Scale-free networks

Last updated: 2020/09/12, 14:01:53 EDT

Principles of Complex Systems, Vol. 1 | @pocsvox
CSYS/MATH 300, Fall, 2020

Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

These slides are brought to you by:

Sealie & Lambie
Productions

PoCS
Principles of
Complex Systems
@pocsvox
What's the Story?

UVM

These slides are also brought to you by:

Special Guest Executive Producer

📷 On Instagram at pratchett_the_cat ↗

Outline

PoCS
Principles of
Complex Systems
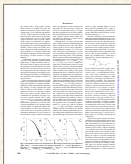@pocsvox
What's the Story?

THE NETWORKS FREE OF SCALE

Scale-free networks

- Networks with power-law degree distributions have become known as scale-free networks.

- Scale-free refers specifically to the degree distribution having a power-law decay in its tail:

$$P_k \sim k^{-\gamma} \text{ for 'large' } k$$

- One of the seminal works in complex networks:

  "Emergence of scaling in random networks" ☑
  Barabási and Albert,
  Science, **286**, 509–511, 1999. [2]

  Times cited: $\sim 23,532$ ☑ (as of October 8, 2015)

- Somewhat misleading nomenclature...

Scale-free networks

- Scale-free networks are not fractal in any sense.
- Usually talking about networks whose links are abstract, relational, informational, ...(non-physical)
- Primary example: hyperlink network of the Web
- Much arguing about whether or networks are 'scale-free' or not...

Some real data (we are feeling brave):

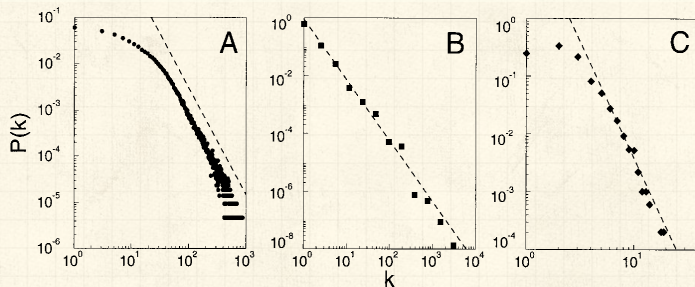From Barabási and Albert's original paper [2]:

Fig. 1. The distribution function of connectivities for various large networks. (**A**) Actor collaboration graph with $N = 212,250$ vertices and average connectivity $\langle k \rangle = 28.78$. (**B**) WWW, $N = 325,729$, $\langle k \rangle = 5.46$ (6). (**C**) Power grid data, $N = 4941$, $\langle k \rangle = 2.67$. The dashed lines have slopes (A) $\gamma_{actor} = 2.3$, (B) $\gamma_{www} = 2.1$ and (C) $\gamma_{power} = 4$.

Random networks: largest components
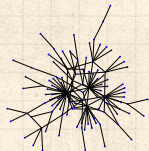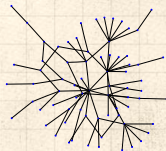
$\gamma = 2.5$
$\langle k \rangle = 1.8$

$\gamma = 2.5$
$\langle k \rangle = 2.05333$

$\gamma = 2.5$
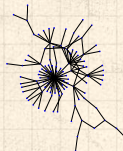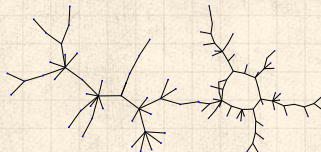$\langle k \rangle = 1.66667$

$\gamma = 2.5$
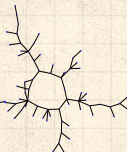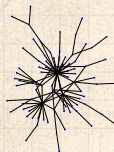$\langle k \rangle = 1.92$

$\gamma = 2.5$
$\langle k \rangle = 1.6$

$\gamma = 2.5$
$\langle k \rangle = 1.50667$

$\gamma = 2.5$
$\langle k \rangle = 1.62667$

$\gamma = 2.5$
$\langle k \rangle = 1.8$

Scale-free networks

The big deal:

🎲 We move beyond describing networks to finding mechanisms for why certain networks are the way they are.

A big deal for scale-free networks:

🎲 How does the exponent γ depend on the mechanism?

🎲 Do the mechanism details matter?

