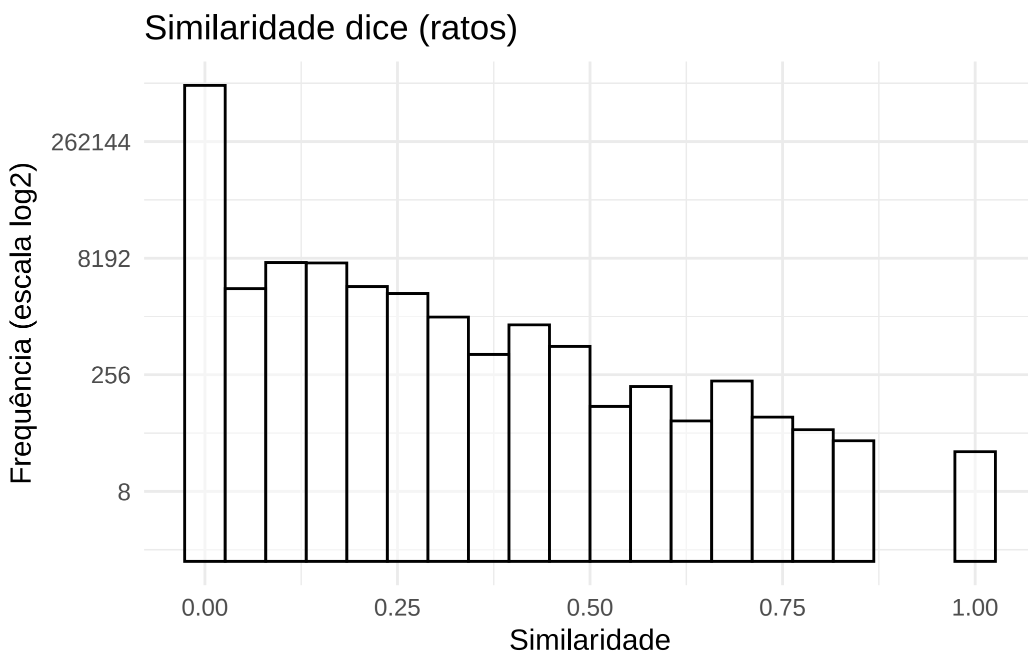
Lista 3 - Similaridade de Redes Complexas

