Universidade Federal do Rio Grande do Sul
Instituto de Física
Programa de Pós-Graduação em Física

Determinação Empírica do Ponto Ótimo de Fragmentação para Redes Modulares

Carolina de Abreu Pereira

Porto Alegre - RS, Novembro de 2019

Carolina de Abreu Pereira

Determinação Empírica do Ponto Ótimo de Fragmentação para Redes Modulares

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Física da UFRGS, como parte dos requisitos necessários para a obtenção do Título de Mestre em Física.

Universidade Federal do Rio Grande do Sul – UFRGS

Instituto de Física

Programa de Pós-Graduação em Física

Orientador: Sebastián Gonçalves

Coorientador: Bruno Requião da Cunha

Porto Alegre - RS

Novembro de 2019

Agradecimentos

Agradeço aos meus pais, Clarivani e Claudio, e aos meus irmãos, Camila e Augusto, pelo apoio incondicional, por todo amor, paciência e compreensão. Também agradeço às minhas tias e tios, que sempre se fizeram presentes e me incentivaram a seguir. Especialmente, agradeço à Sandrinha, por ter continuado me guiando em mais uma jornada; assim como o Guilherme, que foi de importância imprescindível, e a minha amiga Marjana, que esteve comigo nos tempos bons e nos tempos não tão bons assim. Agradeço ao meu orientador, Sebastián Gonçalves, e ao meu co-orientador, Bruno R. da Cunha, pelo conhecimento que me foi transferido, pela disponibilidade e boa vontade em ajudar.

É impossível listar aqui todos os colegas e amigos que fizeram parte da minha vida nestes 2,5 anos, desejo agradecer em especial ao Sandro, João, e Thamiris pelo café de todo dia. Agradeço aos colegas da M201: Fernanda, Paulo, Emanuel, e Calvin, pelas risadas, conversas aleatórias e desabafos. Também agradeço aos colegas de grupo Alexandre e Ben-Hur.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), agradeço pelo apoio financeiro, sem o qual este trabalho nunca teria sido realizado.

Resumo

Muitas redes reais tendem a se organizar em estruturas de comunidades, neste trabalho é explorada a relação entre a existência dessas estruturas e a fragilidade das redes frente a remoções de nodos, também chamado de ataque. Para tanto, realizou-se todas as possíveis remoções de n nodos de redes de tamanho N, então, para cada rede, foi medido o tamanho da maior componente conectada restante depois de cada remoção. Dentre essas medidas, o menor tamanho de componente obtido representa o dano máximo possível à rede, limitado à remoção de n nodos. O conjunto de n nodos que produz tal dano é chamado de conjunto ótimo. Aplicou-se o procedimento em uma série de redes com modularidade controlada e variada, sendo a modularidade uma medida do quão bem um rede pode ser dividida em comunidades. Estes resultados foram comparados com resultados de métodos heurísticos de fragmentação de rede, em i.e., ataque Adaptativo de Alta Intermediação (HBA), Influência Coletiva (IC) e Ataque Baseado em Módulos (MBA). Por questões práticas, foram escolhidos principalmente ataques de tamanho $n = 5$ em redes de tamanho $N = 100$, devido aos limites computacionais para remoção dos nodos, aproximadamente 7.5×10^7 combinações para esse caso. Os resultados mostram que a robustez das redes, tanto para o ataque ótimo quanto para direcionados, tem uma relação inversa com a modularidade. Para modularidades inferiores a um valor crítico, todas as estratégias heurísticas estudadas são muito semelhantes a remoções aleatórias de nodos. Por outro lado, as redes são altamente vulneráveis a ataques heurísticos para modularidades superiores ao valor crítico.

Palavras-chaves: Redes complexas, Teoria de redes, Fragmentação de redes.

Abstract

Many real networks present community structure, this work explores the relationship between the existence of these structures and the fragility of networks in relation to node removals, also called attacks. All possible removals of n nodes from networks of size N were performed, and for each network the size of the largest connected component remaining after each removal was measured. Among these measures, the smallest component size obtained represents the maximum possible network damage, limited to the removal of n nodes. The set of n nodes that produces such damage is called the optimal set. The procedure was applied to a series of networks with controlled and varied modularity, which is a measure of how well defined are the communities within a network. These results were compared with results from heuristic network fragmentation methods, i.e., high betweenness adaptive attack (HBA), Collective Influence (CI), and Module Based Attack (MBA). For practical reasons, attacks of size $n = 5$ on networks of size $N = 100$ were chosen, mainly due to the computational limits for node removals, approximately 7.5×10^7 combinations in this case. The results show that network robustness under the optimal and the heurístic attacks has an inverse relation with the modularity, and that all the analised heuristic strategies are very similar to random removals for modularities below a critical value. On the other hand, networks are highly vulnerable to heuristic attacks for modularities greater than the critical value.

Keywords: Complex networks, Network theory, Network fragmentation.

Sumário

1 Introdução

Atualmente, dependemos de um número considerável de sistemas em nosso cotidiano, os quais podem ser mapeados como um conjunto de pontos interligados, sendo este conjunto o que compõe uma rede. Como exemplo, pode-se considerar uma rede de energia elétrica, que inclui usinas, subestações e linhas de transmissão. Nesse tipo de rede, é muito provável que alguns de seus nodos tenham um papel mais importante que outros. Diferentes fatores podem causar uma falha no seu comportamento usual. Para exemplificar, citamos um evento recente de *blackout* que ocorreu na Argentina e em partes do Uruguai, atingindo aproximadamente 50 milhões de consumidores [1]. Acredita-se que houveram falhas estruturais em pelo menos três centrais de energia e a rede elétrica não foi capaz de suportar tamanho impacto, o que resultou na sua total desativação.

Ao mesmo tempo em que há redes que devem ser protegidas para minimizar os efeitos de falhas em alguns de seus componentes, existem casos em que desejamos o contrário: atacar elementos do sistema para impedir, ou pelo menos reduzir, seu funcionamento. Neste contexto, podemos citar como exemplo a rede de crimes federais no Brasil (CUNHA; GONÇALVES, 2018). Essa rede consiste de indivíduos envolvidos em diversas atividades criminais, como crime organizado transnacional, lavagem de dinheiro, terrorismo, crimes cibernéticos, entre outros, que estão conectados ou relacionados entre si de alguma maneira relevante para a investigação.

Nos dois exemplos citados acima, rede elétrica e rede criminal, usou-se o significado usual do termo "rede". Formalmente, uma rede é definida como um conjunto de nodos (também chamados de vértices ou nós) e *links* (ou arestas), que cumprem o papel de conectar os nodos. Há duas maneiras de propositalmente fragmentar uma rede: pela remoção de nodos ou pela remoção de *links*. No caso da rede criminal, os nodos correspondem aos indivíduos e os *links* às relações entre eles. Existem, portanto, duas formas de fragmentar uma rede criminal: pelo encarceramento desses indivíduos, que perderiam algumas de suas conexões (*links*) em virtude do isolamento; ou pela completa reintegração do mesmo à sociedade (remoção do nodo), abandonando atividades criminais, ou ainda, de forma mais drástica, devido à sua morte (CUNHA; GONÇALVES, 2018).

A Teoria de Redes fornece uma estrutura natural para o estudo sistemático da robustez ou fraqueza de sistemas interligados; em particular, para ajudar a identificar os principais fatores topológicos que desempenham um papel crucial na coesão de sistemas representados por redes. Diferentes redes respondem de maneira diferente à ataques dirigidos ou à aleatórios, podendo ser mais ou menos robustas. A robustez de redes é

[1] https://www.dw.com/en/argentina-uruguay-paraguay-suffer-massive-power-blackout/a-49225070 (acessado em 02/09/2019)

uma questão muito importante em uma ampla variedade de disciplinas aplicadas, como física, biologia, engenharia, sociologia e criminologia (FAN et al., 2018; MURO et al., 2018; BELLINGERI et al., 2018; HU et al., 2018; CUNHA; GONÇALVES, 2018; STAM, 2014) dentre outras. Uma maneira de analisar o problema da robustez de uma rede é encontrar o conjunto de desmantelamento ideal, definido como o conjunto mínimo de nodos que, uma vez removidos, deixam a rede dividida em componentes de tamanho irrelevante para o funcionamento da rede.

A outra maneira de enfrentar o problema é identificar o conjunto que produz o máximo dano possível dado um ataque de tamanho fixo, que corresponde à remoção de um número predefinido de nodos. Em outras palavras, reduzir a maior componente conectada da rede ao menor tamanho possível tendo recursos limitados. O problema ainda é de complexidade combinatória, pois é necessário descobrir qual conjunto de nodos n produz o maior dano em uma rede de tamanho N, então um total de $\binom{N}{n}$ conjuntos deve ser testado. Portanto, este problema pode ser computacionalmente intratável mesmo para redes pequenas, dependendo do número de nodos que se deseja remover.

Como base no exposto acima, este trabalho tem como objetivo explorar como a fragilidade de redes complexas é alterada em função do quão bem elas podem ser divididas em comunidades — grupos de nodos mais conectados entre si do que com o restante da rede, também chamados módulos. Esta escolha é motivada pelo fato de que redes com estrutura de comunidades bem definidas são diferentes de redes puramente aleatórias ou de pequeno mundo (FORTUNATO, 2010). De fato, a característica de modularidade de uma determinada rede sinaliza a presença de outras conexões não apenas aleatórias entre os nós; tal aspecto pode ser encontrado, por exemplo, em redes sociais, onde os indivíduos tendem a formar comunidades. Além disso, sabe-se que a modularidade desempenha um papel fundamental em uma ampla variedade de fenômenos, desde redes de crimes (CUNHA; GONÇALVES, 2018), até redes cerebrais (BERTOLERO; YEO; D'ESPOSITO, 2015; BULLMORE; SPORNS, 2012; FAUST; KENETT, 2014; MEUNIER; LAMBIOTTE; BULLMORE, 2010; GALLEN; D'ESPOSITO, 2019).

Cabe salientar que os ataques baseados em módulos (MBA) (CUNHA; GONZÁLEZ-AVELLA; GONÇALVES, 2015) e de alta intermediação adaptativo (HBA) (HOLME et al., 2002) são conhecidos por atomizar redes modulares com o menor número de remoções, sendo o MBA muito mais viável computacionalmente (CUNHA; GONÇALVES, 2017). No entanto, ainda não está claro quão próximos esses métodos estão do conjunto ótimo, ou mesmo de qual valor de modularidade a propriedade modular se destaca. Tal questionamento é precisamente o que é abordado no presente trabalho, realizando uma análise computacional por força bruta. Este tipo de análise consiste em gerar todas as $\binom{N}{n}$ possíveis listas de remoção para redes de tamanho N e realizar a remoção de cada lista. As estatísticas resultantes são comparadas com os resultados de métodos de ataque heurístico sabidamente

efetivos.

Nas próximas seções, são descritos a metodologia do trabalho, o procedimento de ataques por força bruta e os métodos heurísticos de ataque. Em seguida, apresenta-se os resultados com a estatística de ataques de força bruta em comparação com os resultados heurísticos, assim como uma análise de tamanho finito e outros aspectos relevantes do problema. A seção final é dedicada às conclusões.

2 Teoria de Redes Complexas

Alguns materiais apresentam, no estado sólido, uma estrutura microscópica cristalina formada por uma rede de átomos conectados entre si de forma regular. Em outra palavras, nessas estruturas, todos os átomos estão ordenadamente conectados com um número fixo de outros átomos, como uma grade regular em duas dimensões. Essas redes cristalinas são estruturas periódicas, onde um certo padrão é repetido periodicamente preenchendo o espaço. Outras estruturas teóricas ou redes possíveis são as chamadas redes aleatórias, em que todos os seus nodos têm, em média, mas não exatamente, o mesmo número de conexões, não havendo pontos que se destacam. As redes aleatórias são em certa forma o oposto das redes ordenadas: não apresentam um padrão regular como as redes cristalinas, não tendo, portanto, qualquer simetria. Outrossim, há algo em comum nesses dois extremos. Nas duas estruturas, nenhum elemento é mais importante que outro. Porém, em uma estrutura de rede cristalina monoatômica, todos os átomos são equivalentes, e em uma rede aleatória não há, em princípio, dois nodos que sejam equivalentes entre si, devido ao padrão de ligações diferente para cada um.

Contudo, a maior parte das redes reais não obedece a nenhum destes dois extremos teóricos. Não são nem perfeitamente regulares, nem completamente aleatórias. Redes reais possuem estruturas mais complexas, com certo um certo nível de desordem, mas que também apresentam certa regularidade dentro delas. Uma vez que não são homogêneas como as redes aleatórias ideais, alguns de seus elementos podem ser diferentes do restante. Uma forma que nodos podem se sobressair é pelo seu número maior de conexões, ou por sua posição privilegiada na estrutura. A seguir serão apresentados os conceitos básicos da teoria de grafos, que é usada para estudar tais tipos de redes.