BA model

🦀 Barabási-Albert model = BA model.

🦀 Key ingredients:
Growth and Preferential Attachment (PA).

🦀 Step 1: start with m_0 disconnected nodes.

🦀 Step 2:
1. Growth—a new node appears at each time step $t = 0, 1, 2, ...$.
2. Each new node makes m links to nodes already present.
3. Preferential attachment—Probability of connecting to ith node is $\propto k_i$.

🦀 In essence, we have a rich-gets-richer scheme.

🦀 Yes, we've seen this all before in Simon's model.

BA model

Definition: A_k is the attachment kernel for a node with degree k.

For the original model:

$$A_k = k$$

Definition: $P_{\text{attach}}(k,t)$ is the attachment probability.

For the original model:

$$P_{\text{attach}}(\text{node } i, t) = \frac{k_i(t)}{\sum_{j=1}^{N(t)} k_j(t)} = \frac{k_i(t)}{\sum_{k=0}^{k_{\max}(t)} k N_k(t)}$$

where $N(t) = m_0 + t$ is # nodes at time t
and $N_k(t)$ is # degree k nodes at time t.

Approximate analysis

🎲 When $(N+1)$th node is added, the expected increase in the degree of node i is

$$E(k_{i,N+1} - k_{i,N}) \simeq m \frac{k_{i,N}}{\sum_{j=1}^{N(t)} k_j(t)}.$$

🎲 Assumes probability of being connected to is small.

🎲 Dispense with Expectation by assuming (hoping) that over longer time frames, degree growth will be smooth and stable.

🎲 Approximate $k_{i,N+1} - k_{i,N}$ with $\frac{\mathrm{d}}{\mathrm{d}t} k_{i,t}$:

$$\frac{\mathrm{d}}{\mathrm{d}t} k_{i,t} = m \frac{k_i(t)}{\sum_{j=1}^{N(t)} k_j(t)}$$

where $t = N(t) - m_0$.

Deal with denominator: each added node brings m new edges.

$$\therefore \sum_{j=1}^{N(t)} k_j(t) = 2tm$$

The node degree equation now simplifies:

$$\frac{\mathrm{d}}{\mathrm{d}t} k_{i,t} = m \frac{k_i(t)}{\sum_{j=1}^{N(t)} k_j(t)} = m \frac{k_i(t)}{2mt} = \frac{1}{2t} k_i(t)$$

Rearrange and solve:

$$\frac{\mathrm{d}k_i(t)}{k_i(t)} = \frac{\mathrm{d}t}{2t} \Rightarrow \boxed{k_i(t) = c_i\, t^{1/2}.}$$

Next find c_i ...

Know ith node appears at time

$$t_{i,\text{start}} = \begin{cases} i - m_0 & \text{for } i > m_0 \\ 0 & \text{for } i \leq m_0 \end{cases}$$

So for $i > m_0$ (exclude initial nodes), we must have

$$k_i(t) = m \left(\frac{t}{t_{i,\text{start}}} \right)^{1/2} \text{ for } t \geq t_{i,\text{start}}.$$

All node degrees grow as $t^{1/2}$ but later nodes have larger $t_{i,\text{start}}$ which flattens out growth curve.

First-mover advantage: Early nodes do best.

Clearly, a Ponzi scheme ⟶.

We are already at the Zipf distribution:

❧ Degree of node i is the size of the ith ranked node:

$$k_i(t) = m \left(\frac{t}{t_{i,\text{start}}} \right)^{1/2} \text{ for } t \geq t_{i,\text{start}}.$$

❧ From before:

$$t_{i,\text{start}} = \begin{cases} i - m_0 & \text{for } i > m_0 \\ 0 & \text{for } i \leq m_0 \end{cases}$$

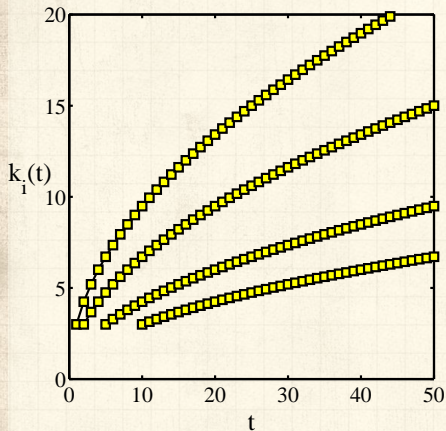so $t_{i,\text{start}} \sim i$ which is the rank.

