Peer-to-Peer and Cloud Computing

Prof. Dr. Jörg Hähner

Part 3:

Network Models

Content

- Motivation
- Definitions
- Random Networks
- Small-world Networks
- Scale-free Networks

Peer-to-Peer system:

- Distributed resources provided by peers
- Overlay network of peers for accessing distributed resources
- Basic functionality of overlay network: resource lookup & routing
- Example: file sharing
  - Lookup peers providing a certain file
  - Find a path within the overlay network to peers providing the file

Suitable overlay network structure is of great importance for an efficient operation of the P2P system!

Desired Network Properties

What are basic principles of such a network?

- Efficient object lookup
  - Small number of overlay network hops to find requested resource
- Scalability
  - Include large number of nodes without performance degradation
    - Peer-to-Peer systems may include millions of nodes!
- Decentralised and self-organised
  - No central control of network
  - Peers do not have global view on network
  - Decentralised message forwarding
  - Network structure emerges through decentralised peer decisions
- Robust despite of dynamic changes
  - Deal efficiently with nodes frequently joining and leaving network

Social network:

- Nodes = people
- Edges = relations between people
  - E.g.: friends, relatives, colleagues
- Paths:
  - Chain of people
  - E.g.: friend of my friend of my friend

Interesting properties:

- Decentralised, self-organising, robust, scalable network
  - E.g.: people are born, die, get to know new people
- Efficient (short) communication paths & decentralised routing
  - Milgram's Experiment

Milgram‘s Experiment: Small World

- Analysis of paths in social networks conducted in 1960s by Stanley Milgram
- Milgram sends a letter to 160 randomly selected persons from Omaha and Nebraska (USA)
- Letter contains task: Deliver letter to a certain stock broker in Boston, Massachusetts, USA
- Constraints: persons must only send letter to someone they know on a first name basis (i.e.: friends, colleagues)
- Results:
  - 44 letters reached the target
  - Average number of “hops”: 6
    I.e., short path in network of 200 million US citizens!

It is definetely a small world!

Small World: Bacon Number

- Network derived from movie database
  - Nodes: actors
  - Edge between two actors if they have acted in same movie
- Bacon number of actor:
  Shortest path between actor and Kevin Bacon
  - Average bacon number: 2.9
  - Via Kevin Bacon, any actor can be linked to any other in 6 “hops”
- 6 is a typical distance between pairs of nodes in such networks → 6 degrees of separation

Yes: it is a small world!

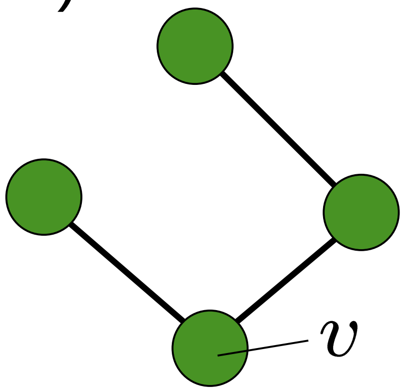
Small World: Various other examples

- "Die Zeit" linked a Turkish Kebab-shop owner in Frankfurt, Germany, to his favorite Actor Marlon Brando in 6 hops:
  - Shop owner has a friend in California who works together with the boyfriend of a woman who is the sorority sister of the daughter of the producer of the movie "Don Juan" starring Marlon Brando.
- Erdös number: distance in graph of paper (co-) authors
  - Paul Erdös: famous mathematician (> 1500 papers with > 500 co-authors)
  - Average Erdös number: 4.7
  - Average distance between authors: 7.3

Which theoretical models can describe such networks?

- Motivation
- Definitions
- Random Networks
- Small-world Networks
- Scale-free Networks

- Overlay network modeled as graph: $G = (V, E)$
  - V : set of vertices (peers)
  - E : set of edges (direct link in overlay network)
    - For simplicity we consider an undirected and unweighted graph
    - $e = \{v, w\}$: undirected edge from v to w i.e., peer v has routing table entry with target w and vice versa
    - All edges have same weight 1

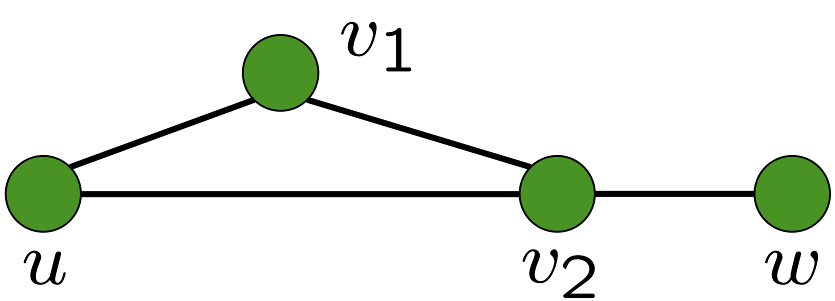


$k_v = 2$

- Degree of vertex v
  - Degree: $k_v = |\{\{v', w'\} \in E | v' = v \lor w' = v\}|$

Definitions (2)

- Overlay network path from u to w: sequence of edges
$$p(u,w) = (\{u,v_1\},\{v_1,v_2\},\dots,\{v_{n-1},v_n\},\{v_n,w\})$$
- Set of Paths connecting vertex u and w
$$P(u,w)$$
- Connected graph: a graph is connected if a path exists between any two vertices
$$\forall u \in V\ \forall w \in V: u = w\ \vee P(u,w) \neq \emptyset$$



$$p_1(u,w) = (\{u,v_1\},\{v_1,v_2\},\{v_2,w\})$$
$$p_2(u,w) = (\{u,v_2\},\{v_2,w\})$$
$$p_3(u,w) = (\{u,v_2\},\{v_2,v_1\},\{v_1,u\},\{u,v_2\},\{v_2,w\})$$

...

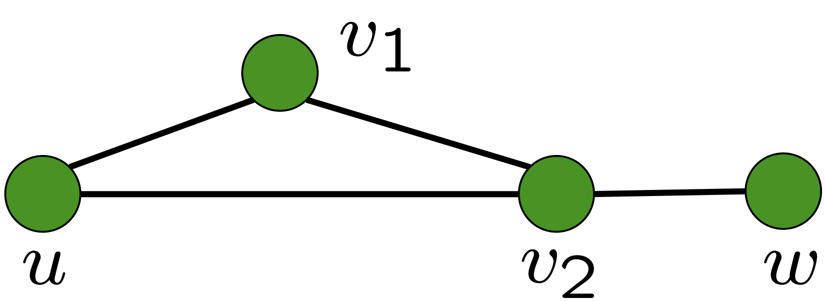
Definitions (3)

- Length of path $p(u,w)$: number of edges in path

$$|p(u,w)|$$

- Distance between two vertices u and w: shortest path between u and w

$$d_{u,w} = \begin{cases} 0 & \text{if } u = w \\ \min_{p \in P(u,w)} |p(u,w)| & \text{if } P(u,w) \neq \emptyset \\ \infty & \text{otherwise} \end{cases}$$



$$p_1(u,w) = (\{u, v_1\}, \{v_1, v_2\}, \{v_2, w\})$$

$$p_2(u,w) = (\{u, v_2\}, \{v_2, w\})$$

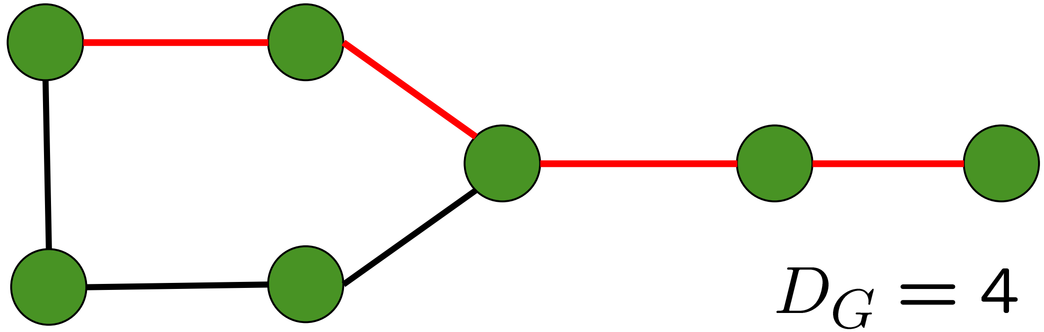
$$d_{u,w} = |p_2(u,w)| = 2$$

- Average path length of graph G:

$$L_G = \frac{\sum_{u,w \in V} d_{u,w}}{|V|(|V|-1)/2}$$

- Diameter of graph G: maximum distance between any two vertices of graph

$$D_G = max_{u,w \in V}\,(d_{u,w})$$



$D_G = 4$

- Neighbourhood of vertex v:

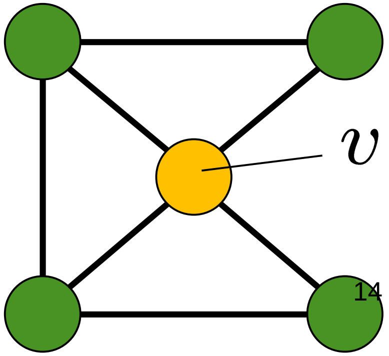
$$N_v = \{v' \in V \mid \exists \{u,w\} \in E : u = v \land w = v'\}$$