2.1 Medidas Básicas

Uma rede, ou grafo, consiste em um conjunto de N vértices e E arestas. Usualmente, representa-se uma rede através da matriz de adjacência, uma matriz quadrada em que cada linha e cada coluna representam um vértice (BARABÁSI, 2016). Na sua versão mais básica, seus elementos representam as ligações entre os vértices como

$$A_{ij} = \begin{cases} 1\,, & \text{se há um link apontando de } j \text{ para } i; \\ 0\,, & \text{caso contrário.} \end{cases} \tag{2.1}$$

Redes podem ser dirigidas quando existe uma ordem específica na conexão, ou não dirigidas quando a conexão é recíproca, o que consequentemente faz a matriz de adjacência ser simétrica. Esta matriz pode ainda ter valores diferentes de 1 em seus

elementos, isto acontece quando as conexões possuem pesos, ou seja, uma conexão pode ser mais importante do que outra, dependendo do seu valor.

A principal característica de um vértice é o número de conexões que ele possui, intitulado grau e denotado por k_i para um vértice i qualquer. Em redes dirigidas são definidos dois tipos de grau: o número de conexões apontando para o nodo – grau de entrada k_{in}, o número de conexões saindo deste vértice e apontando para outros – grau de saída k_{out}. A distribuição de ocorrências de certo grau é uma medida utilizada extensivamente no processo de caracterização de uma rede. Tal distribuição é dada pela probabilidade de um vértice possuir certo grau k:

$$p(k) = N_k/N, \tag{2.2}$$

onde N_k é o número de nodos com grau k.

Uma maneira de definir o modelo de redes aleatórias de Erdös-Rényi (ER) (BOLLOBAS, 1985) é fazendo com que cada um dos N vértices da rede se conectem com outro vértice aleatório com uma probabilidade p, o que resulta em uma distribuição de grau binomial, que para $N >> \langle k \rangle$ torna-se uma distribuição de Poisson, que pode ser descrita somente em função do grau médio da rede:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle}{k!}. \tag{2.3}$$

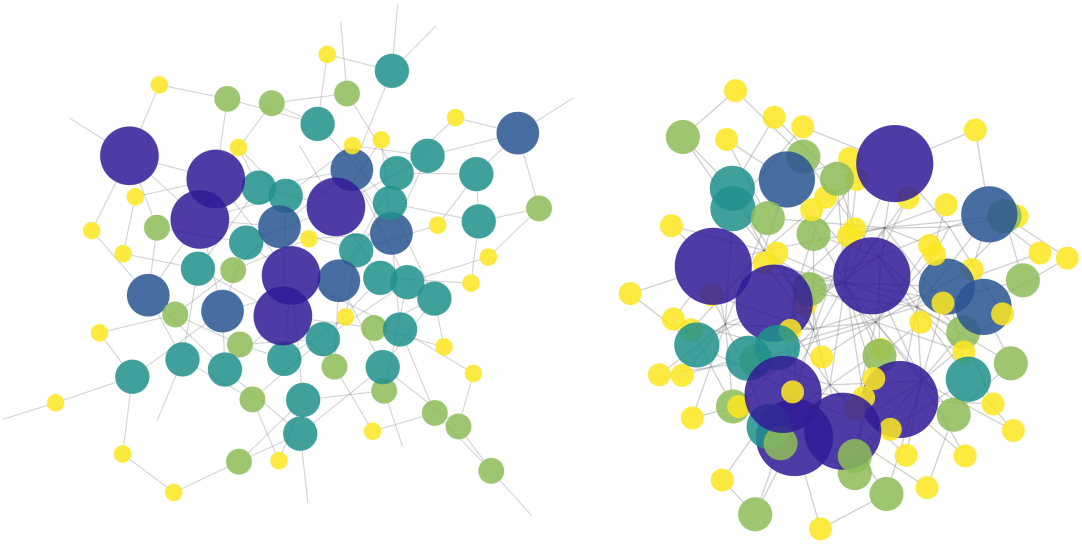
Neste modelo a média e o desvio padrão da distribuição de grau k são bem definidos.

Geralmente, redes reais apresentam distribuições de grau heterogêneas, na forma de uma lei de potência, chamadas redes *scale-free*, ou "invariante em escala" (BARABÁSI; ALBERT, 1999); portanto, a probabilidade de que um nodo tenha grau k é da forma:

$$P(k) \sim k^{-\gamma}, \tag{2.4}$$

onde γ é um expoente característico de cada rede, geralmente entre $2 < \gamma < 3$. O que se destaca neste tipo de rede é a presença de *hubs* – vértices que possuem um grau muito mais alto que a média – e uma grande quantidade de vértices com grau muito baixo. Foram desenvolvidos modelos para tentar obter uma estrutura de distribuição de grau mais complexa do que a de uma rede aleatória. Um modelo que se destaca é o de Barabási-Albert (BA) (BARABÁSI; ALBERT, 1999), que é baseado no conceito de ligação preferencial (*preferential attachment*). Essa expressão significa que a rede se expande pela adição de novas conexões, tendo maior probabilidade de serem geradas nas regiões em que já existem, previamente, nodos bem conectados. Deste modo, o modelo é capaz de reproduzir a estrutura *scale-free* por possibilitar a criação de *hubs*.

Duas redes artificiais com distribuição de grau distintas são apresentadas na Figura 1, Erdös-Rényi e Barabási-Albert, respectivamente. Nessa figura, tanto a escala de tamanho dos nodos quanto as suas cores são proporcionais ao grau de cada um. Na rede



(a) Erdös Rényi (b) Barabási-Albert

Figura 1 – Reapresentações gráficas de redes Erdös Rényi e Barabási-Albert.

(a), a maioria dos nodos possui grau próximo do grau médio $\langle k \rangle$, com apenas alguns nodos se destacando moderadamente. Na figura (b), onde podemos notar um grande número de nodos de grau baixo relativo aos nodos que se destacam mais proeminentemente. De acordo com o esperado, na primeira rede a distribuição é uniforme, com uma diferença suave entre as cores dos nodos, enquanto na figura seguinte há contraste mais notável na coloração devido à distribuição não homogênea de valores de k.

Na análise de redes complexas, muitas vezes desejamos saber quais são os vértices mais centrais da rede. Exemplos de situações em que este tipo de informação nos é útil são no estudo de como epidemias se alastram por uma população, quais pessoas seriam as mais influentes em uma rede social, ou ainda os pontos que seriam um alvo fácil para causar danos em redes elétricas, de roteadores ou em uma rede criminal. Existem muitas medidas que podem ser utilizadas para classificar um vértice quanto à sua centralidade, uma delas é a medida do grau, que aponta os *hubs* claramente como pontos importantes. Outra medida muito usada para caracterizar a importância de um vértice na rede é a contagem de quantos menores caminhos entre dois quaisquer nodos passam por ele (FREEMAN, 1977), o que pode ser encarado como o quão central ele é frente ao processo de transmissão de informação. Esta medida é chamada Centralidade de Intermediação (*Betweenness*). Neste contexto, pode-se definir um menor caminho como a forma de chegar entre um nodo e outro da rede percorrendo o menor número de nodos possível, e que podem haver inúmeros menores caminhos entre quaisquer par de nodos. A medida de betweenness é matematicamente definida para um nodo qualquer v como

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}, \tag{2.5}$$

onde σ_{st} é o número de menores caminhos entre os nodos s e t, e $\sigma_{st}(v)$ é a quantidade

desses caminhos que passa pelo nodo v. Dentre a abundância de medidas que podem ser feitas em nodos e arestas de uma rede, apresentou-se somente as que foram significativas para o desenvolvimento deste trabalho.

2.2 Redes Modulares

Outro aspecto importante que deve ser levado em consideração para caracterizar e distinguir tanto redes reais quanto artificiais é a estrutura de comunidades de uma rede, também denominadas módulos — conjuntos de nodos mais conectados entre si do que com outros nodos da rede. De fato, existem diversas redes que apresentam estrutura de comunidades, como a rede de colaborações entre cientistas no Instituto Santa Fe (Novo México, EUA), onde as comunidades são formadas por pesquisadores da mesma área de estudo que raramente transitam entre diferentes áreas (GIRVAN; NEWMAN, 2002) (Figura 2). Também podemos citar novamente a rede de crime organizado no Brasil, que possui comunidades bem definidas (CUNHA; GONÇALVES, 2018), assim como a rede de escândalos de corrupção no Brasil (RIBEIRO et al., 2018).

Considerando a importância da estrutura de comunidades para o estudo de redes complexas, foram desenvolvidos algoritmos para detectar estas estruturas. Girvan e Newman (GIRVAN; NEWMAN, 2002) desenvolveram um método baseado na intermediação de arestas, mas sua aplicação é somente viável para redes com até 1000 nodos. Para quantificar a qualidade da divisão de comunidades para redes em que esta estrutura não é previamente conhecida, novamente Newman e Girvan (NEWMAN; GIRVAN, 2004) criaram uma quantidade chamada modularidade, que posteriormente tornou-se base para métodos eficientes de detecção de comunidades (CLAUSET; NEWMAN; MOORE, 2004; NEWMAN, 2006).

Dada certa distribuição de conexões, a medida de modularidade quantifica a probabilidade dela existir em uma distribuição aleatória de conexões. Considerando uma rede representada por sua matriz de adjacência A_{ij}, contendo m conexões, o valor esperado de conexões entre os nodos i e j em uma rede aleatória é dado por $k_i k_j / 2m$, definimos a modularidade Q como sendo (CLAUSET; NEWMAN; MOORE, 2004):

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \tag{2.6}$$

na qual c_i é a comunidade designada ao nodo i e, então, $\delta(c_i, c_j) = 1$ se os nós i e j pertencem à mesma comunidade e 0 caso contrário. O valor de Q usualmente varia entre 0 e 1, embora redes com mais de uma conexão entre nodos podem ter valores diferentes. É considerado que $Q = 0$ indica que a rede possui estrutura de comunidades aleatória e 1 se a divisão escolhida é ótima (NEWMAN; GIRVAN, 2004). Algoritmos de detecção de comunidade a partir da modularidade buscam encontrar o valor máximo de modularidade dadas

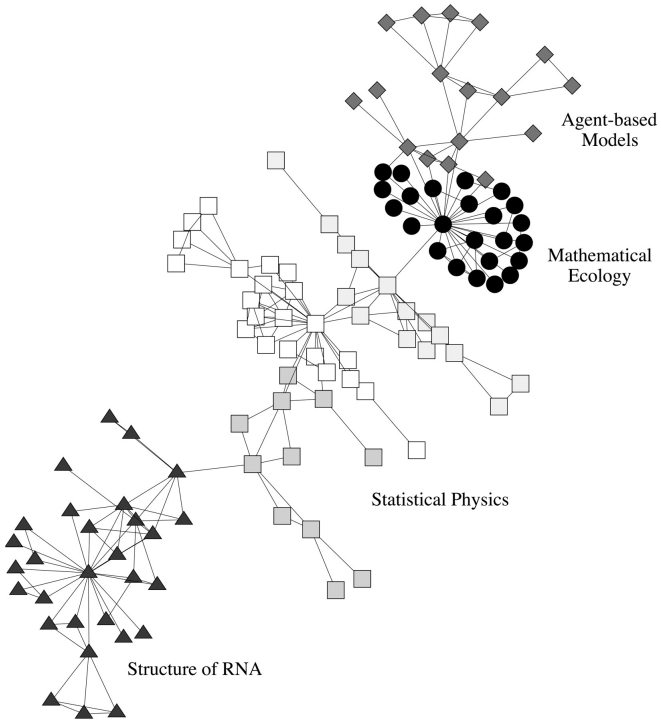


Figura 2 – Rede modular de colaboração científica entre pesquisadores do Instituto Santa Fe (EUA), que mostra a relação entre as diferente áreas de pequisa dentro do instituto. Fonte: (GIRVAN; NEWMAN, 2002)

todas as combinações possíveis de conexões, então este problema é computacionalmente custoso, principalmente para redes grandes. Neste trabalho foi utilizado o algoritmo de Louvain (BLONDEL et al., 2008), que realiza uma boa performance considerando as limitações do problema. Deve ser notado que, recentemente, foi desenvolvido o algoritmo de Leiden (TRAAG; WALTMAN; ECK, 2019), o qual lida com alguns problemas apresentados pelo método de Louvain, como a possibilidade de produzir comunidades desconectadas. Porém, para baixos valores de μ, i.e. comunidades bem definidas, a diferença entre eles é negligenciável.

