Distances in Scale-Free Networks

Social Networks Analysis and Graph Algorithms

Prof. Carlos Castillo — https://chato.cl/teach

Universitat Pompeu Fabra Barcelona

Contents

- Distance distribution of scale-free networks

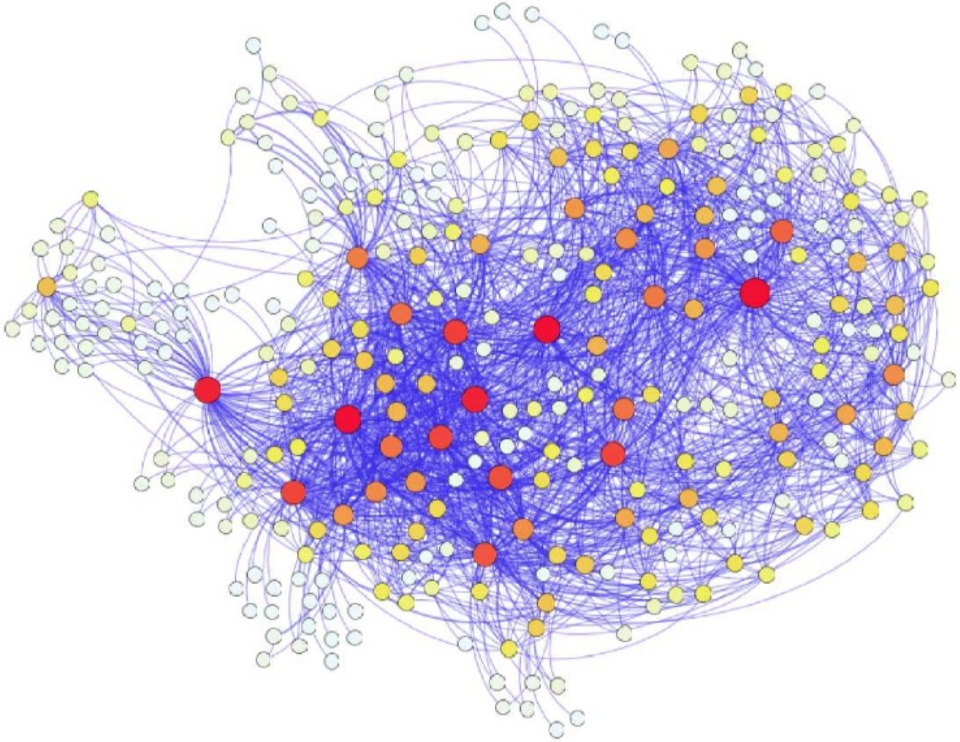
Sources

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

Consequences of having extremely large degree nodes *(also known as "large hubs")*

