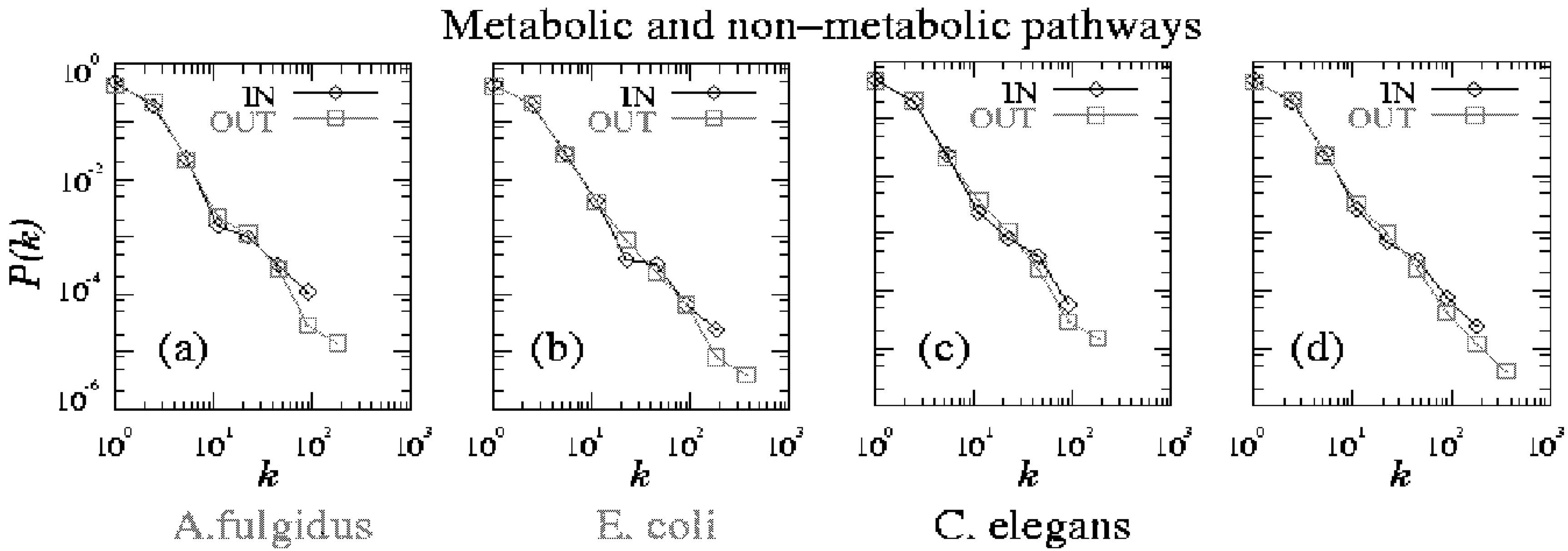
http://www.nd.edu/~networks

Complex Networks &
Biological Applications
Thursday
Room 102
10:30—18:30

Mark Newman
Andrea Levchenko
Lou Pecora
Gabor Forgacs
Allesandro Vespignani
Reka Albert
Gilberto Thomas
Byungnam Kahng
Hawoong Jeong

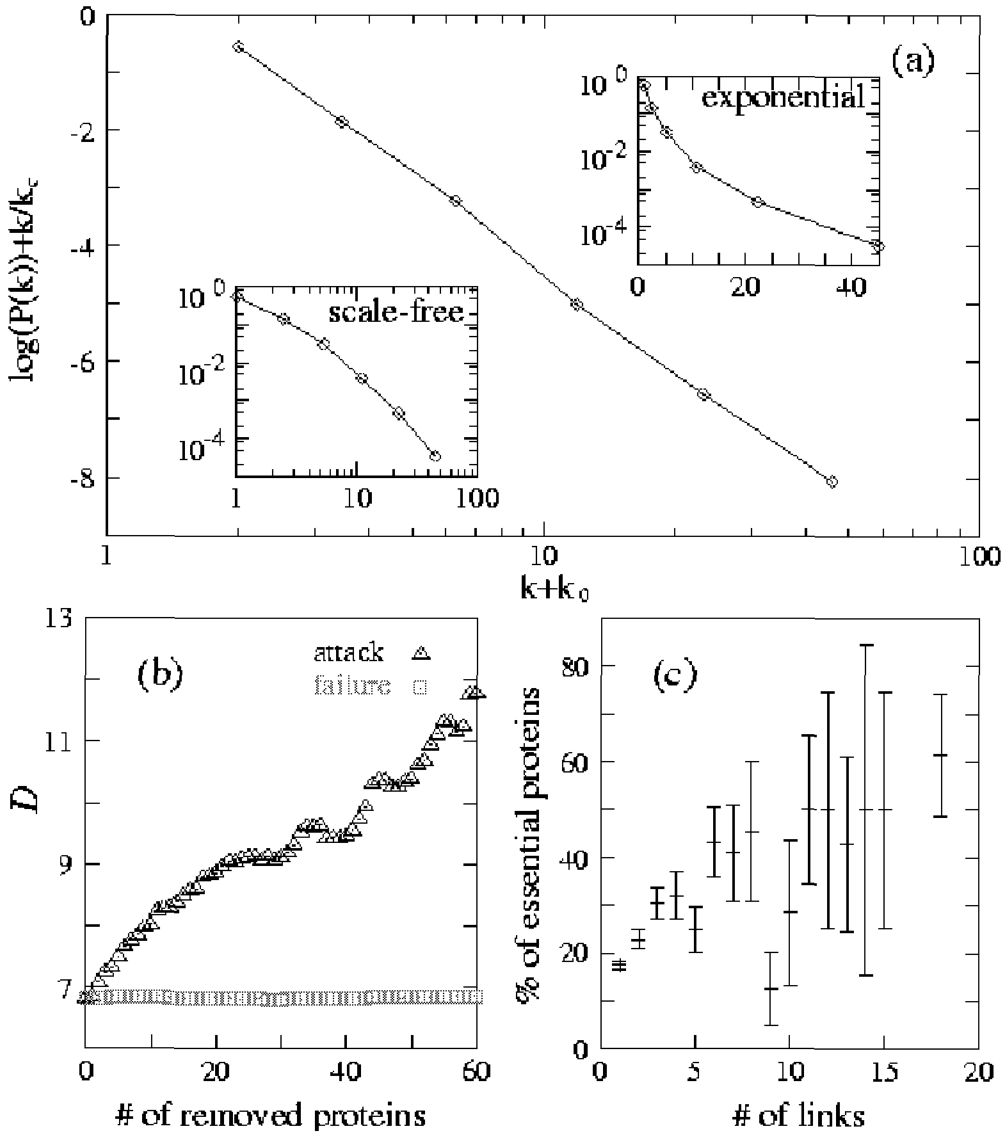


Whole cellular network



Metabolic pathways	Non–metabolic pathways
intermediate metabolism	information pathways
bioenergetics	electron, transmembrane transport
	signal transduction
	structure and function of the cell