❧ We then have:

$$k_i \propto i^{-1/2} = i^{-\alpha}.$$

❧ Our connection $\alpha = 1/(\gamma - 1)$ or $\gamma = 1 + 1/\alpha$ then gives

$$\boxed{\gamma = 1 + 1/(1/2) = 3.}$$

🧊 $m = 3$

🧊 $t_{i,\text{start}} = 1, 2, 5,$ and 10.

Degree distribution

🐾 So what's the degree distribution at time t?

🐾 Use fact that birth time for added nodes is
distributed uniformly between time 0 and t:

$$\mathbf{Pr}(t_{i,\text{start}})\mathrm{d}t_{i,\text{start}} \simeq \frac{\mathrm{d}t_{i,\text{start}}}{t}$$

🐾 Also use

$$k_i(t) = m\left(\frac{t}{t_{i,\text{start}}}\right)^{1/2} \Rightarrow t_{i,\text{start}} = \frac{m^2 t}{k_i(t)^2}.$$

Transform variables—Jacobian:

$$\frac{\mathrm{d}t_{i,\text{start}}}{\mathrm{d}k_i} = -2\frac{m^2 t}{k_i(t)^3}.$$

Degree distribution

$$\mathbf{Pr}(k_i)\mathrm{d}k_i = \mathbf{Pr}(t_{i,\text{start}})\mathrm{d}t_{i,\text{start}}$$

$$= \mathbf{Pr}(t_{i,\text{start}})\mathrm{d}k_i \left| \frac{\mathrm{d}t_{i,\text{start}}}{\mathrm{d}k_i} \right|$$

$$= \frac{1}{t}\mathrm{d}k_i\, 2\frac{m^2 t}{k_i(t)^3}$$

$$= 2\frac{m^2}{k_i(t)^3}\mathrm{d}k_i$$

$$\propto k_i^{-3}\mathrm{d}k_i\,.$$

Degree distribution

- We thus have a very specific prediction of $\mathbf{Pr}(k) \sim k^{-\gamma}$ with $\gamma = 3$.

- Typical for real networks: $2 < \gamma < 3$.

- Range true more generally for events with size distributions that have power-law tails.

- $2 < \gamma < 3$: finite mean and 'infinite' variance (wild)

- In practice, $\gamma < 3$ means variance is governed by upper cutoff.

- $\gamma > 3$: finite mean and variance (mild)

Back to that real data:
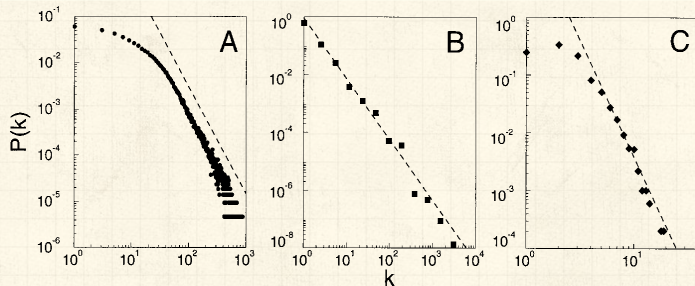
From Barabási and Albert's original paper[2]:

Fig. 1. The distribution function of connectivities for various large networks. (**A**) Actor collaboration graph with $N = 212,250$ vertices and average connectivity $\langle k \rangle = 28.78$. (**B**) WWW, $N = 325,729$, $\langle k \rangle = 5.46$ (6). (**C**) Power grid data, $N = 4941$, $\langle k \rangle = 2.67$. The dashed lines have slopes (A) $\gamma_{actor} = 2.3$, (B) $\gamma_{www} = 2.1$ and (C) $\gamma_{power} = 4$.

Examples

Web	$\gamma \simeq 2.1$ for in-degree
Web	$\gamma \simeq 2.45$ for out-degree
Movie actors	$\gamma \simeq 2.3$
Words (synonyms)	$\gamma \simeq 2.8$

The Internets is a different business...

Things to do and questions

⬡ Vary attachment kernel.

