Scale-Free Networks

Complex Networks
CSYS/MATH 303, Spring, 2011

Prof. Peter Dodds

Department of Mathematics & Statistics
Center for Complex Systems
Vermont Advanced Computing Center
University of Vermont

Outline

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:

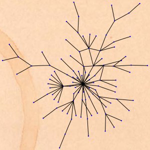
$$P_k \sim k^{-\gamma} \text{ for 'large' } k$$

- One of the seminal works in complex networks: Laszlo Barabási and Reka Albert, Science, 1999: "Emergence of scaling in random networks" [2]
- Somewhat misleading nomenclature...

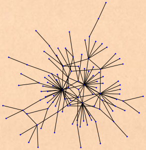
Scale-free networks

- 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. . .

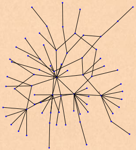
Random networks: largest components



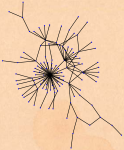
$\gamma = 2.5$
$\langle k \rangle = 1.8$



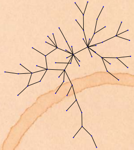
$\gamma = 2.5$
$\langle k \rangle = 2.05333$



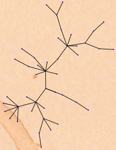
$\gamma = 2.5$
$\langle k \rangle = 1.66667$



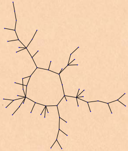
$\gamma = 2.5$
$\langle k \rangle = 1.92$



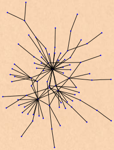
$\gamma = 2.5$
$\langle k \rangle = 1.6$



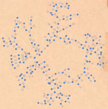
$\gamma = 2.5$
$\langle k \rangle = 1.50667$



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



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



Scale-free networks

The big deal:

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

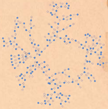
A big deal for scale-free networks:

- How does the exponent γ depend on the mechanism?
- Do the mechanism details matter?

Heritage

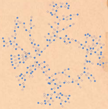
Work that presaged scale-free networks

- 1924: G. Udny Yule [9]:
  # Species per Genus
- 1926: Lotka [4]:
  # Scientific papers per author
- 1953: Mandelbrot [5]):
  Zipf's law for word frequency through optimization
- 1955: Herbert Simon [8, 10]:
  Zipf's law, city size, income, publications, and species per genus
- 1965/1976: Derek de Solla Price [6, 7]:
  Network of Scientific Citations



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 $t = 0, 1, 2, \ldots$.
  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.



BA model

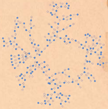
- Definition: A_k is the attachment kernel for a node with degree k.
- For the original model:

$$A_k = k$$

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

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

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



Approximate analysis

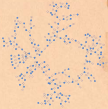
- When $(N+1)$th node is added, the expected increase in the degree of node i is

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

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



The University of Vermont

Approximate analysis

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

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

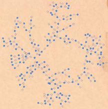
- The node degree equation now simplifies:

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

- 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 ...



The UNIVERSITY of VERMONT

Approximate analysis

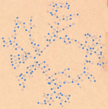
- Know ith node appears at time

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

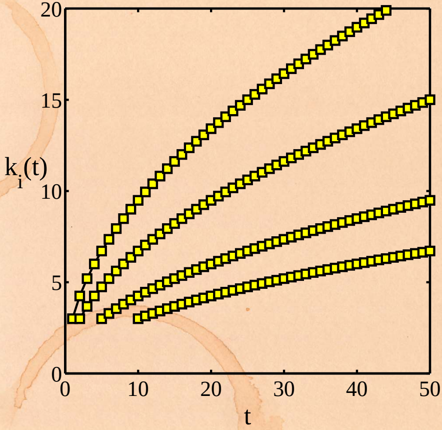
- So for $i > m_0$ (exclude initial nodes), we must have

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

- All node degrees grow as $t^{1/2}$ but later nodes have larger $t_{i,\text{start}}$ which flattens out growth curve.
- Early nodes do best (First-mover advantage).



Approximate analysis



- $m = 3$
- $t_{i,\mathrm{start}} =$ $1, 2, 5,$ and 10.

