Scale-free Networks

Introduction to Network Science

Instructor: Michele Starnini — https://github.com/chatox/networks-science-course

Universitat Pompeu Fabra Barcelona

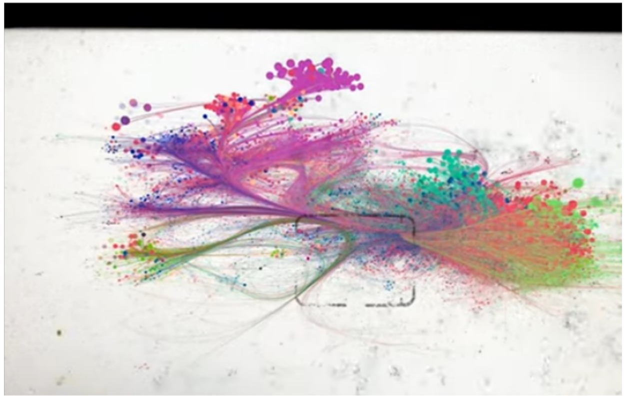
Contents

- Characteristics of scale-free networks
- Degree distribution of scale-free networks
- Distance distribution of scale-free networks

Observed degree distributions

Video (03:15-04:42) by Albert-László Barabási (cont.)





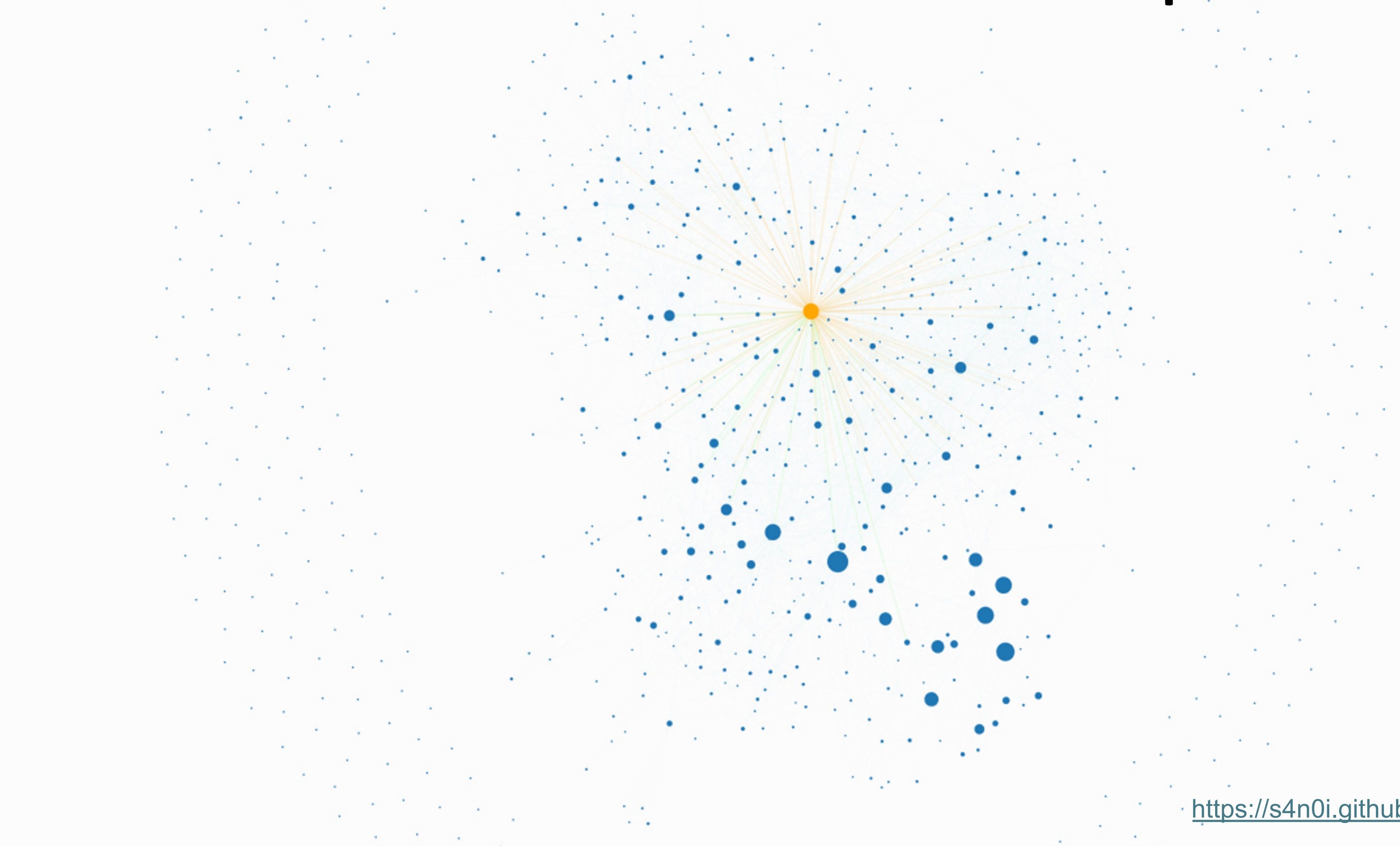
https://www.youtube.com/watch?v=RfgjHoVCZwU
From “After years … I realized I needed to find real data ...”
To “Everything is possible; they are scale-free.”