Properties of the protein network



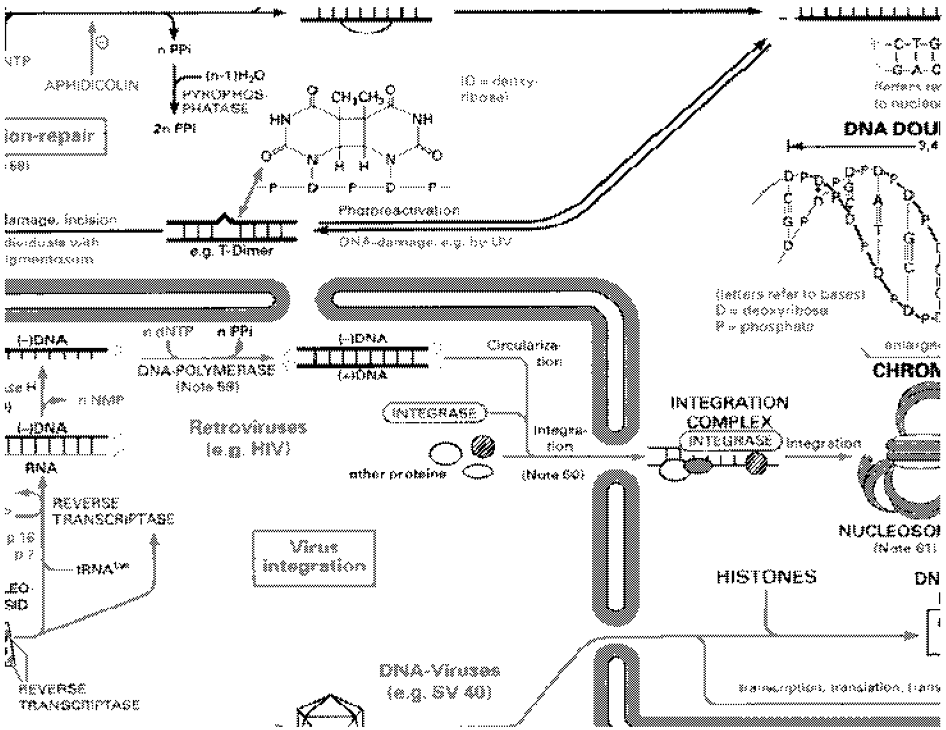
$$P(k) \sim (k + k_0)^{-\gamma} \exp(-\frac{k + k_0}{k_\tau})$$

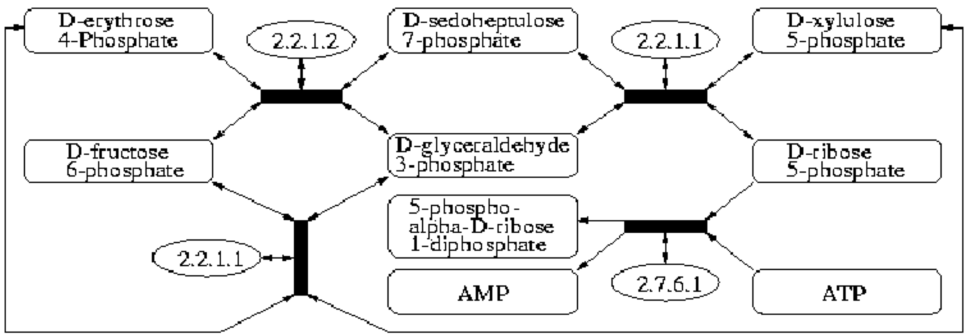
Highly connected proteins are more **essential (lethal)** than less connected proteins.

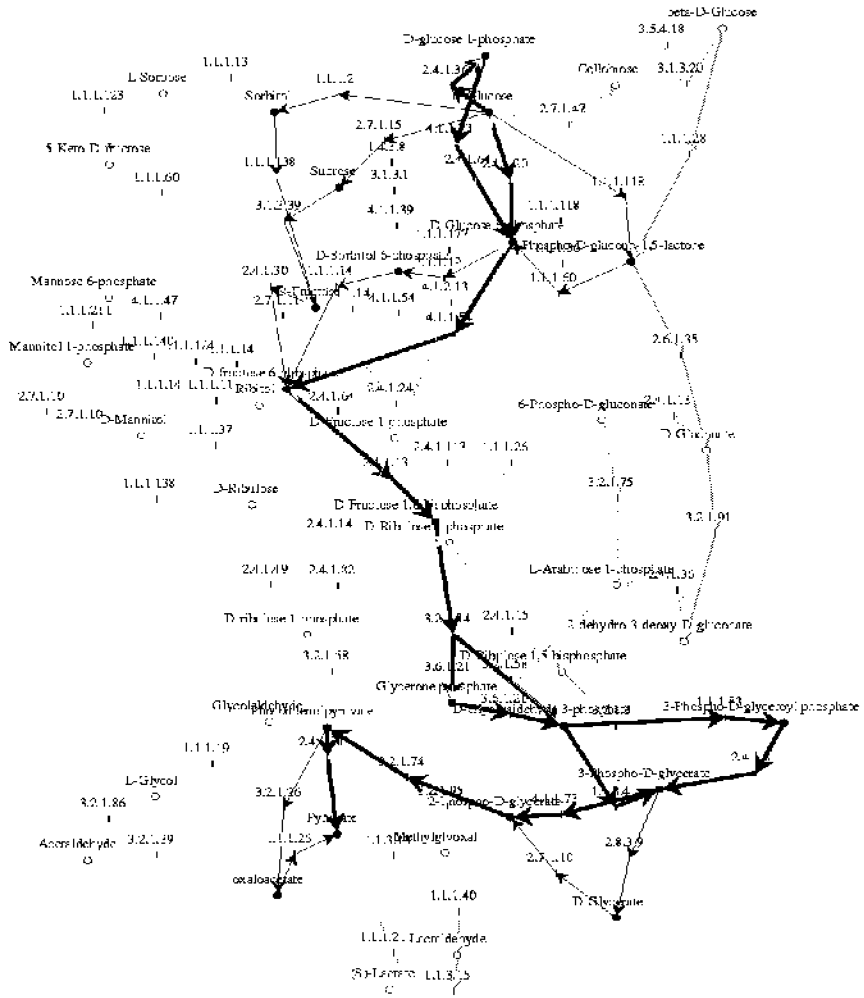
Metabolic Network

Nodes: chemicals (substrates)

