

12º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2021

SIMULAÇÃO COMPUTACIONAL PARA ANÁLISE DE RESILIÊNCIA À ATAQUES EM REDES DE INTERAÇÃO DE PROTEÍNAS DE BACTÉRIAS RESISTENTES

ISABELA RODRIGUES NOGUEIRA FORTI¹, RODRIGO HENRIQUE RAMOS², ANDRÉ DI TOMASO³

¹ Graduanda em Tecnologia de Análise e Desenvolvimento de Sistemas, Bolsista PIBIFSP, IFSP, Campus São Carlos, i.forti@aluno.ifsp.edu.br

² Doutorando em Ciência da Computação, Professor do Instituto Federal de São Paulo, Campus São Carlos, ramos@ifsp.edu.br.

³ Doutor em Ciência da Computação, Professor do Instituto Federal de São Paulo, Campus São Carlos, andredt@ifsp.edu.br

Área de conhecimento (Tabela CNPq): 1.03.03.00-6 Metodologia e Técnicas da Computação

RESUMO: A remoção de nós ou arestas em redes complexas pode ser de grande impacto para o sistema como um todo. Simulações de ataques em redes biológicas podem fornecer *insights* importantes sobre as consequências geradas, uma vez que podemos expandir o conhecimento dessas modelagens em busca de aplicações no mundo real. As redes de interação de proteínas (interactomes) de bactérias resistentes são algumas das redes de grande interesse por acadêmicos, uma vez que tais microrganismos podem exercer comportamentos violentos em sistemas biológicos. Portanto, simulações de ataques em interactomes bacterianos revelam informações relevantes sobre a topologia das redes, possibilitando identificar características estruturais que podem, por exemplo, ser exploradas no desenvolvimento de drogas. Neste estudo foram desenvolvidos algoritmos para avaliar a resiliência de nove redes bacterianas e uma rede humana através de ataques aleatórios e intencionais de nós de interesse (hubs). A remoção sequencial dos alvos mostrou que ataques à hubs são mais agressivos para a rede humana do que para as redes das bactérias estudadas. Diferença essa que identificamos estar associada ao clustering de hubs nas redes de bactéria. Além disso, todos os nove interactomes exibiram comportamentos semelhantes, indicando provavelmente a existência de um padrão quanto à resiliência para bactérias deste tipo.

PALAVRAS-CHAVE: rede complexa; interactome; ataque; resiliência; bactéria;

ANALYSIS OF RESILIENCE TO ATTACKS IN NETWORKS OF PROTEIN INTERACTIONS FROM RESISTANT BACTERIA

ABSTRACT: The removal of nodes or edges in complex networks may be of great impact for the system as a whole. Simulations of attacks in biological networks can provide important insights about the generated consequences, once we can expand the knowledge of these modellings in search for real world applications. The networks of protein interactions (interactomes) of resistant bacteria are some networks with great interest by academics, once such microorganisms can exert violent behavior in biological systems. Therefore, simulations of attacks on bacterial interactomes reveal important information about the topology of the networks, enabling the identification of structural characteristics that can be explored, for example, in the development of drugs. In this study algorithms were developed to evaluate the resilience of nine bacterial networks and one human network through random and intentional attacks of nodes of interest (hubs). The sequential removal of the targets showed that attacks on hubs are more aggressive for the human network than for the bacterial networks studied. Such difference was found to be associated to the clustering of hubs in bacterial networks. In addition, all the nine interactomes exhibited similar behaviors, probably indicating the existence of a pattern relative to resilience for these kinds of bacteria.

KEYWORDS: complex network; interactome; attack; resilience; bacteria.

INTRODUÇÃO

Ataques à sistemas complexos podem causar efeitos devastadores (YAN et al., 2013). O impacto geralmente é maior em nós considerados “importantes” (NIE et al., 2015) tais como hubs, ou seja, nós com uma grande quantidade de vizinhos, principalmente se estes são pouco conectados entre si (baixo coeficiente de cluster), e/ou quando o hub se localiza em um caminho relevante, conectando grupos somente através dele mesmo (alto betweenness), pois pode ocorrer a fragmentação da rede em dois ou mais componentes conexos.

Em redes não-direcionadas, um componente conexo é definido como um grafo que possui ao menos um caminho para qualquer par de nós existentes. De acordo com Yan et al. (2013), uma das formas de se medir o efeito causado por ataques em redes complexas é através do número e tamanho dos componentes conexos. A resiliência, portanto, é a capacidade da rede de se manter, após um ataque ou uma falha, como um único grande componente íntegro no qual qualquer par de nós possui ao menos um caminho de conexão que os ligam. Em outras palavras, a rede é capaz de tolerar a perda de um nó ou uma aresta sem que haja perda significativa do fluxo de informação ou de interações/associações entre os nós envolvidos (REBAZA, 2013).

