

Visitando Redes Complexas

Victor Souza Lima¹, Wladimir Araújo Tavares¹

¹ Universidade Federal do Ceará (UFC) – Campus Quixadá
Caixa Postal 6001 – 63.900-000 – Cedro – Quixadá – CE – Brazil

victorsouza@alu.ufc.br, wladimirufc@gmail.com

Abstract. *We establish relationships everywhere we go and thus form networks all the time. This article introduces the subject of complex networks, giving its formal definition and several examples from everyday life. The main properties of complex networks and some applications that use complex networks are analyzed, being, therefore, a material indicated for researchers beginning in the theme.*

Resumo. *Nós estabelecemos relações em todo lugar que vamos e, assim, formamos redes o tempo todo. Este artigo introduz o assunto de redes complexas, dando sua definição formal e vários exemplos do cotidiano. São analisadas as principais propriedades estruturais de redes complexas e algumas aplicações que utilizam redes complexas, sendo portanto, um material indicado para pesquisadores iniciantes no tema.*

1. Introdução

Atualmente vivemos em um mundo muito conectado, onde estabelecemos várias relações no nosso cotidiano, sejam elas fisicamente ou virtualmente. Essas relações acontecem no trabalho, na faculdade, com amigos, entre outras formas. Podemos representar essas e outras conexões do mundo real com redes complexas, que corresponde a um grafo, onde os vértices - também chamados de nós - são objetos e as arestas indicam relações entre os objetos, logo, dizemos que os vértices que possuem uma aresta entre si são relacionados e, sendo assim, é possível codificar e analisar como as coisas se conectam dentro de uma determinada rede [Metz *et al.* 2007].

É possível codificar vários tipos de relações diferentes com um mesmo conjunto de indivíduos, gerando assim redes distintas. Por exemplo, na Figura 1, temos duas redes diferentes, com o mesmo conjunto de indivíduos, que é o conjunto de pessoas que trabalham em uma determinada empresa. Na rede 1a, a amizade foi o critério estabelecido para as relações, logo, um vértice se conecta a outro caso aquelas pessoas sejam amigas. Já na rede 1b, o critério para se estabelecer uma relação foi trabalhar ou já ter trabalhado no mesmo setor da outra pessoa. Veja que uma rede está relacionada à outra, de certa forma, pois o fato de duas pessoas trabalharem em um mesmo setor da empresa pode contribuir ou não para elas serem amigas.

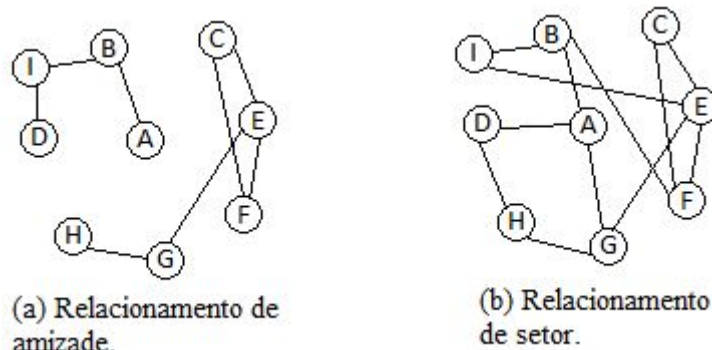


Figura 1. Um mesmo conjunto de pessoas dando origem a duas redes diferentes com dois tipos distintos de relacionamentos.

O exemplo da Figura 1 é de um grafo não direcionado - que representa relacionamentos simétricos - onde as arestas não possuem direção, sendo assim, se o vértice A está relacionado com o vértice B, o vértice B também está relacionado com o vértice A. Temos também grafos que são direcionados - que representam relacionamentos assimétricos - onde as arestas possuem direções. Em um grafo direcionado (dígrafo), cada aresta tem um sentido (direção) que conecta um vértice origem à um vértice destino. Um exemplo de relacionamento assimétrico é a troca de emails eletrônicos entre pessoas.

É importante destacar que pode-se representar muitas outras coisas com redes complexas além de relacionamentos entre pessoas, como o conteúdo das páginas *WEB* - *World Wide Web*, redes neurais, cadeia alimentar e a estrutura física da internet, onde os computadores são os vértices das redes e os meios de transmissão são as arestas, por exemplo. A maioria dessas redes são gigantescas, contendo milhares de vértices e arestas e apresentando propriedades específicas, por isso o nome de redes complexas.