Pareto's Law

- Italian economist Vilfredo Pareto in the 19th century noted 80% of money was earned by 20% of people

- More recently ...
  - 80 percent of links on the Web point to only 15 percent of pages;
  - 80 percent of citations go to only 38 percent of scientists;
  - 80 percent of links in Hollywood are to 30 percent of actors
- A debate that is still open: the wealth of the 1% and the 0.1%

Immanuel Kant in Wikipedia



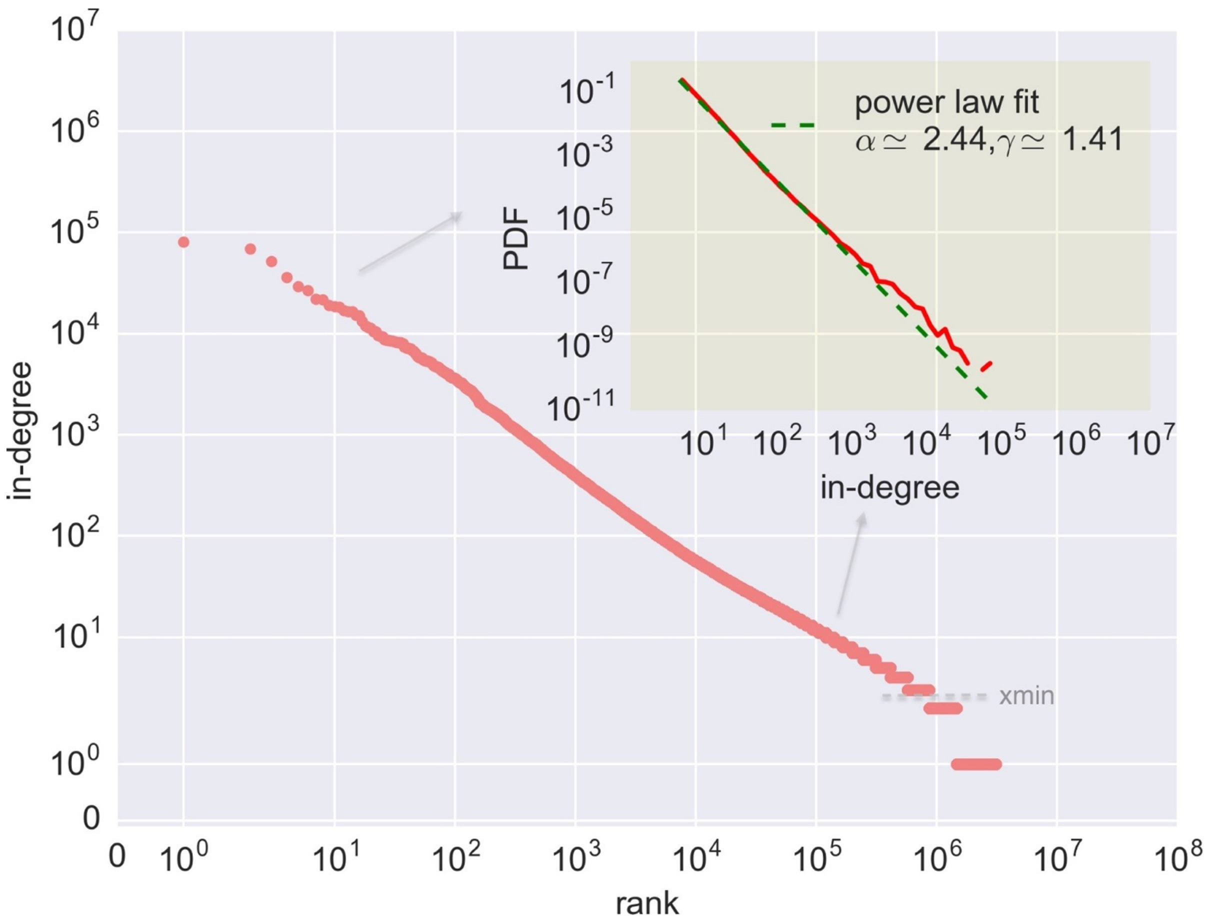
https://s4n0i.github.io/schoolofathens/

Wikipedia

2017; first link only

Article	In-degree
United States	~80 K
village	~70 K
moth	~52 K
Communes of France	~37 K
species	~29 K
...	...

