"Scale-free Networks are Rare"

by Anna Broido & Aaron Clauset

Presented by Cambria Naslund

Sociology Statistics Reading Group - March 8, 2018

Scale-free networks

Emergence of Scaling in Random Networks

Albert-László Barabási* and Réka Albert

Systems as diverse as genetic networks or the World Wide Web are best described as networks with complex topology. A common property of many large networks is that the vertex connectivities follow a scale-free power-law distribution. This feature was found to be a consequence of two generic mechanisms: (i) networks expand continuously by the addition of new vertices, and (ii) new vertices attach preferentially to sites that are already well connected. A model based on these two ingredients reproduces the observed stationary scale-free distributions, which indicates that the development of large networks is governed by robust self-organizing phenomena that go beyond the particulars of the individual systems.

Paper in Science, 1999

Scale-free networks

"A scale-free network is a network whose degree distribution follows a power law." (Barabási, Chapter 4)

For node degrees $k = 0, 1, 2, \ldots,$ the probability p_k that a node has k links is given by

$$p_k = Ck^{-a}$$

Scale-free networks

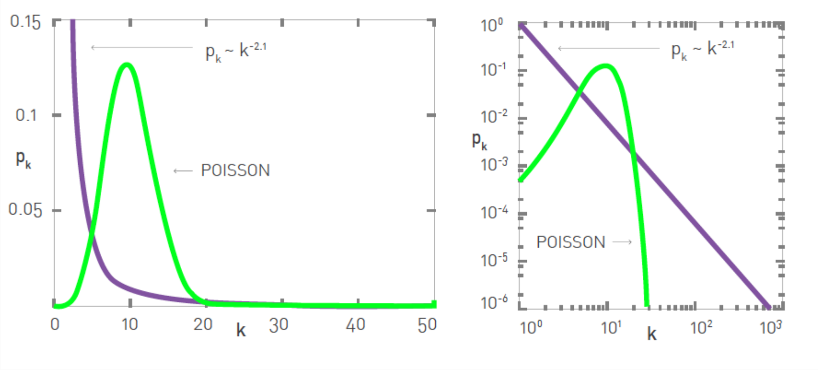


Figure 4.4 in Barabási, Chapter 4.

Broido & Clauset: Definitions

agreed: fraction of nodes with degree k follows a power-law distribution k^{-a}, where $a > 1$

$a \in [2, 3]$ ("ultra-small world")

evolves by preferential attachment mechanism

power law only in the upper tail

power law is only more plausible than other distributions

Broido & Clauset: Data

ICON corpus of 927 network datasets from various domains

4477 simple graphs

exclude graphs with $\langle k \rangle < 2$ and $\langle k \rangle > \sqrt{n}$ mean degrees (7376 extremely sparse and 12,146 extremely dense graphs)

Broido & Clauset: Test

Fit a power-law model for each degree sequence $k_1, k_2, \ldots, k_n$:

$$p(k) = Ck^{-\alpha}$$

$\alpha > 1, k \geq k_{\min} > 0$

$k_{\min}$ is estimated to minimize Kolmogorov-Smirnov (KS) statistic:

$$D = \max_{k \geq k_{\min}} |E(k) - P(k|\hat{\alpha})|$$

α is then estimated using maximum likelihood.

Broido & Clauset: Test

Test for plausibility of fitted model:

p-value estimated using semi-parametric bootstrap approach

Synthetic datasets are generated using model parameters to get null distribution of KS-statistics

$p = \Pr(D \geq D^*)$

If $p < 0.1$, reject power-law model for degree distribution

Broido & Clauset: Test

Test if power-law fit is better than alternatives.

compare fitted power-law to alternative distributions:

exponential, log-normal, power-law w/ exponential cutoff, and stretched exponential (Weibull)

likelihood ratio test: $R = L_{\text{PL}} - L_{\text{Alt}}$

aggregate of test results from all simple graphs in network data set

Broido & Clauset: Criteria

Super-Weak: For at least 50% of graphs, none of the alternative distributions are favored over the power law.

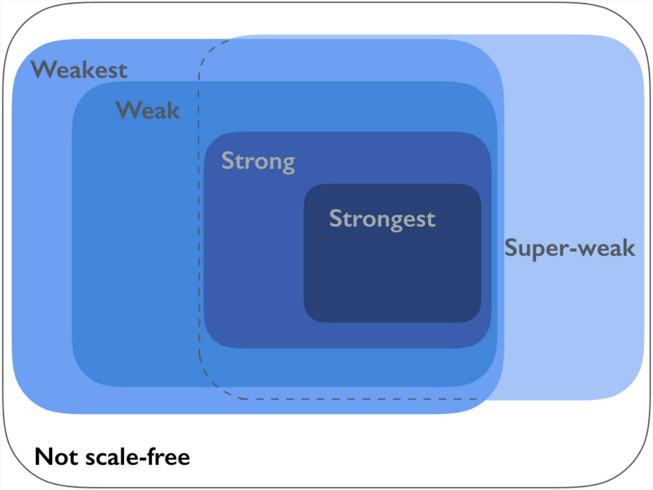
Weakest: For at least 50% of graphs, the power-law hypothesis cannot be rejected ($p \geq 1$).

