Complex Networks
Complexity, Complex Systems
(graphs)

What are they, history, examples, pictorials, current problems
How are they related to Physics, Sociology, Economics…….

graphs ~1930s

networks ~ 1998

Panos Argyrakis

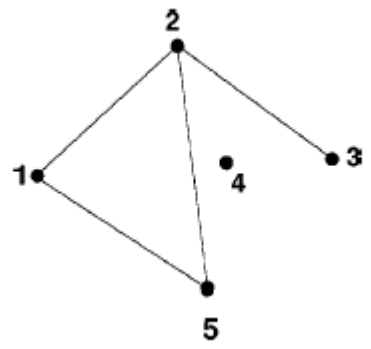
University of Thessaloniki

Characteristics of networks:

- **Structures that are formed by two distinct entities:**

  nodes, sites, vertices, etc.

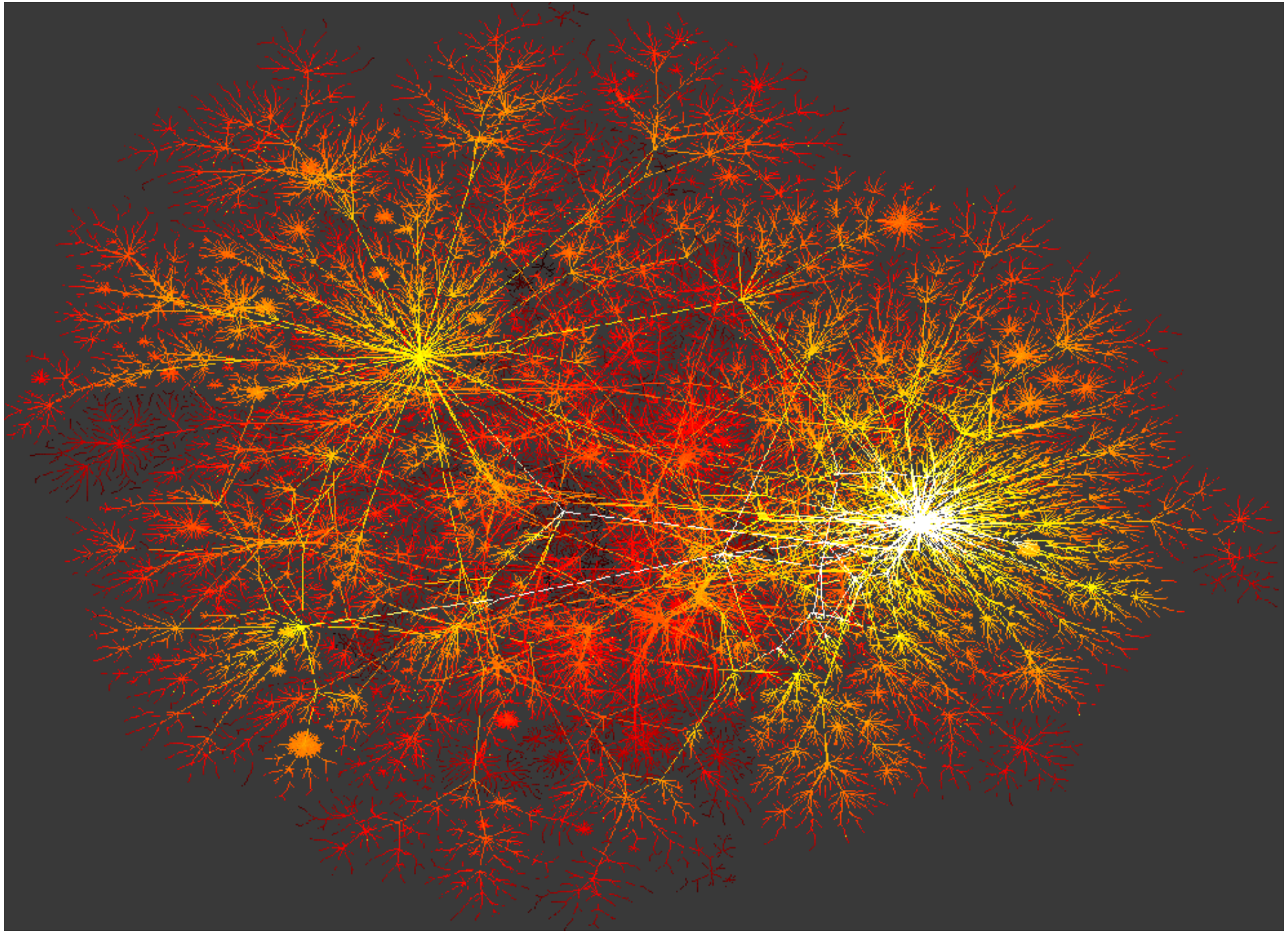
  connections, edges, links, synapses, etc.

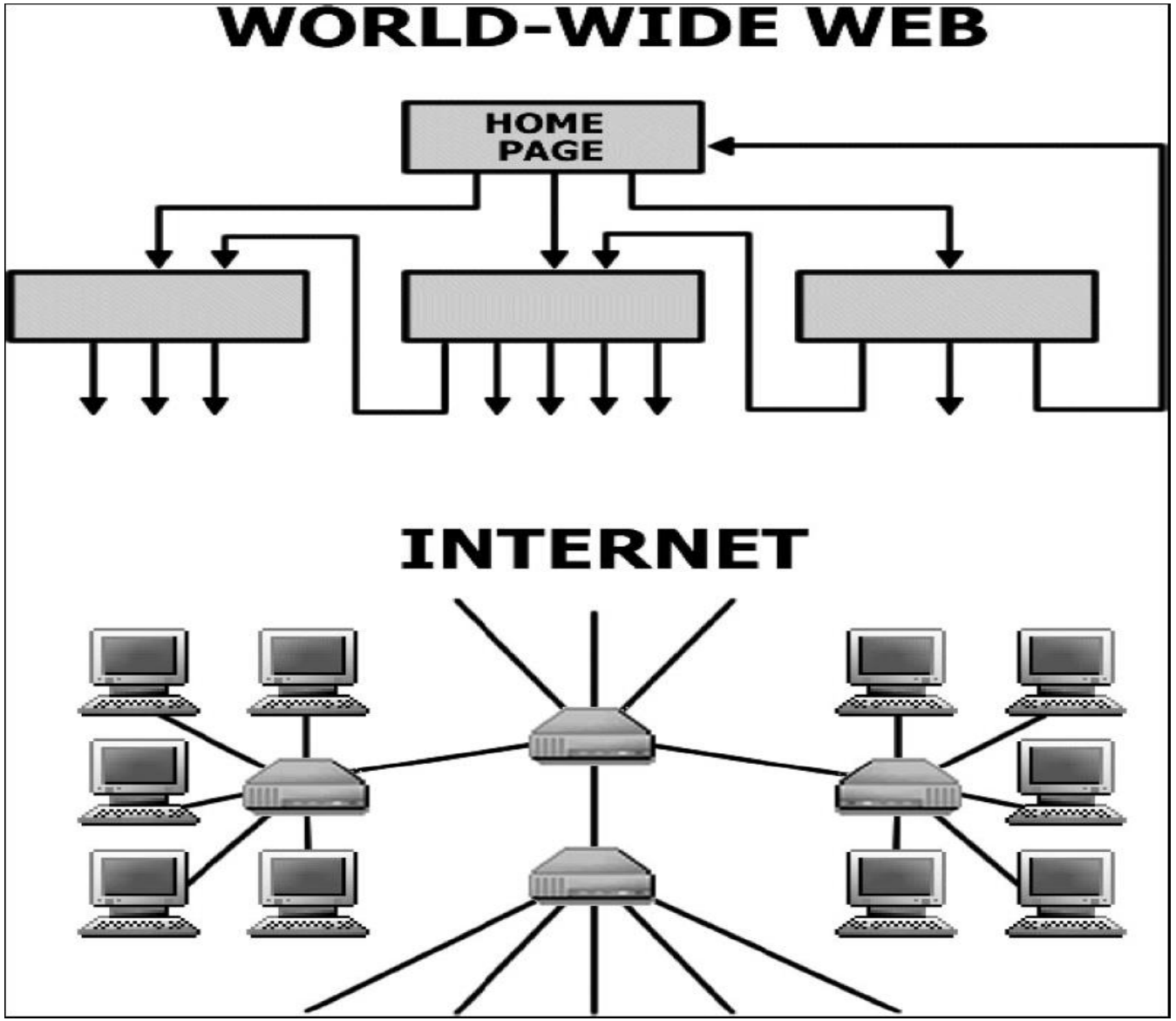


N= 5 nodes, n=4 connections

- **They can be simple structures or very complicated (belonging to the class of complex systems)**

- **How did it all start?**

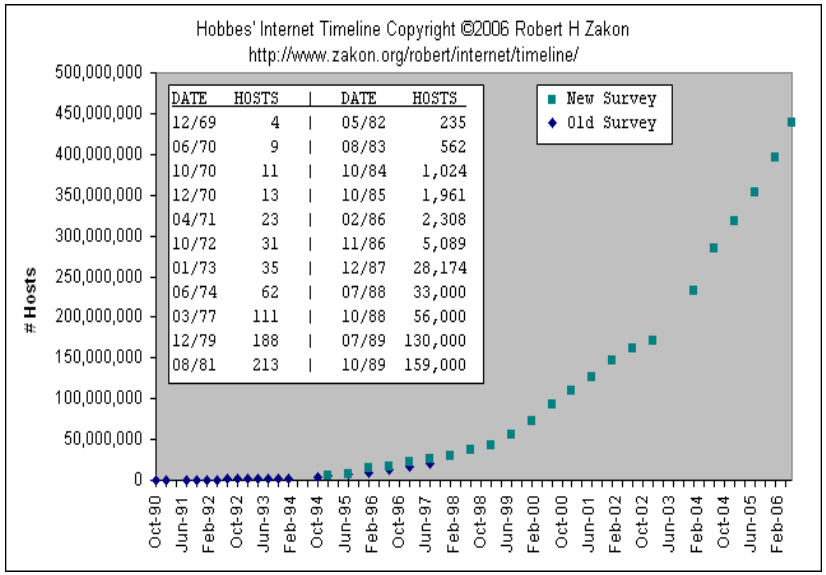



WORLD-WIDE WEB
HOME PAGE
INTERNET

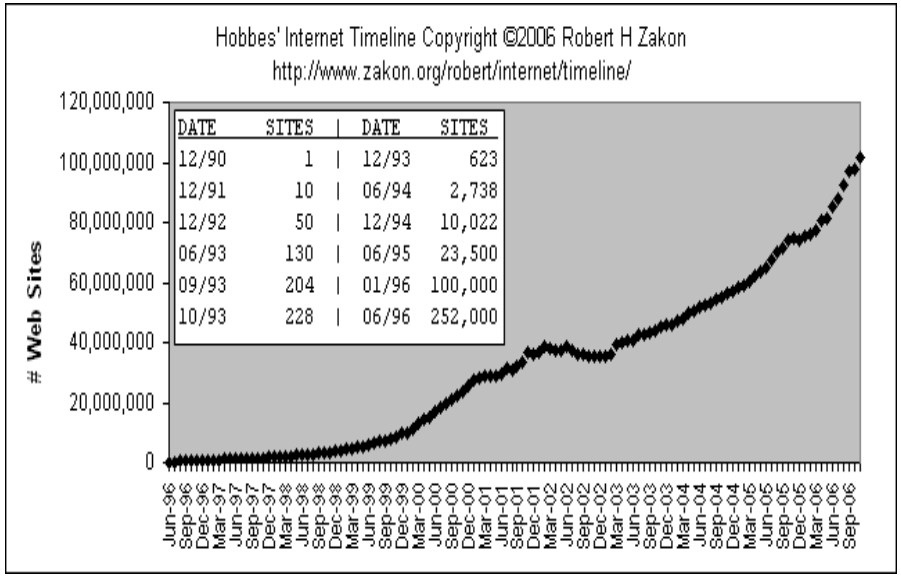
Internet and WWW

1970 – 10 H/Y
1990 – $1.75*10^5$ H/Y
σήμερα – $1.2*10^9$ H/Y

1990 – 1 web site
1996 – 10^5 web sites
σήμερα – 10^9 web sites



DATE	HOSTS	DATE	HOSTS
12/69	4	05/82	235
06/70	9	08/83	562
10/70	11	10/84	1,024
12/70	13	10/85	1,961
04/71	23	02/86	2,308
10/72	31	11/86	5,089
01/73	35	12/87	28,174
06/74	62	07/88	33,000
03/77	111	10/88	56,000
12/79	188	07/89	130,000
08/81	213	10/89	159,000



DATE	SITES	DATE	SITES
12/90	1	12/93	623
12/91	10	06/94	2,738
12/92	50	12/94	10,022
06/93	130	06/95	23,500
09/93	204	01/96	100,000
10/93	228	06/96	252,000

Number of units worldwide

- **Population of Earth ~ 7 billion**
- **Number of computers ~ 1.2 billion**
- **Facebook accounts ~ 1.1 billion**
- **Number of cell phones ~ 8 billion in 2016**

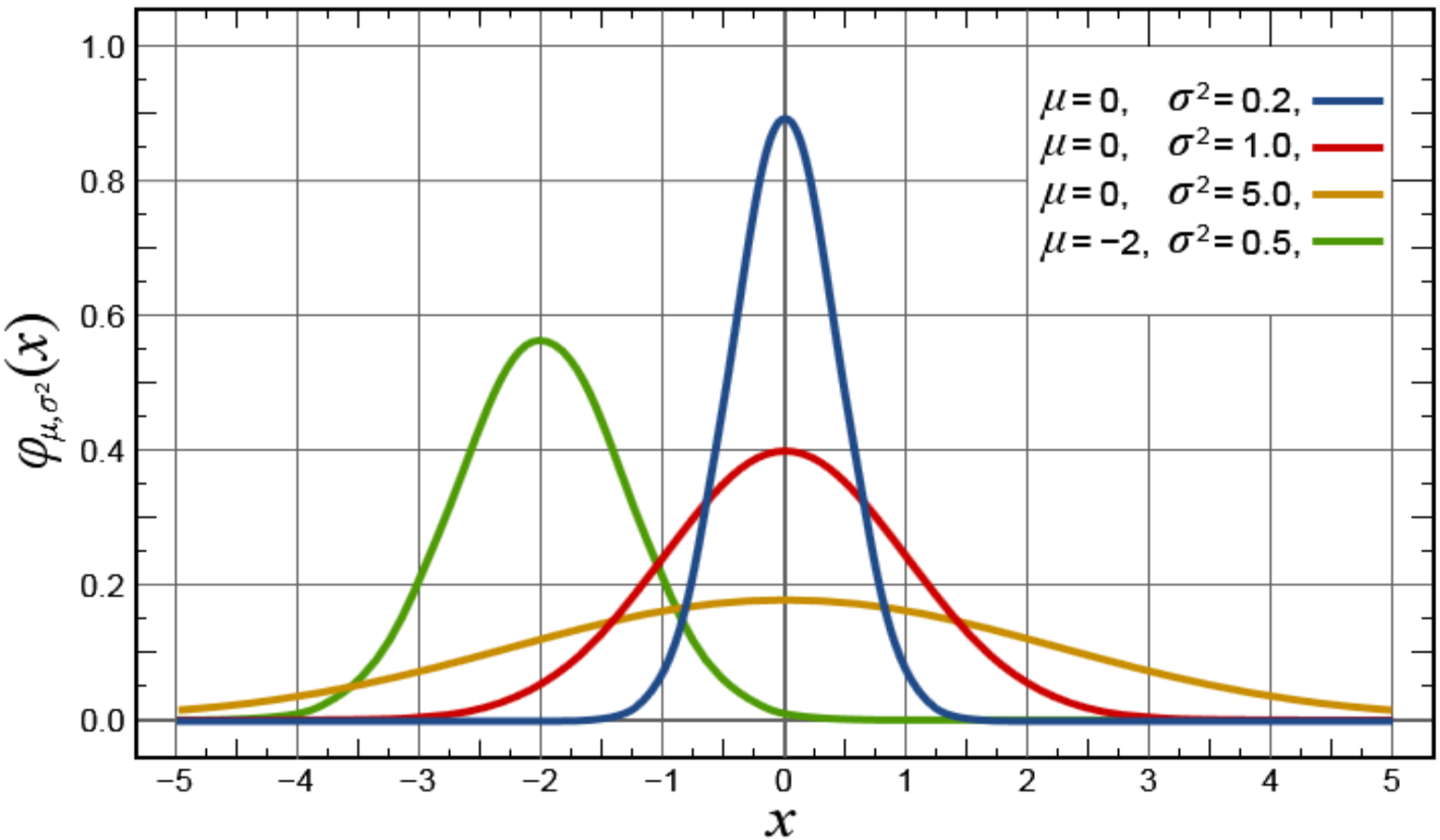
Connectivity between nodes in the Internet

- How are the different nodes in the Internet connected?
- No real top-down approach ever existed
- Has been characterized by a large degree of randomness
- Question of interest: What is the probability distribution function (PDF) of connectivities of all nodes?

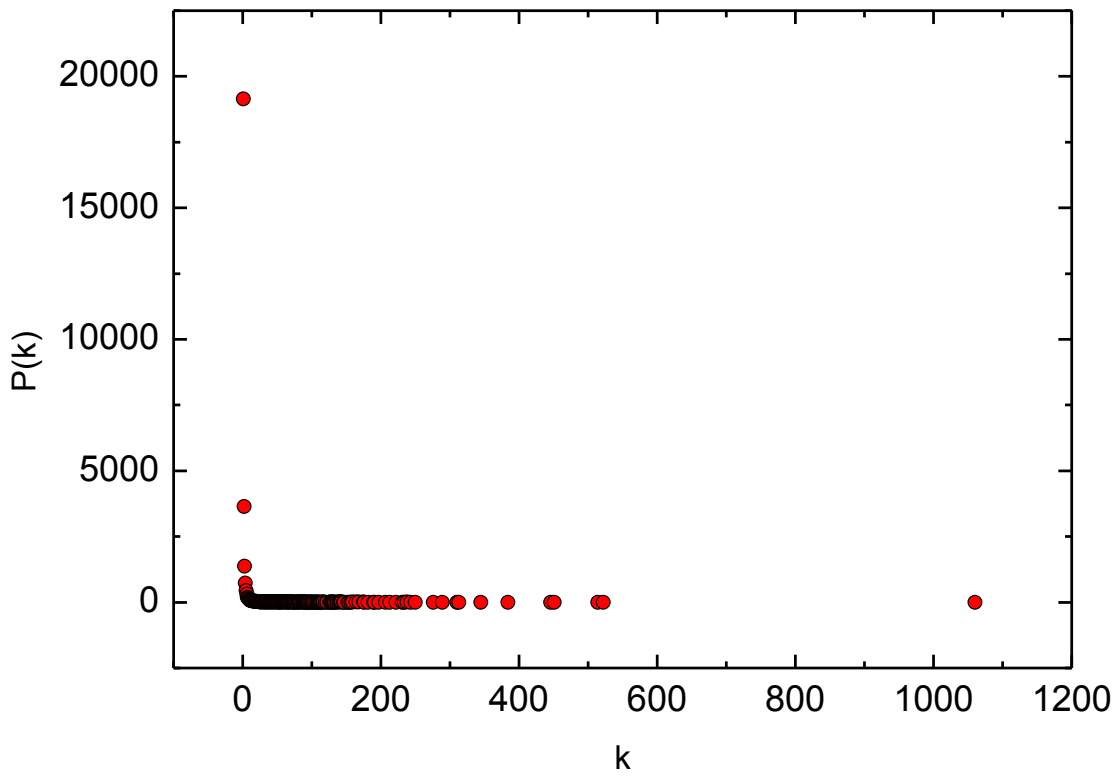
(i.e. how many connections does each node have)?