- Clustering coefficient of vertex v:
  - number of edges between neighbours divided by the maximum number of edges that could exists between neighbours
  - measure for “cliquishness” of neighbourhood (friends of mine are also friends of each other)

$$C_v = \frac{\left|\left\{\{u,w\} \in E : u \in N_v \land w \in N_v\right\}\right|}{k_v(k_v - 1)/2}$$

$$|N_v| = 4$$

$$C_v = \frac{3}{4(4-1)/2} = \frac{3}{6}$$



- Motivation
- Definitions
- Random Networks
- Small-world Networks
- Scale-free Networks

Basic idea: Edges are built at random between a fixed number n of vertices

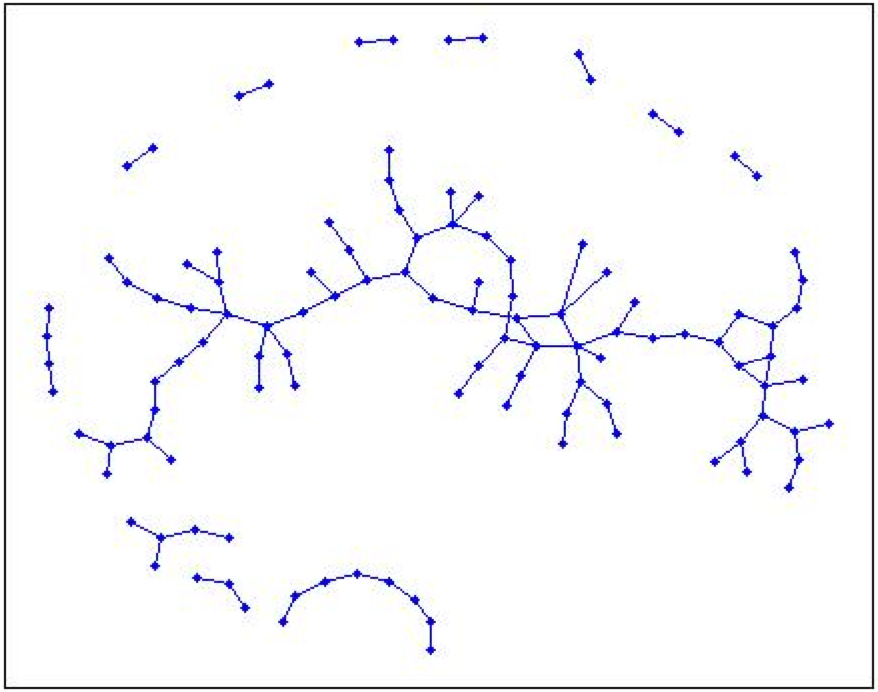
- Different random graph models exist, in particular:
  - Erdös-Renyi random graph
    - Set of all graphs with n vertices and m edges: $\mathcal{G}_{n,m}$
    - Random graph $G_{n,m}$: randomly drawn graph from $\mathcal{G}_{n,m}$
  - Gilbert random graph
    - Probability of an edge between any two vertices: $p \in [0; 1]$
    - Random graph $G_{n,p}$: for any pair $\{v, w\}$ of vertices draw a random number $p_{v,w} \in [0; 1]$ and add an edge if $p_{v,w} \leq p$
- Models interchangeable for $m = \frac{n(n-1)}{2} p$

- Expected property Q of a random graph $G_{n,p}$

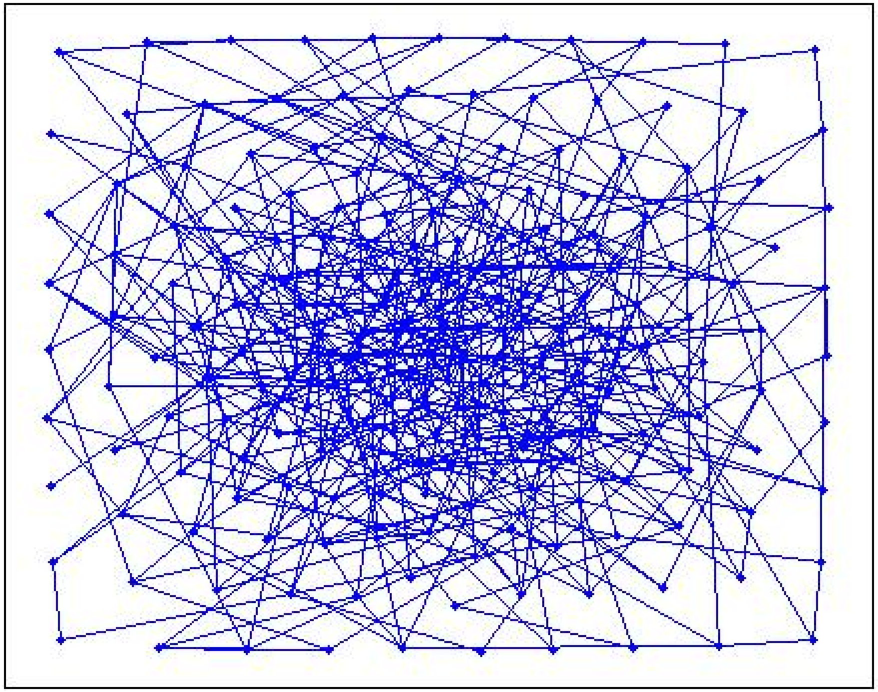
$$P\left(G_{n,p}\ has\ Q\right) \rightarrow 1\ for\ n\ \rightarrow\ \infty$$

- Many interesting properties appear suddenly in random graphs when p crosses a certain threshold value
→ phase transition
  - $pn\ <\ 1$: network expected to consist of isolated tree-shaped components of size O(ln n)
  - $1\ <\ pn\ <\ \ln n$: network expected to have one giant connected component of size O(n) and smaller components of size O(ln n)
  - $pn\ >\ \ln n$: network expected to be fully connected
    - Logarithmic (small) number of links per node sufficient for connected network!

Giant Component



Fully Connected Network

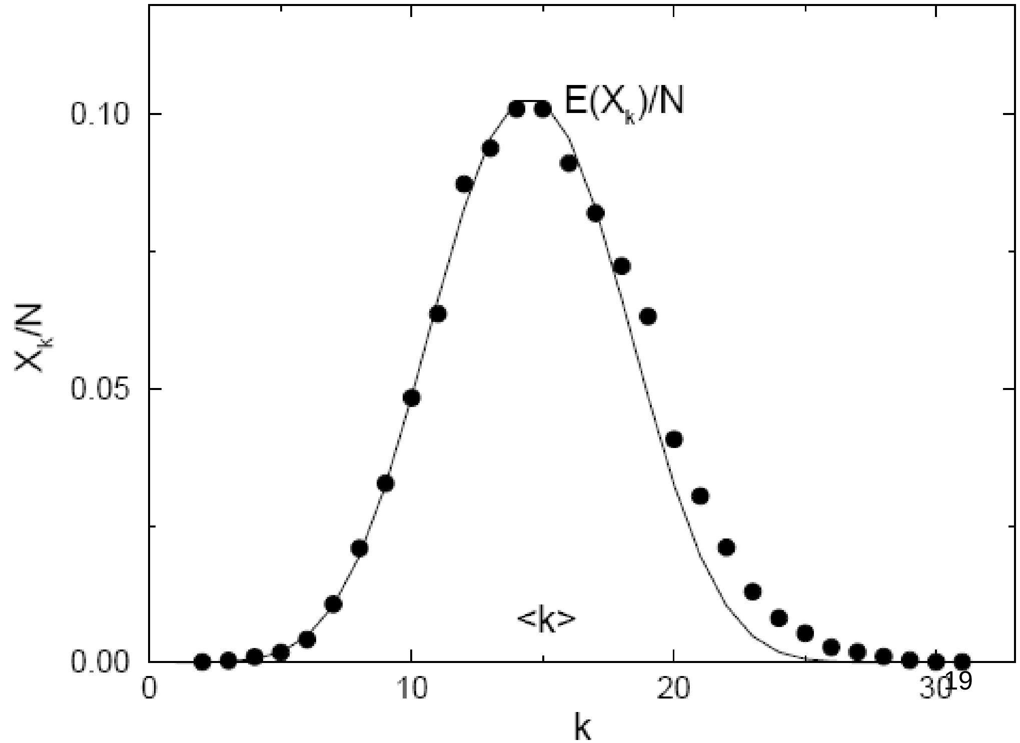


[Source: David Gleich, Stanford University]

- Node degree distribution:
  Poisson distribution for n $\rightarrow \infty$
  - Mean value: pn = <k>

$$\langle k \rangle = p(n-1) = 10000*0.0015 \sim 15$$

$$\mathrm{P}(k) = \frac{(pn)^k \mathrm{e}^{-pn}}{k!}$$



Degree distribution obtained by numerical simulation of random graph with n = 10,000, p = 0.0015

[Albert & Barabasi: Statistical Mechanics of Complex Networks. Reviews of Modern Physics 74, 2002]

Properties of Random Graphs (3)

- Expected diameter $D_{G_{n,p}}$ for $pn > \ln n$ (connected graph) is $O(\frac{\ln n}{\ln pn})$
  - Short (logarithmic) path between any pair of nodes exists with high probability!
  - Similar to real small worlds (social networks, communication networks, biological networks, etc.)
- Expected clustering coefficient of randomly selected vertex v:

$$C_v = p$$

Is this similar to real-world networks?

