Universidade de São Paulo
Instituto de Física de São Carlos

Guilherme Schmidt Domingues

Caminhos mínimos em redes complexas:
estrutura e otimização

São Carlos

2021

Guilherme Schmidt Domingues

Caminhos mínimos em redes complexas:

estrutura e otimização

Dissertação apresentada ao Programa de
Pós-Graduação em Física do Instituto de Física
de São Carlos da Universidade de São Paulo
para obtenção do título de Mestre em Ciências.
Área de Concentração: Física Aplicada
Opção: Física Computacional
Orientador: Prof. Dr. Luciano da Fontoura
Costa

Versão Corrigida

(versão original disponível na Unidade que aloja o Programa)

São Carlos

2021

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Domingues, Guilherme Schmidt

Caminhos mínimos em redes complexas: estrutura e otimização / Guilherme Schmidt Domingues; orientador Luciano da Fontoura Costa - versão corrigida -- São Carlos, 2021.

65 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física Aplicada Computacional) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2021.

1. Ciência de redes. 2. Topologia de redes. 3. Otimização. I. Costa, Luciano da Fontoura, orient. II. Título.

Agradecimentos

Ao Dr. Luciano da F. Costa, pela contribuição e orientação na realização deste trabalho.

Ao Dr. Cesar H. Comin, pela contribuição nas explicações teóricas e interpretações dos resultados deste trabalho.

Ao Instituto de Física de São Carlos, pela oportunidade de realização do curso de mestrado.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de mestrado (131909/2019-3) e pelo apoio financeiro para a realização desta pesquisa. Parte deste trabalho relaciona-se ao projeto temático FAPESP (2015/22308-2).

RESUMO

DOMINGUES, G. S. **Caminhos mínimos em redes complexas:** estrutura e otimização. 2021. 65p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Dentre as várias propriedades topológicas de redes complexas, o caminho mínimo representa uma característica particularmente importante devido ao seu potencial efeito em vários processos dinâmicos. Além disso, várias situações práticas, como o tráfego de veículos nas cidades, por exemplo, podem se beneficiar da redução dos respectivos caminhos mínimos nos sistemas relacionados. No presente trabalho, abordamos o problema da redução do mínimo caminho médio de várias redes complexas teóricas e uma do mundo real, adicionando um determinado número de arestas de acordo com diferentes estratégias e fazemos a comparação do desempenhos destas estratégias. Mais especificamente, consideramos: a adição de novas arestas entre vértices com grau, centralidade de intermediação, centralidade de proximidade e acessibilidade relativamente baixo/baixo, baixo/alto e alto/alto; melhorar a regularidade do grau da rede; e ligação preferencial de acordo com o grau. Também verificamos se a maleabilidade da rede pode ser usada como um meio de prever o potencial desta rede em ser otimizada. Vários resultados interessantes foram obtidos, incluindo a identificação de estratégias baseadas em conectar vértices com valores máximos e mínimos de uma medida como resultante na maior redução do comprimento do mínimo caminho médio em geral e estratégias baseadas em conectar vértices com valores máximos entre si como melhores no caso de redes modulares. Outra descoberta interessante foi que, para vários tipos de redes, os métodos baseados em graus tendem a fornecer melhorias comparáveis àquelas obtidas pelo uso de uma medida muito mais dispendiosa computacionalmente que é a centralidade de intermediação.

Palavras-chave: Ciência de redes. Topologia de redes. Otimização.

ABSTRACT

DOMINGUES, G. S. **Shortest paths in complex networks:** structure and optimization. 2021. 65p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Among the several topological properties of complex networks, the shortest path represents a particularly important characteristic due to its potential impact on several dynamical processes. In addition, several practical situations, such as transit in cities, for example, can benefit from reducing their respective shortest path in the related system. In the present work, we addressed the problem of trying to reduce the average shortest path of several theoretical and one real-world complex networks by adding a given number of links according to different strategies and we also compare the performance of these strategies. More specifically, we considered: placing new links between nodes with relatively low/low, low/high, and high/high degrees, betweenness centralities, closeness centralities and accessibilities; enhancing the degree regularity of the network; and preferential attachment according to the degree. We also checked whether the malleability of the network can be used as a means of predicting the potential of this network to be optimized. Several interesting results have been obtained, including the identification of the strategies based in connect nodes with higher and lower values of a measurement as those providing the largest reduction of the average shortest path length in general and strategies based in connect two nodes with higher values in the case of modular networks. Another interesting finding is that, for several types of networks, the degree-based methods tend to provide improvements comparable to those obtained by using the much more computationally expensive betweenness centrality measurement.

Keywords: Network science. Network topology. Optimization.

LISTA DE FIGURAS

Figura 1	Processo de construção do modelo Watts-Strogatz, onde começamos com (A) uma rede com topologia em anel e (B) arestas são removidas e reconectadas aleatoriamente em sucessivas iterações. . .	24
Figura 2	Nesta figura podemos ver um exemplo de um grafo (a) com um dos caminhos mínimos entre os vértices A e B destacado, com $d_{AB} = 5$. Com a adição de uma única aresta (b, em azul), o caminho mínimo é reduzido para $d_{AB} = 2$	28
Figura 3	Acessibilidade de dois vértices com vizinhança e probabilidades de transição diferentes. O vértice em vermelho (V) indica o vértice de referência, enquanto os vértices em verde e em amarelo formam, respectivamente, a primeira ($h = 1$) e a segunda ($h = 2$) vizinhanças de V . Os valores de p junto a cada vértice representam a probabilidade de transição de serem alcançados por um caminhante aleatório partindo de V . (a) Quando as probabilidades de transição são similares, a acessibilidade é alta, neste caso sendo igual a $A_2(V) = 8$ para a segunda vizinhança de V . (b) Quando as probabilidades de transição são heterogêneas, a acessibilidade de V é menor, sendo igual a $A_2(V) = 1,65$ neste exemplo.	30
Figura 4	Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Barabási-Albert.	34
Figura 5	Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Erdős-Rényi.	34
Figura 6	Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Watts-Strogatz.	35
Figura 7	Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Waxman.	36

Figura 8	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo BA. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.	42
Figura 9	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo ER. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.	42
Figura 10	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo WS. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.	43
Figura 11	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo Waxman. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.	43
Figura 12	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada uma das topologias de redes. Aqui, todas as redes são formadas por $N = 5000$ vértices e grau médio em torno de 6.	44
Figura 13	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada uma das topologias de redes. Aqui, todas as redes são formadas por $N = 8000$ vértices e grau médio em torno de 6.	44
Figura 14	Ordem dos melhores desempenhos das estratégias consideradas em função da taxa de adição de arestas N_a/N para a rede formada por $N = 3000$ vértices seguindo o modelo ER.	46

Figura 15	Exemplo de uma rede, com vértices coloridos de acordo com seus valores de acessibilidade, antes (a) e depois (b) da aplicação da estratégia A_{cp} . O caminho mínimo entre um mesmo par de vértices está indicado em vermelho nas duas redes.	48
Figura 16	Duas possíveis estratégias de se adicionar uma aresta em uma rede com topologia em estrela. (a) Rede original. (b) Um vértice periférico é conectado a um vértice central. (c) Dois vértices periféricos são conectados entre si.	49
Figura 17	Comprimento do mínimo caminho médio em função do número de braços k para as três formas de redes mostradas na Figura 16. . . .	51
Figura 18	Relação entre o ASPL inicial encontrado em cada um dos modelos considerados e a melhoria produzida pelas estratégias consideradas. Aqui, apenas uma estratégias representativa é mostrada.	53
Figura 19	Correlação de Pearson entre a maleabilidade das redes estudadas e os valores inicial e de melhoria encontrados para o ASPL destas redes, quando aplicadas as estratégias propostas. A cor vermelha mais intensa representa uma alta correlação positiva e a cor azul mais intensa representa uma alta correlação negativa. A primeira coluna foi calculada considerando o conjunto de todos os modelos estudados, enquanto a segunda coluna teve seus valores calculados apenas para os modelos BA e a terceira leva em consideração os models ER, WS ou Waxman, que apresentaram comportamentos semelhantes entre si.	54
Figura 20	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada topologia de redes modulares consideradas. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6. As barras de erro são mostradas como linhas verticais com 40% do desvio padrão dos dados e o eixo vertical é normalizado em relação ao valor inicial do ASPL.	55

Figura 21	Visualização típica de uma rede Waxman adaptada para criação de comunidades. Os vértices são coloridos de acordo com suas respectivas acessibilidades, com cores mais claras para acessibilidades menores, para evidenciar as regiões de centro e de periferia de cada comunidade.	55
Figura 22	Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para a topologia de Fortunato com parâmetro de mistura $\mu = 0,03$, $\mu = 0,05$ e $\mu = 0,07$. As barras de erro são mostradas como linhas verticais com 40% do desvio padrão dos dados e o eixo vertical é normalizado em relação ao valor inicial do ASPL.	56
Figura 23	Rede mundial de aeroportos com os vértices posicionados de acordo com a localização geográfica dos aeroportos.	57
Figura 24	Variação do comprimento do mínimo caminho médio, em termos da taxa de adição de arestas N_a/N , obtido pela aplicação das várias estratégias consideradas na rede de aeroportos. O eixo vertical está normalizado em relação ao valor inicial do ASPL.	58

LISTA DE TABELAS

- Figura 1 As três melhores estratégias listadas para cada modelo, com $N = 3000$, $N = 5000$ e $N = 8000$ vértices depois de 20 iterações de cada estratégia proposta. Os números entre parênteses indicam a porcentagem das vezes em que a estratégia resultou na melhor, segunda melhor e terceira melhor redução do ASPL, respectivamente. Resaltamos que este trabalho não apresenta as simulações feitas em redes do modelo BA com $N = 8000$, como explicado nas discussões. 45
- Figura 2 Comparação dos valores de variação percentual do ASPL em relação aos respectivos valores iniciais, depois de 200 iterações de cada estratégia estudada. O valor inicial é calculado antes da aplicação de qualquer estratégia. Também é mostrado o desvio padrão correspondente a 30 realizações para cada estratégia e modelo. A maior variação para cada modelo é destacada com fundo cinza. 47

