Universitat Pompeu Fabra Barcelona

Lecture 10 · Power laws and Hubs; Beyond Erdős–Rényi: Models for Real Networks

Networks, Crowds and Markets

Motivation: What ER misses

- In Lecture 9 we saw that real networks have:
    - high clustering,
    - heavy-tailed degree distributions,
    - and short average distances.
- Erdős–Rényi models explain only the last of these.
- Today we look closer into the power law and start building richer models that match all three.

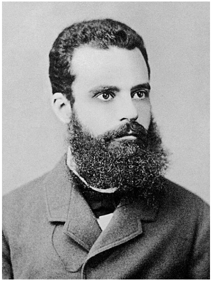
Today's Lecture

1. Power laws and hubs
2. Universality of power laws across networks.
3. Distances: small world vs ultra-small world.
4. Statistic random network models
5. Random network models with prescribed degree distribution.

Power laws and hubs

Historical roots: Pareto and the 80/20 law

Vilfredo Pareto (1848–1923), Italian economist, observed that income distribution in society is very uneven.



- Incomes followed a distribution with a heavy tail: a small fraction of people held most of the wealth.
- This became the well-known "80/20 rule": e.g. 20% of people control 80% of wealth.
- Similar patterns appear in many domains:
    - 80% of web links point to about 20% of webpages.
    - A small number of firms or banks control a large share of markets.
    - A few researchers or papers receive most citations.

Connection: Pareto's law is an early example of a *power law* in economics, closely related to what we now see in network degree distributions.

Power law: Discrete formalism

Assume all degrees are $\geq k_{\mathrm{min}} \geq 1$. If needed, model separately high and low degree nodes. Power law is about high degree nodes.

Power law: Discrete formalism

Assume all degrees are $\geq k_{\text{min}} \geq 1$. If needed, model separately high and low degree nodes. Power law is about high degree nodes.

We model the degree distribution as

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma, k_{\text{min}})}, \qquad k \geq k_{\text{min}},$$

where $\zeta(\gamma, k_{\text{min}}) = \sum_{k=k_{\text{min}}}^{\infty} k^{-\gamma}$ is the normalizing constant.

Power law: Discrete formalism

Assume all degrees are $\geq k_{\min} \geq 1$. If needed, model separately high and low degree nodes. Power law is about high degree nodes.

We model the degree distribution as

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma, k_{\min})}, \qquad k \geq k_{\min},$$

where $\zeta(\gamma, k_{\min}) = \sum_{k=k_{\min}}^{\infty} k^{-\gamma}$ is the normalizing constant.

$\zeta(\gamma, k_{\min})$ is the Hurwitz zeta function; for $k_{\min} = 1$ it reduces to the Riemann zeta $\zeta(\gamma)$.

- The series converges if and only if $\gamma > 1$.
- In many real networks, empirical exponents satisfy $2 < \gamma \leq 3$.

First two moments

If $Z \sim (p_k)$ with $p_k = \dfrac{k^{-\gamma}}{\zeta(\gamma, k_{\text{min}})}$ for $k \geq k_{\text{min}}$, then

$$\mathbb{E}Z = \sum_{k \geq k_{\text{min}}} k\, p_k = \frac{1}{\zeta(\gamma, k_{\text{min}})} \sum_{k \geq k_{\text{min}}} k^{-(\gamma-1)} = \frac{\zeta(\gamma-1, k_{\text{min}})}{\zeta(\gamma, k_{\text{min}})},$$

$$\mathbb{E}Z^2 = \frac{1}{\zeta(\gamma, k_{\text{min}})} \sum_{k \geq k_{\text{min}}} k^{-(\gamma-2)} = \frac{\zeta(\gamma-2, k_{\text{min}})}{\zeta(\gamma, k_{\text{min}})}.$$

First two moments

If $Z \sim (p_k)$ with $p_k = \dfrac{k^{-\gamma}}{\zeta(\gamma, k_{\min})}$ for $k \geq k_{\min}$, then

$$\mathbb{E}Z = \sum_{k \geq k_{\min}} k\, p_k = \frac{1}{\zeta(\gamma, k_{\min})} \sum_{k \geq k_{\min}} k^{-(\gamma-1)} = \frac{\zeta(\gamma-1, k_{\min})}{\zeta(\gamma, k_{\min})},$$

$$\mathbb{E}Z^2 = \frac{1}{\zeta(\gamma, k_{\min})} \sum_{k \geq k_{\min}} k^{-(\gamma-2)} = \frac{\zeta(\gamma-2, k_{\min})}{\zeta(\gamma, k_{\min})}.$$

The regime $2 < \gamma \leq 3$ is special:

- Since $\gamma - 1 > 1$, the mean exists.
- Since $\gamma - 2 \leq 1$, the variance does not!

