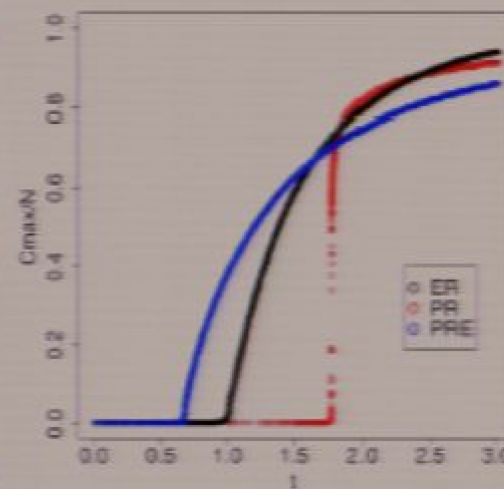
Title: Growing, Jamming and Changing Phase

Date: May 20, 2009 02:00 PM

URL: http://pirsa.org/09050004

Abstract: Key notions from statistical physics, such as "phase transitions" and "critical phenomena", are providing important insights in fields ranging from computer science to probability theory to epidemiology. Underlying many of the advances is the study of phase transitions on models of networks. Starting from the classic ideas of Erdos and Renyi, recent attempts to control and manipulate the nature of the phase transition in network connectivity will be discussed. Next, the influence of self-organization on phase transitions will be presented, as well as connections between the jamming transition in models of granular materials and constraint satisfaction problems in computer science. Finally, turning to network growth, I will show that local optimization can play a fundamental role leading to the mechanism of Preferential Attachment, which previously had been assumed as a basic axiom and, furthermore, resolves a long standing controversy between Herb Simon and Benoit Mandelbrot.

"Growing, Jamming and Changing Phase"



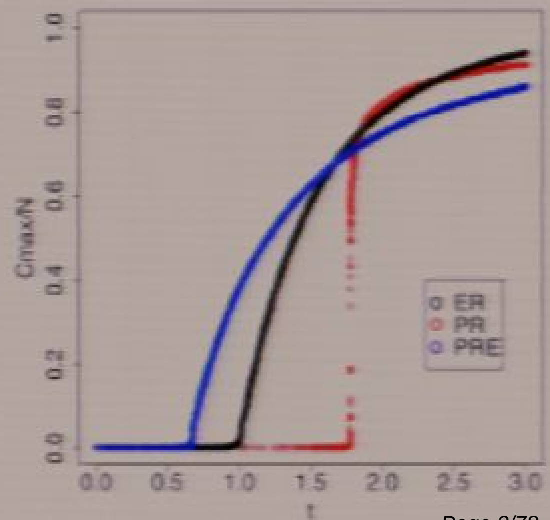
Raissa D'Souza
University of California, Davis
Dept of Mechanical and Aeronautical Eng.
External Professor, **Santa Fe Institute**

Complex systems

- Multiple length and time scales

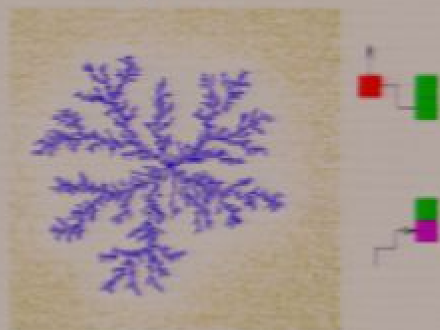
- Irregular connectivity / Networked patterns of interaction

- Emergent behaviors (Collective behaviors not captured by equations of individual parts)
  - self organization
  - phase transtions

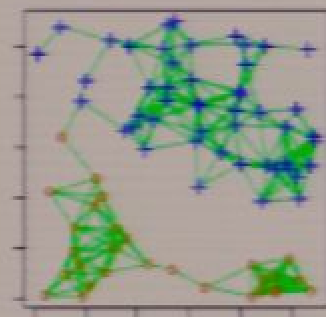


My journey into complex systems

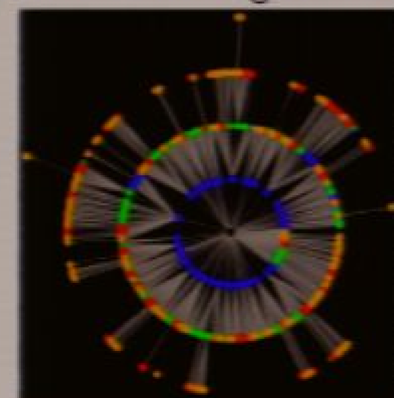
Thermodynamics and computation



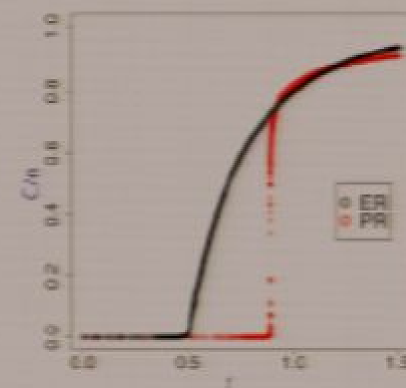
Sensor networks, random walks and timescales



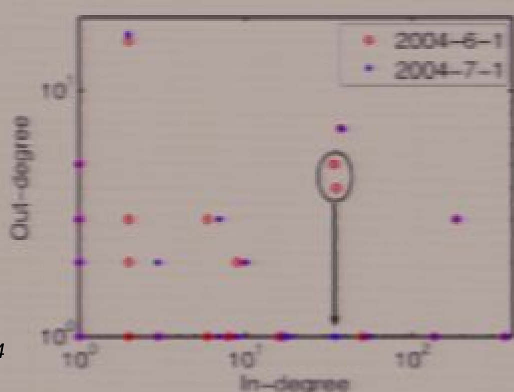
Optimization origins of PA



Self-organized flow and jams

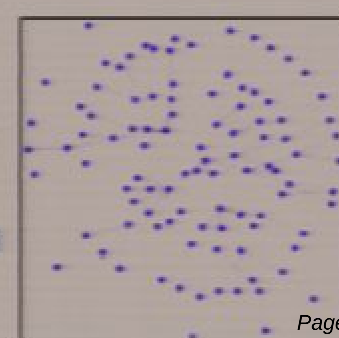


Phase transitions

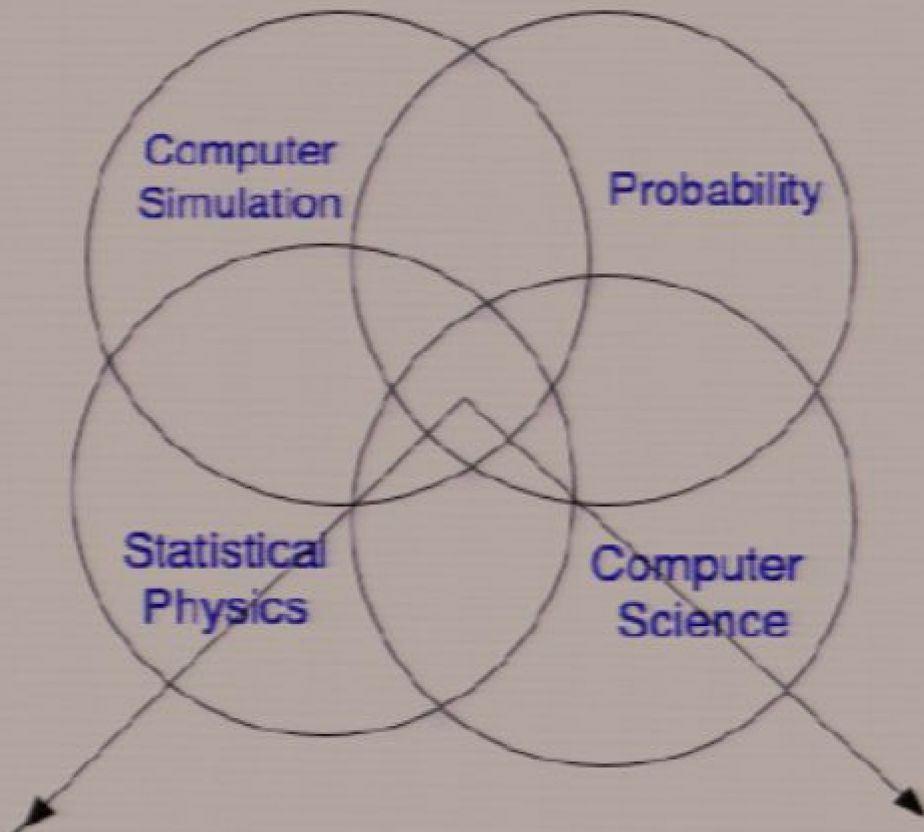


Engineered systems

Feedback and optimization



Fundamental connections



Lattice Models

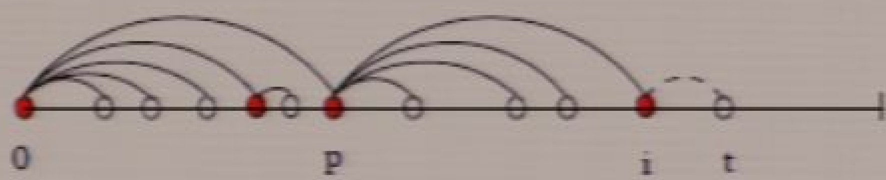
- Ising
- Potts
- DLA
- BML

- Phase transitions
- Percolation
- Comp. Complexity

Network Models

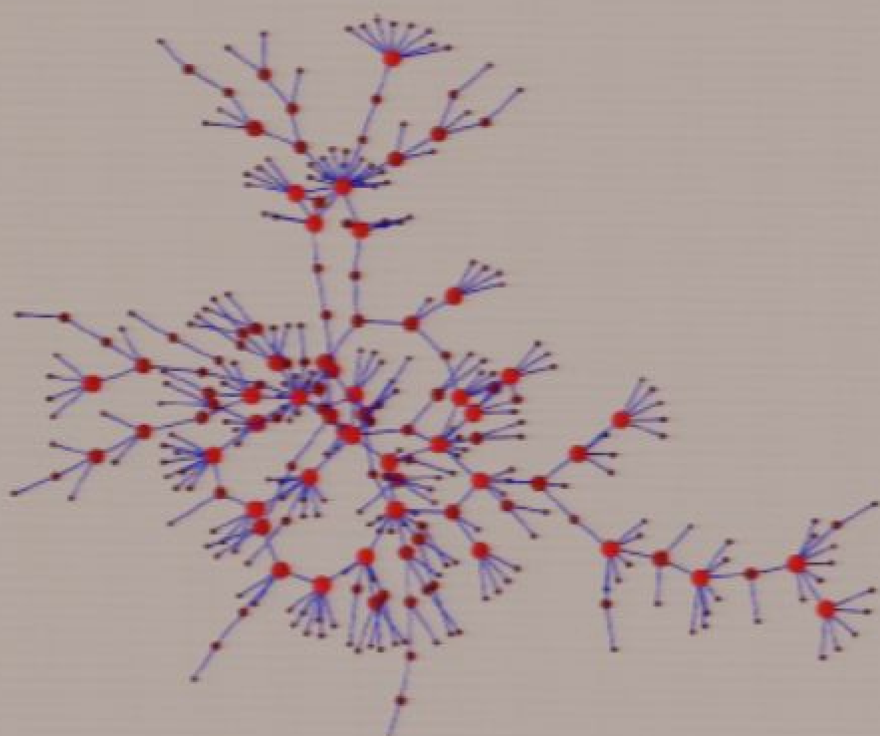
- Boolean
- Circuit
- Random
- Bayesian

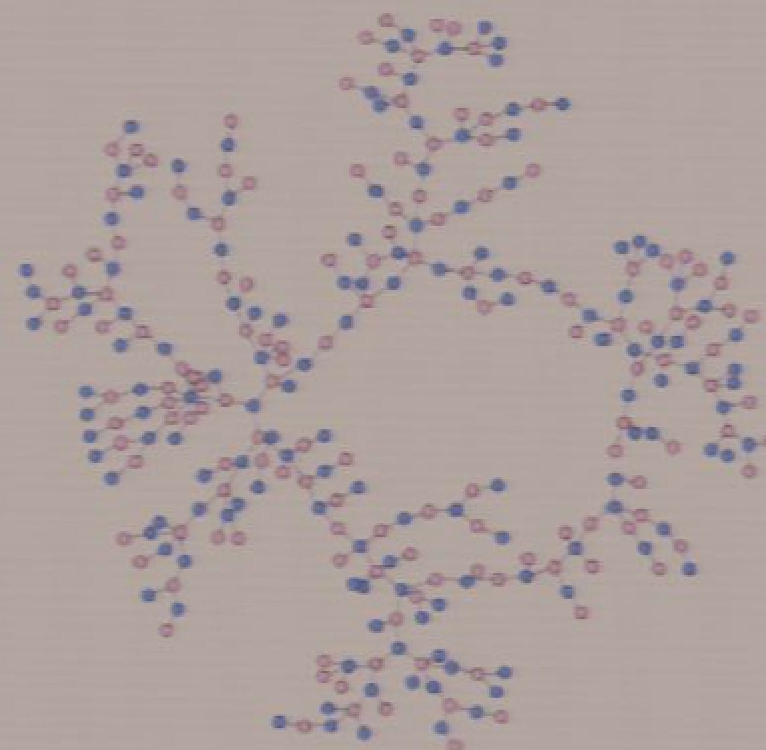
Statistical Physics, Networks and Algorithms



- Locality = decentralized control.
- Self-organization = scalable growth.
- Scaling behaviors, Fluctuations, Correlations (micro to macro)
- Applications: Routing, Load balancing, Distributed computing/ data bases.

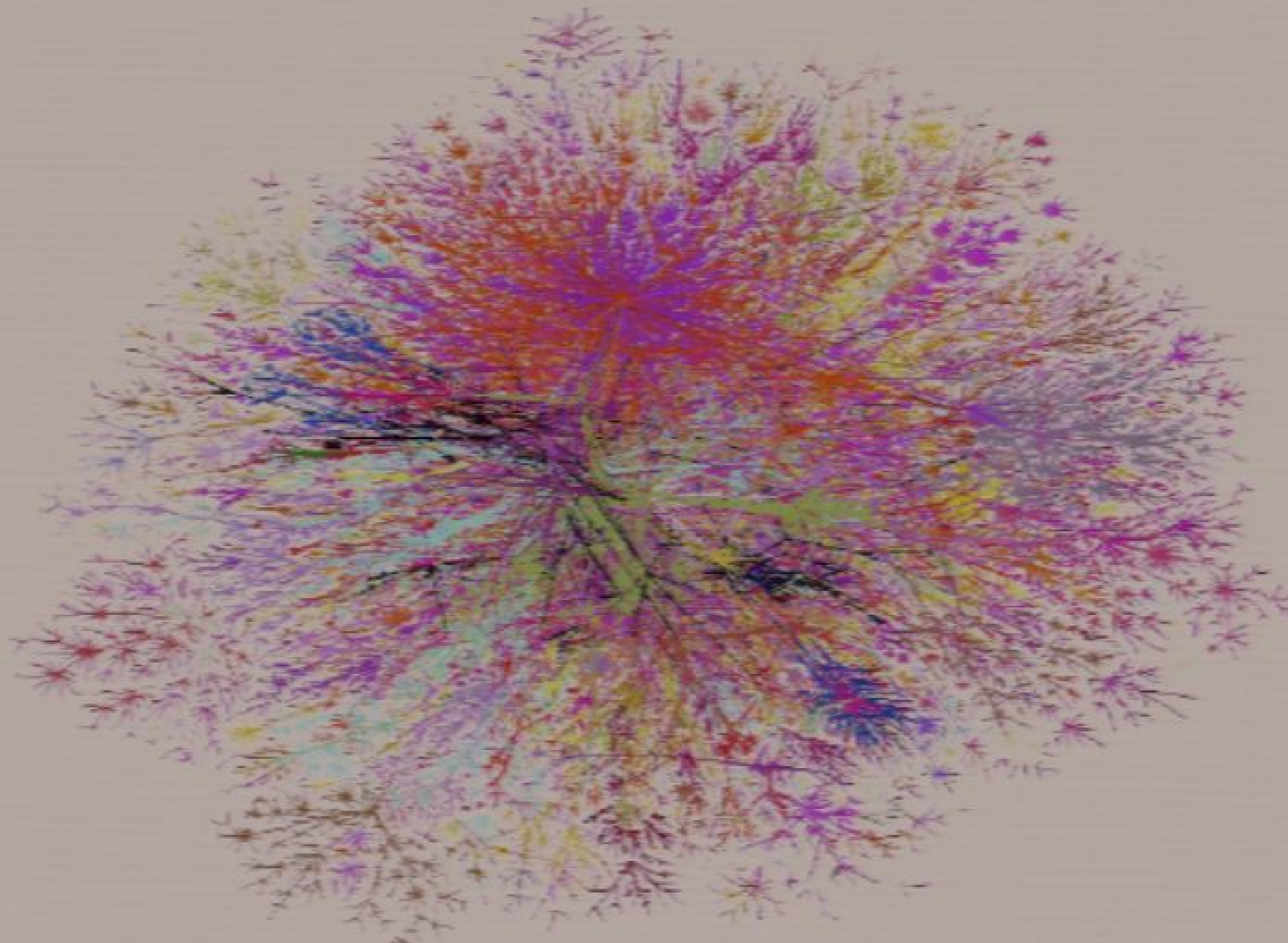
Example social networks
(Immunology; viral marketing)