2.3 Fragmentação de Redes

Um *ataque* a uma rede pode ser definido como a remoção dirigida de vértices ou conexões específicas, ao contrário do que seriam considerados falhas intrínsecas do sistema. O termo remoção pode significar tanto a remoção própria de nós ou arestas quanto a sua desativação, seja ela definitiva ou temporária. Redes aleatórias com distribuição de grau homogênea são facilmente fragmentadas a partir de falhas locais, ou ataques aleatórios – problema que é estudado a partir da teoria de percolação. Redes com distribuição *scale-free* são especialmente vulneráveis a ataques dirigidos a nodos de maior grau e robustas frente a remoções aleatórias (ALBERT; JEONG; BARABÁSI, 2000). Redes reais raramente se encaixam em modelos pré-definidos de redes, fazendo assim necessário um estudo mais

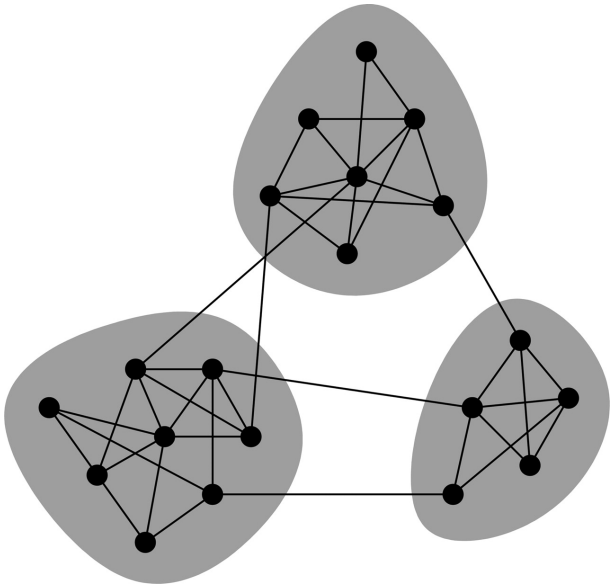


Figura 3 – Ilustração de uma rede com estrutura modular. Existem 3 módulos, ou comunidades, de nodos com mais conexões entre si do que com o restante da rede. Figura retirada de (NEWMAN, 2006).

complexo de suas fragilidades. Um ataque perfeito a uma rede deve destruí-la a ponto da rede deixar de ser funcional, removendo o menor número possível de nodos.

Nos últimos anos, muitos métodos têm sido propostos para aproximar-se o máximo possível do conjunto de desmantelamento ideal — o menor conjunto de nodos que, se retirados, desmantelariam totalmente uma rede (TIAN et al., 2017; REN et al., 2018; KITSAK et al., 2010; WANG et al., 2016; ZDEBOROVÁ; ZHANG; ZHOU, 2016). Em geral, métodos sub-ótimos são baseados em diferentes *rankings* heurísticos e são computacionalmente viáveis, visto que o problema de encontrar o conjunto de desmantelamento ideal é de complexidade *NP-hard* e não possui solução analítica.

Por outra via, foram desenvolvidos alguns algoritmos que buscam encontrar o conjunto ideal de desmantelamento a partir de diferentes abordagens analíticas. Entre estes, pode-se destacar o trabalho de Morone e Makse (MORONE et al., 2016), que propõe-se a identificar o conjunto de "influenciadores ótimos" de grafos aleatórios através do mapeamento deste problema de percolação ótima. Além disso, eles introduzem a medida de Influencia Colectiva, em inglês *Collective Influence* (CI), um algoritmo para lidar com o problema de influenciadores ótimos para grandes conjuntos de dados. Como principal descoberta destaca-se o resultado contraintuitivo de que os "nodos fracos", como chamados pelos autores, seriam muitas vezes nodos de grau baixo ultrapassando a influência de *hubs*. A definição matemática da medida de influência coletiva dá-se por

$$\mathrm{CI}_\ell(i) = (k_i - 1) \sum_{j \in \partial \mathrm{Ball}(i,\ell)} (k_j - 1), \tag{2.7}$$

onde o índice de um nodo qualquer i depende do seu próprio grau e do grau dos seus

vizinhos que se encontram na fronteira de uma bola de raio ℓ ao redor do nodo i.

Seguindo no caminho da procura pelo conjunto ótimo de desmantelamento, Braunstein *et al.* (BRAUNSTEIN et al., 2016) apresentaram um procedimento analítico para determinar de forma exata este conjunto; o qual é baseado em um algoritmo Min-Sum para desmanchar os ciclos, seguido de ataques às estruturas de árvore restantes. O algoritmo é válido em grafos aleatórios com um grande número de nodos e distribuição de grau predefinida. Além disso, considerando CI como o estado da arte, os autores comparam seu método com a estimativa obtida por esta métrica e encontraram resultados que ultrapassam o mesmo. No entanto, seus resultados analíticos não são válidos para grafos gerais, com muitos ciclos pequenos.

Recentemente, Wandelt *et al.* apresentaram uma comparação detalhada e exaustiva do desempenho de 13 métodos de ataque a redes bem estabelecidos com base em vários critérios, incluindo todos os métodos utilizados no presente trabalho (WANDELT et al., 2018). Eles constataram que em geral o melhor método para infringir dano a redes é a estratégia baseada na medida de *betweenness* de cada nodo, pecando apenas no quesito de tempo de processamento, que escala de forma cúbica com o crescimento da rede. Ainda de acordo com o trabalho de Wandelt *et al.*, o método de ataque iterativo baseado na versão aproximada de (*betweenness*) também obteve um ótimo desempenho, aproximando-se muito do método HBA e possui a vantagem de ser processado em tempo quadrático. É importante ressaltar que os conjuntos de nodos selecionados por estes dois métodos nem sempre se interseccionam, mesmo que o ataque tenha obtido resultados semelhantes.

Como mencionado na sessão anterior, uma variedade de redes reais tende a formar comunidades, o que afeta sua robustez. Foi mostrado recentemente que, para redes altamente modulares (CUNHA; GONZÁLEZ-AVELLA; GONÇALVES, 2015), pontes conectando comunidades podem ser muito relevantes para a coesão da rede. Nesse sentido, Dong *et al.* (DONG et al., 2018) mostraram que as redes modulares tendem a se tornar mais resilientes à medida que aumenta-se a fração de nodos que conectam comunidades. Os autores mostram matematicamente que o efeito de aumentar a fração de nós conectando comunidades distintas afeta a transição de fase de percolação da mesma maneira que um campo externo afeta uma transição de fase ferromagnética-paramagnética em sistemas de spin.

Muitas redes empíricas apresentam, em nível microscópico, um grande número de triângulos (NEWMAN, 2010), o que não se encaixa na suposição do algoritmo CI de estrutura de árvore local. Em nível mesoscópico, estas redes podem apresentar estrutura de comunidades, porém a relação entre o algoritmo CI e a existência de comunidades não é bem definida. Neste sentido, Kobayashi e Masuda (KOBAYASHI; MASUDA, 2016) combinaram o algoritmo CI com *coarse graining*, onde um nodo representa uma comunidade, para propor um método de fragmentação de redes modulares visando encontrar os influenciadores

coletivos no nível mesoscópico, e assim provando superar o algoritmo original para redes modulares a um nível de custo computacional razoável.

Redes modulares interdependentes, ou redes de redes, são conhecidas por serem muito mais vulneráveis do que grafos comuns, e rupturas localizadas podem dar origem a poderosas avalanches em redes elétricas, redes biológicas e financeiras (CUNHA; GONZÁLEZ-AVELLA; GONÇALVES, 2015; SHEKHTMAN; DANZIGER; HAVLIN, 2016; SHEKHTMAN; DANZIGER; HAVLIN, 2018). Transições de fase neste tipo de rede indo de uma fase conectada a uma fase fragmentada são mais abruptas. Tendo isso em vista, Shekhtman *et al.* desenvolveram um arcabouço teórico para a ruptura de redes modulares interligadas. Os autores mostraram que tais sistemas podem passar por uma dupla transição de fase de primeira ordem — fragmentações inter e intra-módulos — quando há poucos módulos densamente intra-conectados, ou uma única quando há muitos módulos altamente interconectados (SHEKHTMAN; SHAI; HAVLIN, 2015).

Na próxima sessão serão apresentados os métodos utilizados para a realização do presente trabalho.

3 Métodos

3.1 Análise de Robustez

Este trabalho tem por objetivo principal encontrar os n nodos que fazem o maior dano possível em redes de tamanho N. O método denominado neste trabalho como *força bruta* consiste em gerar todas as possíveis $\binom{N}{n}$ listas de nodos, então remover da rede original os nodos de cada lista, contabilizando o dano produzido pela remoção de cada uma, a fim de encontrar qual, dentre todas elas, é o conjunto de nodos que diminui ao máximo a maior componente conectada da rede. Esse conjunto — ou os conjuntos, no caso de ter mais de um equivalente — é chamado de conjunto ótimo de nodos. Por uma questão de clareza, chamamos de S^* o tamanho da maior componente conectada resultante de um ataque ótimo. Portanto, para um número fixo de nós removidos de n, o menor tamanho de componente conectada possível é dado por S_n^*, que é o ponto de fragmentação ideal.

Além de encontrar o melhor ataque, temos acesso a toda distribuição de possíveis ataques a cada rede, podendo comparar com os resultados de ataques heurísticos, agora possuindo a informação de onde eles se posicionam na distribuição. Este tipo de abordagem gera um resultado exato e somente é possível realizá-la para sistemas de tamanho pequeno devido ao grande custo computacional. Os métodos heurísticos de fragmentação tentam se aproximar do conjunto ótimo com base em diferentes critérios de importância ou de centralidade. Eles podem chegar perto, ou não, do melhor ataque possível dependendo das características da rede em que são aplicados, pois se baseiam em medidas que podem variar drasticamente para grafos de diferentes topologias.

Os três métodos usados como comparação são HBA, MBA e CI, descritos em maior detalhes a seguir:

- **HBA** – High Betweenness Adaptive (Ataque adaptativo por alta intermediação): É baseado na medida de intermediação (*betweenness*) apresentada na Seção 2.2. Neste método calcula-se o valor de intermediação para cada nó da rede e remove-se o de maior valor, então repete-se esse passo o quanto desejado. Portanto, este é dito um método adaptativo, pois recalcula-se a medida depois de cada remoção, este aspecto torna o método muito mais custoso computacionalmente que sua versão não adaptativa. Esta estratégia está entre as mais efetivas desenvolvidas até então, porém é a de maior custo computacional (WANDELT et al., 2018), tornando-o difícil de ser realizado para redes muito grandes.

- **MBA** – Module-Based Attack (Ataque Baseado em Módulos): Neste método, o

primeiro passo é utilizar um algoritmo de detecção de comunidades, então seleciona-se somente os nodos que fazem parte de ligações entre comunidades, que chamam-se "pontes". Dentre estes são retirados os de maior betweenness. Neste trabalho foi utilizada a versão não adaptativa deste método, ou seja, o betweenness é calculado somente uma vez, diminuindo drasticamente o custo de computacional.

- **CI** – Collective Influence (Influência Coletiva): Neste método, os nodos são removidos de acordo com seu nível de Influência Coletiva (CI). Este ataque, quando feito de forma adaptativa, também pode ser muito custoso computacionalmente pelo tempo levado para calcular esse índice para cada nodo, principalmente em redes com milhares de nodos. Para tornar o método possível de ser aplicado, utiliza-se uma versão adaptada para calcular o índice em tempo linear de $O(N \log N)$ (MORONE et al., 2016).

3.2 Algoritmo LFR

Para realizar este experimento é necessário obter um conjunto de redes artificiais que possuam diferentes valores de modularidade e que mantenham outras características fixas. Tal conjunto de redes pôde ser produzido utilizando o algoritmo Lancichinetti–Fortunato–Radicchi (LFR) (LANCICHINETTI; FORTUNATO; RADICCHI, 2008), que foi construído para gerar redes *benchmarks* para testar algoritmos de detecção de comunidades. Tanto a distribuição de grau quanto a de tamanho das comunidades são uma lei de potência com expoentes modificáveis γ e β, respectivamente. Os expoentes padrões são: $\gamma = 2$ e $\beta = 1$. É atribuído a cada nodo um grau proveniente de uma distribuição de expoente γ, e os links são feitos a partir do modelo de configuração (MOLLOY; REED, 1995) para manter a distribuição de grau. Cada nodo possui uma fração μ de conexões com nodos fora da sua comunidade, e uma fração de $\mu - 1$ conexões dentro da comunidade, este parâmetro é o fator que nos permite controlar a modularidade das redes. Para gerar uma rede com o algoritmo LFR é necessário especificar três parâmetros: número de nodos N, grau médio $\langle k \rangle$, grau máximo k_{max} e parâmetro de mistura μ.

