

JOSÉ CARLOS DE MORAES SILVA

**O PAPEL DE SÍTIOS PERIFÉRICOS NO PROCESSO
EPIDÊMICO SIS**

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*.

Orientador: Sílvio da Costa Ferreira Júnior

**VIÇOSA - MINAS GERAIS
2019**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

S586p
2019

Silva, José Carlos de Moraes, 1989-
O papel de sítios periféricos no processo epidêmico SIS /
José Carlos de Moraes Silva. – Viçosa, MG, 2019.
62 f. : il. (algumas color.) ; 29 cm.

Orientador: Silvio da Costa Ferreira Junior.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f. 59-62.

1. Epidemiologia - Modelos matemáticos. 2. Redes
(Matemática). 3. Modelo SIS. I. Universidade Federal de Viçosa.
Departamento de Física. Programa de Pós-Graduação em Física
Aplicada. II. Título.

CDD 22. ed. 614.4

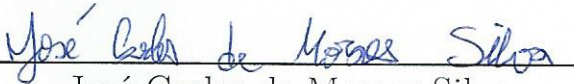
JOSÉ CARLOS DE MORAES SILVA

O PAPEL DE SÍTIOS PERIFÉRICOS NO PROCESSO
EPIDÊMICO SIS

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: 29 de julho de 2019.

Assentimento:



José Carlos de Moraes Silva

Autor



Sílvio da Costa Ferreira Júnior

Orientador

Aos meus pais, Vergínia Cristina e Antônio Carlos

Agradecimentos

Eu gostaria de agradecer primeiramente aos meus pais, Verginia Cristina e Antônio Carlos, por todo apoio e por saber que muito da minha caminhada até aqui se deve a eles. Ao meu primo e amigo Vinícius, pela quase rivalidade que muito tem me inspirado até hoje, aos meus amigos da Escola Técnica Estadual João Luiz do Nascimento, meus amigos da Universidade Federal Rural do Rio de Janeiro, onde me graduei. Agradeço muito à minha querida amiga Ethe Costa, que se tornou fundamental em tão pouco tempo durante meu mestrado, inclusive pela paciência comigo e pelo companheirismo incondicionais. Ao Eliel Leandro pela parceria e momentos de riso. Aos amigos do GISC pelo ambiente agradável e companheirismo - Guilherme, Diogo, Arthur e em especial ao Wesley Cota por toda a colaboração durante o mestrado, a qual foi indispensável. Agradeço ao Professor Sílvio pela orientação e incentivo, sem os quais este trabalho não teria sido realizado, e aos demais professores do Departamento de Física UFV, bem como o suporte fornecido pela Secretaria do Programa de Pós Graduação em Física. Agradeço também ao GISC e ao NSSC, pelo suporte computacional. Agradeço ao CNPq e à FAPEMIG, pelo apoio financeiro e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de financiamento 001, pela concessão da bolsa de estudos.

A vida imita o vídeo, garotos inventam um novo inglês vivendo num país sedento um momento de embriaguez - Somos quem podemos ser. Engenheiros do Hawaii

Resumo

SILVA, José Carlos de Moraes, M.Sc., Universidade Federal de Viçosa, julho de 2019. **O papel de sítios periféricos no processo epidêmico SIS.** Orientador: Silvio da Costa Ferreira Junior.