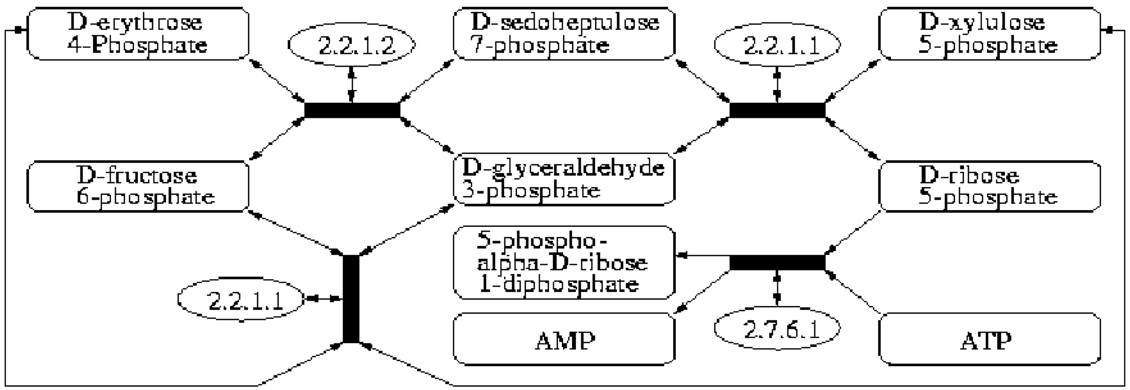
Links: chem. reaction

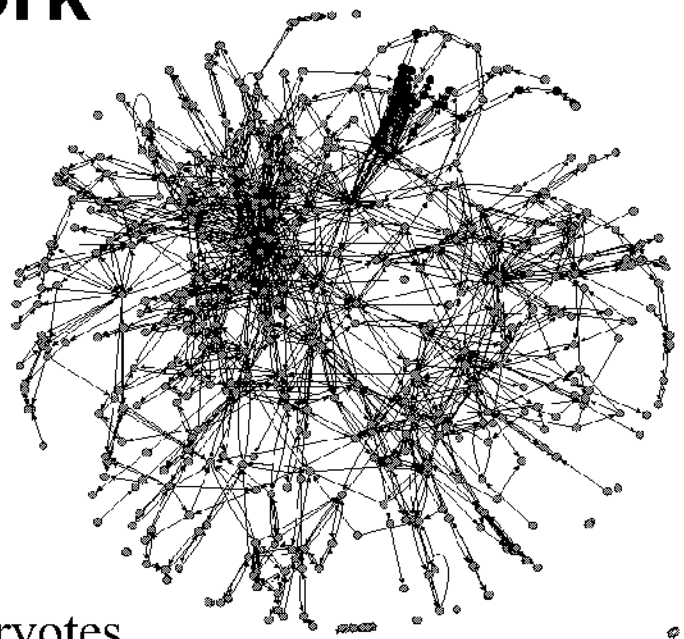


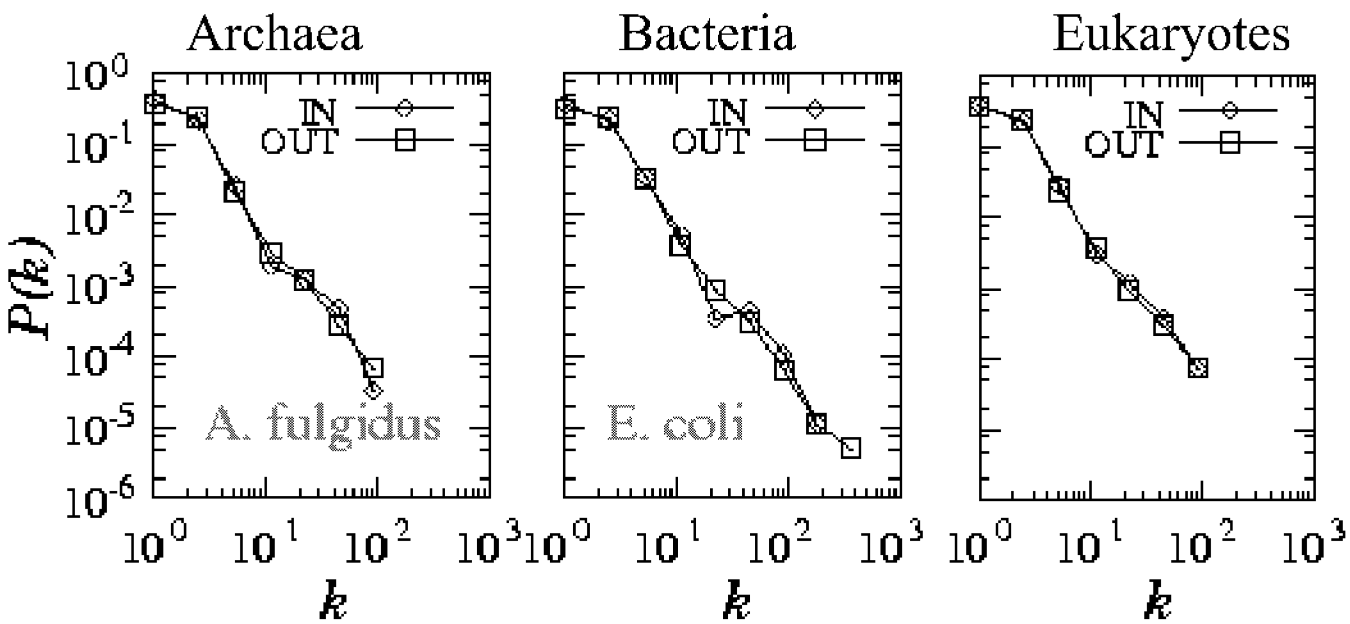




Metabolic network



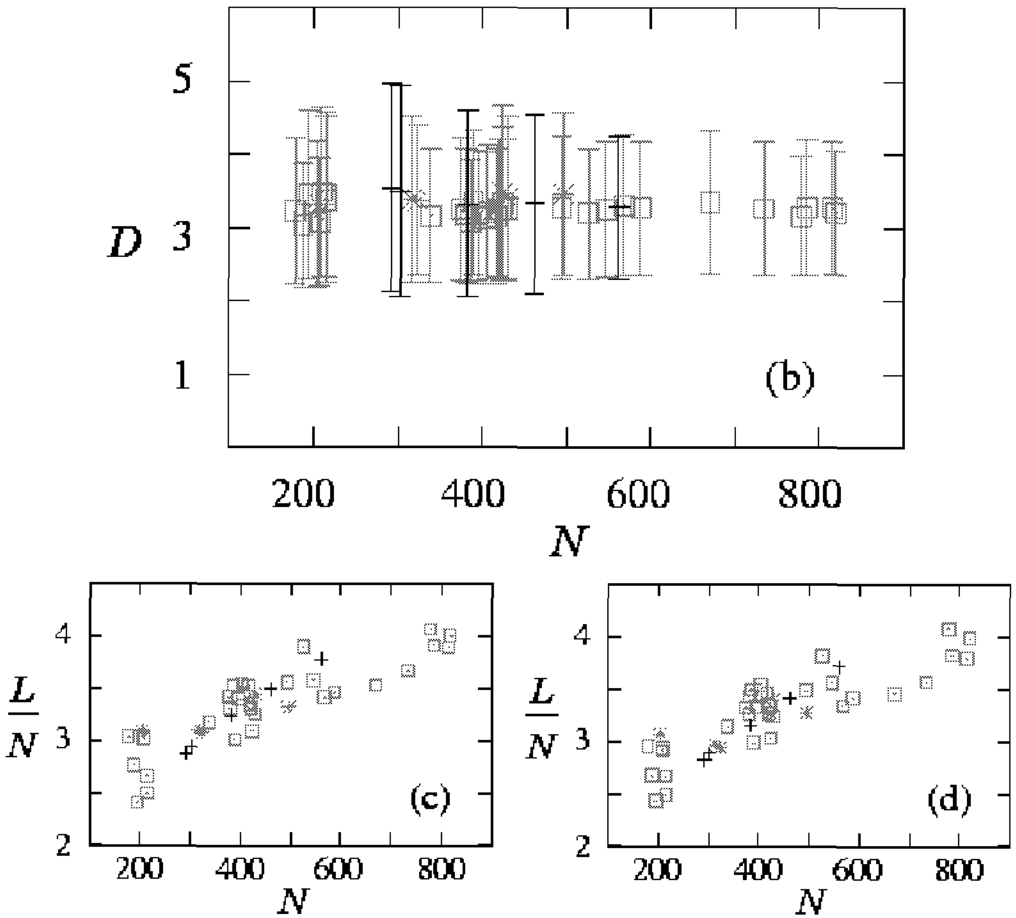




Organisms from all three domains of life are **scale-free** networks!

H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, Nature, **407** 651 (2000)

Properties of metabolic networks



Average distances are independent of organisms!

⇐ by making more links between nodes.

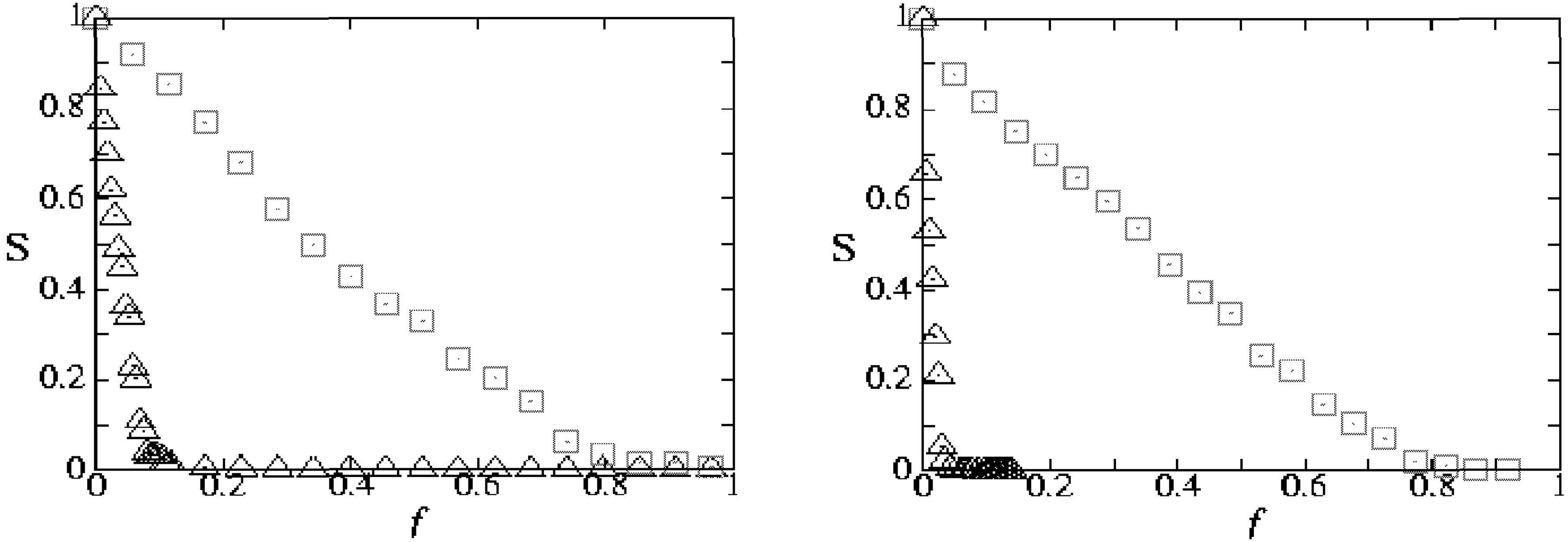
⇐ based on **"design principles"** of the cell through **evolution.**

