# Network Analysis:

## The Hidden Structures behind the Webs We Weave

## 17-213 / 17-668

## Network Inequality

Tuesday, October 26, 2023

Patrick Park & Bogdan Vasilescu

Carnegie Mellon University
School of Computer Science

S3D
Software and Societal Systems Department

# 2-min Quiz, on Canvas

# How do scale-free networks emerge?

# What does “scale-free” actually mean?

Moments in statistics: Quantitative measures that describe the shape of a distribution

- $n=1$: The first moment is the average degree, $\langle k \rangle$.
- $n=2$: The second moment, $\langle k^2 \rangle$, helps us calculate the variance $\sigma 2 = \langle k^2 \rangle - \langle k \rangle^2$, measuring the spread in the degrees. Its square root, $\sigma$, is the *standard deviation*.
- $n=3$: The third moment, $\langle k^3 \rangle$, determines the *skewness* of a distribution, telling us how symmetric is $p_k$ around the average $\langle k \rangle$.

$$\langle k^n \rangle = \sum_{k_{\min}}^{\infty} k^n p_k \approx \int_{k_{\min}}^{\infty} k^n p(k) dk \qquad (4.19)$$

# What does “scale-free” actually mean?

$$\langle k^n \rangle = \int_{k_{\min}}^{k_{\max}} k^n p(k) dk = C \frac{k_{\max}^{n-\gamma+1} - k_{\min}^{n-\gamma+1}}{n-\gamma+1} \qquad (4.20)$$

- If $n - \gamma + 1 \leq 0$ then the first term on the r.h.s. of (4.20), $k_{max}^{n-\gamma+1}$, goes to zero as $k_{max}$ increases.

Therefore all moments that satisfy $n \leq \gamma - 1$ are finite.

- If $n - \gamma + 1 > 0$ then $\langle k^n \rangle$ goes to infinity as $k_{max} \to \infty$. Therefore all moments larger than $\gamma - 1$ diverge.

# What does “scale-free” actually mean?

$$\langle k^n \rangle = \int_{k_{\min}}^{k_{\max}} k^n p(k) dk = C \frac{k_{\max}^{n-\gamma+1} - k_{\min}^{n-\gamma+1}}{n-\gamma+1} \qquad (4.20)$$

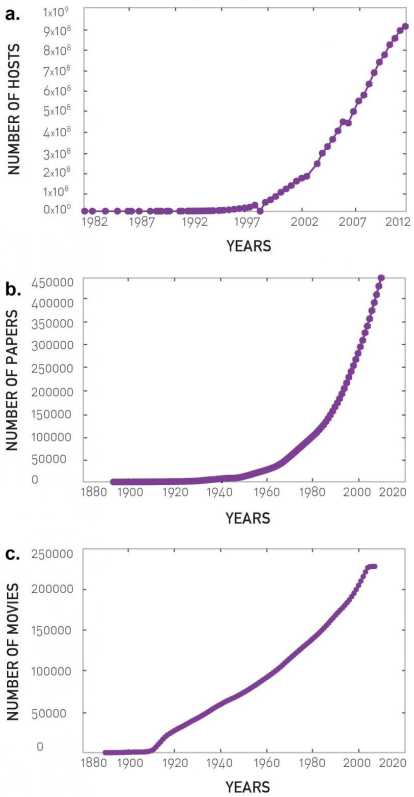
- If $n - \gamma + 1 \leq 0$ then the first term on the r.h.s. of (4.20), $k_{max}^{n-\gamma+1}$, goes to zero as $k_{max}$ increases.

Therefore all moments that satisfy $n \leq \gamma - 1$ are finite.

- If $n - \gamma + 1 > 0$ then $\langle k^n \rangle$ goes to infinity as $k_{max} \to \infty$. Therefore all moments larger than $\gamma - 1$ diverge.