In-degree of k-th page with most (first) in-links

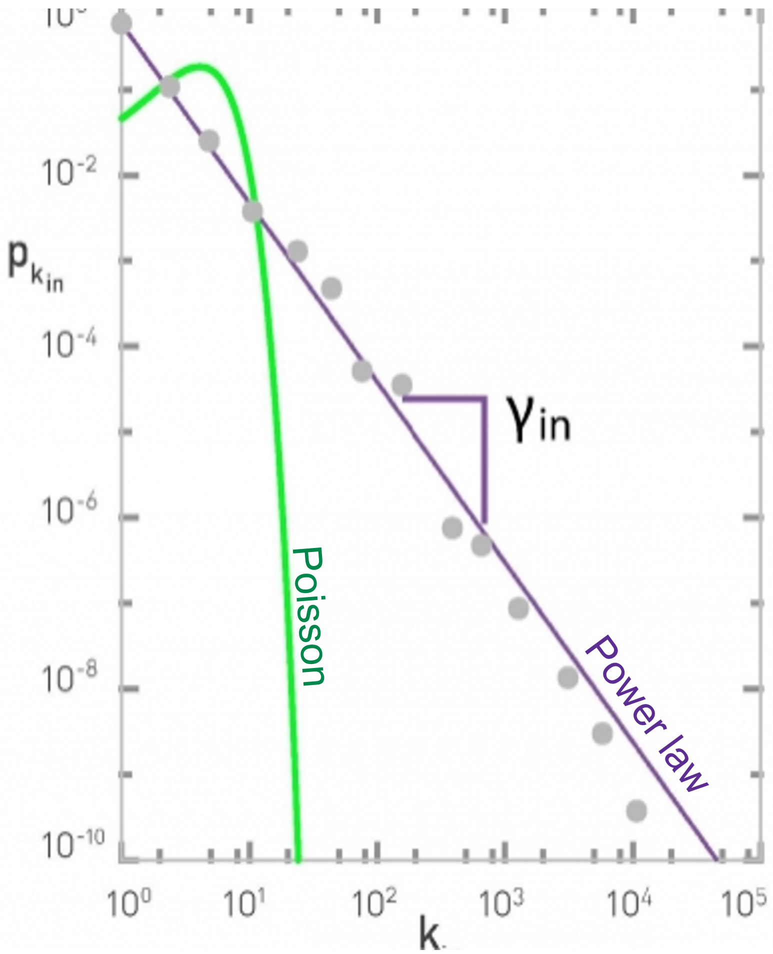


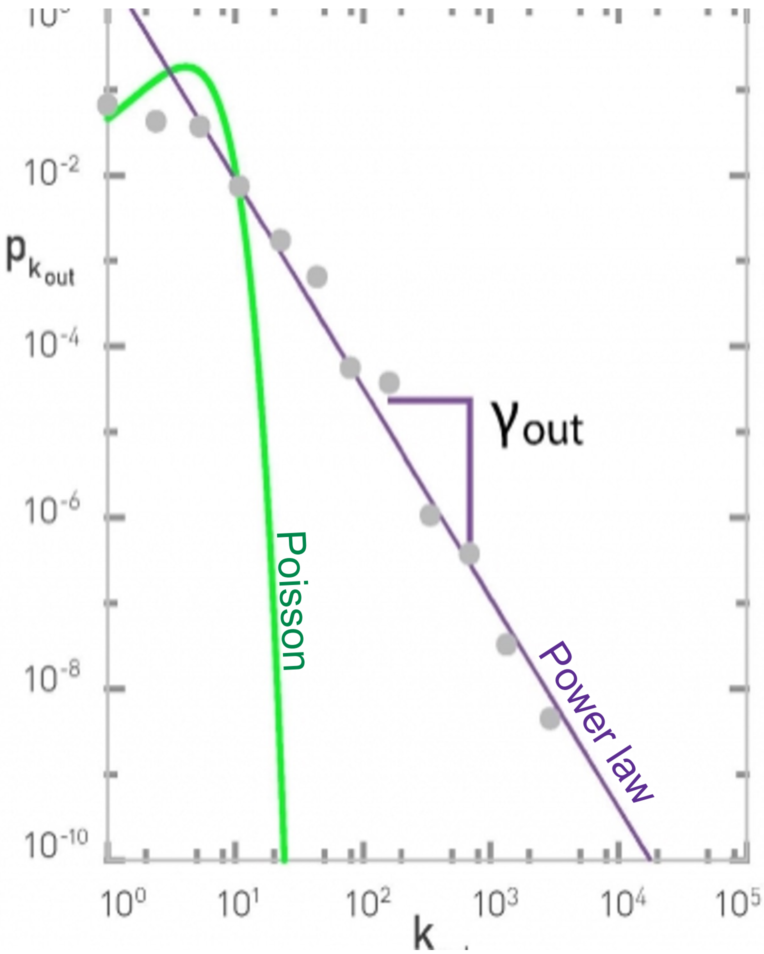
https://doi.org/10.1016/j.jocs.2016.12.001

Degree distributions in Web graphs

- Web graphs have large “hubs”:
  - This is, nodes with very large degree
- This does not happen in a random (ER) graph
- We have already seen the Poisson distribution is a bad approximation of the observed degree distribution

Degree distributions in a web graph [nd.edu]



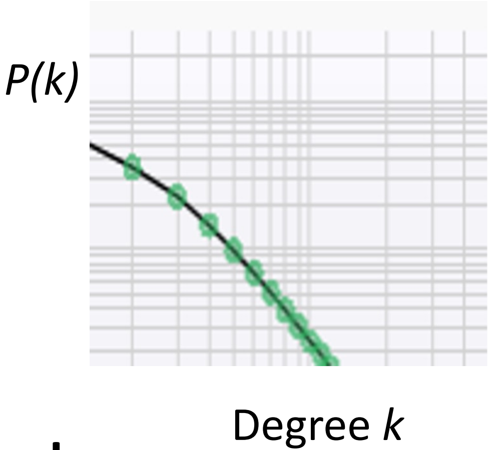


Modeling degree distributions

Degree distribution in real-world networks

- Straight descending line in log-log plot

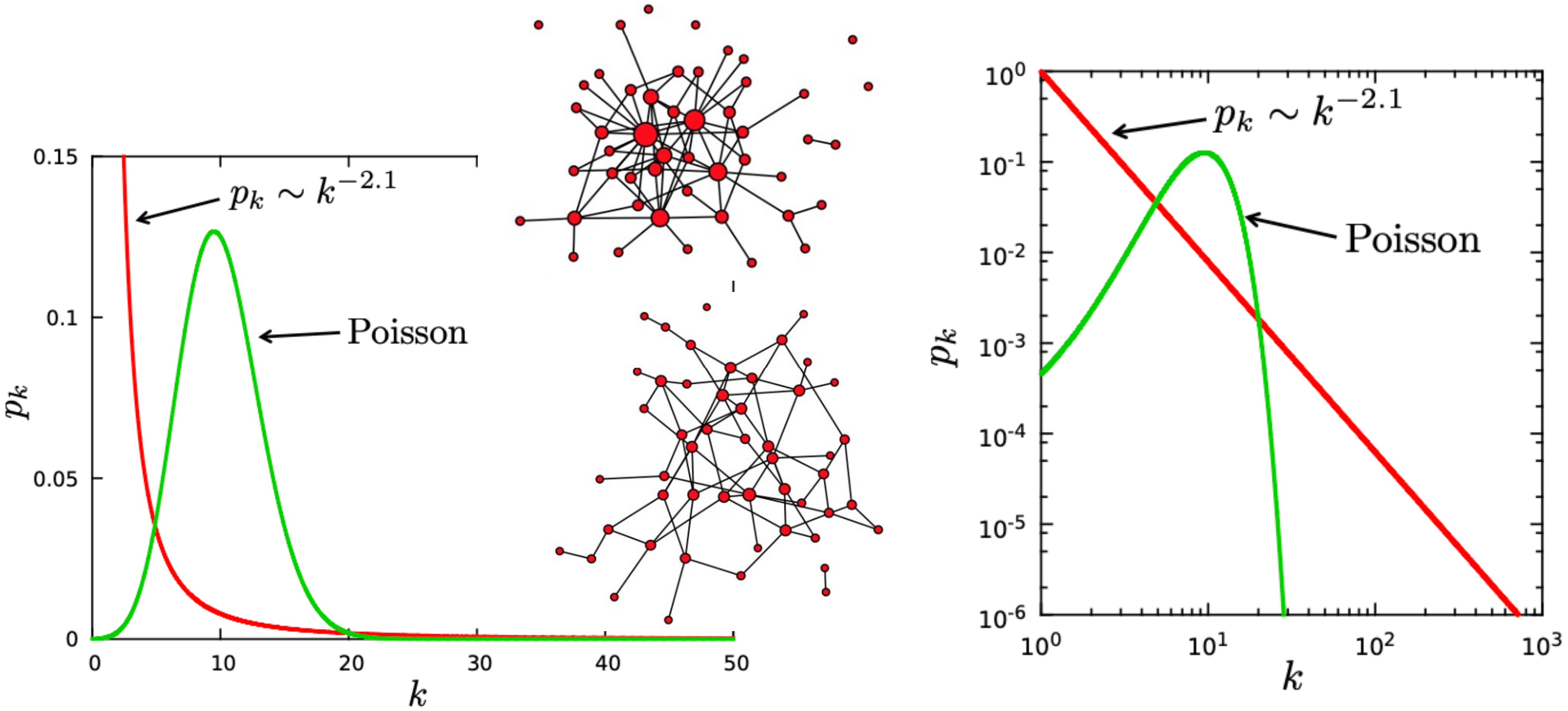
$$\log p_k \sim -\gamma \log k$$
$$p_k \sim k^{-\gamma}$$