Redes Complexas para Ciência da Computação

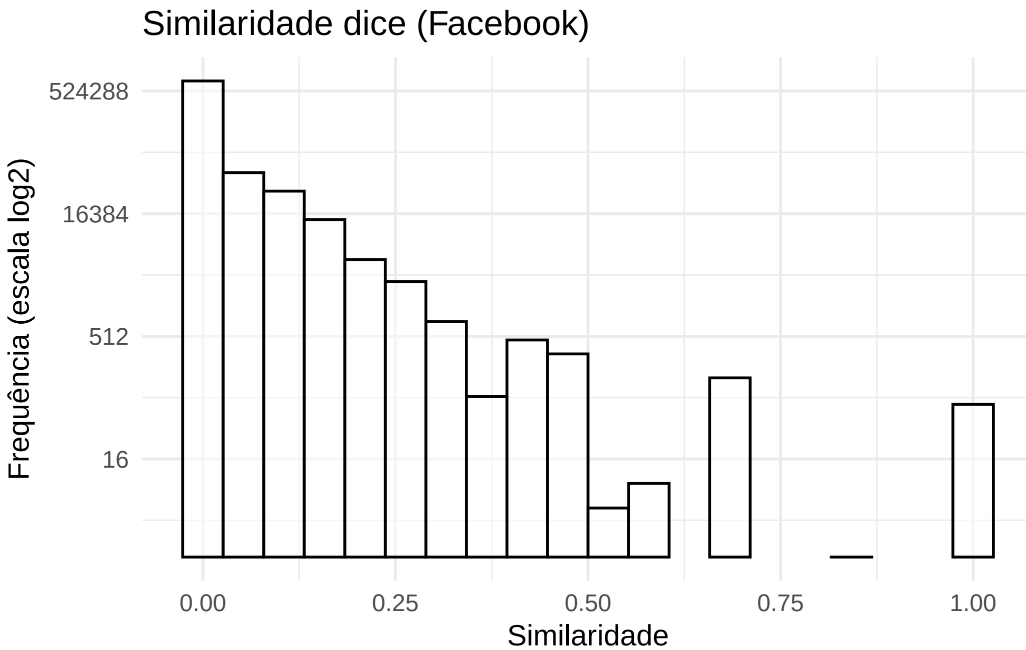
Luben Miguel, Rodrigo Lassance

Exercício 1

Considerando a medida de similaridade obtida a partir da distância euclidiana (similaridade dice), cuja fórmula é $S_{ij} = 2 \cdot \frac{n_{ij}}{(k_i + k_j)}$ para os vértices i e j, obtemos a seguir o histograma das similaridades considerando todos os pares de vértices para as duas redes de interesse: - Rede de ratos



