What’s Going on with Scale-Free Networks?

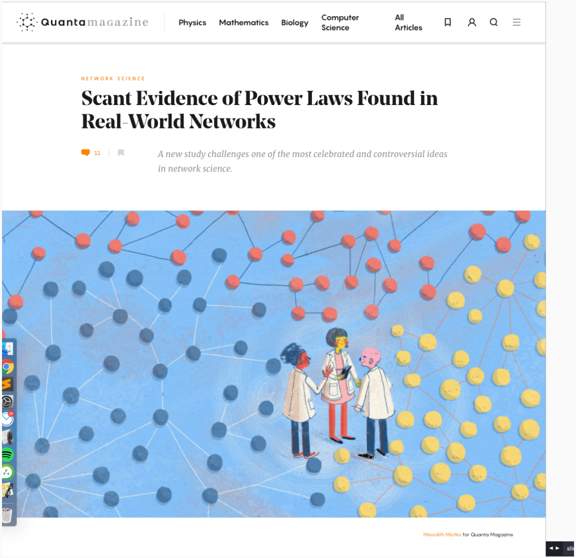
Phil Chodrow

April 26, 2018

MIT Operations Research Center
Human Mobility and Networks Laboratory
Laboratory for Information and Decision Systems

What’s Going On?

Media Coverage



Quanta magazine | Physics | Mathematics | Biology | Computer Science | All Articles

NETWORK SCIENCE

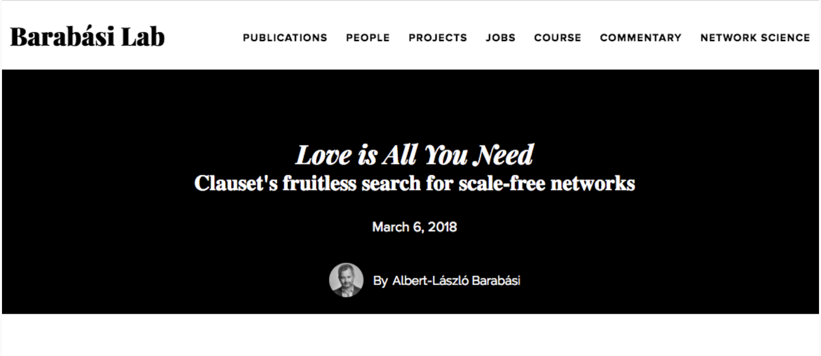
Scant Evidence of Power Laws Found in Real-World Networks

11

A new study challenges one of the most celebrated and controversial ideas in network science.

Meredith Miotke for Quanta Magazine

Weird Blog Posts



Twitter Fights

Laszlo Barabasi @barabasi · 15. jan.

The effort is amazing. The conclusions are less so. The feather falls slower than the rock, yet gravitation is not wrong. We add friction. You need to fit for each system the Pk that is right for it. That is hard, I know. Otherwise you ignore 20 year of work by hundreds.

Oversett fra engelsk

2 4 6

Aaron Clauset @aaronclauset · 15. jan.

It seems easy to get confused here: an empirical power-law degree distribution is evidence for SF structure, but no deviation from the power law can be evidence against SF structure? It is reasonable to believe a fundamental phenomena would require less customized detective work.

Ummmm....What are we talking about?

A Brief History....

The Setting

The Setting

- It's the late 1990s.

The Setting

- It's the late 1990s.
- The work of Erdös and Renyi is still the **dominant paradigm** for thinking about and modeling networks.