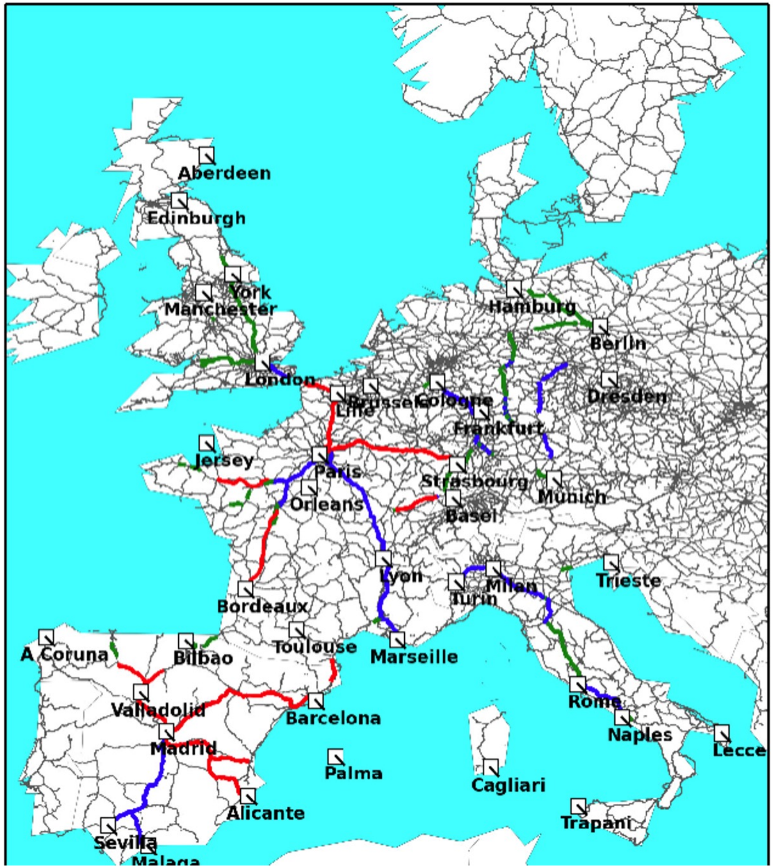
- Parameter ɣ is the exponent of the power law

- **A scale-free network is a network whose degree distribution follows a power law**