3.3 Conjuntos de Redes

Através do algoritmo LFR geramos as 194 redes que compõem o conjunto principal neste estudo. Os parâmetros usados foram: tamanho $N = 100$, grau médio $\langle k \rangle = 3$ e grau máximo $k_{max} = 6$. Usando tais parâmetros fixos foi possível variar o parâmetro μ no intervalo de 0.08 à 0.6, para gerar redes com modularidade entre 0.53 e 0.85. Outros dois conjuntos foram criados com $\langle k \rangle = 4$ (62 redes), e $\langle k \rangle = 6$ (64 redes), com o intuito de estudar o efeito do $\langle k \rangle$. Também geramos um número menor de redes de tamanhos

$N = 80$ (34 redes) e $N = 120$ (30 redes), com os mesmos parâmetros do conjunto principal ($N = 100$, $\langle k \rangle = 3$), para estudar efeitos de tamanho.

Adicionalmente, foram usadas duas redes reais de domínio publico e tamanho comparável com as geradas aqui, tendo como objetivo analisar se os resultados em redes *benchmark* se aplicam aos exemplos reais.

4 Resultados

Nesta seção será apresentada a análise dos resultados obtidos através da aplicação de ataques por força bruta em redes previamente descritas na Seção 3.3, assim como a comparação com métodos baseados em medidas heurísticas. A exposição dos resultados é primeiramente focada nos casos de tamanho $N = 100$ e $n = 5$ remoções, que foram o principal objeto deste estudo. Subsequentemente, são mostradas pequenas variações das redes no tamanho, número de nodos removidos e grau médio. Também são apresentados algumas aplicações do método de força bruta em duas redes reais.

Em um ataque por força bruta, todas as possíveis escolhas de remoção de n nodos de uma rede de tamanho N, $\binom{N}{n}$, são testadas, medidas, organizadas e comparadas. Na prática, o crescimento exponencial do combinatorial N^n, com n fixo, torna inviável o ataque por força bruta até em redes não muito grandes. Por exemplo, aplicar esse tipo de ataque removendo 5% dos nodos de uma rede de $N = 200$, representa avaliar $\approx 2 \times 10^{16}$ combinações diferentes. Portanto, a escolha do conjunto principal para esta dissertação foi de $N = 100$, com remoção de 5% de nodos, que resulta em $\binom{100}{5} \approx 7.5 \times 10^7$ conjuntos diferentes para cada uma das redes analisadas. O tempo de processamento desta quantidade de remoções em uma CPU com processador Intel® Core™ i5-4460 é de, em média, 1h 45min; já para o caso anterior, $\binom{200}{10}$, o tempo chegaria na ordem de 10^8 horas, aproximadamente 11 milênios. A Figura 4 ilustra um exemplo de remoção do conjunto ótimo de 5 nodos em uma rede de tamanho $N = 100$ do conjunto utilizado neste trabalho. A remoção destes nodos reduz a maior componente conectada da rede ao máximo possível, para este caso $S^* = 47$. A rede antes da remoção é apresentada na Fig. 4(a), onde o conjunto ótimo de nodos está destacado, já a Fig. 4(b) ilustra o que ocorre após o ataque ótimo que separa a rede em três componentes, sendo a componente em destaque a maior delas. Este ataque é somente um de todos os possíveis para a mesma rede.

Para cada ataque por força bruta realizado, o tamanho da maior componente remanescente S é medido após a retirada de cada um dos possíveis conjuntos de 5 nodos. Os resultados podem ser ordenados na forma de um histograma, com barras que representam a frequência em que um certo valor de maior componente conectada foi obtido. Tais histogramas, ou distribuições de frequências dos resultados dos ataques, são os principais resultados deste trabalho; são a base para as análises de interesse e serão descritos nesta seção. Exemplos destas distribuições são apresentados na Figura 5 para três redes com valores de modularidade diferentes – baixa ($Q \approx 0.59$), média ($Q \approx 0.76$) e alta ($Q = 0.83$). As três figuras também apresentam linhas verticais posicionando os três ataques heurísticos e a média da distribuição.

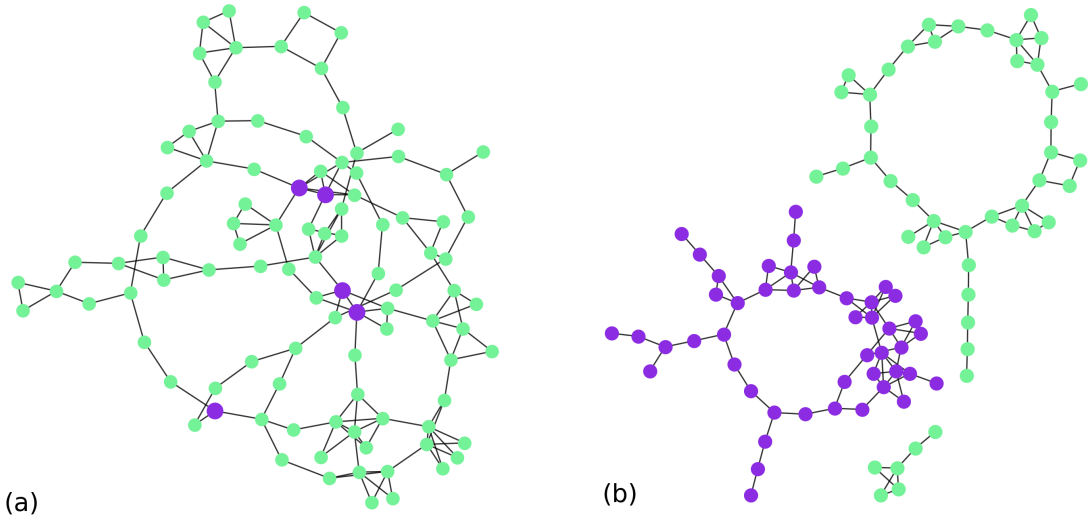


Figura 4 – Exemplo de ataque ótimo em uma rede de tamanho $N = 100$. (a) Nodos do conjunto ótimo aparecem destacados em roxo. (b) Após a remoção desses nodos, a rede está separada em três componentes, a maior componente, com 47 nodos, também aparece destacada em roxo, com as outras duas componentes tendo tamanhos 42 e 6 nodos.

Primeiramente, na Figura 5(a), é mostrada a distribuição de ataques em uma rede de modularidade baixa, $Q \approx 0.59$, assim como a localização dos ataques heurísticos, também chamados de dirigidos, que coincidem com a média, estando todos sobrepostos na mesma linha. É possível notar que esta rede é robusta, visto que nenhuma remoção diminui sua maior componente conectada a menos que 83, seu ponto ótimo S^*. A Figura 5(b) apresenta uma rede de modularidade $Q \approx 0.76$, na qual o melhor ataque, apenas um dos quase 10^8 possíveis, deixa a maior componente da rede com $S^* = 51$ nodos. Esta rede é consideravelmente menos robusta que a anterior, pois a remoção de 5% dos nodos a quebra de forma que o maior fragmento possui 50% do tamanho original. Por outro lado, o ataque médio resulta em praticamente nenhum dano, pois se encontra em $\langle S \rangle = 93$, enquanto os ataques heurísticos se mostram mais eficientes neste caso, com o HBA chegando a $S_{HBA} = 56$, MBA mais longe com $S_{MBA} = 77$ e, por último, CI com $S_{CI} = 83$.

Por último, na Figura 5(c), temos a distribuição de ataques de uma rede com modularidade alta, $Q = 0.83$. Podemos notar prontamente a diferença com as duas distribuições anteriores: a distribuição é aproximadamente uniforme na região $35 < S < 75$, mesmo que em escala logarítmica, o que eleva consideravelmente a probabilidade de obter uma grande redução da maior componente conectada. Essa é uma grande diferença em relação à primeira distribuição apresentada, que concentra praticamente todos ataques em $S = 95$, o pior ataque possível; ou com a segunda distribuição, que também apresenta 10^7 ataques em uma região de dano quase nulo. Além disso, o melhor ataque possível na rede de alta modularidade situa-se em $S^* = 20$, com mais de 10 ocorrências. O método

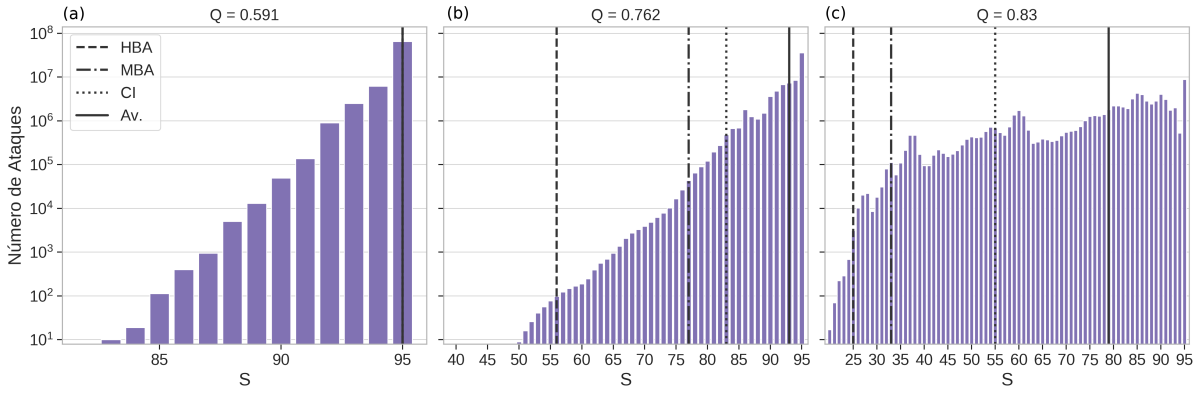


Figura 5 – Três distribuições de ocorrências da maior componente conectada após um ataque por força bruta de retirada de $n = 5$ nodos, onde cada figura corresponde a uma rede distinta de tamanho $N = 100$. São apresentadas redes com modularidade baixa ($Q = 0.591$), média ($Q = 0.762$) e alta ($Q = 0.83$). São também incluídas linhas verticais representando os resultados dos métodos heurísticos e da média de todos ataques.

heurístico HBA continua sendo o mais eficaz, reduzindo a maior componente a $S_{HBA} = 25$, mas MBA fica muito próximo, $S_{MBA} = 33$. Por outro lado, CI situa-se em $S_{CI} = 55$, e a média de todos os ataques possíveis está em $\langle S \rangle = 79$. Com apenas 5% de nodos removidos é possível quebrar a rede de $Q = 0.83$ em fragmentos não maiores do que 20% do tamanho original, entretanto com custo computacional já proibitivo se a rede fosse ligeiramente maior. Porém, com métodos heurísticos, pode-se chegar a uma diminuição da maior componente conectada em até 25% do tamanho original, se o método for adaptativo, ou em 1/3 do tamanho utilizando o método MBA, com custo computacional muito menor. Claramente, utilizar um método heurístico com custo de processamento quase nulo é preferível à um ataque aleatório, principalmente neste último caso.

Podem ser feitas três observações a partir da Figura 5: a rede de modularidades alta ($Q = 0.83$) é evidentemente mais frágil que a de modularidade média ($Q = 0.76$), que por sua vez se mostra mais frágil comparada com a rede de modularidade baixa ($Q = 0.59$). Ou seja, quanto maior a modularidade, mais frágil a rede. Quando a modularidade é alta, a rede se apresenta frágil inclusive frente a ataques sem estratégia nenhuma, visto que um ataque aleatório a 5% dos nodos pode tranquilamente reduzir a rede a 80% do seu tamanho. A terceira observação, que trata especialmente da Figura 5(c), é que, contrariamente ao esperado pela teoria, o método CI não apresenta um bom desempenho neste caso; está longe do ótimo e inclusive abaixo do método MBA. Isso deve-se provavelmente à estrutura modular e com possíveis *loops* das redes pequenas.