(a very heavy-tailed distribution)

Power law: Continuum formalism

Sums like $\sum_{k \geq k_{\text{min}}} k^{-\gamma}$ are hard to handle algebraically. For large networks (and large k_{min}), approximate the sum by an integral:

$$\sum_{k=k_{\text{min}}}^{\infty} k^{-\gamma} \approx \int_{k_{\text{min}}}^{\infty} k^{-\gamma} \, \mathrm{d}k.$$

Power law: Continuum formalism

Sums like $\sum_{k \geq k_{\min}} k^{-\gamma}$ are hard to handle algebraically. For large networks (and large $k_{\min}$), approximate the sum by an integral:

$$\sum_{k=k_{\min}}^{\infty} k^{-\gamma} \approx \int_{k_{\min}}^{\infty} k^{-\gamma} \, \mathrm{d}k.$$

Define a density $p(k) = C\, k^{-\gamma}$ for $k \geq k_{\min}$. Normalize:

$$1 = \int_{k_{\min}}^{\infty} p(k) \, \mathrm{d}k = C \int_{k_{\min}}^{\infty} k^{-\gamma} \, \mathrm{d}k = C \, \frac{k_{\min}^{1-\gamma}}{\gamma - 1} \quad \Rightarrow \quad C = (\gamma - 1) \, k_{\min}^{\gamma - 1}.$$

$$\boxed{p(k) = (\gamma - 1) \, k_{\min}^{\gamma-1} \, k^{-\gamma}, \qquad k \geq k_{\min}.}$$

Power law: Continuum formalism

Sums like $\sum_{k \geq k_{\min}} k^{-\gamma}$ are hard to handle algebraically. For large networks (and large $k_{\min}$), approximate the sum by an integral:

$$\sum_{k=k_{\min}}^{\infty} k^{-\gamma} \approx \int_{k_{\min}}^{\infty} k^{-\gamma} \, \mathrm{d}k.$$

Define a density $p(k) = C\, k^{-\gamma}$ for $k \geq k_{\min}$. Normalize:

$$1 = \int_{k_{\min}}^{\infty} p(k) \, \mathrm{d}k = C \int_{k_{\min}}^{\infty} k^{-\gamma} \, \mathrm{d}k = C \, \frac{k_{\min}^{1-\gamma}}{\gamma - 1} \quad \Rightarrow \quad C = (\gamma - 1)\, k_{\min}^{\gamma-1}.$$

$$\boxed{p(k) = (\gamma - 1)\, k_{\min}^{\gamma-1}\, k^{-\gamma}, \qquad k \geq k_{\min}.}$$

$C \approx 1/\zeta(\gamma, k_{\min})$ with relative error $O(k_{\min}^{-\gamma})$ for fixed $\gamma > 1$.

Extreme value of a power law: scaling of $k_{\max}$

With $p(k) = (\gamma - 1)\, k_{\min}^{\gamma-1}\, k^{-\gamma}$, the survival tail is

$$\mathbb{P}(K \geq k) = \int_k^\infty p(x)\, dx = \left(\frac{k_{\min}}{k}\right)^{\gamma-1}, \qquad k \geq k_{\min}.$$

Extreme value of a power law: scaling of $k_{\max}$

With $p(k) = (\gamma - 1)\, k_{\min}^{\gamma-1}\, k^{-\gamma}$, the survival tail is

$$\mathbb{P}(K \geq k) = \int_k^{\infty} p(x)\, dx = \left(\frac{k_{\min}}{k}\right)^{\gamma-1}, \qquad k \geq k_{\min}.$$

In a network with N nodes, we estimate the max-degree $k_{\max}$ by

$$\Pr(K \geq k_{\max}) \;\approx\; \tfrac{1}{N} \quad\Longrightarrow\quad \left(\frac{k_{\min}}{k_{\max}}\right)^{\gamma-1} \approx \tfrac{1}{N}$$

$$k_{\max} \;\approx\; k_{\min}\, N^{1/(\gamma-1)}.$$

Notes.