⬡ Vary mechanisms:
  1. Add edge deletion
  2. Add node deletion
  3. Add edge rewiring

⬡ Deal with directed versus undirected networks.

⬡ Important Q.: Are there distinct universality classes for these networks?

⬡ Q.: How does changing the model affect γ?

⬡ Q.: Do we need preferential attachment and growth?

⬡ Q.: Do model details matter? Maybe ...

Preferential attachment

- Let's look at preferential attachment (PA) a little more closely.

- PA implies arriving nodes have complete knowledge of the existing network's degree distribution.

- For example: If $P_{\text{attach}}(k) \propto k$, we need to determine the constant of proportionality.

- We need to know what everyone's degree is...

- PA is ∴ an outrageous assumption of node capability.

- But a very simple mechanism saves the day...

Preferential attachment through randomness

- Instead of attaching preferentially, allow new nodes to attach randomly.

- Now add an extra step: new nodes then connect to some of their friends' friends.

- Can also do this at random.

- Assuming the existing network is random, we know probability of a random friend having degree k is

$$Q_k \propto k P_k$$

- So rich-gets-richer scheme can now be seen to work in a natural way.

Robustness

PoCS, Vol. 1
@pocsvox
Scale-free
networks

- Albert et al., Nature, 2000:
  "Error and attack tolerance of complex
  networks"[1]

- Standard random networks (Erdős-Rényi)
  versus Scale-free networks:

Scale-free
networks
Main story
Model details
Analysis
A more plausible
mechanism
Robustness
Krapivsky & Redner's
model
Generalized model
Analysis
Universality?
Sublinear attachment
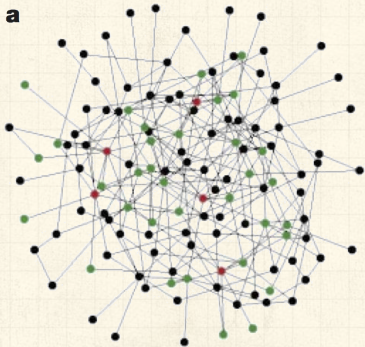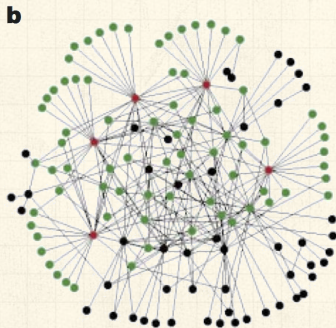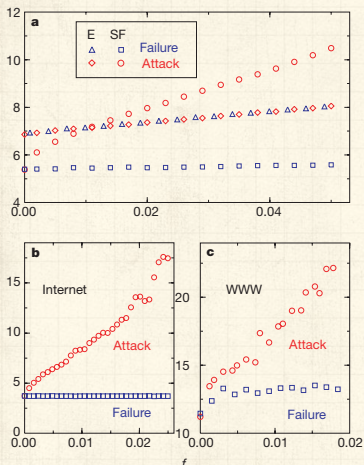kernels
Superlinear attachment
kernels
Nutshell

References

Exponential

Scale-free

from Albert et al., 2000

31 of 57

Robustness

from Albert et al., 2000

- Plots of network diameter as a function of fraction of nodes removed
- Erdős-Rényi versus scale-free networks
- blue symbols = random removal
- red symbols = targeted removal (most connected first)

Robustness

- Scale-free networks are thus robust to random failures yet fragile to targeted ones.
- All very reasonable: Hubs are a big deal.
- But: next issue is whether hubs are vulnerable or not.
- Representing all webpages as the same size node is obviously a stretch (e.g., google vs. a random person's webpage)
- Most connected nodes are either:
  1. Physically larger nodes that may be harder to 'target'
  2. or subnetworks of smaller, normal-sized nodes.
- Need to explore cost of various targeting schemes.

Robustness

Not a robust paper:

"The "Robust yet Fragile" nature of the Internet" ↗
Doyle et al.,
Proc. Natl. Acad. Sci., **2005**, 14497–14502,
2005. [3]

⬢ HOT networks versus scale-free networks

⬢ Same degree distributions, different arrangements.

⬢ Doyle *et al.* take a look at the actual Internet.

