Emergence of Scaling in Random Networks

Albert-László Barabási and Réka Albert (1999)

February 15, 2023

Motivation

Better understanding complex (large) networks

- Living beings: proteins, genes, and the chemical interactions between them
- Nervous system: nerve cells connected by axons
- Social sciences: individuals and organizations, and the social relationships between them
- Internet: Webpages and their links to other sites

Power Law Observed

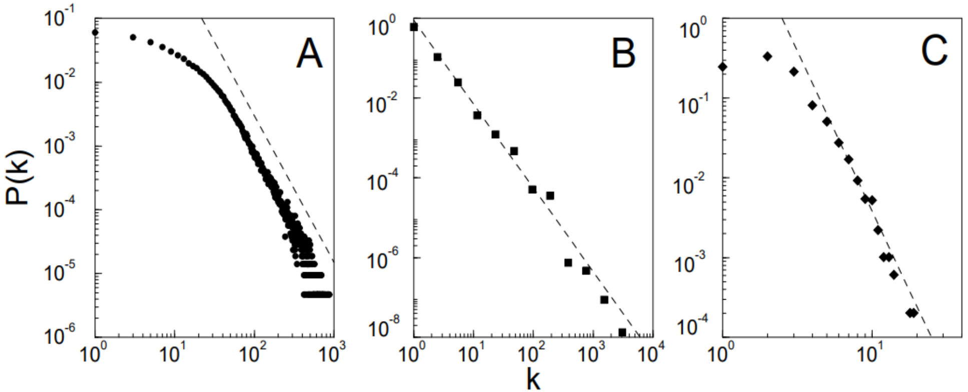
- Power law: $f(x) = ax^{-\gamma}$
- Probability a vertex is of degree $P(k) \sim k^{-\gamma}$
- Scale Invariance: Property of a network that roughly follows this probability in the tail

Topological Data Reviewed

Data Source	N	$\mathbb{E}[k]$	$\hat{\gamma}$
Actor Collaborations	212250	28.78	2.3
Internet Sites	325729	5.46	2.1
Powergrid Configuration	4941	2.67	4

$$P(k) \sim k^{-\gamma}$$

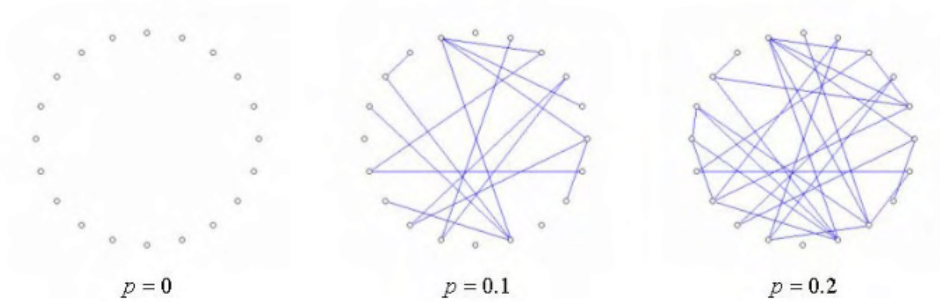
Topological Data Reviewed



How did existing random graph theory reconcile these empirical findings?

Erdös-Rényi-Gilbert Model

Every possible edge is considered for connection with probability p.
There are N total vertices.



https://www.researchgate.net/figure/Erdoes-Renyi-model-of-random-graph-evolution_fig10_313854183

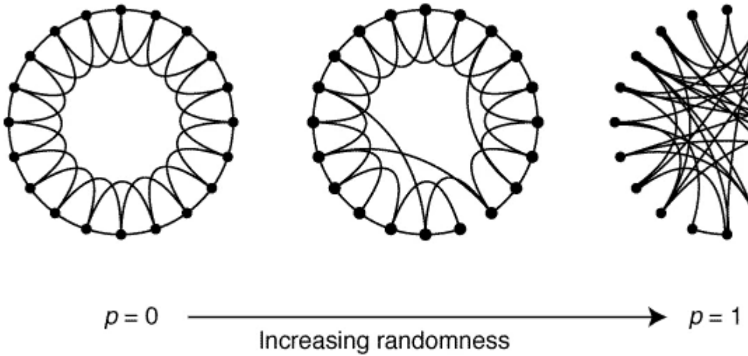
Small World Model

By Watts and Strogatz (1989)

Algorithm:

- Begin with a lattice where every one of the N vertex is connected to z other closest neighboring vertices
- For each original edge, with probability p, disconnect one end of the edge from the vertex and reconnect to any other vertex the original vertex is not connected to

Small World Model



https://www.nature.com/articles/30918

Probability of Vertex Degree k

- ERG: $P(k) \sim$ Poisson
- Small World: $P(k)$ centers around the starting vertex degree z

The power-law tail indicates highly connected vertices have a large chance of occurring, unlike in the existing models of the time.

$$\boxed{P(k) \sim k^{-\gamma}}$$

New Model: Assumptions

The other models assumed the total number of vertices N to be fixed, and that edges and randomly connected or reconnected.

