NETWORK VISUALISATION
(SHORT DIGRESSION)
NETWORK VISUALIZATION

- How to interpret a network drawing?
- What does the position of nodes mean?
- Can we draw conclusions from the drawing alone?
NETWORK VISUALIZATION
NETWORK VISUALIZATION

• Random layout
  ‣ Assign random positions to nodes, draw edges
    - Useless for more than 5-6 nodes

• Geographical layout
  ‣ The position of nodes is fixed apriori, often based on geographical location
  ‣ Variant: position nodes on a circle based on a single, 1D property (age…)

![Network Visualization Diagrams]
NETWORK VISUALIZATION

• Random layout
  ‣ Assign random positions to nodes, draw edges
    - Useless for more than 5-6 nodes

• Geographical layout
  ‣ The position of nodes is fixed apriori, often based on geographical location
  ‣ Variant: position nodes on a circle based on a single, 1D property (age…)

[Diagram images of network visualizations]
NETWORK VISUALIZATION

- Most commonly used: Automatic layout
  - Non deterministic
  - Tries to arrange nodes so that the network is easy to read and understand
    - Minimize edge crossings?
    - Most commonly, tries to put connected nodes close and unconnected nodes far
NETWORK VISUALIZATION

http://kwonoh.net/dgl/
NETWORK VISUALIZATION

• Most common algorithms are variant of the force directed layout
  ‣ Kamada-Kawai
  ‣ Fruchterman-Reingold
  ‣ ...

• Force directed layout: a simple physical model
  ‣ Repulsive forces between nodes
  ‣ Edges are attracting forces
  ‣ There are minimal (to avoid node overlap) and maximal (to avoid connected component drifting out of the figure)
NETWORK VISUALIZATION

• Can we interpret a force layout?
  ‣ Yes…
NETWORK VISUALIZATION

• Can we interpret a force layout?
  ‣ Yes…
  ‣ And no.
Can we interpret a force layout?

- Yes…
- And no.
Scale-free networks
A network is called *Scale-free* when its degree distribution follows (to some extent) a Power-law distribution.

**Power-law distribution:** (PDF)

\[
P(k) \sim Ck^{-\alpha} = C \frac{1}{k^\alpha}
\]

\(\alpha\) (sometimes \(\gamma\)) called the exponent of the distribution.

Positive values

Here, defined as continuous (approximation)
Proper definition

Initial definition: \( P(k) \sim C k^{-\alpha} = \frac{1}{k^\alpha} \)

To have a proper degree distribution, we need
\[
\int P(k) = 1 = \int C k^{-\alpha} = C \int k^{-\alpha}.
\]

We also know that in most cases, there is a lower bound from which the law holds. \( k_{\text{min}} \)

From this, we define the normalisation constant:
\[
C = \frac{1}{\int_{k_{\text{min}}}^{\infty} k^{-\alpha} dk} = (\alpha - 1) k_{\text{min}}^{\alpha-1}
\]
Scale-free distribution

Proper definition

\[ P(k) \sim Ck^{-\alpha} \]

\[ C = \frac{1}{\int_{k_{\text{min}}}^{\infty} k^{-\alpha} dk} = (\alpha - 1)k_{\text{min}}^{\alpha - 1} \]

\[ P(k) = (\alpha - 1)k_{\text{min}}^{\alpha - 1}k^{-\alpha} \]

\[ P(k) = \frac{\alpha - 1}{k_{\text{min}}} \left( \frac{k}{k_{\text{min}}} \right)^{-\alpha} \]
Scale-free networks

Power law plotted with a linear scale, for $k \leq 10$

(100 000 samples)
Scale-free networks

Power law plotted with a linear scale, for $k<100000$
(100 000 samples)
Scale-free networks

**Power law** plotted with a log-log scale, for $k<100000$

(100 000 samples)
Scale-free networks

Comparing a poisson distribution and a power law

\[ \frac{\lambda^k e^{-\lambda}}{k!} \]
Comparing a poisson distribution and a power law

\[ \frac{\lambda^k e^{-\lambda}}{k!} \]
Comparing a poisson distribution and a power law

\[ \frac{\lambda^k e^{-\lambda}}{k!} \]

The “long tail”
Scale-free networks

Comparing an exponential distribution and a power law

\[
\begin{cases}
\lambda e^{-\lambda k} & k \geq 0, \\
0 & k < 0.
\end{cases}
\]
Scale-free networks - first observations


Diameter of the world wide web

Emergence of Scaling in Random Networks
Albert-László Barabási* and Réka Albert
Scale-free networks - other examples

The internet

- Nodes: routers
- Links: Physical wires

Faloutsos, Faloutsos and Faloutsos (1999)
Scale-free networks - other examples

Airline route map network

- Nodes: airports
- Links: airplane connections

Guimera et al. (2004)