LISTA DE SIGLAS

A_{cc}	Acessibilidade centro-centro
A_{cp}	Acessibilidade centro-periferia
A_{pp}	Acessibilidade periferia-periferia
B_{cc}	Centralidade de intermediação centro-centro
B_{cp}	Centralidade de intermediação centro-periferia
B_{pp}	Centralidade de intermediação periferia-periferia
C_{cc}	Centralidade de proximidade centro-centro
C_{cp}	Centralidade de proximidade centro-periferia
C_{pp}	Centralidade de proximidade periferia-periferia
D_{cc}	Grau centro-centro
D_{cp}	Grau centro-periferia
D_{pp}	Grau periferia-periferia
ASPL	Average shortest path length
BA	Modelo Barabási-Albert
ER	Modelo Erdős-Rényi
NTRLA	Network Topology Reconfiguration Optimization with Limited Link Addition
RGG	Random Geometric Graph
WAX	Modelo Waxman
WS	Modelo Watts-Strogatz

SUMÁRIO

1	Introdução	19
2	Ciência de Redes	23
2.1	Pequeno Mundo	24
2.2	Lei de Potência e Hubs	25
2.3	Redes Modulares	25
2.4	Propriedades Topológicas	26
2.5	Mínimo Caminho Médio	28
2.6	Acessibilidade	29
2.7	Maleabilidade	29
3	Metodologia	33
3.1	Modelos de Redes Complexas	33
3.2	Redes Modulares	37
3.3	Estratégias de Otimização	37
4	Resultados e Discussões	41
4.1	Aplicação das Estratégias	41
4.2	Estratégias Baseadas em Acessibilidade	47
4.3	Estratégias Baseadas em Centralidade de Intermediação	51
4.4	Correlação Entre Valores Iniciais e Variações do ASPL	52
4.5	Maleabilidade	52
4.6	Aplicação das Estratégias em Redes Modulares	54
4.7	Um Estudo de Caso: Rede de Aeroportos	57
5	Conclusões	59
6	Trabalhos Futuros	61
	Referências	63

1 Introdução

A ciência de redes é uma área de pesquisa bastante crescente e importante, empregada para caracterizar e modelar uma ampla gama de sistemas complexos reais, ¹ como mercado financeiro, ² sistemas de transmissão de energia elétrica, ³ interações entre proteínas, ⁴ redes de comunicação e telefonia, ⁵ sistemas de ruas de cidades, ^{6, 7} citações em trabalhos científicos ^{8, 9} ou até links entre páginas da web. ¹⁰

Para melhor entender esses sistemas, várias maneiras de estudar sua topologia foram propostas. Uma medida particularmente importante é o comprimento do mínimo caminho médio (Average Shortest Path Length - ASPL), ¹¹ que corresponde à média das distâncias mínimas entre todos os possíveis pares de vértices dentro de uma rede. Em muitos sistemas do mundo real, como nas ruas de cidades ou nas redes da malha elétrica, conhecer o caminho mais curto entre dois elementos é crucial para adotar estratégias que minimizem o tempo e os custos dos agentes que se deslocam dentro desses sistemas. Portanto, também é interessante buscar maneiras de diminuir o mínimo caminho médio de uma rede como forma de melhorar o desempenho de vários tipos de atividades, como a comunicação ou eficiência de transporte destas redes.

Nas últimas décadas vários pesquisadores têm procurado meios de aprimorar a robustez de redes através de métodos de adição de novas arestas, comumente chamados de Network Topology Reconfiguration Optimization with Limited Link Addition (NTRLA - em tradução livre, Otimização por Reconfiguração de Topologia de Redes pela Adição Limitada de Arestas), normalmente usando estratégias baseadas na distribuição de grau dos vértices. ¹² Li ¹³ propôs novos métodos e métricas para otimizar a robustez de redes livre de escala sem perda de eficiência na comunicação, enquanto Hayashi ¹⁴ explorou dinâmicas de robustez em redes geográficas com abordagens distintas. Paiva ¹⁵ investigou recentemente a ação de outras medidas, como a centralidade de proximidade (closeness centrality) e a centralidade de intermediação (betweenness centrality), como guias para a diminuição do caminho mínimo dos modelos Barabási-Albert e Erdős-Rényi, bem como de uma representação de uma rede de energia elétrica, porém estudando apenas redes formadas por poucos vértices e apenas poucas variações de algumas estratégias, por exemplo,

conectando vértices com os maiores e menores valores de centralidade de proximidade ou conectando dois vértices entre os que apresentam os menores valores de centralidade de intermediação, sem explorar todas as combinações de conexão entre valores alto/alto, baixo/baixo e alto/baixo das medidas envolvidas em estratégias desse tipo. Uma das contribuições desta dissertação será, então, a de explorar uma maior gama de medidas, considerando todas as combinações entre valores alto/alto, baixo/baixo e alto/baixo. De acordo com os trabalhos correlatos a estratégia baseada em centralidade de proximidade foi identificada como a que entregou melhores resultados dentre os estudados. A NTRLA também tem sido usada na otimização de diferentes medidas, como a conectividade algébrica.¹⁶ Porém, estes trabalhos consideraram relativamente poucas estratégias para comparação. Aqui, buscamos estudar como diferentes estratégias, baseadas em um maior conjunto de medidas, podem ser usadas para otimizar o mínimo caminho médio de modelos teóricos conhecidos de redes complexas, como o Barabási-Albert (BA), o Erdős-Rényi (ER), o Watts-Strogatz (WS) e o Waxman (WAX), bem como de redes que apresentem comunidades, também chamadas redes modulares, e da rede mundial de aeroportos, como um exemplo de rede real.

Consideramos sete estratégias diferentes: (i) conectar vértices escolhidos aleatoriamente; (ii) aumentar a regularidade do grau; (iii) adicionar arestas entre vértices escolhidos de acordo com seu grau (ligação preferencial); (iv) conectar pares de vértices com baixo e alto valor de grau; (v) conectar pares de vértices com baixa e alta centralidade de intermediação; (vi) ajustar a acessibilidade implementando conexões entre vértices centrais e periféricos; e (vii) conectar pares de vértices com baixa e alta centralidade de proximidade. Vemos que a conexão de vértices centrais e periféricos leva à diminuição mais rápida no comprimento do mínimo caminho médio. Além disso, as estratégias consideradas têm eficácia distinta nos modelos de rede, com o modelo BA resultando na redução mais lenta do ASPL. Também, quando consideramos redes modulares, as estratégias que exploram a conexão de dois vértices centrais se mostraram mais vantajosas.

Outro resultado buscado neste trabalho é a possibilidade de usar o conceito da maleabilidade das redes como ferramenta para prever se determinada rede pode ser otimizada com mais ou menos facilidade, levando em conta a variação do mínimo caminho médio.

A maleabilidade de uma rede é definida como a exponencial da entropia da probabilidade de cada possível nova configuração que pode ser obtida a partir de pequenas alterações feitas nesta rede e, na prática, ela pode ser usada para medir o quão fácil é produzir uma determinada versão da rede original a partir de poucas alterações, como no uso da NTRLA. Contudo, a maleabilidade não se mostrou um bom indicativo do potencial de otimização do ASPL de uma rede.

Este trabalho foi realizado com o grupo do Prof. Dr. Luciano da F. Costa do IFSC, que estuda, além de outros assuntos, modelagem de sistemas e redes complexas com aplicações em sistemas reais, explorando dinâmicas e características topológicas. Nesta dissertação, começamos por apresentar conceitos de ciência de redes que usaremos no decorrer do trabalho, bem como novas medidas recentemente propostas, os modelos de redes adotados e as estratégias de otimização consideradas. Em seguida, os resultados são apresentados e discutidos.

2 Ciência de Redes

A Ciência de redes tem suas origens na teoria de grafos, que envolve representações de conjuntos de vértices e arestas. Uma rede é a aplicação da estrutura de grafos para representar um sistema complexo, sendo formada por um grande conjunto de vértices, ou nós, representando elementos deste sistema e conectados por arestas que, por sua vez, representam as relações entre estes elementos. Podemos dizer que qualquer sistema discreto pode ser representado com o uso de redes, visto que nas últimas décadas diversos sistemas diferentes têm sido modelados com redes complexas, cobrindo uma variedade de assuntos que vão desde infraestrutura de redes de distribuição elétrica, redes sociais e relações entre páginas da Internet, até interação entre proteínas, publicações de artigos científicos e disseminação de doenças ou opiniões. Aos elementos da rede, ou vértices, podemos incluir atributos como pesos ou estados e usá-los para trabalhar dinâmicas do sistema, possibilitando prever o tempo que leva para uma informação percorrer a rede ou para que uma doença se espalhe em uma população, por exemplo. Às arestas, que modelam as interações entre os vértices, podemos atribuir pesos que indicam a força com que estas interações ocorrem ou a distância que separam esses vértices, como no caso de redes geográficas.

Matematicamente, a forma mais simples de representação de uma rede complexa é através de sua matriz de adjacência M_{adj} , de dimensão $N \times N$, sendo N o número de vértices da rede, onde cada entrada da matriz m_{ij} assume um valor diferente de zero quando o vértice i é ligado ao vértice j . O valor de m_{ij} representa o peso da aresta que liga i e j , quando existir, ou simplesmente 1, quando as arestas da rede não contarem com pesos. O modelo de rede mais simples, por apresentar uma maior uniformidade estatística, é o modelo Erdős-Rényi, onde os vértices da rede são conectados de forma aleatória seguindo uma probabilidade de conexão constante.

Diferentes modelos foram propostos para representar a estrutura topológica de sistemas reais e, curiosamente, diversos padrões acerca da formação e da estrutura de redes começaram a emergir, sendo constatado que grande parte destes sistemas compartilha de características em comum, que podem ser modeladas pelos fenômenos de pequeno mundo,

identificado por Watts e Strogatz,¹⁷ e pela distribuição de lei de potências para a distribuição de grau nas redes, como descrito por Barabasi¹⁸ ao estudar a rede da Internet.

