Social and Graph Data Management: Graph Formation Models

Pablo Mollá Chárlez

Contents

1 Introduction to Graph Formation Models

Building on our understanding of fundamental graph concepts such as nodes, edges, degrees, paths, and distances, we now explore graph formation models that simulate the structure of real-world networks. These models, including random networks like the Erdős–Rényi model, aim to replicate the complexity and sparsity observed in social networks using simple, parameter-driven frameworks. By studying these formation models, we can better analyze the inherent properties of social graphs, understand how connections emerge, and apply this knowledge to tasks like influence measurement and link prediction.

2 Random Networks

Random networks are foundational models in graph theory used to understand the properties and behaviors of real-world networks by simplifying their complexity. While **real networks are typically sparse** (having relatively few edges compared to the number of possible connections), sparsity alone does not capture all their intricate characteristics. The **primary objective of graph models** like random networks is to replicate the complexity of real networks using minimal parameters, thereby providing insights into their structural properties.

2.1 Random Network Parameters

Random network models are defined using two key parameters:

- **Number of nodes N**: The total number of vertices in the graph.
- **Edge probability p**: The likelihood that any given pair of nodes is connected by an edge.

These parameters allow the generation of graphs that can mimic various real-world network structures under different conditions.

2.2 Random Network's Algorithm

Random networks were first introduced and extensively studied by **mathematicians Paul Erdős and Alfréd Rényi**, leading to the **Erdős–Rényi (ER) model**.

The **algorithm steps** to generate such networks is straightforward:

1. **Initialize**: Start with N disconnected nodes.
2. **Edge Formation**: For each possible pair of nodes, add an edge between them with probability p.
3. **Completion**: Repeat the edge formation step for all $\binom{N}{2} = \frac{N(N-1)}{2}$ possible node pairs.

There are two primary variants of the **ER model**:

- $G(N, p)$ **Model**: A graph with N nodes where each possible edge is included independently with probability p.
- $G(N, L)$ **Model**: A graph with N nodes and exactly L edges, where the L edges are chosen uniformly at random from all possible pairs.

2.3 Properties

- **Expected Number of Links**:

$$\langle L \rangle = p \cdot \frac{N(N-1)}{2}$$

  This represents the average number of edges expected in the graph.

- **Average Degree**:

$$\langle k \rangle = p(N-1)$$

  The average degree indicates the average number of connections per node.

2.4 Degree Distribution in Random Networks: Binomial Distribution

The degree distribution p_k describes the probability that a randomly selected node has degree k. In the **ER model**, this distribution is governed by the binomial distribution:

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Where the formula is composed of:

- $\binom{N-1}{k}$ is the number of ways to choose k edges from $N-1$ possible connections.
- p^k is the probability of having k edges.
- $(1-p)^{N-1-k}$ is the probability of not having the remaining $N-1-k$ edges.

2.5 Degree Distribution in Very Sparse Random Networks: Poisson Distribution

For very sparse networks like random networks where the average degree $\langle k \rangle$ is much smaller than N ($\langle k \rangle \ll N$), which equivalent to stating that N is large and p small, then the binomial distribution can be approximated by the Poisson distribution:

$$p_k \approx \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!}$$

This simplification arises because, in sparse conditions, the probability of multiple edges between the same pair of nodes becomes negligible, making the Poisson distribution a good fit for modeling the degree distribution.