Outra forma de analisar a grande diferença no ataque por força bruta entre níveis distintos de modularidade é observar a probabilidade cumulativa de obter um valor menor ou igual a certo S. Estes resultados são ilustrados na Figura 6, onde foram consideradas todas as redes do conjunto de $\langle k \rangle = 3$ e $N = 100$. As redes foram separadas em três

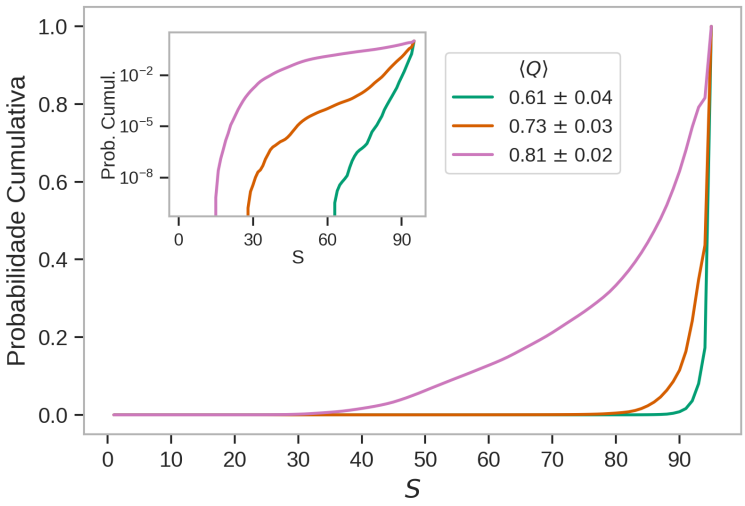


Figura 6 – Probabilidade cumulativa de que se obtenha um dado valor de S depois de um ataque. Cada curva corresponde a um subconjunto de todas as 194 redes agrupadas em torno dos valores médios de modularidade $\langle Q \rangle = 0.61$, 0.73 e 0.81. A figura interna apresenta as mesmas curvas em escala logarítmica. As médias e desvios-padrões são correspondentes às redes dentro dos 3 grupos.

grupos, cada um abrangendo 1/3 do intervalo total de Q. Eles se encontram nos intervalos de modularidade: [0.57, 0.66] (39 redes), [0.66, 0.756] (78 redes), e [0.756, 0.85] (77 redes). Esta figura reforça as observações sobre a Figura 5 em relação à robustez das redes de forma contrária à modularidade. Nesta representação, quanto menor o valor de S em que a probabilidade deixa de ser nula, mais frágil é o conjunto de redes correspondente, e quanto mais tardia e abrupta a subida, mais robustas serão as redes. Pode-se notar que, por exemplo, para o conjunto com $\langle Q \rangle = 0.61$, nenhum dos ataques é capaz de quebrar as redes pela metade, enquanto para $\langle Q \rangle = 0.81$, isso acontece em 6% dos ataques. Outra diferença entre estes grupos é que a mediana dos ataques dos dois grupos de menor modularidade se situam no valor de $S = 93$, ou seja, aproximadamente metade de todas as remoções não fazem efetivamente qualquer dano às redes; porém, para modularidade alta a mediana está em 83. Além disso, para este último grupo, o primeiro quartil situa-se em $S = 72$, diferentemente dos outros dois grupos, que ainda os 25% das remoções que causam o maior dano se encontram abaixo de 90.

A Figura 8 mostra a relação entre o tamanho da maior componente conectada em função da modularidade das redes frente à ataques por força bruta de tamanho $n = 5$. Também é mostrado o comportamento dos três métodos de fragmentação analisados neste trabalho, assim como o valor esperado de um ataque aleatório. Os pontos nesta figura são médias sobre várias redes em intervalos de $\Delta Q = 0.02$, cuja distribuição é mostrada na Figura 7, e suavizados por média móvel numa janela de dois pontos. Esta figura mostra que a diferença entre a curva $S^*(Q)$, do ataque ótimo, e $\langle S \rangle$, ataque aleatório, se acentua com o

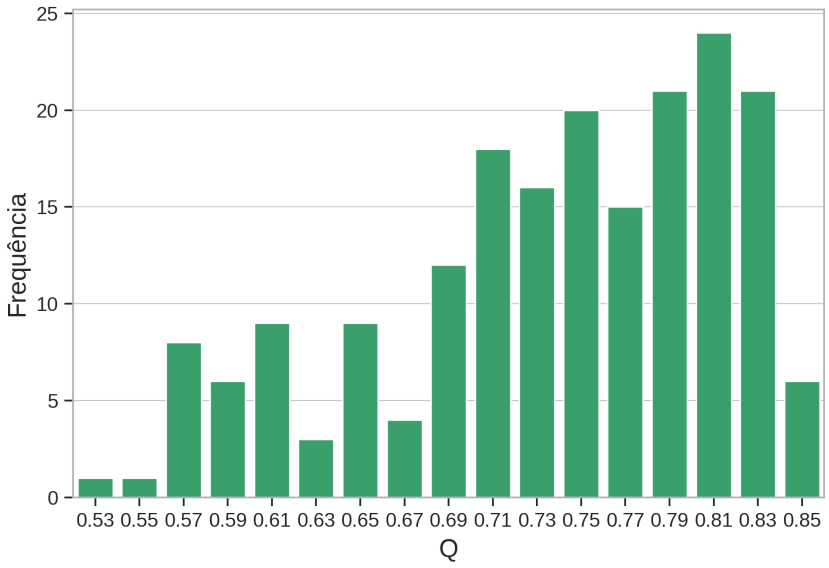


Figura 7 – Distribuição dos valores de modularidade para o conjunto de redes de $N = 100$ e $\langle k \rangle = 3$. Total de 194 redes.

aumento da modularidade, definindo os limites inferior e superior para ataques heurísticos, respectivamente. As três curvas entre estas correspondem aos ataques dirigidos: MBA (linha vermelha e estrelas), HBA (linha verde e triângulos), e CI (linha azul e diamantes). Embora as curvas dos métodos heurísticos e a média dos ataques se separem quando a modularidade aumenta, é importante notar que os valores de S decaem para todos os tipos de ataques, mesmo os aleatórios.

Cada uma das curvas citadas acima pode ser melhor visualizada na Figura 9, onde são apresentadas separadamente e com o desvio padrão correspondente a cada *bin*. As maiores bandas de desvio padrão nos ataques dirigidos são esperadas, visto o número de nodos relativamente pequeno das redes, e essas são mais evidentes a partir de $Q \approx 0.7$. Podemos notar que a inclinação das curvas dos métodos heurísticos aumenta (em valor absoluto) também a partir de aproximadamente $Q = 0.7$, e tende a encontrar a curva ótima, porém o ataque médio não acompanha esta evolução. Pode-se concluir que há uma alta correlação entre $\langle S \rangle$ e Q, assim como uma acentuação na inflexão da curva ótima na região de transição de média à alta modularidade. É possível estimar o ponto de inflexão da curva $S^*(Q)$ através do inverso de $S^*(Q)$, como mostrado na Figura 10. Os pontos foram divididos em dois grupos, e foi realizada regressão linear de $1/S^*$ para cada um deles. A divisão dos pontos foi feita a partir da minimização do erro proveniente das duas regressões lineares. Então, pelo cruzamento das duas retas, o ponto crítico estimado se encontra em $Q_c^* = 0.73$.

A abordagem heurística ao problema de fragmentação de redes mostra dois regimes distintos. Para modularidades menores que o valor crítico Q_c^*, os ataques heurísticos estão

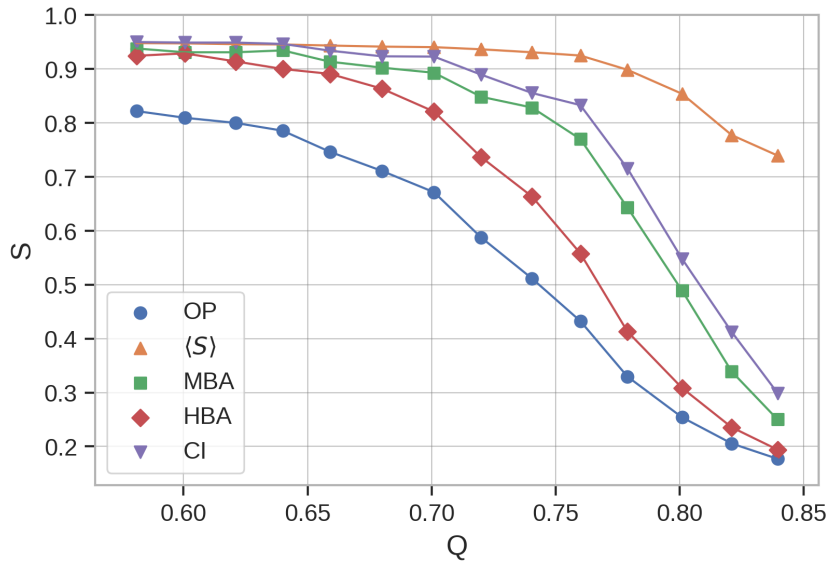


Figura 8 – Tamanho S da maior componente conectada em função da modularidade Q das redes, onde os pontos são calculados em *bins* de $\Delta Q = 0.02$. O eixo ordenado é normalizado pelo número de vértices N, e mostra S do melhor ataque possível OP, da média de ataques $\langle S \rangle$ – o valor esperado para um ataque aleatório –, e dos métodos heurísticos HBA, MBA e CI.

todos juntos, afastados da curva $S^*(Q)$ e bem próximos do ataque médio. Ao se aproximar do ponto de inflexão Q_c^*, os ataques heurísticos se diferenciam entre si e afastam do ataque médio. Passado o ponto de inflexão, as curvas dos ataques heurísticos se distanciam ainda mais da curva média de todos os ataques, e começam a se reaproximar da curva ótima, cada método com seu valor de Q crítico. Como uma forma de localizar o valor crítico Q_c^{heur}, em que os ataques heurísticos se aproximam da curva $S^*(Q)$, calcula-se a razão entre S depois de um ataque dirigido e S^* como função de Q para cada tipo de ataque heurístico, ilustrado na Figura 11. Os pontos máximos encontrados foram: $Q_c^{CI} = 0.78$, $Q_c^{MBA} = 0.78$ e $Q_c^{HBA} = 0.74$; todos estes valores estão destacados como um ponto em preto na Figura do método correspondente. Para $Q < Q_c^*$ (aproximadamente 0.73, vide Fig. 10), os métodos dirigidos acompanham a evolução de S^*, sendo que, nesse regime, qualquer tipo de ataque surte quase o mesmo efeito de um ataque aleatório. No intervalo entre Q_c^* e Q_c^{heur}, o ponto ótimo se distancia das estratégias heurísticas. Por outro lado, para $Q > Q_c^{heur}$, as estratégias heurísticas se aproximam novamente de S^*, consequentemente, mostrando-se muito mais eficientes que um ataque aleatório.

Um aspecto que deve ser considerado é a possibilidade de que a rede tenha uma degenerescência em conjuntos ótimos, ou seja, que possa haver mais de uma lista de nodos que fragmente a rede tanto quanto possível. Na Figura 12 é apresentado o número de diferentes conjuntos de $n = 5$ nodos que, se removidos, reduzem S ao máximo possível. Estes dados são mostrados em função da modularidade, divididos em intervalos de $\Delta Q = 0,02$, as barras de erro são devidas à esta divisão. O número de *sets* que resultam na mesma

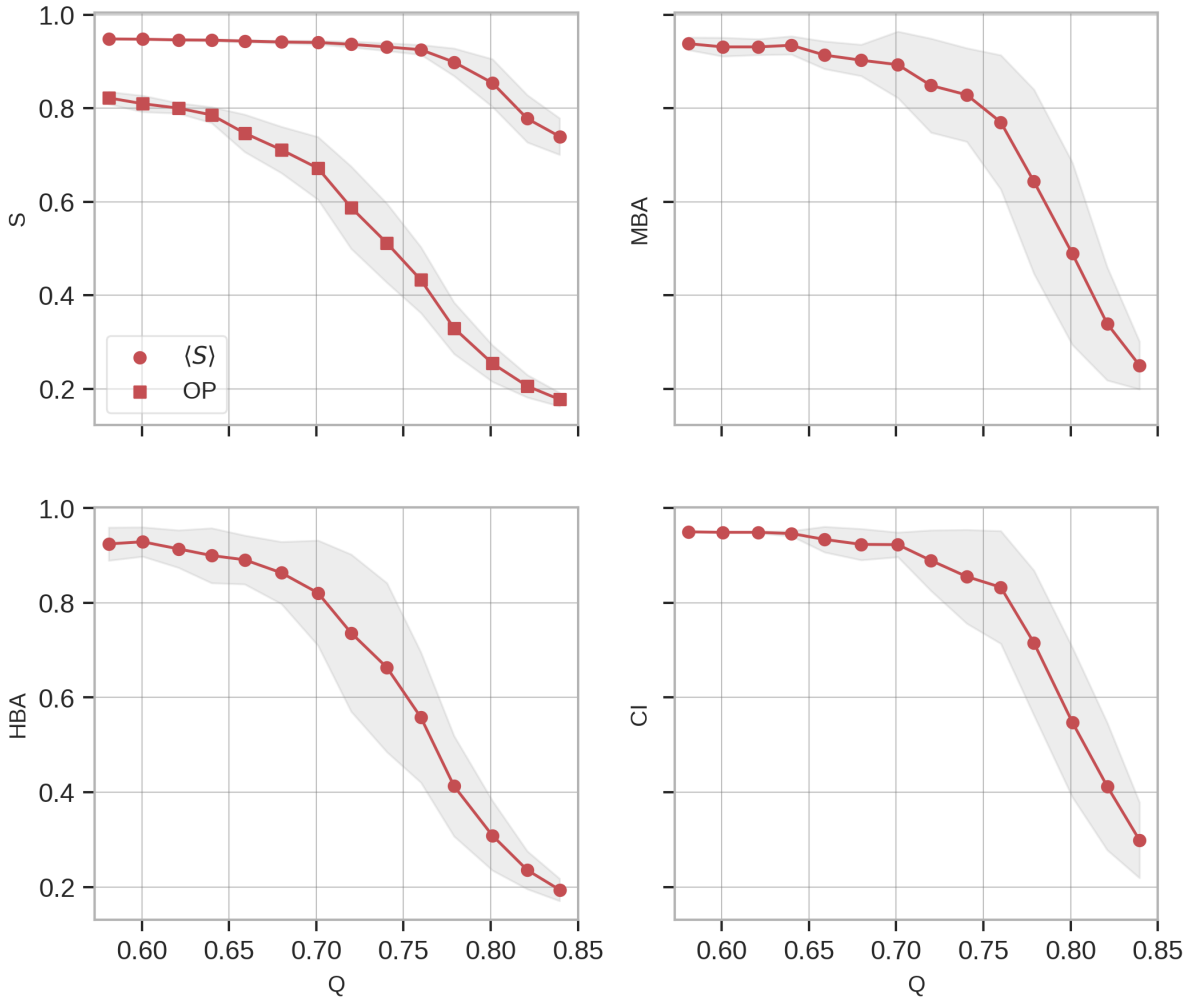


Figura 9 – Expansão da Fig. 8, onde são mostradas as curvas apresentadas anteriormente, mas em quadros separados, incluindo o desvio padrão da aproximação em *bins*.