For n=3 (i.e., skew), when power-law exponent is 2<γ<3, the network’s skew infinitely increases with the size of the network

# Simple Model Explaining Scale-Free Property

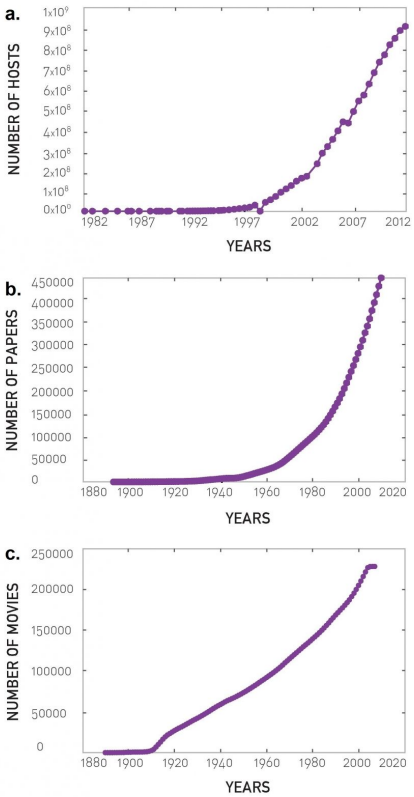



"Preferential attachment" model by Barabasi and Reka Albert

Two assumptions:

- The network infinitely grows, one node added at a time

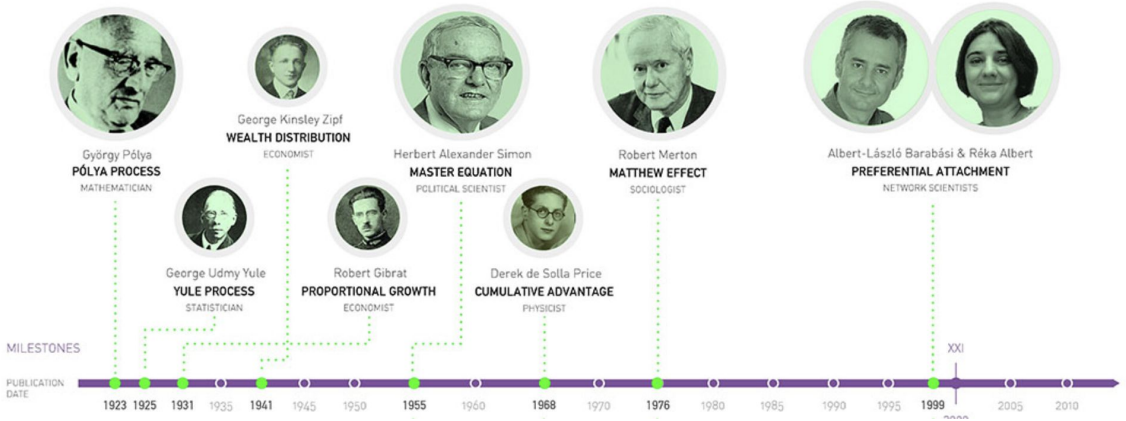
# Simple Model Explaining Scale-Free Property




"Preferential attachment" model by Barabasi and Reka Albert

Two assumptions:

- The network infinitely grows, one node added at a time
- A new node is more likely to link to high degree nodes
  - Rich get richer, "Matthew effect", Zipf's law…



# Simple Model Explaining Scale-Free Property

https://ccl.northwestern.edu/netlogo/models/PreferentialAttachmentSimple

# Degree Distribution and Inequality

# What does γ tell us about inequality?

A social network that is scale-free implies significant social inequality

- few hubs monopolize the edges in a network
- Vast majority of nodes, have degree smaller than <k>

**Q:** From a social justice point of view, which network is closer to an egalitarian, equitable social network: **high γ** or **low γ**?

**Q**: Is it the **extremely high frequency of low-degree nodes** or the **extremely high degree of the few hubs** that determine inequality?

# What does γ tell us about inequality?

Which network is the most unequal?

