Network Analysis:

The Hidden Structures behind the Webs We Weave

17-338 / 17-668

Network Inequality

Thursday, October 31, 2024

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, σ, 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?

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

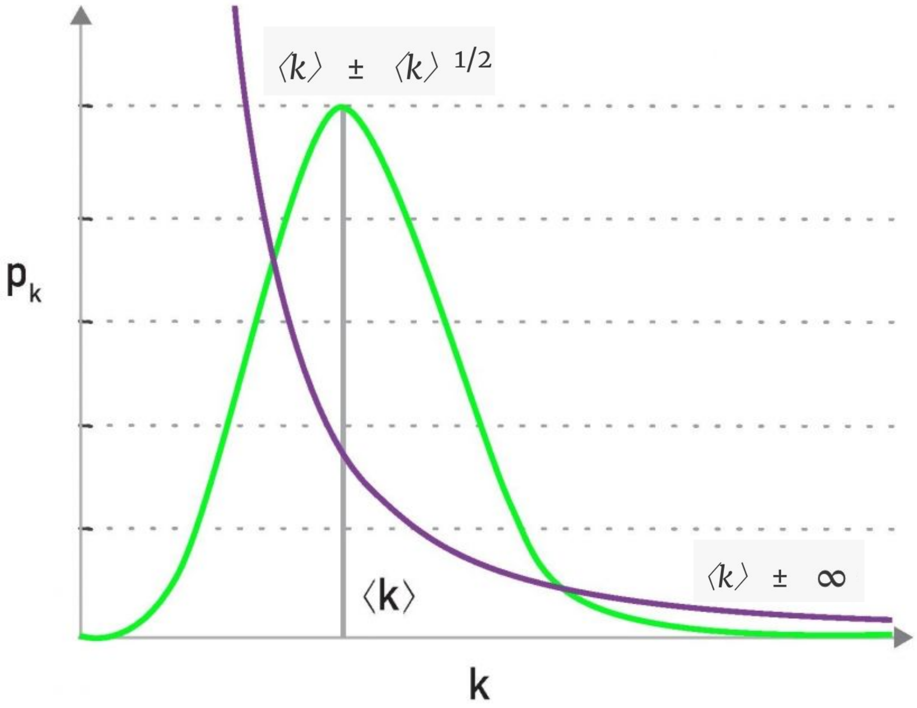
- If $n - \gamma + 1 \leq 0$ then the first term on the r.h.s. of (4.20), ${k_{max}}^{n-\gamma+1}$, goes to zero as k_{max} increases.

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

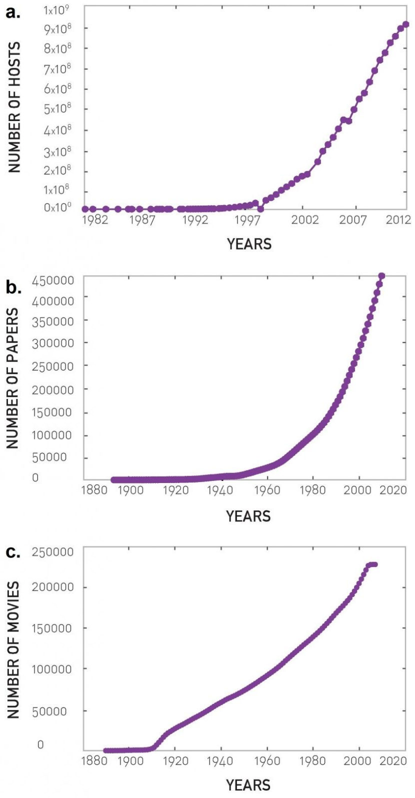
What does “scale-free” actually mean?



Poisson Distribution: Degrees of vast majority of nodes center around <*k*>
→ <*k*> serves as a “scale” that reasonably describes the distribution

Power-law Distribution: Degrees of vast majority of nodes do not center around <k> and some can be arbitrarily large
→ <*k*> is not a reasonable “scale”
→ Hence, “scale-free”

Simple Model Explaining Scale-Free Property



"Preferential attachment" model by Barabasi and Reka Albert

Two assumptions:

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

Simple Model Explaining Scale-Free Property

For to everyone who has will more be given, and he will have abundance;
but from him who has not, even what he has will be taken away.