- Naïve approach: Normal distribution (Gaussian)
- Experimental result?
- Studied ~1999

Gaussian distribution $f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$



Probability distribution P(k) of the No. of connections (degree distribution)

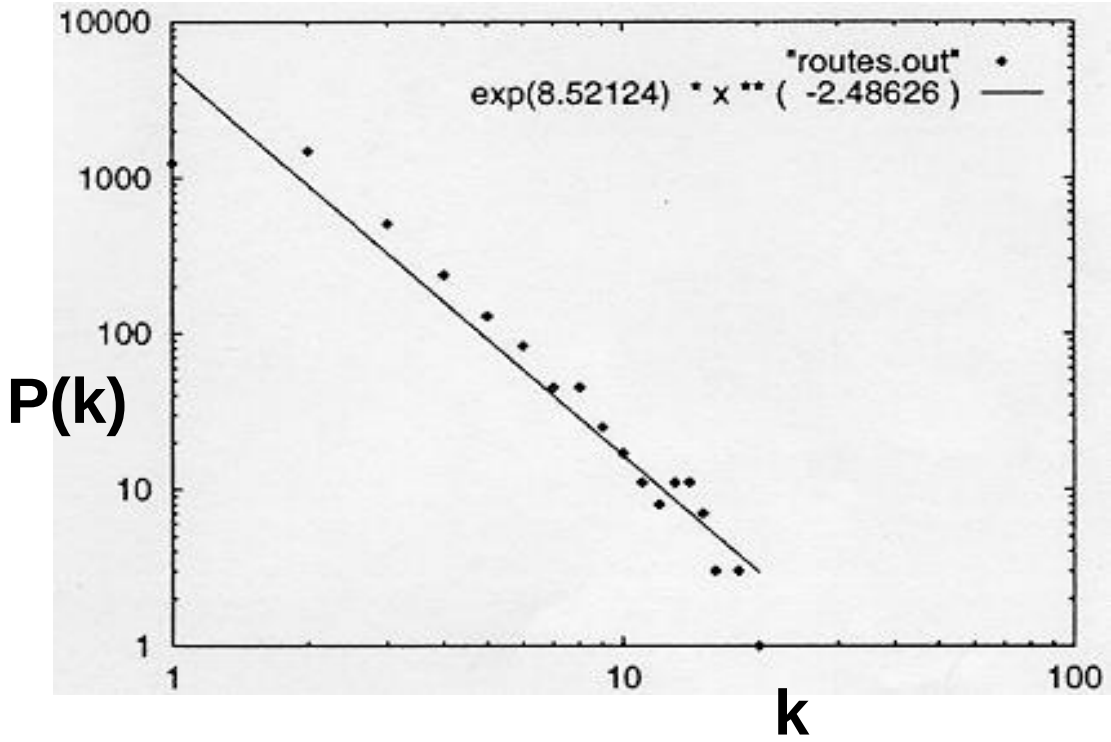


INTERNET BACKBONE

Nodes: computers, routers

Links: physical lines

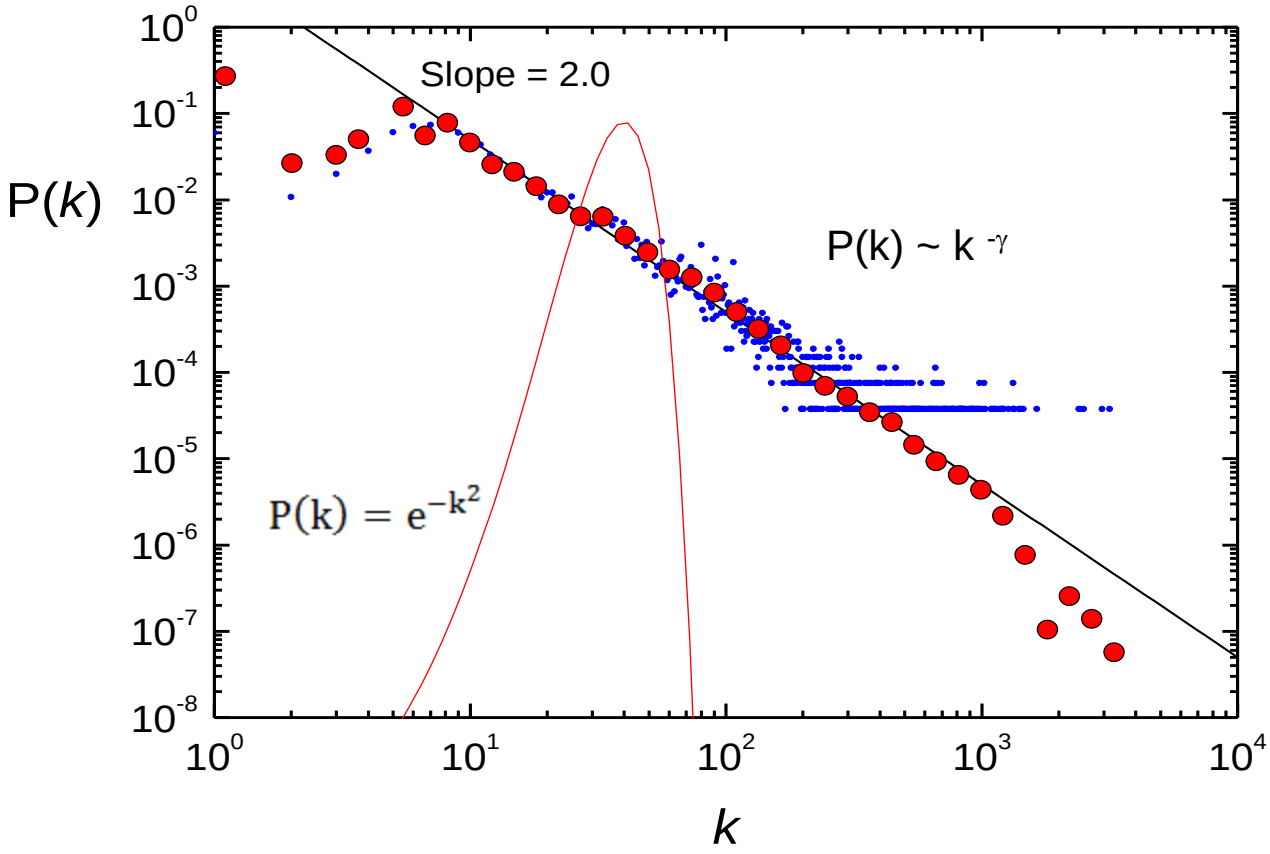
$$P(k) \sim k^{-\gamma}$$



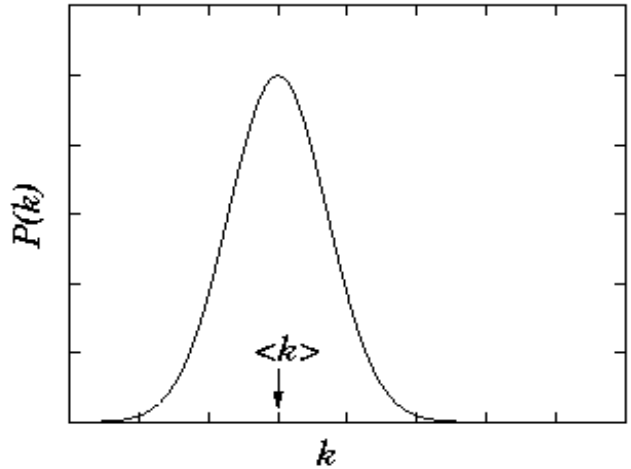
k=number of connections of a node
P(k)= the distribution of k

(Faloutsos, Faloutsos and Faloutsos, 1999)

Degree distribution in a scale-free network



What did we expect?



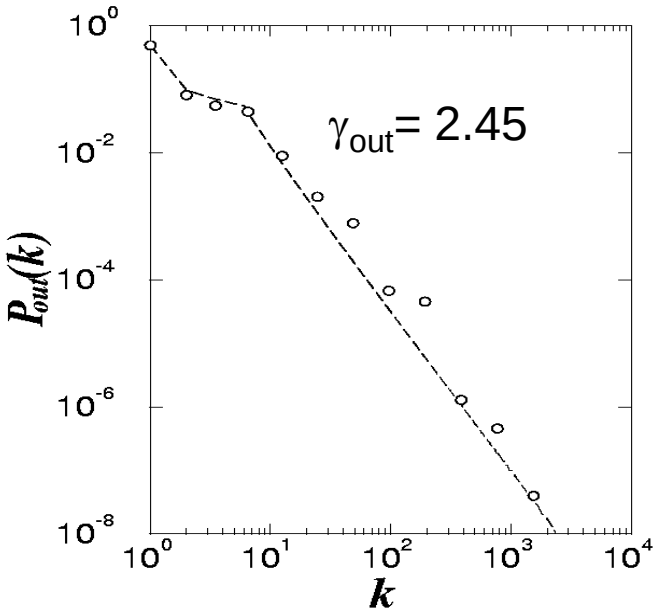
$\langle k \rangle \sim 6$

$P(k=500) \sim 10^{-99}$

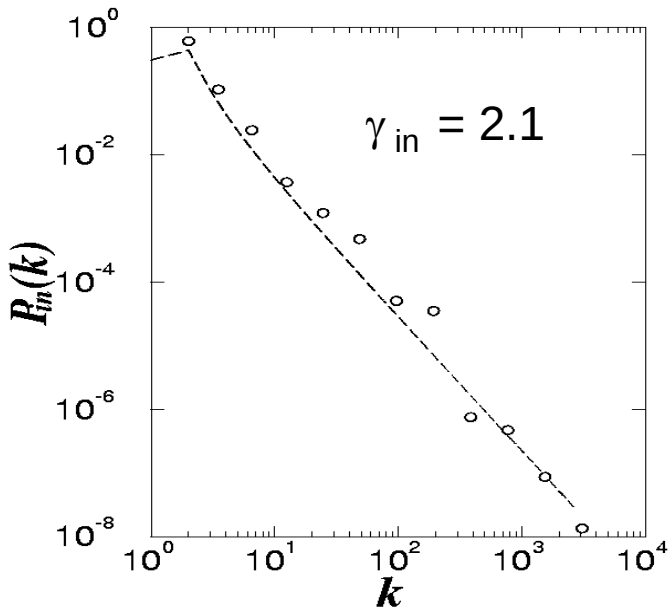
$N_{www} \sim 10^{9}$

$\Rightarrow N(k=500) \sim 10^{-90}$

We find:



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



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

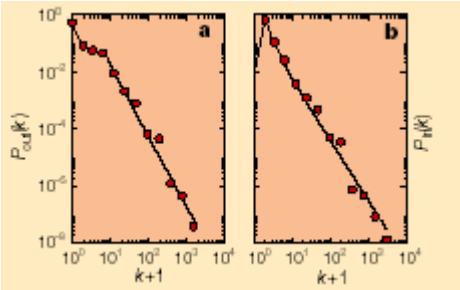
$P(k=500) \sim 10^{-6}$

$N_{www} \sim 10^{9}$

$\Rightarrow N(k=500) \sim 10^{3}$

WWW

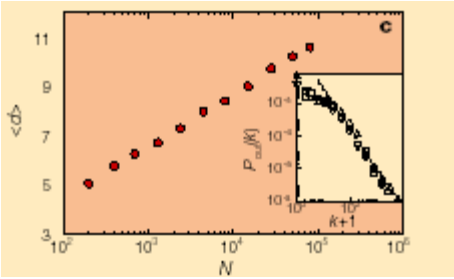
It has a power-law



Diameter of WWW:

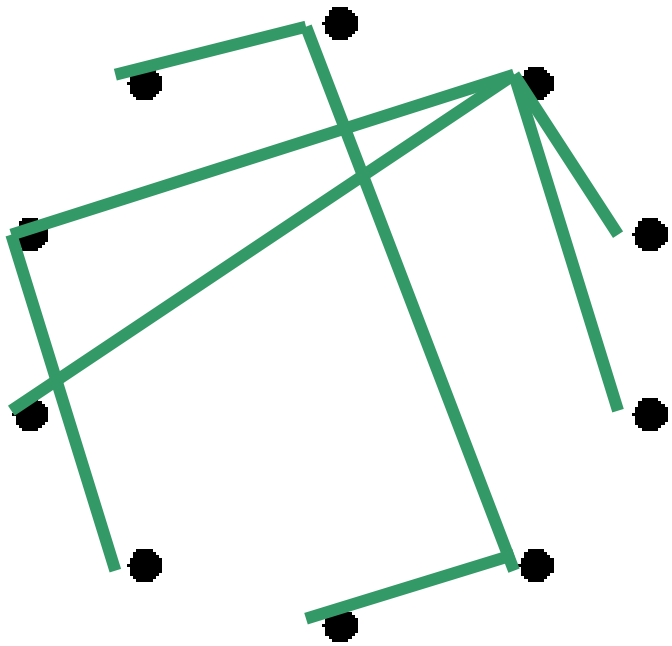
<d>=0.35 + 2.06 logN

For: N=8x10^8 → <d>=18.59



[compare to a square of same size:
edge length=30000
For a cube: edge length=1000]

The model of Erdös-Rényi (1960)

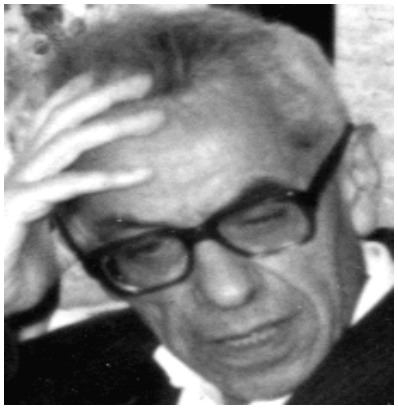


Connections with probability p

p=1/6

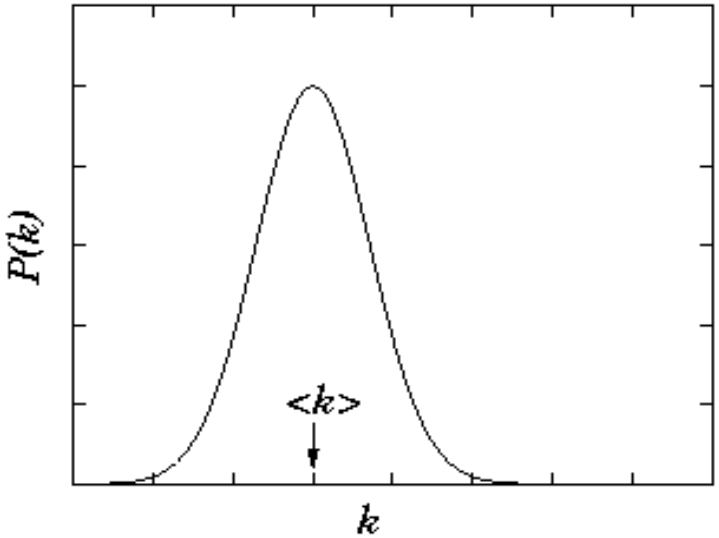
N=10

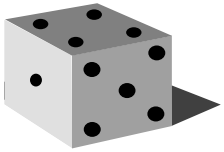
$\langle k \rangle \sim 1.5$



Pál Erdös
(1913-1996)