2.1 Pequeno Mundo

O fenômeno de pequeno mundo, também conhecido popularmente como “seis graus de separação”, pode ser encontrado em diversas redes do mundo real, desde redes de distribuição de energia e redes sociais até redes neurais. Se considerarmos o deslocamento de um agente visitando vértices conectados em uma rede com característica de pequeno mundo, podemos tomar um par de vértices qualquer A e B e encontrar um caminho ligando ambos com apenas alguns passos de distância, ou seja, nosso agente precisaria visitar um número relativamente baixo de vértices para ir de A até B .

Watts e Strogatz¹⁷ propuseram um modelo de construção de redes que reproduzisse este fenômeno. Partindo de uma rede sem ligações de longa distância entre vértices, como em uma topologia de grade ou de círculo, podemos remover uma aresta e reconectá-la à um vértice escolhido aleatoriamente, como ilustrado na Figura 1, criando um atalho na rede e conectando duas regiões que, sem este atalho, estariam topologicamente muito distantes. Com o aumento do número de atalhos, os vértices desta rede estariam “mais próximos” entre si, aumentando consideravelmente a velocidade de transmissão de informações dentre os vértices da rede.

Figura 1 – Processo de construção do modelo Watts-Strogatz, onde começamos com (A) uma rede com topologia em anel e (B) arestas são removidas e reconectadas aleatoriamente em sucessivas iterações.

Fonte: Elaborada pelo autor.

2.2 Lei de Potência e Hubs

O modelo Barabási-Albert ¹⁸ leva em consideração a formação e o crescimento de redes reais e prevê uma distribuição do número de vizinhos k de cada vértice seguindo o que chamamos de lei de potências, ou seja, a probabilidade de encontrarmos um vértice com determinado número de vizinhos tem a forma $P(k) \propto k^{-\gamma}$, sendo γ um parâmetro específico da rede. Esse comportamento é explicado pelo fato de que novos vértices incluídos na rede têm maior probabilidade de se ligar a vértices melhor conectados, com um maior número de vizinhos. Diversas redes reais apresentam essa dinâmica de formação, como nos exemplos da rede de páginas da Internet ou de publicação de artigos, em que novos sites têm maior probabilidade de apontar para sites maiores, mais famosos e um artigo recém-publicado tende a citar artigos mais conhecidos.

A distribuição em lei de potências também prevê que a rede seja livre de escala, ou seja, na mesma rede são encontrados vértices com poucos vizinhos, muitos vizinhos e com todas as possibilidades entre esses dois extremos. Em geral, têm-se muitos vértices com um número baixo de vizinhos e poucos vértices expoentes com um grande número deles, devido à própria dinâmica de formação das redes. Esses vértices expoentes são também chamados *hubs*, funcionando como nós centrais que conectam a maioria dos vértices de sua região, formando vias principais de conexão com outros *hubs*.

2.3 Redes Modulares

Outro fenômeno conhecido da literatura é o da formação de comunidades na topologia das redes, estas chamadas de redes modulares. Quando uma determinada região da rede apresenta um padrão de conexões diferente das demais ou quando os vértices desta região são muito mais conectados entre si do que com vértices de outras regiões podemos identificar estas diferentes regiões como sendo comunidades distintas da rede, ou módulos. Diversos algoritmos para a identificação destas comunidades existem, ¹⁹ como o método Infomap de Rosvall, ²⁰ o método de auto-vetores de Newman ²¹ e o método baseado em caminhadas aleatórias de Latapy e Pons. ²² Há também uma medida dedicada à quantificação da força das estruturas dessas comunidades, que apresenta uma métrica

do número de comunidades em que uma rede é melhor separada, a *modularidade*.²³ Esta medida é importante para o uso de algoritmos supervisionados de identificação de comunidades, onde é preciso saber o número de comunidades para identificá-las.

Diferentes formas de se construir um modelo que reproduza o fenômeno de comunidades em redes têm sido discutidos na literatura, como o Fortunato²⁴ e o Blockmodel.²⁵ Apresentamos o modelo Fortunato na seção 3.2.

2.4 Propriedades Topológicas

A modelagem de sistemas reais se torna interessante à medida que podemos usá-la para entender esses sistemas e fazer previsões sobre seu comportamento e evolução e, para a realização desta tarefa, diferentes propriedades topológicas foram propostas no decorrer dos anos. São chamadas medidas topológicas as que levam em conta a estrutura da rede e a conexão dos vértices com seus vizinhos para caracterizá-la. Aqui, citamos algumas:

- a) Grau: Já usamos o número de vizinhos de um vértice V , também chamado de conectividade ou grau $k(V)$ do vértice, para apresentar outros conceitos acerca da estrutura de redes complexas e, de fato, a distribuição do grau é uma importante medida para sua caracterização, não apenas para a topologia, mas também como uma medida usada na definição de diferentes medidas mais sofisticadas e que influencia diretamente nas dinâmicas que podem acontecer na rede.
- b) Centralidade de intermediação: A centralidade de intermediação²⁶ mede o quão central é um vértice V considerando a fração dos caminhos mínimos que passam por V , isto é, qual a fração do número de caminhos mínimos entre dois outros vértices i e j que passam por V em relação ao número total de caminhos mínimos que ligam i e j . Note que um caminho mínimo entre dois vértices i e j corresponde a um ou mais caminhos que ligam i e j visitando o menor número de vértices possível.

Esta medida pode ser calculada através da equação:

$$B(V) = \sum_{i \neq V \neq j} \frac{\sigma_{ij}(V)}{\sigma_{ij}} \quad (1)$$

onde $\sigma_{ij}(V)$ representa o número de caminhos mínimos entre i e j que passam por V , enquanto σ_{ij} representa o número de caminhos mínimos entre i e j que não necessariamente passam por V .

Vale notar que os vértices de maior centralidade de intermediação de uma rede correspondem aos *hubs* da rede, já que representam os nós que participam dos principais caminhos mínimos que ligam diferentes regiões da rede.

- c) Centralidade de proximidade: De forma semelhante à medida anterior, a centralidade de proximidade²⁷ mede o quão central é um vértice V , porém, aqui consideramos a distância média de V aos demais vértices da rede. Podemos, então, defini-la como:

$$C(V) = \frac{N - 1}{\sum_j d_{jV}} \quad (2)$$

onde d_{jV} é o comprimento do caminho mínimo entre V e j .

- d) Coeficiente de agrupamento: O coeficiente de agrupamento¹ de um dado vértice V mede o quanto seus vizinhos são conectados entre si. Este coeficiente pode ser calculado através da razão entre o número $e(V)$ de arestas existentes entre vértices vizinhos de V e o número do total de arestas que poderiam existir entre estes vértices, ou seja:

$$CC(V) = \frac{2e(V)}{k(V)(k(V) - 1)} \quad (3)$$

Esse tipo de modelagem permite estudar não apenas a participação de um indivíduo no sistema em que se encontra, mas também o comportamento do sistema como um todo, onde a soma das interações pode produzir fenômenos globais. Para quantificar estes efeitos, existem as medidas globais da rede, como as médias do grau ou da centralidade de intermediação de toda a rede, bem como o mínimo caminho médio que introduziremos a seguir.

Figura 2 – Nesta figura podemos ver um exemplo de um grafo (a) com um dos caminhos mínimos entre os vértices A e B destacado, com $d_{AB} = 5$. Com a adição de uma única aresta (b, em azul), o caminho mínimo é reduzido para $d_{AB} = 2$.

Fonte: Elaborada pelo autor.

2.5 Mínimo Caminho Médio

A distância mínima entre dois vértices A e B em uma rede, d_{AB} , corresponde ao número mínimo de arestas que precisam ser percorridas ao mover-se de A para B . O comprimento do mínimo caminho médio $ASPL$ ¹¹ de uma determinada rede é, então, definido como a média das distâncias mínimas entre todos os pares de vértices desta rede.

O caminho mínimo entre dois vértices pode variar substancialmente como consequência de pequenas alterações na topologia da rede, conforme ilustrado na Figura 2.

O $ASPL$ pode ser usado como um indicador de eficiência de certas dinâmicas em algumas redes, influenciando diretamente na habilidade de transmissão de informações em uma rede ou como um agente pode se mover para visitar diferentes vértices. A facilidade com que pedestres se deslocam em uma cidade pode ser considerada inversamente relacionada ao comprimento do mínimo caminho médio da respectiva representação em rede complexa desta cidade. A redução do $ASPL$ por meio de estratégias NTRLA, portanto, pode ser usada para otimizar dinâmicas na rede (como em [28]).

2.6 Acessibilidade

Seja um vértice V , sua acessibilidade $A_h(V)$ ²⁹ é uma medida de quão facilmente acessível são os demais vértices à distância h a partir de V . O conjunto destes vértices é chamado de vizinhança de nível h do vértice V , como pode ser visto na Figura 3. Essa medida considera as probabilidades de transição P_h de um caminhante aleatório auto-excludente, i.e. um caminhante totalmente independente que não pode retornar a um vértice já visitado anteriormente, saindo de V e indo até um vértice destino j à distância h e é definida pelo exponencial da entropia dessas probabilidades, ou seja:

$$A_h(V) = \exp(E_h(V, \Omega)) \quad (4)$$

onde Ω é o conjunto de todos os vértices alcançáveis a partir de V em h passos, N é o tamanho do conjunto Ω e $E_h(V, \Omega)$ é dado por:

$$E_h(V, \Omega) = - \sum_{j=1}^N P_h(V, j) \log(P_h(V, j)) \quad (5)$$

A exponencial da entropia é usada nesta definição para recuperar a correspondência com o grau hierárquico do vértice V no caso de transições equiprováveis, ou seja, quando as probabilidades de um caminhante aleatório alcançar cada vértice forem iguais, a acessibilidade é máxima e numericamente igual ao grau hierárquico de V .