| Network | $N$ | $L$ | $\langle k \rangle$ | $\langle k_{in}^2 \rangle$ | $\langle k_{out}^2 \rangle$ | $\langle k^2 \rangle$ | $\gamma_{in}$ | $\gamma_{out}$ | $\gamma$ |
|---|---|---|---|---|---|---|---|---|---|
| Internet | 192,244 | 609,066 | 6.34 | - | - | 240.1 | - | - | 3.42* |
| WWW | 325,729 | 1,497,134 | 4.60 | 1546.0 | 482.4 | - | 2.00 | 2.31 | - |
| Power Grid | 4,941 | 6,594 | 2.67 | - | - | 10.3 | - | - | Exp. |
| Mobile-Phone Calls | 36,595 | 91,826 | 2.51 | 12.0 | 11.7 | - | 4.69* | 5.01* | - |
| Email | 57,194 | 103,731 | 1.81 | 94.7 | 1163.9 | - | 3.43* | 2.03* | - |
| Science Collaboration | 23,133 | 93,437 | 8.08 | - | - | 178.2 | - | - | 3.35* |
| Actor Network | 702,388 | 29,397,908 | 83.71 | - | - | 47,353.7 | - | - | 2.12* |
| Citation Network | 449,673 | 4,689,479 | 10.43 | 971.5 | 198.8 | - | 3.03* | 4.00* | - |
| E. Coli Metabolism | 1,039 | 5,802 | 5.58 | 535.7 | 396.7 | - | 2.43* | 2.90* | - |
| Protein Interactions | 2,018 | 2,930 | 2.90 | - | - | 32.3 | - | - | 2.89*- |

# Degree Distribution and Social Inequality

In a social network, large degree indicates influence and power

- Degree centrality

The distribution of node degree reflects inequality in power and influence

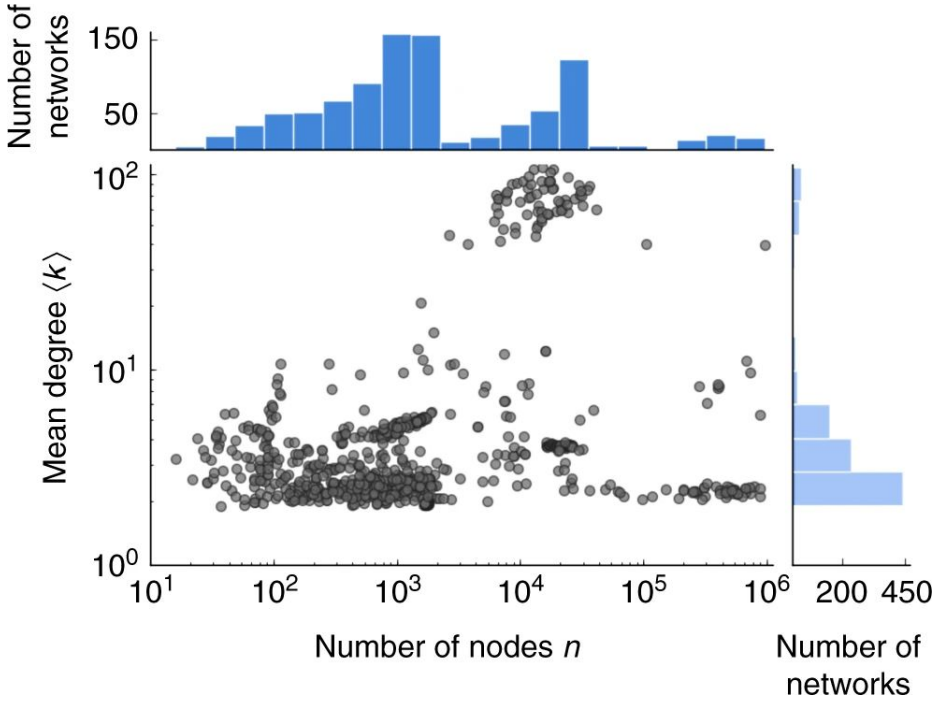
Q: Based on your experience, how extreme is the skew in power and influence?
Q: Does your perception match with the power-law degree distribution?
Q: Is the distribution of power and influence “scale-free”?

Recall, for n=3 (i.e., skew), when power-law exponent is $2<\gamma<3$, the network’s skew infinitely increases with the size of the network

