

“There is no end to education.

It is not that you read a book, pass an examination, and finish with education.

The whole of life, from the moment you are born to the moment you die, is a process of learning.”

Jiddu Krishnamurti

Sumário

7	Glossário	iii
8	1 Introdução	1
9	2 Caracterização	3
10	2.1 Introdução	4
11	2.2 Métricas	7
12	2.2.1 Distribuição dos graus	7
13	2.2.2 Distância média	10
14	2.2.3 Diâmetro	14
15	2.2.4 Eficiência	15
16	2.2.5 Coeficiente de agrupamento	16
17	3 Modelos	19
18	3.1 Introdução	20
19	3.2 Topologias determinísticas	20
20	3.3 Grafos aleatórios	24
21	3.4 Mundo pequeno	26
22	3.5 Livre de escala	28
23	4 Aplicações	31
24	A Resultados complementares	33
25	A.1 Caracterização de redes determinísticas	33
26	A.2 Complexidade das métricas	41
27	Referências Bibliográficas	46
28	Índice Remissivo	47

Glossário

Acrônimos

31	BFS	<i>Breadth-First Search</i>
32	CDF	<i>Cumulative Distribution Function</i>
33	FIFO	<i>First-In First-Out</i>
34	PDF	<i>Probability Distribution Function</i>

Simbologia

C.Q.D. Demarcador contração de ‘como se queria demonstrar’.

□ Demarca fim de Algoritmos, Definições, Teoremas, dentre outros.

Representações

x Letras minúsculas em negrito indicam vetores coluna. É possível parametrizar o vetor, por exemplo, $\mathbf{x}(t) = [x_1(t) \ \cdots \ x_n(t)]^T$ indica que o vetor **x** é variante no tempo.

X Letras maiúsculas em negrito indicam matrizes. Assim como é possível parametrizar vetores, o mesmo é possível com matrizes, por exemplo, uma matriz variante no tempo pode ser representada por **X**(t).

\mathcal{X} Letras maiúsculas caligráficas representam variáveis aleatórias.

$\dot{x}(t)$ Indica a derivada da função $x(\cdot)$ em relação ao tempo t . Também se aplica a funcionais em vetores e matrizes.

$n!$ Operador fatorial, definido recursivamente como $n! = n(n-1)!$ e com caso base $0! = 1$. De forma iterativa também pode ser descrito como

$$n! = \prod_{i=0}^{n-2} (n-i),$$

para $n \geq 2$.

$\binom{n}{k}$ Coeficiente binomial de n dado k , onde $0 \leq k \leq n$, definido como

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

que pode ser computado de forma eficiente utilizando

$$\binom{n}{k} = \prod_{i=1}^k \frac{n - (k - i)}{i},$$

que possui complexidade $\Theta(k)$.

$\delta(t), \delta_{ij}$ A função delta de Kronecker, definida como

$$\delta_{ij} \triangleq \begin{cases} 1 & \text{se } i = j \\ 0 & \text{caso contrário} \end{cases},$$

utilizada como contrapartida discreta da função delta de Dirac. Por conveniência, é possível usar a seguinte representação

$$\delta(t) \triangleq \begin{cases} 1 & \text{se } t = 0 \\ 0 & \text{caso contrário} \end{cases}.$$

Dessa forma temos de forma equivalente que o valor $\delta(i - j)$ é 1 se $i = j$ e 0 caso contrário.

H_n Indica a soma dos n primeiros termos da série harmônica, representada por

$$H_n = \sum_{i=1}^n \frac{1}{i},$$

que diverge no limite quando $n \rightarrow \infty$. Porém, possui a seguinte propriedade assintótica

$$\lim_{n \rightarrow \infty} H_n - \log(n) = \gamma,$$

onde $\gamma \approx 0.57721$ representa a constante de Euler-Mascheroni. Portanto, é possível usar a seguinte igualdade assintótica

$$H_n \simeq \log(n) + \gamma,$$

onde o logaritmo natural é o da base natural e .

$\{x : p(x)\}$ Descrição do conjunto representado pelos elementos x que têm a propriedade, ou predicado, $p(x)$. Adicionalmente, o predicado $p(x)$ pode ser descrito utilizando os operadores da lógica proposicional.

- $(\forall x)(p(x)) \dots$ Quantificação universal em relação aos elementos x que têm a propriedade, ou predicado, $p(x)$. A pertinência dos elementos representados por x também pode ser descrita de forma explícita, por exemplo, $(\forall x \in \mathbb{N})(p(x))$. Que expressa que todos os elementos do conjunto dos números naturais possuem o predicado p . Adicionalmente, o predicado $p(x)$ pode ser descrito utilizando os operadores da lógica proposicional.
- $(\exists x)(p(x)) \dots$ Quantificação existencial em relação aos elementos x que têm a propriedade, ou predicado, $p(x)$. A pertinência dos elementos representados por x também pode ser descrita de forma explícita, por exemplo, $(\exists x \in \mathbb{N})(p(x))$. Que expressa que existe pelo menos um número natural que possui o predicado p . Adicionalmente, o predicado $p(x)$ pode ser descrito utilizando os operadores da lógica proposicional.

Notação assintótica

- $O(\cdot) \dots \dots \dots$ Quando é expresso que $f(n) \in O(g(n))$ ^[i], diz-se que existe uma constante k , tal que a função $f(n)$, para todo valor de $n > n_0$, é sempre limitada superiormente por $kg(n)$.
- $\Omega(\cdot) \dots \dots \dots$ Quando é expresso que $f(n) \in \Omega(g(n))$, diz-se que existe uma constante k , tal que a função $f(n)$, para todo valor de $n > n_0$, é sempre limitada inferiormente por $kg(n)$.
- $\Theta(\cdot) \dots \dots \dots$ Quando é expresso que $f(n) \in \Theta(g(n))$, diz-se que existe uma constante k_1 , tal que a função $f(n)$, para todo valor de $n > n_0$, é sempre limitada inferiormente por $k_1g(n)$, e também existe uma outra constante k_2 , tal que a função $f(n)$, para todo valor de $n > n_0$, é sempre limitada superiormente por $k_2g(n)$. De forma equivalente, define-se que $f(n) \in \Theta(g(n))$ se e somente se

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c,$$

para $g(n)$ diferente de zero ou, pelo menos, sempre maior de que zero a partir de algum ponto e para $0 < c < \infty$.

Igualdades matemáticas

- $\approx \dots \dots \dots$ Valor aproximado.

^[i] Utiliza-se o símbolo de pertinência \in pois interpreta-se que o operador $O(\cdot)$ representa o conjunto das funções que são limitadas superiormente pelo seu argumento, no caso a função $g(\cdot)$. O mesmo princípio pode ser aplicada aos outros operadores assintóticos apresentados em sequência.

\simeq Igualdade assintótica, isto é, se $f(n) \simeq g(n)$ então

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1,$$

para $g(\cdot)$ infinitamente diferente de zero.

\propto Proporcionalidade, isto é, se $f(n) \propto g(n)$, então existe uma constante k tal que $f(n) = kg(n)$. De forma generalista, pode considerar também a igualdade assintótica.

\triangleq Igualdade por definição, por exemplo,

$$\frac{d\mathbf{x}(t)}{dt} \triangleq \left[\frac{dx_1(t)}{dt} \quad \dots \quad \frac{dx_n(t)}{dt} \right]^\top,$$

onde $\mathbf{x}(t)$ é um vetor coluna.

\equiv Equivalência, por exemplo, $x \equiv y$ significa que x é definido como sendo logicamente igual à y .

Notação estatística

\sim Indicador de distribuição de probabilidade, por exemplo $\mathcal{X} \sim N(\mu, \sigma)$ indica que a variável aleatória \mathcal{X} segue uma distribuição de probabilidade normal com média μ e desvio padrão σ .

\mathcal{X}_ζ Resultado ou realização ζ da variável aleatória \mathcal{X} .

$P(\mathcal{X}_\zeta)$ Probabilidade da variável aleatória \mathcal{X} assumir a realização ζ .

$P(\mathcal{X}_\zeta | p)$ Probabilidade da variável aleatória \mathcal{X} assumir a realização ζ dado que o predicado p é verdadeiro.

$E\{\mathcal{X}\}$ Valor esperado da variável aleatória \mathcal{X} . No caso discreto é definido como

$$E\{\mathcal{X}\} = \sum_{\{\zeta \in \mathcal{U}\}} \mathcal{X}_\zeta P(\mathcal{X}_\zeta),$$

onde \mathcal{U} é o conjunto de possíveis realizações da variável aleatória.

$E\{\mathcal{X} | p\}$ Valor esperado da variável aleatória \mathcal{X} dado que o predicado p é verdadeiro. No caso discreto é definido como

$$E\{\mathcal{X}\} = \sum_{\{\zeta \in \mathcal{U}\}} \mathcal{X}_\zeta P(\mathcal{X}_\zeta | p),$$

onde \mathcal{U} é o conjunto de possíveis realizações da variável aleatória.

Operadores matemáticos

$ \cdot $	Se for aplicado a um escalar, indica o seu valor absoluto. Caso seja aplicado a um conjunto, indica sua cardinalidade.
$\lfloor \cdot \rfloor$	O maior valor inteiro menor ou igual ao escalar.
$\lceil \cdot \rceil$	O menor valor inteiro maior ou igual ao escalar.
$\rho(\cdot)$	Posto de uma matriz, por exemplo dada uma matriz identidade $\mathbf{I}_{n \times n}$, $\rho(\mathbf{I}) = n$.
\mathbf{X}^\top	Operação de transposição da matriz \mathbf{X} , isto é, troca dos elementos x_{ij} pelos elementos x_{ji} . Também pode ser aplicada a vetores, no qual transforma vetores coluna em vetores linha, e vice-versa.
$X - Y$	Subtração de elementos de conjuntos. Utilizando a notação de conjuntos pode ser definido por

$$X - Y \triangleq \{z : (z \in X) \wedge (z \notin Y)\},$$

que representa o conjunto resultante da retirada dos elementos em X que também estão em Y .

$X \times Y$	Produto cartesiano entre dois conjuntos X e Y . Utilizando a notação de conjuntos pode ser definido por
--------------------	-------------------------------------------------------------------------------------------------------------

$$X \times Y \triangleq \{(x, y) : (x \in X) \wedge (y \in Y)\},$$

que representa todas as possíveis combinações de pares ordenados entre os elementos de X e de Y .

Operadores lógicos

\neg	Operador unário de negação.
\vee	Operador binário de disjunção, definido como ‘ou inclusivo’.
\wedge	Operador binário de conjunção, definido com valor lógico ‘e’.
\Rightarrow	Operador binário de implicação, por exemplo, $(a \Rightarrow b)$, onde a é denominado antecedente e b consequente. Único operador binário não comutativo.
\Leftrightarrow	Operador binário de bi-implicação. Onde $(a \Leftrightarrow b)$ é logicamente equivalente a representação $((a \Rightarrow b) \wedge (b \Rightarrow a))$.

36 **Capítulo 1**

37 **Introdução**

38 *“If knowledge can create problems,
it is not through ignorance that we can solve them.”*
Isaac Asimov

39 Paragrafo introdutório.
40 Este Capítulo está organizado da seguinte forma...

Capítulo 2

Caracterização

*“Anyone who doesn’t take truth seriously in small matters
cannot be trusted in large ones either.”
Albert Einstein*

O estudo de redes através de sua descrição por meio da Teoria de Grafos é uma abordagem clássica comumente encontrada na literatura. São muitos os casos de algoritmos para solução de problemas em grafos que foram e são utilizados no projeto de redes de computadores, por exemplo, os algoritmos de [Bellman \(1958\)](#) e [Ford e Fulkerson \(1962\)](#) e de [Dijkstra \(1959\)](#), ambos utilizados no projeto e estudo de protocolos de roteamento. Neste Apêndice, estuda-se como é possível representar computacional e matematicamente redes complexas. Especificamente, tem-se como objetivo possibilitar o desenvolvimento de respostas para as seguintes questões: (i) qual tipo de estrutura de dados pode ser utilizada a fim de reduzir a quantidade de dados e diminuir a complexidade de algumas operações sobre grafos e (ii) quais métricas podem ser utilizadas para representar propriedades do grafo. A partir das definições das métricas encontrada na literatura são desenvolvidos algoritmos para que seja possível analisar a complexidade em tempo e memória esperada na utilização de cada métrica.

Este Apêndice está organizado da seguinte forma: na Seção [2.1](#), é feita uma introdução acerca dos requisitos necessários para modelar redes e são descritos os tipos de grafos, suas propriedades e formas de representação. Finalmente, as métricas que podem ser utilizadas para caracterizar grafos são apresentadas na Seção [2.2](#).

2.1 Introdução

Antes de tratar de propriedades e dos modelos de redes complexas, é conveniente definir a descrição matemática e computacional desses sistemas [Costa et al. \(2007\)](#).

Definição 2.1 (Grafo direcionado com pesos). Um grafo direcionado com pesos G é composto por uma tripla ordenada $G = \langle N, E, \omega \rangle$, onde N representa o conjunto de vértices (ou nós) do grafo e E o conjunto de arestas ao qual se atribui as seguintes propriedades: (i) cada aresta é composta por um par ordenado de nós (v_1, v_2) , que indica que existe uma ligação saindo do nó v_1 em direção ao nó v_2 e (ii) para cada aresta $e \in E$ existe um peso que é associado por uma função $\omega(\cdot)$, que realiza o mapeamento dos pesos de cada aresta para um número real, ou seja, $\omega: E \mapsto \mathbb{R}$. \square

Na Definição 2.1 trata-se do tipo mais generalista de grafo, denominado Tipo 1, onde as arestas são direcionadas e possuem um peso associado. Além desse, pode-se definir mais outros três tipos de grafos, fazendo com que se alterne as propriedades de direção e peso das arestas, como apresentado na Figura 2.1^[i].

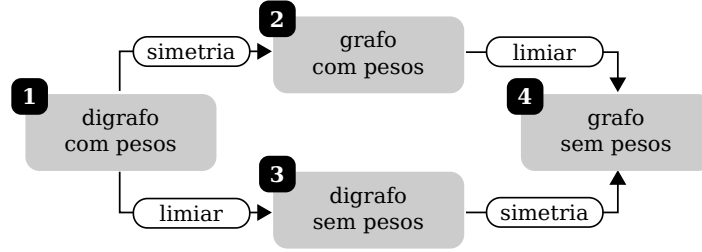


Figura 2.1: Tipos de representação de redes complexas: (1) representação mais generalista de um grafo, onde as arestas são direcionadas e com pesos associados; (2) tipo de grafo onde as arestas são simétricas, ou seja, dada uma aresta $(v_1, v_2) \in E$ tem-se que é possível sair de v_1 em direção à v_2 e vice-versa; (3) onde não há distinção de pesos entre as arestas, portanto, não há necessidade de se definir a função $\omega(\cdot)$ e (4) tipo de grafo simplista, onde as arestas não são direcionadas nem possuem pesos associados. Figura adaptada de [Costa et al. \(2007\)](#).

Além do grafo direcionado com pesos (Tipo 1 na Figura 2.1), existem outros modelos de grafos que podem ser utilizados para representação de redes complexas. Porém, antes de descrever os outros tipos de grafos, é conveniente definir como é construída a matriz de adjacência de cada tipo.

Definição 2.2 (Matriz de adjacência). Um grafo $G = \langle N, E, \omega \rangle$ pode ser representado por uma matriz adjacência \mathbf{A} , que é construída da seguinte forma

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \text{ onde } a_{ij} \triangleq \begin{cases} \omega(e) & \text{se } (v_j, v_i) \in E \text{ (Tipo 1 e 2)} \\ 1 & \text{se } (v_j, v_i) \in E \text{ (Tipo 3 e 4)} \\ 0 & \text{caso contrário} \end{cases}, \quad (2.1)$$

onde $e = (v_j, v_i)$ e para grafos não direcionados dos Tipos 2 e 4 tem-se que e é tratado como um par não ordenado, isto é, $(v_j, v_i) \equiv (v_i, v_j)$. \square

^[i] Neste trabalho, considera-se a inexistência de laços, isto é, arestas cuja origem e destino são o mesmo nó, e de arestas múltiplas, onde tem-se mais de uma aresta com mesma origem e destino.

Em acordo com o descrito na Definição 2.2, para grafos do Tipo 2, a função $\omega(\cdot)$ é definida de forma a ser simétrica, ou seja, se $e_{ij} = (v_i, v_j)$ pertence ao conjunto de arestas E e $e_{ji} = (v_j, v_i)$ não pertence à E , então $\omega(e_{ji})$ é definido como sendo igual à $\omega(e_{ij})$. De forma semelhante, para grafos do Tipo 4, se e somente se, (v_i, v_j) ou (v_j, v_i) pertencem à E , então $a_{ij} = 1$, caso contrário $a_{ij} = 0$. Isso porque, em grafos do Tipo 2, denominados grafos não direcionados com pesos, as arestas são simétricas, ou seja, dada uma aresta $(v_1, v_2) \in E$ tem-se que é possível sair de v_1 em direção a v_2 e vice-versa.

Apesar de bastante útil no desenvolvimento de modelos e na análise de propriedades de redes complexas, a representação por matriz de adjacência nem sempre é a mais adequada para representação computacional da rede. Isso porque a representação em matriz de adjacência de uma rede com n nós sempre ocupa memória em uma ordem de $\Theta(n^2)$, ou seja, quadrática. O que não é escalável em termos de armazenamento da representação do sistema. É conveniente a busca por outro tipo de representação, principalmente quando a matriz é esparsa. Dado que o desempenho da representação das conexões de um grafo está relacionada à sua quantidade de arestas, denotada por $|E|$, e que a quantidade máxima de arestas em um grafo direcionado é de exatamente $|E|_{\max} = n(n-1)$, é possível pensar em uma estrutura de representação mais escalável. Para tanto, é conveniente definir os conceitos de grau de um vértice e de densidade do grafo.

Definição 2.3 (Grau). Em um grafo direcionado com pesos, Tipo 1, o grau de um nó v_i , denotado por $g(v_i)$, representa a quantidade de arestas associadas ao nó v_i . Em grafos direcionados, essa associação pode considerar a quantidade de nós que incidem em v_i ou a quantidade de nós que v_i incide. O primeiro caso, denominado grau de entrada, é quantificado por

$$g_{\text{in}}(v_i) = |\{v_j \in N : (v_j, v_i) \in E\}|, \quad (2.2)$$

e o segundo caso, denominado grau de saída, por

$$g_{\text{out}}(v_i) = |\{v_j \in N : (v_i, v_j) \in E\}|, \quad (2.3)$$

onde considera-se, em ambos os casos, a aresta como um par ordenado. \square

Para o caso de grafos não direcionados, Tipos 2 e 4, o grau de entrada e de saída são iguais já que o par que representa uma aresta é não ordenado. O grau do nó de um grafo não direcionado é quantificado por

$$g(v_i) = |\{v_j \in N : ((v_j, v_i) \in E) \vee ((v_i, v_j) \in E)\}|, \quad (2.4)$$

onde a disjunção é considerada a fim de tratar cada aresta como um par não ordenado.