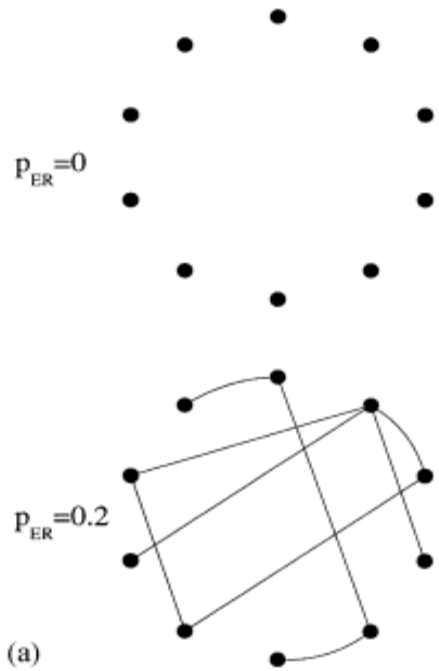
Poisson distribution

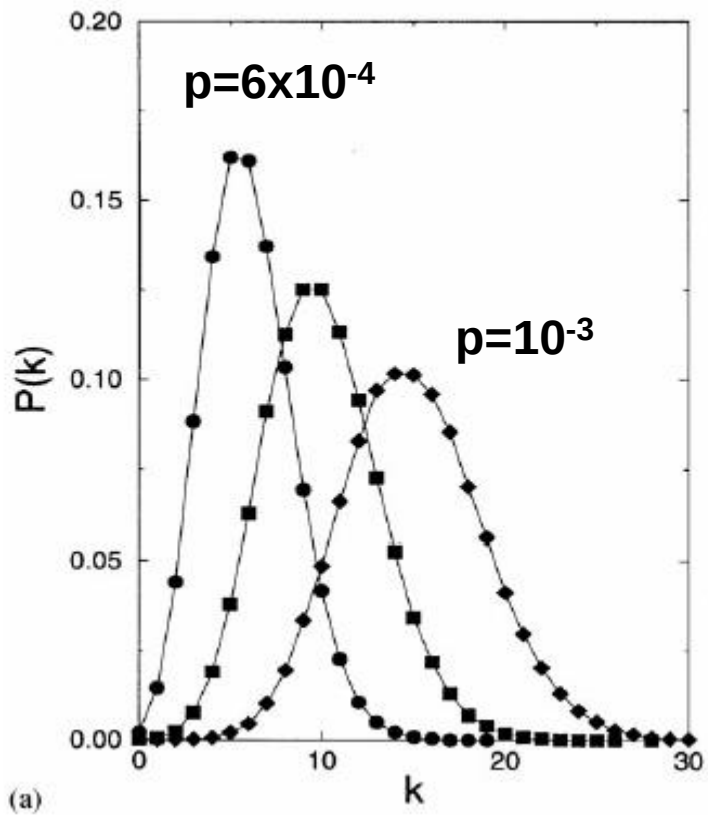




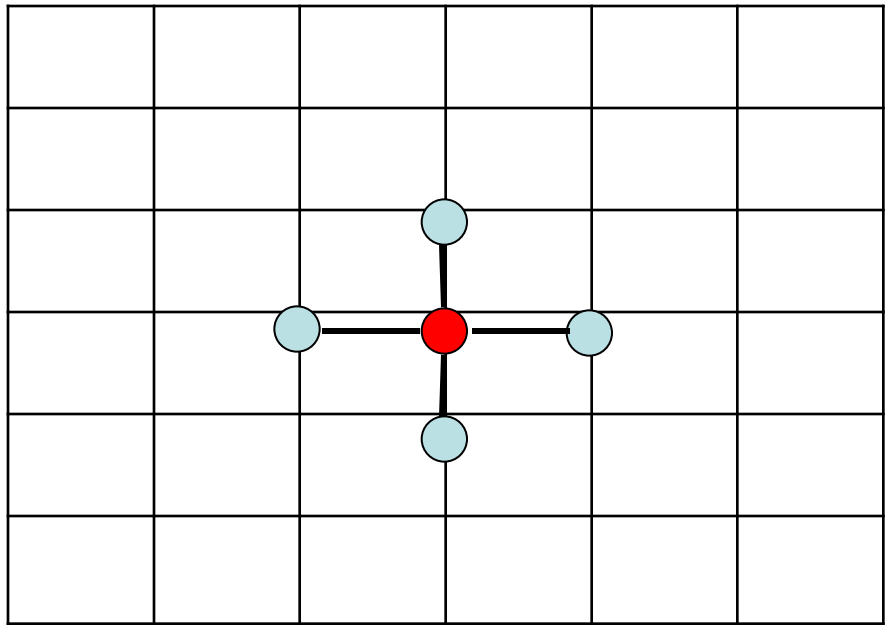
- **Democratic**
- **Random**

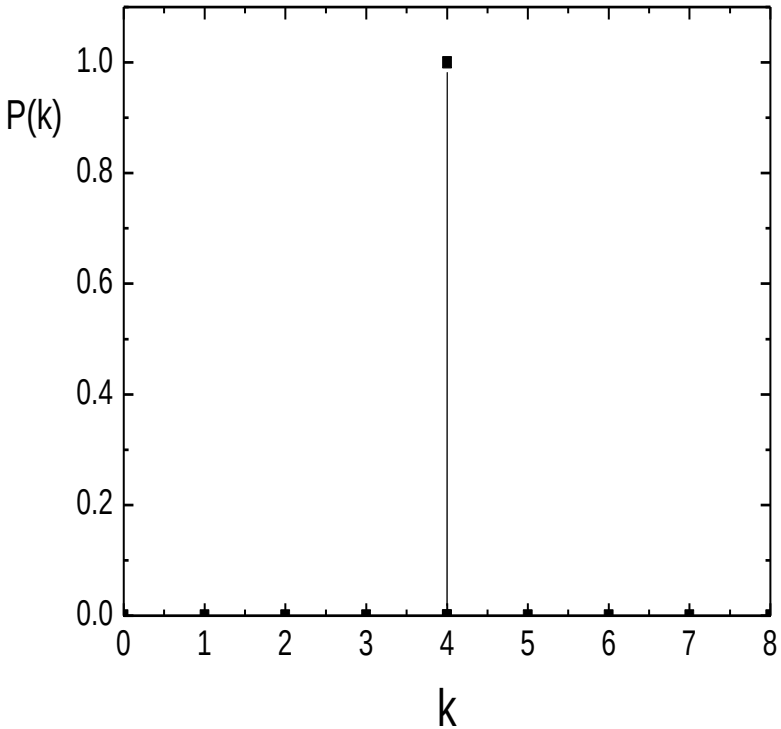
The Erdos-Renyi model



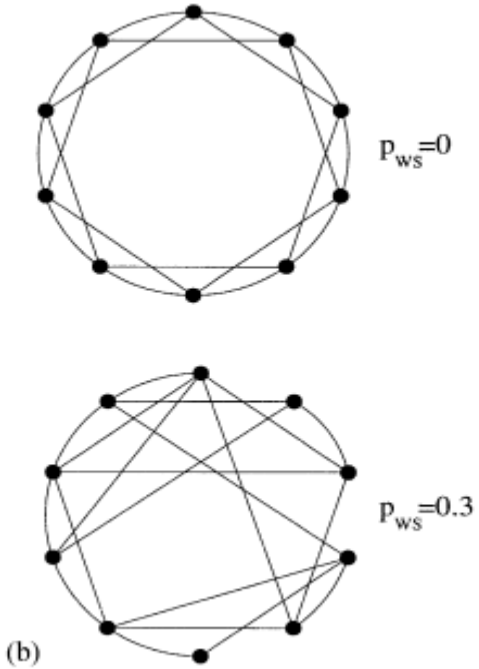


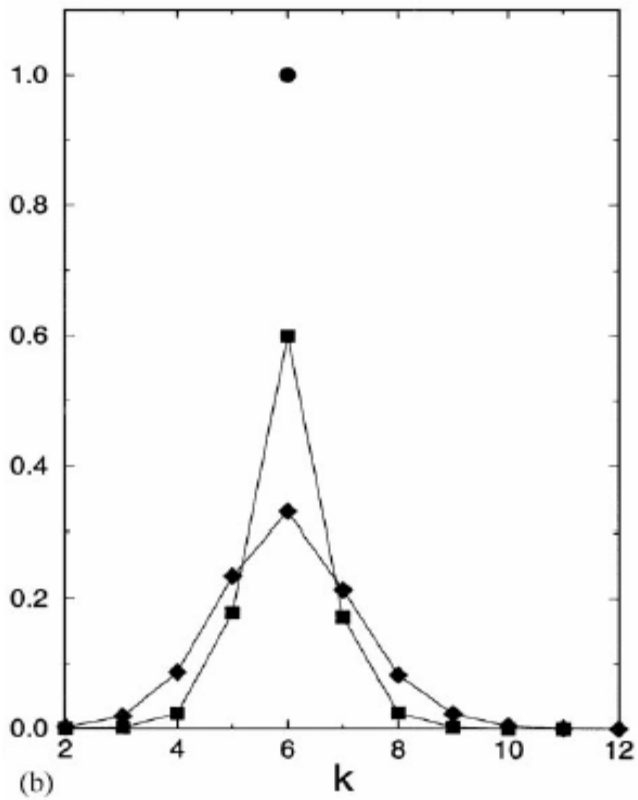
Regular network



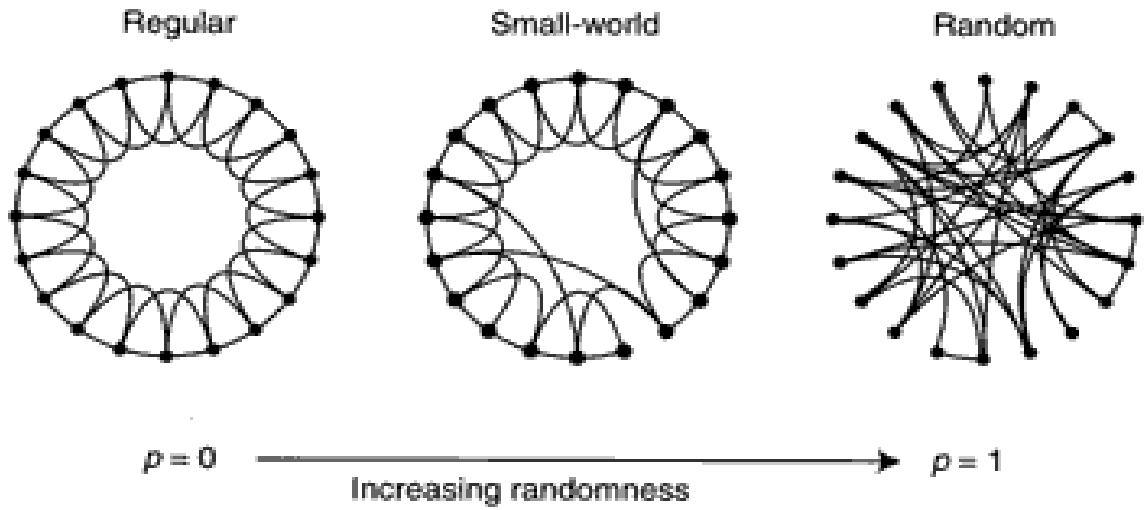


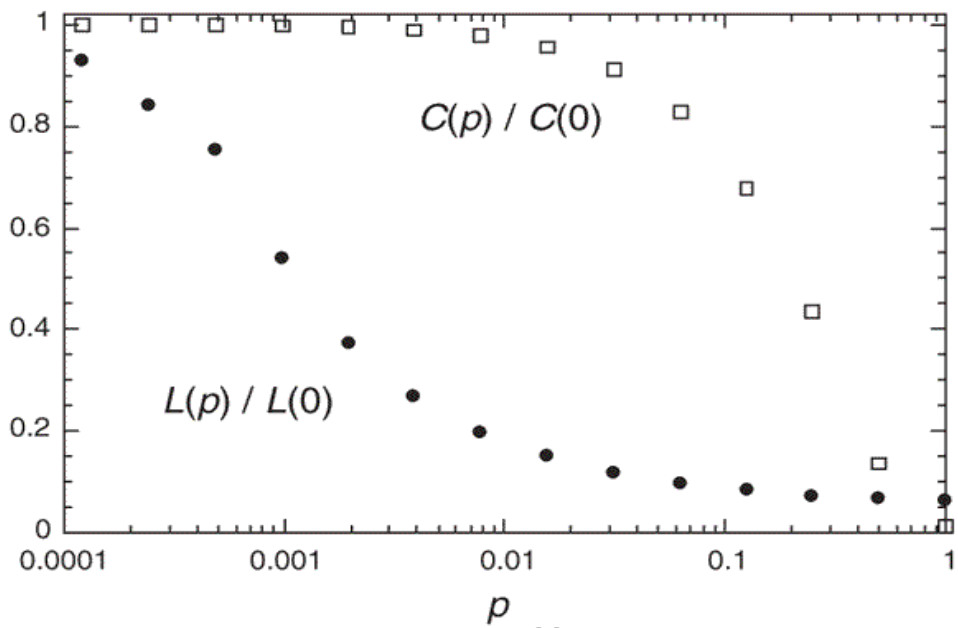
Small world network

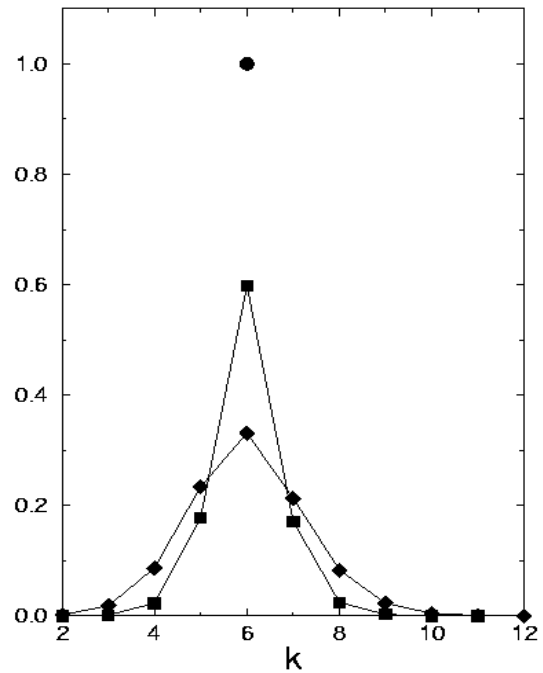




The Watts-Strogatz model





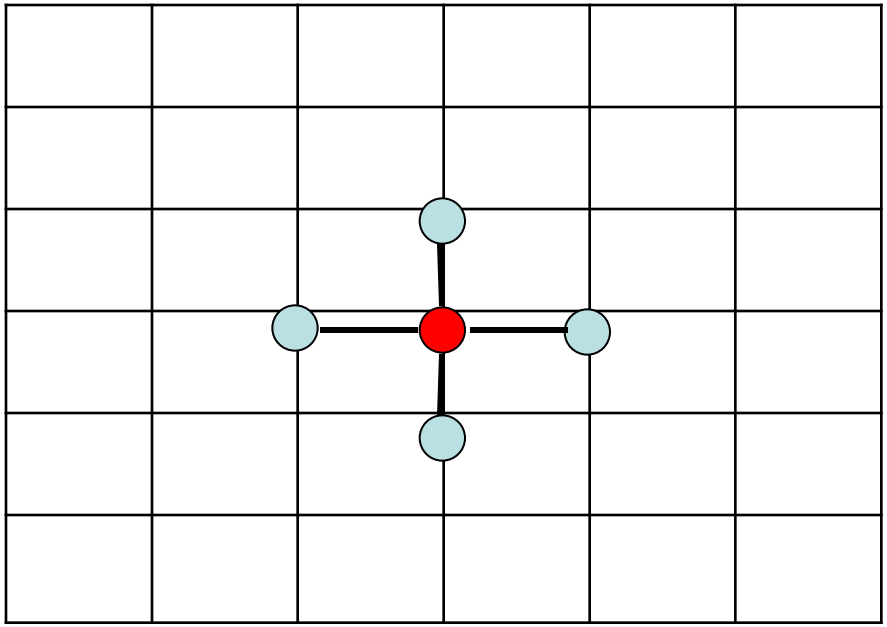


C(p) : clustering coeff.

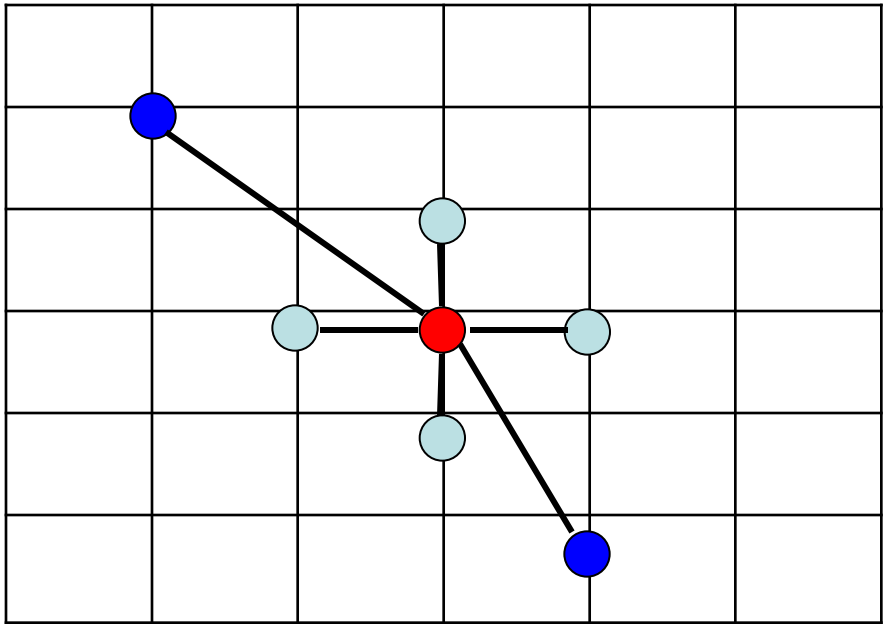
L(p) : average path length

(Watts and Strogatz, Nature **393,** 440 (1998))

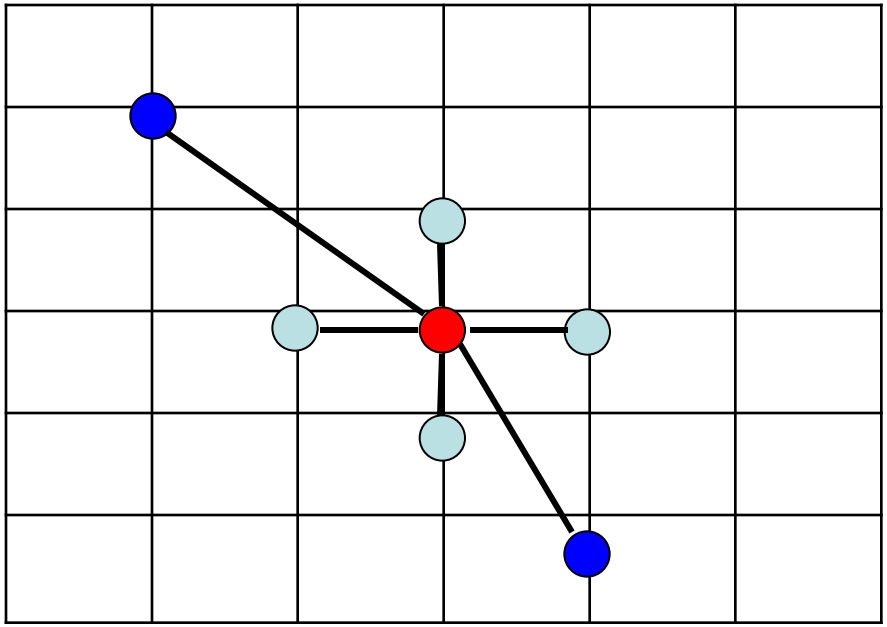
Small world network

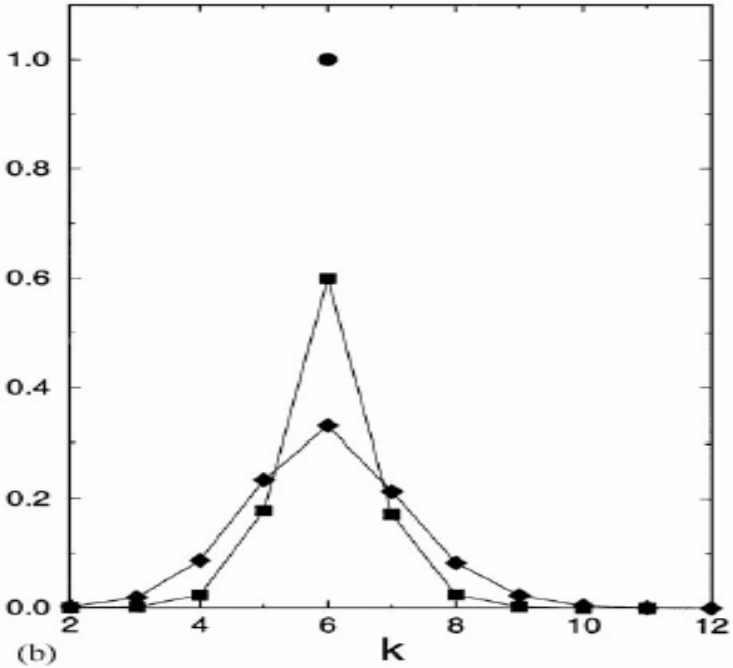


Small world network



Small-world network





Most real networks have similar internal structure:

Scale-free networks

Why?

What does it mean?

Scale-free networks

(1) The number of nodes is not pre-determined

The networks continuously expand
with the addition of new nodes

Example:
WWW : addition of new pages

Citation : publication of new articles

(2) The additions are not uniform

A node that already has a large number of connections
is connected with larger probability than another node.

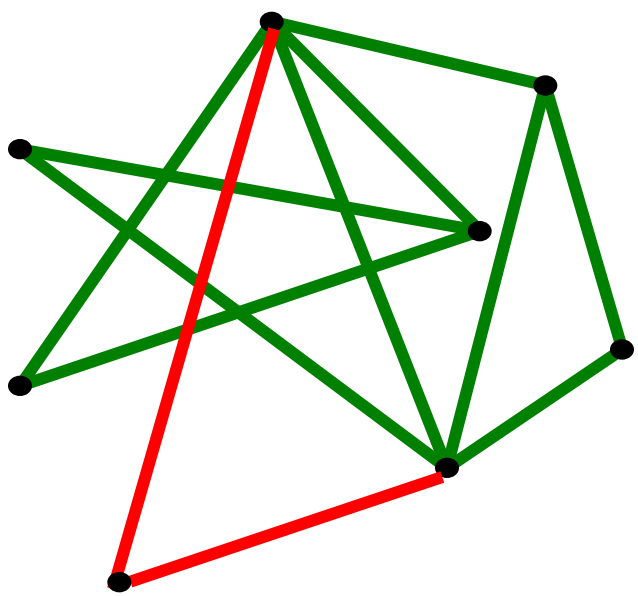
Example:
WWW : new topics usually go to well-known sites
(CNN, YAHOO, NewYork Times, etc)
Citation : papers that have a large number of references are more
probable to be refered again

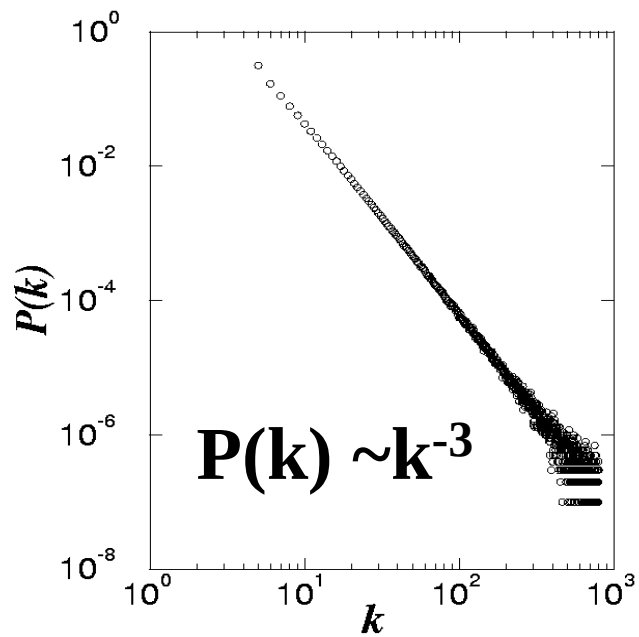
Scale-free networks

(1) Growth: At every moment we add a new node with *m connections* (which is added to the already existing nodes).

