Networks II

Modelling Complex Systems

Some of this lecture is adapted from:
Albert and Barabasi, Reviews of Modern Physics 74 (2002)
M. Barthelemy, Physics Reports 499 (2011)
Newman, Networks (2018) - ebook available Uppsala University Library
-previous slides of David Sumpter.

Modelling Networks with (random) graphs

- Lattice graphs
- Erdos-Renyi random graph/Binomial random graph
- Chung-Lu random graph
- Configuration model
- Preferential attachment model
- Geometric random graph
- Random hyperbolic graph/KPKVB model

How well does the behaviour of each model replicate that in real networks?

Recap-
Five (of many) network measures

- Average degree
- Degree distribution
- Mean path length
- Clustering coefficient *
- Maximum modularity/
  Community partitions

What values do these take in real networks?

Real networks

TABLE I. The general characteristics of several real networks. For each network we have indicated the number of nodes, the average degree $\langle k \rangle$, the average path length ℓ, and the clustering coefficient C. For a comparison we have included the average path length ℓ_{rand} and clustering coefficient C_{rand} of a random graph of the same size and average degree. The numbers in the last column are keyed to the symbols in Figs. 8 and 9.

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 and Strogatz, 1998	3
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman, 2001a, 2001b, 2001c	4
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman, 2001a, 2001b, 2001c	5
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003	Newman, 2001a, 2001b, 2001c	6
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman, 2001a, 2001b, 2001c	7
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.*, 2001	8
Neurosci. co-authorship	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 and Fell, 2000	10
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner and Fell, 2000	11
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya and Solé, 2000	12
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya and Solé, 2000	13
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001	Ferrer i Cancho and Solé, 2001	14
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.*, 2001b	15
Power grid	4941	2.67	18.7	12.4	0.08	0.005	Watts and Strogatz, 1998	16
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts and Strogatz, 1998	17

Degree and average degree

The in in and out degrees are

$$k_i^{in} = \sum_{j=1} A_{ij} \qquad k_i^{out} = \sum_{i=1} A_{ij}$$

The average degree is

$$c = \frac{1}{n} \sum_{i,j} A_{ij}$$

same for in and out degree

Degree distribution

How many people follow you on Twitter.

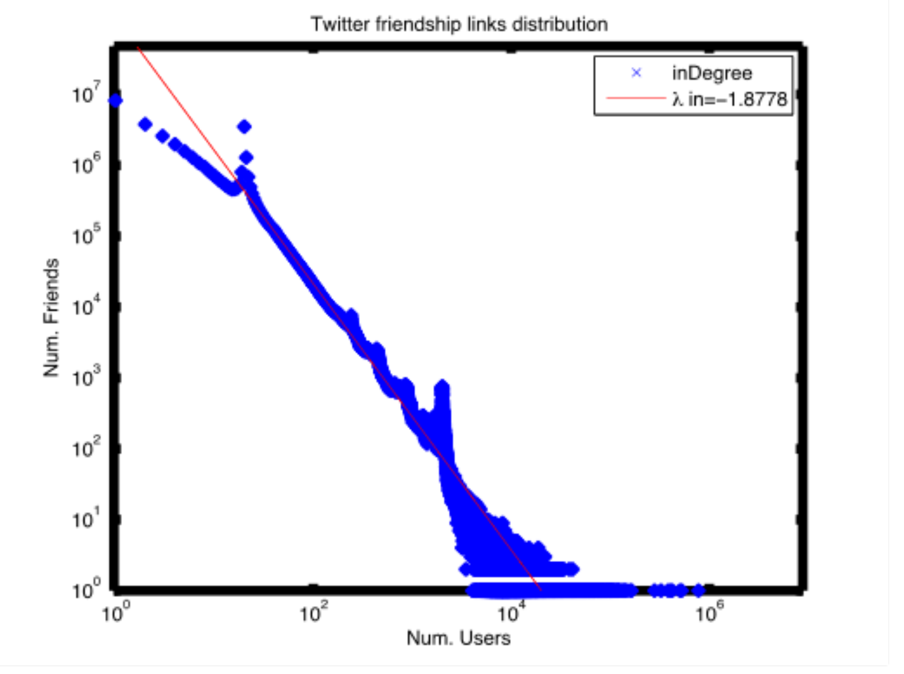


Figure 2. Incoming degree distribution of Twitter's network. As the figure shows, there are a few users with an enormous degree (number of followers). On the contrary, the majority of them have less than 100 followers.

Degree distribution $p(k)$ tells us how the connectedness varies between nodes

Degree distribution

How many people you follow on Twitter.