Path Lengths and Clustering Coefficients of Different Real Networks

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference	Nr.
WWW, site level, undir.	153,127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999	1
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001	2
Movie actors	225,226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998	3
LANL coauthorship	52,909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b	4
MEDLINE coauthorship	1,520,251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b	5
SPIRES coauthorship	56,627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c	6
NCSTRL coauthorship	11,994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b	7
Math coauthorship	70,975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.* 2001	8
Neurosci. coauthorship	209,293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.* 2001	9
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000	10
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000	11
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000	12
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000	13
Words, cooccurence	460.902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001	14
Words, synonyms	22,311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.* 2001	15
Power grid	4,941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998	16
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998	17

[source: Albert & Barabasi:
Statistical Mechanics of Complex Networks.
Reviews of Modern Physics 74, 2002]

Significantly different clustering coefficients!

Content

- Motivation
- Definitions
- Random Networks
- Small-world Networks
- Scale-free Networks

- Small diameter and average path length
  - Comparable to random graphs with same number of nodes and edges

and

- Dense local structure, i.e. high clustering coefficient

- Neither a regular grid-like network model nor a pure random network fulfill both of these properties:
  - Grid-like networks show regularity and locality, but have a high average path length and diameter.
  - Random graphs have a small clustering coefficient of p

- Ring of n vertices
- Each vertex is connected to its neighbours within distance l in the ring

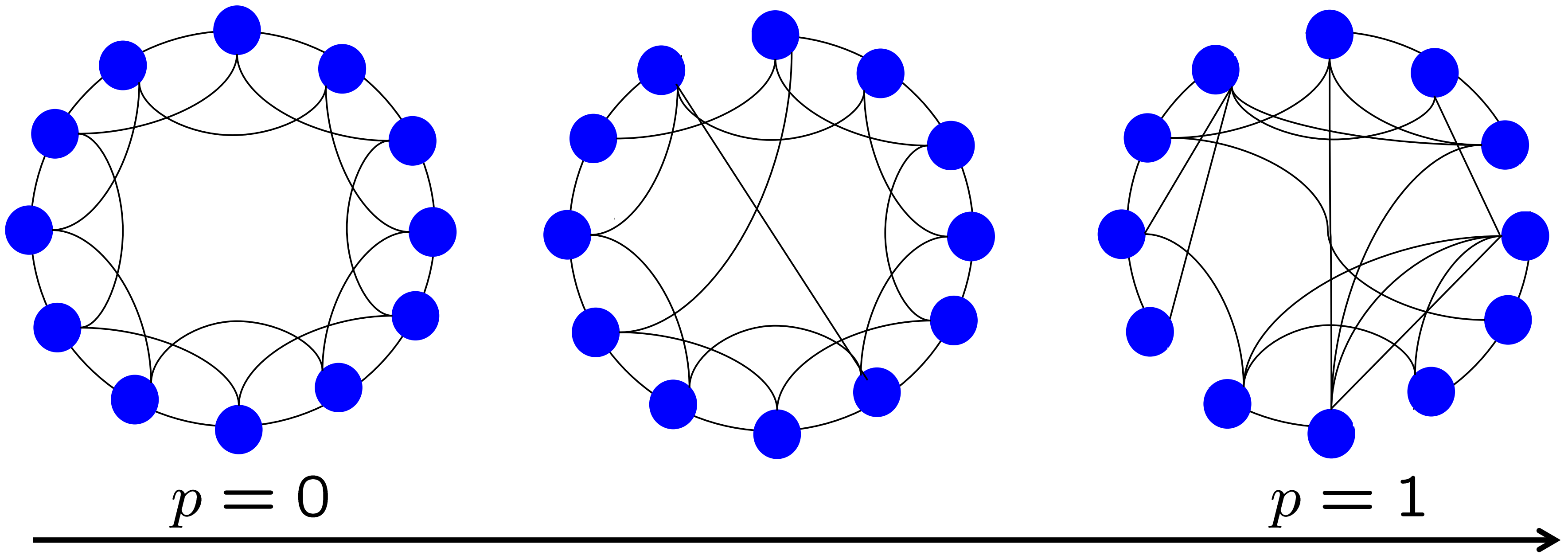
→ regular structure with dense local clusters

Random re-wiring process

- Ring is traversed in clockwise order in multiple rounds
- 1st round considers regular edges to nearest neighbours of each vertex; 2nd round considers regular edges to 2nd-nearest neighbours, etc.
- Each regular edge is replaced by a random (long range) edge with probability p
  - Target of random edge is chosen uniform randomly from V
  - No duplicate edges are allowed

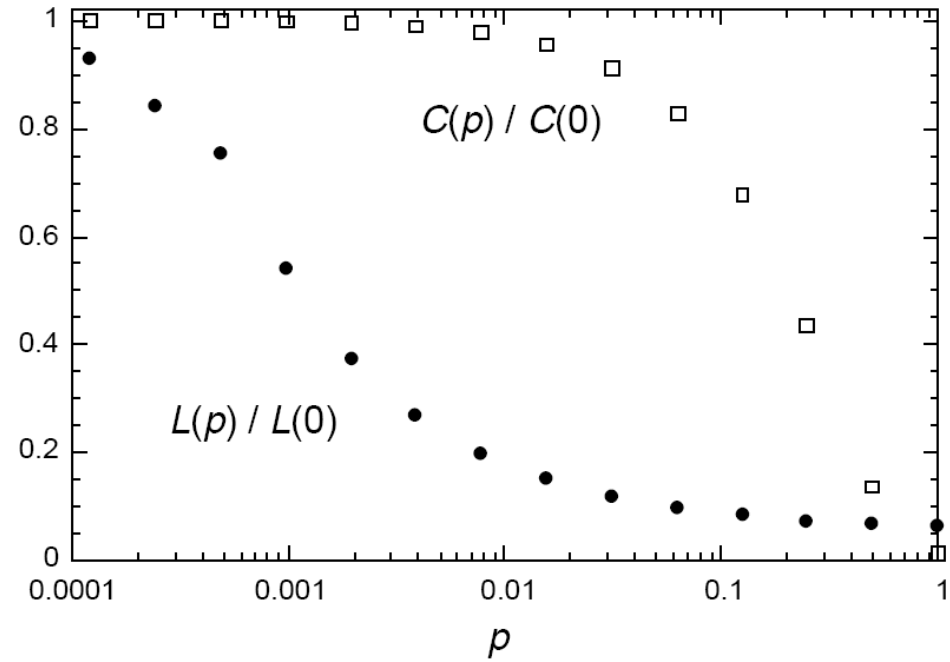
→ random structure

Watts-Strogatz-Model (2)



- p = 0: regular graph
  - Expected diameter is O(n)
- p = 1: random graph
  - Expected diameter is O(ln n)
- What about 0 < p < 1?

Average clustering coefficients and path lengths of rewired graphs
[source:Watts, Strogatz: Collective dynamics of 'small world' networks. Nature 394, 1998]



Values are normalized by the values of regular graphs (p = 0)

For small p (only few long range edges):

- High clustering coefficient comparable to small worlds: $C(p) \gg C_{random}$
- Small path length comparable to random graphs: $L(p) \ll L_{regular}$

Many P2P systems emerge into a small-world network

- Few random links sufficient to achieve connected graph with high probability
  - Unstructured P2P systems like GNUTELLA only need few links to connect the majority of peers
- Random network structure leads to small path lengths
  - Unstructured P2P systems like GNUTELLA have short paths – however, finding them *efficiently* is hard without any structure
- Regular basic structure + few non-regular links increase clustering coefficient
  - Some DHTs like Symphony have a regular structure (i.e. a ring) as basis and add few “random” (long-distance) links

- We have seen: If a P2P system is a small world network, a short path exists
- Task of a P2P system: find a short path to target
- Previous models shows that short paths exist
  - However: Is it possible to find these paths efficiently?
- Wanted: Efficient routing algorithm for finding short paths operating only on local knowledge of the network
  - Nodes do not have global knowledge of the network, i.e. all other nodes and links (otherwise a centralised algorithm can be used)
  - Each node knows its neighbours only
    - Similar to social networks: people only know their friends and still can forward messages to unknown people (→Milgram's experiment)

Simple routing algorithm:

- Forward incoming message to all neighbours except for the node the message has been received from
- Node discards message that it has seen already
  - Avoids circular routing

Properties:

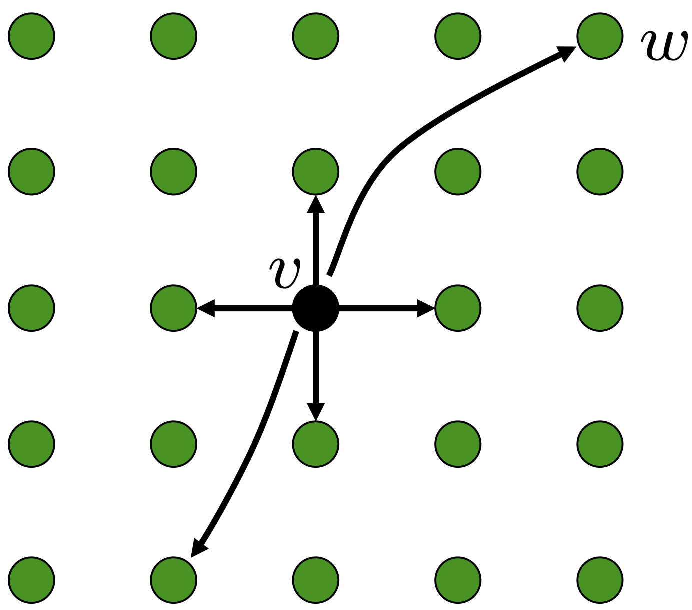
- Flooding finds all paths, including shortest one
- High overhead since all nodes receive the message

Is there an efficient local routing algorithm that forwards messages to only one node at a time?