cf. Other scale-free network: ***D~log(N)***

Achilles' Heel of complex network

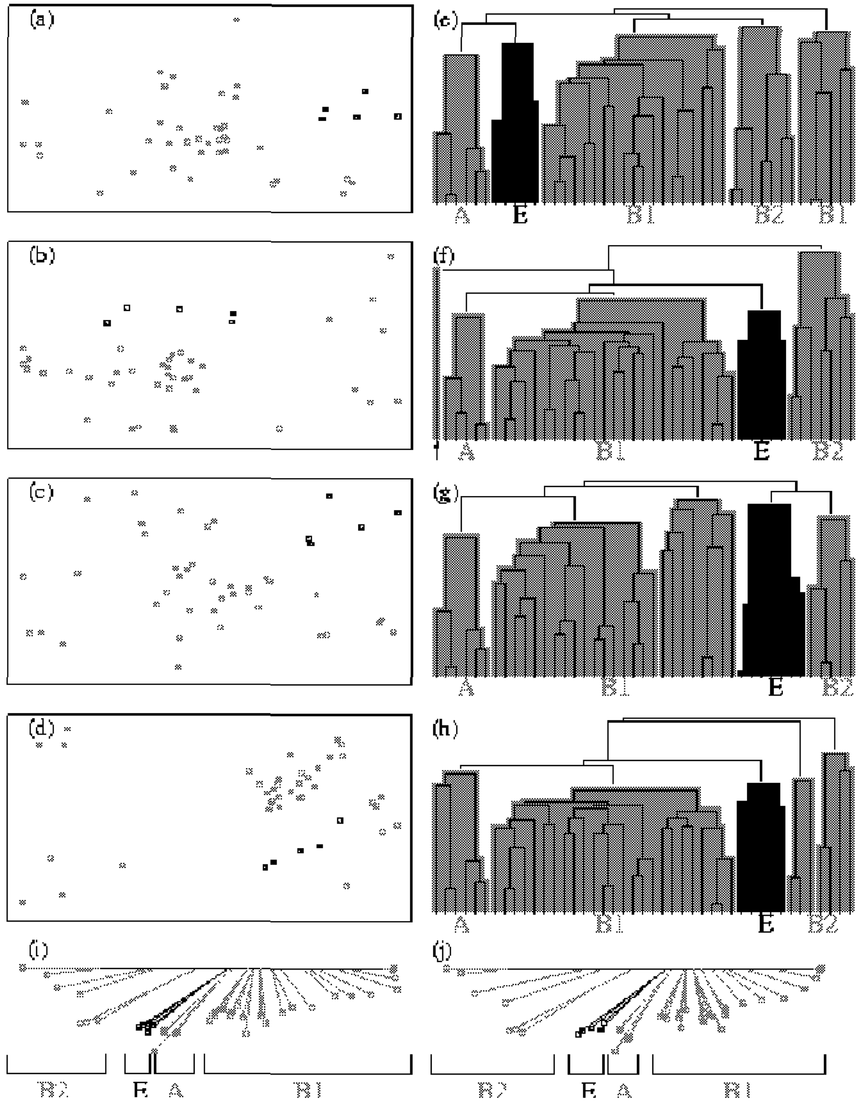
failure
attack

Internet **Protein network**



R. Albert, H. Jeong, A.L. Barabasi, Nature **406** 378 (2000)

Taxonomy using networks



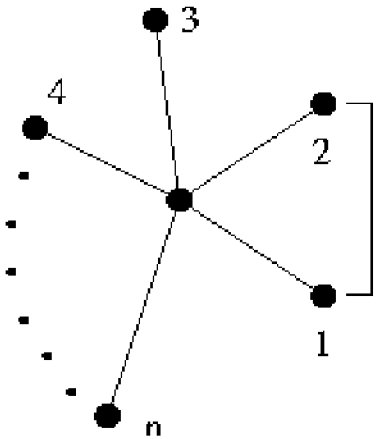
A: Archaea

B: Bacteria

E: Eukaryotes

Watts-Strogatz

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

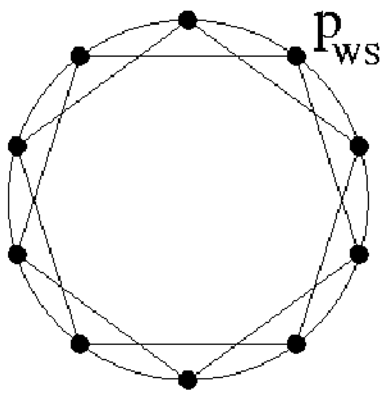


Probability to be connected $C \gg p$

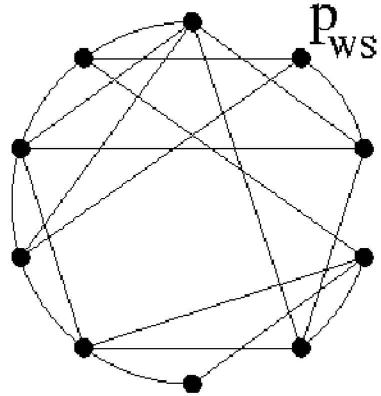
$$C = \frac{\text{\# of links between 1,2,…n neighbors}}{n(n-1)/2}$$

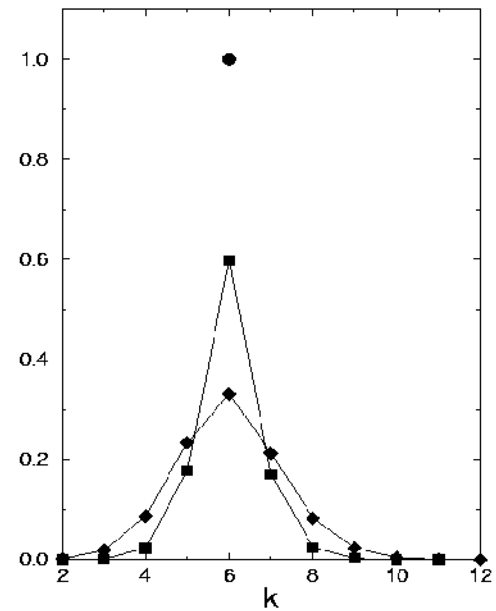
N nodes forms a regular lattice.
With probability p, each
edge is rewired randomly.