- This captures the correct order; fluctuations are smaller-order.
- The same scaling holds for the discrete model up to constants.

Consequences of the $k_{\max}$ scaling

From $k_{\max} \approx k_{\min} N^{1/(\gamma-1)}$:

$$\begin{aligned}
\gamma = 2 \quad &\Rightarrow \quad k_{\max} \sim k_{\min} N \quad \text{(a single hub touches a linear fraction)} \\
2 < \gamma < 3 \quad &\Rightarrow \quad k_{\max} \sim k_{\min} N^{1/(\gamma-1)} \text{ sublinear but large} \\
\gamma = 3 \quad &\Rightarrow \quad k_{\max} \sim k_{\min} N^{1/2} \\
\gamma > 3 \quad &\Rightarrow \quad k_{\max} \text{ grows slowly; tails are lighter}
\end{aligned}$$

Path lenghts in Scale-Free Networks

Average path length in random networks

Let $d(u,v)$ be the distance between two vertices and $\mathbb{E}[d]$ the average distance across all pairs.

- In Erdős–Rényi graphs with mean degree c fixed,

$$\mathbb{E}[d] \sim \frac{\ln N}{\ln c}.$$

- In scale-free networks with degree tail $p_k \sim k^{-\gamma}$,

$$\mathbb{E}[d] \sim \begin{cases} \text{constant}, & \gamma = 2, \\ \ln\ln N, & 2 < \gamma < 3, \\ \frac{\ln N}{\ln\ln N}, & \gamma = 3, \\ \ln N, & \gamma > 3. \end{cases}$$

Idea: The scaling of $\mathbb{E}[d]$ reflects how large the biggest hub can grow, $k_{\text{max}} \approx k_{\text{min}} N^{1/(\gamma-1)}$, and how efficiently hubs act as shortcuts.

Case $\gamma = 2$: hub-and-spoke regime

Here $\mathbb{E}[d] = O(1)$.

- From $k_{\max} \sim k_{\min} N^{1/(\gamma-1)}$, we get $k_{\max} \sim N$: one hub connects to almost all nodes.
- The graph becomes star-like (*hub–and–spoke* structure). Any two peripheral nodes connect via the hub in at most two steps.
- Therefore $\mathbb{E}[d]$ remains bounded independently of N.
- Networks with $\gamma = 2$ are extremely centralized and fragile to hub removal.

Case $2 < \gamma < 3$: ultra-small world

- Here $k_{\text{max}} \sim k_{\text{min}} N^{1/(\gamma-1)}$ grows faster than any power of $\ln N$ but slower than N.
- A few very large hubs act as shortcuts, giving

$$\mathbb{E}[d] \sim \ln\ln N \quad (\text{“ultra-small world”}).$$

- The mean degree $\mathbb{E}[\text{deg}]$ is finite but $\mathbb{E}[\text{deg}^2] = \infty$: variance diverges, so hubs dominate connectivity.
- Most empirical scale-free networks (social, technological, biological) fall in this range.

Case $\gamma = 3$: critical point

- The largest degree scales as $k_{\max} \sim N^{1/2}$.
- The second moment $\mathbb{E}[\deg^2]$ stops diverging but is still large.
- This produces a slower, logarithmically corrected growth:

$$\mathbb{E}[d] \; \sim \; \frac{\ln N}{\ln\ln N}.$$

- Paths are longer than in the $\gamma < 3$ case but still shorter than in Erdős–Rényi graphs.

