

GUILHERME HENRIQUE DA SILVA COSTA

PROPAGAÇÃO DE EPIDEMIAS EM REDES DANIFICADAS

Dissertação apresentada à Universidade Federal de Viçosa como parte das exigências do Programa de Pós-Graduação em Física para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2018

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

C837p
2018

Costa, Guilherme Henrique da Silva, 1995-
Propagação de epidemias em redes danificadas / Guilherme
Henrique da Silva Costa. – Viçosa, MG, 2018.
viii, 63f. : il. ; 29 cm.

Orientador: Sílvio da Costa Ferreira Junior.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f. 60-63.

1. Percolação (Física estatística). 2. Ciberterrorismo.
3. Algoritmos computacionais. 4. Simulação (Computadores).
I. Universidade Federal de Viçosa. Departamento de Física.
Programa de Pós-Graduação em Física Aplicada. II. Título.

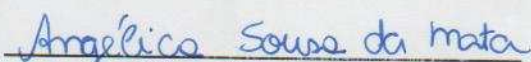
CDD 22. ed. 530.13

GUILHERME HENRIQUE DA SILVA COSTA

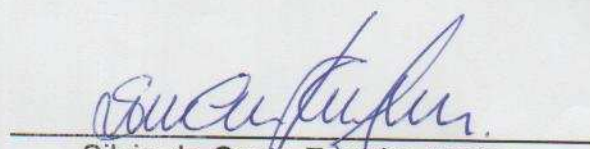
PROPAGAÇÃO DE EPIDEMIAS EM REDES DANIFICADAS

Dissertação apresentada à
Universidade Federal de Viçosa, como
parte das exigências do Programa de
Pós-Graduação em Física, para
obtenção do título de *Magister
Scientiae*.

APROVADA: 22 de fevereiro de 2018.


Angélica Sousa da Mata


Thiago Albuquerque de Assis


Silvio da Costa Ferreira Junior
(Orientador)

Agradecimentos

Aos meus familiares, em particular meus pais Elvio e Sheila e meus irmãos Elvio Jr. e Lucas pelo apoio, pela confiança e pelas lições durante toda minha vida.

Ao meu orientador, Prof. Silvio Ferreira, por todos os ensinamentos, pelo apoio e pela paciência nesses anos de trabalho.

Aos demais professores da UFV, pela minha formação acadêmica.

Aos amigos do GISC, como Wesley, Arthur, Diogo, etc.

Ao GISC e ao NSSC, pelo suporte computacional.

Aos demais colegas de Viçosa, por todas discussões mirabolantes ao longo dos anos, pelas longas conversas no café e pelas jogatinas de RPG.

À Universidade Federal de Viçosa.

À CAPES, pela bolsa de estudos.

Ao CNPq, a CAPES/Programa Ciência sem Fronteiras (projeto nº 88881.030375/2013-01) e à FAPEMIG, pelo apoio financeiro.

“Raramente estou mais feliz do que quando passo um dia inteiro a programar o meu computador para fazer automaticamente uma tarefa que de outra forma demoraria uns bons dez segundos a fazer à mão.”
(Douglas Noel Adams)

Sumário

Lista de abreviaturas e siglas	vi
Resumo	vii
Abstract	viii
I Introdução	1
II Redes complexas	4
2.1 Caracterizando redes complexas	6
2.2 Redes sem escala	13
III Epidemias	17
3.1 Modelagem	17
3.2 Epidemias e transições de fases	19
3.3 Teorias de campo médio	25
3.3.1 Teoria de campo médio homogênea	25
3.3.2 Teoria de campo médio heterogênea	26
3.3.3 Teoria de campo médio microscópica	27
3.3.4 Teorias de campo médio para o modelo SIR	28
3.4 A teoria BCPS	29
IV Percolação	32
4.1 Conceitos de percolação	32
4.2 Ataques em redes	34
4.3 Estratégias de ataque	36
4.3.1 Ataque aleatório	37
4.3.2 Ataque de contatos	38
4.3.3 Ataque de contatos com proteção de hub	38
V Metodologia computacional	40
5.1 Modelo UCM	40
5.2 Algoritmos de simulação de modelos epidêmicos	42
5.2.1 Modelo SIS	42
5.2.2 Modelo SIR	43
5.3 Método quase estacionário	43
5.4 Outros detalhes	45
VI Resultados e discussões	46
6.1 Propriedades das redes atacadas	46
6.2 Efeitos da remoção de vértices no limiar epidêmico do modelo SIS	49
6.3 Comparação com teorias de campo médio	55
6.4 Efeitos da remoção de vértices no λ_p do modelo SIR	56
VII Considerações finais	57

Referências 60

Lista de abreviaturas e siglas

SIS	Suscetível-Infectado-Suscetível
SIR	Suscetível-Infectado-Recuperado
PC	Processo de Contato
SI	Suscetível-Infectado
HMF	<i>Heterogeneous Mean Field</i>
QMF	<i>Quenched Mean Field</i>
BCPS	Boguna, Castellano e Pastor-Satorras
ETF	Escalonamento de Tamanho Finito
RND	Ataque Aleatório
NVA	<i>Neighbor Vertex Attack</i>
NVA-HP	<i>Neighbor Vertex Attack with Hub Protection</i>
UCM	Modelo de Configurações Não Correlacionado, do inglês <i>Uncorrelated Configuration Model</i>
OGA	<i>Optimized Gillespie Algorithm</i>

Resumo

COSTA, Guilherme Henrique da Silva, M.Sc., Universidade Federal de Viçosa, fevereiro de 2018. **Propagação de epidemias em redes danificadas.** Orientador: Silvio da Costa Ferreira Junior.

A sociedade moderna é altamente dependente de estruturas funcionais em rede, como redes de comunicação ou de transporte. Assim, o estudo de ataques em redes e o entendimento de sua robustez/fragilidade são fundamentais na procura de redes mais eficientes. Muitos estudos envolvendo ataques em redes subestimam a importância de processos dinâmicos evoluindo nessas redes atacadas para prevenir danos futuros. A maioria dos estudos em redes atacadas foca na caracterização da estrutura das componentes não danificadas da rede. No entanto, muito menos atenção tem sido dedicada para analisar a influência desses ataques na propagação de epidemias ou informação. Nesse trabalho, investigamos a propagação de epidemias nos modelos suscetível-infectado-suscetível e suscetível-infectado-recuperado em redes após os ataques com diferentes estratégias, enfatizando o estudo dos limiares epidêmicos dos modelos. Nós estudamos duas intensidades de ataques: um brando no qual a componente gigante da rede possui tamanho comparável a rede original e outro mais destrutivo mas ainda apresentando uma componente gigante grande. Ataques aleatórios, que não modificam a distribuição de grau de redes sem escala, não alteram o comportamento do limiar epidêmico em relação às redes intactas e continua indo para zero com o aumento do tamanho da rede. Investigamos outras duas estratégias de ataque de contatos: uma em que um primeiro vizinho de um vértice escolhido ao acaso é removido com probabilidade 1 e outra em que se remove o vizinho com uma probabilidade que protege os vértices de graus mais elevados de serem removidos. Observamos que a dependência do limiar epidêmico com o tamanho da rede se altera dramaticamente, saturando em um valor constante no limite de redes grandes, ou diminuindo mais lentamente do que uma lei de potência ou até mesmo aumentando com o tamanho da rede em alguns casos, a depender da estratégia e da intensidade do ataque. Também concluímos que ambos ataques são mais efetivos quando aumentada a heterogeneidade da rede. Mostramos que mesmo os ataques de contato abaixo do limiar de percolação, que não fragmentam a rede, são capazes de tornar a rede inoperante para a propagação da epidemia no modelo SIS. Além disso, comparamos os resultados das simulações com a teoria de campo médio heterogêneo, conhecida como HMF (do inglês *heterogeneous mean field*), que leva em conta apenas o grau do vértice, e a teoria de campo médio que leva em conta toda a estrutura da rede por meio da matriz de adjacência, conhecida como QMF (do inglês *quenched mean field*) e observamos que a primeira captura as mudanças no limiar epidêmico com maior precisão do que a segunda.

Abstract

COSTA, Guilherme Henrique da Silva, M. Sc., Universidade Federal de Viçosa, February, 2018. **Epidemic spreading on damaged networks.** Adviser: Silvio da Costa Ferreira Junior.

The modern society is highly dependent of functioning networked structures such as communication and transportation networks. Thus, the study of attacks in networks and the understanding of their robustnesses/fragilities are imperative for the search of more efficient networks. Many studies involving network attacks underestimate the importance of dynamic processes evolving on these attacked networks to prevent further damage. Most of the studies on network attack is held on the characterization of the structure of the undamaged components of the networks. However, much less attention has been devoted to the influence of these attacks on the epidemic and information spreading. In this work, we investigate the epidemic spreading of susceptible-infected-susceptible and susceptible-infected-recovered models on networks after the attack according different strategies aiming at the study of the epidemic thresholds. We investigate two attack intensities: a mild one with a giant component with size comparable to the original and other more destructive but still presenting a large giant component. Random attacks, which don't alter the degree distribution on scale free networks, don't change the epidemic threshold behavior in relation to the undamaged network that goes to zero as the network size increases. We investigate other two acquaintance attack strategies: one where a nearest neighbor of a randomly chosen vertex is deleted with probability 1 and other that remove the neighbor with a probability that protect the vertices of higher degree from being deleted. We observe that the threshold scaling changes dramatically, reaching a constant value for the limit of large networks, or decreasing slower than a power law, or even increasing with the size of the network, depending on the model and intensity of the attacks. We also conclude that both attacks are more effective as the heterogeneity of the networks increases. We show that even acquaintance attacks below the percolation threshold, which are not sufficient to fragment the network, are able to make the network ineffective for the epidemic spreading in the SIS model. Furthermore, we compare the results from the simulations with the heterogeneous mean field (HMF) theory, which considers that the vertex degree is the only relevant quantity and the quenched mean field (QMF) theory, which include the whole structure of network through the adjacency matrix. We observed that the HMF theory captures the changes in the epidemic threshold more accurately than the QMF theory.

I Introdução

Muitos sistemas podem ser representados na forma de uma rede, em que os vértices são os constituintes desse sistema e as arestas representam as interações desses constituintes, que podem ser a amizade entre duas pessoas, a conexão física entre computadores, a existência de rotas aéreas entre aeroportos ou a malha rodoviária de um país. Estamos imersos em e interagindo com essas estruturas no dia a dia e o funcionamento da sociedade moderna é altamente dependente delas, sejam redes de transporte, redes de comunicação ou até mesmo a internet.

Estudar as propriedades estruturais dessas redes ajuda a entender os diversos processos dinâmicos (propagação de epidemias, de informação, transportes de pessoas) que ocorrem nesses sistemas e o aumento do poder computacional nos últimos anos possibilitou a coleta de dados e a análise estrutural de redes reais (redes sociais, malhas rodoviárias, tráfego aéreo). Em particular, muitas dessas redes complexas possuem a propriedade de mundo pequeno, a qual indica que dois vértices quaisquer podem se comunicar através de um número bem pequeno de mediadores [1]. Verifica-se também que muitas dessas redes possuem distribuição de conectividade, (pelo menos assintoticamente) como uma lei de potência ($P(k) \sim k^{-\gamma}$), em particular com expoente $2 < \gamma < 3$.

Como toda estrutura, esses sistemas estão sujeitos a mal funcionamento e incidentes como *blackouts*, apagões de internet e *cyberataques*, nos quais as falhas de poucos vértices causam um dano estrutural grande, mostrando a vulnerabilidade dessas redes a certos tipos de falhas e ressaltam a importância de analisar estruturalmente redes complexas no objetivo de construir sistemas mais robustos e minimizar danos em caso de ataques [2]. Além disso, há também o interesse em desenvolver estratégias de ataque que sejam mais efetivas na fragmentação e consequente inutilização de determinadas redes. Por exemplo, a prisão de determinadas pessoas pode inviabilizar uma rede de crime [3], a imunização de pontos chave de uma rede de contatos pode impedir a propagação de uma epidemia [4, 5] ou a atuação de fármacos em pontos específicos de uma cadeia metabólica pode impedir um agente patogênico de se replicar [6].

Dentre os diversos processos dinâmicos que podem ocorrer em uma rede complexa, a propagação de epidemias tem atraído um interesse enorme. Através de modelos simples, estuda-se propriedades de epidemias, como a densidade de sítios infectados, a presença de um surto global, se há surtos endêmicos, limiares epidêmicos, dentre outros. Por exemplo, a partir da modelagem de doenças reais, é possível prever a propagação de uma epidemia entre países ou definir estratégias de imunização de forma a minimizar os danos causados [7, 8, 9]. Modelos epidêmicos podem ser estudados em redes atribuindo estados aos vértices destas, definindo taxas de transição entre esses estados e analisando as propriedades relevantes. Dentre os diversos modelos, um muito estudado é o Suscetível-

Infectado-Suscetível (SIS), no qual os sítios de uma rede assumem os estados infectado ou suscetível e podem espontaneamente se curar ou transmitir a epidemia para sua vizinhança com determinadas taxas. Outro modelo importante é o Suscetível-Infectado-Recuperado, no qual os sítios podem assumir também o estado Recuperado, no qual o sítio não pode ser infectado. Muito já foi estudado sobre as propriedades epidêmicas desses modelos em redes sem escala, cuja distribuição de conectividade segue uma lei de potência ($P(k) \sim k^{-\gamma}$, $2 < \gamma < 3$) de forma que muitos resultados são bem consolidados (*scalings* dos limiares epidêmicos, expoentes críticos, dentre outros).

A maioria dos estudos sobre ataques em redes complexas focam na caracterização estrutural da rede [10, 11], subestimando a importância da análise dos processos dinâmicos após os ataques. Além disso, muitas estratégias de ataque estudadas são localizadas e adaptativas (detalhes na seção IV.2), implicando no conhecimento integral das redes a serem atacadas, o que na prática, é inviável. Na tentativa de minimizar esses problemas, estudamos estratégias de ataques não-localizadas, que não dependem do conhecimento integral da rede e que são facilmente implementadas computacionalmente. Porém, elas são capazes de danificar as redes de forma efetiva. O ataque aleatório é um modelo simples de ataque no qual um vértice é sorteado ao acaso e suas ligações são removidas da rede. Este modelo foi utilizado como ponto de partida na elaboração das outras estratégias, a saber: ataque de contato, no qual um vértice e seu vizinho são sorteados ao acaso e o vizinho é removido e o ataque de contato com proteção de *Hub* (vértice de grau elevado), no qual existe uma probabilidade dependente do grau do vértice não ser removido (detalhes na seção IV.3).

Nesse trabalho, estudamos os efeitos da propagação de epidemias em redes complexas (inicialmente sem escala) em que uma fração de vértices foram removidos segundo uma determinada estratégia. Em particular, estudamos as mudanças no limiar epidêmico dos modelos. Inicialmente, analisamos as estratégias de ataque quanto à mudanças estruturais em redes complexas com maior ou menor ocorrência de *hubs*, caracterizando como as diferentes formas de ataque alteram algumas propriedades dessas redes. Partimos então para a análise do limiar epidêmico dos modelos SIS e SIR nessas redes, com uma fração relativamente pequena de vértices removidos e comparamos os resultados encontrados com as teorias de campo médio adaptadas para redes complexas. De posse dos resultados, análises estruturais mais específicas foram necessárias para poder explicar os resultados encontrados, como o cálculo do maior grau e o número de *hubs* existentes na rede. A partir dessa análise estrutural, pudemos analisar os resultados em termos de uma teoria para o modelo SIS em que o mecanismo de ativação da rede envolve as escalas de tempo de vida de um *hub* e do tempo que um *hub* leva para infectar ou ser infectado por outro.

Partimos então para a análise do modelo SIS em redes danificadas de forma mais agressiva (fração maior de vértices removidos), onde a rede encontra-se mais fragmentada, porém, ainda bem abaixo do limiar de percolação tal que sua componente gigante ainda

é muito grande. A progressão dessas análises se deu da mesma forma que para ataques mais brandos, inicialmente realizando simulações do modelo SIS identificando os limiares epidêmicos para posteriormente realizar a análise envolvendo as teorias.

A organização dessa dissertação se deu da seguinte forma: no Capítulo II, conceitos sobre redes complexas e como caracterizá-las são apresentados, além de uma seção específica sobre redes sem escala. No Capítulo III, apresentamos os fundamentos do estudo de modelos epidêmicos tanto em reticulados quanto em redes complexas, apresentando também as principais teorias de campo médio estudadas, além da teoria BCPS, que prevê um limiar epidêmico nulo para o SIS através da comparação do tempo de atividade de um *hub* e o tempo necessário para reinfecção entre *hubs*.

O Capítulo IV versa sobre percolação, onde apresentamos conceitos de percolação para reticulados, estendendo esses conceitos para redes complexas considerando remoção de vértices nestas. Nesse capítulo, as estratégias de ataque estudadas nessa dissertação também são apresentadas e discutidas. Já no Capítulo V, é apresentada a metodologia computacional e os algoritmos utilizados para realização das simulações dos modelos epidêmicos, criação de redes complexas e aplicação das diferentes estratégias de ataque.

Por fim, no Capítulo VI, alguns resultados obtidos a partir dessas simulações e comparações com as teorias são apresentados e discutidos e algumas considerações finais são feitas no Capítulo VII.

II Redes complexas

Em termos gerais, uma rede é um sistema que admite representação como um grafo, em que os vértices (ou sítios) identificam os elementos do sistema e as arestas representam relações ou interações entre estes vértices. Para exemplificar: em uma malha aérea, os sítios são os aeroportos e os *links* são os vôos entre eles; em uma rede do cérebro, os vértices são neurônios e as arestas são as sinapses. Nesse sentido, redes apresentam um ferramental teórico que permite a representação de relações em sistemas complexos onde sua caracterização implicaria no mapeamento de interações entre uma grande quantidade de indivíduos ou constituintes. Diversos sistemas de interesse em ciência podem ser mapeados em redes, como cadeias metabólicas [12], redes cerebrais [13], redes sociais [14, 15], dentre outros. Na Figura 2.1, temos diversos exemplos de redes.

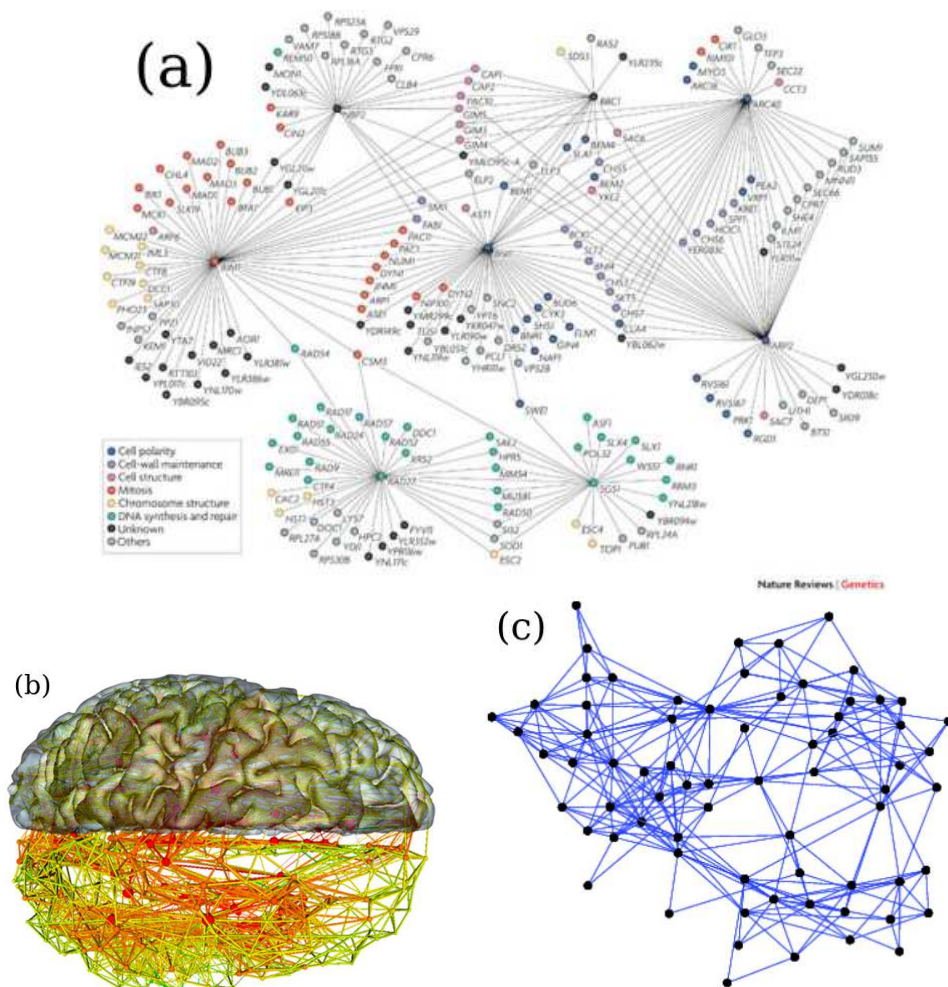


Figura 2.1 – Exemplos de redes. (a) Interação genética da levedura *Saccharomyces cerevisiae* (Retirado de [12]). (b) Figura ilustrando conectomas no cérebro (Retirado de [13]). (c) Rede de amizades entre os alunos de uma escola (Retirado de [16]).

De forma bem simplificada, o conceito de redes "complexas" está intrinsicamente ligado ao fato dessas redes apresentarem características topológicas não triviais, como padrões de conexões que não são nem regulares nem puramente aleatórios, alta heterogeneidade na distribuição de conectividades, alto coeficiente de agregação, padrões de correlação entre vértices, estruturas de comunidade ou hierárquicas (muitos desses conceitos serão apresentados a frente), dentre outros.

Quando estudamos redes reais, simplificamos o problema identificando os principais componentes de sua estrutura. Por exemplo, em uma rede de distribuição elétrica, temos vários constituintes: geradores, linhas de transmissão de alta voltagem, transformadores, linhas de baixa voltagem e consumidores. Identificados os elementos de interesse, muitas vezes descartamos ainda a estrutura espacial da rede e a intensidade das conexões (no caso da rede de distribuição, as diferentes fiações podem ter diferentes voltagens). O grafo resultante é o objeto de estudo da teoria de redes. A Figura 2.2 ilustra o passo a passo dessa construção.

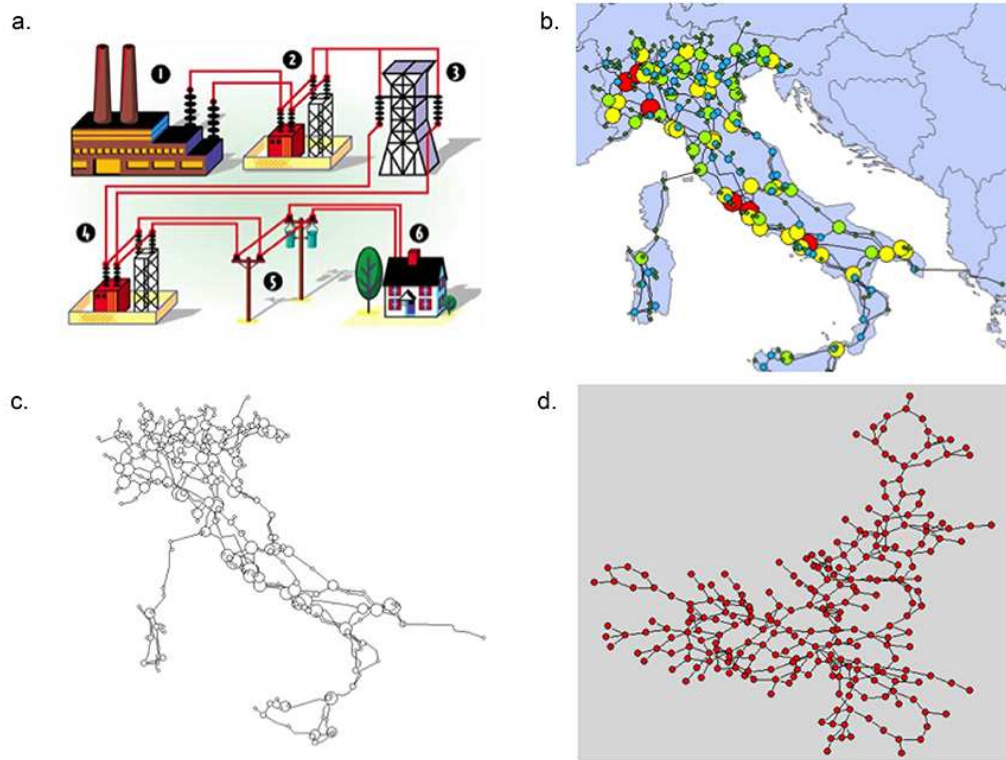


Figura 2.2 – (a) Diferentes constituintes de uma rede de distribuição elétrica: 1-Geradores, 2-Comutadores, 3-Linhas de Alta Voltagem, 4-Transformadores, 5-Linhas de Baixa Voltagem, 6-Consumidores. (b) Rede de distribuição elétrica da Itália, considerando produção e consumo de cada elemento. (c) Rede espacial construída relevando detalhes da rede em (b). (d) Rede complexa construída desconsiderando as informações espaciais da rede de distribuição. Figuras retiradas de [17].

2.1 Caracterizando redes complexas

O tratamento matemático e computacional de redes complexas é comumente feito através da matriz de adjacência A , que indica quais vértices estão conectados. A matriz segue as regras:

- $A_{ij} = 0$, se o vértice i não está conectado com o vértice j .
- $A_{ij} = 1$, se o vértice i está conectado com o vértice j .