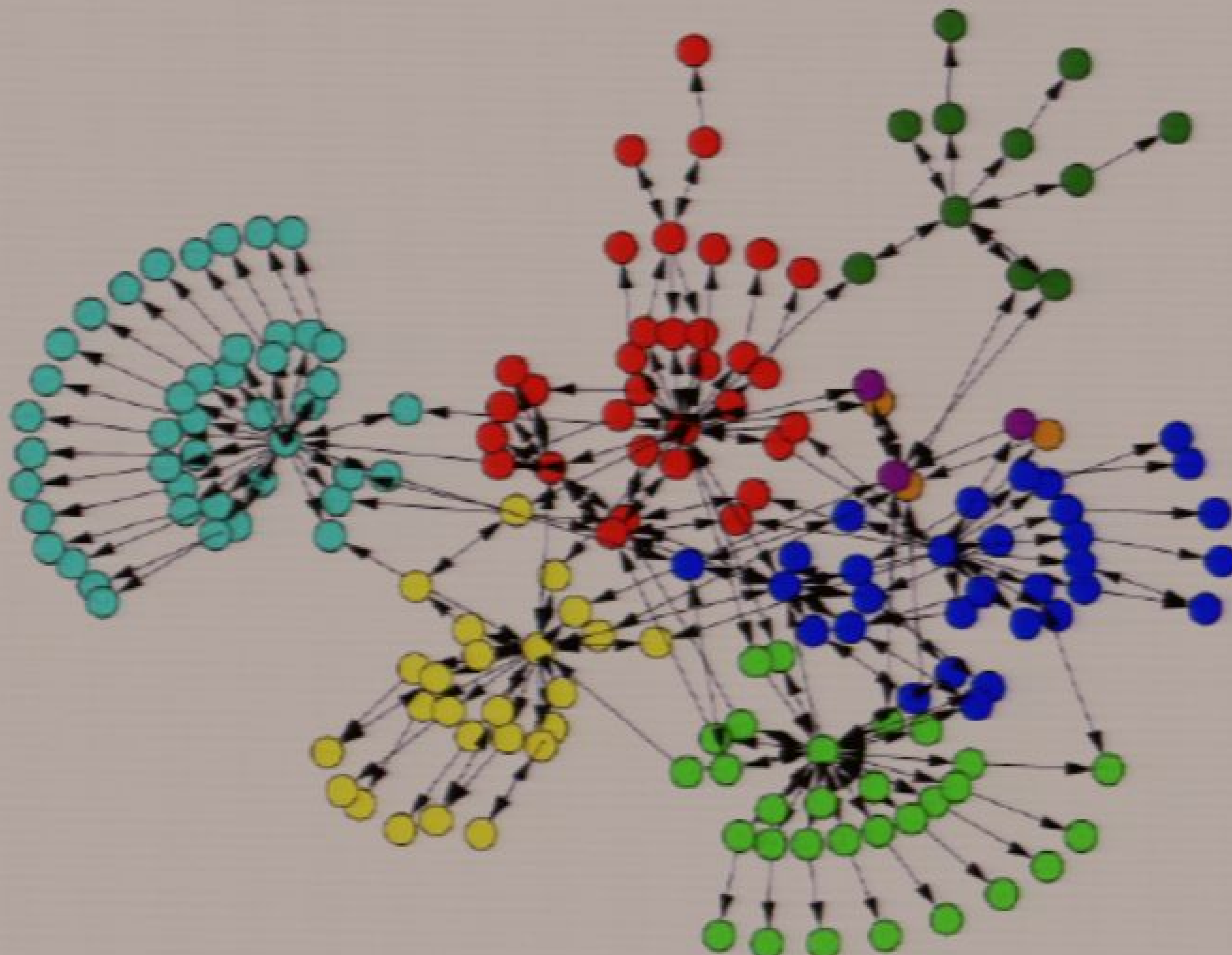
M. E. J. Newman

The Internet
(Robustness to failure; optimizing future growth; testing protocols on sample topologies)



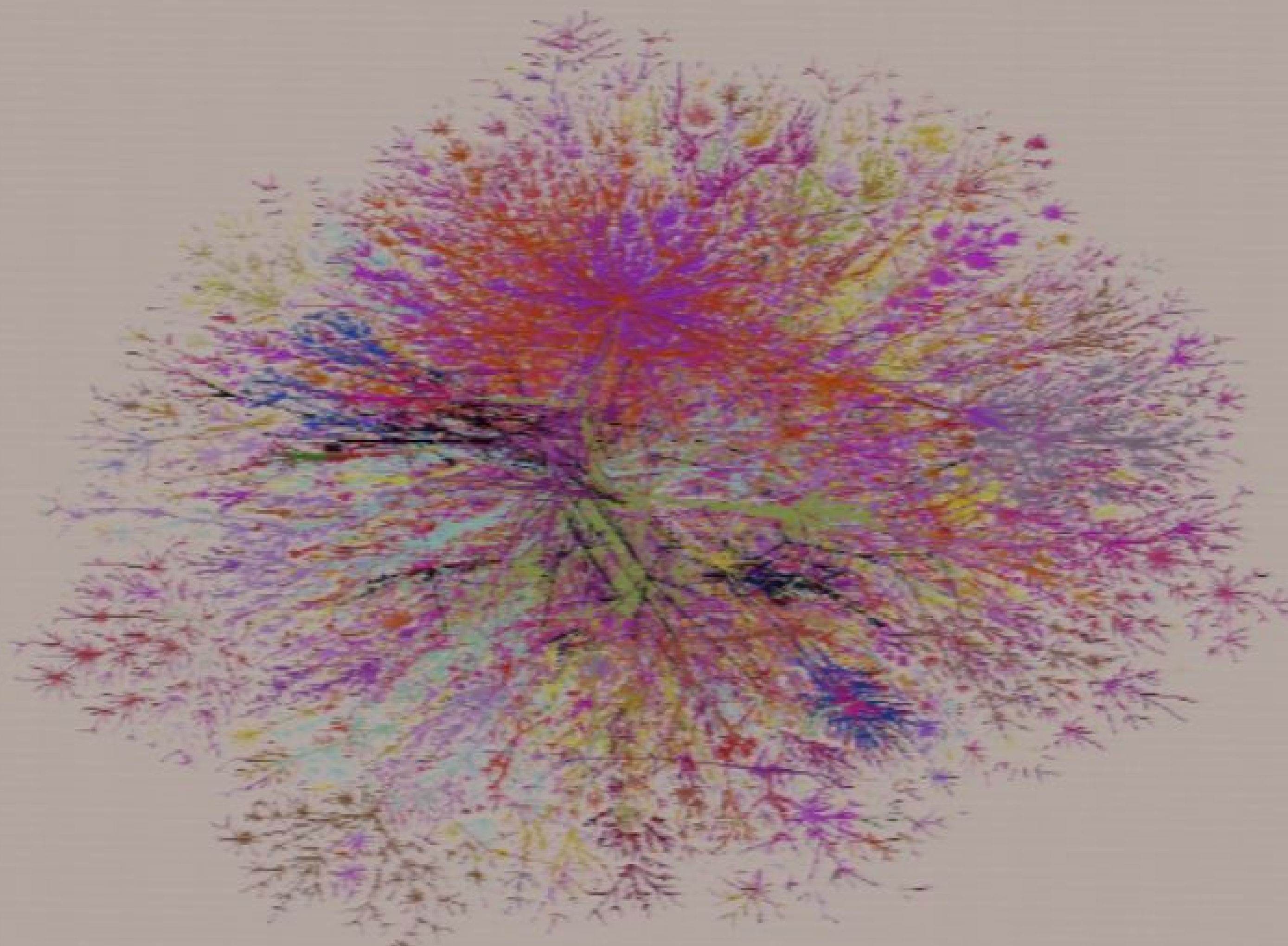
H. Burch and B. Cheswick

A typical web domain
(Web search/organization and growth centralized vs. decentralized protocols)



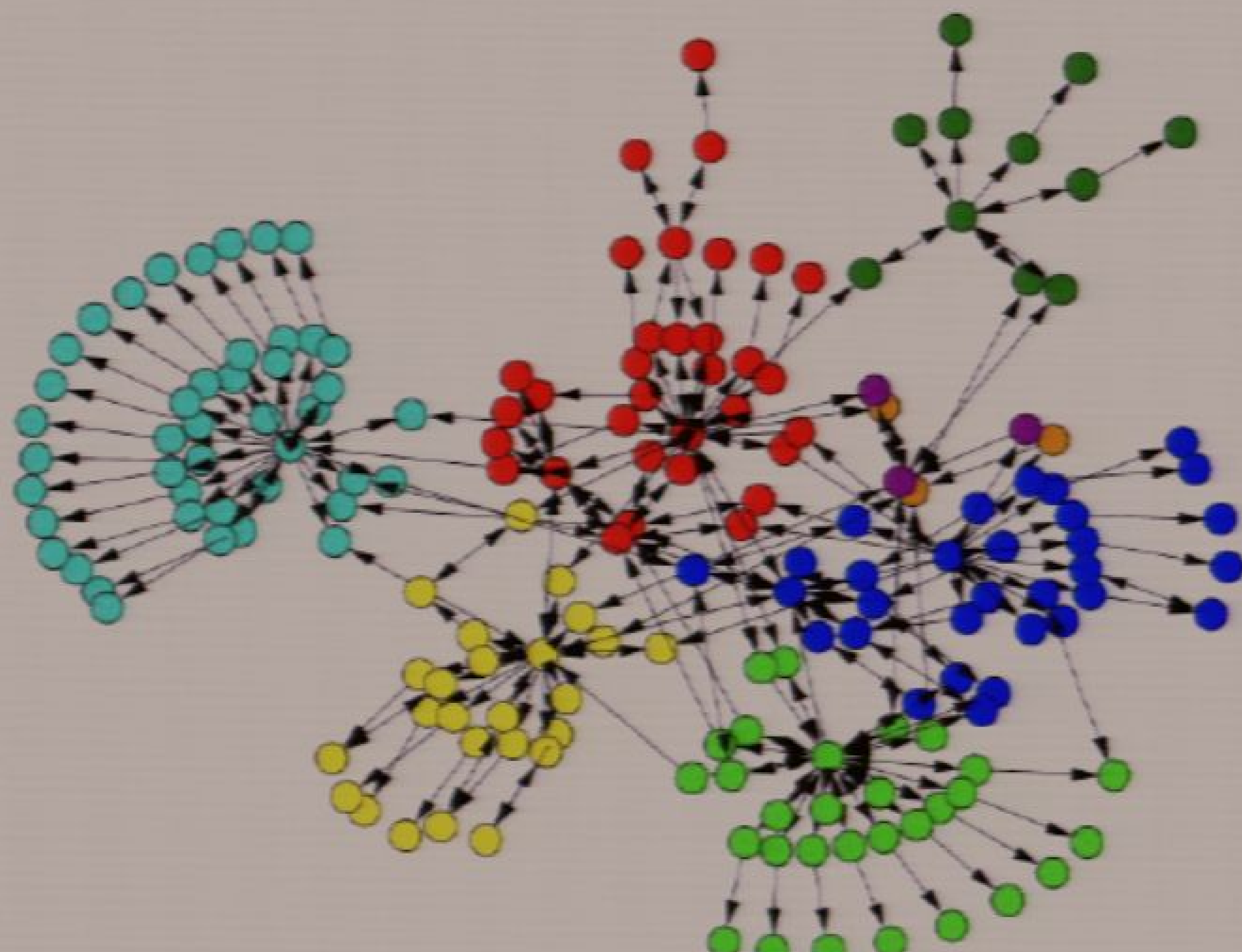
M. E. J. Newman

The Internet
(Robustness to failure; optimizing future growth; testing protocols on sample topologies)



H. Burch and B. Cheswick

A typical web domain
(Web search/organization and growth centralized vs. decentralized protocols)



M. E. J. Newman

The airline network
(Optimization; dynamic external demands; resilience)

k-core decomposition

resilience

[Wuellner, Roy, D'Souza]

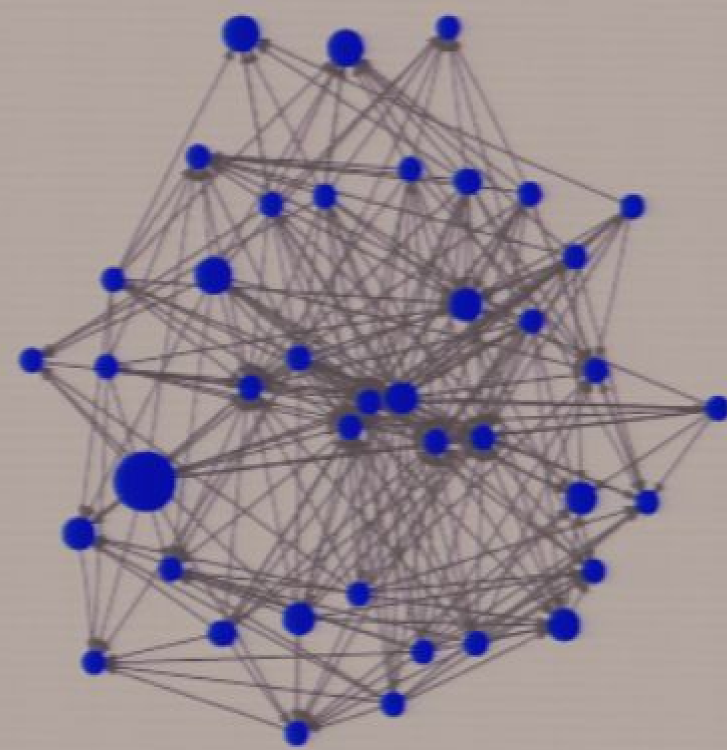
Software systems

(Highly evolveable, modular, robust to mutation, exhibit punctuated eqm)

Open-source software as a "systems" paradigm.

Networks:

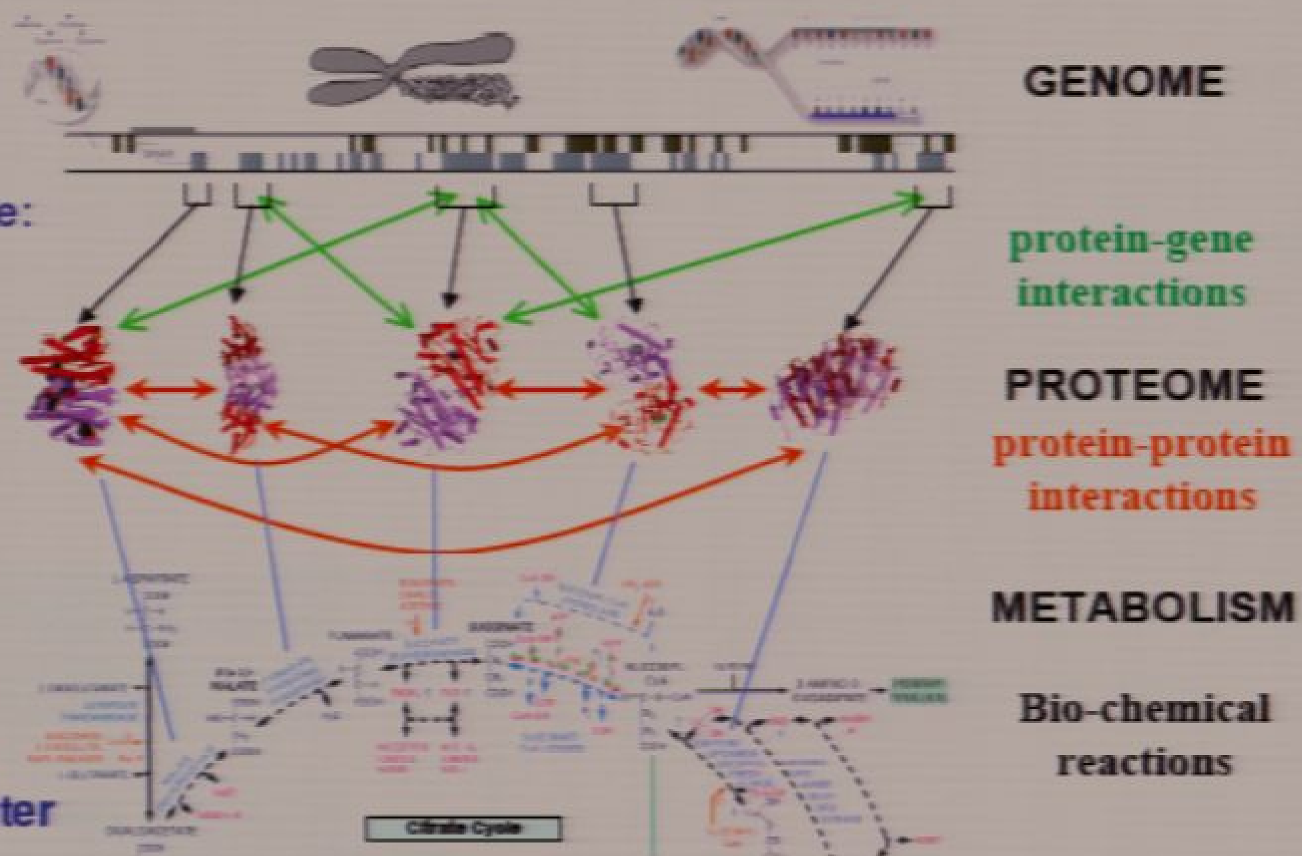
- Function calls
- Email communication
- Socio-Technical congruence



(Devanbu, D'Souza, Filkov: NSF)

Biology: Networks at many levels

Control mechanisms / drug design/ gene therapy / biomarkers of disease



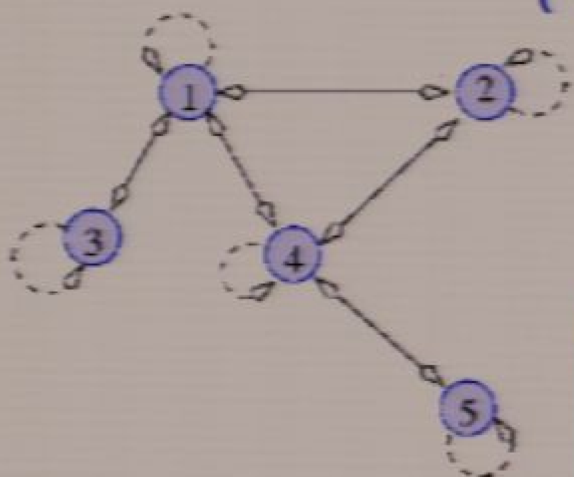
- Genome, Proteome: **Dandekar Lab**
- Metabolome: **Fiehn Lab**
- Data intergration BIOshare **Lin, Genome Center**
- Network structure / search for biomarkers: **D'Souza**

Can we represent a network as a mathematical object?

NETWORK TOPOLOGY

Connectivity matrix, M:

$$M_{ij} = \begin{cases} 1 \text{ if edge exists between } i \text{ and } j \\ 0 \text{ otherwise.} \end{cases}$$



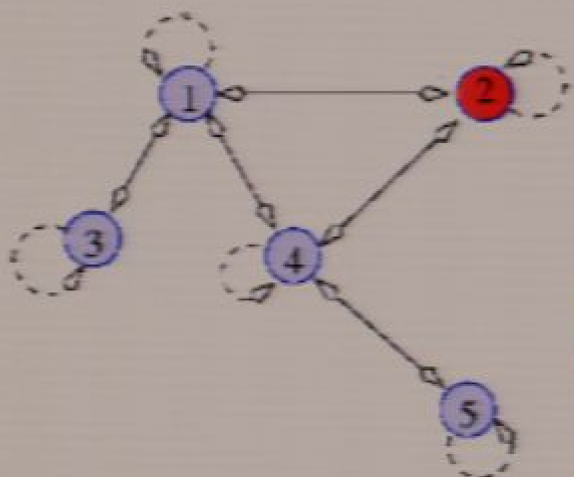
$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} = M$$

Node **degree** is number of links.