Definição 2.4 (Densidade). A densidade de um grafo G , do Tipo 1, denotada por $m(G)$, representa a quantidade relativa do número de arestas do grafo, denotado por $|E|$, em relação ao seu valor máximo. Portanto, tem-se que

$$m(G) = \frac{|E|}{|E|_{\max}}, \quad (2.5)$$

onde, para grafos direcionados, $|E|_{\max} = n(n-1)$, sendo que $n = |N|$. \square

De forma semelhante, é possível definir a densidade de grafos não direcionados, uma vez que sua quantidade máxima de arestas é metade da quantidade máxima em grafos direcionados. A densidade em grafos não direcionados é igual à $2m(G)$.

Se consideramos que um dado grafo G é conexo^[ii], isto é, sempre existe pelo menos uma trajetória entre qualquer par de nós do grafo, então a quantidade mínima de arestas será de, pelo menos, $(n - 1)$. Portanto, considerando grafos conexos do Tipo 1, tem-se que $1/n \leq m(G) \leq 1$. É possível reduzir a quantidade de memória para representação das $n(n - 1)$ possíveis arestas em um fator de $1/n$. A lista de adjacência, ilustrada na Figura 2.2, é uma estrutura que possibilita essa otimização Cormen et al. (2009).

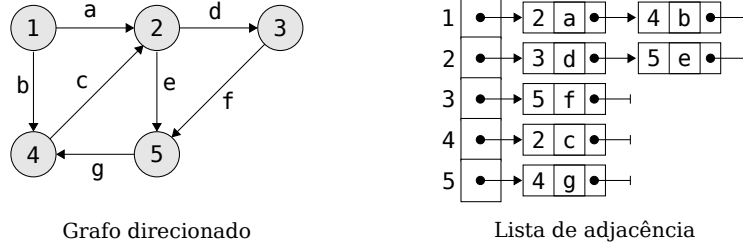


Figura 2.2: Representação de redes por meio de lista de adjacência composta por n listas encadeadas. Cada lista é associada a um nó da rede e armazenam os nós ligados ao nó associado e o respectivo peso da aresta. A quantidade de memória necessária é da ordem de $\Theta(n E\{\mathcal{G}^{\text{out}}\})$.

A representação de redes por meio de lista de adjacência é composta por n listas encadeadas, em que cada lista é associada a um nó da rede. A lista encadeada de um nó v armazena os nós em que v incide e, no caso de grafo com pesos, o peso da respectiva aresta. Obviamente, a quantidade de memória necessária para armazenar essa estrutura depende do tamanho médio das listas. O tamanho de cada uma dessas listas é igual ao grau de saída do respectivo nó, o que implica na seguinte Observação.

Observação 2.1 (Complexidade em memória da lista de adjacência). Utilizando \mathcal{G}^{out} , como sendo a variável aleatória que representa o grau de saída dos nós da rede, tem-se que o tamanho esperado da lista encadeada é dado por $E\{\mathcal{G}^{\text{out}}\}$. A quantidade de memória esperada na representação por lista de adjacência é dada por

$$E\{\mathcal{M}^{\text{list}}\} \in \Theta(n E\{\mathcal{G}^{\text{out}}\}), \quad (2.6)$$

onde $\mathcal{M}^{\text{list}}$ é a variável aleatória que representa a quantidade de memória necessária para armazenar uma lista de adjacência. Considerando o caso de densidade mínima, o valor esperado do grau de saída é unitário e, nesse caso, a quantidade de memória necessária é de ordem linear. A quantidade esperada de memória necessária $E\{\mathcal{M}^{\text{list}}\}$ depende da densidade do grafo. Considerando a densidade mínima e máxima de grafos conexos do Tipo 1, a quantidade de memória é da ordem de $\Omega(n)$ e $O(n^2)$. \square

Para dissertar com mais propriedade sobre a escalabilidade da representação por lista de adjacência, é necessário considerar a natureza de algumas propriedades da rede. Por exemplo, o valor esperado do grau de saída de cada nó. Mesmo assim, considerando o exposto na Observação 2.1, de forma geral, pode-se afirmar que a representação por lista de adjacência é mais escalável que a representação por matriz de adjacência. Mesmo que para alguns algoritmos a representação por matriz de adjacência seja mais eficiente, neste trabalho, considerando a questão da escalabilidade, em todas as análises de eficiência é considerada a representação por lista de adjacência.

^[ii] Considera-se a propriedade do grafo ser conexo porque, a princípio, não faz sentido, no estudo de redes, que existam nós desligados da topologia.

2.2 Métricas

A caracterização de redes complexas pode ser feita por meio da utilização de diferentes métricas. Cada métrica, isoladamente, permite caracterizar a rede em um aspecto bem definido. O estudo conjunto dessas características possibilita um entendimento mais abrangente sobre as propriedades da rede. Por exemplo, ao se criar um conjunto de caracterizações da rede, antes e após a aplicação de uma transformação $T(\cdot)$, denominados $\mathbf{c}^{(1)}$ e $\mathbf{c}^{(2)}$, respectivamente, é possível quantificar sua robustez em relação à transformação em questão. Esse processo é ilustrado pela Figura 2.3.

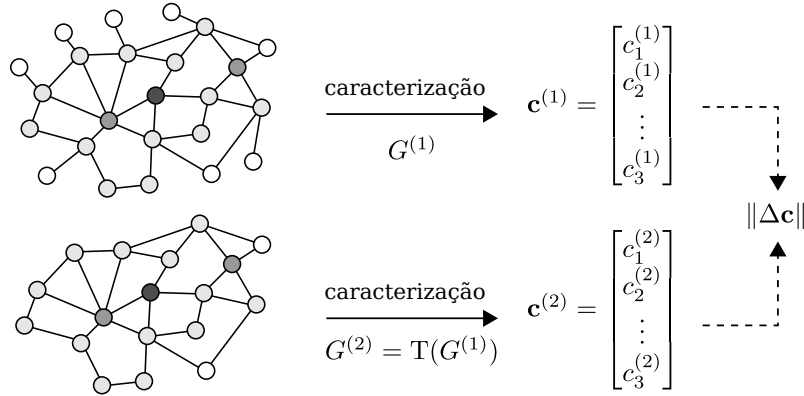


Figura 2.3: Caracterização, transformação e quantificação da robustez de grafos. Quanto menor for $\|\Delta \mathbf{c}\|$, mais robusto é o grafo G a transformação $T(\cdot)$. No caso, a transformação $T(\cdot)$ retira do grafo os nós com grau igual à 1. Figura adaptada de [Costa et al. \(2007\)](#).

Se a norma da diferença entre $\mathbf{c}^{(1)}$ e $\mathbf{c}^{(2)}$, denominada $\|\Delta \mathbf{c}\|$, for significativa, então a rede em questão não é robusta à transformação $T(\cdot)$. No caso da Figura 2.3, a transformação $T(\cdot)$ retira do grafo os nós com grau igual à 1. Se a topologia em questão se tratar de uma rede de computadores e alguma métrica que representasse a capacidade de roteamento não sofresse alterações significativas após a transformação $T(\cdot)$, poderia-se afirmar, por exemplo, que a existência de nós finais não influencia significativamente a capacidade de roteamento da rede. Portanto, é conveniente a definição de métricas que podem ser utilizadas no estudo de redes complexas. A seguir, são descritas algumas dessas métricas, que estão relacionadas com os objetivos deste trabalho^[iii].

2.2.1 Distribuição dos graus

A distribuição dos graus de uma rede é representada por uma função de distribuição de probabilidade [Papoulis e Pillai \(2002\)](#) (ou PDF, do inglês *Probability Distribution Function*), denominada $f(\cdot)$ e indexada por k , que caracteriza a probabilidade de um nó escolhido de forma aleatória possuir exatamente grau k . Para dados experimentais, a função $f(\cdot)$ é discreta, definida para valores de $k \in \mathbb{N}$, e igual à razão entre o número de eventos em que o grau do nó é igual à k e o número total de nós n . Para grafos direcionados, existem duas funções de distribuição, uma relacionada aos graus de entrada e outra aos de saída. A distribuição dos graus de uma rede é definida a seguir.

^[iii] As definições das métricas apresentadas têm como base [Newman \(2003\)](#) e [Costa et al. \(2007\)](#).

Definição 2.5 (Distribuição dos graus da rede). Para dados experimentais, a distribuição dos graus da rede é expressa pela função de densidade de probabilidade que representa a frequência relativa dos graus. Assim, como existem dois tipos de grau, o de entrada e o de saída, são duas as distribuições desses graus. Para cada um deles, tem-se a seguinte forma geral da função de distribuição de probabilidade para grafos direcionados

$$f(k) = \frac{|\{v_i \in N : g(v_i) = k\}|}{|E|}, \quad (2.7)$$

onde $g(v_i)$ é o grau do nó v_i e $|E|$ representa a quantidade de arestas do grafo. Para se obter $f_{\text{in}}(k)$, PDF do grau de entrada, utiliza-se na Equação 2.7 a função $g_{\text{in}}(v_i)$ no lugar de $g(v_i)$. Já no caso de $f_{\text{out}}(k)$, PDF do grau de saída, utiliza-se a função $g_{\text{out}}(v_i)$. \square

No cálculo da função de distribuição de probabilidade, a divisão por $|E|$ justifica-se pelo fato de que o resultado somatório de $f(\cdot)$, considerando todos os possíveis valores de k , deve ser unitário. A probabilidade de um nó, escolhido de forma aleatória ter grau k , é definida por

$$P(\mathcal{G}_k) = f(k), \quad (2.8)$$

onde \mathcal{G} é a variável aleatória que representa o grau de um nó qualquer. De forma geral, tanto a Equação 2.7 quanto a 2.8 também podem ser utilizadas para se obter estatísticas de grafos não direcionados dos Tipos 2 e 4. Para isso, é necessário considerar que a função $g(\cdot)$ utilizada é aquela definida pela Equação 2.4. Para considerar a utilização da função de distribuição de graus para fins de estudo da rede, resta saber se a função $f(\cdot)$ pode ser computada de forma eficiente e escalável. A seguir, é apresentado um algoritmo para cálculo dos graus de entrada e saída de cada nó da rede.

Algoritmo 2.1 (Cálculo dos graus de entrada e saída de cada nó). É possível calcular os graus de entrada e saída de cada nó da rede de forma iterativa com base na representação por lista de adjacência.

algoritmo graus(L)

- 1: {Lista de adjacência L de um grafo direcionado $G = \langle N, E \rangle$.}
- 2: $g_{\text{in}} \leftarrow \text{novovetor}(|N|, 0)$ {Vetor de $|N|$ posições preenchidas com zero.}
- 3: $g_{\text{out}} \leftarrow \text{novovetor}(|N|, 0)$
- 4: **para** i de 1 até $|N|$ **faça**
- 5: **para cada** $(v_j, p) \in L[i]$ **faça**
- 6: {Nó adjacente v_j e peso p da aresta.}
- 7: $g_{\text{out}}[i] \leftarrow g_{\text{out}}[i] + 1$
- 8: $g_{\text{in}}[j] \leftarrow g_{\text{in}}[j] + 1$
- 9: **fim para**
- 10: **fim para**
- 11: **retorne** $\langle g_{\text{in}}, g_{\text{out}} \rangle$ {Vetores com os graus de entrada e saída de cada nó da rede.}

Considera-se que os vetores g_{in} e g_{out} são indexados a partir de 1. A complexidade do algoritmo é da ordem de $\Theta(n E\{\mathcal{G}^{\text{out}}\})$ em tempo e $\Theta(n)$ em memória. \square

No Algoritmo 2.1, considerando os laços de repetição das linhas 4 e 5, tem-se que a complexidade em tempo do algoritmo é proporcional ao produto entre a quantidade de nós n e o tamanho esperado das listas na representação por lista de adjacência. Como descrito na Observação 2.1, o tamanho esperado dessas listas é equivalente ao valor esperado do grau de saída dos nós $E\{\mathcal{G}^{\text{out}}\}$. O tempo esperado de execução do Algoritmo 2.1 é da ordem

216 de $\Theta(n E\{\mathcal{G}^{\text{out}}\})$. Considerando que, no pior caso, o valor máximo de $E\{\mathcal{G}^{\text{out}}\}$ é de $(n-1)$,
 217 o tempo esperado de execução é de ordem $\Theta(n^2)$, ou seja, quadrática. Já no melhor caso,
 218 $E\{\mathcal{G}^{\text{out}}\}$ não depende de n e representa um valor constante^[iv]. Esta configuração torna o
 219 tempo esperado de execução do algoritmo de ordem $\Theta(n)$, ou seja, linear. Em relação à
 220 escalabilidade, o Algoritmo 2.1 necessita de uma quantidade de memória de ordem linear
 221 em relação à quantidade de nós. Especificamente, dois vetores de tamanho n , para o caso
 222 de grafos direcionados, ou apenas um vetor do mesmo tamanho, para o caso de grafos não
 223 direcionados. Com base no Algoritmo 2.1, é possível construir a função de distribuição
 224 de probabilidade dos graus. É apresentado no Algoritmo a seguir os procedimentos para
 225 computação da função $f(\cdot)$, tanto para o grau de entrada quanto o de saída.

226 **Algoritmo 2.2** (Construção da função de distribuição dos graus). É possível construir
 227 a função de distribuição de propriedade dos graus da rede de forma iterativa com base na
 228 representação por lista de adjacência.

229 **algoritmo** distribuicao-graus(L)
 230 1: {Lista de adjacência L de um grafo direcionado $G = \langle N, E \rangle$.}
 231 2: $\langle g_{\text{in}}, g_{\text{out}} \rangle \leftarrow \text{graus}(L)$ {Computação dos graus de entrada e saída de cada nó.}
 232 3: $f_{\text{in}} \leftarrow \text{novo-vetor}(|N|, 0)$
 233 4: $f_{\text{out}} \leftarrow \text{novo-vetor}(|N|, 0)$
 234 5: **para** i de 1 até $|N|$ **faça**
 235 6: $g \leftarrow g_{\text{in}}[i]$ {Utilizando grau como índice no vetor de distribuição.}
 236 7: $f_{\text{in}}[g] \leftarrow f_{\text{in}}[g] + 1$
 237 8: $g \leftarrow g_{\text{out}}[i]$
 238 9: $f_{\text{out}}[g] \leftarrow f_{\text{out}}[g] + 1$
 239 10: **fim para**
 240 11: **para** i de 1 até $|N|$ **faça**
 241 12: $f_{\text{in}}[i] \leftarrow f_{\text{in}}[i]/|E|$ {Normalização dos vetores de distribuição.}
 242 13: $f_{\text{out}}[i] \leftarrow f_{\text{out}}[i]/|E|$
 243 14: **fim para**
 244 15: **retorne** $\langle f_{\text{in}}, f_{\text{out}} \rangle$ {Vetores com a distribuição dos graus de entrada e saída.}

245 A normalização é deixada por último, da linha 11 à 14, a fim de minimizar a propagação
 246 de erros numéricos relacionados à representação de ponto flutuante. A complexidade do
 247 algoritmo é da ordem de $\Theta(n E\{\mathcal{G}^{\text{out}}\})$ em tempo e $\Theta(n)$ em memória. \square

248 Para manter a simplicidade na representação e a consistência em relação às equações,
 249 considera-se que os vetores f_{in} e f_{out} são indexados a partir de 0. A computação realizada
 250 especificamente no Algoritmo 2.2 é menos custosa do que a realizada no Algoritmo 2.1.
 251 Por esse motivo, juntamente com o fato de que o Algoritmo 2.2 faz uso do 2.1, suas
 252 complexidades em tempo são a mesma. O mesmo acontece em relação à escalabilidade.
 253 Como o tempo de execução dos Algoritmos 2.1 e 2.2 depende do valor esperado do grau
 254 de saída, pode ser útil analisar o comportamento desse valor como uma propriedade da
 255 rede. Define-se a seguir como computar esse valor esperado utilizando a função $f(\cdot)$.

^[iv]Nesse caso diz-se que $E\{\mathcal{G}^{\text{out}}\}$ possui a propriedade de ser livre de escala. O que é factível no caso de redes como a Internet, em que a quantidade de conexões entre os nós não tem relação com a quantidade de nós de toda a rede. Isso porque a quantidade de conexões depende apenas da quantidade de portas de comunicação dos dispositivos de interconexão, por exemplo, roteadores, comutadores e repetidores.

Valor esperado do grau O valor esperado do grau pode ser computado com base na função de distribuição de probabilidade $f(\cdot)$. Como a probabilidade de realização k da variável aleatória \mathcal{G} é dada por $P(\mathcal{G}_k) = f(k)$, com base na Equação 2.8, pode-se calcular o valor esperado do grau de um nó pela equação

$$E\{\mathcal{G}\} = \sum_{\zeta=0}^{n-1} \mathcal{G}_{\zeta} P(\mathcal{G}_{\zeta}) = \sum_{k=0}^{n-1} k f(k), \quad (2.9)$$

onde a função $f(\cdot)$ deve ser utilizada em acordo com a variável aleatória utilizada, isto é, a função $f_{\text{in}}(\cdot)$ para a de grau de entrada \mathcal{G}^{in} , ou $f_{\text{out}}(\cdot)$ para a de grau de saída \mathcal{G}^{out} .

Função de distribuição cumulativa A partir da aplicação do Algoritmo 2.2, é possível verificar de forma eficiente e escalável qual é a função de distribuição de probabilidade melhor descreve a forma de $f(\cdot)$, como, por exemplo, a geométrica, a uniforme ou a de Poisson. Isso pode ser de grande utilidade na elaboração de modelos para representação dessas redes. Porém, em alguns casos é mais adequada a utilização da função de distribuição cumulativa $F(\cdot)$, ou CDF (*Cumulative Distribution Function*), pois essa é menos sujeita a problemas de resolução encontrados na visualização da função de distribuição de probabilidade. A função $F(\cdot)$, denominada CDF complementar, é definida por

$$F(k) = \sum_{i=k}^{n-1} f(i), \quad (2.10)$$

onde n representa a quantidade de nós da rede. Considerando os possíveis valores de grau dos nós, tem-se a relação $0 \leq k \leq i \leq (n-1)$. Frequentemente, a distribuição dos graus de uma rede é analisada a partir da visualização da função $F(\cdot)$ em escala logarítmica, pois verifica-se experimentalmente que esse tipo de visualização é mais efetiva que a visualização do histograma da função $f(\cdot)$ Newman (2003).

2.2.2 Distância média

A distância entre dois nós da rede é definida como o menor caminho entre os dois nós. São duas as formas de se quantificar o caminho entre dois nós. Na primeira, no caso de grafos com pesos, essa distância pode ser quantificada como o caminho entre os dois nós cuja soma dos pesos seja mínima. No segundo caso, para grafos sem pesos, ou grafos com pesos em que o peso não indica distância, o menor caminho é quantificado em termos da quantidade mínima de arestas. Neste trabalho é considerado o segundo caso. A distância entre dois nós é definida a seguir.