A acessibilidade pode ser usada para identificar bordas em redes. Mais especificamente, os vértices na região mais central da rede tendem a apresentar valores mais altos de acessibilidade, enquanto os vértices pertencentes à periferia/borda tendem a ter uma menor acessibilidade. ²⁹

2.7 Maleabilidade

A maleabilidade ³⁰ de uma rede quantifica a facilidade com que ela pode ser modificada (para assumir diferentes formas topológicas) como consequência de alterações incrementais na sua topologia (por exemplo, com estratégias NTRLA).

Figura 3 – Acessibilidade de dois vértices com vizinhança e probabilidades de transição diferentes. O vértice em vermelho (V) indica o vértice de referência, enquanto os vértices em verde e em amarelo formam, respectivamente, a primeira ($h = 1$) e a segunda ($h = 2$) vizinhanças de V . Os valores de p junto a cada vértice representam a probabilidade de transição de serem alcançados por um caminhante aleatório partindo de V . (a) Quando as probabilidades de transição são similares, a acessibilidade é alta, neste caso sendo igual a $A_2(V) = 8$ para a segunda vizinhança de V . (b) Quando as probabilidades de transição são heterogêneas, a acessibilidade de V é menor, sendo igual a $A_2(V) = 1,65$ neste exemplo.

Fonte: Elaborada pelo autor.

Dada uma rede, é possível remover ou adicionar arestas, uma por uma, para produzir novas redes, que podem ainda ser isomorfas quando comparadas à rede original. Duas redes são consideradas isomorfas quando elas são iguais a menos de uma renomeação de seus vértices, ou ainda, a menos de uma permutação de colunas (ou linhas) na respectiva matriz de adjacência destas redes, mantendo ao menos uma coluna (ou linha) fixa. Assim, para o cálculo da maleabilidade, redes isomorfas são consideradas como sendo iguais e não uma modificação.

A partir da rede original S_0 , é possível alcançar redes diferentes, chamadas *estados*, e calcular a probabilidade P_{S_0, S_i} de S_0 ser transformada em um determinado estado S_i como consequência de alguma alteração em S_0 (como a criação ou remoção de uma aresta). Diante da alta complexidade computacional de verificar se existe isomorfismo dentre as diferentes redes possíveis a partir de S_0 , sendo o objetivo deste trabalho o de analisar a relação entre a maleabilidade e a evolução de uma única medida (ASLP), podemos nos valer do conceito de *maleabilidade relativa a uma determinada medida*, simplificando a checagem por isomorfismos. Sabendo que duas redes isomorfas resultarão, necessaria-

mente, nos mesmos valores de ASPL, podemos recalcular a medida após cada modificação da rede e considera-las equivalentes quando o ASPL resultante for igual ao encontrado na rede original.

A entropia de todas as transições possíveis, que representa a heterogeneidade da distribuição de probabilidades, pode ser calculada por:

$$E_{S_0} = - \sum_{S_i} P_{S_0, S_i} \log(P_{S_0, S_i}) \quad (6)$$

A maleabilidade é, então, definida como:

$$A_{S_0} = \exp(E_{S_0}) \quad (7)$$

Quando existe apenas um estado possível acessível a partir de S_0 , temos $A_{S_0} = 1$. Se todos os estados forem alcançáveis com a mesma probabilidade, a maleabilidade será igual ao número total de estados possíveis C . Caso as probabilidades de transição sejam diferentes, a maleabilidade assume valores entre 1 e C .

3 Metodologia

Este capítulo apresenta os modelos utilizados em nossas simulações e também define as estratégias que serão utilizadas na adição de arestas à estes modelos.

Os modelos de redes usados no estudo proposto foram gerados e manipulados computacionalmente através do uso das bibliotecas de ferramentas *iGraph* e *NetworkX*, disponíveis para a linguagem *Python*. Também foram usadas essas ferramentas para as definições e aplicações das estratégias de adição de arestas usadas na redução do ASPL.

3.1 Modelos de Redes Complexas

Quatro modelos teóricos de redes complexas já conhecidos foram utilizados neste estudo. Esses modelos foram escolhidos por apresentarem aspectos variados, cobrindo boa parte dos diferentes casos de interesse para nossas comparações. São eles:

- a) o modelo Barabási-Albert (BA),¹ consistindo de uma rede aleatória na qual vértices são adicionados um a um, tendo seus novos vértices conectados a k_{BA} vértices já existentes com uma probabilidade proporcional ao seu grau. Aqui, usamos $k_{BA} = 3$. O modelo BA é um bom representante para redes como a Internet, onde existem *hubs* ligando um grande número de vértices, sendo uma rede livre de escala e com a distribuição de grau dos vértices seguindo uma lei de potências, como explicado na Seção 2.1. A Figura 4 representa um exemplo de visualização de uma rede BA, onde podemos identificar o aparecimento de *hubs* em sua região mais central;
- b) o modelo Erdős-Rényi (ER), onde pares de vértices são escolhidos aleatoriamente e conectados com uma probabilidade fixa p_{ER} . Por ser uniformemente aleatório e ter uma distribuição de grau estatisticamente regular em torno da média, o modelo ER é considerado o mais simples das redes complexas, no sentido de ter menor complexidade, já que a homogeneidade de seus vértices carrega menos informação. A Figura 5 representa um exemplo de visualização de uma rede ER;
- c) o modelo Watts-Strogatz (WS), no qual, a partir de uma malha em anel regular,

Figura 4 – Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Barabási-Albert.

Fonte: Elaborada pelo autor.

Figura 5 – Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Erdős-Rényi.

Fonte: Elaborada pelo autor.

Figura 6 – Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Watts-Strogatz.

Fonte: Elaborada pelo autor.

cada aresta tem uma probabilidade p_{WS} de se reconectar com outro vértice, diferente dos originais, escolhido aleatoriamente, como representado na Figura 1. Adotamos $p_{WS} = 0,01$ para o modelo WS em nossos experimentos. A Figura 6 representa um exemplo de visualização de uma rede WS, onde podemos notar que há poucas conexões de longa distância, em decorrência do baixo valor de p_{WS} , e existem regiões que mantem a forma da topologia original, em anel, da rede;

- d) o modelo Waxman (WAX),³¹ uma rede geométrica construída atribuindo-se aleatoriamente vértices em um determinado espaço e conectando-os com uma probabilidade p_{ij} , conforme:

$$p_{ij} = \alpha \exp^{-d_{ij}/(L\beta)} \quad (8)$$

onde d_{ij} é a distância entre os vértices i e j , L é a distância máxima entre qualquer par de vértices e α e β são parâmetros. A Figura 7 representa um exemplo de visualização de uma rede Waxman, onde vemos que não ocorrem ligações de longa distância.

Figura 7 – Exemplo de uma rede formada com $N = 100$ vértices, seguindo o modelo Waxman.
Fonte: Elaborada pelo autor.

Neste trabalho $\beta = 0,1$ foi usado nos experimentos, significando que as conexões entre vértices são feitas dentro de uma distância típica de $0,1$ do tamanho L da rede, de forma a produzir redes esparsas. O valor de α é ajustado de forma a gerar redes com grau médio em torno de 6.

Redes geométricas ou espaciais são redes em que as posições e adjacências dos vértices influenciam suas conexões no ato da construção da rede. Em geral elas apresentam baixa complexidade e um ASPL alto. Como outros modelos de redes geométricas, podemos citar o modelo Voronoi,³² onde o espaço é particionado de forma equidistante à vértices vizinhos, ligando vértices que representam subpartições adjacentes, e o RGG (Random Geometric Graph),³³ onde os vértices são conectados à todos os demais dentro de uma distância máxima.

Neste trabalho, usaremos redes com $N = 3000$, $N = 5000$ e $N = 8000$ vértices para cada modelo apresentado acima, todas com grau médio em torno de 6.

3.2 Redes Modulares

Para simular o efeito de comunidades em nossas redes usaremos dois modelos distintos: um modelo Waxman adaptado e o modelo proposto por Fortunato, Lancichinetti e Radicchi.²⁴

Nosso modelo Waxman adaptado para redes modulares simplesmente envolve definir N_c centróides, em torno dos quais podemos distribuir N/N_c vértices e construir uma rede Waxman a partir desta configuração, bastando que a distância média dos vértices em torno de seu centróides correspondente seja menor que a distância típica que separa os centróides. Em nossas simulações usaremos $N_c = 4$.

Já o modelo Fortunato envolve criar diferentes “sub-redes” a partir do modelo BA, onde o número de vértices de cada uma destas sub-redes, ou comunidades, é determinado por uma distribuição na forma $n(x) = \exp(x/\delta)$, escolhendo o número de comunidades N_c de forma que a soma dos vértices das comunidades seja N . Assim, garantimos uma variabilidade no tamanho e na quantidade de comunidades em cada rede. Em seguida, uma fração μ (parâmetro de mistura) das arestas de uma comunidade são realocadas para se conectarem com outras comunidades, formando a rede principal. Os parâmetros usados em nossas simulações são $\delta = 1,5$ e μ variando entre os valores 0,03, 0,05, 0,07 e 0,1. As redes são formadas por $N = 3000$ vértices e apresentam grau médio em torno de 6 e, além disso, o tamanho mínimo de uma comunidade foi restrito a 15 vértices.

3.3 Estratégias de Otimização

As abordagens descritas a seguir têm o objetivo de reduzir o comprimento do mínimo caminho médio das redes consideradas, a fim de otimizá-las. Todas as estratégias consideradas envolvem a adição de um mesmo número de novas arestas N_a . No contexto de estratégias NTRLA é desejável manter N_a baixo, visto que ao incluir um grande número de arestas aos modelos teóricos, estes podem sofrer distorções em sua topologia e não mais se comportarem como o esperado nos modelos originais.

Escolha aleatória: É interessante implementarmos uma escolha aleatória de pares

de vértices para a adição de novas arestas a fim de comparar o desempenho de estratégias mais sofisticadas, que definiremos a seguir.