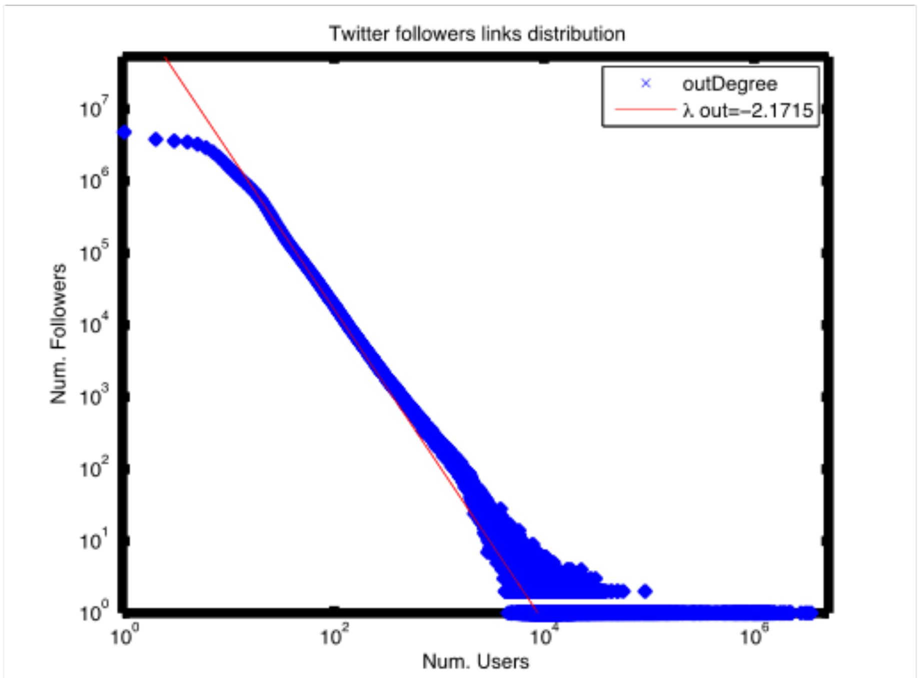


Figure 1. Outgoing degree distribution of Twitter's network. As the figure shows, there are a few users with an enormous degree (number of friends). On the contrary, the majority of them have just at most 1000 friends.

Degree distribution $p(k)$ tells us how the connectedness varies between nodes

Degree distribution

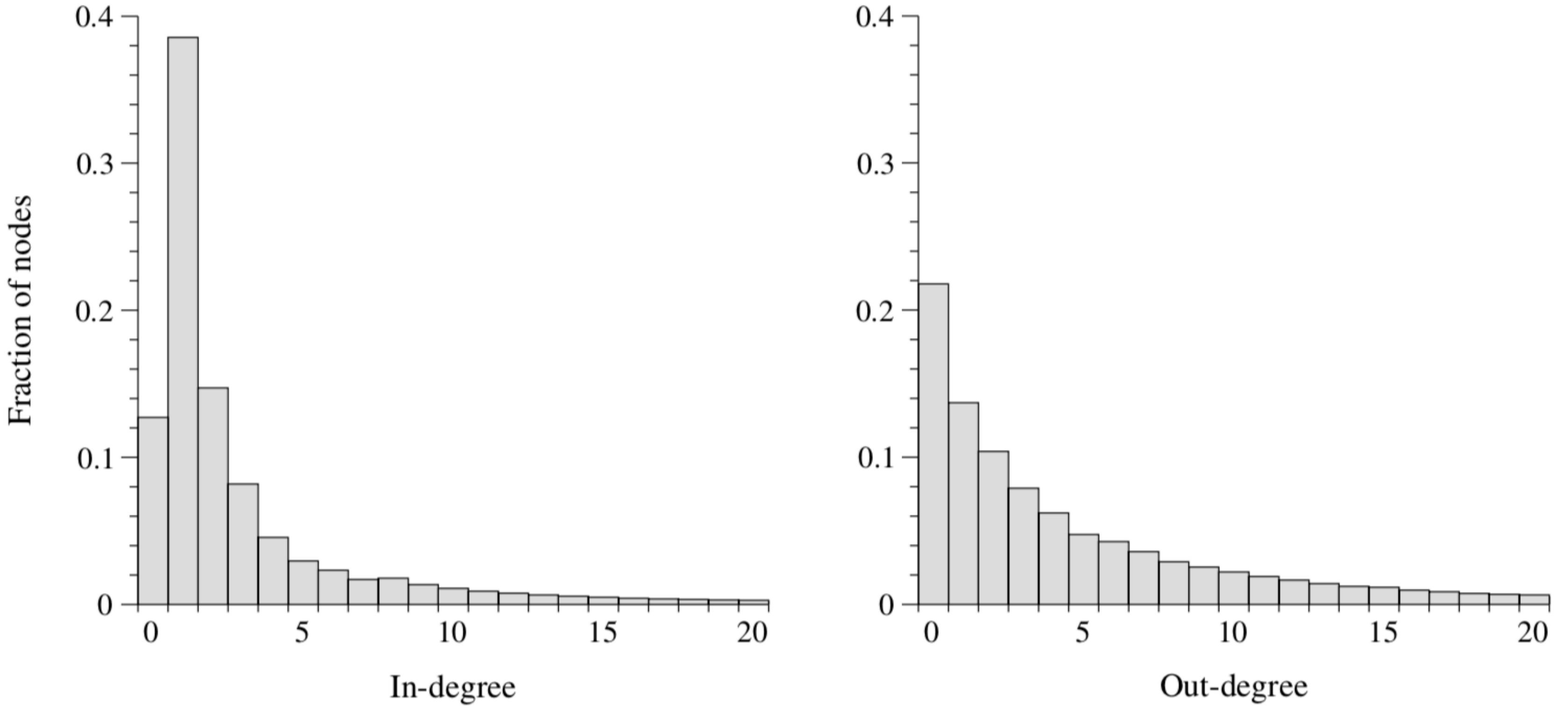


Figure 10.4: The degree distributions of the World Wide Web. Histograms of the distributions of in- and out-degrees of pages on the World Wide Web. Data are from the study by Broder *et al.* [84].

