FGV EMAp

João Pedro Jerônimo

Ciência de Redes

Revisão para A1

Rio de Janeiro

2025

Conteúdo

Grafos

De antemão valhe ressaltar que essa matéria, por mais que seja chamada de **Ciência de Redes**, o termo **rede** se refere a um grafo, não ao tipo específico de grafo que se é visto em **Fluxo em Redes** quando estudamos matemática discreta. Então que já fique esclarecido de antemão que, ao citarmos redes, estamos nos referindo a um grafo no geral, desde que o contrário seja explicitado

Essa sessão será apenas algumas definições que não foram passadas no curso de Matemática Discreta, então conceitos que forem citados sobre grafos e não houver definição nesse resumo, a mesma estará no recap de Matemática Discreta. Aqui segue algumas notações sobre grafos para que não fique confuso:

- $G(V, E)$:= Grafo com conjunto de vértices V e de arestas E (edges)
- $N(v)$:= Vizinhança do vértice v (Neighbourhood)
- $\delta(v)$:= Grau do vértice v
- K_n := Grafo completo com n vértices
- $K_{m,n}$:= Grafo completo bipartido com m vértices no primeiro conjunto e n vértices no segundo
- $\mathrm{X}(G)$:= Número cromático de G
- $\mathrm{X}'(G)$:= Número cromático por arestas de G

Definição 1.1 (Grau Médio): Dado um grafo não-dirigido $G(V, E)$, o grau médio de G é:

$$\delta_{\text{med}}(G) := \frac{1}{|V|} \sum_{v_i \in V} \delta(v_i) \tag{1}$$

Se G é dirigido, podemos definir os graus médios de entrada e saída

$$\delta^{\text{in}}_{\text{med}}(G) := \frac{1}{|V|} \sum_{v_i \in V} \delta^{\text{in}}(v_i) \quad \text{Entrada} \tag{2}$$

$$\delta^{\text{out}}_{\text{med}}(G) := \frac{1}{|V|} \sum_{v_i \in V} \delta^{\text{out}}(v_i) \quad \text{Saída} \tag{3}$$

Definição 1.2 (Distribuição do Grau): A distribuição do grau de um Grafo $G(V, E)$ é a distribuição da variável aleatória X, sendo X o grau do vértice que eu escolho ao acaso

Para os teoremas a seguir e daqui em diante, consideremos a matriz de incidência de forma que $A_{ij} = 1$ se a aresta j se conecta no vértice i e, 0 do contrário (-1 se G for dirigido).

Teorema 1.1: Dado um grafo $G(V, E)$ e sua matriz de incidência A, temos que:

$$\text{nº de ciclos} = |E| - \text{posto}(A) \tag{4}$$

Demonstração:

$$\begin{gathered} \text{posto(A)} + \dim(N(A)) = |E| \\ \Leftrightarrow |E| - \text{posto(A)} = \dim(N(A)) \end{gathered} \tag{5}$$

Porém, a dimensão do núcleo de A é a quantidade de ciclos no grafo, então eu tenho que:

$$\text{nº de ciclos} = |E| - \text{posto}(A) \tag{6}$$

$\square$

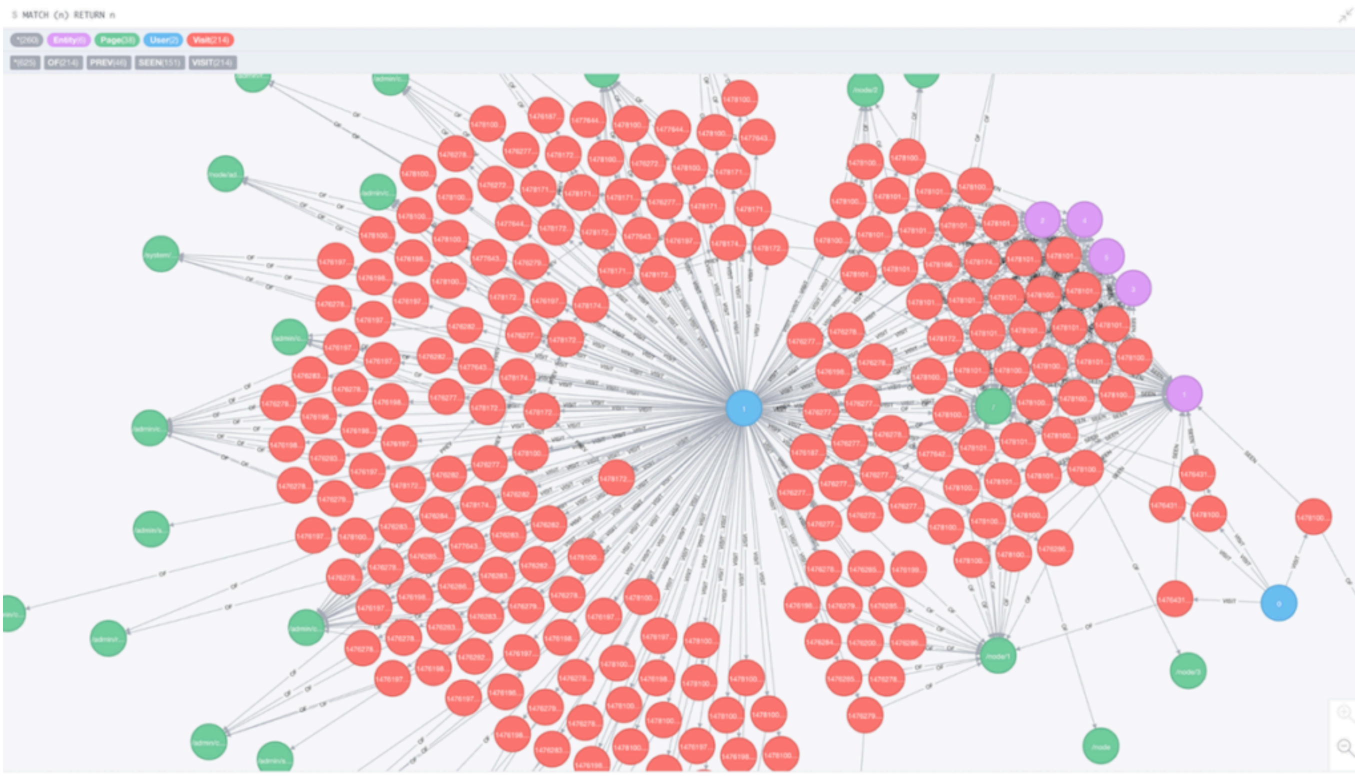
Definição 1.3 (Coeficiente de Clustering): Dado um grafo $G(V, E)$, o coeficiente de clustering de um nó $v \in V$ é:

$$C(v) := \frac{2E_v}{\delta(v)(\delta(v) - 1)} \tag{7}$$

onde E_v é a quantidade de arestas ligadas aos nós vizinhos

Medidas de Centralidade

Quando estamos vendo aplicações reais de grafos, é muito comum querermos ver o "quão importante" um nó é no contexto que estamos analisando. Por exemplo, se nosso grafo representa as conexões entre servidores que um pacote pode percorrer, faz muito sentido querermos ver qual o servidor que quase todos os pactes percorrem



Imagine que esse é o grafo que estávamos falando (Não importa o que ele representa de verdade, só finge que é o caso que falamos), então o nó azul tem uma importância MUITO grande, mas como podemos medir isso? Nem sempre o grafo vai ta arrumadinho assim pra gente. Daí que surgem as medidas de Centralidade.

Definição 2.1 (Farness/'Lonjura'): Dado um grafo $G(V, E)$, a farness de um vértice v_i é dada por

$$L(v_i) := \sum_{v_i \neq v_j \in V} d(v_i, v_j) \tag{8}$$

onde $d(v_i, v_j)$ é o tamanho do menor caminho entre v_i e v_j

Essa medida mede o quão longe o nó está dos outros, de forma que, quanto maior essa medida é, menos importante o meu nó é (Depende do contexto analisado)

Definição 2.2 (Closeness/Proximidade): Dado um grafo $G(V, E)$, a proximidade/closeness do vértice $v_i \in V$ é dada por:

$$C(v_i) := \frac{|V|}{L(v_i)} \tag{9}$$

Por convenção, se v_i e v_j estão em componentes conexas separadas em G, então $d(v_i, v_j) = \infty$, o que torna a definição de antes inútil, então podemos redefinir como:

$$C(v_i) := \frac{1}{|V|} \sum_{v_i \neq v_j \in V} \frac{1}{d(v_i, v_j)} \tag{10}$$

Definição 2.3 (Betweeness/Intermediação): Dado um grafo $G(V, E)$ e $P(v_i, v_j)$ o conjunto de todos os menores caminhos possíveis entre v_i e v_j, então a intermediação de v_i é:

$$B(v_i) := \sum_{v_s, v_t \in V} \frac{|c \in P(v_s, v_t); v_i \in c|}{|\ P(v_s, v_t)\ |} \tag{11}$$

Saindo um pouco dessas definições, vamos tentar pensar em alguma medida mais básica e intuitiva. Uma medida bem padrão que podemos pensar logo de cara é simplesmente o grau do vértice, já que, quanto mais vértices ele se ligar, mais importante ele é! Em muitas literaturas sobre redes o grau do vértice é chamado de **Centralidade de Grau**.

Um outro pensamento que pode surgir a partir desse é: "Poxa, meu vértice tem um grau alto, então ele é importante, mas eu quero valorizar aqueles vértices que se conectam com ele, afinal, se ele é importante, os vértices que estão diretamente ligados nele também são, não é?", e esse pensamento não está errado! É dessa ideia que surge a centralidade por autovetor. Funciona assim: Vamos inicialmente assumir que todos os nossos vértices v_i tem importância $x_i^{(0)} = 1$, o que não me é muito útil agora, porém, vamos tentar fazer uma nova estimativa baseada nos vizinhos, que tal a nova centralidade do vértice v_i ser a soma da centralidade dos vizinhos? Isso faz com que a importância do v_i se baseie no quão importante são seus vizinhos! Eu posso expressar isso com uma fórmula:

$$x_i^{(1)} = \sum_j A_{ij} x_j^{(0)} \tag{12}$$

Onde A é minha matriz de adjacência. Se meu nó v_i não é vizinho de v_j, então $A_{ij} = 0$ o que faz com que minha centralidade $x_j^{(0)}$ não seja somada. Posso reformular isso de forma matricial:

$$x^{(1)} = Ax^{(0)} \tag{13}$$

onde $x^{(k)}$ é o vetor com entradas $x_i^{(k)}$. Se fizermos esse processo várias vezes, depois de k passos, vamos ter algo do tipo:

$$x^{(k)} = A^k x^{(0)} \tag{14}$$

Tomemos a liberdade, então, de escrever $x^{(0)}$ como uma combinação linear dos autovetores w_j de A de forma que

$$x^{(0)} = \sum_{j=1}^{n} c_j w_j \tag{15}$$

Para alguma escolha apropriada de c_j. Então temos:

$$x^{(k)} = A^k \sum_{j=1}^{n} c_j w_j = \sum_{j=1}^{n} c_j \lambda_j v_j = \lambda_1^k \sum_{j=1}^{n} c_j \left(\frac{\lambda_j}{\lambda_1}\right)^k w_j \tag{16}$$