As propriedades de A dependem do tipo de rede a ser considerado. Quanto ao direcionamento das arestas, as redes podem ser **direcionadas**, se a ligação $i \rightarrow j$ não implicar na ligação $j \rightarrow i$ (Twitter, Instagram, Hiperlinks) ou **não-direcionadas**, se a ligação $i \rightarrow j$ implicar na ligação $j \rightarrow i$ (Aeroportos, Rodovias, Internet). Além disso, podemos definir uma matriz peso W que mede a intensidade das conexões entre os vértices. Em redes da Internet, esses pesos podem representar a quantidade de dados que fluem entre dois pontos de acesso, em uma rede social podem representar a frequência de contato entre as pessoas, etc. Nesse trabalho, consideraremos apenas redes não-direcionadas e sem peso, de forma que todas as arestas possuem a mesma intensidade e algumas definições podem não ser aplicáveis no caso de redes direcionadas. Na Figura 2.3, temos exemplos de redes direcionada e não-direcionada.

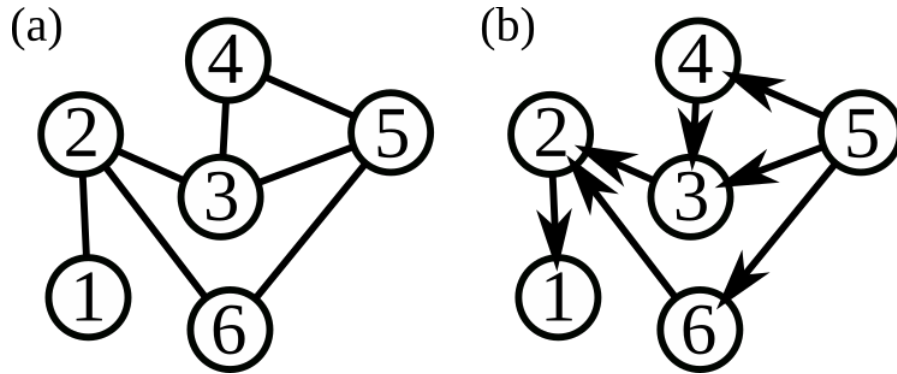


Figura 2.3 – Exemplo de rede com 6 vértices (a) Não-direcionada (b) Direcionada.

Definimos a vizinhança de um vértice (ν_i) como o conjunto de todos os sítios j tais que $A_{ij} = 1$. Exemplificando, na Figura 2.3, o sítio 1 tem como vizinhos $\nu_1 = \{2\}$ e o sítio 3 tem como vizinhos $\nu_3 = \{2, 4, 5\}$.

A partir da matriz de adjacência (considerando redes sem peso), temos toda a informação possível sobre a rede e podemos analisar algumas propriedades importantes que podem ser extraídas desta, como o grau de conectividade de cada sítio (k_i), que representa o número de arestas e é dado por:

$$k_i = \sum_j A_{ij}. \quad (2.1)$$

Duas outras definições são importantes ao se referir ao grau de um vértice. *Hubs* são sítios com alta conectividade (k grande em relação ao resto da rede). Se além disso, a conectividade desse sítio for muito maior que do resto da rede, o sítio é denominado *Outlier*. Considerando uma rede com N vértices, cada um deles com grau k_i , podemos calcular a distribuição de grau ($P(k)$) dos vértices dessa rede

$$P(k) = \frac{N_k}{N}, \quad (2.2)$$

em que N_k é a quantidade de vértices da rede que possui grau k . A partir dessas definições, podemos calcular os momentos da distribuição $P(k)$ facilmente

$$\langle k^n \rangle = \sum_k k^n P(k). \quad (2.3)$$

Alguns momentos da distribuição fornecem informações relevantes sobre a mesma, dentre outros, temos:

- $\langle k \rangle$: valor médio da distribuição.
- $\langle k^2 \rangle$: a partir dele, podemos calcular a variância ($\langle k^2 \rangle - \langle k \rangle^2$) que mede o desvio em torno do valor médio

Para a rede da Figura 2.3, temos que $P(1) = 1/6$; $P(2) = 1/3$ e $P(3) = 1/2$. A partir da definição dos momentos, as redes podem ser classificadas em redes **homogêneas**, se $\langle k \rangle$ é uma boa aproximação para todo k_i (ou seja, se a variância é pequena) ou **heterogêneas**, se a variância de k é grande e existem sítios com grau bem diferente de $\langle k \rangle$.

Uma outra grandeza interessante é o grau médio dos vizinhos de um vértice, definido como

$$K_{nn,i} = \frac{1}{k_i} \sum_{j \in \nu_i} k_j. \quad (2.4)$$

A partir de K_{nn} (do inglês *nearest neighbor*), podemos obter informações sobre as correlações entre os graus e se há ou não tendências dos vértices de se conectarem aos seus semelhantes ou aos seus diferentes. Vamos considerar o valor médio sobre todos os vértices de mesmo grau k dado por

$$K_{nn}(k) = \frac{1}{N_k} \sum_i K_{nn,i} \delta_{k,k_i}, \quad (2.5)$$

em que δ_{k,k_i} é o delta de Kronecker. A partir da análise de $K_{nn}(k)$, podemos tirar informações sobre as correlações entre os vértices da rede [18]. Se os sítios possuem tendência a se conectarem com outros vértices com graus parecidos aos seus, então K_{nn} é uma função crescente de k e dizemos que a rede possui **mistura associativa**. Porém,

se os sítios de grau pequeno possuem tendências a conectarem com sítios de grau grande, então K_{nn} torna-se uma função decrescente de k e, neste caso, dizemos que a rede possui uma **mistura dissociativa**. Se não há essa preferência nas conexões ou se há competição entre ambos mecanismos, dizemos que a rede é não correlacionada e, na média, K_{nn} é constante. A Figura 2.4 ilustra os diferentes casos.

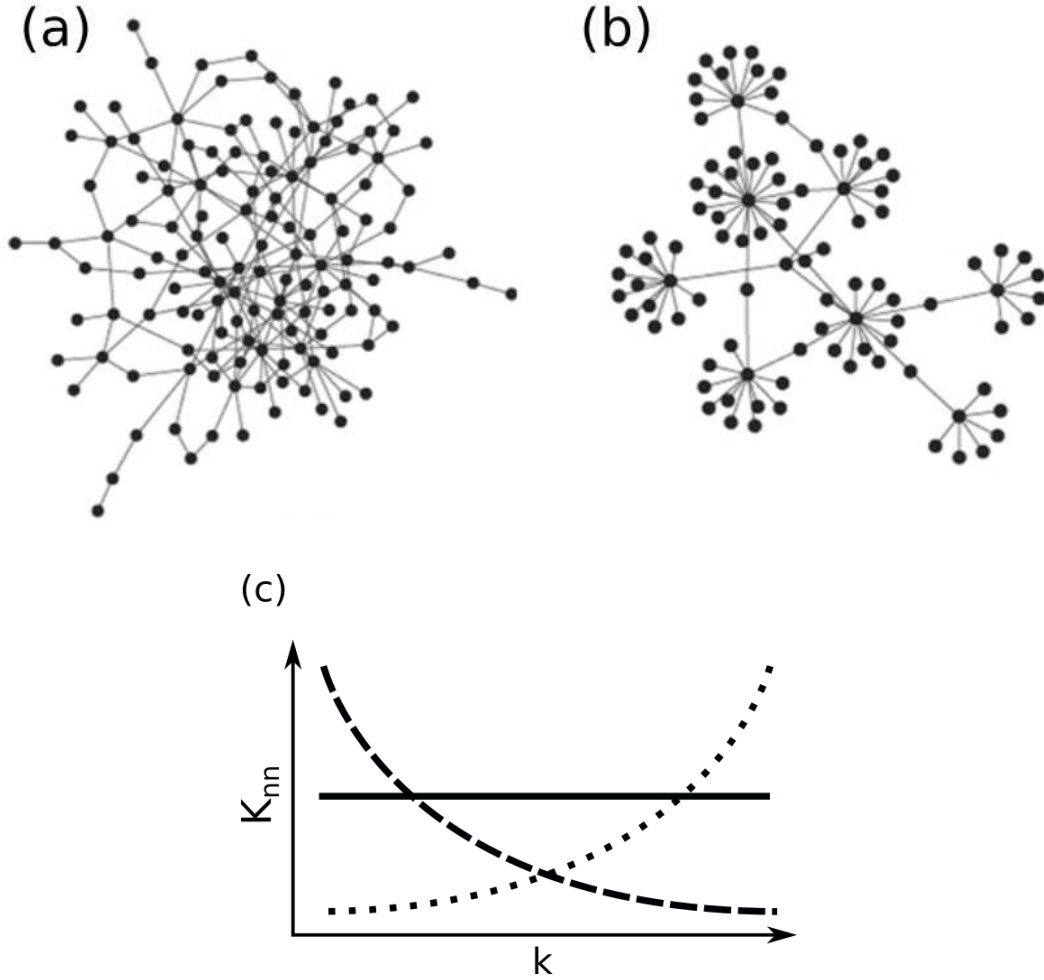


Figura 2.4 – Exemplo de rede (a) Associativa (b) Dissociativa. Figuras retiradas de <http://plus.maths.org> (c) $K_{nn}(k)$ para os três tipos de mistura (pontilhado - associativa ; tracejado - dissociativa ; cheia - não correlacionada).

Podemos escrever K_{nn} de uma forma mais conveniente para fins de cálculo:

$$K_{nn}(k) = \sum_{k'} k' P(k'|k) \quad (2.6)$$

onde $P(k'|k)$ é a probabilidade condicional de um vértice de grau k estar conectado a um vértice de grau k' . Para encontrar uma expressão para $P(k'|k)$, vamos definir a matriz E , cujos elementos satisfazem as relações:

- $E_{kk'}$ é o número de arestas entre vértices de grau k e k' , se $k \neq k'$;
- $E_{kk'}$ é $2 \times$ (número de arestas entre vértices de mesmo grau), se $k = k'$.

A partir dessa definição, a matriz E possui as seguintes propriedades:

$$\sum_{k'} E_{kk'} = kN_k \quad (2.7)$$

$$\sum_k \sum_{k'} E_{kk'} = \langle k \rangle N = 2\mathcal{E} \quad (2.8)$$

com $\mathcal{E}$ o número total de arestas. Assim, $P(k'|k)$ toma a forma

$$P(k'|k) = \frac{E_{kk'}}{kN_k}, \quad (2.9)$$

com kN_k o número de arestas emitidas por todos os sítios de grau k e $E_{kk'}$ a quantidade dessas arestas que estão conectadas a sítios de grau k' . Em seguida, calculamos a probabilidade de, dada uma aresta escolhida aleatoriamente, ela conecte um sítio de grau k a um sítio de grau k' . Essa probabilidade é dada pela probabilidade conjunta

$$P(k', k) = \frac{E_{kk'}}{\mathcal{E}} = \frac{E_{kk'}}{\langle k \rangle N}. \quad (2.10)$$

Rearrajando as Eqs.(2.9) e (2.10), podemos relacionar $P(k'|k)$ com $P(k', k)$

$$P(k'|k) = \frac{\langle k \rangle P(k, k')}{kP(k)}. \quad (2.11)$$

Como não fizemos nenhuma hipótese sobre k' ou k , podemos facilmente fazer a permutação $k' \leftrightarrow k$ e encontrar uma expressão para $P(k|k')$

$$P(k|k') = \frac{\langle k \rangle P(k', k)}{k'P(k')}. \quad (2.12)$$

Como a rede é não direcionada, $P(k', k) = P(k, k')$. Rearrajando essas duas equações, temos:

$$k'P(k')P(k|k') = kP(k)P(k'|k) \quad (2.13)$$

Considerando redes não correlacionadas, temos que $P(k'|k)$ independe de k por definição [1]. Assim, se somarmos em k em ambos os lados da Eq. (2.13), chegamos em uma expressão para $P(k'|k)$

$$P(k'|k) = \frac{k'P(k')}{\langle k \rangle}. \quad (2.14)$$

Substituindo essa expressão na Eq. (2.6), temos

$$K_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}. \quad (2.15)$$

Outro conceito importante em redes é a noção de **caminho**. Considerando uma

rede não direcionada, qualquer sequência de vértices tais que o anterior esteja conectado ao próximo por uma aresta constitui um caminho, uma trajetória que conecta o vértice inicial ao vértice final (vide Figura 2.5). O comprimento desse caminho (ℓ) é dado pelo número de arestas entre os vértices. A partir desse conceito, podemos classificar redes em outros dois tipos: **conectadas**, se para quaisquer dois vértices na rede, existe um caminho entre eles, ou **fragmentadas** caso contrário. Em ambos casos, definimos uma **componente** da rede como os subconjuntos de vértices tais que exista pelo menos um caminho entre quaisquer dois vértices pertencentes a esse grupo. Assim, redes conectadas possuem apenas uma componente, enquanto redes fragmentadas possuem pelo menos duas. Chamamos de **componente gigante** a componente de uma rede que possui maior número de vértices e que seja uma fração finita da rede (no limite termodinâmico). A Figura 2.6 ilustra esses conceitos.

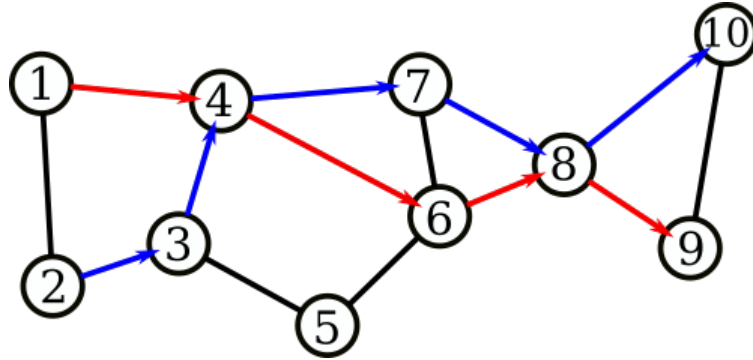


Figura 2.5 – Esquema ilustrativo de caminhos em uma rede. Vermelho: $1 \rightarrow 4 \rightarrow 6 \rightarrow 8 \rightarrow 9$, com comprimento $\ell = 4$; Azul: $2 \rightarrow 3 \rightarrow 4 \rightarrow 7 \rightarrow 8 \rightarrow 10$, com comprimento $\ell = 5$.

A partir disso, podemos definir o menor caminho entre vértices ℓ_{ij} . Dessa forma, podemos definir o menor caminho médio da rede

$$\bar{\ell} = \frac{1}{N(N-1)} \sum_i \sum_j \ell_{ij}. \quad (2.16)$$

O diâmetro de uma rede (D) pode ser calculado em termos de ℓ_{ij}

$$D = \max\{\ell_{ij}\} \quad (2.17)$$

e representa o tamanho da rede.

Em 1967, o americano Stanley Milgram realizou um experimento para tentar medir distâncias em redes sociais [20]. Ele escreveu diversas cartas, com destinatários em Massachusetts e repassou-as aleatoriamente para cidadãos em Omaha e Wichita (cidades nos EUA). Nas cartas, ele pedia as pessoas que repassassem-as para amigos ou conhecidos de forma a tentar chegar ao seu destino com o menor número de conexões. Das 296 cartas, cerca de 60 chegaram ao destino, com um número médio de mediadores 5,2,

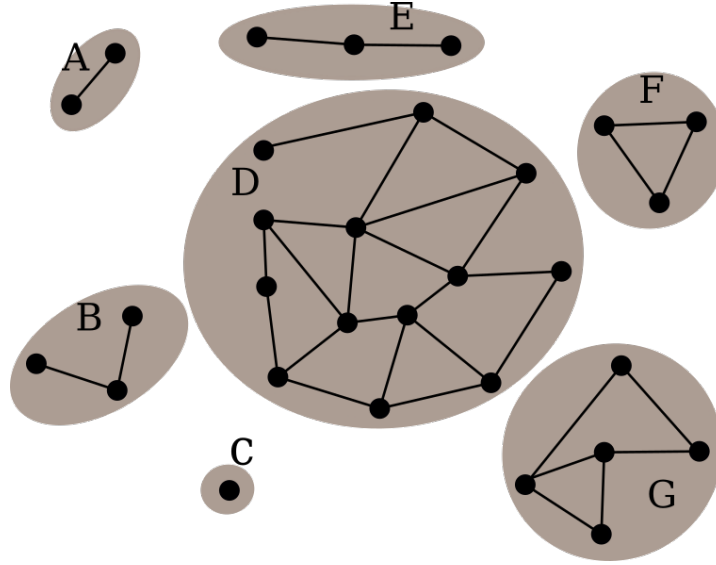


Figura 2.6 – Componentes de uma rede não-direcionada. A componente D é chamada componente gigante da rede, com $N_D = 13$ vértices. Figura adaptada de [19].

muito abaixo do esperado. A partir desse experimento, foi cunhado o termo **seis graus de separação**, que representaria o número médio de conexões necessárias para se chegar a qualquer pessoa no mundo.

Recentemente, estudos envolvendo a rede de usuários do Facebook (≈ 720 milhões de usuários e 69 bilhões de links)[14, 15] mostraram que $\bar{\ell}$ para esta rede é em torno de 4,7 o que representa 3,7 mediadores entre quaisquer usuários da rede social. Um número menor ainda do que o encontrado por Milgram.

Esses experimentos ilustram o conceito de redes de mundo pequeno, em que o menor caminho médio entre os vértices cresce de forma logarítmica com o aumento de N ($\bar{\ell} \sim \ln N$) ou de forma mais lenta. Assim, tomando um vértice como ponto de partida, com um número pequeno de passos é possível alcançar qualquer outro vértice da rede.

Uma maneira simples de formular esse conceito matematicamente encontra-se no Capítulo III de [17]. Considerando uma rede completamente aleatória de N vértices com grau médio $\langle k \rangle$, um sítio dessa rede possui em média $\langle k \rangle$ sítios a distância $l = 1$, $\langle k \rangle^2$ sítios a distância $l = 2$, $\langle k \rangle^3$ sítios a distância $l = 3$ e assim sucessivamente

Assim, o número de sítios que estão a uma distância de até l é dado por

$$N(l) = 1 + \langle k \rangle + \langle k \rangle^2 + \langle k \rangle^3 + \dots + \langle k \rangle^l = \frac{\langle k \rangle^{l+1} - 1}{\langle k \rangle - 1} \quad (2.18)$$

Lembrando que $N(l)$ está limitado ao número de sítios da rede. Dessa forma, $N(D) \approx N$. Realizando algumas manipulações algébricas na Eq. (2.18) temos que

$$N(\langle k \rangle - 1) - 1 = \langle k \rangle^D \quad (2.19)$$

$$\ln \langle k \rangle D + 1 = \ln N + \ln(\langle k \rangle - 1) \quad (2.20)$$

$$D \approx \frac{\ln N}{\ln \langle k \rangle} + \frac{1}{\langle k \rangle \ln \langle k \rangle} + \dots \quad (2.21)$$

A Eq. 2.21 relaciona o diâmetro da rede D com o número de sítios para uma rede completamente aleatória. Se tivermos que $N \gg 1$, podemos aproximar a expressão pelo primeiro termo apenas, de modo que $D \sim \ln N$. Como fizemos um tratamento médio durante esses cálculos e D depende somente dos caminhos de maior comprimento, muitas vezes essa equação é uma melhor aproximação para $\bar{\ell}$ do que para o próprio D . Se considerarmos redes que não sejam completamente aleatórias (como redes reais), algumas correções na equação são necessárias mas, em geral, o comportamento de mundo pequeno é mantido.

Em reticulados, a distância média entre os sítios da rede pode ser inferida geometricamente. Na Figura 2.7(a), temos um reticulado unidimensional com condições de contorno periódicas (anel). Nessa rede, a distância entre dois sítios quaisquer depende linearmente de N ($\bar{\ell} \sim N$). Já na rede quadrada da Figura 2.7(b), a dependência linear é com o tamanho do quadrado (L), e não com N . Porém, $L \sim N^{1/2}$ e assim: $\bar{\ell} \sim N^{1/2}$. Estendendo esse raciocínio para redes hipercúbicas de dimensão d , podemos relacionar $\bar{\ell}$ e N por:

$$\bar{\ell} \sim N^{1/d} \quad (2.22)$$

Em particular para redes de mundo pequeno, temos que $\bar{\ell} \sim \ln N$. Para relacionar com a dimensionalidade da rede, tomemos o limite

$$\lim_{N \rightarrow \infty} \frac{\ln N}{N^{1/\alpha}} = 0 \quad \alpha > 0. \quad (2.23)$$

Dessa forma, esse limite nos diz que não importa o quão grande seja α , o crescimento de $\ln N$ será sempre menor do que $N^{1/\alpha}$. Logo, podemos inferir que redes de mundo pequeno possuem dimensão infinita ($d = \infty$) no limite em que $N \rightarrow \infty$.

Analisando $P(k)$ de diversas redes reais, mostrados na Figura 2.8 e na Tabela 2.1, podemos notar que muitas redes complexas reais possuem $P(k) \sim k^{-\gamma}$ (pelo menos assintoticamente), em particular com $2 < \gamma < 3$, sendo altamente heterogêneas e com sítios com grau muito maior que a média. Redes com essas propriedades são chamadas de redes sem escala. Além disso, por muitas vezes também são redes de mundo pequeno. O fato de muitas redes reais terem essas propriedades em comum nos motiva a estudar redes sintéticas que capturem esses elementos.

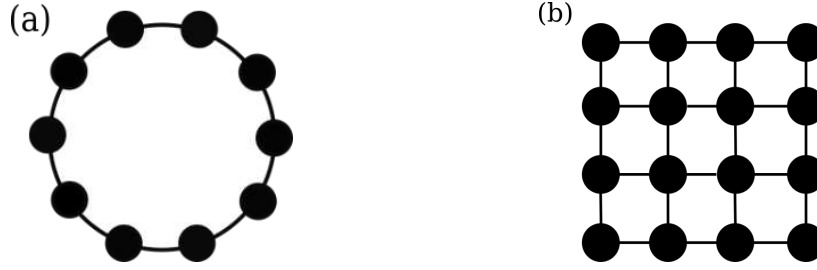


Figura 2.7 – (a) Exemplo de reticulado 1D com condições de contorno periódicas. (b) Exemplo de reticulado 2D.

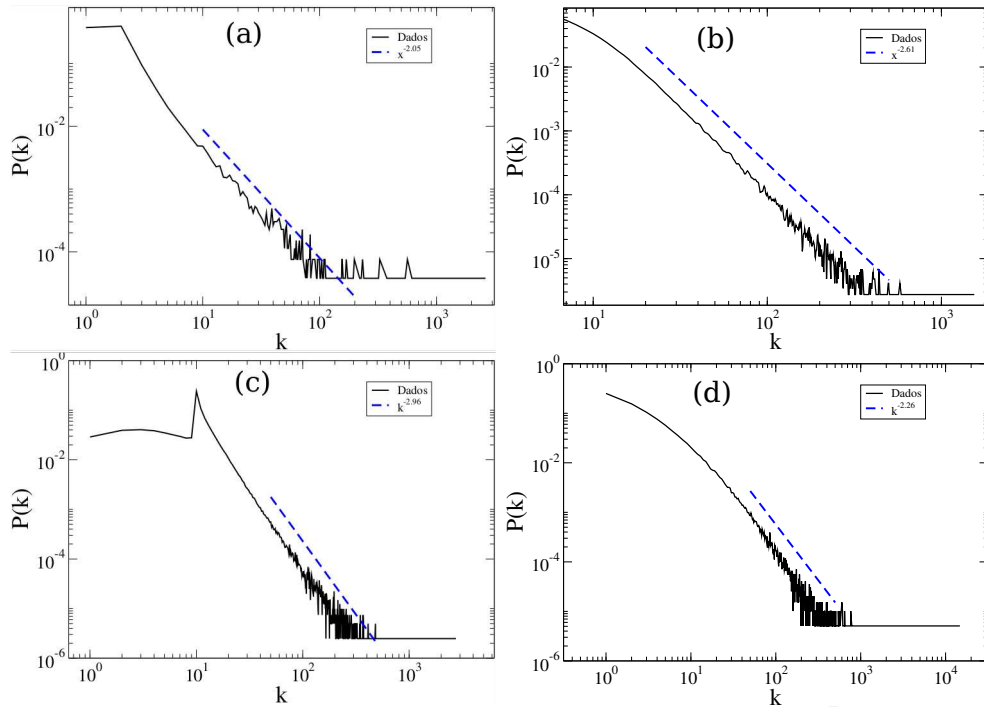


Figura 2.8 – Distribuição de conectividade para diversas redes reais. Retiradas de [21]. (a)ASCaida (b)CiteSeer (c)Amazon (d)Gowalla

	N	$\mathcal{E}$	D	$\langle l \rangle$	$\langle k \rangle$	$\langle k^2 \rangle$	$\langle k^2 \rangle / \langle k \rangle$	K_{nn}	k_{max}
ASCaida	26475	53381	17	3.88	4.03	1130	280.4	471	2628
CiteSeer	365154	1721981	34	6.46	9.43	457	48.5	54	1739
Amazon	403364	2443311	25	6.43	12.1	370	30.6	36	2752
Gowalla	196591	950327	15	4.63	9.66	2964	480	306.8	14730

Tabela 2.1 – Algumas propriedades topológicas e estatísticas de diversas redes reais.

2.2 Redes sem escala

Conforme visto na Figura 2.8, diversas redes reais apresentam $P(k)$ como uma lei de potência, sendo denominadas redes sem escala. Apesar da enorme diversidade em termos de constituintes, interações, natureza e origem dessas redes, grande parte delas apresenta estruturas similares. Dessa forma, podemos teorizar que existem mecanismos

fundamentais por trás da formação dessas redes.

Em 1999, Barabási e Albert [22] propuseram um modelo que seria capaz de gerar redes sem escala a partir de duas hipóteses simples:

- **Crescimento da rede:** Essa hipótese se baseia no fato de que N , em uma rede real, não é estacionário. Ele aumenta com o passar do tempo a medida que novos constituintes são adicionados. Em uma rede do Facebook, por exemplo, um novo vértice é adicionado à rede sempre que uma conta é criada e *links* são adicionados/removidos sempre que duas pessoas fazem/desfazem amizade.
- **Ligação preferencial:** Novos vértices adicionados à rede possuem uma preferência de se conectarem com vértices que possuam alta conectividade. Para ilustrar, consideremos uma rede de citações. Cientistas não conseguem ter acesso a todos os artigos que saem no mundo, focando muitas vezes nos mais citados ou nos autores com mais prestígio. Dessa forma, quanto maior o número de citações de um artigo, maior a chance deste ser lido e, por conseguinte, ser citado novamente em outros trabalhos.