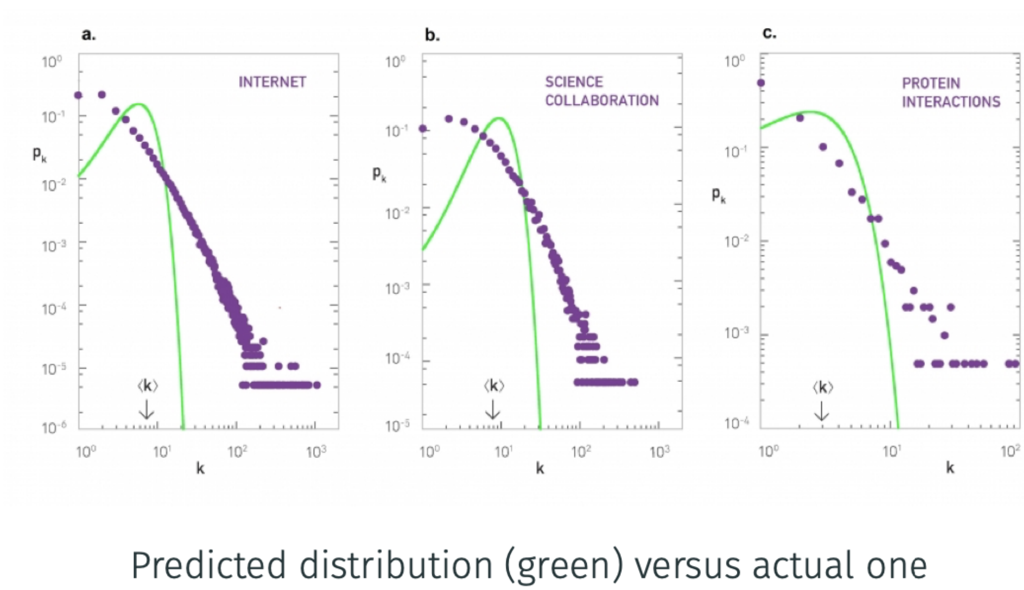

Figure 1: Real Networks don't follow Poisson Distributions

2.6 Regimes of Random Networks

The behavior of random networks varies significantly depending on the edge probability p relative to $\frac{1}{N}$ (where N is the number of nodes) or the average degree $\langle k \rangle$ relative to 1. These variations define distinct regimes within random networks:

1. $\boxed{\textbf{Subcritical Regime: } p < \frac{1}{N} \text{ or } \langle k \rangle < 1}$

   The network consists of numerous small, disconnected components. The size of the largest connected component N_G grows logarithmically with N, i.e., $N_G \in O(\ln N)$. The structure contains clusters which are predominantly tree-like, lacking cycles.

2. $\boxed{\textbf{Critical Point: } p = \frac{1}{N} \textbf{ or } \langle k \rangle = 1}$

   This is the threshold where a phase transition occurs. A large component emerges with $N_G \sim N^{2/3}$, alongside many smaller tree-like clusters. The large cluster may contain loops, while smaller clusters remain mostly trees.

3. $\boxed{\textbf{Supercritical Regime: } p > \frac{1}{N} \textbf{ or } \langle k \rangle > 1}$

   A single dominant connected component, known as the giant component, forms and $N_G \sim (p - p_c)N$, where $p_c \approx \frac{1}{N}$ is the critical probability. The giant component contains **numerous loops**, whereas **other smaller clusters are typically tree structures**.

4. $\boxed{\textbf{Connected Regime: } p > \frac{\ln(N)}{N} \textbf{ or } \langle k \rangle \geq \ ln(N)}$

   The network becomes fully connected. N_G approaches N, meaning almost all nodes are part of a single connected component. The network is richly interconnected with **multiple loops**.

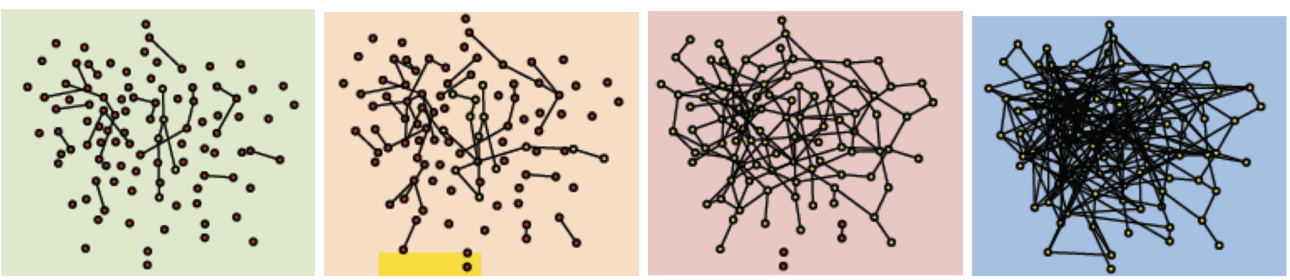

Figure 2: Sucritical, Critical, Supercritial and Connected

Empirical observations show that real-world networks often operate in the supercritical regime, characterized by a large, dominant connected component. However, it can be observed that unlike random networks in the supercritical regime, which predict the existence of multiple connected components, real networks typically maintain a single, extensive connected component, aligning with their observed supercritical nature.

2.7 Clustering Coefficient in Random Networks

Reminder: The clustering coefficient measures the degree to which nodes in a graph tend to cluster together. In random networks, it is calculated as follows:

- **Expected Number of Links Among Neighbors for node i with k_i neighbours ($\langle L_i \rangle$)**:

$$\langle L_i \rangle = p \cdot \frac{k_i(k_i - 1)}{2}$$

  Here, k_i is the degree of node i, and p is the probability of edge formation.

- **Clustering Coefficient (C_i)**:

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p$$

  This indicates that, in random networks, the clustering coefficient is solely dependent on p.