Note: the cumulative distribution of a power law is also a line on a log-log plot.
Scale-free networks - other examples

Scientific collaborations
- Nodes: scientists (here geo-localised)
- Links: common papers

![Diagram of scientific collaborations](image)

Map of scientific collaborations from 2005 to 2009
Computed by Olivier H. Beauchesne @ Science-Metrix, Inc.
Data from Scopus, using books, trade journals and peer-reviewed journals

Newman (2001)
Sexual-interaction networks

- Nodes: individuals
- Links: sexual incursion

Bearman et. al. (2004)

The Structure of Romantic and Sexual Relations at "Jefferson High School"
Scale-free networks - other examples

Online social networks

- Nodes: individuals
- Links: online interactions

Social network of Steam
http://85.25.226.110/mapper
Degree fluctuations have no characteristic scale (scale invariant)
Scale-free networks

Idea of scale free

In contrast to Fig. 4.1, numerous small-degree nodes coexist with a few hubs, nodes with an exceptionally large number of links. The purpose of this chapter is to show that these hubs are not unique to the Web, but we encounter them in many real networks. They represent a signature of a deeper organizing principle that we call the scale-free property.
Interesting properties of power law distributions

\[ P(k) \sim C k^{-\alpha} \]

\[ C = \frac{1}{\int_{k_{\text{min}}}^{\infty} k^{-\alpha} \, dk} = (\alpha - 1)k_{\text{min}}^{\alpha - 1} \]

\[ P(k) = (\alpha - 1)k_{\text{min}}^{\alpha - 1}k^{-\alpha} \]

\[ P(k) = \frac{\alpha - 1}{k_{\text{min}}} \left( \frac{k}{k_{\text{min}}} \right)^{-\alpha} \quad k \geq k_{\text{min}} \]
Scale-free distribution

Moments

Distribution:

\[ P(k) = (\alpha - 1)k_{\text{min}}^{\alpha - 1}k^{-\alpha} \]

(central) Moments:

\[ \langle k^m \rangle = \int_{k_{\text{min}}}^{\infty} k^m p(k)dk \]

Reminder:

\[ \langle k^1 \rangle \quad \text{Average} \]
\[ \langle k^2 \rangle \quad \text{Variance} \]
\[ \langle k^3 \rangle \quad \text{Skewness} \]
\[ \ldots \]
Scale-free distribution

Moments

Distribution:

\[ P(k) = (\alpha - 1)k_{\min}^{\alpha-1}k^{-\alpha} \]

(central) Moments:

\[ \langle k^m \rangle = \int_{k_{\min}}^{\infty} k^m p(k)dk \]

\[ \langle k^m \rangle = (\alpha - 1)k_{\min}^{\alpha-1} \int_{k_{\min}}^{\infty} k^{-\alpha+m}dk \]

\[ \langle k^m \rangle = k_{\min}^m \left( \frac{\alpha - 1}{\alpha - 1 - m} \right) \]

Defined for \( \alpha > m + 1 \),
Otherwise diverge (+inf)

\[ \int x^n dx = \frac{1}{n+1}x^{n+1} + c, \; n \neq -1 \]

Scale-free distribution

Moments

Distribution:

\[ P(k) = \frac{\alpha - 1}{k_{\text{min}}} \left( \frac{k}{k_{\text{min}}} \right)^{-\alpha} \]

(cenral) Moments:

\[ \langle k^m \rangle = k_{\text{min}}^m \left( \frac{\alpha - 1}{\alpha - 1 - m} \right) \]

Defined for \( \alpha > m + 1 \), Otherwise diverge (+inf)

\[ \Rightarrow \text{Mean: } \langle k \rangle = \frac{\alpha - 1}{\alpha - 2} k_{\text{min}} \]

(But diverges for \( \alpha \leq 2 \))

\[ \Rightarrow \text{Variance: } \langle k^2 \rangle = \frac{\alpha - 1}{\alpha - 3} k_{\text{min}}^2 \]

(But diverges for \( \alpha \leq 3 \))
Moments

What does divergence means in practice?

We can always compute the mean and variance, given samples of a distribution (e.g., an observe degree distribution)

=>The value computed depends on the size of the sample, it is not a characteristic of the distribution.

Moments are dominated by elements in the long tail. Some events are rare, but they have so large values, that if observed, they are strong enough to modify substantially the corresponding moment. And they appear frequently enough so that the mean will continue to shift when increasing the sample size.
Scale-free distribution

Moments

\( \alpha < 2 \)
Mean diverge

\( 2 < \alpha < 3 \)
Mean well defined,
Variance diverge

\( \alpha > 3 \)
Mean and variance
defined

\[ p(cx) \propto p(x) \]

Further, it can be shown that a power law form is the only function that has this property.