Case $\gamma > 3$: small-world regime

- Both mean and variance of deg are finite: hubs are limited in size.
- $k_{\max} \sim N^{1/(\gamma-1)}$ grows slowly, producing no global shortcuts.
- The average distance recovers the classic small-world scaling:

$$\mathbb{E}[d] \sim \ln N.$$

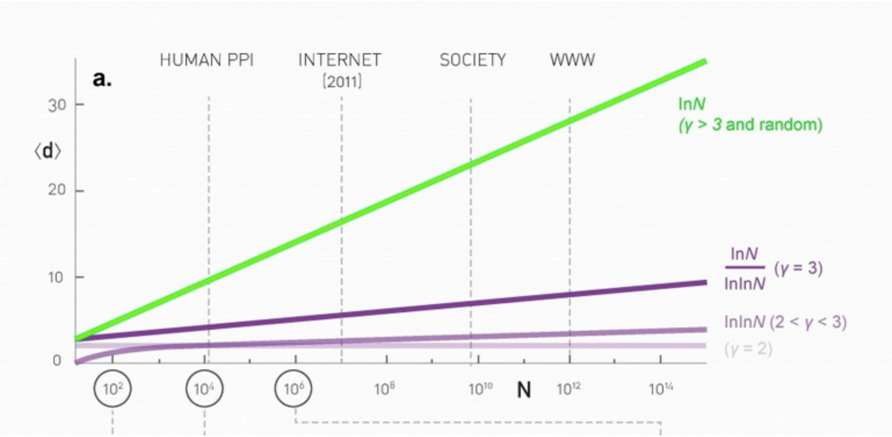
- This regime behaves similarly to Erdős–Rényi graphs in terms of average distance.

When $\gamma < 2$: nonphysical limit

- Then $1/(\gamma - 1) > 1$, so

$$k_{\mathsf{max}} \;\sim\; k_{\mathsf{min}}\, N^{1/(\gamma-1)}$$

  grows faster than N.
- This would require nodes of degree larger than the entire network — impossible in a simple graph.
- Moreover $\mathbb{E}[\deg]$ diverges: even the mean degree is infinite.
- $\Rightarrow$ Infinite scale-free networks with $\gamma < 2$ cannot exist; finite networks must have an effective cutoff.



Note that for large networks the difference in average degrees between the four regimes is much larger than for small networks.

Conclusions

In summary, the effects on distances in scale-free networks are:

- They **shrink average path lengths**. Most scale-free networks of practical interest are "ultra-small", because hubs act as bridges linking many low-degree nodes.
- They **change the scaling of $\mathbb{E}[d]$ with system size**: the smaller the exponent γ, the shorter the distances between nodes.
- Only for $\gamma > 3$ do we recover the $\mathbb{E}[d] \sim \ln N$ scaling — the *small-world* behavior characteristic of Erdős–Rényi graphs.

Next: we explore richer models that explain how such networks emerge.

Need for more sophisticated models

Erdős–Rényi: clean benchmark for randomness in networks.

- Degrees: Binomial $\to$ Poisson in sparse regime, sharply concentrated (Hoeffding).
- Sharp thresholds: giant component at $p \sim 1/N$, full connectivity at $p \sim (\log N)/N$.

Analytic power: every property can be studied precisely—gives language for thresholds, asymptotics, and "with high probability" results.

But realism is limited:

- Clustering $\mathbb{E}[C_v] = p \to 0$ as $N \to \infty$ (in the sparse regime).
- Degree distribution thin-tailed: no hubs or communities.
- Real social, financial, and web networks are way more structured.

This motivates a study of other random graph models.

Static random graph models

Graphs as random objects

Consider an undirected graph $G = (V, E)$.

Order all pairs of elements in V: $\{1, 2\}, \{1, 3\}, \dots, \{N - 1, N\}$.

Each graph is uniquely identified by a vector $\boldsymbol{y} = (y_{ij}) \in \{0, 1\}^{\binom{N}{2}}$:

- $y_{ij} = 1$ if and only if $ij \in E$.

Graphs as random objects

Consider an undirected graph $G = (V, E)$.

Order all pairs of elements in V: $\{1,2\}, \{1,3\}, \dots, \{N-1, N\}$.

Each graph is uniquely identified by a vector $\boldsymbol{y} = (y_{ij}) \in \{0,1\}^{\binom{N}{2}}$:

- $y_{ij} = 1$ if and only if $ij \in E$.