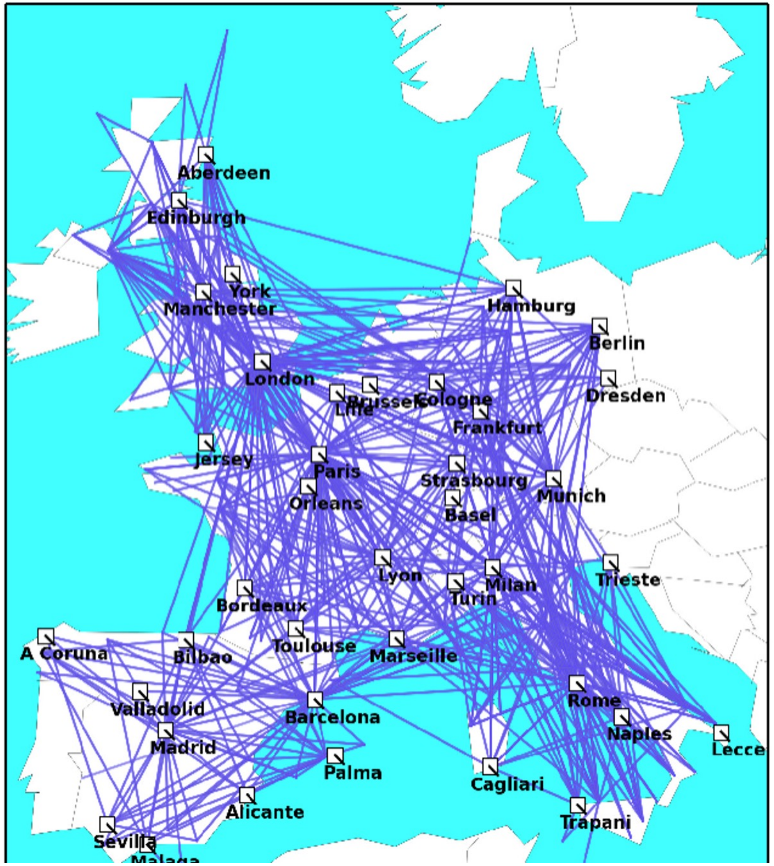
Comparing Poisson to power law



Random vs scale-free networks

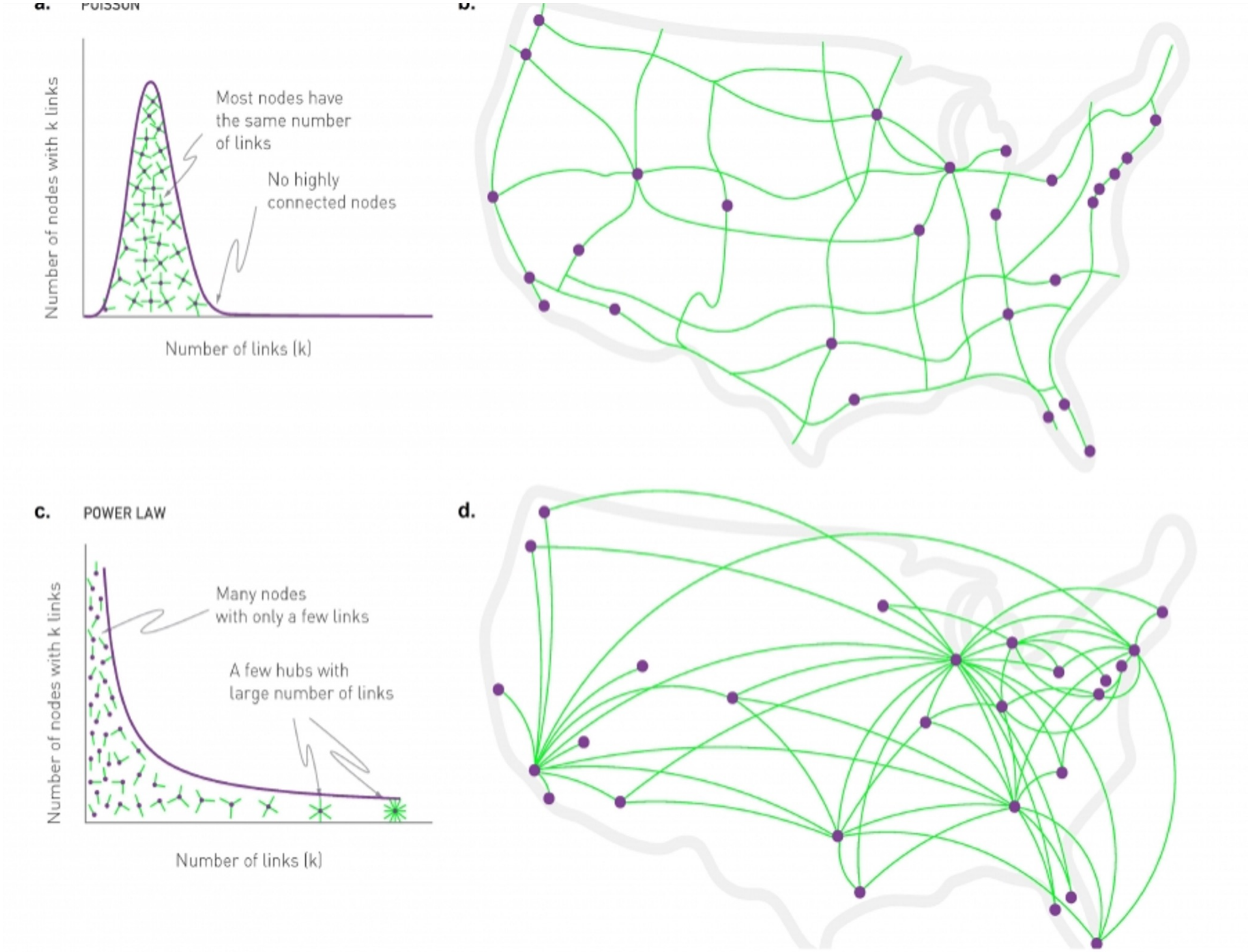


High-speed trains



Air transportation

Random vs scale-free networks (cont.)



Ground transportation

Air transportation

Formally (discrete)

$$p_k = Ck^{-\gamma}$$

$$\sum_{k=1}^{\infty} p_k = 1$$

$$C = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta(\gamma)}$$

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

Riemann's zeta

This formalism assumes there are no nodes with degree zero

Formally (continuous approx.)

$$p_k = Ck^{-\gamma}$$

$$\int_{k=k_{\min}}^{\infty} p_k = 1 \qquad C = \frac{1}{\int_{k=k_{\min}}^{\infty} k^{-\gamma}} = (\gamma - 1)k_{\min}^{\gamma-1}$$

$$p_k = (\gamma - 1)k_{\min}^{\gamma-1}k^{-\gamma}$$

k_{min} is the smaller degree found in the network

Exercise

In the actor network, N=702,388, γ=2.12

1. How many actors do we expect to have …

1 other co-star?

10 other co-stars?

100 other co-stars?

https://www.wolframalpha.com/
recognizes `x*y, x/y, Zeta(x), x^(-y)`, etc.