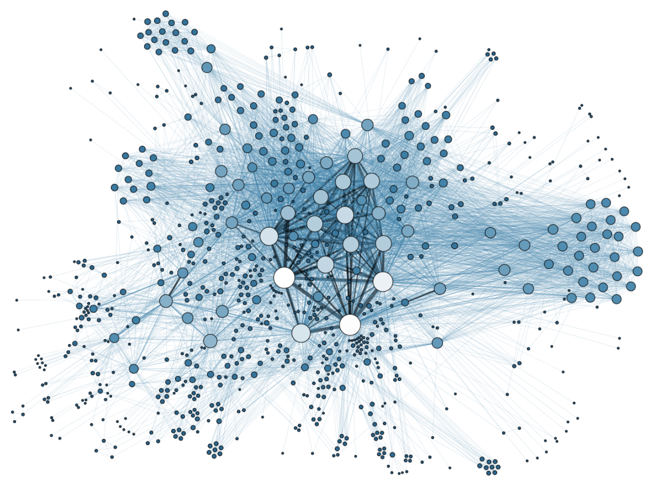
The Setting

- It's the late 1990s.
- The work of Erdös and Renyi is still the **dominant paradigm** for thinking about and modeling networks.
- But **digital infrastructure** and **computing power** are growing...

The Setting

- It’s the late 1990s.
- The work of Erdös and Renyi is still the **dominant paradigm** for thinking about and modeling networks.
- But **digital infrastructure** and **computing power** are growing…
- For the first time, we can **observe** and **analyze** large networks in the real world.

Network Analysis



Grandjean, Martin (2014). ”La connaissance est un réseau”. Les Cahiers du Numérique 10 (3): 37-54. DOI:10.3166/LCN.10.3.37-54.

Erdös-Renyi Graphs Don't Fit!

[1] S.H. Strogatz and D.J. Watts. "Collective dynamics of 'small-world' networks". In: *Nature* 393.June (1998), pp. 440–442.

Erdös-Renyi Graphs Don't Fit!

1. Real networks have high clustering coefficients $\implies$ Watts + Strogatz small world model.[1]

[1] S.H. Strogatz and D.J. Watts. "Collective dynamics of 'small-world' networks". In: *Nature* 393.June (1998), pp. 440–442.

Erdös-Renyi Graphs Don't Fit!

1. Real networks have high clustering coefficients $\implies$ Watts + Strogatz small world model.[1]
2. Real networks have highly heterogeneous *degree distributions* – heavier tails than Poisson.

[1] S.H. Strogatz and D.J. Watts. "Collective dynamics of 'small-world' networks". In: *Nature* 393.June (1998), pp. 440–442.

Emergence of Scaling in Random Networks


Albert-Laszlo Barabasi + Reka Albert (1999-2002); R.A.'s PhD thesis work.

Emergence of Scaling in Random Networks

Albert-Laszlo Barabasi + Reka Albert (1999-2002); R.A.'s PhD thesis work.

Central claim: **many real-world networks are scale-free.**

Mathematical Definition

Discrete random variable K has a **power law tail** if

$$\mathbb{P}(K = k) \approx ck^{-\gamma}$$

for some $\gamma > 1$ and "sufficiently large" k. Formally,

$$\lim_{k\to\infty} k^{\gamma}\mathbb{P}(K = k) = c\ .$$

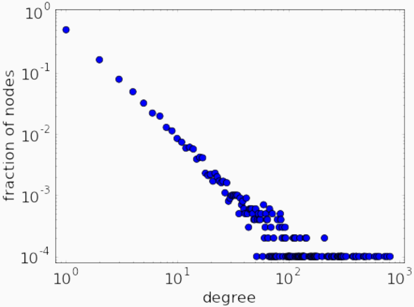
Note that $\Theta(k^{-\gamma}) \gg \Theta(e^{-k}) \gg \Theta(e^{-k^2})$.

The Log-Log Plot

Power law tails look **linear** on **log-log** axes.

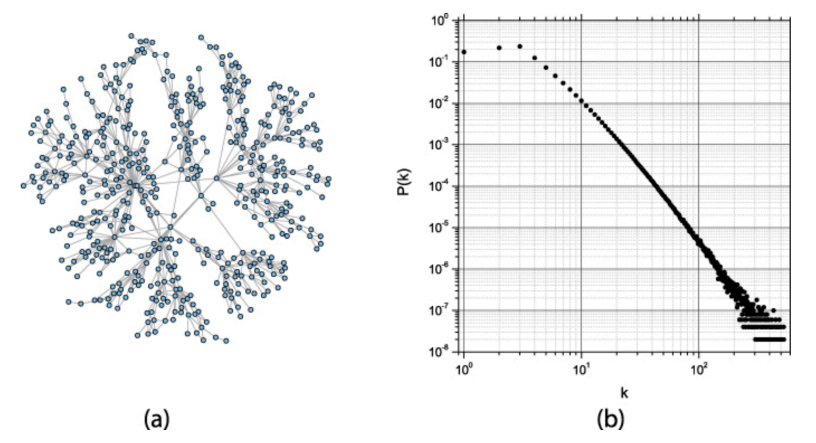
$$\mathbb{P}(K = k) \approx ck^{-\gamma}$$