In this sense, every distribution for a random binary vector in $\{0,1\}^{\binom{N}{2}}$ gives a distribution of a random graph with N nodes.

e.g. $(p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) = (\frac{1}{2}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14})$ gives a distribution over 3-node graphs.

Graphs as random objects

Consider an undirected graph $G = (V, E)$.

Order all pairs of elements in V: $\{1,2\},\{1,3\},\ldots,\{N-1,N\}$.

Each graph is uniquely identified by a vector $\boldsymbol{y} = (y_{ij}) \in \{0,1\}^{\binom{N}{2}}$:

- $y_{ij} = 1$ if and only if $ij \in E$.

In this sense, every distribution for a random binary vector in $\{0,1\}^{\binom{N}{2}}$ gives a distribution of a random graph with N nodes.

e.g. $(p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) = (\frac{1}{2}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14})$ gives a distribution over 3-node graphs.

Every family of distributions over $\{0,1\}^{\binom{N}{2}}$ gives a statistical model for random graphs with N nodes.

Replacing with $\{0,1\}^{N(N-1)}$ gives models for directed graphs.

Erdős–Rényi model as an example

Recall: Every family of distributions over $\{0,1\}^{\binom{N}{2}}$ gives a statistical model for random graphs with N nodes.

Erdős-Rényi model: for $\boldsymbol{y} = (y_{ij}) \in \{0,1\}^{\binom{N}{2}}$ consider distribution

$$p(\boldsymbol{y}) \;=\; \prod_{i<j} (1-p)^{1-y_{ij}} p^{y_{ij}}.$$

Denote $s = \sum_{i<j} y_{ij}$ (the number of edges) then

$$p(\boldsymbol{y}) \;=\; (1-p)^{\binom{N}{2}-s} p^{s} \;=\; (1-p)^{\binom{N}{2}} \left(\frac{p}{1-p}\right)^{s}.$$

Quick recall: exponential families

Let $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^n$, $T : \mathbb{R}^n \to \mathbb{R}^d$, $\theta \in \mathbb{R}^d$.

Definition (Eponential family)

A family of probability distributions on $\mathcal{X}$ is an *exponential family* if the pms/densities take the form

$$p_\theta(\boldsymbol{x}) = h(\boldsymbol{x}) \exp\left(\theta^T T(\boldsymbol{x}) - \psi(\theta)\right).$$

- $T(\boldsymbol{x})$ = sufficient statistics (counts of edges, triangles, ...).
- θ = natural parameter.
- $\psi(\theta)$ = log-partition function (ensures normalization).

Bernoulli, binomial, Poisson, Ising models, multivariate Gaussian, and many other popular statistical models are exponential families.

Exponential Random Graph Models

Definition (Exponential Random Graph Models (ERGMs):)

$$\mathbb{P}(G = g) \propto \exp\{\theta_1 \cdot \#\text{edges}(g) + \theta_2 \cdot \#\text{triangles}(g) + \cdots\}.$$

- The parameters: θ_1 tunes density, θ_2 tunes clustering, etc.

Erdős-Rényi model is a special case of ERGM:

$$\mathbb{P}(G = g) \; = \; (1-p)^{\binom{N}{2}} \left(\frac{p}{1-p}\right)^{s} \propto \exp(\theta \cdot s),$$

where $\theta = \log\left(\frac{p}{1-p}\right)$

Example: the p_2 model for undirected networks

Extension of Erdős–Rényi that introduces *node-specific propensities* to form ties.

Model: All edges are independent with

$$\Pr(Y_{ij} = 1 \mid \alpha_i, \alpha_j) = \frac{\exp(\mu + \alpha_i + \alpha_j)}{1 + \exp(\mu + \alpha_i + \alpha_j)}.$$

Interpretation:

- μ — overall network density.
- α_i — sociability of node i (a random effect).

Example: the p_2 model for undirected networks

Extension of Erdős–Rényi that introduces *node-specific propensities* to form ties.

Model: All edges are independent with

$$\Pr(Y_{ij} = 1 \mid \alpha_i, \alpha_j) = \frac{\exp(\mu + \alpha_i + \alpha_j)}{1 + \exp(\mu + \alpha_i + \alpha_j)}.$$

Interpretation:

- μ — overall network density.
- α_i — sociability of node i (a random effect).