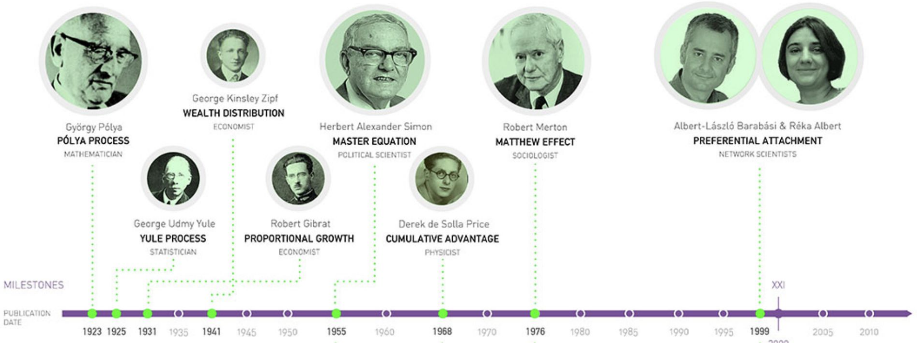
–Matthew 25:29

i.e., **The rich get richer and the poor get poorer**

"Preferential attachment" model by Barabasi and Reka Albert

Two assumptions:

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



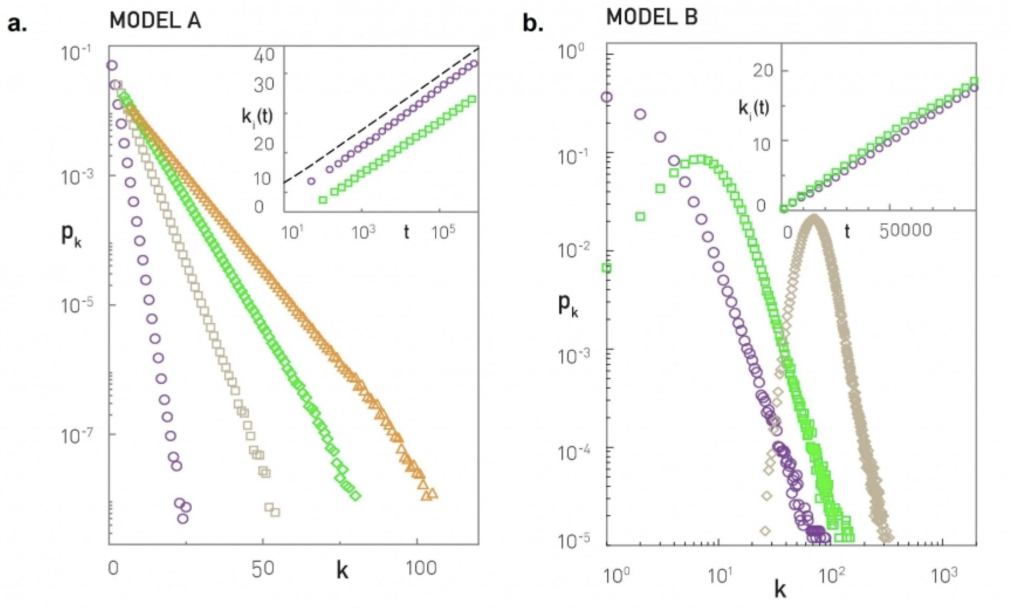
Simple Model Explaining Scale-Free Property