Definição 2.6 (Distância entre dois nós). Em um grafo sem pesos $G = \langle N, E \rangle$, ou com pesos, mas em que os pesos não discriminam uma relação de proximidade, a distância de um nó $v_i \in N$ até um nó $v_j \in N$ é definida como a quantidade de arestas que compõem o caminho mais curto de v_i até v_j , denominada s_{ij} . \square

Dado o fato de que a distância entre dois nós é definida em termos da quantidade de arestas do caminho mais curto, é direta a associação da construção de sua solução algorítmica com base em algum algoritmo de caminho mais curto em grafos. Como foi considerado na Definição 2.6 que o peso das arestas não caracterizam relação de proximidade, o problema de caminho mais curto pode ser resolvido com a estratégia de busca em largura, ou BFS (*Breadth-First Search*), como apresentado no algoritmo a seguir.

Algoritmo 2.3 (Distância em relação a um nó). É possível calcular a distância de um nó v_i para todos os outros nós do grafo G de forma iterativa com base na representação por lista de adjacência L , percorrendo as listas e armazenando nós não visitados em uma fila q , implementada segundo uma política de acesso do tipo FIFO (*First-In First-Out*).

```

297 algoritmo distancia( $L, v_i$ )
298   1: {Lista de adjacência  $L$  de um grafo  $G = \langle N, E \rangle$ , e nó  $v_i$  raiz da árvore de busca.}
299   2:  $s \leftarrow$  novo-vetor( $|N|, \infty$ ) {Vetor de  $|N|$  posições preenchidas com infinito.}
300   3:  $p \leftarrow$  novo-vetor( $|N|$ , nulo) {Vetor de  $|N|$  posições preenchidas com referências nulas.}
301   4:  $q \leftarrow$  nova-fila() {Cria uma nova fila vazia.}
302   5:  $s[i] \leftarrow 0$ 
303   6:  $c \leftarrow 0$ 
304   7: marcar( $v_i$ )
305   8: enfileirar( $q, v_i$ ) {Enfileira o nó  $v_i$  na fila vazia  $q$ .}
306   9: enquanto  $q$  não estiver vazia faça
307     10:    $c \leftarrow c + 1$ 
308     11:    $q_{\text{local}} \leftarrow$  nova-fila() {Cria uma nova fila vazia.}
309     12:   enquanto  $q$  não estiver vazia faça
310       13:      $v_k \leftarrow$  desenfileirar( $q$ )
311       14:     para cada  $(v_j, w) \in L[k]$  faça
312         15:       se  $v_j$  não está marcado então
313           16:          $s[j] \leftarrow c$  {Distância  $c$  de  $v_i$  até  $v_j$ .}
314           17:          $p[j] \leftarrow k$  {Nó  $v_k$  é pai de  $v_j$  na árvore de busca.}
315           18:         marcar( $v_j$ )
316           19:         enfileirar( $q_{\text{local}}, v_j$ )
317       20:     fim se
318     21:   fim para
319     22:   fim enquanto
320     23:   mover( $q_{\text{local}}, q$ ) {Esvazia fila  $q_{\text{local}}$  em  $q$ .}
321   24: fim enquanto
322   25: retorne  $\langle s, p \rangle$  {Vetor  $s$  com as distâncias a partir de  $v_i$  e  $p$  com a árvore de busca.}

```

Como a estratégia BFS acha o menor caminho em número de arestas, o contador c é a distância de v_i para cada nó v_j . A complexidade é de $O(n E\{\mathcal{G}^{\text{out}}\})$ em tempo e de $\Theta(n)$ em memória, onde $n = |N|$. \square

O Algoritmo 2.3 percorre a árvore de busca de forma a expandir um nível da árvore por iteração do laço de repetição da linha 9. Por esse motivo, a cada entrada no laço o contador c é incrementado, pois, dessa forma, esse contador representa a profundidade de busca na árvore. Obviamente, a profundidade de uma árvore de busca corresponde a menor distância entre o nó raiz e os nós naquela profundidade. Além disso, na linha 17, a cada nó é atribuído uma referência ao pai na árvore, o que torna possível reconstruir o menor caminho posteriormente.

Em relação à complexidade do Algoritmo 2.3 tem-se que a linha 10 do algoritmo irá iterar sobre cada nó da rede, no caso em que o grafo seja conexo e, para cada nó, a linha 12 irá iterar sobre cada elemento da lista de adjacência do nó. Utilizando $n = |N|$ para representar a quantidade de nós do grafo e $E\{\mathcal{G}^{\text{out}}\}$ para o valor esperado do tamanho da lista de adjacência, tem-se que a complexidade em tempo do Algoritmo 2.3 é da ordem de $O(n E\{\mathcal{G}^{\text{out}}\})$. Considera-se que as operações de enfileirar, desenfileirar e marcar são

de tempo constante. Além disso, tem-se que a operação de mover os elementos da fila q_{local} para a fila q depende apenas do tamanho da fila q_{local} . Somando todos os tamanhos de q_{local} obtém-se a quantidade de nós $|N|$, portanto, a linha 23 executa a mesma ordem de quantidade de operações que o laço de repetição da linha 12. Esses fatores contribuem para que a complexidade em tempo do algoritmo seja de $O(n E\{\mathcal{G}^{\text{out}}\})$. Utilizou-se a notação $O(\cdot)$, pois caso o grafo não seja conexo, o tempo de execução esperado não é limitado inferiormente pela função. Caso o grafo seja conexo, a complexidade em tempo é da ordem de $\Theta(n E\{\mathcal{G}^{\text{out}}\})$. Em relação à quantidade de memória necessária, deve-se considerar o tamanho do vetor s , da ordem de $\Theta(n)$, e o da fila q , da ordem de $O(n)$. Portanto, a complexidade em memória do algoritmo é da ordem de $\Theta(n)$.

Além de computar as distâncias, o Algoritmo 2.3 também cria um vetor p que é capaz de representar a árvore de busca em que o nó v_i é raiz. Logo, a partir desse vetor, é possível reconstruir o menor caminho entre v_i e qualquer nó v_j da rede em tempo linear. Para isso, basta iterar sobre as referências do vetor p , partindo de $p[j]$ até se chegar em v_i , se o grafo for conexo, ou até se encontrar uma referência nula, o que indica que não há caminho partindo de v_i até v_j . A seguir, é apresentado um algoritmo que efetua esse procedimento.

Algoritmo 2.4 (Caminho mais curto). A partir da descrição da árvore de busca computada pelo Algoritmo 2.3 é possível reconstruir o menor caminho entre um nó v_i e qualquer outro nó v_j da rede.

```

algoritmo menor-caminho( $L, v_i, v_j$ )
1: {Lista de adjacência  $L$  de um grafo  $G = \langle N, E \rangle$ , e nó  $v_i$  raiz da árvore de busca.}
2:  $l \leftarrow \text{nova-lista}()$  {Cria uma nova lista vazia.}
3:  $\langle s, p \rangle \leftarrow \text{distancia}(L, v_i)$ 
4:  $r \leftarrow j$ 
5: inserir( $l, v_j$ ) {Insere o nó  $v_j$  no início da lista  $l$ .}
6: enquanto  $p[r] \neq \text{nulo}$  faça
7:    $r \leftarrow p[r]$ 
8:   inserir( $l, v_r$ ) {Insere pai não nulo no início da lista.}
9: fim enquanto
10:  $c \leftarrow \text{nulo}$ 
11: se  $\text{inicio}(l) = v_i$  então
12:    $c \leftarrow l$  {Se o início de  $l$  é  $v_i$ , então o menor caminho existe e é descrito por  $l$ .}
13: fim se
14: retorne  $c$  {Caminho mais curto de  $v_i$  em direção a  $v_j$ .}

```

Considerando as complexidades do procedimento ‘distancia()’, tem-se que a complexidade é de $O(n E\{\mathcal{G}^{\text{out}}\})$ em tempo e de $\Theta(n)$ em memória, onde $n = |N|$. \square

Em relação às complexidades, a construção do caminho no laço da linha 6 é de $O(n)$, porém predominam as complexidades do procedimento ‘distancia()’. Portanto, a complexidade é de $O(n E\{\mathcal{G}^{\text{out}}\})$ em tempo e de $\Theta(n)$ em memória, onde $n = |N|$. A partir da Definição 2.6 e do Algoritmo 2.3, é possível definir e computar a distância média entre todos os pares de nós da rede. A soma dessas distâncias é, então, dividida pelo total de distâncias, que, em uma rede com n nós, é de exatamente n^2 . Quando o grafo não é conexo, há que se analisar qual valor deve ser considerado quando não há um caminho entre dois nós. Para evitar esse problema, neste trabalho considera-se apenas a distância média para o caso grafos conexos, matematicamente definida a seguir.

Definição 2.7 (Distância média). A distância média em um grafo conexo $G = \langle N, E \rangle$ é definida como sendo o valor esperado da distância entre cada par ordenado de nós (v_i, v_j) pertencentes ao produto cartesiano $N \times N$, que pode ser escrito como

$$E\{\mathcal{S}\} = \frac{1}{n^2} \sum_{\{(v_i, v_j) \in N \times N\}} s_{ij}, \quad (2.11)$$

onde \mathcal{S} é a variável aleatória que indica a distância entre dois nós quaisquer da rede. A quantidade de elementos em $N \times N$ é de $|N|^2$. Assim, o valor esperado representa a média aritmética das distâncias. \square

A partir da análise de valor mínimo e máximo da Equação 2.11, para o caso de grafos conexos, é possível verificar a relação de desigualdade

$$1 - \frac{1}{n} \leq E\{\mathcal{S}\} \leq \frac{n-1}{2}, \quad (2.12)$$

onde o termo à esquerda corresponde à distância média do grafo completo, isto é, o grafo possui todas as arestas possíveis. Já o termo à direita corresponde à distância média no grafo linear. Se o grafo, além de conexo, possui a propriedade em que a sentença $(\forall e_{ij} \in E)((\exists e_{ji} \in E)((e_{ij} = (v_i, v_j)) \wedge (e_{ji} = (v_j, v_i))))$ é sempre satisfeita, então, é possível reduzir a quantidade de cálculos de distância quase que pela metade. Isso porque, nessas condições, tem-se que $s_{ij} = s_{ji}$, mesmo que a sentença $(\exists e_{ij} \in E)(\omega(e_{ij}) \neq \omega(e_{ji}))$ seja verdadeira. Portanto, nessas condições, é possível redefinir a distância média como

$$E\{\mathcal{S}\} = \frac{2}{n(n+1)} \sum_{\{(v_i, v_j) \in N \times N : i \geq j\}} s_{ij}, \quad (2.13)$$

que pode reduzir o tempo de cálculo de $E\{\mathcal{S}\}$ praticamente pela metade. Porém, por se tratar de uma redução em um fator constante, essa otimização não impacta de forma significativa a complexidade teórica do cálculo da distância média. Na Equação 2.13, a restrição do somatório exclui os pares simétricos mas mantém os casos em que $i = j$ a fim de estabelecer equivalência com a equação original. A partir da Equação 2.11, que é mais geral e possui a mesma complexidade assintótica da Equação 2.13, é possível elaborar um algoritmo para calcular a distância média em um grafo conexo.

Algoritmo 2.5 (Distância média). É possível calcular a distância média em um grafo a partir de sua representação em lista de adjacência com base no Algoritmo 2.3 e na Equação 2.11, como descrito pelo procedimento a seguir.

```

algoritmo distancia-media( $L$ )
1: {Lista de adjacência  $L$  de um grafo direcionado  $G = \langle N, E \rangle$ .}
2:  $m \leftarrow 0$ 
3: para cada  $v \in N$  faça
4:    $\langle s, p \rangle \leftarrow \text{distancia}(L, v)$  {Calcula as distâncias de  $v$  para os outros nós do grafo.}
5:   para  $i$  de 1 até  $|N|$  faça
6:      $m \leftarrow s[i]$ 
7:   fim para
8: fim para
9: retorne  $m/|N|^2$  {Soma das distâncias sobre o quadrado da quantidade de nós.}

```

Considerando que o grafo é conexo e que $n = |N|$, a complexidade em tempo do algoritmo é da ordem de $\Theta(n^2 E\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é da ordem de $\Theta(n)$. \square

A complexidade em tempo do Algoritmo 2.5 seria de ordem quadrática se a quantidade de execuções da linha 5 fosse o ponto crítico. Porém, como a complexidade do Algoritmo 2.3 (a função ‘distancia()’ na linha 4) é da ordem de $\Theta(n E\{\mathcal{G}^{\text{out}}\})$, para grafos conexos, e este algoritmo é utilizado $n = |N|$ vezes, então a complexidade em tempo do Algoritmo 2.5 é de $\Theta(n^2 E\{\mathcal{G}^{\text{out}}\})$. Já a complexidade em memória necessita de considerações acerca da política de alocação e liberação de memória do vetor s . Na implementação do Algoritmo 2.5, se o bloco de memória associado ao vetor sempre é desalocado antes de uma nova associação, tem-se que sua complexidade em memória é da ordem de $\Theta(n)$, caso contrário, será da ordem de $\Theta(n^2)$. Porém, a falta da implementação de uma estratégia de desalocação da memória associada ao vetor s está mais para um problema de implementação de que para algo relacionado à natureza do algoritmo. Por esse motivo, estabelece-se que sua complexidade em memória é da ordem de $\Theta(n)$.

2.2.3 Diâmetro

A distância máxima entre dois nós de uma rede representa o pior caso de eficiência na propagação de informação através dela. No projeto de redes complexas, pode ser desejável que essa distância máxima não ultrapasse um determinado limiar. O conceito de diâmetro da rede representada por grafos é apresentado a seguir.

Definição 2.8 (Diâmetro). O diâmetro de um grafo $G = \langle N, E \rangle$ é definido como a maior distância entre dois nós quaisquer do grafo. Para grafo direcionados, esse conceito pode ser quantificado da seguinte forma

$$d(G) = \max_{\{(v_i, v_j) \in N \times N\}} s_{ij}, \quad (2.14)$$

onde s_{ij} representa a distância partindo do nó v_i em direção ao nó v_j . \square

Com base na Definição 2.8, é possível identificar que, assim como não há sentido em quantificar a distância média em grafos não conexos, a função de diâmetro $d(G)$ não tem utilidade em quantificar o diâmetro de grafos não conexos. Por esse motivo, o seguinte algoritmo é definido apenas para grafos conexos.

Algoritmo 2.6 (Diâmetro). É possível calcular o diâmetro de um grafo a partir de sua representação em lista de adjacência com base no Algoritmo 2.3 e na Equação 2.14, como descrito pelo procedimento a seguir.

```

algoritmo diametro( $L$ )
1: {Lista de adjacência  $L$  de um grafo direcionado  $G = \langle N, E \rangle$ .}
2:  $d \leftarrow 0$ 
3: para cada  $v \in N$  faça
4:    $\langle s, p \rangle \leftarrow \text{distancia}(L, v)$  {Calcula as distâncias de  $v$  para os outros nós do grafo.}
5:   para  $i$  de 1 até  $|N|$  faça
6:     se  $d < s[i]$  então
7:        $d \leftarrow s[i]$ 
8:     fim se
9:   fim para
10: fim para
11: retorne  $d$  {Maior distância entre os nós do grafo.}
```

Considerando que o grafo é conexo e que $n = |N|$, a complexidade em tempo do algoritmo é da ordem de $\Theta(n^2 E\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é da ordem de $\Theta(n)$. \square

464 Devido à similaridade entre os Algoritmos 2.3 e 2.5, suas complexidades, tanto em
 465 tempo quanto em espaço, são iguais, pelas mesmas justificativas. Assim, a complexidade
 466 em tempo do algoritmo é da ordem de $\Theta(n^2 E\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é
 467 da ordem de $\Theta(n)$.

468 2.2.4 Eficiência

469 Com as Equações 2.11 e 2.13, além do Algoritmo 2.5, é possível quantificar a velocidade
 470 de propagação da informação na rede. Quando $E\{\mathcal{S}\}$ é pequeno em relação à quantidade
 471 de nós $|N|$, são necessárias poucas transmissões para que uma mensagem seja propagada
 472 em toda a rede. A velocidade de transmissão na rede é inversamente proporcional à
 473 sua distância média. Essa relação é explorada no trabalho de [Latora e Marchiori \(2001\)](#)
 474 e é denominada eficiência média da rede. A eficiência média depende da eficiência da
 475 propagação da informação entre os nós da rede. A eficiência entre dois nós é definida a
 476 seguir.

477 **Definição 2.9** (Eficiência). Em um grafo $G = \langle N, E \rangle$, a medida de eficiência associada
 478 ao trajeto partindo de um nó v_i para um nó v_j , denotada por f_{ij} , é dada em função do
 479 inverso da distância entre os dois nós, ou seja

$$f_{ij} = \frac{1}{s_{ij}}, \quad (2.15)$$

480 onde s_{ij} é a distância do nó v_i em direção ao nó v_j e $i \neq j$, pois $s_{ii} = 0$. \square

481 Apesar de não ser definida para $i = j$, a medida f_{ij} é definida quando não há um
 482 caminho entre os dois nós v_i e v_j , nesse caso, admitindo $s_{ij} = \infty$, tem-se que $f_{ij} = 0$.
 483 A eficiência média não exige que o grafo G seja conexo, diferente do que acontece com a
 484 distância média. Com base nessas afirmações, a eficiência média de um grafo é definida a
 485 seguir.

486 **Definição 2.10** (Eficiência média). A eficiência média de um grafo $G = \langle N, E \rangle$, deno-
 487 tada por $E\{\mathcal{F}\}$, consiste na média aritmética da eficiência entre todos os pares de nós
 488 distintos da rede, ou seja, todo par do conjunto $\{(v_i, v_j) \in N \times N : i \neq j\}$. Portanto, a
 489 eficiência média pode ser calculada como

$$E\{\mathcal{F}\} = \frac{1}{n(n-1)} \sum_{\{(v_i, v_j) \in N \times N : i \neq j\}} f_{ij}, \quad (2.16)$$

490 onde $n = |N|$ representa a quantidade de nós e o denominador se deve ao fato de que a
 491 quantidade de medidas de eficiência f_{ij} , nesse caso, é de $n(n-1)$. \square

492 A partir da análise da Equação 2.16 é possível verificar a seguinte relação de desigual-
 493 dade $0 \leq E\{\mathcal{F}\} \leq 1$. Porém, considerando o grafo conexo menos eficiente (topologia em
 494 anel direcionada em apenas um sentido), essa desigualdade pode ser reescrita como

$$\frac{H_{n-1}}{n-1} \leq E\{\mathcal{F}\} \leq 1, \quad (2.17)$$

495 onde H_n representa o n -ésimo termo da série harmônica (definida na Seção de Simbologia
 496 do Glossário). Mesmo assim, como H_n é da ordem de $\Theta(\log n)$, mesmo considerando que
 497 o grafo é conexo, a eficiência ainda pode ser muito baixa. Isso porque o limite inferior

tende a 0 quando n tende para infinito. A seguir, é descrito o algoritmo utilizado para se computar a eficiência de um grafo.