Generalized model

Fooling with the mechanism:

🪲 2001: Krapivsky & Redner (KR) [4] explored the general attachment kernel:

$$\mathbf{Pr}(\text{attach to node } i) \propto A_k = k_i^\nu$$

where A_k is the attachment kernel and $\nu > 0$.

🪲 KR also looked at changing the details of the attachment kernel.

🪲 KR model will be fully studied in CoNKS.

Generalized model

🎲 We'll follow KR's approach using rate equations☐.

🎲 Here's the set up:

$$\frac{\mathrm{d}N_k}{\mathrm{d}t} = \frac{1}{A}\left[A_{k-1}N_{k-1} - A_k N_k\right] + \delta_{k1}$$

where N_k is the number of nodes of degree k.

1. One node with one link is added per unit time.
2. The first term corresponds to degree $k-1$ nodes becoming degree k nodes.
3. The second term corresponds to degree k nodes becoming degree $k-1$ nodes.
4. A is the correct normalization (coming up).
5. Seed with some initial network (e.g., a connected pair)
6. Detail: $A_0 = 0$

Generalized model

❧ In general, probability of attaching to a specific node of degree k at time t is

$$\textbf{Pr}(\text{attach to node } i) = \frac{A_k}{A(t)}$$

where $A(t) = \sum_{k=1}^{\infty} A_k N_k(t)$.

❧ E.g., for BA model, $A_k = k$ and $A = \sum_{k=1}^{\infty} k N_k(t)$.

❧ For $A_k = k$, we have

$$A(t) = \sum_{k'=1}^{\infty} k' N_{k'}(t) = 2t$$

since one edge is being added per unit time.

❧ Detail: we are ignoring initial seed network's edges.

Generalized model

So now

$$\frac{dN_k}{dt} = \frac{1}{A} \left[A_{k-1} N_{k-1} - A_k N_k \right] + \delta_{k1}$$

becomes

$$\frac{dN_k}{dt} = \frac{1}{2t} \left[(k-1) N_{k-1} - k N_k \right] + \delta_{k1}$$

As for BA method, look for steady-state growing solution: $N_k = n_k t$.

We replace dN_k/dt with $dn_k t/dt = n_k$.

We arrive at a difference equation:

$$n_k = \frac{1}{2t} \left[(k-1) n_{k-1} t - k n_k t \right] + \delta_{k1}$$

Universality?

As expected, we have the same result as for the BA model:

$$N_k(t) = n_k(t)t \propto k^{-3}t \text{ for large } k.$$

Now: what happens if we start playing around with the attachment kernel A_k?

Again, we're asking if the result $\gamma = 3$ universal ☞?

KR's natural modification: $A_k = k^\nu$ with $\nu \neq 1$.

But we'll first explore a more subtle modification of A_k made by Krapivsky/Redner [4]

Keep A_k linear in k but tweak details.

Idea: Relax from $A_k = k$ to $A_k \sim k$ as $k \to \infty$.

Universality?

📦 Recall we used the normalization:

$$A(t) = \sum_{k'=1}^{\infty} k' N_{k'}(t) \simeq 2t \text{ for large } t.$$

📦 We now have

$$A(t) = \sum_{k'=1}^{\infty} A_{k'} N_{k'}(t)$$

where we only know the asymptotic behavior of A_k.

📦 We assume that $A = \mu t$

📦 We'll find μ later and make sure that our assumption is consistent.

📦 As before, also assume $N_k(t) = n_k t$.

Universality?

For $A_k = k$ we had

$$n_k = \frac{1}{2}\left[(k-1)n_{k-1} - kn_k\right] + \delta_{k1}$$

This now becomes

$$n_k = \frac{1}{\mu}\left[A_{k-1}n_{k-1} - A_k n_k\right] + \delta_{k1}$$

$$\Rightarrow (A_k + \mu)n_k = A_{k-1}n_{k-1} + \mu\delta_{k1}$$

Again two cases:

$$k = 1 : n_1 = \frac{\mu}{\mu + A_1}; \qquad k > 1 : n_k = n_{k-1}\frac{A_{k-1}}{\mu + A_k}.$$

Universality?

- Time for pure excitement: Find asymptotic behavior of n_k given $A_k \to k$ as $k \to \infty$.

