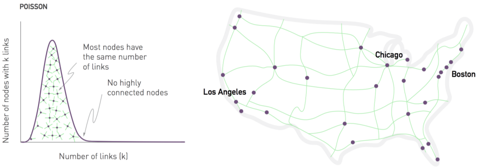
scale-free networks

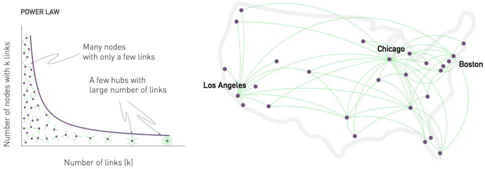
advanced topics in *network science* (*ants*)

Lovro Šubelj & Jure Leskovec
University of Ljubljana
spring 2019/20

scale-free *property*

- *random graphs Poisson degree distribution p_k* [ER59]
- *real networks* contain *highly linked hubs* [Pri65, FFF99]
- *scale-free networks power-law degree distribution p_k* [BA99]



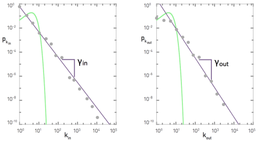


scale-free *power-law*

— *power-law degree distribution* p_k with *exponent* $\gamma > 1$

$$p_k \sim k^{-\gamma}$$

$$\log p_k \sim -\gamma \log k$$



— *theoretically correct discrete power-law* p_k for $k \geq 1$

$$\sum_{k=1}^{\infty} p_k = C \sum_{k=1}^{\infty} k^{-\gamma} = C\zeta(\gamma) = 1$$

$$p_k = Ck^{-\gamma} = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

— *analytically convenient continuos power-law* $p(k)$ for $k \geq k_{min}$

$$\int_{k_{min}}^{\infty} p(k)\,\mathrm{d}k = C \int_{k_{min}}^{\infty} k^{-\gamma}\,\mathrm{d}k = C \left.\frac{k^{-\gamma+1}}{-\gamma+1}\right|_{k_{min}}^{\infty} = C\frac{k_{min}^{-\gamma+1}}{\gamma-1} = 1$$

$$p(k) = Ck^{-\gamma} = (\gamma - 1)k_{min}^{\gamma-1} k^{-\gamma}$$

scale-free *hubs*

- for *small* $k \ll \langle k \rangle$ *power-law above Poisson*

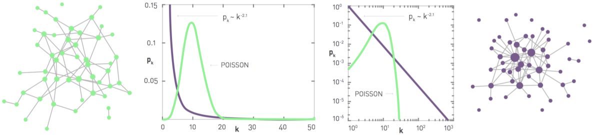
  many *small degree nodes* in *scale-free network*

- for *average* $k \approx \langle k \rangle$ *power-law below Poisson*

  most *nodes similar degree* in *random graph*

- for *large* $k \gg \langle k \rangle$ *power-law above Poisson*

  existence of hubs in *scale-free network*



- *random graph* with $n \approx 10^{12}$ and $\langle k \rangle = 4.6$ then $n_{k \geq 100} \approx 10^{-82}$
- *scale-free network* with $n \approx 10^{12}$ and $\gamma = 2.1$ then $n_{k \geq 100} \approx 4 \cdot 10^9$

scale-free *cutoff*

- *maximum degree* k_{max} by *upper natural cutoff* of $p(k)$
- for *random graph* with *exponential* $p(k) = \lambda e^{\lambda k_{min}} e^{-\lambda k}$

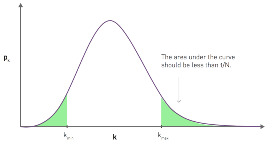
$$\int_{k_{max}}^{\infty} p(k)\,\mathrm{d}k = \lambda e^{\lambda k_{min}} \left.\frac{e^{-\lambda k}}{-\lambda}\right|_{k_{max}}^{\infty} = e^{\lambda k_{min}} e^{-\lambda k_{max}} = n^{-1}$$

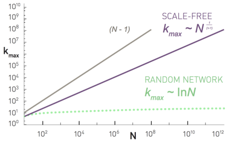
$$k_{max} = k_{min} + \frac{\ln n}{\lambda}$$