Algoritmo 2.7 (Eficiência média). É possível calcular a eficiência média em um grafo G a partir de sua representação em lista de adjacência com base no Algoritmo 2.3 e na Equação 2.16, como descrito pelo procedimento a seguir.

```

algoritmo eficiencia-media( $L$ )
1: {Lista de adjacência  $L$  de um grafo direcionado  $G = \langle N, E \rangle$ .}
2:  $f \leftarrow 0$ 
3: para cada  $v_i \in N$  faça
4:    $\langle s, p \rangle \leftarrow \text{distancia}(L, v_i)$  {Distâncias de  $v_i$  para os outros nós do grafo.}
5:   para  $j$  de 1 até  $|N|$  faça
6:     se  $i \neq j$  então
7:        $f \leftarrow 1/s[j]$  {Quando  $s[j]$  for diferente de zero.}
8:     fim se
9:   fim para
10: fim para
11: retorne  $f/(|N|^2 - |N|)$  {Eficiência média do grafo  $G$ .}
```

Considerando que o grafo é conexo e que $n = |N|$, a complexidade em tempo do algoritmo é da ordem de $\Theta(n^2 E\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é da ordem de $\Theta(n)$. \square

Devido a similaridade entre os Algoritmos 2.7 e 2.5, suas complexidades, tanto em tempo quanto em espaço, são iguais, pelas mesmas justificativas. Assim, a complexidade em tempo do algoritmo é da ordem de $\Theta(n^2 E\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é da ordem de $\Theta(n)$.

2.2.5 Coeficiente de agrupamento

É de se esperar que, se os vizinhos do nó v_i possuem uma quantidade suficiente de arestas entre eles, a remoção de v_i pouco deve influenciar na eficiência que o nó representa para a rede. Nessas condições, a remoção de v_i não influenciaria de forma significativa a distância média entre os nós da rede. Para quantificar essa influência, a seguir define-se o conceito de vizinhança.

Definição 2.11 (Vizinhança). A vizinhança de um nó v_i , denominada $h(v_i)$, consiste em um subconjunto de nós de N formado pelos nós em que v_i incide, ou seja

$$h(v_i) = \{v_j \in N : (v_i, v_j) \in E\}, \quad (2.18)$$

onde $|h(v_i)| = g_{\text{out}}(v_i)$ e o valor esperado de vizinhos é equivalente a $E\{\mathcal{G}^{\text{out}}\}$. \square

A quantidade de vizinhos do nó v_i , representada por $|h(v_i)|$, é igual à definição do grau de saída do nó v_i , como apresentado na Equação 2.3 da Definição 2.3. Utilizando a Definição 2.11, é possível analisar o quão denso é o subgrafo dos vizinhos de v_i . Quanto mais denso é esse subgrafo, maior é a suficiência do número de arestas e menor será a influência da remoção de v_i . É possível quantificar a suficiência do número de arestas entre os vizinhos como a razão entre a quantidade de arestas existentes sobre a quantidade total de possíveis arestas. A quantidade de arestas entre os vizinhos de v_i pode ser escrita como

$$t_i = |\{(v_j, v_k) \in E : (v_j \in h(v_i)) \wedge (v_k \in h(v_i))\}|, \quad (2.19)$$

onde $0 \leq t_i \leq g_{\text{out}}(v_i)(g_{\text{out}}(v_i) - 1)$, cujo valor mínimo representa os casos onde $h(v_i) = \emptyset$, ou quando não há arestas entre os vizinhos, e o valor máximo quando existem todas as $g_{\text{out}}(v_i)(g_{\text{out}}(v_i) - 1)$ arestas entre os vizinhos. É possível quantificar a densidade desse subgrafo, ou seja, o quão agrupados são os vizinhos de um nó, por meio da seguinte definição e do algoritmo subsequente.

Definição 2.12 (Coeficiente de agrupamento). O coeficiente de agrupamento de um nó v_i é definido como a razão entre a quantidade de arestas entre os vizinhos, denotada t_i , e a total de arestas possível, da seguinte forma

$$c_i = \frac{t_i}{g_{\text{out}}(v_i)(g_{\text{out}}(v_i) - 1)}, \quad (2.20)$$

onde é possível identificar que $0 \leq c_i \leq 1$. \square

A partir da Equação 2.20 é possível verificar que só é possível computar o coeficiente de agrupamento de nós em que $g_{\text{out}}(v_i) > 1$, caso contrário teria-se uma divisão por 0. Desse forma, no caso em que $g_{\text{out}}(v_i) \leq 1$ pode-se assumir que o coeficiente de agrupamento é nulo, isto é, $((g_{\text{out}}(v_i) \leq 1) \Rightarrow (c_i = 0))$.

Algoritmo 2.8 (Coeficiente de agrupamento). É possível calcular o coeficiente de agrupamento de um nó da rede representada por um grafo G a partir da representação em lista de adjacência, com base na Definição 2.12, como descrito pelo procedimento a seguir.

```

algoritmo agrupamento( $L, v_i$ )
1: {Lista de adjacência  $L$  de um grafo  $G = \langle N, E \rangle$ , e nó  $v_i \in N$ .}
2:  $t_i \leftarrow 0$ 
3:  $g_{\text{out}} \leftarrow 0$ 
4: para cada  $(v_j, p_{ij}) \in L[v_i]$  faça
5:     marcar( $v_j$ )
6:      $g_{\text{out}} \leftarrow g_{\text{out}} + 1$ 
7: fim para
8: para cada  $(v_j, p_{ij}) \in L[v_i]$  faça
9:     para cada  $(v_w, p_{jw}) \in L[v_j]$  faça
10:        se  $v_w$  está marcado então
11:             $t_i \leftarrow t_i + 1$ 
12:        fim se
13:    fim para
14: fim para
15: retorne  $t_i / (g_{\text{out}}(g_{\text{out}} - 1))$  {Coeficiente de agrupamento do nó  $v_i$ .}
```

Considerando que $n = |N|$ representa a quantidade de nós do grafo G , a complexidade em tempo do algoritmo é da ordem de $\Theta(E^2\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é da ordem de $\Theta(E\{\mathcal{G}^{\text{out}}\})$. \square

Inicialmente, o Algoritmo 2.8 marca os nós que são vizinhos de v_i e realiza a contagem da quantidade de vizinhos. Essa marcação acarreta em uma complexidade em tempo da ordem de $\Theta(E\{\mathcal{G}^{\text{out}}\})$. É possível eliminar a necessidade de memória, porém isso acarreta no aumento da complexidade em tempo. Opcionalmente, em vez de marcar os nós vizinhos, é possível verificar se v_w está dentre os vizinhos de v_i no Algoritmo 2.8. Essa tarefa adicional é de complexidade $\Theta(E\{\mathcal{G}^{\text{out}}\})$ e acarretaria no aumento da complexidade em tempo do algoritmo, de forma que o algoritmo resultante seria da ordem de $\Theta(E^3\{\mathcal{G}^{\text{out}}\})$.

A fim de não aumentar mais a complexidade em tempo do algoritmo, que já é maior que a complexidade em memória, neste trabalho adota-se a estratégia expressa no Algoritmo 2.8. Posteriormente, é realizada a contagem de arestas entre os vizinhos de v_i . Assim, espera-se percorrer $E\{\mathcal{G}^{\text{out}}\}$ listas, cujo tamanho esperado é $E\{\mathcal{G}^{\text{out}}\}$. A complexidade esperada em tempo do algoritmo é da ordem de $\Theta(E^2\{\mathcal{G}^{\text{out}}\})$. A seguir, o conceito de coeficiente de agrupamento de um nó, juntamente com o Algoritmo 2.8, é utilizado para calcular o coeficiente de agrupamento da rede, definido a seguir.

Definição 2.13 (Coeficiente de agrupamento da rede). O coeficiente de agrupamento da rede é definido como a média aritmética dos coeficientes de agrupamento individuais de cada nó da rede

$$E\{\mathcal{C}\} = \frac{1}{n} \sum_{\{v_i \in N\}} c_i, \quad (2.21)$$

onde $n = |N|$ representa a quantidade de nós da rede. \square

Com base na Definição 2.13 tem-se que $E\{\mathcal{C}\}$ representa o coeficiente de agrupamento da rede ou valor esperado do coeficiente de agrupamento dos nós da rede, já que \mathcal{C} é a variável aleatória que expressa o coeficiente de agrupamento de um nó qualquer da rede. A seguir é apresentado o algoritmo para calcular o coeficiente de agrupamento da rede.

Algoritmo 2.9 (Coeficiente de agrupamento da rede). É possível calcular o coeficiente de agrupamento da rede representada por um grafo G a partir com base na média dos coeficientes de agrupamento de cada nó da rede, Definição 2.13, como descrito pelo procedimento a seguir.

```

algoritmo agrupamento-medio( $L$ )
1:  $\{ \text{Lista de adjacência } L \text{ de um grafo } G = \langle N, E \rangle. \}$ 
2:  $c \leftarrow 0$ 
3: para cada  $v_i \in N$  faça
4:    $c \leftarrow c + \text{agrupamento}(L, v_i)$ 
5: fim para
6: retorne  $c/|N|$   $\{ \text{Coeficiente de agrupamento da rede.} \}$ 

```

Considerando que $n = |N|$ representa a quantidade de nós do grafo G , a complexidade em tempo do algoritmo é da ordem de $\Theta(n E^2\{\mathcal{G}^{\text{out}}\})$ e sua complexidade em memória é da ordem de $\Theta(E\{\mathcal{G}^{\text{out}}\})$. \square

As complexidades do Algoritmo 2.9 dependem de forma direta das complexidades do Algoritmo 2.8. Por utilizar o algoritmo de cálculo do coeficiente de agrupamento para cada nó, a complexidade em tempo do Algoritmo 2.9 é da ordem de $\Theta(n E^2\{\mathcal{G}^{\text{out}}\})$. Já a complexidade em memória, depende da do Algoritmo 2.8, portanto, a complexidade em memória do Algoritmo 2.9 é da ordem de $\Theta(E\{\mathcal{G}^{\text{out}}\})$. Se o valor esperado do grau de saída dos nós depender da quantidade de nós, tem-se que $E\{\mathcal{G}^{\text{out}}\} \in O(n)$. Nesse caso, pode-se considerar que a complexidade em tempo e memória, no pior caso, é da ordem de $O(n^3)$ e $O(n)$, respectivamente. Porém, em redes reais, espera-se que o valor de $E\{\mathcal{G}^{\text{out}}\}$ seja constante, ou, pelo menos, não dependa de n linearmente.

618 Capítulo 3

619 Modelos

620

*“It is not enough to be in the right place at the right time.
You should also have an open mind at the right time.”
Paul Erdős*

621 A Teoria de Redes Complexas apresenta-se como um modelo efetivo para representação
622 de Sistemas Complexos. Estes Sistemas podem ser representados por elementos (nós ou
623 vértices) e suas relações (ligações ou arestas). A adoção dessas teorias serve de base
624 para a criação de modelos eficazes para o entendimento de redes de grande magnitude e
625 complexidade.

626 Neste Apêndice, são apresentadas representações matemáticas e computacionais de
627 redes complexas. Especificamente, tem-se como objetivo possibilitar o desenvolvimento
628 de experimentos com base em algoritmos que podem ser utilizados para gerar modelos de
629 grafos. Como contribuição desta tese, tem-se a demonstração de propriedades topológicas
630 de topologias em anel, em estrela e linear.

631 Este Apêndice está organizado da seguinte forma: na Seção 3.1 é feita uma introdução
632 acerca de modelos topológicos de redes complexas. Nas Seções 3.2, 3.3, 3.4 e 3.5, são apre-
633 sentados modelos e propriedades de grafos determinísticos, aleatórios, de mundo pequeno
634 e livre de escala, respectivamente.

3.1 Introdução

Nesta Seção, são explorados modelos de topologias que servem como referência no estudo de redes complexas. Inicialmente, serão apresentadas três topologias determinísticas. O estudo dessas três estruturas topológicas têm como objetivo introduzir um método de utilização das métricas de caracterização para o estudo das propriedades da estrutura topológica de uma rede. Posteriormente, são apresentados modelos de estruturas topológicas não determinísticas fundamentais na teoria de redes complexas. Esses modelos fundamentais servem como referência no estudo das propriedades topológicas dessas redes.

3.2 Topologias determinísticas

Nesta Seção, com o objetivo de ilustrar a capacidade de caracterização das métricas apresentadas no Capítulo 2, é feita uma análise de três tipos de topologias determinísticas^[i] definidas como grafos direcionados conexos e simétricos. São elas: (i) topologia em anel, onde os nós são conectados a fim de formar um único ciclo que pode ser percorrido em ambos os sentidos, horário e anti-horário; (ii) topologia em estrela, onde há um nó central que se liga diretamente aos outros, e todos os outros nós ligam-se apenas a ele e (iii) topologia em linha, ou linear, onde os nós são dispostos como uma lista permitindo comunicação em ambos os sentidos.

A definição formal de cada topologia é apresentada na Seção A.1. Em todos os casos, considera-se que se existe uma aresta partindo de v_i em direção ao nó v_j , então também existe uma aresta partindo de v_j em direção a v_i , ou seja, $((v_i, v_j) \in E) \Leftrightarrow ((v_j, v_i) \in E)$. As topologias são ilustradas na Figura 3.1.

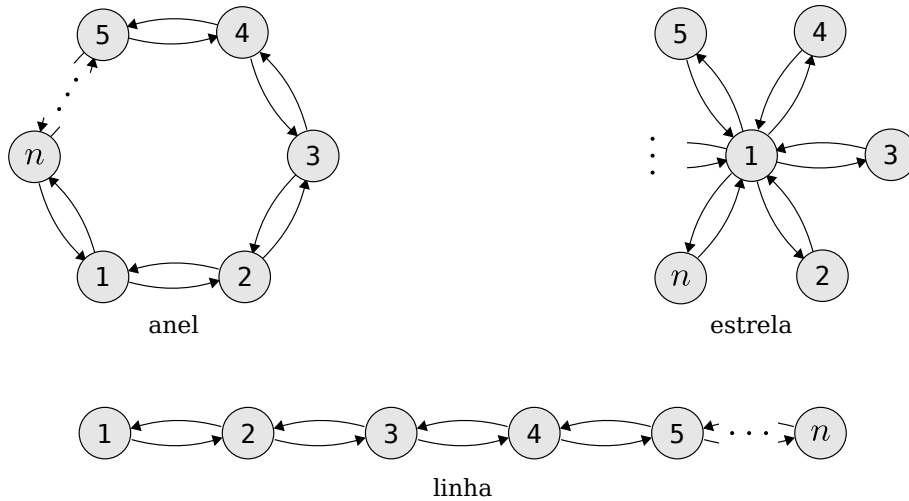


Figura 3.1: Ilustração das topologias em anel, em estrela e linear com n nós. A quantidade de arestas é igual para as topologias em linha e estrela, enquanto a em anel é maior em duas unidades.

A seguir as métricas de caracterização apresentadas no Capítulo 2 são aplicadas a cada uma das três topologias determinísticas apresentadas na Figura 3.1. Apenas o coeficiente de agrupamento não é aplicado pois ele é nulo para as três topologias em questão. Além

^[i]Essas três topologias foram escolhidas porque fazem parte das topologias básicas desenvolvidas no estudo de infraestrutura de comunicação de redes de computadores.

dessa propriedade em comum, outro fato relevante é que essas três topologias têm aproximadamente a mesma relação entre a quantidade de arestas, ou seja, aproximadamente a mesma densidade $m(G)$, como definida pela Equação 2.5. Essa propriedade em comum é importante porque, para uma mesma quantidade de nós, são utilizadas praticamente a mesma quantidade de canais de comunicação, isto é, a mesma quantidade de arestas. De fato, apenas a topologia em anel possui uma quantidade de arestas maior em duas unidades. Porém, assintoticamente essa quantidade torna-se irrelevante. É possível afirmar que a implementação de uma rede com n nós, em qualquer uma das três topologias, necessita da mesma quantidade de recursos. Isso permite qualificar de forma relativa cada uma das topologias.

Distribuição dos graus Nas três topologias determinísticas apresentadas, o grau de entrada de cada nó é igual ao grau de saída. Por esse motivo, a seguir, o termo grau é utilizado de forma genérica. A distribuição dos nós na topologia em anel possui a característica peculiar de que apenas existem nós com grau 2. Esse fato implica que a função de distribuição dos graus $f(\cdot)$ consiste em um pico unitário em $k = 2$, ou seja,

$$f_{\text{ring}}(k) = \delta(k - 2), \quad (3.1)$$

onde $\delta(\cdot)$ é a função delta de Kronecker. Já na topologia em linha, os dois nós das extremidade possuem grau unitário enquanto os outros $(n - 2)$ nós têm grau 2. O que resulta em uma função de distribuição com dois picos: o primeiro em $k = 1$ com amplitude $2/n$ e outro em $k = 2$ com amplitude $(n - 2)/n$, ou seja,

$$f_{\text{line}}(k) = \frac{2}{n}\delta(k - 1) + \frac{(n - 2)}{n}\delta(k - 2), \quad (3.2)$$

onde $n = |N|$ representa a quantidade de nós da rede. Na topologia em estrela, o nó central possui grau $(n - 1)$ e os demais nós grau unitário. A função de distribuição dos graus de uma topologia em estrela possui dois picos, um em $k = 1$ de amplitude $(n - 1)/n$ e outro em $k = (n - 1)$ de amplitude $1/n$, ou seja,

$$f_{\text{star}}(k) = \frac{(n - 1)}{n}\delta(k - 1) + \frac{1}{n}\delta(k - (n - 1)). \quad (3.3)$$

Na Tabela 3.1, são apresentadas cada uma das funções de distribuição.

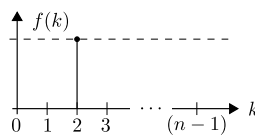
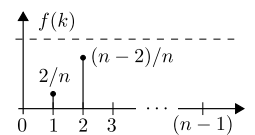
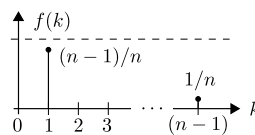
anel	linha	estrela
$m(G) = \frac{2}{(n - 1)}$	$m(G) = \frac{2}{n}$	$m(G) = \frac{2}{n}$
$f(k) = \delta(k - 2)$ 	$f(k) = \frac{2}{n}\delta(k - 1) + \frac{(n - 2)}{n}\delta(k - 2)$ 	$f(k) = \frac{(n - 1)}{n}\delta(k - 1) + \frac{1}{n}\delta(k - (n - 1))$ 
$E\{\mathcal{G}\} = 2$	$E\{\mathcal{G}\} = 2\frac{(n - 1)}{n}$	$E\{\mathcal{G}\} = 2\frac{(n - 1)}{n}$

Tabela 3.1: Densidade, distribuição e valor esperado dos graus das topologias determinísticas.