Tornando a topologia mais regular: A primeira estratégia que propomos será tentar tornar o grau dos vértices mais uniforme, ou seja, reduzir o desvio padrão do grau, produzindo uma nova versão da rede, mais próxima de um grafo regular (todos os vértices com o mesmo grau). A motivação para esta estratégia é que as redes regulares têm a tendência de apresentar um ASPL menor.³⁴ A estratégia adotada consiste em conectar dois vértices escolhidos aleatoriamente que tenham grau menor que o atual grau médio da rede $\langle k \rangle$.

Ligação Preferencial: A segunda estratégia é executar o procedimento de *ligação preferencial*³⁵ na rede, conectando vértices escolhidos aleatoriamente a outros vértices i , de acordo com uma probabilidade p_i dada como $p_i \propto k(i)$, onde $k(i)$ é o grau de i . A estratégia de ligação preferencial neste trabalho é motivada pelo fato de que muitas redes do mundo real têm uma distribuição de grau dada por uma lei de potência, que pode ser modelada pelo conceito de construção por ligação preferencial.

Abordagem baseada em acessibilidade: A característica da acessibilidade de identificar bordas na rede sugere que podemos tentar otimizar seu mínimo caminho médio, melhorando a acessibilidade dos vértices das regiões de pior interconexão da rede, considerando as seguintes estratégias: (a) adicionar arestas entre vértices centrais e periféricos, conectando o vértice com o maior valor de acessibilidade ao vértice com o menor valor, depois o vértice com o segundo maior valor ao vértice com segundo menor valor, e assim sucessivamente, até que Na novas arestas sejam adicionadas – chamaremos a estratégia de *Acessibilidade centro-periferia* ou A_{cp} ; (b) interconectar Na pares de vértices periféricos, ligando dois vértices escolhidos aleatoriamente dentre o conjunto de vértices com os menores valores de acessibilidade – chamaremos de *Acessibilidade periferia-periferia* ou A_{pp} ; e (c) interconectar Na pares de vértices centrais, ligando dois vértices escolhidos aleatoriamente dentre o conjunto de vértices com os maiores valores de acessibilidade – chamaremos de *Acessibilidade centro-centro* ou A_{cc} .

Dado que a fração de vértices visitados pelo caminhante aleatório é influenciada pelo

tamanho da rede, é interessante que o parâmetro h seja escolhido proporcionalmente ao número de vértices N . Para a aplicação das estratégias propostas, tomamos $h = N/1000$.

Conectando vértices por grau: Outra estratégia usada para otimizar o mínimo caminho médio de uma rede envolve adicionar novas arestas entre vértices com graus relativamente pequenos e grandes, como uma tentativa de fornecer rotas alternativas, possivelmente mais curtas, para os caminhos que passam pelos vértices de grau pequeno. Definimos as estratégias D_{cp} , D_{pp} e D_{cc} de forma similar às adotadas para a acessibilidade.

Abordagem baseada em centralidade de intermediação: Medidas de centralidade ²⁶ são importantes para identificar vértices que podem ser potencialmente críticos, por exemplo, no sentido de transportar mais informações ou de serem mais visitados por agentes viajantes. Vértices mais centrais também participam de um maior número de caminhos mínimos em comparação a vértices menos centrais. Assim, uma estratégia que pode ser usada na tentativa de aprimorar a centralidade é conectar vértices com os menores e maiores valores de centralidades de intermediação. Definimos as estratégias B_{cp} , B_{pp} e B_{cc} de forma similar às adotadas para a acessibilidade.

Abordagem baseada em centralidade de proximidade: Centralidade de proximidade ²⁷ apresenta alguns comportamentos similares à acessibilidade no uso de detecção de bordas em redes e também consiste em uma medida de centralidade, como a centralidade de intermediação, sendo portanto mais uma candidata para basear uma estratégia para diminuir o ASPL. Novamente, podemos definir as estratégias C_{cp} , C_{pp} e C_{cc} de forma semelhante às adotadas para a acessibilidade.

4 Resultados e Discussões

Este capítulo começa apresentando os resultados das aplicações das estratégias propostas aos modelos de rede considerados e discute comportamentos que possam explicar os desempenhos encontrados. A correlação entre valor inicial e variação do *ASPL* é apresentada e discutida, bem como sua relação com a maleabilidade das redes. Por fim, redes com presença de comunidades e uma rede real são consideradas nas simulações.

Redes de diferentes tamanhos foram geradas, usando $p_{ER} = 0,0021$ e $\alpha = 0,0122$ em redes com $N = 3000$ vértices; $p_{ER} = 0,0012$ e $\alpha = 0,0071$ para $N = 5000$ vértices; e $p_{ER} = 0,00075$ e $\alpha = 0,0045$ quando $N = 8000$, com estes parâmetros escolhidos de forma a produzir redes com grau médio em torno de 6, sendo que todas as redes do modelo Barabási-Albert são construídas com $k_{BA} = 3$. Um total de 30 instâncias de cada modelo foi considerado e $N_a = 200$ iterações de cada uma de nossas estratégias foram aplicadas, com o mínimo caminho médio recalculado a cada iteração.

4.1 Aplicação das Estratégias

A variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada uma das topologias de redes estudadas é mostrada nas Figuras 8 a 13 para redes construídas com $N = 3000$, $N = 5000$ e $N = 8000$ vértices nas quatro topologias adotadas. As barras de erro apresentadas são mostradas como linhas verticais com 40% do desvio padrão dos dados e o eixo vertical é normalizado em relação ao valor inicial do *ASPL*. Vale notar que os eixos verticais dos gráficos correspondentes ao modelo WS mostram um intervalo diferente dos demais gráficos. Dado que o cálculo da acessibilidade com $h = 8$ em uma rede BA é consideravelmente mais custoso que para as demais redes, decidimos não considerar o modelo BA com $N = 8000$ vértices em nossas análises.

Com exceção do modelo WS, as estratégias C_{cp} , A_{cp} , B_{cp} e D_{cc} apresentaram os melhores resultados quando comparadas às demais estratégias, estando em concordância também com resultados de trabalhos correlatos de outros autores, sendo superadas ape-

Figura 8 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo BA. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.

Fonte: Elaborada pelo autor.

Figura 9 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo ER. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.

Fonte: Elaborada pelo autor.

Figura 10 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo WS. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.

Fonte: Elaborada pelo autor.

Figura 11 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para o modelo Waxman. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6.

Fonte: Elaborada pelo autor.

Figura 12 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada uma das topologias de redes. Aqui, todas as redes são formadas por $N = 5000$ vértices e grau médio em torno de 6.

Fonte: Elaborada pelo autor.

Figura 13 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada uma das topologias de redes. Aqui, todas as redes são formadas por $N = 8000$ vértices e grau médio em torno de 6.

Fonte: Elaborada pelo autor.

Tabela 1 – As três melhores estratégias listadas para cada modelo, com $N = 3000$, $N = 5000$ e $N = 8000$ vértices depois de 20 iterações de cada estratégia proposta. Os números entre parênteses indicam a porcentagem das vezes em que a estratégia resultou na melhor, segunda melhor e terceira melhor redução do ASPL, respectivamente. Resaltamos que este trabalho não apresenta as simulações feitas em redes do modelo BA com $N = 8000$, como explicado nas discussões.

	BA	ER	Wax	WS
N=3000	1.Ccp (67%, 30%, 3%)	1.Ccp (93%, 7%, 0%)	1.Ccp (93%, 7%, 0%)	1.Dcc (70%, 17%, 0%)
	2.Acp (3%, 50%, 30%)	2.Acp (3%, 87%, 10%)	2.Acp (7%, 93%, 0%)	2.Acp (0%, 30%, 33%)
	3.Dcc (16%, 3%, 23%)	3.Bcp (0%, 0%, 43%)	3.Bcp (0%, 0%, 63%)	3.Acc (3%, 17%, 20%)
N=5000	1.Ccp (100%, 0%, 0%)	1.Ccp (97%, 3%, 0%)	1.Ccp (100%, 0%, 0%)	1.Bcc (53%, 43%, 0%)
	2.Acp (0%, 100%, 0%)	2.Acp (3%, 97%, 0%)	2.Acp (0%, 100%, 0%)	2.Acc (40%, 37%, 17%)
	3.Dcc (0%, 0%, 40%)	3.Dcc (0%, 0%, 40%)	3.Bcp (0%, 0%, 57%)	3.Dcc (7%, 7%, 27%)
N=8000		1.Ccp (100%, 0%, 0%)	1.Ccp (100%, 0%, 0%)	1.Bcc (63%, 37%, 0%)
		2.Acp (0%, 100%, 0%)	2.Acp (0%, 100%, 0%)	2.Acc (33%, 33%, 20%)
		3.Dcc (0%, 0%, 73)	3.Bcp (0%, 0%, 47%)	3.Dcc (3%, 20%, 60%)

Fonte: Elaborada pelo autor.

nas pela estratégia A_{cc} no caso dos modelos WS. Possivelmente, a diferença encontrada no modelo WS é consequência do fato de que redes WS, quando considerados valores pequenos de p_{WS} , possuem poucos atalhos, não sendo suficientes para uma divisão clara entre regiões de centro e de periferia, que é uma característica necessária para as estratégias que se utilizam desta configuração. Também uma possível consequência do baixo número de atalhos na rede é o resultado da estratégia de conectar um par de vértices escolhidos aleatoriamente ter tido um alto desempenho para o modelo WS com $N = 3000$ vértices, já que qualquer abordagem de como escolher este par de vértices implicaria em uma diminuição considerável do ASPL. O modelo WS foi o único caso em que nenhuma das estratégias produziram curvas retilíneas para a redução do ASPL, provavelmente uma consequência de os valores iniciais do ASPL no modelo WS, serem bem maiores que os encontrados nos outros modelos, implicando em uma redução mais rápida. Em todos os modelos adotados, a pior redução do mínimo caminho médio foi observada para as estratégias com interação periferia-periferia.