No que se refere ao estudo de processos dinâmicos em redes complexas, destaca-se a pesquisa envolvendo modelos epidêmicos, com ênfase para o modelo *suscetível-infectado-suscetível* (SIS). Muitos artigos científicos tem se ocupado no estudo do papel de sítios altamente conectados e sítios altamente centrais na dinâmica do modelo SIS. Dentre os temas desses estudos, podemos citar os mecanismos de ativação, como a ação coletiva de hubs sobre a epidemia, em redes com distribuição de graus em lei de potência, $P(k) \propto k^{-\gamma}$, com $\gamma > \frac{5}{2}$. Outros estudos apontam ainda para o papel dos sítios mais centrais no desencadeamento do estado ativo, como investigado por meio da decomposição *k-core* em redes com distribuição em lei de potência com $\gamma < \frac{5}{2}$. No entanto, há uma desproporcional falta de estudos sistemáticos no esforço de discriminar os impactos de sítios periféricos na dinâmica do modelo SIS. Dentre os efeitos esperados, destacamos a influência sobre a atividade na vizinhança dos hubs e a adição de atalhos, ligando hubs e sítios altamente centrais, além de caminhos ligando diferentes partes da rede. Neste trabalho, investigamos principalmente os efeitos de sítios com índice $k_S = 1$ e $k_S = 2$, definidos através da decomposição *k-core*. Com este objetivo, usamos diferentes mecanismos de ligação destes sítios a redes sintéticas com distribuição de grau k em lei de potência, $P(k) \propto k^{-\gamma}$, para diferentes valores de γ . Os mecanismos usados foram a *ligação preferencial linear*, na qual um sítio é escolhido ao acaso e aceito com probabilidade proporcional ao seu grau e a *ligação quadrática*, na qual um sítio é escolhido ao acaso e aceito com probabilidade proporcional ao quadrado de seu grau. Nós observamos que, quando a dinâmica é ativada pela componente mais densamente conectada, determinada pela decomposição *k-core*, a adição de sítios periféricos não altera a transição; mesmo no cenário mais extremo (ligação quadrática), em que muitos hubs mais conectados que os originais surgem. Ainda, quando a ativação é feita via ação combinada de hubs, observamos que a regra de ligação linear não muda o ponto de transição. Por outro lado, o modelo de conexão preferencial quadrática produz comportamentos novos e desvios dos comportamentos de escala usuais, observados em redes sem nós periféricos. É importante destacar que analisamos o modelo SIS também em redes reais. Estas evidenciam alta densidade de sítios periféricos ($k_S = 1$ e 2 , via decomposição *k-core*) e observamos que seu padrão de conexão assemelha-se muito mais ao modelo quadrático, mostrando, portanto, a importância do nosso estudo.

Palavras-chave: Epidemiologia - Modelos matemáticos. Redes (Matemática). Modelo SIS.

Abstract

SILVA, José Carlos de Moraes, M.Sc., Universidade Federal de Viçosa, July, 2019. **The role of peripheral vertices in SIS epidemic processes.** Adviser: Silvio da Costa Ferreira Junior.

As far as dynamical processes in complex networks started to be investigated, notable attention can be drawn to the research involving epidemic models, in special to the *susceptible-infected-susceptible* (SIS) model. Many scientific works have studied the role of highly connected and highly central nodes in the dynamics of the model, including systematic studies on the activation of the epidemics. These studies take into account the long-range interaction among hubs in triggering and sustaining the endemic state for networks with power-law degree distribution, $P(k) \propto k^{-\gamma}$ with $\gamma > \frac{5}{2}$. Other kind of activation involves nodes belonging to the innermost core of the network, investigated by means of the k -core decomposition, in triggering the active state in networks with power-law degree distribution with $\gamma < \frac{5}{2}$. Nonetheless, much less systematic studies are dedicated to investigate the effects of peripheral nodes in the dynamics of the SIS model. Among their influences, we highlight their action in enhancing the lifespan of activity in hubs and their neighborhood, or even in setting shortcuts linking hubs and highly central nodes and connecting different parts of the network. Therefore, we investigate in this work the effects of $k_S = 1$ and $k_S = 2$ nodes, defined by k -core decomposition. With this purpose, we use different kinds of attachment patterns of these nodes to networks with power-law degree distribution. To the proposed connection schemes account *the linear preferential attachment*, in which a node is randomly chosen and accepted with a probability proportional to its degree, and the *quadratic preferential attachment*, in which a node is randomly chosen and accepted with a probability proportional to the square of its degree. Results suggest that when activation takes place in the innermost core, addition of peripheral nodes does not alter the transition, even in the most extreme scenario considered, *i.e.*, the quadratic attachment scheme, in which many new hubs are created. Yet, when the activation is triggered by the interplay of hubs, linear preferential attachment scheme does not change significantly the transition. On the other hand, quadratic attachment scheme produces new behaviors and deviations from the scale-laws observed in the dynamics on top of the original networks. It is important to stress that we analyzed the SIS model in real world networks. These manifest high peripheral nodes density ($k_S = 1$ and 2), whose attachment patterns are similar to the quadratic scheme, showing, therefore, the importance of our studies.