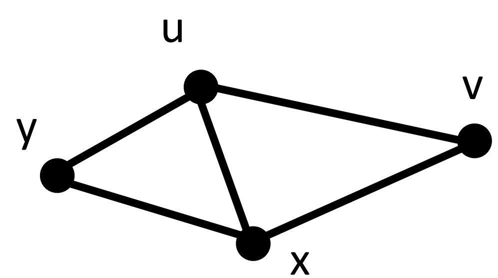
Degree distribution power law - $p(k)$ = k^{-\alpha}

Newman 'Networks' 2018

Network Measures

Clustering coefficient

- $$C = \frac{6 \times \text{number of triangles}}{\text{number of paths of length two}}$$



Graph has 2 triangles and 16 paths of length two.

C = 12 / 16 = 3 / 4

= probability that nodes **a** and **b** are connected if both have a common neighbour **c**

- High in social networks. You are friends with your friends friends.

Modelling Networks with (random) graphs

- Lattice graphs
- Erdos-Renyi random graph/Binomial random graph
- Chung-Lu random graph
- Configuration model
- Preferential attachment model
- Geometric random graph
- Random hyperbolic graph/KPKVB model

How well does the behaviour of each model replicate that in real networks?

Lattice networks

- All internal nodes have the same degree
- High C (~ constant)
- High mean path length (increases as $n^{1/d}$)

Erdös-Rényi Random graph

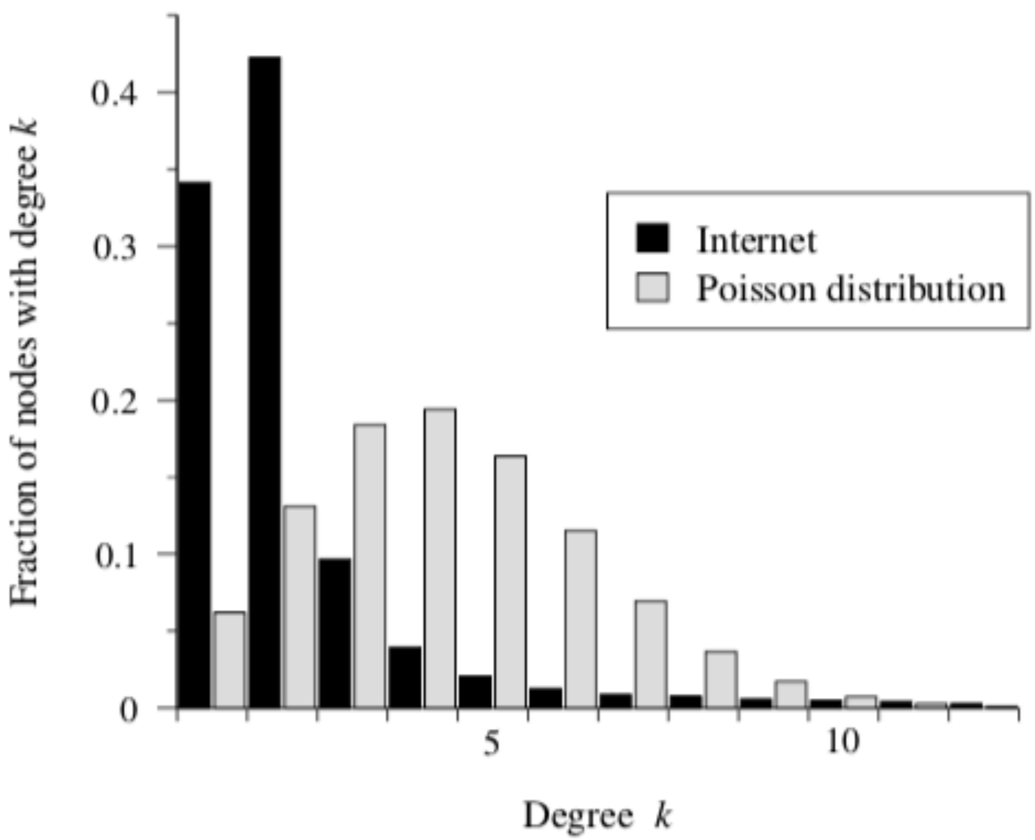
Every pair of nodes i,j is connected with probability p*. Total of n* nodes

- Binomial degree distribution, c = $p(n\text{-}1)$
- Low C = c/n
- Low mean path length l ~ log(n)

Random graph process
Start with n vertices with 0 edges. Each step add a missing edge.

Erdös-Rényi Random graph

- Degree distribution ….



Erdös-Rényi Random graph

- Not a realistic model but good toy model

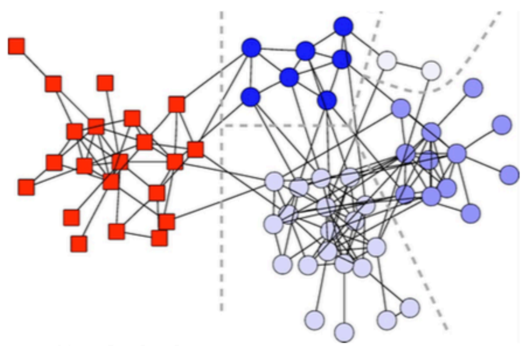
- Serves as a null model

A differentiation between graphs which are truly modular and those which are not can ... only be made if we gain an understanding of the intrinsic modularity of random graphs.