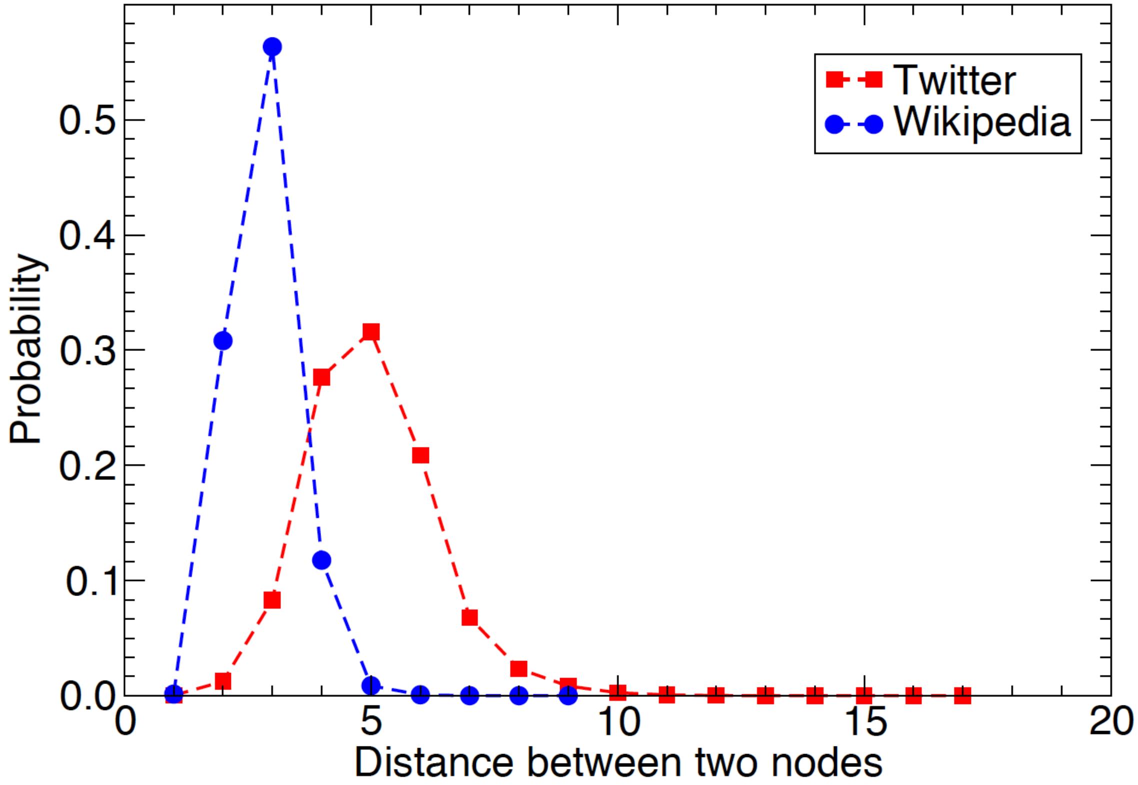
Air travel

- You can travel between almost all pairs of European airports directly or (most of the time) with at most one stop
- All you have to do is **go to a well connected airport**
- This is because there are large degree airports



Cardillo, A et al. (2013). Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. Euro. Phys. J. Special Topics, 215(1), 23-33. [DOI]

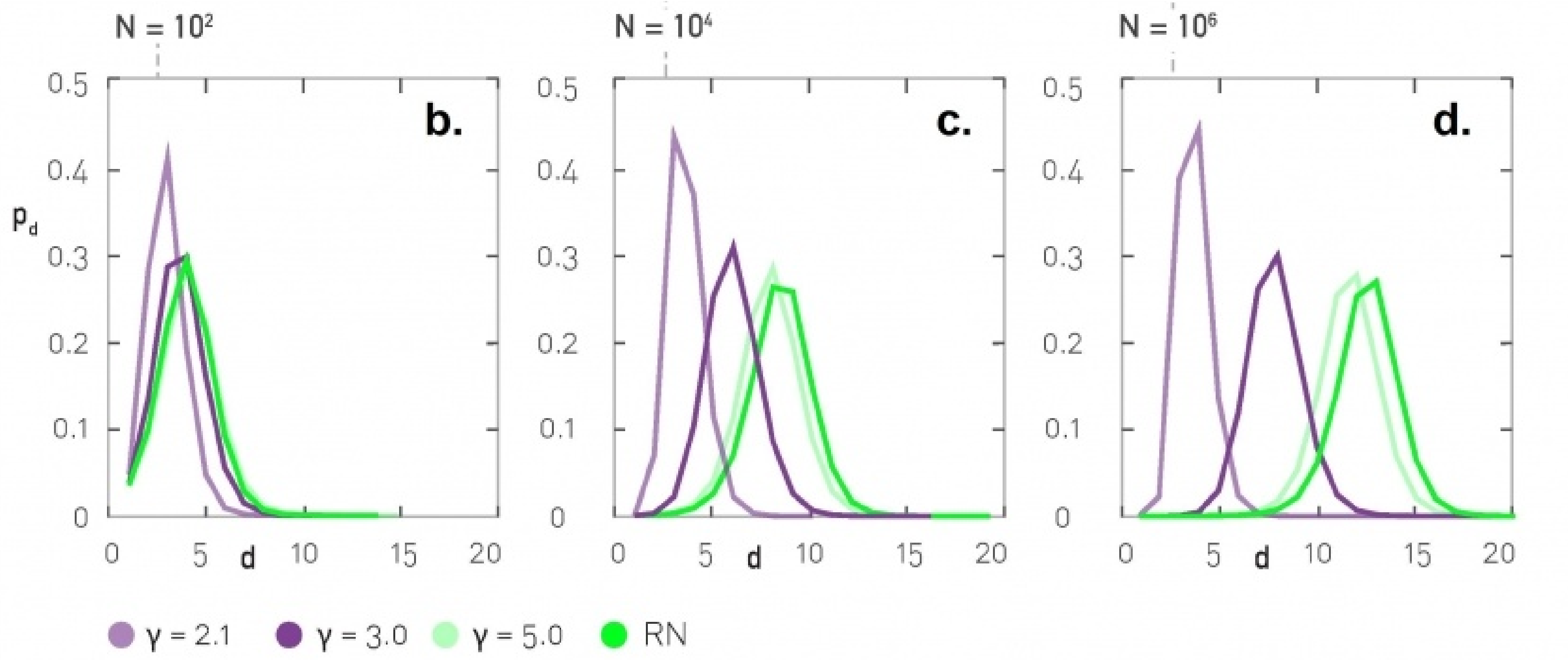
In general, having “hubs” or large degree nodes reduces distances