Em resumo, o modelo de Barabási-Albert indica que esses dois mecanismos simples podem explicar a emergência de redes sem escala. A origem da lei de potência em $P(k)$ e o aparecimento de *hubs* são induzidas pela coexistência desses dois fenômenos. Na Figura 2.9 temos uma comparação entre redes completamente aleatórias e redes sem escala como as geradas pelo modelo de Barabási-Albert. Pode-se notar que enquanto no grafo aleatório os sítios possuem aproximadamente o mesmo grau $\langle k \rangle \approx 2$, no grafo sem escala há a presença de sítios mais conectados (na figura representados em cinza). Detalhes matemáticos sobre os mecanismos e a formação de redes sem escala podem ser encontrados no Capítulo IV de [17].

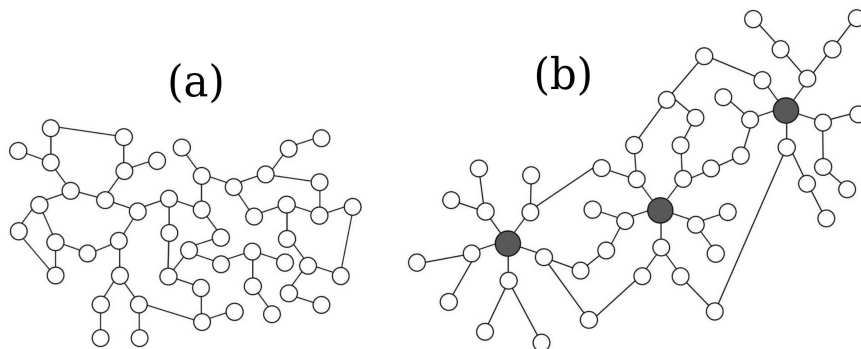


Figura 2.9 – Ilustração e comparação entre redes complexas (a) Modelo de Erdős-Rényi (Aleatória) e (b) Modelo de Barabási-Albert (Sem escala). Ambas redes possuem o mesmo número de *links* e sítios.

A partir das definições da seção anterior, podemos calcular algumas grandezas importantes para redes sem escala. Considerando uma distribuição de graus $P(k) = Bk^{-\gamma}$, temos pela condição de normalização

$$\sum_{k_{\min}}^{k_{\max}} P(k) = 1 \rightarrow B \sum_{k_{\min}}^{k_{\max}} k^{-\gamma} = 1. \quad (2.24)$$

Se considerarmos $k_{\max} \gg k_{\min}$, podemos substituir o somatório por uma integral

$$\sum_{k_{\min}}^{k_{\max}} k^{-\gamma} \simeq \int_{k_{\min}}^{k_{\max}} k^{-\gamma} dk = \frac{k_{\max}^{-\gamma+1} - k_{\min}^{-\gamma+1}}{-\gamma + 1}. \quad (2.25)$$

Dessa forma, encontramos um valor para B

$$B = \frac{-\gamma + 1}{k_{\max}^{-\gamma+1} - k_{\min}^{-\gamma+1}} \simeq (\gamma - 1)k_{\min}^{\gamma-1}. \quad (2.26)$$

Partindo da Eq. (2.3), podemos encontrar expressões para os momentos dessa distribuição, de forma parecida ao cálculo de B:

$$\langle k^m \rangle = \sum_{k_{\min}}^{k_{\max}} k^m P(k) \simeq \int_{k_{\min}}^{k_{\max}} k^{-\gamma+m} dk = B \frac{k_{\max}^{-\gamma+m+1} - k_{\min}^{-\gamma+m+1}}{-\gamma + m + 1}$$

Considerando $k_{\max} \gg k_{\min}$, podemos escrever:

$$\langle k^m \rangle \sim k_{\max}^{-\gamma+m+1} \quad (2.27)$$

Outra grandeza importante de ser calculada é o corte estrutural da rede (k_c) que é a cota superior para o grau dos vértices da rede que surge de sua própria estrutura ou imposto intencionalmente. Dada uma rede com N vértices, o número de nós com $k \geq k_c$ é

$$\phi(k_c) = N \int_{k_c}^{\infty} P(k) dk. \quad (2.28)$$

Dessa forma, impondo que $\phi(k_c) \simeq 1$ e resolvendo a Eq. (2.28) para k_c , encontramos

$$k_c \sim N^{\frac{1}{\gamma-1}} \quad (2.29)$$

Assim, o grau máximo observado será da mesma ordem de grandeza que k_c a menos que outros cortes mais rígidos sejam impostos. De posse dessa informação, podemos analisar os momentos de $P(k)$ para essas redes sem escala no limite termodinâmico em que $k_{\max} \rightarrow \infty$ para $N \rightarrow \infty$. Analisando a Eq. (2.27), temos :

- Para $\gamma > m + 1$, $\langle k^m \rangle$ é finito no limite termodinâmico.

- Para $\gamma < m + 1$, $\langle k^m \rangle$ diverge no limite termodinâmico.

Para o intervalo $2 < \gamma < 3$, temos que $\langle k \rangle$ é finito e todos momentos de ordem superior divergem. Dessa forma, a desvio padrão ($\sigma = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}$), medida que quantifica flutuações de k , também diverge. Portanto a distribuição de graus de uma rede sem escala é altamente heterogênea e a rede possui um número grande de *hubs*.

Comparativamente, em redes completamente aleatórias temos tipicamente que $\sigma_{\text{rnd}} \sim \langle k \rangle^{1/2}$ [17]. Dessa forma, k para essas redes encontra-se essencialmente no intervalo $\langle k \rangle \pm \langle k \rangle^{1/2}$. Assim, os vértices possuem graus comparáveis e $\langle k \rangle$ nos fornece uma estimativa quantitativa da escala de grau de um vértice qualquer da rede. Por outro lado, considerando redes com $P(k) \sim k^{-\gamma}$ com $2 < \gamma < 3$, as flutuações em k divergem e, escolhendo sítios ao acaso, podemos encontrar vértices com graus arbitrariamente grandes com uma probabilidade que é desprezível. Devido a esse fenômeno, surge a ideia de rede sem escala.

A presença de *hubs* em uma rede sem escala cria um questionamento interessante. Como esses vértices altamente conectados influenciam na propriedade de mundo pequeno em redes complexas? Conforme mostrado no Capítulo IV de [17], a presença de *hubs* na rede diminui ainda mais a distância média entre dois sítios na rede. De acordo com o valor de γ , essa distância possui diferentes dependências com N :

- **Regime Anômalo:** Esse regime ocorre quando $\gamma = 2$, e nesse caso, o corte cresce linearmente com N (vide Eq. 2.29). Dessa forma a rede consiste basicamente de um vértice altamente conectado ao centro e diversos vértices conectados a esse central e fracamente conectados entre si. Isso faz com que a distância média entre os vértices se torne pequena e constante com o aumento de N . Em outras palavras, praticamente todos os vértices são conectados ao vértice central, logo, a distância entre quaisquer pares de vértices está entre 1 e 2.
- **Ultra Small World:** Esse regime ocorre para $2 < \gamma < 3$. A existência de um alto número de *hubs* diminui a distância entre muitos pares de vértices, mas não de forma tão abrupta quanto no regime anômalo. Assim, a dependência encontrada é $\langle k \rangle \sim \ln \ln N$. Um crescimento muito mais lento do que redes aleatórias.
- **Ponto Crítico:** Esse caso ocorre para $\gamma = 3$ e é denominado ponto crítico pois divide a região dependente de $\ln N$ da região onde essa dependência cresce mais lentamente com N . Cálculos mostram que nesse regime, há uma correção na dependência de $\langle d \rangle$ comparado com redes aleatórias: $\langle d \rangle \sim \frac{\ln N}{\ln \ln N}$. Assim, a existência de poucos *hubs* diminui um pouco a distância entre os vértices.
- **Small World:** Para $\gamma > 3$, $\langle k^2 \rangle$ torna-se finito. Dessa forma, a distância entre os vértices volta a ter a dependência logarítmica com N e a propriedade de mundo pequeno usual.

III Epidemias

A modelagem matemática de epidemias é um campo ativo de pesquisa com interdisciplinaridade entre várias áreas. Epidemiologistas, cientistas da computação, físicos, matemáticos, cientistas sociais, dentre outros compartilham um interesse em estudar fenômenos de propagação de epidemias ou de informação e se apoiam em modelos matemáticos para descrevê-los. Normalmente, os modelos epidêmicos estão preocupados com a evolução do número e a localização dos vértices infectados na população em função do tempo ou a forma como a epidemia se propaga a partir de um foco inicial. Além disso, procura-se entender as propriedades da epidemias em tempos longos, como a densidade de sítios infectados, surtos de epidemia, pandemias, dentre outros.

Em termos de segurança pública, analisar as propriedades de uma epidemia é crucial para prever a forma de propagação desta, projetar cenários e discutir as melhores estratégias para contenção de uma doença. A partir de simulações de modelos epidêmicos utilizando dados reais (taxas de transmissão, redes de aeroportos, etc) durante o surto de Ebola em 2014, Vespignani *et al* [7] analisou a eficácia das políticas públicas implementadas pelos EUA e concluiu que elas seriam ineficientes a longo prazo. Dessa forma, uma reestruturação foi feita e após novos estudos, a propagação da doença foi contida. Além disso, diversos outros estudos [8, 9] analisam e discutem os surtos epidêmicos de doenças reais.

A partir desses exemplos, salientamos a importância de se estudar epidemias em redes complexas de forma a prever e conter surtos epidêmicos, determinar melhores estratégias de imunização e entender os mecanismos de propagação de doenças reais. Apesar de originalmente terem sido propostos para estudar epidemias [23], os modelos aqui estudados podem ser mapeados em diversos outros fenômenos fora do ramo epidemiológico, como propagação de informação ou redes neuronais [24].

3.1 Modelagem

Geralmente, descrevemos um sistema em uma rede atribuindo uma variável de estado σ_i a cada vértice i , com essa variável representando uma determinada característica do vértice. No contexto de epidemias, essa variável pode indicar se o indivíduo está infectado, em que o vértice é capaz de espalhar a epidemia ou suscetível, em que o sítio pode tornar-se infectado. Agrupando os estados de todos os sítios da rede, temos um estado microscópico do sistema, que pode ser representado por $\sigma(t) = (\sigma_1(t), \sigma_2(t), \dots, \sigma_N(t))$.

Podemos estudar a dinâmica desses sistemas através das transições espontâneas entre microestados $\sigma \rightarrow \sigma'$ determinadas por taxas $w_{\sigma \rightarrow \sigma'}$. Se definirmos $P(\sigma, t)$ como a probabilidade de encontrar o sistema no estado σ no tempo t e considerando $W_{\sigma \rightarrow \sigma'}$ a

taxa com que ocorre a transição de um estado σ para um estado σ' , a dinâmica do sistema pode ser escrita na forma de uma equação mestra

$$\frac{\partial P(\sigma, t)}{\partial t} = \sum_{\sigma'} W_{\sigma' \rightarrow \sigma} P(\sigma') - \sum_{\sigma} W_{\sigma \rightarrow \sigma'} P(\sigma), \quad (3.1)$$

que descreve o fluxo de probabilidades saindo ou entrando em um determinado micro-estado σ . A partir da solução dessa equação, pode-se calcular os valores esperados de qualquer grandeza relevante no sistema. Considerando uma grandeza $B(t)$, temos que

$$\langle B(t) \rangle = \sum_{\sigma} B(\sigma) P(\sigma, t). \quad (3.2)$$

Porém, na grande maioria dos casos, não é possível encontrar a solução formal da equação mestra, de forma que abordagens alternativas aproximadas são necessárias. Dentre estas, uma abordagem muito utilizada é a simulação computacional, na qual as taxas são convertidas em probabilidades, que são implementadas utilizando números aleatórios. A partir disso, é possível calcular numericamente a distribuição de estados de um processo associando às transições entre estados uma determinada probabilidade. No limite de tempos longos, espera-se que a distribuição de estados convirja para a solução estacionária da equação mestra.

Nesse trabalho, consideramos três estados possíveis para os sítios da rede: Suscetível (S), onde o vértice está saudável e pode tornar-se infectado, Infectado (I), onde o vértice possui a doença e pode disseminar a infecção para seus vizinhos e Recuperado (R), estado no qual o sítio não pode ser infectado e nem transmite a doença. A partir desses três estados, definimos as taxas de transição entre estes:

- **Infecção:** Um vértice j no estado I infecta cada um de seus vizinhos suscetíveis independentemente com taxa λ_j .
- **Recuperação:** Um vértice j no estado I torna-se espontaneamente R com taxa μ_j
- **Perda de Imunidade:** Um vértice j no estado R torna-se espontaneamente S com taxa α_j

Na Figura 3.1, temos exemplos das transições entre os estados. A partir dessas taxas, podemos definir diferentes modelos epidêmicos. Entre os mais conhecidos e estudados estão:

- **Suscetível-Infectado-Suscetível (SIS) com taxas uniformes:** $\lambda_i = \lambda \forall i$, $\alpha_i = \infty$ (perda instantânea de imunidade) e $\mu_i = 1$. Esse modelo é utilizado para o estudo de doenças infecciosas que possuem um estado endêmico. Nesse modelo, os *hubs* têm papel fundamental, já que eles são capazes de disseminar a epidemia para um grande número de sítios.

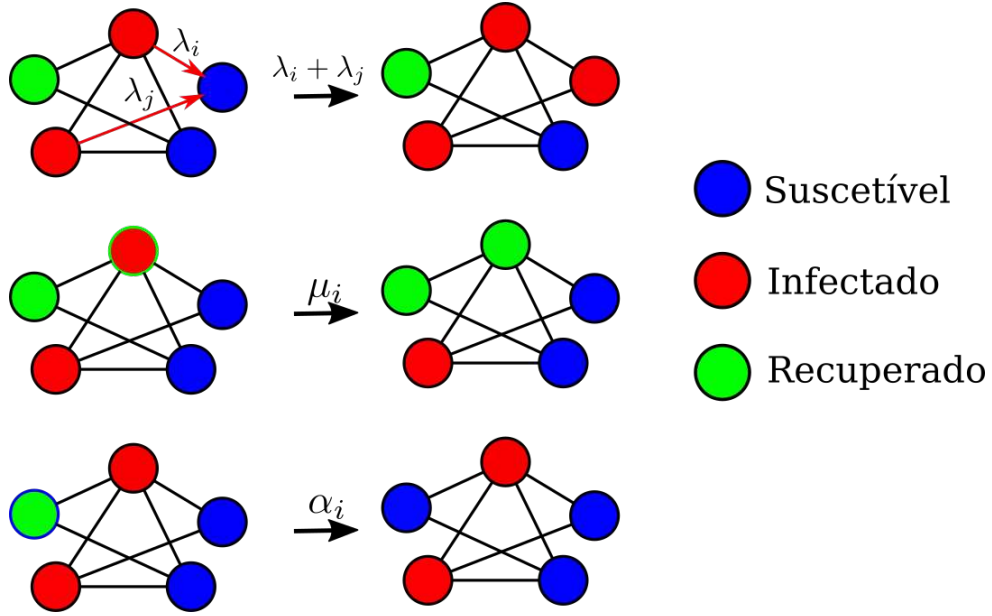


Figura 3.1 – Ilustração dos tipos de transição entre os diferentes estados em uma epidemia (a) Sítios infectados infectam um vizinho em comum com taxa $\lambda_i + \lambda_j$. (b) Sítio infectado espontaneamente torna-se recuperado com taxa μ_i . (c) Sítio recuperado espontaneamente torna-se suscetível com taxa α_i .

- **Processo de Contato (PC):** $\lambda_i = \lambda/k_i$, $\alpha_i = \infty$ e $\mu_i = \mu$. Modelo clássico proposto por T. E. Harris em 1974 [25] para propagação epidêmica, consolidando-se como um dos processos dinâmicos mais simples a apresentar transições de fase com estado absorvente.
- **Suscetível-Infectado-Recuperado (SIR):** $\lambda_i = \lambda$, $\alpha_i = 0$ e $\mu_i = \mu$. Esse modelo leva em conta a possibilidade do indivíduo ser removido da dinâmica através de morte ou imunidade adquirida [26].

Dentre esses diversos modelos, estudamos principalmente o modelo SIS [27], que apresenta fenômenos interessantes em sua dinâmica. Nesse modelo, temos que a taxa com que um vértice propaga a infecção é λn_s , em que n_s é o número de contatos suscetíveis do vértice. Dessa forma, quanto maior o grau de um vértice, maior será a propagação da epidemia através dele. Por isto, *hubs* apresentam papel fundamental nesse modelo e podem levar a surtos de epidemias que são localizados (apenas um *hub* e sua vizinhança estão ativos) mas que atuando coletivamente podem levar a ativação de fases endêmicas. Esses surtos serão mais explorados na próxima seção.

3.2 Epidemias e transições de fases

Em geral, modelos epidêmicos possuem uma mudança de comportamento conforme variamos algum parâmetro de controle. Considere o modelo SIS com taxa de cura fixa

$\mu = 1$ (determina a escala de tempo) e taxa de infecção λ em uma rede que possua um corte superior para o grau. Para λ suficientemente pequeno, a probabilidade de um surto epidêmico é muito baixa e, efetivamente, a densidade de sítios infectados (ρ) tende a zero. Porém, com o aumento de λ , a probabilidade de um surto epidêmico e da epidemia se sustentar por longos períodos cresce, até ρ tornar-se uma fração finita da rede (estamos considerando o limite termodinâmico, onde $N \rightarrow \infty$). Essa mudança de uma fase inativa, em que $\rho \rightarrow 0$ para uma fase ativa, com ρ finito caracteriza uma transição de fase, com λ sendo o parâmetro de controle dos modelos e ρ um possível parâmetro de ordem. O valor de λ onde ocorre essa transição é denominado limiar epidêmico e representado por λ_c . Uma representação da mudança de comportamento em ρ pode ser vista na Figura 3.2.

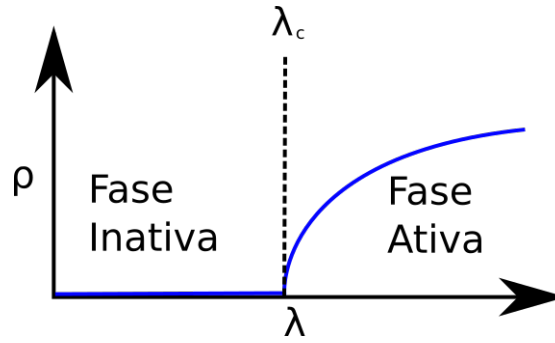


Figura 3.2 – Esquema da transição de fase que ocorre no modelo SIS. Note a mudança de comportamento ao variar o parâmetro de controle λ .

Um detalhe importante sobre esses modelos epidêmicos é o fato da infecção não surgir de forma espontânea na rede. Assim, uma configuração onde nenhum sítio se encontra no estado I não possui atividade. Esse tipo de estado é chamado estado absorvente. No geral, estados absorventes referem-se a estados de um sistema que aprisionam a dinâmica deste. Uma vez alcançado, o sistema fica aprisionado neste estado indefinidamente. Como exemplo, consideremos a Figura 3.3. Neste sistema, temos três estados possíveis, dois deles são estados transientes e um estado é absorvente. O estado 1 pode transitar para os estados 2 e 3, o estado 2 pode transitar para os estados 1 e 3, mas o estado 3 não transita para nenhum dos outros dois estados. Assim, podemos dizer que o estado 3 absorveu a dinâmica desse sistema e é denominado estado absorvente. Voltando para epidemias, em particular para os modelos SIS e SIR, qualquer estado que não apresente nenhum vértice no estado infectado é um estado absorvente, visto que as transições $S \rightarrow I$ e $R \rightarrow I$ não são possíveis em tais estados.

Ao se estudar transições de fase em fenômenos de não equilíbrio em redes regulares (como redes hipercúbicas), as correlações espaciais e temporais desempenham papel importante na caracterização e identificação das transições. De forma simplificada, as correlações espaciais indicam como o estado de um vértice influencia o estado de outros vértices enquanto as correlações temporais representam a influência do estado de um sítio em tempos diferentes. As escalas dessa influência são denominados comprimentos de

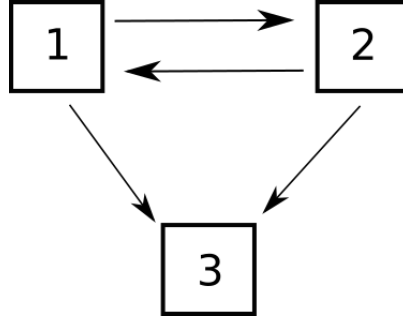


Figura 3.3 – Esquema de um sistema com 3 estados, sendo um desses absorvente. Os estados 1 e 2 possuem fluxo de entrada e de saída (representado pelas setas) e são chamados estados transientes. Já o estado 3 possui apenas fluxo de entrada, uma vez neste estado, o sistema permanece ali para sempre, caracterizando um estado absorvente.

correlação temporal $\xi_{\parallel}$ e espacial $\xi_{\perp}$. É verificado que nos pontos críticos (onde ocorrem as transições de fase), as influências são de longo alcance, sem um comprimento característico no limite termodinâmico e, efetivamente, há uma divergência dos comprimentos de correlação. Na Figura 3.4, temos uma ilustração visual dos comprimentos de correlação típicos no modelo SIS. As dimensões dos diferentes agregados de vértices (infectados ou suscetível) representam os comprimentos de correlação, tanto espacialmente quanto temporalmente.

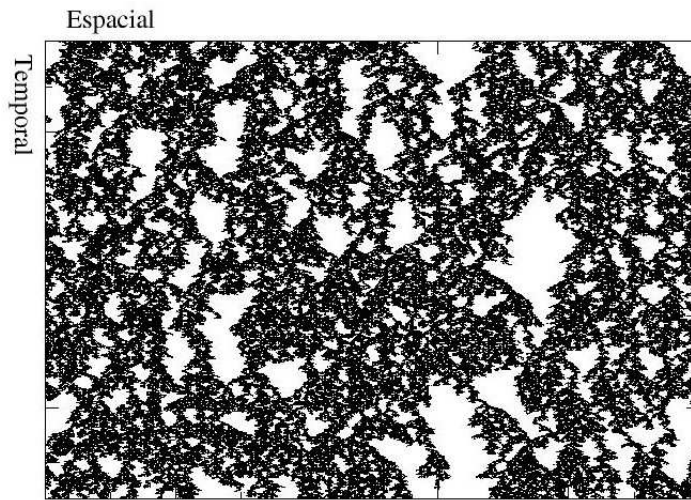


Figura 3.4 – Representação esquemática dos comprimentos de correlação para uma simulação do modelo SIS no seu ponto crítico $\lambda_c = 1.6389$ em um reticulado unidimensional com condições de contorno periódicas. Sítios em preto são infectados e sítios brancos suscetíveis. As dimensões dos diferentes *clusters* representam os comprimentos de correlação, em seus respectivos eixos.

O conceito de universalidade é uma ideia importante no estudo de transições de fase de qualquer tipo. No âmbito de física estatística, refere-se a observação de sistemas com interações microscópicas distintas mas compartilhando simetrias que não dependem dessas

interações. A partir disso, visamos associar o comportamento crítico dessas transições a um número finito de classes, denominadas classes de universalidade. Como exemplo, podemos citar a classe de universalidade Kardar-Parisi-Zhang (KPZ), que engloba diversos fenômenos de deposição e crescimento em superfícies, sistemas de partículas interagentes e polímeros em meios aleatórios [28]. No contexto de transições de fase, um ponto crítico é um valor do parâmetro de controle do sistema no qual ocorre efetivamente a transição de fases contínua que são caracterizadas pela ausência de comprimentos característicos e seu comportamento crítico é descrito por leis de escala definidas próximos da criticalidade. Logo, o conhecimento destes expoentes é fundamental.

Considerando estados estacionários ($t \rightarrow \infty$) próximos ao ponto crítico λ_c , observamos um comportamento para a densidade de sítios ativos (ρ) dado por

$$\bar{\rho} \sim |\lambda - \lambda_c|^\beta, \quad (3.3)$$

com $\bar{\rho} = \rho(\infty)$ e β um expoente crítico associado ao parâmetro de ordem $\bar{\rho}$.

Além disso, os comprimentos de correlação espacial e temporal seguem as respectivas leis de potência

$$\xi_\perp \sim |\lambda - \lambda_c|^{-\nu_\perp} \quad (3.4)$$

e

$$\xi_\parallel \sim |\lambda - \lambda_c|^{-\nu_\parallel}, \quad (3.5)$$

em que novamente, $\nu_\perp$ e $\nu_\parallel$ são expoente críticos associados a classe de universalidade dos modelos.

Devido a finitude dos sistemas simulados, os estados absorventes são sempre visitados se tempos suficientemente longos são considerados (na fase super-crítica esse tempo cresce exponencialmente com o tamanho). Dessa forma, precisamos de estratégias para contornar esses problemas. Uma abordagem a ser utilizada é a dos **estados quase estacionários**, definidos em um contexto geral como perturbações na dinâmica do sistema de forma a evitar o estado absorvente (detalhes na seção V.3). Dessa forma, podemos construir séries temporais de ρ sem o empecilho do estado absorvente ser visitado. Médias durante o estado quase estacionário serão representados por $\langle \dots \rangle$.

A suscetibilidade (variância da densidade de sítios no estado estacionário) também é uma grandeza útil e definida por

$$\chi_N \equiv N(\langle \rho^2 \rangle - \langle \rho \rangle^2). \quad (3.6)$$

χ_N nos fornece uma medida das flutuações de ρ . Na Figura 3.5, mostramos uma série temporal quase estacionária de ρ e a diferença em suas flutuações a medida que saímos de um regime subcrítico ($\lambda < \lambda_c$), passamos por um regime crítico ($\lambda \approx \lambda_c$) e chegamos em um regime supercrítico ($\lambda > \lambda_c$). O aumento das flutuações em ρ faz com que o gráfico de

χ versus λ apresente um pico em torno do valor λ_c .