Remarks:

- Reduces to Erdős–Rényi when $\alpha_i \equiv 0$.
- Adds degree heterogeneity while preserving tractability.
- Foundation for later hierarchical and latent-space models.

Why is it an ERGM? What are the sufficient statistics?

Latent space random graphs

Definition

Each node i has a position z_i in a latent feature space (e.g. $\mathbb{R}^d$). The probability of an edge depends on distance:

$$\mathbb{P}(i \sim j) = f\left(\|z_i - z_j\|\right), \quad f \text{ decreasing.}$$

As an example, imagine an interaction network in a big company. Apart from the usual topology that follows the company's structure, unexpected links may occur (e.g. among smokers etc).

Example: latent space in economic networks

- Think of banks, firms, or households as nodes.
- Each actor has a position in a **latent space**:
  - Geography (local vs. international).
  - Sector (energy, tech, manufacturing).
  - Risk profile or credit rating.
- Links (e.g. loans, partnerships, trade) are more likely between **nearby** nodes in this space.
- But a few "long-distance" links (large international banks, global supply chains) can connect distant clusters and **dramatically reduce path lengths**.

Takeaway

Latent space models explain why real networks show both **clustering** (local ties) and **small-world shortcuts** (rare global ties).

Dynamic random graph models

From Static to Growing Models

All previous models assumed a fixed number of nodes and edges.

But real networks *grow* over time: new users, new webpages, new firms.

Let's study a simple dynamic rule that explains why hubs emerge:
preferential attachment.

Recursive growth: preferential attachment

Preferential attachment: New node attaches to existing node v with probability proportional to $\deg(v)$.

- "Rich get richer" $\rightarrow$ hubs emerge.

Recursive growth: preferential attachment

Preferential attachment: New node attaches to existing node v with probability proportional to $\deg(v)$.

- "Rich get richer" $\rightarrow$ hubs emerge.

Result: degree distribution follows a *power law*.

- Few very large hubs.
- Many low-degree nodes.
- Matches data: web, citation networks, finance.

Preferential Attachment: Formal Definition

We construct a growing sequence of graphs $G_m, G_{m+1}, G_{m+2}, \ldots$.

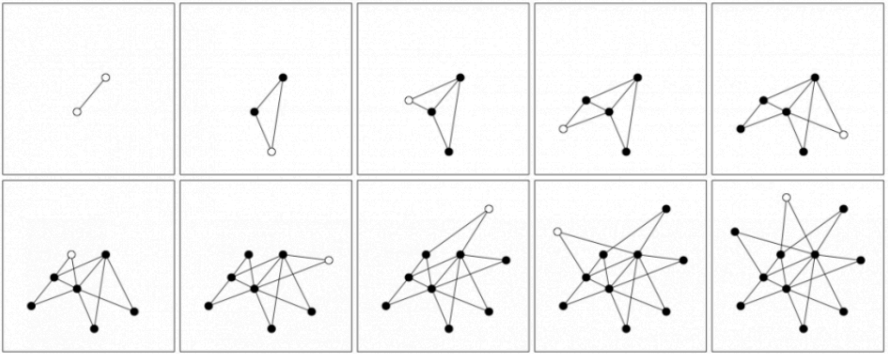
1. **Initialization:** Start from a complete graph G_m on m nodes (so each node initially has degree $m-1$).
2. **Growth rule:** For each step $t = m+1, m+2, \ldots$:
   - Add a new node v_t.
   - Connect v_t to exactly m existing nodes.
   - Each existing node u is chosen with probability

   $$\mathbb{P}(v_t \to u) \;=\; \frac{\deg(u, t-1)}{\sum_w \deg(w, t-1)}.$$

   Thus, high-degree nodes are more likely to receive new links.

This process defines the **Barabási–Albert (BA) model**.

Evolution of the Barabási-Albert model



Expected degree growth in the BA model

At time $t \geq m$, the network has $L_t = \binom{m}{2} + m(t - m)$ edges. Up to constants depending only on m, we may write $2L_t \approx 2mt$.