De forma que λ_j são os autovalores de A e λ_1 pode ser, sem perca de generalização, o maior de todos em módulo. Como $\lambda_i/\lambda_1 < 1 \ \forall \lambda_i \text{ com } i \neq j$, então:

$$\lim_{k\to\infty} \sum_{j=1}^{n} c_j \lambda_j^k w_j = c_1 \lambda_1 w_1 \tag{17}$$

Ou seja, o vetor de centralidades que limita as centralidades que eu fiz antes é proporcional ao autovetor associado ao maior autovalor de A, que é equivalente a dizer que o vetor de centralidades x satisfaz:

$$Ax = \lambda_1 x \tag{18}$$

> **Definição 2.4** (Centralidade Autovalor): Seja r um vetor com as centralidades dos vértices v_i de uma rede G de forma que $r_i = \text{centralidade de } v_i$, então:
>
> $$Ar = \lambda_1 r \tag{19}$$
>
> onde λ_1 é o maior autovalor de A

Agora temos outro problema. Quando temos um grafo dirigido, essa medida de centralidade autovalor já não funciona, já que se um nó não tem nenhuma aresta apontando para ele (Apenas saem arestas dele), ele não terá sequer uma centralidade, e isso afeta não só esse vértice como os vértices que ele aponta, que não terão nenhuma "pontuação" adicionada por serem apontados por esse vértice, e isso não pode ocorrer, já que não faz muito sentido na maioria das aplicações práticas. O que podemos fazer para contornar isso? Então entra a solução a seguir:

$$x_i = \alpha \sum_j A_{ij} x_j + \beta \tag{20}$$

Onde α e β são constantes positivas. O primeiro termo é a centralidade autovetor que vimos antes, porém o termo β garante que os nós que comentei anteriormente (Sem grau de entrada) possuam uma pontuação e possam contribuir para a pontuação dos nós que eles apontam. Essa medida é interessante por conta do termo α que balanceia o termo constante e a medida de centralidade autovetor. Podemos expressar isso de forma matricial:

$$x = \alpha A x + \beta \mathbf{1} \tag{21}$$

Onde $1 = (1, ..., 1)$. Se rearranjarmos para x, obtemos:

$$x = \beta (I - \alpha A)^{-1} \mathbf{1} \tag{22}$$

Normalmente colocamos $\beta = 1$ pois não estamos interessados em saber o valor exato das centralidades, mas saber quais vértices são ou não mais ou menos centrais.

$$x = -\alpha \left(A - \frac{1}{\alpha} I\right)^{-1} \tag{23}$$

Perceba que eu quero que $A - \frac{1}{\alpha} I$ seja invertível, e isso acontece quando $\frac{1}{\alpha} \neq \lambda_j$ onde λ_j são os autovalores de A. Ou seja, o meu α não é completamente arbitrário, eu vou ter que analisar o contexto da minha aplicação. Porém, muito comumente, se é utilizado $\alpha = \frac{1}{\lambda_1}$ com λ_1 sendo o maior autovalor

Definição 2.5 (Centralidade de Katz): Dado uma rede $G(V, E)$ e duas contantes $\alpha, \beta > 0$, o vetor de centralidades de katz de todos os nós em V é:

$$K(V) = \beta(I - \alpha A)^{-1}\mathbb{1} \tag{24}$$

Onde A é a matriz de adjacência de G. ($K(V) \in \mathbb{R}^{|V|}$)

Um outro tipo de medida surge quando queremos responder a questão: "Se eu estou navegando entre meus nós, ao longo prazo, qual é o nó que eu mais vou percorrer/parar nele?". Um exemplo são páginas na internet que referenciam entre si, daí surge o nome da medida: **PageRank**. O que fazemos essencialmente é transformar a rede em uma cadeia de markov. Por exemplo:

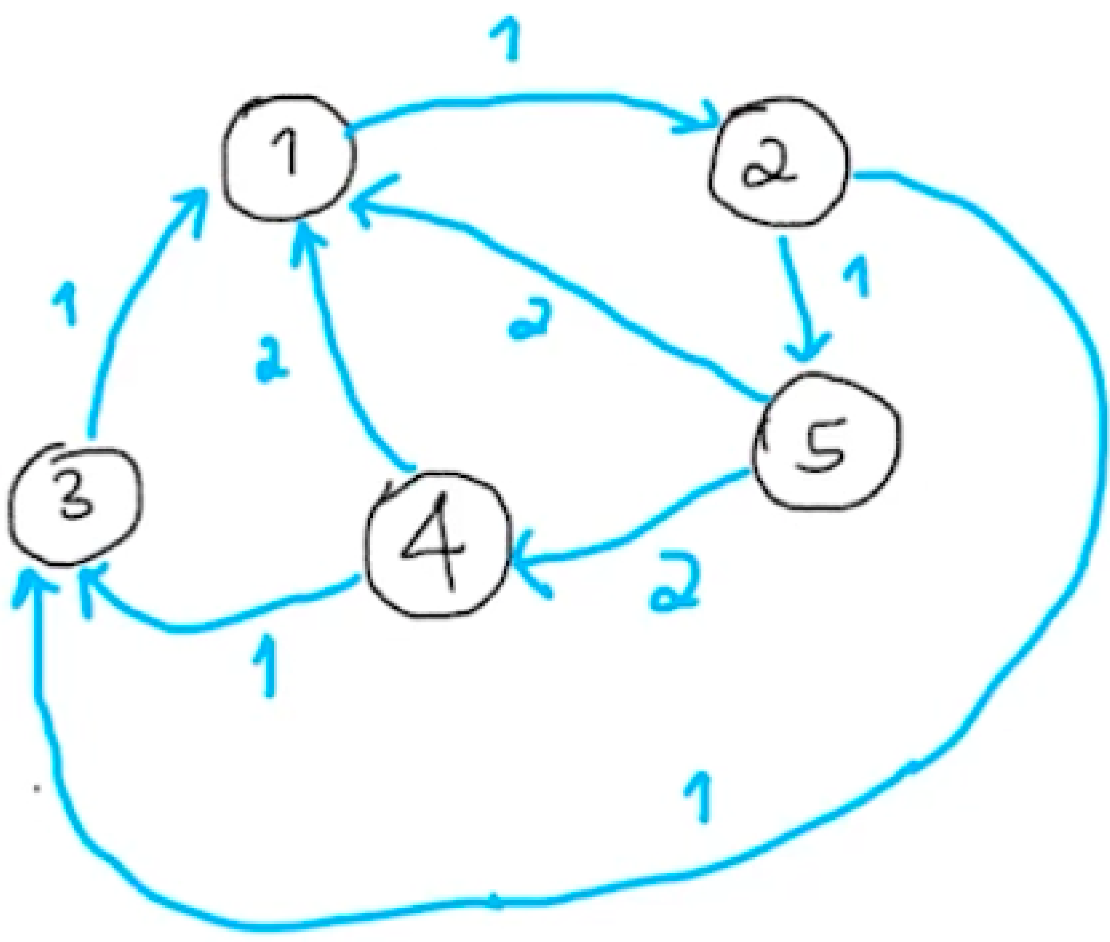


Figura 1: Grafo de Exemplo 1

Vamos supor que estamos no nó 4 e queremos escolher aleatoriamente entre os nós 3 e 1 para irmos, como podemos ver na distribuição dos pesos (Nesse exemplo, isso indica que a página 4 tem 2 links referenciando a página 1 e apenas 1 link referenciando a página 3), então teríamos:

$$\begin{aligned} \mathbb{P}(4 \to 3) &= \frac{1}{3} \\ \mathbb{P}(4 \to 1) &= \frac{2}{3} \end{aligned} \tag{25}$$

E fazemos isso definindo uma matriz estocástica H de tal forma que:

$$H_{ij} = \frac{A_{ij}}{\sum_k^n A_{ik}} \tag{26}$$

Com A sendo a matriz de adjacência. De forma que a soma de todos os elementos de uma coluna dê 1. Agora que vem o truque interessante. Dado um vetor $p \in \mathbb{R}^n$ de forma que cada entrada de p_i representa a chance de eu ir do nó que eu estou para o nó v_i (Ou seja, p tem que ser alguma coluna de H), ao fazer a operação:

$$Hp \tag{27}$$

Eu estou ponderando as probabilidades de p com os seus respectivos nós, ou seja, $(Hp)_k$ representa a probabilidade esperada de que, ao sair do nó v_i, eu vá para o nó v_k. Se isso é verdade e, como eu defini antes, eu quero saber qual nó é mais visitado conforme se passa o tempo, faz sentido eu refazer esse processo inúmeras vezes, então eu tenho uma centralidade do vértice v_i:

$$r = \lim_{t \to \infty} H^t p \tag{28}$$

Com p sendo a i-ésima coluna de H. Porém isso ainda nos trás um problema, veja essa outra rede:

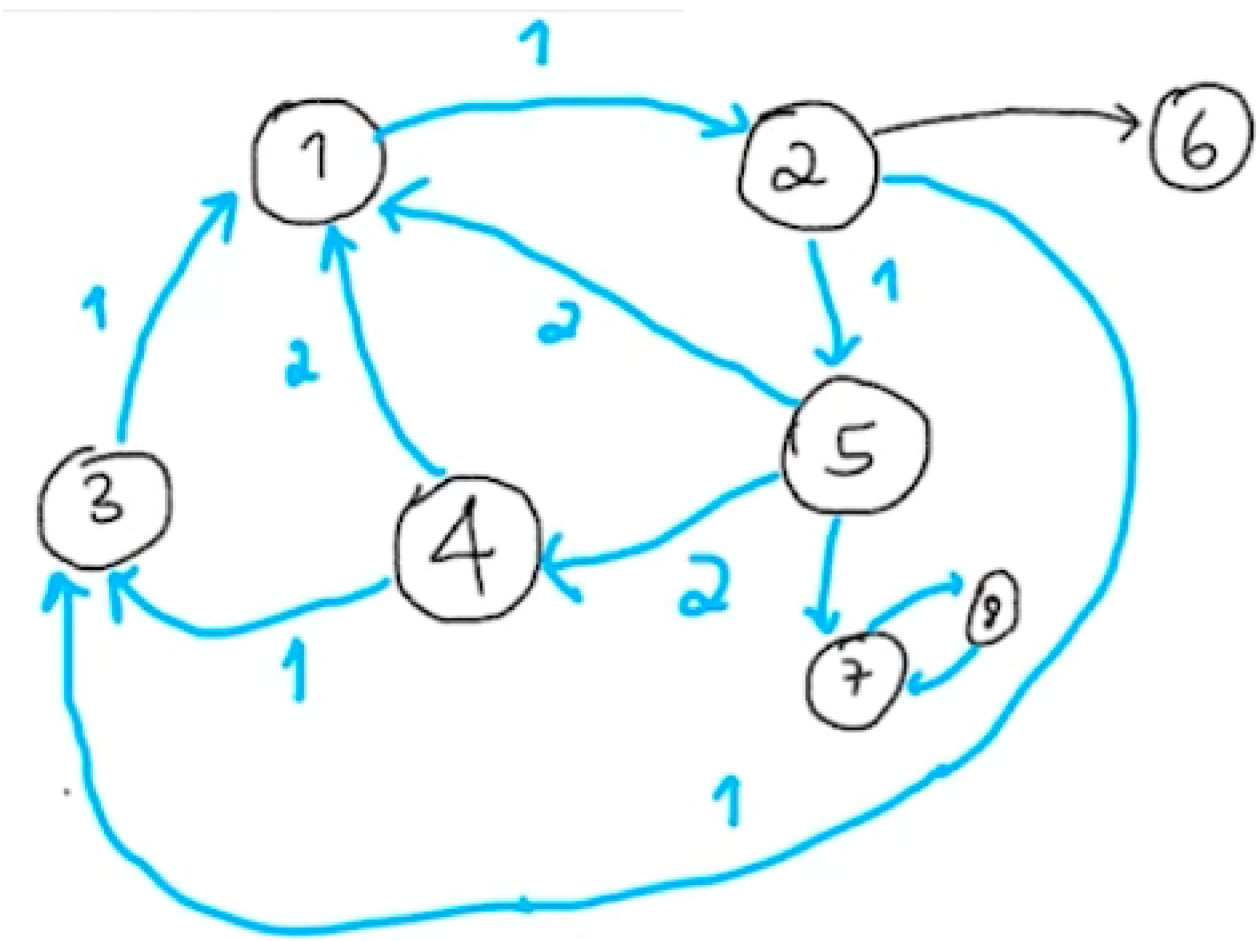


Figura 2: Grafo de Exemplo 2

Veja que, por conta do nó 6, eu não posso transformar meu esquema em uma cadeia de markov, pois eu teria uma coluna de 0, e no caso dos nós 7 e 8 eu teria um problema por conta que eles sempre vão um para o outro. Como podemos resolver isso? O PageRank vem para resolver isso. Vamos pensar no caso da internet, você navegador aleatório, uma hora, pode se cansar de estar onde estar, e visitar uma página aleatoriamente dentro da sua rede, e é nessa ideia que trabalhamos em cima.

Definimos um $\alpha \in (0,1)$, onde podemos interpretar α como a chance do meu navegador permanecer no meu nó. Definimos então nossa nova matriz de chances da seguinte forma:

$$\mathbb{G} = \alpha H + (1-\alpha)C \tag{29}$$

De forma que C é uma matriz $n \times n$ com todas as entradas iguais a $1/n$ para representar um dirigido onde todos os nós apontam para todos os outros nós (Representando a ideia de que eu posso ir para o nó que eu quiser). Porém, há uma propriedade que, se eu tenho uma combinação convexa entre duas matrizes estocásticas/markovianas, então o resultado é uma matriz markoviana. Ou seja, eu ainda posso aplicar a mesma ideia de antes do vetor p_0 inicial e aplicar o limite, assim, eu vou obter meu vetor de centralidades r, de tal forma que

$$\lim_{t\to\infty} \mathbb{G}^t p_0 = r \tag{30}$$

Definição 2.6 (PageRank): Sejam a matriz $\mathbb{G}$ como definida anteriormente e o vetor inicial p_i sendo a i-ésima coluna de $\mathbb{G}$, então o vetor de centralidades PageRank r onde a k-ésima entrada é a centralidade de v_k, então:

$$r = \lim_{t\to\infty} \mathbb{G}^t p_0 \tag{31}$$

Redes Aleatórias

3.1 Ideia Inicial

Também chamadas de **Redes Erdös-Renyi** ou **Redes de Poisson**, são tipos de redes que vão se montando aleatoriamente. Por exemplo, imagine que você está em uma festa e o anfitrião está fornecendo um vinho da melhor qualidade, mas ele não avisou ninguém. Um convidado curioso, por acidente, provou desse vinho e **adorou**, então ele vai contar para as pessoas da festa. A pergunta é, para quem ele vai falar? Ele vai falar para todos? Vai sobrar vinho para você?

Em cima disso conseguimos montar as redes aleatórias, onde cada par de nós (Aresta) é formado de acordo com uma **probabilidade**

> **Definição 3.1.1** (Rede Aleatória): Uma rede aleatória é um grafo $G(V, E)$ de $|V| = N$ nós onde cada par de nós é conectado por uma probabilidade $\boldsymbol{p}$

Considere, agora, uma rede aleatória $G(V, E)$ com $|V| = N$. Sendo L a variável aleatória que representa a quantidade de arestas em E, queremos descobrir sua distribuição. Como cada aresta tem uma probabilidade p de aparecer, podemos interpretar como ela aparecer ou não sendo uma variável indicadora, de forma que o número total de arestas segue uma distribuição binomial (Soma de variáveis de bernoulli independentes). Ou seja, a probabilidade a quantidade de arestas ser $L = l$ é:

$$\mathbb{P}(L = l) = \binom{\binom{N}{2}}{l} p^l (1-p)^{\frac{N(N-1)}{2} - l} \tag{32}$$

Podemos aplicar a mesma ideia para o grau de um vértice também. Vamos definir que K é a variável aleatória que representa o **grau de um vértice arbitrário**, então:

$$\mathbb{P}(K = k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} \tag{33}$$

Já que meu vértice pode se ligar a $N - 1$ vértices com probabilidade p, então isso vira a soma das variáveis indicadores que são 1 quando o meu vértice se liga com outro vértice ($\mathbb{P}(\mathbb{I} = 1) = p$), de forma que eu tenho a soma de $N - 1$ variáveis de bernoulli independentes

Com isso, nós podemos definir o grau médio de G como $\mathbb{E}[K]$:

$$\delta_{\text{med}}(G) = \mathbb{E}[K] = (N-1)p \tag{34}$$

E podemos obter também a variância dos graus

$$\mathbb{V}(K) = (N-1)p(1-p) \tag{35}$$

Então, apenas para resumir, temo que:

$$\begin{gathered} L \sim \text{Bin}\left(\binom{N}{2}, p\right) \\ K \sim \text{Bin}(N-1, p) \end{gathered} \tag{36}$$

Porém, em redes reais, elas são **esparsas**, ou seja, eu tenho **muitos** nós e graus pequenos ($N \gg \mathbb{E}[K]$ notação que diz que N é **muito maior** que $\mathbb{E}[K]$). E lembra qual é a distribuição que é a binomial com n muito grande? Exato, a **Poisson**! Essas redes aleatórias também são chamadas de **redes de poisson**. Vamos, a partir de agora, denotar $\delta_{\text{med}}(G) = \mathbb{E}[K] = \hat{k}$

$$\mathbb{P}(\delta(v) = k) = e^{-\hat{k}} \frac{\hat{k}^k}{k!} \tag{37}$$

Ou seja, para N muito grande e k pequeno com relação a N, podemos estimar de forma que:

$$K \sim \text{Poisson}\left(\hat{k}\right) \tag{38}$$

E isso tudo nos dá um resultado bem condizente e intuitivo, que é que, conforme nós aumentamos a probabilidade p de uma aresta existir, então a rede vai ficando cada vez mais densa

3.2 Evolução das Redes Aleatórias

Conforme iniciamos um grafo com um grau médio 0 e vamos aumentando ele aos poucos, nós percebemos que a partir de um ponto chave, os nós começam a se agrupar em algo que chamamos de **componente gigante**, que seria a maior componente conexa da rede.

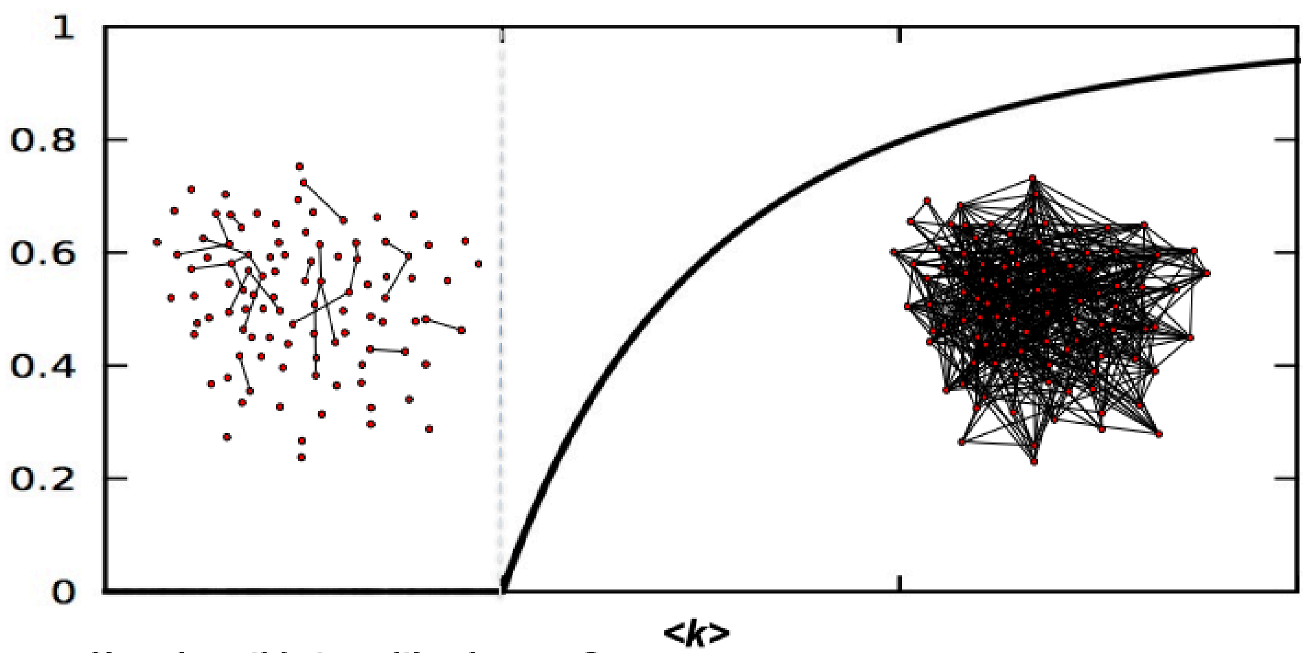


Figura 3: Gráfico que mostra a fração de nós dentro de uma grande componente conexa em função do grau médio

Quanto $\hat{k} < 1$, então a quantidade de nós na componente gigante é desprezível em relação à quantidade de nós na rede, porém, a partir de $\hat{k} = 1$, isso indica que temos, pelo menos, $\frac{n}{2}$ componentes conexas, o que já começa a fazer uma diferença no gráfico. Esse é um argumento utilizado por Erdös e Renyi em um paper por eles publicado

Teorema 3.2.1 (Ponto Crítico): Temos uma componente gigante $\Leftrightarrow \mathbb{E}[K] \geq 1$

Demonstração: Dado uma rede $G(V, E)$, vamos definir a fração de nós que **não está** na componente gigante como:

$$u = 1 - \frac{N_G}{|V|} \tag{39}$$

De forma que N_G é a quantidade de nós dentro dessa componente gigante, vamos definir essa componente como $\Psi \subseteq V$. Se um nó $v_i \in \Psi$, então ele deve estar interligado com outro nó v_j, que também deve satisfazer $v_j \in \Psi$. Por isso, se $v_i \notin \Psi$, então isso pode ocorrer por duas razões:

- $\{v_i, v_j\} \notin E$. A probabilidade de isso acontecer é $1 - p$
- $\{v_i, v_j\} \in E$, porém $v_j \notin \Psi$. A probabilidade de isso acontecer é pu

Então temos:

$$\mathbb{P}(v_i \notin \Psi) = 1 - p + pu \tag{40}$$

Então a probabilidade de que v_i não esteja linkado a Ψ por qualquer nó é de $(1 - p + pu)^{|V| - 1}$, já que temos outros $|V| - 1$ nós que poderiam fazer com que v_i se interligasse a componente gigante.

