

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS**

Roberto Hiroshi Matos Furuta

**Maleabilidade em redes complexas: efeitos de
sucessivas remoções de arestas**

São Carlos

2023

Roberto Hiroshi Matos Furuta

Maleabilidade em redes complexas: efeitos de sucessivas remoções de arestas

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

2023

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.

Furuta, Roberto Hiroshi Matos
Maleabilidade em redes complexas: efeitos de
sucessivas remoções de arestas / Roberto Hiroshi Matos
Furuta; orientador Luciano da Fontoura Costa - versão
corrigida -- São Carlos, 2023.
48 p.

Dissertação (Mestrado - Programa de Pós-Graduação em
Física Aplicada Computacional) -- Instituto de Física de
São Carlos, Universidade de São Paulo, 2023.

1. Maleabilidade. 2. Grafos. 3. Redes complexas. 4.
Ciência das redes. I. Costa, Luciano da Fontoura, orient.
II. Título.

Este trabalho é dedicado a meus pais Silvia e Rubens.

AGRADECIMENTOS

Agradeço ao Prof. Dr. Luciano da Fontoura Costa, pela orientação atenciosa desse o tempo da graduação, com muito apoio e paciência. Agradeço ao Instituto de Física de São Carlos, pela oportunidade de realização do curso de mestrado, assim como o de bacharelado e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), pela concessão da bolsa de mestrado e pelo apoio financeiro para a realização desta pesquisa.

Digo um grande muito obrigado aos meus pais Silvia e Rubens, assim como meu irmão Takeo, pelo incentivo, apoio e carinho ao longo desse processo.

Agradeço imensamente aos meus colegas de grupo que, mesmo estando majoritariamente à distância, tentamos nos apoiar mutuamente. Um abraço especial aos meus colegas de mestrado Éverton, Guilherme e Renan; e aos veteranos que me ajudaram tanto: Eric, Messias, Burke, Paulo Ventura, Thomas e Henrique!

Agradeço aos meus colegas de apartamento Kauê, Alê e Nathan por estarem lá este tempo todo carregando uns aos outros nestes últimos anos incertos e desafiadores. Agradeço, claro, aos demais “Batatas” também: Estevão, Bretas e Gabriela.

Um enorme abraço aos meus colegas desde a época da graduação, que também compartilharam muitos dos apertos e alegrias desses últimos anos. Em especial: Zika, Luana, Maju, Lucas, Camilla, Gian, Polegs e Giane.

Agradeço, de coração, ao meu amado parceiro Mani, que não me deixou desistir, que me dá tanto carinho e sem o qual eu não consigo imaginar como estes anos teriam sido. Te amo, amor.

Compartilho, por último, minha gratidão a todas as organizações e indivíduos que se dedicam à democratização da ciência, reduzindo barreiras de acessos.

“O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001, PROEX 88887.482594/2020-00”.

“Don’t forget to be awesome.”
John and Hank Green

RESUMO

FURUTA, R. H. M. **Maleabilidade em redes complexas: efeitos de sucessivas remoções de arestas**. 2023. 48p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

A disseminação recente de tecnologias de monitoramento e simulação tem levado a um aumento na captação e geração de dados. Para modelar estes sistemas, é comum se utilizar redes complexas por poder incorporar conexões entre os elementos em sua estrutura. Este estudo aborda uma medida recém-introduzida, a maleabilidade de redes complexas, com foco nos modelos Erdős-Rényi, Watts-Strogatz e Barabási-Albert, estimando-a com base no coeficiente de agrupamento. A maleabilidade pode ser interpretada como a quantidade efetiva de estados isomórficos que uma rede pode assumir diante de modificações específicas. No entanto, calculá-la por meio da determinação de isomorfismos é custoso e não fornece um valor intuitivo quando o interesse é quantificar a variabilidade de uma medida particular. Para superar essas limitações, é possível estimá-la considerando-se uma ou mais medidas. Neste estudo, foca-se na maleabilidade estimada em relação ao coeficiente de agrupamento médio, medida selecionada por avaliar o grau de interconexão de um nó com seus vizinhos e boa eficiência computacional. A pesquisa concentra-se na remoção de arestas como uma modificação incremental nas redes, escolhidas com base em três critérios: aleatório, maximização e minimização da maleabilidade estimada. Nos experimentos, foi observado que, nas redes do tipo Watts-Strogatz, a maleabilidade mostra um aumento acentuado até atingir um pico, seguido por um declínio suave. O mesmo padrão é observado na curva de maximização da mesma, porém otimizado. Em contraste, a curva de minimização da maleabilidade apresenta um decréscimo suave e uniforme. Também nota-se que, em todos os modelos, as remoções que minimizam a maleabilidade resultaram no maior valor médio do coeficiente de agrupamento observado. Para investigar o impacto das remoções de arestas nas redes em simulações de dinâmicas, realizamos caminhadas aleatórias em cada uma das redes mencionadas. A ativação de nós em cada caminhada e a correlação entre essa ativação e os graus de entrada e saída ao longo das remoções foram determinadas. Essa correlação, denominada *steering coefficient*, foi comparada entre os três critérios de remoção de arestas mencionados anteriormente. Observa-se que a remoção aleatória de arestas tende a diminuir o *steering coefficient*, enquanto as remoções que minimizam a maleabilidade preservam essa correlação. As remoções que maximizam a maleabilidade mostram um perfil intermediário. Um achado notável é a alta robustez do *steering coefficient* em relação às remoções que minimizam a maleabilidade nas redes do tipo Watts-Strogatz. Acreditamos que, embora focado em uma combinação específica de modelos e critérios, este trabalho apresenta um avanço na compreensão das redes analisadas por meio dessa nova medida.

Palavras-chave: Maleabilidade. Grafos. Redes complexas. Ciência das redes.

ABSTRACT

FURUTA, R. H. M. **Malleability in complex networks: effects of successive removals of edges**. 2023. 48p. Dissertation (Master in Science) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

The recent dissemination of surveillance and simulation technologies has led to an increase in data acquisition and generation. To model these systems, it is common to use complex networks as they can incorporate connections between elements into their structure. This study addresses a recently introduced measurement, the malleability of complex networks, focusing on the Erdős-Rényi, Watts-Strogatz, and Barabási-Albert models, estimating it based on the clustering coefficient. Malleability can be interpreted as the effective quantity of isomorphic states a network can unfold to under specific modifications. However, calculating the malleability by isomorphic state determination is costly and it does not provide an intuitive value when the interest is in quantifying the variability of a particular measurement. To overcome these limitations, it is possible to estimate it by calculating it with respect to one or more measurements. In this study, the focus is on the malleability estimated in respect to the average clustering coefficient, a measure chosen to assess the degree of interconnection of a node with its neighbors and for offering good computational efficiency. This research focuses on edge removal as an incremental modification in the networks, chosen based on three criteria: random, maximization, and minimization of estimated Malleability. In the experiments, it was observed that in Watts-Strogatz networks, malleability shows a sharp increase until reaching a peak, followed by a gentle decline. The same pattern was observed in the maximization curve of Malleability, though optimized. In contrast, the minimization curve of malleability presents a smooth and uniform decrease. It is also noticeable that, in all models, the removals that minimize malleability resulted in the highest average clustering coefficient observed. To investigate the impact of edge removals on networks in dynamic simulations, random walks were performed on each of the mentioned networks. The activation of nodes in each walk and the correlation between its activation and the in-degree and out-degree along the removals were determined. This correlation, referred to as the "steering coefficient," was compared among the three aforementioned edge removal criteria. It is observed that random edge removal tends to decrease the steering coefficient, while removals that minimize malleability preserve this correlation. Removals that maximize malleability show an intermediate profile. A notable finding is the high robustness of the steering coefficient regarding removals that minimize Malleability in Watts-Strogatz-type networks. We believe that, although focused on a specific combination of models and criteria, this work represents an advancement in understanding the analyzed networks through this new measure.

Keywords: Malleability. Graphs. Complex networks. Network science.

LISTA DE FIGURAS

<p>Figura 1 – Cada sub-figura representa um modelo de rede: modelo Erdős-Rényi (2.2.1); modelo de Watts e Strogatz (2.2.2); e modelo de Barabási e Albert (2.2.3) respectivamente. No topo, temos uma representação da rede de 12 nós e grau médio próximo de 6. Em seguida temos a média da distribuição de grau de 100 redes (cada uma de $N = 100$, $\bar{k} \approx 6$) e seus respectivos caminho mínimo médio $\bar{\ell}$ e coeficiente de agrupamento médio \bar{C}.</p>	26
<p>Figura 2 – Maleabilidade topológica devido à remoção de uma única aresta. Para cada exemplo, os diagramas de rede são acompanhados pela tabela de probabilidades de transição de S_0 a S_i e pelos valores de entropia \mathcal{E}_{S_0} e maleabilidade A_{S_0} associados a estas transições. (A) Representa uma rede onde todas as possíveis remoções de bordas levam à mesma topologia, resultando, portanto, em um valor de maleabilidade (número efetivo de estados) de 1. (B) Representa uma rede onde diferentes remoções de bordas podem levar a três estruturas diferentes, resultando em um valor mais alto de 2,8284 para a maleabilidade.</p>	29
<p>Figura 3 – Estimativa da maleabilidade topológica, devido à remoção de uma única aresta, no mesmo grafo da Figura 2-B. O espaço de medida tridimensional inclui as seguintes medidas: grau médio $\langle K \rangle$, coeficiente de agrupamento médio $\langle C \rangle$ e <i>degree assortativity coefficient</i> (r). Para cada, a probabilidade de transição $P_{S_0 S_i}^{M_i}$, a entropia associada $\tilde{\mathcal{E}}_{S_0}^{M_i}$ e a maleabilidade estimada $\tilde{A}_{S_0}^{M_i}$ foram obtidas. Finalmente, a maleabilidade geral estimada \tilde{A}_{S_0} foi determinada como sendo 2,8284, o maior valor encontrado, o qual corresponde ao valor de A_{S_0} obtido na Figura 2.17-B.</p>	30
<p>Figura 4 – Exemplo ilustrativa de uma caminhada aleatória, representada em cor laranja, sobre os nós de um grafo, representado em verde.</p>	32
<p>Figura 5 – Ilustrações de gráficos de dispersão (<i>scatterplots</i>) de duas variáveis x e y com diferentes valores correlação de Pearson r. Na primeira linha, em verde, vemos exemplos de correlações positivas; e de valores negativos representados em vermelho na segunda linha. A cor dos pontos de dispersão são pintados com menor saturação (mais acinzentados) à medida que r se aproxima de 0 (quando suas variáveis apresentam baixa correlação).</p>	33
<p>Figura 6 – Modelos de rede. (A) A rede é gerada como um grafo não direcionado. (B) A rede gerada é reinterpretada como um grafo simétrico dirigido, do qual uma conexão dirigida pode ser removida, sem quebrar sua correspondente conexão inversa. (C) A remoção consecutiva das arestas é realizada. As conexões pontilhadas representam as bordas removidas, onde a seta laranja é a mais recentemente removida.</p>	36