- Combination of local (short-range) and long-range connections (so far similar to Watts-Strogatz model)
- Regular structure for set-up of short range connections: grid
  - Vertex v has links to each vertex w within distance $d_{v,w} \leq l$ for constant l
- Random set-up of q long range connections $\{e_{v,1},\ldots,e_{v,q}\}$ for each vertex v:
  - Target of edge $e_{v,i}$ is w with probability

$$P\left(e_{v,i} = (v,w)\right) = \frac{{d^{-r}}_{v,w}}{\sum_{u \in V} {d^{-r}}_{v,u}}$$

$d_{v,w} = 4$



$l = 1 \quad q = 2$

Routing algorithm A only operates on local knowledge, i.e. each node v only knows

- v's grid coordinates and coordinates of its contacts
- Grid coordinates of target node (from message)
- Regular grid structure

Performance of A:

Expected path length from message source to target node

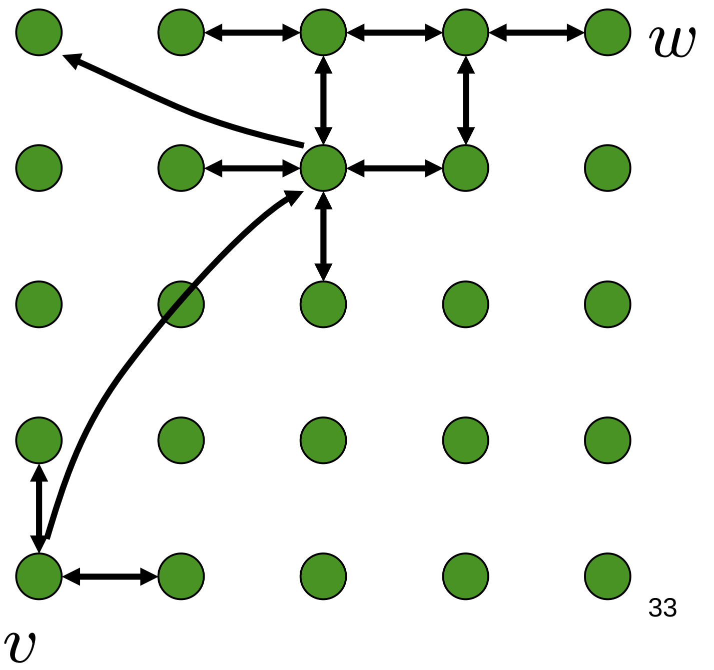
Performance of Decentralised Routing Algorithms (2)

- Expected path length dependent on r:
  - $r = 2$: for $l = q = 1$ an A exists with path length of $O(\log^2 n)$
  - $r \neq 2$: no A exists with path length of $O(\log^2 n)$
- In general, if grid has dimension dim:
  A achieving path length of $O(\log^2 n)$ only exists if $r = dim$!
- Idea of proof:
  - $0 \leq r < 2$: most shortcuts too long and distant to target when message closes in on target
  - $r > 2$: most "shortcuts" too short to shorten path significantly compared to path in grid

Decentralised forwarding algorithm operating only on local knowledge: Greedy forwarding

- Forward message to contact with shortest grid distance (Manhattan distance) to target node

$l = q = 1$



Expected path length for q $\geq$ 1 on Kleinberg-like models using greedy routing (e.g., Symphony, Random Chord):

- O(($\log^2$ n) / q)
  - O(log n) for q = log n

Different structured P2P systems based on distributed hash tables apply the ideas of the Kleinberg model:

- Hash function for mapping peers and resources to (1D) coordinates
  - For instance, hash of file and peer IP address
  - Distance metric defined on hash values
- Regular basic structure, e.g. ring
- Small number of long-range contacts for short network paths
- Greedy routing strategy for forwarding messages to peer providing resource

More about this in section on structured P2P systems.

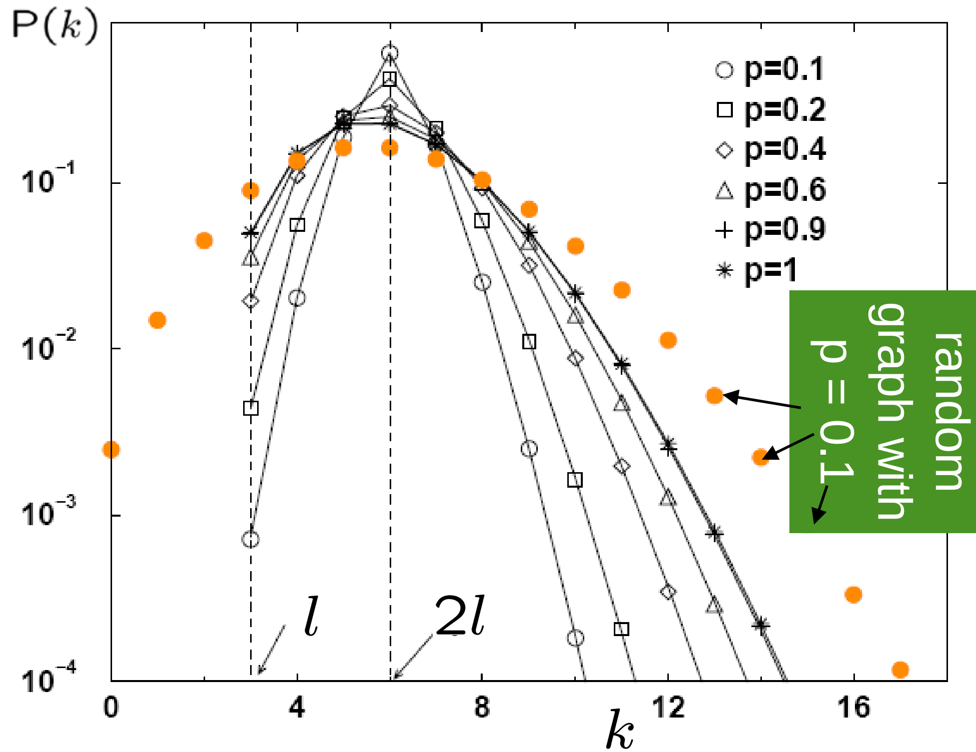
Content

- Motivation
- Definitions
- Random Networks
- Small-world Networks
- Scale-free Networks

Degree distributions of presented random and small world models:

- Homogeneous distribution, i.e., most vertices have similar degrees
- Random model & Watts-Strogatz model: Poisson distribution

$$P(k) = \frac{(pn)^k e^{-pn}}{k!}$$



Degree distribution of Watts-Strogatz Model [Source: Barrat & Weigt: On the properties of small-world network models. Europ. Phys. J. B 13, 547 (2000)]

What about real networks?

The degree distribution of vertices in many networks (social, communication, biological, etc.) follows a power law: $P_\lambda(k) \sim k^{-\lambda}$ for some exponent $\lambda > 1$

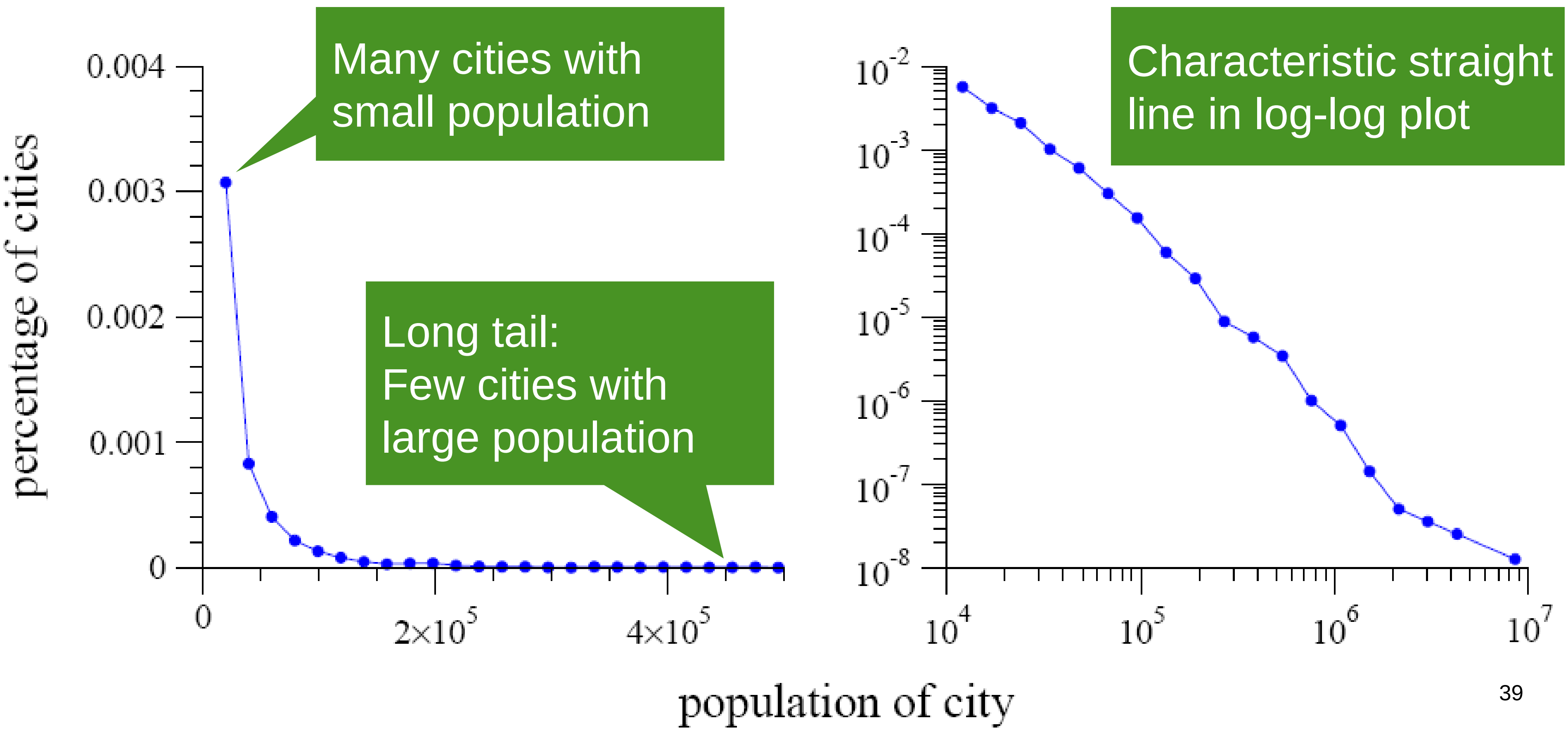
- Degree distribution is right-skewed (“rechts-verdreht”) / heavy-tailed rather than homogeneous
  - Many vertices with low degree
  - Few vertices with high degree (hubs)
    - Probability of nodes with high degree decreases polynomial
    - In contrast: random & Watts-Strogatz graphs: exponential decrease
- Such networks are called scale-free networks or power-law networks.