$$\log \mathbb{P}(K = k) \approx -\gamma \log k + \log c\ .$$



Is everything that looks linear on log-log axes a power law?...

Scale Free Networks Have Power Law Degree Distributions



A. Lublóy and M. Szenes. “The network of corporate clients: Customer attrition at commercial banks”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.12 (2008)

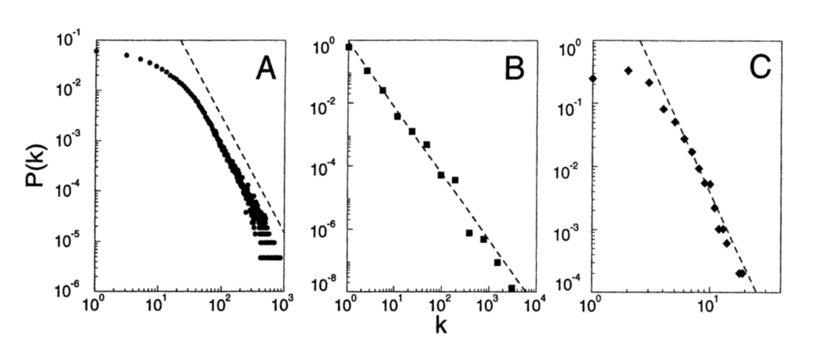
Emergence of Scaling in Random Networks

Albert-Laszlo Barabasi + Reka Albert (1999-2002); R.A.'s PhD thesis work.

Central claim: **many real-world networks are scale-free.**

Emergence of Scaling in Random Networks[2]



(A) Actor collaboration graph $\gamma \approx 2.3$, (B) World Wide Web $\gamma \approx 2.1$, (C) Power grid network, $\gamma \approx 4$.

---

[2] Albert-László Barabási and Reka Albert. "Emergence of scaling in random networks". In: *Science* 286.5439 (1999), p. 11.

Empirical Claim

Many real networks have power law degree distributions with $2 \leq \gamma \leq 3$.[3]

1. Actor collaboration: $\gamma \approx 2.3$.
2. World Wide Web: $\gamma \approx 2.1$.
3. Protein interaction: $\gamma \approx 2.2 - 2.4$.
4. Synonyms of words: $\gamma \approx 2.8$.
5. *Power grid:* $\gamma \approx 4$.

[3] Reka Albert and Albert-László Barabási. “Statistical mechanics of complex networks”. In: *Reviews of Modern Physics* 74.January (2002), p. 47.

Preferential Attachment Model (Barabasi-Albert)

1. Initialize with a fixed graph $\mathcal{G}_0$.
2. At each time-step, add a node v_t to $\mathcal{G}_{t-1}$.
3. Generate m links between v_t and $\mathcal{G}_{t-1}$, *proportional to the degree of* $\mathcal{G}$.

$$\mathbb{P}(E = (v_t, v)) = \frac{d_v}{\sum_{u \in \mathcal{G}_{t-1}} d_u}$$

"The rich get richer."

Properties of PA

For large t,

[4] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010, p. 720.

[5] S. N. Dorogovtsev, J. F. F. Mendes, and a. N. Samukhin. "Structure of Growing Networks: Exact Solution of the Barabasi–Albert's Model". In: (2000), p. 4.

Properties of PA

For large t,

1. $\mathcal{G}_t$ has $t + |\mathcal{G}_0|$ nodes.

[4] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010, p. 720.

[5] S. N. Dorogovtsev, J. F. F. Mendes, and a. N. Samukhin. “Structure of Growing Networks: Exact Solution of the Barabasi–Albert’s Model”. In: (2000), p. 4.