(2) Preferential Attachment: The probability **Π** that a new node will be connected to node *i* depends on the number k_i , the number of connections of this node

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





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

Network categories:

- **(1) Random network (Erdos-Renyi)**
- **(2) Network on a regular lattice**
- **(2) Small world network (Strogatz-Watts)**
- **(3) Network with a power-law**

Human Brain

Cells: Neurons

Human Brain:

- **Nodes: cells=neurons**
- **Links: dendrites**
- **No. of nodes: 10,000,000,000, i.e. 10 billion**
- **All neurons are different, i.e. no two neurons are identical**
- **Each neuron has ~k= 10,000 connections (synapses)**
- **10 billion nodes x 10,000 links each = total 10^{+14} links**
- **100 trillion connections**

Dendrimer molecules

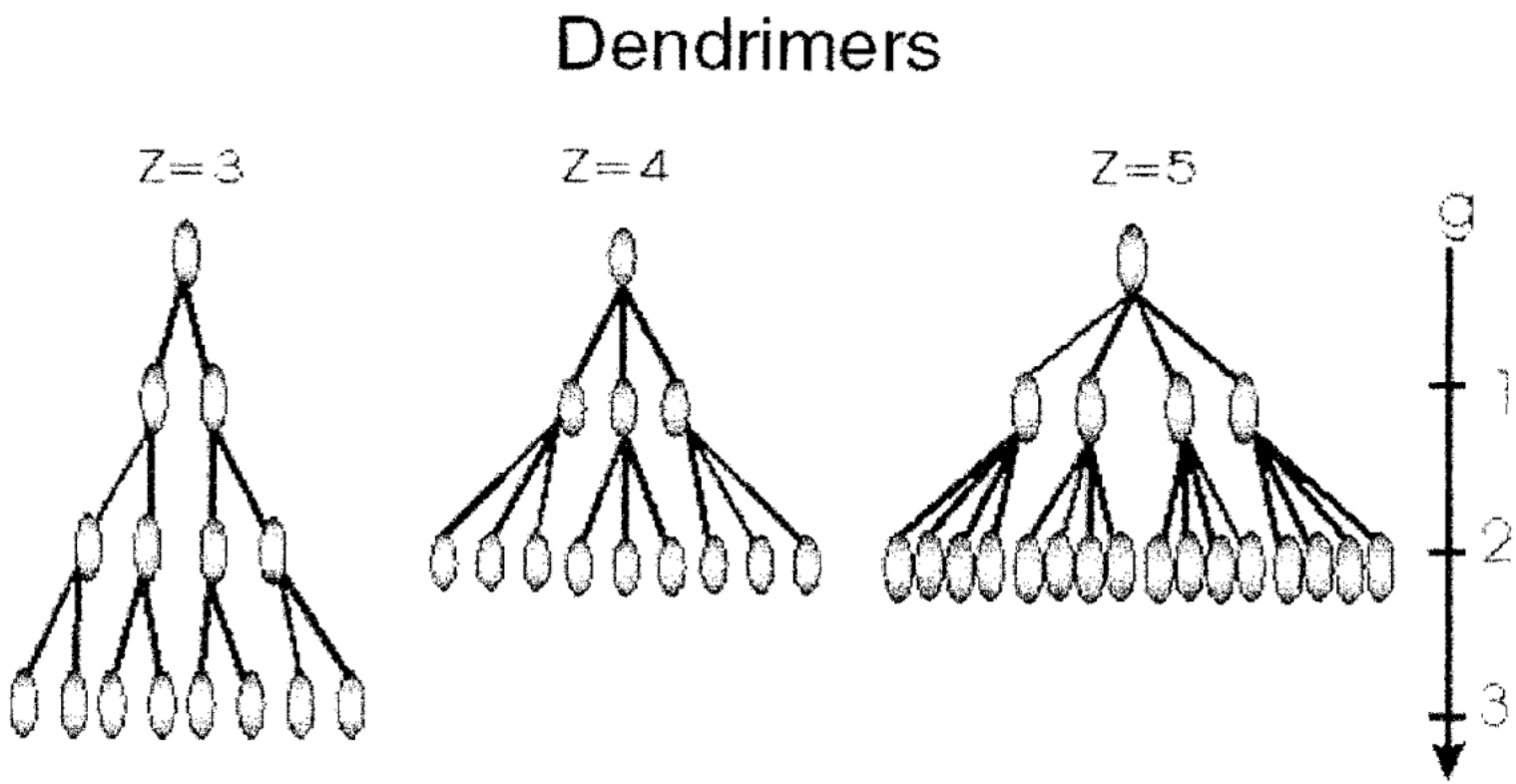
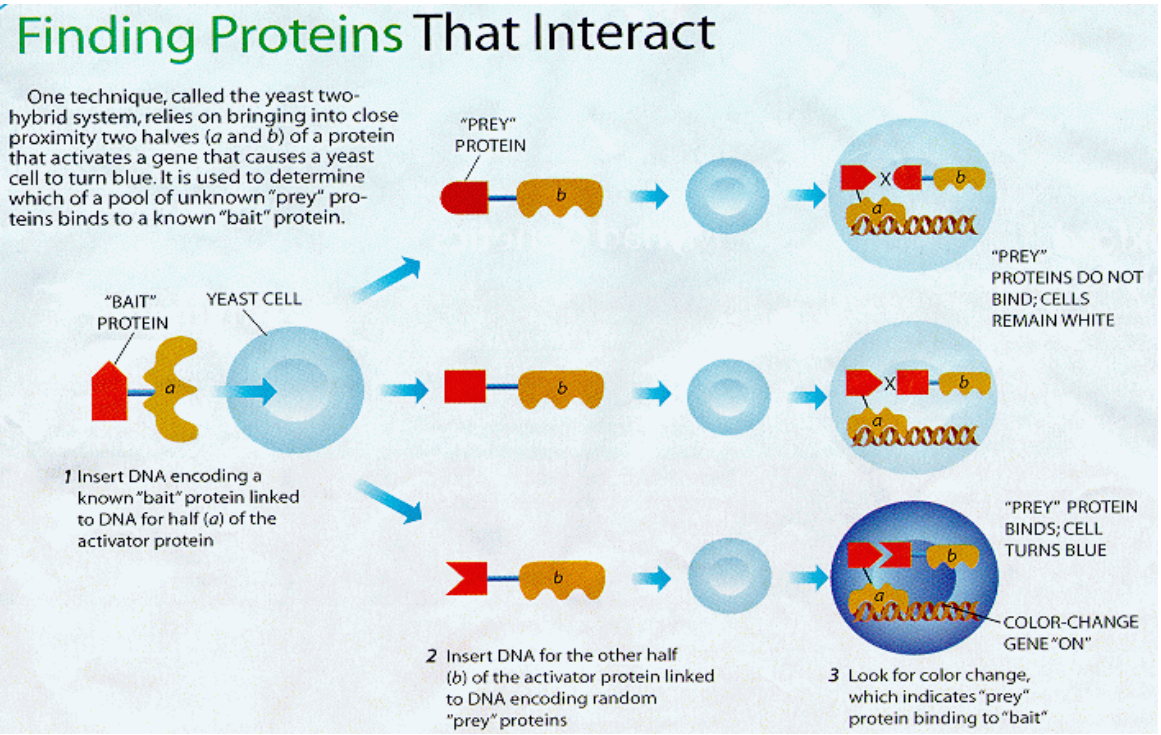


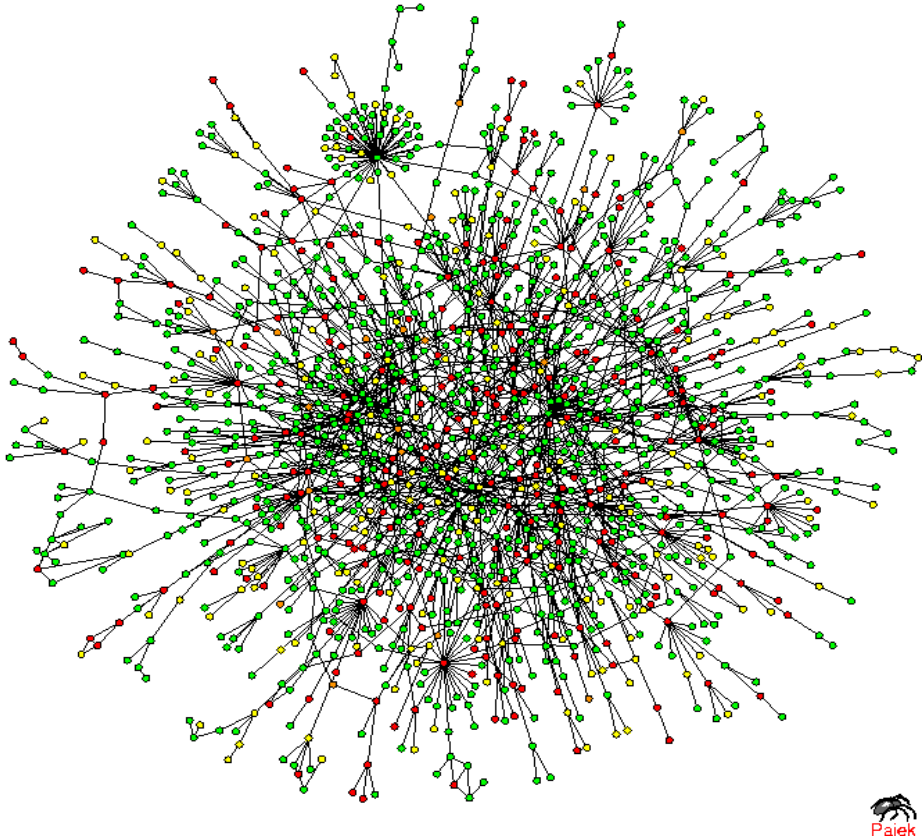
Fig. 1. Schematic of three dendrimers, with $z = 3$, 4, and 5. The generation order, g, is as shown, $g = 3$ (for $z = 3$), and $g = 2$ (for $z =$ 4 and 5).

Yeast protein network

Nodes: proteins

Links: physical interactions (binding)





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

Communication networks

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

-computers

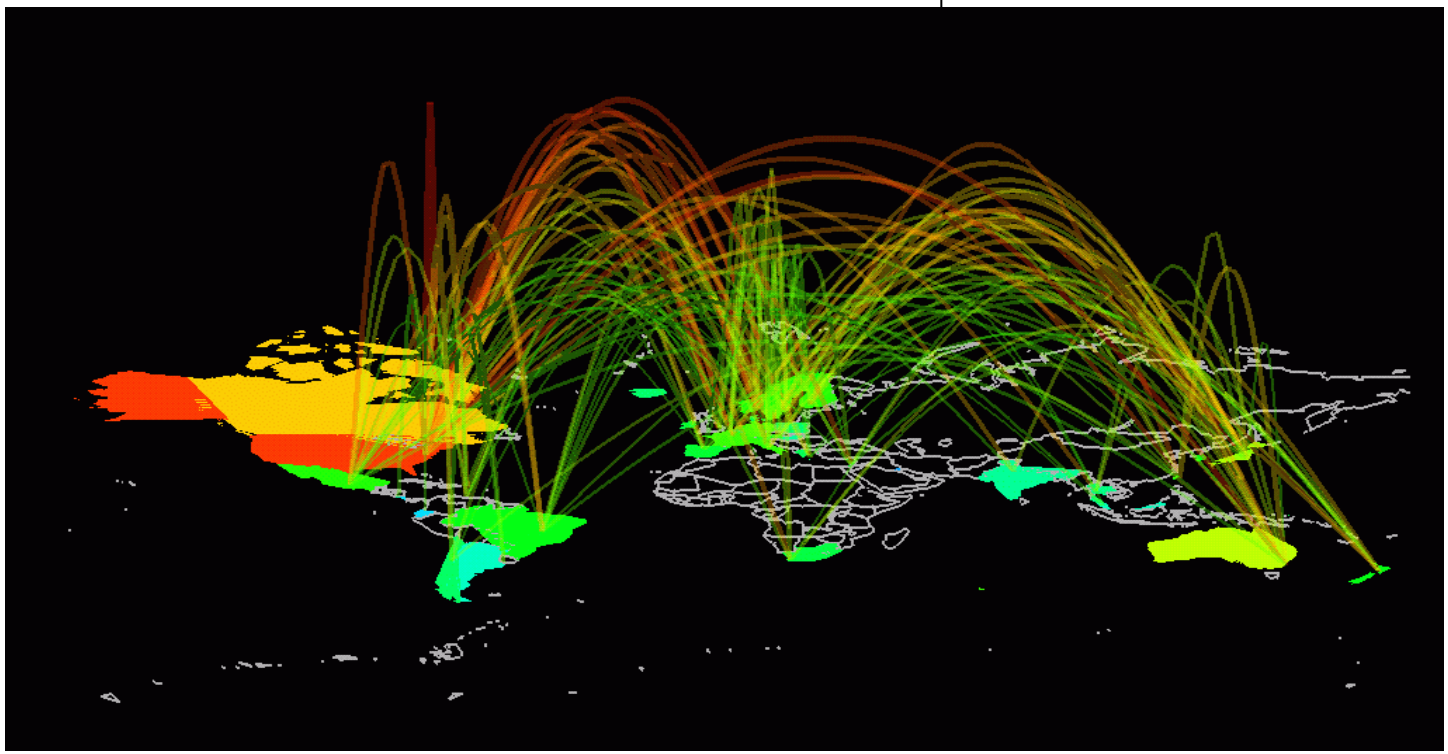
-routers

-satellites

-phone lines

-TV cables

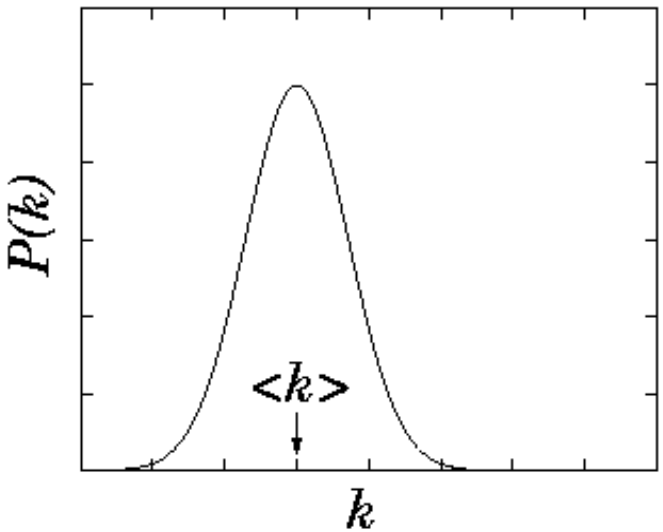
-EM waves



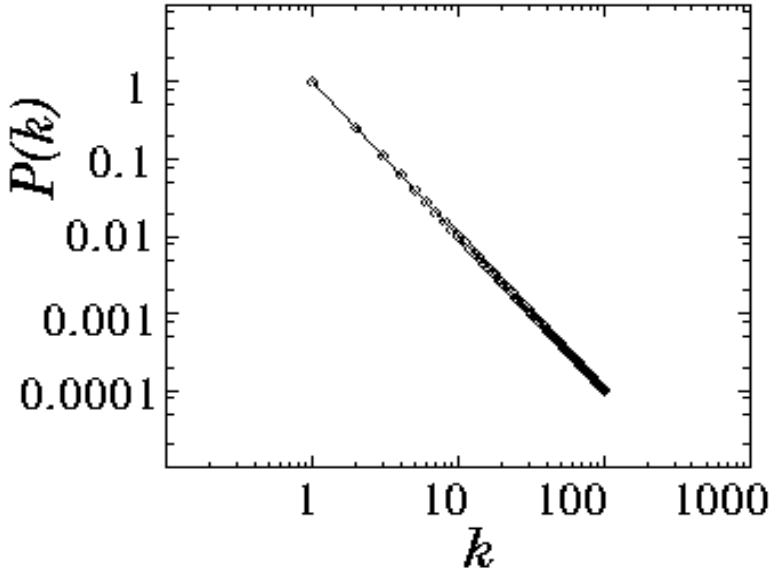
Communication networks: Many non-identical components with diverse connections between them.

What does it mean?

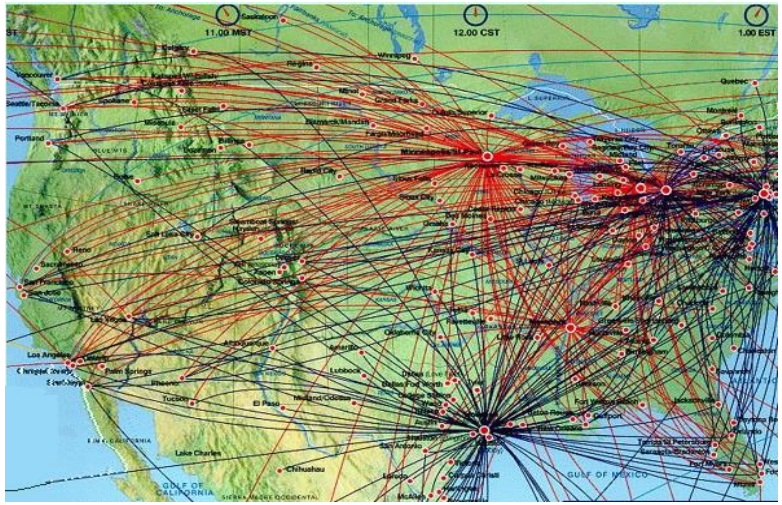
Poisson distribution



Power-law distribution



Exponential Network

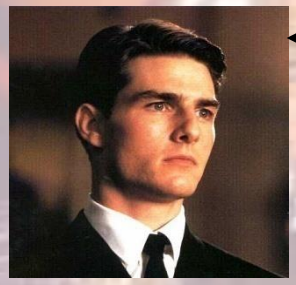


Scale-free Network

ACTOR CONNECTIVITIES

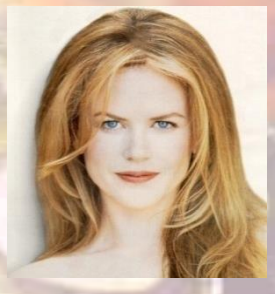
Nodes: actors
cast jointly

Links:



Days of Thunder (1990)

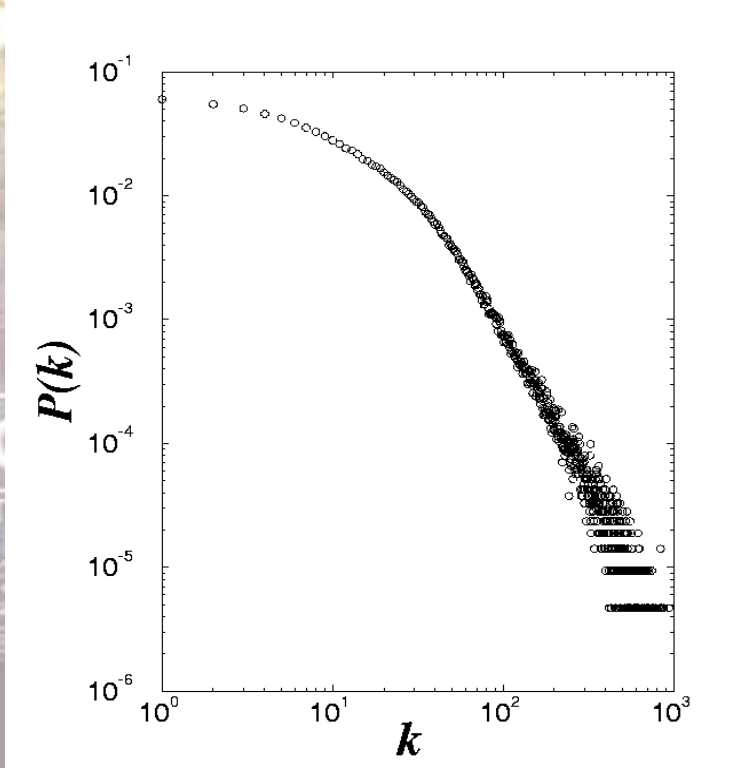
Far and Away (1992)
Eyes Wide Shut (1999)