What does „scale-free“ mean?

- Distribution looks similar, no matter at which scale we look at it
- Ratio $\frac{P(sk)}{P(k)}$ only depends on s, not on k
- Example (not based on real numbers):
  - Ratio of 2 KB files to 1 KB files (i.e., s = 2, k = 1 KB): ¼
  - Ratio of 2 MB files to 1 MB files (i.e., s = 2, k = 1 MB): ¼
- “Scale“ does not come from scalability with respect to number of network nodes!

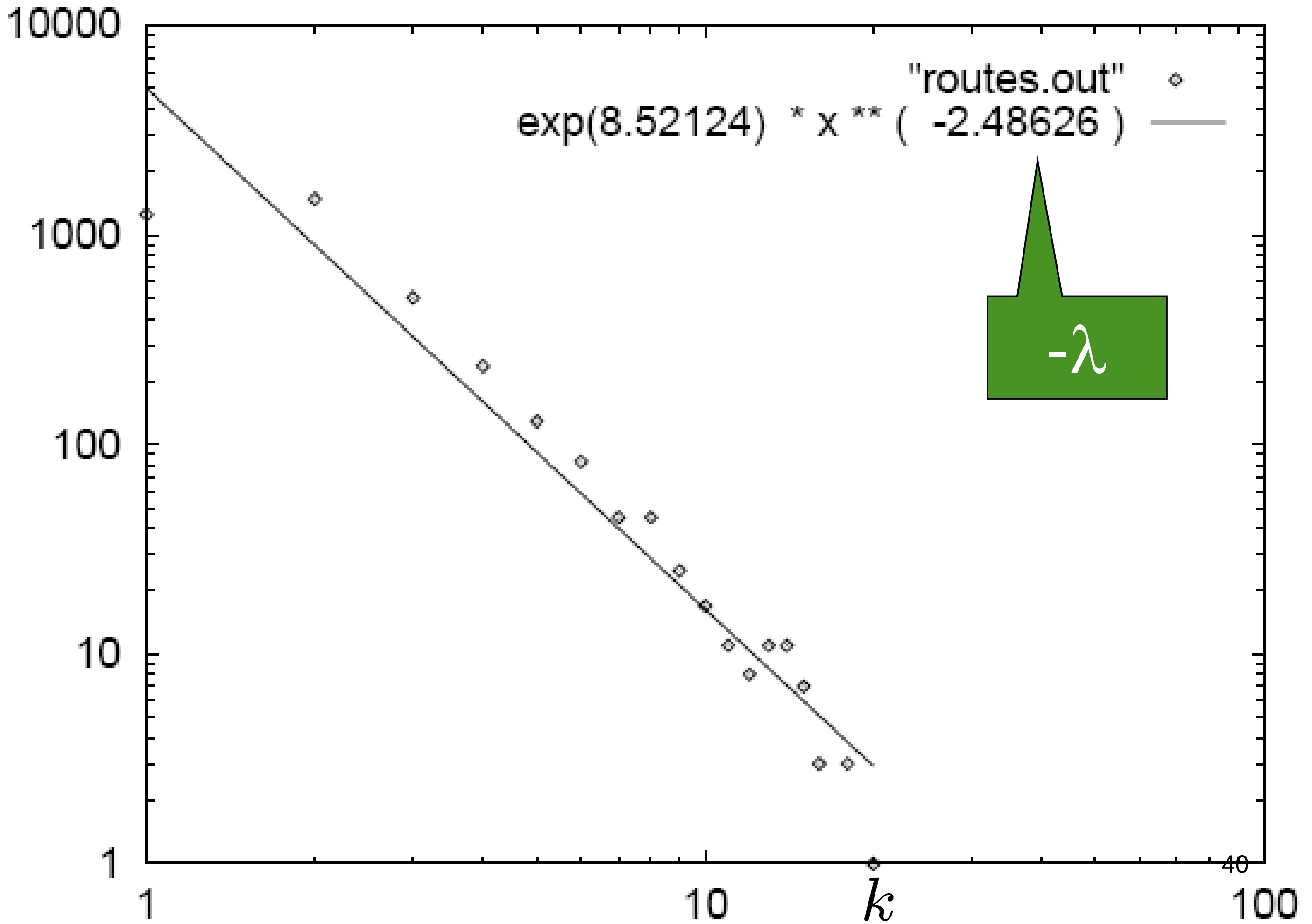
Population distribution of US cities

[source: Newman: *Power laws, Pareto distributions and Zipf's law*. University of Michigan]



Degree distribution of Internet Routers in 1995

[Source:
Faloutsos: *On Power Law Relationships of the Internet Topology*.
Proceedings of SIGCOMM 1999]



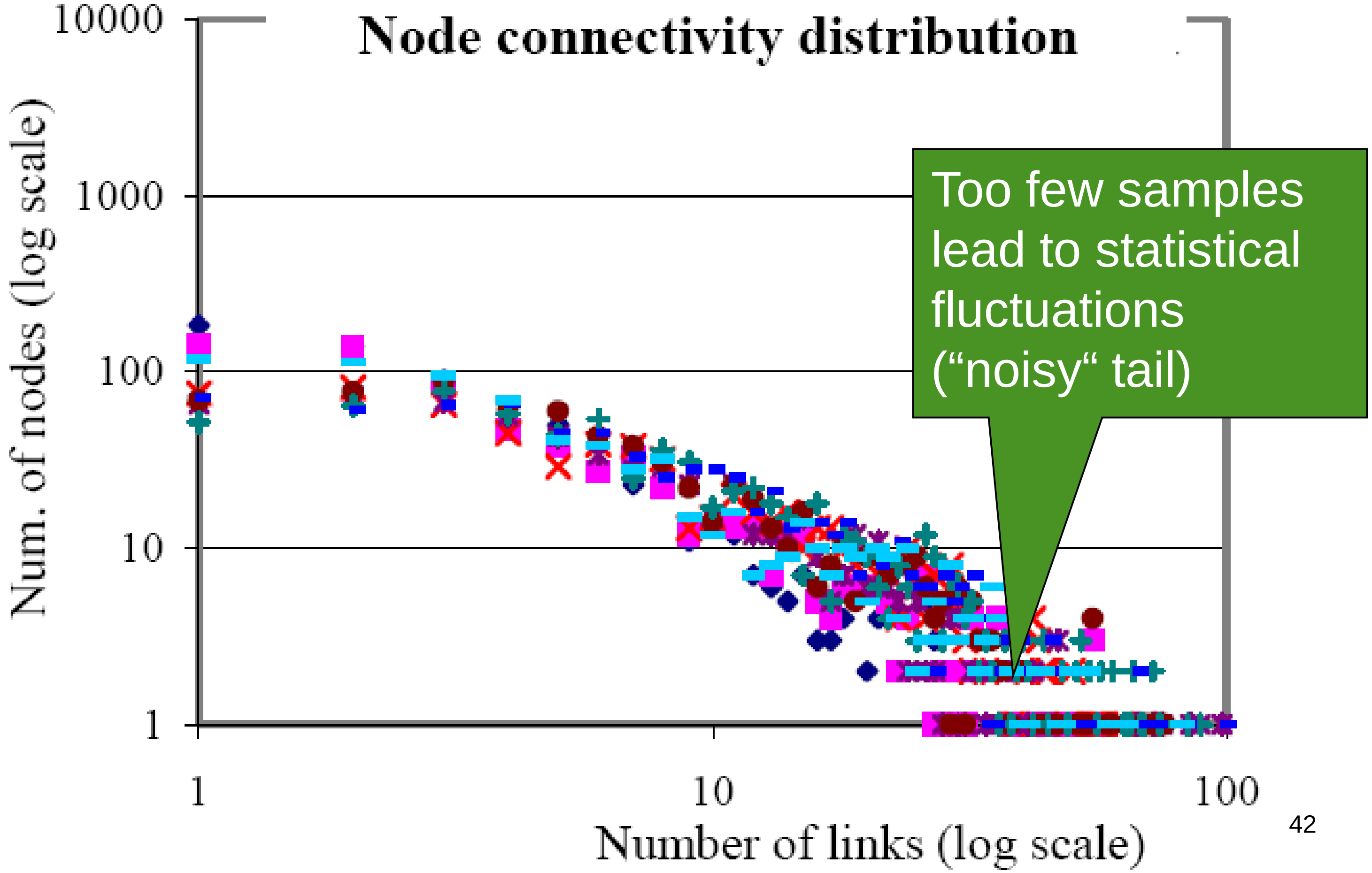
Various Real Power-law Distributions

[Source: Albert & Barabasi:
Statistical Mechanics of Complex Networks.
Reviews of Modern Physics 74, 2002]