Weak: The requirements of the *Weakest* set, and there are at least 50 nodes in the distribution's tail ($n_{tail} > 50$).

Strong: The requirements of the *Weak* set, and that both $2 < \hat{a} < 3$, and for at least 50% of graphs none of the alternative distributions are favored over the power-law.

Strongest: The requirements of the *Strong* set for at least 90% of graphs, rather than 50%, and for at least 95% of graphs none of the alternative distributions are favored over the power-law.

Broido & Clauset: Criteria

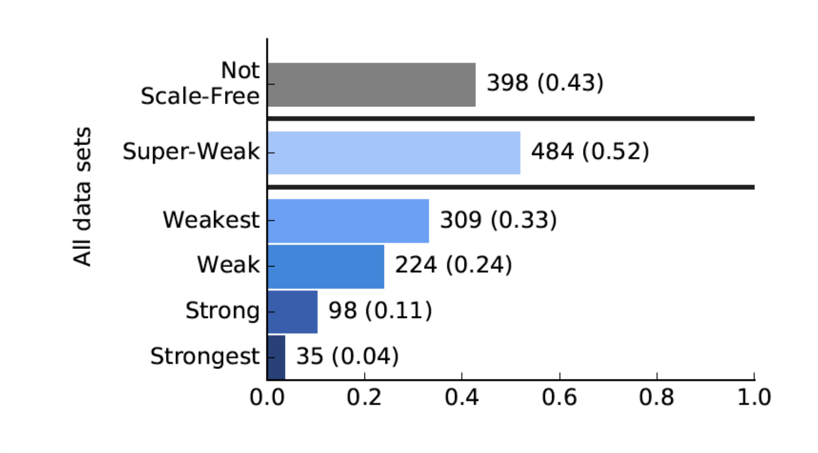


Broido & Clauset, Fig. 3

Broido & Clauset: Method Validation

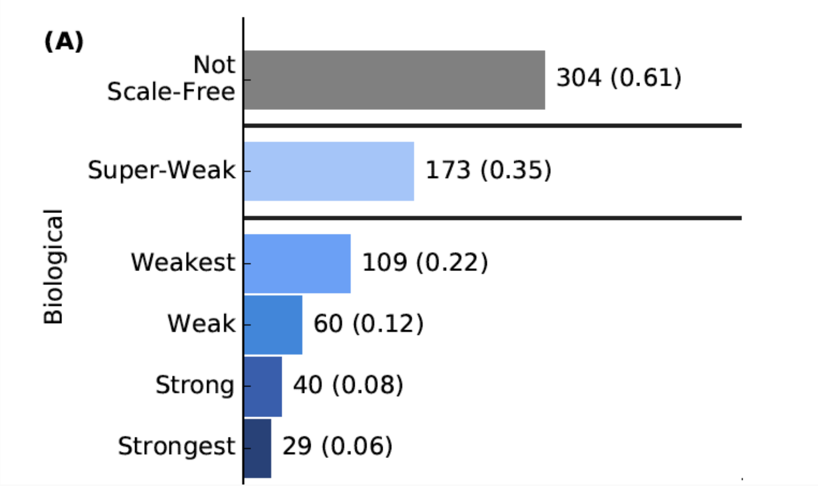
We'll come back to this…

Broido & Clauset: Results



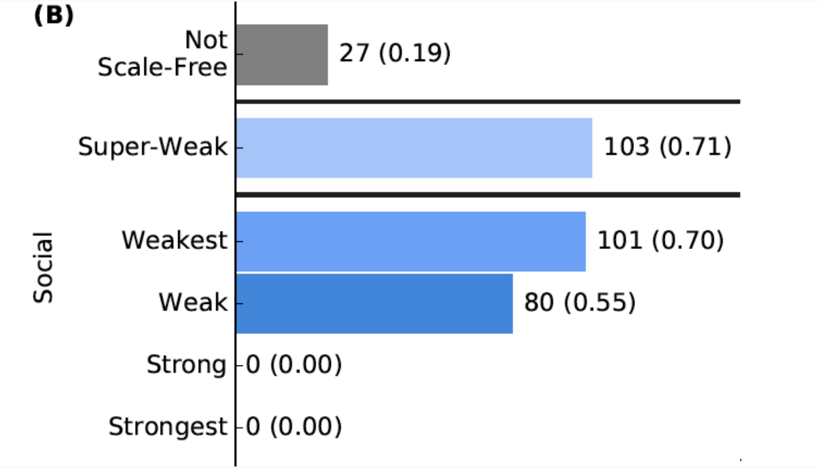
Broido & Clauset, Fig. 6

Broido & Clauset: Results



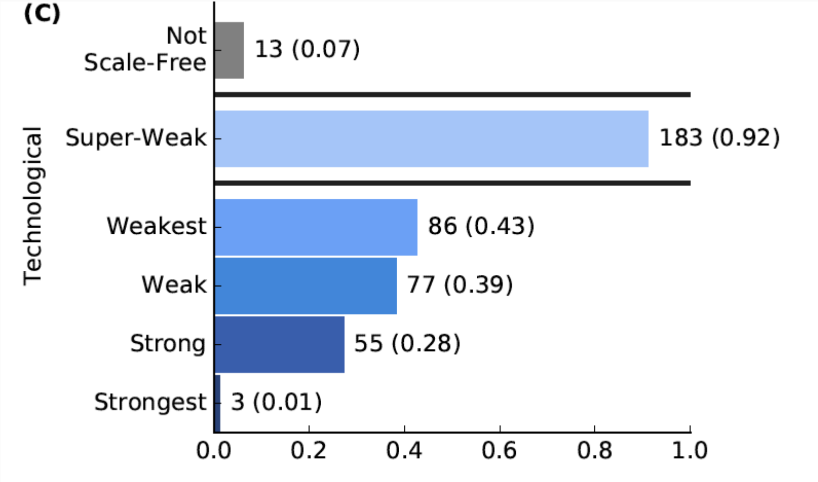
Broido & Clauset, Fig. 7

Broido & Clauset: Results



Broido & Clauset, Fig. 7

Broido & Clauset: Results



Broido & Clauset, Fig. 7

Broido & Clauset: Results

"… across a large and diverse corpus, we find that it is remarkably rare for a network data set to exhibit the strongest form of direct evidence of scale-free structure…"

Broido & Clauset: Results

"… we find essentially no empirical evidence to support the special status that the power law has held in network science as a starting point for modeling and analyzing the structure of real networks. Instead, it is an empirical fact that real-world networks exhibit a rich variety of degree structures, relatively few of which are convincingly scale free."

Barabási's Response

Love is All You Need

Clauset's fruitless search for scale-free networks

by Albert-László Barabási March 6, 2018

https://www.barabasilab.com/post/love-is-all-you-need

Barabási’s Response

“by 2001 it was pretty clear that there is no one-size-fits all formula for the degree distribution for scale-free networks. A pure power law only emerges in simple idealized models, driven only by growth and preferential attachment, and free of any additional effects.”