Here's another way of seeing this behavior. If we take the logarithm of both sides of Eq. (1), we get an expression for \( \ln p(x) \) that's linear in \( \ln x \). That is,

\[ \ln p(x) = \ln C - \alpha \ln x. \]

This shows another of the more well-known properties of a power-law distribution: it's a straight line on a log-log plot. This is in contrast to the strongly curved behavior of, say, an exponential distribution, as in Fig. 1.

An exercise left to the reader.

\[ \Rightarrow \text{Even when well defined, moments converge very slowly} \]
Scale-free networks

Computing the exponent of an observed network

**Method 1**: find the slope of the line of the log-log plot

Problem: most of data is on first value, so we overfit based on a few values in the long tail

**More advanced method**: Maximum Likelihood Estimation (MLE)

[Fitting to the Power-Law Distribution, Goldstein et al.]
### Scale-free networks

**Exponents of real-world networks are usually between 2 and 3**

<table>
<thead>
<tr>
<th>Network</th>
<th>Size</th>
<th>\langle k \rangle</th>
<th>\kappa</th>
<th>\gamma_{out}</th>
<th>\gamma_{in}</th>
<th>\ell_{real}</th>
<th>\ell_{rand}</th>
<th>\ell_{pow}</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWW</td>
<td>325 729</td>
<td>4.51</td>
<td>900</td>
<td>2.45</td>
<td>2.1</td>
<td>11.2</td>
<td>8.32</td>
<td>4.77</td>
<td>Albert, Jeong, and Barabási 1999</td>
</tr>
<tr>
<td>WWW</td>
<td>4 \times 10^{7}</td>
<td>7</td>
<td>2.38</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Kumar et al., 1999</td>
</tr>
<tr>
<td>WWW</td>
<td>2 \times 10^{8}</td>
<td>7.5</td>
<td>4000</td>
<td>2.72</td>
<td>2.1</td>
<td>16</td>
<td>8.85</td>
<td>7.61</td>
<td>Broder et al., 2000</td>
</tr>
<tr>
<td>WWW, site</td>
<td>260 000</td>
<td>1.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Huberman and Adamic, 2000</td>
</tr>
<tr>
<td>Internet, domain*</td>
<td>3015–4389</td>
<td>3.42–3.76</td>
<td>30–40</td>
<td>2.1–2.2</td>
<td>2.1–2.2</td>
<td>4</td>
<td>6.3</td>
<td>5.2</td>
<td>Faloutsos, 1999</td>
</tr>
<tr>
<td>Internet, router*</td>
<td>3888</td>
<td>2.57</td>
<td>30</td>
<td>2.48</td>
<td>2.48</td>
<td>12.15</td>
<td>8.75</td>
<td>7.67</td>
<td>Faloutsos, 1999</td>
</tr>
<tr>
<td>Internet, router*</td>
<td>150 000</td>
<td>2.66</td>
<td>60</td>
<td>2.4</td>
<td>2.4</td>
<td>11</td>
<td>12.8</td>
<td>7.47</td>
<td>Govindan, 2000</td>
</tr>
<tr>
<td>Movie actors*</td>
<td>212 250</td>
<td>28.78</td>
<td>900</td>
<td>2.3</td>
<td>2.3</td>
<td>4.54</td>
<td>3.65</td>
<td>4.01</td>
<td>Barabási and Albert, 1999</td>
</tr>
<tr>
<td>Co-authors, SPIRES*</td>
<td>56 627</td>
<td>173</td>
<td>1100</td>
<td>1.2</td>
<td>1.2</td>
<td>4</td>
<td>2.12</td>
<td>1.95</td>
<td>Newman, 2001b</td>
</tr>
<tr>
<td>Co-authors, neuro.*</td>
<td>209 293</td>
<td>11.54</td>
<td>400</td>
<td>2.1</td>
<td>2.1</td>
<td>6</td>
<td>5.01</td>
<td>3.86</td>
<td>Barabási et al., 2001</td>
</tr>
<tr>
<td>Co-authors, math.*</td>
<td>70 975</td>
<td>3.9</td>
<td>120</td>
<td>2.5</td>
<td>2.5</td>
<td>9.5</td>
<td>8.2</td>
<td>6.53</td>
<td>Barabási et al., 2001</td>
</tr>
<tr>
<td>Sexual contacts*</td>
<td>2810</td>
<td>3.4</td>
<td>120</td>
<td>3.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Liljeros et al., 2001</td>
</tr>
<tr>
<td>Metabolic, E. coli</td>
<td>778</td>
<td>7.4</td>
<td>110</td>
<td>2.2</td>
<td>2.2</td>
<td>3.2</td>
<td>3.32</td>
<td>2.89</td>
<td>Jeong et al., 2000</td>
</tr>
<tr>
<td>Protein, S. cerev.*</td>
<td>1870</td>
<td>2.39</td>
<td>35</td>
<td>2.4</td>
<td>2.4</td>
<td>2.43</td>
<td>2.26</td>
<td>1.71</td>
<td>Jeong, Mason, et al., 2001</td>
</tr>
<tr>
<td>Ythan estuary*</td>
<td>134</td>
<td>8.7</td>
<td>35</td>
<td>1.05</td>
<td>1.05</td>
<td>2.43</td>
<td>2.26</td>
<td>1.71</td>
<td>Montoya and Solé, 2000</td>
</tr>
<tr>
<td>Silwood Park*</td>
<td>154</td>
<td>4.75</td>
<td>27</td>
<td>1.13</td>
<td>1.13</td>
<td>3.4</td>
<td>3.23</td>
<td>2</td>
<td>Montoya and Solé, 2000</td>
</tr>
<tr>
<td>Citation</td>
<td>783 339</td>
<td>8.57</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Redner, 1998</td>
</tr>
<tr>
<td>Phone call</td>
<td>53 \times 10^{6}</td>
<td>3.16</td>
<td></td>
<td>2.1</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td>Aiello et al., 2000</td>
</tr>
<tr>
<td>Words, co-occurrence*</td>
<td>460 902</td>
<td>70.13</td>
<td></td>
<td>2.7</td>
<td>2.7</td>
<td></td>
<td></td>
<td></td>
<td>Ferrer i Cancho and Solé, 2001</td>
</tr>
<tr>
<td>Words, synonyms*</td>
<td>22 311</td>
<td>13.48</td>
<td></td>
<td>2.8</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
<td>Yook et al., 2001b</td>
</tr>
</tbody>
</table>
Average values are not reliable since the convergence is very slow.