Source: Menczer, Fortunato, Davis: A First Course on Networks Science. Cambridge, 2020.

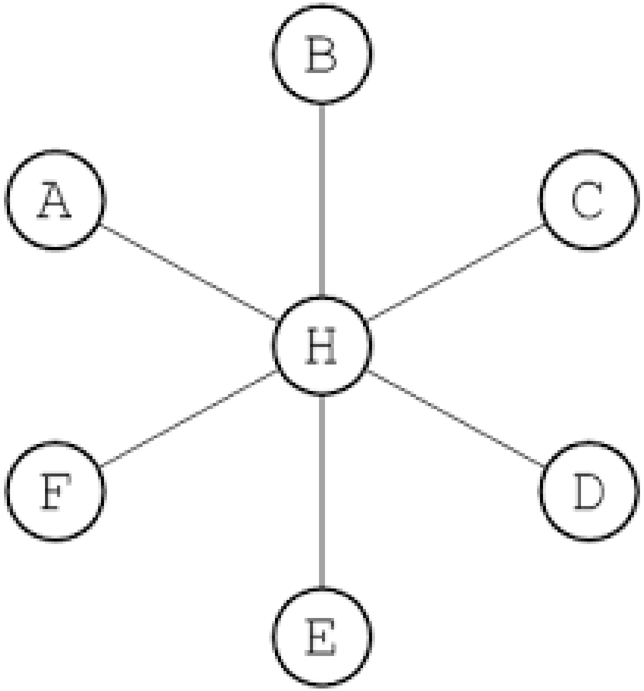
Distance distributions: simulation results

Scale-free networks of increasing size, $<k> = 3$



Distance regimes

Anomalous regime $\gamma = 2$



Ultra-small world $2 < \gamma < 3$

- Average distance follows log(log(N))
- Example (humans):

$$N \approx 7 \times 10^9$$

$$\log N \approx 22.66$$

$$\log \log N \approx 3.12$$

Average distance

Scale-free network
$p_k \propto k^{-\gamma}$

- Depends on γ and N

$$\langle d \rangle = \begin{cases} \text{const.} & \text{if } \gamma = 2 \\ \log \log \mathrm{N} & \text{if } 2 < \gamma < 3 \\ \log \mathrm{N} / \log \log \mathrm{N} & \text{if } \gamma = 3 \\ \log \mathrm{N} & \text{if } \gamma > 3 \end{cases}$$