Sabemos que u é a fração de nós que não está em Ψ, para qualquer p e $|V|$, a solução da equação

$$u = (1 - p + pu)^{|V| - 1} \tag{41}$$

nos dá o tamanho da componente gigante por meio de $N_G = |V|(1 - u)$. Usando $p = \frac{\hat{k}}{|V|-1}$ e tirando $\log$ de ambos os lados, para $\hat{k} \ll |V|$ (Grau médio **muito** menor que $|V|$), obtemos:

$$\begin{gathered} \ln(u) \approx (|V| - 1) \ln\left[1 - \frac{\hat{k}}{|V| - 1}(1 - u)\right] \\ \text{Tiramos exponencial e obtemos:} \\ u \approx \exp\left\{-\frac{\hat{k}}{1 - u}\right\} \end{gathered} \tag{42}$$

Se denotarmos $S = \frac{N_G}{|V|}$, obtemos que:

$$S = 1 - e^{-\hat{k} \cdot S} \tag{43}$$

Agora obtemos o tamanho da componente gigante em função do **grau médio**. O ponto crítico ocorre na mudança de fase do sistema (Tópico mais complicado que não compreendo, estou apenas falando o que o livro do Barabas fala), que é quando os dois lados da igualdade tem a mesma derivada, então:

$$\begin{gathered} \frac{\mathrm{d}}{\mathrm{d}S}\left(1 - e^{\hat{k}S}\right) = 1 \\ \hat{k}e^{-\hat{k}S} = 1 \end{gathered} \tag{44}$$

Onde, colocando $S = 0$, descobrimos que o ponto crítico é $\hat{k} = 1$ □

Na verdade esse resultado é bem intuitivo. Faz sentido dizer que para que uma componente gigante exista, todos os nós precisam ter pelo menos grau 1, já que eles precisam estar conectados com algum outro nó, porém, o que não é muito intuitivo, é que todos eles terem grau 1 é **suficiente** para que a componente gigante apareça

A gente pode reescrever $\mathbb{E}[K] = 1$ como:

$$\mathbb{E}[K] = 1 \Leftrightarrow p(N - 1) = 1 \Leftrightarrow p = \frac{1}{N - 1} \approx N \tag{45}$$

E o que isso significa? Isso mostra outro resultado intuitivo. Quanto maior é minha rede, **menos probabilidade eu preciso para que uma componente gigante apareça**

Algo interessante que podemos fazer é analisar como a proporção $\frac{N_G}{N}$ (Porcentagem de nós dentro da componente gigante) se comporta conforme nós aumentamos $\mathbb{E}[K]$. Nós fazemos isso dividindo esse processo em 4 fases (Ou 4 **regimes**), veja a imagem abaixo:

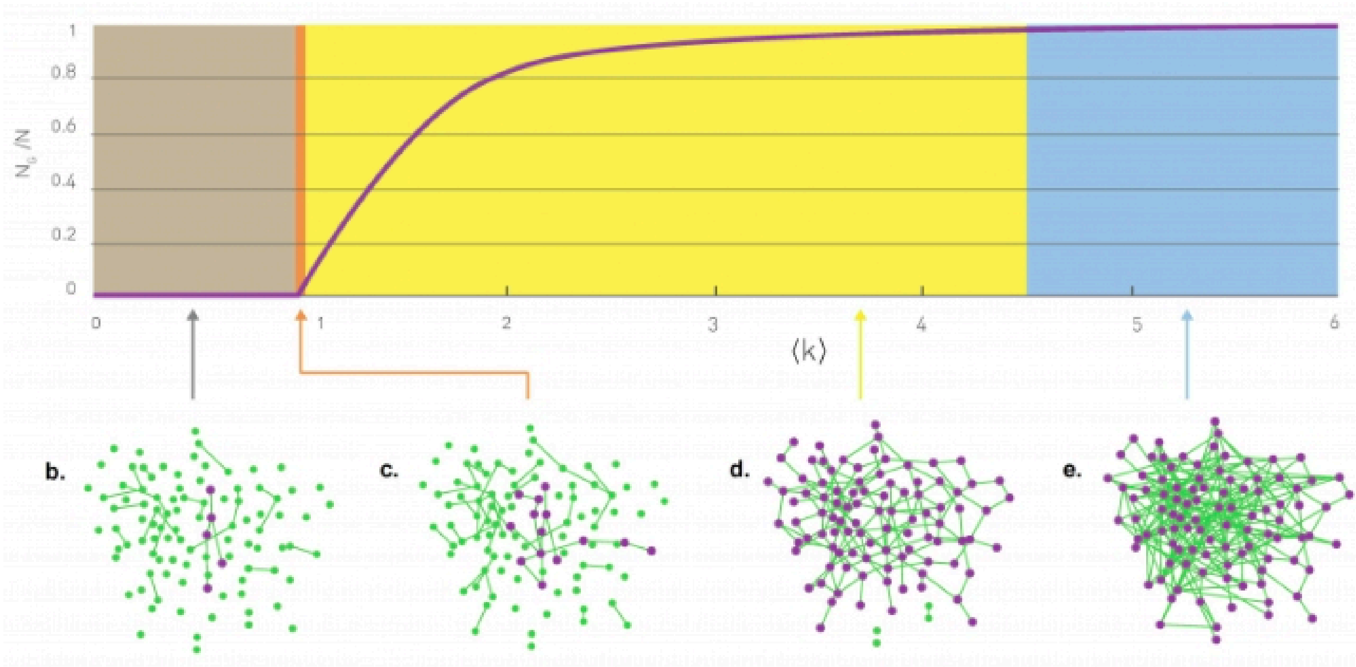


Figura 4: Crescimento da componente conexa em função do grau médio

3.2.1 Regime Subcrítico ($0 < \hat{k} < 1$)

Quando $\hat{k} = 0$, temos N nós soltos na rede e conforme aumentamos $\hat{k}$, mas mantemos ele menor que 1, temos a formação de vários nós soltos e pequenos agrupamentos (Coisa pouca mesmo). Dessa forma, mesmo escolhendo a componente gigante como o maior desses agrupamentos, a proporção N_G/N ainda vai ser muito baixa. O Barabás aproxima essa relação como

$$\frac{N_G}{N} \approx \frac{\ln(N)}{N} \to 0 \text{ quando } N \to \infty \tag{46}$$

Pois podemos considerar essas componentes menores como várias árvores (Também pequenas)

3.2.2 Ponto Crítico

É a transição do momento onde não há uma componente gigante para o momento que há uma. Porém o tamanho relativo dela (N_G/N) ainda é muito próximo de 0. O livro do Barabás afirma que $N_G \approx N^{\frac{2}{3}}$, então N_G cresce muito mais devagar se comparado a N, logo:

$$\frac{N_G}{N} \approx N^{-\frac{1}{3}} = O(N) \tag{47}$$

Porém, perceba que o salto de diferença de tamanho pode ser enorme dependendo da rede. Se pegarmos uma rede de tamanho $N = 7 \times 10^9$ (Parecido com a rede mundial), para $\hat{k} < 1$, a gente teria que o tamanho da componente gigante era de ordem:

$$N_G \approx \ln(N) \approx 22.7 \tag{48}$$

Em contraste, se $\hat{k} = 1$, então teriamos que

$$N_G \approx N^{\frac{2}{3}} \approx 3 \times 10^6 \tag{49}$$

Que é uma diferença notável no tamanho das componentes gigantes

3.2.3 Regime Supercrítico ($\hat{k} > 1$)

Esse regime tem mais relevância para redes reais, já que a componente gigante começa a se parecer realmente com uma rede. Aqui, o tamanho N_G pode ser dado como:

$$N_G = (p - p_c)N \tag{50}$$

Onde $p_c = 1/(N)$. Ou seja, conforme eu aumentar meu grau médio, menor vai ficar meu p e maior será a fração de nós que pertencem à componente gigante. Em resumo, nesse regime, várias componentes conexas coexistem junto da componente gigante, onde a componente gigante é uma rede comum, enquanto as outras componentes conexas são mais prováveis de serem árvores

3.2.4 Regime Conexo ($\hat{k} > \ln(N)$)

Agora, nesse regime, temos que o grafo é (ou quase) conexo, logo, todos os nós fazem parte da componente conexa (Ou a maioria, logo $N_G \approx N$)

Teorema 3.2.4.1: Se $N_G \approx N$, o valor de $\hat{k}$ que satisfaz a propriedade de **a maior parte dos nós estarem na componente gigante** é:

$$\hat{k} = \ln(N) \Rightarrow p = \frac{\ln(N)}{N} \tag{51}$$

Demonstração: Para determinar o valor de $\hat{k}$ no qual a maior parte dos nós fazem parte da componente gigante, temos que saber a probabilidade de que um **nó aleatório não tenha um link para a componente gigante**, e isso é:

$$(1-p)^{N_G} \approx (1-p)^N \tag{52}$$

Já que eu tenho exatamente N_G nós na componente gigante e eu não quer me ligar com nenhum deles. Novamente, tomando $\mathbb{I}_k$ sendo a variável indicadora de que um nó k **não** na componente gigante (1 quando ele não está), temos que a **quantidade de nós que não estão na componente gigante** tem uma distribuição **binomial** com parâmetros N, $(1-p)^N$, então, se considerarmos L_G sendo essa quantidade, temos que:

$$\mathbb{E}[L_G] = N(1-p)^N = N\left(1 - \frac{Np}{N}\right)^N \approx Ne^{-Np} \tag{53}$$

Queremos então chegar no ponto em que temos, para um p suficientemente próximo de 1, que apenas 1 único nó esteja fora da componente conexa, então gostaríamos de analisar em que ponto:

$$\mathbb{E}[L_G] = 1 \Leftrightarrow Ne^{-Np} = 1 \tag{54}$$

Logo, tirando $\ln$ em ambos os lados, chegamos que:

$$p = \frac{\ln(N)}{N} \tag{55}$$

Ou seja

$$\hat{k} = \ln(N) \tag{56}$$

□

Esse resultado é de grande impacto! Quando analisamos muitas das redes reais, a maioria segue esse padrão de $\hat{k} = \ln(N)$, logo, as **redes reais são supercríticas**. Veja a tabela presente no livro do Barabás:

Rede	**N**	**L**	$\mathbb{E}[K]$	$\ln(N)$
Internet	192244	609066	6.34	12.17
Power Grid	$4,941$	$6,594$	2.67	8.51
Science Collaboration	$23,133$	$94,437$	8.08	10.05
Actor Network	$702,388$	$29,397,908$	83.71	13.46
Protein Interactions	$2,018$	$2,930$	2.90	7.61

Figura 1: Tabela de redes no Barabás

3.3 Distribuição de tamanhos de Cluster

Queremos também ter uma noção da probabilidade de um nó v_i qualquer estar em um cluster (Grupo de nós na rede) de tamanho s. No livro do Newman, ele nos mostra que essa probabilidade é:

$$\mathbb{P}\left(v_i \in \Psi_{|\Psi|=s}\right) = e^{-\delta_{\text{med}}(G)\cdot s} \frac{\left(\delta_{\text{med}}(G) \cdot s\right)^{s-1}}{s!} \tag{57}$$

3.4 Mundos pequenos

Mundos pequenos (Small worlds) são grafos em que, independente da quantidade de vértices, a distância entre dois nós aleatórios costuma ser muito pequeno. Um exemplo é um modelo que cada nó representa todas as pessoas do mundo e as arestas indicam se elas já interagiram e se conhecem ou não (Impressionantemente), tanto que existe a teoria dos 6 graus de distância entre as pessoas

Vídeo sobre o assunto (Clique aqui)

E se quisermos ter uma noção de o quão **não-relacionadas** duas pessoas são em uma rede social? Podemos calcular sua distância, obviamente, mas alguns algoritmos ficam computacionalmente inviáveis. Podemos então estimar uma distância média entre dois nós selecionados aleatoriamente no grafo.