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

Simple Model Explaining Scale-Free Property

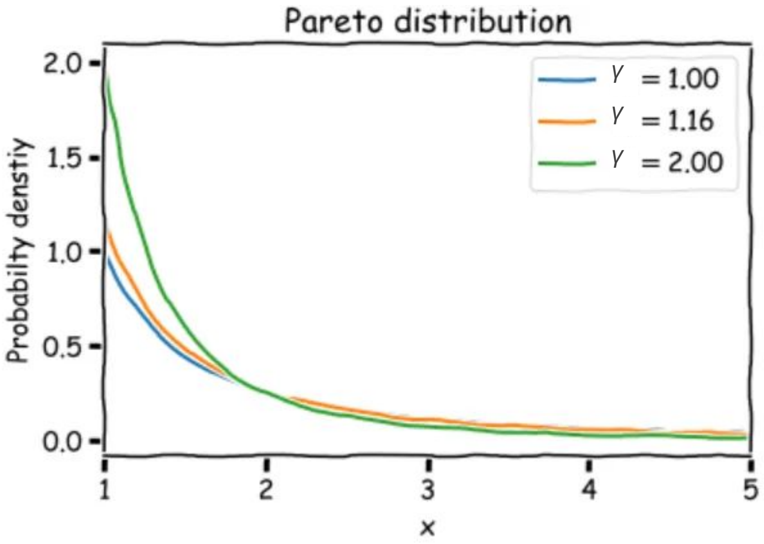
Both conditions are necessary

- Model A: No growth
- Model B: No preferential attachment



Degree Distribution and Inequality

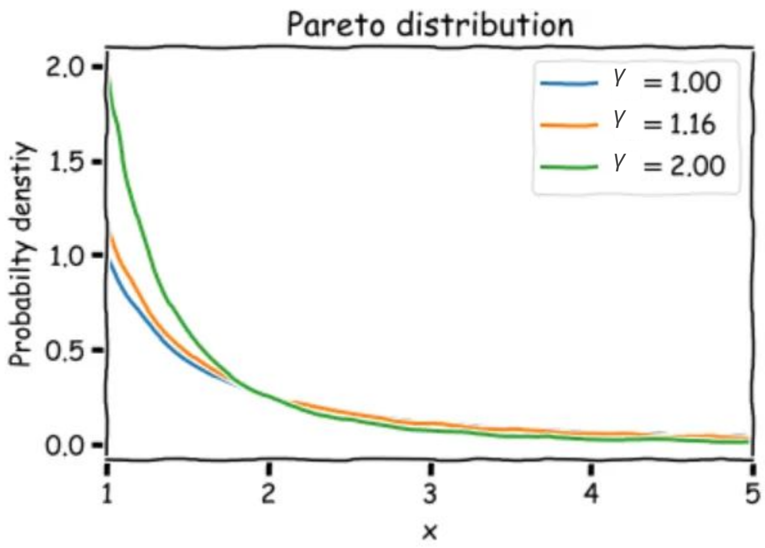
What does γ tell us about inequality?



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*>

What does γ tell us about inequality?



Q: Which 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?

Q: From a social justice perspective, which is preferable: **impoverished society that is egalitarian** vs. **affluent society under dictatorship**?

What does γ tell us about inequality?

Which network is the most unequal?

Network	N	L	$\langle k \rangle$	$\langle k_{in}^2 \rangle$	$\langle k_{out}^2 \rangle$	$\langle k^2 \rangle$	γ_{in}	γ_{out}	γ
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?

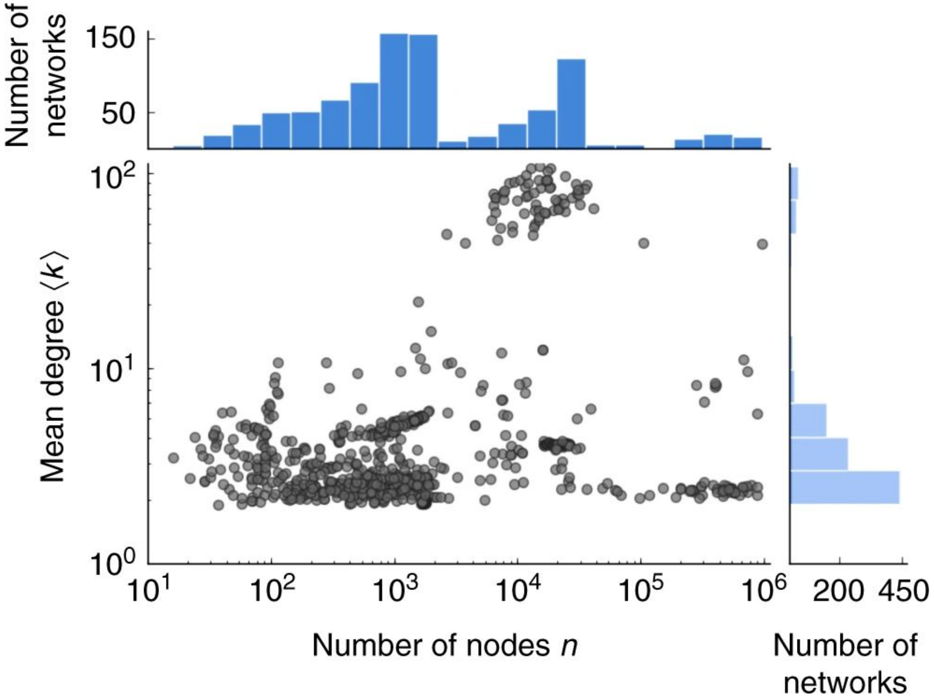
Q: Is the distribution of power and influence “scale-free”?