Neste sentido, este trabalho tem como objetivo a criação e execução de algoritmos de ataque para exploração da resiliência de redes de interação de proteínas de bactérias resistentes a fim de extrair informações relevantes sobre os mecanismos de defesa bem como inferir possíveis desdobramentos em genes associados à suas virulências.

MATERIAL E MÉTODOS

O ambiente de trabalho escolhido para a criação dos algoritmos foi o Google Colaboratory no qual a linguagem Python foi utilizada para o desenvolvimento. As bibliotecas NetworkX, Matplotlib, Numpy, Glob e Random foram importadas para utilização de funcionalidades específicas.

Nove redes de interação de proteínas das bactérias resistentes consideradas (*Acinetobacter baumannii*, *Burkholderia cepacia*, *Clostridium difficile*, *Escherichia coli*, *Klebsiella pneumoniae*, *Mycobacterium tuberculosis*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* e *Streptococcus pyogenes*) foram extraídas através do banco de dados STRING¹, que contém redes de interações físicas e funcionais de proteínas de mais de 14000 organismos até o momento. Essas bactérias foram escolhidas a partir de uma pesquisa prévia quanto à sua agressividade e virulência. O número de nós das redes bacterianas foi filtrado de forma que consideramos apenas as arestas com grau de confiança de ao menos 80%, ou seja, somente as arestas com 80% de chance de realmente existirem no mundo real foram consideradas. Além disso, uma rede de interação de proteínas humanas (HRPD) também foi submetida sob o mesmo processo de ataque para observação e comparação.

Após a importação das redes biológicas para o ambiente de desenvolvimento, os algoritmos foram aplicados de forma a remover sequencialmente os hubs de maior grau, ou seja, as proteínas altamente conectadas, mas também de forma aleatória. Para os ataques aleatórios, foram repetidos 30 ciclos de remoção de nós sorteados aleatoriamente, dos quais os resultados referentes aos impactos foram armazenados em uma matriz de ataque. No caso de ataques à hubs, 5 ciclos foram realizados e os nós eram reordenados de maneira decrescente quanto ao seu grau a cada ciclo. Então, a média e o desvio padrão foram extraídos para cada coluna da matriz.

A resiliência das redes foi calculada através da divisão do tamanho do componente conexo máximo pelo número de nós inicial do componente gigante (subgrafo do tamanho da rede completa, NIE et al., 2015) para cada remoção de nó. Comparações foram feitas entre as redes utilizadas para examinar as diferenças e semelhanças encontradas. Gráficos foram gerados para facilitar a análise e compreensão dos resultados e serão apresentados a seguir. Características descritivas de medidas de redes complexas como o número de nós e o cluster coeficiente foram examinadas em paralelo.

¹ <https://string-db.org/>

RESULTADOS E DISCUSSÃO

Os gráficos abaixo resumem os resultados encontrados neste estudo (Figuras 1 e 2).

No eixo das abscissas tem-se a quantidade relativa de nós removidos durante o processo de aplicação do algoritmo de ataque (definido pela razão entre nós removidos sobre a quantidade de nós iniciais), enquanto que as ordenadas refletem o impacto correspondente. Tanto para o ataque aleatório quanto para o ataque intencional há um maior impacto no início do processo, porém a queda torna-se mais brusca a medida que as remoções avançam e mais hubs ou nós são removidos das redes, o que implica em uma maior desestruturação de suas formações. Isso ocorre devido a fragmentação da rede, diminuindo cada vez mais o tamanho do maior componente conexo existente, e consequentemente impactando a topologia até que o impacto torna-se nulo, pois houve um total colapso da estrutura da rede e não há mais ligações entre os nós.

As nove bactérias exibiram o mesmo padrão de resiliência entre si. Esse fato induz o questionamento de quais fatores ou características são responsáveis pela consistência dos resultados semelhantes. Além disso, as estruturas das redes de bactérias se mostraram mais resistente aos ataques do que a rede de proteína humana. Comparando o ataque à hubs das Figuras 1 e 2 (linha vermelha) percebe-se que este ocasiona uma total remoção de arestas nas redes de bactéria aproximadamente entre 35% e 45%. Já na rede humana (HRPD) este efeito ocorre antes dos 25%. Esse fenômeno acontece pela diferença de clustering, medida que indica as interconexões entre os vizinhos de cada vértice.