$\boxed{\textbf{Implications}}$

1. **Network Size Effect**: For a constant average degree $\langle k \rangle$, as the network size N increases, the clustering coefficient C_i decreases because $p = \frac{\langle k \rangle}{N-1}$.
2. **Degree Independence**: The clustering coefficient C_i is independent of the node degree k_i, meaning all nodes have the same clustering tendency regardless of their connectivity.

$\boxed{\textbf{Limitation}}$ Real-world networks often exhibit higher clustering coefficients than predicted by random network models. This discrepancy arises because **random networks fail to capture the inherent clustering and community structures** present in real social, biological, and technological networks.

3 Six Degrees of Separation and Small-World Phenomenon

Six Degrees of Separation refers to the small-world phenomenon, which posits that any two individuals in a large network can be connected through a short chain of neighbours, typically around six steps. In graph terms, this means that the distance between any two nodes in a social network is **surprisingly short**, facilitating quick information flow and connectivity.

3.1 Average and Maximum Distance in Networks

In a graph with an **average degree** $\langle k \rangle$, each node typically connects to $\langle k \rangle^d$ nodes at a distance d. To encompass the entire network of N nodes, the maximum distance $d_{\max}$ can be approximated by:

$$d_{\max} \approx \frac{\ln N}{\ln \langle k \rangle}$$

This relationship shows that as the number of nodes N increases, the maximum distance grows logarithmically, rather than linearly, highlighting the efficiency of connectivity in large networks.

3.2 Small Worlds in Random Networks

For most networks, especially random networks, the average distance $\langle d \rangle$ between nodes follows a similar logarithmic relationship:

$$\langle d \rangle \approx \frac{\ln N}{\ln \langle k \rangle} = \frac{\ln N}{\ln(p(N-1))}$$

This indicates that even in very large networks, the typical path length between nodes remains short. The term $\frac{1}{\ln \langle k \rangle}$ implies that denser networks (with higher $\langle k \rangle$) have even shorter average distances. Importantly, this estimation holds true not only for theoretical random networks but also for many **real-world networks**, albeit with slight adjustments to account for additional structural nuances.

4 Real Networks

4.1 Hubs and Degree Distributions in Real Networks

4.1.1 Hubs in Real Networks

In many real-world networks, certain nodes, known as hubs, possess an exceptionally high number of connections compared to typical nodes. These hubs play a crucial role in the network's structure and functionality by:

1. **Facilitating Connectivity**: Hubs act as major connectors, linking disparate parts of the network and enhancing overall connectivity.

2. **Influence and Robustness**: They often hold significant influence, controlling information flow and network dynamics. Additionally, the presence of hubs can make networks more robust against random failures but more vulnerable to targeted attacks on these key nodes.

In contrast, random network models like the **Erdős–Rényi (ER) model** typically lack such highly connected nodes. The **probability of any node becoming a hub is exceedingly low** because edges are distributed uniformly and independently across all node pairs. This uniformity results in a degree distribution where **most nodes have degrees close to the average**, preventing the emergence of prominent hubs.

4.1.2 Degree Distributions in Real Networks: Power-Law Distribution

Reminder: Degree distribution describes the probability p_k that a randomly selected node has degree k. In real networks, these distributions often exhibit distinct patterns:

4.1.2.1 Power-Law Distribution & Normalization with Riemann Zeta Function

Real networks frequently follow a power-law distribution:

$$p_k \sim k^{-\gamma}$$

where γ is a positive constant typically between 2 and 3. Power-law distributions have heavy tails, meaning the probability is large enough to be considered important (or worthy of attention) of finding nodes with very high degrees (hubs). Networks whose degree distributions follow a power law are termed scale-free networks because the same power-law behavior persists regardless of the network's size.

To ensure the probabilities sum to one, the degree distribution is normalized using the **Riemann zeta function** $\zeta(\gamma)$:

$$p_k = \frac{C}{k^\gamma}, \quad \text{where} \quad C = \frac{1}{\zeta(\gamma)} = \left(\sum_{k=1}^{\infty} k^{-\gamma}\right)^{-1}$$

This normalization ensures that $\sum_{k=1}^{\infty} p_k = 1$, maintaining the validity of the probability distribution.

4.1.2.2 Poisson vs. Power-Law Degree Distributions

As we mentioned earlier, Poisson distribution is typically associated with very sparse random networks, where most nodes have **degrees really small compared to the number of nodes**, and **hubs are rare**. On the other hand, Power-Law distribution is common in real-world networks (e.g., social, biological, technological), where a **few nodes have very high degrees (hubs)**, and **many have low degrees**. Therefore, is interesting to comparing Poisson and power-law **degree distributions** which highlights fundamental differences in network structures:

- **Poisson Distribution (Very Sparse Random Networks)** The distribution has the form:

$$p_k = \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!}$$

  where $\langle k \rangle$ is the average degree. The distribution is characterized by a sharp peaked around the average degree $\langle k \rangle$ and probabilities decrease exponentially for large k, making the presence of high-degree nodes rare.