Barabási's Response

"… fit a pure power law to every network, and ignoring what the theory predicts for any of them. As it is difficult to find real systems that are free of additional effects, it makes no sense to fit indiscriminately a power law to all of them. One must fit the distribution that the theory predicts, which is predictably different for each system."

Barabási's Response

"… the theory predicts that in many real networks driven by growth and preferential attachment, the degree distribution should follow a stretched exponential (power law with an exponential cutoff). If you look at Table II of Ref[1], BC find that 51% of the networks they explored favor this distribution. In other words, their measurements validate the theory, contradicting their central claim."

Broido & Clauset: Table II

Alternative	$p(x) \propto f(x)$	Test Outcome M_{PL}	Inconclusive	M_{Alt}
Exponential	$e^{-\lambda x}$	**37%**	27%	36%
Log-normal	$\frac{1}{x} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}$	**12%**	43%	45%
Weibull	$e^{-\left(\frac{x}{b}\right)^a}$	**33%**	25%	42%
Power law with cutoff	$x^{-\alpha} e^{-\lambda x}$	—	49%	51%

Broido & Clauset, Table II

Barabási's Response

"Even the exact model of scale-free networks, following a pure power law, fails their test."

Broido & Clauset: Appendix E

	PA	VC	E-R
Super-Weak	87%	88%	16%
S-W w/o cutoff alternative	98%	99%	31%
Weakest & Weak	62%	74%	51%, 50%
Strong	60%	70%	0
Strongest	0	0	0

Broido & Clauset: Appendix E

Strongest category requires that 90% of the simple-graphs follow a power law.

Out-degree sequences do not follow power law, so we can get at best 2/3.

Example from Barabási's response: citation networks

Discussion

References

Broido, A. D., & Clauset, A. (2018). Scale-free networks are rare. arXiv preprint arXiv:1801.03400.

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286(5439), 509-512.

Barabási, A. L. (2016). Network science. Cambridge University Press.

Barabási, A. L. (2018, March 6). Love is All You Need [Blog post]. https://www.barabasilab.com/post/love-is-all-you-need