N = 212,250 actors
⟨k⟩ = 28.78

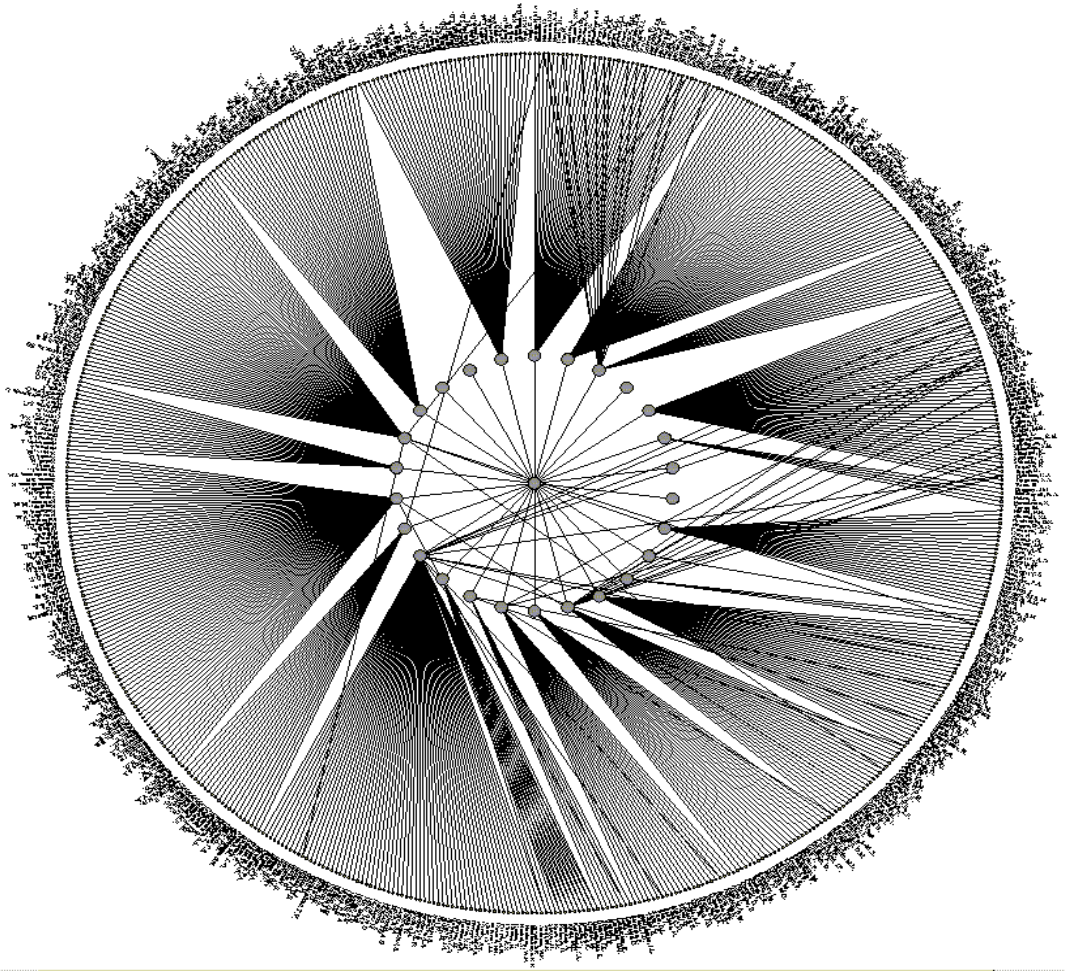
$P(k) \sim k^{-\gamma}$

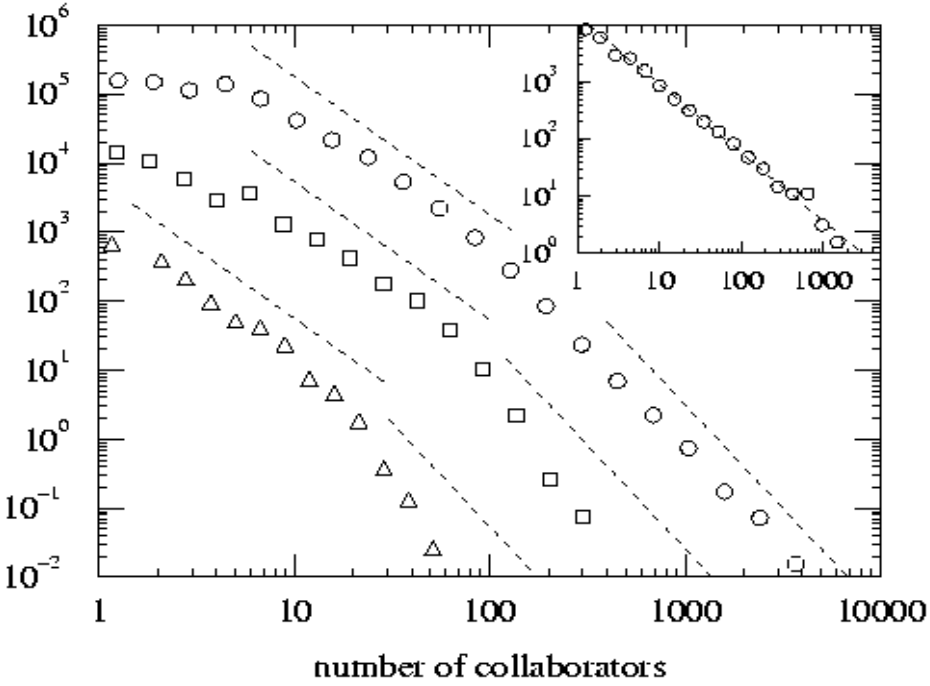
$\gamma = 2.3$



SCIENCE COAUTHORSHIP

Nodes: scientist (authors)
Links: write paper together





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

SCIENCE CITATION INDEX

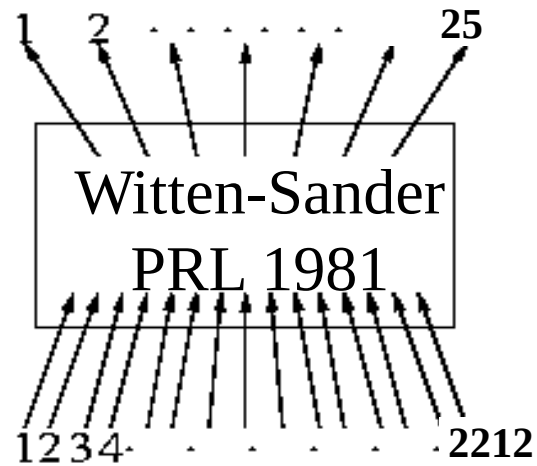
1,000 Most Cited Physicists, 1981-June 1997
Out of over 500,000 Examined
(see http://www.sst.nrel.gov)

Author name		Institute	Country	Field	avg. cites	total art.	total cites	rank by total cit.
Witten	E	Princeton (U)	USA, NJ	High-energy (T)	168	138	23235	1
Gossard	AC	UCSB (U)	USA, CA	Sem				2
Cava	RJ	Bell Labs (I)	USA, NJ	Supe				3
Batlogg	B	[illegible]	USA, NJ	Supe				4
Ploog	K	Max Planck (NL)	Germany	Sem				5
Ellis	J	Euro Nuclear Cent.	Switzerland	Astr				6
Fisk	Z	[illegible]	USA, FL	Solid				7
Cardona	M	Max Planck (NL)	Germany	Sem				8
Nanopoulos	DV	Texas A&M (U)	USA, TX	High				9
Heeger	AJ	UCSB (U)	USA, CA	Poly				10
Lee*	PA							11
Suzuki*	T							12
Anderson			NJ	Solid				13
Suzuki*								14
Freeman			IL	Solid				15
Tanaka*	S							16
Muller			nd	Supe				17
Schnee				Supe				18
Chemk				Optics (E)	60	162	9668	19
Morko				Semiconductors (E)	20	477	9668	19
Miller				Semiconductors (E)	67	144	9652	21
Chu				Superconductivity (E)	[illegible]	213	9453	22
Bedno			nd	Superconductivity (E)	[illegible]	85	9311	23
Cohen				Solid State (T)	33	284	9311	23
Meng				Superconductivity (E)	86	108	9300	25
Waszc				Superconductivity (E)	[illegible]	162	9170	26
Shiran				Superconductivity (E)	[illegible]	269	8841	27
Wiegm				Semiconductors (E)	85	104	8822	28
Vando				Magnetism (E)	67	129	8686	29
Uchida					28	301	8520	30
Hor				Superconductivity (E)	72	119	8512	31
Murph				Astronomy (E)	111	76	8439	32
Birgeneau	RJ	MIT (U)	USA, MA	Superconductivity (E)	41	286	8375	33
Jorgensen	JD	Argonne (NL)	USA, IL	Superconductivity (E)	[illegible]	[illegible]	[illegible]	34
Hinks	DG	Argonne (NL)	USA, IL	Superconductivity (E)	[illegible]	225	8263	35

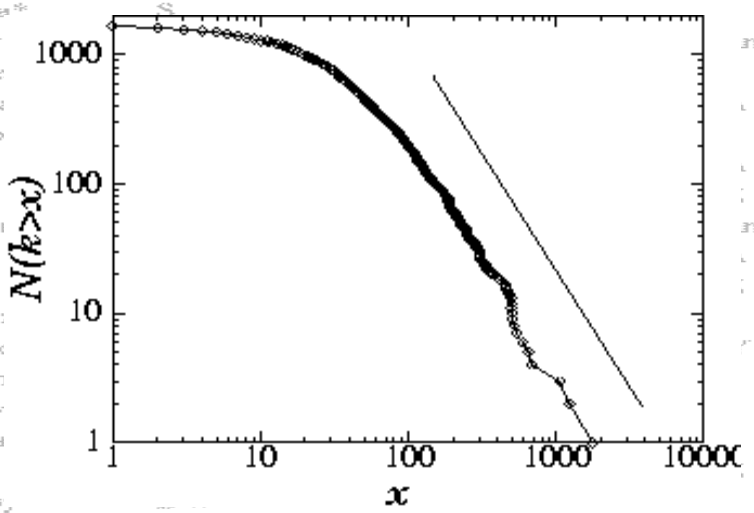
* citation total may be skewed because of multiple authors with the same name

Nodes: papers

Links: citations



1736 PRL papers (1988)



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

$(\gamma = 3)$

(S. Redner, 1998)

Society

Nodes: individuals

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

S. Milgram (1967)

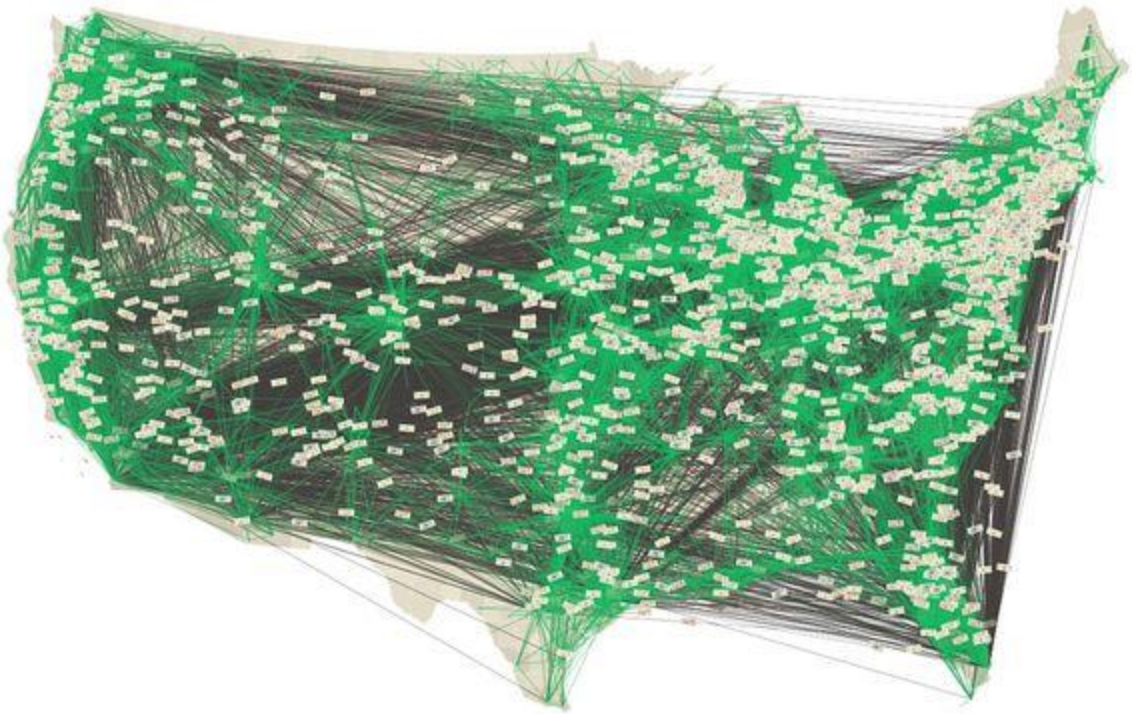
John Guare

How many (n) connections are needed so that an individual is connected with any other person in the world?
N=6 billion people

Result: n~6

Conclusion: We live in a small world
Six Degrees of Separation!!

Where is George?



http://www.facebook.com/ | Welcome to Facebook - Lo...

Edit View Favorites Tools Help

Google facebook Search More

Download Download Manager Browse Player N/A

rt phys um auth TWS ox ox My Yahoo! k-k booking Q&A cosmote comm pre_m OA a3 G GM CF 3B A-B MASOUTIS pfizer gas prices Newsbeast mb marketbeast Get more Add-ons Welcome to the UM Wirel...

facebook

Email Password Log In

Keep me logged in Forgot your password?

Facebook helps you connect and share with the people in your life.

Sign Up

It's free and always will be.

First Name:

Last Name:

Your Email:

Re-enter Email:

New Password:

I am: Select Sex:

Birthday: Month: Day: Year:

Why do I need to provide my birthday?

By clicking Sign Up, you agree to our Terms and that you have read and understand our Data Use Policy.

Sign Up

Create a Page for a celebrity, band or business.

English (US) Ελληνικά Español Português (Brasil) Français (France) Deutsch Italiano العربية हिन्दी 中文(简体) ...

Facebook © 2012 · English (US)

Mobile · Find Friends · Badges · People · Pages · About · Advertising · Create a Page · Developers · Careers · Privacy · Terms · Help

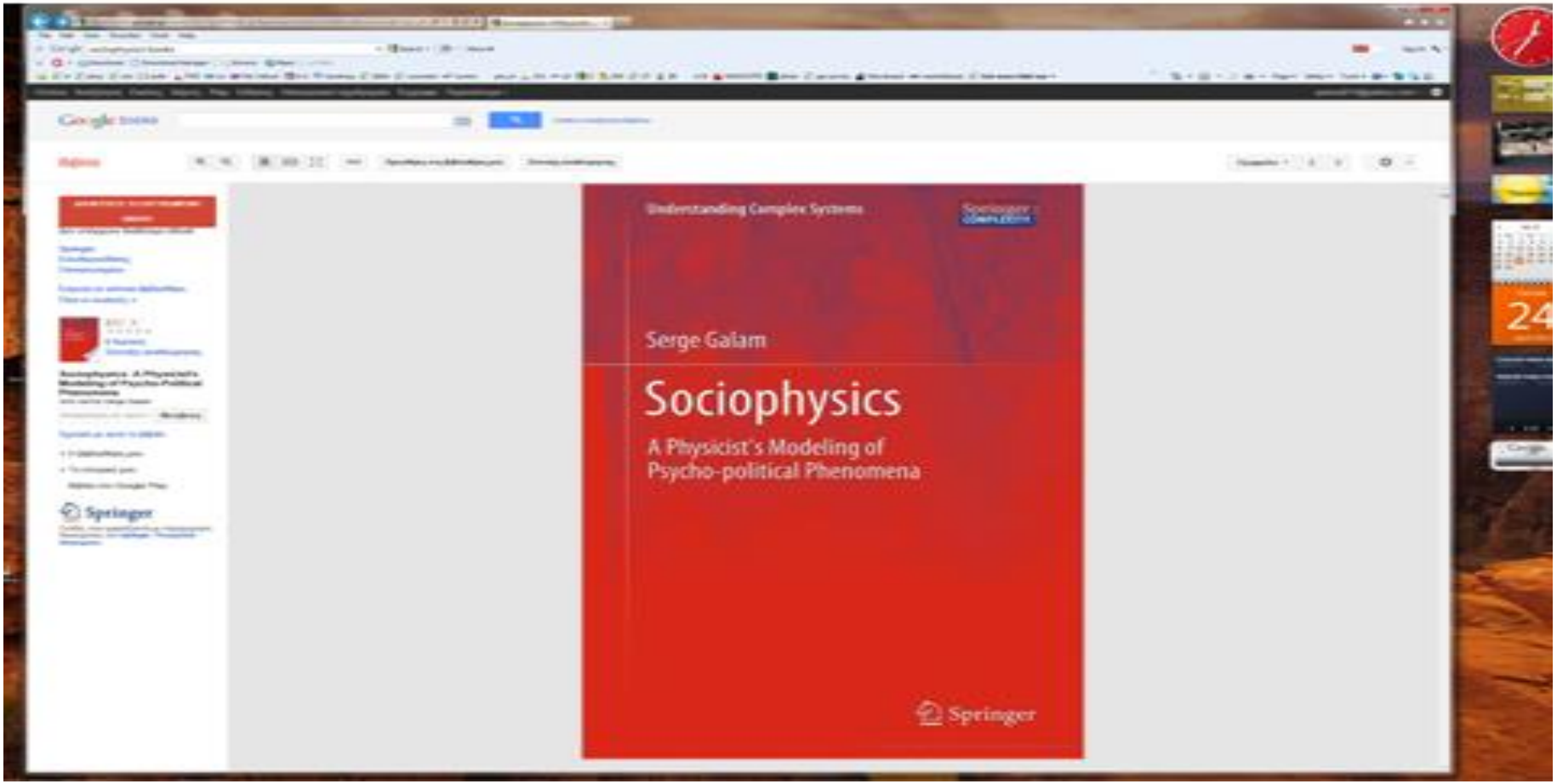
Facebook:

- 1,000,000,000 registered users
- 50,000,000 active users
- 5,000,000 generate 95% of the traffic

Questions needing answers:

- How many people communicate ?
- How many connections does one have?
- How often does he communicate?
- How long does it last?