683 Dada a função de distribuição dos graus de cada topologia é possível computar o valor
 684 esperado do grau a partir da Equação 2.9. Para o caso da topologia em anel tem-se que
 685 $E\{\mathcal{G}\} = 2$ e para as topologias em estrela e linear o valor esperado tende a 2 quando
 686 a quantidade de nós tende ao infinito. Esse fato indica uma redução nas complexidades
 687 apresentadas na Tabela A.1, uma vez que $E\{\mathcal{G}^{\text{out}}\}$ pode ser considerado constante.

688 **Distância média** A distância média para cada uma das topologias determinísticas apre-
 689 sentadas possui um comportamento assintótico distinto. Para que fosse possível analisar
 690 tal comportamento, foi necessário elaborar para cada topologia, um equação que descre-
 691 vesse a distância média em relação à quantidade de nós de cada topologia. Essa análise
 692 é apresentada na Seção A.1. Além dos resultados analíticos descritos nessa Seção, foram
 693 também realizados experimentos computacionais. A distância média verificada experi-
 694 mentalmente para cada tipo de topologia é apresentada na Figura 3.2.

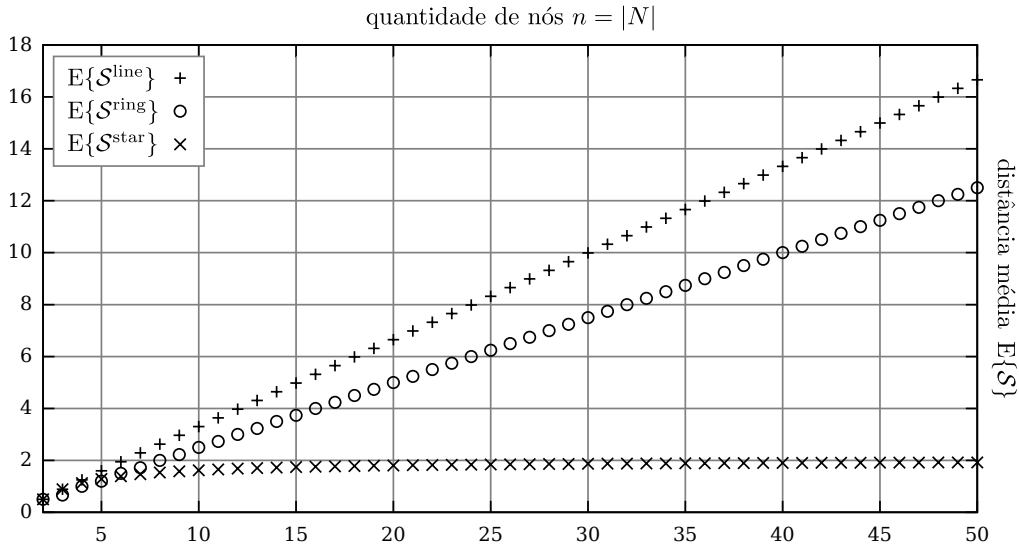


Figura 3.2: Distância média das topologias determinísticas. As variáveis aleatórias S^{line} , S^{ring} e S^{star} indicam a distância média nas topologias em linha, em anel e em estrela, respectivamente.

695 Na Figura 3.2, é possível verificar que a distância média na topologia linear é propor-
 696 cional à razão da quantidade de nós da rede por 3, por exemplo, para $n = 30$ tem-se que
 697 $E\{S^{\text{line}}\} \approx 10$. O que confirma o Corolário A.4, que trata do comportamento assintótico
 698 da distância média na topologia linear, que afirma que $E\{S^{\text{line}}\} \simeq n/3$. Em relação a to-
 699 pologia em anel tem-se que a distância média é proporcional à razão da quantidade de nós
 700 da rede por 4, por exemplo, para $n = 40$ tem-se que $E\{S^{\text{ring}}\} \approx 10$. O que está em acordo
 701 com o Corolário A.2 que define $E\{S^{\text{ring}}\} \simeq n/4$. Finalmente, em relação à topologia em
 702 estrela, vê-se na Figura 3.2 que quanto maior a quantidade de nós o valor esperado de
 703 S^{star} se aproxima de 2, por exemplo, para $n = 50$ tem-se que $E\{S^{\text{star}}\} \approx 1.92$. O que
 704 também é confirmado pelo Corolário A.3, onde $E\{S^{\text{star}}\} \simeq 2$.

705 **Diâmetro** A distância máxima em uma topologia representa seu diâmetro, como mos-
 706 trado na Equação 2.14. O diâmetro em uma rede linear é diretamente proporcional à sua
 707 quantidade de nós, pois a maior distância é aquela entre os nós extremos da rede, ou seja,

708

$$d(G_{\text{line}}) = n - 1. \quad (3.4)$$

709 No caso da topologia em anel, o diâmetro depende da quantidade de nós da rede assim como
 710 de sua paridade. Quando a quantidade de nós é ímpar, tem-se que a distância máxima é
 711 definida como $(n + 1)/2$, já para uma quantidade par, tem-se $n/2$. Com a composição dos
 712 dois casos, pode-se representar o diâmetro na topologia em anel da seguinte forma

$$d(G_{\text{ring}}) = \begin{cases} (n + 1)/2 & \text{se } n \text{ é ímpar} \\ n/2 & \text{se } n \text{ é par} \end{cases}. \quad (3.5)$$

713 No caso da topologia em estrela, a maior distância é aquela entre qualquer par de nós que
 714 não inclua o nó central, isto é,

$$d(G_{\text{star}}) = 2. \quad (3.6)$$

715 Portanto, assim como acontece com a métrica da distância média, o diâmetro das topo-
 716 logias em linha e anel são de ordem linear em relação à quantidade de nós, enquanto a
 717 topologia em estrela converge para a constante 2.

718 **Eficiência média** A eficiência média verificada experimentalmente para cada tipo de
 719 topologia é apresentada na Figura 3.3.

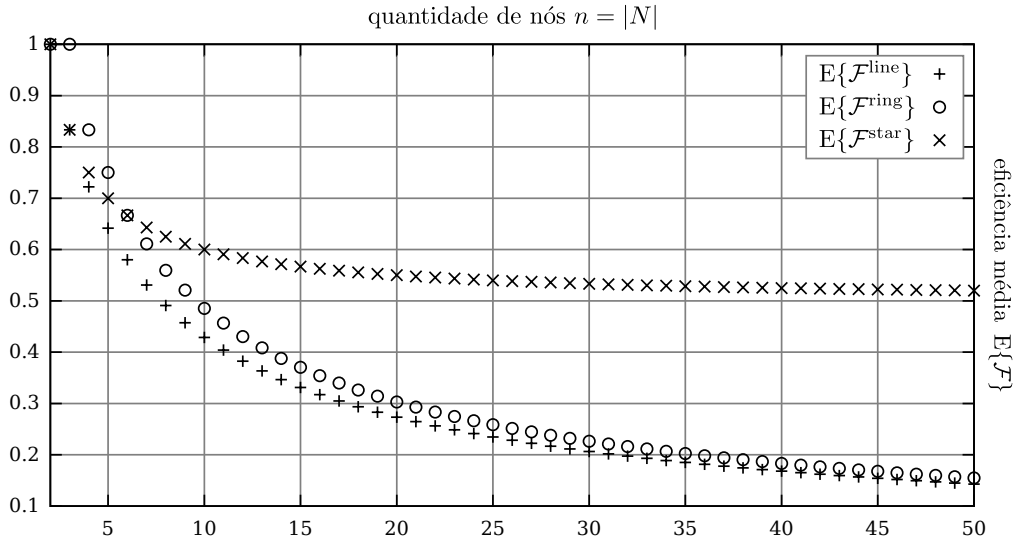


Figura 3.3: Eficiência média das topologias básicas determinísticas. As variáveis aleatórias $\mathcal{F}^{\text{line}}$, $\mathcal{F}^{\text{ring}}$ e $\mathcal{F}^{\text{star}}$ indicam a eficiência média nas topologias em linha, em anel e em estrela, respectivamente.

720 Em relação às topologias em anel e linear, há uma relação assintótica definida pela
 721 seguinte relação

$$E\{\mathcal{F}^{\text{ring}}\} \simeq E\{\mathcal{F}^{\text{line}}\} \simeq \frac{2 \log n}{n}. \quad (3.7)$$

722 Esses resultados podem ser verificados nos Corolários A.6 e A.8, respectivamente. A
 723 eficiência dessas duas topologias convergem para 0 quando n tende para o infinito, o que
 724 não ocorre no caso da topologia em estrela, cujo comportamento assintótico é definido
 725 como

$$E\{\mathcal{F}^{\text{star}}\} \simeq \frac{1}{2}, \quad (3.8)$$

ou seja, converge para um grau de eficiência de 50%, como apresentado pelo Corolário A.7. Considerando o exposto sobre a distância média dessas topologias, esse resultado é esperado, pois enquanto a distância média diverge, a eficiência média converge para zero. Nesse sentido, das três topologias determinísticas, a topologia em estrela, devido ao fato de sua distância média não divergir, assim como seu diâmetro, é a mais eficiente.

Considerações Nesta Seção, foram caracterizadas três topologias determinísticas quanto à sua distribuição dos graus, distância média, diâmetro e eficiência média. Foram apresentadas relações entre cada uma das métricas. Essas relações, associadas ao estudo de modelos representativos, podem auxiliar no projeto e o no entendimento da estrutura de redes. Isso pode ser realizado por meio da verificação de características gerais desses modelos. A seguir, são apresentados três modelos de redes complexas que são utilizados para criação de topologias não-determinísticas.

3.3 Grafos aleatórios

O modelo tradicional utilizado no estudo de redes complexas foi introduzido por Erdős e Rényi (1959) e compreende o estudo de grafos aleatórios. O procedimento proposto para criação desses grafos tem como parâmetros a quantidade de nós $n = |N|$ e a probabilidade de dois nós quaisquer estarem ligados por uma aresta, denominada p .

Algoritmo 3.1 (Criação de uma rede aleatória). Criação da lista de nós e de arestas de um grafo aleatório a partir da quantidade de nós n e da probabilidade de ligação p .

```

algoritmo grafo-aleatorio( $n, p$ )
1: {Quantidade de nós  $n$  e probabilidade de conexão  $p$ .}
2:  $N \leftarrow$  lista-de-nos( $n$ ) {Cria uma lista com  $n$  nós.}
3:  $E \leftarrow$  lista-de-arestas() {Cria uma nova lista vazia.}
4: para cada possível aresta  $e$  faça
5:   se distribuicao-uniforme(0, 1) <  $p$  então
6:     inserir( $E, e$ ) {Inserção condicionada por sorteio em distribuição uniforme.}
7:   fim se
8: fim para
9: retorne  $G = \langle N, E \rangle$  {Grafo aleatório resultante.}

```

Considerando o laço de repetição da linha 4, a complexidade do algoritmo em tempo é de $\Theta(|E|_{\max})$ e a complexidade esperada em memória de $\Theta(n + p|E|_{\max})$. \square

O Algoritmo 3.1 constrói o grafo aleatório G com $n = |N|$ nós com grau inicial nulo. A cada iteração, cada possível aresta é adicionada com probabilidade p . Como é necessário verificar todas as possibilidades de conexão, sua complexidade em tempo é da ordem de $\Theta(|E|_{\max})$, onde $|E|_{\max}$ representa a quantidade máxima de arestas no grafo. De forma geral, tem-se que a quantidade esperada de arestas em um grafo aleatório é dada por

$$E\{\mathcal{E}^{\text{rand}}\} = p|E|_{\max}, \quad (3.9)$$

onde o termo $\mathcal{E}^{\text{rand}}$ representa a variável aleatória que indica a quantidade de arestas em um grafo aleatório qualquer. Como o novo grafo aleatório possui n nós e a quantidade esperada de arestas é $p|E|_{\max}$, a quantidade necessária de memória é da ordem de $\Theta(n + p|E|_{\max})$. O valor de $|E|_{\max}$ depende do tipo de grafo a ser considerado. Esse valor é $n(n - 1)$ se o grafo for direcionado e $n(n - 1)/2$ se for não direcionado.

767 **Distribuição dos graus** Bollobás (1981) demonstra que a distribuição dos graus em
 768 um grafo aleatório segue a seguinte distribuição binomial

$$P(\mathcal{G}_k^{\text{rand}}) = \binom{n-1}{k} p^k (1-p)^{n-1-k}, \quad (3.10)$$

769 onde $n = |N|$ representa a quantidade de nós da rede e k o grau de um nó. Considerando
 770 o valor esperado de uma distribuição binomial tem-se que o valor esperado do grau em um
 771 grafo aleatório é dado por Papoulis e Pillai (2002)

$$E\{\mathcal{G}^{\text{rand}}\} = p(n-1), \quad (3.11)$$

772 ou seja, se $p = 1$, então o valor esperado do grau é igual à $(n-1)$. Um fato que pode
 773 ser utilizado tanto em demonstrações quanto na simulação de grafos aleatórios, é que para
 774 valores grandes de n , a Equação 3.10 pode ser aproximada pela seguinte distribuição de
 775 Poisson Albert e Barabási (2002)

$$P(\mathcal{G}_k^{\text{rand}}) \simeq \frac{z^k}{k!} e^{-z}, \quad (3.12)$$

776 onde $z = E\{\mathcal{G}^{\text{rand}}\}$. Por esse motivo, é comum na literatura de redes complexas que grafos
 777 aleatórios desse tipo sejam denominados grafos aleatórios de Poisson.

778 **Diâmetro e distância média** O distância máxima em um grafo aleatório é uma prop-
 779riedade importante para o entendimento de sua eficiência. Um resultado verificado por
 780 vários autores (ver Luczak (1998); Bollobás (2001); Chung e Lu (2001)) é que a distância
 781 média em um grafo aleatório, para valores de p não muito pequenos, possui a seguinte
 782 relação de proporcionalidade

$$E\{d(G_{\text{rand}})\} \propto \frac{\log n}{\log E\{\mathcal{G}^{\text{rand}}\}}, \quad (3.13)$$

783 que possibilita a seguinte conjectura sobre valor esperado da distância em um grafo
 784 aleatório Albert e Barabási (2002)

$$E\{\mathcal{S}^{\text{rand}}\} \propto \frac{\log n}{\log E\{\mathcal{G}^{\text{rand}}\}}. \quad (3.14)$$

785 Este fato possibilita concluir que grafos aleatórios possuem uma distância média pequena
 786 em relação à quantidade de nós. Especificamente, o distância média é logarítmica em
 787 função da quantidade de nós, o que indica que grafos aleatórios também são eficientes.

788 **Coefficiente de agrupamento da rede** Em um grafo aleatório, o coeficiente de agrupa-
 789mento esperado pode ser obtido com base no valor esperado de vizinhos e na probabilidade
 790 de conexão p . O valor esperado de ligações entre os vizinhos de qualquer nó do grafo pode
 791 ser definido como

$$E\{\mathcal{D}^{\text{rand}}\} = p E\{\mathcal{G}^{\text{rand}}\} (E\{\mathcal{G}^{\text{rand}}\} - 1), \quad (3.15)$$

792 onde $\mathcal{D}^{\text{rand}}$ representa a variável aleatória que indica a quantidade arestas entre os vizi-
 793nhos de um nó qualquer em um grafo aleatório e o fator multiplicado por p representa
 794 a quantidade máxima de arestas entre os vizinhos de um nó qualquer. Para encontrar o

coeficiente de agrupamento, é suficiente dividir $E\{\mathcal{D}^{\text{rand}}\}$ pela quantidade total possível de arestas entre os vizinhos, ou seja, $E\{\mathcal{G}^{\text{rand}}\}(E\{\mathcal{G}^{\text{rand}}\} - 1)$, que resulta em

$$E\{\mathcal{C}^{\text{rand}}\} = p. \quad (3.16)$$

Dessa forma, é possível verificar que o agrupamento da rede depende diretamente, e unicamente, da probabilidade de conexão p .

Considerações A teoria de grafos aleatórios consistiu no estado da arte do estudo de redes complexas por décadas até a descoberta de alguns princípios de organização em redes reais, que naturalmente contrastam com sua natureza puramente aleatória. Porém, a teoria de grafos aleatórios ainda é significativa, uma vez que algumas propriedades desses novos modelos também são presentes em grafos aleatórios.

3.4 Mundo pequeno

Watts e Strogatz (1998) propuseram um modelo, denominado modelo de mundo pequeno (ou *small-world*)^[ii], em que é possível, por meio do controle de um parâmetro p , interpolar entre a construção de uma estrutura regular e a estrutura de um grafo aleatório. Nesse modelo, é construído um grafo regular inicial com n nós, onde existe uma aresta entre cada nó e seus h vizinhos. Considerando os extremos, para $p = 0$ não há alteração na estrutura regular inicial e para $p = 1$ a estrutura se comporta de forma similar à um grafo aleatório. Porém, mesmo para $p = 1$ algumas propriedades do grafo de mundo pequeno difere da contrapartida aleatória discutida na Seção 3.3. O algoritmo que descreve a construção desse tipo de grafo é apresentado a seguir.

Algoritmo 3.2 (Criação de uma rede de mundo pequeno). Criação da lista de nós e de arestas de um grafo de mundo pequeno a partir da quantidade de nós n , da quantidade de vizinhos h e da probabilidade de religação p .

```

algoritmo grafo-mundo-pequeno( $n, h, p$ )
1: {Quantidade de nós  $n$ , grau inicial  $h$  e probabilidade de reconexão  $p$ .}
2:  $G \leftarrow$  grafo-regular( $n, h$ ) {Cria grafo regular  $G = \langle N, E \rangle$ .}
3: para cada aresta  $e \in E$  faça
4:   se distribuicao-uniforme(0, 1) <  $p$  então
5:      $e \leftarrow$  redefinir( $e$ ) {Redefinir conexão evitando laços e arestas múltiplas.}
6:   fim se
7: fim para
8: retorne  $G = \langle N, E \rangle$  {Grafo de mundo pequeno resultante.}
```

Considerando o laço de repetição da linha 3 e a quantidade de arestas do grafo, tanto a complexidade em tempo como a em memória do algoritmo é da ordem de $\Theta(nh)$. \square

O Algoritmo 3.2 possui complexidade assintótica em tempo dependente da quantidade de iterações do procedimento ‘grafo-regular()’, que têm como princípio criar os n nós e as $n(h/2)$ arestas (considerado h par). A complexidade em tempo desse procedimento é da ordem de $\Theta(nh)$. O mesmo acontece com o laço de repetição da linha 3, que deve iterar sobre todo o conjunto de arestas, cujo tamanho é da ordem de $n(h/2)$, considerando que

^[ii]Esse modelo possui esse nome porque é baseado no trabalho de Milgram (1967), que apresenta uma tendência natural da separação de indivíduos por uma pequeno quantidade de conhecidos em comum.

o procedimento ‘redefinir()’ é de ordem constante. Portanto o algoritmo possui complexidade em tempo da ordem de $\Theta(nh)$. A complexidade em memória do algoritmo depende da quantidade de arestas, uma vez que se valor mínimo é de $(n - 1)$. Dessa forma, a complexidade em memória é também da ordem de $\Theta(nh)$. Na Figura 3.4, é apresentada uma ilustração da construção de uma rede desse tipo.