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)} \qquad \zeta(2.12) \approx 1.5452$$

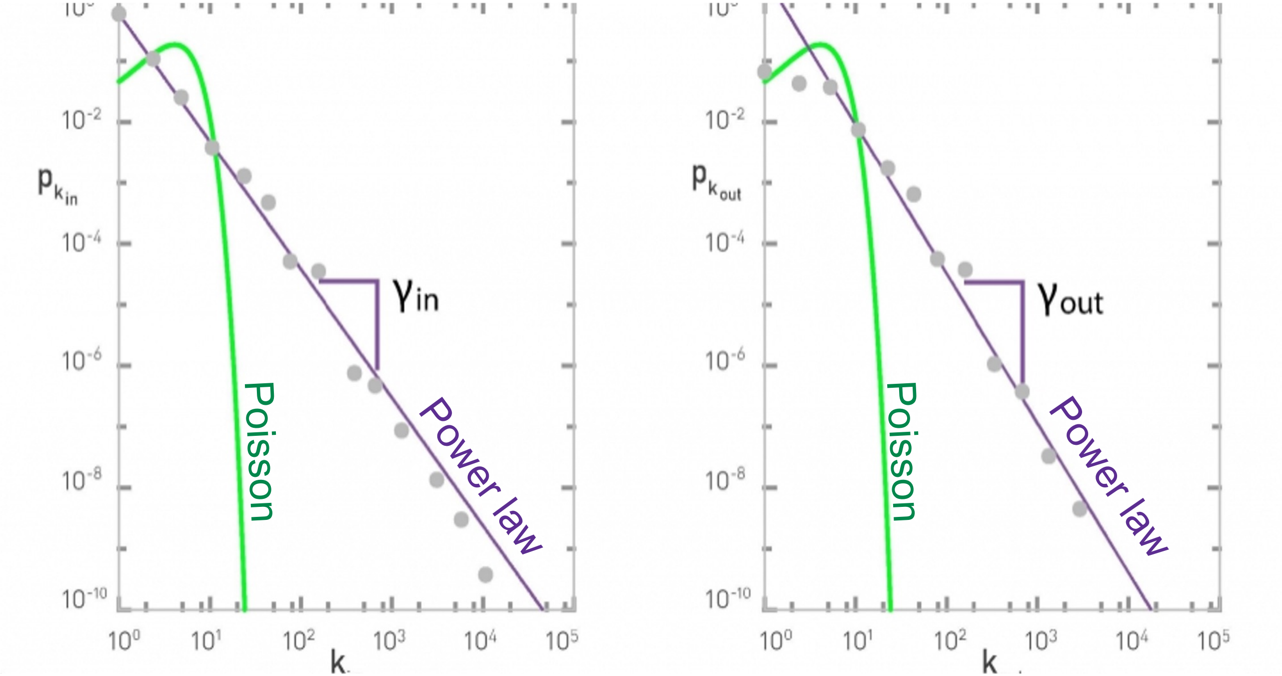
2. For how many co-stars do we still expect to have one actor that has that many co-stars?

Directed graphs

In-Degree vs Out-Degree

In TikTok or Instagram (or any online social networking site), which distribution is likely to have a larger variance, that of in-degree or out-degree?

In directed graphs the exponent might be different for in-degree and out-degree



$$p_k \sim k^{-\gamma}$$

What kind of values of gamma make the power law "tail" look more like the Poisson "tail"?

In directed networks ...

- Each node has two degrees: k_{in} and k_{out}
- In general they may differ, hence

$p_{k_{in}} \sim k^{-\gamma_{in}}$

$p_{k_{out}} \sim k^{-\gamma_{out}}$

- In nd1998, $\gamma_{in} \simeq 2.1$, $\gamma_{out} \simeq 2.4$

Dispersion of the degree distribution

What does it mean “scale-free”?

- A distribution **has a “scale”** if values are close to each other, for instance in a ER network $\sigma_k = \langle k \rangle^{1/2}$
- Hence, most nodes are in the range $\langle k \rangle \pm \langle k \rangle^{1/2}$
- However in scale-free network with $\gamma < 3$

$$\sigma_k \to \infty$$

Moments of degree distribution

- Moments of degree distribution

$$\langle k^n \rangle = \int_{k_{\min}}^{k_{\max}} k^n p_k dk = C \frac{k_{\max}^{n-\gamma+1} - k_{\min}^{n-\gamma+1}}{n-\gamma+1}$$

$$C = (\gamma - 1) k_{\min}^{\gamma-1}$$

Variance in SF networks

$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2$$

- In a SF network $\langle k^2 \rangle = \int_{k_{\min}}^{k_{\max}} k^2 p_k dk = C \frac{k_{\max}^{3-\gamma} - k_{\min}^{3-\gamma}}{3-\gamma}$

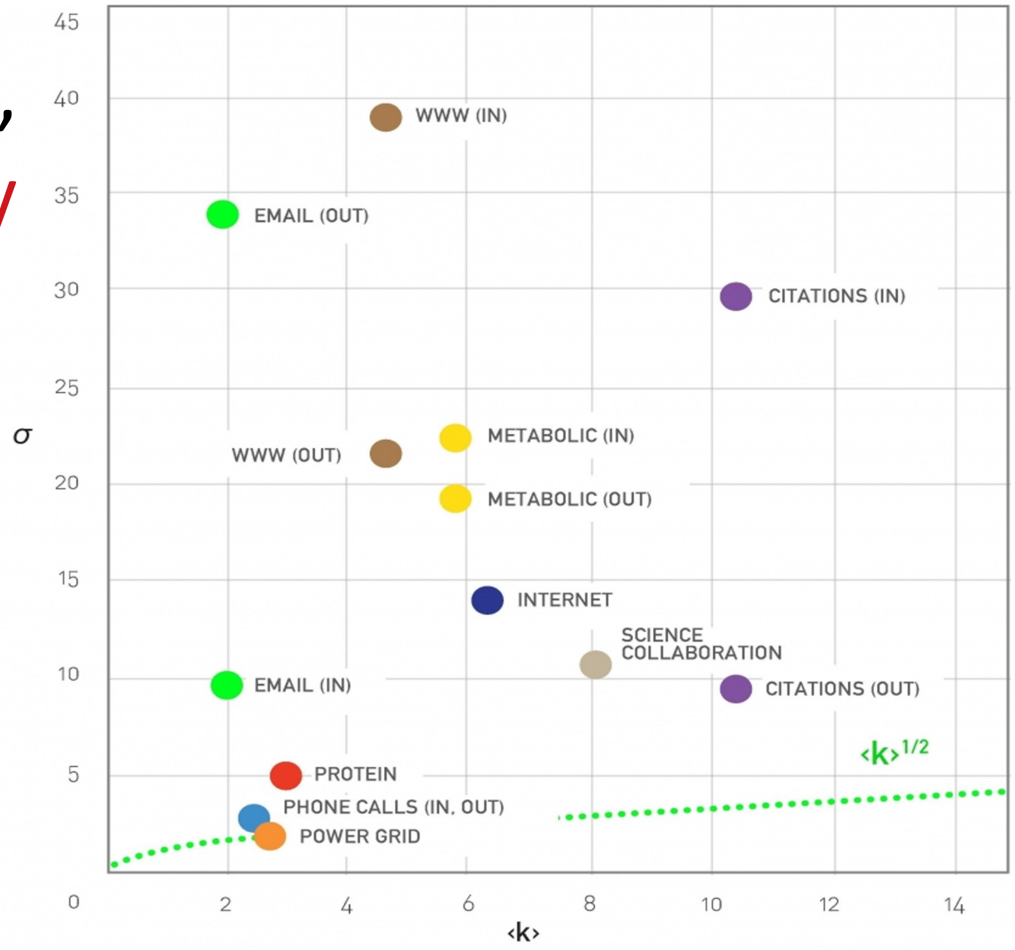
- This diverges as $k_{\max} \to \infty$ if $\gamma < 3$