Keywords: Epidemiology - Mathematical models. Networks (Mathematics). SIS model.

Lista de abreviaturas e siglas

SIS	Suscetível-Infetado-Suscetível
HMF	<i>Heterogeneous Mean Field</i>
QMF	<i>Quenched Mean Field</i>
UCM	Modelo de Configurações Não Correlacionado, do inglês <i>Uncorrelated Configuration Model</i>
OGA	<i>Optimized Gillespie Algorithm</i>

Sumário

I	Introdução	10
II	Teoria de redes complexas	13
2.1	A matriz de adjacência	13
2.2	Medidas em redes	14
2.3	Caminho e menor distância	15
2.4	Coeficiente de aglomeração	17
2.4.1	Correlação de grau	18
2.4.2	Modelo de Erdős-Rényi	21
2.4.3	Grau médio e a componente gigante	22
2.4.4	Modelo de Barabási-Albert	23
2.4.5	Corte natural de uma rede sem escala	25
2.4.6	Conflito entre falta de escala e correlação neutra	27
2.5	Decomposição k-core	28
III	Redes sintéticas e redes reais	30
3.1	O modelo de configurações e o modelo de configurações descorrelacionado	30
3.2	Redes com nós periféricos	32
3.2.1	Redes reais	33
3.2.2	Criação de redes com folhas	34
3.2.3	Criação de redes com tubos	36
IV	Processos epidêmicos em redes complexas	38
4.1	Modelos epidêmicos	38
4.1.1	A equação mestra	39
4.1.2	O modelo SIS	40
4.1.3	Teoria HMF	41
4.1.4	Teoria QMF	42
4.2	O método quase-estacionário	43
4.3	Mecanismos de ativação no modelo SIS	45
V	Resultados	47
5.1	Dinâmica SIS em redes reais	47
5.2	Resultados para redes sintéticas	48
5.2.1	Dinâmica em redes com folhas	49
5.2.2	Dinâmica em redes com tubos	53
VI	Considerações finais	57
	Referências	59

I Introdução

O estudo de redes tem ganhado bastante força nos últimos anos. Isto decorre da disponibilidade de enormes massas de dados associados a sistemas complexos reais e do aumento da capacidade de processamento e armazenamento de computadores aptos para lidar com esses dados [1]. O mundo moderno é permeado por redes complexas. Dentre elas destacam-se a rede mundial de computadores, representada pela *world wide web*, *redes sociais* e *redes de aeroportos*, para citar apenas alguns exemplos [2]. Com estas redes vem a sensação de encurtamento de distâncias - apenas alguns “apertos de mão” e estamos conectados a qualquer pessoa do globo, um número pequeno de roteadores nos separa de qualquer computador ou servidor conectado à internet ou, ainda, um pequeno número de conexões aéreas e/ou baldeações na rede ferroviária e chegamos a quase qualquer parte do mundo. Distâncias nunca foram tão curtas na história da humanidade. Com esta sensação de conexão e *mundo pequeno*, efeito homônimo conhecido na ciência de redes complexas, vem seus efeitos colaterais: uma doença suficientemente contagiosa pode também se espalhar pelo globo em tempo recorde [2]. Prever como as estruturas do mundo moderno facilitam a propagação de agentes contagiosos não é, portanto, um tópico de interesse teórico apenas, mas uma questão da garantia de mecanismos de isolamento de agentes contagiosos, a fim de que eles não se estabeleçam em um regime endêmico [3].

A difusão de informação e *softwares* maliciosos envolve mecanismos de propagação semelhantes àqueles envolvidos na propagação de doenças. Portanto, redes de contato, associadas à transmissão de doenças infecto-contagiosas, passam a dar lugar a redes sociais e redes de computadores neste cenário [4, 5, 6]. Portanto, modelos epidêmicos ganham uma importância capital e a busca por novos métodos de soluções das equações envolvidas, bem como o emprego de métodos computacionais na previsão de estados endêmicos são tarefas que podem ser encaradas com importância até mesmo vital.