S_5^* cresce quando a modularidade passa aproximadamente 0.7. Este resultado indica que redes com maior modularidade são mais fracas, não somente porque o melhor ataque é devastador, mas também porque há mais conjuntos de nodos que chegam no mesmo resultado. As quantidades de *sets* ótimos das redes com modularidade a partir de $Q \approx 0.82$ encontram-se na zona que poderia ser considerada de *strong outliers* (valores maiores que 13), o que indica um comportamento estatisticamente diferente do que foi encontrado para outros intervalos de modularidade.

A limitação inevitável do presente estudo frente a complexidade combinatorial, como já mencionado, fez com que fosse necessário escolher o melhor conjunto custo-benefício, considerando a relação entre o custo computacional e a validade dos resultados. A escolha ficou em sistemas de tamanho $N = 100$, remoção de $n = 5$ nodos (5%) e $\langle k \rangle = 3$. Diante dessa restrição, é natural procurar saber o que pode-se esperar ao variar esses três parâmetros.

O primeiro efeito a ser estudado foi o de remover números diferentes de nodos,

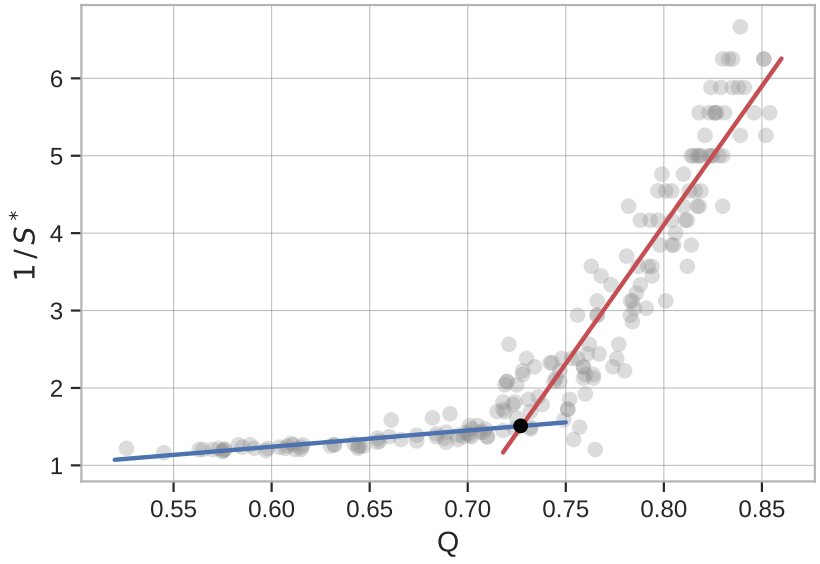


Figura 10 – Inverso do ponto ótimo S^* como função da modularidade. As duas retas mostradas são ajustes lineares em intervalos diferentes de Q, onde encontramos a indicação de um ponto crítico $Q_c \approx 0.73$, onde há uma mudança do comportamento de S^*. Retas possuem coeficiente de correlação 0.92 e 0.77.

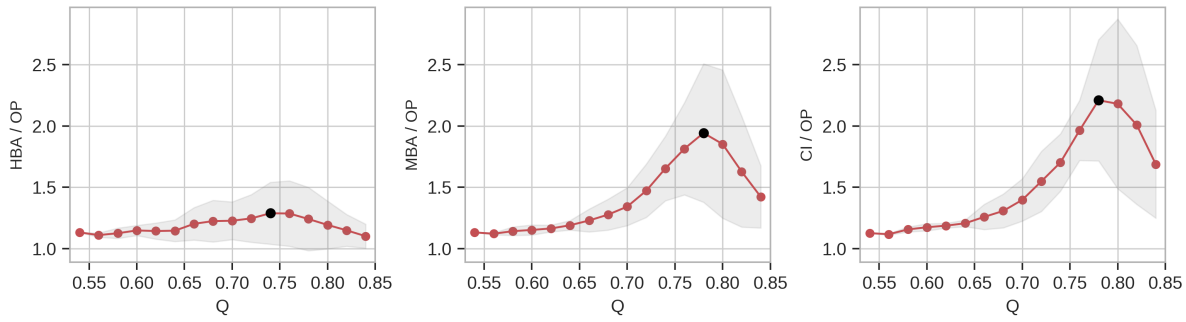


Figura 11 – Razão entre tamanho da maior componente depois de um ataque por método heurístico e do tamanho depois do ataque ótimo em função de Q. Localizou-se o ponto crítico em cada método pelo posição do máximo da relação: (a) $Q_c^{HBA} = 0.74$, (b) $Q_c^{MBA} = 0.78$, e (c) $Q_c^{CI} = 0.78$.

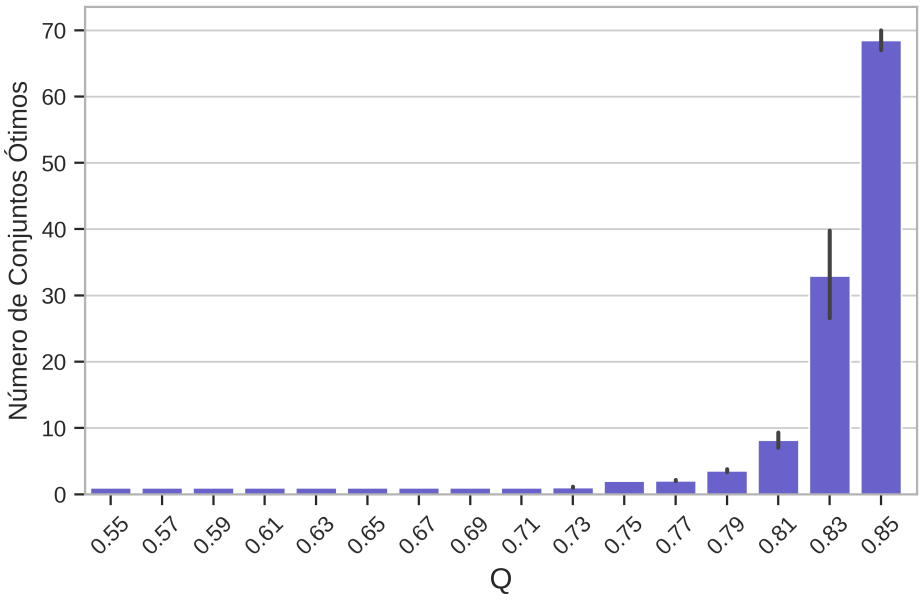


Figura 12 – Número de conjuntos de $n = 5$ nodos que levam ao valor ótimo de S. As barras são a média em intervalos de $\Delta Q = 0.02$ com seu respectivo desvio padrão. Três redes que apresentaram valores maiores que centenas de conjuntos foram descartadas desta imagem.

considerando conjuntos de tamanho $n = 3, 4, 6$ nodos, ou seja, um pouco abaixo e um pouco acima do valor escolhido. A região de $Q > 0.7$ foi utilizada baseado nos pontos de inflexão encontrados anteriormente. Para $n = 6$ foi utilizado um subconjunto de 28 redes devido ao tempo maior de execução. A Figura 13 mostra o comportamento do ponto ótimo S_n^* em função da modularidade. Apesar das curvas de $S_n^*(Q)$ para cada n começarem, como esperado, separadas – quanto menor o número de nodos removidos, menor o dano – com o aumento da modularidade elas começam a convergir na região entorno de $S < 0.3$ em $Q \approx 0.81$. Além disso, é esperada a atenuação na variação de S_n^* em relação a Q à medida que aumenta-se n, visto que quanto mais nodos são removidos o dano é naturalmente maior. Porém, percebemos que tanto o resultado geral da fragilidade como função crescente da modularidade quanto a existência de um valor critico de modularidade em torno de $Q = 0.8$ se confirmam.

A seguir estudamos o tamanho da rede, que pode ser a maior limitação do trabalho, pois a pergunta mais relevante talvez seja como estes resultados podem ser estendidos a sistemas maiores. Na Figura 14 apresentamos uma estimativa modesta de escalonamento de tamanho finito em volta de $N = 100$, mostrando S^* *vs* Q para redes de tamanho um pouco menor ($N = 80$) e um pouco maior ($N = 120$), mantendo fixa a porcentagem de nodos removidos em 5% ($n = 4$ e 6, respectivamente). Claramente, encontra-se a mesma tendência das Figuras 9 e 10. Então, pode-se dizer que os resultados encontrados são razoavelmente robustos. Há, entretanto, outras questões em relação a efeitos de tamanho

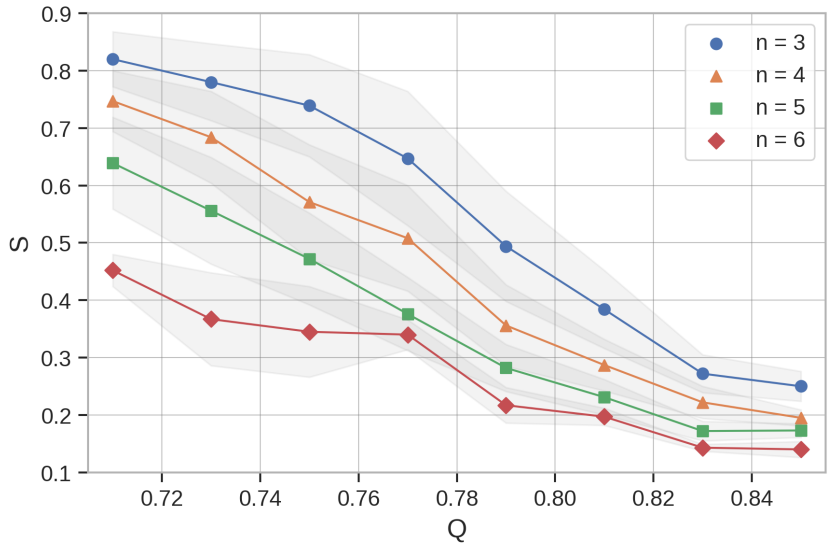


Figura 13 – Tamanho da maior componente conectada correspondente ao ataque ótimo depois de remover $n = 3,\ 4,\ 5$ e 6 vértices das redes de tamanho $N = 100$. Para $n = 6$ foi usado um subconjunto menor de redes. Os ataques estão limitados ao intervalo de médias a altas modularidades, onde observa-se a maior queda de S^*.

finito, como *cutoffs* estruturais da distribuição de grau, impactando a robustez da rede em geral. Neste sentido, o *cutoff* estrutural limita o grau máximo, o que em consequência restringe tanto a correlação de grau quanto o alcance de modularidade que o modelo LFR pode gerar.