New books:


Understanding Complex Systems
Serge Galam
Sociophysics
A Physicist's Modeling of
Psycho-political Phenomena
Springer

Real-world phenomena related to communication patterns

- ***Crowd behavior*: strategies to evacuate people and stop panic**
- ***Search strategies*: efficient networks for searching objects and people**
- ***Traffic flow*: optimization of collective flow**
- ***Dynamics of collaboration*: human relationship networks such as collaboration, opinion propagation and email networks**
- ***Spread of epidemics*: efficient immunization strategies**
- ***Patterns in economics and finance*: dynamic patterns in other disciplines, such as Economics and Finance, and Environmental networks**

Spreading phenomena

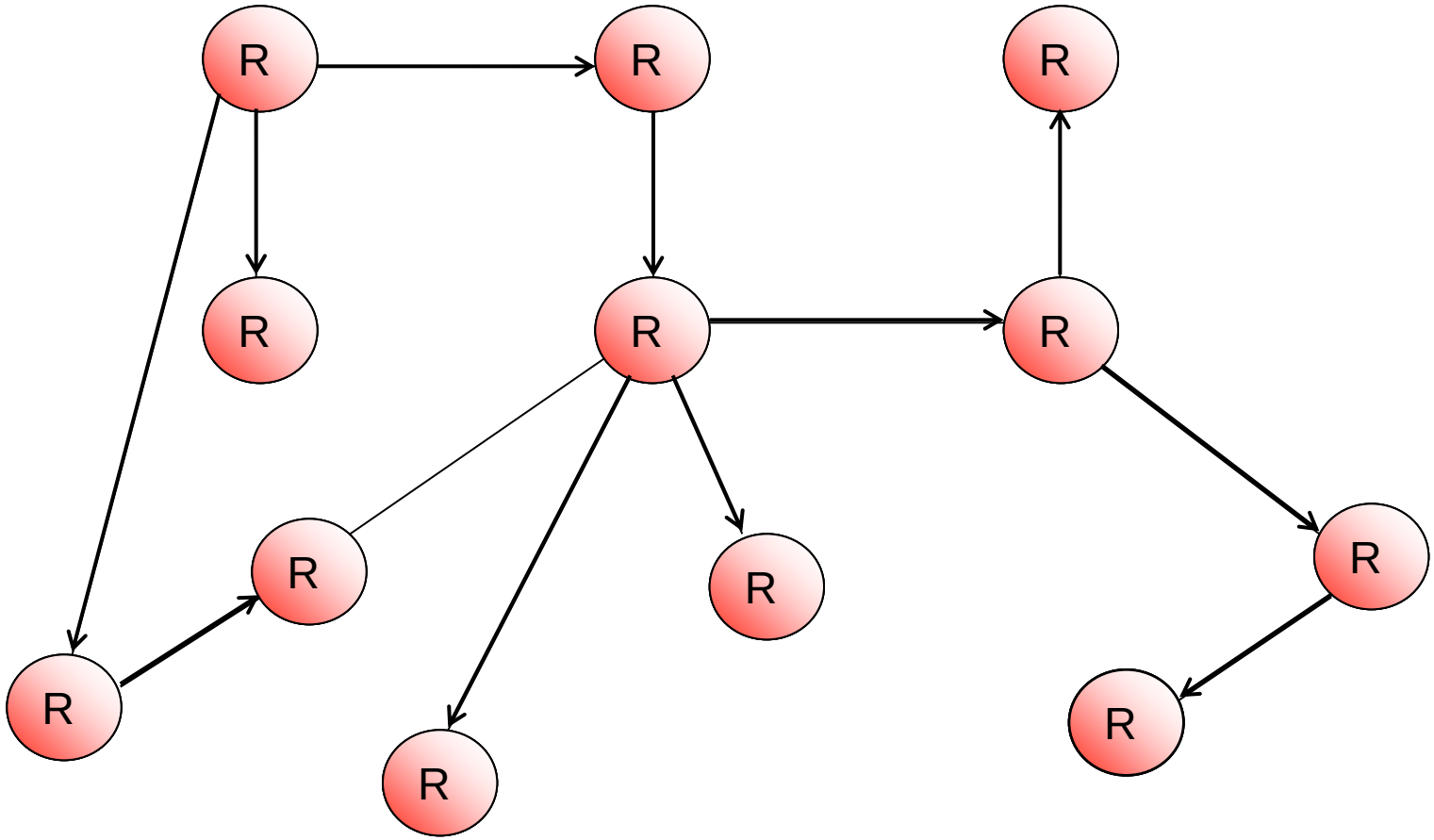
- **Starting at a random node**
- **Moving to nearest neighbors**

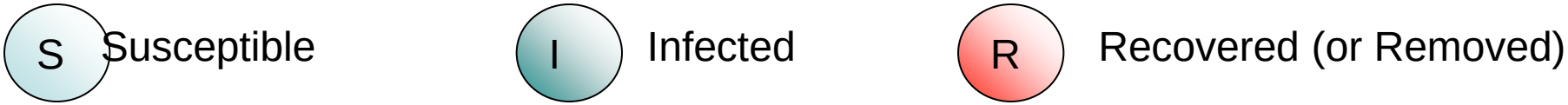
  use: Breadth First Search Algorithm
- **SIR (susceptible-infected-removed/recovered)**
- **SIS (susceptible-infected-susceptible)**
- **SIRS (susceptible-infected-refractory-susceptible)**
- **Applications: Forest fires, epidemics, rumor spreading, virus spreading, etc.**

The SIR model

- SIR (Susceptible, Infected, Removed),
- q=probability of infection
- Initially all nodes are susceptible (S)
- Then, a random node is infected (I)
- The virus is spread in the network, all I nodes become R
- This process continues until the virus either
  - has been spread in the entire network, or
  - has been totally eliminated
- M=infected mass
- Duration

SIR model


R
R
R
R
R
R
R
R
R
R
R
R

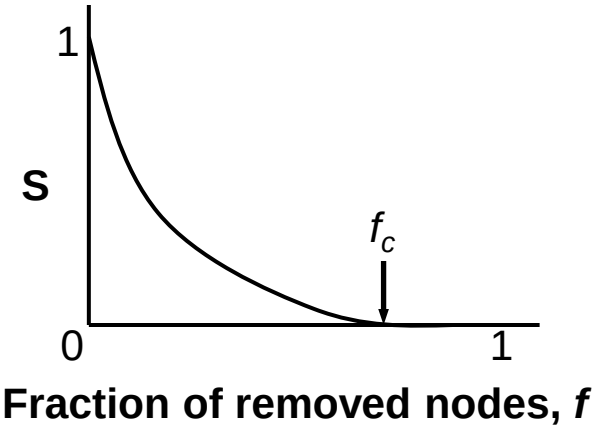

S
Susceptible
I
Infected
R
Recovered (or Removed)

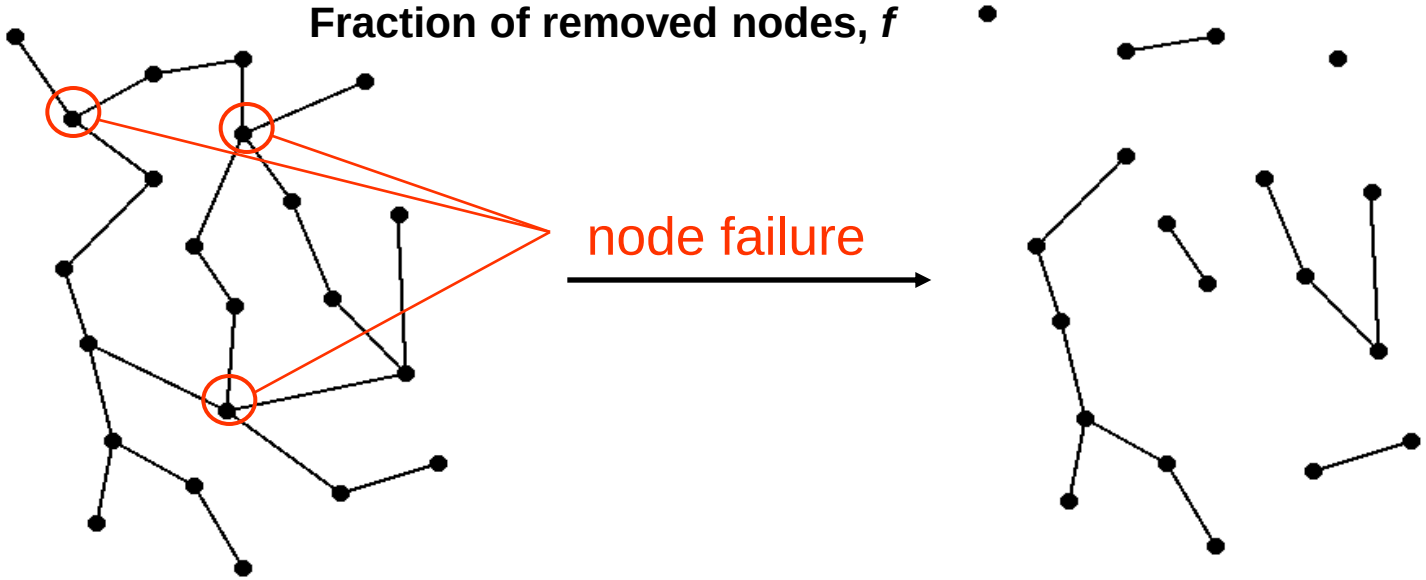
Immunization
p_c =critical point for transmission

- **Complete immunization of populations is not feasible (p_c can not be made significantly lower than 1)**
- **Try to find an efficient method of immunization**
- **We seek strategic ways of immunizing, e.g. try to isolate via immunization the largest possible portion of the population instead**
  - **New strategies may be needed**

Robustness of networks

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





Attack tolerance

- **The stability of the networks under failure or attack is very important.**
- **In general, the integrity is destroyed after a critical percentage p_c of the nodes has been removed (no giant cluster).**
- **Scale-free networks are:**

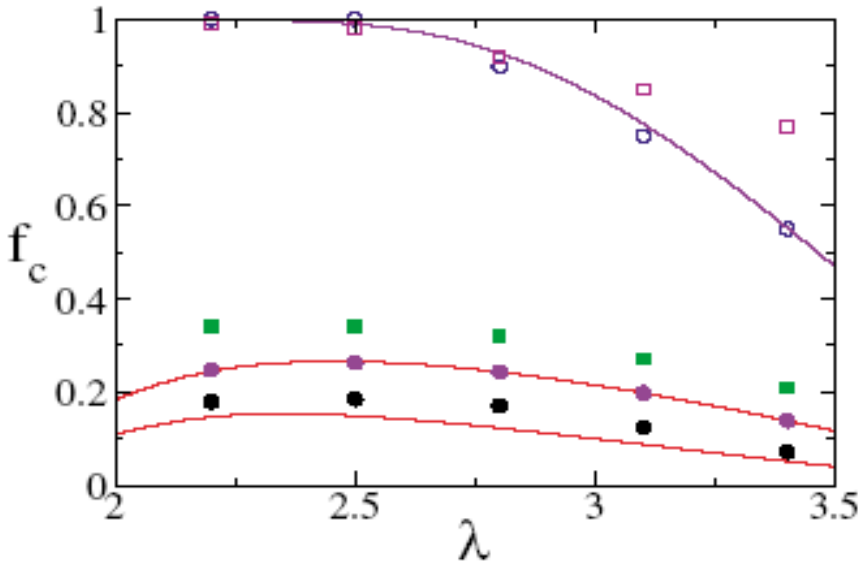
  extremely robust under random failure (p_c=0.99),

  very vulnerable under targeted attacks (p_c=0.07).
- **The existence of a spanning cluster is based on the criterion**

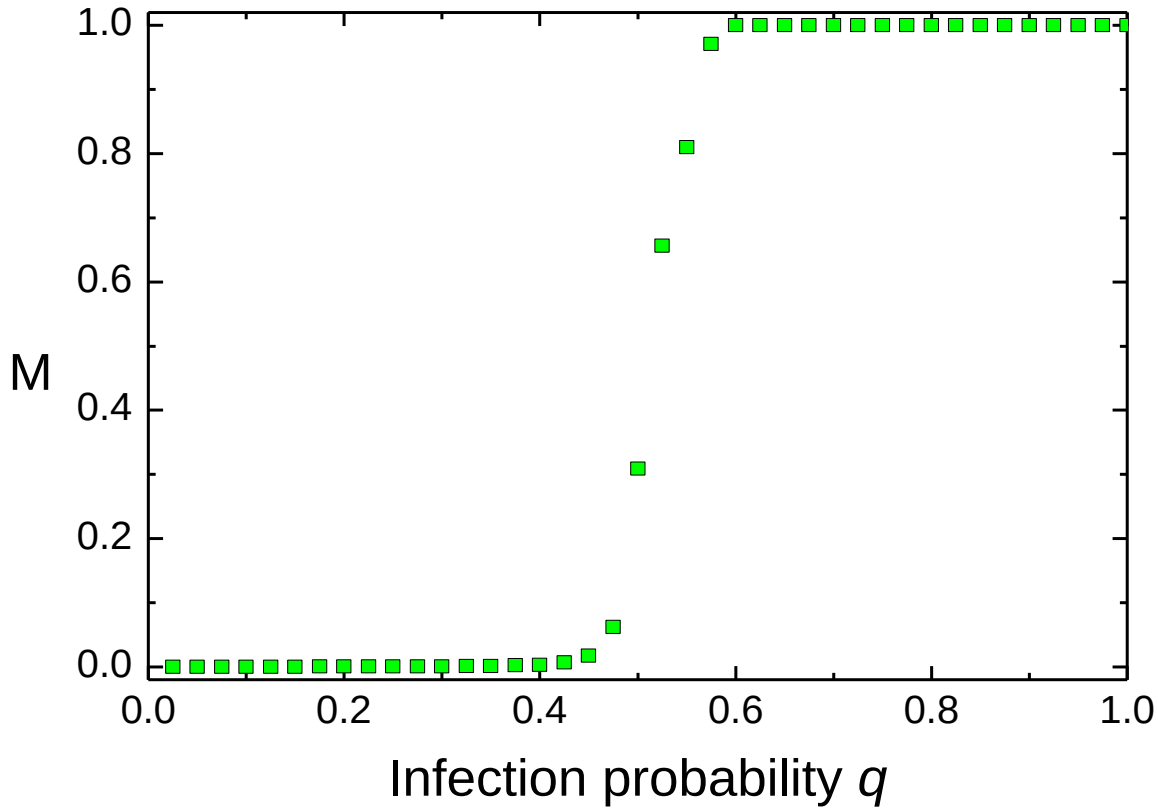
$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

Acquaintance immunization

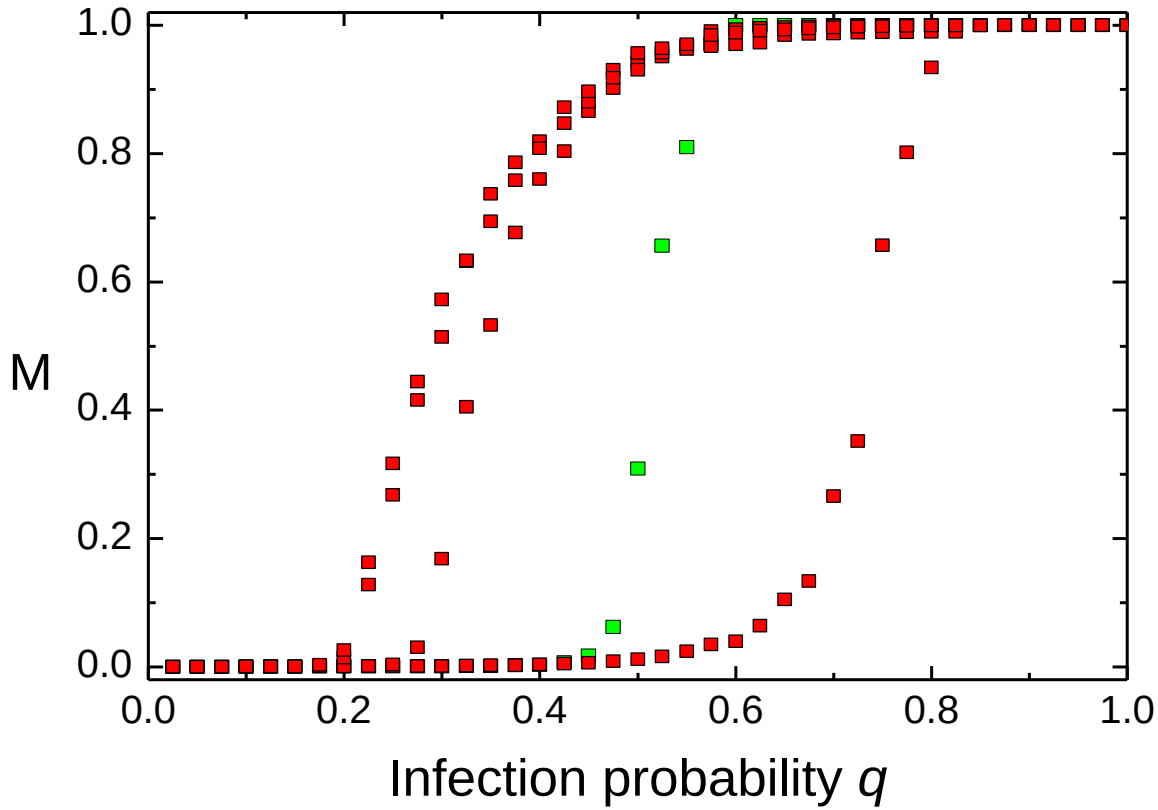
- **Suggested by Cohen, Havlin, ben-Avraham (PRL, 2003)**
- **Strategy: Randomly choose a node and immunize a random neighbor of this node.**