$p_{ws}=0$



$p_{ws}=0.3$

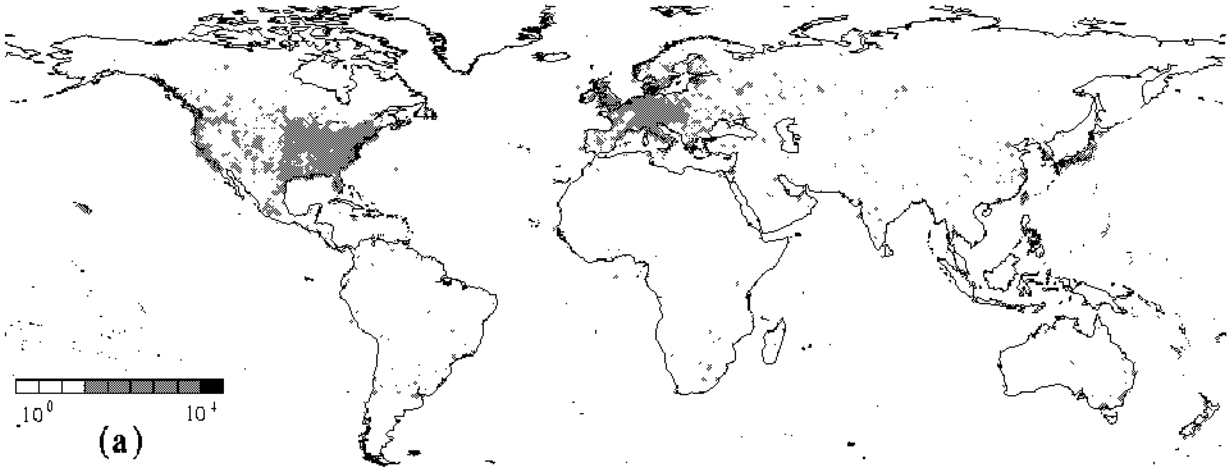




(Nature **393,** 440 (1998))

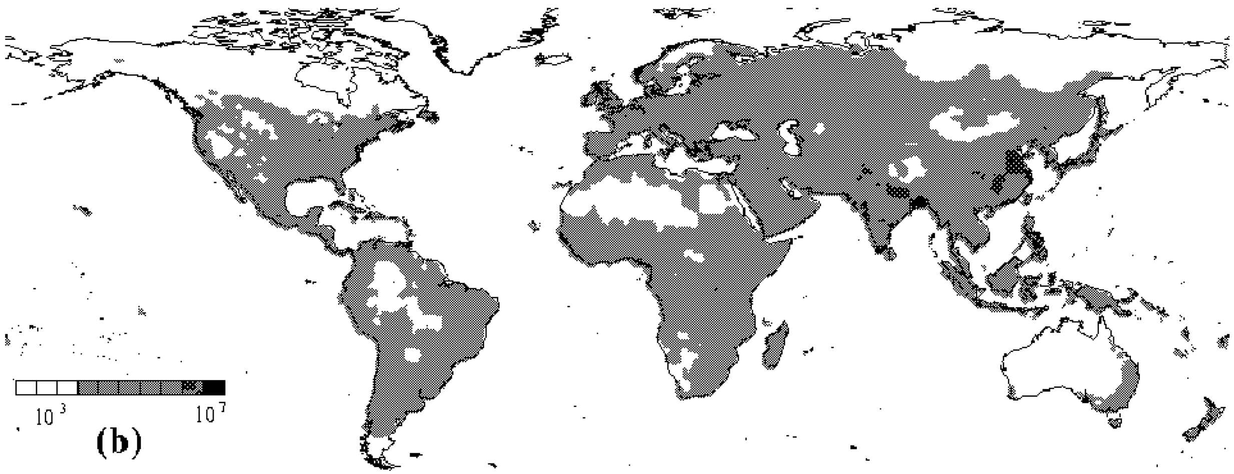
Spatial Distributions

Router density



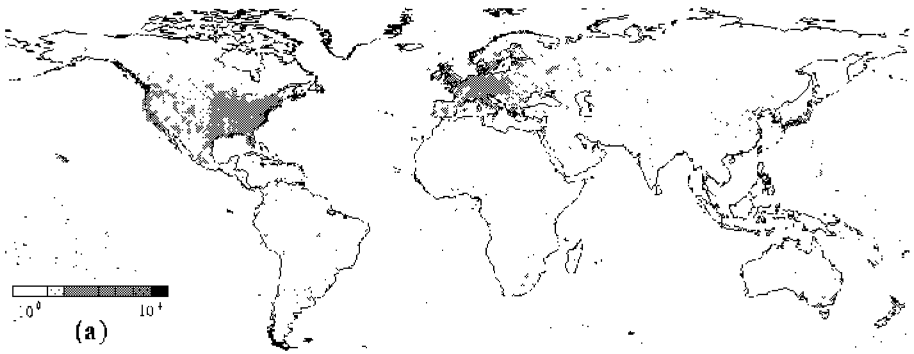
(a)

Population density



(b)