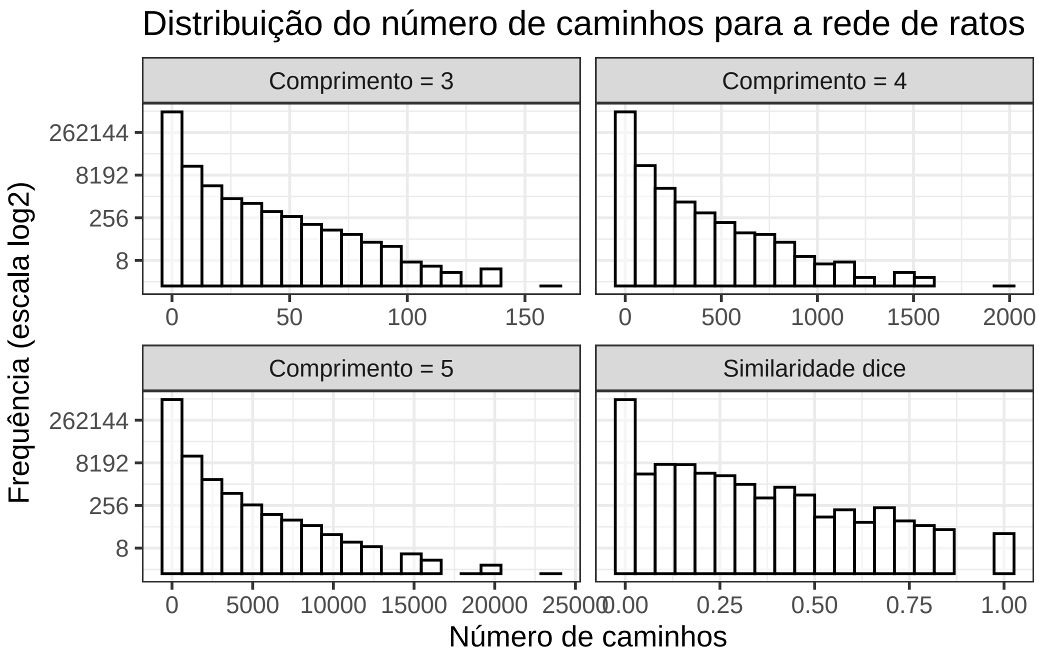
- Rede de mensagens do facebook



Exercício 2

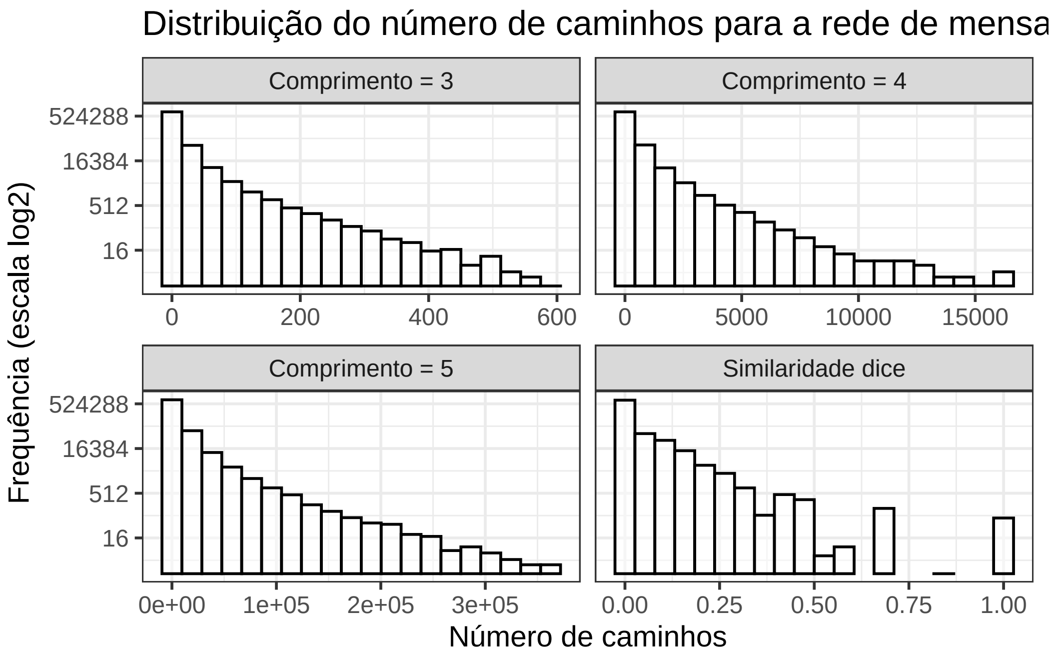
A seguir obtemos o número de caminhos de comprimento 3, 4 e 5 para cada par de vértice (i, j) para cada rede de interesse: - Rede de ratos

```
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
```



- Rede de mensagens no Facebook

```
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
```



Vemos que tanto para a rede de ratos quanto para a rede de mensagens a distribuição do número de caminhos é decrescente e suave para todos os comprimentos tomados sendo muito parecidas à distribuição da similaridade, tendo apenas certas diferenças na cauda. Isso pode estar relacionado ao fato de que, dois vértices muito similares tendem a ter um número maior de vizinhos compartilhados e por consequência, tem um número maior de possíveis caminhos a serem tomados nos diversos comprimentos. Por outro lado, para vértices não muito similares, o número de caminhos não será tão grande quanto para vértices mais similares pelo fato de terem vizinhos mais distantes e/ou poucos vizinhos compartilhados. Assim, é esperado que ocorra certo espelhamento da distribuição do número de caminhos de diferentes comprimentos na distribuição de similaridade entre vértices. Porém, nota-se como principal diferença entre as distribuições dos caminhos e da similaridade o decrescimento mais rapido do número de caminhos em comparação a um decrescimento mais lento da similaridade.