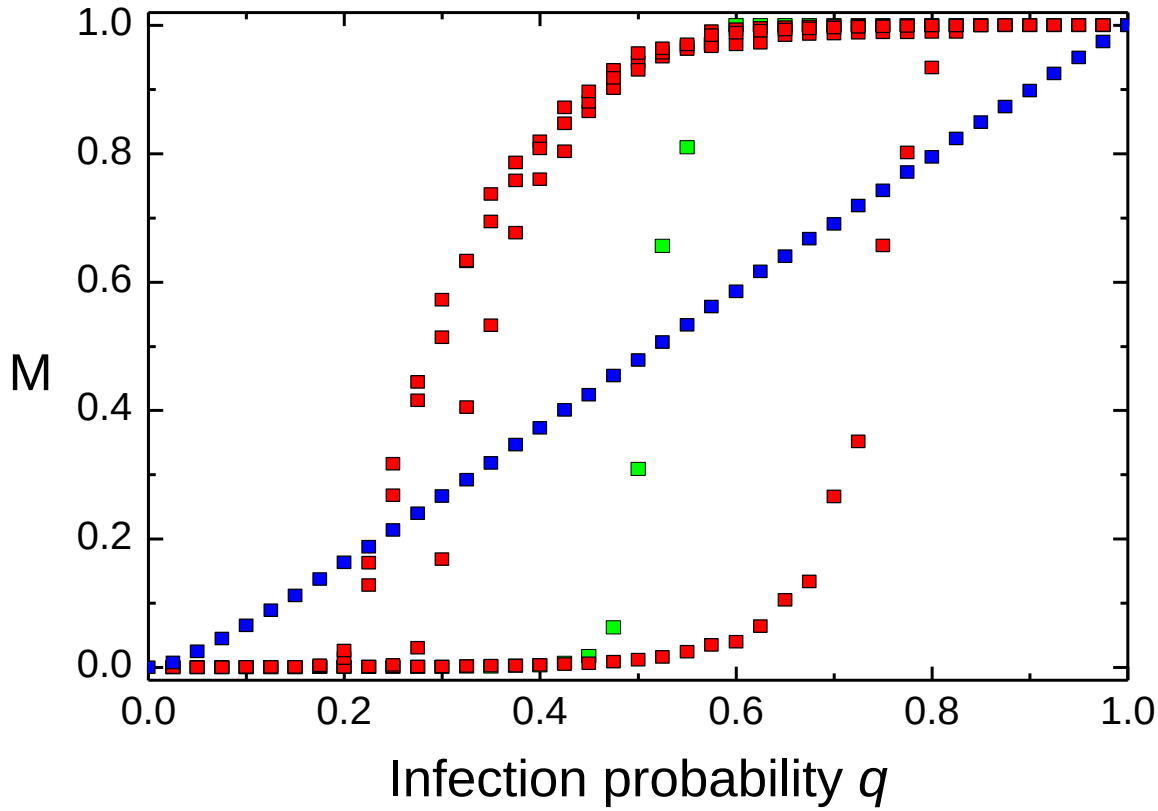
Comparison of the 3 network types



Comparison of the 3 network types

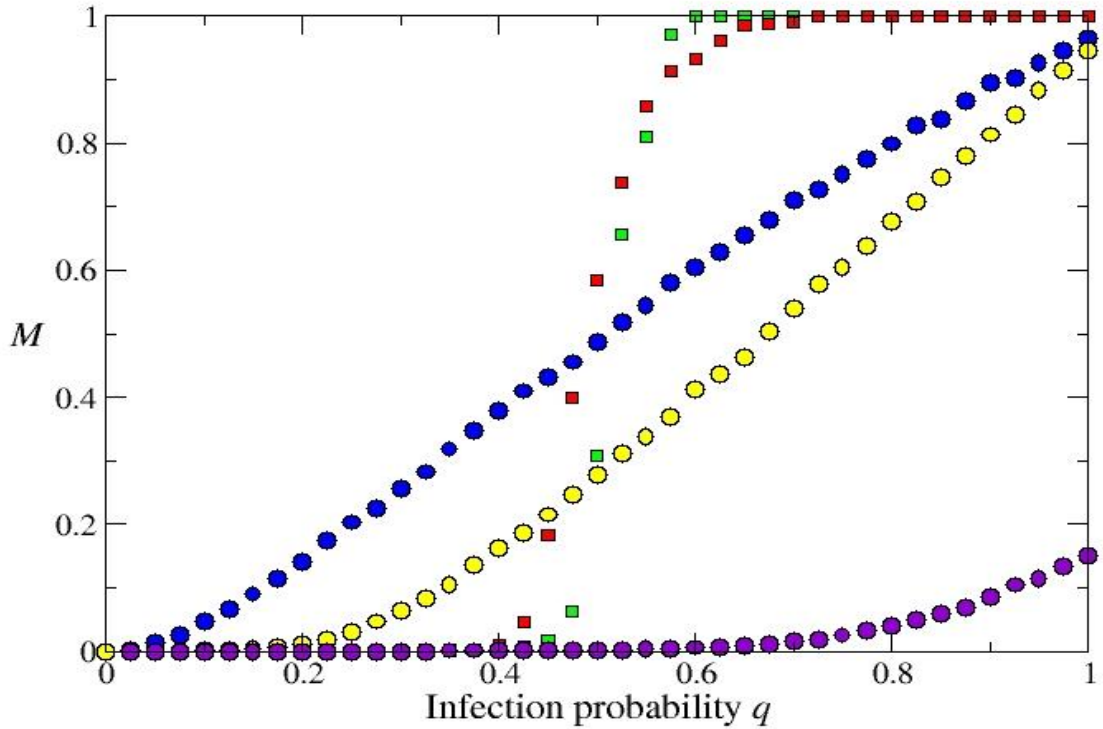


Comparison of the 3 network types



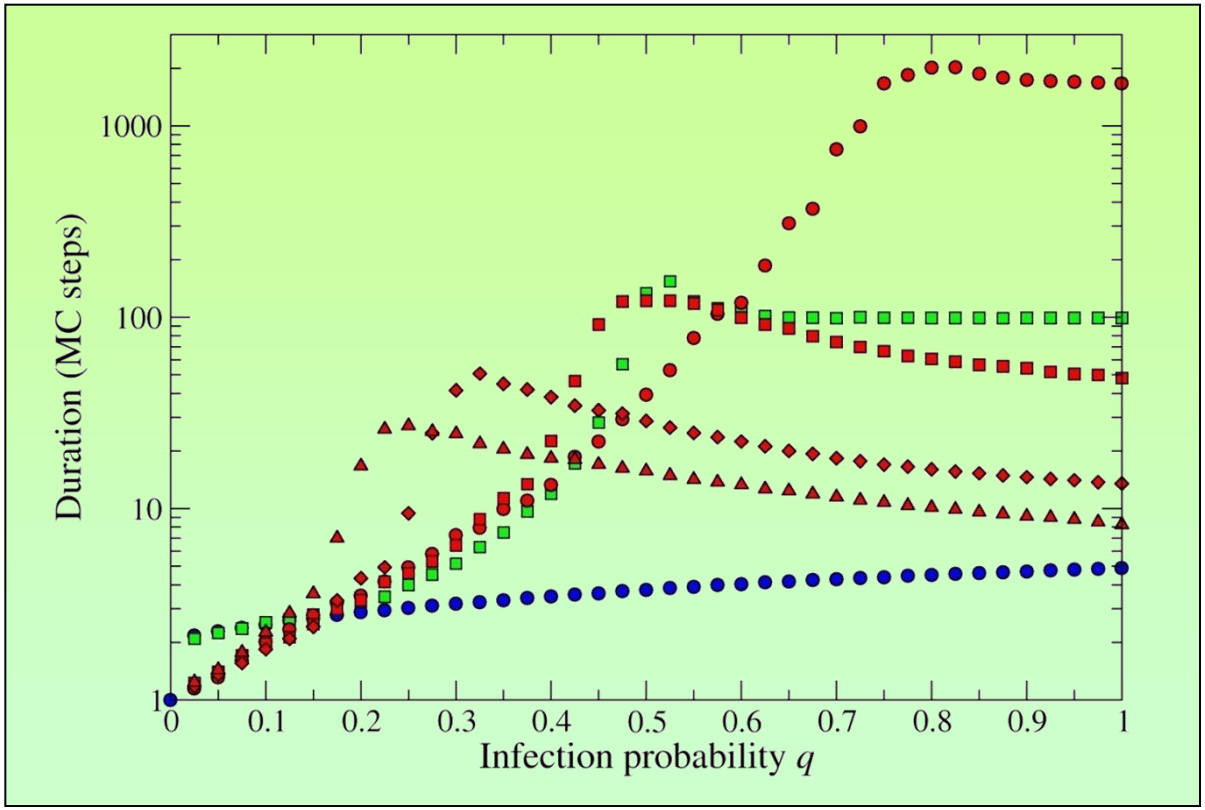
CASE 2

Comparison of different network types



- Lattice
- Small-world network
- Scale-free network, γ=2.0
- Scale-free network, γ=2.5
- Scale-free network, γ=3.0

Duration of epidemics



DISEASE SPREADING:

The dynamical consequences of a scale-free network are drastically different than those on lattice and small-world networks.

There is no threshold in the infected mass as a function of the infection transmission probability.

The starting point of the disease is important, since it determines whether the disease will spread or die out.

The spreading is rapid and manifests the small diameter of the network.

ATTACK TOLERANCE:

The tolerance of a network depends on its connectivity.

The random node removal is the marginal case where the critical threshold moves from p_c=0 to p_c=1.

A small bias in the probability of selecting nodes either retains or destroys the compactness of the cluster.

SIR model:

- **The network with a power-law structure is a more realistic representation than all the other network categories**
- **It does NOT show critical behavior**

What is Econophysics?

" Econophysics is the application of typical methods from physics to the study of the financial markets, seen as a complex system."

H. E. Stanley, Boston University, Boltzmann Medal 2004:

"For his influential contribution to several areas of statistical physics…"

Physica A, Vol. 285, p. 1 (2000)

Exotic statistical physics with applications to biology, medicine and economics.

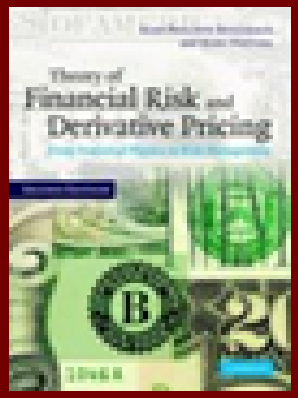
However…it is nothing really that new!

E. Majorana, ***Scientia***, Vol. 36, 58 (1942)

"On the value of the statistical laws in physics and in social sciences."

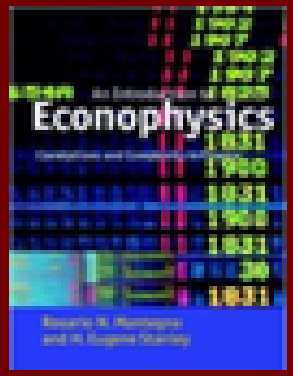
Textbooks in Econophysics

J.P. Bouchaud and M. Potters – Theory of Financial Risk and Derivative Pricing: from Statistical Physics to Risk Management, Cambridge University Press (2003)

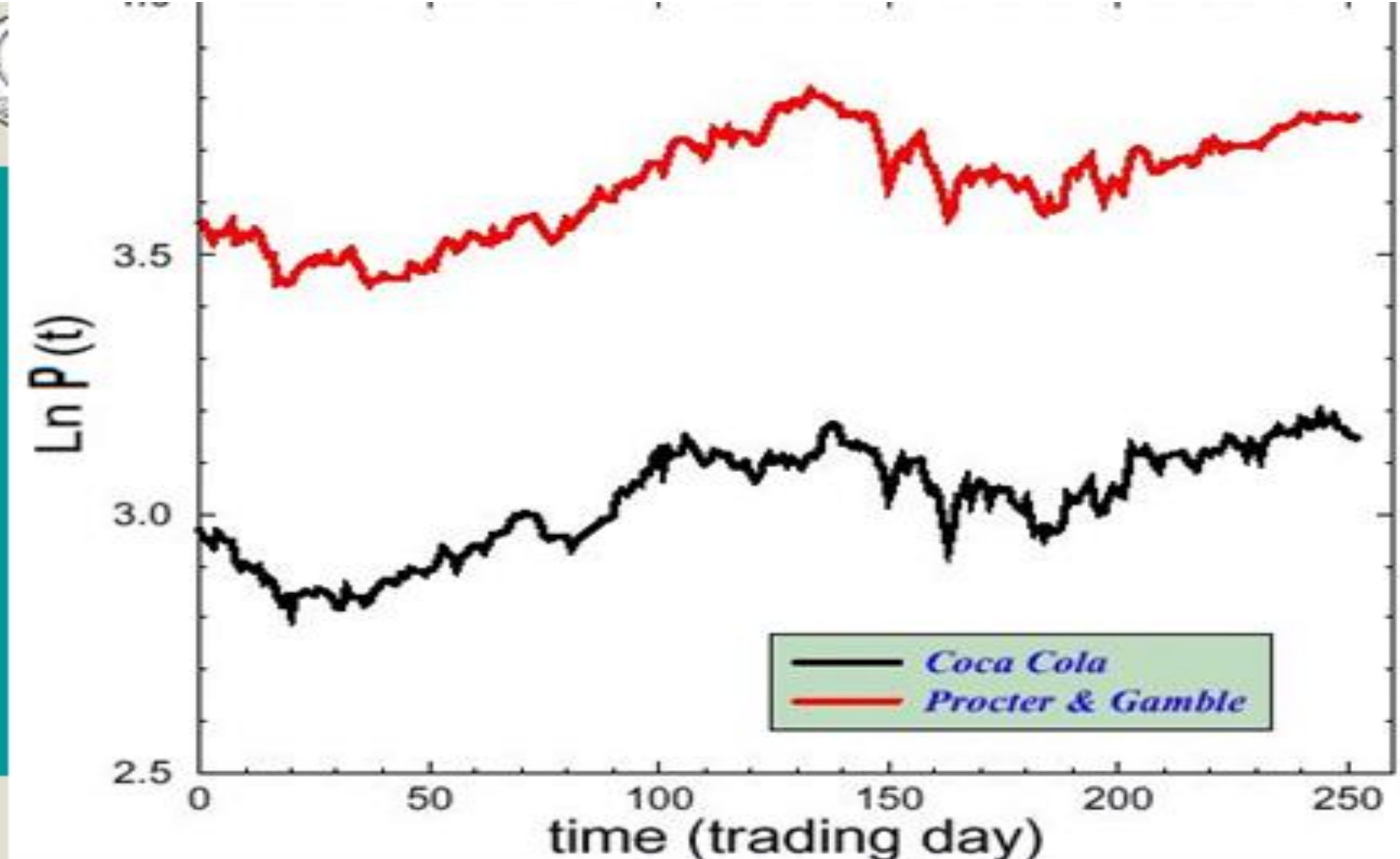


J. Voit – The Statistical Mechanics of Financial Markets, Springer (2005)

R.N. Mantegna and H.E. Stanley – An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press (2000)



Stock price changes



Financial Data

Price Index

$$P(t)$$

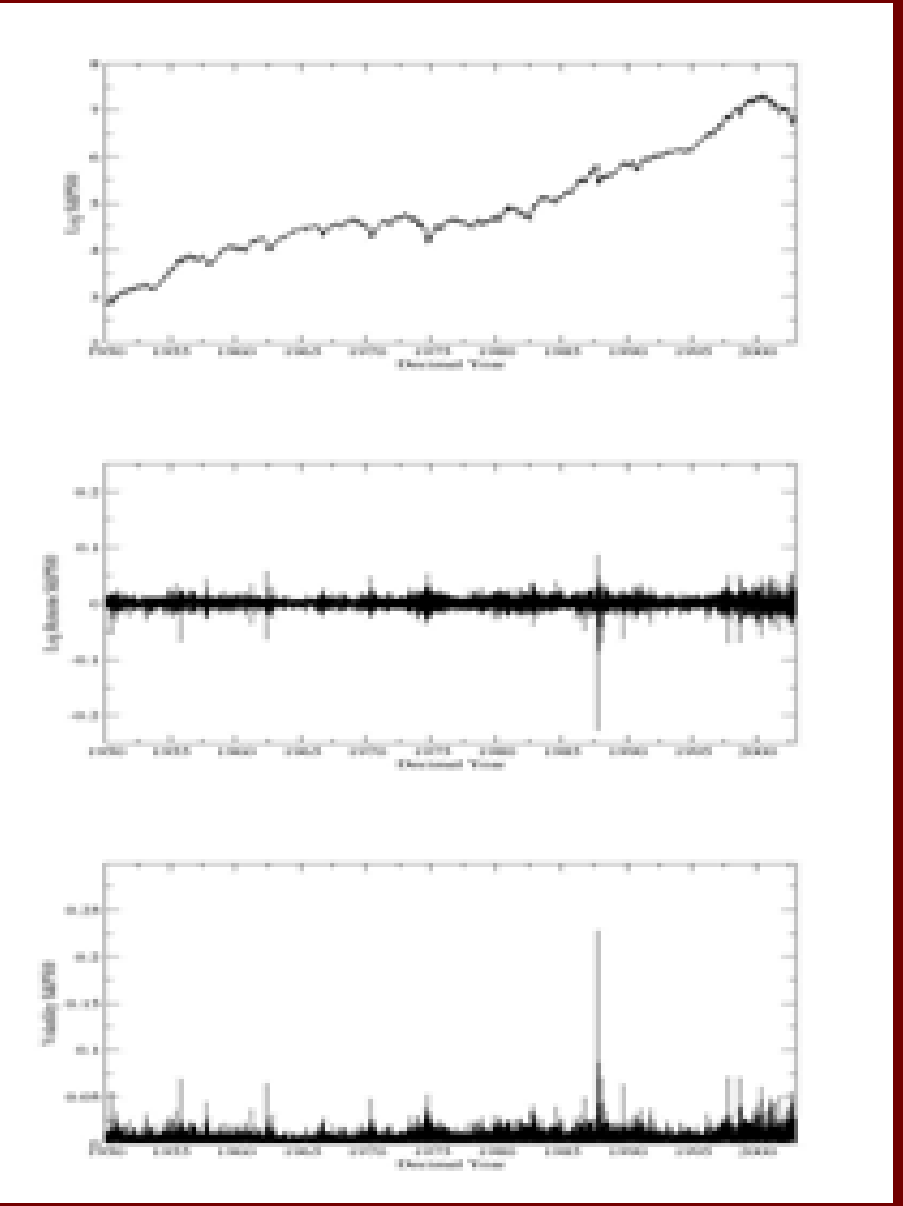
Logarithmic Price Returns

$$R(t) = \ln[P(t+1)] - \ln[P(t)]$$

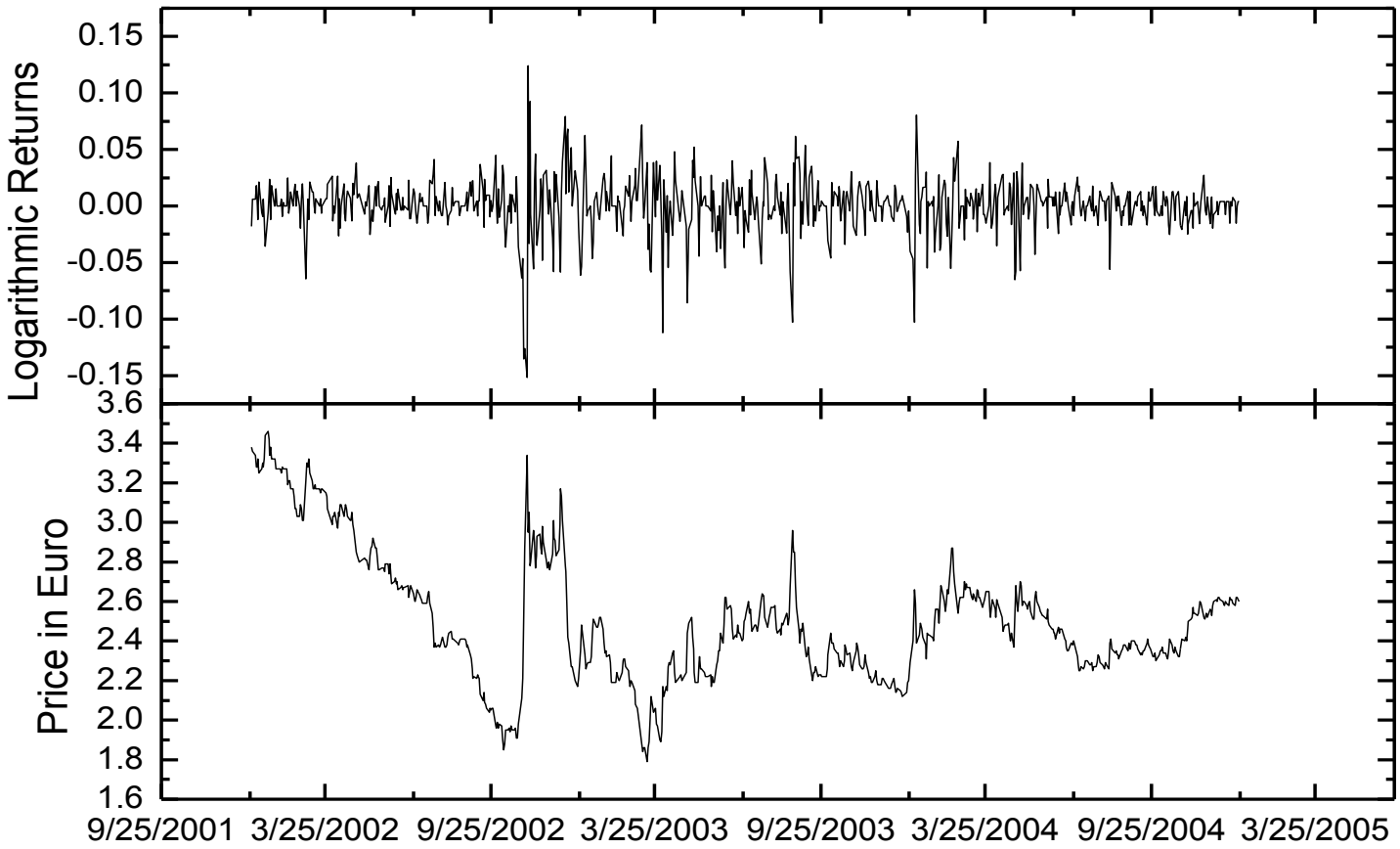
Volatility

$$v(t) = |R(t)|$$

Standard & Poor 500 (S&P500) data set from 3/1/1950 to 18/7/2003. N=13468.



