

# Love is All You Need

## Clauset's fruitless search for scale-free networks

by Albert-László Barabási, March 6, 2018

In the past weeks, I have received several requests to address the merits of the Anna D. Broido and Aaron Clauset (BC) preprint [1] and their fruitless search for scale-free networks in nature. The preprint's central claim is deceptively simple: It starts from the textbook definition of a scale-free network as a network with a power law in the degree distribution [2]. It then proceeds to fit a power law to 927 networks, reporting that only 4% are scale-free. The author's conclusion that 'scale-free networks are rare,' is turned into the title of the preprint, helping it to get maximal attention. It worked—*Quanta* magazine accepted its conclusions without reservations. After the *Atlantic* carried the article, the un-refereed preprint received a degree of media exposure that the original discovery of scale-free networks never enjoyed.

While I saw the conceptual problems with the manuscript, I was convinced that the paper must be technically proficient. Yet, once I did dig into it, it was a real ride. If you have the patience to get to the end of this commentary, you will see where it fails at the conceptual level. But, we will learn that it also fails, repeatedly, at the technical level.

### The Conceptual Problem

Let me start by summarizing an enormous body of work on scale-free networks as briefly as possible, essential to understand the moment when the BC paper arrives on the scene.

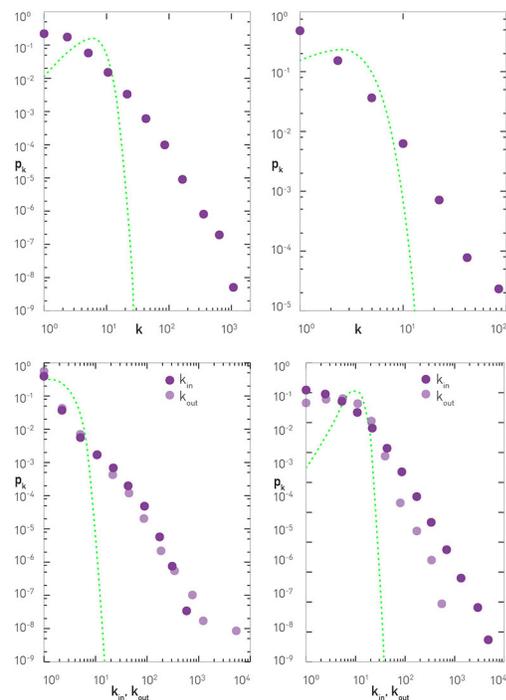
*Empirical Discovery.* The scale-free property was observed in 1999, independently in the WWW [3] and the internet at the autonomous systems level [4]. The empirical observation was simple: the degree distribution of these networks is well approximated with a power law,

$$P(k) \sim k^{-\gamma} \quad (1)$$

Back then this was an unexpected finding, given that the prevailing model was the random network model of Erdős and Rényi [5] which predicted a Poisson degree distribution, easily differentiable from a power law (Figure 1). We named these networks *scale-free*, inspired by the scale-free nature of power laws observed in the vicinity of phase transitions.

*Mechanistic Model.* In 1999 Réka Albert and I proposed a mechanism to explain the origin of the observed power law [2], finding that it is rooted in growth and preferential attachment. A simple model based on these two mechanisms, known today as the BA or the *scale-free model*, generated a network whose degree distribution follows (1), with degree exponent  $\gamma=3$ . The scale-free model is a mechanistic model, meaning that it is not a model of the Internet, or the WWW or the cell—it only aims to explain the mechanism behind the scale-free nature of a network. Real networks are far more complicated, indicated by the fact that their degree exponent  $\gamma$  varies from 2 to 8. Growth and preferential attachment *alone* cannot explain that variation.

*Real Networks.* Months after the 1999 paper several key discoveries helped the community understand the origins of the different exponents. The rate-equation based continuum theory, developed by Mendes, Dorogovtsev,



**Figure 1. How hard is to distinguish random from scale-free networks?** To show how different are the predictions of the two modeling paradigms, the scale-free and that of the random network models, I show the degree distribution of four systems: Internet at the router level; Protein-protein interaction network of yeast; Email network; Citation network, together with the expected best Poisson distribution fit. It takes no sophisticated statistical tools to notice that the Poisson does not fit.