(Diameter, Clustering, Assortativity, Betweenness centrality)

Network Activity: FLOWS on NETWORKS
(Spread of disease, routing data, materials transport/flow)

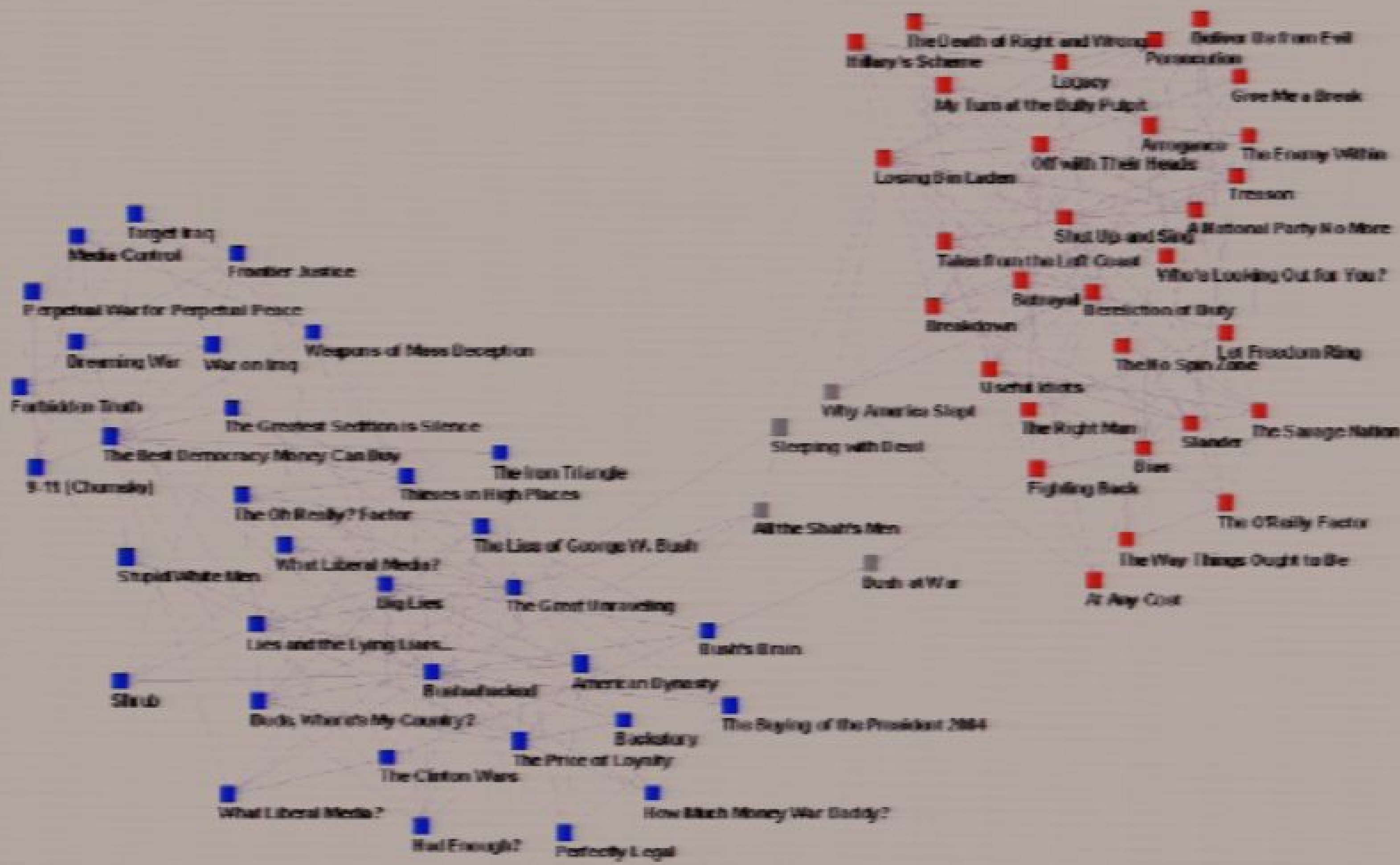
Random walk on the network has state transition matrix, P:



$$\begin{pmatrix} 1/4 & 1/3 & 1/2 & 1/4 & 0 \\ 1/4 & 1/3 & 0 & 1/4 & 0 \\ 1/4 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/3 & 0 & 1/4 & 1/2 \\ 0 & 0 & 0 & 1/4 & 1/2 \end{pmatrix} = P$$

- Eigenvectors → **Page Rank**, Communities/modules.
- Eigenvalues → timescales.

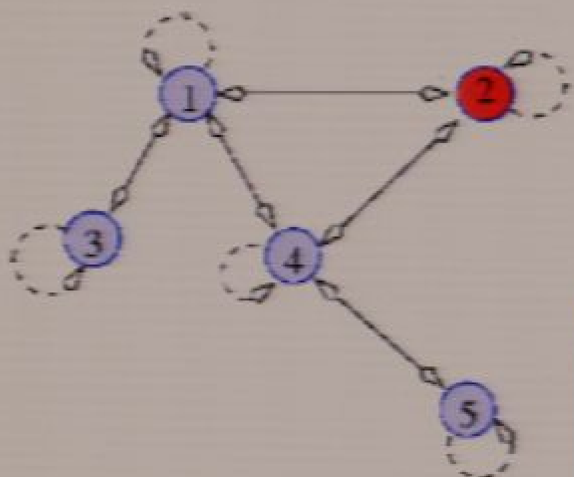
Example Eigen-technique: Community structure (Political Books 2004)



M. Girvan and M. E. J. Newman

Network Activity: FLOWS on NETWORKS
(Spread of disease, routing data, materials transport/flow)

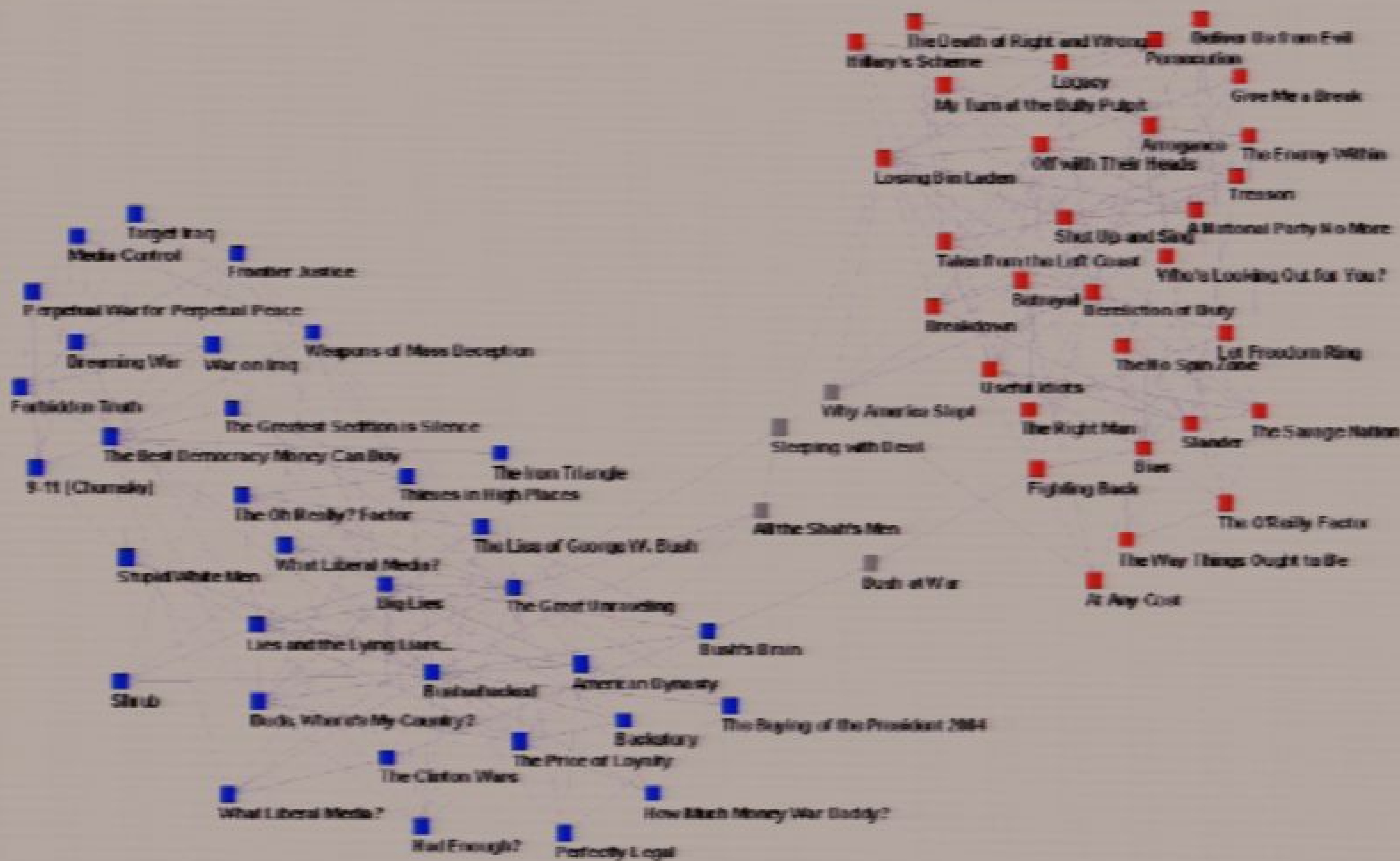
Random walk on the network has state transition matrix, P:



$$\begin{pmatrix} 1/4 & 1/3 & 1/2 & 1/4 & 0 \\ 1/4 & 1/3 & 0 & 1/4 & 0 \\ 1/4 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/3 & 0 & 1/4 & 1/2 \\ 0 & 0 & 0 & 1/4 & 1/2 \end{pmatrix} = P$$

- Eigenvectors → **Page Rank**, Communities/modules.
- Eigenvalues → timescales.

Example Eigen-technique: Community structure (Political Books 2004)



M. Girvan and M. E. J. Newman

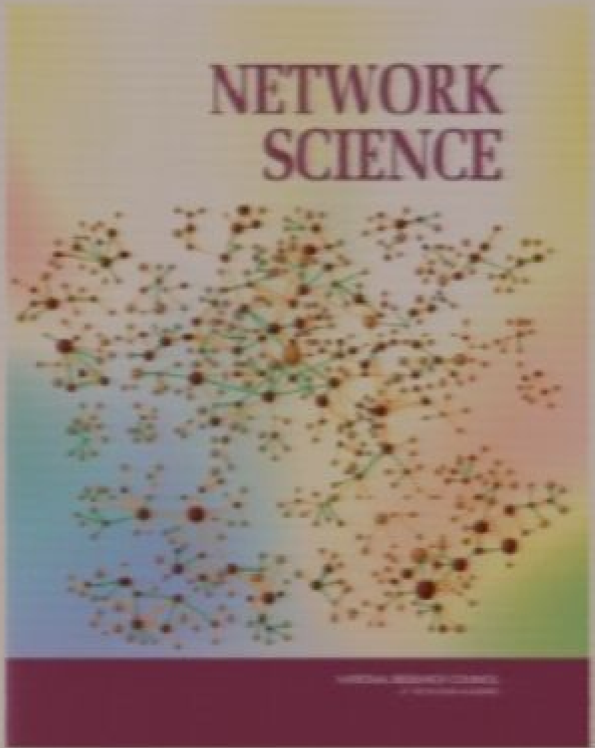
Networks: Physical, Biological, Social

- **Geometric** versus **virtual** (Internet versus WWW).
- **Natural** /spontaneously arising versus **engineered** /built.
- Each network **optimizes** something unique.
- Identifying **similarities** and fundamental **differences** can guide future design/understanding/control.
- Interplay of **topology** and **function** ?
- One unifying feature: **Broad heterogeneity in node degree.**

"Network Science"

Explosion of work and tools in past decade

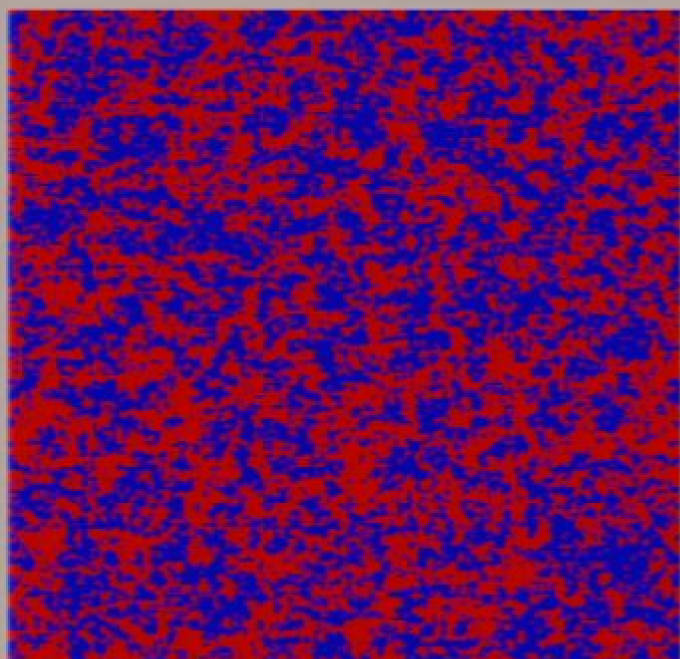
- New tools: R, Graphviz, Pajek, igraph, Network Workbench, NetworkX, Netdraw, UCInet, Bioconductor....



- 2005 NAS/NRC study calling for a cross-cutting science of networks

But focus today on **topology** and **phase transitions** ...

The Ising Model (Lenz 1920, Ising 1924)



(High T)

- 2-D lattice, at each site $s_i \in \{-1, +1\}$.

- Nearest neighbor interaction:

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J_{ij} s_i s_j - \sum_i h s_i$$

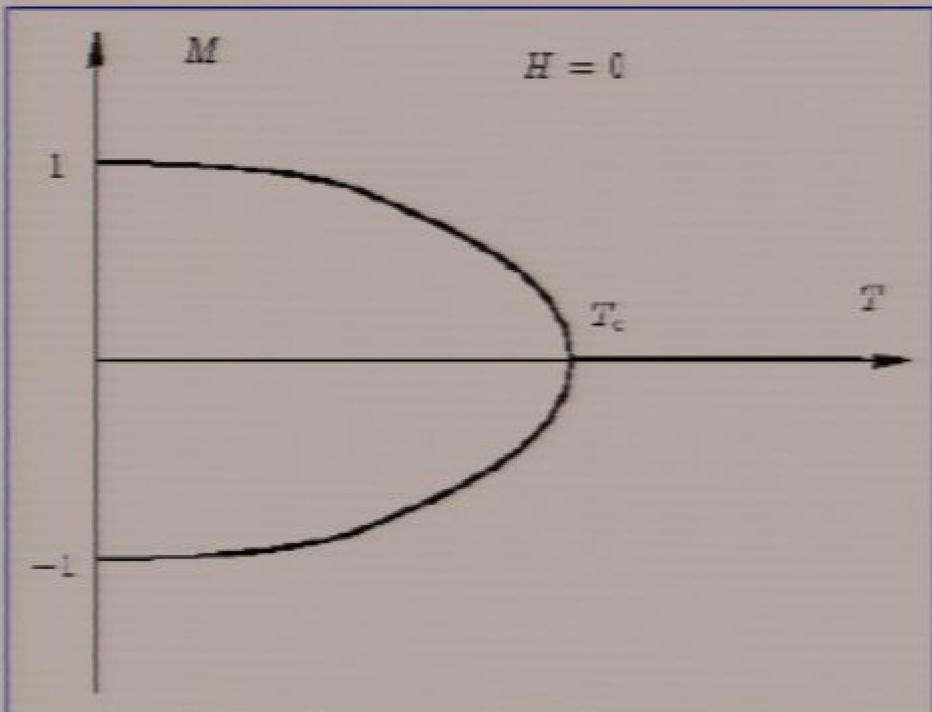
- Probability of configuration with $\mathcal{H} = E_i$

$$p(E_i) = e^{(-E_i/kT)}/z$$

(High energy configuration more likely at high T)

Phase transition in M as function of T

Magnetization: $M = \frac{1}{N}\sum_i s_i$ (# Spin up - # Spin down)/N



- Peierls (1936), gave a non-rigorous proof that spontaneous magnetization must exist for the 2-D Ising model.
- Onsager (1944), gave a complete analytic solution.

Phase transitions in Epidemiology

Kermack and McKendrick, Proc. Roy. Soc. Lond, 1927

Three coupled ordinary differential equations:
"**S**" susceptible, "**I**" infected, "**R**" recovered

Infection rate "β", recovery rate "γ"

1. $\frac{dS}{dt} = S_0 - \beta IS,$

2. $\frac{dI}{dt} = \beta IS - \gamma I,$ ("mean field" eqns)

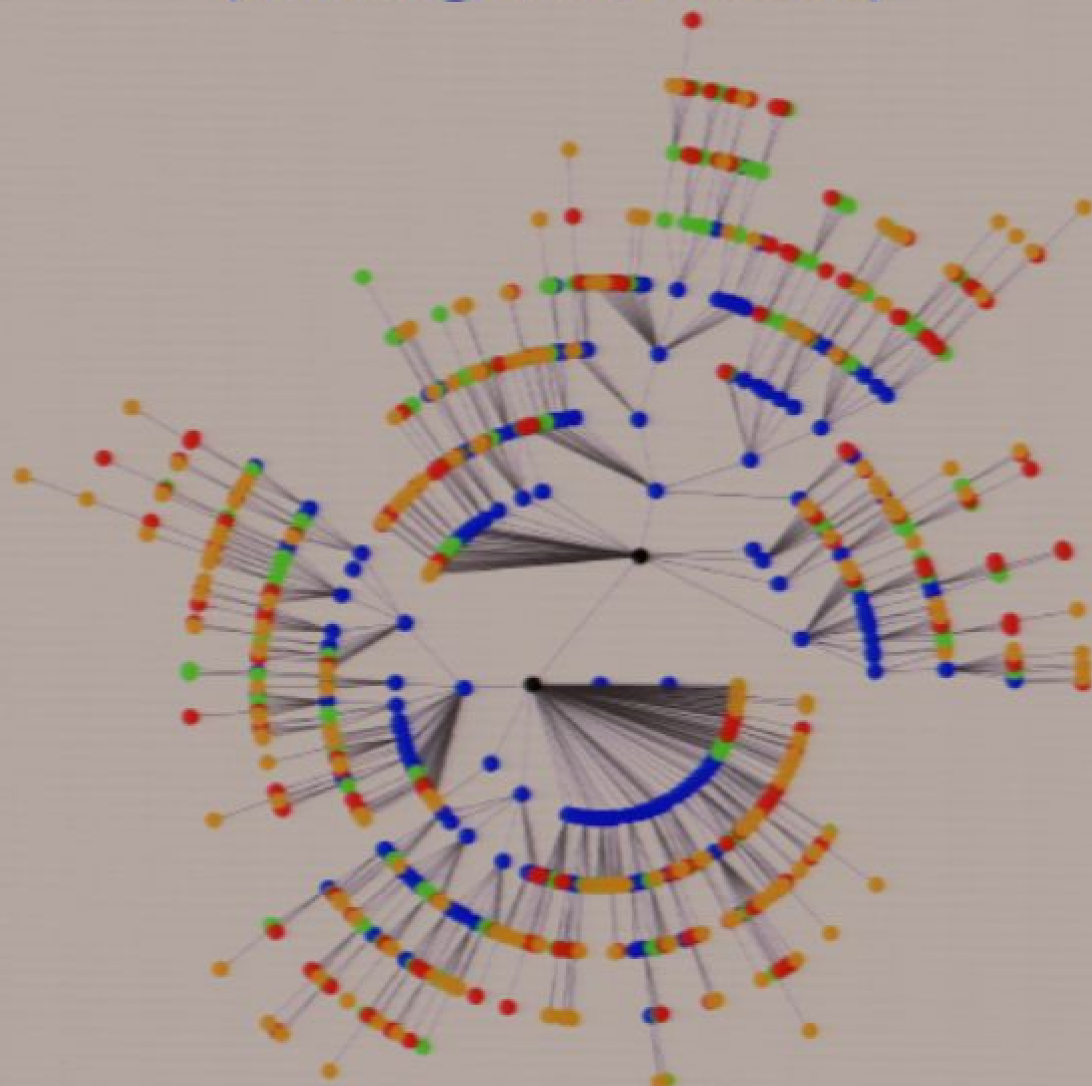
3. $\frac{dR}{dt} = \gamma I.$

Threshold: $\boxed{T_c = \frac{\beta S_0}{\gamma} > 1}$

(Epidemic outbreak if $T_c > 1$, disease dies out if $T_c < 1$.)