Same as in ER graphs

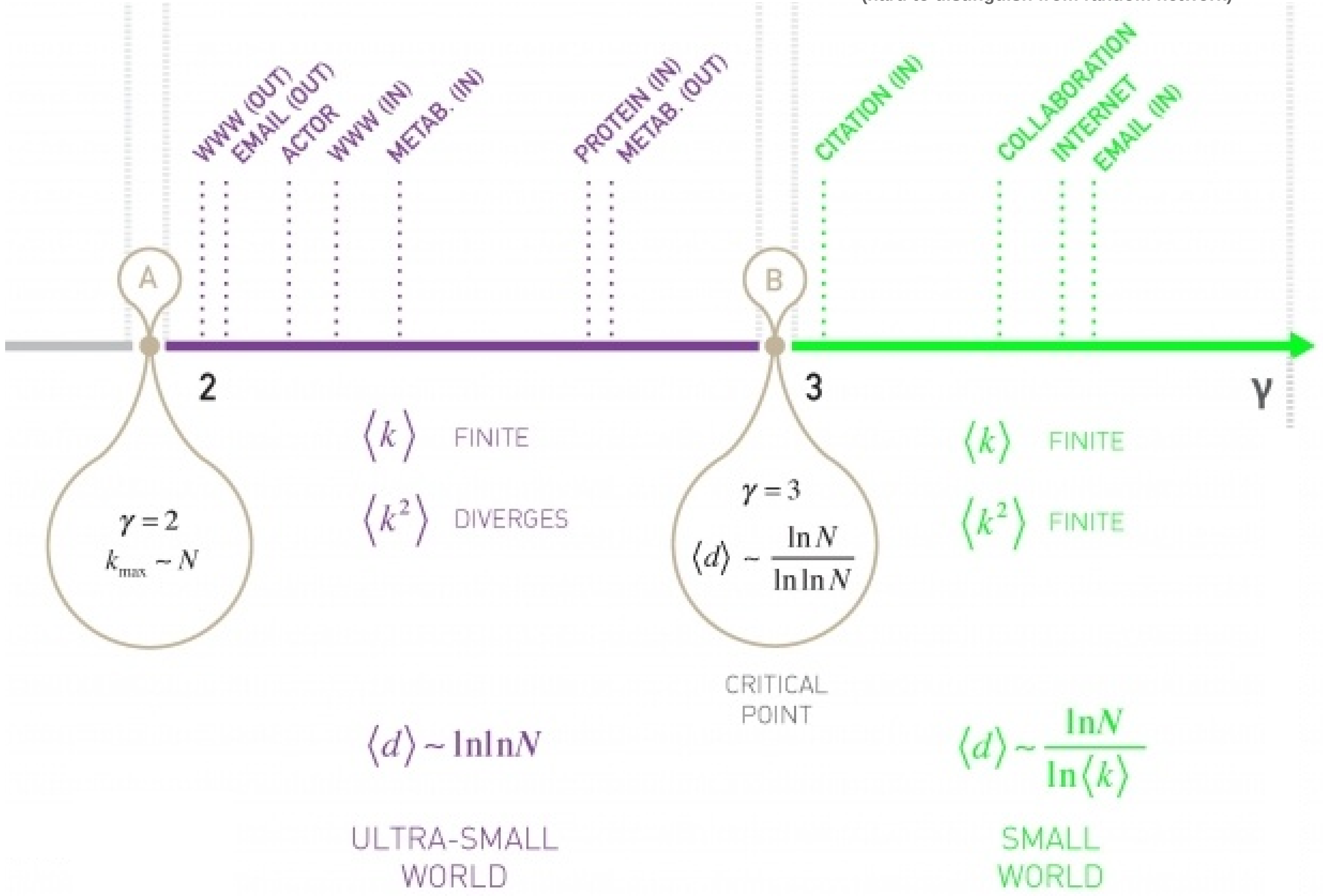
Small world $\gamma > 3$

- Average distance follows log(N)
- Similar to ER graphs where it followed log(N)/log($<$k$>$)

Scale-free regime

Random regime

(hard to distinguish from random network)



When $\gamma > 3$

- In this case it is hard to distinguish this case from an ER graph
- In most real complex networks (but not all)

$$2 < \gamma < 3$$

When $\gamma > 3$

- Remember $$k_{\max} = k_{\min} N^{\frac{1}{\gamma-1}} \qquad N = \left(\frac{k_{\max}}{k_{\min}}\right)^{\gamma-1}$$