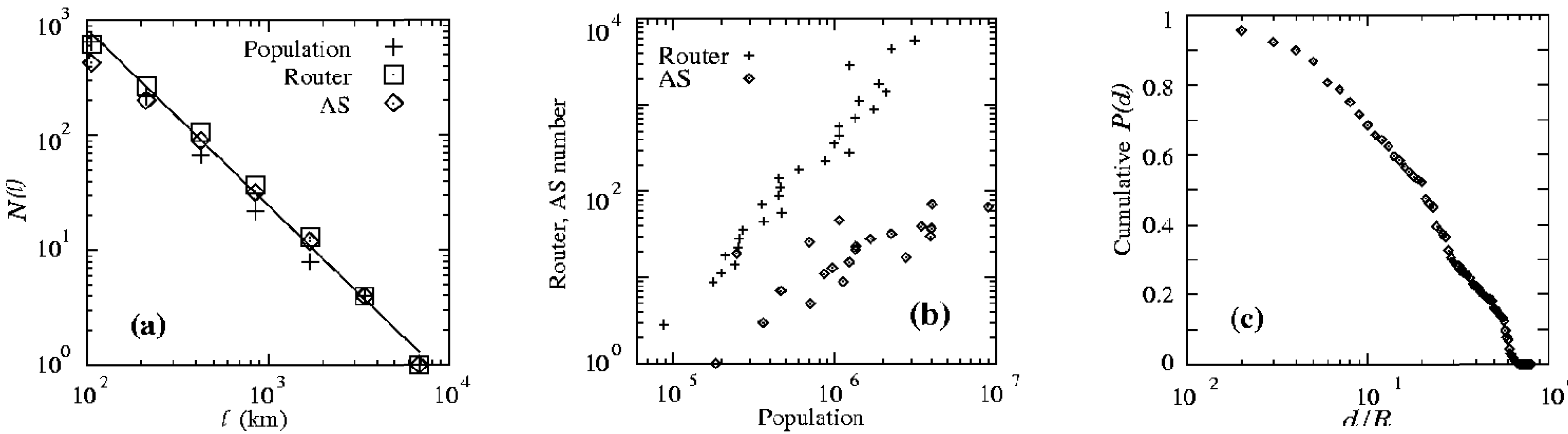
Spatial Distribution of Routers



Fractal set

Box counting: $N(\ell) \equiv$ No. of boxes of size ℓ that contain routers

$N(\ell) \sim \ell^{-D_f}$ $D_f=1.5$

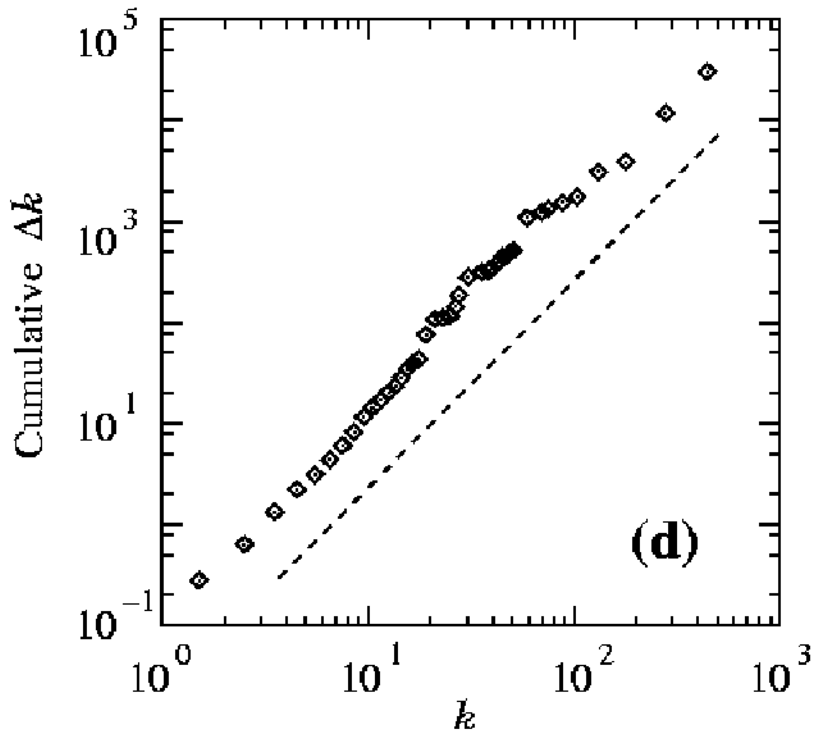


Preferential Attachment

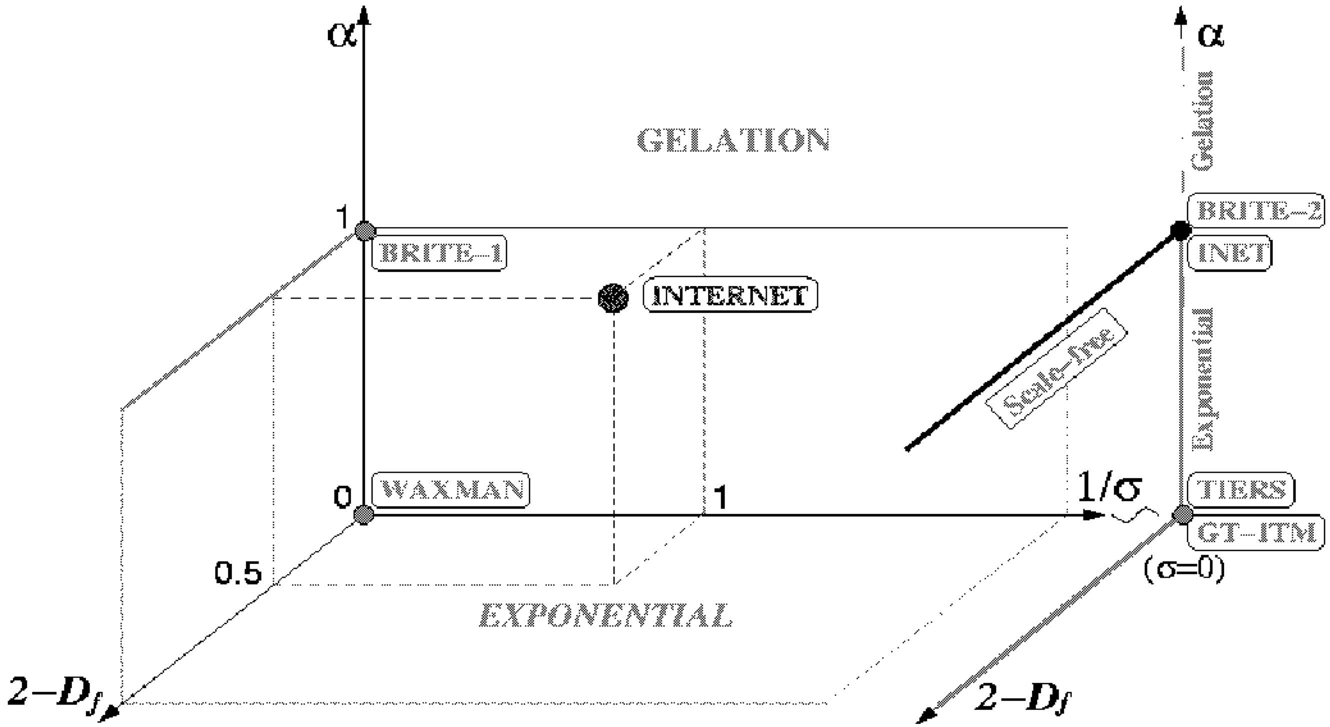
- Compare maps taken at different times (Δt = 6 months)
- Measure $\Delta k(k)$, increase in No. of links for a node with k links

Preferential Attachment:

$$\Delta k(k) \sim k$$



INTERNET



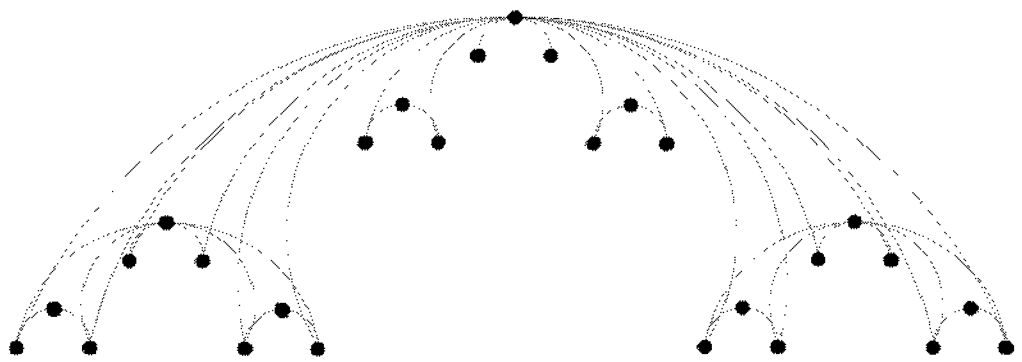
$N(\ell) \sim \ell^{-D_f}$ D_f=1.5

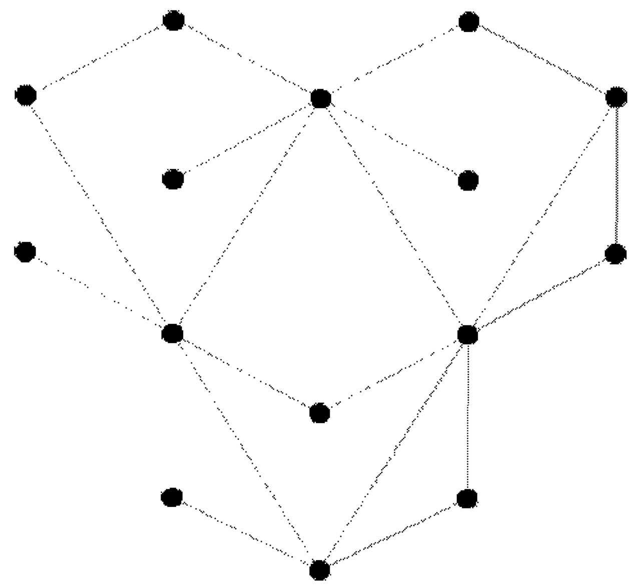
$\Delta k(k) \sim k^{\alpha}$ α=1

$P(d) \sim d^{-\sigma}$ σ=1

Combining Modularity and the Scale-free Property

Deterministic Scale-Free Networks

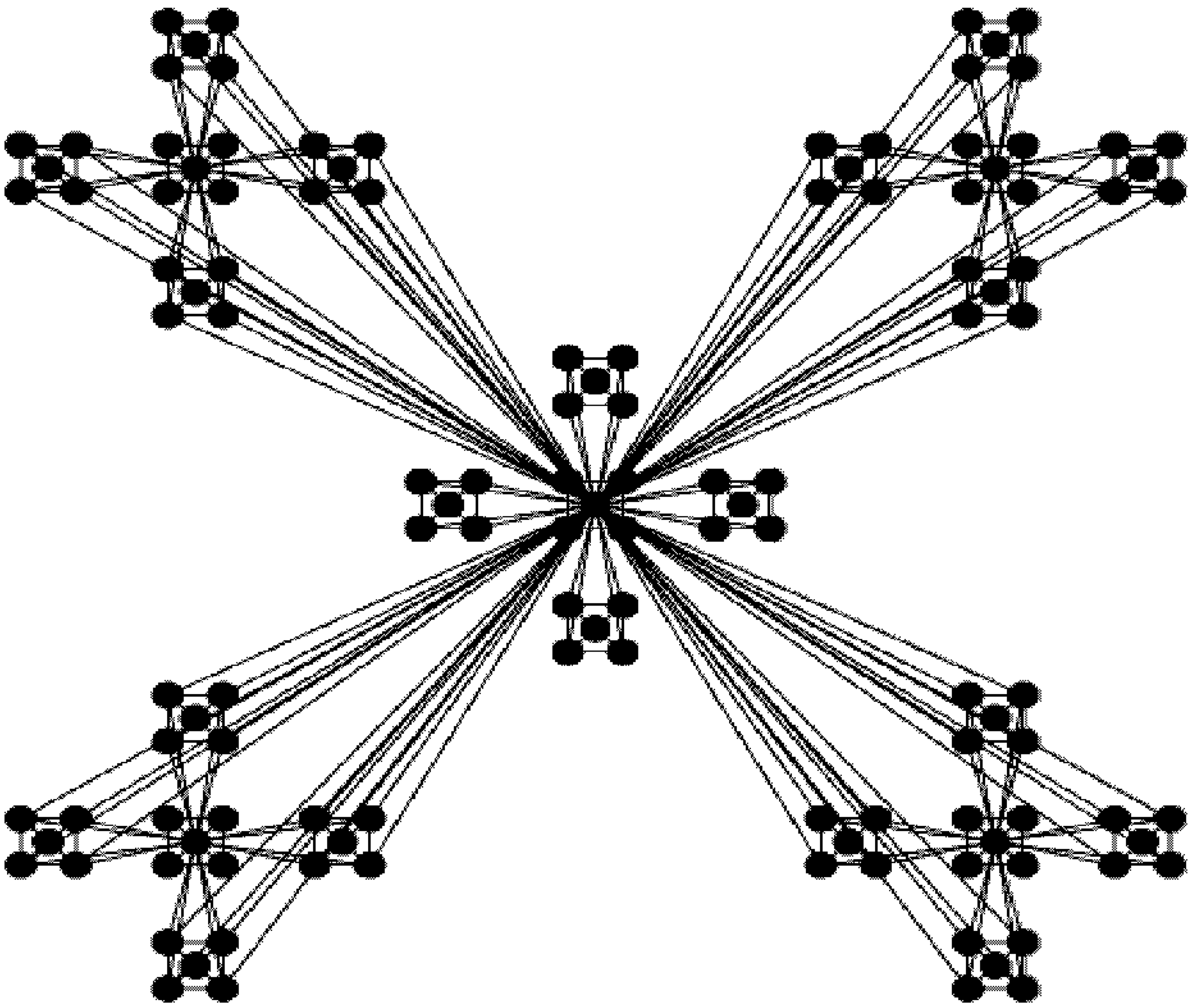




Barabási, A.-L., Ravasz, E., & Vicsek, T. (2001) *Physica A* **299**, 559.

Dorogovtsev, S. N., Goltsev, A. V., & Mendes, J. F. F. (2001) cond-mat/0112143. (DGM)

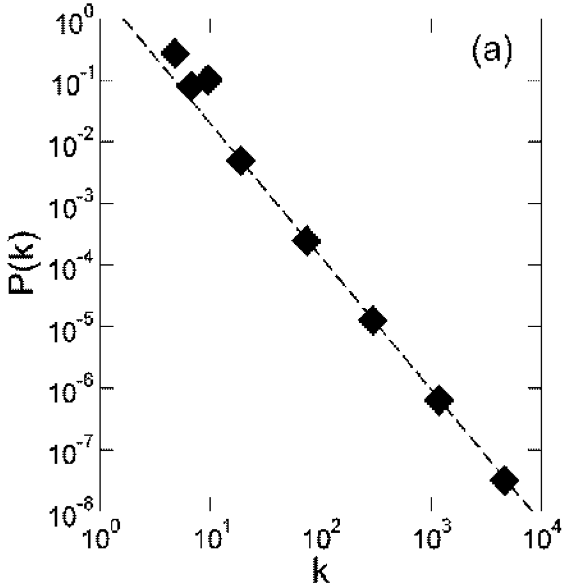
Hierarchical Networks



Properties of hierarchical networks

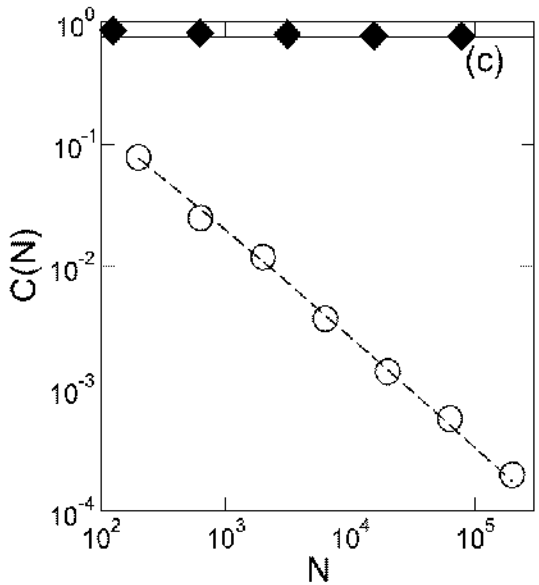
1. Scale-free

$$\gamma = 1 + \frac{\ln 5}{\ln 4} = 2.161$$



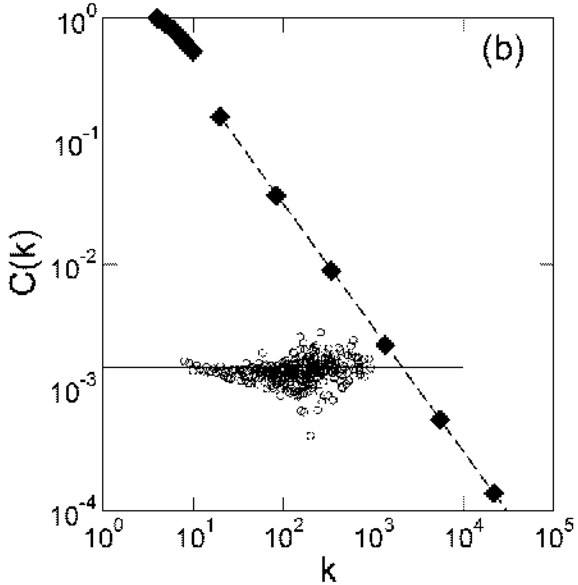
2. Clustering coefficient independent of N