-- Reichardt and Bornholdt

Erdös-Rényi Random graph

- Serves as a null model

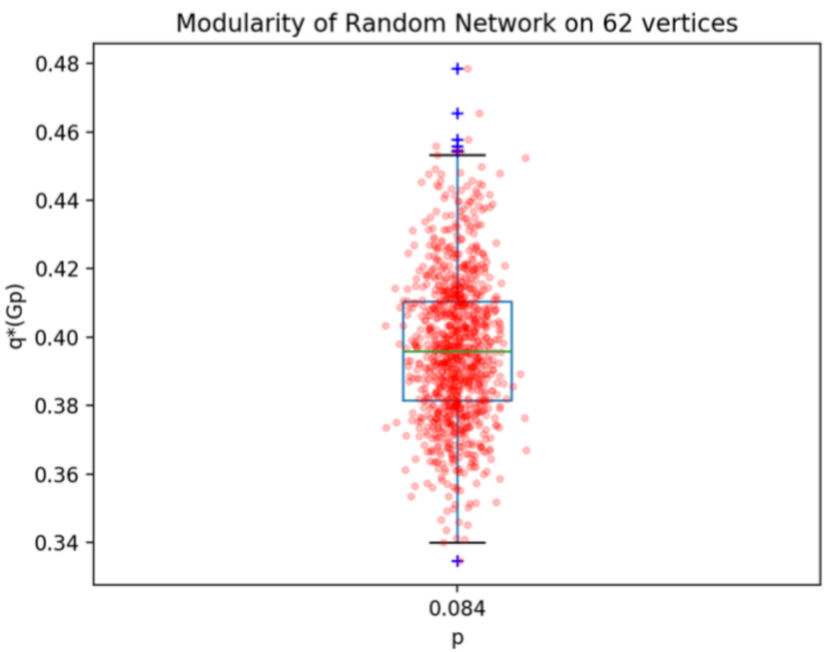


$\#$ dolphins $=62$
$\#$ edges $=159$

$q^* = 0.52$

8.4% of possible edges

$q^*(\text{dolphins}) > q^*(\text{random network})??$



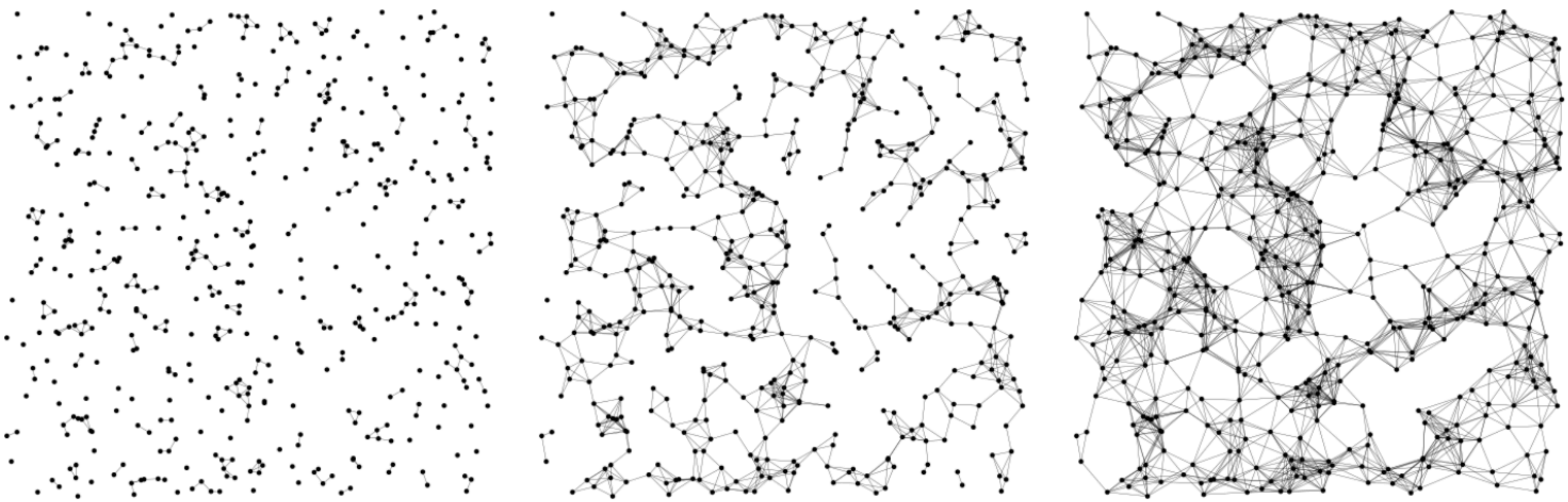
Configuration Model

Start with degree sequence d_1, … d_n
Place d_i half edges on each node
Choose a random matching of half edges

Serves as a null model.

Random Geometric Graph

- Place n points uniformly. Join any two vertices with distance less than r.