Para o modelo SIS em redes complexas, é necessário uma adaptação na Eq. (3.6), devido aos efeitos de localização da rede [29]. Essa nova suscetibilidade é denominada suscetibilidade dinâmica e é definida por

$$\chi_N = N \frac{\langle \rho^2 \rangle - \langle \rho \rangle^2}{\langle \rho \rangle}. \quad (3.7)$$

A partir de agora, deixaremos de lado a definição tradicional da suscetibilidade e utilizaremos somente a suscetibilidade dinâmica.

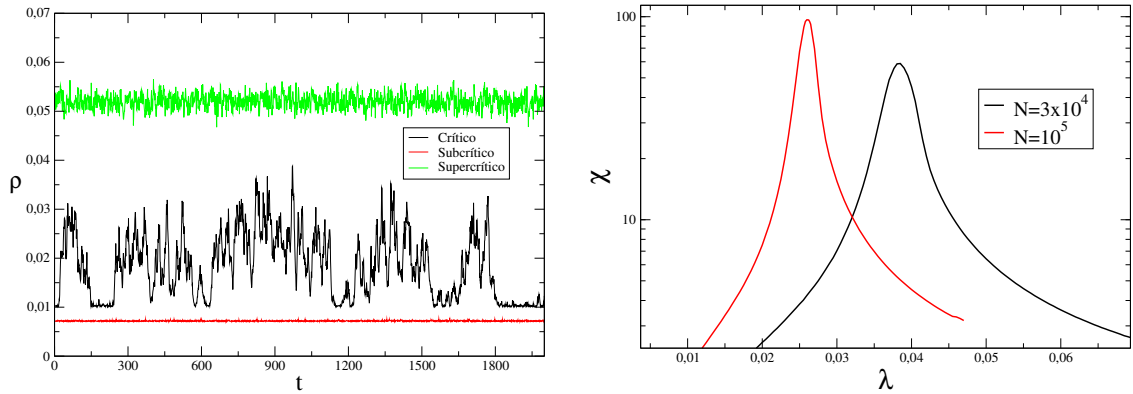


Figura 3.5 – (a) Série temporal de ρ mostrando as flutuações dessa grandeza estacionária em diferentes regimes. (b) Suscetibilidade dinâmica ao variarmos o parâmetro de controle para duas redes sem escala de diferentes tamanhos ($\gamma = 2.80$). Note o pico nos gráficos, indicando os pontos onde ocorrem as transições.

Ao estudar transições de fase, observa-se que os comprimentos de correlação $\xi_{\perp}$ e $\xi_{\parallel}$ divergem no limiar epidêmico. Conforme já foi dito, o comprimento de correlação espacial está diretamente ligado às correlações entre os vértices da rede. Como fazemos simulações em sistemas de tamanho finito, há um valor limite para $\xi_{\perp}$, que seria o comprimento da rede L . Dessa forma, o limiar epidêmico encontrado em simulações de tamanho finito (λ_p) não é o ponto crítico do sistema no limite termodinâmico. Porém, o comportamento de $\xi_{\perp}$ nos provê uma maneira de extrapolar λ_c . Partindo da Eq. (3.4), quando $\lambda \rightarrow \lambda_p$ temos $\xi_{\perp} \approx L$ e isso implica em

$$L \sim (\lambda_p - \lambda_c)^{\nu_{\perp}} \rightarrow \lambda_p - \lambda_c \sim L^{-1/\nu_{\perp}} \quad (3.8)$$

Assim, conhecendo $\nu_{\perp}$ podemos encontrar λ_c a partir do valor de λ_p para sistemas de tamanhos diferentes. Essa técnica é chamada de análise de tamanho finito ou escalonamento de tamanho finito (ETF) e pode ser utilizada para encontrar a dependência de diversas grandezas de interesse com o tamanho

$$\rho_p(L) \sim L^{-\beta/\nu_{\perp}} \quad (3.9)$$

e

$$\chi_p(L) \sim L^{-\gamma/\nu_\perp}. \quad (3.10)$$

Conforme foi dito na seção anterior, a presença de outliers em uma rede pode acarretar em fenômenos interessantes no modelo SIS. Esses outliers e seus vizinhos formam uma estrutura que pode ser aproximada por uma rede do tipo estrela composta por um vértice altamente conectado ao centro e diversos sítios com k pequeno conectados neste (folhas). A medida que variamos o parâmetro de controle, pode existir um valor $\lambda_{\text{star}} < \lambda_p$ no qual uma estrela consegue se manter ativa por um longo período enquanto a rede está globalmente inativa. Dessa forma, as flutuações de ρ podem capturar essa ativação local e χ apresentará um pico referente a esta. Esse fenômeno não se restringe somente a uma única estrela por rede. Sob determinadas condições várias regiões podem se ativar localmente em diferentes valores λ_{star} e a suscetibilidade apresentar múltiplos picos [30]. Esse tipo de fenômeno é comum em redes em lei de potência com $\gamma > 3$, onde apesar da baixa probabilidade de aparecimento de hubs, eventualmente um sítio muito mais conectado que a maioria aparece e é responsável por esta localização. Na Figura 3.6(a) temos uma distribuição de graus em lei de potência para uma rede complexa sintética com $\gamma = 3.5$ e outliers (em destaque). Já na Figura 3.6(b), temos χ em função de λ para esta mesma rede ilustrando os múltiplos picos.

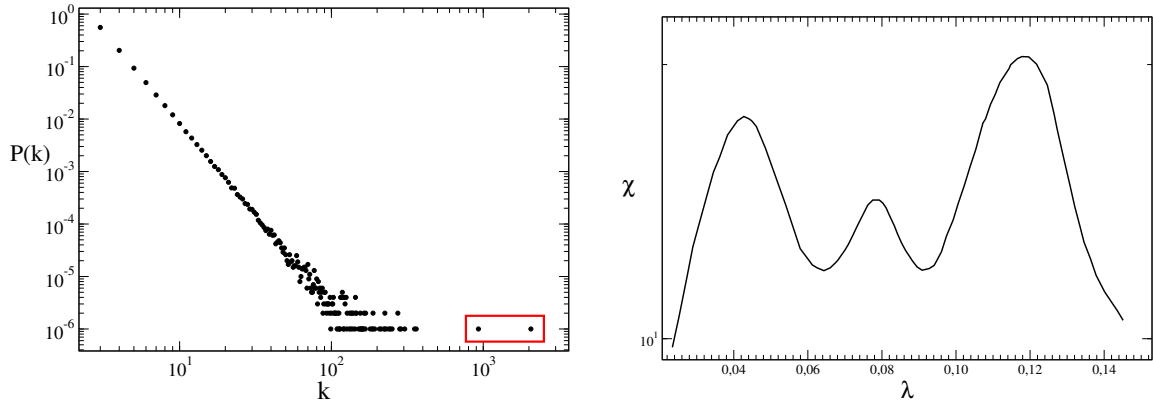


Figura 3.6 – (a) Distribuição de graus para uma rede com $\gamma = 3.5$. Em destaque, temos dois *hubs*. (b) Suscetibilidade em função de λ para a rede mostrada em (a). Note a existência de múltiplos picos, representando as ativações locais (a esquerda) e global (mais a direita)

Para o modelo SIR, que não possui estado estacionário ativo, outra abordagem é utilizada. Começando com apenas um sítio infectado, podemos analisar como a epidemia se propaga. Se λ for baixo, poucos sítios serão infectados antes da epidemia se extinguir. Lembre que no SIR qualquer estado composto apenas por sítios S e R é um estado absorvente. Por conseguinte, a densidade de sítios recuperados também será baixa. Porém, a medida que aumentamos λ , há a possibilidade da infecção se espalhar para uma fração finita da rede. Assim, a densidade de sítios recuperados também aumenta. Dessa forma,

a densidade final ρ_r pode ser utilizada como parâmetro de ordem do modelo e possui comportamento análogo ao da Figura 3.2 (no limite termodinâmico).

De forma similar, ρ_r também apresenta flutuações que crescem próximo ao limiar epidêmico. Nesse caso, utilizamos a variabilidade epidêmica [31] para encontrar λ_p

$$\Delta = \frac{\sqrt{\langle \rho_r^2 \rangle - \langle \rho_r \rangle^2}}{\langle \rho_r \rangle}. \quad (3.11)$$

A princípio, os conceitos mostrados aqui foram originalmente desenvolvidos para se estudar modelos epidêmicos em reticulados. Porém, as ideias podem ser estendidas para redes complexas através de pequenas adaptações (como feito na suscetibilidade) como utilizar o número de vértices N em vez do tamanho do reticulado L . Com o ferramental teórico obtido pela análise da transição de fase e a partir da ETF, a identificação das transições utilizando simulações é feita analisando grandezas relevantes ($\rho, \rho_r, \chi, \Delta$). Cada uma dessas quantidades possui um comportamento em torno do ponto crítico, conforme discutido anteriormente e, quando utilizadas em conjunto, contribuem para uma maior precisão na identificação dos limiares epidêmicos dos modelos.

3.3 Teorias de campo médio

Uma forma de estudar processos dinâmicos em redes complexas é através de **teorias de campo médio**. No contexto de redes complexas, correlações dinâmicas entre os vértices da rede são desprezadas por essas aproximações. Diversas teorias de campo médio foram desenvolvidas, cada uma levando diferentes fatores em consideração e, aplicando-as aos modelos epidêmicos, podemos obter aproximações para o limiar epidêmico.

3.3.1 Teoria de campo médio homogênea

Uma primeira aproximação para estudarmos processos epidêmicos em redes complexas é considerar todos os vértices como possuindo o mesmo grau $\langle k \rangle$ e assumindo que a densidade de sítios infectados independe das propriedades da rede, ou seja, estamos considerando um sistema homogêneo. A partir dessas considerações, podemos montar uma equação dinâmica para o modelo SIS

$$\frac{d\rho}{dt} = -\rho + \lambda \langle k \rangle \rho (1 - \rho) \quad (3.12)$$

em que do lado direito da equação temos o termo de cura espontânea ($-\rho$) com $\mu = 1$ e o termo relacionado a infecção de vértices suscetíveis via contato com um vértice infectado. Apesar dessa equação possuir solução exata, vamos ilustrar o método de análise para as próximas teorias. Faremos uma análise de estabilidade linear para determinar o comportamento dessa equação na vizinhança da solução estacionária $\rho = 0$. Dessa forma,

linearizando a equação em torno desse ponto fixo, temos

$$\frac{d\rho}{dt} = \rho(-1 + \lambda\langle k \rangle),$$

cuja solução é $\rho(t) \sim e^{(-1+\lambda\langle k \rangle)t}$. A solução será estável se $-1 + \lambda\langle k \rangle < 0$ e instável se $-1 + \lambda\langle k \rangle > 0$. Dessa forma, o limiar epidêmico de acordo com essa aproximação ocorre em $\lambda_c^{\text{SIS}} = 1/\langle k \rangle$.

3.3.2 Teoria de campo médio heterogênea

Na aproximação de campo médio heterogênea (HMF, do inglês *heterogeneous mean field*), tratamos todos os sítios de mesmo grau como sendo equivalentes, desprezando correlações dinâmicas e a localização explícita do vértice na rede [23]. Dessa forma, a densidade de sítios infectados possui dependência somente em k . A partir disso, podemos escrever a equação dinâmica para o modelo SIS [32]

$$\frac{d\rho_k}{dt} = -\rho_k + \lambda k(1 - \rho_k) \sum_{k'} P(k'|k) \rho_{k'}, \quad (3.13)$$

em que o primeiro termo do lado direito continua representando a cura espontânea de vértices infectados enquanto o segundo termo representa a infecção de vértices suscetíveis em contato com vizinhos infectados. A presença do termo $P(k'|k)$ inclui a correlação de graus (seção II.1) existente entre os vértices. Novamente, realizando uma análise de estabilidade linear em torno da solução estacionária da equação ($\rho_k \approx 0$), encontramos

$$\frac{d\rho_l}{dt} = \sum_{k'} J_{kk'} \rho_{k'}, \quad (3.14)$$

em que J é a matriz Jacobiana do sistema é

$$J_{k'k} = -\delta_{k'k} + \lambda k P(k'|k). \quad (3.15)$$

Sabemos de teoria de equações lineares que o ponto fixo será estável se o maior autovalor da matriz Jacobiana for negativo e instável caso contrário [33]. Dessa forma, precisamos encontrar o maior autovalor da matriz de correlação $C_{kk'} = kP(k'|k)$ para determinar o limiar epidêmico. Considerando redes não correlacionadas, $P(k'|k)$ assume a forma da Eq. (2.14) e a matriz $C_{kk'}$ tem como um de seus autovetores $v_k = k$, com autovalor associado $\Lambda = \langle k^2 \rangle / \langle k \rangle$ [32]. Como a matriz $C_{kk'}$ é positiva e simétrica, podemos utilizar o **teorema de Perron-Frobenius** [19] segundo o qual uma matriz quadrada irredutível não negativa sempre possui um autovalor real positivo Λ_{max} e o módulo de nenhum outro autovalor excede Λ_{max} . Além disso, o autovetor correspondente a Λ_{max} possui todas as entradas positivas.

Portanto, $C_{kk'}$ satisfaz as condições de utilização do teorema e conseguimos garantir que $\Lambda = \langle k^2 \rangle / \langle k \rangle$ é o maior autovalor dessa matriz. Assim, conseguimos determinar o

limiar epidêmico $\lambda_c = \langle k \rangle / \langle k^2 \rangle$.

3.3.3 Teoria de campo médio microscópica

Nessa aproximação, a estrutura da rede aparece de forma explícita ao introduzirmos a matriz de adjacência na equação dinâmica. Ela é conhecida na literatura em inglês por QMF (*quenched mean field*), sigla que será usada no restante da dissertação pois não encontramos uma tradução adequada para este termo em português. Dessa forma, para o modelo SIS, podemos analisar a densidade de infectados para um sítio i

$$\frac{d\rho_i}{dt} = -\rho_i + \lambda(1 - \rho_i) \sum_j A_{ij} \rho_j. \quad (3.16)$$

Realizando uma análise de estabilidade linear em torno da solução estacionária ($\rho_i = 0$), temos

$$\frac{d\rho_i}{dt} = \sum_j J_{ij} \rho_j, \quad (3.17)$$

em que a matriz Jacobiana é

$$J_{ij} = -\delta_{ij} + \lambda A_{ij}. \quad (3.18)$$

Note que para redes não direcionadas e conectadas, A satisfaz as condições de aplicação do Teorema de Perron-Frobenius. Usando novamente análise de estabilidade linear, temos que o limiar epidêmico será dado pelo inverso do maior autovalor da matriz de adjacência $\Lambda_{\max}$. Para redes aleatórias não correlacionadas com $P(k) \sim k^{-\gamma}$, $\Lambda_{\max}$ foi calculado em [34], levando às seguintes estimativas para λ_c [35]:

$$\lambda_c = \frac{1}{\Lambda_{\max}} \simeq \begin{cases} \langle k \rangle / \langle k^2 \rangle & \text{se } 2 < \gamma < 5/2 \\ 1/\sqrt{k_c} & \text{se } 5/2 < \gamma \end{cases}$$

O limiar epidêmico encontrado pela teoria HMF é $\lambda_c = \langle k \rangle / \langle k^2 \rangle$. Na seção II, calculamos os valores de $\langle k \rangle$ e $\langle k^2 \rangle$ e encontramos que para $\gamma < 3$, $\langle k^2 \rangle$ diverge no limite termodinâmico, enquanto $\langle k \rangle$ permanece finito. Dessa forma, para $2 < \gamma < 3$, temos que $\lambda_c^{HMF} \rightarrow 0$ a medida que $N \rightarrow \infty$ enquanto que para $\gamma > 3$, λ_c^{HMF} torna-se finito. Para $2 < \gamma < 5/2$, o resultado para a teoria QMF é equivalente ao da teoria HMF. Dessa forma, o comportamento de λ_c^{QMF} é idêntico ao λ_c^{HMF} nessa faixa de γ . Para $\gamma > 5/2$, calculamos a dependência de k_c com N na seção II e vimos que $k_c \sim N^{\frac{1}{\gamma-1}}$ para redes não correlacionadas. Dessa forma

$$\lambda_c^{QMF} \sim N^{-\frac{1}{2(\gamma-1)}} \quad (3.19)$$

e no limite termodinâmico, $\lambda_c^{\text{QMF}} \rightarrow 0$.

Assim, para $\gamma > 3$, há uma divergência marcante entre os resultados das teorias HMF e QMF. Enquanto a primeira prevê um limiar epidêmico finito no limite termodinâmico, a segunda prevê um limiar indo para zero. Além disso, mesmo para $5/2 < \gamma < 3$, há divergências em relação a dependência de λ_c com o tamanho N , apesar de ambas preverem limiares indo para zero.

Na Figura 3.7, temos um comparativo entre os limiares epidêmicos para o modelo SIS em redes UCM (seção V) com $\gamma = 2.30$ e $\gamma = 2.80$ e diferentes tamanhos esperados pelas diferentes teorias e os valores encontrados através de simulações do modelo. Pode-se observar que o limiar previsto pela teoria de campo médio homogênea possui leis de escala diferentes do encontrado pelas simulações enquanto os limiares previstos pela HMF e QMF possuem leis de escala próximas, convergindo para um valor comum considerando $\gamma = 2.3$, onde as análises mostram que essas teorias são equivalente e divergindo para as redes com $\gamma = 2.80$, onde as análises mostram que as teorias diferem.

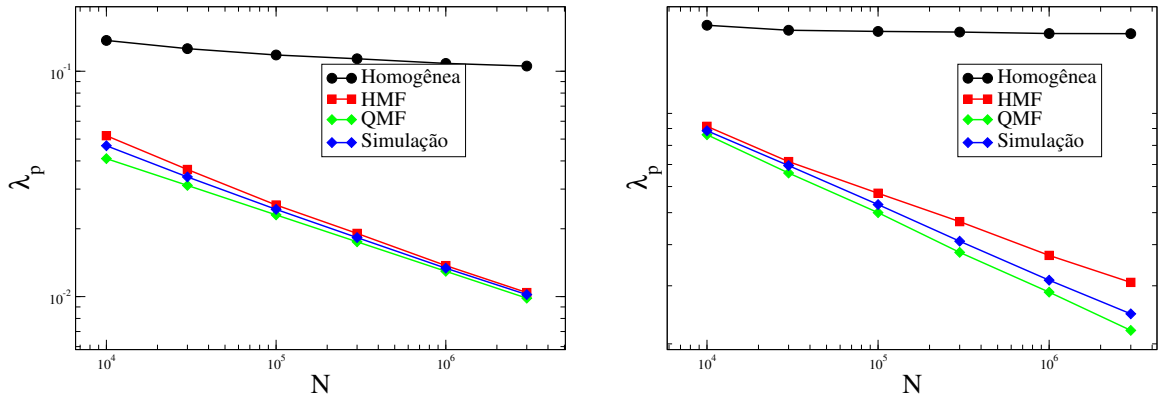


Figura 3.7 – Limiares epidêmicos encontrados pelas diferentes teorias de campo médio em redes complexas com diferentes tamanhos e expoentes (a) $\gamma = 2.30$ (b) $\gamma = 2.80$.

3.3.4 Teorias de campo médio para o modelo SIR

Podemos formular teorias de campo médio HMF e QMF para o modelo SIR de forma análoga ao modelo SIS, porém, agora temos que escrever equações para a densidade de sítios infectados ρ^I e recuperados ρ^R . Assim, pela teoria HMF, chegamos às equações

$$\frac{d\rho_k^I}{dt} = -\rho_k^I + \lambda k \rho_k^S \sum_k' \frac{k' - 1}{k'} P(k'|k) \rho_{k'}^I \quad (3.20)$$

$$\frac{d\rho_k^R}{dt} = \rho_k^I \quad (3.21)$$

em que o termo $\frac{k'-1}{k'}$ representa o fato de um vértice não poder infectar o nó que o infectou. Dessa forma é necessário acrescentar esse fator na taxa de infecção de vértices. Linear-

zando essa equação seguindo os métodos já discutidos nessa dissertação e lembrando que temos um sistema de equações, chegamos ao resultado do limiar epidêmico [36]

$$\lambda_c^{\text{HMF}} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \quad (3.22)$$

Além disso, podemos realizar um cálculo similar utilizando a teoria QMF e chegar a um resultado idêntico ao SIS para o limiar epidêmico [36]

$$\lambda_c^{\text{QMF}} = \frac{\langle k \rangle}{\langle k^2 \rangle}. \quad (3.23)$$

3.4 A teoria BCPS

Em 2013, Bogunã *et al* [37] propuseram um mecanismo por trás do limiar nulo do modelo SIS. Esse mecanismo se baseia na comparação entre o tempo de vida/atividade de um *hub* de grau k (τ_k^{rec}) com a escala de tempo com a qual um *hub* infectado de grau k transmite a infecção para outro *hub* suscetível de grau k' ($\tau_{k,k'}^{\text{inf}}$). Essa teoria ficou conhecida como teoria BCPS (em homenagem aos sobrenomes de seus desenvolvedores: Boguna, Castellano e Pastor-Satorras) e será brevemente discutida nessa seção da forma como apresentada em [38].

Inicialmente, vamos calcular aproximadamente o tempo de vida de um *hub* na rede. Para fazer isso, vamos tratar o *hub* e sua vizinhança como um grafo estrela. Além disso, estudaremos o modelo SIS um pouco modificado: o processo de cura será tratado de forma discreta (a cada passo de tempo $t = 1/\mu$, todos os sítios infectados se tornam suscetíveis) enquanto que o processo de infecção será contínuo, segundo uma distribuição de Poisson (a densidade de probabilidade de infecção em cada aresta no tempo t é dada por $\rho_\lambda(t) = \lambda e^{-\lambda t}$).

No instante inicial $t_0 = 0$, consideremos que apenas o vértice central esteja infectado, no passo de tempo $t_1 = 1/\mu$, o centro torna-se suscetível e com certa probabilidade, infecta alguns dos vértices periféricos. No próximo passo de tempo, $t_2 = 2/\mu$, os vértices periféricos tornam-se suscetíveis e possuem a chance de reinfetar o centro, retornando a configuração inicial (ver Figura 3.8). Assim, estamos interessados em calcular a probabilidade de ocorrência desses ciclos e encontrar o tempo de atividade do *hub* a partir dela.

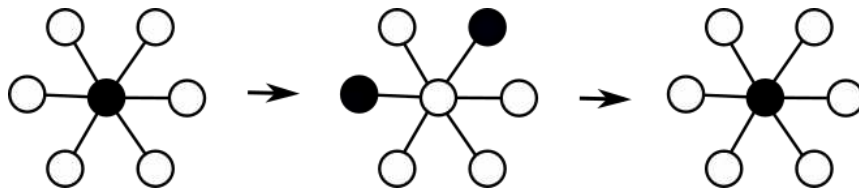


Figura 3.8 – Esquema da dinâmica SIS em uma estrela.

Entre os passos de tempo t_0 e t_1 , a probabilidade do vértice central infectar uma das folhas é

$$q = \int_0^{1/\mu} \rho_\lambda(t) dt = 1 - \exp\left(-\frac{\lambda}{\mu}\right) \approx \frac{\lambda}{\mu}, \quad (3.24)$$

considerando $\lambda/\mu \ll 1$. Considerando que a estrela possui $N = k + 1$ vértices, com k sendo o grau do vértice central, a probabilidade do centro infectar n vértices periféricos é

$$P(n|k) = \binom{k}{n} q^n (1 - q)^{k-n}. \quad (3.25)$$

Já entre os passos t_1 e t_2 , desejamos determinar a probabilidade de que algum dos vértices periféricos infectados consiga reinfectar o centro. A probabilidade de reinfeção do centro por cada folha é a mesma da infecção da folha pelo centro e dada pela Eq. (3.24). Assim, a probabilidade de que uma folha não consiga reinfectar o centro é $(1 - q)$. Logo, a probabilidade de que pelo menos uma das n folhas reinfecte o vértice central é

$$Q(n|k) = [1 - (1 - q)^n]. \quad (3.26)$$

Portanto, a probabilidade de que a estrela se mantenha ativa por um ciclo (centro infecta folhas que reinfectam o centro) é

$$P_c = \sum_{n=1}^k Q(n|k) P(n|k) = \sum_{n=1}^k [1 - (1 - q)^n] \binom{k}{n} q^n (1 - q)^{k-n}. \quad (3.27)$$

Após algum algebrismo, encontramos $P_c \approx 1 - \exp\left(\frac{-k\lambda^2}{\mu^2}\right)$ dado $q \approx \lambda/\mu \ll 1$. A partir de P_c , podemos calcular a probabilidade da estrela sobreviver por S ciclos: $P_s = P_c^{S-1}(1 - P_c)$ e, por fim, calcular o tempo de vida médio:

$$\tau_k^{rec} = \frac{2}{\mu} \langle S \rangle = \frac{2}{\mu} \sum_{S=1}^{\infty} S P_c^{S-1} (1 - P_c) = \frac{2}{\mu} \exp\left(\frac{k\lambda^2}{\mu^2}\right). \quad (3.28)$$

Dessa forma, temos que o tempo de atividade de uma estrela diverge exponencialmente com o seu tamanho. O próximo passo é calcular o tempo que um *hub* de grau k leva para infectar um *hub* de grau k' , separados a uma distância d (ver Figura 3.9). Para isso, vamos considerar que uma estrela com k folhas está sempre ativa (conforme a Eq. (3.28), se $k \gg \frac{\mu^2}{\lambda^2}$, temos que $\tau_k^{rec} \rightarrow \infty$). Além disso, vamos considerar a epidemia onde os vértices da esquerda infectam os vértices da direita antes de se tornarem suscetíveis. Podemos mostrar que a probabilidade de transmissão entre cada aresta é $q = \lambda/(\lambda + \mu)$. Assim, a probabilidade da infecção partir do vértice $i = 1$ e chegar ao centro da estrela inativa é q^{d-1} e a taxa de transmissão é $\lambda_{kk'} = \lambda q^{d-1}$. Para uma rede aleatória não

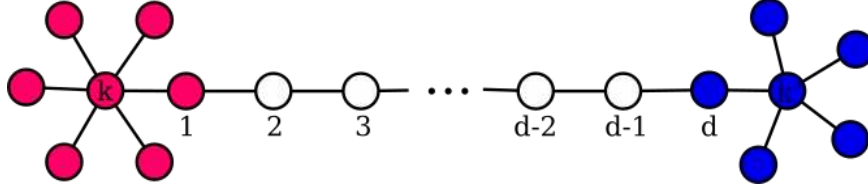


Figura 3.9 – Estrela infectada de grau k (em vermelho) a uma distância d de uma estrela suscetível de grau k' (em azul).

correlacionada com N vértices, a distância média entre vértices de grau k e k' é [39]:

$$d = 1 + \frac{\ln(N\langle k \rangle / kk')}{\ln \kappa}, \quad (3.29)$$

em que $\kappa = \langle k^2 \rangle / \langle k \rangle - 1$. Dessa forma, temos que o tempo de infecção $\tau_{kk'}^{inf}$ é:

$$\tau_{kk'}^{inf} = \frac{1}{\lambda_{kk'}} = \frac{1}{\lambda} \left(\frac{N\langle k \rangle}{kk'} \right)^{b(\lambda)}, \quad (3.30)$$

com $b(\lambda) = \ln(1 + \mu/\lambda) / \kappa$. Para calcular esse tempo de infecção, consideramos apenas o menor caminho entre as estrelas mas existem diversos outros caminhos por onde a infecção pode se propagar. Dessa forma, a estimativa para $\tau_{kk'}^{inf}$ encontrada é uma cota superior para o tempo real de infecção entre os *hubs*.