- **Power-Law Distribution (Real Networks)** The distribution has the form:

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}, \ \zeta(\gamma) = \sum_{k=1}^{\infty} k^{-\gamma}$$

  The main features are a high probabilities for large k compared to Poisson, allowing for the existence of hubs and it can model a wide range of real-world scenarios where some nodes are disproportionately connected.

Comparative Insights

- **Small k (Low Degrees):** $\boxed{\textbf{Power-Law} > \textbf{Poisson}}$

  This implies that, for small k low degrees, a Power-Law distribution assigns a higher probability to nodes having these degrees compared to a Poisson distribution. This results in more nodes with low degrees in networks modeled by a Power-Law, leading to a broader and more heterogeneous connectivity base than what is seen with Poisson-distributed random networks.

- **Around $\langle k \rangle$ (Mean Degree):** $\boxed{\textbf{Poisson} > \textbf{Power-Law}}$

  For degrees around the average degree $\langle k \rangle$, a Poisson distribution assigns a higher probability to nodes having these degrees compared to a Power-Law distribution. This results in a significant concentration of nodes around the mean degree in PoissonPoisson-distributed random networks, leading to more uniform connectivity. In contrast, Power-Law networks have fewer nodes near the average degree, contributing to a less uniform and more variable connectivity structure.

- **Large k (High Degrees):** $\boxed{\textbf{Power-Law} > \textbf{Poisson}}$

  For large k (high degrees), a Power-Law distribution assigns a higher probability to nodes having these degrees compared to a Poisson distribution. This leads to the presence of hubs in Power-Law modeled networks. These hubs are critical for the network's robustness and efficiency, enabling rapid information flow and connectivity. In contrast, Poisson-distributed random networks have a much lower likelihood of high-degree nodes, resulting in fewer or no significant hubs. Which wouldn't represent correctly the data observed.

Example 1 **Professional Networking Platform**

Imagine the following **scenario**. Suppose you are tasked with **modeling the structure of a professional networking platform** to understand how information spreads and to assess its resilience to node failures. Upon examining the network data from the professional networking platform, you notice that a few users have exceptionally high numbers of connections (hubs), while most users have relatively few connections. In such case, recognizing that a Power-Law distribution better accounts for the presence of these hubs, you decide to model the network as a Power-Lawscale-free network rather than using a Poisson distribution typical of random networks.

Example 2 **Road vs. Flight Networks in USA**

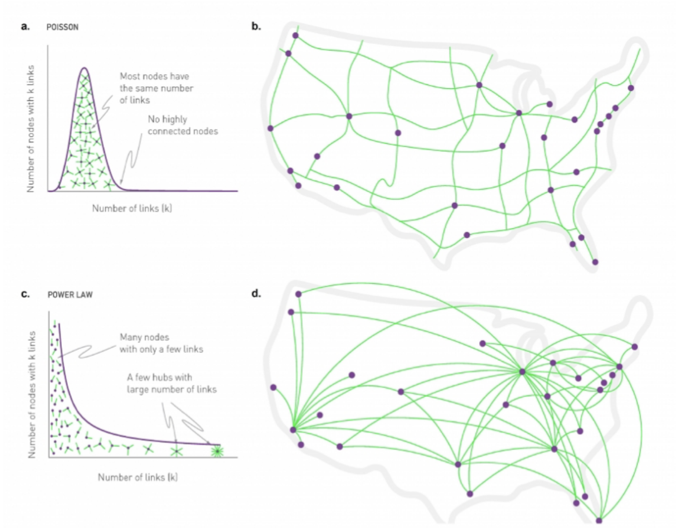

Figure 3: Road vs. Flight Networks

The previous image illustrates the difference between Poisson and Power-Law degree distributions by comparing road networks **(b)** and flight networks **(d)** across the United States.

- **Road Network (b)**: This network represents a Poisson distribution (shown in graph a). Most nodes (cities) have a similar number of connections (roads), with few or no high-degree nodes (hubs). This structure results in a sparse and uniform network where each location is directly connected to only a few others.
- **Flight Network (d)**: This network represents a Power-Law distribution (shown in graph **c**), where there are hubs (major airports with many flights) connected to numerous smaller nodes (regional airports). This structure creates a network with a few highly connected nodes and many with only a few connections, enabling efficient, long-distance travel through hubs.

In essence, this comparison demonstrates how Power-Law networks, like flight routes, are more efficient for long-range connectivity and are more similar to real-world networks, whereas Poisson-distributed road networks reflect more localized, uniform connections without major hubs.