Fix a node u with current degree $d_t = \deg(u, t)$. When a new node arrives at step $t+1$, it creates m new edges, each connecting to an existing node v with probability proportional to its degree:

$$\mathbb{P}(\text{edge connects to } u) \approx \frac{d_t}{2mt}.$$

Expected degree growth in the BA model

At time $t \geq m$, the network has $L_t = \binom{m}{2} + m(t-m)$ edges. Up to constants depending only on m, we may write $2L_t \approx 2mt$.

Fix a node u with current degree $d_t = \deg(u, t)$. When a new node arrives at step $t{+}1$, it creates m new edges, each connecting to an existing node v with probability proportional to its degree:

$$\mathbb{P}(\text{edge connects to } u) \approx \frac{d_t}{2mt}.$$

Expected increment:

$$\mathbb{E}[d_{t+1} - d_t \mid d_t] \approx m \cdot \frac{d_t}{2mt} = \frac{d_t}{2t}.$$

Expected degree growth in the BA model

At time $t \geq m$, the network has $L_t = \binom{m}{2} + m(t - m)$ edges. Up to constants depending only on m, we may write $2L_t \approx 2mt$.

Fix a node u with current degree $d_t = \deg(u, t)$. When a new node arrives at step $t+1$, it creates m new edges, each connecting to an existing node v with probability proportional to its degree:

$$\mathbb{P}(\text{edge connects to } u) \approx \frac{d_t}{2mt}.$$

Expected increment:

$$\mathbb{E}[d_{t+1} - d_t \mid d_t] \approx m \cdot \frac{d_t}{2mt} = \frac{d_t}{2t}.$$

Recursion for the expected degree:

$$\mathbb{E}[d_{t+1}] \approx \mathbb{E}[d_t]\Big(1 + \frac{1}{2t}\Big), \qquad t \geq m.$$

Expected degree growth in the BA model

At time $t \geq m$, the network has $L_t = \binom{m}{2} + m(t - m)$ edges. Up to constants depending only on m, we may write $2L_t \;\approx\; 2mt$.

Fix a node u with current degree $d_t = \deg(u, t)$. When a new node arrives at step $t{+}1$, it creates m new edges, each connecting to an existing node v with probability proportional to its degree:

$$\mathbb{P}(\text{edge connects to } u) \;\approx\; \frac{d_t}{2mt}.$$

Expected increment:

$$\mathbb{E}[d_{t+1} - d_t \mid d_t] \;\approx\; m \cdot \frac{d_t}{2mt} \;=\; \frac{d_t}{2t}.$$

Recursion for the expected degree:

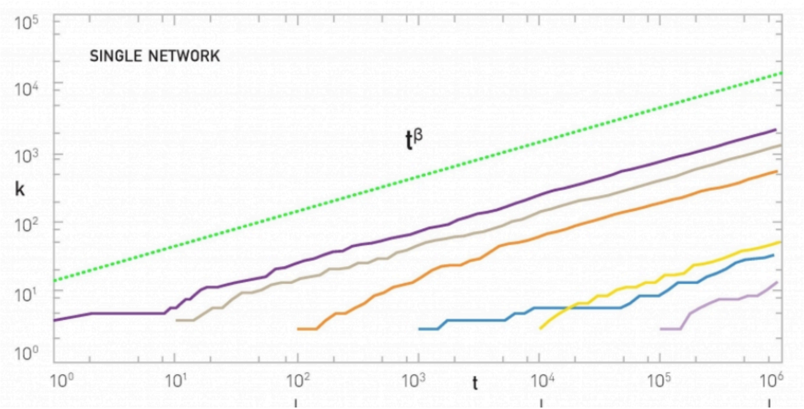
$$\mathbb{E}[d_{t+1}] \;\approx\; \mathbb{E}[d_t]\Bigl(1 + \frac{1}{2t}\Bigr), \qquad t \geq m.$$

Solution: If t_u is the time when u appears, we can show

$$\boxed{\mathbb{E}[d_t] \;\approx\; m\Bigl(\frac{t}{t_u}\Bigr)^{1/2}.}$$

Evolution of the degree

In the log – log scale:



Quick comparison (who becomes a hub?)