Figura 7 – Maleabilidade aproximada em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max \tilde{A}** , **Min \tilde{A}** e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20. 37

Figura 8 – Coeficiente de agrupamento médio em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max \tilde{A}** , **Min \tilde{A}** e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20. 38

Figura 9 – *Steering coefficient* respectivo aos graus de entrada em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max \tilde{A}** , **Min \tilde{A}** e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20. 40

Figura 10 – *Steering coefficient* respectivo aos graus de saída em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max \tilde{A}** , **Min \tilde{A}** e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20. 41

SUMÁRIO

1	INTRODUÇÃO	19
2	FUNDAMENTOS	23
2.1	Medidas de Rede	23
2.1.1	Grau	23
2.1.1.1	Grau de saída	23
2.1.1.2	Grau de entrada	23
2.1.1.3	Grau médio	24
2.1.2	Coeficiente de agrupamento	24
2.1.2.1	Coeficiente de agrupamento de redes dirigidas	24
2.2	Modelos de Rede	25
2.2.1	Grafo não dirigido de Erdős-Rényi	25
2.2.2	Grafo não dirigido de Watts-Strogatz	26
2.2.3	Grafo não dirigido de Barabási-Albert	27
2.3	Maleabilidade de redes complexas	27
2.3.1	Maleabilidade no espaço de transição original	28
2.3.2	Maleabilidade estimada no espaço de medidas	28
2.4	Caminhada aleatória	31
2.5	<i>Steering coefficient</i>	32
2.5.1	Coeficiente de correlação de Pearson	32
3	DESENVOLVIMENTO, RESULTADOS E DISCUSSÃO	35
3.1	Estudo 1 - Maleabilidade de redes modelos ao longo de sucessivas modificações	35
3.2	Estudo 2 - Efeito do critério de remoção de aresta condicionado por maleabilidade no coeficiente de direcionamento de redes modelos	39
4	CONCLUSÃO	43
	REFERÊNCIAS	47

1 INTRODUÇÃO

A cada ano que se passa, presenciamos uma tendência de aumento da captação e geração de dados, por meio da ampla disseminação de tecnologias de monitoramento e simulação.¹ Para uma modelagem mais completa desses dados, é de grande importância que sejam representados como sistemas do mundo real, em termos de grafos e redes complexas, com o objetivo de se entender melhor as propriedades topológicas desses sistemas, permitindo assim a obtenção de modelos efetivos que possam ser utilizados para compreensão e previsão do comportamento do sistema.² Assim, torna-se evidente a importância de se desenvolver e aplicar conceitos e métodos que permitam a caracterização abrangente das propriedades estruturais dos sistemas a serem estudados,³ assim como possíveis implicações em dinâmicas desempenhadas nestes sistemas.⁴

Enquanto existem várias medidas topológicas de redes complexas descritas na literatura, poucas focam na quantificação de como uma mesma rede se modifica incrementalmente como consequência de adição/remoção de nós e/ou arestas.

Recentemente, foi proposto o conceito de maleabilidade de um determinado grafo ou rede complexa.⁵ Esta medida quantifica a quantidade de estados efetivos isomórficos que em que uma rede pode se desdobrar quando é sujeita a uma modificação específica. Embora a quantificação da maleabilidade possa ser abordada através da identificação de isomorfismo entre grafos,^{6,7} é também possível considerar as modificações da rede em relação a um determinado conjunto de medidas topológicas, sumarizando as propriedades das redes a serem estudadas ao longo das respectivas modificações, determinando assim uma maleabilidade estimada.

Cada uma dessas duas abordagens possuem suas respectivas vantagens e limitações. No caso do uso de medidas topológicas para se representar as redes, temos um mapa não necessariamente biunívoco, ou seja, que não necessariamente discerne entre todos os isomorfismos. Entretanto, o estudo da maleabilidade em termos de um conjunto específico de medidas possui uma vantagem particularmente interessante, em que: a análise se dá efetivamente sobre as medidas de interesse em cada problema.

O presente trabalho tem como principal objetivo o estudo da maleabilidade de três modelos padrão de redes complexas (nominalmente Erdős-Rényi,⁸ Watts-Strogatz⁹ e Barabási-Albert¹⁰) sob a perspectiva de uma medida topológica específica: o coeficiente de agrupamento.^{3,9,11}

Esta medida foi especificamente escolhida na análise a ser reportada devido a ser uma quantificação de uma propriedade importante das redes complexas, indicando como os vizinhos de cada nós se interconectam. Um outro aspecto justificando esta escolha é que esta medida possui custo computacional relativamente baixo, viabilizando desta forma a estimação efetiva da maleabilidade, que tem uma demanda computacional bastante elevada devido ao seu aspecto

combinatório. Além disso, o coeficiente de agrupamento é uma medida naturalmente quantizada, permitindo assim uma estimaco mais estvel e acurada no clculo das entropias necessrios para a anlise de maleabilidade.

Quanto s alteraces a serem consideradas para a obteno incremental de novas verses das redes, o foco foi colocado na operao especfica de remoo de arestas. Isso se justifica pois, em grafos esparsos como nos modelos adotados, esta operao pode ser implementada com um custo computacional menor do que a incluso de arestas ou manipulao de ns (em uma rede densa, em contrapartida, a adio de conexes seria menos custoso que a anlise de remoes). Outra vantagem importante  que, desta forma, pudemos fazer um paralelo com resultados preliminares respectivos descritos no trabalho de Silva e colaboradores.⁵

Assim, estudamos o valor da maleabilidade estimada em relao ao coeficiente de agrupamento, nas redes modelos, ao longo de sucessivas remoes de arestas. As arestas a serem removidas foram selecionadas utilizando trs critrios diferentes a fim de se observar seu efeito nos valores de maleabilidade em si. Os trs critrios so: remoes aleatrias, remoes que maximizam a maleabilidade a cada remoo e remoes que minimizam a maleabilidade a cada remoo (ambas de forma gulosa).¹²

Nesse estudo, observamos, conforme pressupomos que as curvas de maximizao gulosa de maleabilidade apresentaram a maior maleabilidade possvel, que as curvas de minimizao gulosa levam  menor maleabilidade observada e que representa remoes aleatrias levaram a um comportamento intermedirio para todos os tipos de modelos. Todavia,  notvel que as redes do tipo **WS** apresentam um padro interessante, em que a curva de remoes aleatrias tem seu valor acrescido de forma acentuada durante as suas primeiras remoes at atingir um mximo, a partir do qual, segue decrescendo de forma aparentemente assinttica. A curva que maximiza para a maleabilidade segue um padro bastante similar, porm alcana tal assntota de forma otimizada. J a curva que a minimiza, apresenta um perfil completamente distinto, apresentando um decrescimento suave e quase uniforme.

Dado o nosso interesse em estudar o efeito de tais remoes em uma medida de dinmica possvel nessas redes, simulamos caminhadas aleatrias sobre cada uma das redes apresentadas anteriormente (a cada passo de remoo de arestas).¹³ A ativao dos ns em cada caminhada foi registrada e sua correlao com os graus de entrada e sada foram monitoradas ao longo das sucessivas remoes. Esta correlao, em particular, chamamos de *steering coefficient* em concordncia com os artigos onde foi introduzido.^{14,15}

Ao compararmos o valor de *steering coefficient* obtidos a partir de remoes aleatrias e remoes que visam minimizar ou maximizar a maleabilidade da rede, observamos que a remoo aleatria de arestas tende a reduzir o *steering coefficient*. Em contrapartida, as remoes que minimizam a maleabilidade da rede preservam o *steering coefficient*. J as que maximizam a maleabilidade apresentam um perfil intermedirio, mas ainda mais semelhante s remoes aleatrias.

Dito isso, o resultado mais notável é a robustez do *steering coefficient* em relação às remoções que minimizam a maleabilidade nas redes do tipo Watts-Strogatz, com uma correlação alta, enquanto outros critérios de remoção levaram a uma perda abrupta de correlação, mesmo com a mesma quantidade de arestas removidas.

Este trabalho desenvolve-se dentro do seguinte plano de apresentação. Após a introdução, no capítulo 2, apresentamos os fundamentos necessários para o acompanhamento desta dissertação. Nela definimos medidas de redes relevantes para nossos estudos, modelos de redes notáveis, a medida central para este estudo: a maleabilidade entre outros conceitos. No capítulo 3 descrevemos cada um dos estudos que realizamos e apresentamos seus resultados seguidos das discussões dos respectivos experimentos. No capítulo 4, enfim, fazemos uma conclusão recapitulando nosso trabalho e resultados, seguida de sugestões de possíveis desenvolvimentos futuros.

2 FUNDAMENTOS

Neste capítulo, fundamentos necessários para compreensão dos estudos conduzidos nesta dissertação são apresentados. O material aqui abordado corresponde a conceitos e métodos já existentes, descritos em respectiva literatura.

2.1 Medidas de Rede

Redes complexas possuem características topológicas que dependem de disposição de suas conexões; e estas, por sua vez, influenciam processos dinâmicos que nela se desenvolvem. Com o fim de caracterizar estas redes, foram desenvolvidas *medidas topológicas* que são, em essência, cálculos realizados utilizando como argumento a conectividade da rede, com o objetivo de quantificar suas características topológicas.^{3,11} Aqui, introduzimos algumas destas medidas que são essenciais para nossos estudos.