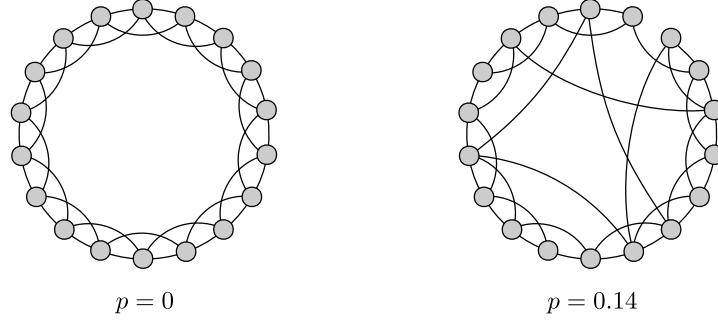


Figura 3.4: Exemplo de rede de mundo pequeno com $n = 18$ nós. Em que o número de vizinhos $h = 4$ e a probabilidade de reconexão $p = 0.14$.

No modelo representado pela Figura 3.4 é construído um grafo regular inicial com $n = 18$ e $h = 4$. O grafo regular é apresentado à esquerda da Figura. Posteriormente, com probabilidade $p = 0.14$, cada uma das arestas tem seu destino alterado.

Propriedades Inicialmente, antes da religação das arestas do grafo, todos os nós possuem o mesmo grau h . Dessa forma, a distribuição dos graus pode ser representada por um impulso unitário em h , ou seja,

$$P(\mathcal{G}_k^{\text{small}} \mid p = 0) = \delta(k - h). \quad (3.17)$$

Porém, quando p se aproxima da unidade a distribuição dos graus se aproxima da mesma distribuição de graus encontrada em grafos aleatórios [Albert e Barabási \(2002\)](#). Já a distância média quando $p = 0$ em um grafo de mundo pequeno pode ser definido assintoticamente como [Newman \(2003\)](#)

$$E\{\mathcal{S}^{\text{small}} \mid p = 0\} \simeq \frac{n}{4h}, \quad (3.18)$$

que está em desacordo com o conceito matemático de mundo pequeno, onde espera-se que $E\{\mathcal{S}\} \propto \log n$. Porém, [Barrat e Weigt \(2000\)](#) descobriram a existência da seguinte relação de transição de regime de distância média em função do produto (pnh)

$$E\{\mathcal{S}^{\text{small}}\} \propto \begin{cases} n & \text{se } pnh \ll 1 \\ \log n & \text{se } pnh \gg 1 \end{cases}, \quad (3.19)$$

que resulta no fato de que, para valores não muito pequenos de p , a distância média em um grafo de mundo pequeno é logarítmica em relação à quantidade de nós^[iii]. Finalmente, em relação ao coeficiente de agrupamento, tem-se que o grafo de mundo pequeno apresentada a seguinte relação [Barrat e Weigt \(2000\)](#)

$$E\{\mathcal{C}^{\text{small}}\} \simeq \frac{3(h-1)}{2(2h-1)}(1-p)^3, \quad (3.20)$$

^[iii] A partir desse conceito tem-se que grafos aleatórios também são de mundo pequeno, considerando valores não muito pequenos da probabilidade de conexão em grafos aleatórios.

que mostra que, para $p = 0$, o agrupamento tende a $3/4$ para valores grandes de h . Já para valores de p próximo de 1, o coeficiente de agrupamento se aproxima de 0. Assim, para valores não muito pequenos de p , um grafo de mundo pequeno possui uma distribuição dos graus similar ao de grafos aleatórios, uma distância média de ordem logarítmica e um coeficiente de agrupamento relativamente alto.

3.5 Livre de escala

Barabási e Albert (1999) verificaram que grande parte das redes reais possuem uma distribuição dos graus que parecia seguir uma lei de potência, ou seja, a probabilidade de um nó possuir grau k é proporcional à potência $k^{-\alpha}$, para alguma constante α . Isso não acontece com os modelos de mundo pequeno e de grafo aleatório, onde essa distribuição se aproxima de uma exponencial, por exemplo, a da Equação 3.12. Uma das principais consequências desse comportamento é que a probabilidade de existirem nós com grau elevado é significativamente maior que nos outros dois modelos. Para se criar um modelo que reproduzisse esse resultado, criou-se um procedimento incremental de inserção de nós na rede, em que há uma ligação preferencial (*preferential attachment*) dos novos nós em relação aos nós existentes com maior grau.

A cada iteração do processo de criação do grafo, deve-se considerar a quantidade de arestas que serão criadas para o novo nó, representada por m . Para $m = 1$, tem-se a criação de estruturas do tipo árvore, como a apresentada na Figura 3.5. O processo de criação de um grafo livre de escala é representado pelo algoritmo a seguir.

Algoritmo 3.3 (Criação de uma rede livre de escala). Criação da lista de nós e de arestas de um grafo livre de escala a partir da quantidade de nós n , da quantidade de arestas por inserção m .

```

algoritmo grafo-livre-de-escala( $n, m$ )
1: {Quantidade de nós  $n$  e quantidade de arestas por inserção  $m$ .}
2:  $N \leftarrow$  lista-de-nos( $m$ ) {Cria uma lista com  $m$  nós com grau nulo.}
3:  $E \leftarrow$  lista-de-arestas() {Cria uma nova lista vazia.}
4: enquanto  $|N| < n$  faça
5:    $h \leftarrow$  sortear-vizinhos( $N, m$ ) {Escolha preferencial de vizinhos com maior grau.}
6:    $i \leftarrow |N| + 1$ 
7:   inserir( $N, v_i$ )
8:   para cada  $v_j \in h$  faça
9:     inserir( $E, (v_i, v_j)$ )
10:  fim para
11: fim enquanto
12: retorne  $G = \langle N, E \rangle$  {Grafo livre de escala resultante.}

```

Considerando o procedimento de sorteio da linha 5, a complexidade em tempo é da ordem de $\Theta(mn^2)$. Já a complexidade em memória depende da quantidade de arestas, que é da ordem de $\Theta(mn)$. \square

A complexidade em tempo do Algoritmo 3.3 depende diretamente da escolha das arestas considerando a ligação preferencial, ou seja, do procedimento ‘sortear-vizinhos()’. Esse procedimento deverá sortear m vizinhos, considerando a escolha de um dentre os $(n - 1)$ possíveis. Desconsiderando a possibilidade de sorteios com resultados iguais, a complexidade do procedimento ‘sortear-vizinhos()’ deve ser da ordem de $\Theta(mn)$. Essa complexi-

899 dade é associada à do laço de repetição da linha 4, que executa $(n - m + 1)$ vezes, e faz com
 900 que a complexidade em tempo do algoritmo seja da ordem de $\Theta(mn^2)$. Já a complexidade
 901 em memória depende exclusivamente da quantidade de arestas do grafo. Considerando
 902 grafos não direcionados livre de escala, criados com base no Algoritmo 3.3, essa quantidade
 903 é de exatamente

$$|E_{\text{free}}| = m(n - m) \quad (3.21)$$

904 arestas. Como $m \ll n$, a complexidade em memória é da ordem de $\Theta(mn)$.

905 Para ilustrar a evolução da criação de uma rede livre de escala a partir do Algoritmo 3.3,
 906 um exemplo da criação de um grafo é apresentado na Figura 3.5.

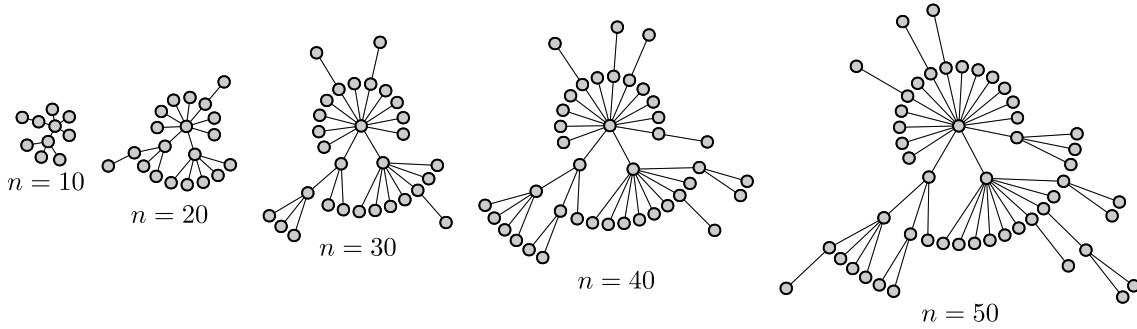


Figura 3.5: Exemplo da evolução de uma rede livre de escala. Representação da rede para valores de $n = \{10, 20, 30, 40, 50\}$, considerando $m = 1$.

907 Na Figura 3.5, a rede livre de escala possui parâmetro $m = 1$. É possível verificar
 908 com a evolução da rede como os nós de maior grau possuem uma ligação preferencial. As
 909 demais propriedades desse modelo em relação às métricas são apresentadas a seguir.

910 **Distribuição de lei de potência** A característica que denomina uma rede livre de
 911 escala é a sua distribuição dos graus dos nós. Essa distribuição segue uma lei de potência,
 912 ou seja, $\mathcal{G}_k^{\text{free}} \sim k^{-\alpha}$, onde a probabilidade de nó qualquer de uma rede livre de escala
 913 possuir grau k é definida por

$$P(\mathcal{G}_k^{\text{free}}) \simeq ck^{-\alpha}, \quad (3.22)$$

914 onde α é uma constante que deve, necessariamente, ser maior que a unidade^[iv], a fim de
 915 garantir a convergência da distribuição e c é uma constante de normalização, necessária,
 916 uma vez que, por se tratar de uma distribuição discreta de probabilidade, deve-se verificar
 917 $\sum_{k=1}^{\infty} ck^{-\alpha} = 1$. O valor mínimo e máximo de k é m e $(n - 1)$, respectivamente, o que
 918 simplifica a restrição da distribuição para

$$c \sum_{k=m}^{n-m} k^{-\alpha} = 1, \quad (3.23)$$

919 que permite inferir que c deve ser o inverso multiplicativo do somatório. Com o objetivo de
 920 apresentar um forma comumente utilizada para se verificar se uma rede é livre de escala,
 921 são apresentadas as frequências dos graus de três redes livre de escala na Figura 3.6.

922 Por causa da dependência do valor inicial de k em função de m , na Figura 3.6, as
 923 seqüências de frequência para cada valor de m iniciam e terminam em valores distintos

^[iv]Para redes reais, geralmente é verificada a relação $2 < \alpha < 3$ Albert e Barabási (2002).

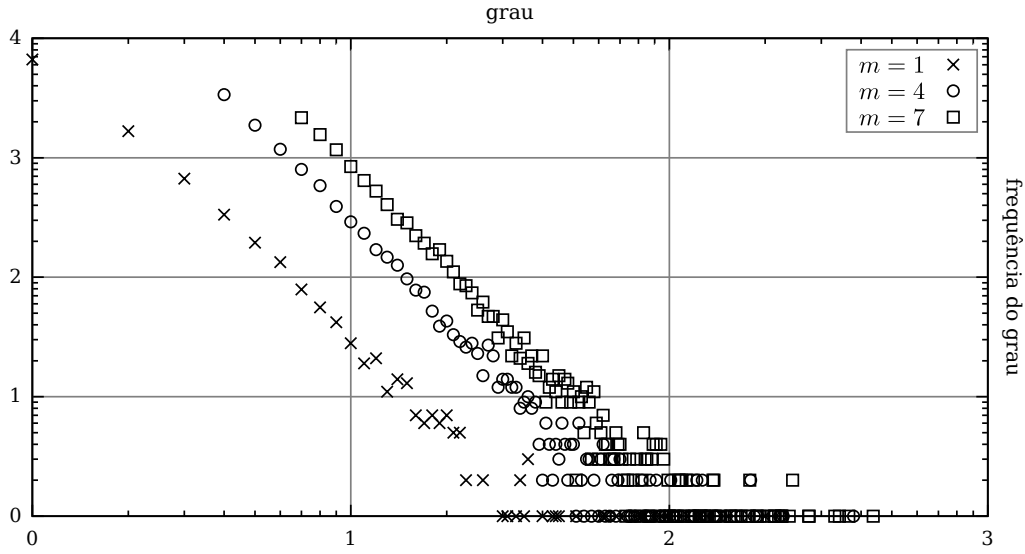


Figura 3.6: Distribuição dos graus dos nós em uma rede livre de escala considerando valores de $m = \{1, 4, 7\}$ em uma rede com 10^4 nós. As marcações indicam o expoente na base 10.

de k . Além disso, é possível verificar que, utilizando a escala logarítmica, distribuições que obedecem uma lei de potência possuem um comportamento linear. É verificando esse comportamento que, comumente, se classifica uma rede real como livre de escala^[v]. Em relação à distância média, tem-se que, pelo trabalho de [Bollobás e Riordan \(2004\)](#), em um grafo livre de escala a seguinte relação de proporcionalidade é verificada

$$E\{\mathcal{S}^{\text{free}}\} \propto \frac{\log n}{\log \log n}, \quad (3.24)$$

o que demonstra que grafos livre de escala também são coerentes com o conceito de mundo pequeno. Porém, o coeficiente de agrupamento da rede difere tanto do modelo de mundo pequeno como do de grafo aleatório. Isso porque seu valor é estabelecido pela seguinte relação

$$E\{\mathcal{C}^{\text{free}}\} \propto \frac{\log n}{n}, \quad (3.25)$$

que depende da quantidade de nós, como pode ser verificado no trabalho de [Eggemann e Noble \(2011\)](#), e converge para 0 para valores assintóticos de n .

^[v]Um estudo sobre verificação de distribuição de leis de potência é apresentado em [Clauset et al. \(2009\)](#).

Capítulo 4

Aplicações

*“If we can really understand the problem,
the answer will come out of it,
because the answer is not separate from the problem.”
Jiddu Krishnamurti*

Neste trabalho...

939 Apêndice A

940 Resultados complementares

941 Neste Apêndice, são apresentadas Definições, Teoremas, Lemas e Corolários desenvol-
942 vidos durante o trabalho e utilizados na tese. Nas Seções A.1 e A.2 são apresentados
943 resultados da caracterização e complexidades dos algoritmos para as métricas das redes
944 determinísticas, respectivamente.

945 A.1 Caracterização de redes determinísticas

946 As Definições a seguir servem de base para a demonstração dos resultados subsequentes.

947 **Definição A.1** (Grafo direcionado em anel). Um grafo direcionado $G = \langle N, E \rangle$ é dito
948 em anel se é possível rotular seus nós de forma a satisfazer as sentenças:

- 949 1. $((v_1, v_n) \in E) \wedge ((v_n, v_1) \in E)$,
- 950 2. $(\forall v_i \in N)((g_{\text{in}}(v_i) = 2) \wedge (g_{\text{out}}(v_i) = 2))$,
- 951 3. $(\forall v_i \in N)((1 \leq i < n) \Leftrightarrow ((v_i, v_{i+1}) \in E))$ e
- 952 4. $(\forall v_i \in N)((1 < i \leq n) \Leftrightarrow ((v_i, v_{i-1}) \in E))$;

953 onde $n = |N|$ representa a cardinalidade do conjunto de nós $N = \{v_1, v_2, \dots, v_n\}$ e $g_{\text{in}}(v_i)$
954 e $g_{\text{out}}(v_i)$ representam o grau de entrada e saída de v_i , respectivamente. \square

955 **Definição A.2** (Grafo direcionado em estrela). Um grafo direcionado $G = \langle N, E \rangle$ é dito
956 em estrela se é possível rotular seus nós de forma a satisfazer as sentenças:

- 957 1. $(\forall v_i \in N)((1 < i \leq n) \Leftrightarrow (((v_1, v_i) \in E) \wedge ((v_i, v_1) \in E)))$ e
- 958 2. $(\forall v_i \in N)((1 < i \leq n) \Leftrightarrow ((g_{\text{in}}(v_i) = 2) \wedge (g_{\text{out}}(v_i) = 2)))$;

959 onde $n = |N|$ representa a cardinalidade do conjunto de nós $N = \{v_1, v_2, \dots, v_n\}$ e $g_{\text{in}}(v_i)$
960 e $g_{\text{out}}(v_i)$ representam o grau de entrada e saída de v_i , respectivamente. \square

961 **Definição A.3** (Grafo direcionado em linha). Um grafo direcionado $G = \langle N, E \rangle$ é dito
962 em linha se é possível rotular seus nós de forma a satisfazer as sentenças:

- 963 1. $((v_1, v_n) \notin E) \wedge ((v_n, v_1) \notin E)$,
- 964 2. $(\forall v_i \in N)((1 \leq i < n) \Leftrightarrow ((v_i, v_{i+1}) \in E))$,
- 965 3. $(\forall v_j \in N)((1 < j \leq n) \Leftrightarrow ((v_j, v_{j-1}) \in E))$ e
- 966 4. $(\forall v_i \in N)((1 < i < n) \Leftrightarrow ((g_{\text{in}}(v_i) = 2) \wedge (g_{\text{out}}(v_i) = 2)))$;

967 onde $n = |N|$ representa a cardinalidade do conjunto de nós $N = \{v_1, v_2, \dots, v_n\}$ e $g_{\text{in}}(v_i)$
968 e $g_{\text{out}}(v_i)$ representam o grau de entrada e saída de v_i , respectivamente. \square

969 Distância média na topologia em anel

970 **Lema A.1** (Distância média na topologia em anel). A distância média de uma topo-
 971 logia descrita por um grafo direcionado $G = \langle N, E \rangle$ em anel, em conformidade com a
 972 Definição A.1, é dado por

$$E\{\mathcal{S}^{\text{ring}}\} = \frac{1}{2n} \left[\left\lfloor \frac{n}{2} \right\rfloor \left(\left\lfloor \frac{n}{2} \right\rfloor + 1 \right) + \left\lfloor \frac{(n-1)}{2} \right\rfloor \left(\left\lfloor \frac{(n-1)}{2} \right\rfloor + 1 \right) \right], \quad (\text{A.1})$$

973 onde $n = |N|$ representa a quantidade de nós do grafo. \square

974 *Demonstração.* Devido ao carácter circular da topologia em anel, qualquer nó do grafo
 975 terá a mesma soma total de distâncias para os demais nós da rede. Considerando uma
 976 rede com $n = |N|$ nós, a sequência de distâncias de cada nó pode ser representada por
 977 $(1+2+\dots+(n-1)/2+(n-1)/2+\dots+2+1)$ se n for ímpar e $(1+2+\dots+n/2+\dots+2+1)$
 978 se n for par. Essas sequências podem ser representadas pelo seguinte par de somatórios

$$\left(\sum_{i=1}^{\lfloor n/2 \rfloor} i + \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} i \right),$$

979 uma vez que $\lfloor n/2 \rfloor = \lfloor (n-1)/2 \rfloor$ se n é ímpar e $\lfloor n/2 \rfloor = \lfloor (n-1)/2 \rfloor + 1$ se n é par.
 980 Considerando a Definição 2.7 e que a rede possui n nós, a distância média na topologia
 981 em anel é representada por

$$E\{\mathcal{S}^{\text{ring}}\} = \frac{1}{n^2} \left[n \left(\sum_{i=1}^{\lfloor n/2 \rfloor} i + \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} i \right) \right].$$