Exercício 3

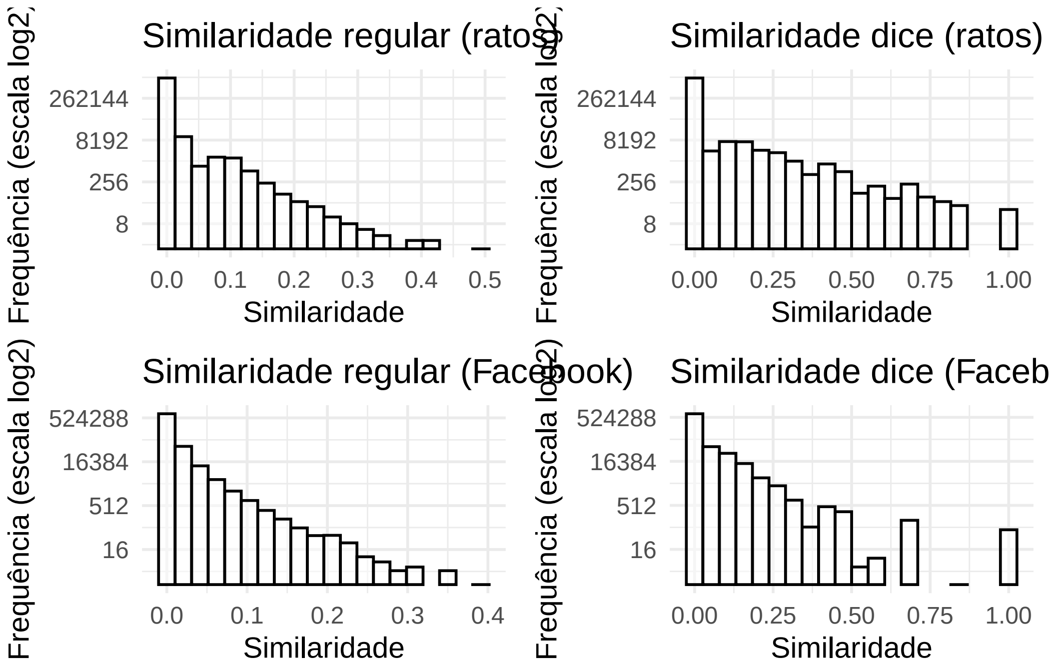
Para o calcular uma medida de similaridade regular, consideraremos nossa medida $\sigma = \sum_{m=1}^{\infty} (\alpha A)^m$. Para tal, tomaremos $\alpha < \frac{1}{\lambda_1}$ sendo λ_1 o maior autovalor de A. Os maiores autovalores para as redes de ratos e do facebook são:

Podemos tomar assim $\alpha_{ratos} = 0.07$ e $\alpha_{fb} = 0.035$ para cada rede e computar a matriz σ iterativamente. Tomamos como critério de convergência $\max |\sigma^{(n)} - \sigma^{(n-1)}| \leq \varepsilon$, fixando $\varepsilon = 10^{-6}$.

Tabela 1: Limiares para ambas redes

Rede	Limiares
Ratos	0.0825
Facebook	0.0377

```
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
<sparse>[ <logic> ] : .M.sub.i.logical() maybe inefficient
```



Notamos pelo gráfico, para a rede de ratos, que a distribuição da similaridade regular decresce mais rapidamente do que a regularidade estrutural. Ou seja, mesmo em casos que ambos vértices de interesse compartilham vizinhos entre si, os vizinhos de um dos vértices tendem a ser menos similares ao outro. O mesmo argumento pode ser aplicado para a rede de mensagens do Facebook, visto que novamente a similaridade regular tem maior descrescimento na distribuição do que a similaridade estrutural.

Exercício 4

Exercício 5

Consideraremos neste caso os pares de vértices $(400, 434)$ e $(675, 615)$ da rede de ratos. Para obter o primeiro tempo de passagem médio de i para j, consideraremos que a probabilidade de ir de um vértice para o outro é $1/k_i$, sendo i o vértice de saída e k_i seu grau. Simularemos $B = 1000$ caminhos do vértice i ao j e armazenaremos o tamanho de cada caminho realizado. Ao final, calcularemos o caminho/tempo médio de chegar ao vértice j partindo do vértice i. Depois, repetiremos o processo de j para i. A tabela abaixo resume os resultados:

Tabela 2: MFPT(i,j) e MFPT(j, i) para dois pares de vértices distintos

Vértice.i	Vértice.j	MFPT.i.j.	MFPT.j.i.
400	434	391.235	2937.692
675	615	1088.248	1830.095

Vemos que $MFPT(i, j) \neq MFPT(j, i)$ neste exemplo da rede de ratos. De fato, essa desigualdade é esperada. Em um exemplo em que i e j compartilham apenas um vizinho, porém o número de vizinhos de i é menor que o de j, é esperado que, caso não se passe para o vizinho compartilhado em um primeiro passo, o processo demorará muito mais para voltar a j do que voltar para i. Neste caso, visualiza-se isso pelo número médio de passos de j para i em comparação ao de i para j