Furthermore, average values are meaningless since the fluctuations are infinitely large (diverging variance).

The topology of the Internet is studied at two different levels. At the router level, the nodes are the routers, and edges are the physical connections between them. At the interdomain (or autonomous system) level, each node is a router or a computer, and the edges are the wires and cables that physically connect them.

### TABLE II: The scaling exponents characterizing the degree distribution of several scale-free networks

<table>
<thead>
<tr>
<th>Network</th>
<th>Size</th>
<th>(\langle k \rangle)</th>
<th>(\kappa)</th>
<th>(\gamma_{out})</th>
<th>(\gamma_{in})</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWW</td>
<td>325 729</td>
<td>4.51</td>
<td>900</td>
<td>2.45</td>
<td>2.1</td>
</tr>
<tr>
<td>WWW</td>
<td>(4 \times 10^7)</td>
<td>7</td>
<td></td>
<td>2.38</td>
<td>2.1</td>
</tr>
<tr>
<td>WWW</td>
<td>(2 \times 10^8)</td>
<td>7.5</td>
<td>4000</td>
<td>2.72</td>
<td>2.1</td>
</tr>
<tr>
<td>WWW, site</td>
<td>260 000</td>
<td></td>
<td></td>
<td></td>
<td>1.94</td>
</tr>
<tr>
<td>Internet, domain*</td>
<td>3015–4389</td>
<td>3.42–3.76</td>
<td>30–40</td>
<td>2.1–2.2</td>
<td>2.1–2.2</td>
</tr>
<tr>
<td>Internet, router*</td>
<td>3888</td>
<td>2.57</td>
<td>30</td>
<td>2.48</td>
<td>2.48</td>
</tr>
<tr>
<td>Internet, router*</td>
<td>150 000</td>
<td>2.66</td>
<td>60</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Movie actors*</td>
<td>212 250</td>
<td>28.78</td>
<td>900</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>Co-authors, SPIRES*</td>
<td>56 627</td>
<td>173</td>
<td>1100</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Co-authors, neuro.*</td>
<td>209 293</td>
<td>11.54</td>
<td>400</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>Co-authors, math.*</td>
<td>70 975</td>
<td>3.9</td>
<td>120</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Sexual contacts*</td>
<td>2810</td>
<td></td>
<td></td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>Metabolic, <em>E. coli</em></td>
<td>778</td>
<td>7.4</td>
<td>110</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Protein, S. cerev.*</td>
<td>1870</td>
<td>2.39</td>
<td></td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Ythan estuary*</td>
<td>134</td>
<td>8.7</td>
<td>35</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td>Silwood Park*</td>
<td>154</td>
<td>4.75</td>
<td>27</td>
<td>1.13</td>
<td>1.13</td>
</tr>
<tr>
<td>Citation</td>
<td>783 339</td>
<td>8.57</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Phone call</td>
<td>(53 \times 10^6)</td>
<td>3.16</td>
<td></td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>Words, co-occurrence*</td>
<td>460 902</td>
<td>70.13</td>
<td></td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Words, synonyms*</td>
<td>22 311</td>
<td>13.48</td>
<td></td>
<td>2.8</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Scale-free networks

Why do most of the real networks have degree exponent between 2 and 3?