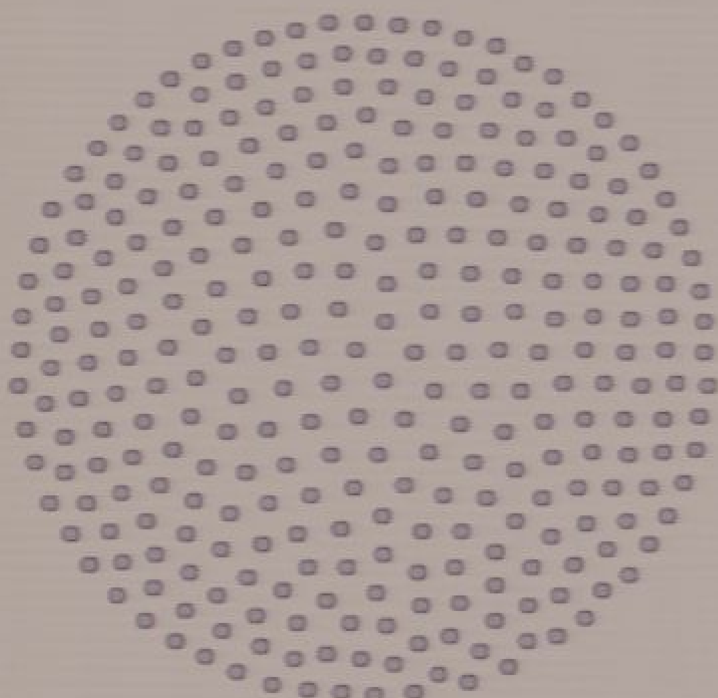
Branching process

Disease spread on a network
(No longer mean field)



When w.h.p. is there a giant cluster on a random graph?

- P. Erdõs and A. Rényi, "On random graphs", *Publ. Math. Debrecen.* 1959.
- P. Erdõs and A. Rényi, "On the evolution of random graphs", *Publ. Math. Inst. Hungar. Acad. Sci.* 1960.
- E. N. Gilbert, "Random graphs", *Annals of Mathematical Statistics*, 1959.

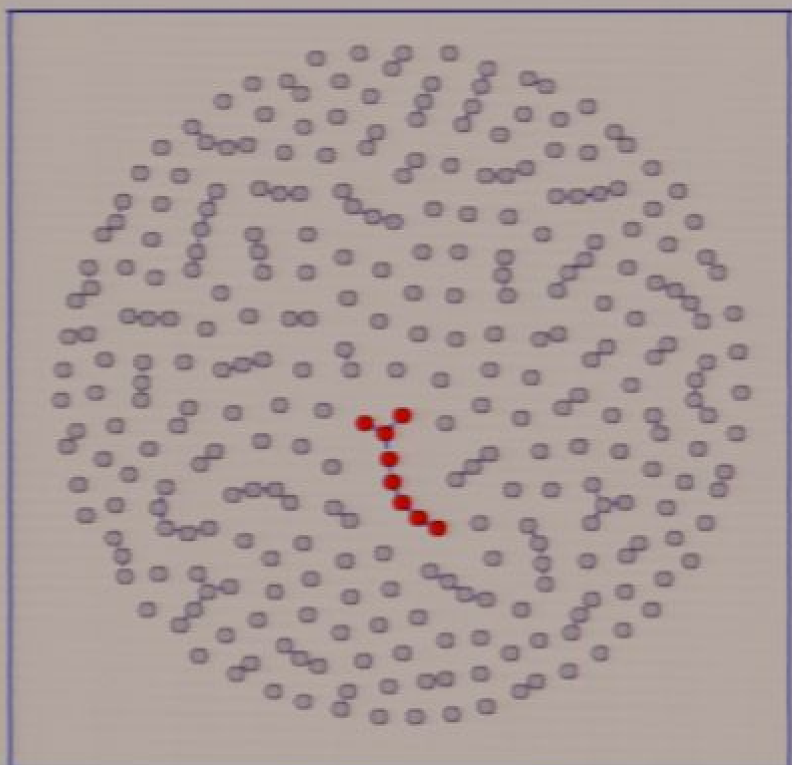


- Start with N isolated vertices.
- Add random edges one-at-a-time.
  $E = N(N-1)/2$ total edges possible.
- After $\mathcal{E}$ edges, probability p of any edge is $p = \mathcal{E}/E = 2\mathcal{E}/N(N-1)$

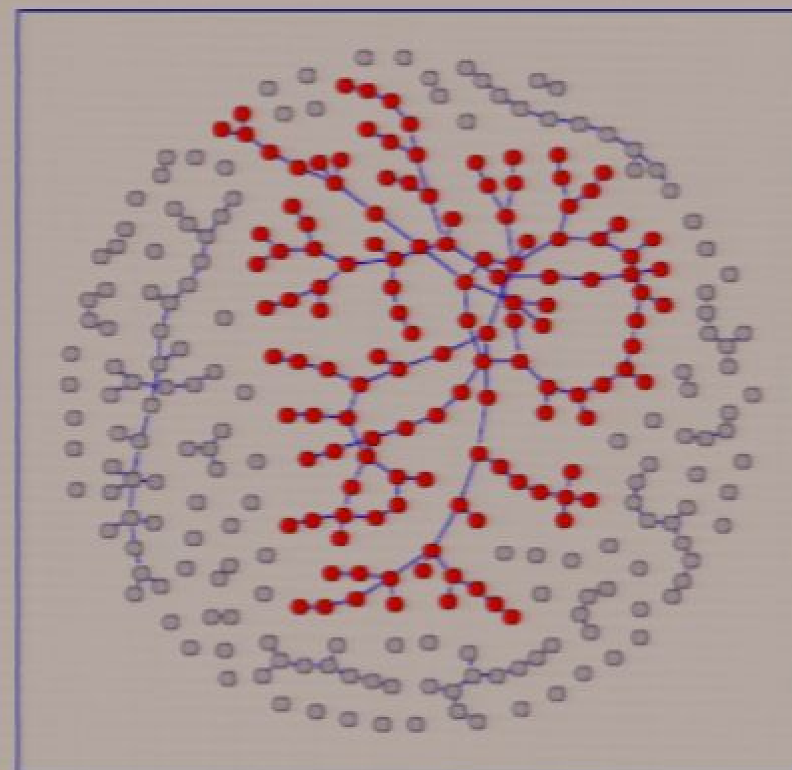
What does the resulting graph look like?

(Typical member of the ensemble)

N=300

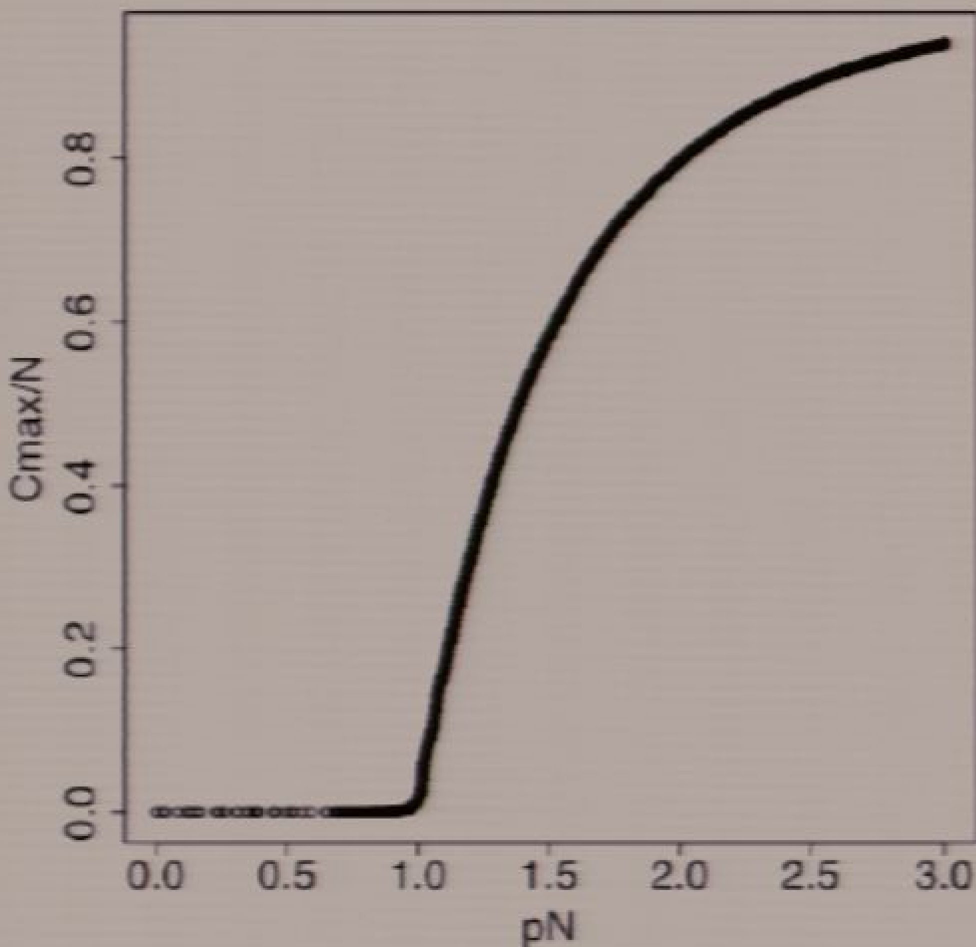


$p = 1/400 = 0.0025$



$p = 1/200 = 0.005$

Emergence of a "giant component"



- $p_c = 1/N$.
- $p < p_c$, $C_{\max} \sim \log(N)$
- $p > p_c$, $C_{\max} \sim A \cdot N$

(Ave node degree $t = pN$ so $t_c = 1$.)

Branching process (Galton-Watson); "tree"-like at $t_c = 1$.

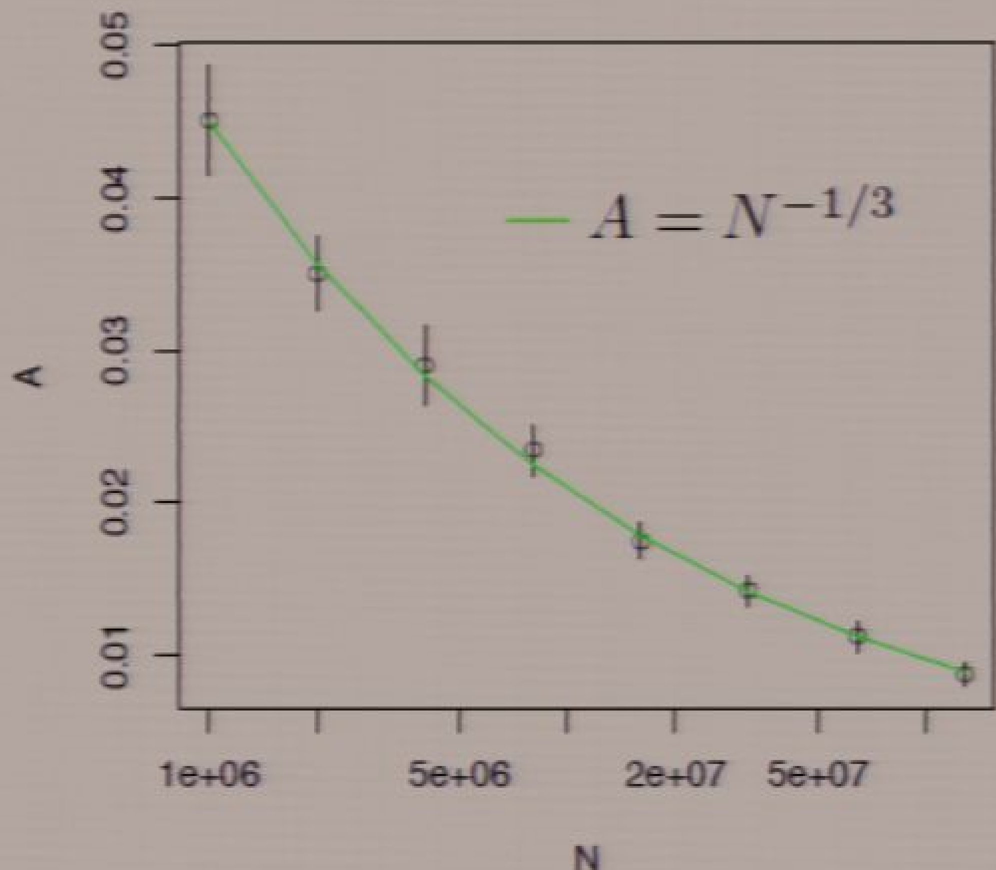
A continuous phase transition

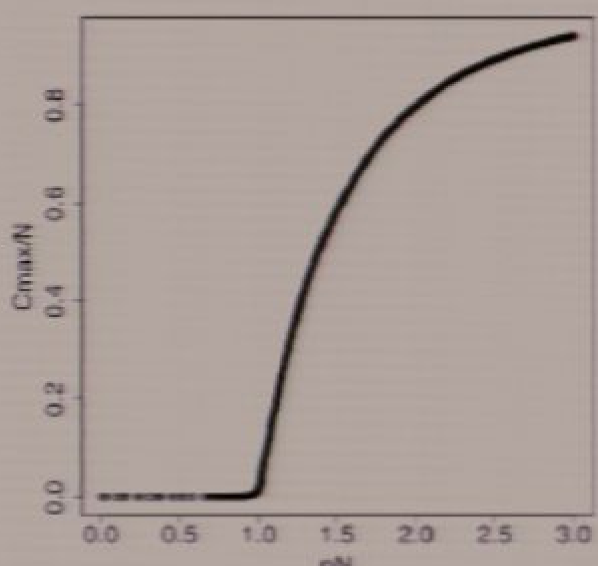
$$C_{\max}(t_c - \epsilon) = N^{2/3}$$

$$C_{\max}(t_c + \epsilon) = N^{2/3} + A$$

$$A = N^{-1/3}$$

$$\lim N \to \infty, A \to 0$$





$$C_{\max}(t_c + \epsilon) - C_{\max}(t_c) \to 0$$

Can any limited perturbation change the phase transition?

(D'Souza, Achlioptas, Spencer, 2009)

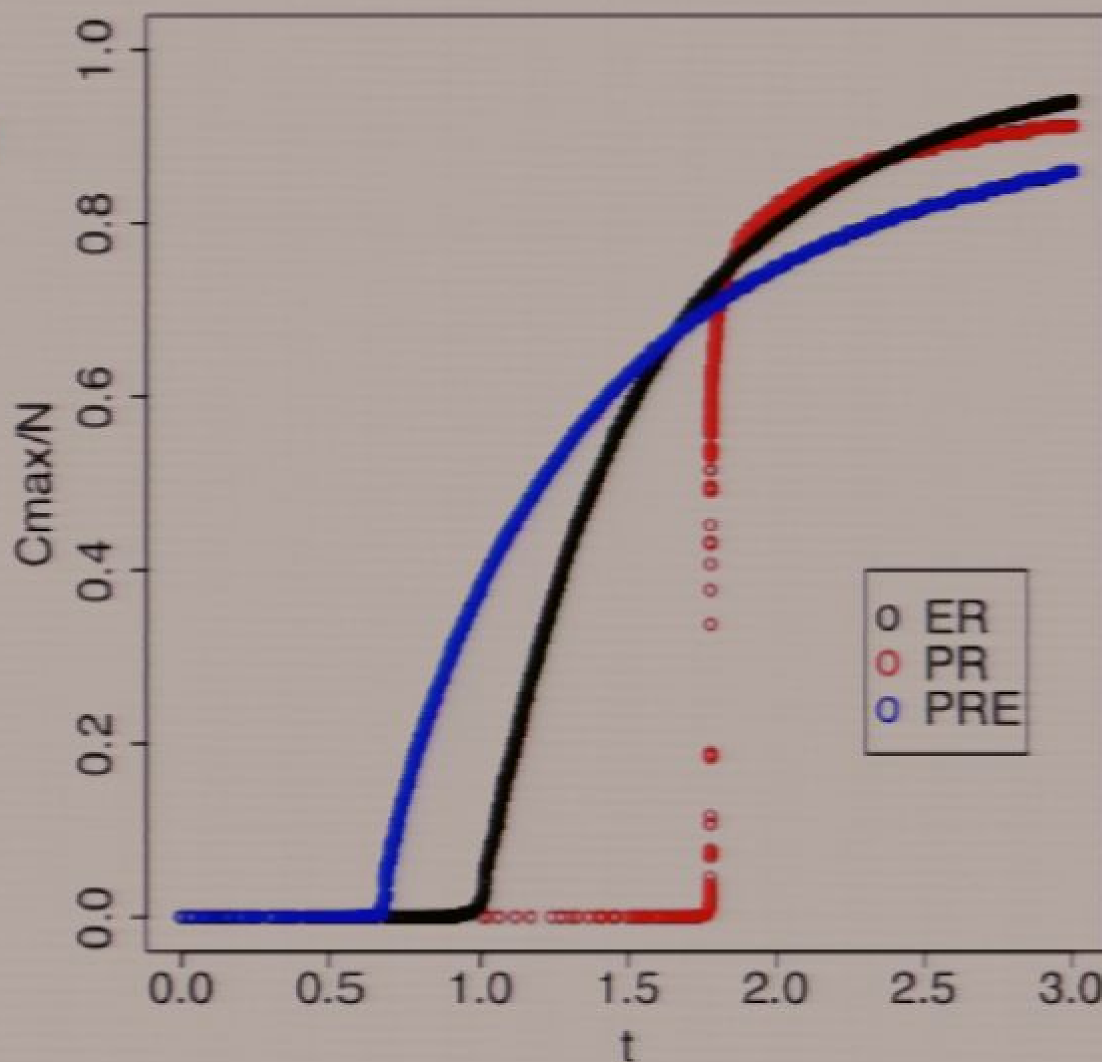
- Possible to **Enhance** or **Delay** the onset?
- The "**Product Rule**"
  - Choose ***two*** edges at random each step.
  - Add only the desirable edge and discard the other.