O artigo será organizado da seguinte maneira. Na Seção 2, apresentamos as principais propriedades das redes complexas, entre elas o grau dos vértices e o caminho e distância entre vértices. Na Seção 3, apresentamos os principais tipos de redes complexas. Na Seção 4, apresentamos algumas aplicações que usam redes complexas.

2. Principais propriedades de Redes Complexas

Existem inúmeras propriedades em redes complexas. Essas propriedades mostram características (estrutura) do grafo que podem ou não serem importantes no estudo sobre como a rede opera (funcionalidade), então, cada propriedade vai ser observada de acordo com o objetivo da análise. O mais importante sobre as propriedades não é apenas obter o seu valor e sim saber interpretar o que aquele valor significa para a rede.

Serão abordadas a seguir apenas algumas das propriedades de uma rede complexa, tendo assim, muitas outras propriedades importantes a serem consideradas para o estudo de uma rede, como reciprocidade, assortatividade, centralidade etc.

Uma rede é definida por $R = (V, E)$, onde V é o conjunto de todos os vértices da rede e E é o conjunto de todas as arestas da rede, logo, cada elemento do conjunto E

corresponde a dois vértices, indicando que há uma aresta entre eles. O número de objetos (vértices) é dado por n e o número de relacionamentos (arestas) é dado por m . Usando o número de vértices e arestas de uma rede, podemos definir a densidade da rede, ρ , que representa a fração de arestas que a rede possui [Figueiredo 2011].

2.1. Grau

O **grau** é a quantidade de arestas que um determinado vértice possui. Pode ser visto também como a quantidade de relações que um objeto da rede tem. Na figura 1a, o grau do vértice I é 2 (dois), já que ele estabelece relação com dois outros vértices da rede.

Em um grafo direcionado, em que as arestas possuem vértice de origem e vértice de destino, o grau é dividido em grau de saída e grau de entrada. O grau de saída de um vértice se refere à quantidade de arestas que esse mesmo vértice é a origem (arestas que saem dele), já o grau de entrada de um vértice corresponde à quantidade de arestas que esse mesmo vértice é o destino (arestas que chegam nele). O valor de grau de saída e de grau de entrada de um grafo são iguais.

É importante notar que a soma do grau de todos os vértices é igual ao dobro de número de arestas da rede [West, 1996]. Podemos ainda obter o grau médio de uma rede, que é dada pela média aritmética do grau de todos os vértices [Figueiredo 2011].

2.2. Caminho e Distância

Caminho e distância são duas propriedades muito importantes no estudo de redes complexas. Podemos definir como **um caminho entre dois vértices** A e B , uma sequência de vértices que pode-se passar para sair de A e chegar a B , sem repetição. Por exemplo, na figura 1b as sequências CFEG e CEIBAG são dois caminhos de C para G . O **tamanho do caminho** é dado pela quantidade de arestas que o mesmo possui ou a quantidade de vértices menos 1 (um), logo, os tamanhos dos caminhos anteriormente citados são 3 (três) e 5 (cinco), respectivamente.

Observe que podem existir muitos caminhos entre dois vértices. Podemos definir a **distância entre dois vértices** A e B , como o tamanho de um menor caminho entre eles. Voltando ao exemplo da figura 1b, a distância dos vértices C e G é 2 (dois), pois o menor caminho entre eles é dado pela sequência CEG.

Sabendo disso, podemos definir a distância média de uma rede, que é a média aritmética das distâncias entre todos os pares de vértices dela e o diâmetro, que é dado pela maior distância presente nela. Entretanto, muitas redes não são conexas, ou seja, elas não possuem caminhos entre todos os pares de vértices. Logo, a distância entre estes pares de vértices não está definida. Neste caso, o cálculo da distância média e do diâmetro devem desconsiderar estes pares ou considerar apenas os pares que pertencem a maior componente conexa, ou seja, ao maior pedaço conectado [Figueiredo 2011]. Por exemplo, na figura 1a a distância média da rede é 1.7, um valor relativamente baixo, até pelo tamanho da rede ser pequena. Já o diâmetro tem valor igual a 3 (três). Veja que foi considerada nos cálculos apenas a maior componente conexa presente na rede, que é formada pelos vértices C, E, F, G e H .

