# Redes Mundo Pequeno

Um dos tópicos mais populares na teoria de redes aleatórias é o estudo de quão conectadas elas são. Seis graus de separação é a frase comumente usada para descrever a natureza interconectada dos relacionamentos humanos: vários estudos, feitos de formas pouco controladas, mostraram que quaisquer duas pessoas escolhidas aleatoriamente no mundo podem ser conectadas, uma à outra, por uma pequena cadeia de pessoas (que normalmente contém cerca de seis pessoas), sendo que cada uma destas conhece bem a próxima pessoa da cadeia. Se representarmos as pessoas como nós e os conhecidos como vizinhos, reduzimos o problema ao estudo da rede de relacionamentos. Muitos problemas interessantes surgem do estudo de propriedades de redes geradas aleatoriamente. Uma rede é uma coleção de nós e arestas, com cada aresta conectada a dois nós, mas com cada nó potencialmente conectado a qualquer número de arestas (Fig. 1.5). Uma rede aleatória é construída probabilisticamente de acordo com algumas regras definidas; estudo de redes aleatórias geralmente é feito estudando todo o conjunto de redes, cada uma ponderada pela probabilidade de ter sido construída. Assim, esses problemas naturalmente se enquadram no amplo escopo da mecânica estatística.

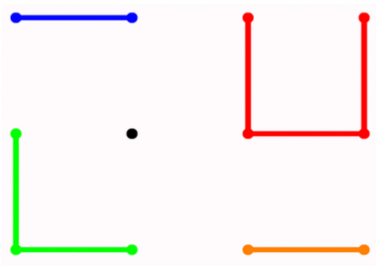

**Fig. 1.5 Network.** A network is a collection of nodes (circles) and edges (lines between the circles).

Neste exercício vamos gerar algumas redes aleatórias e calcular a distribuição de distâncias entre pares de pontos. Estudaremos redes de mundo pequeno (*small world networks*) [2,3], um modelo teórico que sugere como um pequeno número de atalhos (amizades internacionais e interculturais incomuns, por exemplo) pode encurtar drasticamente os comprimentos típicos da cadeia.

**Construindo uma rede de mundo pequeno.** Os $N$ nós em uma rede de mundo pequeno estão dispostos em torno de um círculo. Existem dois tipos de arestas. Cada nó tem $Z$ arestas curtas conectando-o aos seus vizinhos mais próximos ao redor do círculo (até uma distância $Z/2$). Além disso, existem $p\,N\,Z/2$ atalhos adicionados à rede, que conectam nós aleatoriamente (veja a Fig. 1.6). (Esta é uma versão mais tratável [2] do modelo original [3], onde uma fração p das $NZ/2$ arestas foram religadas de forma aleatória. Note que da forma proposta aqui $p$ pode assumir valores maiores que 1 sem, com isso, conectar todos os nós, já que há uma probabilidade de adicionarmos aleatoriamente um atalho entre um par de nós que já estava conectado ou até mesmo entre um nó e ele mesmo.)

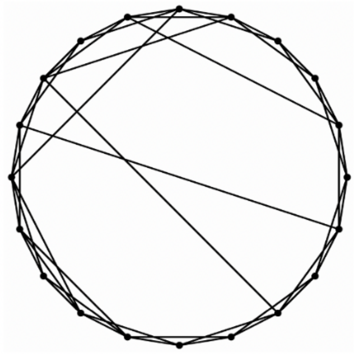

**Fig. 1.6 Small world network** *with* $L = 20$, $Z = 4$, *and* $p = 0.2$.[11]

(a) Para facilitar nossa análise utilizaremos a biblioteca python NetworkX (https://networkx.org/documentation/stable/tutorial.html). Utilizando esta biblioteca, podemos definir uma rede (ou equivalentemente um grafo), utilizando o comando G = nx.Graph() (veja mais detalhes no link acima). Para representar uma rede precisamos, no mínimo, os seguintes elementos:
(1) G.add_node(node), que adiciona um novo nó ao sistema;
(2) G.add_edge (node1, node2), que adiciona uma nova aresta ao sistema;
(3) list(G.nodes) que retorna uma lista de nós existentes; e
(4) list(G.neighbors(node)), que retorna os vizinhos de um nó existente.

Escreva uma rotina para construir uma rede de mundo pequeno, que (dados $N$, $Z$ e $p$) adiciona os nós e as arestas curtas e, em seguida, adiciona aleatoriamente os atalhos. Use o a biblioteca NetworkX para desenhar este grafo de mundo pequeno e verifique se você implementou as condições de contorno periódicas corretamente (cada nó $i$ deve ser conectado aos nós $(i - Z/2) mod\ N$,...,$(i + Z/2\ ) mod\ N$). Para plotar vocês podem usar o comando nx.draw_circular(G).

**Medindo as distâncias mínimas entre nós.** A propriedade mais estudada dos grafos de mundo pequeno é a distribuição dos caminhos mais curtos entre os nós. Sem as arestas longas, o caminho mais curto entre $i$ e $j$ será dado saltando em passos de comprimento $Z/2$ ao longo do menor dos dois arcos ao redor do círculo; não haverá caminhos de comprimento maior que N/Z (meio caminho ao redor do círculo), e a distribuição $\rho(d)$ de comprimentos de caminho $d$ será constante para $0 < d < N/Z$. Quando adicionamos atalhos, esperamos que a distribuição seja deslocada para caminhos mais curtos.