- for *scale-free network* with *power-law* $p(k) = (\gamma - 1) k_{min}^{\gamma-1} k^{-\gamma}$

$$\int_{k_{max}}^{\infty} p(k)\,\mathrm{d}k = (\gamma - 1) k_{min}^{\gamma-1} \left.\frac{k^{-\gamma+1}}{-\gamma+1}\right|_{k_{max}}^{\infty} = k_{min}^{\gamma-1} k_{max}^{-\gamma+1} = n^{-1}$$

$$k_{max} = k_{min} n^{\frac{1}{\gamma-1}}$$





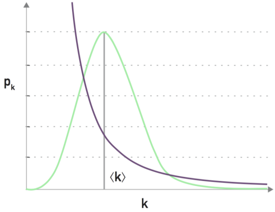
- *random graph* with $n \approx 3 \cdot 10^5$ and $\lambda = 1$ then $k_{max} \approx 14$
- *scale-free network* with $n \approx 3 \cdot 10^5$ and $\gamma = 2.1$ then $k_{max} \approx 10^5$

scale-free *moments*

- *x-th moment* $\langle k^x \rangle$ of *power-law* $p_k \sim k^{-\gamma}$
  - $\langle k^2 \rangle = \sigma_k^2 + \langle k \rangle^2$ determines *spread* and $\langle k^3 \rangle$ is *skewness*

$$\langle k^x \rangle = \sum_{k=1}^{\infty} k^x p_k \approx \int_{k_{min}}^{k_{max}} k^x p(k)\, \mathrm{d}k \sim \frac{k_{max}^{x-\gamma+1} - k_{min}^{x-\gamma+1}}{x-\gamma+1}$$

- *moments* $x \leq \gamma - 1$ *finite* whereas *moments* $x > \gamma - 1$ *diverge*

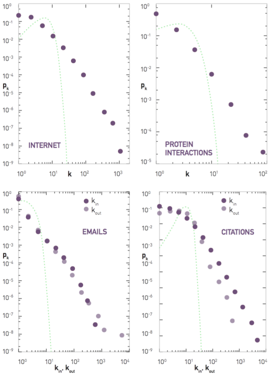


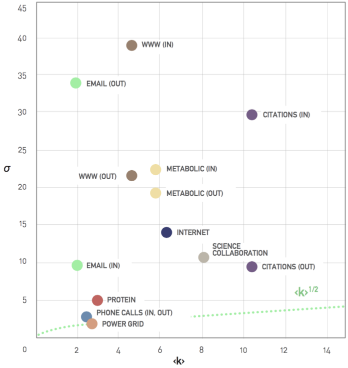
- *scale-free network* $\gamma < 3$ *lacks scale* with $k = \langle k \rangle \pm \infty$
- *random graph has scale* with $k = \langle k \rangle \pm \sqrt{\langle k \rangle}$

scale-free *networks*

— *heavy-tail* p_k of real networks [Bar16]

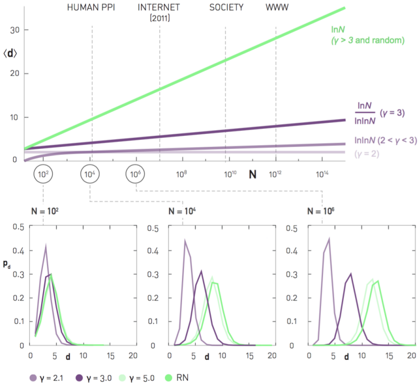
— *spread* $\sigma_k = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}$ in real networks



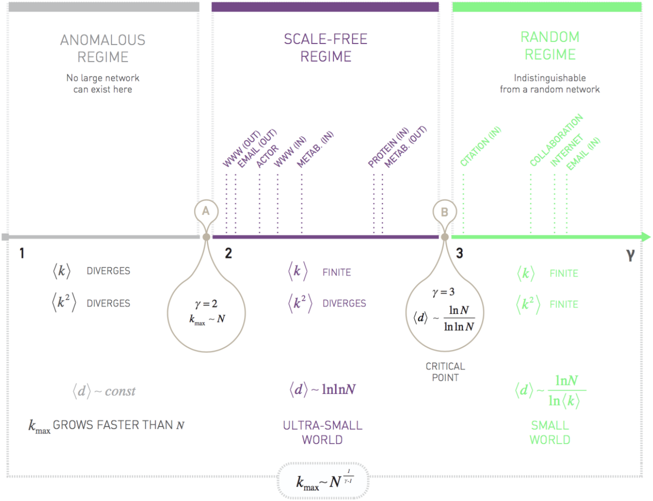


scale-free *"small-world"*

- *random graph* is *"small-world"* with $\langle d \rangle \simeq \frac{\ln n}{\ln \langle k \rangle}$
- *scale-free network* $\gamma > 3$ is *"small-world"* with $\langle d \rangle \sim \ln n$
- *scale-free network* $\gamma < 3$ *"ultrasmall-world"* with $\langle d \rangle \sim \ln \ln n$