Comparando os valores encontrados para τ_k^{rec} e para $\tau_{kk'}^{inf}$, podemos mostrar que a medida que N cresce, o crescimento do tempo de atividade de uma estrela é muito maior do que o tempo de infecção entre as estrelas. Dessa forma, em uma rede complexa com muitos *hubs*, essas estrelas são capazes de sustentar a epidemia na rede por tempos muito longos. Dessa forma, efetivamente, o limiar epidêmico vai a zero no limite termodinâmico.

IV Percolação

Em termos simples, percolação comumente refere-se a fenômenos que envolvam análise de fluidos e filtragem em meios porosos. Porém, esses processos podem ser mapeados através de modelos simples para estudar diversos outros sistemas em diversas áreas científicas tais como geologia, física, química, matemática, ciência da computação, ciência dos materiais, dentre outros.

Historicamente, os primeiros estudos sobre fenômenos de percolação datam da Segunda Guerra Mundial, onde Paul Flory e Walter Stockmayer estudaram como a ramificação de pequenas moléculas formavam moléculas cada vez maiores, à medida que mais ligações químicas eram formadas entre estas. Esse processo de polimerização pode levar a formação de uma rede de ligações químicas que coalesce todo o sistema, conhecida como fase Gel [40]. A partir daí, passou-se a estudar diversos outros fenômenos utilizando essa mesma base de pensamento com a ativação de constituintes do sistema levando a formação de agregados que se estendem por toda sua extensão. Podemos citar como exemplos a modelagem de incêndios, propagação de epidemias, análise de rugosidade de superfícies, transporte em meios porosos, dentre outros [41, 42, 43].

Desde sua consolidação como um novo ramo da ciência, o interesse na caracterização e análise estrutural de redes complexas encontrou na teoria de percolação uma base para seus estudos envolvendo robustez e fragilidade das redes [17], considerando a ideia de percolação de forma invertida, ou seja, um agregado se fragmentando devido a remoção de seus constituintes.

4.1 Conceitos de percolação

O ferramental para estudar e analisar problemas envolvendo ataques em redes pode ser importado do estudo de problemas clássicos envolvendo percolação. Entre os diversos modelos estudados, dois deles se destacam pela sua simplicidade e grande utilidade: percolação por sítio e percolação por ligação.

No modelo de percolação por sítio, considere um reticulado (como os da Figura 2.7) com N sítios. Visitamos os sítios do reticulado e, com probabilidade p , ele é ativado. Sítios vizinhos que estejam simultaneamente ativados formam agregados. Nesse caso, o interesse consiste em estudar a estatística desses agregados: distribuição de tamanhos, tamanho do maior agregado, distância média entre dois agregados, entre outros. Além disso, há o interesse em estudar a dependência dessas grandezas com p .

No modelo de percolação por ligação, visitamos cada ligação entre os vértices e as ativamos com probabilidade p . Assim, sítios vizinhos que possuam *links* ativados formam agregados. As quantidades de interesse são as mesmas do caso de percolação de sítio. Na

Figura 4.1, temos uma ilustração dos modelos.

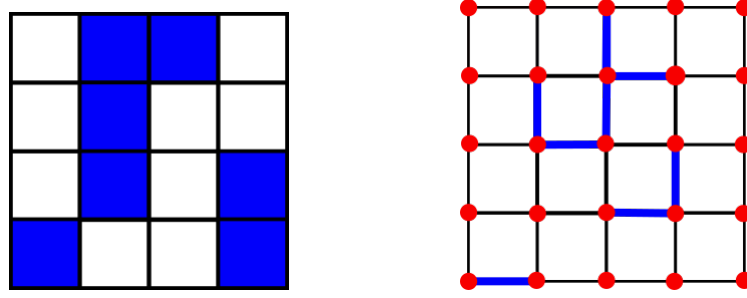


Figura 4.1 – (a) Ilustração do modelo de percolação por sítio em reticulados. Vértices em azul representam sítios ativados. Sítios ativos que sejam vizinhos forma agregados. (b) Ilustração do modelo de percolação por ligação. As arestas em azul representam ligações ativas. Sítios vizinhos que possuam uma ligação ativa formam agregados.

Em ambos os casos, a forma dos agregados depende de p . Para p suficientemente pequeno, há a formação de inúmeros agregados pequenos. À medida que aumentamos p , o tamanho dos agregados vai aumentando até o ponto em que todos coalescem em uma grande estrutura, denominada agregado de percolação. O valor de p no qual ocorre essa transição é chamado limiar de percolação p_c . Resumindo, em p_c , observamos uma transição de fase de apenas agregados pequenos para o aparecimento de um agregado de percolação. Note que esta transição é de origem puramente geométrica, ao contrário dos modelos epidêmicos em que a transição tem origem dinâmica.

Dessa forma, podemos estudar percolação em termos de uma transição de fase, onde p é o parâmetro de controle do sistema e a probabilidade de um sítio pertencer ao maior agregado da rede P_∞ é o parâmetro de ordem. Outras grandezas de interesse são o tamanho médio dos agregados $\langle s \rangle$, excluindo o maior da rede, que possui uma divergência em torno do limiar de percolação e o comprimento de correlação que é a distância média entre sítios em um mesmo agregado e diverge no ponto crítico. Essas grandezas obedecem determinadas leis de escala em torno do ponto crítico resumidas abaixo:

- Tamanho médio dos agregados: $\langle s \rangle \sim (p - p_c)^{-\gamma}$;
- Parâmetro de ordem: $P_\infty \sim (p - p_c)^\beta$;
- Comprimento de correlação: $\xi \sim (p - p_c)^{-\nu}$,

em que γ , β e ν são expoentes críticos que caracterizam o comportamento do sistema próximo do ponto crítico. Na Figura 4.2, temos representações das dependências de P_∞ e $\langle s \rangle$ com p .

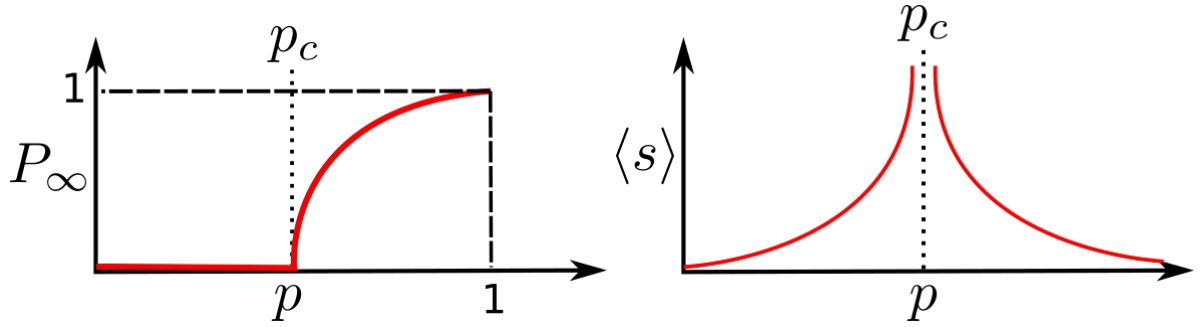


Figura 4.2 – Comportamentos de (a) P_∞ e (b) $\langle s \rangle$ ao variarmos o parâmetro de controle p .

4.2 Ataques em redes

Todo sistema é suscetível a falhas ou ataques em maior ou menor grau. Por exemplo, um pequeno componente elétrico danificado em um computador pode comprometer a integridade total da máquina, o rompimento de cabos de fibra ótica em uma cidade pode desligar completamente a internet em sua vizinhança, a atuação de um fármaco em determinada fase do ciclo biológico de uma bactéria pode destruí-la completamente, a imunização de determinados indivíduos pode impedir uma doença local de se espalhar globalmente, dentre outros.

Acontecimentos como o *blackout* da Carolina do Norte em Julho de 2017, onde somente um cabo danificado deixou a área de *Outer Banks* sem energia por 8 dias [44] ou o *ransomware* (um tipo de software nocivo que restringe o acesso ao sistema infectado e cobra um resgate para que o acesso possa ser restabelecido), conhecido como *Wanna Decryptor* ou *WannaCry* [45, 46], um *worm* de encriptação de dados que afetou cerca de 300 mil aparelhos em diversos países, inclusive invadindo diversos hospitais na Inglaterra e na Escócia levando a desativação de diversos serviços e causando um prejuízo de milhões de dólares, ilustram a importância de se estudar e analisar a estrutura de redes complexas na tentativa de entender suas fragilidades de forma a minimizar ataques ou falhas e construir sistemas mais robustos.

No contexto de redes complexas, os ataques representam a desativação de vértices ou ligações nas redes, de forma que eles não interagem com o resto da rede. Essa desativação pode depender do processo dinâmico a ser considerado (sítios imunizados em uma epidemia, rompimentos de cabos em redes de transmissão, o fechamento de aeroportos em uma malha aérea) e podem representar a remoção física desses vértices ou ligações (exclusão de um perfil em uma rede social, por exemplo). Nessa dissertação, estudaremos apenas ataques a vértices da rede e todas as suas arestas.

Espera-se que ao removermos poucos vértices, o dano causado não seja suficiente para afetar a integridade da rede. Porém, a medida que o número de vértices removidos aumenta, a rede pode fragmentar-se em diversas componentes pequenas e perder sua

funcionalidade [17]. A Figura 4.3 ilustra um processo de ataque em uma rede até a sua fragmentação. Para estudar esses fenômenos, podemos utilizar o ferramental de teoria de percolação. De fato, esses ataques em rede podem ser vistos como um mecanismo de percolação inversa [17], onde inicialmente temos uma componente conectada que é fragmentada em agregados pequenos.

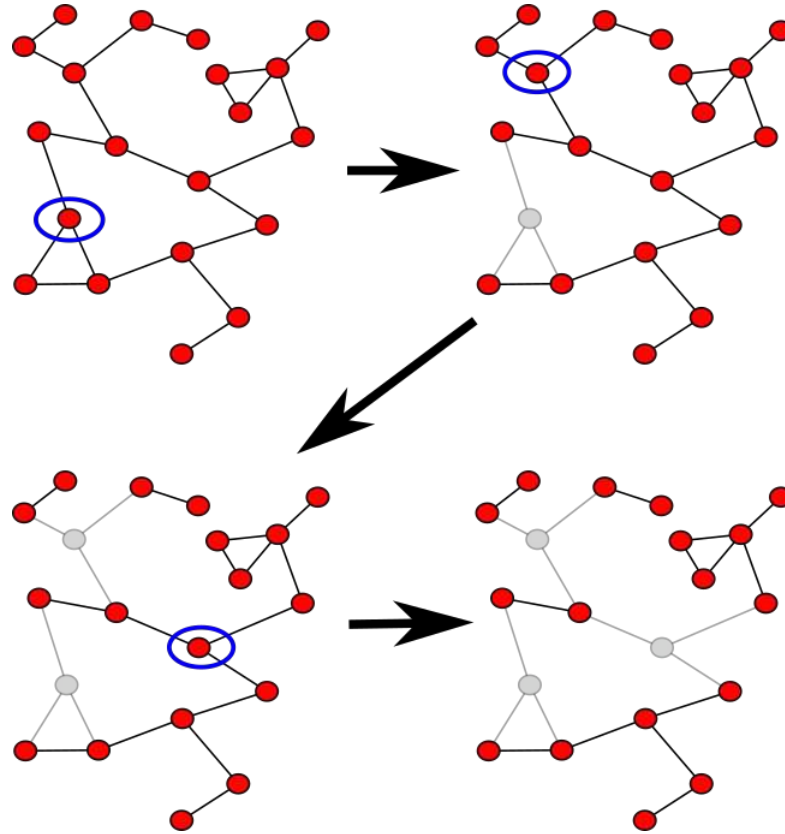


Figura 4.3 – Ilustração mostrando a fragmentação de uma rede complexa a medida que removemos vértices (circulados em azul). Com a remoção do primeiro vértice a rede continua conectada, removendo o segundo, cria-se duas pequenas componentes e ao remover o terceiro a rede se fragmenta em cinco componentes. Figura baseada em [17].

Da mesma forma que no caso de percolação usual, o fenômeno de fragmentação de uma rede é caracterizado por uma fração de vértices removidos (f) crítica, que separa as regiões onde existe ou não a componente gigante, como ilustra a Figura 4.4.

Assim, para analisar a robustez de redes complexas, grande interesse reside na determinação de f_c para um determinado tipo de ataque (seção IV.3). Em 1995, Molloy e Reed desenvolveram um critério simples para determinar se uma rede aleatória possui ou não uma componente gigante [47] segundo o qual uma rede possuirá componente gigante se

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} > 2. \quad (4.1)$$

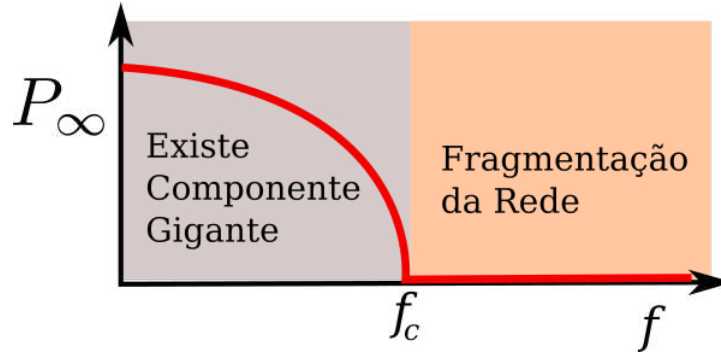


Figura 4.4 – Probabilidade de um vértice pertencer a componente gigante de uma rede (P_∞) em função da fração de vértices removidos. As duas regiões são separadas pelo limiar de percolação f_c .

Uma visualização desse critério pode ser observado para redes complexas da seguinte forma: Para a rede ser conectada, é necessário que o grau médio da rede seja pelo menos 2. Dessa forma, cada vértice possuirá pelo menos uma aresta entrando e uma aresta saindo deste. Assim, caminhos são criados entre os vértices e na prática ocorre a emergência da componente gigante a partir dessa propriedade. Esse critério conecta a integridade da rede (representada pela existência ou não da componente gigante) à distribuição de graus desta (representada pelos momentos $\langle k \rangle$ e $\langle k^2 \rangle$) e é válido para qualquer $P(k)$ dado que a rede é aleatória [47].

4.3 Estratégias de ataque

Diferentes estratégias de ataque podem ser utilizadas de acordo com o objetivo que se deseja alcançar. Se desejamos inutilizar a rede por inteiro, atacar pontos chaves, transformar a rede de uma forma específica, dentre outros. No geral, essas estratégias podem ser organizadas em duas categorias:

- **Ataques localizados:** Requerem o conhecimento integral da rede, a importância dos vértices é quantificada segundo alguma propriedade (grau, K_{nn} , modularidade, etc) e eles são removidos ordenadamente.
- **Ataques não localizados:** Não requerem o conhecimento total da estrutura da rede, apenas de propriedades locais. Os vértices são escolhidos ao acaso e a partir desse sorteio inicial as estratégias são implementadas. Por exemplo, pode-se remover o próprio vértice sorteado ou um de seus vizinhos.

Considerando uma estratégia de ataque localizada, a remoção dos vértices pode alterar a propriedade utilizada para quantificar a importância destes. Assim, podemos fazer esse ordenamento antes de iniciar o ataque e segui-lo independente das mudanças na rede. Nesse caso, denominamos o ataque de **não adaptativo**. Se a importância

dos vértices é recalculada a cada vértice removido dizemos que o ataque é **adaptativo**. Por exemplo, vamos supor que uma determinada estratégia de ataque remova os sítios em ordem decrescente de grau. Assim, a remoção dos vértices altera o grau dos sítios remanescentes na rede. Se o grau dos vértices é recalculado a cada ataque e a partir daí os próximos alvos são determinados, a estratégia é adaptativa. Do contrário, é não adaptativa.

Estratégias localizadas e adaptativas usualmente são muito efetivas para fragmentar redes porém, adquirir o conhecimento da estrutura de uma rede por completo é muito difícil e muitas vezes, inviável. O mais comum é possuir conhecimentos de partes da rede. Além disso, os constantes cálculos necessários para adaptar o ataque a cada remoção de vértice podem fazer essas estratégias custosas computacionalmente. Dessa forma, buscamos estratégias de ataque que sejam não localizadas e ainda assim capazes de danificar efetivamente as redes.

Visto que muitas redes reais são sem escala e apresentam propriedades de mundo pequeno, inicialmente estudamos modelos de ataque que tentem preservar essas propriedades (pelo menos assintoticamente):

- Ataque aleatório;
- Ataque de contato, denominado NVA do inglês *neighbor vertex attack*;
- Ataque de contato com proteção de *hub*, denominado NVA-HP do inglês *neighbor vertex attack with hub protection*.

4.3.1 Ataque aleatório

O ataque aleatório é geralmente o ponto de partida para elaboração e o estudo de ataques mais complexos. Geralmente é utilizado para modelar falhas não intencionais em redes complexas. Nesse modelo, escolhemos um vértice da rede ao acaso e o removemos juntamente com todas suas arestas. A Figura 4.5 ilustra esse processo.

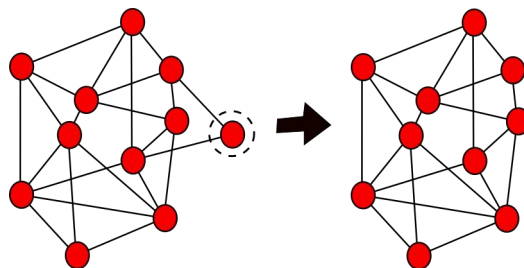


Figura 4.5 – Ilustração esquemática de um ataque aleatório.

Considerando redes sem escala, é esperado que essa estratégia não seja muito eficaz, já que a probabilidade de remover vértices de grau pequeno é muito grande e a remoção desses sítios não comprometem a integridade da rede.

Partindo da definição do critério de Molloy-Reed (Eq. (4.1)), podemos encontrar uma expressão para o limiar de percolação dos ataques aleatórios [17]

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}. \quad (4.2)$$

A partir desta expressão, encontramos um resultado interessante. Lembrando que para redes sem escala com $2 < \gamma < 3$ há divergência de $\langle k^2 \rangle$ no limite termodinâmico. Assim, temos que $f_c \rightarrow 1$ sendo necessário remover todos os vértices da rede para que esta se fragmente. Dessa forma, dizemos que redes sem escala são robustas a ataques aleatórios. [17].

4.3.2 Ataque de contatos

A estratégia NVA é utilizada na tentativa de determinar os vértices mais conectados da rede. Um sítio e um de seus vizinhos são sorteados aleatoriamente e o vizinho e todas suas conexões são removidas. Uma vez que o grau médio dos vizinhos (K_{nn}) é maior do que k , o NVA danifica a rede de forma mais eficaz que o ataque aleatório ao remover vértices com grau maior. A Figura 4.6 ilustra essa estratégia.

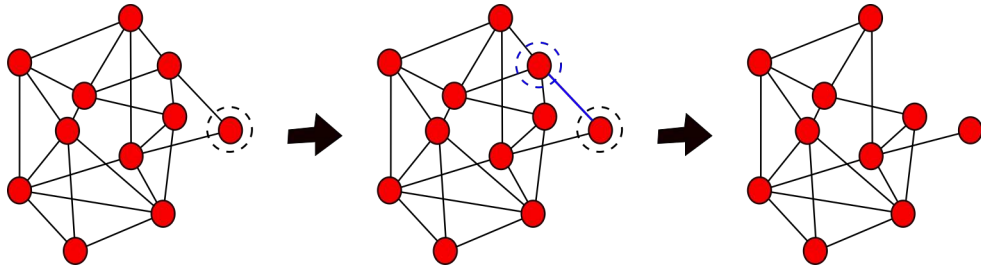


Figura 4.6 – Ilustração esquemática de um ataque de vizinho

Estudando uma estratégia de imunização denominada de imunização por contato (*acquaintance immunization*) na qual um vértice é escolhido ao acaso e um de seus vizinhos é imunizado, Cohen *et al* [48] conseguiram estimativas para o limiar de percolação desse modelo e concluíram que para redes sem escala f_c é finito. Dessa forma essa estratégia de imunização foi capaz de fragmentar a rede após a remoção de uma fração finita de vértices. Com exceção de um detalhe em que os vértices imunizados continuam fazendo parte da rede e podem futuramente serem sorteados novamente, ou seja, é não adaptativa, essa estratégia é similar ao NVA, que também deve possuir f_c finito.

4.3.3 Ataque de contatos com proteção de hub

Em um contexto de redes tecnológicas, quanto maior a importância de um vértice, mais difícil é removê-lo da rede. Por exemplo, servidores dos grandes provedores de acesso a internet devem possuir uma segurança maior que os distribuidores locais e dos

computadores domésticos. Para levar isso em conta, vamos estudar o método NVA-HP, que dá aos vértices uma probabilidade de não serem removidos, representando a capacidade do sítio de se defender. Esse modelo considera que sítios de grau maior são mais difíceis de serem atacados. Utilizamos a seguinte função de ataque:

$$\Phi(k) = \frac{\langle k \rangle^s}{\langle k \rangle^s + k^s} \quad (4.3)$$

em que s é um parâmetro do modelo. Dessa forma, escolhemos ao acaso um vértice e um vizinho deste e removemos o vizinho e todas suas conexões com probabilidade $\Phi(k)$. Na figura 4.7 temos a função $\Phi(k)$ plotada para alguns valores de s . Note que o caso particular $s \rightarrow 0$ corresponde à nenhuma proteção (todos os sítios com igual chance de serem removidos) enquanto que o caso $s \rightarrow \infty$ equivale a utilizar uma função degrau em que apenas sítios com grau menor que $\langle k \rangle$ poderão ser removidos. Nesse trabalho, utilizamos $s = 0.5$.

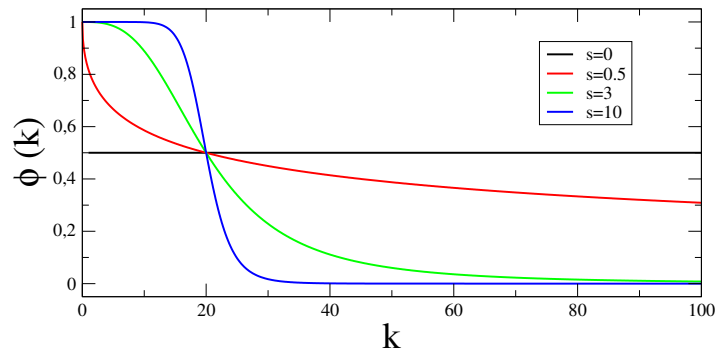


Figura 4.7 – Probabilidade de remoção de um vértice ($\Phi(k)$) em função de k para alguns valores de s e considerando $\langle k \rangle = 20$.

Os impactos dessas diferentes estratégias de ataque em redes complexas serão analisados na seção VI.1.

V Metodologia computacional

5.1 Modelo UCM

Apesar de redes reais possuírem correlações entre seus vértices, redes não correlacionadas possuem um papel teórico fundamental, visto que muitos cálculos analíticos são válidos somente na ausência de correlação entre os vértices. Dessa forma, geramos redes aleatórias não correlacionadas utilizando o modelo UCM, do inglês *Uncorrelated Configuration Model*[49] que deriva de um modelo utilizado para gerar redes com qualquer distribuição de graus $P(k)$ denominado somente modelo de configurações [47].

No modelo de configurações, são atribuídas k_i pontas (*stubs*) seguindo $P(k)$ previamente estabelecido a cada um dos N vértices da rede, com a restrição $k_{\min} \leq k_i < N$ em que $k_{\min}$ será fixado nesse trabalho como igual a 3. Além disso, $\sum_i k_i$ deve ser par, garantindo assim que não sobre nenhuma conexão vazia ao final da montagem da rede. A partir daí, duas pontas são sorteadas ao acaso e os vértices são conectados formando uma aresta, excluindo-se auto-conexões (um vértice conectado a si mesmo) e múltiplas conexões (um vértice ligado mais de uma vez a outro vértice). Na Figura 5.1, temos uma ilustração esquemática da conexão entre dois vértices segundo esse algoritmo, além da ilustração de auto e múltiplas conexões. Essa abordagem é amplamente utilizada para gerar distribuições nas quais $\langle k^2 \rangle$ seja finito. Nesse caso, a finitude de $\langle k^2 \rangle$ naturalmente introduz um corte na rede ($k_{\max} \ll N$). Dessa forma, as redes não possuem *hubs* e a chance de ocorrência de auto e múltiplas conexões é pequena.