Degree distribution

- So what's the degree distribution at time t?
- Use fact that birth time for added nodes is distributed uniformly:

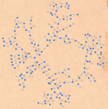
$$P(t_{i,\text{start}})\mathrm{d}t_{i,\text{start}} \simeq \frac{\mathrm{d}t_{i,\text{start}}}{t + m_0}$$

- Using

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

and by understanding that later arriving nodes have lower degrees, we can say this:

$$\mathbf{Pr}(k_i < k) = \mathbf{Pr}(t_{i,\text{start}} > \frac{m^2 t}{k^2}).$$



The UNIVERSITY of VERMONT

Degree distribution

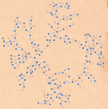
- Using the uniformity of start times:

$$\mathbf{Pr}(k_i < k) = \mathbf{Pr}(t_{i,\mathrm{start}} > \frac{m^2 t}{k^2}) \simeq \frac{t - \frac{m^2 t}{k^2}}{t + m_0}.$$

- Differentiate to find $\mathbf{Pr}(k)$:

$$\mathbf{Pr}(k) = \frac{\mathrm{d}}{\mathrm{d}k}\mathbf{Pr}(k_i < k) = \frac{2m^2 t}{(t + m_0)k^3}$$

$$\sim 2m^2 k^{-3} \text{ as } m \to \infty.$$



The University of Vermont

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)

Examples

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

The Internets is a different business...

Real data

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

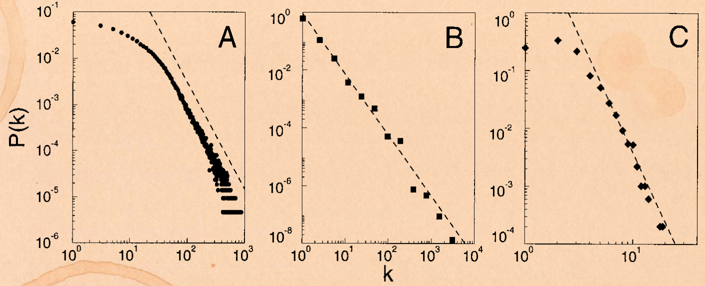
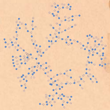
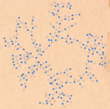


Fig. 1. The distribution function of connectivities for various large networks. (**A**) Actor collaboration graph with $N = 212{,}250$ vertices and average connectivity $\langle k \rangle = 28.78$. (**B**) WWW, $N = 325{,}729$, $\langle k \rangle = 5.46$ (6). (**C**) Power grid data, $N = 4941$, $\langle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



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?
- The answer is (surprisingly) yes.



Preferential attachment

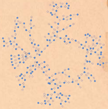
- Let's look at preferential attachment (PA) a little more closely.
- PA implies arriving nodes have complete knowledge of the existing network's degree distribution.
- For example: If $P_{\mathrm{attach}}(k) \propto k$, we need to determine the constant of proportionality.
- We need to know what everyone's degree is...
- PA is $\therefore$ 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.
- We know that friends are weird...
- Assuming the existing network is random, we know probability of a random friend having degree k is

$$Q_k \propto kP_k$$

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

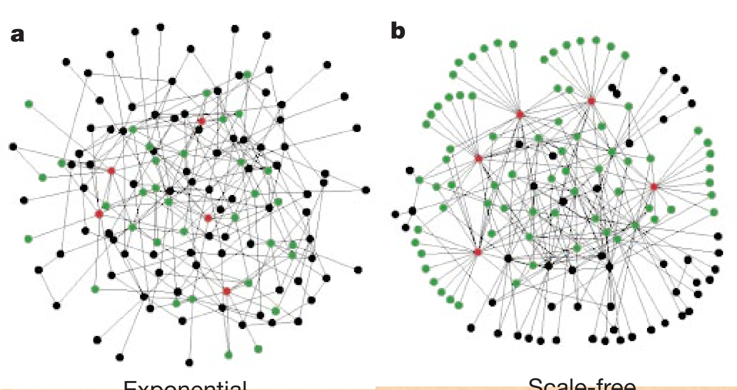


Robustness

- We've looked at some aspects of contagion on scale-free networks:
  1. Facilitate disease-like spreading.
  2. Inhibit threshold-like spreading.
- Another simple story concerns system robustness.
- Albert et al., Nature, 2000:
  "Error and attack tolerance of complex networks" [1]

Robustness

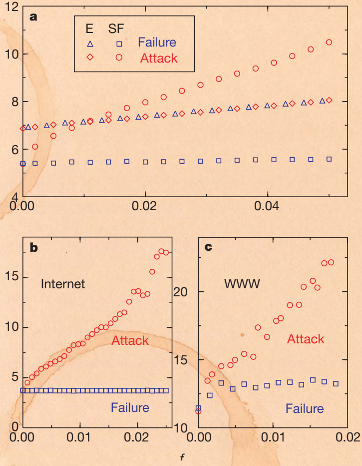
- Standard random networks (Erdős-Rényi) versus Scale-free networks



Exponential Scale-free

Albert et al., 2000

Robustness



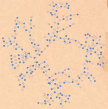
from

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.