Tendo uma rede $G(V, E)$ com grau médio $\hat{k} = \mathbb{E}[K]$, é intuitivo pensar que cada nó tem, em média:

$$\begin{aligned} &\hat{k} \text{ nós a 1 unidade de distância} \\ &\hat{k}^2 \text{ nós a 2 unidades de distância} \\ &\vdots \\ &\hat{k}^d \text{ nós a } d \text{ unidades de distância} \end{aligned} \tag{58}$$

Então é plausível dizer que a quantidade média de nós presentes até uma distância d de um nó qualquer é expresso como:

$$N(d) = \sum_{i=0}^{d} \hat{k}^i = \frac{\hat{k}^{d+1} - 1}{\hat{k} - 1} \tag{59}$$

Sabemos que esse valor não pode ter valores arbitrários, ele não passa de $|V| = N$, então podemos encontrar o grau médio que satisfaz esse o valor. Assim, fazemos:

$$\frac{\hat{k}^{d+1} - 1}{\hat{k} - 1} \approx N \tag{60}$$

Assumindo que $\hat{k} \gg 1$, podemos desprezar os termos -1, assim vamo obter que:

$$d_{\max} \approx \frac{\ln(N)}{\ln\left(\hat{k}\right)} \tag{61}$$

Que é a representação matemática do problema dos minimundos. Porém, isso também traz uma interpretação muito interessante.

Porém, aqui a gente ta vendo o **diâmetro** da rede, e nós comentamos anteriormente sobre **a distância entre dois nós aleatórios**. Impressionantemente, essa aproximação também é válida para essa ocasião. Denotando essa distância média, temos que a característica dos minimundos é:

$$\hat{d} \approx \frac{\ln(N)}{\ln\left(\hat{k}\right)} \tag{62}$$

Mas por que isso acontece? Falando de um jeito mais intuitivo, essa aproximação de $d_{\max}$ costuma funcionar mais para a média do caminho entre dois nós aleatórios pois, em redes reais, o $d_{\max}$ é dado por um único caminho ou pouquissimos caminhos daquele tamanho, enquanto $\hat{d}$ é ponderado em todos os nós. Além de que essa fórmula traz algumas intuições interessantes. Ela mostra que a distância média entre os nós aumenta conforme aumentamos o tamanho da rede, mesmo que não linearmente ou exponencialmente. E mostra também com o termo $1/\ln\left(\hat{k}\right)$ que, quanto mais densa é minha rede, menor vai ser a distância média

3.5 Coeficiente de Clustering

Indica o quão agrupado um nó está dentro de uma rede. O grau de um nó não fala nada sobre a relação entre seus vizinhos, e é aí que o coeficiente de clustering entra

Definição 3.5.1 (Coeficiente de Clustering): Dado uma rede $G(V, E)$, o coeficiente de clustering de um nó $v_i \in V$ é definido como:

$$\text{Cluster}(v_i) := \frac{2 \cdot \mathbb{L}(v_i)}{\delta(v_i)(\delta(v_i) - 1)} \tag{63}$$

Onde $\mathbb{L}(v_i)$ é quantas arestas **entre si** os **vizinhos** de v_i possuem e $\frac{\delta(v_i)(\delta(v_i)-1)}{2}$ é a quantidade **máxima** de arestas que poderiam estar interligando os vizinhos de v_i (Quantidade de arestas em um grafo completo $K_{\delta(v_i)}$)

Vamos tomar $\mathbb{L}(v_i)$ como sendo a variável aleatória que indica quantas arestas os vizinhos de v_i tem entre si. Novamente, como sempre, tomamos a variável indicadora $\mathbb{I}_k$ como sendo a variável indicadora que diz se a aresta k faz parte desse grupo de links entre os vizinhos do nó v_i. Sabemos que $\mathbb{P}(I_k = 1) = p$, então $\mathbb{L}(v_i)$ seria uma binomial, mas qual seria o parâmetro da quantidade? Quantas variáveis indicadoras $\mathbb{I}_k$ eu tenho que somar? Se pararmos para pensar, o **máximo** de links que podem existir entre os vizinhos de v_i é o grafo completo formado por todos eles, então, no final, temos que:

$$\mathbb{L}(v_i) \sim \text{Bin}\left(\binom{\hat{k}}{2}, p\right) \quad (64)$$

Então, no final, vamos ter que:

$$\mathbb{E}[\mathbb{L}(v_i)] \approx p\frac{\delta(v_i)(\delta(v_i)-1)}{2} \Rightarrow \text{Cluster}(v_i) = p = \frac{\mathbb{E}[K]}{|V|} \quad (65)$$

Só que sabemos que, em redes aleatórias, para que esse número seja alto, a probabilidade em si das arestas tem que ser alto, porém, se p é alto, então a rede aleatória em si será um grande aglomerado, seria um único cluster enorme. Essa característica é um forte indicativo, por exemplo, de que redes como as **redes sociais não são** redes aleatórias. O livro do Barabás mostra um experimento e mostra que, em redes reais, o coeficiente de clustering é muito maior do que o esperado em redes aleatórias, de forma que, em redes reais, esse coeficiente costuma ser bastante independente de N, diferente do que encontramos agora há pouco

3.6 Grau Máximo e Grau Mínimo

Dependendo do contexto analisado, pode ser de grande interesse saber os valores esperados do **maior grau** de uma rede e do **menor grau**. Para descobrir o **maior grau**, precisamos que, na rede, tenhamos **no máximo** um nó com grau maior que $k_{\max}$. Isso significa que a área do gráfico da distribuição **em frente** a $k_{\max}$ é aproximadamente 1:

$$\begin{gathered} N \cdot \mathbb{P}(K \geq k_{\max}) \approx 1 \\ N \cdot (1 - \mathbb{P}(K < k_{\max})) \approx 1 \end{gathered} \quad (66)$$

E podemos usar um argumento análogo, afirmando que deveríamos ter, no máximo, apenas um nó com grau menor que $k_{\min}$, então teríamos:

$$N \cdot \mathbb{P}(K \leq k_{\min} - 1) = 1 \quad (67)$$

Assim resolvemos as duas equações para achar $k_{\min}$ e $k_{\max}$

3.7 Conclusão

Como conclusão, temos que redes aleatórias **não representam bem as redes da vida real**. Não existem redes na natureza que são corretamente descritas como **redes aleatórias**. Então por que estudar elas? Na verdade, veremos posteriormente que, mesmo elas sendo erradas e irrelevantes, elas são **muito úteis**

Evoluções de Redes

Vimos as redes aleatórias onde os graus dos nós tinham distribução de Poisson. Mas e se eu quisesse fazer uma rede com distribuição diferente? Muitos pacotes de grafos e redes utilizam de **configuration models**, que são funções que recebem a quantidade de nós da rede e um **vetor** que representa a **função de distribuição** dos graus dos nós

Voltando ao assunto sobre **evoluções**, eu estou interessado em pensar um jeito intuitivo/natural de como as redes vão evoluir com o passar do tempo.

Então vamos imaginar o seguinte cenário. Eu tenho uma rede inicial $G_0(V_0, E_0)$ e a cada unidade de tempo t eu vou ter uma nova rede $G_t(V_t, E_t)$, de forma que a cada unidade de tempo, eu vou adicionar um novo nó em V_{t-1} e novas arestas em E_{t-1}. Qual é a distribuição do grau médio desses nós? O que podemos inferir dessa rede?

4.1 Anexação Uniforme

Vamos imaginar uma **anexação uniforme**. Nesse caso, cada nó inserido sempre terá um grau de m. Ou seja, a **probabilidade** de um link do meu novo nó inserido se interligar ao vértice v_i é igual a m/i (A chance de ele se ligar com uma das arestas é $1/i$, logo, como eu posso me ligar com m arestas diferentes, todas independentes entre si, a probabilidade total vai ser m/i), logo:

$$\delta\left(v_j, t = i\right) := \text{Grau de } v_j \text{ no momento } i \tag{68}$$

Com isso, podemos interpretar esse grau como uma **variável aleatória**. Temos que o grau de v_i no momento inicial i é fixa como m. Então a quantidade de arestas no momento $i + 1$ pode ser escrita como:

$$\delta(v_i, i + 1) = m + \mathbb{I}_{i+1}(1) + \mathbb{I}_{i+1}(2) + ... + \mathbb{I}_{i+1}(m) \tag{69}$$

Onde $\mathbb{I}_j(k)$ é a variável indicadora que diz se, no momento j, a aresta k do **novo nó que está sendo adicionado na rede** foi adicionado ou não no nosso nó. Podemos reescrever como a soma de uma única variável aleatória de distribuição binomial também. Você pode ter reparado que eu utilizei i tanto no v_i quanto no i. Vou utilizar isso pois eu estou supondo que, na nossa análise, estamos saindo do último nó adiconado (Uma aproximação razoável do modelo real, obviamente que nem todos os nós vão ser adicionados com essa anexação, já que antes de eu iniciar essa abordagem, já vai ter uma rede "preexistente")

Porém, queremos ter uma **noção** de como isso vai ser ao longo prazo, podemos então tirar a esperança disso.

$$\mathbb{E}[\delta(v_i, i + 1)] = m + \frac{m}{i} \tag{70}$$

Porém, isso é apenas para um único passo, queremos generalizar para vários passos. Vamos supor então que estamos saindo do i-ésimo nó adiconado e estamos no momento t:

$$\delta(v_i, t) = m + \sum_{k=i+1}^{t} \sum_{j=1}^{m} \mathbb{I}_k(j) \tag{71}$$

$$\begin{aligned} \mathbb{E}[\delta(v_i, t)] &= m + \sum_{k=i+1}^{t} \sum_{j=1}^{m} \mathbb{E}[\mathbb{I}_k(j)] \\ &= m + \sum_{k=i+1}^{t} \sum_{j=1}^{m} \frac{j}{k-1} \end{aligned} \tag{72}$$

$$
\begin{aligned}
\mathbb{E}[\delta(v_i,t)] &= m \cdot \sum_{k=i+1}^{t} \frac{1}{k-1} \\
\mathbb{E}[\delta(v_i,t)] &\approx m + m\ln\left(\frac{t-1}{i-1}\right) \\
\frac{\mathbb{E}[\delta(v_i,t)]}{m} - 1 &\approx \ln\left(\frac{t-1}{i-1}\right) \\
\exp\left(\frac{\mathbb{E}[\delta(v_i,t)] - m}{m}\right) &\approx \frac{t-1}{i-1} \\
\exp\left(-\frac{\mathbb{E}[\delta(v_i,t)] - m}{m}\right) &\approx \frac{i-1}{t-1} \\
\exp\left(-\frac{\mathbb{E}[\delta(v_i,t)] - m}{m}\right) &\approx \frac{i}{t}
\end{aligned}
\tag{73}
$$