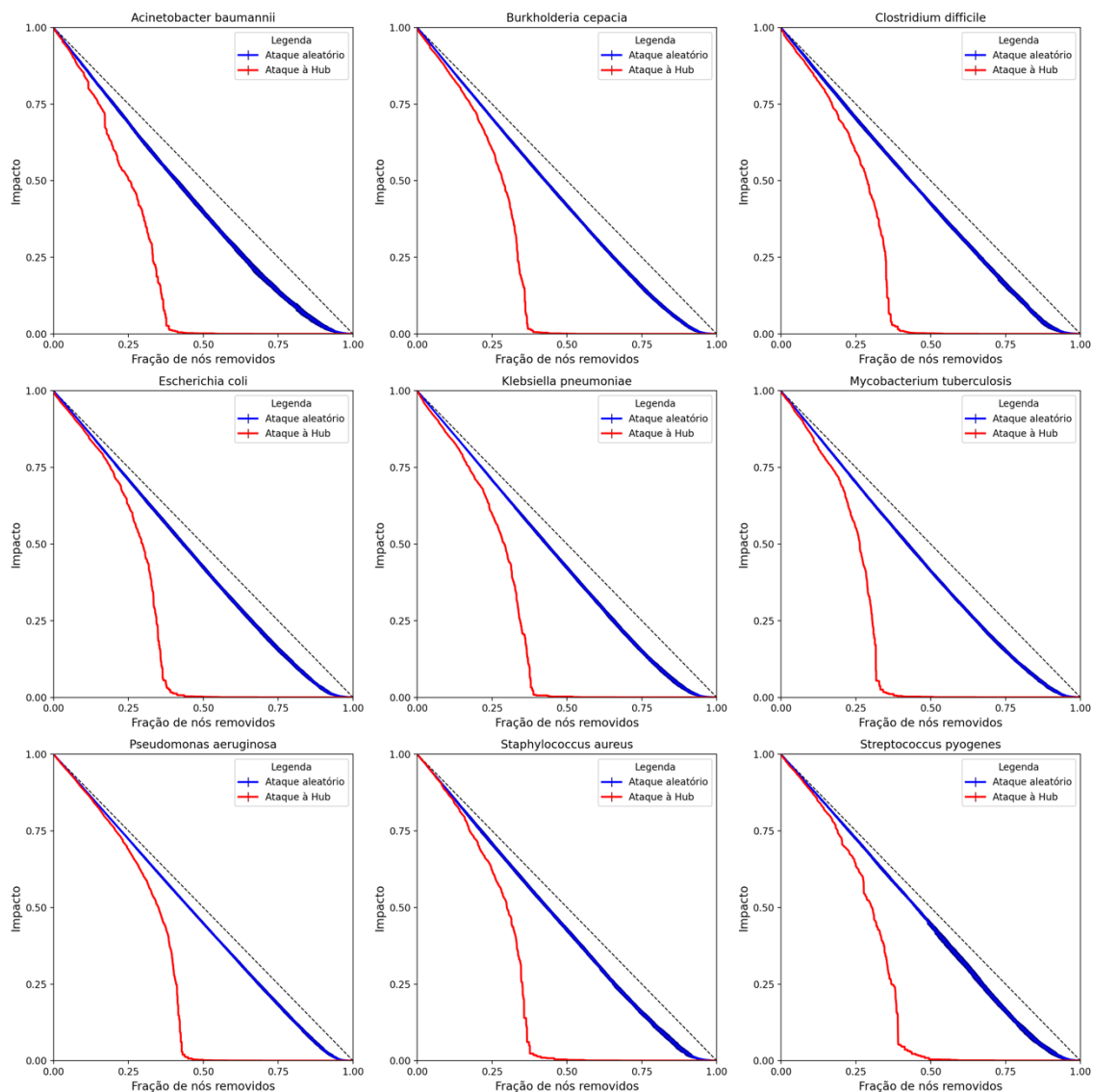


FIGURA 1. Impacto causado por ataques aleatórios e intencionais em redes de interação de proteínas de bactérias resistentes.

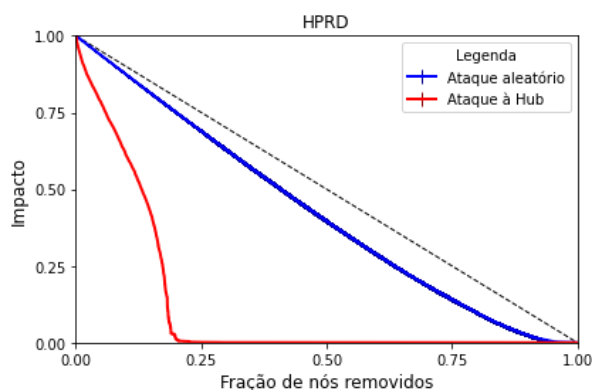


FIGURA 2. Impacto causado por ataques aleatórios e intencionais em rede de interação de proteínas humana.