4.2 Scale-Free Networks

While **random networks** modeled by Poisson or binomial **degree distributions** provide a foundational understanding of network structures, they fall short in capturing the complexity observed in real-world networks. The presence of **hubs** and **heterogeneous connectivity** in actual social, biological, and technological networks necessitates more sophisticated models. This leads us to the concept of scale-free networks, which offer a more accurate representation by incorporating Power-Law degree distributions.

4.3 Why Scale-Free?

The concept of scale-free networks originates from the **study of phase transitions in physics**, where Power-Law distributions play a significant role. To comprehend scale-free behavior, we analyze the moments of a distribution, which provide insights into its shape and variability. The general formula to determine the moments of the Power-Law distributions is:

$$\langle k^n \rangle = \sum_k k^n \cdot p_k$$

The three first moments provide interesting information:

1. **Mean Degree** ($\langle k \rangle$):
$$\langle k \rangle = \sum_k k \cdot p_k$$
Represents the average number of connections per node.
2. **Variance** (σ_k^2):
$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2$$
Measures the spread of the degree distribution.
3. **Skewness** ($\langle k^3 \rangle$):
$$\langle k^3 \rangle = \sum_k k^3 \cdot p_k$$
Indicates the asymmetry of the degree distribution.

4.4 Major Differences Between Random and Scale-Free Networks

In random networks, node degrees are concentrated around the mean degree $\langle k \rangle$. The distribution decays rapidly, making high-degree nodes very rare. The second moment $\langle k^2 \rangle$ is calculated as:

$$\langle k^2 \rangle = \sum_{k=0}^{\infty} k^2 \cdot p_k = \sum_{k=0}^{\infty} k^2 \cdot \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!} < \infty$$

Since random networks have a bounded, exponential degree distribution, $\langle k^2 \rangle$ is finite and does not grow as the network size increases. **Most nodes have degrees close to the mean, and the degree variance remains small**.

In scale-free networks (with a Power-Law distribution where $p_k \sim k^{-\gamma}$ and $\gamma < 3$), the degree distribution has a "heavy tail." This means there is a non-negligible probability of finding nodes with very high degrees (hubs). For these networks:

$$\langle k^2 \rangle = \sum_{k=1}^{\infty} k^2 \cdot \frac{C}{k^\gamma} = C \sum_{k=1}^{\infty} k^{2-\gamma} = \frac{\sum_{k=1}^{\infty} k^{2-\gamma}}{\zeta(\gamma)} = \frac{\sum_{k=1}^{\infty} k^{2-\gamma}}{\sum_{k=1}^{\infty} k^{-\gamma}}$$

When $\gamma < 3$, the exponent from the numerator $2 - \gamma$ is negative (because $\gamma \in [2, 3]$), which causes the sum to diverge as k grows. On the other hand, the **Riemann zeta function** converges for $\gamma > 1$ and diverges for $\gamma \leq 1$. Therefore, for $\gamma \in [2, 3]$, the second moment $\langle k^2 \rangle$ becomes infinite in the theoretical limit of an infinitely large network. This divergence indicates that there are many nodes with much higher degrees than the average, **allowing the existence of hubs**. The presence of these high-degree nodes creates a network structure without a clear "scale," in contrast to random networks, where node degrees are more uniformly distributed around the mean.

4.5 Scale-Free Properties

Scale-free networks exhibit unique properties where the average distance $\langle d \rangle$ between nodes depends on the network size N and the Power-Law exponent γ:

1. **Anomalous Regime** ($\gamma = 2$): The largest hub has a degree that grows linearly with N and the average distance:

   $$\langle d \rangle \sim \text{constant}$$

   The structure is a **Hub-and-spoke model** with minimal average distance.

2. **Ultra-Small World** ($2 < \gamma < 3$): The average distance if:

   $$\langle d \rangle \sim \ln \ln N$$

   The growth of the distance is slower than random networks, maintaining extremely short paths even as N increases. Most real-world networks fall into this category.

3. **Critical Point** ($\gamma = 3$): The second moment $\langle k^2 \rangle$ ceases to diverge and the average distance is:

   $$\langle d \rangle \sim \frac{\ln N}{\ln \ln N}$$

   Marks the boundary between scale-free and random network behaviors.

4. **Small World** ($\gamma > 3$): The average distance is:

   $$\langle d \rangle \sim \ln N$$

   It mirrors that of random networks, with average distances growing logarithmically with N.