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

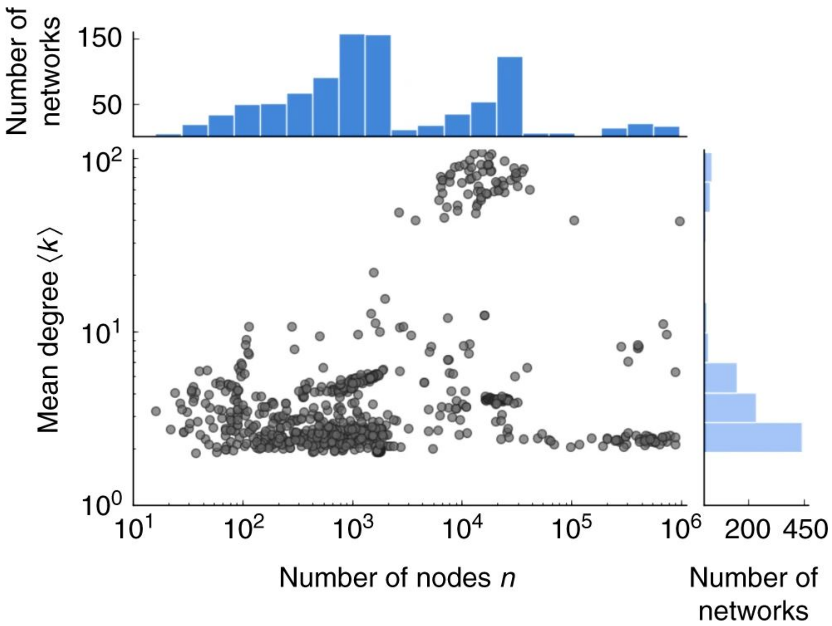
How common are scale-free networks?: Sample of 928 networks



Broido and Clauset 2019

Rarity of scale-free social networks

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



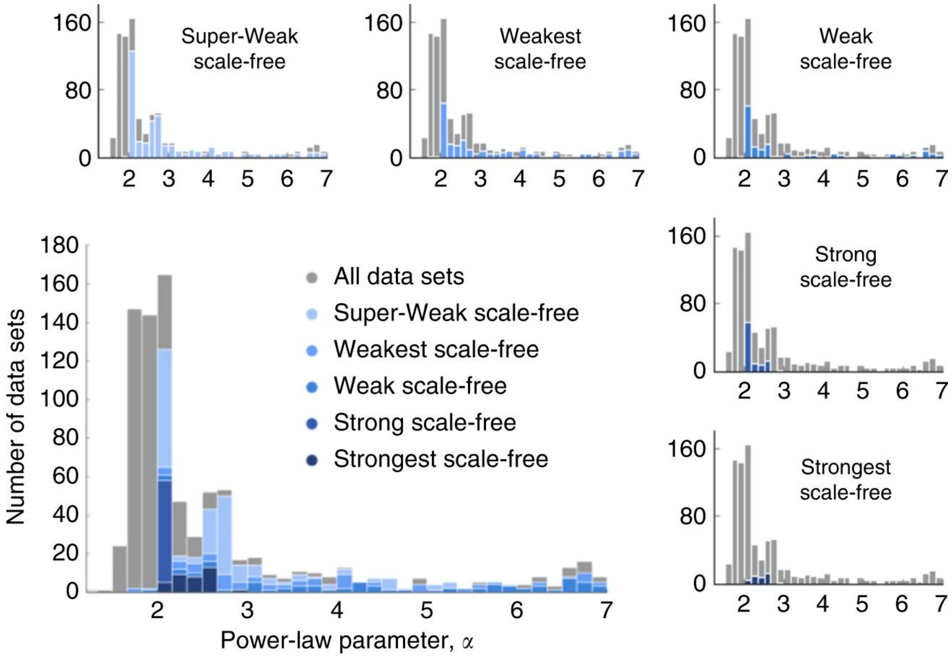
Criterion for judging “scale-freeness”