500 points. r=0.03, r=0.06, r=0.09

Tobias Muller

KPKVB model - random hyperbolic graph

- Hyperbolic plane curvature -alpha^2

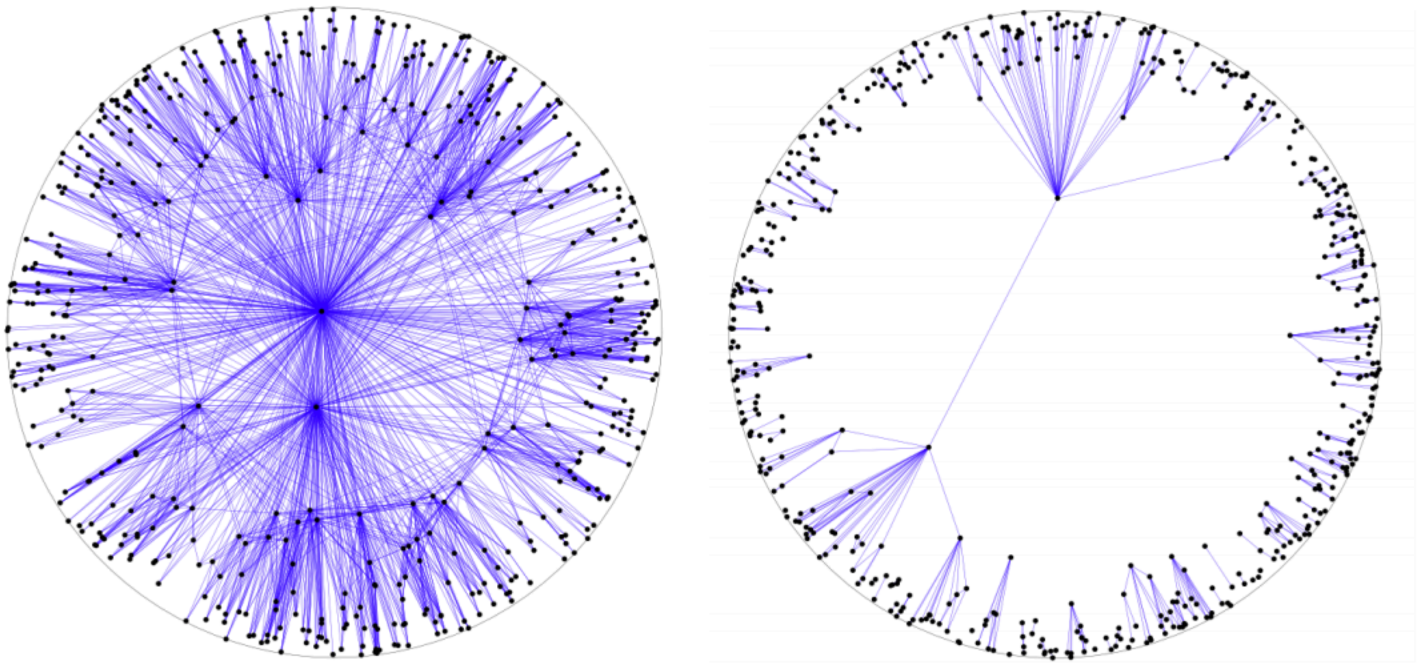


Figure 1: The random graph $G(N; \alpha, \nu)$ with $N = 500$ vertices, $\nu = 2$ and $\alpha = 0.7$ and $3/2$.

Müller and Fountoulakis, Law of large numbers for the largest component in a hyperbolic model of complex networks, 2018

- KPKVB model - random hyperbolic graph
  - Krioukov-Papadopoulos-Kitsak-Vahdat-Boguñá
  - Power law degree distribution
  - Clustering coefficient
  - Hard to prove results in this model

Müller and Fountoulakis, Law of large numbers for the largest component in a hyperbolic model of complex networks, 2018