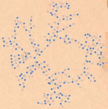
Generalized model

Fooling with the mechanism:

- 2001: Redner & Krapivsky (RK) [3] 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 $\nu > 0$.
- RK also looked at changing the details of the attachment kernel.
- We'll follow RK's approach using rate equations (⊞).



Generalized model

- Here's the set up:

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

where N_k is the number of nodes of degree k.

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

Generalized model

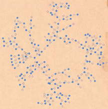
- In general, probability of attaching to a specific node of degree k at time t is

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

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

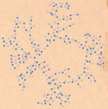
$$\frac{\mathrm{d}N_k}{\mathrm{d}t} = \frac{1}{A}\left[A_{k-1}N_{k-1} - A_k N_k\right] + \delta_{k1}$$

becomes

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

- As for BA method, look for steady-state growing solution: $N_k = n_k t$.
- We replace $\mathrm{d}N_k/\mathrm{d}t$ with $\mathrm{d}n_k t/\mathrm{d}t = n_k$.
- We arrive at a difference equation:

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



Generalized model

- Rearrange and simply:

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

$$\Rightarrow (k+2)n_k = (k-1)n_{k-1} + 2\delta_{k1}$$

- Two cases:

$$k = 1 : n_1 = 2/3 \text{ since } n_0 = 0$$

$$k > 1 : n_k = \frac{(k-1)}{k+2} n_{k-1}$$

Generalized model

- Now find n_k:

$$k > 1 : n_k = \frac{(k-1)}{k+2} n_{k-1} = \frac{(k-1)}{k+2}\frac{(k-2)}{k+1} n_{k-2}$$

$$= \frac{(k-1)}{k+2}\frac{(k-2)}{k+1}\frac{(k-3)}{k} n_{k-3}$$

$$= \frac{(k-1)}{k+2}\frac{(k-2)}{k+1}\frac{(k-3)}{k}\frac{(k-4)}{k-1} n_{k-4}$$

$$= \frac{\cancel{(k-1)}}{k+2}\frac{\cancel{(k-2)}}{k+1}\frac{\cancel{(k-3)}}{k}\frac{\cancel{(k-4)}}{\cancel{(k-1)}}\frac{\cancel{(k-5)}}{\cancel{(k-2)}}\frac{\cdots}{\cdots}\frac{\cancel{5}}{\cancel{8}}\frac{\cancel{4}}{\cancel{7}}\frac{3}{\cancel{6}}\frac{2}{\cancel{5}}\frac{1}{\cancel{4}} n_1$$

$$\Rightarrow n_k = \frac{6}{k(k+1)(k+2)} n_1 = \frac{4}{k(k+1)(k+2)} \sim k^{-3}$$

Universality?

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

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

- Now: what happens if we start playing around with the attachment kernel A_k?
- Again, is the result $\gamma = 3$ universal (⊞)?
- Natural modification: $A_k = k^{\nu}$ with $\nu \neq 1$.
- But we'll first explore a more subtle modification of A_k made by Redner/Krapivsky [3]
- 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:

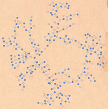
$$A(t) = \sum_{k'=1}^{\infty} k' N_{k'}(t) \simeq 2t \text{ for large } t.$$

- 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

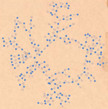
$$n_k = \frac{1}{\mu}\left[A_{k-1}n_{k-1} - A_k n_k\right] + \delta_{k1}$$

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

- Again two cases:

$$k = 1 : n_1 = \frac{\mu}{\mu + A_1}.$$

$$k > 1 : n_k = n_{k-1}\frac{A_{k-1}}{\mu + A_k}$$



Universality?

- Dealing with the $k > 1$ case:

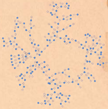
$$n_k = n_{k-1}\frac{A_{k-1}}{\mu + A_k} = n_{k-1}\frac{A_{k-1}}{A_k}\frac{1}{1+\frac{\mu}{A_k}}$$

$$= n_{k-2}\frac{A_{k-2}}{\cancel{A_{k-1}}}\frac{1}{1+\frac{\mu}{A_{k-1}}}\frac{\cancel{A_{k-1}}}{A_k}\frac{1}{1+\frac{\mu}{A_k}}$$

$$= n_1\frac{A_1}{A_k}\prod_{j=2}^{k}\frac{1}{1+\frac{\mu}{A_j}}$$

$$= n_1\frac{A_1}{A_k}\left(1+\frac{\mu}{A_1}\right)\prod_{j=1}^{k}\frac{1}{1+\frac{\mu}{A_j}}$$

$$= \frac{\mu}{A_k}\prod_{j=1}^{k}\frac{1}{1+\frac{\mu}{A_j}} \text{ since } n_1 = \mu/(\mu + A_1)$$



Universality?