No intervalo $[0,t]$, temos a seguinte estruturação:

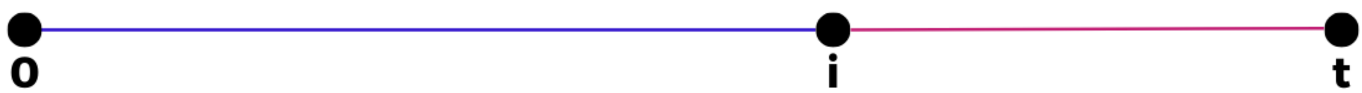


Figura 5: Intervalo de i/t

Então, o que encontramos foi a fração de nós que tem grau maior que v_i. Então temos que:

$$\mathbb{P}(\delta_t(v_i) \leq k) = 1 - e^{-\frac{k-m}{m}} \tag{74}$$

Logo, temos uma distribuição **Exponencial**

4.2 Anexação Preferencial

Na anexação **uniforme**, cada novo nó podia se ligar com um dos nós anteriores com mesma probabilidade. Nessa abordagem, os nós de **maior grau** terão uma **maior preferência** para serem escolhidos (Não é uma obrigatoriedade). Podemos expressar, então da seguinte forma. Antes, vamos fazer duas definições rápidas:

> **Definição 4.2.1**: Dada uma rede $G(V,E)$, o conjunto E_t é definido como o conjunto de arestas da rede no momento t

> **Definição 4.2.2**: Dado uma rede $G(V,E)$, a função $\delta_t : V \to \mathbb{N}$ é a função que retorna o grau de um vértice em um momento t do tempo

Voltando, queremos então, antes de tudo, saber qual que é a probabilidade do nó que vai ser adicionado se ligar com um vértice v_i, então:

$$\mathbb{P}(\{v_i, v_{t+1}\} \in E_{t+1}) = \frac{\delta_t(v_i)}{\sum_{j=1}^{t} \delta_t(v_j)} \tag{75}$$

Queremos achar uma distribuição para os graus dos nós. Vamos tentar achar, então, uma taxa de crescimento do grau dos nós:

$$\frac{\delta_{t+1}(v_i) - \delta_t(v_i)}{\Delta t} = m \cdot \frac{\delta_t(v_i)}{\sum_{j=1}^{t} \delta_t(v_j)} \approx \frac{\mathrm{d}(\delta_t(v_i))}{\mathrm{d}t} \tag{76}$$

Porém, sabemos que $\sum_{j=1}^{t} \delta_t(v_j) = 2\ |E_t|$, então vamos obter:

$$\frac{\mathrm{d}(\delta_t(v_i))}{\mathrm{d}t} = m \cdot \frac{\delta_t(v_i)}{2\ |E_t|} = m \cdot \frac{\delta_t(v_i)}{2tm} = \frac{\delta_t(v_i)}{2t} \tag{77}$$

Logo, obtemos uma EDO para resolver. Vamos chamar $\delta_t(v_i)$ de k apenas para facilitar a visualização:

$$\begin{aligned} \frac{\mathrm{d}k}{\mathrm{d}t} &= \frac{k}{2t} \qquad (\delta_i(v_i) = m) \\ \frac{\mathrm{d}k}{k} &= \frac{\mathrm{d}t}{2t} \\ \int \frac{\mathrm{d}k}{k} &= \int \frac{\mathrm{d}t}{2t} \\ \ln k &= \frac{1}{2} \ln t + C \\ \delta_t(v_i) &= t^{1/2} \cdot D \end{aligned} \tag{78}$$

Resolvendo para o caso $\delta_i(i) = m$, temos:

$$\begin{aligned} m = i^{1/2} D &\Rightarrow D = m \cdot i^{-1/2} \\ &\Rightarrow \delta_t(v_i) = m \left(\frac{t}{i}\right)^{1/2} \end{aligned} \tag{79}$$

Queremos então calcular $\mathbb{P}(\delta_t(v_i) \leq k)$. Na média, todos os nós **posteriores** ao nó v_i tem grau menor do que k, então precisamos apenas inverter aquela equação de antes, assim, vamos obter:

$$\begin{aligned} \delta_t(v_i)^2 &= m^2 \frac{t}{i} \\ \Leftrightarrow i &= \frac{m^2 \cdot t}{\delta_t(v_i)^2} \end{aligned} \tag{80}$$

Assim, conseguimos obter a **fração de nós com grau maior que v_i**, que são justamente os nós anteriores a ele (i/t) que é $m^2 \delta_t(v_1)^{-2}$. Temos então que:

$$\mathbb{P}(\delta_t(v_i) \leq k) = 1 - m^2 k^{-2} \tag{81}$$

Temos também que a densidade vai ser:

$$f_K(k) = 2m^2 k^{-3} \tag{82}$$

Percebemos, então, que a variável aleatória K, que representa o grau de um nó na rede, tem a distribuição $\mathbf{Paretto(2, m)}$. O que, na verdade, faz bastante sentido. A distribuição de Paretto é bastante usada para descrever a concentração de riquezas e, como sabemos bem, o dinheiro costuma se concentrar sempre em quem tem mais dinheiro, então é só imaginar que o grau de um nó representa o quão rica uma pessoa é e bingo, faz todo sentido essa distribuição!

Redes Livres de Escala

São as redes geradas após um processo de **Anexação Preferencial**. Um grande exemplo é a rede da internet (WWW). Quando olhamos ela de relance, ela aparenta ser uma rede aleatória, porém, é notável que certos nós ficam agrupados em regiões com outros nós de grau **muito grande**. Veja essa representação em rede dos documentos da Internet gerada por Hawoong Jeong na Universidade de Notre-Dame

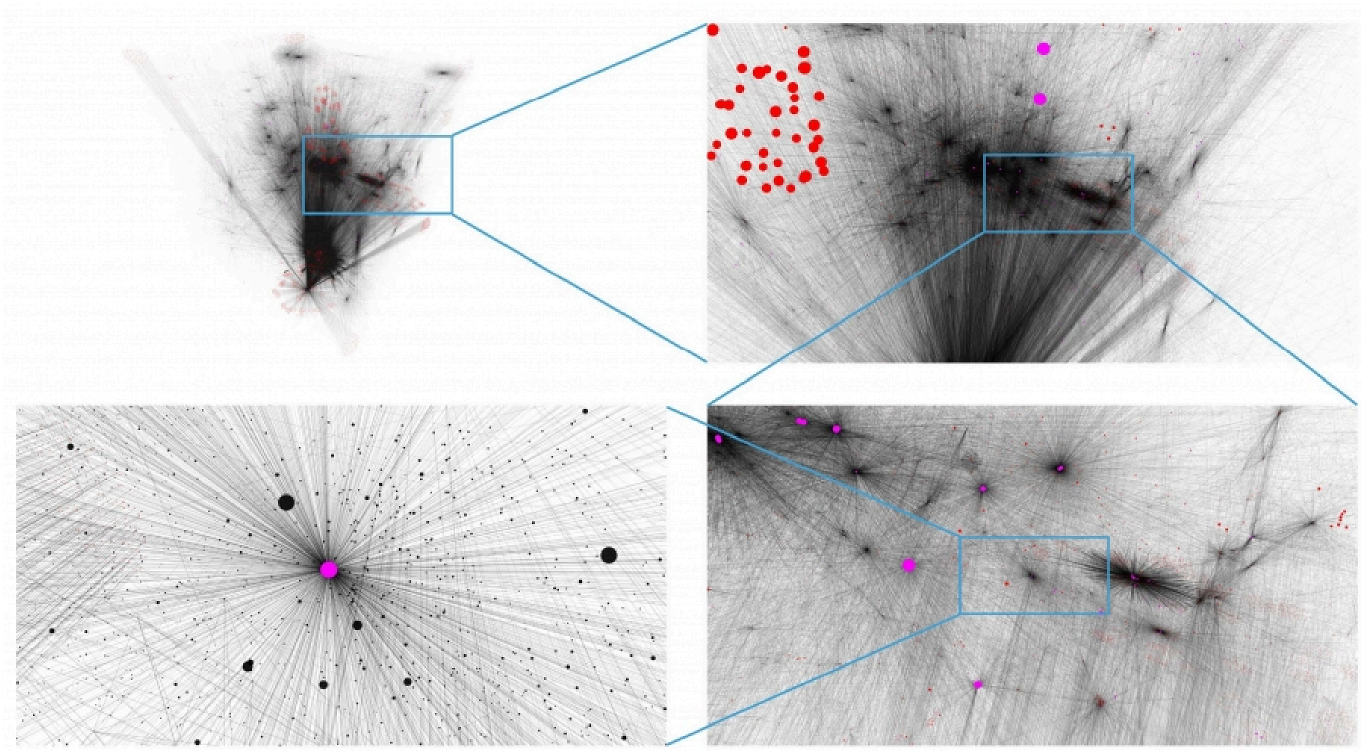


Figura 6: Rede WWW

Se a rede WWW fosse uma rede aleatória, os graus teriam uma distribuição Poisson, porém, como a imagem a seguir mostra, isso não ocorre:

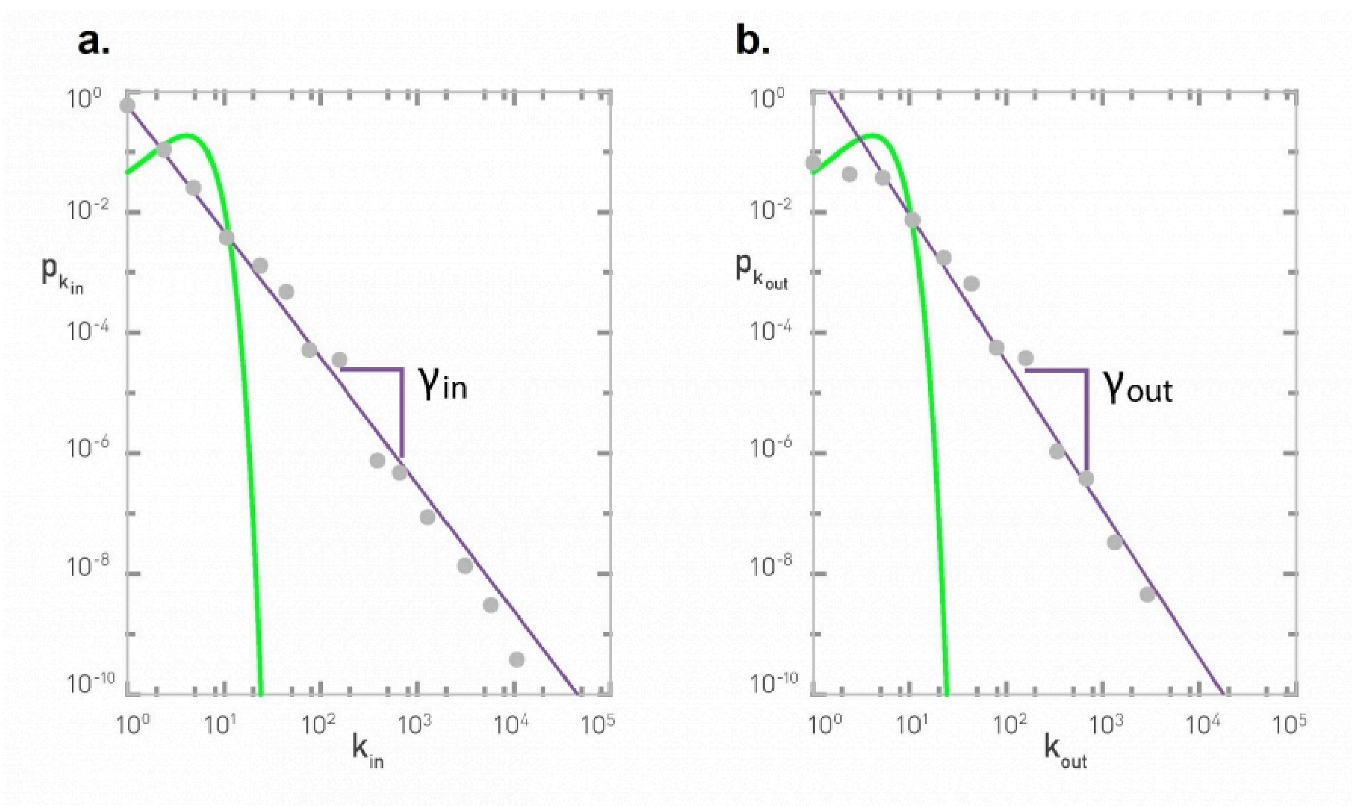


Figura 7: Distribuição dos graus em Log-Log

Como podemos ver, os graus, no gráfico log-log, são bem aproximados por uma reta. Isso é um indicativo de que a sua distribuição é algo parecido com:

$$\mathbb{P}(K = k) = k^{-\gamma} \tag{83}$$

Isso é chamado de **Distribuição de Lei de Potência**, e γ é o **expoente do grau**. A internet é uma rede direcionada, então todo documento tem um grau de entrada e de saída. Com esse contexto em mente, podemos finalmente definir:

> **Definição 5.1** (Redes livres de Escala): Uma rede é dita livre de escala quando a distribuição do grau de seus vértices segue uma forma exponencial

5.1 Formalismo Discreto

Para cálculos analíticos, é interessante deixar que os graus possam assumir qualquer tipo de valor real (Mesmo que apenas os naturais sejam possíveis). Seja K a variável aleatória que indica o grau de um vértice escolhido aleatoriamente, temos que:

$$\mathbb{P}(K = k) = Ck^{-\gamma} \tag{84}$$

Normalizando, temos:

$$\begin{gathered}\int_{k_{\min}}^{\infty} \mathbb{P}(K = k)\, \mathrm{d}k = 1 \\ \Rightarrow C = (\gamma - 1)k_{\min}^{\gamma-1}\end{gathered} \tag{85}$$

Então temos que a distribuição segue a P.M.F:

$$\mathbb{P}(K = k) = (\gamma - 1)k_{\min}^{\gamma-1}k^{-\gamma} \tag{86}$$

5.2 Centros

Vamos analisar a seguinte imagem:

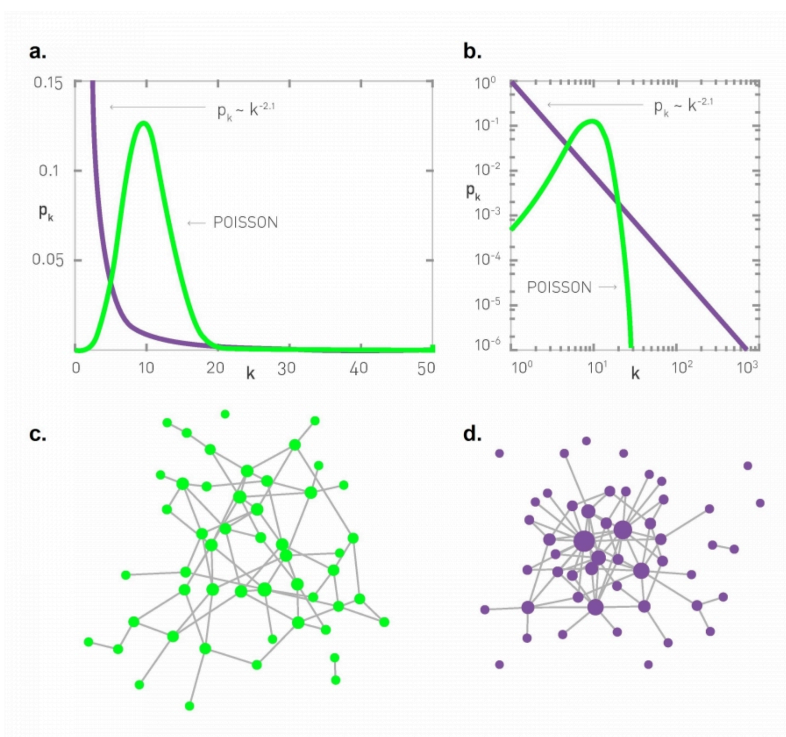


Figura 8: Distribuição de Poisson e Distribuição de Potência

A gente pode analisar em 3 pontos principais:

- Antes de $\hat{k}$ a rede livre de escala é maior, ou seja, há mais nós com grau pequeno nela do que na poisson
- Na vizinhança de $\hat{k}$, a poisson é maior, logo, existe um excesso de nós com grau $\hat{k}$ na rede Poisson
- Depois a rede livre de escala volta a ser maior

Isso significa que, em redes livre de escala, temos altas chances, ou de obter **um nó com grau muito grande (Hub)** ou obter vários nós com grau pequeno

5.2.1 Maior Centro

Também chamados de **hubs**, são os nós mais centrais, aqueles que representam uma maior importância dependendo do contexto, aqui, sendo aqueles com o maior grau. Podemos querer saber como eles se comportam nessas redes livres de escala! Para isso, temos que calcular qual é o maior grau da distribuição $k_{\max}$, também chamado de corte natural da distribuição. Representa o tamanho esperado do maior hub. Antes de partir para o caso complicado geral de $p(k) = Ck^{-\gamma}$, vamos primeiro fazer um caso mais simples, vamos fazer para a **exponencial**:

$$p(k) = Ce^{-\lambda k} \tag{87}$$

Para uma rede com grau mínimo $k_{\min}$, temos que a normalização vai ficar:

$$\int_{k_{\min}}^{\infty} p(k)\,\mathrm{d}k = 1 \Rightarrow C = \lambda e^{\lambda k_{\min}} \tag{88}$$

Agora, para saber $k_{\max}$, fazemos o mesmo processo que vimos antes, vamos supor que, em uma rede com N nós, o valor esperado do grau para o regime $(k_{\max}, \infty)$ seja 1, ou seja:

$$\begin{gathered}\mathbb{E}[K|k > k_{\max}] = 1 \Rightarrow N \cdot \mathbb{P}(K \geq k_{\max}) = 1 \\ \Leftrightarrow \int_{k_{\max}}^{\infty} p(k)\,\mathrm{d}k = \frac{1}{N}\end{gathered} \tag{89}$$

Resolvendo a integral, vamos obter:

$$k_{\max} = k_{\min} + \frac{\ln(N)}{\lambda} \tag{90}$$

Essa equação nos indica algo interessante. $\ln(N)$ é uma função que cresce devagar conforme $N \to \infty$, já que a sua derivada tende a 0, então quanto maior o N, mais devagar a função vai crescer. Ou seja, isso indica que, conforme o N cresce, o grau máximo e mínimo não diferem tanto!

Esse cálculo pra distribuição de Poisson é um pouquinho mais evoluído, mas a gente chega que o resultado é muito parecido e que N cresce mais lentamente ainda

Agora, para as redes livre de escala, resolvendo (89), a gente obtém:

$$k_{\max} = k_{\min} \cdot N^{\frac{1}{\gamma-1}} \tag{91}$$

Ou seja, quanto maior é minha rede, maior vai ser o tamanho do meu centro (Maior é o grau do nó com mais graus). Isso é um resultado bem intuitivo, na verdade! Lembra que nós começamos dando o contexto da rede da internet (WWW)? Se pararmos para pensar, conforme as pessoas criam páginas na internet, elas tendem a colocar links para páginas famosas na internet, ou que tem alguma relevância em **comunidades**, ou seja, quanto mais links referenciando uma página, mais páginas vão referenciar ela, de forma que, quanto mais páginas vão sendo criadas, maior vai ser a quantidade de links referenciando páginas famosas ou reconhecidas!

5.3 Significado de Livre de Escala

Antes de entender o significado desse termo, vamos nos familiarizar com alguns conceitos. Vimos em probabilidade o conceito de **momentos**. O n-ésimo momento da distribuição dos graus (Levando em conta a variável aleatória K que é o grau de um vértice aleatório) é:

$$\mathbb{E}[K^n] = \sum_{i=k_{\min}}^{\infty} k^n \cdot \mathbb{P}(K = k) = \int_{k_{\min}}^{\infty} k^n p(k) \, \mathrm{d}k \tag{92}$$

Resolvendo a integral, vamos obter:

$$\mathbb{E}[K^n] = C \frac{k_{\max}^{n-\gamma+1} - k_{\min}^{n-\gamma+1}}{n - \gamma + 1} \tag{93}$$

Sabemos que, normalmente, $k_{\min}$ é fixo enquanto $k_{\max}$ aumenta confirme $N \to \infty$. Então vamos fazer uma análise mais detalhada sobre essa fórmula para o n-ésimo momento

- Se $n - \gamma + 1 \leq 0$, então $k_{\max}^{n-\gamma+1} \to 0$ quando $N \to \infty$ (Ou 1 quando a equação é igual a 0). Então todos os momentos que satisfazem $n < \gamma - 1$ são **finitos**
- Do contrário, se $n - \gamma + 1 > 0$, então $k_{\max}^{n-\gamma+1} \to \infty$ quando $N \to \infty$, então os momentos que satisfazem $n > \gamma - 1$ **divergem**

Agora a gente pode tentar entender melhor o que esse **sem escala** significa. Vamos pegar uma rede de Poisson, sabemos que $\mathbb{E}[K] = \hat{k}$ e que $\sigma_k = \sqrt{\hat{k}}$ (Desvio padrão dos graus). Pela desigualdade de Chebyshev:

$$\mathbb{P}\Big(|K - \hat{k}| \geq h\sigma_k\Big) \leq \frac{1}{h^2} \tag{94}$$

Que que isso quer dizer? O que quero dizer é que, em redes de Poisson, a chance de os graus estárem a h desvios padrões da média é **no máximo** $1/h^2$. Isso é um indicativo grande de que a média dos graus serve como uma “escala”, de forma que temos uma noção do quão longe desse valor podemos estar caso escolhemos um nó aleatório.