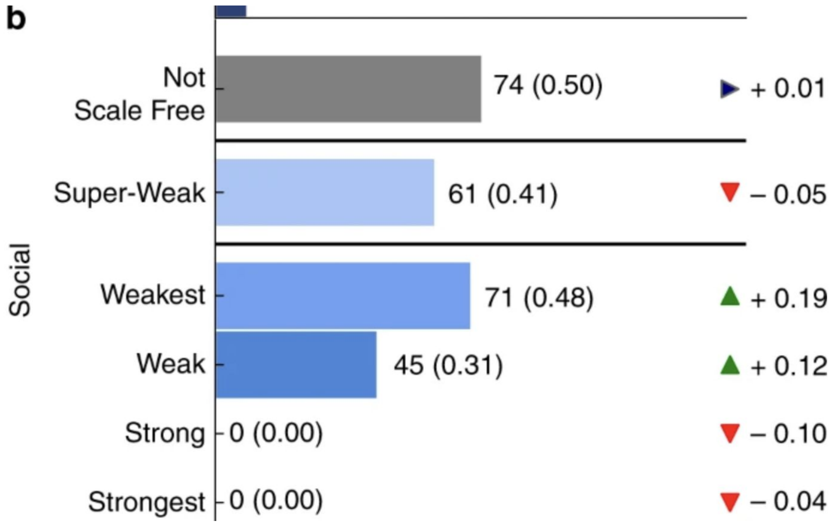
- **Super-Weak**: For at least 50% of graphs, no alternative distribution is favored over the power law.
- **Weakest**: For at least 50% of graphs, a power-law distribution cannot be rejected ($p \geq 0.1$).
- **Weak**: Requirements of Weakest, and the power-law region contains at least 50 nodes (ntail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



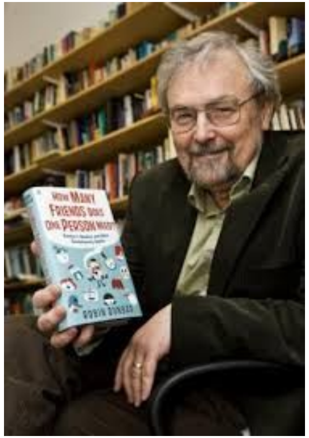


Broido and Clauset 2019

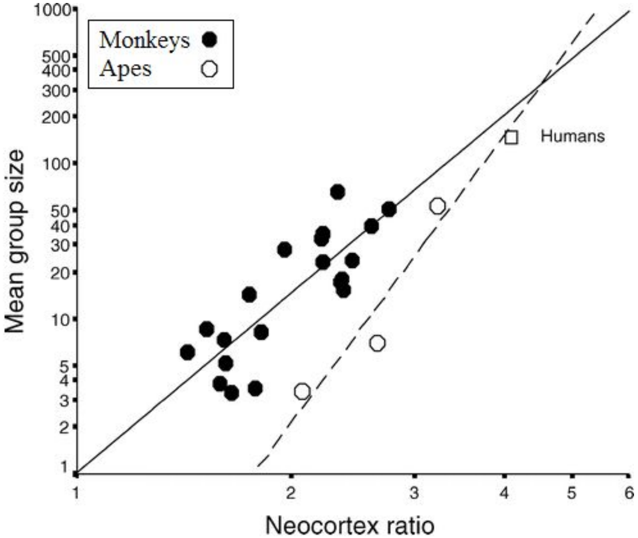
Why are many social networks not scale-free?