- The Power of Two Choices
  (Recent theme in optimization/randomized algorithms):
  Azar, Broder, Mitzenmacher, Upfal, Karlin, D'Souza/Krapivsky/Moore

Product Rule

- **Enhance** – similar to **ER** but with earlier onset.

- **Delay** – changes from ***continuous*** to ***discontinuous*** transition!



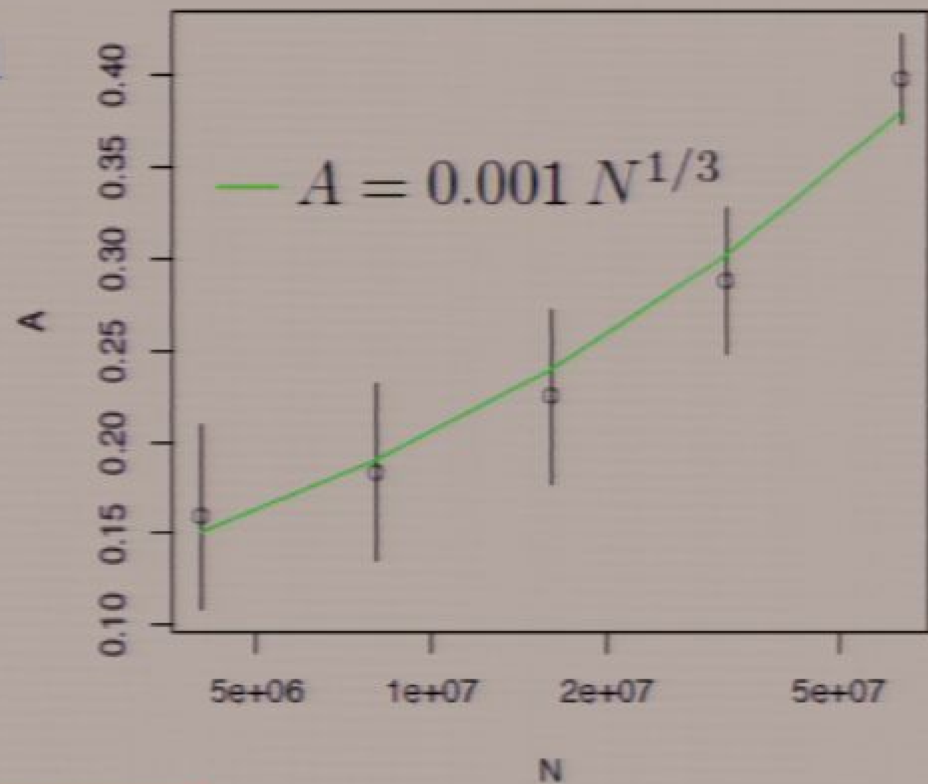
Delayed Product Rule: Discontinuous change

$C_{\max}(t_c - \epsilon) = N^{\gamma}$, with $\gamma < 1$

$C_{\max}(t_c + \epsilon) = A \cdot N$

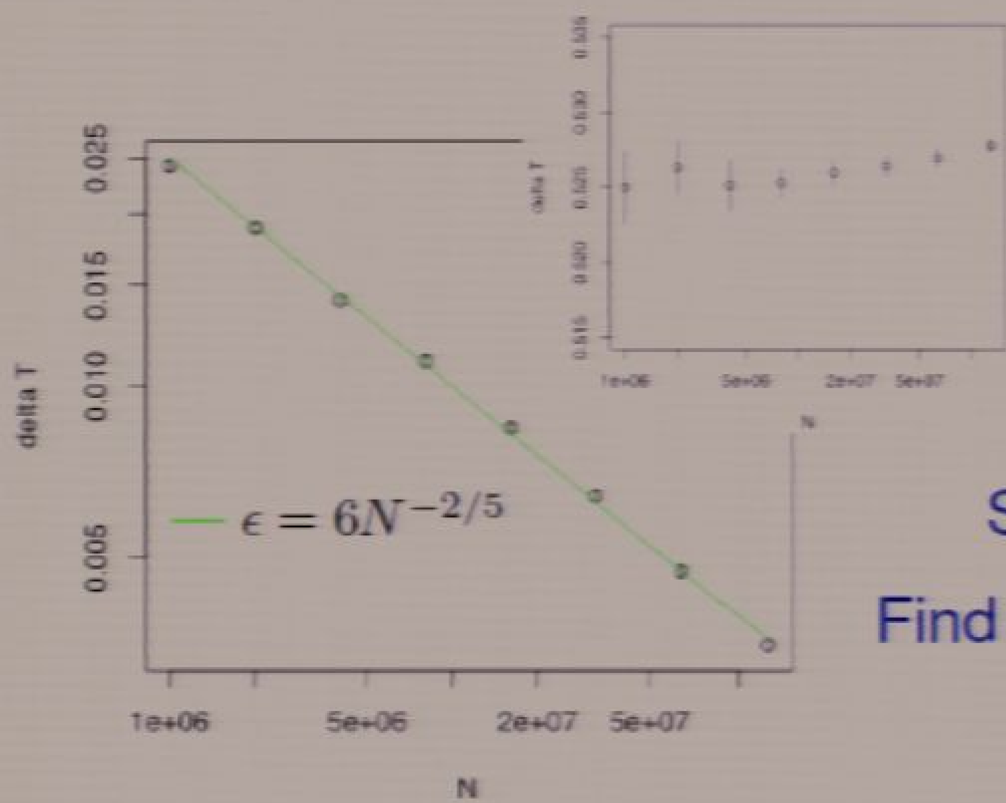
A increases with N!

$\epsilon = 0.0005$



$$C_{\max}(t_c + \epsilon) - C_{\max}(t_c) \to A$$

Can A = 0.5?
(How long to go from $C_{\max}$ sublinear n^{γ} to $\geq 0.5n$?)



(ER)

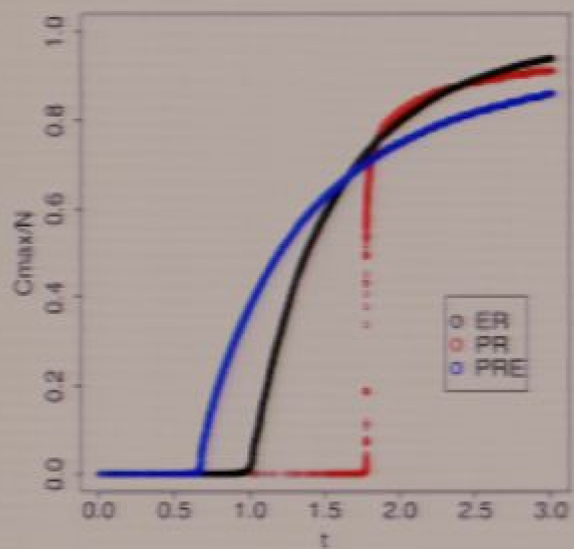
Set $\gamma = 1/2$, $A = 0.5$.

Find $\epsilon \propto N^{-2/5}$, $\lim_{N\to\infty} \epsilon \to 0$.

Jumps instantaneously from $C_{\max} = N^{1/2}$ to $0.5N$!

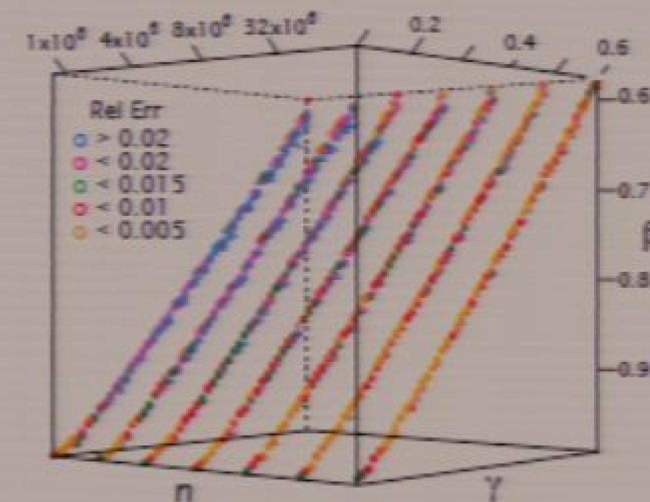
Delayed Product Rule

$C_{\max}$ jumps from sublinear n^{γ} to $\geq 0.5n$ in time n^{β}.



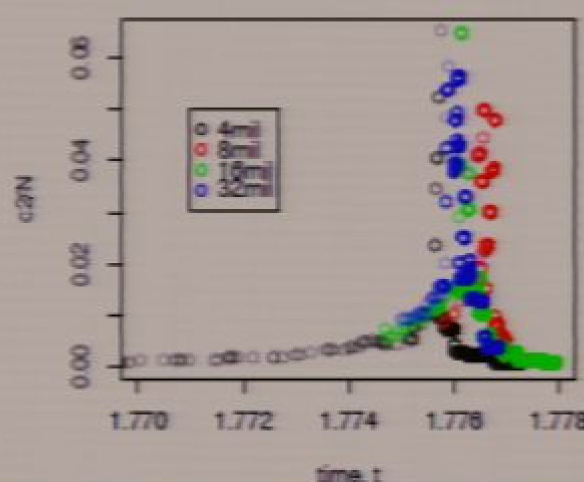
Nontrivial Scaling behaviors

$\gamma + 1.2\beta = 1.3$ for $A \in [0.1, 0.6]$



Signatures of divergences:

Slope $dC_2/dt \to \infty$



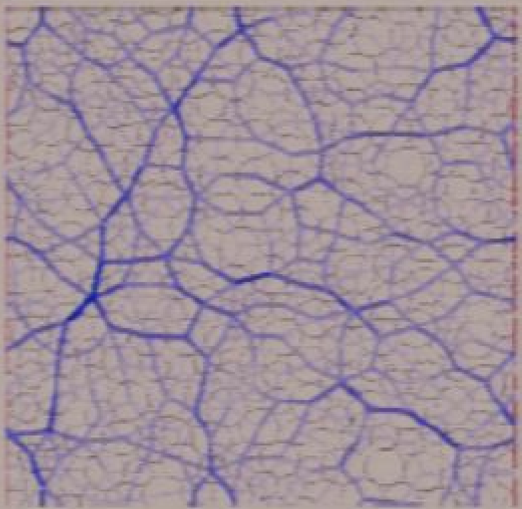
Achlioptas, D'Souza, Spencer, *Science*, **323**, 2009

"Mixed" transitions

Discontinuous transition but diverging correlation lengths

- **k-sat** (constraint satisfaction) for $k \geq 3$, Infinite dimensional
  (Monasson, Zecchina, Kirkpatrick, Selman, Troyansky, *Nature*, 1999)

- **Jamming in models of granular materials**
  Finite dimensional / spatial constraints
  (OHern, Langer, Liu, Nagel, *Phys. Rev. Lett.* 2002)
  (Henkes, Chakraborty, *Phys. Rev. Lett.* 2005)
  (Toninelli, Biroli, Fisher, *Phys. Rev. Lett.* 2006)
  (Schwarz, Liu, Chayes, *Europhys. Lett.* 2006)



- **Spin glasses** glassy systems, slow relaxation time
  (Kirkpatrick and Thirumalai, *Phys. Rev. Lett.* 1987)

k-sat: Constraint satisfaction

- Consider n Boolean variables $\{x_1, x_2, \cdots x_n\}; x_i \in \{+1, -1\}$.
- Consider a "k-clause" : (x_i "or" $not(x_j)$ "or" x_k) $\equiv (x_i \vee \overline{x}_j \vee x_k)$
- And finally, a collection of m such clauses:

$$(x_i \vee \overline{x}_l \vee x_j) \wedge \ (\overline{x}_j \vee x_h \vee x_y) \wedge \ (x_j \vee \overline{x}_k \vee \overline{x}_y) \cdots$$

- **Is there any assignment of x_i's that will make the clause TRUE? (A Yes/No question)**

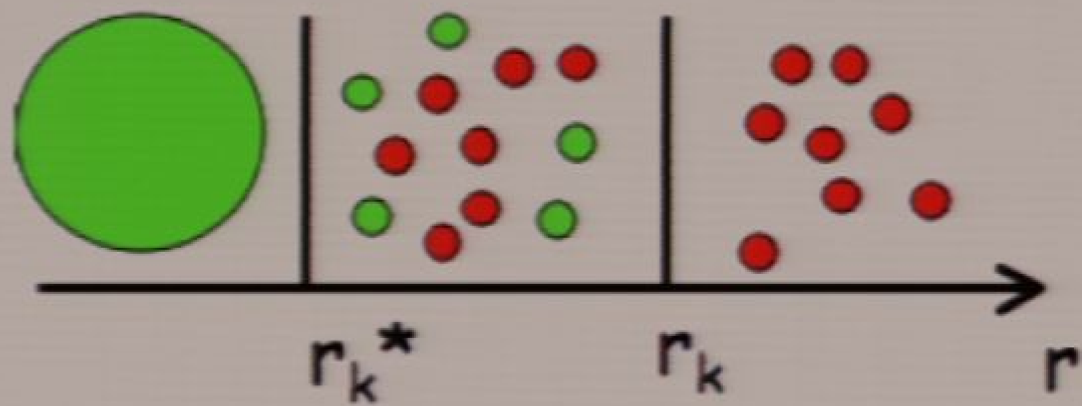
Applications: – Scheduling / Manufacturing / Package delivery

– Transport on networks / Flows

– Optimization (Traveling salesman/Hamiltonian paths)

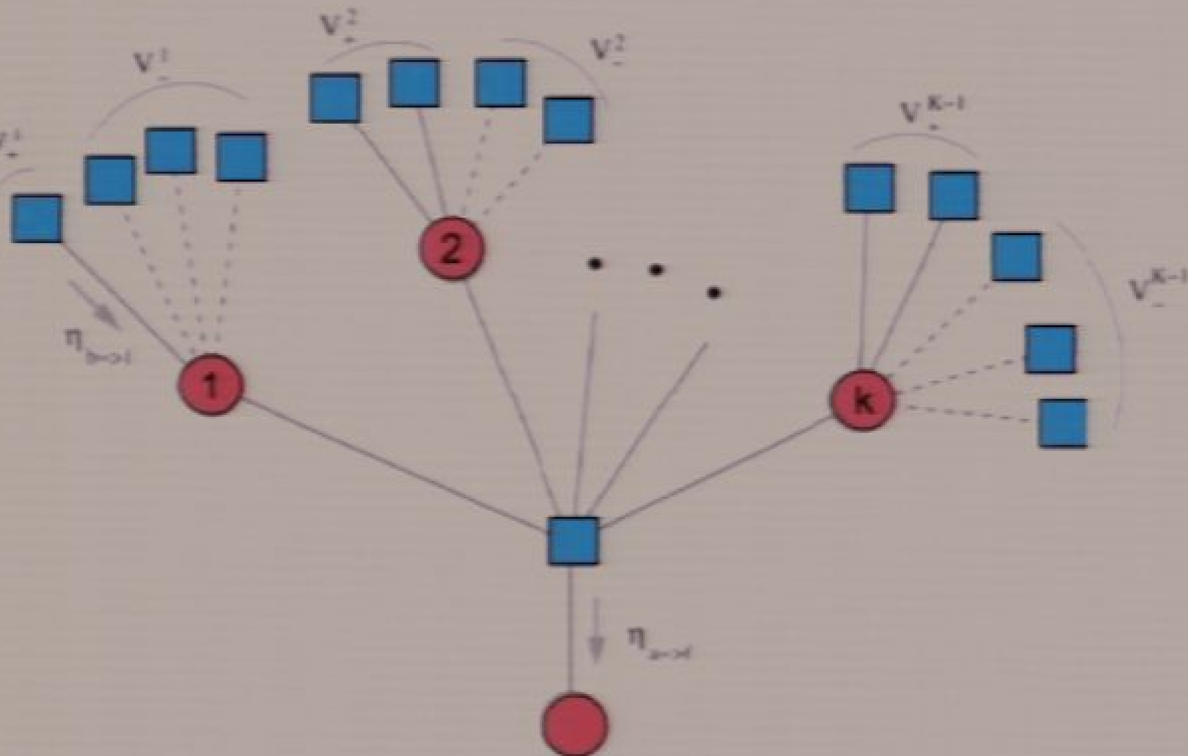
– An "NP complete" problem, for $k \geq 3$.

Phase transition as function of $r = m/n$



- For $m/n < r_k^*$, few clauses, many variables – easy to find a "YES".
- For $m/n > r_k$, many clauses, few variables – easy to find a "NO".
- For $r_k^* < m/n < r_k$, "YES" solns exist, **but difficult to find!**

The "factor graph" / network of a k-sat instance: (Insights from statistical physics)

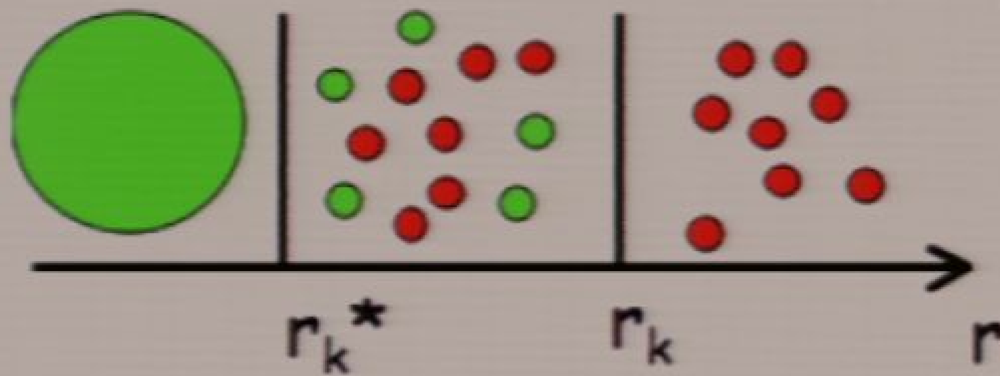


- Each k-clause is a box
- Each variable is a circle

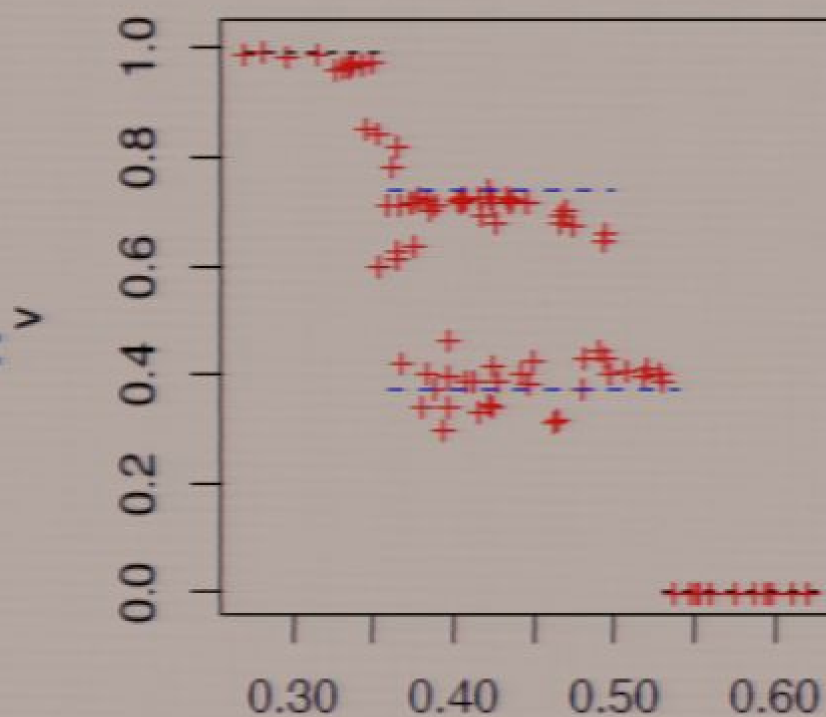
- Fastest known algorithm: Mézard, Parisi, Zecchina. *Science*, 2002. (Replica symmetry breaking / spin glasses / locality of constraint chains). Run time, $O(n \log n)$.
- Standard approach (Davis-Putnam backtracking). Run time, $O(2^n)$.

k-sat and JAMMING phase portraits

k-sat:



Jamming:

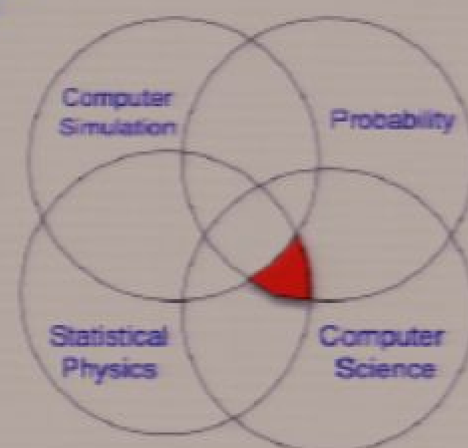


Mixed transitions: geometry, disorder, computation

- Critical slowing down and computational complexity?
- *Nature* 1999: 2-sat continuous $\in$ P, 3-sat discontinuous $\in$ NP
- Hard instances, applications to Cryptography?