Redner, Krapivsky, and many others [6], helped incorporate into the scale-free model many effects that naturally occur in real networks, like the disappearance of nodes, the addition of new links between previously existing nodes, link deletion, aging, and so on. These papers have shown analytically that the presence of the additional effects alter the degree distribution in two ways. First, they change the degree exponent—successfully explaining the diverse exponents observed in real networks. Equally important, *they induce corrections to the degree distribution, making  $P(k)$  deviate in predictable ways from a pure power law.* Some of the most common corrections include [7]:

- *Small degree saturation:* In any system where two nodes that are already in the network can choose to connect to each other (internal links), the analytical form of the degree distribution becomes  $P(k)=C(k+A)^{-\gamma}$ , leading to a small degree saturation. In social networks and the WWW, most links are such internal links. A similar small-degree saturation can also be induced by initial attractiveness, well documented in citation networks [8].
- *High degree cutoffs:* If preferential attachment is sublinear, *the degree distribution follows a stretched exponential, or a power law with an exponential cutoff.* Similar high degree cutoffs also appear under node removal.
- *Fitness:* Nodes have different abilities to complete for the links, a diversity that can be modelled by giving each of them a unique and different fitness  $\eta$ . In the presence of fitness, the precise form of  $P(k)$  depends on the fitness distribution  $\rho(\eta)$ . For example, a uniform fitness distribution induces a logarithmic correction in  $P(k)$  and other forms of  $\rho(\eta)$  can lead to rather exotic forms for  $P(k)$ .

In real networks many of the elementary processes discussed above appear together, and so do many others. In other words, by 2001 it was pretty clear *that there is no one-size-fits all formula for the degree distribution of networks driven by the scale-free mechanism. A pure power law only emerges in simple idealized models, driven by only growth and preferential attachment, and free of any additional effects.* The theory was predicting that in real networks, if the scale-free mechanism is present, the power law tends to coexist with some corrective function, expecting power laws with exponential cutoffs, stretched exponentials, logarithmic corrections to the power law, and so on. These corrections are so important, that I devoted a full chapter in the *Network Science* [9] textbook to them. We also learned that there are multiple ways of analyzing the presence of the scale-free property, as I explain in Box 1. The conclusion is simple, well understood in the literature: *if we wish to obtain an accurate fit to the degree distribution of a real network, we first must build a generative model that analytically predicts the functional form of  $P(k)$ .*

So it is time to go back to Ref [1] and its key claim that “Scale-free networks are rare.” How exactly did it arrive to this conclusion? By insisting to 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.*

Interestingly, the theory predicts that in many real networks driven by growth and preferential attachment, the degree distribution should follow a 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, the measurements of BC validate the theory, contradicting the authors' central claim.

### ***Box 1: How do we analyze real networks?***

Given the complexity of real systems, how can we detect the scale-free property of a real network? We can do better than blindly fitting of a power law to it. The procedure we choose depends on what question we wish to address:

1. If our goal is to obtain an accurate fit of the *degree distribution*, then we must build a continuum model of the network we wish to fit. This has been done successfully in several well-studied systems, finding that in each case the scale-free mechanism, captured by growth and preferential attachment, are necessary to describe their topology.
2. If our goal is to obtain *direct evidence of the scale-free mechanism*, we can do that without doing the hard work in point 1, but only if we have dynamical data (i.e. the time when each node or link was added). In that case we can measure both growth and preferential attachment directly [10].
3. Finally, if we want to see the consequences of the scale-free property, we can bypass both 1 and 2 and measure the variance of the degree distribution. This helps us ask, why does the scale-free property matter in the first place? I discuss this in Box 2.

## The Technical Problems

Getting this far, you may ask yourself, if it is so difficult to fit the degree distribution, why do thousands of papers claim that a large number of real networks, from the Internet to protein interaction and social networks, are scale-free? The answer is simple—despite the many processes shaping their topology, for many real networks the fat tailed nature is so obvious that it's hard to miss (see Figure 1). Indeed, Clauset's most cited work found that many of the large networks that the community does consider scale-free, like the Internet, protein interaction network, citation network, co-authorship networks, do pass statistical significance without even considering the necessary corrections (see Table 6.1 in Ref [12]).

So the puzzling technical question is this: Why do the authors fail to find the scale-free networks, that everyone else in the literature does, including Clauset himself in his earlier work? This is where the first technical surprise comes: **they fail to see the scale-free property because they invent a new criterion of weak and strong scale-free networks.** And the real surprise? **Even the exact model of scale-free networks, following a pure power law, fails their test.** As incredible as this sounds, all the evidence is in Appendix E, on the very last page. So let's dig into it.

You would think that the nomenclature of 'weak' and 'strong' scale-free networks BC uses has to do with statistical significance. Indeed, one could plausibly call some of the large and well mapped real networks 'strong scale-free', implying that for them the statistical significance is exceptional. And one could coin the term "weak scale-free network" for a network for which a power law has lower statistical significance. But that is not what BC do. Instead, they take each network, and generate from the original *multiple synthetic networks*. This way their set of 927 real networks turns into 23,999 *synthetic networks*. Then they toss out 81% of them, deeming them 'unlikely to be scale-free' (pg 3, BC). They proceed with the remaining 4,477 synthetic networks, which is the corpus they study.