Properties of PA

For large t,

1. $\mathcal{G}_t$ has $t + |\mathcal{G}_0|$ nodes.
2. $\mathcal{G}_t$ mean degree m.

[4] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010, p. 720.

[5] S. N. Dorogovtsev, J. F. F. Mendes, and a. N. Samukhin. "Structure of Growing Networks: Exact Solution of the Barabasi–Albert's Model". In: (2000), p. 4.

Properties of PA

For large t,

1. $\mathcal{G}_t$ has $t + |\mathcal{G}_0|$ nodes.
2. $\mathcal{G}_t$ mean degree m.
3. The degree distribution p_k is asymptotically power-law with exponent $\gamma = 3$ (can be generalized):

[4] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010, p. 720.

[5] S. N. Dorogovtsev, J. F. F. Mendes, and a. N. Samukhin. "Structure of Growing Networks: Exact Solution of the Barabasi–Albert's Model". In: (2000), p. 4.

Properties of PA

For large t,

1. $\mathcal{G}_t$ has $t + |\mathcal{G}_0|$ nodes.
2. $\mathcal{G}_t$ mean degree m.
3. The degree distribution p_k is asymptotically power-law with exponent $\gamma = 3$ (can be generalized):

$$\mathbb{P}(K_t = k) \propto k^{-3}. \qquad \text{(asymptotic)}$$

[4] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010, p. 720.

[5] S. N. Dorogovtsev, J. F. F. Mendes, and a. N. Samukhin. "Structure of Growing Networks: Exact Solution of the Barabasi–Albert's Model". In: (2000), p. 4.

Properties of PA

For large t,

1. $\mathcal{G}_t$ has $t + |\mathcal{G}_0|$ nodes.
2. $\mathcal{G}_t$ mean degree m.
3. The degree distribution p_k is asymptotically power-law with exponent $\gamma = 3$ (can be generalized):

$$\mathbb{P}(K_t = k) \propto k^{-3}. \qquad \text{(asymptotic)}$$

Proofs: master equations[4], method of moments.[5]

[4] M. E. J. Newman. *Networks: An Introduction.* Oxford University Press, 2010, p. 720.

[5] S. N. Dorogovtsev, J. F. F. Mendes, and a. N. Samukhin. "Structure of Growing Networks: Exact Solution of the Barabasi–Albert's Model". In: (2000), p. 4.

What's interesting about these networks?

Empirical Claim

Many real networks have power law degree distributions with $2 \leq \gamma \leq 3$.

Recall:

$$\mathbb{E}[K] < \infty \iff \gamma > 2 \tag{1}$$

$$\mathrm{var}(K) < \infty \iff \gamma > 3 \tag{2}$$

What does this mean for phenomena governed by power laws?

What’s interesting about these networks?

- Scale-free networks are “**ultrasmall**”[6]: $D \sim \log\log n$.
- Scale-free networks have no **epidemic threshold**: Any rumor, no matter how silly, has nonzero probability to spread through a large proportion of a scalefree network.[7]
- Scale-free networks are **robust** to random failure.[8]

---

[6] Reuven Cohen and Shlomo Havlin. “Scale-Free Networks Are Ultrasmall”. In: *Physical Review Letters* 90.5 (2003), p. 4.

[7] Romualdo Pastor-Satorras and Alessandro Vespignani. “Epidemic spreading in scale-free networks”. In: *Physical Review Letters* 86.14 (2001), pp. 3200–3203.

[8] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi. “Error and attack tolerance of complex networks”. In: *Nature* 406 (2000), pp. 378–382.

More Models of Scale-Free Networks

- Preferential attachment model.
- Configuration model (exact degree sequence).
- Chung-Lu model (expected degrees).[9]
- Vertex-copying model.[10]

[9] F. Chung and L. Lu. "The average distances in random graphs with given expected degrees". In: *Proceedings of the National Academy of Sciences* 99.25 (2002), pp. 15879–15882.