Two nodes joined at $t_u = 10$ and $t_v = 100$. After $t = 1000$ with $m = 3$:

$$\frac{\deg(u, 1000)}{\deg(v, 1000)} = \sqrt{\frac{1000/10}{1000/100}} = \sqrt{10} \approx 3.16,$$

$$\deg(u, 1000) = 3\sqrt{100} = 30, \qquad \deg(v, 1000) = 3\sqrt{10} \approx 9.48.$$

Earlier arrival systematically advantages degree.

Heuristic derivation of the degree distribution

From the recursion we found:

$$\mathbb{E}[\deg(u,t)] \;\approx\; m\left(\frac{t}{t_u}\right)^{1/2}.$$

Hence, older nodes (small t_u) have larger expected degree.

To find the degree distribution at time t, note that

$$\deg(u,t) \approx m\left(\frac{t}{t_u}\right)^{1/2} \quad\Longleftrightarrow\quad t_u \approx t\,\frac{m^2}{\deg(u,t)^2}.$$

Heuristic derivation of the degree distribution

From the recursion we found:

$$\mathbb{E}[\deg(u,t)] \;\approx\; m\left(\frac{t}{t_u}\right)^{1/2}.$$

Hence, older nodes (small t_u) have larger expected degree.

To find the degree distribution at time t, note that

$$\deg(u,t) \approx m\left(\frac{t}{t_u}\right)^{1/2} \quad\Longleftrightarrow\quad t_u \approx t\,\frac{m^2}{\deg(u,t)^2}.$$

Since arrival times t_u are roughly *uniform* on $\{1,2,\ldots,t\}$, we can compute

$$\mathbb{P}(\deg(u,t) \geq k) \;=\; \mathbb{P}\left(t_u \leq t\,\frac{m^2}{k^2}\right) \;\approx\; \frac{m^2}{k^2}.$$

Discrete exact formula and asymptotic tail

The prob. that a node has degree $\geq k$ decreases quadratically in k:

$$\Pr\{\deg \geq k\} \; \propto \; k^{-2} \quad \Longrightarrow \quad p_k = \Pr\{\deg = k\} \; \propto \; k^{-3}.$$

Hence: the Barabási–Albert model produces a *power-law* degree distribution with

$$\boxed{\gamma = 3.}$$

Discrete exact formula and asymptotic tail

The prob. that a node has degree $\geq k$ decreases quadratically in k:

$$\Pr\{\deg \geq k\} \propto k^{-2} \implies p_k = \Pr\{\deg = k\} \propto k^{-3}.$$

Hence: the Barabási–Albert model produces a *power-law* degree distribution with

$$\boxed{\gamma = 3.}$$

Takeaways

- The tail exponent $\gamma = 3$ is universal for the BA model (independent of m).
- The discrete formula matches simulations closely, including small-k corrections.
- This result follows directly from preferential attachment, not from a continuum limit.

Take-home Messages

- Growth + preferential attachment $\Rightarrow$ power-law degrees with $\gamma = 3$.
- BA networks are short and heterogeneous, but almost unclustered.
- Real systems combine multiple effects: growth, fitness, and local closure.
- These ideas form the foundation for modern network science.

Final Exercise (preferential attachment probabilities)

Given the degree multiset $\{1,1,1,1,2,3,3,4,4,5,5,8\}$, let a new node add *one* link using BA attachment $\Pr\{u\} = \deg(u)/(2L)$.

a) Probability it attaches to the highest–degree node:

$$\Pr\{\text{choose } k=8\} \;=\; \frac{8}{\sum \deg} \;=\; \frac{8}{1+1+1+1+2+3+3+4+4+5}$$

b) Probability it attaches to a node of degree 1: there are four such nodes, each with probability $1/38$: $4 \times \frac{1}{38} = \frac{4}{38}$.

(Here $2L = \sum \deg = 38$.)

Final Exercise

Given a Network with the following degree sequence:

$$\{1, 1, 1, 1, 2, 3, 3, 4, 4, 5, 5, 8\}$$

If we consider the preferential attachment of the BA model:

a) What is the probability that a new node attaches to the bigger node (higher degree).

b) What is the probability that a new node attaches to the smallest node(s) (lower degree).

Take into account that it will attach to the network with one link only.