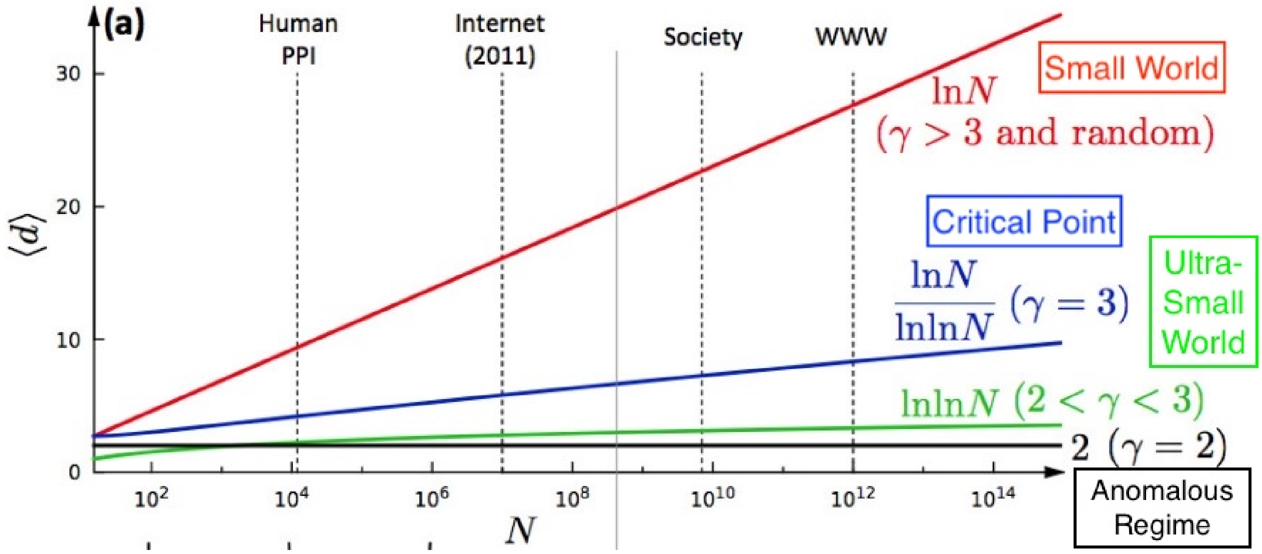

Figure 4: Scale-Free Properties

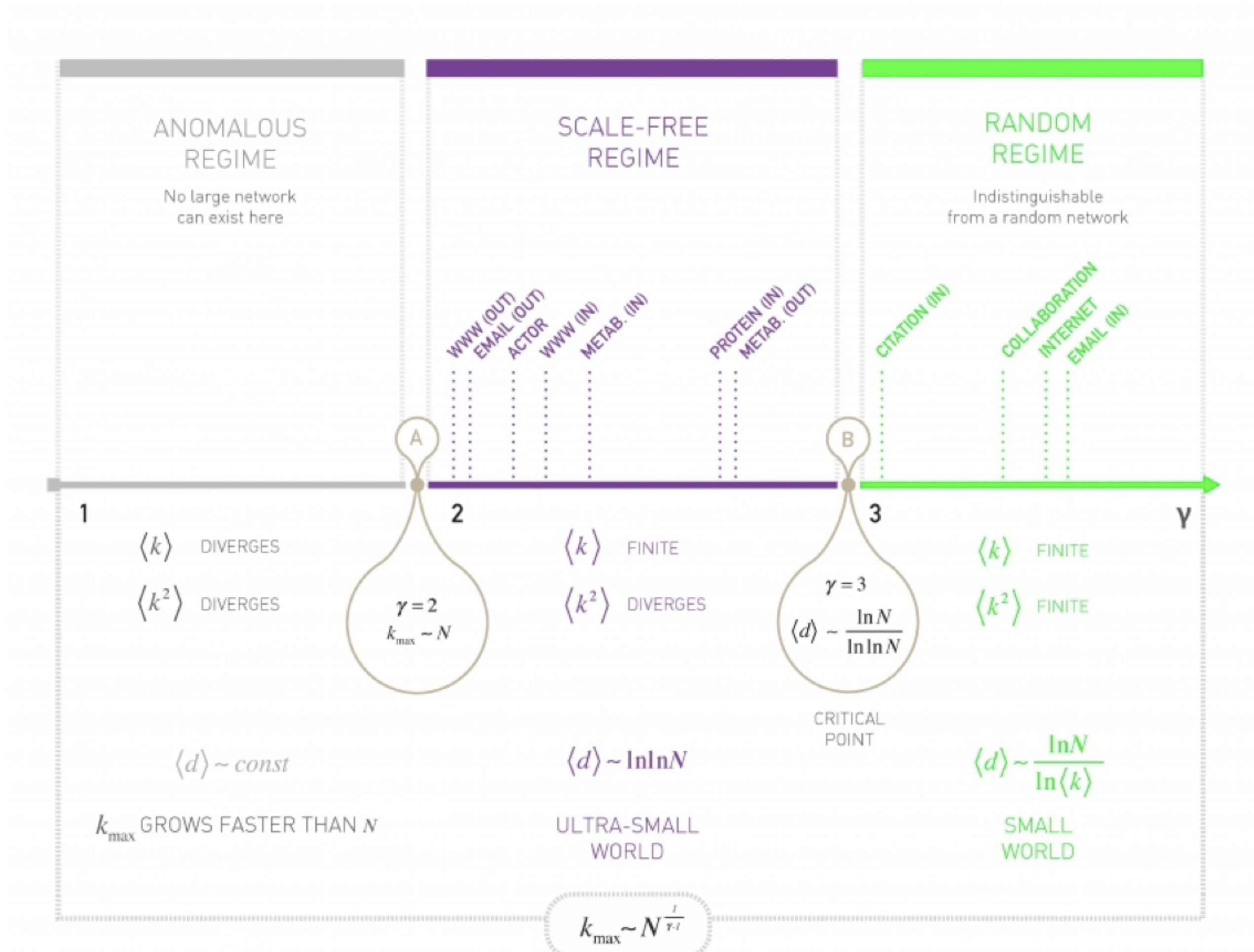

Figure 5: Role of the Degree Exponent

5 Preferential Attachment Model

In our previous discussion, we saw that random networks assume a **fixed number of nodes N** with **each pair connected independently with a probability** p. This approach, while useful for studying certain network properties, does not accurately reflect real-world networks. Real networks, such as social networks or the web, do not have a fixed number of nodes; they grow continuously over time. Moreover, new nodes in real networks often prefer to connect to well-connected nodes, a phenomenon known as preferential attachment. For example, when people join Twitter, they are more likely to follow popular accounts, or when new pages are created on the web, they tend to link to already well-known websites. Therefore, we need to capture those 2 main characteristics in a model.

5.1 Barabási–Albert Model

The **Barabási–Albert (BA) model** is a widely used model that incorporates both growth and preferential attachment to generate scale-free networks, which are networks with a Power-Law degree distribution. The **BA model** uses a **single parameter, m**, to determine how new nodes are added and connected. Let's see how the Barabási–Albert model works:

1. **Starting Configuration**: Begin with an initial network of m_0 nodes connected in some arbitrary way.