Financial Time Series



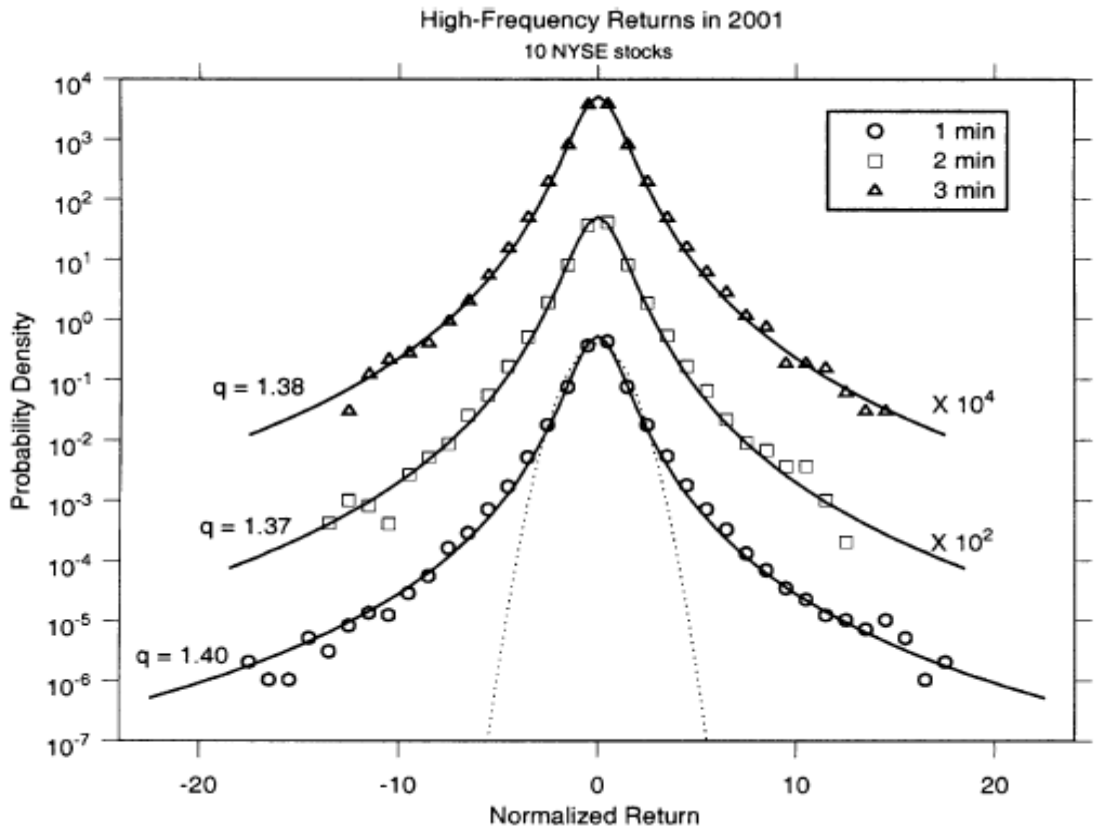
Theoretical Economics is dominated by pure mathematics:

- **Lemma/theorem style is required**
- **Little effort to compare theoretical predictions to "experimental data" - say, price record from real stock markets**
- **Bulk of papers are inaccessible and of no interest to "experimentalists" - practitioners of the field**

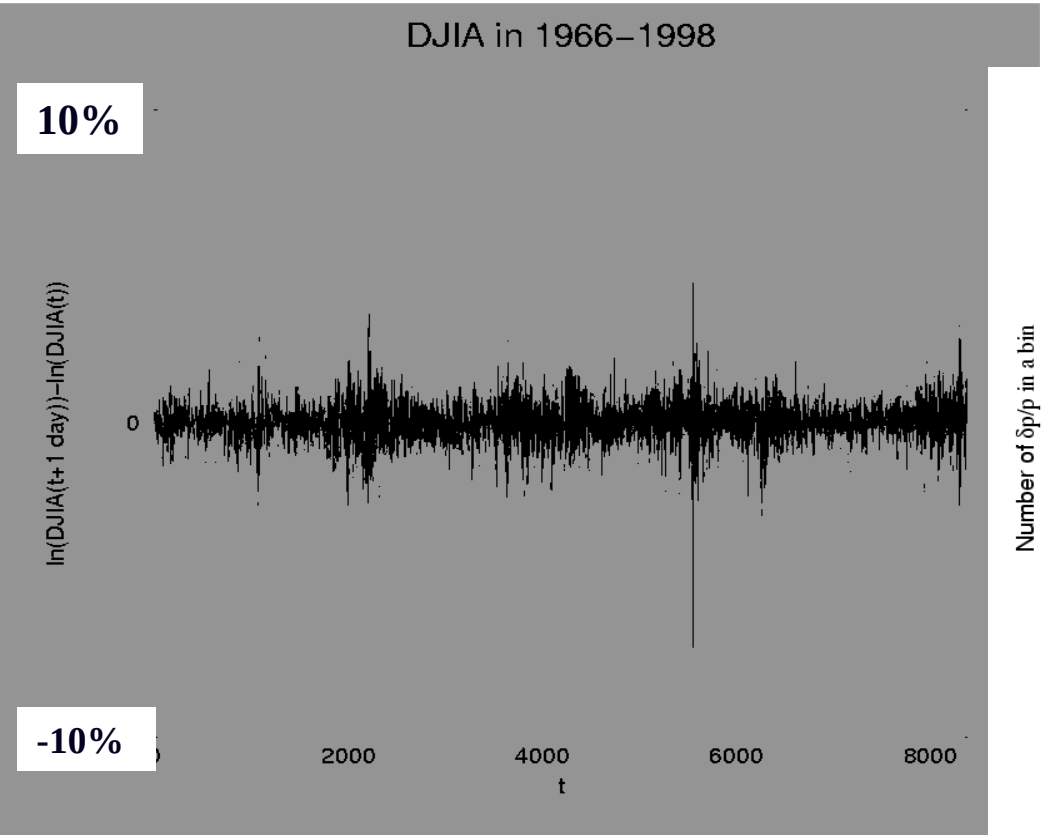
George Soros on theoretical economics

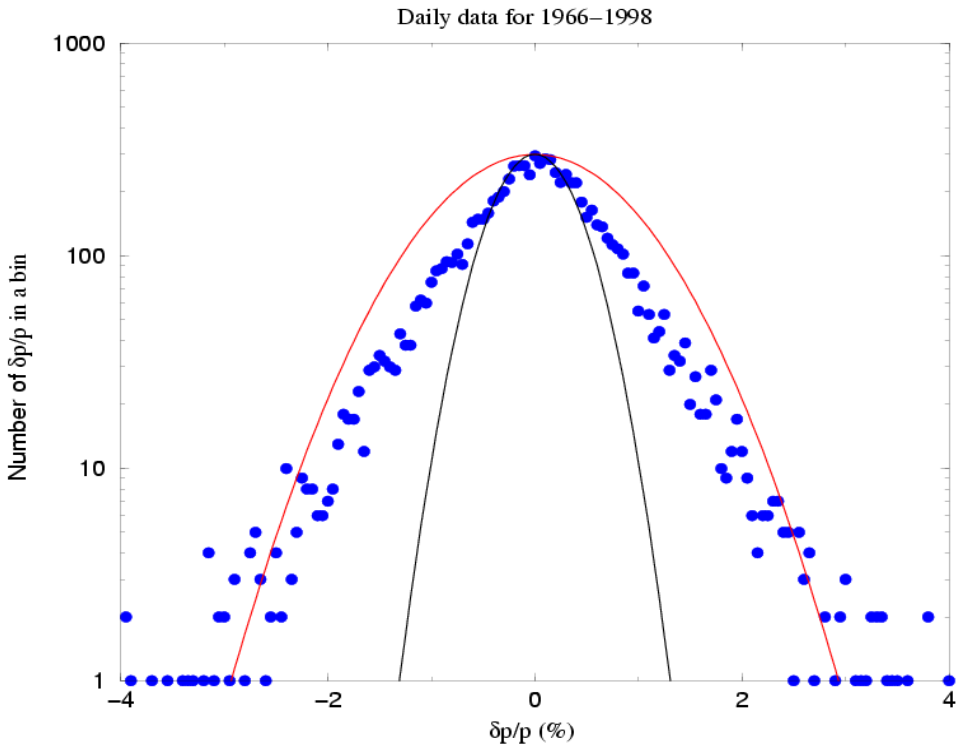
“Existing theories about the behavior of stock prices are remarkably inadequate. They are of so little value to the practitioner that I am not even fully familiar with them. The fact that I could get by without them speaks for itself.”

G. Soros, “Alchemy of Finance” 1994


High-Frequency Returns in 2001
10 NYSE stocks
1 min
2 min
3 min
q = 1.38
q = 1.37
q = 1.40
X 10^4
X 10^2
Probability Density
Normalized Return
10^4
10^3
10^2
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
-20
-10
0
10
20

Quick experiment: free data from www.nyse.com/marketinfo/nysestatistics.html



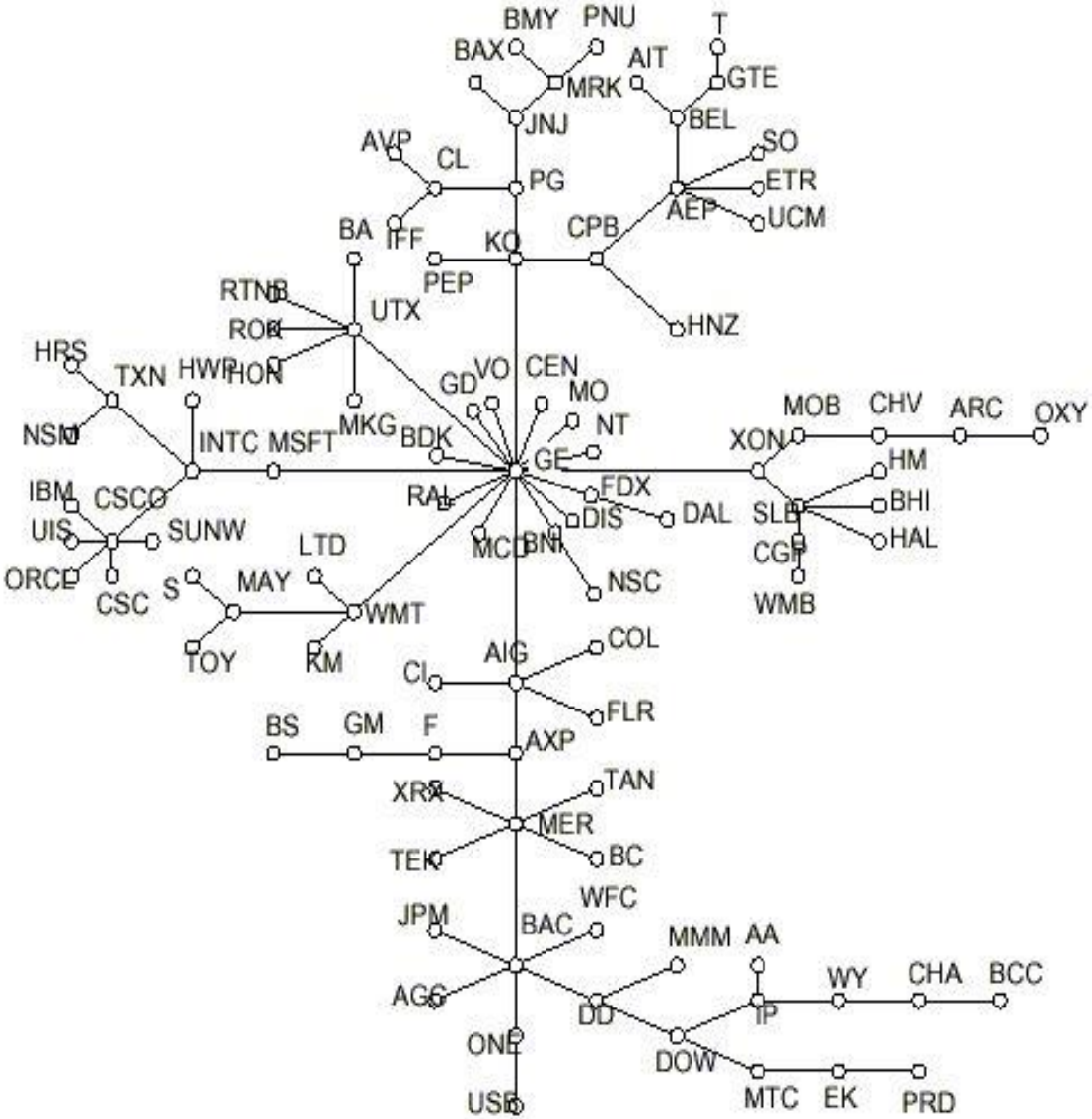


In a gaussian world the probability of the October 1987 crash would be 10^{-135}!

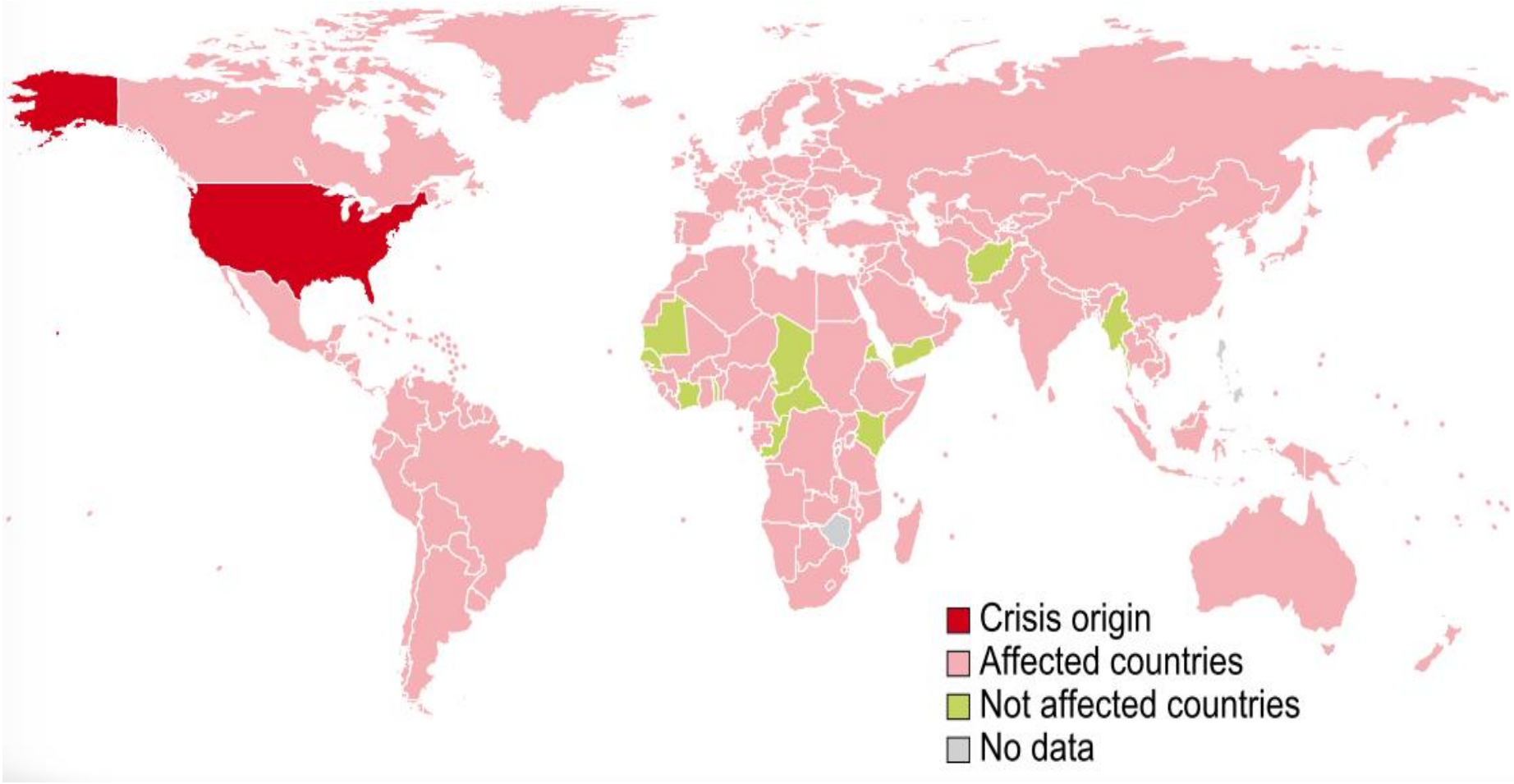
Filtering of the correlation matrix

•**Minimum Spanning Tree of the 100 most capitalized US stocks in 1998 (R.N.Mantegna).**

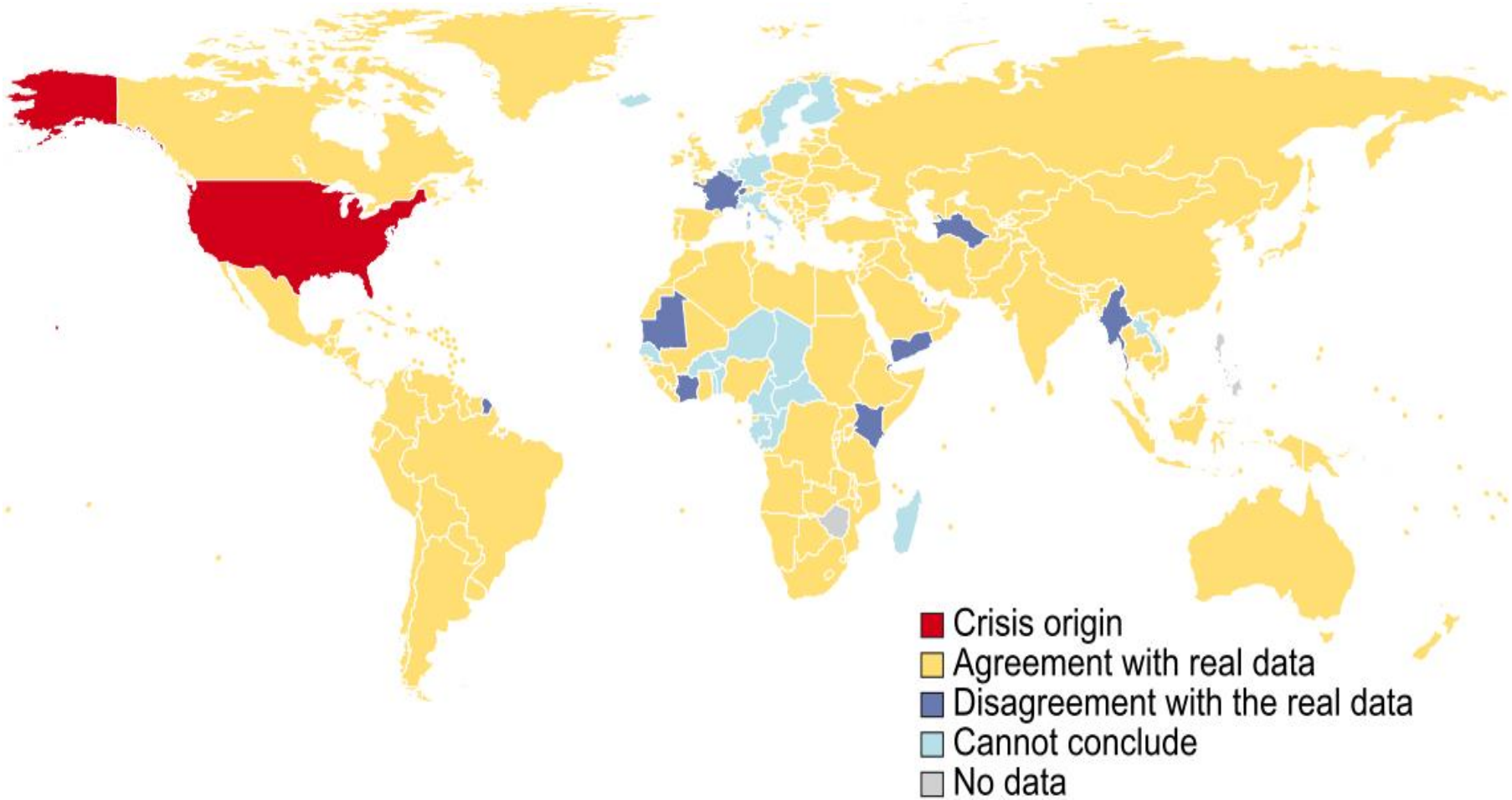
•**From n(n-1)/2 connections only n-1 survive.**



Countries affected by the 2008 economic crisis



Model results: a crisis starts in US and spreads through CON



Summary- Conclusions

- **Networks** is a new area in sciences which started from Physics, but pertains to ALL sciences today
- It shows rich dynamical behavior
- It has many-many applications in everyday life
- Will influence directly the way we live and act