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



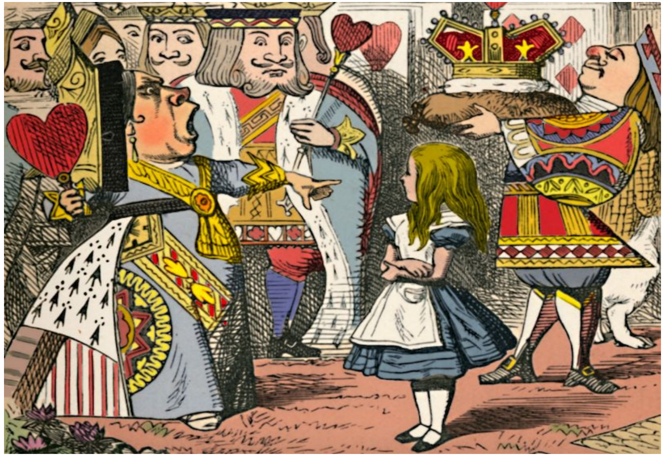
Source: Robin Dunbar



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?

Table 3: Assortativity for BA graph, N=1000

Graph definition		*Assortativity*		
N_0	m	r_{min}	r_{max}	$r_{average}$
3	2	-0.147	-0.038	-0.092
4	2	-0.158	-0.038	-0.089
5	2	-0.135	-0.038	-0.084
10	2	-0.116	-0.006	-0.064
10	3	-0.093	-0.018	-0.055
10	5	-0.078	-0.008	-0.046

Source: Noldus and Mieghem

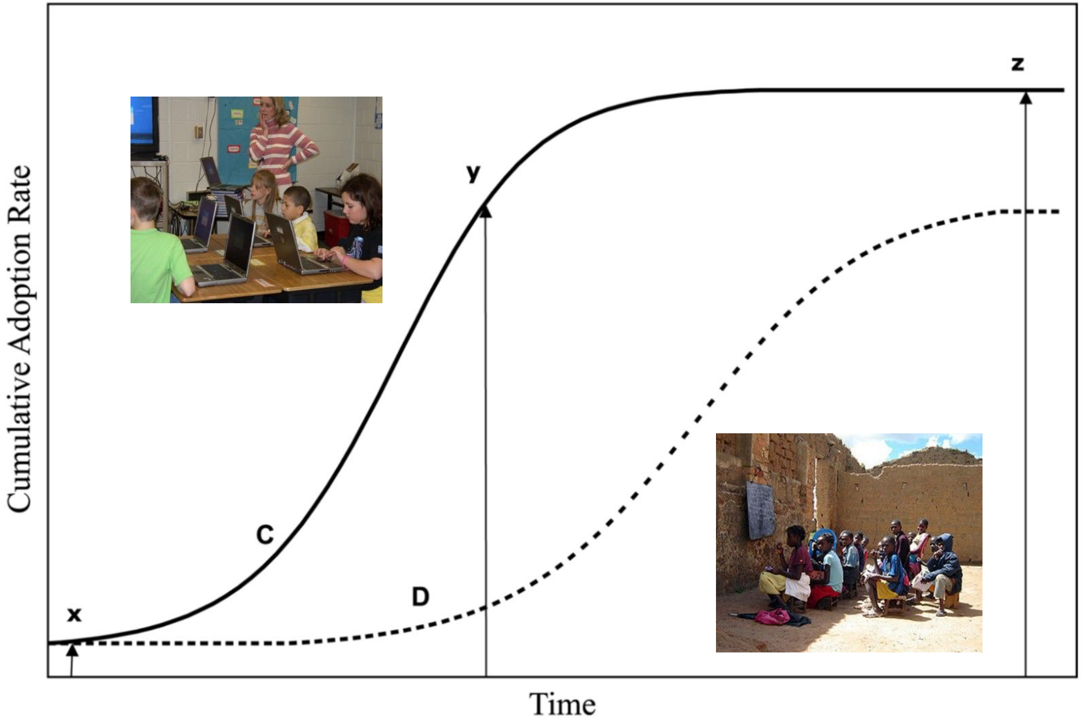
Social networks show positive assortativity
Scale-free networks generated by the BA model are not assortative