This is not realistic for social networks

# Rarity of scale-free social networks

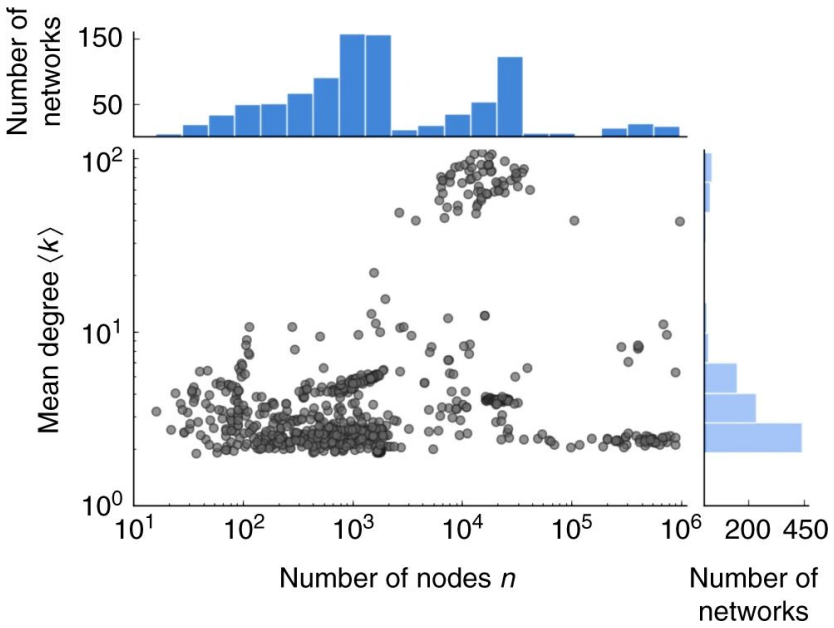
How common are scale-free networks?: Sample of 928 networks




Broido and Clauset 2019

# Rarity of scale-free social networks

How common are scale-free networks?: Sample of 928 networks



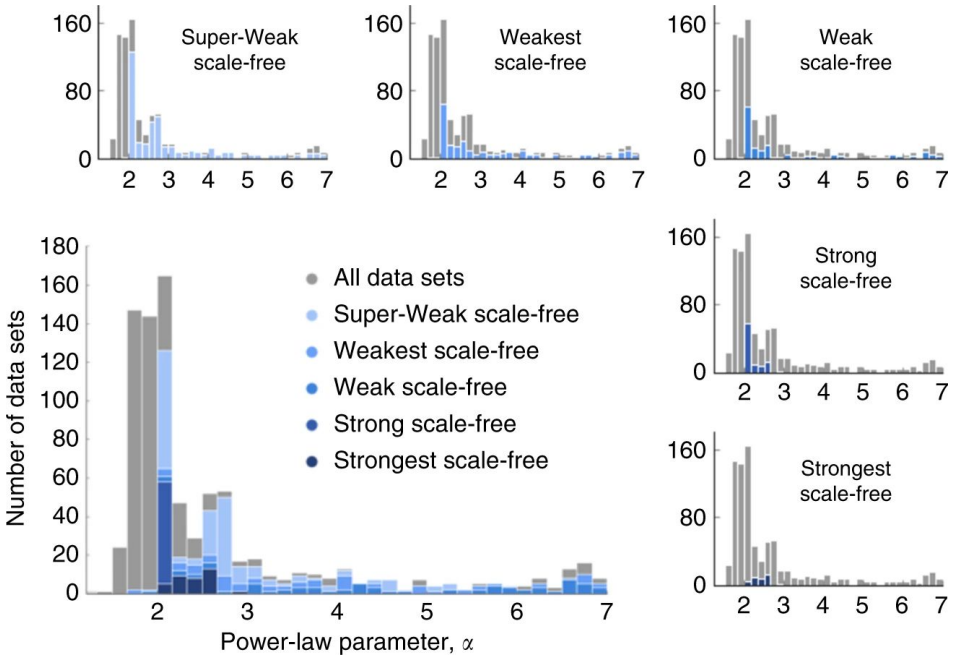

Criterion for judging “scale-freeness”