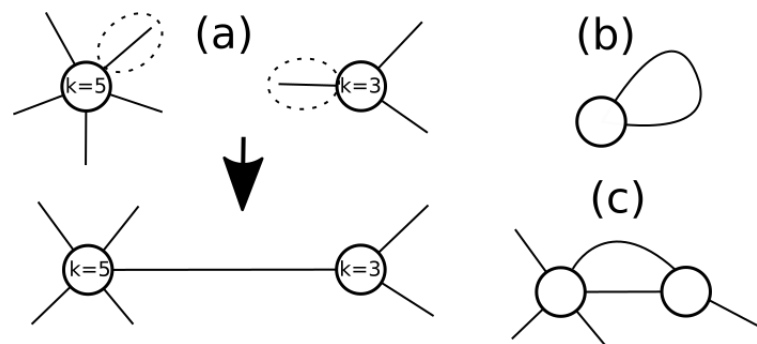


Figura 5.1 – (a) Esquema ilustrativo da conexão entre dois vértices pelo algoritmo CM. Exemplos de (b) auto-conexão (c) múltipla conexão.

No caso de redes sem escala com $2 < \gamma < 3$, em que $\langle k^2 \rangle$ diverge, esse algoritmo apresenta problemas. Já que em redes livres de escala existem muitos *hubs* com um número muito grande de pontas para conectar. Dessa forma, a tentativa de auto-conexões e múltiplas conexões não são desprezíveis e geram correlações entre os graus dos vizinhos. Ao rejeitar conexões entre *hubs*, eles tendem a conectar com vértices de graus mais baixo,

introduzindo correlações dissociativas na rede. Além disso, é muito difícil gerar uma rede grande nessa situação, pois sobram pontas desconectadas e o restante da rede se conecta de tal modo a esgotar a possibilidade de não haver múltiplas ou auto conexões.

Boguñá *et al* [50] mostraram que uma condição suficiente para destruir a correlação entre os vértices no modelo de configurações para redes sem escala era modificar o corte do modelo. Em [50], é mostrado que se a nova restrição para k_i for $k_{\min} < k_i < N^{1/2}$, excluindo auto e múltiplas conexões, então a rede gerada não é correlacionada. Esse novo modelo ficou conhecido como UCM [49] e pode ser implementado utilizando os seguintes passos:

- i) Sortear os graus dos sítios seguindo a distribuição $P(k)$ desejada impondo as condições: $k_{\min} < k < N^{1/2}$ e $\sum_i k_i$ par.
- ii) Atribuir a cada sítio um número de pontas referentes ao grau deste.
- iii) Sortear aleatoriamente duas pontas evitando auto e múltiplas conexões.
- iv) Conectar os vértices sorteados.
- v) Repetir os passos iii e iv até que a rede inteira esteja conectada. Caso sobre pontas a rede é descartada e o processo reiniciado.

Na Figura 5.2(a), temos distribuições $P(k)$ geradas pelo algoritmo UCM, sendo todos leis de potência com expoentes diferentes. Conforme comentado na seção II.1, para redes não correlacionadas temos que $K_{nn}(k)$ é aproximadamente constante, resultado mostrado na Figura 5.2(b).

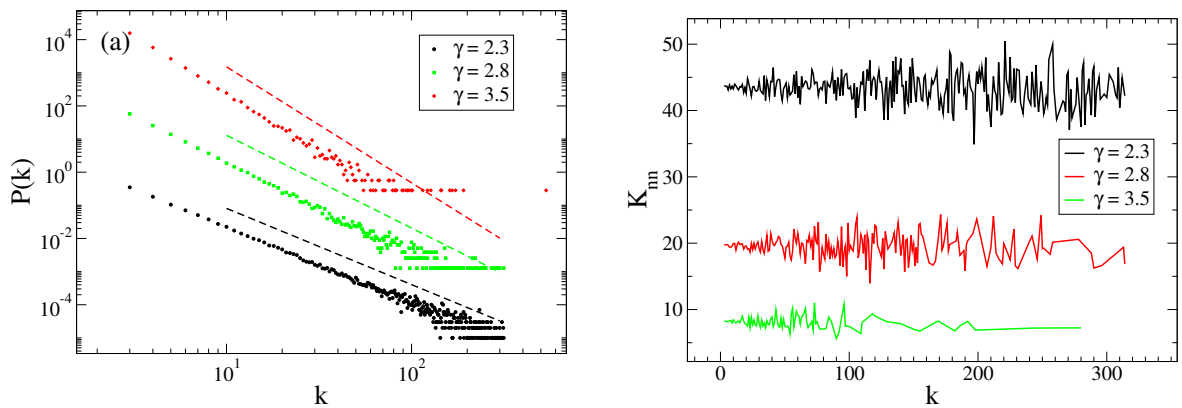


Figura 5.2 – (a) Distribuições de grau para redes complexas com $N = 10^5$ vértices geradas pelo algoritmo UCM com diferentes expoentes (os gráficos foram reescalados verticalmente para evitar sobreposição). Linhas tracejadas representam as leis de potência assintoticamente. (b) K_{nn} em função de k para as redes apresentadas em (a). Note que a função é aproximadamente constante, mostrando que as redes não possuem correlação entre os graus dos vizinhos.

No caso de redes com $\gamma > 3$, o corte natural da distribuição de graus é menor do que o corte estrutural imposto no modelo UCM. Assim, na prática, os cortes são equivalentes e isso gera *outliers*, que possuem papel importante na dinâmica do modelo SIS.

5.2 Algoritmos de simulação de modelos epidêmicos

Um algoritmo clássico utilizado para simulação de processos markovianos de tempo contínuo é o algoritmo de Gillespie (GA) [51]. Nesse algoritmo, decompomos a dinâmica do sistema em processos espontâneos independentes (infecção, cura e perda de imunidade, no caso de epidemias) e executamos uma mudança de estado por passo de tempo, que não é fixo. No entanto, a cada mudança de estado, uma lista com todos os processos espontâneos possíveis deve ser atualizada. Para redes com N muito grande, essa tarefa é computacionalmente inviável. Dessa forma, otimizações desses algoritmos foram propostas e em [52] os métodos são discutidos e as eficiências comparadas. O algoritmo utilizado para simulação dos modelos epidêmicos é chamado de OGA, do inglês *Optimized Gillespie Algorithm* e apresenta desempenho computacional muito maior do que o algoritmo padrão de Gillespie ao passo que sua complexidade não é elevada.

5.2.1 Modelo SIS

A implementação do algoritmo OGA para o modelo SIS se dá com os seguintes passos:

- i) Partindo de uma configuração inicial, calculamos o número de vértices infectados N_i e a soma dos graus desses vértices N_n . Além disso, criamos uma lista com todos os vértices infectados.
- ii) Com probabilidade $\frac{N_i}{N_i + \lambda N_n}$, um sítio Infectado, escolhido ao acaso, torna-se Suscetível. Com probabilidade complementar $\frac{\lambda N_n}{N_i + \lambda N_n}$, um sítio Infectado é escolhido proporcionalmente ao seu grau e uma aresta desse sítio é escolhida aleatoriamente. Caso a aresta aponte para um sítio suscetível, este se torna infectado, do contrário, a simulação continua.
- iii) N_i , N_n e a lista de vértices são atualizados.
- iv) O tempo t é incrementado: $t \rightarrow t + \Delta t$, com $\Delta t = \frac{1}{N_i + \lambda N_n}$
- v) O processo é iterado por um determinado tempo $t_{\max}$.
- vi) Caso o sistema visite o estado absorvente, regras envolvendo o método quase estacionário são aplicadas e o sistema volta a fase ativa (ver seção V.3).

5.2.2 Modelo SIR

A implementação do modelo SIR segue a maioria dos passos do modelo SIS. A diferença é que quando um sítio infectado é curado, ele não retorna ao estado suscetível, mas a um estado recuperado. Uma lista com as posições dos sítios recuperados também é mantida. Ela será usada para reiniciar as condições iniciais que, em geral, correspondem a um único sítio infectado e o restante suscetível. Isso era um grande ganho computacional pois muitos surtos infectam somente uma pequena fração da rede. Assim, ao reiniciar os estados de todos os vértices como suscetível, precisa-se apenas visitar aqueles que estavam recuperados.

5.3 Método quase estacionário

Em simulações de sistemas de tamanho finito que possuam estados absorventes, é possível mostrar que o sistema sempre visitará esse estado [53, 54] e, por conseguinte, ficará estagnado nele. Dessa forma, se queremos estudar efeitos que ocorram em intervalos de tempo longos, temos que contornar esse problema de alguma forma. Uma abordagem utilizada são as chamadas simulações quase-estacionárias. O conceito de estados quase estacionários refere-se à perturbações na dinâmica do sistema de forma a evitar que este visite o estado absorvente, porém sem afetar as grandezas estacionárias no limite termodinâmico [55]. Essa seção será baseada em [56], trabalho onde discutimos diversas formas diferentes de amostragem utilizando estados quase estacionários.

Para compreender alguns conceitos necessários para desenvolver os métodos quase estacionários, precisamos definir o que sejam processos de um passo [57]. Em termos gerais, processos de um passo são processos markovianos em tempo contínuo em que os estados são representados por uma variável inteira e as transições ocorrem somente entre estados com $|n - n'| = 1$, n e n' denotando o número de vértices infectados no caso de um processo de propagação de epidemia. Definimos X_t como um processo estocástico de um passo em que o número de sítios infectados determina o estado do sistema, incluindo o estado absorvente com $n = 0$ indivíduos infectados. Partindo disso, considere X_t^* um processo similar a X_t , porém sem os estados absorventes. A dinâmica de X_t^* é idêntica a X_t , salvo perturbações durante a fase ativa, desprezíveis no limite termodinâmico. Assim, o processo X_t^* retorna para um estado ativo após um tempo finito visitando o estado absorvente.

Denotando P_n e P_n^* a probabilidade de que existem n indivíduos infectados na dinâmica original e na modificada, respectivamente, a equação mestra para a dinâmica original é:

$$\frac{dP_n}{dt} = \sum_m w_{n,m} P_m - \sum_n w_{m,n} P_n \quad (5.1)$$

em que $w_{n,m}$ são transições de estados m para n . Para o processo X_t^* , a equação dinâmica

possui um termo perturbativo, relacionado a remoção do estado absorvente dado por

$$\frac{dP_n^*}{dt} = \sum_m w_{n,m} P_m^* - \sum_n w_{m,n} P_n^* + F(P_0^* \dots P_n^* \dots), \quad (5.2)$$

em que F depende do método quase estacionário utilizado.

Neste trabalho, utilizaremos o método quase estacionário padrão [55, 56], que consiste em fazer médias somente sobre amostras que sobreviveram (não visitaram o estado absorvente) até um tempo t . Dessa forma, temos que a distribuição quase-estacionária da dinâmica é

$$\bar{P}_n = \lim_{t \rightarrow \infty} \frac{P_n(t)}{P_s(t)}, \quad (5.3)$$

em que $P_s = \sum_{i \geq 1} P_i(t) = 1 - P_0(t)$ é a probabilidade de sobrevivência até um determinado tempo t . Essa estratégia apresenta um problema para simulações nas porções crítica e subcrítica: como o sistema visita o estado absorvente constantemente nesses regimes, isso proporciona intervalos de dados curtos e ruidosos para análises estatísticas. Para contornar esse empecilho, em [55] foi proposta uma nova estratégia. Cada vez que o sistema visita o estado absorvente, ele é redirecionado para um estado ativo selecionado segundo a probabilidade quase estacionária. Dessa forma, o termo perturbativo na Eq. (5.2) torna-se $F = w_0 P_n^*$ onde w_0 é a taxa total de transições para o estado absorvente. A Figura 5.3 ilustra esse processo.

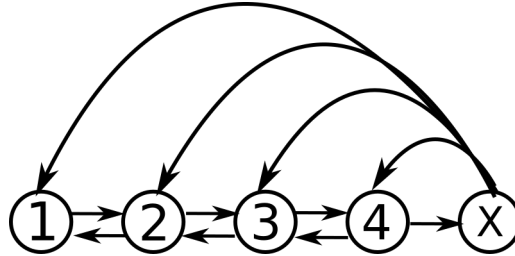


Figura 5.3 – Esquema representando o método de simulação proposto em [55]. Vértices numerados representam estados ativos do sistema enquanto o vértice com um X representa o estado absorvente.

Computacionalmente, a implementação é feita seguindo os seguintes passos:

- i) Durante a simulação, enquanto o sistema ainda não visitou o estado absorvente, com probabilidade p_{rep} a configuração do sistema é salva em uma lista.
- ii) A lista de configurações ($\simeq 100$) é constantemente atualizada substituindo um elemento da lista pela configuração atual com probabilidade p_{rep} durante a simulação. Dessa forma, configurações mais prováveis aparecerão mais frequentemente na lista e eliminará a dependência da condição inicial.

- iii) Caso o sistema visite o estado absorvente, uma das configurações da lista é escolhida com igual chance e é utilizada como condição para reativar o sistema. A partir daí, a simulação prossegue até que novas visitas ao estado absorvente ocorram.

5.4 Outros detalhes

Durante as simulações, geramos as redes complexas utilizando o UCM com diferentes valores de γ . De posse da rede, removemos uma fração f dos vértices utilizando alguma estratégia de ataque (Aleatório, NVA, NVA-HP). As componentes da rede atacada são categorizadas e a componente gigante identificada. Utilizamos somente esta componente para simular os processos epidêmicos.

Para comparações dos limiares epidêmicos entre as redes danificadas e as redes intactas, estas devem possuir tamanho relativamente próximo. Como a remoção dos vértices altera o tamanho das redes, foi necessário gerar redes maiores de forma que, após o ataque, a componente gigante possuía tamanho entre $0,95N$ e $1,05N$ (N sendo o tamanho da rede intacta para comparação). Para $f = 0.1$, essa condição é desnecessária, já que a componente gigante da rede danificada é cerca de 99% do restante. Assim, apenas geramos redes com tamanho $N' = \frac{N}{1-f}$. Porém, para $f = 0.4$, a componente gigante das redes varia muito de tamanho e representam de 30 – 50% da rede resultante e, dessa forma, a estratégia foi necessária.

Durante as simulações do modelo SIS, uma fração dos sítios pertencentes a componente gigante é infectado e o sistema evolui a partir daí. As grandezas quase estacionárias relevantes são calculadas durante um tempo de média t_{av} da ordem de 10^7 , após um tempo de relaxação $t_{rlx} \simeq 10^6$. A probabilidade $\bar{P}_n$ é calculada durante t_{av} de forma que cada configuração com n sítios ativos contribui para $\bar{P}_n$ proporcionalmente ao tempo de vida desta configuração. A partir de $\bar{P}_n$, podemos extrair os momentos de densidade de sítios infectados e calcular a suscetibilidade dinâmica χ via Eq.(3.7). Repetimos esse processo para diferentes valores de λ e extraímos o limiar epidêmico através do pico de χ .

Para o modelo SIR, infectamos um único vértice da componente gigante e analisamos como a epidemia se propaga a partir daí. Quando o sistema cai no estado absorvente, calculamos ρ_r . Médias foram feitas sobre diversas realizações (a infecção começa aleatoriamente em um vértice da componente gigante). Partindo dessa amostragem de ρ_r , calculamos a variabilidade epidêmica Δ em função do parâmetro de controle e a partir do pico de Δ , encontramos o limiar epidêmico para este modelo.

VI Resultados e discussões

6.1 Propriedades das redes atacadas

Inicialmente, estudamos os efeitos das diferentes estratégias de ataque nas propriedades estruturais das redes complexas. Vamos estudar redes geradas pelo algoritmo UCM, que permite variar a quantidade de *hubs* por meio do expoente de grau γ . Dessa forma, escolhemos dois valores: $\gamma = 2.3$ (para um número maior de *hubs*) e 2.8 (para um número menor de *hubs*). A Figura 6.1 mostra o parâmetro de ordem do processo de ataque em função da fração de vértices removidos para três estratégias considerando redes de vários tamanhos e os dois valores de γ . Podemos notar que o limiar de percolação é bem próximo de 1 e não se altera muito com o aumento da rede no caso do ataque aleatório, confirmando que redes sem escala são robustas a esse tipo de ataque. Considerando os ataques NVA-HP e NVA, pode-se notar que o limiar de percolação é bem menor que 1, indicando que essas redes não são tão robustas a estas estratégias. Além disso, apesar de haver um pequeno deslocamento do limiar para as redes referentes a $\gamma = 2.30$, pode-se notar que o limiar é próximo de 0.5 para os tamanhos analisados. Dessa forma, podemos identificar valores de f que sejam subcríticos a todos os casos.

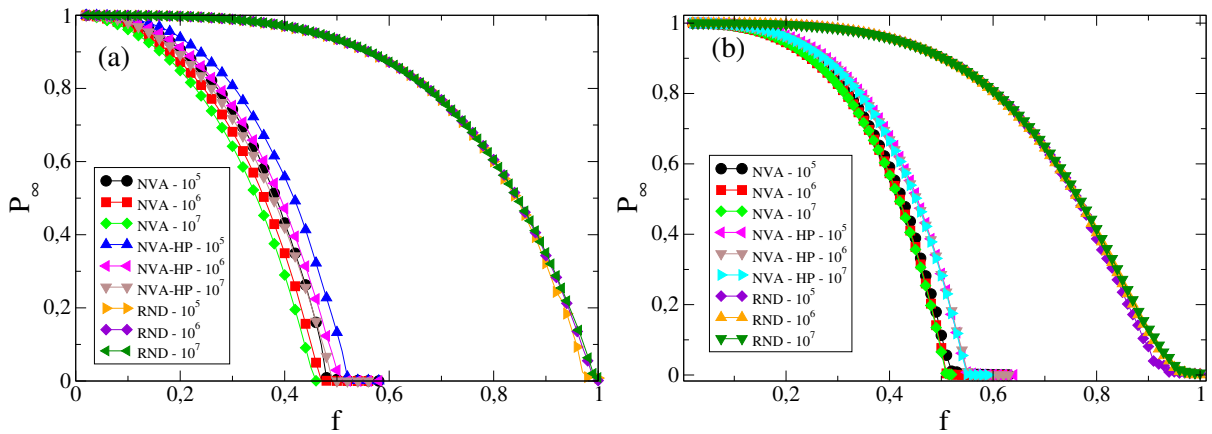


Figura 6.1 – P_∞ em função de f para redes UCM de diferentes tamanhos indicados nas legendas (a) $\gamma = 2.30$ (b) $\gamma = 2.80$. RND é um acrônimo para aleatório

Para aplicações em epidemias, é importante analisar como as estratégias de ataque afetam a distribuição de graus das redes. Esses resultados podem ser encontrados na Figura 6.2. Pode-se notar que ataques aleatórios (Figura 6.2(a)) realmente não alteram $P(k)$ significativamente e a lei de potência é preservada, conforme já conhecido [17]. Já a Figura 6.2(b), referente ao NVA, temos que a distribuição de graus passa a ter um corte em $P(k)$ a medida que removemos mais vértices. Por fim, a estratégia NVA-HP (Figura 6.2(c)) altera principalmente a cauda da distribuição, aumentando a inclinação desta.

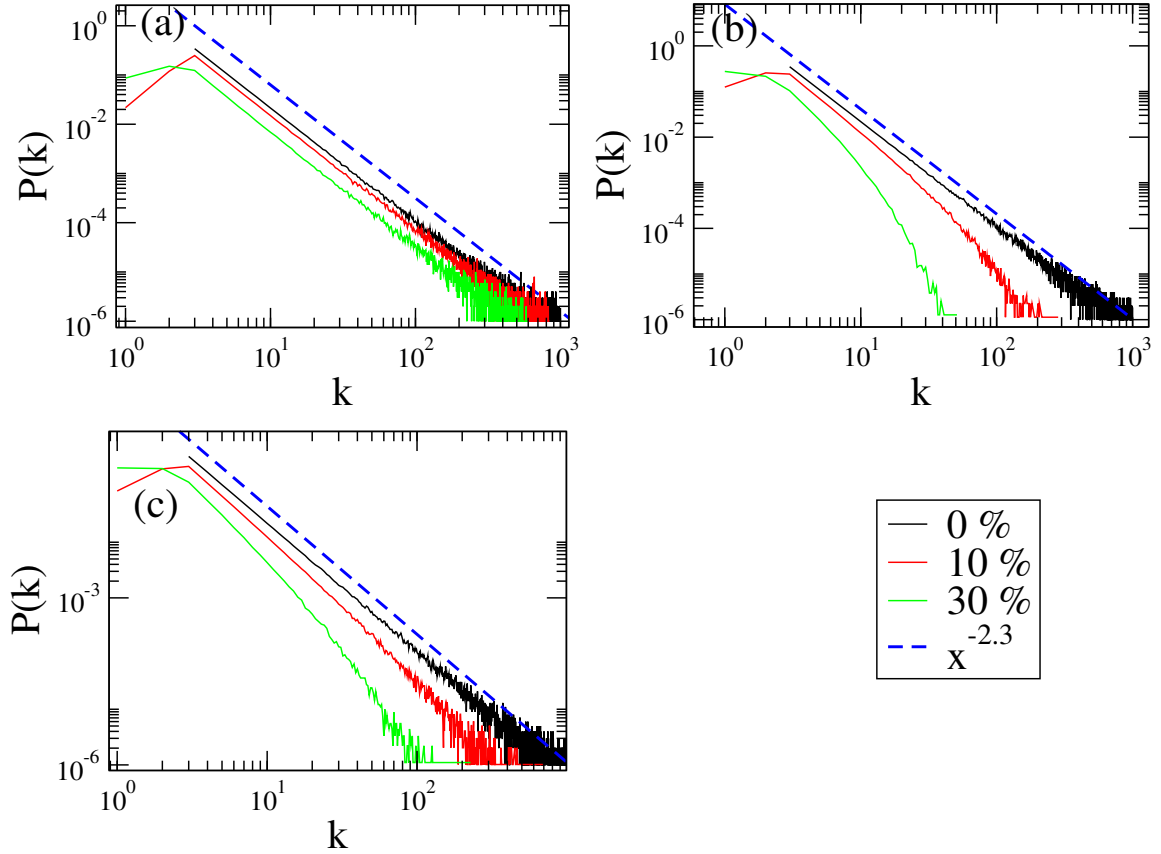


Figura 6.2 – Distribuição de probabilidades de grau para as redes atacadas ($N = 10^6$, $\gamma = 2.80$ e média sobre 10 redes) considerando as três estratégias (a) Aleatório (b) NVA (c) NVA-HP.

A partir desses resultados, escolhemos os valores de f que utilizaremos para o restante da análise estrutural e para as simulações dos processos epidêmicos. Escolhemos $f = 0.1$ por estar bem abaixo do limiar de percolação em todos os modelos e pela componente gigante ser praticamente toda a rede restante.

Como a teoria BCPS [37](III.4) compara o tempo de infecção entre *hubs* e o tempo de vida de uma estrela, duas grandezas são importantes para se analisar os possíveis efeitos das estratégias de ataque no limiar epidêmico dos modelos: o menor caminho médio $\langle \ell \rangle$ e o grau máximo da rede ($k_{\max}$). A partir da análise conjunta dessas duas grandezas podemos tentar entender como os ataques afetarão o limiar epidêmico.

Nas Figuras 6.3(a) e 6.3(b), temos $\langle \ell \rangle$ em função de $\ln N$ para redes UCM com $\gamma = 2.3$ e $\gamma = 2.8$ respectivamente. Fizemos um ajuste do tipo

$$\langle \ell \rangle = C_1 + C_2(\ln N)^b \quad (6.1)$$

para caracterizar e verificar o comportamento de mundo pequeno. Conforme esperado, a dependência entre o menor caminho médio e N é sublogarítmico no caso das redes intactas (resultado descrito na seção II.2), dessa forma, temos que o expoente do ajuste não linear é menor do que 1. Para as redes atacadas, nota-se inicialmente que o ataque aleatório

não altera essa propriedade da rede, como esperado. Porém, para as outras estratégias há um aumento em $\langle \ell \rangle$ e uma mudança na dependência com N . As redes atacadas pela estratégia NVA-HP possuem um expoente um pouco maior do que 1 enquanto as redes afetadas pelo NVA possuem um expoente bem maior que 1 (os valores de b encontrados estão nas legendas das figuras). Contudo, um crescimento $\sim (\ln N)^b$ será sempre menor do que um crescimento algébrico, de forma que ainda podemos considerar as redes como redes de mundo pequeno dentro do intervalo de tamanhos investigados.

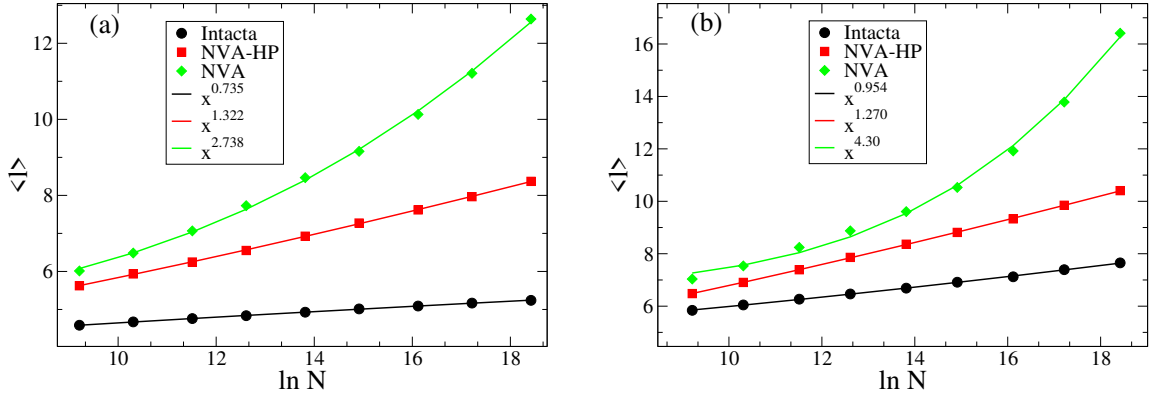


Figura 6.3 – Menor caminho médio em função do tamanho para redes danificadas com $f = 0.1$ (Média sobre 100 amostras). Símbolos representam as simulações enquanto que as linhas cheias representam ajustes não lineares do tipo $\langle \ell \rangle = C_1 + C_2(\ln N)^b$. (a) $\gamma = 2.30$ (b) $\gamma = 2.80$.