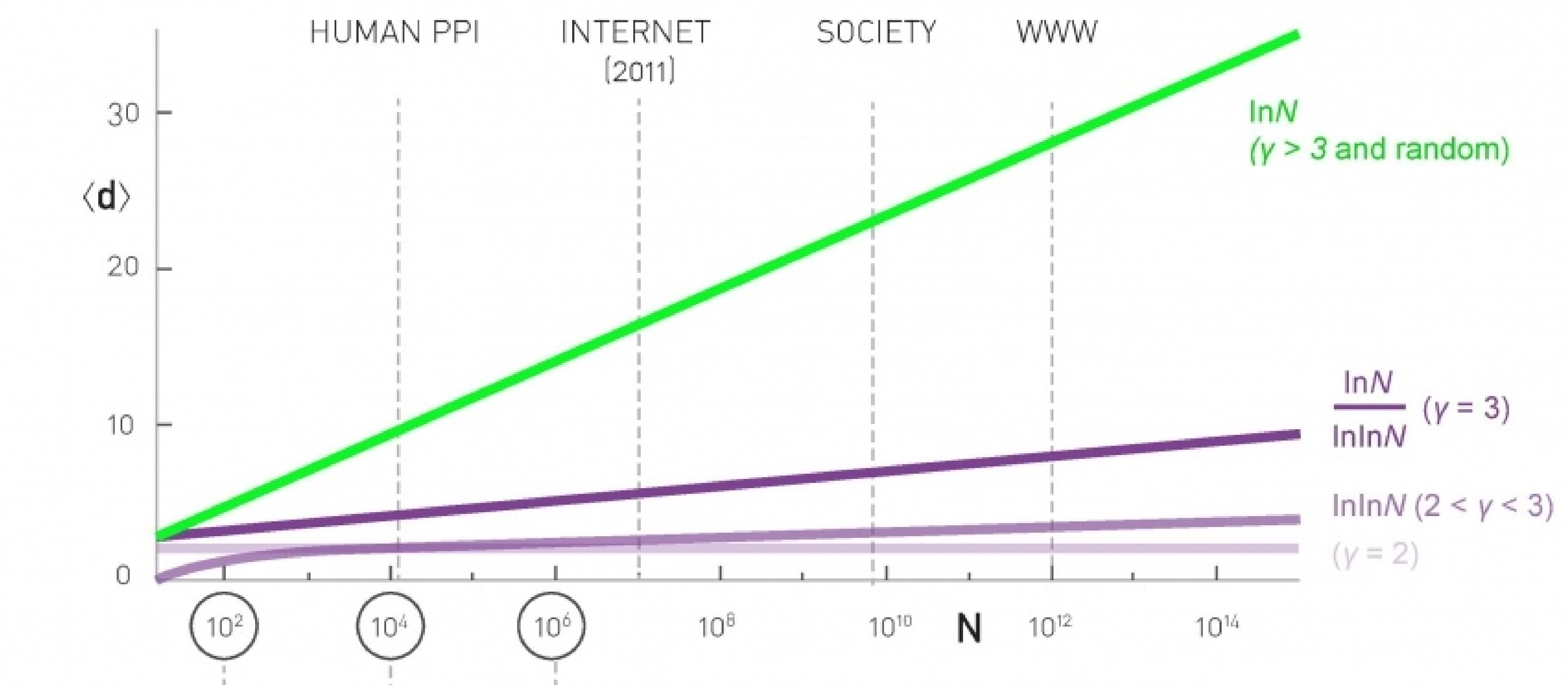
- Observing the scale-free properties requires that