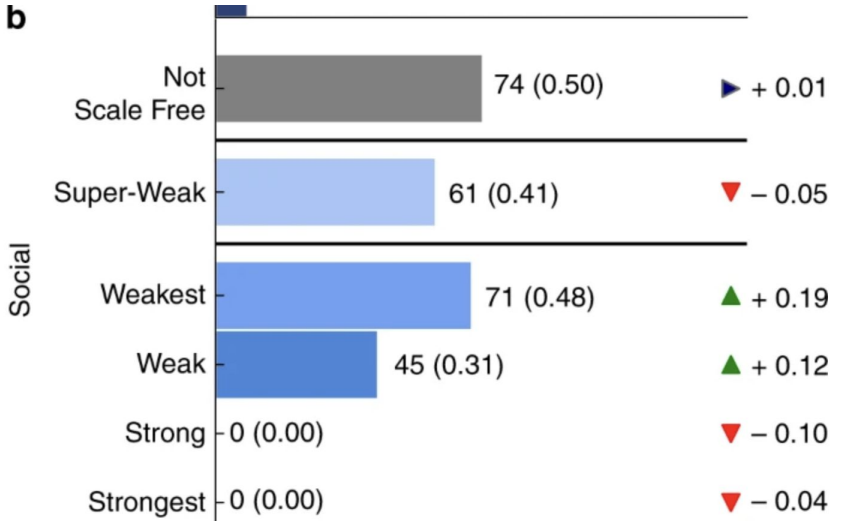
- **Super-Weak**: For at least 50% of graphs, no alternative distribution is favored over the power law.
- **Weakest**: For at least 50% of graphs, a power-law distribution cannot be rejected ($p \geq 0.1$).
- **Weak**: Requirements of Weakest, and the power-law region contains at least 50 nodes ($ntail \geq 50$).
- **Strong**: Requirements of Weak and Super-Weak, and for at least 50% of graphs.
- **Strongest**: Requirements of Strong for at least 90% of graphs, and requirements of Super-Weak for at least 95% of graphs.

Broido and Clauset 2019

# Rarity of scale-free social networks

Most social networks are not scale-free



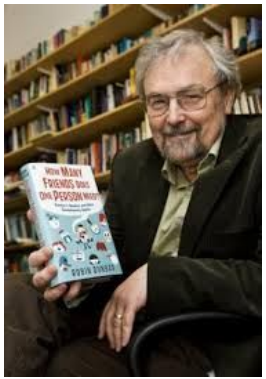




Broido and Clauset 2019

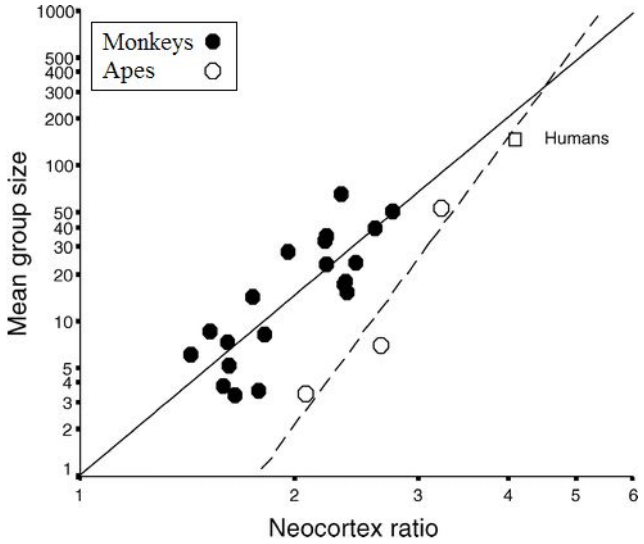
# Why are many social networks not scale-free?

Maintaining a large network is cognitively costly!

- Dunbar’s number: A species group size correlates with brain size
- Human groups have been about 120 people