- If the exponent is smaller than 2, the distribution is so skewed that we expect to find nodes with a degree \textit{larger} than the size of the network \Rightarrow not possible in finite networks
Scale-free networks

Why do most of the real networks have degree exponent between 2 and 3?

- To detect a scale-free network its degree distribution needs to span through several (at least 2-3) orders of magnitude $\Rightarrow K_{\text{max}} \sim 10^3$

- If the exponent is large (>3), large degrees become so rare that the size of the sample (i.e., size of observed network) must be enormous to indeed observe such an edge

- Example: let’s choose $\gamma = 5$, $K_{\text{min}} = 1$ and $K_{\text{max}} \sim 10^3$

$$K_{\text{max}} = K_{\text{min}} N^\frac{1}{\gamma - 1}$$

$$N = \left( \frac{K_{\text{max}}}{K_{\text{min}}} \right)^{\gamma - 1} \approx 10^{12}$$

We need to observe $10^{12}$ nodes to observe a node of degree 1000 for exponent=5

$\Rightarrow$ Forget about (single planet) social networks…
Scale-free networks - distances

Size of the biggest hub is of order $O(N)$. Most nodes can be connected within two layers of it, thus the average path length will be independent of the system size.

The average path length increases slower than logarithmically. In a random network all nodes have comparable degree, thus most paths will have comparable length. In a scale-free network the vast majority of the path go through the few high degree hubs, reducing the distances between nodes.

Some key models produce $\gamma = 3$, so the result is of particular importance for them. This was first derived by Bollobas and collaborators for the network diameter in the context of a dynamical model, but it holds for the average path length as well.

The second moment of the distribution is finite, thus in many ways the network behaves as a random network. Hence the average path length follows the result that we derived for the random network model earlier.

Scale-free networks - summary

ANOMALOUS REGIME
No large network can exist here

SCALE-FREE REGIME
\[ \langle k \rangle \text{ DIVERGES} \]
\[ \langle k^2 \rangle \text{ DIVERGES} \]
\[ \gamma = 2 \]
\[ k_{\text{max}} \sim N \]
\[ \langle d \rangle \sim \text{const} \]
\[ k_{\text{max}} \text{ grows faster than } N \]

Random REGIME
Indistinguishable from a random network

\[ \langle k \rangle \text{ FINITE} \]
\[ \langle k^2 \rangle \text{ DIVERGES} \]
\[ \gamma = 3 \]
\[ \langle k \rangle \sim \frac{\ln N}{\ln \ln N} \]
\[ \langle d \rangle \sim \frac{\ln N}{\ln \langle k \rangle} \]

ULTRA-SMALL WORLD

\[ k_{\text{max}} \sim N^\frac{1}{\gamma} \]
Scale-free networks

- Are real networks really Scale Free?
- In most real networks, the scale free stands only for a range of degrees, i.e., between a minimum degree and maximum degree different than those observed (cut-offs)
- Some other distributions, in particular log-normal distributions, might “look like” power-law

Emergence of scaling in random networks (1999)

Scale-free networks are rare (2018)

Love is All You Need - Clauset’s fruitless search for scale-free networks (2018)

Rare and everywhere: Perspectives on scale-free networks (2019)
Scale-free networks

Comparing a log-normal distribution and a power law

\[
\frac{1}{k \sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln k - \mu)^2}{2\sigma^2}\right) \quad k^{-\alpha}
\]

Log-normal distribution = Probability distribution of a random variable whose logarithm is normally distributed

\(\mu, \sigma\) Mean, std of the log of the variable
Scale-free networks

@aaronaaronclauset Every 5 years someone is shocked to rediscover that a pure power law does not fit many networks. True: Real networks have predictable deviations. Hence forcing a pure power law on these is like...fitting a sphere to the cow. Sooner or later the hoof will stick out.

4.13.4 Systematic Fitting Issues

The procedure described above may offer the impression that determining the degree exponent is a cumbersome but straightforward process. In reality, these fitting methods have some well-known limitations: Network Science, Chapter 4, pg 159

1. A pure power law is an idealized distribution that emerges in its form (4.1) only in simple models (Chapter 5). In reality, a whole range of processes contribute to the topology of real networks, affecting the precise shape of the degree distribution. These processes will be discussed in Chapter 6. If $p_k$ does not follow a pure power law, the methods described above, designed to fit a power law to the data, will inevitably fail to detect statistical significance. While this finding can mean that the network is not scale-free, it most often means that we have not yet gained a proper understanding of the precise form of the degree distribution. Hence we are fitting the wrong functional form of $p_k$ to the dataset.
Rigorous statistical tests show that observed degree distributions are not compatible with a power law distribution (high p-values).

Compared with different distributions, in particular log-normal, most degree distributions are more likely to be generated by something else than power laws.

Networks are real objects, not mathematical abstraction, therefore they are sensible to noise (real life limits...).

Power law is a good, simple model of degree distributions of a class of networks.