  $k_{max} >> k_{min}$, e.g. $k_{max} = 10\ k_{min}$

- Then if $\gamma = 5, N > 10^8$

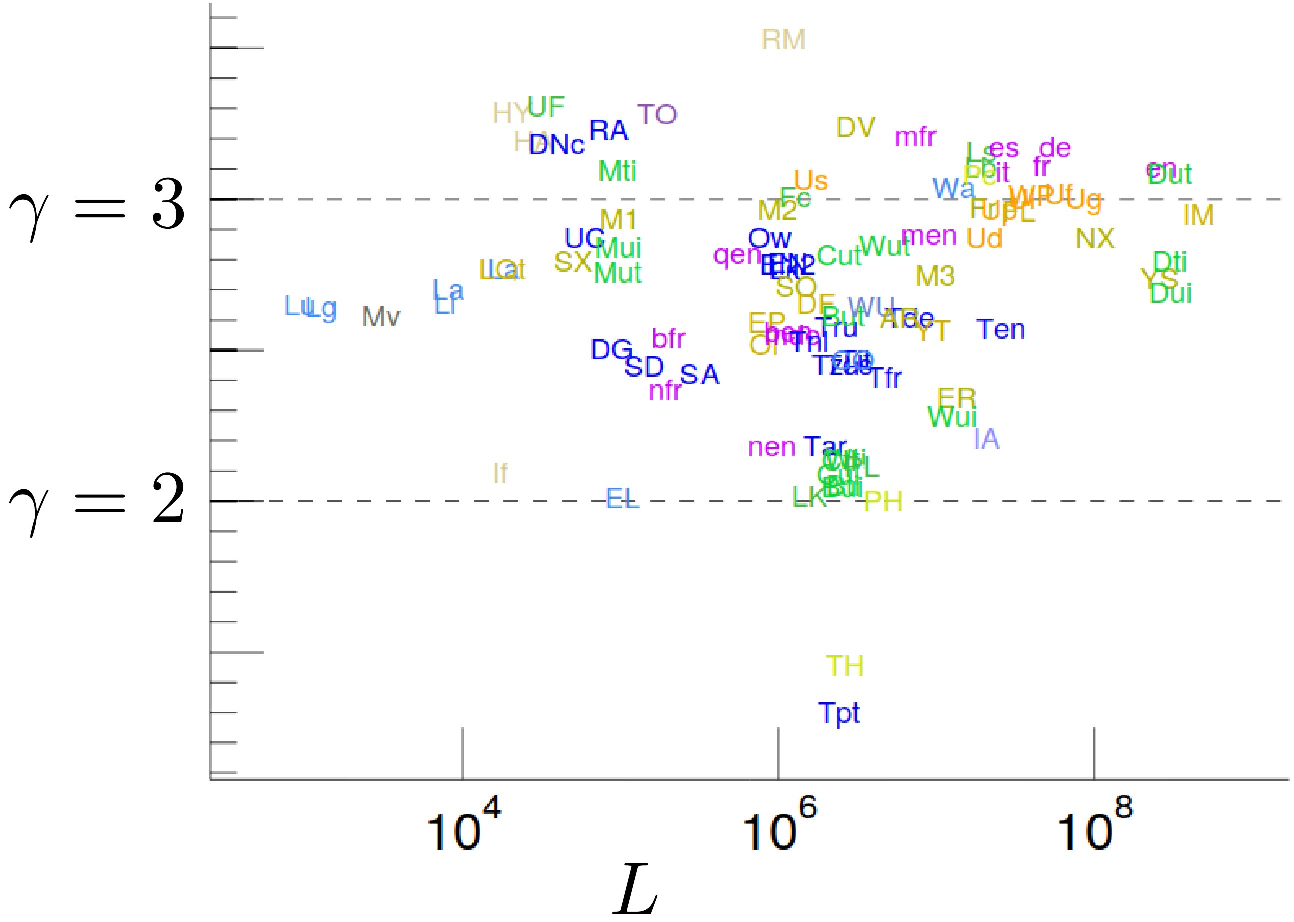
- There are not many such networks for which we have available data

Average distance and N



Examples

http://konect.uni-koblenz.de/statistics/prefatt



EL	Wikipedia elections
LK	Linux kernel mailing list threads
Bui	BibSonomy u-i
Bti	BibSonomy t-i
Cui	CiteULike u-i
If	Infectious
PL	Prosper loans
Cti	CiteULike t-i
Wti	Twitter t-i
nen	Wikinews (en)
Tar	Wikipedia talk, Arabic
Wui	Twitter u-i
ER	Epinions
nfr	Wikinews (fr)
Tfr	Wikipedia talk, French
SD	Slashdot
Tzh	Wikipedia talk, Chinese
Tes	Wikipedia talk, Spanish

Etc.

Exercise: average distance

	Network	N	‹k›	‹d›	lnN/ln‹k›
$\gamma > 3$	Internet	192,244	6.34	6.98	6.58
$2 < \gamma < 3$	WWW	325,729	4.60	11.27	8.31
$\gamma > 3$	Email	57,194	1.81	5.88	18.4
$\gamma > 3$	Science Collaboration	23,133	8.08	5.35	4.81
$2 < \gamma < 3$	Actor Network	702,388	83.71	3.91	3.04
$\gamma > 3$	Citation Network	449,673	10.43	11.21	5.55
$2 < \gamma < 3$	E. Coli Metabolism	1,039	5.58	2.98	4.04
$2 < \gamma < 3$	Protein Interactions	2,018	2.90	5.61	7.14

Pick 4 of these networks and compare the approximation of average distance assuming a scale-free regime ...

$$\langle d \rangle = log(log(N))$$

vs assuming a random regime ...

$$\langle d \rangle = \frac{logN}{\log \langle k \rangle}$$

Pin board: https://upfbarcelona.padlet.org/chato/tt14-average-distance-38m66yhjwvvh9q4a

Summary

Things to remember

- Regimes of distance and connectivity

Practice on your own

- Remember the regimes of a graph given γ

  (It is useful to know this by heart)

- Estimate distance distributions for some graphs