Além disso, calculamos também o maior grau de um vértice da rede. Nas Figuras 6.4(a) e 6.4(b), temos o valor médio de $\langle k_{\max} \rangle$ para redes UCM atacadas pelas três estratégias de ataque, com $\gamma = 2.3$ e $\gamma = 2.8$ respectivamente. Primeiramente, podemos notar que novamente os ataques aleatórios não alteram $\langle k_{\max} \rangle$ tal que esta estratégia é similar ao corte imposto pelo modelo UCM. Considerando as outras duas estratégias, é introduzido um corte no valor de $\langle k_{\max} \rangle$, sendo este corte bem evidente no caso NVA mesmo para tamanhos pequenos enquanto que para a NVA-HP temos que o corte, se realmente existir, ocorre para tamanhos muito maiores. Dessa forma, esperamos que a convergência de $\langle k_{\max} \rangle$ nesse caso possa ocorrer para tamanhos muito maiores que os estudados.

Analisando conjuntamente em termos da teoria BCPS, temos que $\langle \ell \rangle$ aumenta consideravelmente para a estratégia NVA, assim a comunicação entre os *hubs* é dificultada e conseqüentemente, o tempo de infecção entre eles aumenta. Além disso, há uma diminuição no grau de conectividade dos *hubs*. Dessa forma, haverá diminuição no tempo de atividade médio dos *hubs*. Lembre que na teoria BCPS o mecanismo de reinfeção centro $\rightarrow$ folhas $\rightarrow$ centro é o responsável por manter a estrela ativa, logo, se a estrela possui menos folhas, a probabilidade da reinfeção acontecer é menor e, na média, o *hub* fica ativo por menos tempo. Assim, é esperado que essas duas alterações causadas por essa estratégia de ataque modifiquem consideravelmente o limiar epidêmico. Já para a estratégia NVA-HP, esses efeitos são mais brandos, visto que $\langle \ell \rangle$ se altera de maneira

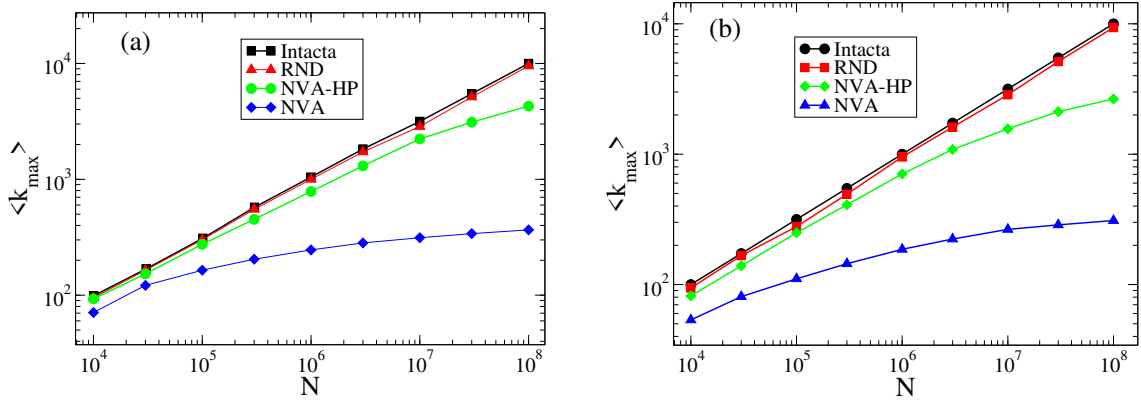


Figura 6.4 – Valor médio do grau do vértice mais conectado das redes intactas e danificadas por diferentes estratégias (média sobre 100 redes). Dois valores dos expoente foram usados: (a) $\gamma = 2.30$ (b) $\gamma = 2.80$.

mais branda assim como o valor de $k_{\max}$. A princípio, esperaríamos mudanças pequenas no limiar epidêmico. Entretanto, como veremos a seguir, mesmo no caso NVA-HP as mudanças em λ_p são grandes.

6.2 Efeitos da remoção de vértices no limiar epidêmico do modelo SIS

A partir da análise de como as diferentes estratégias de ataque alteram a estrutura da rede, estudamos efetivamente o efeitos desses ataques na propagação dos modelos epidêmicos. Inicialmente, faremos a análise dos modelos epidêmicos considerando $f = 0.1$. Conforme já visto, esse valor de f não é suficiente para fragmentar a rede. Na Figura 6.5, temos χ em função de λ para redes UCM (médias sobre 35 redes) e $\gamma = 2.30$, atacadas via método NVA-HP com $f = 0.1$. Pode-se notar que, em comparação com as redes do grupo de controle (mesmo N e γ mas sem ataques), o limiar epidêmico aumenta consideravelmente e converge para um valor finito a medida que N aumenta. Na Figura 6.6, temos o limiar epidêmico em função de N para os dois tipos de redes. Pode-se notar que, enquanto para as redes intactas λ_p converge para zero seguindo uma lei de potência como predito pela Eq. (3.8), para as redes danificadas há uma mudança drástica nesse comportamento e uma aparente convergência para um valor maior que zero.

Analizamos também redes com γ maior e, por conseguinte, menor número de *hubs*. Na Figura 6.7, temos χ em função de λ para redes UCM com $\gamma = 2.80$ danificadas via estratégia NVA-HP com $f = 0.1$. De forma similar aos resultados para $\gamma = 2.3$, há uma aparente convergência do limiar epidêmico para um valor maior que zero. Pela Figura 6.8, que mostra o limiar epidêmico tanto para as redes de controle quanto para as redes atacadas, temos uma alteração drástica na dependência de λ_p com N , mas a convergência para um valor finito é muito mais lenta do que no caso anterior.

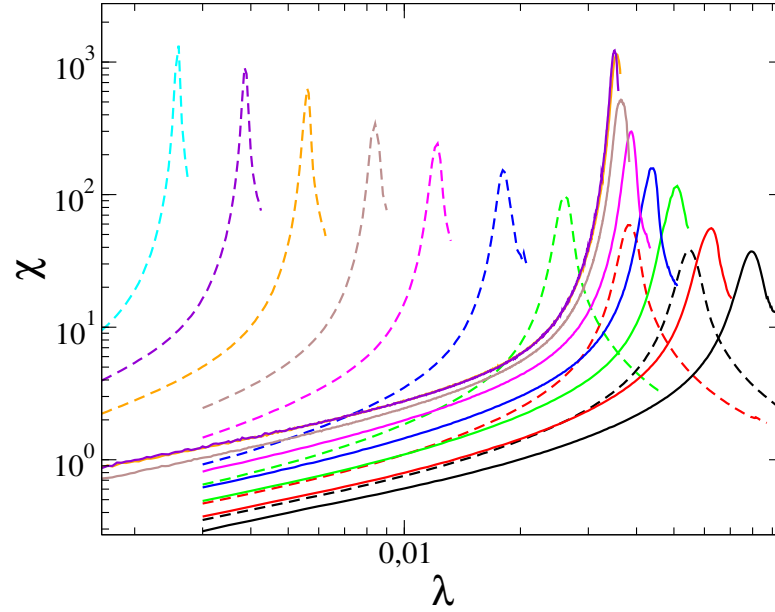


Figura 6.5 – Suscetibilidade em função de λ para redes UCM ($\gamma = 2.30$; $k_0 = 3$) atacadas pelo método NVA-HP (média sobre 35 redes) e diferentes tamanhos. Da direita para a esquerda: $N = 10^4$, 3×10^4 , 10^5 , 3×10^5 , 10^6 , 3×10^6 , 10^7 , 3×10^7 , 10^8 . Linhas cheias representam as redes atacadas e linhas tracejadas representam redes intactas.

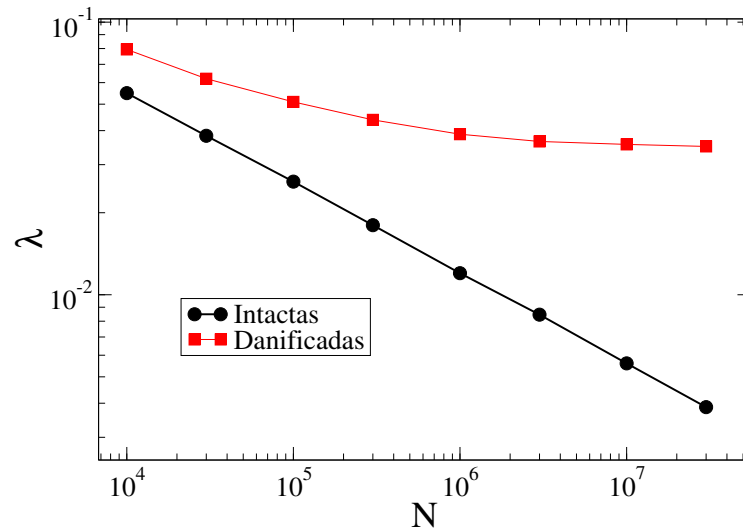


Figura 6.6 – Análise de tamanho finito do limiar epidêmico para redes UCM com $\gamma = 2.3$ e $k_0 = 3$ (média sobre 35 redes) intactas e danificadas via NVA-HP com $f = 0.1$.

Assim, temos um resultado interessante: apesar da estratégia NVA-HP aparentemente não danificar a rede de forma muito severa já que a rede danificada é composta quase inteiramente pela componente gigante e não alterar drasticamente algumas propriedades analisadas (menor caminho médio, grau máximo e $P(k)$), há uma grande mudança no limiar epidêmico do modelo SIS nessas redes. Assim, análises de outras grandezas são necessárias de forma a entender melhor esse resultado.

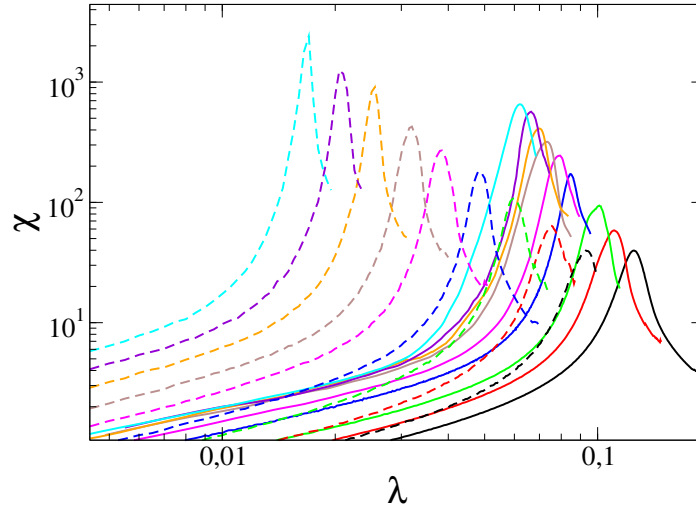


Figura 6.7 – Suscetibilidade em função de λ para redes UCM ($\gamma = 2.80$; $k_0 = 3$) atacadas (média sobre 35 redes) e diferentes tamanhos. Da direita para a esquerda: $N = 10^4$, 3×10^4 , 10^5 , 3×10^5 , 10^6 , 3×10^6 , 10^7 , 3×10^7 , 10^8 . Linhas cheias representam as redes atacadas e linhas tracejadas representam redes com mesmo N mas intactas (controle).

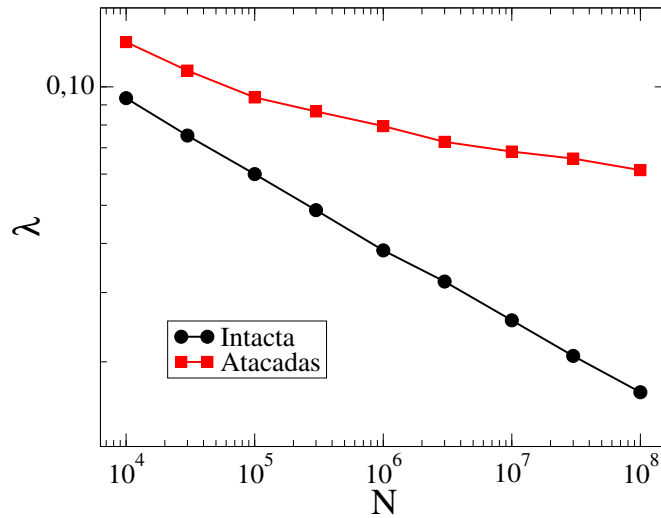


Figura 6.8 – Análise de tamanho finito do limiar epidêmico para redes UCM com $\gamma = 2.8$ e $k_0 = 3$ (média sobre 35 redes) intactas e para as redes danificadas via NVA-HP com $f = 0.1$.

Pela teoria BCPS, a sustentação da epidemia no modelo SIS é dada principalmente pela comunicação entre os *hubs* ativos da rede. Dessa forma, calculamos a quantidade de *hubs* que existem na rede, tanto antes dos ataques quanto depois. Porém, como definir exatamente se um determinado vértice é um *hub* ou não? Formalmente, um *hub* é qualquer vértice que possua grau da ordem de $k \sim N^\alpha$ para $\alpha > 0$, no limite termodinâmico. A partir disso, utilizamos o seguinte critério para identificar *hubs* em redes finitas: qualquer vértice cujo grau tenha probabilidade de ocorrência na rede menor que 1. Assim, $NP(k_{\text{hub}}) = 1 \rightarrow k_{\text{hub}} = (BN)^{1/\gamma}$ em que B é dado pela Eq. (2.26). Na Figura 6.9, temos

gráficos ilustrando os valores de k_{hub} calculados a partir desse critério para redes UCM com $\gamma = 2.30$ e 2.80 e $N = 10^6$.

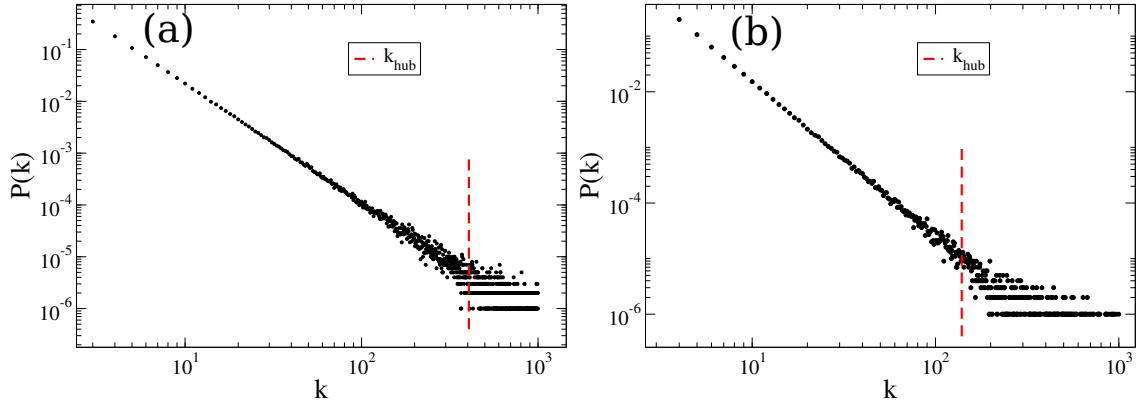


Figura 6.9 – Distribuição de graus para redes UCM intactas com $N = 10^6$ e indicações de k_{hub} . (a) $\gamma = 2.30$; $k_{\text{hub}} \approx 400$ (b) $\gamma = 2.80$; $k_{\text{hub}} \approx 130$.

Utilizando esse critério, calculamos a quantidade de *hubs* nas redes atacadas pelas diferentes estratégias. Os resultados estão mostrados na Figura 6.10. A partir dessas figuras vemos que, utilizando esse critério, a quantidade de *hubs* praticamente não se altera para ataques aleatórios, enquanto que para a estratégia NVA-HP há um comportamento não monotônico, com um crescimento no número de *hubs* para tamanhos pequenos sucedida por uma queda nessa quantidade, para tamanho grandes. Já para o NVA, o número de *hubs* sempre diminui com o aumento do tamanho. Pode-se notar também que os efeitos são mais brandos para redes que possuem naturalmente menos *hubs* ($\gamma = 2.80$), conforme mostra a Figura 6.10(b).

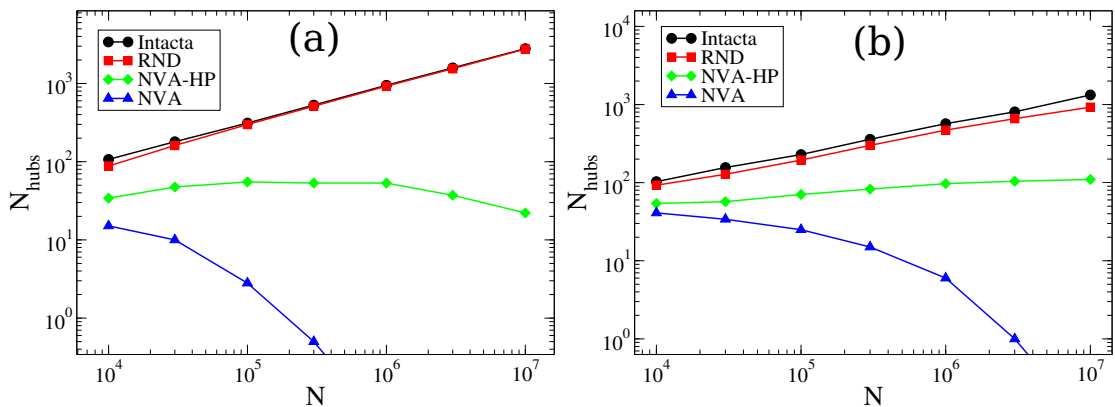


Figura 6.10 – Número de *hubs* em função do tamanho das redes (média sobre 100 amostras) considerando diferentes estratégias de ataque (a) $\gamma = 2.3$ (b) $\gamma = 2.8$.

Dessa forma, não somente o grau máximo das redes diminui quando danificadas pelas estratégias NVA e NVA-HP, mas o número de *hubs* diminui muito mais drasticamente quando comparado com as redes intactas. Essa diluição de *hubs* pela rede contribui ainda mais para dificultar a comunicação entre eles, aumentando ainda mais o tempo de

infecção entre *hubs*. Essa análise conjuntamente com as análises feitas a partir do cálculo do menor caminho médio e do $k_{\max}$, em termos da teoria BCPS, explicaria qualitativamente os resultados encontrados para o limiar epidêmico das Figuras 6.6 e 6.8. Além disso, esses resultados mostram que as estratégias de ataque NVA e NVA-HP são menos efetivas em redes que possuem menor número de *hubs* (o limiar epidêmico para redes com $\gamma = 2.80$ é menos alterado do que para redes com $\gamma = 2.30$).

Também analisamos a estratégia NVA nas redes atacadas e os resultados encontram-se na Figura 6.11. Nesse caso, o efeito da remoção dos vértices é mais evidente. Para $\gamma = 2.30$, o limiar epidêmico aumenta com o tamanho (vide Figuras 6.11(a) e (b)) após um determinado N enquanto que para $\gamma = 2.80$ a convergência para o valor finito é rápida. Pela análise estrutural realizada, é esperado que as variações no limiar sejam mais evidentes para esta estratégia do que para o NVA-HP já que as variações no menor caminho médio, no número de *hubs* e no grau máximo também são mais evidentes.

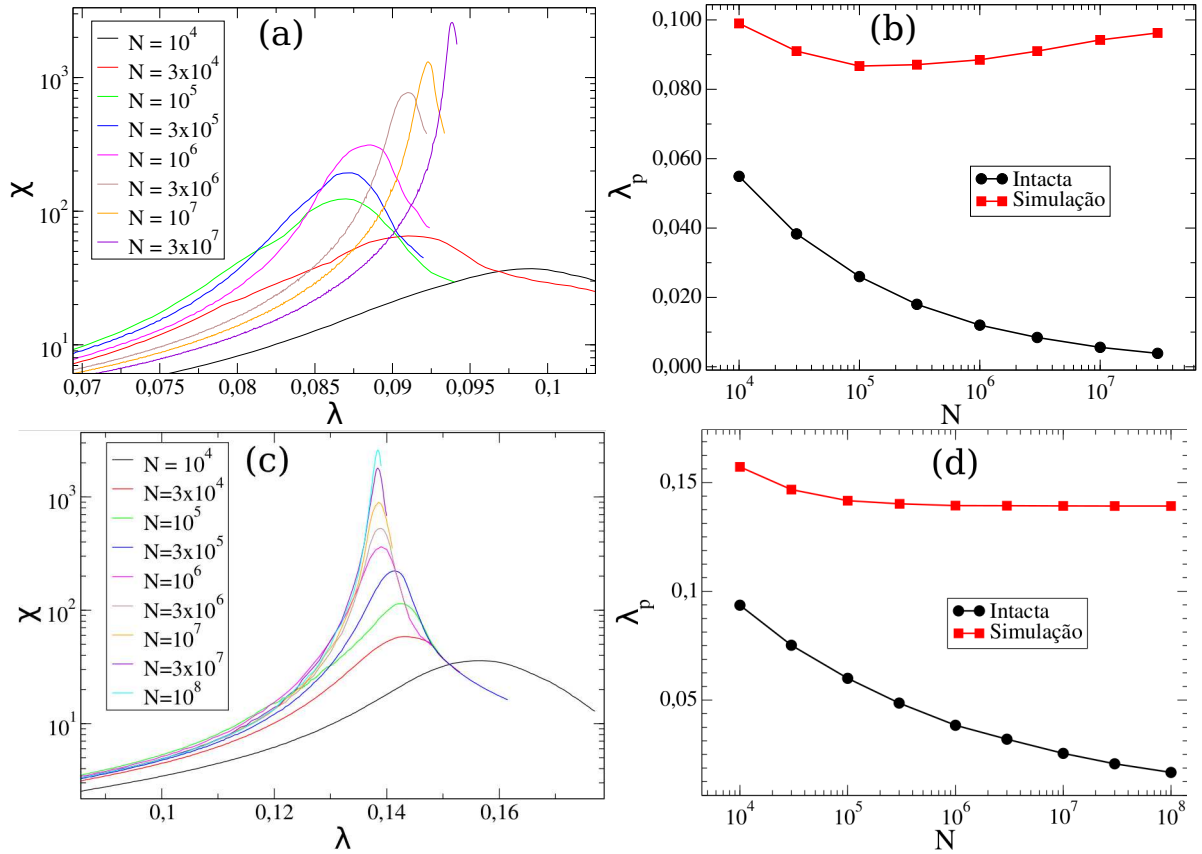


Figura 6.11 – (a) χ em função de λ para redes UCM ($\gamma = 2.30$) danificadas via estratégia NVA ($f = 0.1$). (b) análise de tamanho finito do limiar epidêmico considerando λ_p da figura (a). (c) χ em função de λ para redes UCM ($\gamma = 2.80$) danificadas via estratégia NVA ($f = 0.1$). (d) análise de tamanho finito do limiar epidêmico considerando λ_p da figura (c).

Estudamos ataques com outras frações de vértices removidos. Para fração 0.02, não foram encontradas variações para o limiar epidêmico nos tamanhos estudados (resultados não mostrados nessa dissertação). Como as estratégias visam a remoção de *hubs*, a

medida que aumentamos o tamanho da rede, os *hubs* vão ficando cada vez maiores e o dano causado pela remoção destes também. Dessa forma, conjecturamos que qualquer fração finita de vértices removidos irá fragilizar a rede e modificar o limiar epidêmico se esta possuir um tamanho suficientemente grande. Também analisamos ataques mais destrutivos ($f = 0.4$), cujos resultados estão mostrados nas Figuras 6.12 e 6.13. Podemos notar o efeito esperado, no qual um ataque mais destrutivo modifica o limiar epidêmico de forma mais brusca do que os mais brandos. Devido a limitações computacionais, não foi possível analisar redes maiores do que 3×10^6 vértices para $f = 0.4$, já que precisávamos gerar redes muito maiores para garantir que a componente gigante após o ataque fosse do tamanho desejado.

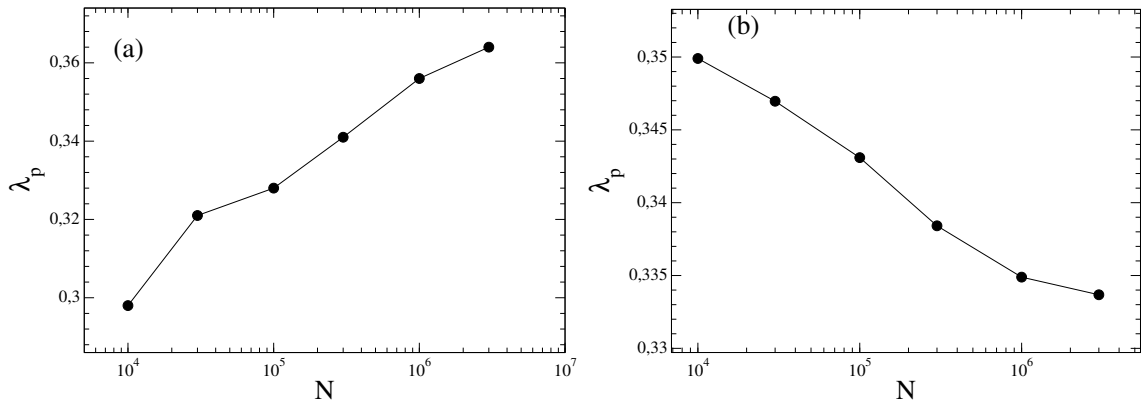


Figura 6.12 – Análise de tamanho finito do limiar epidêmico para o modelo SIS em redes UCM (média sobre 10 redes) atacadas pela estratégia NVA-HP ($f = 0.4$). (a) $\gamma = 2.30$ (b) $\gamma = 2.80$.