Muitos estudos tem se ocupado até agora no estudo sistemático do papel de sítios altamente conectados na propagação de epidemias em redes complexas [7, 8], incluindo sua importância nos mecanismos de ativação do estado endêmico no modelo SIS - *suscetível-infectado-suscetível*. Na dinâmica do modelo, um sítio infectado transmite a infecção a um vizinho a uma taxa λ e torna-se espontaneamente suscetível a uma taxa μ . Dentre os mecanismos de ativação estudados, destacam-se a ativação via interação *hub-hub*, predominante em redes com distribuição de grau $P(k) \propto k^{-\gamma}$ com $\gamma > \frac{5}{2}$, e a ativação no núcleo densamente conectado da rede, predominante em redes com $\gamma < \frac{5}{2}$ [9, 10, 11]. No entanto, há uma desproporcional falta de estudos sistemáticos sobre os efeitos de sítios periféricos, como o caso de sítios em camadas com índice $k_S = 1$ e $k_S = 2$, definidas via *decomposição k-core*. Dentre os efeitos esperados estão a promoção de hubs com elevado tempo de atividade em suas vizinhanças e a criação de caminhos redundantes e atalhos

entre sítios de grau elevado ou comparavelmente centrais. Para investigar os efeitos das camadas $k_S = 1$ e $k_S = 2$ no modelo SIS, empregamos o uso do mecanismo de *ligação preferencial linear*, onde um sítio é escolhido aleatoriamente e aceito com uma probabilidade proporcional ao seu grau e a *ligação preferencial quadrática*, onde um sítio é escolhido aleatoriamente e aceito com uma probabilidade proporcional ao quadrado de seu grau.

Nossos resultados mostram a importância dos sítios periféricos na dinâmica, especialmente em regime de ligação quadrática em redes com $\gamma = 2.7$ e 3.5 . Os comportamentos não triviais das medidas *quase-estacionárias* e leis de escala evidenciam o papel central destes sítios na dinâmica. Em redes de tamanho elevado com $\gamma = 3.5$, a ligação linear mostra-se efetiva, mostrando que sítios periféricos podem ser fator decisivo na competição entre os mecanismos de ativação.

Essa tese possui mais 5 capítulos organizados da seguinte maneira. No Capítulo II é feita uma revisão de conceitos de teoria de grafos e redes complexas, passando pela descrição das principais propriedades de grafos em geral. São apresentados também alguns dos modelos clássicos em teorias de redes complexas, como o *Modelo de Barabási-Albert* e o *Modelo de Erdős-Rényi*, devido ao uso didático destes modelos na abordagem de conceitos importantes no desenvolvimento do presente trabalho. Ainda no Capítulo II, descrevemos brevemente a decomposição *k-core*, técnica que será importante na descrição dos *mecanismos de ativação* do modelo SIS, discutidos na seção 4.3. Por meio desta decomposição, a rede passa a ser enxergada como composta por *camadas*, ou *shells*, e um *núcleo*, ou *core*

No Capítulo III é apresentado o modelo *UCM*, utilizado na criação de redes em lei de potência sem correlação de graus. Na seção 3.2.1, apresentamos alguns exemplos de redes reais que motivaram as propriedades que fomentamos nas redes criadas através do nosso modelo, cuja descrição e propriedades estruturais básicas serão apresentadas e discutidas nas seções 3.2.2 e 3.2.3. A metodologia empregada em redes reais segue uma linha paralela à utilizada em redes sintéticas - ao invés de sítios de grau $k = 1$ ou 2 , imunizamos sítios com medida de centralidade baseada na decomposição *k-core*. Baseado nessa técnica, sítios de grau $k = 1$ e 2 dão lugar a sítios de índice $k_S = 1$ e 2 , respectivamente. Em contraposição, em redes sintéticas, criamos um núcleo com distribuição de graus em lei de potência e adicionamos os sítios periféricos de grau $k = 1$ e $k = 2$.