scale-free *exponent*

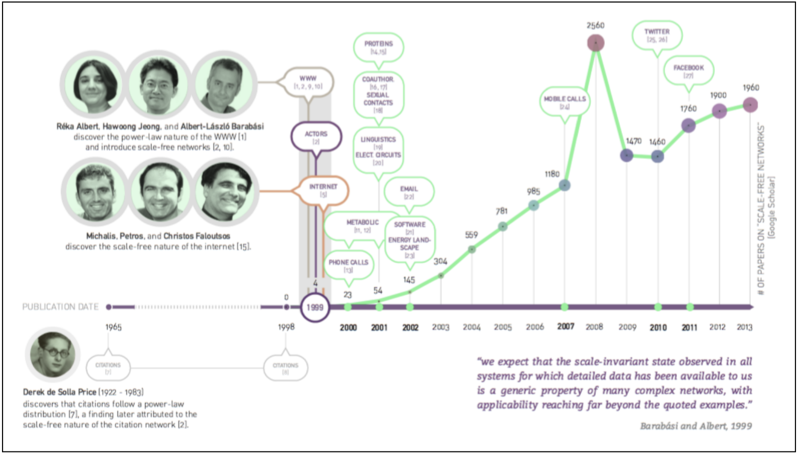


no graphical $\{k\}$ for $\gamma < 2$ $\qquad$ $n = (k_{max}/k_{min})^{\gamma-1}$ *nonexistent* for $\gamma \gg 3$

scale-free *distributions*

NAME	$p_x/p(x)$	$\langle x\rangle$	$\langle x^2\rangle$
Poisson (discrete)	$e^{-\mu}\mu^x/x!$	μ	$\mu(1+\mu)$
Exponential (discrete)	$(1-e^{-\lambda})e^{-\lambda x}$	$1/(e^\lambda-1)$	$(e^\lambda+1)/(e^\lambda-1)^2$
Exponential (continuous)	$\lambda e^{-\lambda x}$	$1/\lambda$	$2/\lambda^2$
Power law (discrete)	$x^{-\alpha}/\zeta(\alpha)$	$\begin{cases}\zeta(\alpha-2)/\zeta(\alpha), & \text{if } \alpha>2\\ \infty, & \text{if } \alpha\leq 1\end{cases}$	$\begin{cases}\zeta(\alpha-1)/\zeta(\alpha), & \text{if } \alpha>1\\ \infty, & \text{if } \alpha\leq 2\end{cases}$
Power law (continuous)	$\alpha x^{-\alpha}$	$\begin{cases}\alpha/(\alpha-1), & \text{if } \alpha>2\\ \infty, & \text{if } \alpha\leq 1\end{cases}$	$\begin{cases}\alpha/(\alpha-2), & \text{if } \alpha>1\\ \infty, & \text{if } \alpha\leq 2\end{cases}$
Power law with cutoff (continuous)	$\frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha)}x^{-\alpha}e^{-\lambda x}$	$\lambda^{-1}\frac{\Gamma(2-\alpha)}{\Gamma(1-\alpha)}$	$\lambda^{-2}\frac{\Gamma(3-\alpha)}{\Gamma(1-\alpha)}$
Stretched exponential (continuous)	$\beta\lambda^\beta x^{\beta-1}e^{-(\lambda x)^\beta}$	$\lambda^{-1}\Gamma(1+\beta^{-1})$	$\lambda^{-2}\Gamma(1+2\beta^{-1})$
Log-normal (continuous)	$\frac{1}{x\sqrt{2\pi\sigma^2}}e^{-(\ln x-\mu)^2/(2\sigma^2)}$	$e^{\mu+\sigma^2/2}$	$e^{2(\mu+\sigma^2)}$
Normal (continuous)	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/(2\sigma^2)}$	μ	$\mu^2+\sigma^2$

scale-free *history*



scale-free models

advanced topics in *network science* (*ants*)