O principal efeito observado quando aumentamos o número de vértices N é uma redução na velocidade de diminuição do $ASPL$ quando comparamos algumas estratégias, fazendo com que a ordem das melhores estratégias para determinada rede sofra alterações ao considerar maiores valores de N_a .

Figura 14 – Ordem dos melhores desempenhos das estratégias consideradas em função da taxa de adição de arestas N_a/N para a rede formada por $N = 3000$ vértices seguindo o modelo ER.

Fonte: Elaborada pelo autor.

Para resumir, a Tabela 1 mostra a classificação das estratégias de melhor desempenho para cada modelo, considerando redes de $N = 3000$, $N = 5000$ e $N = 8000$ vértices, após a adição de 20 arestas, para evidenciar o desempenho imediato de cada estratégia. Podemos ver na Figura 14 que a ordem dos melhores desempenhos das estratégias consideradas sofre alterações devido à adição de um número alto de arestas, possivelmente ocasionando na alteração da estrutura topológica particular de cada modelo e, portanto, escolhemos avaliar o desempenho relativo das estratégias, na Tabela 1, em um instante que precede uma maior variação na ordem dos melhores desempenhos e que corresponde a adição da vigésima aresta. Embora as barras de erro das Figuras 8 a 13 levantem incertezas quanto a qual estratégia se saiu melhor, a ordem com que as estratégias aparecem, em função do seu desempenho, se manteve consistente para as diversas realizações dos experimentos, independente de flutuações nos valores de ASPL. Para verificar esta consistência nos resultados podemos analisar as porcentagens das vezes em que determinada estratégia resultou em uma dada posição, também mostradas na Tabela 1.

A Tabela 2 apresenta a redução do ASPL relativa ao seu valor inicial, ou seja, antes

Tabela 2 – Comparação dos valores de variação percentual do ASPL em relação aos respectivos valores iniciais, depois de 200 iterações de cada estratégia estudada. O valor inicial é calculado antes da aplicação de qualquer estratégia. Também é mostrado o desvio padrão correspondente a 30 realizações para cada estratégia e modelo. A maior variação para cada modelo é destacada com fundo cinza.

	BA	ER	WS	WAX
Valor inicial	$3,87 \pm 0,03$	$4,56 \pm 0,02$	21 ± 1	$4,70 \pm 0,03$
Escolha aleatoria	$-0,42 \pm 0,02$	$-1,02 \pm 0,03$	-46 ± 3	$-1,28 \pm 0,03$
Topologia regular	$-0,34 \pm 0,02$	$-0,81 \pm 0,02$	-21 ± 1	$-1,27 \pm 0,05$
Lig. preferencial	$-0,52 \pm 0,06$	$-1,04 \pm 0,02$	-27 ± 1	$-1,66 \pm 0,07$
D_{cp}	$-0,84 \pm 0,04$	-1.47 ± 0.04	-22 ± 1	-2.37 ± 0.09
D_{pp}	-0.33 ± 0.02	-0.73 ± 0.02	-15 ± 1	-1.07 ± 0.05
D_{cc}	-1.27 ± 0.09	-1.82 ± 0.05	-27 ± 1	-2.5 ± 0.2
A_{cp}	-1.43 ± 0.03	-1.58 ± 0.04	-45 ± 2	-2.33 ± 0.08
A_{pp}	-0.36 ± 0.01	-0.75 ± 0.02	-27 ± 4	-0.94 ± 0.03
A_{cc}	-0.84 ± 0.09	-1.58 ± 0.05	-43 ± 3	-1.49 ± 0.08
C_{cp}	-1.67 ± 0.06	-1.62 ± 0.04	-31 ± 2	-2.39 ± 0.09
C_{pp}	-0.374 ± 0.009	-0.75 ± 0.02	-10 ± 1	-0.93 ± 0.04
C_{cc}	-0.7 ± 0.1	-1.55 ± 0.05	-21 ± 2	-1.52 ± 0.08
B_{cp}	-1.02 ± 0.05	-1.51 ± 0.03	-26 ± 1	-2.46 ± 0.09
B_{pp}	-0.26 ± 0.02	-0.72 ± 0.02	-17 ± 2	-1.07 ± 0.05
B_{cc}	-1.21 ± 0.09	-1.79 ± 0.06	-32 ± 1	-2.4 ± 0.1

Fonte: Elaborada pelo autor.

da aplicação de qualquer estratégia de otimização, para cada modelo e estratégia. De um modo geral, as maiores melhorias foram observadas para o modelo WS, sendo o modelo BA o que produziu as piores melhorias, mostrando mais resistência contra as estratégias empregadas.

4.2 Estratégias Baseadas em Acessibilidade

A acessibilidade, bem como a centralidade de proximidade, pode ser usada para discriminar a posição dos vértices, da região central da rede e da sua borda, já que vértices mais centrais tendem a apresentar valores mais altos de acessibilidade em comparação aos vértices de sua borda, que apresentam acessibilidade menor.^{29, 36} Portanto, como vértices na borda da rede tendem a estar mais mal conectados entre si, o caminho mínimo entre dois vértices periféricos tende a atravessar a região central da rede, ou seja, o caminho

Figura 15 – Exemplo de uma rede, com vértices coloridos de acordo com seus valores de acessibilidade, antes (a) e depois (b) da aplicação da estratégia A_{cp} . O caminho mínimo entre um mesmo par de vértices está indicado em vermelho nas duas redes.

Fonte: Elaborada pelo autor.

mais curto entre dois vértices com baixa acessibilidade tende a passar por vértices de alta acessibilidade, conforme ilustrado na Figura 15(a). As estratégias C_{cp} e A_{cp} produziram os melhores desempenhos para as simulações consideradas, pois essas estratégias, ao conectar vértices de alta e baixa acessibilidade/centralidade de proximidade, tendem a fornecer atalhos entre os vértices periféricos (ver Figura 15(b)).

O resultado da maior efetividade destas estratégias foi estudado analiticamente pelo Prof. Dr. Cesar Henrique Comin, em colaboração na realização deste trabalho, usando a rede com topologia em estrela mostrada na Figura 16(a). Esta rede é formada por k braços, cada braço sendo formado por dois vértices. O vértice central é um *hub* na rede, com grau k . Usando os índices mostrados na figura, a média dos comprimentos dos caminhos mínimos partindo do vértice 1 é dado por

$$l_1 = \frac{3k}{2k}. \quad (9)$$

De maneira similar, as médias dos comprimentos dos caminhos mínimos partindo dos vértices 2 a $k + 1$ podem ser calculadas como

Figura 16 – Duas possíveis estratégias de se adicionar uma aresta em uma rede com topologia em estrela. (a) Rede original. (b) Um vértice periférico é conectado a um vértice central. (c) Dois vértices periféricos são conectados entre si.

Fonte: Elaborada pelo autor.

$$l_{2..k+1} = \frac{5k - 3}{2k}. \quad (10)$$

Os vértices de índice $k + 2$ até $2k + 1$ têm uma média dos comprimentos dos caminhos mínimos de

$$l_{k+2..2k+1} = \frac{7k - 4}{2k}. \quad (11)$$

Portanto, o comprimento do mínimo caminho médio da rede pode ser calculado como

$$L = \frac{l_1 + kl_{2..k+1} + kl_{k+2..2k+1}}{2k(2k + 1)} \quad (12)$$

$$L = \frac{12k^2 - 4k}{2k(2k + 1)}. \quad (13)$$

A Figura 16(b) mostra a situação em que uma nova aresta é adicionada entre um vértice periférico e um vértice central. Neste caso, o comprimento médio dos caminhos mínimos partindo de cada vértice nesta rede pode ser expresso como

$$l_1 = \frac{3k}{2k}$$

$$l_{2,3} = \frac{5k-4}{2k}$$

$$l_{4..k+1} = \frac{5k-3}{2k}$$

$$l_{k+2..k+3} = \frac{7k-8}{2k}.$$

O que resulta em

$$L_{cp} = \frac{12k^2 - 8k + 2}{2k(2k+1)}. \quad (14)$$

Na Figura 16(c) uma nova conexão é criada entre dois vértices periféricos. O comprimento médio dos caminhos mínimos partindo de cada vértice da rede pode ser calculado como

$$l_1 = \frac{3k-1}{2k}$$

$$l_2 = \frac{5k-3}{2k}$$

$$l_{3..k+1} = \frac{5k-4}{2k}$$

$$l_{k+2} = \frac{5k-3}{2k}$$

Figura 17 – Comprimento do mínimo caminho médio em função do número de braços k para as três formas de redes mostradas na Figura 16.

Fonte: Elaborada pelo autor.

$$l_{k+3..2k+1} = \frac{7k - 5}{2k}$$

Resultando em

$$L_{pp} = \frac{12k^2 - 4k - 10}{2k(2k + 1)}. \quad (15)$$

Usando as Equações 14 e 15, nós temos que $L_{cp} < L_{pp}$ para $k > 3$. Portanto, conectar vértices periféricos com centrais nos leva a um comprimento de mínimo caminho médio menor ao obtido quando vértices periféricos são conectados entre si. A Figura 17 mostra a variação ASPL em função de k para as três situações mostradas na Figura 16.

4.3 Estratégias Baseadas em Centralidade de Intermediação

Sabe-se que a centralidade de intermediação permite a identificação de *hubs* - vértices com alta centralidade de intermediação - em redes.¹ *Hubs* representam vértices centrais que conectam a maioria dos vértices ao seu redor com *hubs* de outras regiões da rede, atu-

ando como pontes entre essas regiões. Portanto, o que causou um impacto considerável nos caminhos mínimos da rede pode ter sido o fato de que a estratégia B_{cc} atuou aprimorando as conexões de *hubs*, facilitando o fluxo de informações entre diferentes regiões.

Ao comparar as melhorias produzidas pelas estratégias de centralidade de intermediação e baseada no grau, que apresentaram desempenhos bastante parecidos, conclui-se que, na prática, poderia ser mais interessante favorecer a abordagem baseada no grau para esta atividade, devido ao seu custo computacional substancialmente menor, comparado ao que é implicado pelo cálculo da centralidade de intermediação.