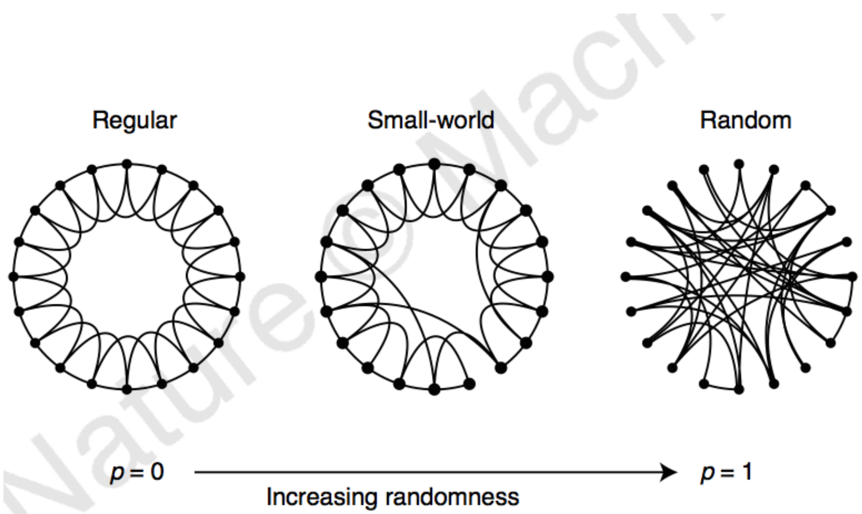
Real networks

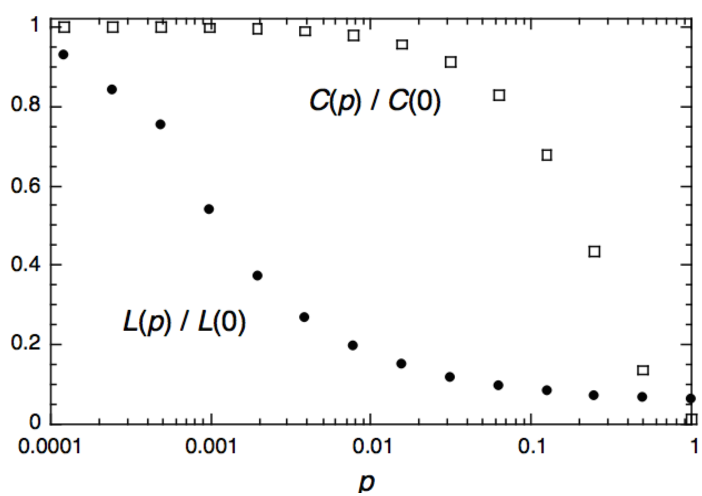
TABLE I. The general characteristics of several real networks. For each network we have indicated the number of nodes, the average degree $\langle k \rangle$, the average path length ℓ, and the clustering coefficient C. For a comparison we have included the average path length ℓ_{rand} and clustering coefficient C_{rand} of a random graph of the same size and average degree. The numbers in the last column are keyed to the symbols in Figs. 8 and 9.

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 and Strogatz, 1998	3
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman, 2001a, 2001b, 2001c	4
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman, 2001a, 2001b, 2001c	5
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003	Newman, 2001a, 2001b, 2001c	6
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman, 2001a, 2001b, 2001c	7
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.*, 2001	8
Neurosci. co-authorship	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 and Fell, 2000	10
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner and Fell, 2000	11
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya and Solé, 2000	12
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya and Solé, 2000	13
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001	Ferrer i Cancho and Solé, 2001	14
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.*, 2001b	15
Power grid	4941	2.67	18.7	12.4	0.08	0.005	Watts and Strogatz, 1998	16
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts and Strogatz, 1998	17

Small world network

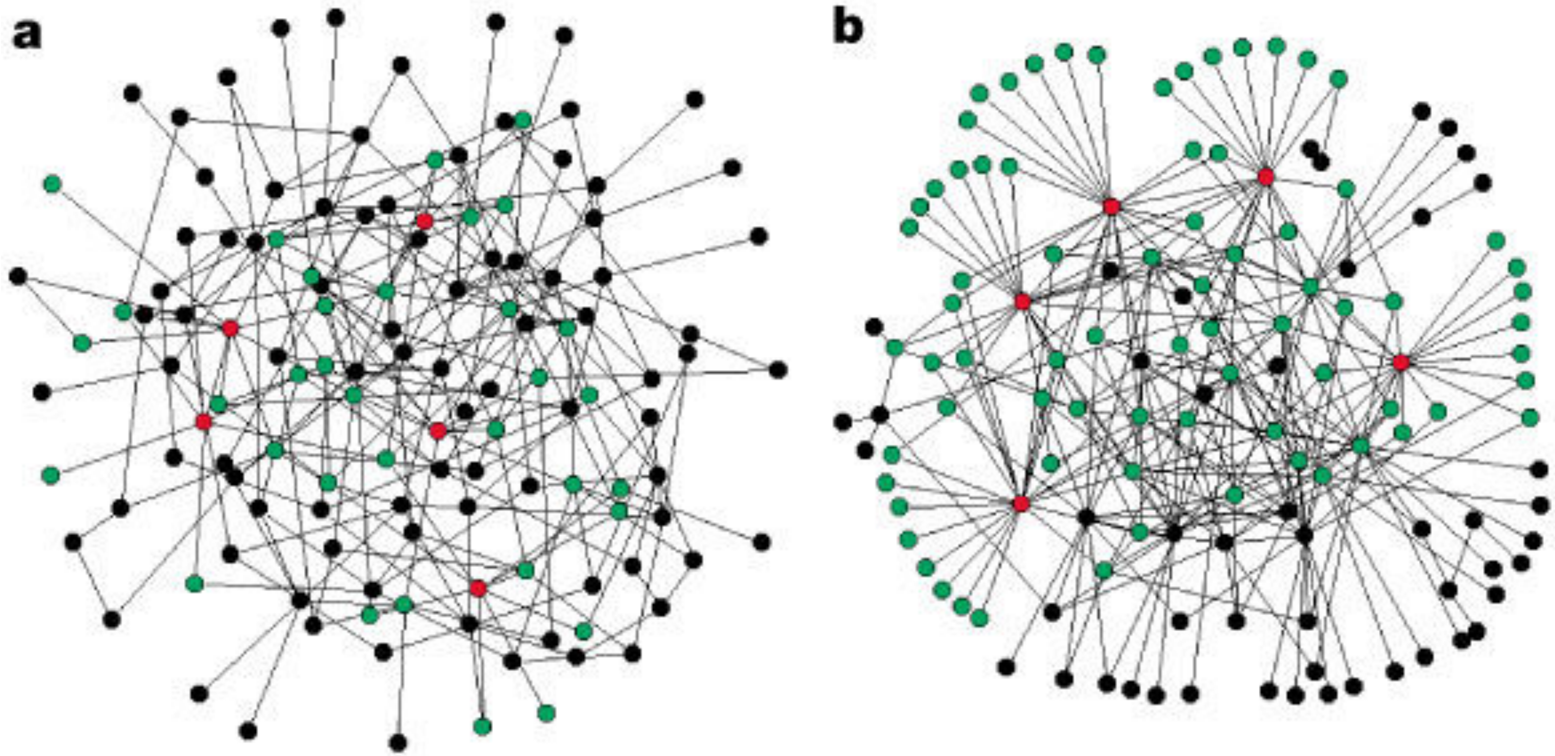
- Watts & Strogatz model interpolates between a structured and random network
- Low diameter + high clustering = small world