982 Como os somatórios caracterizam uma progressão aritmética de ordem 1, é possível rees-
 983 crever cada somatório utilizando a relação $\sum_{i=1}^x i = x(x+1)/2$, de forma que

$$E\{\mathcal{S}^{\text{ring}}\} = \frac{1}{2n} \left[\left\lfloor \frac{n}{2} \right\rfloor \left(\left\lfloor \frac{n}{2} \right\rfloor + 1 \right) + \left\lfloor \frac{(n-1)}{2} \right\rfloor \left(\left\lfloor \frac{(n-1)}{2} \right\rfloor + 1 \right) \right],$$

984 onde o termo $n/2$ de cada somatório foi colocado em evidência. **C.Q.D.**

985 **Corolário A.1** (Influência da paridade de $|N|$ em $E\{\mathcal{S}^{\text{ring}}\}$). Considerando a possibili-
 986 dade de valores pares e ímpares para a quantidade de nós, tem-se que a distância média
 987 em uma topologia em anel definida no Lema A.1 pode ser expressa por

$$E\{\mathcal{S}^{\text{ring}} \mid n \text{ é par}\} = \frac{n}{4} \quad (\text{A.2})$$

$$E\{\mathcal{S}^{\text{ring}} \mid n \text{ é ímpar}\} = \frac{n}{4} - \frac{1}{4n}, \quad (\text{A.3})$$

988 onde $n = |N|$ representa a quantidade de nós do grafo. \square

989 *Demonstração.* Pode ser demonstrado a partir das seguintes substituições na Equação A.1:
 990 (i) $\lfloor n/2 \rfloor$ por $n/2$ e $\lfloor (n-1)/2 \rfloor$ por $(n-2)/2$ quando n é par; e (ii) $\lfloor n/2 \rfloor$ por $(n-1)/2$
 991 e $\lfloor (n-1)/2 \rfloor$ por $(n-1)/2$ quando n é ímpar. **C.Q.D.**

992 **Corolário A.2** (Comportamento assintótico de $E\{\mathcal{S}^{\text{ring}}\}$). Em uma topologia em anel,
 993 a distância média possui um comportamento assintótico linear dado pela expressão

$$E\{\mathcal{S}^{\text{ring}}\} \simeq \frac{n}{4} \quad (\text{A.4})$$

994 onde $n = |N|$ representa a quantidade de nós do grafo. \square

995 *Demonstração.* Quando n é par tem-se a igualdade de fato e, portanto, a igualdade as-
 996 sintótica também é válida. Já para o caso em que n é ímpar deve-se mostrar que o
 997 valor esperado $E\{\mathcal{S}^{\text{ring}} \mid n \text{ é ímpar}\}$ é assintoticamente igual à $n/4$. Isso é verdade se
 998 $\lim_{n \rightarrow \infty} 4 E\{\mathcal{S}^{\text{ring}} \mid n \text{ é ímpar}\}/n = 1$. Utilizando a Equação A.3 tem-se que

$$\lim_{n \rightarrow \infty} \frac{4}{n} \left(\frac{n}{4} - \frac{1}{4n} \right)$$

999 deve ser unitário. Distribuindo o fator em evidência e o limite, é possível reescrever a
 1000 expressão como

$$\lim_{n \rightarrow \infty} \frac{4}{n} \left(\frac{n}{4} - \frac{1}{4n} \right) = 1 - \lim_{n \rightarrow \infty} \frac{1}{n^2},$$

1001 onde o limite a direita converge para zero, fazendo com que $E\{\mathcal{S}^{\text{ring}}\} \simeq n/4$. **C.Q.D.**

1002 Distância média na topologia em estrela

1003 **Lema A.2** (Distância média na topologia em estrela). A distância média de uma to-
 1004 pologia descrita por um grafo direcionado $G = \langle N, E \rangle$ em estrela, como apresenta a
 1005 Definição A.2, é dado por

$$E\{\mathcal{S}^{\text{star}}\} = \frac{2}{n} \left(n - 2 + \frac{1}{n} \right), \quad (\text{A.5})$$

1006 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1007 *Demonstração.* As distâncias entre os nós de uma topologia em estrela podem ser dividas
 1008 em relação ao nó central e entre os demais nós. Nesse sentido, tem-se $2(n-1)$ distâncias
 1009 de tamanho 1 em direção e a partir do nó central e $(n-1)(n-2)$ distâncias de tamanho
 1010 2. Utilizando a definição de distância média é possível descrevê-la no caso da topologia
 1011 em estrela da seguinte forma

$$E\{\mathcal{S}^{\text{star}}\} = \frac{1}{n^2} [2(n-1) + 2(n-1)(n-2)],$$

1012 onde, ao se evidenciar a constante, desenvolver o produto na parte interna dos colchetes e
 1013 distribuir o denominador em evidência na parte interna, é possível reescrever a expressão
 1014 como

$$E\{\mathcal{S}^{\text{star}}\} = 2 \left(1 - \frac{2}{n} + \frac{1}{n^2} \right), \quad (\text{A.6})$$

1015 de onde pode-se obter a Equação A.5 ao se colocar o termo $1/n$ em evidência. **C.Q.D.**

1016 **Corolário A.3** (Comportamento assintótico de $E\{\mathcal{S}^{\text{star}}\}$). Em uma topologia em estrela
 1017 a distância média possui um comportamento assintótico linear dado pela expressão

$$E\{\mathcal{S}^{\text{star}}\} \simeq 2, \quad (\text{A.7})$$

1018 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1019 *Demonstração.* Pode ser demonstrado a partir da aplicação do limite quando n tende ao
 1020 infinito na Equação A.6, onde na parte interna dos parênteses resta a unidade, que mul-
 1021 tiplicada com o termo em evidência resulta em $E\{\mathcal{S}^{\text{star}}\} \simeq 2$. **C.Q.D.**

1022 Distância média na topologia em linha

1023 **Lema A.3** (Distância média na topologia em linha). A distância média de uma topologia
 1024 descrita por um grafo direcionado $G = \langle N, E \rangle$ em linha, como apresenta a Definição A.3,
 1025 é dado por

$$E\{\mathcal{S}^{\text{line}}\} = \frac{1}{3} \left(n - \frac{1}{n} \right), \quad (\text{A.8})$$

1026 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1027 *Demonstração.* Na topologia em linha, a distância de cada nó em relação aos demais pode
 1028 ser escrita como um somatório sobre a quantidade de nós à esquerda e à direita do nó em
 1029 questão. Dado um nó qualquer v_i , a quantidade de nós à esquerda é de $(i - 1)$, e portanto
 1030 o somatório das distâncias à esquerda pode ser representado por $\sum_{k=1}^{i-1} k$. Por sua vez, a
 1031 quantidade de nós à direita é de $(n - i)$ e, portanto, o somatório das distâncias à direita
 1032 pode ser representado por $\sum_{j=1}^{n-i} j$. Considerando a soma das distâncias para cada um dos
 1033 nós, a distância média em uma topologia linear pode ser representada por

$$E\{\mathcal{S}^{\text{line}}\} = \frac{1}{n^2} \sum_{i=1}^n \left(\sum_{k=1}^{i-1} k + \sum_{j=1}^{n-i} j \right),$$

1034 que pode ser reescrito da seguinte forma a partir da solução dos somatórios interiores

$$E\{\mathcal{S}^{\text{line}}\} = \frac{1}{n^2} \sum_{i=1}^n [(n - i)^2 + (n - i) + (i - 1)^2 + (i - 1)].$$

1035 A partir da solução de cada termo do somatório e da combinação dos resultados, tem-se
 1036 a seguinte expressão resultante

$$E\{\mathcal{S}^{\text{line}}\} = \frac{1}{n^2} \left[\frac{1}{3}(n^3 - n) \right],$$

1037 que pode ser reescrita como

$$E\{\mathcal{S}^{\text{line}}\} = \frac{1}{3} \left(n - \frac{1}{n} \right),$$

1038 a partir da distribuição do fator mais à esquerda. **C.Q.D.**

1039 **Corolário A.4** (Comportamento assintótico de $E\{\mathcal{S}^{\text{line}}\}$). Em uma topologia em linha,
 1040 a distância média possui um comportamento assintótico linear dado por

$$E\{\mathcal{S}^{\text{line}}\} \simeq \frac{n}{3} \quad (\text{A.9})$$

1041 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1042 *Demonstração.* Se $\lim_{n \rightarrow \infty} 3E\{\mathcal{S}^{\text{line}}\}/n = 1$, então a Equação A.9 é verdadeira. Utili-
 1043 zando a Equação A.8, tem-se que

$$\lim_{n \rightarrow \infty} \frac{3}{n} \left[\frac{1}{3} \left(n - \frac{1}{n} \right) \right],$$

1044 deve ser unitário. Distribuindo o fator em evidência e o limite, é possível reescrever a
 1045 expressão como

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(n - \frac{1}{n} \right) = 1 - \lim_{n \rightarrow \infty} \frac{1}{n^2},$$

1046 onde o limite à direita converge para zero, fazendo com que $E\{\mathcal{S}^{\text{line}}\} \simeq n/3$. **C.Q.D.**

1047 Eficiência média na topologia em anel

1048 **Lema A.4** (Eficiência média na topologia em anel). A eficiência média de uma topologia
 1049 descrita por um grafo direcionado $G = \langle N, E \rangle$ em anel, como apresenta a Definição A.1,
 1050 é dado por

$$E\{\mathcal{F}^{\text{ring}}\} = \frac{1}{(n-1)} (H_{\lfloor n/2 \rfloor} + H_{\lfloor (n-1)/2 \rfloor}), \quad (\text{A.10})$$

1051 onde $n = |N|$ representa a quantidade de nós do grafo e H_n a soma dos n primeiros termos
 1052 da série harmônica (definida na Seção de Simbologia do Glossário). \square

1053 *Demonstração.* Devido ao carácter circular da topologia em anel, qualquer nó do grafo
 1054 tem a mesma soma total de eficiências para os demais nós da rede. Considerando uma
 1055 rede com $n = |N|$ nós, a sequência de distâncias de cada nó pode ser representada por
 1056 $(1+2+\dots+(n-1)/2+(n-1)/2+\dots+2+1)$ se n for ímpar e $(1+2+\dots+n/2+\dots+2+1)$ se
 1057 n for par. Assim, é possível definir a eficiência média, de forma geral, com essas sequências
 1058 representadas pelo par de somatórios

$$E\{\mathcal{F}^{\text{ring}}\} = \frac{1}{n(n-1)} \left[n \left(\sum_{i=1}^{\lfloor n/2 \rfloor} \frac{1}{i} + \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} \frac{1}{i} \right) \right],$$

1059 de onde, a partir da evidenciação do termo n e utilizando para representação do somatório
 1060 a notação de série harmônica, $H_{\lfloor n/2 \rfloor}$ e $H_{\lfloor (n-1)/2 \rfloor}$, respectivamente, têm-se como resultado
 1061 a Equação A.10. **C.Q.D.**

1062 **Corolário A.5** (Influência da paridade de $|N|$ em $E\{\mathcal{F}^{\text{ring}}\}$). Considerando a possibi-
 1063 lidade de valores pares e ímpares para a quantidade de nós na topologia em anel, tem-se
 1064 que a eficiência média da topologia em anel, definida no Lema A.4, pode ser escrita, para
 1065 cada caso, como

$$E\{\mathcal{F}^{\text{ring}} \mid n \text{ é ímpar}\} = \frac{2}{(n-1)} (H_{((n-1)/2)}) \quad (\text{A.11})$$

$$E\{\mathcal{F}^{\text{ring}} \mid n \text{ é par}\} = \frac{1}{(n-1)} (H_{(n/2)} + H_{((n-2)/2)}), \quad (\text{A.12})$$

1066 onde $n = |N|$ representa a quantidade de nós do grafo e H_n a soma dos n primeiros termos
 1067 da série harmônica (definida na Seção de Simbologia do Glossário). \square

1068 *Demonstração.* A demonstração é feita a partir da substituição de $H_{\lfloor n/2 \rfloor}$ e $H_{\lfloor (n-1)/2 \rfloor}$, por
 1069 $H_{((n-1)/2)}$ quando n é ímpar. E da substituição de $H_{\lfloor n/2 \rfloor}$ por $H_{(n/2)}$, e $H_{\lfloor (n-1)/2 \rfloor}$, por
 1070 $H_{((n-2)/2)}$, quando n é par. **C.Q.D.**

1071 **Corolário A.6** (Comportamento assintótico de $E\{\mathcal{F}^{\text{ring}}\}$). Em uma topologia em anel,
 1072 a eficiência média possui um comportamento assintótico que pode ser escrita como

$$E\{\mathcal{F}^{\text{ring}}\} \simeq \frac{2 \log n}{n}, \quad (\text{A.13})$$

1073 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1074 *Demonstração.* Utilizando a relação $H_n \simeq \log n + \gamma$, é possível reescrever a Equação A.10
 1075 da seguinte forma

$$E\{\mathcal{F}^{\text{ring}}\} \simeq \frac{1}{(n-1)} \left(\log \left\lfloor \frac{n}{2} \right\rfloor + \gamma + \log \left\lfloor \frac{n-1}{2} \right\rfloor + \gamma \right). \quad (\text{A.14})$$

1076 Se $\lim_{n \rightarrow \infty} E\{\mathcal{F}^{\text{ring}}\}/f(n) = 1$, então $E\{\mathcal{F}^{\text{ring}}\} \simeq f(n)$. Como $E\{\mathcal{F}^{\text{ring}}\}$ depende da
 1077 paridade de n , serão considerados os dois casos. Considerando que a quantidade de nós
 1078 $n = |N|$ é par, tem-se que

$$E\{\mathcal{F}^{\text{ring}} \mid n \text{ é par}\} \simeq \frac{1}{(n-1)} \left[\log \left(\frac{n}{2} \right) + \log \left(\frac{n-2}{2} \right) + 2\gamma \right],$$

1079 onde é possível aplicar a identidade $\log x/y = \log x - \log y$, resultando em

$$E\{\mathcal{F}^{\text{ring}} \mid n \text{ é par}\} \simeq \frac{1}{(n-1)} (\log n + \log(n-2) + 2(\gamma - \log 2)).$$

1080 Para se verificar a igualdade assintótica, considera-se $f(n) = (2 \log n)/n$. Portanto o limite
 1081 pode ser descrito como

$$\lim_{n \rightarrow \infty} \frac{n}{2 \log n} \frac{1}{(n-1)} (\log n + \log(n-2) + 2(\gamma - \log 2)),$$

1082 que pode ser representado pela soma dos limites

$$\lim_{n \rightarrow \infty} \frac{n}{2 \log n} \frac{\log n}{n-1} + \lim_{n \rightarrow \infty} \frac{n}{2 \log n} \frac{\log(n-2)}{n-1} + \lim_{n \rightarrow \infty} \frac{n}{2 \log n} \frac{2(\gamma - \log 2)}{n-1} = \frac{1}{2} + \frac{1}{2} + 0 = 1.$$

1083 Enquanto os dois primeiros limites convergem para $1/2$, o terceiro converge para 0. De
 1084 forma que o resultado é a unidade, confirmando a igualdade assintótica. Para o segundo
 1085 caso, quando n é ímpar, temos que

$$E\{\mathcal{F}^{\text{ring}} \mid n \text{ é ímpar}\} \simeq \frac{2}{(n-1)} (\log(n-1) + \gamma - \log 2),$$

1086 onde a igualdade assintótica depende do limite que pode ser descrito como

$$\lim_{n \rightarrow \infty} \frac{n}{2 \log n} \frac{2}{(n-1)} (\log(n-1) + (\gamma - \log 2)),$$

1087 que pode ser representado pela soma dos limites

$$\lim_{n \rightarrow \infty} \frac{n}{\log n} \frac{\log(n-1)}{(n-1)} + \lim_{n \rightarrow \infty} \frac{n}{\log n} \frac{\gamma - \log 2}{(n-1)} = 1 + 0 = 1,$$

1088 onde o primeiro limite converge para 1 enquanto o segundo converge para 0. Portanto,
 1089 $E\{\mathcal{F}^{\text{ring}}\} \simeq (2 \log n)/n$. **C.Q.D.**

1090 Eficiência média na topologia em estrela

1091 **Lema A.5** (Eficiência média na topologia em estrela). A eficiência média de uma to-
 1092 pologia descrita por um grafo direcionado $G = \langle N, E \rangle$ em estrela, como apresenta a
 1093 Definição A.2, é dada por

$$E\{\mathcal{F}^{\text{star}}\} = \frac{1}{2} + \frac{1}{n}, \quad (\text{A.15})$$

1094 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1095 *Demonstração.* As eficiências entre os nós de uma topologia em estrela podem ser dividas
 1096 em relação ao nó central e dentre os demais nós já que tem-se $2(n-1)$ eficiências unitárias
 1097 em direção e a partir do nó central e $(n-1)(n-2)$ eficiências de valor $1/2$. Utilizando
 1098 a definição de eficiência média é possível descrevê-la, no caso da topologia em estrela, da
 1099 seguinte forma

$$E\{\mathcal{F}^{\text{star}}\} = \frac{1}{n(n-1)} \left[2(n-1) + \frac{1}{2}(n-1)(n-2) \right],$$

1100 onde é possível por em evidência o termo $(n-1)$, cancelando com o respectivo denominador
 1101 na fração a esquerda, resultando em

$$E\{\mathcal{F}^{\text{star}}\} = \frac{1}{n} \left[2 + \frac{1}{2}(n-2) \right],$$

1102 que, a partir da distribuição dos termos em evidência, pode ser reescrito como

$$E\{\mathcal{F}^{\text{star}}\} = \frac{1}{n} \left(1 + \frac{n}{2} \right),$$

1103 de onde pode-se verificar a equivalência com a Equação A.15. **C.Q.D.**

1104 **Corolário A.7** (Comportamento assintótico de $E\{\mathcal{F}^{\text{star}}\}$). Em uma topologia em es-
 1105 trela, a eficiência média possui um comportamento assintótico que pode ser escrito como

$$E\{\mathcal{F}^{\text{star}}\} \simeq \frac{1}{2}, \quad (\text{A.16})$$

1107 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1108 *Demonstração.* Ao se considerar valores assintóticos de n o segundo termo da Equação A.15
 1109 tende a 0, restando apenas a fração $1/2$. **C.Q.D.**

1110 Eficiência média na topologia em linha

1111 **Lema A.6** (Eficiência média na topologia em linha). A distância média de uma topologia
 1112 descrita por um grafo direcionado $G = \langle N, E \rangle$ em linha, como apresenta a Definição A.3,
 1113 é dada por

$$E\{\mathcal{F}^{\text{line}}\} = \frac{2}{(n-1)} (H_n - 1), \quad (\text{A.17})$$