4.4 Correlação Entre Valores Iniciais e Variações do ASPL

Para entender melhor a relação entre o valor inicial do ASPL encontrado em cada modelo e sua variação relativa, ou melhoria, resultante das estratégias de otimização, podemos analisar a Figura 18, que considera a estratégia A_{cp} como exemplo. De forma geral, a distribuição de pontos nos gráficos apresenta uma correlação relativamente alta entre os valores iniciais e a melhoria, indicando que um comprimento de mínimo caminho médio mais alto tem maior potencial de ser reduzido proporcionalmente, o que explica o fato de o modelo WS apresentar uma melhoria maior em comparação aos demais modelos, bem como o modelo BA apresentando as piores melhorias.

4.5 Maleabilidade

Por fim, para verificarmos se existe uma relação entre a maleabilidade das redes e a diminuição encontrada no ASPL após a aplicação das estratégias propostas, podemos analisar a correlação de Pearson ³⁷ entre estes valores, apresentada na Figura 19.

Visto que estratégias NTRLA operam de maneira similar à abordagem usada no cálculo da maleabilidade de redes complexas, ou seja fazendo pequenas alterações na topologia da rede buscando medir variações em suas propriedades, parece natural procurar uma relação entre a maleabilidade e um *potencial de otimização* de uma rede em ter o seu ASPL reduzido através da aplicação das estratégias que estudamos. Contudo, ao compa-

rar estas grandezas, apenas correlações baixas foram encontradas, com os maiores valores próximos a 30% de correlação, indicando que a maleabilidade não se mostrou ser um bom indicador para prever o *potencial de otimização* das redes consideradas neste trabalho.

De fato, uma rede com maleabilidade alta sugere que existem muitas variações diferentes de sua topologia que podem ser alcançadas a partir de pequenas alterações em sua topologia, tornando possível a escolha de uma configuração que apresente maior redução do ASPL. Porém, uma maleabilidade baixa, que indicaria a existência de uma variação da topologia de tal rede que pode ser alcançada com uma probabilidade muito maior em comparação às demais variações, não significa que esta variação “preferencial” apresente um ASPL maior que a rede original, visto que o ASPL pode ser facilmente reduzido a partir da adição de novas arestas na rede. Assim, vemos que a facilidade com que o ASPL pode variar está pouco relacionada com a maleabilidade da rede.

Figura 18 – Relação entre o ASPL inicial encontrado em cada um dos modelos considerados e a melhoria produzida pelas estratégias consideradas. Aqui, apenas uma estratégia representativa é mostrada.

Fonte: Elaborada pelo autor.

Figura 19 – Correlação de Pearson entre a maleabilidade das redes estudadas e os valores inicial e de melhoria encontrados para o ASPL destas redes, quando aplicadas as estratégias propostas. A cor vermelha mais intensa representa uma alta correlação positiva e a cor azul mais intensa representa uma alta correlação negativa. A primeira coluna foi calculada considerando o conjunto de todos os modelos estudados, enquanto a segunda coluna teve seus valores calculados apenas para os modelos BA e a terceira leva em consideração os modelos ER, WS ou Waxman, que apresentaram comportamentos semelhantes entre si.

Fonte: Elaborada pelo autor.

4.6 Aplicação das Estratégias em Redes Modulares

A Figura 20 apresenta os resultados encontrados pela aplicação das estratégias propostas às redes modulares no modelo Waxman e Fortunato com $\mu = 0, 1$. Podemos ver que, para o caso da configuração que usamos para representar redes modulares a partir do modelo Waxman, as estratégias que utilizam a dinâmica centro-centro, ou seja, que conectam vértices com os maiores valores de determinada medida, proporcionaram uma diminuição do ASPL superior às encontradas com as demais estratégias, principalmente para poucas adições de arestas, com destaque para a estratégia B_{cc} que se manteve a mais vantajosa inclusive para valores mais altos de N_a . Esse modelo produz um isolamento bastante expressivo das comunidades criadas na rede, como pode ser visto na Figura 21, o

Figura 20 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para cada topologia de redes modulares consideradas. Aqui, todas as redes são formadas por $N = 3000$ vértices e grau médio em torno de 6. As barras de erro são mostradas como linhas verticais com 40% do desvio padrão dos dados e o eixo vertical é normalizado em relação ao valor inicial do ASPL.

Fonte: Elaborada pelo autor.

que pode estar favorecendo as estratégias de dinâmica centro-centro, já que estas regiões isoladas da rede são mais dependentes de *hubs* nas composições de caminhos mínimos que ligam duas comunidades distintas.

Figura 21 – Visualização típica de uma rede Waxman adaptada para criação de comunidades. Os vértices são coloridos de acordo com suas respectivas acessibilidades, com cores mais claras para acessibilidades menores, para evidenciar as regiões de centro e de periferia de cada comunidade.

Fonte: Elaborada pelo autor.

Já o modelo Fortunato, com $\mu = 0,1$, mostrou um comportamento bastante similar aos modelos de redes não modulares considerados, porém com as estratégias C_{cp} e A_{cp} apresentando desempenhos um pouco menos eficientes quando comparadas às demais estratégias, fazendo com que as estratégias com dinâmica centro-centro B_{cc} e D_{cc} sejam as estratégias que possibilitaram maiores reduções do ASPL neste modelo, provavelmente pelas mesmas razões que este comportamento também foi verificado para o modelo Waxman adaptado, porém com menos intensidade no modelo Fortunato, já que as comunidades aqui são menos isoladas.

Figura 22 – Variação do comprimento do mínimo caminho médio em termos da taxa de adição de arestas N_a/N para a topologia de Fortunato com parâmetro de mistura $\mu = 0,03$, $\mu = 0,05$ e $\mu = 0,07$. As barras de erro são mostradas como linhas verticais com 40% do desvio padrão dos dados e o eixo vertical é normalizado em relação ao valor inicial do ASPL.

Fonte: Elaborada pelo autor.

Para verificar esta hipótese, consideramos também os valores $\mu = 0,03$, $\mu = 0,05$ e $\mu = 0,07$ para o parâmetro de mistura do modelo Fortunato, de forma a controlar o isolamento das comunidades. Verificamos que quanto menor o valor de μ , maior foi o desempenho das estratégias B_{cc} e D_{cc} em comparação com as demais estratégias, como

pode ser visto na Figura 22, confirmando o comportamento esperado, embora as demais estratégias com dinâmica centro-centro não tenham apresentado um ganho expressivo.

4.7 Um Estudo de Caso: Rede de Aeroportos

Para estudar como as aplicações das estratégias propostas podem afetar sistemas do mundo real, consideramos uma rede mundial de aeroportos ³⁸ que representa vários aeroportos, conectados quando há vôos diretos entre eles. A Figura 23 apresenta esta rede com os vértices posicionados de acordo com a localização geográfica dos aeroportos.

Figura 23 – Rede mundial de aeroportos com os vértices posicionados de acordo com a localização geográfica dos aeroportos.

Fonte: Elaborada pelo autor.

Os resultados da aplicação das estratégias a essa rede podem ser vistos na Figura 24. A estratégia com o melhor desempenho foi, novamente, a C_{cp} , com o segundo melhor desempenho obtido pela estratégia A_{cp} também apresentando destaque em comparação à terceira melhor estratégia, B_{cc} , visto que esta rede de aeroportos apresenta aspectos de

uma rede esparsa e livre de escala, para as quais as estratégias A_{cp} e C_{cp} apresentaram os melhores resultados na Seção 4.1.

A estratégia B_{cc} também resultou em uma diminuição significativa do ASPL, apresentando uma redução de 4,04% depois de 200 adições de arestas, sendo consideravelmente maior do que as reduções obtidas para a maioria dos modelos teóricos, mostrando um maior potencial de otimização do mínimo caminho médio para esta rede real.

Figura 24 – Variação do comprimento do mínimo caminho médio, em termos da taxa de adição de arestas N_a/N , obtido pela aplicação das várias estratégias consideradas na rede de aeroportos. O eixo vertical está normalizado em relação ao valor inicial do ASPL.

Fonte: Elaborada pelo autor.

5 Conclusões

O comprimento do mínimo caminho médio (ASPL) é uma importante propriedade de redes complexas, pois tende a influenciar muitas das dinâmicas que ocorrem nas redes. Por exemplo, a troca de informações em uma rede social tende a ser mais rápida quando o ASPL é pequeno.³⁹ Assim, é interessante não apenas identificar estratégias capazes de otimizar essa propriedade, mas também verificar a eficácia dessas estratégias entre diferentes modelos de redes.

Neste trabalho, consideramos 15 estratégias para otimização do ASPL, todas envolvendo a adição de arestas à rede, mas usando critérios diferentes para escolher pares de vértices. Os valores para grau, centralidade de intermediação, centralidade de proximidade e acessibilidade dos vértices foram levados em consideração para a inclusão das novas conexões. As estratégias foram aplicadas aos modelos Barabasi-Albert, Erdős-Rényi, Watts-Strogatz e Waxman e também à modelos de redes modulares e à uma rede de aeroportos.

Os resultados indicam que, entre os modelos considerados, três estratégias resultaram em um maior potencial em diminuir o ASPL de uma rede: (i) conectar pares de vértices que tenham, respectivamente, valores altos e baixos de centralidade de proximidade; (ii) ligar pares de vértices que tenham, respectivamente, valores altos e baixos de acessibilidade e; (iii) conectar pares de vértices com alta centralidade de intermediação. O modelo BA apresentou a maior resistência à redução do ASPL. Provavelmente porque as redes geradas pelo modelo BA já possuem ASPL pequeno.⁴⁰

Para redes modulares em que as comunidades que formam estas redes são bastante isoladas umas das outras, vimos que estratégias com dinâmicas centro-centro, como B_{cc} e D_{cc} podem ser mais eficientes por melhor explorarem a participação de *Hubs* nas redes. Já o ASPL da rede de aeroportos apresentou uma variação relativamente pequena, com a notável exceção das estratégias C_{cp} , A_{cp} e B_{cc} que proporcionaram uma grande redução no ASPL desta rede.