- Aspen Center for Physics: "**Complexity, Disorder, and Algorithms**"
  Organizers: S. Coppersmith, A. Middleton, J. Machta, C. Moore
  May 25 - June 22, 2008.

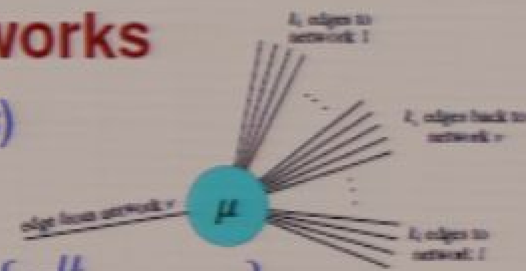
- American Institute for Mathematics: "**Phase Transitions**" Aug 21-25, 2006
  Organizers: P. Diaconis, D. Fisher, C. Moore, C. Radi-



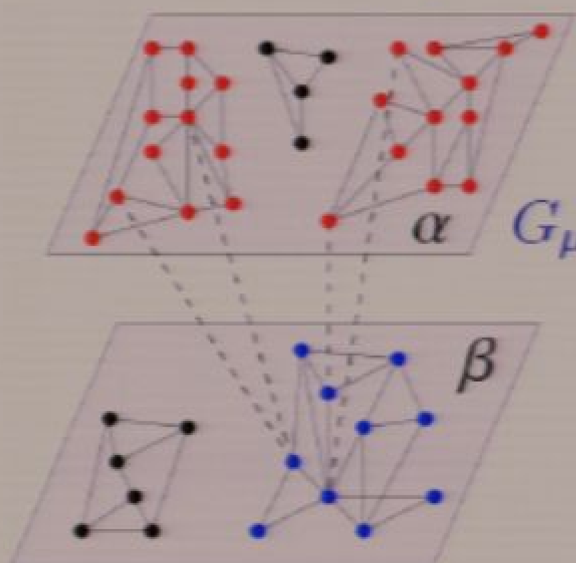
- Mathematical Sciences Research Institute (MSRI)
  "**Probability, Algorithms and Statistical Physics**"
  Organizers: Y. Peres, A. Sinclair, D. Aldous,
  C. Kenyon, H. Kesten, J. Kleinberg, F. Martinelli,
  A. Sokal, P. Winkler, U. Zwick
  Jan 3 - May 15, 2005.

Percolation on interacting networks

(D'Souza and Leicht, *in progress*)

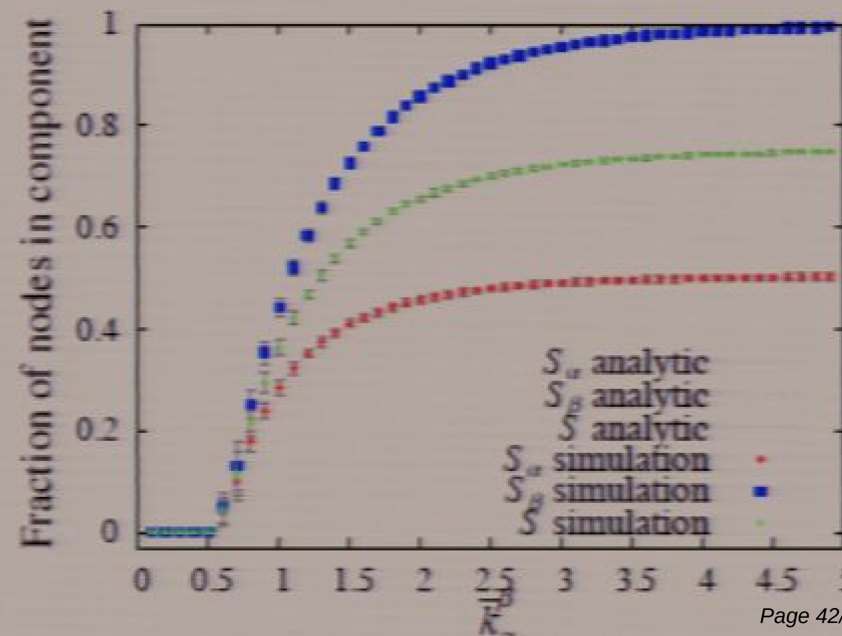


Given connection probabilities for each network μ: $\{p^{\mu}_{k_1 k_2 \cdots k_l}\}$

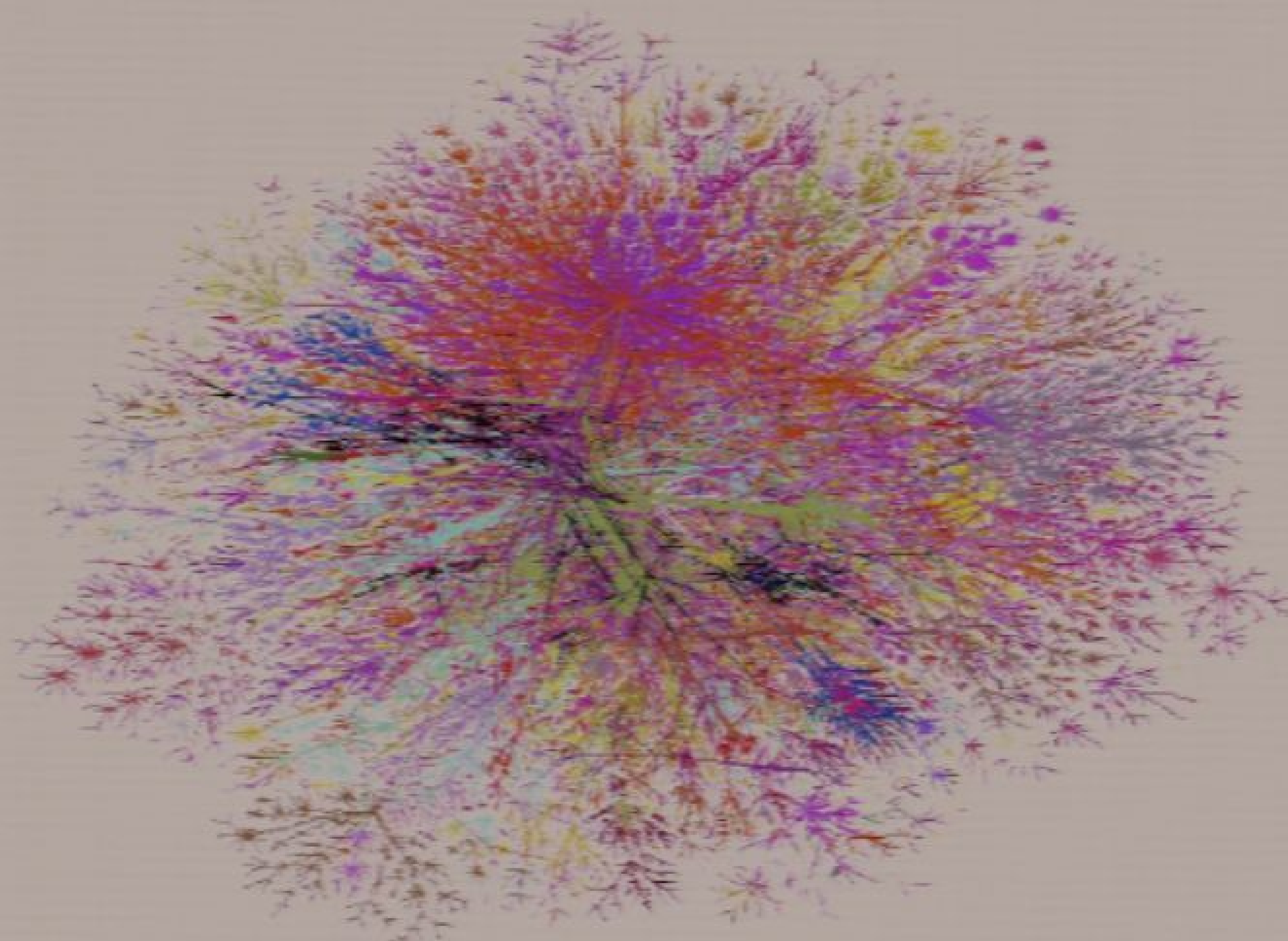


$$G_{\mu}(x_1, \ldots, x_l) = \sum_{k_1, \cdots, k_l = 0}^{\infty} p^{\mu}_{k_1 \cdots k_l} x_1^{k_1} \cdots x_l^{k_l}$$

$$\langle s_{\alpha} \rangle_{\alpha} = 1 + \frac{\overline{k}^{\alpha}_{\alpha} + \overline{k}^{\alpha}_{\beta} \overline{k}^{\beta}_{\alpha} - \overline{k}^{\alpha}_{\alpha} \overline{k}^{\beta}_{\beta}}{(1 - \overline{k}^{\alpha}_{\alpha})(1 - \overline{k}^{\beta}_{\beta}) - \overline{k}^{\alpha}_{\beta} \overline{k}^{\beta}_{\alpha}}$$



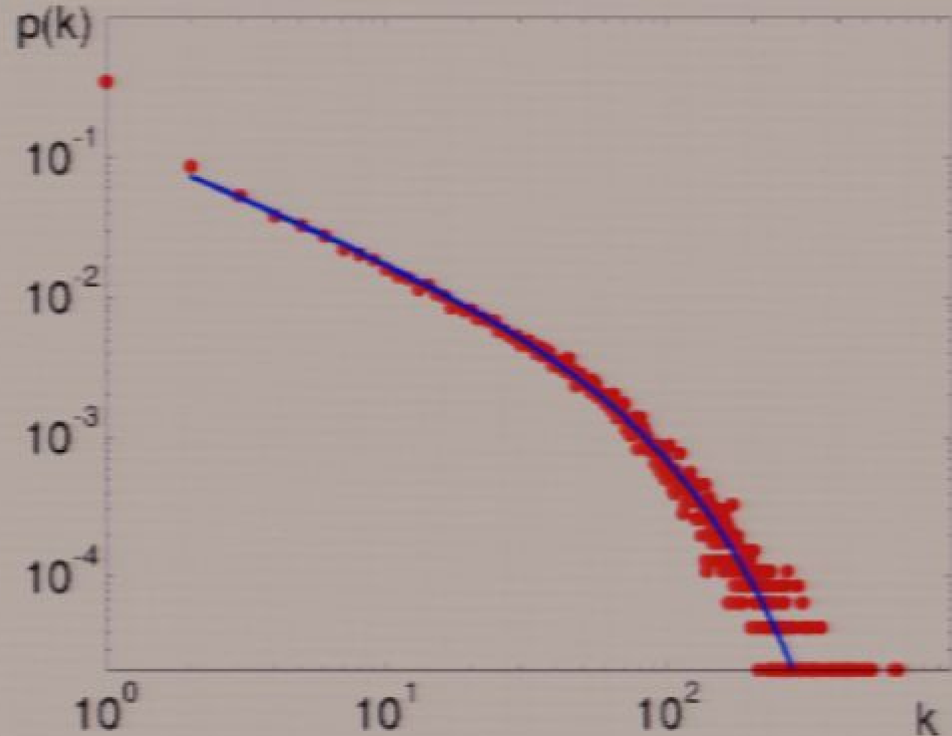
Back to “real-world” networks



Extremely broad range of node degree observed: from biological, to technological, to social.

The "Who-is-Who" network in Budapest

(Analysis by Balázs Szendröi and Gábor Csányi)



Bayesian curve fitting $\rightarrow p(k) = ck^{-\gamma}e^{-\alpha k}$

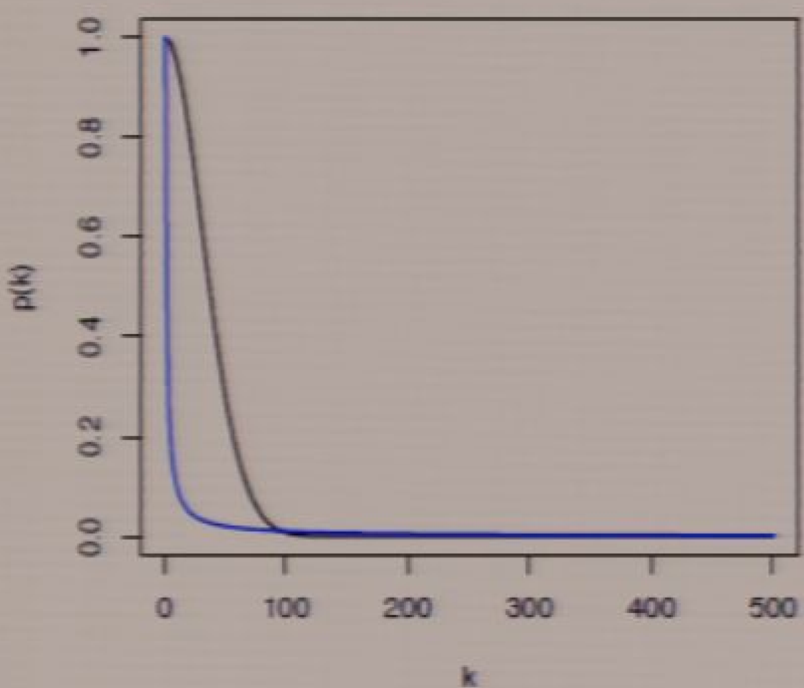
Power law with exponential tail

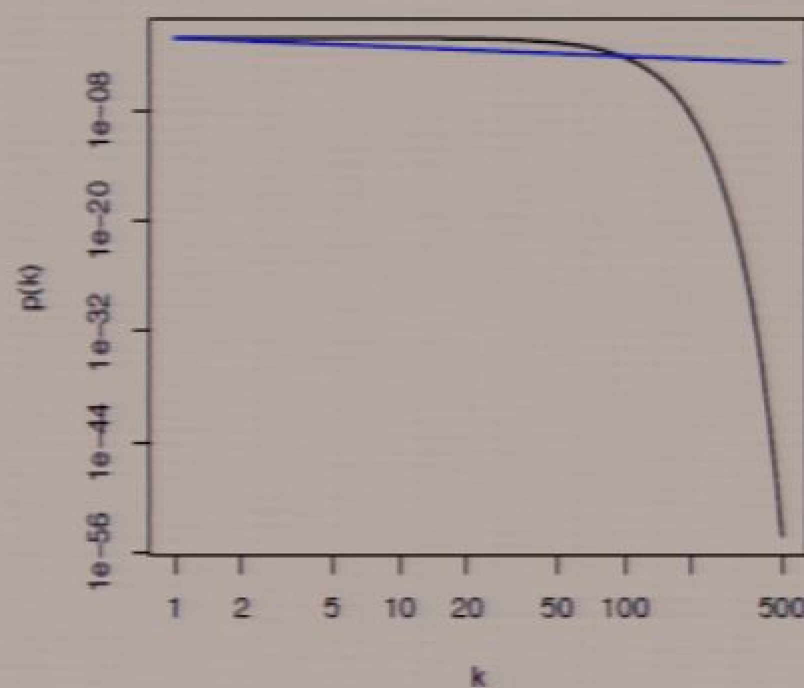
Ubiquitous empirical measurements:

System with: $p(x) \sim x^{-B} \exp(-x/C)$	B	C
Full protein-interaction map of *Drosophila*	1.20	0.038
High-confidence protein-interaction map of *Drosophila*	1.26	0.27
Gene-flow/hydridization network of plants as function of spatial distance	0.75	10^5 m
Earthquake magnitude	1.35 - 1.7	$\sim 10^{21}$ Nm
Avalanche size of ferromagnetic materials	1.2 - 1.4	$L^{1.4}$
ArXiv co-author network	1.3	53
MEDLINE co-author network	2.1	$\sim$ 5800
PNAS paper citation network	0.49	4.21

Power Laws versus Bell Curves: "Heavy tails"

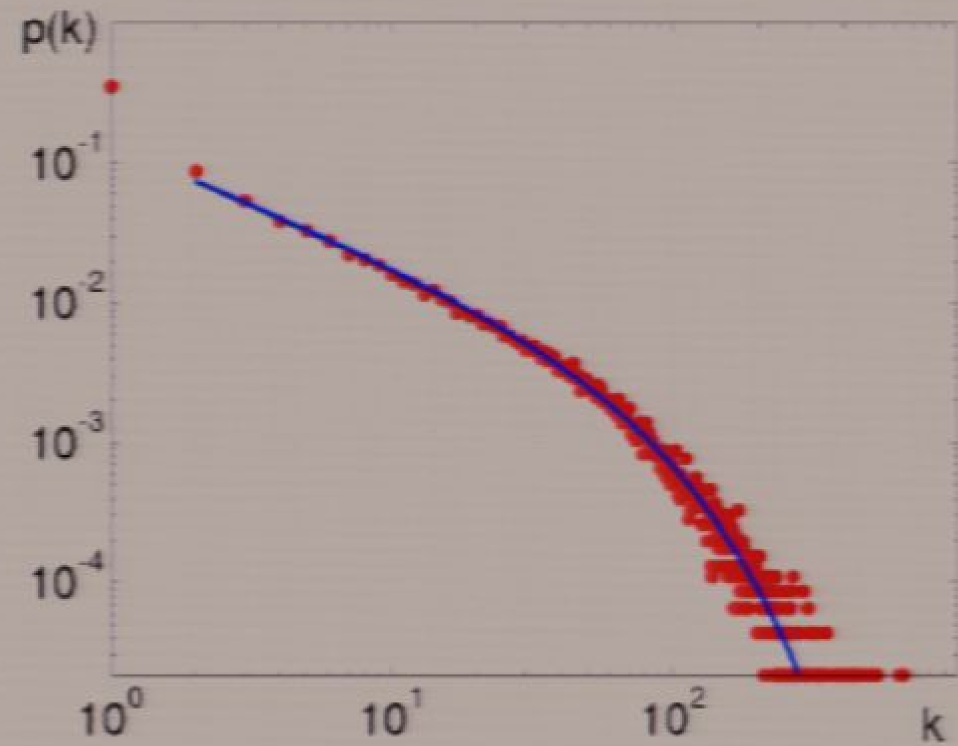
- Power law distribution: $p_k \sim k^{-\gamma}$.
- Gaussian distribution: $p_k \sim \exp(-k^2/2\sigma^2)$.