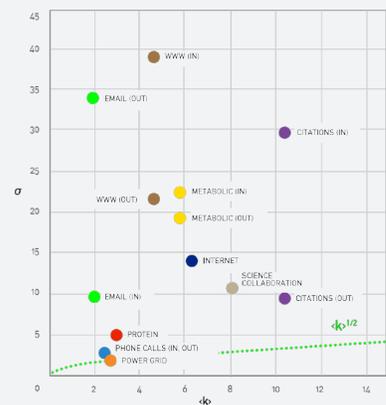
For example, an unweighted directed network like the WWW turns into three different degree sequences: the incoming, the outgoing, and the total degree. Instead of asking which of these synthetic networks may be well fitted with a power law, they ask, what *fraction of these synthetic networks* passes the power law criteria. Then they propose the following naming convention, taken from their paper:

- *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 ( $p \geq 0.1$ ).
- *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  $2 < \alpha < 3$ , 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.

To help understand what this is, let us take the citation graph, a well studied scale-free network. It is a directed network, whose incoming degrees (the number of citations a paper gets) is known since the 1960s to be well fitted with a power law [13]. However, the outgoing degrees follow a different distribution, as those degrees capture how many different papers a given publication cites,

## Box 2: Why do scale-free networks matter?

Why did the scale-free property get so much attention in the past two decades? The main reason is that *the scale-free property does matter*. Indeed, several groundbreaking discoveries by Cohen, Havlin, Pastor-Satorras, Vespignani and others [11], showed that networks with high  $\langle k^2 \rangle$  have a series of unusual properties, like robustness to failures, fragility to attacks, anomalous spread of viruses, anomalous synchronization, and so on. For scale-free networks with a pure power law degree distribution  $\langle k^2 \rangle$  diverges for  $\gamma < 3$ , so they display all these intriguing properties. Yet, *you do not need a pure power law to witness the impact of the high  $\langle k^2 \rangle$* , as systems driven by the scale-free mechanism tend to have anomalously high  $\langle k^2 \rangle$ . You can easily test that by comparing the variance of a real network with that of a random network of similar size (see Figure 2)—if it is higher, the intriguing properties of the scale-free models will manifest themselves in your network, whether or not the degree distribution follows a pure power law.



**Figure 2. High degree variations in real networks**  
For a random network with Poisson degree distribution the standard deviation of the degrees follows  $\sigma = \langle k \rangle^{1/2}$  shown as a green dashed line on the figure. The symbols show  $\sigma$  for nine reference networks. For each network  $\sigma$  is larger than the value expected for a random network with the same  $\langle k \rangle$ . The only exception is the power grid, which is not scale-free. While the phone call network is scale-free, it has a large  $\gamma$ , hence it is well approximated by a random network.

and that number is determined by journal policy, having artificial cutoffs. So, of the three graphs they generate from a well-studied scale-free network, only one can pass their criteria, the one defined by the incoming degrees. But that is not enough to make it even super-weak scale-free in the language of BC.

In some cases, their methodology generates as many as 900 synthetic networks from a single real system. If at least 50% of these synthetic networks pass the power law test they would call it super-weak. They require 90% to pass to place a network in the strongest category. Now, the truth is that typically **only one of these networks matters: the one that captures the purpose or the function of the original system**. But they offer an equal vote to each fictional synthetic network, letting them decide whether the original system is scale-free or not (Box 3).

The classification BC choose to use has no precedent in network science, and has no relevance to the role of the network they study. They offer no explanation of the need for these made-up criteria, nor do they explain how they choose the 50% and 90% percentages. They are arbitrary.

But let's focus on what really matters. And those are controls. That is, whatever criteria we choose, we must test it on networks whose degree distribution is well understood. So if we feed into our methodology a scale-free network, like the one produced by the scale-free model with a pure power law degree distribution, free of any corrections, the method should predict that that network is in the strongest category. Indeed, *if the gold standard fails the strongest class, who would pass?* In other words, the scale-free model should not be super-weak, weakest, or weak. It should be in the strongest class, correct? After all, we have a formal exact proof of the power law nature of the scale-free model [14].