2.1.1 Grau

O grau k_i de um nó n_i da rede de N nós consiste na quantidade de arestas que este nó estabelece com seus vizinhos. Em uma rede dirigida, o grau pode ser discernido entre *grau de saída* (descrito em 2.1.1.1) e *grau de entrada* (descrito em 2.1.1.2). Numa rede não dirigida (ou simétrica) esses dois valores são iguais.³

2.1.1.1 Grau de saída

O grau de saída k_i^{out} de um nó n_i é a quantidade de arestas direcionadas que saem de nó n_i . Em um grafo representado por sua matriz de adjacência, convencionamos aqui que seu valor é obtido pela soma dos elementos da linha i .

$$k_i^{\text{out}} = \sum_j a_{ij} \quad (2.1)$$

2.1.1.2 Grau de entrada

O grau de entrada k_i^{in} de um nó n_i é a quantidade de arestas direcionadas que chegam ao nó n_i . Em um grafo representado por sua matriz de adjacência, convencionamos aqui que seu valor é obtido pela soma dos elementos da coluna i .

$$k_i^{\text{in}} = \sum_j a_{ji} \quad (2.2)$$

2.1.1.3 Grau médio

Frequentemente, é de nosso interesse sabermos o valor médio do grau de todos os nós da rede, para isso utilizamos a notação \bar{k}_i para quando temos uma rede não dirigida, e para redes dirigidas temos \bar{k}_i^{out} e \bar{k}_i^{in} .

$$\bar{k}^{\text{out}} = \sum_i \frac{k_i^{\text{out}}}{N}; \bar{k}^{\text{in}} = \sum_i \frac{k_i^{\text{in}}}{N}; \text{ e} \quad (2.3)$$

$$\bar{k} = \sum_i \frac{k_i}{N} = \bar{k}^{\text{in}} = \bar{k}^{\text{out}} \text{ se o grafo não for dirigido.} \quad (2.4)$$

2.1.2 Coeficiente de agrupamento

Consideremos um grafo não dirigido G . Suponha que um vértice i tenha k_i vizinhos; então, podem existir no máximo $N_3(i)$ conexões entre eles, que fechariam triângulos com i (isso ocorre quando todos os vizinhos de i estão conectados a todos os outros vizinhos de i). Seja $N_{\Delta}(i)$ a quantidade desses triângulos possíveis que realmente existem. O coeficiente de agrupamento C_i é definido como a fração*:⁹

$$C_i = \frac{N_{\Delta}(i)}{N_3(i)}, \text{ sendo} \quad (2.5)$$

$$N_{\Delta}(i) = \sum_j \sum_{h>j} a_{ij} a_{ih} a_{jh} = \frac{k_i(k_i - 1)}{2}, \text{ e} \quad (2.6)$$

$$N_3(i) = \sum_j \sum_{h>j} a_{ij} a_{ih}. \quad (2.7)$$

O coeficiente de agrupamento médio da rede \bar{C} é definido como a média de C_i em todos os seus vértices:

$$\bar{C} = \frac{1}{N} \sum_i C_i. \quad (2.8)$$

Para redes de amizade, seu significado é intuitivo: C_i reflete o grau em que os amigos de i também são amigos uns dos outros e, portanto, \bar{C} mede o grau de formação de cliques em um círculo de amizade típico.

2.1.2.1 Coeficiente de agrupamento de redes dirigidas

Adotamos a definição de coeficiente de agrupamento para redes dirigidas sem pesos, aqui representados como C_i , como descrito por Fagiolo,¹⁶ independente dos padrões de triângulos dirigidos. Entre todas as definições propostas neste trabalho, decidimos usar aquela que calcula todas as configurações possíveis de triângulos de arestas possíveis, uma vez que não tínhamos razão para priorizar nenhum padrão em particular.

* Neste trabalho, adotamos o índice h , além dos índices convencionais i, j , em vez do mais comum índice k , para evitar que este seja confundido com o símbolo de grau k definido em 2.1.1.

Intuitivamente, C_i é a fração do quantidade de triângulos que o nó i forma com seus vizinhos t_i , sobre o número máximo de triângulos que i poderia fazer parte, se todos os seus vizinhos estivessem mutuamente ligados T_i (como representado na primeira igualdade na equação 2.13, que será elaborada abaixo).

Para calcular os C_i da matriz de adjacência A , primeiramente, precisamos determinar quatro medidas de grau da rede:

$$d_i^{\text{in}} = \sum_{j \neq i} a_{ji} = (A^T)_i \mathbf{1}, \quad (2.9)$$

$$d_i^{\text{out}} = \sum_{j \neq i} a_{ij} = (A)_i \mathbf{1}, \quad (2.10)$$

$$d_i^{\text{tot}} = d_i^{\text{in}} + d_i^{\text{out}} = (A^T + A)_i \mathbf{1} \text{ e} \quad (2.11)$$

$$d_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji} = A_{ii}^2; \quad (2.12)$$

onde d_i^{in} é a quantidade de conexões de entrada, d_i^{out} é a quantidade de conexões de saída, d_i^{tot} é a soma destas e d_i^{\leftrightarrow} é o número de conexões mútuas i com seus vizinhos. Nesta construção, nenhuma auto-interação (conexão do nó i com o nó i) está presente. Então, estes graus são usados para calcular o coeficiente de agrupamento generalizado C_i :

$$\begin{aligned} C_i(A) &= \frac{t_i}{T_i} = \frac{\sum_j \sum_h (a_{ij} + a_{ji}) (a_{ih} + a_{hi}) (a_{jh} + a_{hj})}{2 [d_i^{\text{tot}} (d_i^{\text{tot}} - 1) - 2d_i^{\leftrightarrow}]} \\ &= \frac{(A + A^T)_{ii}^3}{2 [d_i^{\text{tot}} (d_i^{\text{tot}} - 1) - 2d_i^{\leftrightarrow}]} \end{aligned} \quad (2.13)$$

Tal como no caso da rede não dirigida, O coeficiente de agrupamento médio da rede \bar{C} é definido como como a média de C_i em todos os seus vértices:

$$\bar{C} = \frac{1}{N} \sum_i C_i. \quad (2.14)$$

2.2 Modelos de Rede

Como objetos de nossos estudos, adotamos três modelos de rede amplamente utilizados na literatura, os quais possuem diferentes propriedades de conectividade, menor caminho e agrupamento. São eles: o modelo de rede estatisticamente uniforme introduzido por Erdős e Rényi,⁸ o modelo de pequeno mundo elaborado por Watts e Strogatz⁹ e o modelo proposto por Barabási e Albert para representar redes livre de escala.¹⁰

A Figura 1 ilustra algumas das principais características comparativas dessas redes.

2.2.1 Grafo não dirigido de Erdős-Rényi

Este modelo, abreviado como *modelo ER*, é construído ao definir-se uma quantidade inicial de vértices N , os quais possuem uma probabilidade p constante de se conectarem. Uma

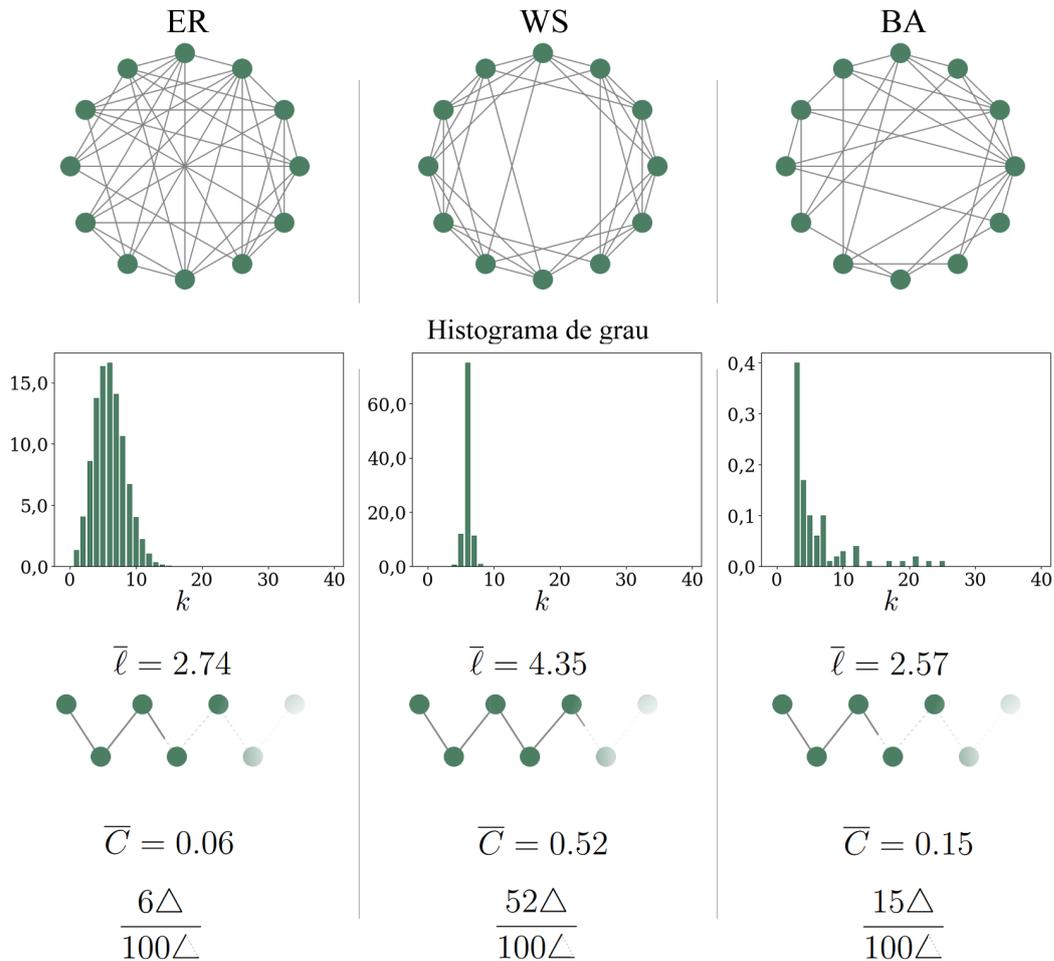


Figura 1 – Cada sub-figura representa um modelo de rede: modelo Erdős-Rényi (2.2.1); modelo de Watts e Strogatz (2.2.2); e modelo de Barabási e Albert (2.2.3) respectivamente. No topo, temos uma representação da rede de 12 nós e grau médio próximo de 6. Em seguida temos a média da distribuição de grau de 100 redes (cada uma de $N = 100$, $\bar{k} \approx 6$) e seus respectivos caminho mínimo médio $\bar{\ell}$ e coeficiente de agrupamento médio \bar{C} .

Fonte: Elaborada pelo autor.

característica deste modelo é seu caminho mínimo médio ser curto e sua distribuição de grau ser binomial, centrada em $(N - 1)p$ (que com N grande, tende a Np).⁸

Devido aos graus de seus nós serem similares aos valores de grau de um grafo regular (no qual todos os nós têm a mesma quantidade de conexões), o modelo ER é considerado como um dos modelos mais simples (de menor complexidade).

2.2.2 Grafo não dirigido de Watts-Strogatz

O *modelo WS* foi proposto como um modelo capaz de representar de maneira mais satisfatória alguns tipos de redes organicamente presentes no mundo real, especialmente redes sociais. Historicamente, os modelos estudados até então eram ou completamente regulares ou

completamente aleatórios. Por meio de um algoritmo que usa como base uma rede regular e a modifica introduzindo irregularidades, foi possível se obter uma rede com características mistas entre redes regulares e redes aleatórias, apresentando propriedades similares às daquelas de diversas redes reais, como curto caminho mínimo médio e alto coeficiente de agrupamento médio.⁹

O algoritmo de construção da versão mais comum das redes WS consiste em se criar uma rede regular do tipo anel com N vértices, cada um com k arestas e transpor cada aresta para outro sítio aleatório da rede com a probabilidade p .