Lovro Šubelj & Jure Leskovec
University of Ljubljana
spring 2019/20

scale-free *models*

- *graph models* are *ensembles* of random graphs
- *generative models* reason about *network evolution*
- *cumulative advantage* process of *Price model* [Pri76]
- *preferential attachment* or *Barabási-Albert model* [BA99]

Pólya process Yule process Zipf's law Matthew effect
rich-get-richer proportional growth cumulative advantage

see preferential attachment model NetLogo demo

Derek de Solla Price

Albert-László Barabási

Réka Albert

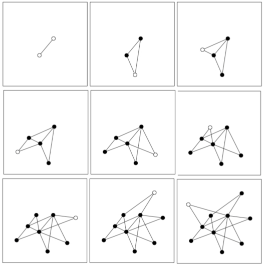
scale-free $G(n, c, a)$ *model*

- $G(n, c, a)$ *cumulative advantage* model [Pri76]
- each new node i *forms* $k_i^{out} = c > 0$ *directed links*
- node j *receives link with probability* $\sim k_j^{in} + a = q_j + a > 0$

n, c, a given $\qquad$ p_q *unknown*

input parameters n, c, a
output *directed* graph G

$G \leftarrow\ \geq c$ isolated nodes
while not G has n nodes **do**
  add node i to G
  for c **times do**
    add link (i, j) with $\sim q_j + a$
  end for
end while
return G



scale-free $G(n, c, a)$ *equation*

- *master equation* for *in-degree distribution* $p_q(n)$
  - $p_q(n)$ is in-degree distribution p_q *at time* n

$$\frac{q_i+a}{\sum_i q_i+a} = \frac{q_i+a}{n(c+a)} \qquad cnp_q(n)\frac{q+a}{n(c+a)} = \frac{c(q+a)}{c+a}p_q(n)$$

$$(n+1)p_q(n+1) = np_q(n) + \frac{c(q-1+a)}{c+a}p_{q-1}(n) - \frac{c(q+a)}{c+a}p_q(n)$$

$$(n+1)p_0(n+1) = np_0(n) + 1 - \frac{ca}{c+a}p_0(n)$$

- *power-law in-degree distribution* $p_q \sim q^{-\gamma}$ with $\gamma > 2$
  - p_q is in-degree distribution *in limit* $n \to \infty$

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t \qquad B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \simeq x^{-y}\Gamma(y)$$

$$p_q = \frac{q+a-1}{q+a+1+a/c}p_{q-1} = \cdots = \frac{B(q+a,2+a/c)}{B(a,1+a/c)} \sim q^{-2-a/c}$$

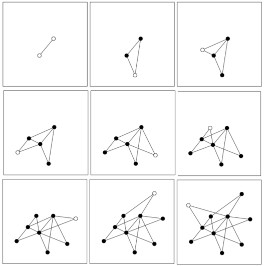
$$p_0 = \frac{1+a/c}{a+1+a/c}$$

scale-free $G(n, c)$ *model*

- $G(n, c)$ *preferential attachment* model [BA99]
- each new node i *forms* $c > 0$ *undirected links*
- node j *receives links with probability* $\sim k_i$

n, c given $\qquad$ p_k *unknown*

```
input  parameters n, c
output undirected graph G
  G ← c connected nodes
  while not G has n nodes do
      add node i to G
      for c times do
          add link {i, j} with ~ k_j
      end for
  end while
  return G
```



scale-free $G(n, c)$ *equation*

- *undirected* $G(n, c)$ is *directed* $G(n, c, c)$ for $k_i = q_i + c$
- *same master equation* for *in-degree distribution* p_q
  - p_q is in-degree distribution *in limit* $n \to \infty$

$$p_q = \frac{B(q+c,2+c/c)}{B(c,1+c/c)} = \frac{B(q+c,3)}{B(c,2)} \sim q^{-3}$$

- *power-law degree distribution* $p_k \sim k^{-3}$
  - p_k is degree distribution *in limit* $n \to \infty$

$$p_k = \frac{B(k,3)}{B(c,2)} = \cdots = \frac{2c(c+1)}{k(k+1)(k+2)} \sim k^{-3}$$