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}
WWW	$325,729$	4.51	900	2.45	2.1
WWW	4×10^7	7		2.38	2.1
WWW	2×10^8	7.5	$4,000$	2.72	2.1
WWW, site	$260,000$				1.94
Internet, domain*	$3,015$ - $4,389$	3.42 - 3.76	$30 - 40$	2.1 - 2.2	2.1 - 2.2
Internet, router*	$3,888$	2.57	30	2.48	2.48
Internet, router*	$150,000$	2.66	60	2.4	2.4
Movie actors*	$212,250$	28.78	900	2.3	2.3
Coauthors, SPIRES*	$56,627$	173	$1,100$	1.2	1.2
Coauthors, neuro.*	$209,293$	11.54	400	2.1	2.1
Coauthors, math*	$70,975$	3.9	120	2.5	2.5
Sexual contacts*	2810			3.4	3.4
Metabolic, E. coli	778	7.4	110	2.2	2.2
Protein, S. cerev.*	1870	2.39		2.4	2.4
Ythan estuary*	134	8.7	35	1.05	1.05
Silwood park*	154	4.75	27	1.13	1.13
Citation	$783,339$	8.57			3
Phone-call	53×10^6	3.16		2.1	2.1
Words, cooccurence*	$460,902$	70.13		2.7	2.7
Words, synonyms*	$22,311$	13.48		2.8	2.8

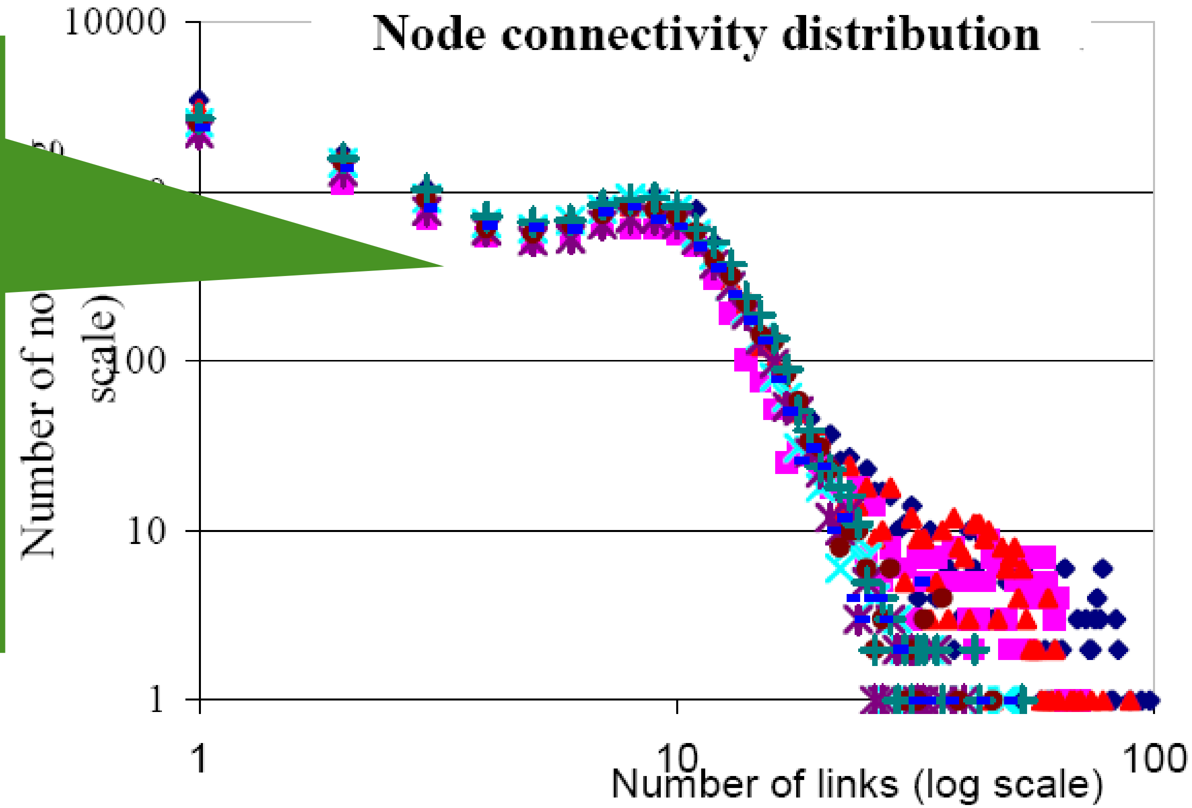
Degree distribution of Gnutella Peers in 2000

[Source: Ripeanu, Foster, Iamnitchi: Mapping the Gnutella Network: *Properties of Large-Scale Peer-to-Peer Systems and Implications for System Design*. IEEE Internet Computing Journal, 2002]



Degree distribution of Gnutella Peers in 2001

- For small k: deviation from power law (too few users with small number of connections)
- Users might think: The more connnections, the better the download rate
- Gnutella users are power users with good connections



[Source: Ripeanu, Foster, Iamnitchi: Mapping the Gnutella Network: *Properties of Large-Scale Peer-to-Peer Systems and Implications for System Design*. IEEE Internet Computing Journal, 2002]

- Network Growth
  - Nodes successively enter the network
  - In contrast (e.g. random graph & Watts-Strogatz models): fixed number of nodes
- Preferential attachment
  - New nodes preferentially attach to nodes that have many neighbours
  - In contrast: random graph & Watts-Strogatz choose neighbours without preference w.r.t. node degree

→ Network growth and preferential attachment lead to scale free-networks!

- Start with small number $\mathrm{n}_0 = |\mathrm{V}_0|$ of nodes
- At each time step t add one node v to network
  - v establishes links to $\mathrm{m} \leq \mathrm{n}_0$ nodes using preferential attachment
  - **Preferential attachment**
    edge $e = \{v, w\}$ is added with probability

$$P(e = \{v, w\}) = \frac{k_w}{\sum_{u \in V_t} k_u}$$

    - New nodes will link to nodes with high degree with higher probability
    - *“The rich get richer”*

- Power-law degree distribution
  - $P(k) \sim k^{-3}$
  - $\lambda = 3$ independent of n_0
  - Real scale-free networks typically $\lambda \in [1,3]$

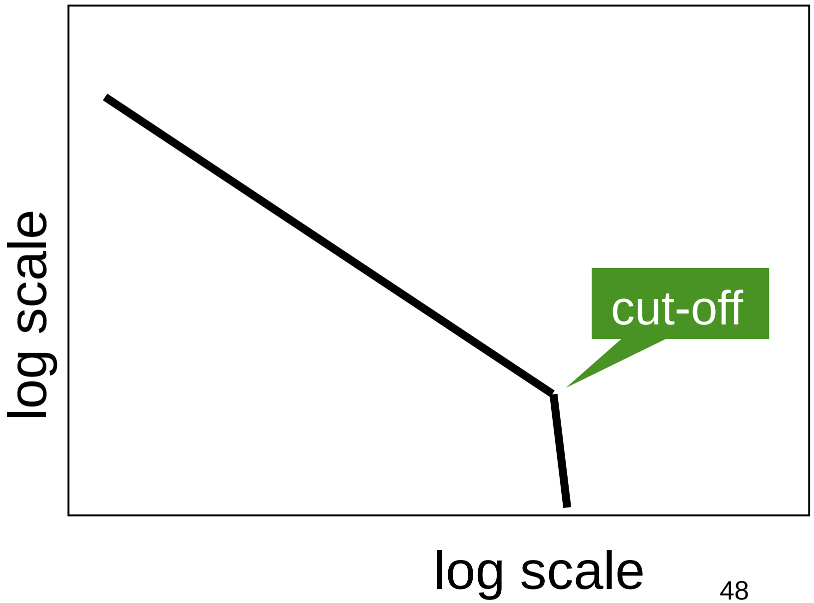
Network	γ_{out}	γ_{in}
WWW	2.45	2.1
WWW	2.38	2.1
WWW	2.72	2.1
WWW, site		1.94
Internet, domain*	2.1 - 2.2	2.1 - 2.2
Internet, router*	2.48	2.48
Internet, router*	2.4	2.4
Movie actors*	2.3	2.3
Coauthors, SPIRES*	1.2	1.2
Coauthors, neuro.*	2.1	2.1
Coauthors, math*	2.5	2.5
Sexual contacts*	3.4	3.4
Metabolic, E. coli	2.2	2.2
Protein, S. cerev.*	2.4	2.4
Ythan estuary*	1.05	1.05
Silwood park*	1.13	1.13
Citation		3
Phone-call	2.1	2.1
Words, cooccurence*	2.7	2.7
Words, synonyms*	2.8	2.8