Por último, analisamos a sensibilidade dos resultados frente a distribuição de grau. O limite crítico para a existência de uma componente gigante diante falhas ou ataques usualmente depende da distribuição de grau (BARABÁSI, 2016). Para acessar esta provável dependência gerou-se dois pequenos conjuntos de redes (aproximadamente 60 para cada caso) com grau médio $\langle k \rangle = 4$ e $\langle k \rangle = 6$, no mesmo intervalo de modularidade utilizado anteriormente. Foi observado que realmente há influência do grau médio no ponto de modularidade crítica Q_c, e em ambos os casos este valor desloca-se para a direita, com Q mais alto. A relação entre os diferentes Q_c obtidos pode ser observada na Figura 15(b), com valores de 0.71, 0.78 e 0.80 respectivamente para $\langle k \rangle = 3$, 4, e 6. Também foi observada a distribuição de probabilidade cumulativa de resultados de ataques na Figura 15(a). As duas novas curvas encontram-se sempre abaixo da curva de $\langle k \rangle = 3$, novamente indicando uma maior robustez ao mesmo tamanho de ataque. Estas se mostram menos suaves, com mais ruido que a primeira, possivelmente devido ao tamanho reduzido do conjunto de redes. Esse resultado é esperado pelo fato de que, no LFR, o grau médio é a quantidade que controla o número de arestas existentes na rede, então um maior número de arestas indica redes mais robustas independentemente de modularidade.

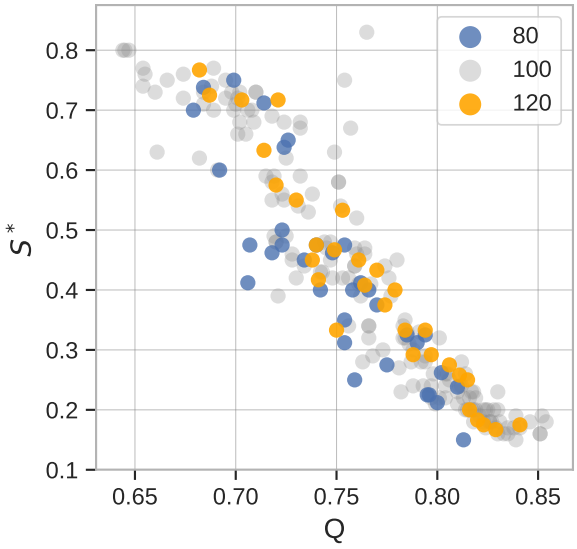


Figura 14 – Ataque ótimo em função da modularidade para diferentes tamanhos de rede (80, 100 e 120), mantendo a proporção de retirar 5% dos nodos.

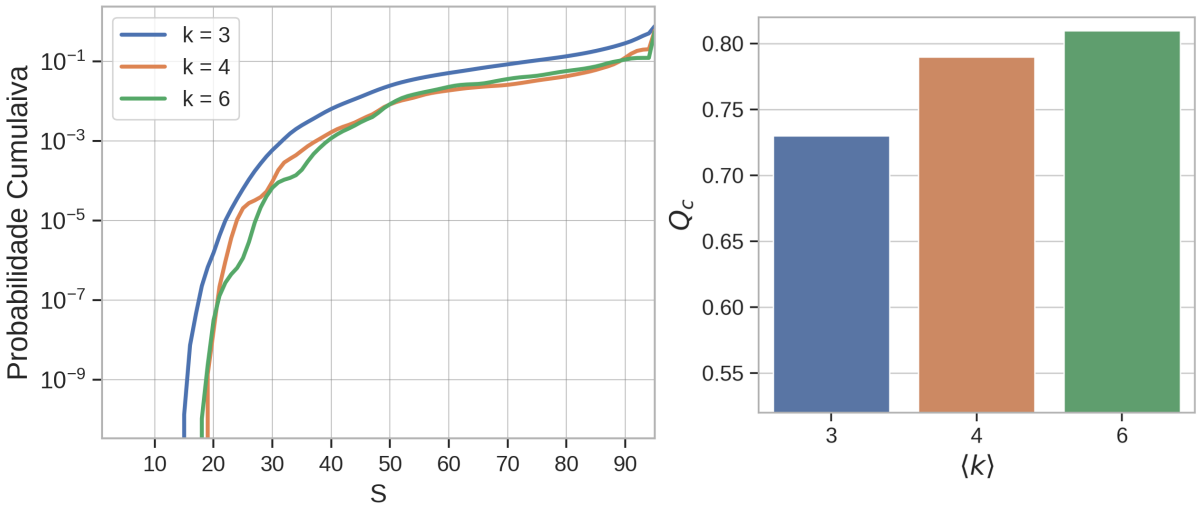


Figura 15 – (a) Probabilidade cumulativa de que, para qualquer rede, seja obtido um dado valor de S em um ataque qualquer, independente da modularidade. São apresentadas 3 curvas para $\langle k \rangle = 3$, 4, e 6. (b) Valores críticos de modularidade Q_c, em que o ponto ótimo começa a decair mais intensamente, para 3 diferentes valores de grau médio. Estes foram obtidos por meio da interpolação de duas retas nos diferentes regimes de S^*. Coeficientes de correlação para cada caso: $\langle k \rangle = 3$, $r_1 = 0.92, r_2 = 0.77$, $\langle k \rangle = 4$, $r_1 = 0.92, r_2 = 0.65$ e $\langle k \rangle = 6$, $r_1 = 0.94, r_2 = 0.72$.

4.1 Redes Reais

Quanto à validade geral dos resultados deste estudo, isto é, a aplicabilidade em redes reais, ela não está garantida, pois as redes estudadas, obtidas por um *benchmark*, foram geradas focando o controle da modularidade, e não necessariamente representam muitos tipos de redes, incluindo redes reais. Então, tendo como motivação essa situação, foram analisadas duas redes reais de tamanho compatível com as redes artificiais estudadas. Aplicou-se tanto o método de força bruta, quanto os três métodos heurísticos: HBA, MBA, e CI. Estas duas redes foram obtidas em (JJATT, 2009) e são formadas por indivíduos envolvidos em ataques terroristas e suas relações.

Na Figura 16 são apresentados resultados do ataque por força bruta em uma rede da organização fundamentalista islâmica Al Qaeda, obtida a partir de investigações entre os anos 1993 e 2003. Sua maior componente conectada possui $N = 218$, $\langle k \rangle = 5.5$, e $Q = 0.82$; na qual foram realizadas remoções de $n = 4$ nodos, devido ao maior tamanho da rede. Os resultados da Figura correspondem à distribuição de tamanhos da maior componente remanescente depois de serem realizados todas as possíveis remoções de 4 nodos, assim como os resultados dos ataques por métodos heurísticos. Nesta distribuição, 65% deles estão situados no maior valor de S possível, $S = 214$, e somente 1% dos ataques podem fazer qualquer dano abaixo de $S = 200$. Por outro lado, todos os ataques heurísticos se encontram dentro desses 1%: $S_{MBA} = 195$, $S_{CI} = 190$, $S_{HBA} = 157$, sendo que $S^* = 142$, e mais uma vez o método HBA é o que mais se aproxima do ataque ótimo.

A Figura 17 é outro exemplo de rede real de terrorismo, desta vez com tamanho bem próximo do estudado nesta dissertação; com $N = 108$, corresponde ao levantamento de investigações de uma série de ataques terroristas na Indonésia em 2005 (JJATT, 2009). Ela possui uma distribuição de ataques peculiar, visto que é claramente bimodal, com uma lacuna ou *gap* considerável na região intermediária. Todavia, 99.8% dos ataques estão localizados no segundo "grupo", entre $S = 83$ e o valor máximo $S = 103$. Já os ataques heurísticos, HBA e MBA, fazem uso dessa divisão da rede, com ataques caindo no primeiro grupo: $S_{MBA} = 55$, e $S_{HBA} = 57$, apesar da baixa modularidade da rede $Q = 0.52$. Notavelmente, nesta rede real, o método CI mostra-se ineficiente ($S_{CI} = 101$), tanto quanto um ataque aleatório.

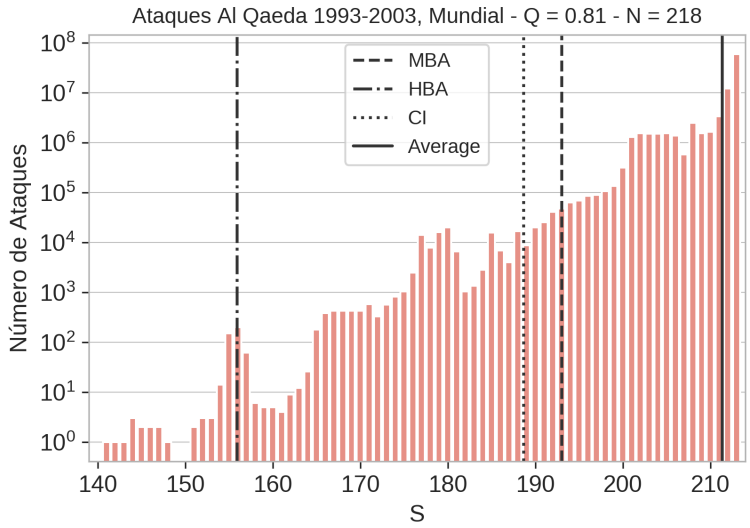
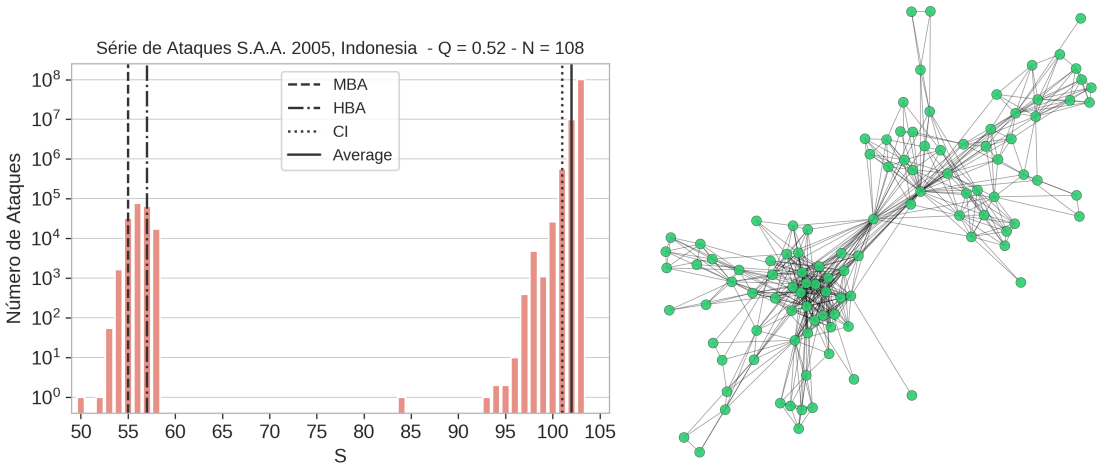


Figura 16 – Rede de ataques terroristas da Al Qaeda entre os anos de 1993 e 2003. Tamanho da rede: $N = 218$; número de remoções: $n = 4$; Ponto ótimo: $S^* = 142$; $S_{HBA} = 157$, $S_{MBA} = 195$ e $S_{CI} = 190$; Grau médio: $\langle k \rangle = 5.5$.



(a) Histograma de ataques (b) Representação da rede

Figura 17 – Rede de ataques terroristas no sul da Indonésia em 2005. Tamanho da rede: $N = 108$; número de remoções: $n = 5$; Ponto ótimo: $S^* = 50$; $S_{HBA} = 57$, $S_{MBA} = 55$ e $S_{CI} = 101$; Grau médio: $\langle k \rangle = 10.5$.

5 Conclusões

A noção de que redes modulares são mais frágeis que outras redes não modulares, mas com outras características similares, é uma ideia que paira em trabalhos recentes sobre desmantelamento de redes. Porém, uma constatação precisa e detalhada da relação entre modularidade e fragilidade era algo inexistente até o momento. Esta é justamente a lacuna na literatura que este trabalho se propôs a preencher. Para tanto, foram realizados experimentos computacionais para comparar o quanto métodos heurísticos se aproximam do ponto ótimo de fragmentação S^* para um ataque de tamanho fixo. Devido ao grande custo computacional, foram geradas redes de tamanho $N = 100$ e realizou-se sobre cada uma delas todos os $\binom{100}{n}$ ($\sim 10^7$) possíveis ataques de tamanhos $n = 3$, 4, 5, e $n = 6$ (neste último caso, em apenas um subconjunto de redes), sendo o foco do trabalho o valor de $n = 5$ (5% dos nodos). Outras duas questões que foram abordadas são a variação do ponto ótimo com pequenas alterações no tamanho da rede, e, mantendo $N = 100$, a sensibilidade da variação no ponto ótimo para diferentes valores de grau médio $\langle k \rangle$. Todas redes foram geradas utilizando o algoritmo LFR, sendo que o conjunto principal de redes possui $\langle k \rangle = 3$ e se encontram no intervalo de modularidade $0.53 < Q < 0.85$.