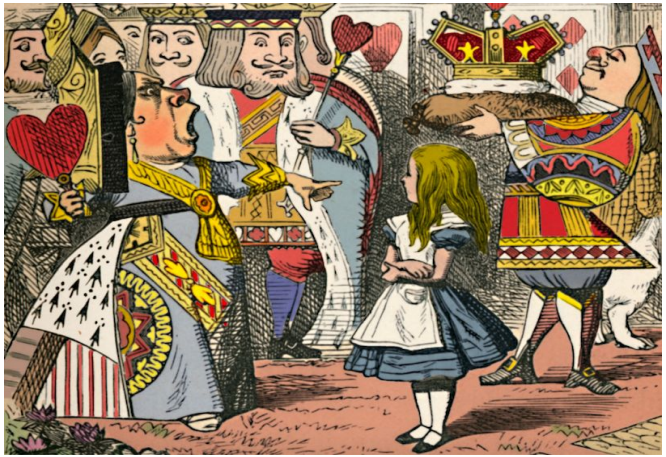
Robin Dunbar



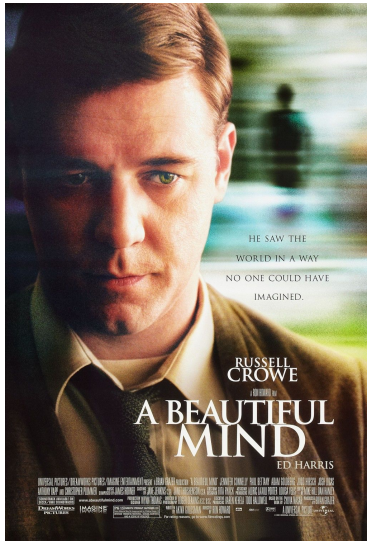
# Why are many social networks not scale-free?

Status distinction in social groups

- Status homophily (Remember degree assortativity?)
- Avoidance of status contamination



# Why are many social networks not scale-free?




Beautiful Mind (2001)

**Individual level:** Low degree nodes have incentive to avoid humiliation / reminder of lower status

**Collective level:** Trying to connect to the highest degree node is not always optimal due to competition (“Adam Smith was wrong”)

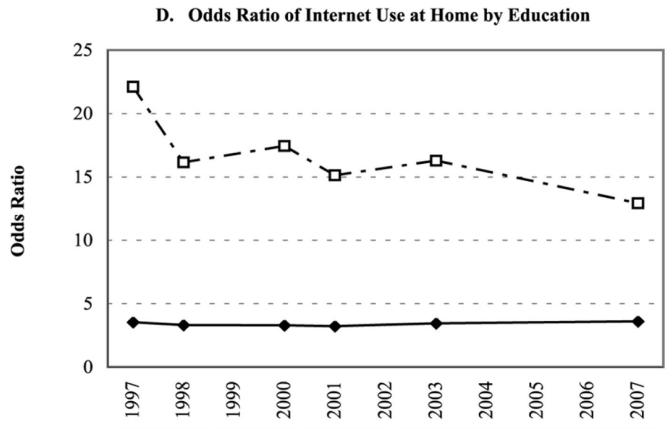
# Other Mechanisms of Network Inequality

# Homophily and Intergroup Inequality

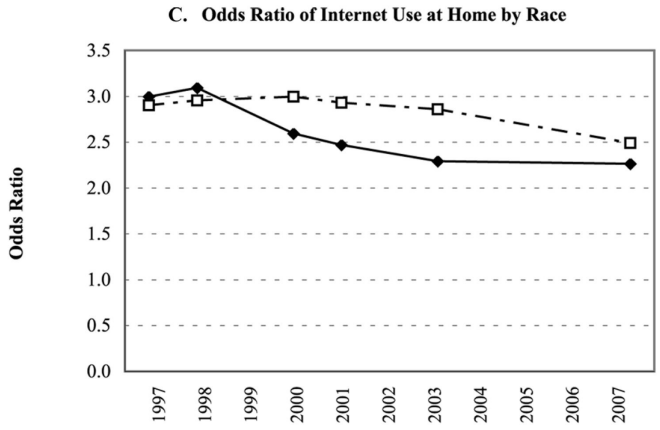
A society with high homophily:

- Beneficial practices / technology diffuse quickly in the already advantaged group, but slowly in the disadvantaged group, leading to intergroup inequality (e.g., internet adoption)

Dimaggio and Garip, 2011




| | 1997 | 1998 | 2000 | 2001 | 2003 | 2007 |
|---|---|---|---|---|---|---|
| College+ vs. HS | 3.529 | 3.325 | 3.291 | 3.233 | 3.456 | 3.611 |
| College+ vs. LHS | 22.104 | 16.138 | 17.437 | 15.123 | 16.291 | 12.916 |




| | 1997 | 1998 | 2000 | 2001 | 2003 | 2007 |
|---|---|---|---|---|---|---|
| White vs. Black | 2.994 | 3.091 | 2.594 | 2.469 | 2.293 | 2.264 |
| White vs. Hispanic | 2.901 | 2.955 | 2.995 | 2.931 | 2.858 | 2.490 |

# Summary

Mechanism of scale-free networks

Social networks often do not follow power-law degree distributions

Scale-free networks → network inequality

Cost and social dynamics matter for the degree distribution (i.e., social inequality)