Nesse tipo de rede, o que se observa é que ao introduzir mesmo que poucas conexões transpostas ($0 < p \ll 1$), o caminho mínimo médio da rede se reduz drasticamente, pois as conexões transpostas servem como atalhos que interligam nós que outrora estariam distantes. Este processo ocorre, sem que haja quase efeito algum no valor do agrupamento médio da rede, sendo assim, uma rede com características de pequeno-mundo. Quanto maior for p , as características da rede se aproximam da rede ER, porém o coeficiente de agrupamento médio possui robustez em relação às reconexões, tendo seu valor significativamente reduzido somente em quando ($0 \ll p < 1$).

2.2.3 Grafo não dirigido de Barabási-Albert

Este modelo, abreviado como *modelo BA*, foi introduzido em¹⁰ com o objetivo de reproduzir uma propriedade de muitas redes complexas reais: a distribuição de grau de seus vértices obedece uma lei de potência (propriedade também descrita como *livre de escala*), assim favorecendo o aparecimento de uma quantidade relativamente grande de nós com grau elevado (os assim chamados *hubs*). Esta propriedade foi reproduzida ao incorporar duas características de redes reais no processo de construção das redes modelos: crescimento e ligação preferencial.

Na construção de uma rede deste modelo, cria-se uma rede original com m_0 nós conectados entre si. A cada passo temporal, insere-se um novo nó que conectar-se-á a $m < m_0$ vértices existentes, dando preferência para aqueles com mais conexões. Isso é obtido ao definir a probabilidade do novo vértice conectar-se ao nó i como:

$$P_i = \frac{k_i}{\sum_j k_j}, \quad (2.15)$$

onde k_i é o grau do nó i e j são todos os vértices pré-existentes da rede, incluindo i . Na prática, intuitivamente, podemos pensar que o novo nó formará conexões com m nós, os quais serão sorteados entre os nós existentes, sendo que a chance de cada nó ser escolhido é diretamente proporcional ao seu grau.

2.3 Maleabilidade de redes complexas

Adotamos a definição de maleabilidade proposta em.⁵ Esta medida quantifica a diversidade topológica das redes possíveis quando a rede original é sujeita a uma regra de modificação

(como a remoção de arestas). Tanto em seu trabalho como nos exemplos desta seção são estudados apenas grafos não direcionados, mas os princípios são aplicáveis também aos grafos direcionados, como aqueles descritos na seção anterior.

2.3.1 Maleabilidade no espaço de transição original

Vamos definir a topologia original de uma rede como o estado S_0 . Quando esta rede estiver sujeita à regra de modificação escolhida (remoção de arestas, neste exemplo), ela se desdobrará em uma coleção de todas as redes resultantes. Algumas delas podem ser isomórficas umas às outras, portanto somos capazes de construir um conjunto menor dos estados topológicos únicos $S = \{S_1, S_2, \dots, S_C\}$. Na figura 2, podemos ver dois exemplos contrastantes de redes, onde (A) representa uma rede que se desdobra em poucos estados (neste caso, um único) e (B) uma rede que se desdobra em muitos (três) estados. Para cada S_i , podemos calcular P_{S_0, S_i} , a probabilidade de transição de S_0 para S_i , (indicada na tabela abaixo de cada exemplo ilustrado) e usá-las para determinar a entropia associada a estas transições,

$$\mathcal{E}_{S_0} = - \sum_{S_i \in \mathcal{S}} P_{S_0, S_i} \log(P_{S_0, S_i}). \quad (2.16)$$

Esta entropia indica a heterogeneidade da distribuição de probabilidades de transição. Finalmente, a maleabilidade é definida como o exponencial da entropia

$$A_{S_0} = e^{\mathcal{E}_{S_0}}. \quad (2.17)$$

Esta quantidade (A_{S_0}), pode ser considerada como um número efetivo de estados, devido ao fato de que, quando todos os estados são alcançados com probabilidades iguais, a maleabilidade é máxima e igual ao número de estados ($P_{S_0, S_i} = P_{S_0, S_j} \forall i, j \implies A_{S_0} = C$).

Para cada exemplo na Figura 2, podemos ver seus respectivos valores de entropia e maleabilidade. A rede (A) tem a maleabilidade (número efetivo de estados) de 1. Por outro lado, a rede (B) tem a maleabilidade de aproximadamente 2,8, que é próxima, mas inferior ao número de estados (3 estados), uma vez que as probabilidades de transição não são todas iguais.

2.3.2 Maleabilidade estimada no espaço de medidas

De preferência, gostaríamos de calcular a maleabilidade, conforme definido na seção anterior. Entretanto, a identificação se dois ou mais gráficos são isomórficos ou não, uma etapa necessária para a construção do conjunto de estados, é uma tarefa computacional particularmente cara. Portanto, um método de estimativa da maleabilidade que requer menos poder computacional precisa ser considerado.⁵

A maleabilidade estimada \tilde{A}_{S_0} é calculada ao aplicar as equações 2.16 e 2.17 às probabilidades de transição em um *espaço de medidas*.

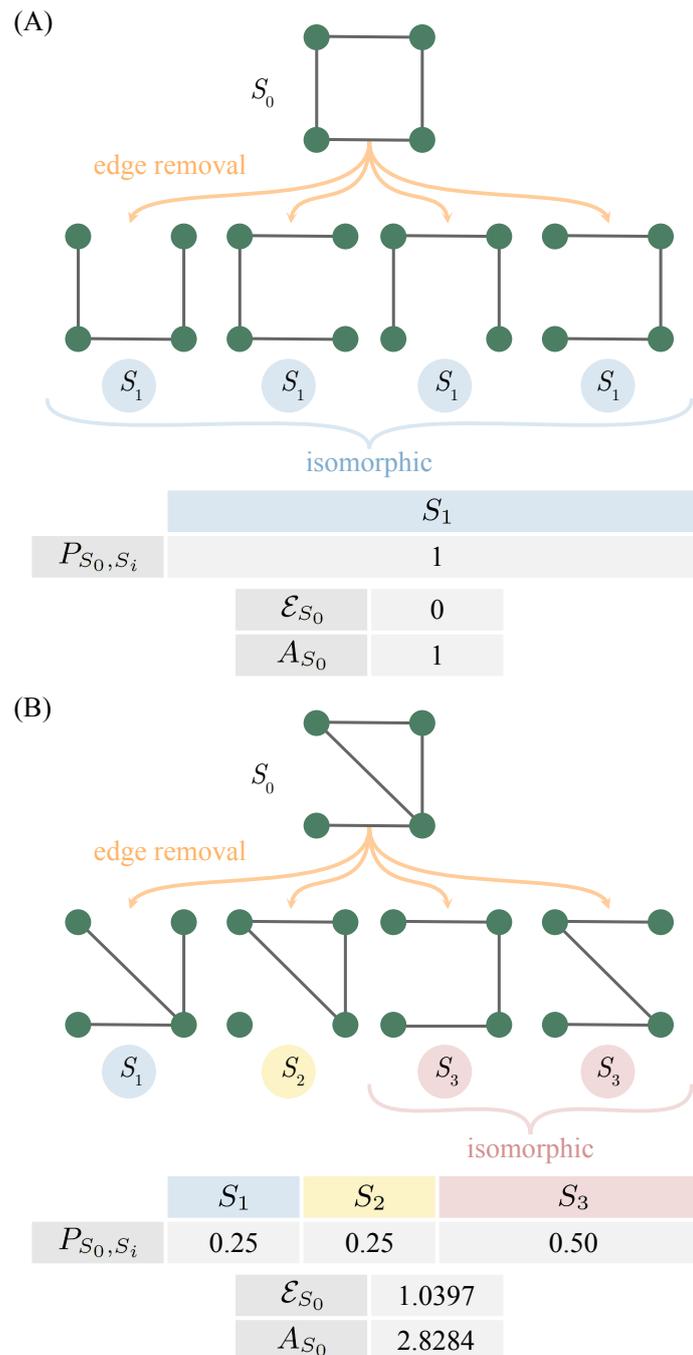


Figura 2 – Maleabilidade topológica devido à remoção de uma única aresta. Para cada exemplo, os diagramas de rede são acompanhados pela tabela de probabilidades de transição de S_0 a S_i e pelos valores de entropia \mathcal{E}_{S_0} e maleabilidade A_{S_0} associados a estas transições. (A) Representa uma rede onde todas as possíveis remoções de bordas levam à mesma topologia, resultando, portanto, em um valor de maleabilidade (número efetivo de estados) de 1. (B) Representa uma rede onde diferentes remoções de bordas podem levar a três estruturas diferentes, resultando em um valor mais alto de 2,8284 para a maleabilidade.

Fonte: Elaborada pelo autor.

O espaço de medidas \mathcal{M} é definido como $\mathcal{M} = M_1 \times M_2 \times \dots \times M_l$, onde M_k é uma propriedade topológica, como o grau médio ou o coeficiente de agrupamento médio, e, neste espaço, um estado de rede pode ser representado como um ponto com as coordenadas (m_1, m_2, \dots, m_l) . Assim, $\tilde{S}_i = (m_1, m_2, \dots, m_l)$. Para determinarmos a maleabilidade estimada, calcularemos a probabilidade de transição associada a cada medida em \mathcal{M} . Isto significa que, para cada medida, obteremos um valor de maleabilidade $\tilde{A}_{S_0}^{M_k}$ que será específico de uma medida e de uma regra de modificação.. Depois de calcular todos os valores de $\tilde{A}_{S_0}^{M_k}$, a maleabilidade estimada de \tilde{A}_{S_0} será considerada como aquela com o valor mais alto.