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

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

$$= \frac{\mu}{\cancel{A_k}} \frac{A_1}{(A_1 + \mu)} \frac{A_2}{(A_2 + \mu)} \cdots \frac{k-1}{(k-1+\mu)} \frac{\cancel{k}}{(k+\mu)}$$

$$\propto \frac{\Gamma(k)}{\Gamma(k+\mu+1)} \sim \frac{\sqrt{2\pi} k^{k+1/2} e^{-k}}{\sqrt{2\pi} (k+\mu+1)^{k+\mu+1+1/2} e^{-(k+\mu+1)}}$$

$$\propto k^{-\mu-1}$$

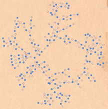
- Since μ depends on A_k, details matter...

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:

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

- Closed form expression for μ.
- We can solve for μ in some cases.
- Our assumption that $A = \mu t$ is okay.

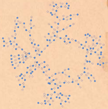


Universality?

- Amazingly, we can adjust A_k and tune γ to be anywhere in $[2, \infty)$.
- $\gamma = 2$ is the lower limit since

$$\mu = \sum_{k=1}^{\infty} A_k n_k \sim \sum_{k=1}^{\infty} k n_k$$

must be finite.
- Let's now look at a specific example of A_k to see this range of γ is possible.



Universality?

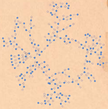
- Consider $A_1 = \alpha$ and $A_k = k$ for $k \geq 2$.
- Find $\gamma = \mu + 1$ by finding μ.
- Expression for μ:

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

$$1 = \frac{1}{1 + \frac{\mu}{A_1}} + \sum_{k=2}^{\infty} \prod_{j=1}^{k} \frac{1}{1 + \frac{\mu}{A_j}}$$

$$1 - \frac{1}{1 + \frac{\mu}{A_1}} = \frac{1}{1 + \frac{\mu}{A_1}} \sum_{k=2}^{\infty} \prod_{j=2}^{k} \frac{1}{1 + \frac{\mu}{A_j}}$$

$$\frac{\frac{\mu}{\alpha}}{1 + \frac{\mu}{\alpha}} = \frac{1}{1 + \frac{\mu}{\alpha}} \sum_{k=2}^{\infty} \prod_{j=2}^{k} \frac{1}{1 + \frac{\mu}{A_j}} \quad \text{since } A_1 = \alpha$$



Universality?

- Carrying on:

$$\frac{\frac{\mu}{\alpha}}{\cancel{1+\frac{\mu}{\alpha}}} = \frac{1}{\cancel{1+\frac{\mu}{\alpha}}} \sum_{k=2}^{\infty} \prod_{j=2}^{k} \frac{1}{1+\frac{\mu}{A_j}}$$

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

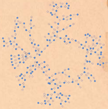
- Now use result that [3]

$$\sum_{k=2}^{\infty} \frac{\Gamma(a+k)}{\Gamma(b+k)} = \frac{\Gamma(a+2)}{(b-a-1)\Gamma(b+1)}$$

with $a = 1$ and $b = \mu + 1$.

- $$\mu = \alpha \frac{\Gamma(3)}{(\mu+1-1-1)\Gamma(2+\mu)} \Gamma(2+\mu)$$

$$\Rightarrow \mu(\mu-1) = 2\alpha$$



Universality?

- $$\mu(\mu - 1) = 2\alpha \Rightarrow \mu = \frac{1 + \sqrt{1 + 8\alpha}}{2}.$$
- 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: [3]

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

- 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:

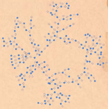
- For $1/2 < \nu < 1$:

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

- 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.

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

[4] A. J. Lotka.
The frequency distribution of scientific productivity.
Journal of the Washington Academy of Science, 16:317–323, 1926.

References II

[5] B. B. Mandelbrot.
An informational theory of the statistical structure of languages.
In W. Jackson, editor, Communication Theory, pages 486–502. Butterworth, Woburn, MA, 1953. pdf (⊞)

[6] D. J. d. S. Price.
Networks of scientific papers.
Science, 149:510–515, 1965. pdf (⊞)

[7] D. J. d. S. Price.
A general theory of bibliometric and other cumulative advantage processes.
J. Amer. Soc. Inform. Sci., 27:292–306, 1976.

[8] H. A. Simon.
On a class of skew distribution functions.
Biometrika, 42:425–440, 1955. pdf (⊞)

References III

[9] G. U. Yule.
A mathematical theory of evolution, based on the conclusions of Dr J. C. Willis, F.R.S.
Phil. Trans. B, 213:21–, 1924.

[10] G. K. Zipf.
Human Behaviour and the Principle of Least-Effort.
Addison-Wesley, Cambridge, MA, 1949.