As estratégias baseadas em acessibilidade e em centralidade de proximidade resultaram nos melhores desempenhos, provavelmente porque os vértices com baixa acessibili-

dade/centralidade de proximidade tendem a ser associados à periferia, ou borda,^{29, 36} da rede. Esses vértices tendem a ser mal interconectados, tendo sua comunicação dependente de caminhos que passam por vértices mais centrais, caracterizados por alta acessibilidade/centralidade de proximidade. Quando novas conexões são adicionadas entre vértices de acessibilidade/centralidade de proximidade baixa e alta (estratégias A_{cp} e C_{cp}), os comprimentos dos caminhos mínimos entre os vértices da periferia diminuem rapidamente, em comparação à aplicação das demais estratégias. A correlação entre os valores iniciais e finais de ASPL também sugerem que um comprimento de mínimo caminho médio mais alto tem maior potencial de ser reduzido proporcionalmente.

Já a maleabilidade não se mostrou uma boa ferramenta para prever o potencial de otimização do ASPL de uma rede, seguindo as estratégias propostas, visto que as maiores correlações entre maleabilidade e a variação do ASPL ficaram em torno de 30%.

6 Trabalhos Futuros

Trabalhos futuros podem abordar a otimização do ASPL de modelos de redes espaciais, bem como de outras redes do mundo real, incluindo uma análise de custo de adicionar conexões de longo alcance à rede. Outra possibilidade seria investigar a influência do grau médio sobre as estratégias de otimização, já que essa medida é conhecida por ter um grande efeito na distribuição de caminhos mínimos, bem como investigar mais de perto a relação entre o caminho mínimo e regiões de baixa acessibilidade de diferentes redes, ou seja, os caminhos mínimos que ligam vértices de borda.

Referências

- 1 ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, APS, v. 74, n. 1, p. 47, 2002.
- 2 PERON, T. K. D.; COSTA, L. da F.; RODRIGUES, F. A. The structure and resilience of financial market networks. **Chaos**, American Institute of Physics, v. 22, n. 1, p. 013117, 2012.
- 3 PAGANI, G. A.; AIELLO, M. The power grid as a complex network: a survey. **Physica A**, Elsevier, v. 392, n. 11, p. 2688–2700, 2013.
- 4 CIAFRÈ, S. A.; GALARDI, S. micornas and rna-binding proteins: a complex network of interactions and reciprocal regulations in cancer. **RNA Biology**, Taylor & Francis, v. 10, n. 6, p. 934–942, 2013.
- 5 ONNELA, J.-P. *et al.* Structure and tie strengths in mobile communication networks. **Proceedings of the National Academy of Sciences**, National Academy Sciences, v. 104, n. 18, p. 7332–7336, 2007.
- 6 STRANO, E. *et al.* Urban street networks, a comparative analysis of ten european cities. **Environment and Planning B**, SAGE Publications, v. 40, n. 6, p. 1071–1086, 2013.
- 7 DOMINGUES, G. S. *et al.* Topological characterization of world cities. **Journal of Statistical Mechanics**, IOP Publishing, v. 2018, n. 8, p. 083212, 2018.
- 8 AMANCIO, D. R. *et al.* Using complex networks concepts to assess approaches for citations in scientific papers. **Scientometrics**, Springer, v. 91, n. 3, p. 827–842, 2012.
- 9 GOLOSOVSKY, M.; SOLOMON, S. Growing complex network of citations of scientific papers: modeling and measurements. **Physical Review E**, APS, v. 95, n. 1, p. 012324, 2017.
- 10 COSTA, L. *et al.* Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, Taylor & Francis, v. 60, n. 3, p. 329–412, 2011.
- 11 PRIM, R. C. Shortest connection networks and some generalizations. **The Bell System Technical Journal**, Nokia Bell Labs, v. 36, n. 6, p. 1389–1401, 1957.
- 12 ZHAO, J.; XU, K. Enhancing the robustness of scale-free networks. **Journal of Physics A**, IOP Publishing, v. 42, n. 19, p. 195003, 2009.
- 13 LI, L. *et al.* Enhancing the robustness and efficiency of scale-free network with limited link addition. **KSII Transactions on Internet & Information Systems**, v. 6, n. 5, 2012. DOI: 10.3837/tiis.2012.05.005.

- 14 HAYASHI, Y.; MATSUKUBO, J. Improvement of the robustness on geographical networks by adding shortcuts. **Physica A**, Elsevier, v. 380, p. 552–562, 2007. DOI: 10.1016/j.physa.2007.02.080.
- 15 PAIVA, W. R.; MARTINS, P. S.; ANGELIS, A. F. Unsupervised strategies to network topology reconfiguration optimization with limited link addition. *In*: BARBOSA, H. *et al.* (ed.). **Complex networks XI**, Berlin: Springer, 2020. p. 51–59.
- 16 WANG, H.; MIEGHEM, P. V. Algebraic connectivity optimization via link addition. *In*: CONFERENCE ON BIO-INSPIRED MODELS OF NETWORK, INFORMATION AND COMPUTING SYTEMS,3.,2008, Hyogo. **Proceedings** [...], Hyogo: Bionetics, 2008. p. 1–8.
- 17 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. **Nature**, Nature Publishing Group, v. 393, n. 6684, p. 440–442, 1998.
- 18 BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, American Association for the Advancement of Science, v. 286, n. 5439, p. 509–512, 1999.
- 19 NEWMAN, M. E. Detecting community structure in networks. **The European Physical Journal B**, Springer, v. 38, n. 2, p. 321–330, 2004.
- 20 KHEIRKHAHAZADEH, M.; LANCICHINETTI, A.; ROSVALL, M. Efficient community detection of network flows for varying markov times and bipartite networks. **Physical Review E**, APS, v. 93, n. 3, p. 032309, 2016.
- 21 NEWMAN, M. E. Finding community structure in networks using the eigenvectors of matrices. **Physical Review E**, APS, v. 74, n. 3, p. 036104, 2006.
- 22 PONS, P.; LATAPY, M. Computing communities in large networks using random walks. *In*: YOLUM, P.; GÜNGÖR, T.; GÜRGEN, F.; ÖZTURAN, C. (ed.). **Computer and information sciences**, ISCIS 2005. Berlin: Springer, 2005. p. 284–293. (Lectures notes in computer science, v. 3733).
- 23 NEWMAN, M. E.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E**, APS, v. 69, n. 2, p. 026113, 2004.
- 24 LANCICHINETTI, A.; FORTUNATO, S.; RADICCHI, F. Benchmark graphs for testing community detection algorithms. **Physical Review E**, APS, v. 78, n. 4, p. 046110, 2008.
- 25 GIRVAN, M.; NEWMAN, M. E. Community structure in social and biological networks. **Proceedings of the National Academy of Sciences**, National Academy Sciences, v. 99, n. 12, p. 7821–7826, 2002.
- 26 BRANDES, U. A faster algorithm for betweenness centrality*. **Journal of Mathematical Sociology**, Taylor & Francis, v. 25, n. 2, p. 163–177, 2001.

- 27 BAVELAS, A. Communication patterns in task-oriented groups. **The Journal of the Acoustical Society of America**, Acoustical Society of America, v. 22, n. 6, p. 725–730, 1950.
- 28 AHUJA, R. K. *et al.* Applications of network optimization. **Handbooks in Operations Research and Management Science**, Elsevier, v. 7, p. 1–83, 1995. DOI: 10.1016/S0927-0507(05)80118-5.
- 29 TRAVENÇOLO, B. A. N.; COSTA, L. da F. Accessibility in complex networks. **Physics Letters A**, Elsevier, v. 373, n. 1, p. 89–95, 2008.
- 30 SILVA, F. N.; COMIN, C. H.; COSTA, L. da F. Malleability of complex networks. **Journal of Statistical Mechanics**, IOP Publishing, v. 2019, n. 8, p. 083203, 2019.
- 31 WAXMAN, B. M. Routing of multipoint connections. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 6, n. 9, p. 1617–1622, 1988.
- 32 COSTA, L. da F. Voronoi and fractal complex networks and their characterization. **International Journal of Modern Physics C**, World Scientific, v. 15, n. 01, p. 175–183, 2004.
- 33 DALL, J.; CHRISTENSEN, M. Random geometric graphs. **Physical Review E**, APS, v. 66, n. 1, p. 016121, 2002.
- 34 BUHL, J. *et al.* Topological patterns in street networks of self-organized urban settlements. **The European Physical Journal B**, Springer, v. 49, n. 4, p. 513–522, 2006.
- 35 COOPER, C.; FRIEZE, A.; PRAŁAT, P. Some typical properties of the spatial preferred attachment model. **Internet Mathematics**, Taylor & Francis, v. 10, n. 1-2, p. 116–136, 2014.
- 36 TRAVENÇOLO, B. A.; VIANA, M. P.; COSTA, L. da F. Border detection in complex networks. **New Journal of Physics**, IOP Publishing, v. 11, n. 6, p. 063019, 2009.
- 37 FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A. d. Desvendando os mistérios do coeficiente de correlação de pearson (r). **Revista Política Hoje**, v. 18, n. 1, p. 115, 2009.
- 38 OPENFLIGHTS.ORG. Disponível em: <https://openflights.org/>. Acesso em: 20 jul. 2020.
- 39 KWAK, H. *et al.* What is twitter, a social network or a news media? *In*: CONFERENCE ON WORLD WIDE WEB, 19., 2010, Raleigh. **Proceedings [...]**, Raleigh: ACM, 2010. p. 591–600.
- 40 BARABÁSI, A.-L.; ALBERT, R.; JEONG, H. Scale-free characteristics of random networks: the topology of the world-wide web. **Physica A**, Elsevier, v. 281, n. 1-4, p. 69–77, 2000.