Watts and Strogatz, *Nature* 393 (1998)

Power law network



Exponential

Scale-free

Power law network

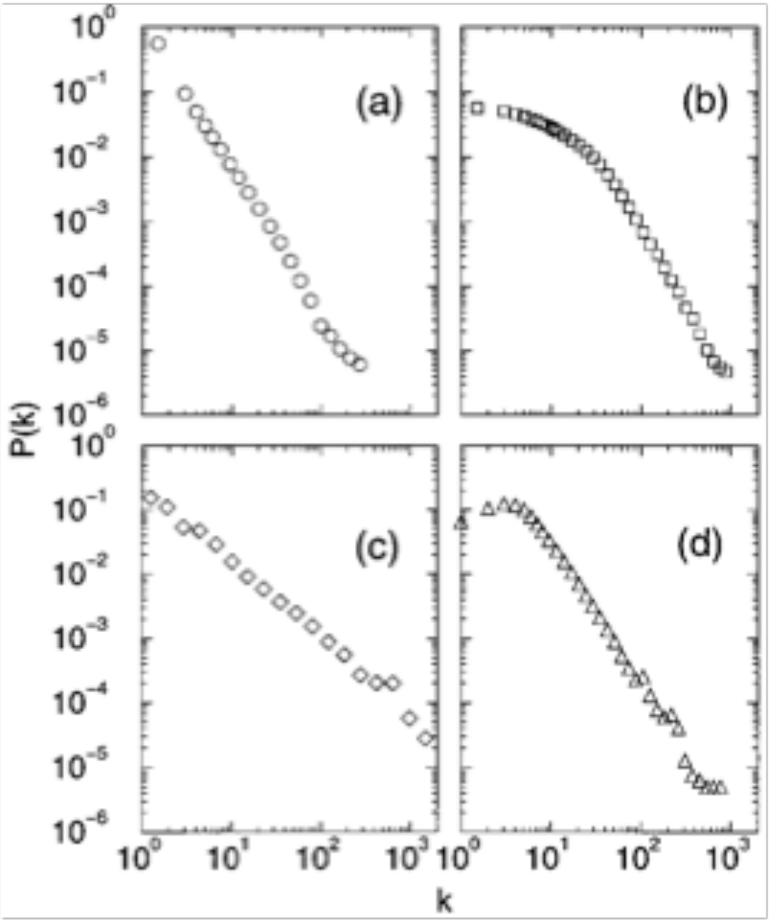
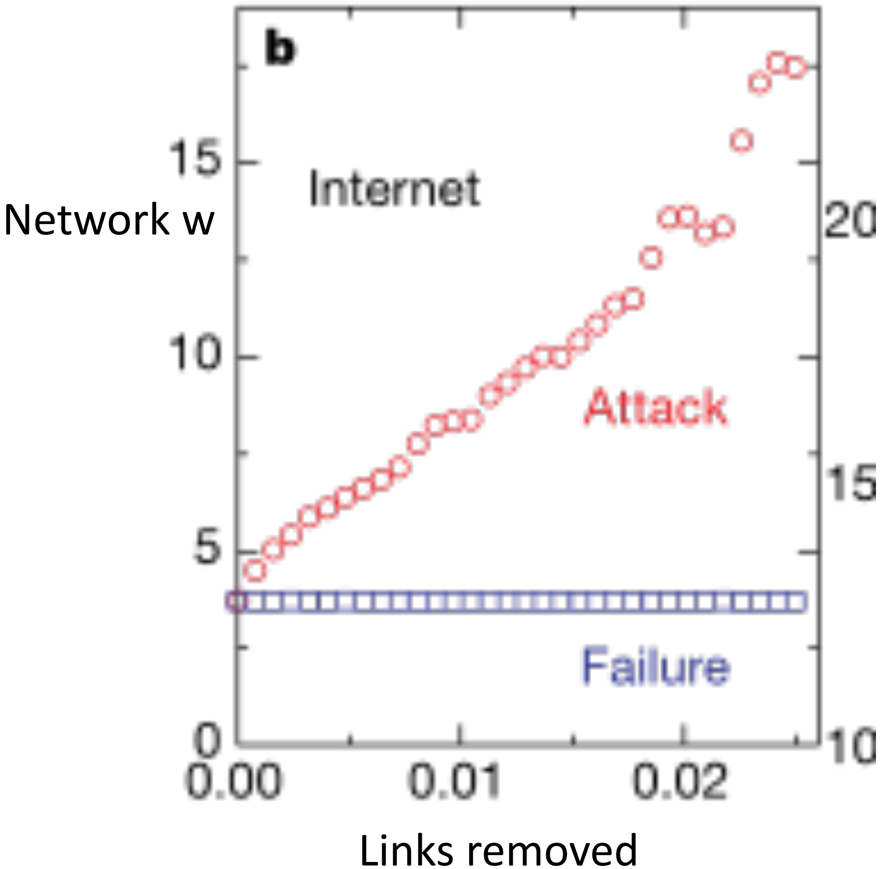