 © Šubelj & Leskovec

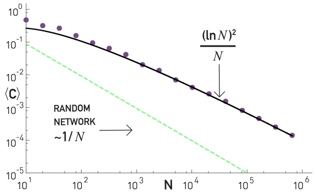
preferential ¬*small-world*

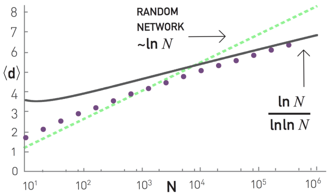
— *random graphs* are *"small-world"* as $\langle d \rangle \simeq \frac{\ln n}{\ln \langle k \rangle}$

— *random graphs* are *not small-world* as $\langle C \rangle = \frac{\langle k \rangle}{n-1}$

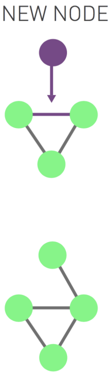
— *scale-free networks* $\gamma = 3$ are *"small-world"* as $\langle d \rangle \sim \frac{\ln n}{\ln \ln n}$

— $G(n, c)$ *scale-free model* is *not small-world* as $\langle C \rangle \simeq \frac{(\ln n)^2}{n}$

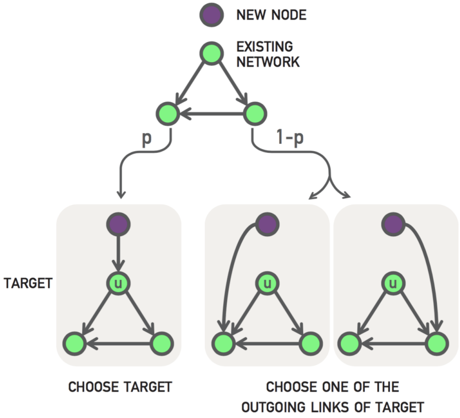




preferential *models*

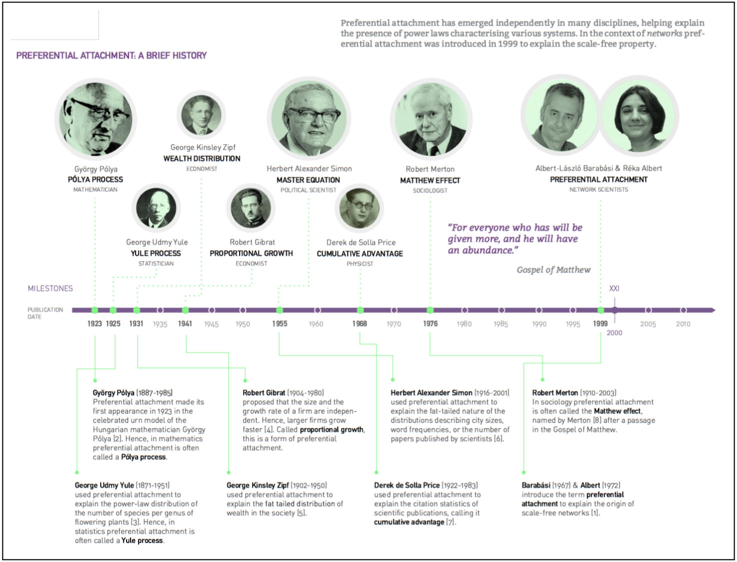


link selection [DM02]



random *link copying* model [KKR+99]

scale-free *history*



scale-free *references*

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

A.-L. Barabási.
Network Science.
Cambridge University Press, Cambridge, 2016.

S. N. Dorogovtsev and J. F. F. Mendes.
Evolution of networks.
Adv. Phys., 51(4):1079–1187, 2002.

David Easley and Jon Kleinberg.
Networks, Crowds, and Markets: Reasoning About a Highly Connected World.
Cambridge University Press, Cambridge, 2010.

P. Erdős and A. Rényi.
On random graphs I.
Publ. Math. Debrecen, 6:290–297, 1959.

Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos.
On power-law relationships of the Internet topology.
Comput. Commun. Rev., 29(4):251–262, 1999.

Jon M. Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S. Tomkins.
The web as a graph: Measurements, models, and methods.
In *Proceedings of the International Conference on Computing and Combinatorics*, pages 1–17, Tokyo, Japan, 1999.

scale-free *references*

Mark E. J. Newman.
Networks: An Introduction.
Oxford University Press, Oxford, 2010.

D. J. de Solla Price.
Networks of scientific papers.
Science, 149:510–515, 1965.

Derek De Solla Price.
A general theory of bibliometric and other cumulative advantage processes.
J. Am. Soc. Inf. Sci., 27(5):292–306, 1976.