The "Who-is-Who" network in Budapest

(Analysis by Balázs Szendrői and Gábor Csányi)



Bayesian curve fitting $\rightarrow p(k) = ck^{-\gamma}e^{-\alpha k}$

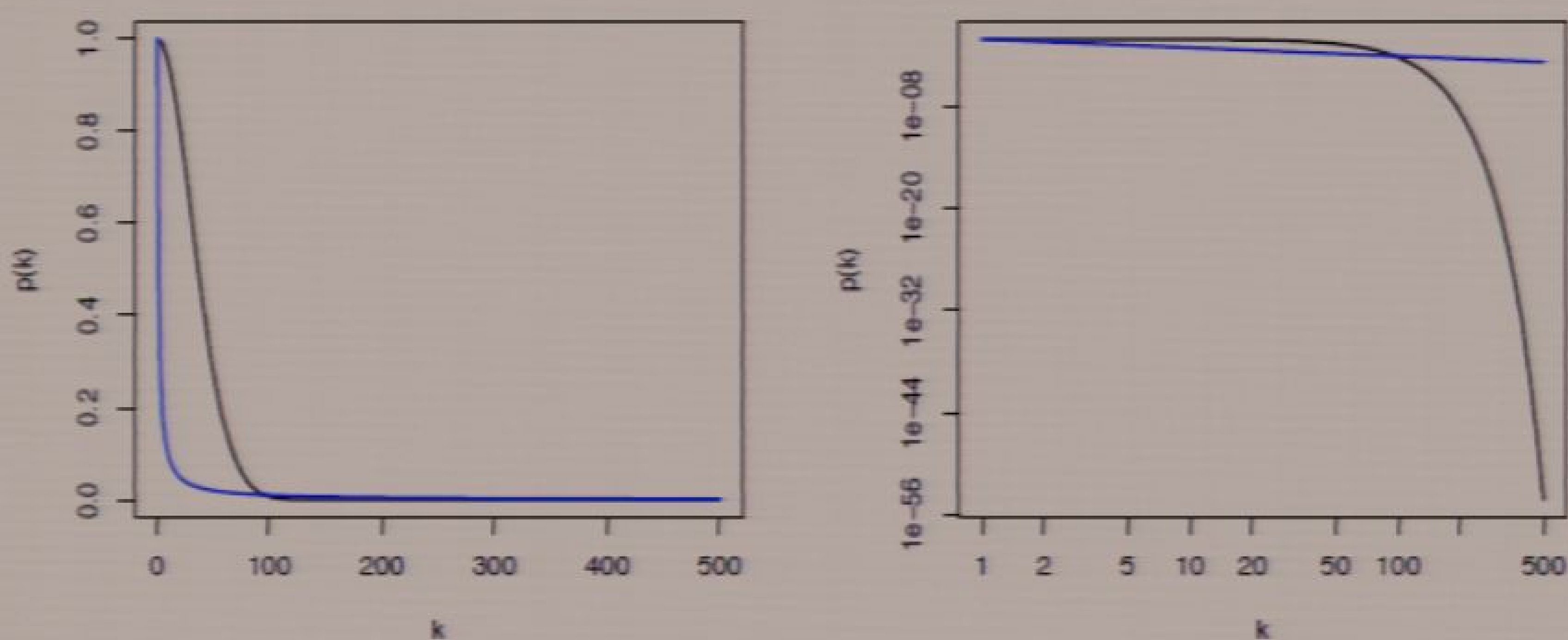
Power law with exponential tail

Ubiquitous empirical measurements:

System with: $p(x) \sim x^{-B} \exp(-x/C)$	B	C
Full protein-interaction map of *Drosophila*	1.20	0.038
High-confidence protein-interaction map of *Drosophila*	1.26	0.27
Gene-flow/hydridization network of plants as function of spatial distance	0.75	10^5 m
Earthquake magnitude	1.35 - 1.7	$\sim 10^{21}$ Nm
Avalanche size of ferromagnetic materials	1.2 - 1.4	$L^{1.4}$
ArXiv co-author network	1.3	53
MEDLINE co-author network	2.1	~ 5800
PNAS paper citation network	0.49	4.21

Power Laws versus Bell Curves: "Heavy tails"

- Power law distribution: $p_k \sim k^{-\gamma}$.
- Gaussian distribution: $p_k \sim \exp(-k^2/2\sigma^2)$.



Degree distribution and Network Growth

- **Heterogeneity** in real networks.
- Concentrated, **Poisson Distribution in Erdös-Rényi**:
  - Probability to connect to k nodes is p^k.
  - Probability to be disconnected from remaining $(n-k)$ is $(1-p)^{(n-k)}$.
  - Probability for a vertex to have degree k follows a binomial distribution:

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}.$$

- **Seek alternate mechanisms...**

Known Mechanisms for Power Laws

- Phase transitions (singularities)
- Random multiplicative processes (fragmentation)
- Combination of exponentials (e.g. word frequencies)
- **Preferential attachment / Proportional attachment**
  (Polya 1923, Yule 1925, Zipf 1949, Simon 1955, Price 1976, Barabási and Albert 1999)

Attractiveness is proportional to size:

$$\boxed{\frac{dP(s)}{dt} \propto s}$$

- Add in **saturation** [Amaral 2000, Börner 2004], **get power laws with exponential decay** .

Origins of preferential attachment

- 1923 — **Polya** , urn models.
- 1925 — **Yule** , explain genetic diversity.
- 1949 — **Zipf** , distribution of city sizes (*1/f*).
- 1955 — **Simon** , distribution of wealth in economies.
  (**“The rich get richer”**).

- [Interesting note, in sociology this is referred to as the ***Matthew effect*** after the biblical edict, “For to every one that hath shall be given ... ” (Matthew 25:29)]

Preferential attachment in networks

- D. J. de S. Price, "Networks of scientific papers" *Science*, 1965.
  First observation of power laws in a network context.
  Studied paper co-citation network.

- D. J. de S. Price, "A general theory of bibliometric and other cumulative advantage processes" *J. Amer. Soc. Info. Sci.*, 1976.

The rate at which a paper gains citations is proportional to the number it already has. (Probability to learn of a paper proportional to number of references it currently has).

- A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks" *Science*, 1999.

(Citations: ~ 1000 in 2006
~ 2000 in 2007
~ 3400 in April, 2008
~ 4300 in Feb, 2009)

Origins of preferential attachment

- 1923 — **Polya** , urn models.
- 1925 — **Yule** , explain genetic diversity.
- 1949 — **Zipf** , distribution of city sizes (*1/f*).
- 1955 — **Simon** , distribution of wealth in economies.
  (**“The rich get richer”**).

- [Interesting note, in sociology this is referred to as the ***Matthew effect*** after the biblical edict, “For to every one that hath shall be given ... ” (Matthew 25:29)]

Preferential attachment in networks

- D. J. de S. Price, "Networks of scientific papers" *Science*, 1965.
  First observation of power laws in a network context.
  Studied paper co-citation network.

- D. J. de S. Price, "A general theory of bibliometric and other cumulative advantage processes" *J. Amer. Soc. Info. Sci.*, 1976.

The rate at which a paper gains citations is proportional to the number it already has. (Probability to learn of a paper proportional to number of references it currently has).

- A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks" *Science*, 1999.

(Citations: $\sim$ 1000 in 2006
$\sim$ 2000 in 2007
$\sim$ 3400 in April, 2008
$\sim$ 4300 in Feb, 2009)

An alternate view, Mandelbrot, 1953: optimization

(Information theory of the statistical structure of language)

- **Goal:** Optimize information conveyed for unit transmission cost
- Consider an alphabet of d characters, with n distinct words
- Order all possible words by length (A,B,C,....AA,BB,CC....)
- "Cost" of j-th word, $C_j \sim \log_d j$
- Ave information per word: $H = -\sum p_j \log p_j$
- Ave cost per word: $C = \sum p_j C_j$
- Minimize: $\frac{d}{dp_j}\left(\frac{C}{H}\right) \implies p_j \sim j^{-\alpha}$

Optimization versus Preferential Attachment origin of power laws

Mandelbrot and Simon's heated public exchange

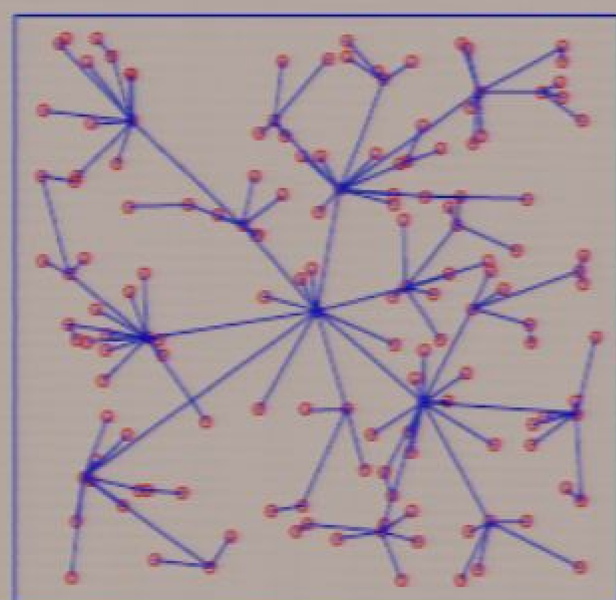
- A series of six letters between 1959-61 in *Information and Control*.

Optimization on hold for many years, but recently resurfaced:

- Calson and Doyle, "HOT" (PRE 1999, PRL 2000, PNAS 2002).
- Fabrikant, Koutsoupias, and Papadimitriou (ICALP 2002).
- Valverde, Ferrer Cancho, and Solé (Europhys. Lett. 2002).

FKP (Fabrikant, Koutsoupias, and Papadimitriou, 2002)

- Nodes arriving sequentially at random in a unit square.
- Upon arrival, each node connects to an already existing node that minimizes "cost": $\alpha d_{ij} + h_j$



Tempered Preferential Attachment

[Berger, Borgs, Chayes, D'Souza, Kleinberg, *ICALP* 2004.]

[Berger, Borgs, Chayes, D'Souza, Kleinberg, *CPC*, 2005.]

[D'Souza, Borgs, Chayes, Berger, Kleinberg, *Proc Natn Acad Sci*, 2007.]

- **Optimization**
  Like FKP, start with linear tradeoffs, but consider a scale-free metric. (Plus will result in local model.) Gives rise to:

$\rightarrow$ **PA**

$\rightarrow$ **Saturation**

$\rightarrow$ **Viability**

(Not all children have equal fertility, not all spin-offs equally fit, etc).

Competition-Induced Preferential Attachment

Consider points arriving sequentially, uniformly at random along the unit line:

Each incoming node, t, attaches to an existing node j (where $j < t$), which minimizes the function:

$$F_{tj} = \min_j [\alpha_{tj} d_{tj} + h_j]$$

Where $\alpha_{tj} = \alpha \rho_{tj} = \alpha n_{tj} / d_{tj}.$

The "cost" becomes: $\boxed{F_{tj} = \min_j [\alpha n_{tj} + h_j]}$

$$\boxed{F_{tj} = \min_j \left[\alpha n_{tj} + h_j\right]}$$

- $\alpha_{tj} = \alpha \rho_{tj}$ geometric cost proportional to local density
- Reduces to n_{tj} — number of points in the interval between t and j
- "Transit domains" — captures realistic aspects of Internet costs (i.e. AS/ISP-transit requires BGP and peering).

"Fertility"/Viability

Node 1 becomes "fertile" at time $t = 3$.

- Define $A = \lceil 1/\alpha \rceil$
- A node must have $A - 1$ "infertile" children before giving birth to a "fertile" child.

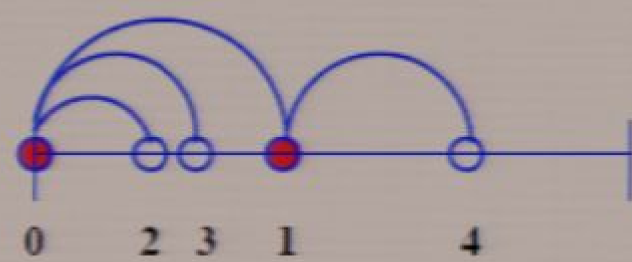
From line to tree

Integrating out the dependence on interval length from the conditional probability:

$$Pr\left[x_{t+1} \in I_k \,|\pi(t)\right] = \int Pr\left[x_{t+1} \in I_k \,|\pi(t), \vec{s}(t)\right] dP(\vec{s}(t))$$
$$= \int s_k(t) dP(\vec{s}(t)) = \frac{1}{t+1},$$

i.e., **The probability to land in the k-th interval is uniform over all intervals.**

Preferential attachment with a cutoff



Let $d_j(t)$ equal the degree of fertile node j at time t.

The number of intervals contributing to j's fertility is $\min(d_j(t), A)$.

Probability node $(t+1)$ attaches to node j is:

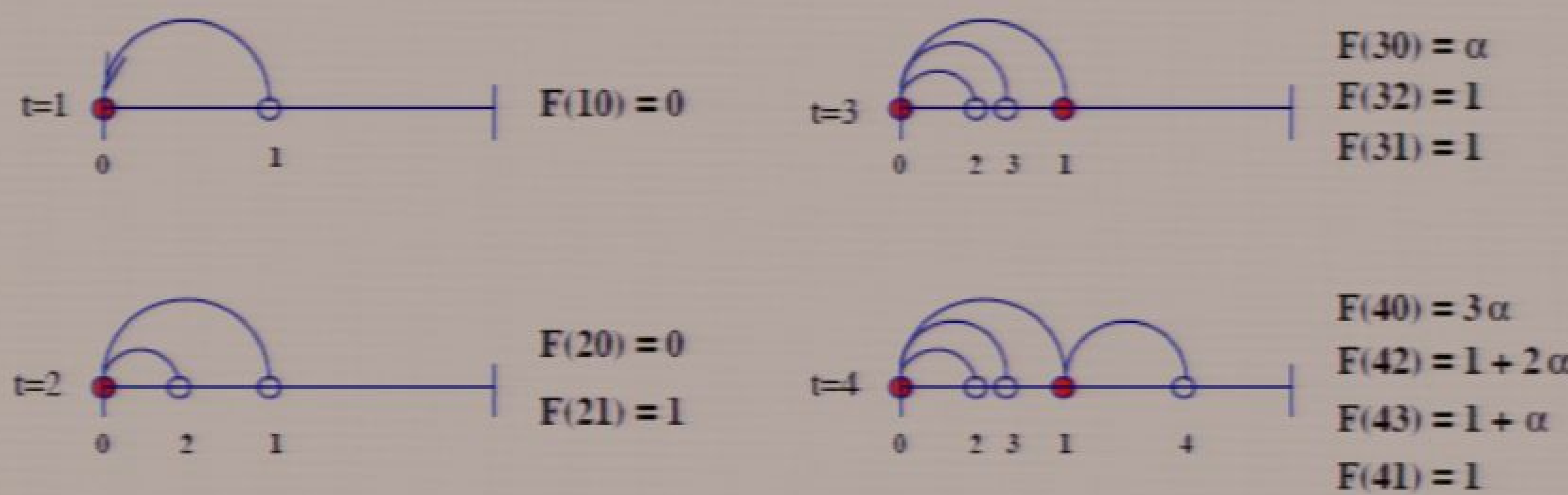
$$Pr(t+1 \rightarrow j) = \min(d_j(t), A)/(t+1).$$

Standard PA: $Pr(t+1 \rightarrow j) = d_j(t)/(t+1)$.

The process on the line (for $1/3 < \alpha < 1/2$)

"Border Toll Optimization Problem" (BTOP)

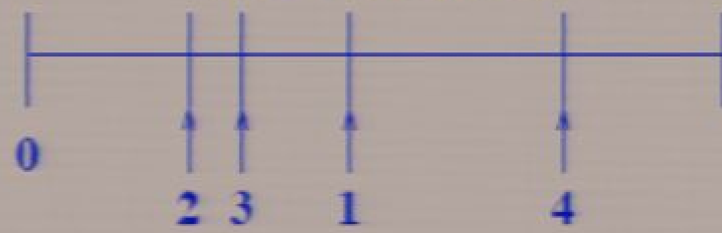
$$F_{tj} = \min_j [\alpha n_{tj} + h_j]$$



(A **local** model – connect either to closest node, or its parent.)