Observando a Tabela 1, nota-se que a maior parte dos hubs da rede de interação de proteínas humana, representada pelos 5% dos maiores nós, possui coeficiente de cluster menor do que o cluster geral, diferentemente das redes bacterianas, que possuem coeficientes de cluster consideravelmente maior entre os hubs.

TABELA 1. Descrição do número de nós (N), valores do coeficiente de cluster geral (Clustering N), valores do coeficiente de cluster dos hubs que correspondem a 5% do total de nós (Clustering 5% hubs) e razão entre a diferença dos valores de cluster e o cluster geral (Diferença).

	N	Clustering N	Clustering 5% hubs	Diferença
<i>Acinetobacter baumannii</i>	1287	0,47	0,81	0,72
<i>Burkholderia cepacia</i>	2489	0,49	0,70	0,43
<i>Clostridium difficile</i>	1375	0,50	0,74	0,48
<i>Escherichia coli</i>	2193	0,48	0,67	0,40
<i>Klebsiella Pneumoniae</i>	2192	0,51	0,63	0,24
<i>Mycobacterium tuberculosis</i>	1577	0,51	0,74	0,45
<i>Pseudomonas aeruginosa</i>	3451	0,47	0,58	0,23
<i>Staphylococcus aureus</i>	1203	0,49	0,81	0,65
<i>Streptococcus pyogenes</i>	886	0,48	0,79	0,65
HPRD	9219	0,11	0,06	-0,45

CONCLUSÕES

A partir das comparações entre as redes estudadas, conclui-se que o ataque direcionado à hubs impacta a topologia das redes de forma mais agressiva que o ataque aleatório, ou seja, a remoção de um nó altamente conectado prejudica significativamente mais a estrutura e o funcionamento das redes, já que estes concentram grande parte das conexões existentes e muitos outros nós dependem deste ponto de concentração para terem acesso à outros nós (CRUCITTI et al, 2004). Portanto, a característica "scale-free" das redes reais, neste contexto, tornam as redes mais susceptíveis a sofrerem grandes consequências com a perda de um nó por um ataque intencional (NIE et al., 2015; ALBERT, JEONG,

BARABÁSI, 2000). Em contrapartida, redes "small-world" costumam ser mais robustas, ou seja, possuem forte resiliência, pois existe uma maior conectividade entre os vizinhos com alto grau, isto é, há uma rede de suporte entre os nós que se adapta melhor caso haja algum problema (NIE et al., 2015). Redes "small-world" possuem alto coeficiente de cluster que pode explicar esta resiliência.

Além disso, os resultados obtidos sugerem que as redes de interação de proteínas das bactérias resistentes analisadas possuem maior resiliência que a rede de interação de proteínas humanas, como é possível visualizar através dos gráficos criados. Assim como nas redes "small-world", as redes das bactérias consideradas neste artigo também possuíam coeficientes cluster mais altos do que a rede humana, explicando a maior resiliência encontrada, já que o impacto no sistema como um todo é menor quando se há alternativas aos hubs removidos.

O gráfico da rede HPRD exibiu uma queda brusca com valores menores de razão de nós removidos que das redes bacterianas. Isso demonstra uma maior vulnerabilidade e menor tolerância aos ataques intencionais direcionados à hubs. São necessárias investigações mais aprofundadas sobre as causas dessas diferenças de resiliência para que haja uma maior compreensão do assunto, bem como a utilização de diferentes medidas de resiliência para uma análise mais ampla da eficiência das redes em permanecerem robustas por mais tempo.

AGRADECIMENTOS

A aluna é financiada pelo Programa Institucional de Bolsas de Iniciação Científica e Tecnológica do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (PIBIFSP).

REFERÊNCIAS

ALBERT, Réka; JEONG, Hawoong; BARABÁSI, Albert-László. Error and attack tolerance of complex networks. **nature**, v. 406, n. 6794, p. 378-382, 2000.

CRUCITTI, Paolo et al. Error and attack tolerance of complex networks. **Physica A: Statistical mechanics and its applications**, v. 340, n. 1-3, p. 388-394, 2004.

NIE, Tingyuan et al. New attack strategies for complex networks. **Physica A: Statistical Mechanics and its Applications**, v. 424, p. 248-253, 2015.

REBAZA, Jorge Carlos Valverde. **Predição de links em redes complexas utilizando informações de estruturas de comunidades**. 2013. Tese de Doutorado. Universidade de São Paulo.

YAN, Shu et al. Resilience to Intentional Attacks of Complex Networks. In: **Applied Mechanics and Materials**. Trans Tech Publications Ltd, 2013. p. 647-651.