1114 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1115 *Demonstração.* Na topologia em linha, a eficiência de cada nó em relação aos demais pode
 1116 ser expressa como um somatório do inverso multiplicativo das quantidade de nós à esquerda
 1117 e à direita do nó em questão. Dado um nó qualquer v_i , a quantidade de nós à esquerda é
 1118 de $(i - 1)$ e, portanto, o somatório das eficiências à esquerda pode ser representado por
 1119 $\sum_{k=1}^{i-1} 1/k$. Por sua vez, a quantidade de nós à direita é de $(n - i)$ e, portanto, o somatório
 1120 das distâncias à direita pode ser representado por $\sum_{j=1}^{n-i} 1/j$. Considerando a soma das
 1121 eficiências para cada um dos nós, a eficiência média em uma topologia linear pode ser
 1122 representada por

$$E\{\mathcal{F}^{\text{line}}\} = \frac{1}{n(n-1)} \sum_{i=1}^n \left(\sum_{k=1}^{i-1} \frac{1}{k} + \sum_{j=1}^{n-i} \frac{1}{j} \right),$$

1123 que, utilizando a representação por série harmônica dos somatórios interiores, pode ser
 1124 reescrito da seguinte forma

$$E\{\mathcal{F}^{\text{line}}\} = \frac{1}{n(n-1)} \sum_{i=1}^n (H_{(i-1)} + H_{(n-i)}),$$

1125 onde os somatórios podem ser reescritos utilizando a relação $\sum_{i=1}^n H_{(i-1)} = \sum_{i=1}^{n-1} H_i$ e
 1126 $\sum_{i=1}^n H_{(n-i)} = \sum_{i=1}^{n-1} H_i$, resultando em

$$E\{\mathcal{F}^{\text{line}}\} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} H_i,$$

1127 que, a partir da relação $\sum_{i=1}^{n-1} H_i = n(H_n - 1)$ pode ser definido como

$$E\{\mathcal{F}^{\text{line}}\} = \frac{2}{(n-1)} (H_n - 1),$$

1128 de forma que o denominador n da fração é cancelado. **C.Q.D.**

1129 **Corolário A.8** (Comportamento assintótico de $E\{\mathcal{F}^{\text{line}}\}$). Em uma topologia em linha,
 1130 a eficiência média possui um comportamento assintótico que pode ser dado por

$$E\{\mathcal{F}^{\text{line}}\} \simeq \frac{2 \log n}{n}, \quad (\text{A.18})$$

1131 onde $n = |N|$ representa a quantidade de nós do grafo. \square

1132 *Demonstração.* Se $\lim_{n \rightarrow \infty} E\{\mathcal{F}^{\text{line}}\}/f(n) = 1$, então $E\{\mathcal{F}^{\text{ring}}\} \simeq f(n)$. Para se verificar a
 1133 igualdade assintótica, considera-se $f(n) = (2 \log n)/n$. Portanto, dada a representação de
 1134 $E\{\mathcal{F}^{\text{line}}\}$ descrita pela Equação A.17, o limite pode ser definido como

$$\lim_{n \rightarrow \infty} \frac{n}{2 \log n} \frac{2}{(n-1)} (\log n + \gamma - 1),$$

1135 que pode ser representado pela soma dos limites

$$\lim_{n \rightarrow \infty} \frac{n}{\log n} \frac{\log n}{(n-1)} + \lim_{n \rightarrow \infty} \frac{n}{\log n} \frac{\gamma - 1}{(n-1)} = 1 + 0 = 1,$$

1136 onde o primeiro limite converge para 1 enquanto o segundo para 0. **C.Q.D.**

1137 A.2 Complexidade das métricas

métrica (nó)	definição	algoritmo	tempo	memória
grau	2.3	2.1	$\Theta(n \mathbb{E}\{\mathcal{G}^{\text{out}}\})$	$\Theta(n)$
distância	2.6	2.3	$\Theta(n \mathbb{E}\{\mathcal{G}^{\text{out}}\})$	$\Theta(n)$
agrupamento	2.12	2.8	$\Theta(\mathbb{E}^2\{\mathcal{G}^{\text{out}}\})$	$\Theta(\mathbb{E}\{\mathcal{G}^{\text{out}}\})$

(a)

métrica (rede)	definição	algoritmo	tempo	memória
distribuição dos graus	2.5	2.2	$\Theta(n \mathbb{E}\{\mathcal{G}^{\text{out}}\})$	$\Theta(n)$
distância média	2.7	2.5	$\Theta(n^2 \mathbb{E}\{\mathcal{G}^{\text{out}}\})$	$\Theta(n)$
diâmetro	2.8	2.6	$\Theta(n^2 \mathbb{E}\{\mathcal{G}^{\text{out}}\})$	$\Theta(n)$
eficiência média	2.10	2.7	$\Theta(n^2 \mathbb{E}\{\mathcal{G}^{\text{out}}\})$	$\Theta(n)$
agrupamento	2.13	2.9	$\Theta(n \mathbb{E}^2\{\mathcal{G}^{\text{out}}\})$	$\Theta(\mathbb{E}\{\mathcal{G}^{\text{out}}\})$

(b)

Tabela A.1: Desempenho das métricas para caracterização de nós (a) e de redes (b).

Referências Bibliográficas

- Albert, R. e A.L. Barabási (2002), ‘Statistical mechanics of complex networks’, *Reviews of Modern Physics* **74**(1), 47–97.
(Citado nas páginas 25, 27 e 29)
- Anton, Howard e Robert C. Busby (2003), *Contemporary Linear Algebra*, John Wiley & Sons.
- Barabási, A.L. (2005), ‘Taming Complexity’, *Nature Physics* **1**(2), 68–70.
- Barabási, A.L. e R. Albert (1999), ‘Emergence of Scaling in Random Networks’, *Science* **286**(5439), 509–512.
(Citado na página 28)
- Barrat, A. e M. Weigt (2000), ‘On the properties of small-world networks’, *The European Physical Journal B* **13**, 547–560.
(Citado na página 27)
- Battista, Giuseppe Di, Peter Eades, Roberto Tamassia e Ioannis G. Tollis (1998), *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall.
- Bellman, Richard (1958), ‘On a routing problem’, *Quarterly of Applied Mathematics* **16**, 87–90.
(Citado na página 3)
- Bollobás, Béla (1981), ‘Degree sequences of random graphs’, *Discrete Mathematics* **33**(1), 1–19.
(Citado na página 25)
- Bollobás, Béla (2001), *Random Graphs*, 2ª edição, Cambridge University Press.
(Citado na página 25)
- Bollobás, Béla e Oliver Riordan (2004), ‘The diameter of a scale-free random graph’, *Combinatorica* **24**(1), 5–34.
(Citado na página 30)
- Brassard, G. e P. Bratley (1996), *Fundamentals of Algorithmics*, Prentice Hall.
- Chandy, K. Mani e Leslie Lamport (1985), ‘Distributed snapshots: determining global states of distributed systems’, *ACM Transactions on Computer Systems (TOCS)* **3**(1), 63–75.

- 1168 Chen, Chi-Tsong (2009), *Linear System Theory and Design*, 3ª edição, Oxford University
1169 Press.
- 1170 Chung, Fan e Linyuan Lu (2001), ‘The diameter of sparse random graphs’, *Advances in*
1171 *Applied Mathematics* **26**, 257–279.
1172 (Citado na página 25)
- 1173 Clauset, Aaron, Cosma Rohilla Shalizi e M.E.J. Newman (2009), ‘Power-Law Distributions
1174 in Empirical Data’, *SIAM Review* **51**(4), 661–703.
1175 (Citado na página 30)
- 1176 Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest e Clifford Stein (2009),
1177 *Introduction to Algorithms*, 3ª edição, The MIT Press.
1178 (Citado na página 6)
- 1179 Costa, L.F., F.A. Rodrigues, G. Travieso e P.R.V. Boas (2007), ‘Characterization of com-
1180 plex networks: A survey of measurements’, *Advances in Physics* **56**(1), 167–242.
1181 (Citado nas páginas 4 e 7)
- 1182 Dijkstra, Edsger Wybe (1959), ‘A note on two problems in connexion with graphs’, *Nu-*
1183 *merische Mathematik* **1**(1), 269–271.
1184 (Citado na página 3)
- 1185 Edmonds, Jack (1965), ‘Paths, trees and flowers’, *Canadian Journal of Mathematics*
1186 **17**, 449–467.
- 1187 Eggemann, N. e S. D. Noble (2011), ‘The clustering coefficient of a scale-free random
1188 graph’, *Discrete Applied Mathematics* **159**(10), 953–965.
1189 (Citado na página 30)
- 1190 Erdős, P. e A. Rényi (1959), ‘On Random Graphs I’, *Publ. Math. Debrecen* **6**, 290–297.
1191 (Citado na página 24)
- 1192 Ford, Jr., Lestor R. e D. R. Fulkerson (1962), *Flows in Networks*, Princeton University
1193 Press.
1194 (Citado na página 3)
- 1195 Goodman, J. David (2011), ‘China: Censorship Suit Filed Against Web Company’, *The*
1196 *New York Times* **maio**(20), A8.
- 1197 Hartmeier, Daniel (2002), Design and performance of the OpenBSD stateful packet filter
1198 (pf), em ‘Proceedings of the FREENIX Track: USENIX Annual Technical Conference’,
1199 pp. 171–180.
- 1200 Haykin, Simon (2008), *Neural Networks and Learning Machines*, 3ª edição, Prentice Hall.
- 1201 Hopcroft, John E. e Richard M. Karp (1973), ‘An $n^{5/2}$ algorithm for maximum matchings
1202 in bipartite graphs’, *SIAM Journal on Computing* **2**(4), 225–231.
- 1203 Kalman, Rudolph Emil (1960a), ‘A new approach to linear filtering and prediction pro-
1204 blems’, *Transactions of the ASME–Journal of Basic Engineering* **82**(Series D), 35–45.

- 1205 Kalman, Rudolph Emil (1960*b*), On the general theory of control systems, *em* ‘Proceedings
1206 of First IFAC Congress Automatic Control’, pp. 481–493.
- 1207 Kalman, Rudolph Emil (1963), ‘Mathematical Description of Linear Dynamical Systems’,
1208 *SIAM Journal on Control and Optimization* **1**(2), 152–192.
- 1209 Kohonen, Teuvo (2000), *Self-Organizing Maps*, 3ª edição, Springer.
- 1210 Kotelnikov, Vladimir A. (1933), On the transmission capacity of the ‘ether’ and of cables
1211 in electrical communications, *em* ‘Proceedings of the first All-Union Conference on the
1212 technological reconstruction of the communications sector and the development of low-
1213 current engineering’, Moscow, Russian.
- 1214 Kruskal, Joseph B. (1956), ‘On the shortest spanning subtree of a graph and the traveling
1215 salesman problem’, *Proceedings of the American Mathematical Society* **7**(1), 48–50.
- 1216 Latora, Vito e Massimo Marchiori (2001), ‘Efficient Behavior of Small-World Networks’,
1217 *Physical Review Letters* **87**(19), 198701.
1218 (Citado na página [15](#))
- 1219 Luczak, Tomasz (1998), ‘Random trees and random graphs’, *Random Structures Algo-*
1220 *rithms* **13**(3–4), 485–500.
1221 (Citado na página [25](#))
- 1222 Markov, Andrey (1913), ‘An Example of Statistical Investigation of the Text Eugene
1223 Onegin Concerning the Connection of Samples in Chains’. Lecture at the Physical-
1224 mathematical Faculty, Royal Academy of Sciences, St. Petersburg, 23 January 1913.
- 1225 Micali, Silvio e Vijay V. Vazirani (1980), An $O(\sqrt{|V|}|E|)$ algorithm for finding maximum
1226 matching in general graphs, *em* ‘Proceedings of 21st Annual Symposium on Foundations
1227 of Computer Science’, pp. 17–27.
- 1228 Milgram, Stanley (1967), ‘The small world problem’, *Psychology Today* **2**, 60–67.
1229 (Citado na página [26](#))
- 1230 Newman, M.E.J. (2003), ‘The Structure and Function of Complex Networks’, *SIAM Re-*
1231 *view* **45**, 167–256.
1232 (Citado nas páginas [7](#), [10](#) e [27](#))
- 1233 Nyquist, Harry Theodor (1928), ‘Certain topics in telegraph transmission theory’, *Trans.*
1234 *American Institute of Electrical Engineers* **47**(2), 617–644.
- 1235 Papoulis, Athanasios e S. Unnikrishna Pillai (2002), *Probability, Random Variables and*
1236 *Stochastic Process*, 4ª edição, McGraw-Hill.
1237 (Citado nas páginas [7](#) e [25](#))
- 1238 Shannon, Claude Elwood (1949), ‘Communication in the presence of noise’, *Proc. Institute*
1239 *of Radio Engineers* **37**(1), 10–21.
- 1240 Ware, Colin (2004), *Information Visualization: Perception for Design*, 2ª edição, Morgan
1241 Kaufmann.

- 1242 Watts, D.J. e S.H. Strogatz (1998), ‘Collective dynamics of ‘small-world’ networks’, *Nature*
1243 **393**(6684), 440–442.
1244 (Citado na página [26](#))
- 1245 Whittaker, Edmund Taylor (1915), ‘On the functions which are represented by the expan-
1246 sions of the interpolation theory’, *Proc. Royal Soc. Edinburgh* **35**(A), 481–493.
- 1247 Willinger, Walter, David Alderson e John C. Doyle (2009), ‘Mathematics and the Internet:
1248 A Source of Enormous Confusion and Great Potential’, *Notices of the AMS* **56**(5), 586–
1249 599.

Índice Remissivo

Símbolos

1251	$\Omega(\cdot)$	v
1252	$\Theta(\cdot)$	v
1253	\approx	v
1254	$\delta(t), \delta_{ij}$	iv
1255	\equiv	vi
1256	$E\{\mathcal{X}\}$	vi
1257	H_n	iv
1258	$O(\cdot)$	v
1259	$P(\mathcal{X}_\zeta)$	vi
1260	$P(\mathcal{X}_\zeta p)$	vi
1261	\propto	vi
1262	$\rho(\cdot)$	vii
1263	\simeq	vi
1264	\square	iii
1265	\triangle	vi
1266	C.Q.D.	iii

A

1268	agrupamento	
1269	grafo aleatório	25
1270	grafo de mundo pequeno	27
1271	grafo livre de escala	30
1272	algoritmo	
1273	agrupamento()	17
1274	agrupamento-medio()	18
1275	diametro()	14
1276	distancia()	11
1277	distancia-media()	13
1278	distribuicao-grau()	9
1279	eficiencia-media()	16
1280	grafo-aleatorio()	24
1281	grafo-livre-de-escala()	28
1282	grafo-mundo-pequeno()	26
1283	graus()	8
1284	menor-caminho()	12
1285	average shortest path	10

B

1287	BFS	10
1288	bi-implicação	vii

C

1290	cardinalidade	vii
1291	clustering coefficient	17
1292	coeficiente binomial	iv
1293	complexidade	
1294	agrupamento()	17
1295	agrupamento-medio()	18
1296	diametro()	15
1297	distancia()	11
1298	distancia-media()	14
1299	distribuicao-graus()	9
1300	eficiencia-media()	16
1301	grafo-aleatorio()	24
1302	grafo-livre-de-escala()	28
1303	grafo-mundo-pequeno()	26
1304	graus()	8
1305	menor-caminho()	12
1306	métricas	41
1307	conjunção	vii
1308	constante de Euler-Mascheroni	iv

D

1310	degree distribution	7
1311	delta de Dirac	iv
1312	delta de Kronecker	iv
1313	derivada	iii
1314	desvio padrão	vi
1315	diâmetro	
1316	grafo aleatório	25
1317	disjunção	vii
1318	distância média	
1319	grafo aleatório	25
1320	grafo de mundo pequeno	27
1321	grafo livre de escala	30

1323	F	1369	custo em memória 6
1324	fatorial iii	1370	livre de escala 28
1325	FIFO 11		
1326	G	1371	M
1327	grafo	1372	matriz iii
1328	aleatório 24	1373	de adjacência 4
1329	caracterização 7	1374	posto da vii
1330	definição 4	1375	transposta vii
1331	densidade 5	1376	média vi
1332	direcionado com pesos 4	1377	métrica
1333	livre de escala .. <i>veja</i> livre de escala	1378	agrupamento 16, 17
1334	métricas 7	1379	<i>average shortest path</i> 10
1335	mundo pequeno 26	1380	<i>clustering coefficient</i> 17
1336	robustez 7	1381	<i>degree distribution</i> 7
1337	topologia em anel 20	1382	densidade 5
1338	topologia em estrela 20	1383	diâmetro 14
1339	topologia em linha 20	1384	distância média 10
1340	grafo aleatório 24	1385	distribuição dos graus 7
1341	agrupamento 25	1386	eficiência 15
1342	diâmetro 25	1387	métricas 7
1343	distância média 25	1388	N
1344	grau esperado 25	1389	negação vii
1345	grafo de mundo pequeno 26	1390	P
1346	agrupamento 27	1391	política
1347	distância média 27	1392	de acesso 11
1348	grau esperado 27	1393	posto vii
1349	grafo livre de escala	1394	<i>preferencial attachment</i> 28
1350	agrupamento 30	1395	probabilidade vi
1351	distância média 30	1396	condicional vi
1352	grau esperado 29	1397	produto cartesiano vii
1353	grau 5	1398	Q
1354	CDF 10	1399	quantificador
1355	de entrada 5	1400	existencial v
1356	de saída 5	1401	universal v
1357	distribuição dos 8	1402	R
1358	PDF 8	1403	<i>random graphs</i> 24
1359	valor esperado 9	1404	recorrência <i>veja</i> recursividade
1360	grau esperado	1405	recursividade <i>veja</i> recorrência
1361	grafo aleatório 25	1406	rede livre de escala .. <i>veja</i> livre de escala
1362	grafo de mundo pequeno 27	1407	S
1363	grafo livre de escala 29	1408	<i>scale free</i> 28
1364	I	1409	série harmônica iv, 15, 37, 40
1365	igualdades v	1410	sistemas complexos 19
1366	implicação vii	1411	<i>small world</i> 26
1367	L		
1368	lista de adjacência 6		

1412	T	1430	definição 33
1413	topologia	1431	diâmetro 22
1414	em anel 20	1432	distância média 22, 36
1415	em estrela 20	1433	distribuição dos graus 21
1416	em linha 20	1434	eficiência média 23, 39
1417	topologia em anel	1435	topologias determinísticas 20
1418	definição 33	1436	V
1419	diâmetro 22	1437	valor absoluto vii
1420	distância média 22, 34	1438	valor esperado vi
1421	distribuição dos graus 21	1439	agrupamento 18
1422	eficiência média 23, 37	1440	distância 13
1423	topologia em estrela	1441	eficiência 15
1424	definição 33	1442	grau 9
1425	diâmetro 22	1443	variável aleatória iii
1426	distância média 22, 35	1444	realização vi
1427	distribuição dos graus 21	1445	vetor iii
1428	eficiência média 23, 39	1446	vizinhança 16
1429	topologia em linha		