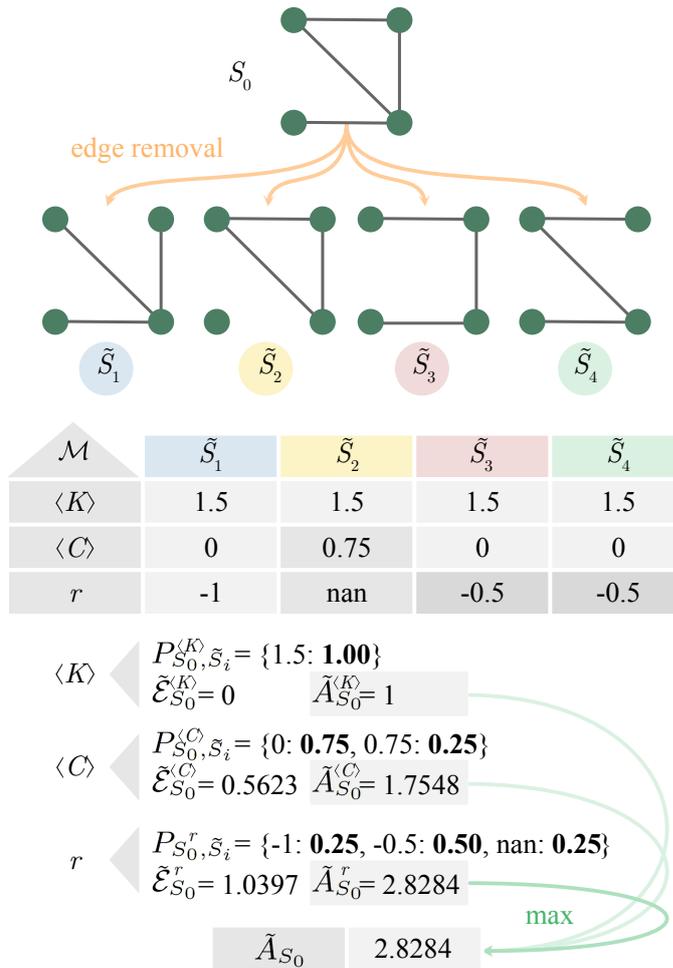


Figura 3 – Estimativa da maleabilidade topológica, devido à remoção de uma única aresta, no mesmo grafo da Figura 2-B. O espaço de medida tridimensional inclui as seguintes medidas: grau médio $\langle K \rangle$, coeficiente de agrupamento médio $\langle C \rangle$ ⁹ e *degree assortativity coefficient* (r).¹¹ Para cada, a probabilidade de transição $P_{S_0, \tilde{S}_i}^{M_k}$, a entropia associada $\tilde{\mathcal{E}}_{S_0}^{M_k}$ e a maleabilidade estimada $\tilde{A}_{S_0}^{M_k}$ foram obtidas. Finalmente, a maleabilidade geral estimada \tilde{A}_{S_0} foi determinada como sendo 2,8284, o maior valor encontrado, o qual corresponde ao valor de A_{S_0} obtido na Figura 2.17-B.

Fonte: Elaborada pelo autor.

Na Figura 3, podemos ver este método de estimativa aplicado no mesmo grafo apresentado na Figura 2.17-B. Cada tipo de medida produz um valor de maleabilidade diferente,

e o maior é igual ao obtido no espaço de transição original. Mas cuidado: em vários casos, a maleabilidade estimada pode ser menor do que a maleabilidade real, devido à seleção de medidas que são incapazes de distinguir entre todos os estados topológicos. Isto porque duas ou mais redes não isomórficas podem ser mapeadas para o mesmo ponto no estado de medidas. Assim, a adoção de mais medidas pode, em princípio, proporcionar uma melhor discriminação entre os gráficos não isomórficos.

A escolha do método acima baseado em medições para a estimativa da maleabilidade, embora seja suscetível a algum nível de degeneração pode, na prática, ser preferível pois, tipicamente, o interesse de quem estuda essas redes está interessado na maleabilidade com relação às propriedades específicas da rede, e não no problema abstrato de quantificar quão maleável é uma rede, levando em conta apenas os isomorfismos.⁶ Como exemplo, pode-se estar interessado em quão maleável é uma rede em relação aos seus caminhos mais curtos, caso em que a estimativa da maleabilidade levando em conta apenas esta medida fornecerá uma solução mais intuitiva e compreensiva do que o valor de maleabilidade obtido a partir das probabilidades de transição que distingue perfeitamente entre todos os estados não isomórficos.

Neste trabalho, para estudar a maleabilidade específica da medição, quisemos selecionar uma propriedade bem conhecida. Portanto, levamos em consideração: o grau médio $\langle K \rangle$, o coeficiente de agrupamento médio $\langle C \rangle$ ¹⁶ e a média dos menores caminhos $\langle L \rangle$. Entre estes, escolhemos o coeficiente de agregação médio. Isto porque a maleabilidade em relação ao grau médio seria trivial (a remoção de qualquer borda leva ao mesmo valor de $\langle K \rangle$, sempre resultando na maleabilidade estimada 1) e porque, embora o comprimento médio do caminho mais curto seja uma medida de alto interesse, seria computacionalmente muito caro para fazer nossa análise pretendida.

2.4 Caminhada aleatória

Caminhadas aleatórias surgem em diversos modelos matemáticos e físicos, desde o mapeamento das posições de jogos de baralho até movimentos brownianos de partículas de poeira. Caminhadas aleatórias são abordagens poderosas para simular e compreender a evolução de estados de um sistema.¹³

Em essência, uma caminhada aleatória é definida da seguinte forma: dado um grafo e um nó inicial, selecionamos um vizinho aleatoriamente e caminhamos até esse novo nó; então selecionamos um vizinho deste vértice aleatoriamente, e caminhamos até essa posição; e assim por diante. A sequência (aleatória) de posições percorridas desta maneira é uma *caminhada aleatória* realizada neste grafo. Na figura 4 é apresentado um exemplo ilustrativo.

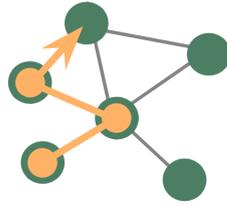


Figura 4 – Exemplo ilustrativa de uma caminhada aleatória, representada em cor laranja, sobre os nós de um grafo, representado em verde.

Fonte: Elaborada pelo autor.

2.5 Steering coefficient

Em 2007, foi realizado um estudo sobre caminhadas aleatórias^{17,18} em grafos dirigidos e sua correlação com a sua distribuição de graus.¹⁴ Nesse estudo, foi demonstrado que o **coeficiente de correlação de Pearson** r (2.5.1) entre o vetor de graus de uma rede dirigida G , com todos os nós apresentando grau de entrada igual ao grau de saída, e o vetor de frequência de visitação da caminhada aleatória que nela ocorre apresenta valor tendendo a 1 (alta correlação positiva). Aqui, em particular, vamos nos referir ao coeficiente de correlação de Pearson aplicado neste contexto como **steering coefficient** S (tradução: coeficiente de direção), tal como no trabalho.^{15,19} Este pode ser calculado em relação aos graus de entrada out em relação aos graus de saída. A nomenclatura foi dada devido à sua capacidade de mensurar quanto a topologia consegue dirigir a propriedade da dinâmica que ocorre no grafo analisado.

2.5.1 Coeficiente de correlação de Pearson

O coeficiente de correlação de Pearson ($-1 \leq r \leq 1$) é uma das medidas de correlação linear entre dois conjuntos de dados mais utilizadas. É dada pela razão entre a covariância de duas variáveis e o produto de seus desvios padrões, sendo assim uma medida normalizada da covariância.^{20,21} Quando o valor é próximo de +1, indica uma forte variação conjunta, com uma variável aumentando (ou diminuindo) da mesma forma que a outra. Já quando o valor é próximo de -1, a relação também é altamente correlacionada, porém de forma inversa. Quando o valor é próximo de 0, indica que não há uma variação conjunta entre as variáveis, como ilustrado na figura 5.

O coeficiente de correlação de Pearson entre duas variáveis aleatórias X e Y , quando aplicado a uma amostra, pode ser expresso como

$$r_{x,y} = \frac{cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2.18)$$

sendo cov a covariância entre as variáveis e σ o desvio padrão de cada uma delas.

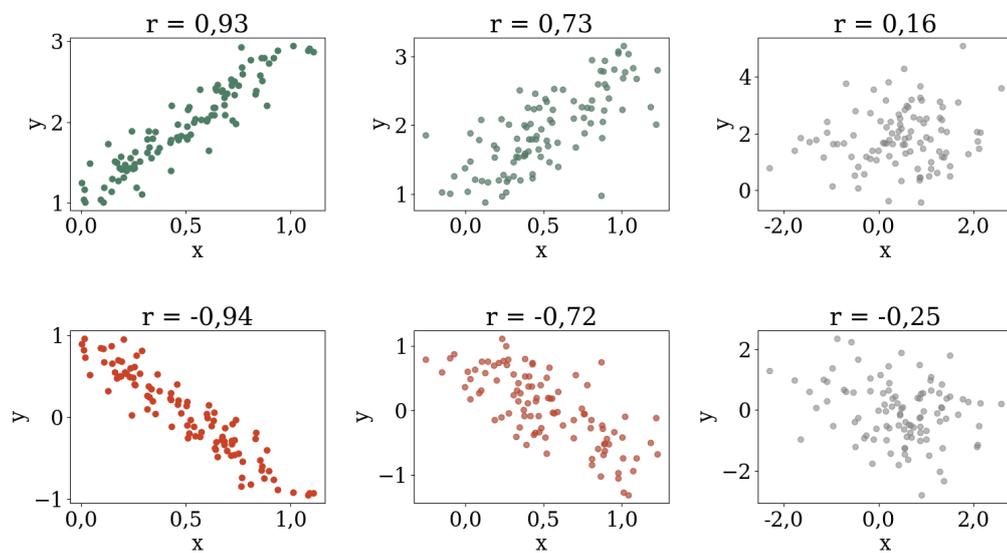


Figura 5 – Ilustrações de gráficos de dispersão (*scatterplots*) de duas variáveis x e y com diferentes valores correlação de Pearson r . Na primeira linha, em verde, vemos exemplos de correlações positivas; e de valores negativos representados em vermelho na segunda linha. A cor dos pontos de dispersão são pintados com menor saturação (mais acinzentados) à medida que r se aproxima de 0 (quando suas variáveis apresentam baixa correlação).

Fonte: Elaborada pelo autor.

3 DESENVOLVIMENTO, RESULTADOS E DISCUSSÃO

Neste trabalho, tivemos como objetivo principal estudar as propriedades associadas à maleabilidade em grafos sujeitos a consecutivas modificações, mais especificamente remoções de arestas.

Para isso, delineamos dois estudos, utilizando os mesmos grafos e sequência de modificações como objetos de análise.

No primeiro estudo, descrito na seção 3.1, tínhamos a intenção de aferir maleabilidade de redes modelos ao longo de sucessivas remoções de arestas, as quais foram escolhidas seguindo critérios variados, com o objetivo de comparar o efeito desses critérios no perfil de valores de maleabilidade.

Já no segundo estudo, descrito na seção 3.2, o objetivo principal foi o de analisar quais seriam os efeitos dos critérios de remoção de arestas em uma medida de dinâmica, mais especificamente caminhadas aleatórias, que acontecem nas redes estudadas. Para isso, aferimos o coeficiente de direcionamento nas redes modelos modificadas.

Estes dois estudos são reportados respectivamente nas seções seguintes.