Competition-Induced Preferential Attachment

Consider points arriving sequentially, uniformly at random along the unit line:



Each incoming node, t, attaches to an existing node j (where $j < t$), which minimizes the function:

$$F_{tj} = \min_j \left[\alpha_{tj} d_{tj} + h_j\right]$$

Where $\alpha_{tj} = \alpha\rho_{tj} = \alpha n_{tj}/d_{tj}.$

The "cost" becomes: $$\boxed{F_{tj} = \min_j \left[\alpha n_{tj} + h_j\right]}$$

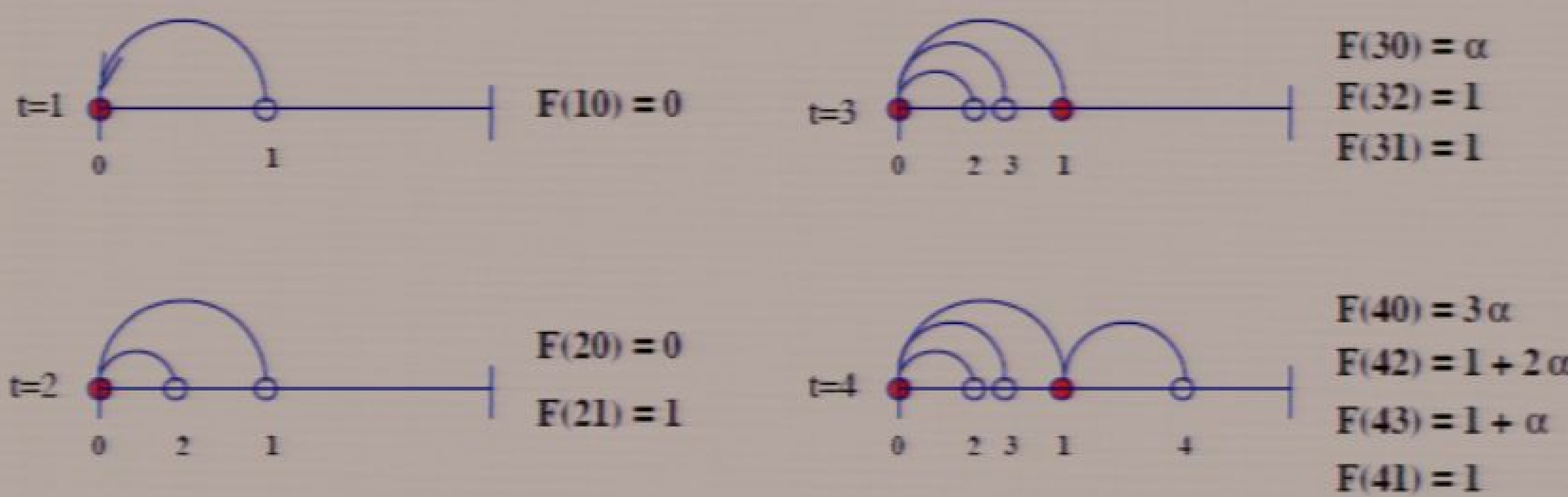
$$F_{tj} = \min_j \left[\alpha n_{tj} + h_j\right]$$

- $\alpha_{tj} = \alpha \rho_{tj}$ geometric cost proportional to local density
- Reduces to n_{tj} — number of points in the interval between t and j
- "Transit domains" — captures realistic aspects of Internet costs (i.e. AS/ISP-transit requires BGP and peering).

The process on the line (for $1/3 < \alpha < 1/2$)

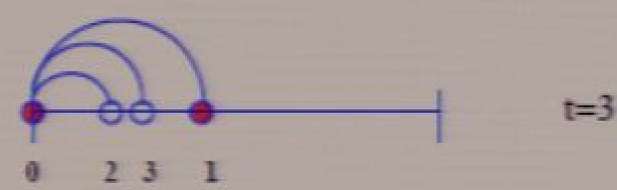
"Border Toll Optimization Problem" (BTOP)

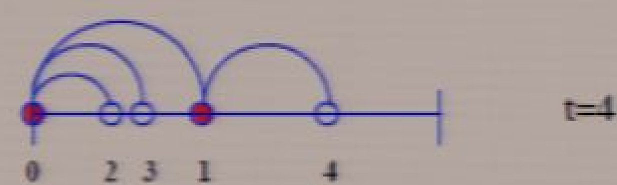
$$F_{tj} = \min_j \left[\alpha n_{tj} + h_j\right]$$



(A **local** model – connect either to closest node, or its parent.)

“Fertility”/Viability





Node 1 becomes “fertile” at time $t = 3$.

- Define $A = \lceil 1/\alpha \rceil$
- A node must have $A - 1$ “infertile” children before giving birth to a “fertile” child.

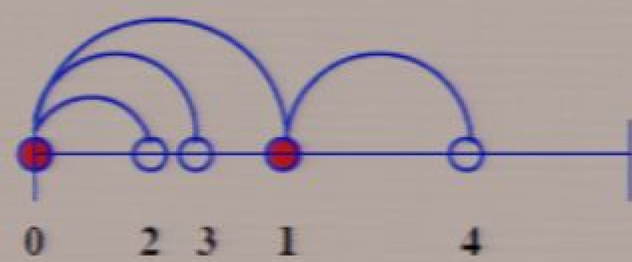
From line to tree

Integrating out the dependence on interval length from the conditional probability:

$$\begin{aligned} Pr\left[x_{t+1} \in I_k \,|\pi(t)\right] &= \int Pr\left[x_{t+1} \in I_k \,|\pi(t), \vec{s}(t)\right] dP(\vec{s}(t)) \\ &= \int s_k(t) dP(\vec{s}(t)) = \frac{1}{t+1}. \end{aligned}$$

i.e., **The probability to land in the k-th interval is uniform over all intervals.**

Preferential attachment with a cutoff



Let $d_j(t)$ equal the degree of fertile node j at time t.

The number of intervals contributing to j's fertility is $\min(d_j(t), A)$.

Probability node $(t+1)$ attaches to node j is:

$$Pr(t+1 \rightarrow j) = \min(d_j(t), A)/(t+1).$$

Standard PA: $Pr(t+1 \rightarrow j) = d_j(t)/(t+1)$.

The process on degree sequence

Let $N_0(t) \equiv$ number of infertile vertices.

Let $N_k(t) \equiv$ number of fertile vertices of degree k (for $1 \leq k < A$).

Let $N_A(t) \equiv$ number of fertile vertices of degree $k \geq A$
(i.e. $N_A(t) = \sum_{k=A}^{\infty} N_k(t)$ "the tail")

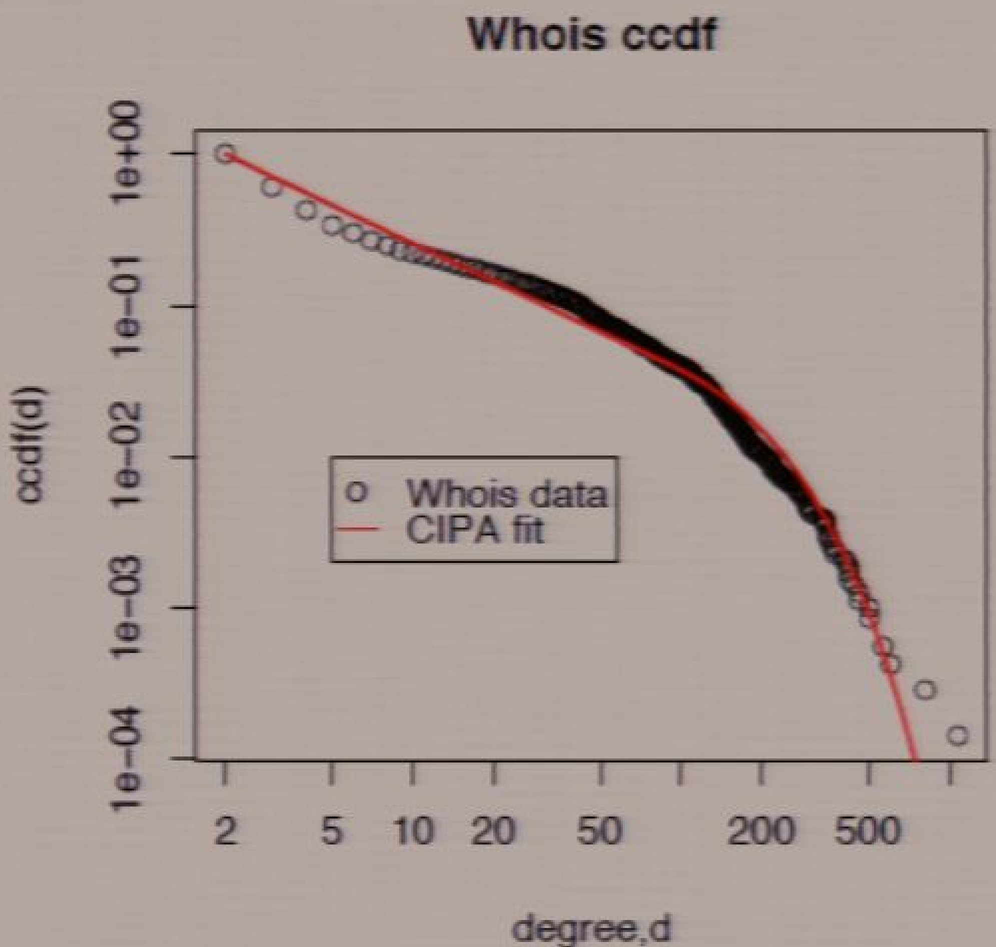
Rigorous Proofs for

- Power law for $d < A$, with $1 < \gamma < 3$.
- Exponential decay for $d > A$.

$$p_k = c_1 k^{-\gamma} \quad \text{for} \quad k < A.$$

$$p_k = c_2 \exp[-k/(A+1)] \quad \text{for} \quad k > A.$$

Using TPA to fit the AS-level Internet

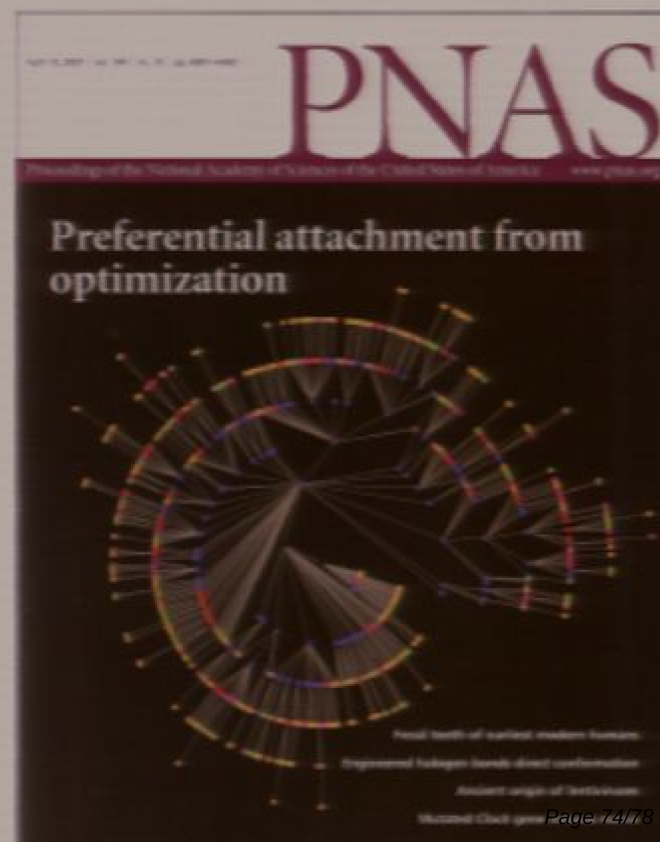


Extensions: Optimization and Network Growth

- Different cost functions and geometries:
  - Biological choices? (modularity versus efficiency)
  - Open-source software ("systems' motifs")
  - Economics/financial trades (trust versus value)

- Hierarchy and feedback
  (D'Souza, Roy *PRE* **78** 045101(R) 2008.)

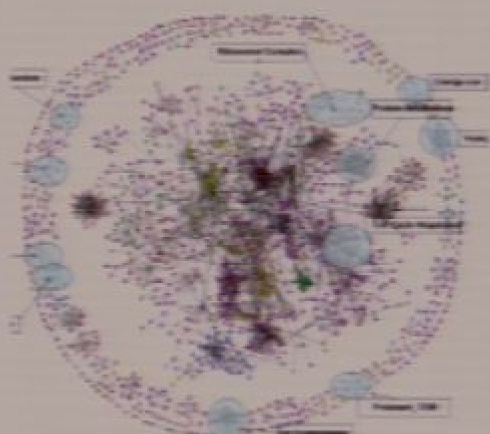
- Interacting networks.



Extensions to networked systems
(The physics of networks)

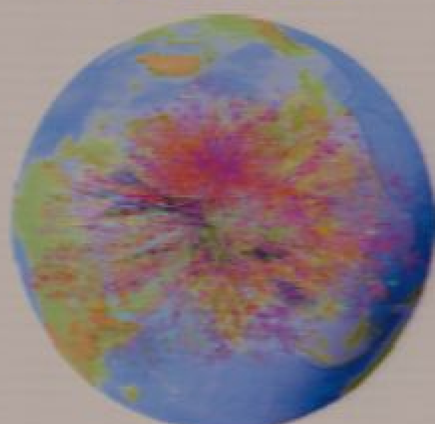
Networks:

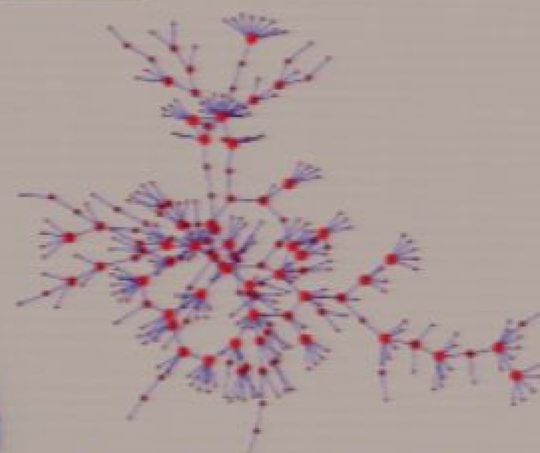
Transportation Networks/ Power grid
(distribution/ collection networks)



Biological networks
- protein interaction
- genetic regulation
- drug design

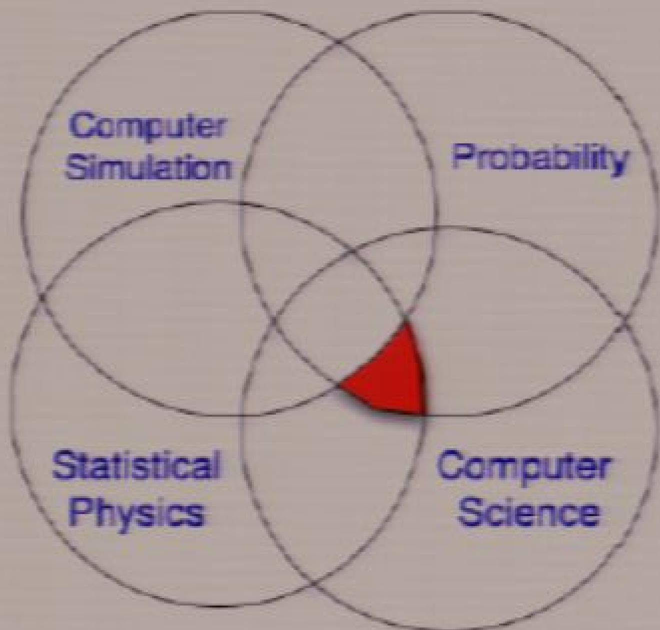
Computer networks





Social networks
- Immunology
- Information
- Commerce

Growing, Jamming and Changing Phase: Conclusions



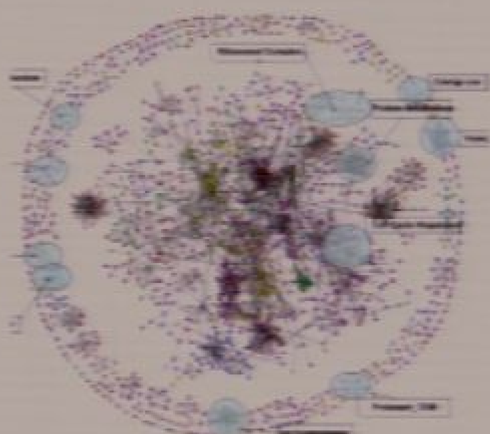
Geometry, Disorder and Algorithms
(Networks, Lattices)

- **Network Science**: Physicists contributing substantially to its foundation.
- **Phase transitions**: Understanding, manipulating, controling
- **Fruitful interplay between the fields**: e.g., connecting run-time of algorithms and critical slowing down

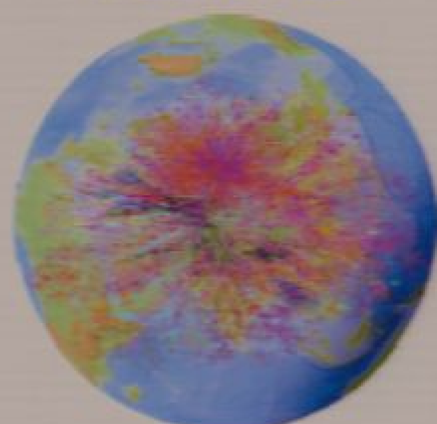
Extensions to networked systems
(The physics of networks)

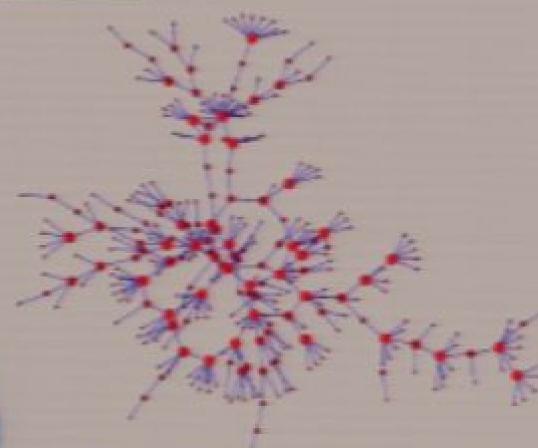
Networks:

Transportation Networks/ Power grid
(distribution/ collection networks)



Computer networks

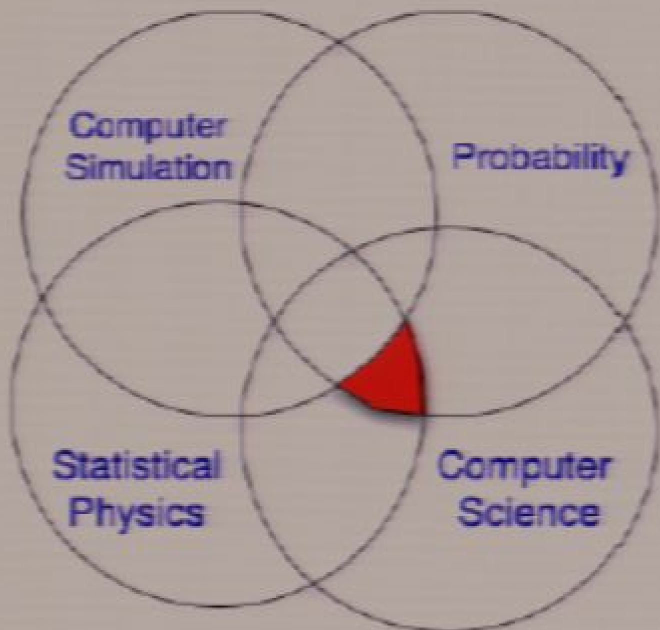




Biological networks
- protein interaction
- genetic regulation
- drug design

Social networks
- Immunology
- Information
- Commerce

Growing, Jamming and Changing Phase: Conclusions



Geometry, Disorder and Algorithms
(Networks, Lattices)

- **Network Science**: Physicists contributing substantially to its foundation.
- **Phase transitions**: Understanding, manipulating, controling
- **Fruitful interplay between the fields**: e.g., connecting run-time of algorithms and critical slowing down