BC are truly aware of this crucial criteria, hence they do assure the reader on pg 5 that their method passes this obvious but essential test:

*“Two mechanisms produce scale-free networks (a directed version of preferential attachment [42] and a directed vertex copy model [43]), and one does not (simple Erdős-Rényi random graphs). Applied to synthetic networks generated by these mechanisms, our methodology correctly placed the synthetic data sets into scale-free categories suitable for their generating parameters, i.e., preferential attachment and vertex-copy data sets were categorized as scale-free networks, while Erdős-Rényi random graphs were not (see Appendix E).”*

So it's all good. Yet, curious, I did look into Appendix E, which, I assumed that would simply offer the evidence. And it did. But not the evidence we would all expect based on the text above. Here is word by word what we find there:

*“When we consider the plausibility of the power-law fit, we see fewer networks. 62% of the preferential attachment graphs fall into the Weakest and Weak categories,”*

They also report for the Erdős-Rényi network, that

*“51% and 50% of these networks fall into the Weakest and Weak categories, respectively.”*

In other words, according to the criteria the authors invented, **38% of the time the scale-free model is not even 'weak scale-free.'** In contrast, **51% of the time the ER model is classified as 'weak scale-free'!**

You may ask, how hard is to distinguish an Erdős-Rényi model from a scale-free

### **Box 3: All we need is love**

If you have difficulty understanding the need for the super-weak, weakest, weak, strong and strongest classification, you are not alone. It took me several days to get it. So let me explain it in simple terms.

Assume that we want to find the word Love in the following string: "Love". You could of course simply match the string and call it mission accomplished. That, however, would not offer statistical significance for your match.

BC insist that we must use a rigorous algorithm to decide if there is Love in Love. And they propose one, that works like this: Take the original string of letters, and break it into all possible sub-strings:

{L,o,v,e,Lo,Lv,Le,ov,oe,ve,Lov,Loe,ove,Love}.

They call the match super-strong if at least 90% of these sub-strings matches Love. In this case we do have Love in the list, *but it is only one of the 14 possible sub-strings, so Love is not super strong.*

They call the match super-weak if at least 50% of the strings matches the search string. Love is obviously not super-weak either.

At the end Clauset's algorithm arrives to the inevitable conclusion: *There is no Love in Love.*

The rest of us: *Love is all you need.*

one? In their test BC used networks of 5,000 nodes. In Figure 3 we can see how different are the degree distribution of a scale-free network and an Erdős-Rényi network at this size. It needs no sophisticated statistical test to capture the difference.

But wait, it gets even more interesting. If we read further (pg 7, BC), this is what we find about their gold standard, the scale-free model (they refer to it as 'preferential attachment graph'):

*“When we consider the plausibility of the power-law fit, we see fewer networks. 62% of the preferential attachment graphs fall into the Weakest and Weak categories, 60% in the Strong category, and 0 in the Strongest category.”*

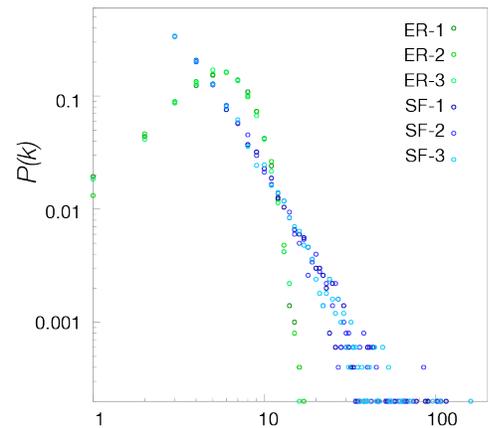
That is, not a single realization of the scale-free model is deemed strongly scale-free. **According to the criteria they invented, not even the scale-free model is scale-free any longer.**

At the end the preprint's main result, prominently featured in the press, rests on a simple claim: the authors fail to find the numerous scale-free networks in nature. If we actually read the paper, we find that they refuse to see them: Table II in Ref [1] shows that a high fraction of networks have degree distributions in agreement with the predictions of the continuum theory describing scale-free networks. The true failure is their methodology: it fails to detect that the gold standard is scale-free. Now, if you invent an arbitrary criterion that not even the mathematically proven models can pass, what exactly do you expect to get?

What really puzzles me, after all this, is that 4% of real networks did pass their criteria, finding them to be strongly scale-free. Which are those networks that are more scale-free than the scale-free model? They must be walking on water.

It is crucial for network science to constantly test the solidity of its pillars, like the scale-free nature of real networks, and the mechanisms responsible for it. We must continue therefore to constantly rule out competing hypotheses, a process that is essential for science to advance. Cutting corners will not get us there, however.