3.1 Estudo 1 - Maleabilidade de redes modelos ao longo de sucessivas modificações

No estudo 1, quatro modelos de rede foram utilizados para nossas análises. Todos eles incluem 100 nós e grau médio de aproximadamente 6. Para melhor comparação, todos foram gerados utilizando da biblioteca NetworkX de Python 3.^{22,23} Os quatro modelos adotados são:

- **ER**: grafo não dirigido de Erdős-Rényi ($n = 100, p = 0.06$), descrito em 2.2.1;
- **WS1**: grafo não dirigido de Watts-Strogatz ($n = 100, k = 6, p = 0,05$), como em 2.2.2;
- **WS2**: grafo não dirigido de Watts-Strogatz ($n = 100, k = 6, p = 0,20$) e
- **BA**: grafo não dirigido de Barabási-Albert ($n = 100, m = 3, m_0 = 4$), descrito em 2.2.3.

Após a criação destes modelos (Figura 6 - A), todos os grafos foram re-interpretados como grafos bidirecionais simétricos (Figura 6 - B), em preparação para a realização do estudo de correlação pretendido, que envolverá a remoção de arestas direcionadas (Figura 6 - C). Isto significa que, inicialmente, para cada conexão do nó i ao nó j , existe também uma conexão dirigida do nó j ao nó i ($a_{ij} = a_{ji} = 1$), mas caso haja o rompimento de uma conexão direcionada (como de i a j), isso não implica na remoção da conexão recíproca j a i (é possível que $a_{ij} = 0$, enquanto $a_{ji} = 1$).

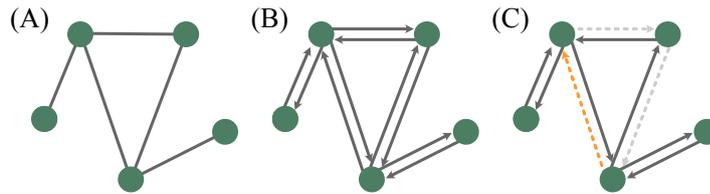


Figura 6 – Modelos de rede. (A) A rede é gerada como um grafo não direcionado. (B) A rede gerada é reinterpretada como um grafo simétrico dirigido, do qual uma conexão dirigida pode ser removida, sem quebrar sua correspondente conexão inversa. (C) A remoção consecutiva das arestas é realizada. As conexões pontilhadas representam as bordas removidas, onde a seta laranja é a mais recentemente removida.

Fonte: Elaborada pelo autor.

Em nosso estudo, O tipo de modificação escolhida a ser estudada foi a remoção de aresta. Escolhemos esta modificação, pois em modelos de grafos esparsos adotados, a realização dessa operação apresenta um custo computacional mais baixo quando comparada à adição ou transposição de arestas ou quando manipulamos vértices. Além disso, este é o mesmo tipo de modificação estudado por SILVA.⁵

Nosso interesse está também em analisamos sequências de remoções sucessivas. Para isso, em cada rede modelo, realizamos remoções consecutivas de 100 arestas, escolhidas por diferentes critérios:

- **Aleat**: aresta selecionada aleatoriamente*;
- **Max \tilde{A}** : aresta que, se removida, resulta em maior crescimento da maleabilidade estimada \tilde{A} (conforme definida em 2.3.2), sendo assim uma otimização (maximização) gulosa da medida \tilde{A} dentro da sequência de remoções*;
- **Min \tilde{A}** : aresta que, se removida, resulta em maior decréscimo da maleabilidade estimada \tilde{A} (conforme definida em 2.3.2), sendo assim uma minimização gulosa da medida \tilde{A} dentro da sequência de remoções*.

A determinação de quais são as arestas que, se removidas, levam a um valor máximo ou mínimo de maleabilidade foi realizada considerando todas as redes possíveis consequentes de uma única remoção de aresta.

Com os modelos e protocolos enunciados, repetimos o estudo em 20 redes para cada tipo de modelo, para que pudéssemos saber o perfil médio dos valores do coeficiente de agrupamento médio e maleabilidade estimada para cada tipo de rede. Apresentamos, então, a seguir, os resultados resultantes desses estudos.

Observamos na figura 7 que, conforme o pressuposto intuitivo, a curva **Max \tilde{A}** que maximiza de forma gulosa para a medida \tilde{A} apresentou a maleabilidade máxima observada; que a

* Desde que o grafo permaneça fortemente conectado.

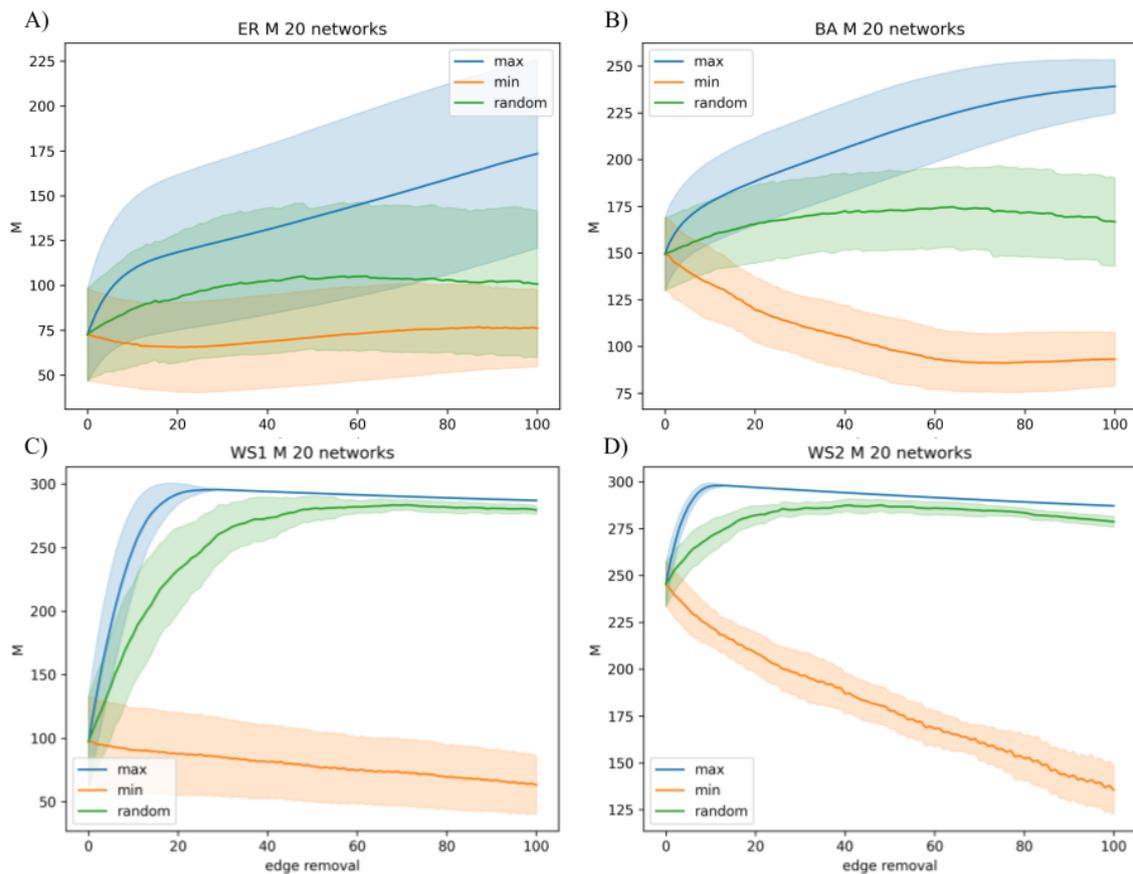


Figura 7 – Maleabilidade aproximada em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max** \tilde{A} , **Min** \tilde{A} e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20.

Fonte: Elaborada pelo autor.

curva **Min** \tilde{A} que minimiza de forma gulosa para a medida \tilde{A} apresentou a maleabilidade mínima observada e que a curva **Aleat** de remoções aleatória apresentou comportamento intermediário para todos os tipos de modelos.

Entretanto, é notável que as redes do tipo **WS** apresentam um padrão interessante, em que a curva **Aleat** tem seu valor acrescido de forma aguda durante as suas primeiras remoções até atingir um máximo, a partir do qual, segue decrescendo de forma aparentemente assintótica a uma curva com dispersão cada vez menor. A curva **Max** \tilde{A} segue um padrão bastante similar, porém alcança tal assíntota de forma otimizada. Já a curva **Min** \tilde{A} apresenta um perfil completamente distinto das curvas decorrente dos outros critério de remoção mencionados anteriormente.

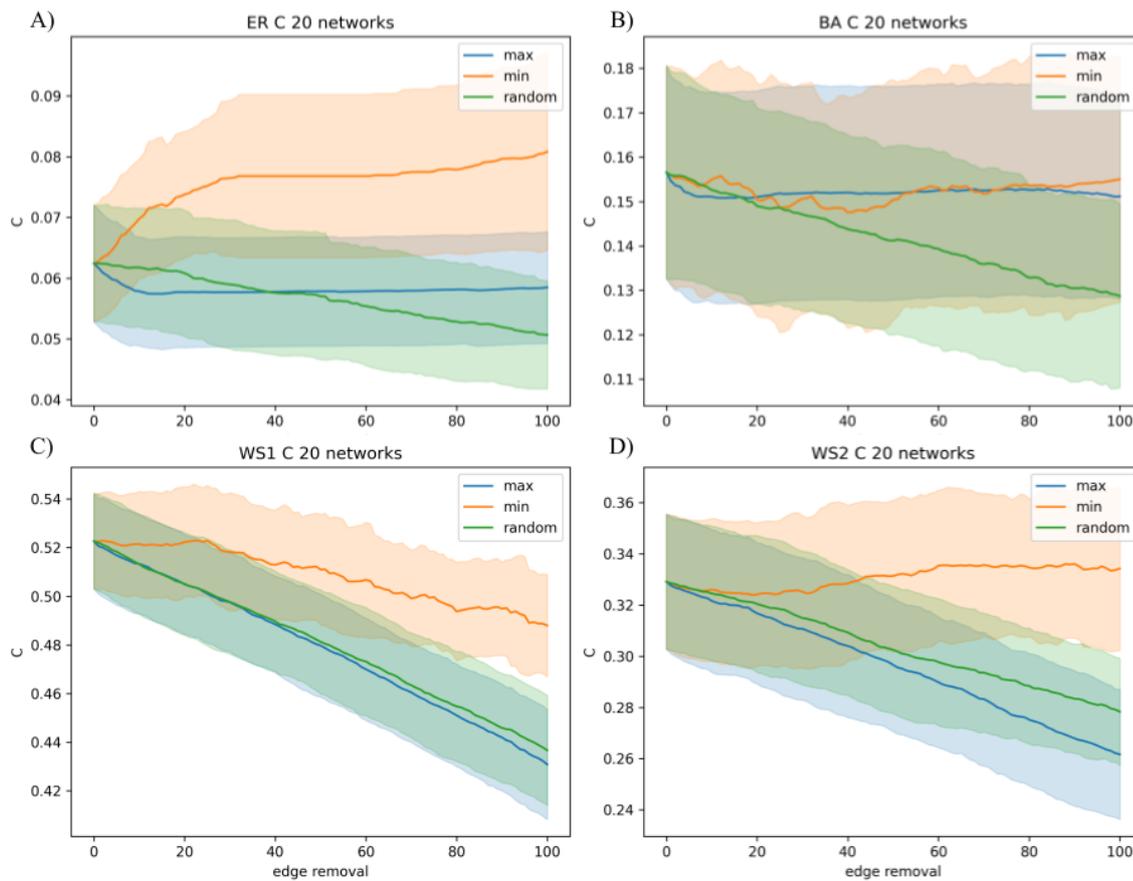


Figura 8 – Coeficiente de agrupamento médio em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max** \tilde{A} , **Min** \tilde{A} e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20.