- For large k, we find:

$$n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \frac{1}{1 + \frac{\mu}{A_j}} \propto k^{-\mu-1}$$

- Since μ depends on A_k, details matter...

Universality?

- Now we need to find μ.
- Our assumption again: $A = \mu t = \sum_{k=1}^{\infty} N_k(t)A_k$
- Since $N_k = n_k t$, we have the simplification $\mu = \sum_{k=1}^{\infty} n_k A_k$
- Now subsitute in our expression for n_k:

$$1\mu = \sum_{k=1}^{\infty} \frac{\mu}{A_k} \prod_{j=1}^{k} \frac{1}{1 + \frac{\mu}{A_j}} A_k$$

- Closed form expression for μ.
- We can solve for μ in some cases.
- Our assumption that $A = \mu t$ looks to be not too horrible.

Universality?

- Consider tunable $A_1 = \alpha$ and $A_k = k$ for $k \geq 2$.
- Again, we can find $\gamma = \mu + 1$ by finding μ.
- Closed form expression for μ:

$$\frac{\mu}{\alpha} = \sum_{k=2}^{\infty} \frac{\Gamma(k+1)\Gamma(2+\mu)}{\Gamma(k+\mu+1)}$$

#mathisfun

$$\mu(\mu - 1) = 2\alpha \Rightarrow \mu = \frac{1 + \sqrt{1 + 8\alpha}}{2}.$$

- Since $\gamma = \mu + 1$, we have

$$0 \leq \alpha < \infty \Rightarrow 2 \leq \gamma < \infty$$

- Craziness...

Sublinear attachment kernels

- Rich-get-somewhat-richer:

$$A_k \sim k^\nu \text{ with } 0 < \nu < 1.$$

- General finding by Krapivsky and Redner: [4]

$$n_k \sim k^{-\nu} e^{-c_1 k^{1-\nu}} + \text{correction terms}.$$

- Stretched exponentials (truncated power laws).
- aka Weibull distributions.
- Universality: now details of kernel do not matter.
- Distribution of degree is universal providing $\nu < 1$.

Sublinear attachment kernels

Details:

  🎲 For $1/2 < \nu < 1$:

$$n_k \sim k^{-\nu} e^{-\mu \left(\frac{k^{1-\nu} - 2^{1-\nu}}{1-\nu} \right)}$$

  🎲 For $1/3 < \nu < 1/2$:

$$n_k \sim k^{-\nu} e^{-\mu \frac{k^{1-\nu}}{1-\nu} + \frac{\mu^2}{2} \frac{k^{1-2\nu}}{1-2\nu}}$$

  🎲 And for $1/(r+1) < \nu < 1/r$, we have r pieces in exponential.

Superlinear attachment kernels

Rich-get-much-richer:

$$A_k \sim k^\nu \text{ with } \nu > 1.$$

Now a winner-take-all mechanism.

One single node ends up being connected to almost all other nodes.

For $\nu > 2$, all but a finite # of nodes connect to one node.

Nutshell:

Overview Key Points for Models of Networks:

🎲 Obvious connections with the vast extant field of graph theory.

🎲 But focus on dynamics is more of a physics/stat-mech/comp-sci flavor.

🎲 Two main areas of focus:
   1. Description: Characterizing very large networks
   2. Explanation: Micro story $\Rightarrow$ Macro features

🎲 Some essential structural aspects are understood: degree distribution, clustering, assortativity, group structure, overall structure,...

🎲 Still much work to be done, especially with respect to dynamics... #excitement

References I

[1] R. Albert, H. Jeong, and A.-L. Barabási.
Error and attack tolerance of complex networks.
Nature, 406:378–382, 2000. pdf

[2] A.-L. Barabási and R. Albert.
Emergence of scaling in random networks.
Science, 286:509–511, 1999. pdf

[3] J. Doyle, D. Alderson, L. Li, S. Low, M. Roughan,
S. S., R. Tanaka, and W. Willinger.
The "Robust yet Fragile" nature of the Internet.
Proc. Natl. Acad. Sci., 2005:14497–14502, 2005.
pdf

[4] P. L. Krapivsky and S. Redner.
Organization of growing random networks.
Phys. Rev. E, 63:066123, 2001. pdf