20 years of fruitful research based on this model.
Scale-free networks
Robustness against failures and attacks
Scale-free networks - role of hubs

Network robustness and attack tolerance

- How network topology is resistant against failure and targeted attacks

(a) Poisson random graph
(b) Scale-free network
Both networks have the same parameters:
- \( N=130 \)
- \( \langle k \rangle=3.3 \)

Numerical experiment:
1. Take a connected network
2. Remove nodes one at a time
3. Observe the size of the LCC

Node removal strategies:
Remove nodes randomly ("failures")

Example: Random failure of routers on the internet
Removal of nodes from networks

Inverse percolation problem

\[ f = \text{fraction of removed nodes} \]

(Graph) Component structure

(Inverse Percolation phase transition)
Malloy-Reed criteria for giant components

A giant cluster exists if each node is connected to at least two other nodes.

The average degree of a node i linked to the GC, must be 2.

Can be shown to correspond to the following relation:

\[ \kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2 \]

\( \kappa > 2 \): Giant component exist
\( \kappa < 2 \): Many disconnected cluster

Random node removal changes

- The degree of individual nodes \([k \rightarrow k' \leq k]\) decrease by losing links via node removal
- A node with degree \(k\) becomes a node with degree \(k'\) with probability:
  \[
  \binom{k}{k'} f^{k-k'} (1-f)^{k'}
  \]
  where \(k' \leq k\)

- the degree distribution \([P(k) \rightarrow P'(k')]\) after random removal of \(f\) fraction of nodes

  \[
  P'(k') = \sum_{k=k'}^{\infty} P(k) \binom{k}{k'} f^{k-k'} (1-f)^{k'}
  \]

Breakdown threshold for arbitrary $P(k)$

$$P'(k') = \sum_{k=k'}^{\infty} P(k) \binom{k}{k'} f^k (1-f)^{k'}$$

$$\langle k' \rangle_f = (1-f)\langle k \rangle$$
$$\langle k'^2 \rangle_f = (1-f)^2 \langle k^2 \rangle + f(1-f)\langle k \rangle$$

We know:

$$\langle k' \rangle_f = (1-f)\langle k \rangle$$
$$\langle k'^2 \rangle_f = (1-f)^2 \langle k^2 \rangle + f(1-f)\langle k \rangle$$
$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

Breakdown threshold:

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle}} - 1$$

$$\kappa \equiv 1 - CN \frac{3-\gamma}{\gamma-1}$$

$k > 2$: Giant component exist

$k < 2$: Many disconnected clusters
Robustness of scale-free networks

Scale-free random graph with

\[ P(k) = A k^{-\gamma} \quad \text{with} \quad k = m, \ldots, K \]

\[ \kappa = 1 - CN^{\frac{3-\gamma}{\gamma-1}} \]

Internet:

Router level map, \( N = 228,263; \gamma = 2.1 \pm 0.1; \kappa = 28 \rightarrow f_c = 0.962 \)

AS level map, \( N = 11,164; \gamma = 2.1 \pm 0.1; \kappa = 264 \rightarrow f_c = 0.996 \)

**Robustness of scale-free networks**

**Scale-free networks do not appear to break apart under random failures.**
Reason: the likelihood of removing a hub is small.  

The robustness of scale-free networks is due to the hubs, which are difficult to hit by chance.

Node removal strategies:
Remove nodes in descending order of their degrees, i.e. hubs first (“attacks”)

Examples: Terrorist attacks, efficient vaccination in epidemics

\[ \gamma \leq 3 : f_c = 1 \]

(R. Cohen et al PRL, 2000)
**Attack threshold for scale-free networks**

**Attack problem:** we remove a fraction $f$ of the hubs.

At what threshold $f_c$ will the network fall apart (no giant component)?

**Critical threshold for scale-free networks:**

\[
\frac{2-\gamma}{f_c^{\frac{1}{1-\gamma}}} = 2 + \frac{2-\gamma}{3-\gamma} K_{\text{min}} \left( f_c^{\frac{3-\gamma}{1-\gamma}} - 1 \right)
\]

- $f_c$ depends on $\gamma$; it reaches its max for $\gamma < 3$
- $f_c$ depends on $K_{\text{min}}$ ($m$ in the figure)
- **Most important:** $f_c$ is tiny. Its maximum reaches only 6%, i.e. the removal of 6% of nodes can destroy the network in an attack mode.
- **Internet:** $\gamma = 2.1$, so 4.7% is the threshold.

*Pastor-Satorras, Vespignani, Evolution and Structure of the Internet (Cambridge University Press, 2004)*
Scale-free networks - role of hubs

Network robustness and attack tolerance

**Poisson random graph**
- Both removal methods give the same result
- The network falls apart after a finite fraction of nodes are removed ($S \to 0$)

**Scale-free network**
- Robust against random removal (blue)
- Vulnerable against targeted attacks

Consequences:
- Internet is still working even several servers are out of service
- Random vaccination is not effective in case of epidemic spreading

- $S$: relative size of the LCC
- $\langle s \rangle$: average size of components other than LCC
The Barabási-Albert model of scale-free networks
Emergence of hubs

What did we miss with the earlier network models?