- Graph Diameter D_G
  - $n_0 = 1$: $D_G \in O(\log n)$
  - $n_0 \geq 2$: $D_G \in O(\frac{\log n}{\log \log n})$
  - Slightly underestimates path length of real scale-free graphs

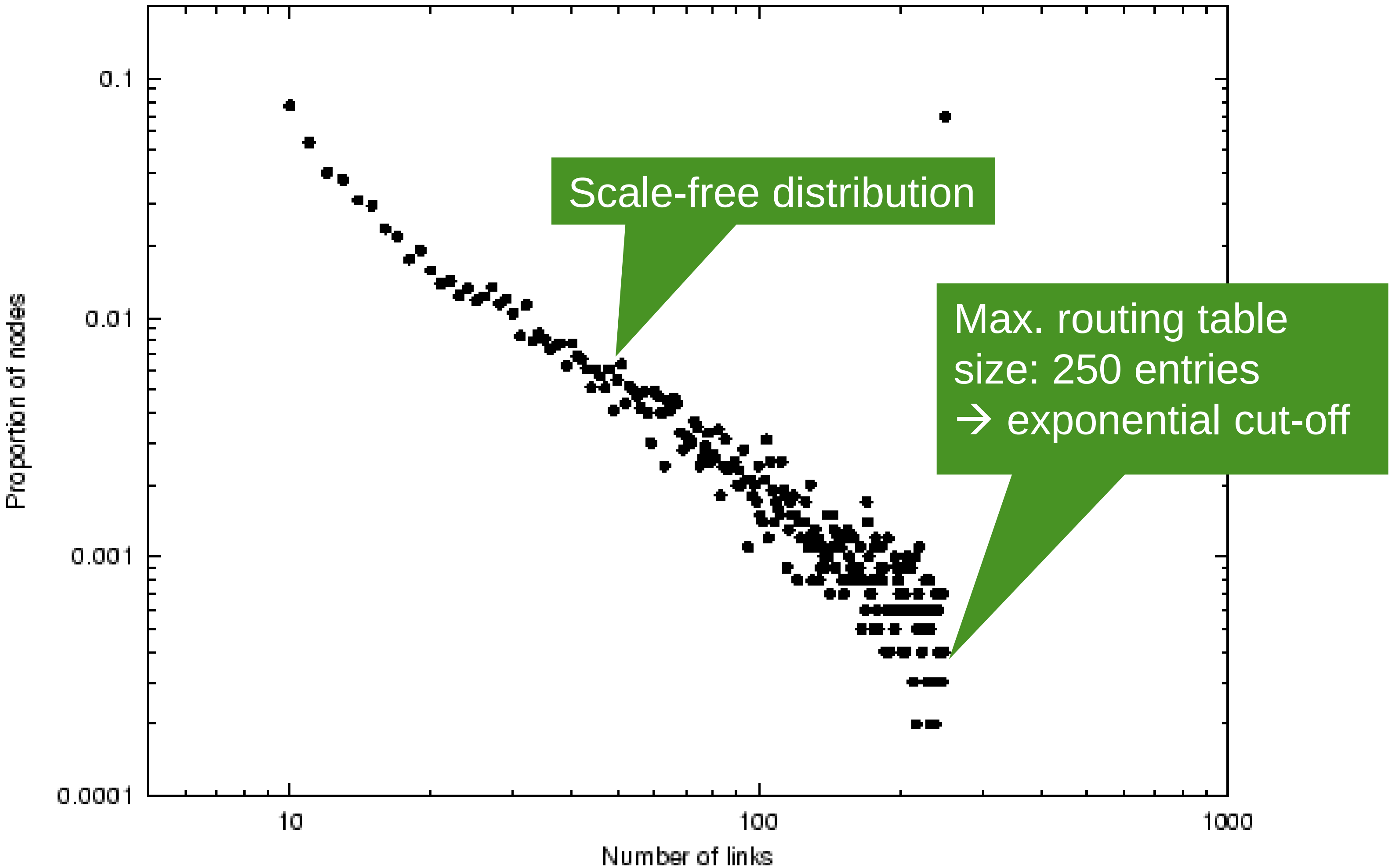
Network	ℓ_{real}	ℓ_{rand}	ℓ_{pow}
WWW	11.2	8.32	4.77
WWW			
WWW	16	8.85	7.61
WWW, site			
Internet, domain∗	4	6.3	5.2
Internet, router∗	12.15	8.75	7.67
Internet, router∗	11	12.8	7.47
Movie actors∗	4.54	3.65	4.01
Coauthors, SPIRES∗	4	2.12	1.95
Coauthors, neuro.∗	6	5.01	3.86
Coauthors, math∗	9.5	8.2	6.53
Sexual contacts∗			
Metabolic, E. coli	3.2	3.32	2.89
Protein, S. cerev.∗			
Ythan estuary∗	2.43	2.26	1.71
Silwood park∗	3.4	3.23	2
Citation			
Phone-call			
Words, cooccurence∗			
Words, synonyms∗			

Characteristics of the Barabasi-Albert Model (3)

- Clustering coefficient
  - Decreases with network size following a power law
  - No high degree of cliquishness
    - In contrast to Watts-Strogatz model and real networks
- Does not model some properties such as exponential cut-off
  - Exponential distribution for very large degrees
  - Example: limited peer capacity
    - Real peers cannot handle huge numbers of TCP connections



Node Degree Distribution in Freenet P2P System



[Source: Clarke, Sandberg, Wiley, Hong: *Freenet: A Distributed Anonymous Information Storage and Retrieval System*. In Proc. of the ICSI Workshop on Design Issues in Anonymity and Unobservability, 2000]

We consider two different kinds of node break-downs:

- Failures
  - Random node is removed from the network
- Attacks
  - Most important nodes are removed from network
    - Most important = nodes with highest degree (hubs)

Efficiency E_G of Graph G:

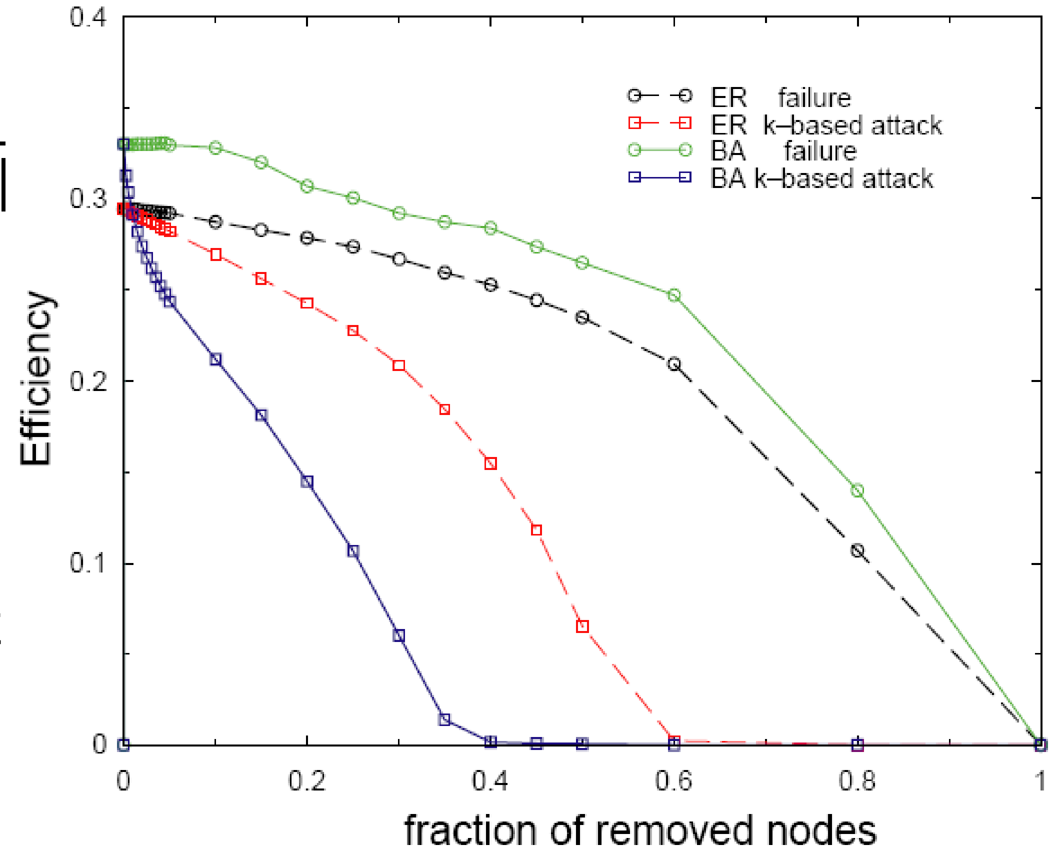
$$E_G = \frac{1}{|V|(|V|-1)} \sum_{v \neq w} \frac{1}{|d_{v,w}|}$$