Proposed model assumes:

- Growth: N increases over the network's lifetime
- Preferential Attachment: new vertices prefer to connect with high-degree vertices

A model built with these assumption shows the desired power-law decay result.

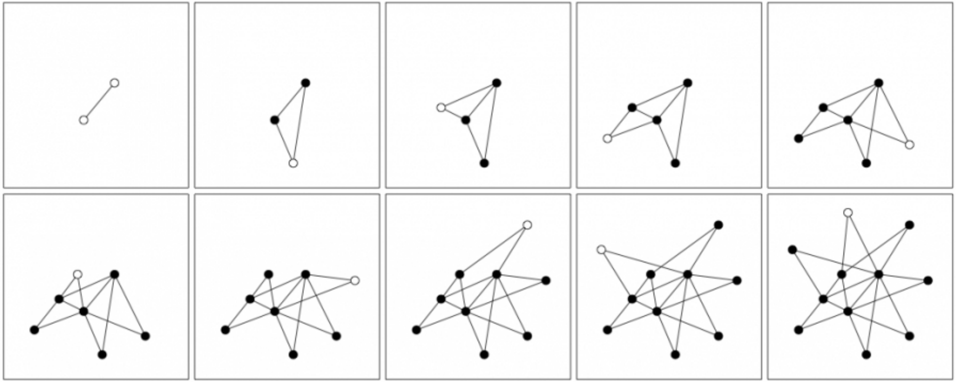
New Model

- Begin with m_o vertices (with positive degrees), fix some $m \leq m_o$
- Every timestep, add a new vertex and connect it to m existing vertices preferentially
- Probability the new vertex will be connected to vertex $i = \Pi(k_i)$ where k_i is the degree of vertex i

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

After t timesteps, yields a model with $t + m_o$ vertices and tm edges.

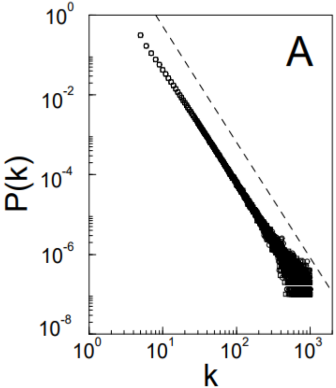
Visualization, $m_o = m = 2$



http://networksciencebook.com/chapter/5barabasi-model

Model Implications

They found: $P(k) \sim k^{-\gamma}$ $\gamma_{model} = 2.9 \pm 0.1$



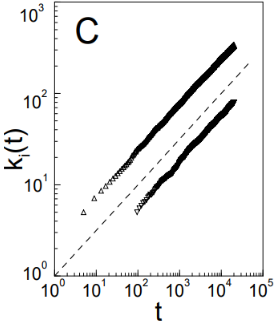
Distribution at t = 150,000 (o) and t = 200,000 (square)

Model Implications

When either assumption is violated, this result does not occur

- If preferential connectivity is not enforced, $P(k) \sim exp(-\beta k)$
- If N is fixed, eventually all vertices will be directly connected

Also note the degree of a vertex depends on when it was added



Deriving Gamma

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_j k_j} = m\frac{k_i}{2mt-m} \approx \frac{k_i}{2t}$$

$$k_i(t) = m(t/t_i)^{0.5}$$

Where t_i represents the time vertex i was added to the network

$$\begin{aligned} P(k_i(t) < k) &= P(t_i > m^2t/k^2) = 1 - P(t_i \leq m^2t/k^2) \\ &= 1 - P(t_i \leq m^2t/k^2) = 1 - m^2t/[k^2(t+m_o)] \end{aligned}$$

Under the assumption that vertices are added at equal time intervals

https://barabasi.com/f/622.pdf

Deriving Gamma

$$
\begin{aligned}
P(k) &= \partial P(k_i(t) < k)/\partial k \\
&= \frac{\partial}{\partial k}(1 - m^2 t/[k^2(t + m_o)]) \\
&= \frac{2m^2 t}{(t + m_o)k^3}
\end{aligned}
$$

Over long times, they found

$$\boxed{P(k) = \frac{2m^2}{k^3} \sim k^{-3}}$$

$$\boxed{P(k) \sim k^{-\gamma}}$$

Discussion

- Considered nonlinear preferential attachment $\Pi(k_i) = \frac{k_i^\alpha}{\sum_j k_j^\alpha}$ but chose $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$ as simulations showed scaling for only $\alpha = 1$
- In the empirical networks they looked at, γ ranged from 2.1 to 4
- If p of the edges are directed $\gamma(p) = 3 - p$ is reasonable and supported by numerical simulation

Conclusion

- Growth and preferential attachment are mechanisms common to many complex systems
- Large networks exhibit power-law decay in the probability a vertex has a particular degree (scale-free)
- Networks with very different origins can display similar behavior (self-organization)