1. Networks are evolving
   • Networks are not static but growing in time as new nodes are entering the system

2. Preferential attachment
   • Nodes are not connected randomly but tends to link to more attractive nodes
     • Pólya urn model (1923)
     • Yule process (1925)
     • Zipf’s law (1941)
     • Cumulative advantage (1968)
     • Preferential attachment (1999)
     • Pareto’s law - 80/20 rule
     • The rich get richer phenomena
     • etc.
The Barabási-Albert model

1. Start with $m_0$ connected nodes

2. At each timestep we add a new node with $m$ ($\leq m_0$) links that connect the new node to $m$ nodes already in the network.

3. The probability $\pi(k)$ that one of the links of the new node connects to node $i$ depends on the degree $k_i$ of node $i$ as

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

• The emerging network will be scale-free with degree exponent $\gamma = 3$ independently from the choice of $m_0$ and $m$
The degree of each node increases as a power-law with exponent $\beta=1/2$.

Earlier a node was added larger its degree due to its earlier arrival and not because it grows faster.

Rich-get-richer mechanism
The BA model - degree distribution

- The degree exponent is independent of $m$
- The degree exponent is stationary in time and the degree distribution is time independent
- The exponent is compatible to the exponents of real networks

Algorithm 1

```
BarabásiAlbertModel($m_0$, $m$, $t$)

1: Starting from a fully connected graph with $m_0$ vertices
2: for all timestep $t$ do
3:   add a new node with $m (\frac{m}{m_0})$ edges
4:   for all $m$ number of edges proceed from the new node do
5:     choose a node $i$ to which the new node connects with a probability rational to its degree
       such that: $\gamma(k_i) = k_i \sum_j k_j$.
6:   end for
7: end for
```

This model was defined on undirected graphs and gave a degree exponent $\gamma = 3$ (Figure 3.6), which is a good approach for the exponents of real world networks. It can be solved in $-6$.

Figure 3.6: Degree distribution of Barabási-Albert network for various $m$ with system size $N = 300000$.

The slope of the skew line is $\gamma = 3$ and gives the power-law degree exponent. In the inset we demonstrate, that the degree distributions are independent of $m$ after rescaling as $k \sim P(k)/m^2$. The fitted line gives the expected exponent $\gamma = 3$.

The large size limit with an approximate solution given by Barabási and Albert [15, 125], and it can be proved exactly with two another equivalent methods: the master equation approach published by Dorogovtsev et.al. [126,127] and rate-equation approach introduced by Krapivsky et.al. [128]. For these solutions see Appendix B.

Types of correlation

There are two types of correlation in evolving Barabási-Albert networks found by analytical studies [129]. The first nontrivial correlation is between the age and the degree of a node. The vertices, which were added earlier to the network had better chance to acquire edges, so these nodes have higher degrees. For the case $m = 1$, when the evolving graph is a tree, the probability distribution of the degree of vertex $i$, with age $a$ (the elapsed time since $i$ was given)
The BA model - path length

\[ \langle l \rangle = \frac{\ln N}{\ln \ln N} \]

Ultra Small World

- \begin{align*}
\text{const.} & \quad \gamma = 2 \\
\frac{\ln \ln N}{\ln(\gamma - 1)} & \quad 2 < \gamma < 3 \\
\ln N & \quad \gamma > 3
\end{align*}

Size of the biggest hub is of order \( O(N) \). Most nodes can be connected within two layers of it, thus the average path length will be independent of the system size.

The average path length increases slower than logarithmically. In a random network all nodes have comparable degree, thus most paths will have comparable length. In a scale-free network the vast majority of the path go through the few high degree hubs, reducing the distances between nodes.

Some key models produce \( \gamma = 3 \), so the result is of particular importance for them. This was first derived by Bollobás and collaborators for the network diameter in the context of a dynamical model, but it holds for the average path length as well.

The second moment of the distribution is finite, thus in many ways the network behaves as a random network. Hence the average path length follows the result that we derived for the random network model earlier.