Porém, em redes livres de escala em que o segundo momento diverge? Isso significa que, quando eu pego um nó aleatoriamente nessa rede, eu não sei o que esperar, a diferença dele para a média pode ser arbitrariamente grande ou pequena, não temos como ter ideia, ou seja, **não há uma escala para comparação**

É claro que a divergência de $\mathbb{E}[K^2]$ só acontece no limite $N \to \infty$, mas isso ainda tem uma relevância para redes finitas. Vamos pegar o caso da rede de internet novamente, sabemos que a quantidade de documentos (Nós) está na casa dos bilhões ou trilhões, o que indica que temos uma variância MUITO GRANDE, ou seja, mesmo tendo uma variância finita e, no concreto, tenhamos uma escala, ela é quase irrelevante, já que, ao pegarmos um documento aleatório, ele pode estar sendo citado por apenas dois outros documentos, ou ser citado por bilhões de documentos (Como google, facebook, etc.)

5.4 Propriedade *Ultra Small*

Essa propriedade dos centros faz levantar uma pergunta: Será que os centros afetam a propriedade dos minimundos? (Distância média). Se formos parar para tentar ter uma visão intuitiva, faz sentido dizer que elas afetam. Se eu tenho nós que se ligam em **MUITOS** outros nós (Os centros), então faz sentido dizer que a probabilidade de a distância entre dois outros nós quaisquer ser pequena é bem alta. Na verdade essa visão intuitiva está **correta**. As distâncias em uma rede **livre de escala** são menores do que em redes aleatórias equivalentes. Nós temos a seguinte relação: Seja D a variável aleatória que representa a distância entre dois nós aleatórios na rede

$$\mathbb{E}[D] = \begin{cases} \text{const} & \gamma = 2 \\ \ln(\ln(N)) & 2 < \gamma < 3 \\ \frac{\ln(N)}{\ln(\ln(N))} & \gamma = 3 \\ \ln(N) & \gamma > 3 \end{cases} \tag{95}$$

Vamos falar um pouco sobre cada um desses *regimes*

5.4.1 Regime Anômalo ($\gamma = 2$)

De acordo com a equação (91), quando $\gamma = 2$, o maior hub (Maior centro) vai crescer linearmente com relação a N, ou seja, o tamanho do caminho entre dois nós aleatórios não depende de N já que essa relação linear indica que todos os nós vão estar conectados ao mesmo hub central

5.4.2 Super minimundo ($2 < \gamma < 3$)

Nesse regmie, como previsto pela relação (95), a dsitância fica em relação a $\ln(\ln(N))$, que é um crescimento absurdamente lento comparado a $\ln(N)$ obtido em redes aleatórias. Essas redes são chamadas de **Ultra Small** por que os hubs reduzem o tamanho dos caminhos **muito**, já que eles se ligam com milhares de nós com baixo grau

5.4.3 Ponto Crítico ($\gamma = 3$)

Aqui o segundo momento já não diverge mais, então é um ponto teórico de bastante interesse. Aqui o termo $\ln(N)$ encontrado nas redes aleatórias volta, mas tem uma correção com $\ln(\ln(N))$ ainda

5.4.4 Minimundo ($\gamma > 3$)

Aqui o termo $\ln(N)$ volta! Isso mostra um indicativo que para essas redes, mesmo a presença de hubs ainda existindo, eles não são grandes o suficiente para influenciar na distância entre os nós, sendo desprezíveis praticamente ao afetarem a distância

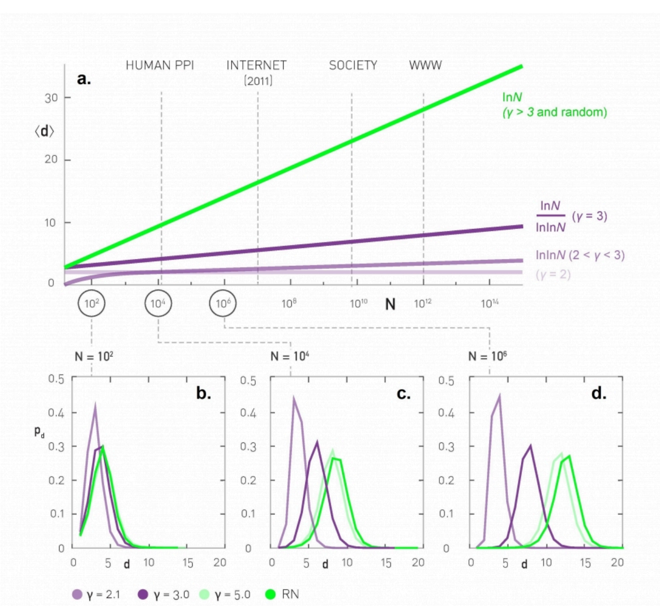


Figura 9: Distância média em função da quantidade de nós e distribuição das distâncias para $N = 10^2$, $N = 10^4$ e $N = 10^6$

Essa imagem presente no livro do barabasi mostra a prograssão das distâncias médias conforme aumentamos a quantidade de nós. Perceba que, para N não muito grandes, como $N = 10^4$, as distribuições (E a distância média) não tem tanta diferença assim, porém com $N = 10^6$, ja da para notar diferenças atenuadas. Isso também é um indicativo de que, quanto maior o expoente da rede livre de escala, maior é a distância média entre dois nós

5.5 O Papél do Expoente do Grau

Se pararmos para analisar redes na vida real, vamos perceber que γ varia de rede para rede, isso nos leva a intuitivamente querer saber como γ influencia nas redes reais. Na maioria das redes reais, temos que $\gamma > 2$, o que também gera a pergunta: Por quê?

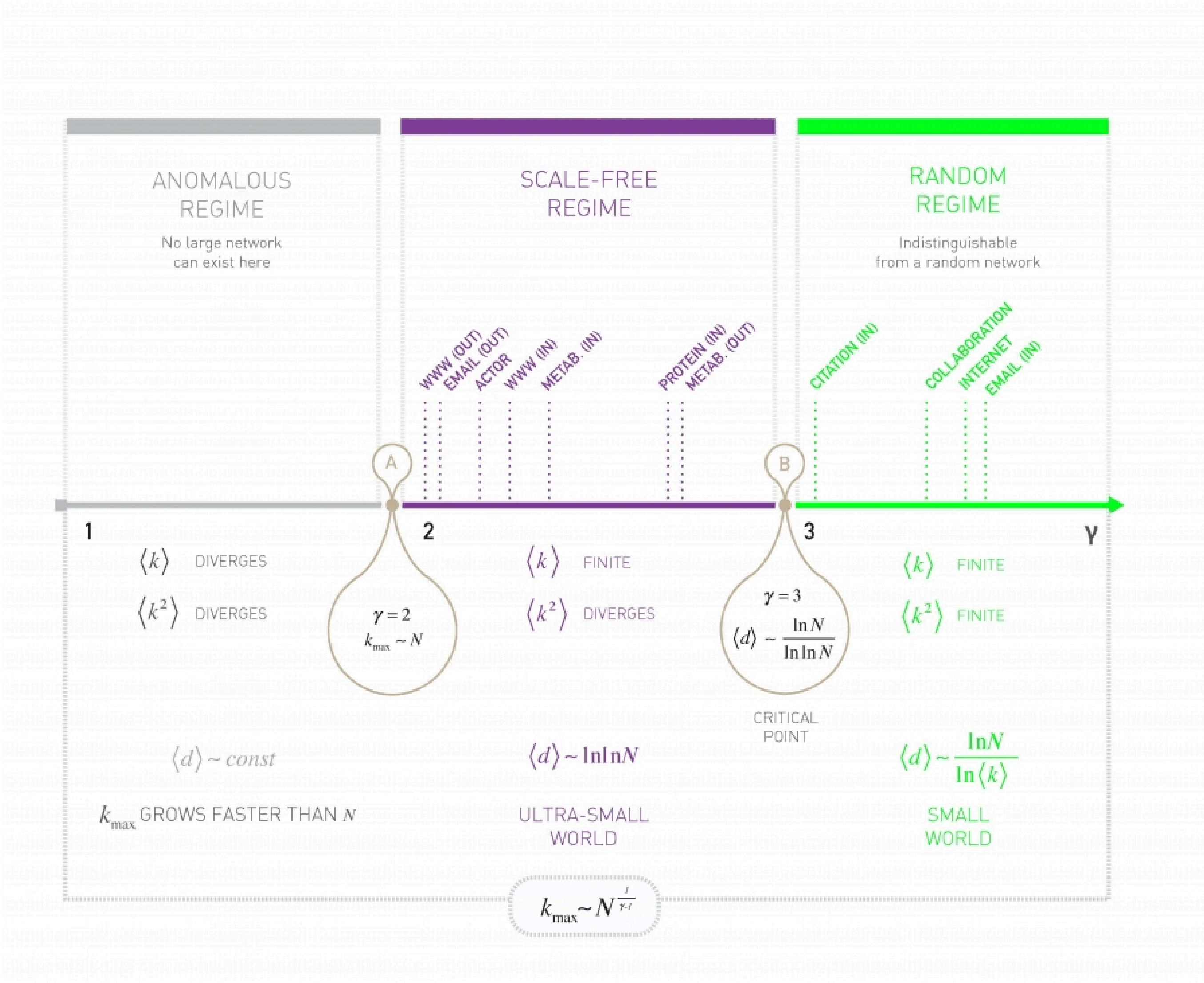


Figura 10: Regimes de γ

5.5.1 Regime Anômalo ($\gamma \leq 2$)

Nesse regime, o expoente $1/(\gamma - 1)$ é maior que 1, ou seja, o número de links conectados ao maior hub cresce **mais rápido que o próprio número de links em si**, além de que $\mathbb{E}[K]$, sendo K a variável aleatória do grau dos nós, também diverge. O que isso indica? Isso mostra que, redes livre de escala **sem links múltiplos**, ou seja, a existência de várias arestas que ligam os mesmos dois nós, **não podem existir**

5.5.2 Regime Livre de Escala ($2 < \gamma < 3$)

Aqui, o primeiro momento converge enquanto o segundo diverge, o que faz a gente cair na situação que ja comentei anteriormente de os graus serem arbitrariamente grandes, porém, que a distância entre dois nós cresce **muito** devagar, os já mencionado **Ultra Small Worlds**

5.5.3 Regime de Rede Aleatória ($\gamma > 3$)

Como indicado relação (95), e por motivos práticos também, nesse regime, as propriedades das redes livres de escala não são muito diferentes das propriedades das redes aleatórias. Isso pois, como ja comentado, o grau dos nós decaem rapido o suficiente para que os hubs, mesmo os maiores, não sejam tão numerosos ao ponto de que afetem muito a distância média entre os nós

Na prática, costuma-se observar que, para que os hubs venham a influenciar na distância média, $k_{\max}$ tem que ser, pelo menos, umas 10^2, 10^3 vezes maior que $k_{\min}$. Na prática a gente pode reformular a relação (91) como:

$$N = \left(\frac{k_{\max}}{k_{\min}}\right)^{\gamma-1} \tag{96}$$

E isso daria uma relação de quantos nós precisamos para que começássemos a registrar a propriedade da rede livre de escala. Por exemplo, vamos supor que queremos saber quantos nós precisamos para começar a ver essa propriedade em redes de $\gamma = 5$ (E, por exemplo, $k_{\min} = 1$ e $k_{\max} = 10^2$), então deveríamos ter $N > 10^8$, e são poucas as redes, na prática, com um tamanho absurdo desses!