2. **Growth**: At each step, a new node is added to the network. This node will form m links to existing nodes.

3. **Preferential Attachment**: Each of the m links from the **new node connects to an existing node** i **with a probability proportional to the degree of** i. The probability $P(i)$ that the new node connects to node i is given by:

$$P(i) = \frac{k_i}{\sum_j k_j}$$

   where k_i is the degree of node i, and $\sum_j k_j$ is the sum of degrees of all existing nodes. This rule ensures that nodes with higher degrees (more connections) are more likely to attract new links, leading to the **emergence of hubs**.

The combination of growth and preferential attachment in the Barabási–Albert model naturally leads to a Power-Law degree distribution with **an exponent γ typically around 3**.

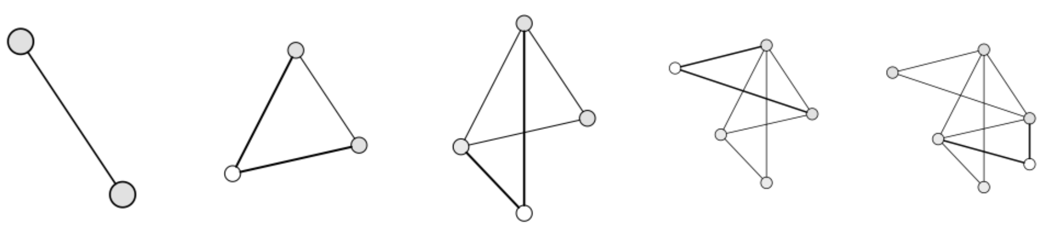

Figure 6: 5-Steps Barabási-Albert Model ($m_0 = 2, m = 2$)

5.2 Degree Coefficient and Distribution

In the **BA model**, new nodes prefer to connect to existing nodes with higher degrees. This tendency is captured by the probability $P(i)$ that a new node will connect to an existing node i, which is proportional to the degree k_i of that node:

$$P(i) = \frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

where $\sum_{j=1}^{N-1} k_j$ is the total degree of all existing nodes in the network at that time.