Scale-free Graph

- Resists failures well
- Strong impact of attacks

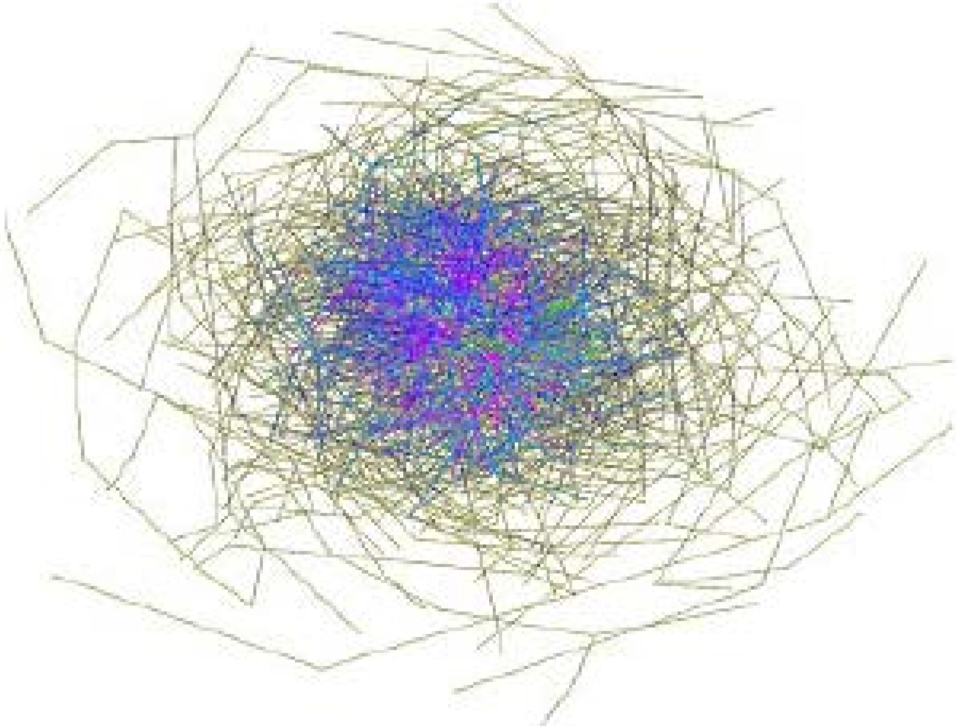
Random Graph

- Similar robustness against failures
- Better robustness against attacks

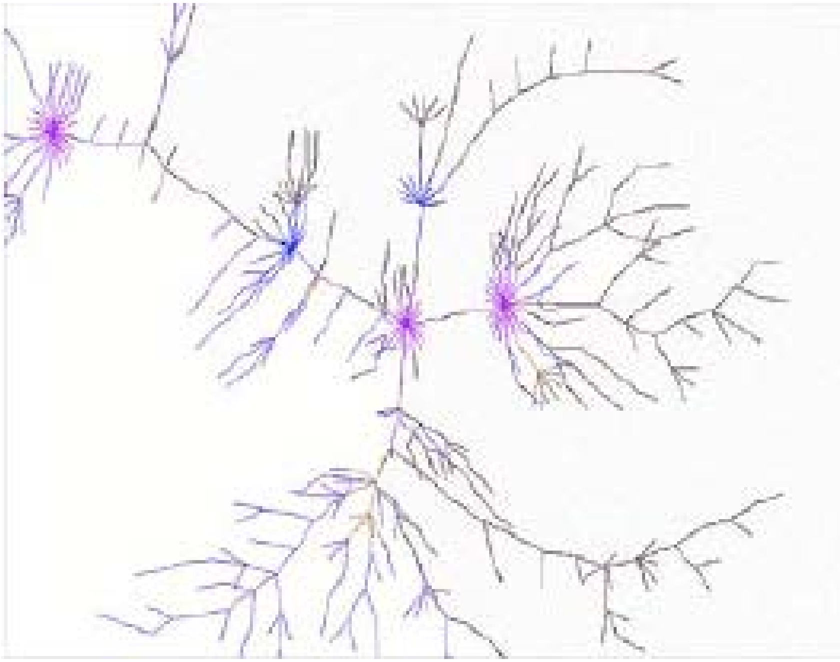


[Source: Crucitty, Latora, Marchiori, Rapisarda: *Error and attack tolerance of complex networks*. Physica A 340, 2004]

Connected part of a Gnutella network: 1771 peers



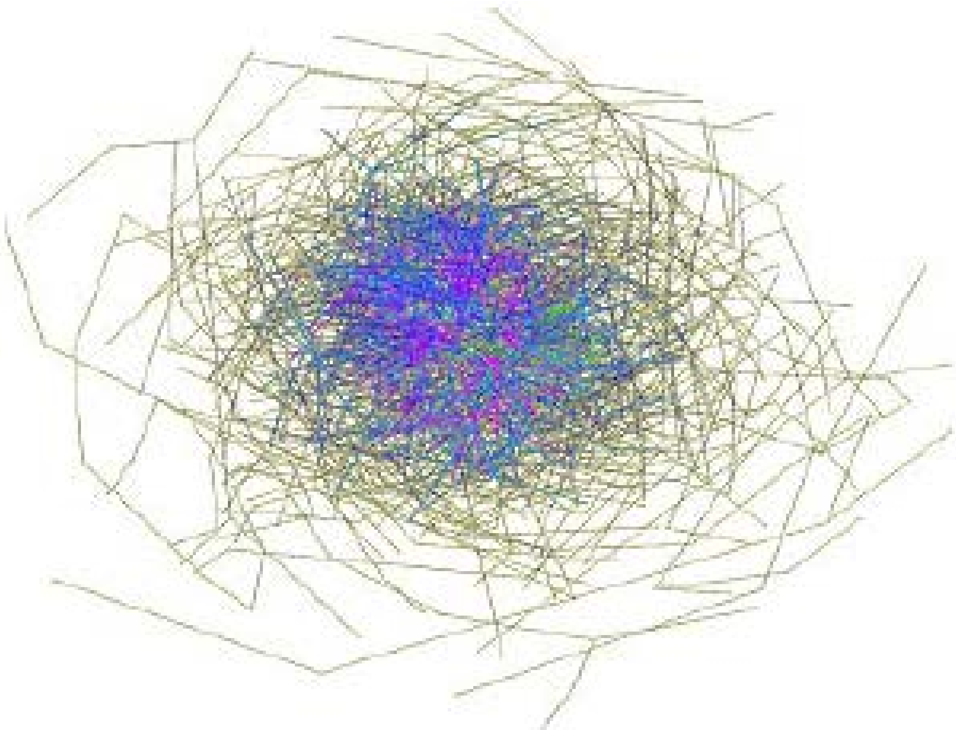
30% of nodes removed randomly: Largest connected component has 1106 nodes



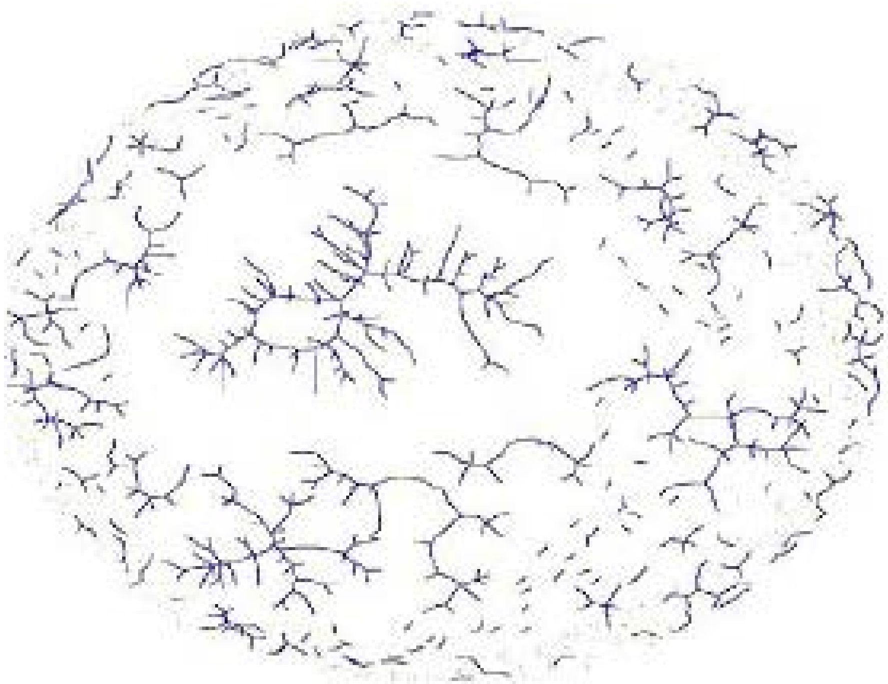
[Source: Saroiu, Gummadi, Gribble:
A Measurement Study of Peer-to-Peer File Sharing Systems.
In Proceedings of Multimedia Computing and Networking, 2002]

Robustness of a Gnutella Network (2)

Connected part of a Gnutella network: 1771 peers



Removed 4% of highest degree nodes



[Source: Saroiu, Gummadi, Gribble:
A Measurement Study of Peer-to-Peer File Sharing Systems.
In Proceedings of Multimedia Computing and Networking, 2002]

- Preferential attachment and network growth of a P2P system may lead to a scale-free network
- For instance, Gnutella and Freenet show scale-free properties
- Kleinberg-like P2P systems usually do not have scale-free properties since long-range contact selection does not involve node degree
- If a P2P system is a scale-free network, it is
  - robust to random failures
  - vulnerable to attacks of high-degree nodes