Os resultados do presente trabalho mostram claramente que qualquer tipo de ataque, inclusive remover nodos aleatoriamente, possui uma correlação negativa com a modularidade das redes, *e.g.* a modularidade é relacionada inversamente com a robustez. De fato, foi mostrado que, conforme a modularidade aumenta, também se eleva a probabilidade cumulativa de ocorrer um grande dano na rede realizando qualquer ataque. Considerando somente o ponto ótimo, estimou-se o ponto crítico da inflexão da curva que representa o ataque ótimo S^*, o qual localiza-se em $Q = 0.73$. O comportamento da curva ótima também foi comparado com outros dois tamanhos de rede $N = 80$ e $N = 120$, e foi possível observar o mesmo efeito de queda do valor de S^* com a elevação de Q. Outra variação no mesmo sentido foi alterar o número de nodos removidos para as redes de $N = 100$, a mesma tendência surgiu. Porém, para valores de n menores que 5, o tamanho da maior componente remanescente é naturalmente menor, com uma inflexão mais tardia da curva. O contrário ocorre para $n = 6$, que infringe grande dano às redes em todo intervalo de Q analisado neste cenário. Para explorar os efeitos da variação do grau médio sobre o ponto ótimo, a modularidade crítica foi medida para os valores $\langle k \rangle = 3$ e 4. O valor crítico para a inflexão de S^* ainda está presente, porém aumentou conforme o grau médio foi elevado, indicando maior robustez das redes quanto maior o seu grau médio. Portanto, as variações em tamanho, número de nodos removidos e grau médio dão suporte à observação primária de que há uma forte relação inversa do ponto ótimo com a modularidade.

Além disso, em geral, observou-se que uma melhor performance com modularidades

mais altas não é exclusiva do método MBA, que tem como premissa valer-se da existência de módulos bem definidos, mas os métodos HBA e CI também apresentam-na. Por outro lado, para modularidades mais baixas que 0.7, ataques heurísticos apresentaram desempenho próximo ao de ataques aleatórios, e todos eles obtém resultados praticamente equivalentes quando $Q \lessapprox 0.6$. Estes valores de Q são equivalentes aos de redes aleatórias – devido a flutuações – e, nestes casos a única forma de infringir um dano mais significativo é através de força bruta. Portanto, redes nesse intervalo de modularidade estão relativamente protegidas de ataques maliciosos de tamanho limitado, por exemplo, remover 5% dos nodos.

Ao se tratar da performance dos métodos de ataque heurísticos e considerando que foram testados três dos métodos mais relevantes, podemos concluir que HBA é sempre o melhor dentre os três, chegando muito perto do melhor ataque possível para altos valores de modularidade. Espera-se encontrar a mesma tendência em redes de qualquer tamanho, mesmo que o custo computacional seja alto. Esta conclusão está de acordo com os achados de Wandelt *et al.* (WANDELT et al., 2018), mencionados anteriormente. Os autores também apontam que o método que utiliza a versão aproximada de *betweenness* chega a resultados similares, porém removendo conjuntos de nodos diferentes, implicando em uma possível degenerescência no conjunto ótimo de nodos. Este aspecto foi observado no presente trabalho, visto que geralmente há mais de um conjunto de nodos que, se removidos, chegam no valor ótimo de menor componente conectada. Com performance similar à do HBA, mas com custo computacional inferior, os métodos MBA e CI podem ser considerados mais eficientes. Destacando que MBA, sendo não adaptativo, na maioria dos casos se mostra ligeiramente superior ao CI, sendo que este último método deveria dar o ponto ótimo no limite termodinâmico em redes sem ciclos. Portanto, em redes reais, MBA mostra vantagem sobre CI, como ficou evidente no exemplo da rede terrorista de Singapura. Outros métodos, baseados em centralidade de grau, autovalores, entre outros, não foram testados pois há evidência consistente do seu desempenho ser inferior aos três métodos heurísticos aqui estudados (WANDELT et al., 2018).

Até então, métodos heurísticos ainda não haviam sido comparados com listas de remoção ótimas, devido às limitações computacionais mencionadas anteriormente. Neste sentido, a relação entre os métodos de ataque heurísticos e o ponto ótimo conforme varia-se a modularidade é a principal contribuição deste estudo. Além disso, foi possível estimar o ponto crítico em que as curvas heurísticas aproximam-se da curva ótima S^*, com o método HBA se sobressaindo e os métodos MBA e CI aproximando-se de S^* para Q um pouco mais alto. Tal informação podendo servir de guia para a escolha do método mais eficiente de ataque para redes modulares.

Existem ainda algumas limitações do trabalho devido ao custo computacional. A limitação de tamanho também implica em efeitos de tamanho finito, como *degree cut-off*

da distribuição de grau, o que pode impactar a robustez das redes. Contudo, mesmo considerando estas restrições, os resultados aqui expostos mostram a existência de uma vulnerabilidade intrínseca às redes modulares, indicando a modularidade como o calcanhar de Aquiles de redes reais. Assim, este trabalho apresenta um embasamento sólido para resultados prévios de pesquisas com diversas redes que apontavam nessa direção.

Referências

ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Error and attack tolerance of complex networks. *Nature*, v. 406, n. 6794, p. 378–382, 2000. Citado na página 9.

BARABÁSI, A.-L. *Network Science*. University Printing House, Cambridge, United Kingdom: Cambridge University Press, 2016. Citado 2 vezes nas páginas 5 e 26.

BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, n. 5439, p. 509–512, 1999. Citado na página 6.

BELLINGERI, M. et al. Efficacy of local attack strategies on the beijing road complex weighted network. *Physica A: Statistical Mechanics and its Applications*, v. 510, p. 316 – 328, 2018. Citado na página 2.

BERTOLERO, M. A.; YEO, B. T. T.; D'ESPOSITO, M. The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, v. 112, n. 49, p. E6798–E6807, 2015. Citado na página 2.

BLONDEL, V. D. et al. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, v. 2008, n. 10, p. P10008, 2008. Citado na página 9.

BOLLOBAS, B. *Random Graphs*. [S.l.]: Cambridge University Press, 1985. Citado na página 6.

BRAUNSTEIN, A. et al. Network dismantling. *Proceedings of the National Academy of Sciences*, National Acad Sciences, p. 201605083, 2016. Citado na página 11.

BULLMORE, E.; SPORNS, O. The economy of brain network organization. *Nature Reviews Neuroscience*, v. 13, n. 5, p. 336, 2012. Citado na página 2.

CLAUSET, A.; NEWMAN, M. E. J.; MOORE, C. Finding community structure in very large networks. *Physical Review E*, v. 70, p. 066111, Dec 2004. Citado na página 8.

CUNHA, B. R. da; GONÇALVES, S. Performance of attack strategies on modular networks. *Journal of Complex Networks*, v. 5, n. 6, p. 913–923, 2017. Citado na página 2.

CUNHA, B. R. da; GONÇALVES, S. Topology, robustness, and structural controllability of the brazilian federal police criminal intelligence network. *Applied Network Science*, v. 3, n. 1, p. 36, Aug 2018. Citado 3 vezes nas páginas 1, 2 e 8.

CUNHA, B. R. da; GONZÁLEZ-AVELLA, J. C.; GONÇALVES, S. Fast fragmentation of networks using module-based attacks. *PLoS ONE*, v. 10, n. 11, p. e0142824, 2015. Citado 3 vezes nas páginas 2, 11 e 12.

DONG, G. et al. Resilience of networks with community structure behaves as if under an external field. *Proceedings of the National Academy of Sciences*, v. 115, n. 27, p. 6911–6915, 2018. Citado na página 11.

FAN, J. et al. Structural resilience of spatial networks with inter-links behaving as an external field. *New Journal of Physics*, v. 20, n. 9, p. 093003, 2018. Citado na página 2.

FAUST, M.; KENETT, Y. N. Rigidity, chaos and integration: hemispheric interaction and individual differences in metaphor comprehension. *Frontiers in Human Neuroscience*, Frontiers, v. 8, p. 511, 2014. Citado na página 2.

FORTUNATO, S. Community detection in graphs. *Physics Reports*, v. 486, n. 3–5, p. 75 – 174, 2010. Citado na página 2.

FREEMAN, L. C. A set of measures of centrality based on betweenness. *Sociometry*, JSTOR, p. 35–41, 1977. Citado na página 7.

GALLEN, C. L.; D'ESPOSITO, M. Brain modularity: A biomarker of intervention-related plasticity. *Trends in Cognitive Sciences*, Elsevier, 2019. Citado na página 2.

GIRVAN, M.; NEWMAN, M. E. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, v. 99, n. 12, p. 7821–7826, 2002. Citado 2 vezes nas páginas 8 e 9.

HOLME, P. et al. Attack vulnerability of complex networks. *Physical Review E*, v. 65, n. 5, p. 056109, 2002. Citado na página 2.

HU, Y. et al. Local structure can identify and quantify influential global spreaders in large scale social networks. *Proceedings of the National Academy of Sciences*, 2018. Citado na página 2.

JJATT. *John Jay ARTIS Transnational Terrorism Database*. 2009. Disponível em: <http://doitapps.jjay.cuny.edu/jjatt/data.php>. Acesso em: 28/11/2019. Citado na página 28.

KITSAK, M. et al. Identification of influential spreaders in complex networks. *Nature Physics*, v. 6, n. 11, p. 888, 2010. Citado na página 10.

KOBAYASHI, T.; MASUDA, N. Fragmenting networks by targeting collective influencers at a mesoscopic level. *Scientific Reports*, v. 6, 2016. Citado na página 11.

LANCICHINETTI, A.; FORTUNATO, S.; RADICCHI, F. Benchmark graphs for testing community detection algorithms. *Physical Review E*, v. 78, n. 4, p. 046110, 2008. Citado na página 14.

MEUNIER, D.; LAMBIOTTE, R.; BULLMORE, E. T. Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience*, v. 4, p. 200, 2010. Citado na página 2.

MOLLOY, M.; REED, B. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, v. 6, n. 2-3, p. 161–180, 1995. Citado na página 14.

MORONE, F. et al. Collective influence algorithm to find influencers via optimal percolation in massively large social media. *Scientific Reports*, v. 6, 2016. Citado 2 vezes nas páginas 10 e 14.

MURO, M. A. D. et al. Multiple outbreaks in epidemic spreading with local vaccination and limited vaccines. *New Journal of Physics*, v. 20, n. 8, p. 083025, 2018. Citado na página 2.

NEWMAN, M. *Networks: An introduction*. Great Clarendon Street, Oxford, United Kingdom: Oxford University press, 2010. ISBN 0199206651. Citado na página 11.

NEWMAN, M. E.; GIRVAN, M. Finding and evaluating community structure in networks. *Physical Review E*, v. 69, n. 2, p. 026113, 2004. Citado na página 8.

NEWMAN, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, v. 103, n. 23, p. 8577–8582, 2006. Citado 2 vezes nas páginas 8 e 10.

REN, X.-L. et al. Generalized network dismantling. *arXiv preprint arXiv:1801.01357*, 2018. Citado na página 10.

RIBEIRO, H. V. et al. The dynamical structure of political corruption networks. *Journal of Complex Networks*, v. 6, n. 6, p. 989–1003, 01 2018. ISSN 2051-1329. Citado na página 8.

SHEKHTMAN, L. M.; DANZIGER, M. M.; HAVLIN, S. Recent advances on failure and recovery in networks of networks. *Chaos, Solitons & Fractals*, Elsevier, v. 90, p. 28–36, 2016. Citado na página 12.

SHEKHTMAN, L. M.; DANZIGER, M. M.; HAVLIN, S. Spreading of failures in interdependent networks. In: *Diffusive Spreading in Nature, Technology and Society*. [S.l.]: Springer, 2018. p. 397–410. Citado na página 12.

SHEKHTMAN, L. M.; SHAI, S.; HAVLIN, S. Resilience of networks formed of interdependent modular networks. *New Journal of Physics*, v. 17, n. 12, p. 123007, 2015. Citado na página 12.

STAM, C. J. Modern network science of neurological disorders. *Nature Reviews Neuroscience*, v. 15, n. 10, p. 683, 2014. Citado na página 2.

TIAN, L. et al. Articulation points in complex networks. *Nature Communications*, v. 8, p. 14223, 2017. Citado na página 10.

TRAAG, V. A.; WALTMAN, L.; ECK, N. J. van. From louvain to leiden: guaranteeing well-connected communities. *Scientific rReports*, v. 9, 2019. Citado na página 9.

WANDELT, S. et al. A comparative analysis of approaches to network-dismantling. *Scientific Reports*, v. 8, n. 1, p. 13513, 2018. Citado 3 vezes nas páginas 11, 13 e 32.

WANG, Z. et al. Fast ranking influential nodes in complex networks using a k-shell iteration factor. *Physica A: Statistical Mechanics and its Applications*, v. 461, p. 171–181, 2016. Citado na página 10.

ZDEBOROVÁ, L.; ZHANG, P.; ZHOU, H.-J. Fast and simple decycling and dismantling of networks. *Scientific Reports*, v. 6, p. 37954, 2016. Citado na página 10.