We know that in the **BA model**, each new node that joins the network creates m links. Given preferential attachment, the probability that any one of these new links connects to node i is $P(i)$. Therefore, the rate of change of the degree k_i of node i is:

$$\frac{\partial k_i}{\partial t} = mP(i) = m\ [\frac{k_i}{\sum_{j=0}^{N-1} k_j}]$$

This equation states that the degree of node i grows proportionally to its existing degree, reflecting the "rich-get-richer" effect. The total degree for a given step t of the **BA model** network is given by:

$$\sum_{j=0}^{N-1} k_j = m(2t-1)$$

If you don't believe this result, just check the previous example:

1. $\boxed{\textbf{t=1}}$ For the creation of the network ($m_0 = N = 2$), the formula is directly:

$$\sum_{j=0}^{N-1} k_j = k_0 + k_1 = 1 + 1 = (2) \cdot (2(1) - 1) \checkmark$$

2. $\boxed{\textbf{t=2}}$ For the next step, we have $N = 3$ nodes:

$$\sum_{j=0}^{N-1} k_j = \sum_{j=0}^{3-1} k_j = k_0 + k_1 + k_2 = 2 + 2 + 2 = 6 = (2) \cdot (2(2) - 1) \checkmark$$

3. $\boxed{\textbf{t=3}}$ Here, we don't actually care who exactly has a given degree, the formula is still satisfied:

$$\sum_{j=0}^{N-1} k_j = \sum_{j=0}^{4-1} k_j = 10 = (2) \cdot (2(3) - 1) \checkmark$$

4. $\boxed{\textbf{t=4}}$ Checking we obtain:

$$\sum_{j=0}^{N-1} k_j = \sum_{j=0}^{5-1} k_j = 14 = (2) \cdot (2(4) - 1) \checkmark$$

5. $\boxed{\textbf{t=5}}$ Finally:

$$\sum_{j=0}^{N-1} k_j = \sum_{j=0}^{6-1} k_j = 18 = (2) \cdot (2(5) - 1) \checkmark$$

After making sure that the previous formula holds, we can therefore notice that:

$$\frac{\partial k_i}{\partial t} = mP(i) = m\ [\frac{k_i}{\sum_{j=0}^{N-1} k_j}] \longleftrightarrow m\ [\frac{k_i}{m(2t-1)}] \approx m\frac{k_i}{m(2t)} = \frac{k_i}{2t} \longleftrightarrow \frac{\partial k_i}{\partial t} = \frac{k_i}{2t}$$

In the last step, we approximate the partial derivative assuming t is large enough that -1 can be ignored for simplicity. Therefore, integrating both sides with respect to t, from the time t_i (when node i was introduced) to some later time t, and correspondingly, the degree from the initial degree $k_i(t_i) = m$ (where m is the initial number of links the new node makes) to $k_i(t)$:

$$\int_{k_i(t_i)=m}^{k_i(t)} \frac{1}{k_i}\, \mathrm{d}k_i = \int_{t_i}^{t} \frac{1}{2t}\, \mathrm{dt} \longleftrightarrow \ln\left(\frac{k_i(t)}{m}\right) = \frac{1}{2}\ln\left(\frac{t}{t_i}\right) \longleftrightarrow e^{\ln\left(\frac{k_i(t)}{m}\right)} = e^{\frac{1}{2}\ln\left(\frac{t}{t_i}\right)}$$

Then, simplyfing we obtain the following result where $\beta = \frac{1}{2}$ (β is the dynamical exponent).

$$\frac{k_i(t)}{m} = \left(\frac{t}{t_i}\right)^{\frac{1}{2}} \longleftrightarrow k_i(t) = m\left(\frac{t}{t_i}\right)^{\frac{1}{2}} = m\left(\frac{t}{t_i}\right)^{\beta}$$

We need to retain that:

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}$$

This time-dependent degree formula leads to a degree distribution:

$$p_k \approx 2m^{\frac{1}{\beta}} k^{-\gamma},\ \gamma = \frac{1}{\beta} + 1 = 3$$

For large k, the distribution $p_k \sim k^{-3}$ results in a scale-free network. The model predicts a stable/stationaly , scale-free structure over time.

5.3 Average Distance & Clustering Coefficient

The average distance is given by:

$$\langle d \rangle \sim \frac{\ln N}{\ln \ln N}$$

The distances grow slower than in random networks, aligning more closely with real network behavior. On the other hand, the clustering coefficient follows:

$$\langle C \rangle \sim \frac{(\ln N)^2}{N}$$

The model predicts higher local clustering than random networks.

5.4 Shortcomings of the Barabási-Albert Model

- The model consistently produces $\gamma = 3$, while real networks often have $2 \leq \gamma \leq 5$.
- It only applies to undirected networks.
- It does not account for link formation between existing nodes or node removal.
- The model does not distinguish between nodes with different attributes or characteristics.

In summary, the Barabási–Albert model successfully captures the scale-free nature and local clustering of real networks but has limitations in flexibility and accuracy for complex, real-world structures.