So before accepting the attention-grabbing claims of the BC paper, you must peek under its hood. And if you take the time to do that, you will find that the study **is oblivious to 18 years of knowledge accumulated in network science**. You will also find a fictional criterion of scale-free networks. Most important, **you will find that their central criterion fails the most elementary tests**. And those are only the big problems - if you are willing to devote time to it, you will discover additional technical lapses - but this piece must have an end. What you will not find, is a careful and unbiased statistical study that upends a paradigm.



**Figure 3. Differentiating model systems**

Curious about the reason the method adopted by BC cannot distinguish the Erdős-Rényi and the scale-free model, we generated the degree distribution of both models for  $N=5,000$  nodes, the same size BC use for their test. We have implemented the scale-free model described in Appendix E of Ref [1], a version of the original scale-free model (their choice is problematic, btw, but let us not dwell on that now). In the plot we show three different realizations for each network, allowing us to see the fluctuations between different realisations, which are small at this size. The differences between the two models are impossible to miss: The largest nodes in any of the Erdős-Rényi networks have degree less than 20, while the scale-free model generate hubs with hundreds of links. Even a poorly constructed statistical test could tell the difference. Yet, 38% of the time the method used by BC does not identify the scale-free model to be even 'weak scale-free,' while 51% of the time it classifies the ER model to be 'weak scale-free.'

#### **Box 4: Falling Like a Stone**

The situation encountered by BC might be familiar to anyone who has taken a mechanics class. The textbook tells us that according to Newton's theory of gravity doped objects will accelerate with  $g$ . Yet, if I drop a rock and a feather from my 11th floor window of the Network Science Institute, and measure their acceleration, I would quickly conclude that Newton's theory is flawed. First, I would only observe a short period of acceleration, at a rate that will be smaller than  $g$ . Second, after this initial period, both the rock and the feather will fall with constant speed. Both observations violate Newton's theory. If, however, I jump to the conclusion that gravity is wrong, I will flunk my mechanics class. A correct answer, that will get me a pass, is to acknowledge that the trajectory of the falling object is affected by both gravity and friction. Once I add friction to my equation, I will be able to fit the full trajectory of both the feather and the rock

## References

- [1] AD Broido, A Clauset. Scale-free networks are rare. arXiv:1801.03400
- [2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509-512, 1999.
- [3] R. Kumar, P. Raghavan, S. Rajalopagan, and A. Tomkins. Extracting Large-Scale Knowledge Bases from the Web. *Proceedings of the 25th VLDB Conference, Edinburgh, Scotland*, pp.639-650, 1999; H. Jeong, R. Albert and A. L. Barabási. Internet: Diameter of the world-wide web. *Nature*, 401:130-131, 1999;
- [4] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Proceedings of SIGCOMM. Comput. Commun. Rev.* 29: 251-262, 1999.
- [5] P. Erdős and A. Rényi. On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6:290-297, 1959.
- [6] P. L. Krapivsky, S. Redner, and F. Leyvraz, "Connectivity of Growing Random Networks", *Phys. Rev. Lett.* 85, 4629-4632 (2000); S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin. Structure of growing networks with preferential linking, *Physical Review Letters*, 85: 4633, 2000.
- [7] For more details, see Chapter 6 of *Network Science*.
- [8] Y.-H. Eom and S. Fortunato. Characterizing and Modeling Citation Dynamics. *PLoS ONE*, 6: e24926, 2011.
- [9] A.-L. Barabási *Network Science*. Cambridge University Press, 2017.
- [10] Section 5.7 in Ref. [9].
- [11] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin. Resilience of the Internet to random breakdowns. *Physical Review Letters*, 85:4626, 2000; B. Bollobás and O. Riordan. Robustness and Vulnerability of ScaleFree Random Graphs. *Internet Mathematics*, 1, 2003; R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scalefree networks. *Physical Review Letters*, 86:3200–3203, 2001.; R. Albert, H. Jeong, A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406: 378–482, 2000;
- [12] A Clauset, CR Shalizi, MEJ Newman. Power-law distributions in empirical data. *SIAM review* 51 (4), 661-703, 5933 (2009).
- [13] In some cases, depending how the data was collected, the log-normal also offers a reasonable fit to citation networks. For the power law, we have generative methods; the mechanistic origin of the log normal are less understood.
- [14] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*, 18:279-290, 2001.
- [15] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin. Resilience of the Internet to random breakdowns. *Physical Review Letters*, 85:4626, 2000; B. Bollobás and O. Riordan. Robustness and Vulnerability of ScaleFree Random Graphs. *Internet Mathematics*, 1, 2003; R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scalefree networks. *Physical Review Letters*, 86:3200–3203, 2001.; R. Albert, H. Jeong, A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406: 378–482, 2000;