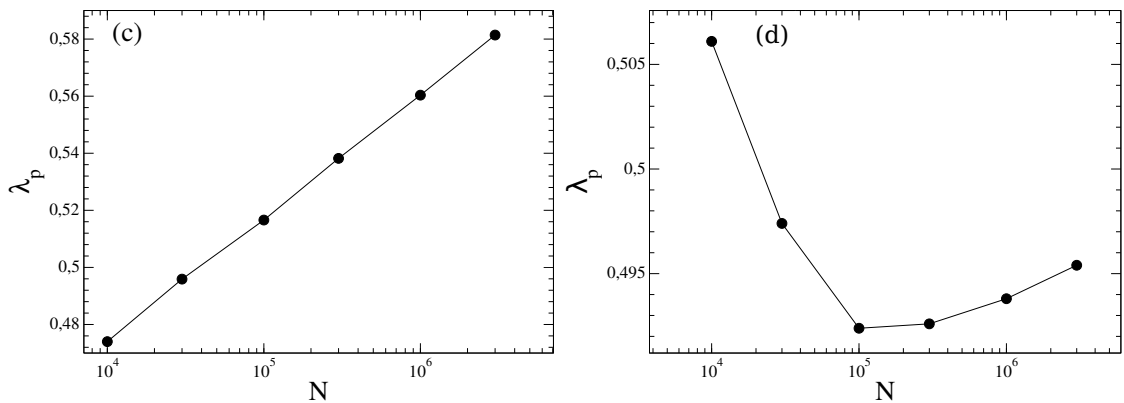


Figura 6.13 – Análise de tamanho finito do limiar epidêmico para o modelo SIS em redes UCM (média sobre 10 redes) atacadas pela estratégia NVA ($f = 0.4$). (a) $\gamma = 2.30$ (b) $\gamma = 2.80$.

6.3 Comparação com teorias de campo médio

Após a obtenção dos resultados para os limiares epidêmicos da seção anterior, comparamos esses resultados obtidos com as teorias de campo médio apresentadas na seção III.3. Na Figura 6.14, temos o limiar epidêmico em função de N para redes UCM com $\gamma = 2.3$ e 2.8 atacadas pelos métodos NVA-HP e NVA, em comparação com os limiares previstos pelas teorias HMF e QMF. Pode-se notar em todos os casos que a teoria HMF consegue capturar o comportamento do limiar epidêmico enquanto a teoria QMF falha em capturar nas redes atacadas pela estratégia NVA. Além disso, em todos os casos a teoria HMF foi quantitativamente mais precisa que a QMF.

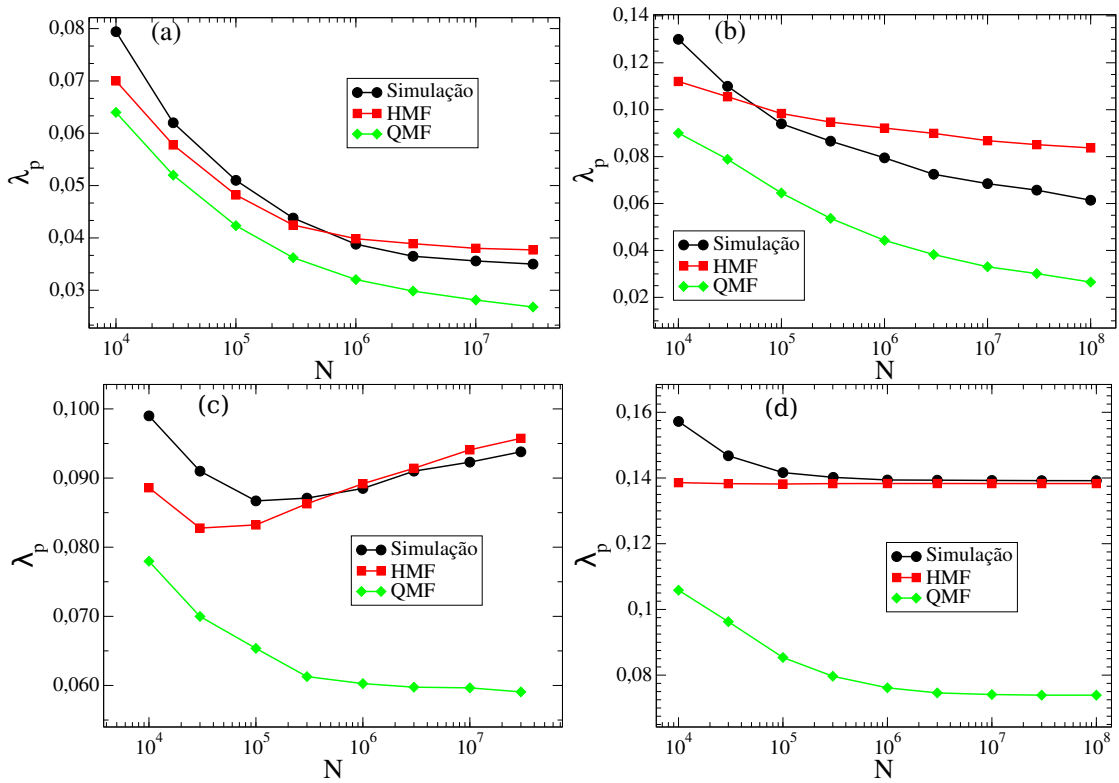


Figura 6.14 – Análise de tamanho finito do limiar epidêmico do modelo SIS em redes UCM danificadas além de comparação com as teorias HMF e QMF. (a) $\gamma = 2.3$ e NVA-HP (b) $\gamma = 2.8$ e NVA-HP (c) $\gamma = 2.3$ e NVA (d) $\gamma = 2.8$ e NVA.

A teoria QMF, ao incorporar explicitamente a estrutura da rede no cálculo do limiar epidêmico através da matriz de adjacência, acaba por capturar em suas equações os efeitos locais do modelo SIS, que são dominantes quando há *hubs* na rede. Enquanto isso, a teoria HMF assume uma interação média dos vértices de mesmo grau na rede e acaba por não capturar esses efeitos. Dessa forma, é esperado que em modelos epidêmicos que apresentam ativação de *hub* a teoria QMF se sobressaia sobre a HMF. Porém, a medida que atacamos a rede pelas diferentes estratégias, diminuimos o número de *hubs* e, portanto, deixamos a rede menos heterogênea e sujeita a menos efeitos de localização.

Dessa forma, assumir uma interação média torna-se uma aproximação melhor quanto mais homogênea for a rede, assim, é esperado uma melhora na teoria HMF a medida que removemos mais *hubs*.

6.4 Efeitos da remoção de vértices no λ_p do modelo SIR

Estudamos também o modelo epidêmico SIR, cuja transição de fase não é ativada por *hub* e diferentemente do SIS, não possui tantas controvérsias e complicações. Para o valor de γ estudado (2.30), a transição de fase no modelo SIS também não é ativada por *hubs*, assim, é esperado que as mudanças nos limiares sejam parecidas para ambos modelos. Na Figura 6.15, temos a variabilidade epidêmica em função de λ para redes UCM com $\gamma = 2.30$ atacadas via NVA-HP ($f = 0.1$), além do limiar epidêmico encontrado a partir dos picos de Δ . Pode-se notar um comportamento parecido ao do modelo SIS, com um limiar tendendo a zero nas redes intactas (resultado condizente com o mostrado na seção III.3) e convergindo para um valor finito maior que zero no caso das redes danificadas.

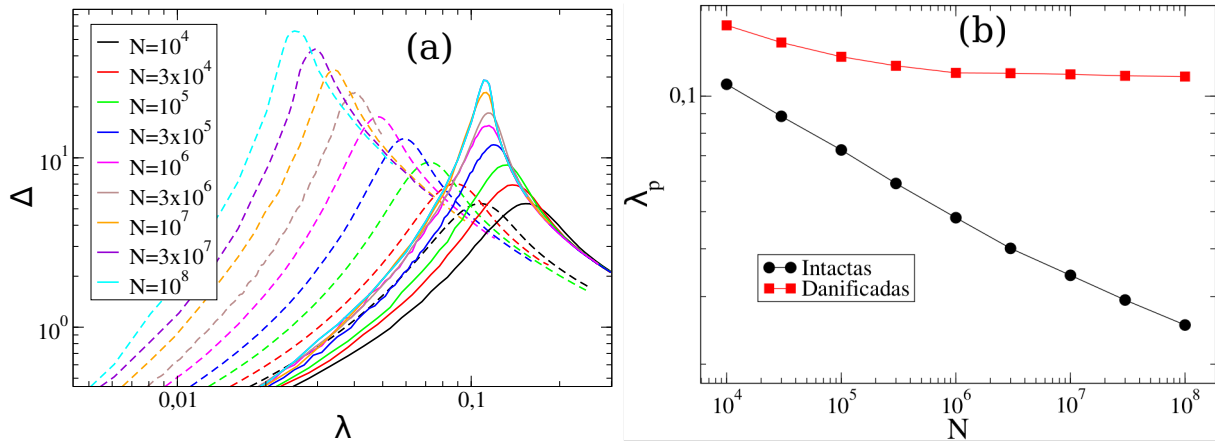


Figura 6.15 – (a) Δ em função de λ para redes UCM ($\gamma = 2.30$) danificadas via estratégia NVA-HP ($f = 0.1$). (b) análise de tamanho finito do limiar epidêmico considerando λ_p da figura (a).

VII Considerações finais

A construção de redes que sejam robustas a falhas e ataques e o desenvolvimento de estratégias de imunização ou de contenção de epidemias que sejam viáveis e eficientes são problemas fundamentais nos âmbitos social, econômico e, dentro de teoria de redes, científico. Aliados à estudos teóricos, as simulações constituem uma parte importante da validação de teorias e da compreensão de fenômenos que seriam muito difíceis de serem feitas de outro modo, seja pelo grande número de interações entre seus constituintes ou pela inexistência de soluções analíticas fechadas para as equações. O avanço da computação também é um fator crucial para o desenvolvimento da ciência de redes pois com um maior poder de processamento, mais facilmente podemos trabalhar com sistemas cada vez maiores.

No âmbito de teoria de redes, o estudo da fragilidade dessas a ataques é uma área importante no entendimento e prevenção de fenômenos como *blackouts*, *cyber attacks*, apagões de internet, entre outros. Além disso, o estudo de estratégias de ataques capazes de inutilizar a comunicação em uma rede é de vital importância na tentativa de desmantelar redes de crime [3], por exemplo. Porém, muitos estudos focam exclusivamente na caracterização estrutural [10, 11, 58] da rede após os danos, sem dar atenção aos processos dinâmicos que ocorrem nessas redes, como propagação de epidemias ou de informação. A partir disso, estudamos diretamente o impacto da remoção de vértices na propagação de epidemias.

Fizemos uma revisão geral sobre redes complexas, apresentando as principais medidas utilizadas para caracterizar estruturalmente uma rede, além de um tópico específico sobre redes sem escala, que aparecem constantemente ao estudar redes reais. Além disso, uma extensa revisão sobre processos epidêmicos, em particular sobre identificação e caracterização da transição de fase nos modelos SIS e SIR, além de comparações com as principais teorias de campo médio e a recente teoria BCPS [37]. As divergências e discussões acerca das teorias HMF e QMF, que apresentam resultados diferentes em alguns casos importantes também é um ponto central e foi levantado nessa dissertação.

Iniciamos os estudos através de três estratégias de ataque diferentes: o ataque aleatório, um modelo simples e utilizado como base para formular outras estratégias, em que um vértice é escolhido ao acaso e removido, o ataque de contato, que busca encontrar os nós mais conectados da rede, em que o vizinho de um vértice escolhido ao acaso é removido e o ataque de contato com proteção de *hub*, que adiciona aos vértices uma chance de se defender do ataque, chance que aumenta com o grau deste. Pensando em termos de um *cyberattack*, quando mais conectado o servidor, mais difícil será invadí-lo e derrubá-lo.

A partir da análise de como essas diferentes estratégias alteram a estrutura de uma

rede sem escala, podemos concluir que ataques aleatórios são ineficientes contra esse tipo de rede, já que é necessário remover praticamente toda a rede para fragmentá-la. O mesmo não ocorre com as outras duas estratégias, já que redes sem escala usualmente possuem muitos *hubs*, que são os mediadores da propagação da infecção no modelo SIS, sendo que essas estratégias visam justamente encontrá-los e removê-los. Assim, observamos que as redes são mais vulneráveis ao NVA e menos vulneráveis ao NVA-HP. Analisando o limiar de percolação, escolhemos valores de f para continuar estudando a rede estruturalmente. Escolhemos $f = 0.1$ por estar bem abaixo do limiar de percolação para ambas estratégias e pela componente gigante da rede danificada ser praticamente a rede toda. Além disso, estudamos o efeito de ataques mais destrutivos, com $f = 0.4$, valor no qual a componente gigante possui tamanho bastante inferior ao da rede ($\sim 30\%$) porém ainda grande. Por fim, a análise do menor caminho médio $\langle \ell \rangle$ nos levou a concluir que as redes continuam sendo de mundo pequeno, apesar do aumento significativo de $\langle \ell \rangle$ em todos os casos. Em ordem crescente de aumento temos o ataque aleatório, o NVA-HP e o NVA.

Estudamos também os efeitos das estratégias de ataque no maior grau da rede e concluímos novamente que o ataque aleatório não altera essa grandeza enquanto a estratégia NVA introduz um corte significativo em $k_{\max}$ e a estratégia NVA-HP introduz um corte mais brando. Dessa forma, analisando conjuntamente com as medidas de $\langle \ell \rangle$, pudemos estimar os efeitos que ocorreriam no limiar epidêmico através da teoria BCPS. Com o corte introduzido pela estratégia NVA aliado ao aumento de $\langle \ell \rangle$, o mecanismo de reinfeção de *hubs* fica prejudicado e é esperado que o limiar convirja para um valor não nulo no limite termodinâmico. Já para o NVA-HP, visto que o corte em $k_{\max}$ é mais brando e não há um aumento tão substancial em $\langle \ell \rangle$, é esperado que hajam mudanças mais leves no comportamento do limiar epidêmico.

Ao realizar as simulações do modelo SIS em redes atacadas de forma branda ($f = 0.1$), notamos que para o método NVA-HP há uma convergência do limiar epidêmico para um valor finito tanto no caso $\gamma = 2.3$ quanto para $\gamma = 2.8$. Dessa forma, a análise de $\langle \ell \rangle$ e $k_{\max}$ não é capaz de explicar esse fenômeno e análises posteriores tiveram de ser feitas. Assim, escolhemos um critério para definir *hubs* e calculamos o número de *hubs* antes e depois do ataque. Percebemos que no ataque aleatório não há diminuição significativa desse número enquanto que nos casos NVA-HP e NVA esse número diminui assintoticamente com o tamanho, com esse efeito sendo maior no NVA do que no NVA-HP. Assim, essa diluição dos *hubs* pela rede conjuntamente com a diminuição de $k_{\max}$ e o aumento de $\langle \ell \rangle$ levam a um limiar epidêmico maior do que zero. Para a estratégia NVA, o limiar epidêmico converge para um valor finito rapidamente para $\gamma = 2.8$ e apresenta um comportamento não monotônico para $\gamma = 2.3$, aumentando após determinado tamanho. Esses fenômenos também são condizentes com a análise estrutural feita. Estudamos também os efeitos de ataques mais destrutivos na rede, com $f = 0.4$ e constatamos que o comportamento do limiar epidêmico é condizente com o resto da análise, convergindo para um valor finito no caso de $\gamma = 2.8$ e crescendo com o tamanho no caso $\gamma = 2.3$.

Realizamos também simulações com o modelo SIR e constatamos que os resultados são similares.

Comparando os resultados das simulações com as teorias HMF e QMF, obtivemos um resultado interessante. Para o modelo SIS ambas teorias foram capazes de capturar o comportamento do limiar epidêmico (com exceção do caso NVA para redes com $\gamma = 2.3$, em que o limiar epidêmico cresce após determinado tamanho e a teoria QMF prevê um limiar convergindo monotonicamente), com a teoria HMF mais precisa quantitativamente em todos os casos. Isso levanta uma discussão sobre o limite de validade da teoria QMF, que em muitos casos para o modelo SIS é mais precisa do que a HMF. Concluimos então que ao incorporar a estrutura da rede de forma explícita nas equações, a teoria QMF é capaz de capturar efeitos locais, dominantes no modelo SIS e esses efeitos locais acabam por ajudar a teoria a ter um melhor desempenho. Quando essas localidades são removidas, por exemplo, ao removermos *hubs*, as aproximações da teoria QMF tornam-se piores do que as HMF.

Assim, estudamos estratégias de ataque que são simples de se implementar computacionalmente e viáveis em serem aplicadas a casos reais e que mesmo assim possuem a capacidade de danificar a rede de forma severa, alterando o limiar epidêmico de forma a dificultar a propagação de epidemias nessas redes através da remoção de uma fração relativamente pequena dos vértices. Além disso, mostramos que apesar dos ataques estudados não serem suficientes para fragmentar a rede, não é necessário que haja essa fragmentação para a rede se tornar inoperante do ponto de vista de propagação no modelo SIS. A rede mantém-se conectada porém de tal forma que a estrutura resultante inviabiliza a propagação da epidemia de forma eficiente.

Como perspectivas, pretendemos estudar os modelos epidêmicos em redes reais sujeitas aos ataques estudados nessa dissertação além do estudo de redes complexas que apresentem estruturas de comunidades ou redes *multiplex*, que são redes diferentes como *Twitter* e *Facebook*, interconectadas. Além disso, pretendemos analisar o efeito da remoção de vértices em outras grandezas envolvendo epidemias.

Referências

- [1] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008.
- [2] A.-L. B. Réka Albert, Hawoong Jeong, “Error and attack tolerance of complex networks,” *Nature*, vol. 406, 2000.
- [3] B. Requiao and S. Gonçalves, “The web of federal crimes in brazil: topology, weaknesses, and control,” *arXiv:1706.03153 [physics.soc-ph]*, 2017.
- [4] L. K. Gallos, F. Liljeros, P. Argyrakis, A. Bunde, and S. Havlin, “Improving immunization strategies,” *Phys. Rev. E*, vol. 75, p. 045104, 2007.
- [5] R. Cohen, S. Havlin, and D. ben Avraham, “Efficient immunization strategies for computer networks and populations,” *Phys. Rev. Lett.*, vol. 91, p. 247901, 2003.
- [6] V. Ágoston, P. Csermely, and S. Pongor, “Multiple weak hits confuse complex systems: A transcriptional regulatory network as an example,” *Phys. Rev. E*, vol. 71, p. 051909, 2005.
- [7] A. V. et al, “Assessing the international spreading risk associated with the 2014 west african ebola outbreak,” *PLOS Currents Outbreaks*, 2014.
- [8] V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, and A. Vespignani, “Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions,” *PLoS*, vol. 4, p. e13, 2007.
- [9] L. Medeiros, C. Castilho, C. Braga, W. Souza, L. Regis, and A. Monteiro, “Modeling the dynamic transmission of dengue fever: Investigating disease persistence,” *PLoS*, vol. 5, p. e942, 2011.
- [10] L. K. Gallos and P. Argyrakis, “Scale-free networks resistant to intentional attacks,” *Europhysics Letters*, vol. 80, no. 5, p. 58002, 2007.
- [11] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, “Attack vulnerability of complex networks,” *Phys. Rev. E*, vol. 65, p. 056109, 2002.
- [12] C. Boone, H. Bussey, and B. J Andrews, “Exploring genetic interactions and networks in yeast,” *Nature Reviews Genetics*, vol. 8, pp. 437–49, 2007.
- [13] M. P. van den Heuvel and O. Sporns, “Rich-club organization of the human connectome,” *The Journal of Neuroscience*, vol. 31, pp. 15775–86, 2011.

-
- [14] P. B. M. R. J. U. L. Backstrom and S. Vigna, “Four degrees of separation,” *Proceedings of the 4th Annual ACM Web Science Conference*, 2012.
 - [15] L. B. J. Ugander, B. Karrer and C. Marlow, “The anatomy of the facebook social graph,” *arXiv:1111.4503 [cs.SI]*, 2011.
 - [16] <http://konect.uni-koblenz.de>, “Konect network dataset.” Acessado em Agosto de 2017.
 - [17] A. Barabási, *Network Science*. Cambridge University Press, 2016.
 - [18] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, “Large-scale topological and dynamical properties of the internet,” *Physical Review E*, vol. 65, p. 066130, 2002.
 - [19] M. Newman, *Networks: An Introduction*. OUP Oxford, 2010.
 - [20] J. Travers and S. Milgram, “An experimental study of the small world problem,” *Sociometry*, 1969.
 - [21] F. Radicchi and C. Castellano, “Breaking of the site-bond percolation universality in networks,” *Nature Communications*, vol. 6, p. 10196, 2015.
 - [22] A.-L. Barabasi and R. Albert, “Emergence of scaling in random networks.,” *Science*, vol. 286, pp. 509–12, 1999.
 - [23] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” *Rev. Modern Phys.*, vol. 87, pp. 925–979, 2015.
 - [24] P. Moretti and M. Muñoz, “Griffiths phases and the stretching of criticality in brain networks,” *Nature Communications*, vol. 4, p. 2521, 2013.
 - [25] T. E. Harris, “Contact interactions on a lattice,” *Ann. Probab.*, vol. 2, no. 6, pp. 969–988, 1974.
 - [26] R. Anderson and R. May, *Infectious Diseases of Humans: Dynamics and Control*. OUP Oxford, 1992.
 - [27] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Phys. Rev. Lett.*, vol. 86, pp. 3200–3203, 2001.
 - [28] I. Corwin, “The kardar–parisi–zhang equation and universality class,” *Random matrices: Theory and applications*, vol. 1, no. 01, p. 1130001, 2012.
 - [29] S. C. Ferreira, C. Castellano, and R. Pastor-Satorras, “Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results,” *Physical Review E*, vol. 86, no. 4, p. 041125, 2012.

- [30] A. S. Mata and S. C. Ferreira, “Multiple transitions of the susceptible-infected-susceptible epidemic model on complex networks,” *Phys. Rev. E*, vol. 91, p. 012816, 2015.
- [31] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, “Critical phenomena in complex networks,” *Rev. Mod. Phys.*, vol. 80, pp. 1275–1335, 2008.
- [32] M. Boguñá and R. Pastor-Satorras, “Epidemic spreading in correlated complex networks,” *Physical Review. E*, vol. 66, p. 047104, 2002.
- [33] S. Strogatz, *Nonlinear Dynamics And Chaos*. Studies in nonlinearity, Sarat Book House, 2007.
- [34] F. Chung, L. Lu, and V. Vu, “Spectra of random graphs with given expected degrees,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 11, pp. 6313–6318, 2003.
- [35] C. Castellano and R. Pastor-Satorras, “Thresholds for epidemic spreading in networks,” *Phys. Rev. Lett.*, vol. 105, p. 218701, 2010.
- [36] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” *Rev. Mod. Phys.*, vol. 87, pp. 925–979, 2015.
- [37] M. Boguñá, C. Castellano, and R. Pastor-Satorras, “Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks,” *Phys. Rev. Lett.*, vol. 111, p. 068701, 2013.
- [38] S. Ferreira, R. S. Sander, and R. Pastor-Satorras, “Collective versus hub activation of epidemic phases on networks,” *Physical Review E*, vol. 93, 2015.
- [39] J. Hołyst, J. Sienkiewicz, A. Fronczak, P. Fronczak, and K. Suchecki, “Universal scaling of distances in complex networks,” *Physical review. E*, vol. 72, p. 026108, 2005.
- [40] D. Stauffer and A. Aharony, *Introduction to percolation theory*. Taylor and Francis London, 2nd ed. ed., 1992.
- [41] J. L. Cardy and P. Grassberger, “Epidemic models and percolation,” *Journal of Physics A*, vol. 18, no. 6, p. L267, 1985.
- [42] T. Beer and I. G. Enting, “Fire spread and percolation modelling,” *Mathematical and Computer Modelling*, vol. 13, pp. 77–96, 1990.
- [43] M. Sahimi, *Applications of Percolation Theory*. Taylor and Francis, 1994.
- [44] http://edition.cnn.com/2017/07/28/us/nc-outer-banks-power_outages/index.html, “Cnn.” Acessado em Janeiro de 2018.

- [45] <https://www.theguardian.com/technology/2017/may/12/global-cyber-attack-ransomware-nsa-uk-nhs>, “The guardian,” Acessado em Janeiro de 2018.
- [46] <https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/>, “Cbs,” Acessado em Janeiro de 2018.
- [47] M. Molloy and B. Reed, “A critical point for random graphs with a given degree sequence,” *Random Structures and Algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995.
- [48] R. Cohen, S. Havlin, and D. ben Avraham, “Efficient immunization strategies for computer networks and populations,” *Physical Review Letters*, vol. 91, p. 247901, 2013.
- [49] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, “Generation of uncorrelated random scale-free networks,” *Physical review. E*, vol. 71, p. 027103, 2005.
- [50] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, “Cut-offs and finite size effects in scale-free networks,” vol. 38, pp. 205–209, 2004.
- [51] D. T. Gillespie, “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions,” *Journal of Computational Physics*, vol. 22, no. 4, pp. 403 – 434, 1976.
- [52] W. Cota and S. C. Ferreira, “Optimized gillespie algorithms for the simulation of markovian epidemic processes on large and heterogeneous networks,” *Computer Physics Communications*, vol. 219, pp. 303 – 312, 2017.
- [53] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models*. Cambridge University Press, 2005.
- [54] R. Dickman and R. Vidigal, “Quasi-stationary distributions for stochastic processes with an absorbing state,” *Journal of Physics A*, vol. 35, 2001.
- [55] M. M. de Oliveira and R. Dickman, “How to simulate the quasistationary state,” *Phys. Rev. E*, vol. 71, p. 016129, 2005.
- [56] R. S. Sander, G. S. Costa, and S. C. Ferreira, “Sampling methods for the quasistationary regime of epidemic processes on regular and complex networks,” *Phys. Rev. E*, vol. 94, p. 042308, 2016.
- [57] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. Elsevier, 1992.
- [58] P. Van Mieghem, D. Stevanović, F. Kuipers, C. Li, R. Bovenkamp, D. Liu, and H. Wang, “Decreasing the spectral radius of a graph by link removals,” *Physical Review. E*, vol. 84, p. 016101, 2011.