Fonte: Elaborada pelo autor.

Ao analisar a consequência dessas remoções no valor do coeficiente de agrupamento médio, apresentado na figura 7, percebemos que em todos os casos, a curva **Min** \tilde{A} causou uma maximização do valor de $\langle C \rangle$ no fim do experimento (ou pelo menos um decaimento reduzido). Isso é relevante, já que observamos uma tendência natural de que, em todos os casos, as curvas **Aleat** de remoções aleatórias levaram o coeficiente de agrupamento médio a diminuir ao longo de remoções. Tal redução significa que triângulos de arestas foram rompidos no processo. A conservação do $\langle C \rangle$ nas curvas **Min** \tilde{A} implicam numa predileção de remoções de conexões que não participam da formação de triângulos (comunidades).

O caso do modelo **BA** é particularmente interessante, pois a curva **Max** \tilde{A} também apresenta resultado similar à curva **Min** \tilde{A} , e não um perfil similar às remoções aleatórias, como

ocorreu nos demais casos.

Assim, percebemos que as curvas de remoções de arestas que minimizam gulosamente a maleabilidade, e a curva **Max \tilde{A}** do modelo **BA** em particular, apresentam robustez (resistência à modificações) em relação ao valor do coeficiente de agrupamento.

3.2 Estudo 2 - Efeito do critério de remoção de aresta condicionado por maleabilidade no coeficiente de direcionamento de redes modelos

Nós desejávamos avaliar o efeito de remoções consecutivas de arestas condicionadas pelo valor da maleabilidade em uma medida de dinâmica nessas redes. Para isso, optamos por uma dinâmica bem estabelecida: uma caminhada aleatória através da rede direcionada. As propriedades da caminhada foram definidas como:

- O andarilho começa a partir de um nó aleatório;
- O andarilho caminha em direção a um dos vizinhos de saída, escolhido aleatoriamente, tendo cada vizinho igual probabilidade de ser selecionado e
- O passeio dura 10^6 passos.

Como medida para avaliar a caminhada, adotamos o *steering coefficient* r (respectivo aos graus de entrada e respectivo aos graus de saída), conforme definido em 2.5. Em cada etapa de remoção, ao calcar o *steering coefficient*, foram realizadas 100 simulações de caminhadas, para que pudéssemos ter os perfis médios com os respectivos desvios padrões.

As redes modelos utilizadas para o Estudo 2 foram as mesmas definidas no Estudo 1 (3.1), tendo duas conexões consecutivamente removidas seguindo também os mesmos critérios nele descritos.

Apresentamos, então, os resultados decorrentes das simulações.

Percebemos que os resultados demonstrados nas figuras 9 e 10 são bastante semelhantes, com a diferença de que os efeitos são ampliados e menos ruidosos quando analisamos a correlação com o grau de saída. Portanto, faremos aqui a discussão majoritariamente focada no que pode ser observado na figura 10, porém os comentários são aplicáveis aos resultados da correlação respectiva aos graus de entrada apresentados na figura 9 também.

A característica comum a todos é que as remoções aleatórias tendem a reduzir o *steering coefficient*, sendo o efeito muito mais fortemente notável nas redes do tipo Watts-Strogatz dentro do intervalo observado, onde a correlação se perde, aproximando-se de 0.

As remoções que minimizam de forma gulosa a maleabilidade tendem a preservar o *steering coefficient* e aquelas que maximizam a maleabilidade apresentam perfil intermediário, mas com maior semelhança com as remoções aleatórias.

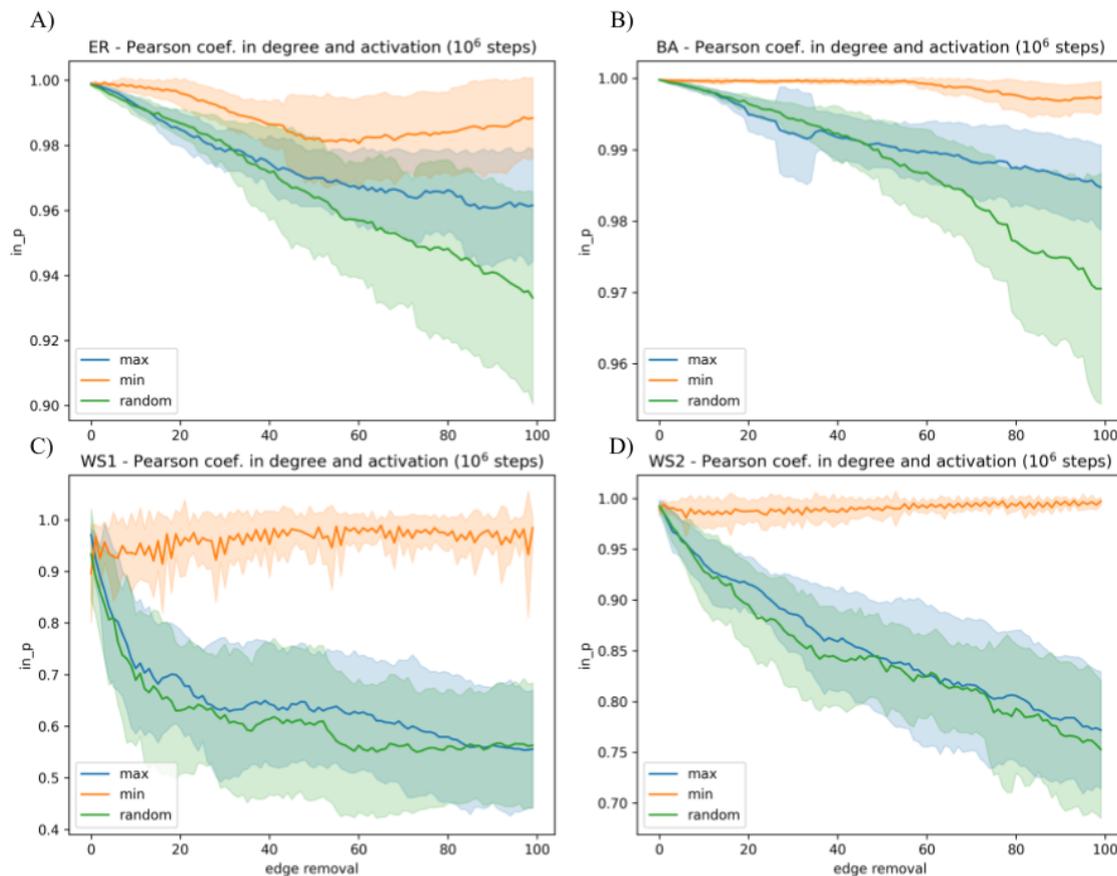


Figura 9 – *Steering coefficient* respectivo aos graus de entrada em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **Max**, **Min** e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20.

Fonte: Elaborada pelo autor.

O resultado mais marcante foi o forte contraste de perfis de r nas redes do tipo Watts-Strogatz. Nessas redes, o r teve fortíssima robustez em relação a remoções que minimizam a maleabilidade, apresentando correlação acima de 0,9, enquanto os outros critérios de remoção acarretaram numa perda abrupta da correlação, mesmo que a mesma quantidade de arestas tenham sido removidas.

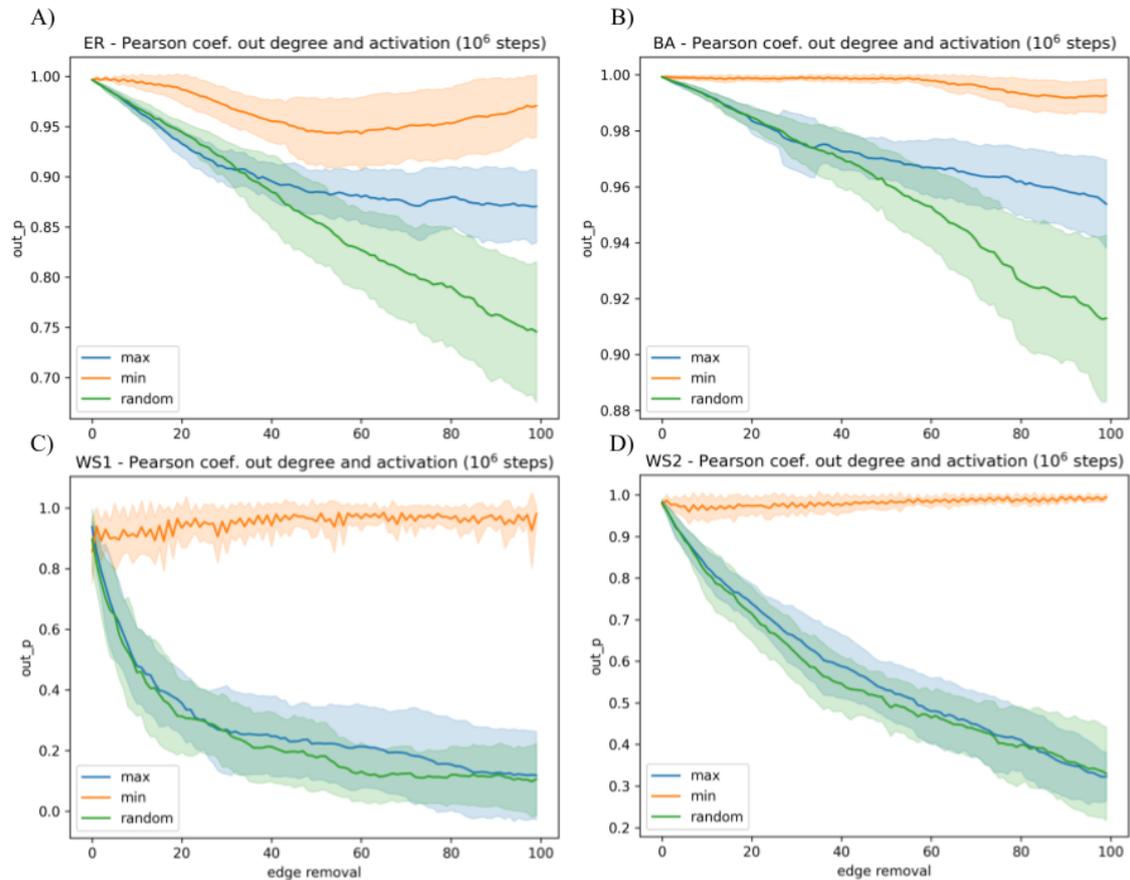


Figura 10 – *Steering coefficient* respectivo aos graus de saída em função da quantidade de arestas consecutivamente removidas. Todos os sub-gráficos apresentam três curvas de remoção de arestas: **MaxÃ**, **MinÃ** e **Aleat**, conforme enunciadas em 3.1. A linha sólida representa o valor médio das 20 redes geradas para cada rede modelo, a área semi-transparente representa o desvio padrão. A) modelo Erdős-Rényi (2.2.1); B) modelo de Barabási e Albert (2.2.3); C) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,05; D) modelo de Watts e Strogatz (2.2.2) construído com probabilidade de reconexão 0,20.