- Hence there is no "typical" scale

Note that the average degree <k> diverges if $\gamma < 2$

Real network are SF

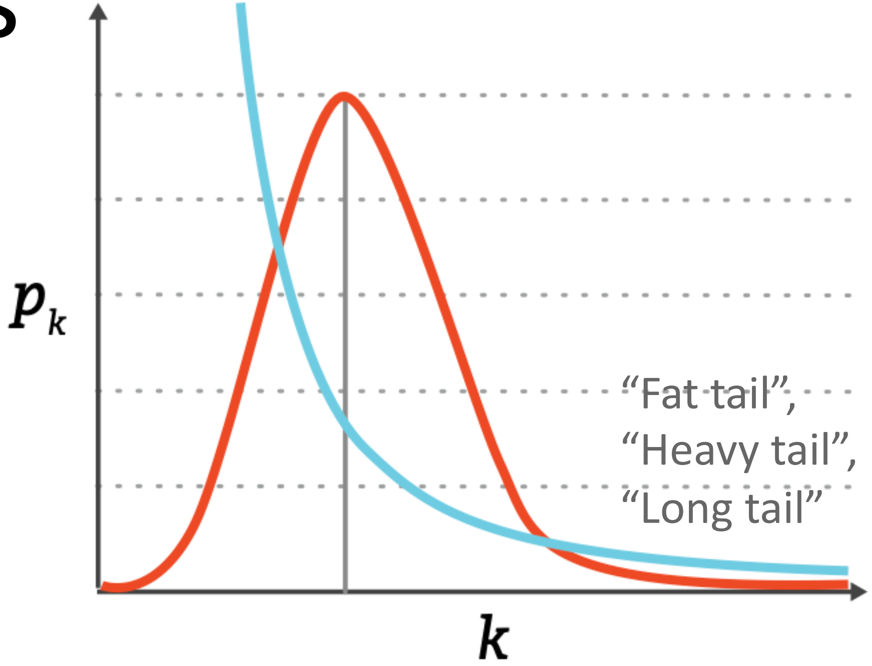
Real networks are finite,
variance cannot be infinity



Citation networks

$$k_{\text{in}} \approx 10 \pm 900$$

In general, the average degree is not very informative in scale-free networks



Random network
Randomly chosen node: $k = \langle k \rangle \pm \langle k \rangle^{1/2}$
Scale: $\langle k \rangle$

Scale-free network
Randomly chosen node: $k = \langle k \rangle \pm \infty$
$\langle k \rangle$ is meaningless as 'scale'

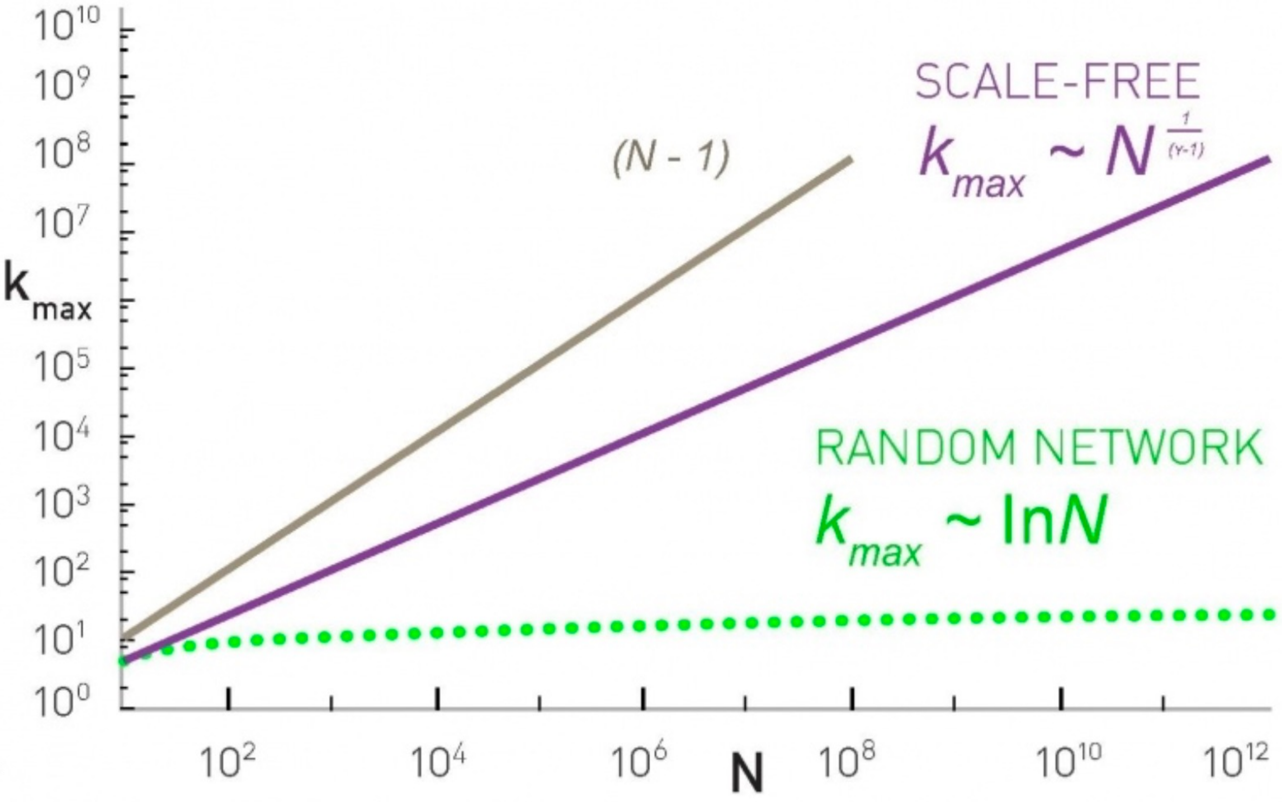
There is a natural cut-off of the degree