[10] Jon M Kleinberg et al. "The Web as a Graph: Measurements, Models, and Methods Computing and Combinatorics". In: *Computing and Combinatorics: 5th Annual International Conference, COCOON'99, Tokyo, Japan, July 1999. Proceedings* 1627 (1999), pp. 1–17.

A Theory of Everything for Networks?

"Yet, probably the most surprising discovery of modern network theory is the *universality of the network topology*: Many real networks, from the cell to the Internet, independent of their age, function, and scope, converge to similar architectures. *It is this universality that allowed researchers from different disciplines to embrace network theory as a common paradigm.*"

[11] A. L. Barabási. "Scale-free networks: A decade and beyond". In: *Science* 325.5939 (2009), pp. 412–413.

Emergence of Citations of “Emergence of Scaling”

Imagine writing your PhD thesis…

Emergence of Citations of “Emergence of Scaling”

Imagine writing your PhD thesis…

- …and having it published in *Nature*, *Science*, and *Physical Review Letters*…

Emergence of Citations of “Emergence of Scaling”

Imagine writing your PhD thesis...

- ...and having it published in *Nature*, *Science*, and *Physical Review Letters*...
- ...and being cited **70,000** times....

Emergence of Citations of “Emergence of Scaling”

Imagine writing your PhD thesis...

- ...and having it published in *Nature*, *Science*, and *Physical Review Letters*...
- ...and being cited **70,000** times....
- ...10 times per day, for almost 20 years.

Emergence of Citations of “Emergence of Scaling”

Imagine writing your PhD thesis...

- ...and having it published in *Nature*, *Science*, and *Physical Review Letters*...
- ...and being cited **70,000** times....
- ...10 times per day, for almost 20 years.
- In expectation, there will be a paper published that cites their work ***during this talk.***

Emergence of Citations of “Emergence of Scaling”

Imagine writing your PhD thesis...

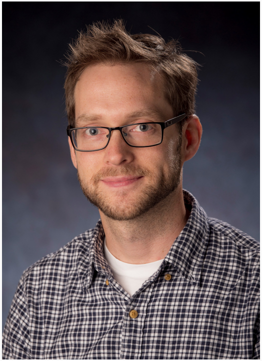
- ...and having it published in *Nature*, *Science*, and *Physical Review Letters*...
- ...and being cited **70,000** times....
- ...10 times per day, for almost 20 years.
- In expectation, there will be a paper published that cites their work ***during this talk.***

R. Albert is now a Distinguished Prof. at Penn State. Barabasi is the **5th most cited physicist** on Google Scholar.

A Tidy Picture

- **Theory**: We have many models of how scale-free networks emerge.
- **Theory**: We have many interesting theoretical properties of scale-free networks.
- **Data**: We have many papers observing power laws in real-world data…right?

Well...



Aaron Clauset, UC Boulder

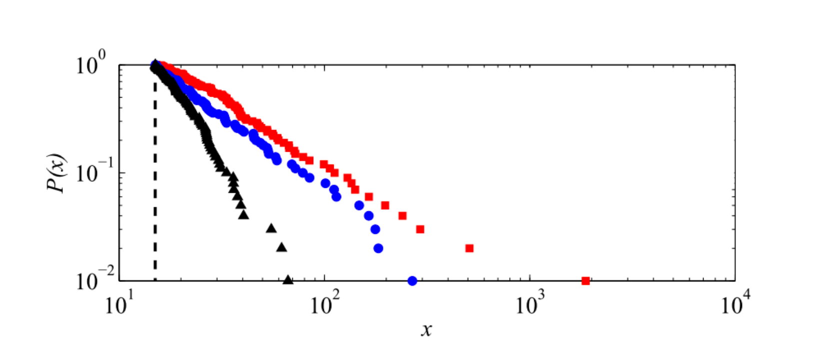
"There must be a thousand papers in which people plot the degree distribution, put a line through it and say it's scale-free without really doing the careful statistical work."

[12] Quanta: "Scant Evidence of Power Laws Found in Real-World Networks."

Measuring Power Laws

The Problem with Log-Log Plots...