2.3. Coeficiente de aglomeração

Antes de definirmos o coeficiente de aglomeração, precisamos saber o que são os vizinhos de um vértice. Os **vizinhos de um vértice** é o conjunto de vértice que faz relação com ele, ou seja, um vértice é vizinho de outro caso tenha uma aresta entre eles.

O **coeficiente de aglomeração** se refere à quantidade de relações triangulares existentes na rede. Essas relações triangulares - também chamadas de transitividade - dizem respeito ao conjunto de 3 (três) vértices conectados uns aos outros. Essa propriedade também pode ser vista como a porção de arestas que os vizinhos de um vértice possuem entre si.

Para vértices que possuem grau 0 (zero) ou 1 (um), o coeficiente de aglomeração não está definido. O valor do coeficiente de aglomeração da rede é a média aritmética de todos os coeficientes de aglomeração que estão definidos.

3. Tipos de redes complexas

Foram estudados vários tipos de redes ao longo dos anos, mas vamos abordar brevemente apenas 3 (três) neste artigo, que são as redes aleatórias, as redes de pequeno mundo e as redes livres de escala.

A **rede aleatória** - também chamada de modelo $G(n,p)$ - é uma das mais antigas e foi proposta por Paul Erdős e Alfréd Rényi. É considerada uma rede bem simples, pois contém apenas dois parâmetros. O parâmetro n define quantos vértices a rede terá e o parâmetro p determina a probabilidade de uma certa aresta pertencer à rede. Sendo assim, duas redes aleatórias com os mesmos parâmetros podem - e muito provavelmente serão - ser geradas com estruturas e quantidade de arestas diferentes, pois a probabilidade de uma determinada aresta pertencer à rede é independente da probabilidade de outras arestas da rede também pertencerem, dando jus ao seu nome de rede aleatória. Por exemplo, ao gerar-se uma rede $G(4, 0.5)$ ela poderá ter qualquer estrutura de um rede com 4 (quatro) vértices, sendo que cada possível aresta tem 50% de chance de pertencer à rede.

A **rede de pequeno mundo** foi estudada e intitulada por Watts e Strogatz no fim da década de 90, que observaram um padrão muito conectado em algumas redes, fazendo com que vértices próximos sejam muito conectados. Nesse modelo, a distância média entre quaisquer dois vértices de uma rede muito grande não ultrapassa um número pequeno de vértices [Metz *et al.* 2007]. Um experimento feito em 1960 por Stanley Milgram nos ajuda a perceber o efeito desse tipo de rede, onde foram entregues cartas a várias pessoas e a carta tinha que chegar até o destinatário, mas a pessoa que estivesse com a carta só poderia entregar para quem ela conhecesse pessoalmente. Percebeu-se que rapidamente as cartas chegavam ao seu destino, em média entre 5 (cinco) e 6 (seis) passagens. As redes sociais são exemplos de redes de pequeno mundo.

A **rede livre de escala** foi proposta por Albert-László Barabási e Réka Albert também no fim da década de 90 e possui uma característica bem específica. Nesse tipo de rede, é comum ter poucos vértices com muitas conexões e muitos vértices com poucas conexões. Isso se dá pelo motivo de que quando um novo vértice na rede é criado, ele tem tendência a se conectar com vértices que já possuem um grande número de conexões. Um exemplo desse tipo de rede é a estrutura de páginas da *Web*, onde

novas páginas quando criadas tendem a ter hiperlinks (sendo essa a forma de conexão) para páginas “famosas”, que já possuem muitas conexões.

4. Aplicações com redes complexas

Como dito na seção anterior, é interessante modelar muitas redes reais como redes complexas, a fim de se obter algumas informações estruturais de como aquela rede se comporta, de acordo com as características encontradas nela. Nesse cenário, podemos citar algumas aplicações que utilizam redes complexas, como: detecção de comunidades ([Félix *et al.* 2018] usa essa técnica para analisar quais países fazem mais transferências entre si no mercado do futebol e o motivo disso), avaliação de sumários (apresentado em [Metz *et al.* 2007]), congestionamento em redes ([Metz *et al.* 2007] verifica os principais motivos de uma rede congestionar e propõe alternativas para amenizar o congestionamento), avaliação da qualidade de textos (abordado e aprofundado em [Antiqueira *et al.* 2005]) etc. Bem, vamos abordar aqui brevemente esse último tema.