(b) Escreva as três funções a seguir para encontrar e analisar a distribuição do comprimento do caminho.

(1) FindPathLengthsFromNode(graph, node), que retorna para cada node2 no grafo a distância mais curta de node a node2. Um algoritmo eficiente é uma busca em largura do grafo, trabalhando para fora do nó em camadas (shells).

Haverá um currentShell de nós cuja distância será definida como $d$ a menos que eles já tenham sido visitados, e um nextShell que será considerado após o término do atual (olhando de lado antes de avançar, largura primeiro), como segue.

– Inicialize $d = 0$, a distância do nó a si mesmo para zero e currentShell = [node].
– Enquanto houver nós no novo currentShell:
  - $*$ iniciar um novo nextShell vazio;
  - $*$ para cada vizinho de cada nó no shell atual, se a distância até o vizinho não tiver sido definida, adicione o nó ao nextShell e defina a distância como $d + 1$;
  - $*$ adicione um a $d$ e defina o shell atual para nextShell.

– Retorne as distâncias.

Isso varrerá para fora do nó, medindo a distância mais curta para todos os outros nós da rede. (Dica: verifique seu código com uma rede com N pequeno e p pequeno, comparando alguns caminhos com cálculos manuais da imagem do gráfico gerada como na parte (a).) Quem já domina o algoritmo de busca em largura pode usar a implementação disponível no pacote networkx.

(2) FindAllPathLengths(graph), que gera uma lista de todos os comprimentos (um por par de nós no gráfico) usando repetidamente FindPathLengthsFromNode. Verifique sua função testando se o histograma de comprimentos de caminho em $p = 0$ é constante para $0 < d < N/Z$, conforme esperado. Gere gráficos para $N = 1.000$ e $Z = 2$ para $p = 0{,}02$ e $p = 0{,}2$; exiba os gráficos circulares e trace o histograma de comprimentos de caminho. Amplie o histograma; o quanto ele muda com p? Que valor de p você precisaria para obter “seis graus de separação”?

(3) FindAveragePathLength(graph), que calcula a média $\langle d \rangle$ sobre todos os pares de nós. Calcule $d$ para $Z = 2$, $N = 100$ e $p = 0{,}1$ algumas vezes; sua resposta deve ser em torno de $d = 10$. Observe que existem flutuações estatísticas substanciais no valor de amostra para amostra. Aproximadamente quantas arestas longas existem neste sistema? Você esperaria flutuações nas distâncias?

(c) Plote o comprimento médio do caminho entre os nós $d(p)$ dividido por $d(p = 0)$ para $Z = 2$, $N = 50$, em função de $p$ em um gráfico semi-log com valores entre $p = 0{,}001$ e $p = 1000$. (Dica : Sua curva deve ser semelhante à de Watts e Strogatz [3, Fig. 2], com os valores de $p$ deslocados por um fator de 100; veja a discussão do limite do contínuo abaixo. Para visualizar melhor, gere valores de $p$ igualmente espaçados entre $10^{-3}$ e $10^{3}$.) Por que o gráfico é fixado em um para $p$ pequeno?

**Grande $N$ e a emergência de um limite contínuo.** Podemos entender o deslocamento em $p$ da parte (c) como um limite contínuo do problema. No limite onde o número de nós, $N$, torna-se grande e o número de atalhos $pNZ/2$ permanece fixo, este problema de rede tem um bom limite onde a distância é medida em radianos $\Delta\theta$ ao redor do círculo. Dividir $d$ por $d(p = 0) \approx N/(2Z)$ essencialmente faz isso, pois $\Delta\theta = \pi Z d/N$.

(d) Crie e exiba um gráfico circular de sua geometria da parte (c) ($Z = 2, N = 50$) em $p = 0{,}1$; crie e exiba gráficos circulares da geometria de Watts e Strogatz ($Z = 10, N = 1.000$) em $p = 0{,}1$ e $p = 0{,}001$ (veja o comando nx.watts_strogatz(N,k,p) no NetworkX). Qual dos sistemas deles parece estatisticamente mais semelhante ao seu? Plote o comprimento médio do caminho redimensionado, $\pi Z d/N$, versus o número total de atalhos, $pNZ/2$, para um intervalo de $0{,}001 < p < 1000$, para $N = 100$ e $200$ e para $Z = 2$ e $4$ .
Neste limite, o comprimento médio de ligação $\langle \Delta\theta \rangle$ deve ser uma função apenas de $M$. Como Watts e Strogatz [3] funcionaram em um valor de $ZN$ um fator de 100 maior que o nosso, nossos valores de $p$ são um fator de $100$ maior para obter o mesmo valor de $M = pNZ/2$. Newman e Watts [2] derivam esse limite contínuo com uma análise de grupo de renormalização.

(e) Redes reais. Procure por uma rede real e encontre a distância média e o histograma das distâncias entre os nós. (Veja, por exemplo, https://networkrepository.com/network-data.php)

## Referências


[1] J. Sethna, Statistical Mechanics: Entropy, order parameters and complexity, Oxford University Press (2020). Disponível em: https://www.lassp.cornell.edu/sethna/StatMech//

[2] Newman, M. E. J. (2000). Models of the small world. Journal of Statistical Physics, 101, 819–41.

[3] Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. Nature, 393, 440–42.

[4] Albert-László Barabási, Network Science, Cambridge University Press (2016). Disponível em http://networksciencebook.com/