Ultra Small World network

Bollobás, Riordan (2001)
The BA model - clustering coefficient

• The clustering coefficient decreases with the system size as
  \[ C = \frac{m}{4} \frac{(\ln N)^2}{N} \]

• It is 5 times more than for random graphs

Degree correlations:

• The BA model is inducing non-trivial degree correlations due to its definition
  \[ n_{kl} \approx k^{-2} l^{-2} \]
## ER Random Network - catch up

<table>
<thead>
<tr>
<th>Network</th>
<th>Degree distribution</th>
<th>Path length</th>
<th>Clustering coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real world networks</td>
<td>broad</td>
<td>short</td>
<td>large</td>
</tr>
<tr>
<td>Regular lattices</td>
<td>constant</td>
<td>long</td>
<td>large</td>
</tr>
<tr>
<td>ER random networks</td>
<td>Poissonian</td>
<td>short</td>
<td>small</td>
</tr>
<tr>
<td>WS small-world networks</td>
<td>exponential</td>
<td>short</td>
<td>large</td>
</tr>
<tr>
<td>BA scale-free networks</td>
<td>power-law</td>
<td>short</td>
<td>Rather small</td>
</tr>
</tbody>
</table>
(some)
Other random models
Other scale-free models

The vertex-copying model

• Motivation:
  • Citations network or WWW where links are often copied
  • Local explanation to preferential attachment

1. Take a small seed network
2. Pick a random vertex
3. Make a copy of it
4. With probability $p$, move each edge of the copy to point to a random vertex
5. Repeat 2.-4. until the network has grown to desired size of $N$ vertices

• Asymptotically scale-free with exponent $\gamma \geq 3$
Other scale-free models

The Holme-Kim model

• Motivation: more realistic clustering coefficient

1. Take a small seed network

2. Create a new vertex with \( m \) edges

3. Connect the first of the \( m \) edges to existing vertices with a probability proportional to their degree \( k \) (just like BA)

4. With probability \( p \), connect the next edge to a random neighbour of the vertex of step 3., otherwise do 3. again

5. Repeat 2.-4. until the network has grown to desired size of \( N \) vertices

\[
C(k) \propto \frac{1}{k^\gamma}
\]

for large \( N \), ie clustering more realistic! This type of clustering is found in many real-world networks.
Other scale-free models

The forest fire model

• Models generating power laws are usually based on the idea of a growing network.

• It has been shown that real growing networks have two properties: growing average degree, shrinking average shortest path.

Other scale-free models

The forest fire model

Consider a node $v$ joining the network at time $t > 1$, and let $G_t$ be the graph constructed thus far. ($G_1$ will consist of just a single node.) Node $v$ forms out-links to nodes in $G_t$ according to the following process.

(i) $v$ first chooses an ambassador node $w$ uniformly at random, and forms a link to $w$.

(ii) We generate a random number $x$ that is binomially distributed with mean $(1 - p)^{-1}$. Node $v$ selects $x$ links incident to $w$, choosing from among both out-links and in-links, but selecting in-links with probability $r$ times less than out-links. Let $w_1, w_2, \ldots, w_x$ denote the other ends of these selected links.

(iii) $v$ forms out-links to $w_1, w_2, \ldots, w_x$, and then applies step (ii) recursively to each of $w_1, w_2, \ldots, w_x$. As the process continues, nodes cannot be visited a second time, preventing the construction from cycling.

Other scale-free models

The forest fire model

- **Heavy-tailed in-degrees.** Our model has a “rich get richer” flavor: highly linked nodes can easily be reached by a newcomer, no matter which ambassador it starts from.
- **Communities.** The model also has a “copying” flavor: a newcomer copies several of the neighbors of his/her ambassador (and then continues this recursively).
- **Heavy-tailed out-degrees.** The recursive nature of link formation provides a reasonable chance for a new node to burn many edges, and thus produce a large out-degree.
- **Densification Power Law.** A newcomer will have a lot of links near the community of his/her ambassador; a few links beyond this, and significantly fewer farther away. Intuitively, this is analogous to the Community Guided Attachment, although without an explicit set of communities.
- ** Shrinking diameter.** It is not a priori clear why the Forest Fire Model should exhibit a shrinking diameter as it grows. Graph densification is helpful in reducing the diameter, but it is important to note that densification is certainly not enough on its own to imply shrinking diameter. For example, the Community Guided Attachment model obeys the Densification Power Law, but it can be shown to have a diameter that slowly increases.

Rigorous analysis of the Forest Fire Model appears to be quite difficult.

Figure 6: Degree distribution of a sparse graph with decreasing diameter (forward burning probability: 0.37, backward probability: 0.32).

Other scale-free models

Generalized random models

• Exponential Random Graphs
• Random graphs with latent variables
• Graphons
• …

• Nodes have some **known** properties $\theta$ (degrees, attributes…).
• We would like to study **another** property $x$
• We can compare the observed $x$ to the distribution of $x$ among graphs in the space of all possible graphs respecting $\theta$.
• How to efficiently analyse this graph space?
Complex models can have all three properties, but what is the point if they are themselves quite complex?
• “All models are wrong, but some are useful”

• ER models and Configuration models are used as reference models in a very large number of applications

• WS, BA models are more “making a point” type models: simple processes can explain some non-trivial properties of networks, unfound in random networks.

• Correlation is not causation. Are these simple processes the “cause”? Maybe, maybe not, sometimes…