Fonte: Elaborada pelo autor.

4 CONCLUSÃO

A crescente disponibilidade de dados do mundo real permite uma oportunidade sem precedentes para a respectiva análise e construção de modelos que possam levar a um melhor entendimento dos sistemas que produzem esses dados. Embora diversas abordagens interessantes tenham sido propostas para a caracterização de dados representados como grafos, uma quantidade menor de trabalhos têm abordado a quantificação e estudo de alterações da topologia desses grafos quando submetidos a modificações incrementais progressivas e como essas alterações podem impactar na respectiva mudança ou maleabilidade dos grafos.

O presente trabalho teve como principal objetivo estudar a recentemente introduzida abordagem para estudo da maleabilidade de redes⁵ considerando-se três modelos teóricos de redes complexas, com foco na medida de coeficiente de agrupamento. Em nosso trabalho, porém, demos ênfase no estudo de como esta medida de maleabilidade é afetada, quando as arestas são gradualmente removidas. Para selecionar as arestas a serem rompidas, usamos três abordagens distintas: escolha aleatória, a seleção que maximiza a maleabilidade a cada etapa e, por último, a seleção que minimiza a maleabilidade a cada etapa (ambas de maneira gulosa).

Ao analisarmos as curvas de maleabilidade resultantes, percebemos que as curvas de maximização gulosa da maleabilidade atingem o valor máximo possível, enquanto as curvas de minimização gulosa levam à menor maleabilidade observada. As remoções aleatórias apresentam um comportamento intermediário para todos os tipos de redes, que são comportamentos intuitivos. Destacamos, porém, que um padrão intrigante surgiu nas redes do tipo **WS**, nas quais as remoções aleatórias inicialmente aumentam rapidamente a maleabilidade até atingir um pico, a partir do qual diminui gradualmente, seguindo uma tendência aparentemente assintótica. O padrão da curva que maximiza a maleabilidade é semelhante, mas atinge a assíntota de forma mais otimizada. A curva de minimização, por sua vez, exibe um perfil distinto: quase uniforme e decrescente.

Para explorar como essas remoções afetam dinâmicas em redes, conduzimos caminhadas aleatórias em cada uma das redes mencionadas, a cada aresta removida. Registramos quais nós foram ativados em cada caminhada e determinamos sua correlação com os graus de entrada e saída ao longo das remoções. Conforme os artigos que a introduziram, chamamos essa correlação de *steering coefficient*.^{14,15}

Comparamos os *steering coefficients* obtidos de remoções aleatórias com aqueles de remoções que visam minimizar ou maximizar a maleabilidade e notamos que a remoção aleatória tende a diminuir o seu valor conforme o esperado, já que na primeira simulação de seu valor máximo (1.0), e as simulações subsequentes têm valores menores, pois têm menor simetria entre graus de entrada e saída. Por outro lado, as remoções que diminuem a maleabilidade

da rede levam a uma redução do *steering coefficient* muito mais suave. As remoções que maximizam a maleabilidade apresentam um perfil intermediário, mas ainda mais próximo das remoções aleatórias. Entre todos os modelos, as redes Watts-Strogatz foram as que apresentaram o perfil mais interessantemente distinto, demonstrando resiliência deste coeficiente em relação às remoções que minimizam a maleabilidade, conservando significativamente sua correlação em 0.9 ou maior. Em contraste, as outras curvas consequentes de remoções que seguiram outros critérios resultaram em uma queda abrupta na correlação.

Assim, os resultados mais interessantes decorrentes deste trabalho foram:

- em redes do tipo **WS**, remoções que maximizam a maleabilidade* de forma gulosa levam a valores de maleabilidade* similares a remoções simplesmente aleatórias, ambas com tendência de manter seu valor elevado. Por outro lado, remoções que minimizam a maleabilidade* de forma gulosa levam a uma redução consistente e quase suave da maleabilidade*;
- em todos os modelos, as remoções que minimizam a maleabilidade* levaram ao maior valor de coeficiente de agrupamento médio observado, indicando uma conservação das trincas de conexões na rede;
- em todos os modelos, as remoções que minimizam a maleabilidade* levaram a uma maior conservação do *steering coefficient*, indicando uma maior simetria entre os graus de entrada e saída dos nós da rede. Este efeito foi mais evidente nas redes do tipo **WS** e
- na maioria dos casos, para as medidas aferidas, as remoções que maximizam o valor da maleabilidade* apresentaram resultados similares àqueles das remoções de arestas aleatórias.

Nossa pesquisa indica como a medida maleabilidade estimada por meio de uma medida específica, coeficiente de agrupamento médio, se comporta ao longo de sucessivas modificações, do tipo remoção de aresta. Acreditamos que, apesar de limitar-se a uma combinação específica de modelos e critérios, apresentamos aqui um passo significativo para um melhor entendimento de redes observadas por meio desta medida. Alguns dos possíveis desenvolvimentos futuros nesta linha de pesquisa são:

- conduzir estudos similares, mas em redes reais de características similares;
- estudar o padrão da maleabilidade quando estimada por meio de outras medidas relevantes, como o menor caminho médio e
- avaliar os efeitos de outras perturbações de redes, como remoções de nós.

* Maleabilidade estimada pela medida: coeficiente de agrupamento médio.

Tais estudos complementares tem potencial para levar a um melhor entendimento das propriedades e eventual controle da maleabilidade de redes complexas.

REFERÊNCIAS

- 1 ABDALLA, H. B. A brief survey on big data: technologies, terminologies and data-intensive applications. **Journal of Big Data**, Springer Science and Business Media LLC, v. 9, n. 1, nov. 2022.
- 2 COSTA, L. da F. *et al.* Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, Informa UK Limited, v. 60, n. 3, p. 329–412, June 2011.
- 3 COSTA, L. da F. *et al.* Characterization of complex networks: a survey of measurements **Advances in Physics**, Informa UK Limited, v. 56, n. 1, p. 167–242, jan. 2007.
- 4 BOCCALETTI, S. *et al.* Complex networks: structure and dynamics. **Physics Reports** , v. 424, n. 4-5, p. 175–308, 2006.
- 5 SILVA, F. N.; COMIN, C. H.; COSTA, L. da F. Malleability of complex networks. **Journal of Statistical Mechanics: theory and experiment** , IOP Publishing, v. 2019, n. 8, p. 083203, Aug. 2019.
- 6 BABAI, L.; ERDOS, P.; SELKOW, S. M. Random graph isomorphism. **SIAM Journal on Computing** v. 9, n. 3, p. 628–635, 1980.
- 7 GROHE, M.; SCHWEITZER, P. The graph isomorphism problem. **Communications of the ACM**, v. 63, n. 11, p. 128–134, Oct 2020.
- 8 ERDÖS, P.; RÉNYI, A. On random graphs i. **Publicationes Mathematicae Debrecen**, v. 6, p. 290–297, 1959.
- 9 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. **Nature**, v. 393, n. 6684, p. 440–442, June 1998.
- 10 BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, n. 5439, p. 509–512, Oct. 1999.
- 11 NEWMAN, M. E. J. Mixing patterns in networks. **Physical Review E**, v. 67, n. 2, Feb. 2003.
- 12 PENG, S. *et al.* Influence analysis in social networks: a survey. **Journal of Network and Computer Applications**, v. 106, p. 17–32, 2018.
- 13 LÁSZLÓ, L.; LOV, L.; ERDÖS, O. **Random Walks on Graphs: a survey**. 1993. 46 p.
- 14 COSTA, L. F. *et al.* Correlations between structure and random walk dynamics in directed complex networks. **Applied Physics Letters**, v. 91, n. 5, p. 054107, July 2007.
- 15 ARRUDA, H. F. de *et al.* Network information science. 2017. Disponível em: <http://arxiv.org/abs/1704.03091>. Acesso em: 23 jan. 2023.
- 16 FAGIOLO, G. Clustering in complex directed networks. **Physical Review E**, v. 76, n. 2, Aug. 2007.
- 17 MASUDA, N.; PORTER, M. A.; LAMBIOTTE, R. Random walks and diffusion on networks. **Physics Reports** v. 716, p. 1–58, 2017.

18 NOH, J. D.; RIEGER, H. Random walks on complex networks. **Physical Review Letters**, APS, v. 92, n. 11, p. 118701, 2004.

19 SOUZA, P. J. P.; COMIN, C. H.; COSTA, L. F. **Topology and dynamics in complex networks**: the role of Edge reciprocity. 2017. Disponível em: <https://arxiv.org/abs/1711.08838>. Acesso em: 23 jan. 2023.

20 PEARSON, K. **Mathematical contributions to the theory of evolution**. London: Dulau, 1907.

21 WRIGHT, S. Correlation and causation. **Journal of Agricultural Research**, v. 20, n. 7, p. 557–585, 1921.

22 HAGBERG, A. A.; SCHULT, D. A.; SWART, P. J. Exploring network structure, dynamics, and function using networkx. *In*: PYTHON IN SCIENCE CONFERENCE, 7., 2008, Pasadena. **Proceedings** [...]. Pasadena: Los Alamos, 2008. p. 11 – 15.

23 ROSSUM, G. V. **Python reference manual** . Amsterdam: Centre for Mathematics and Computer Science, 1995.