Prob. of observing only one large hub with k_{max} (that's why it's the max!):

$$\int_{k_{max}}^{\infty} p(k)dk = \frac{1}{N}.$$

$$p_k = Ck^{-\gamma}$$

$$k_{\max} = k_{\min} N^{\frac{1}{\gamma - 1}}$$

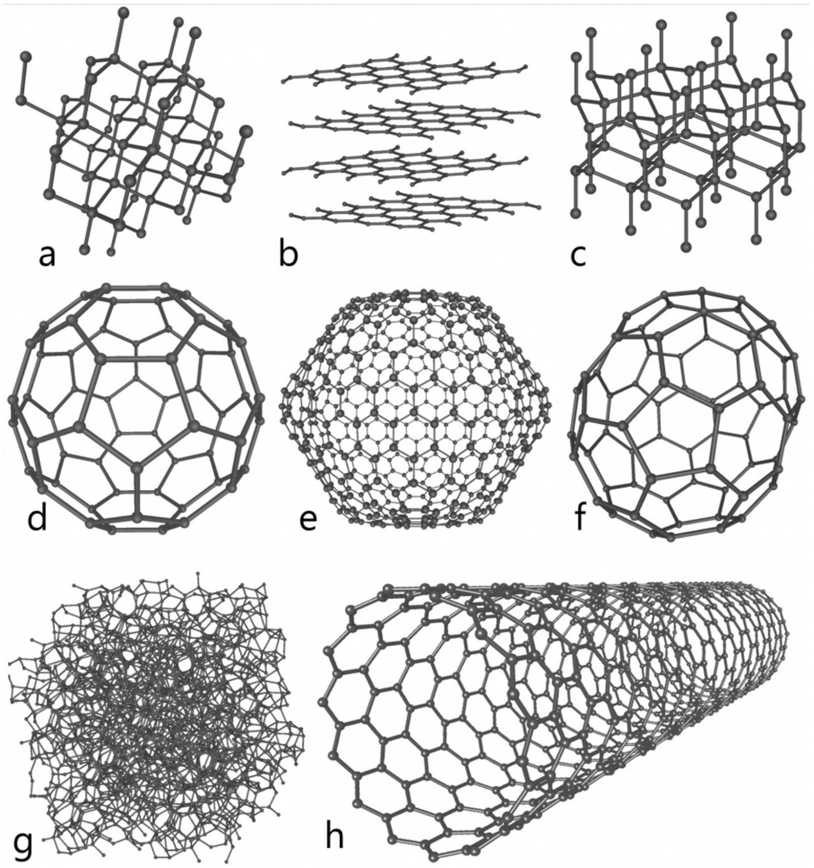


Real network examples

Network	N	L	$\langle k \rangle$	$\langle k_{in}^2 \rangle$	$\langle k_{out}^2 \rangle$	$\langle k^2 \rangle$	γ_{in}	γ_{out}	γ
Internet	192,244	609,066	6.34	-	-	240.1	-	-	3.42*
WWW	325,729	1,497,134	4.60	1546.0	482.4	-	2.00	2.31	-
Power Grid	4,941	6,594	2.67	-	-	10.3	-	-	Exp.
Mobile-Phone Calls	36,595	91,826	2.51	12.0	11.7	-	4.69*	5.01*	-
Email	57,194	103,731	1.81	94.7	1163.9	-	3.43*	2.03*	-
Science Collaboration	23,133	93,437	8.08	-	-	178.2	-	-	3.35*
Actor Network	702,388	29,397,908	83.71	-	-	47,353.7	-	-	2.12*
Citation Network	449,673	4,689,479	10.43	971.5	198.8	-	3.03*	4.00*	-
E. Coli Metabolism	1,039	5,802	5.58	535.7	396.7	-	2.43*	2.90*	-

Not all networks are SF!

- In general, when there is a limit to k_{max}
- Out-degree in some social networks
- Materials/crystals



Strictly speaking, these are not "complex" networks

Summary

Things to remember

- Power law
- Formulas for degree distribution
  - Discrete formula
  - Continuous formula
- Definition of scale-free

Practice on your own

- **(Somewhat) difficult, try to solve it ON YOUR OWN**
- Imagine a connected scale-free graph with 1 million nodes and average degree 5

If we draw 100 nodes from this graph, how many will have degree 1?

Remember, if the graph is connected, $k_{min}=1$

If you cannot clear the unknown in a formula, plot it

- Solution in next slide (shown only in .odp, not .pdf)

$$\langle k \rangle = C \frac{k_{\max}^{2-\gamma} - k_{\min}^{2-\gamma}}{2-\gamma}$$

$$C = (\gamma - 1) k_{\min}^{\gamma-1}$$

$$k_{\max} = k_{\min} N^{\frac{1}{\gamma-1}}$$

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

Sources

- A. L. Barabási (2016). Network Science – Chapter 04
- URLs cited in the footer of specific slides