Antes de avaliar a qualidade dos textos, é preciso modelar a rede, que por sua vez, precisa de um pré-processamento e depois de uma construção da rede propriamente em si. Na fase de pré-processamento dos textos são retiradas as *stopwords* (palavras pouco significativas, como preposições e conjunções), ocorre posteriormente uma lematização das palavras (junção de significados da mesma forma canônica, mas com flexões distintas) e por último, os textos são etiquetados morfossintaticamente. Após a fase de pré-processamento, ocorre a construção da rede, que é um grafo direcionado, os vértices são as palavras que restaram e as arestas indicam quantas vezes as palavras aparecem uma depois da outra, naquele sentido. A estrutura pode ser representada por uma matriz de adjacências com pesos entre os vértices.

Foram computadas a média dos graus de entrada, dos graus de saída e do coeficiente de aglomeração para todos os nós, e também a média dos caminhos mínimos entre todos os pares de nós da rede (com exceção das auto-conexões) [Antiqueira *et al.* 2005].

Foram realizados alguns experimentos com juízes humanos sempre com textos de dois gêneros diferentes, com o objetivo deles avaliarem textos em relação a alguns critérios, como: complexidade, clareza, legibilidade ou qualidade dos textos. No experimento 1 (um), foram pegos textos do caderno Esporte e do caderno Dinheiro de um mesmo jornal. No experimento 2 (dois), foram avaliados 5 (cinco) textos considerados bons e 5 (cinco) textos considerados ruins, de um mesmo assunto. Nos dois primeiros experimentos, as medidas entre os dois grupos de textos não tiveram uma boa diferença e, assim, não foi possível tirar grandes conclusões.

Já no experimento 3 (três) - que direcionou a avaliação apenas no critério qualidade - foram avaliados 10 (dez) textos do gênero informativo (considerados bons) e 10 (dez) redações (consideradas ruins). O resultado da avaliação dos juízes confirmou que os textos informativos são melhores do que as redações. Dessa forma, é possível tirar algumas conclusões observando as propriedades das redes geradas. [Antiqueira *et al.* 2005] relata algumas conclusões, como: quando todos os textos são considerados, a qualidade tende a cair na medida em que os valores do grau de saída aumentam; a

relação entre o coeficiente de aglomeração e as notas revelou que a qualidade dos textos diminui com o coeficiente de aglomeração; a qualidade é praticamente uniforme para os textos bons, sendo que, em média, seus caminhos mínimos são maiores do que os dos textos ruins.

5. Conclusão e trabalhos futuros

Este trabalho introduziu o assunto redes complexas, dando sua definição formal, relacionando com o cotidiano, principais propriedades estruturais, diferentes tipos de redes complexas e algumas aplicações com redes complexas. Trabalhos futuros devem estudar outras propriedades aqui não expostas, aplicar redes complexas em uma das áreas aqui abordadas ou até em outras, podendo até continuar na avaliação de textos, mudando a categoria dos textos, por exemplo.

Referências

- METZ, Jean; CALVO, Rodrigo; SENO, Eloize Rossi Marques; ROMERO, Roseli Aparecida Francelin; LIANG, Zhao. Redes complexas: conceitos e aplicações. [S.l: s.n.], 2007.
- Daniel R. Figueiredo, Introdução a Redes Complexas. Jornada de Atualização em Informática (JAI), 2011.
- Félix, Lucas G. S.; BARBOSA, Carlos M.; CARVALHO, Iago A.; VIEIRA, Vinícius da F.; XAVIER, Carolina Ribeiro. Uma análise das seleções da copa utilizando uma rede de transferências de jogadores entre países. In: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING (BRASNAM), 7. , 2018, Natal. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2018 . ISSN 2595-6094. DOI: <https://doi.org/10.5753/brasnam.2018.3588>.
- Antiqueira, Lucas; NUNES, Maria das Graças Volpe; OLIVEIRA JUNIOR, Osvaldo Novais de; COSTA, Luciano da Fontoura. Modelando textos como redes complexas. Anais.. [S.l.]: [s.n.], 2005.
- West, Douglas Brent. *Introduction to graph theory*. Vol. 2. Upper Saddle River, NJ: Prentice hall, 1996.