$$C(N) = \text{const.}$$

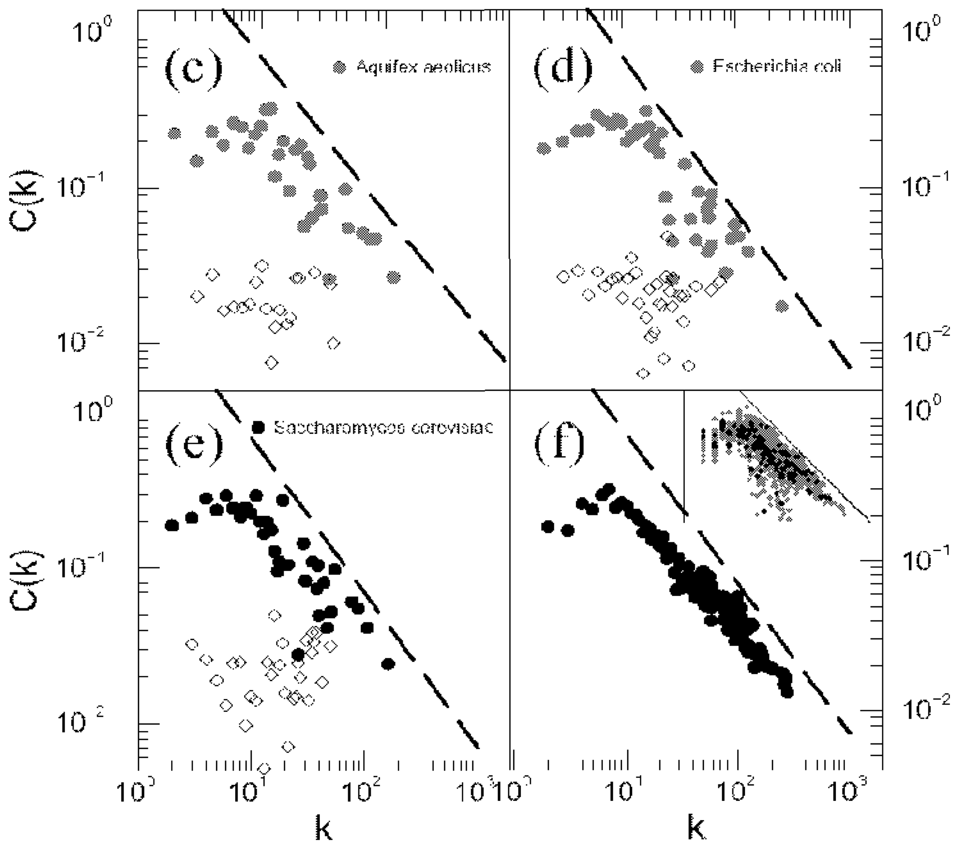


3. Scaling clustering coefficient (DGM)

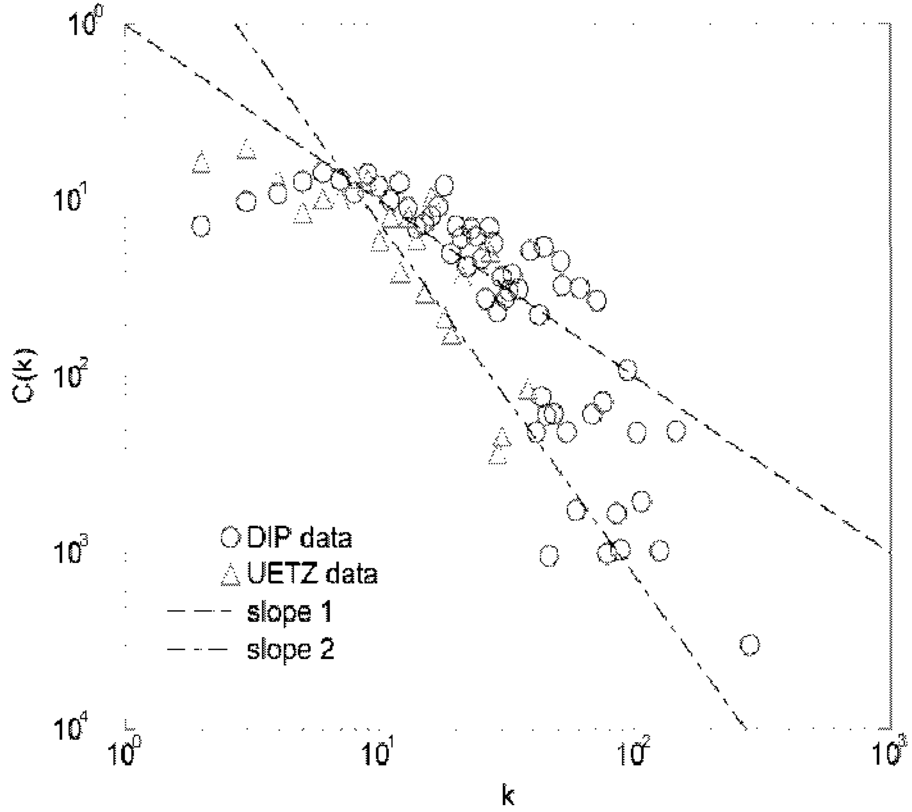
$$C(k) \sim k^{-1}$$



Hierarchy in biological systems



Metabolic networks

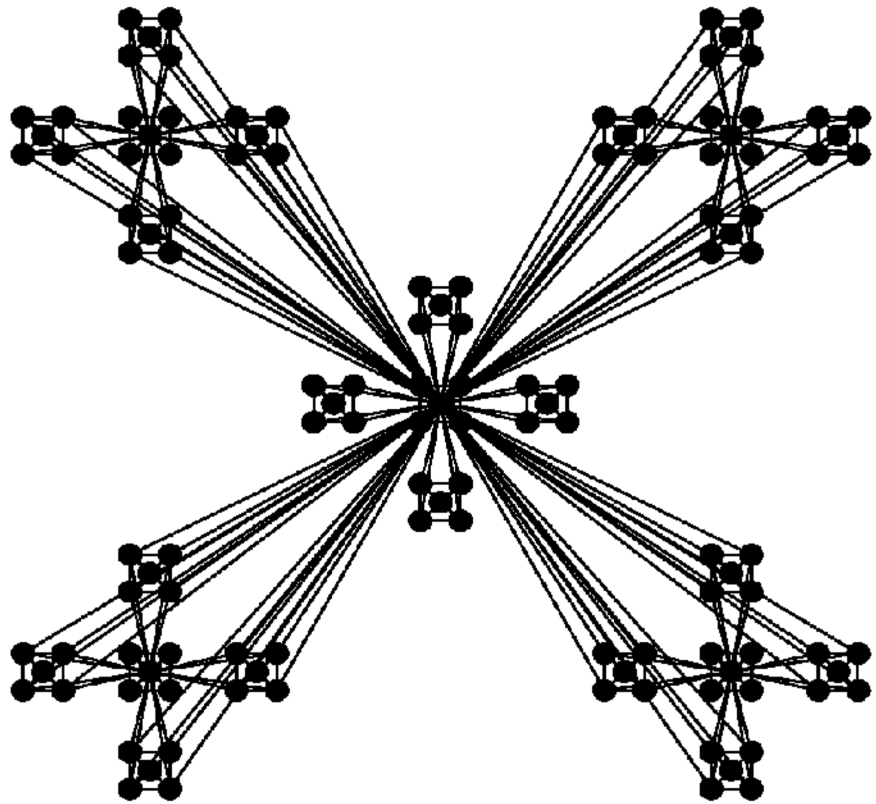


Protein networks

Real Networks Have a Hierarchical Topology

What does it mean?

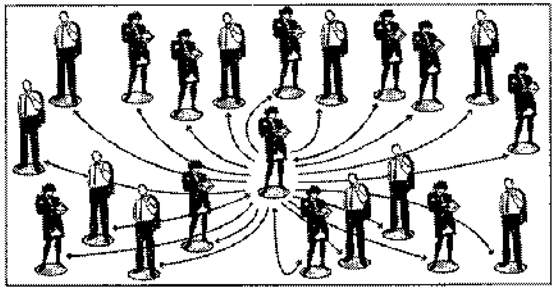
Many highly connected small clusters
combine into
few larger but less connected clusters
combine into
even larger and even less connected clusters



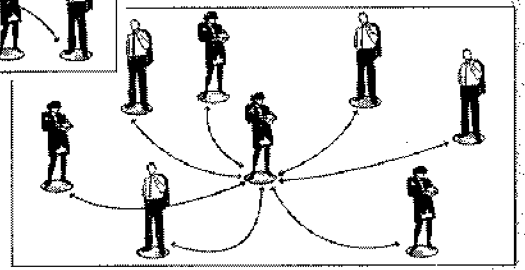
- The degree of clustering follows:

$$C(k) \sim k^{-\beta}$$

Society



Mega-Hub. An MTV veejay spreads the word to thousands or millions of people through one-way links.



Hub. This undergraduate has spread the word to seven other people through two-way links.

Internet