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

Other Mechanisms of Network Inequality

Case: Digital Divide



Technology adoption occurs at different rates for different groups

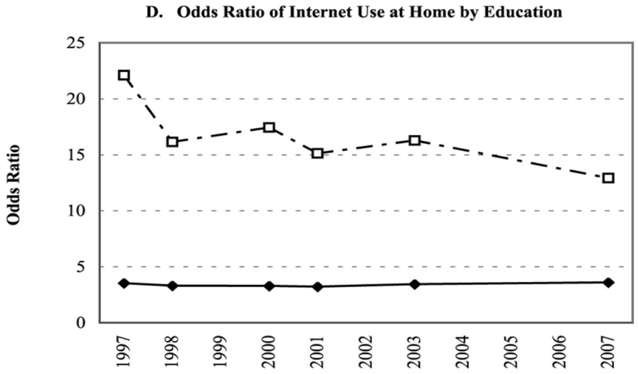
Network structure → Adoption dynamics → Intergroup Inequality

Dimaggio and Garip, 2011

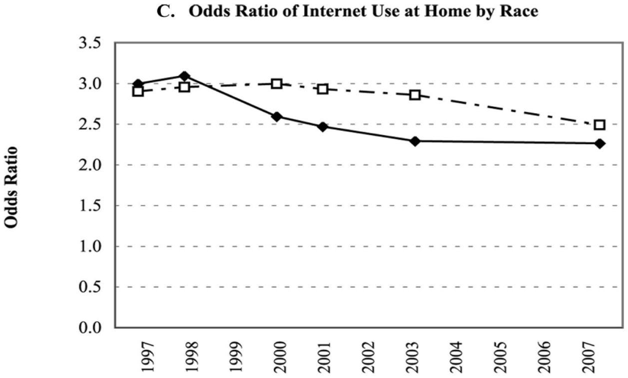
Homophily, Externalities, and Intergroup Inequality

Network externalities
- Value of the technology increases with adoption
- The more your friends use it, the more value to you (e.g., Twitter vs. Mastodon)

Homophily
- Adoption rate difference between groups is greater when social network is very homophilous
- Strong homophily means sparse intergroup connections → adoption is slow in the disadvantaged group



	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

Dimaggio and Garip, 2011

Homophily, Externalities, and Intergroup Inequality

Network externalities

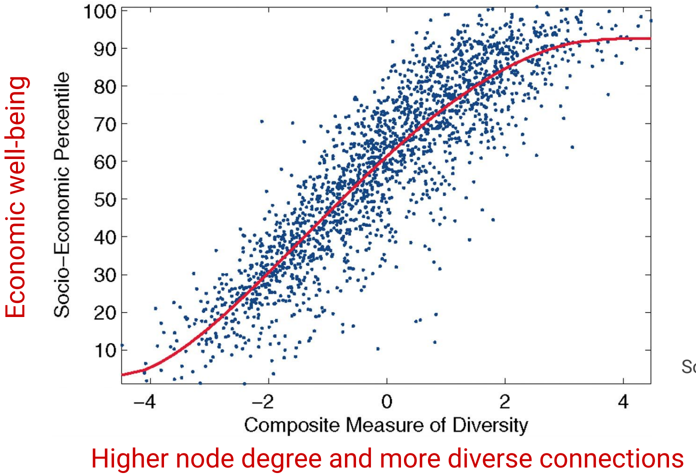
- Value of the technology increases with adoption
- The more your friends use it, the more value to you (e.g., Twitter vs. Mastodon)

Homophily

- Adoption rate difference between groups is greater when social network is very homophilous
- Strong homophily means sparse intergroup connections → adoption is slow in the disadvantaged group

Feedback

- Benefits of adoption can lead to network inequality



Source: Eagle et al. 2009

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)