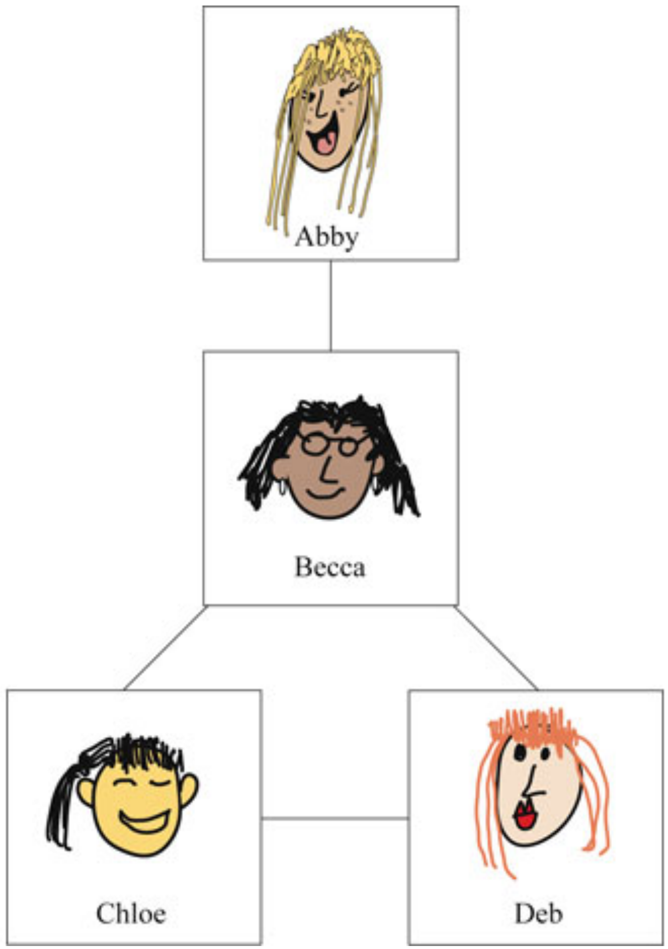
FIG. 3. The degree distribution of several real networks: (a) Internet at the router level. Data courtesy of Ramesh Govindan; (b) movie actor collaboration network. After Barabási and Albert 1999. Note that if TV series are included as well, which aggregate a large number of actors, an exponential cut-off emerges for large *k* (Amaral *et al.*, 2000); (c) co-authorship network of high-energy physicists. After Newman (2001a, 2001b); (d) co-authorship network of neuroscientists. After Barabási *et al.* (2001).

Power law network



- Robust to random failures.
- Susceptible to attack

Friendship Paradox



https://opinionator.blogs.nytimes.com/2012/09/17/friends-you-can-count-on/

Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media

Friendship Paradox Redux: Your Friends Are More Interesting Than You

Nathan O. Hodas
USC Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292
nhodas@isi.edu

Farshad Kooti
USC Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292
kooti@usc.edu

Kristina Lerman
USC Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292
lerman@isi.edu

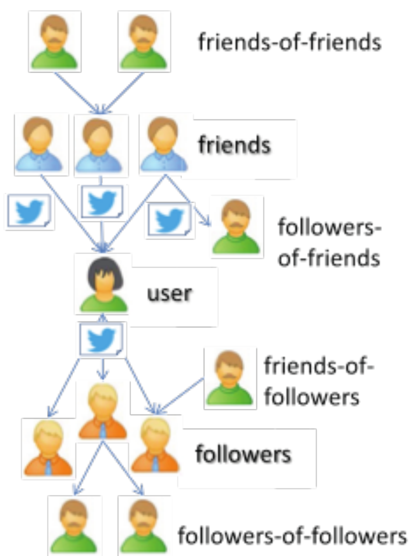
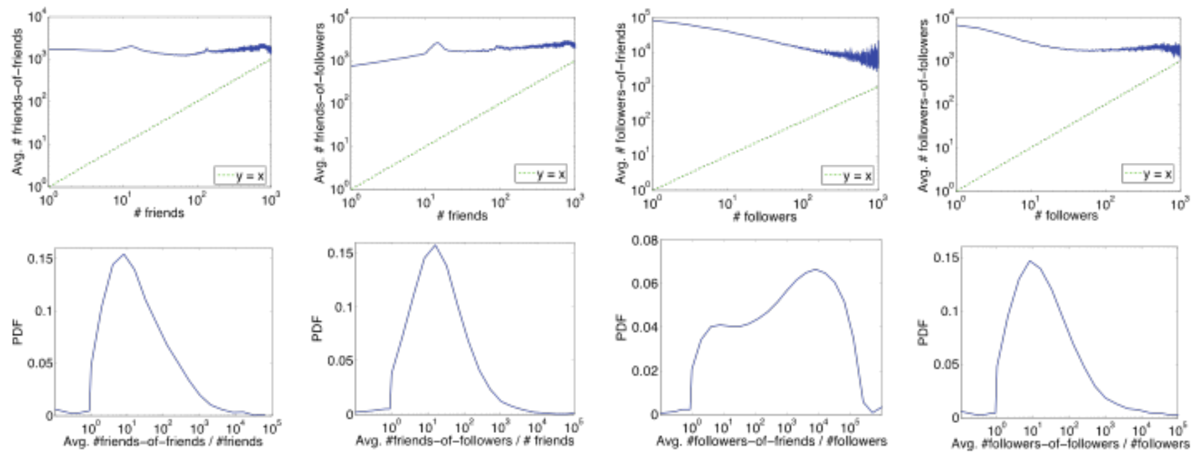


Figure 1: *An example of a directed network of a social media site with information flow links. Users receive information from their friends and broadcast information to their followers.*



(i) (ii) (iii) (iv)