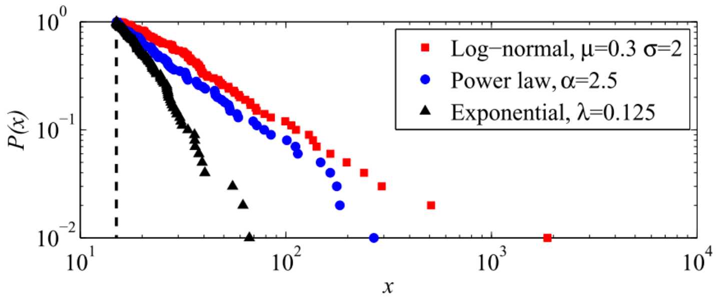
Empirical CDF of 100 samples from a degree distribution. How many are power laws?



Aaron Clauset, C R Shalizi, and M E J Newman. “Power-law distributions in empirical data”. In: *SIAM Review* 51.4 (2009), pp. 661–703

The Problem with Log-Log Plots...

Empirical CDF of 100 samples from a degree distribution. How many are power laws?



Aaron Clauset, C R Shalizi, and M E J Newman. “Power-law distributions in empirical data”. In: *SIAM Review* 51.4 (2009), pp. 661–703

We need statistics...

Inference

What are the power law parameters that best fit the data?

Model Evaluation

Is the power law a plausible model of the data at all?

“Scale Free Networks are Rare”


[13] Anna Broido and Aaron Clauset. “Scale-free networks are rare”. In: *arXiv:1801.03400* (2018), pp. 1–14.

Methodology

1. Collect large set of network data sets.[14]
2. For each data set, perform inference and model evaluation as described above.
3. **Compare** power law to a set of alternative models (exponential, log-normal, etc) using an *likelihood ratio test*.

[14] https://icon.colorado.edu/

Some Startling Findings

1. For $573/882 = 65\%$ of data sets, power law is either implausible or worse than a simple alternative model.
2. For only 11% of data sets is a power law plausible, clearly favored over simple alternatives, and displays $2 \leq \gamma \leq 3$.

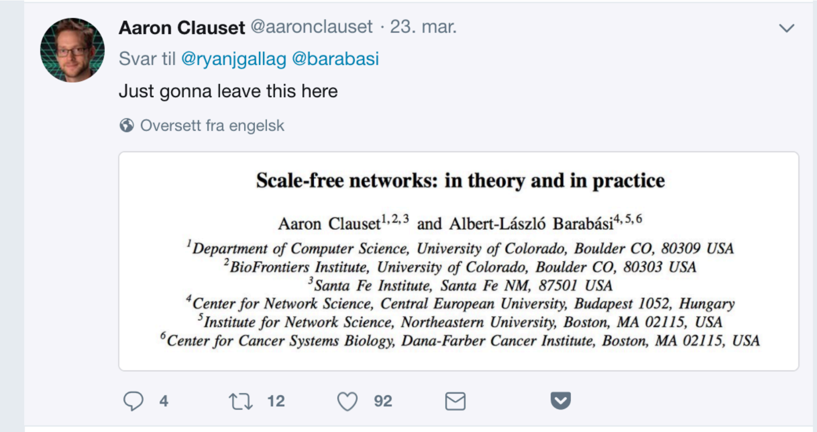
But...

Even ground-truth preferential attachment networks fail many of Broido + Clauset's tests...

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.

Ongoing Debate...

"...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 of scale-freeness fails the most elementary tests. And those are only the big problems..."


Aaron Clauset @aaronclauset · 23. mar.
Svar til @ryanjgallag @barabasi
Just gonna leave this here
Oversett fra engelsk
Scale-free networks: in theory and in practice
Aaron Clauset[1,2,3] and Albert-László Barabási[4,5,6]
[1]Department of Computer Science, University of Colorado, Boulder CO, 80309 USA
[2]BioFrontiers Institute, University of Colorado, Boulder CO, 80303 USA
[3]Santa Fe Institute, Santa Fe NM, 87501 USA
[4]Center for Network Science, Central European University, Budapest 1052, Hungary
[5]Institute for Network Science, Northeastern University, Boston, MA 02115, USA
[6]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA
4
12
92

Thanks!