Uma breve apresentação dos principais processos epidêmicos em redes complexas, com atenção especial ao modelo SIS, é feita no Capítulo IV. Nele apresentamos a equação mestra e as principais teorias de campo médio usadas em seu estudo. Na seção 4.2, apresentamos o *método quase-estacionário*, importante ferramenta na amostragem de soluções *estacionárias* da equação mestra - ou mais precisamente, *quase-estacionárias* [12]. Na seção 4.3, dedicamos nossa atenção à discussão dos principais mecanismos de ativação da epidemia no modelo SIS.

Nossos resultados dinâmicos são apresentados no Capítulo V, onde nossa metodologia de investigação e os mecanismos conhecidos de ativação do modelo SIS são discutidos. Dentre os resultados, discorreremos sobre medidas realizadas em redes reais na seção 5.1 e na seção 5.2 dedicamo-nos à apresentação e discussão dos principais resultados deste trabalho, sobre redes sintéticas.

Terminamos então este trabalho discutindo nossos resultados e deixando a pedra fundamental para trabalhos futuros, bem como discutindo os possíveis empregos dos resultados obtidos na investigação dos efeitos de sítios periféricos na dinâmica do modelo SIS.

II Teoria de redes complexas

Uma forma bastante útil de abstração de um sistema físico pode ser obtida através de sua representação por meio de uma *rede* ou *grafo* $G(N,E)$ (na literatura matemática). No contexto dessa abstração, um conjunto N de *nós* ou *vértices* representará os constituintes do sistema. As relações ou interações existentes entre os nós da rede serão então representadas por um conjunto E de *ligações* ou *arestas*, dependendo da nomenclatura adotada pela área de estudo. Neste trabalho serão usados ambos os modos alternadamente.

Muitos são os sistemas que já são representados usando a linguagem de redes para alguma finalidade, indo da análise de qualidade de um texto [13], sincronização das componentes da rede elétrica de energia [14] e estratégias otimizadas de imunização de populações [15], para citar alguns.

2.1 A matriz de adjacência

Com o propósito de se representar matematicamente a estrutura de uma rede, costuma-se definir a respectiva *matriz de adjacência*, definida por

$$A_{ij} = \begin{cases} 1, & \text{se } i \text{ e } j \text{ possuem uma ligação} \\ 0, & \text{caso não possuam ligação,} \end{cases} \quad (2.1)$$

com i e j variando entre 1 e N , onde N é o número de nós da rede, tornando a matriz de adjacência uma matriz $N \times N$ [16]. A matriz de adjacência apresentada na equação (2.1) representa o caso especial de uma rede *não direcionada*, onde uma ligação saindo do sítio i para o sítio j implica na existência de um link saindo de j para o sítio i . Além disso, representa também uma rede *sem pesos*, na qual não são atribuídos pesos aos links da rede, ou ainda, na qual os pesos dos links da rede são iguais a 1. A fim de termos uma situação exemplo, consideremos a rede representada na figura 2.1.

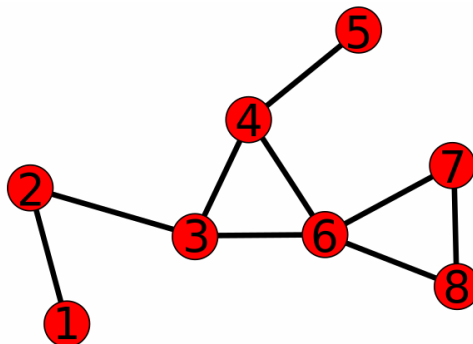


Figura 2.1 – Exemplo de rede

Ela possuirá a seguinte matriz de adjacência

$$A_{ij} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

Como pode-se observar da figura (2.1), o número de vizinhos de cada nó é representado através do vetor linha

$$k_i = [1 \ 2 \ 3 \ 3 \ 1 \ 4 \ 2 \ 2],$$

que resulta da soma de todos elementos de uma coluna i da matriz de adjacências A_{ij} .

Algumas observações gerais devem ser feitas sobre as matrizes de adjacências consideradas neste trabalho. O fato de haver apenas 0 ou 1 na matriz assim definida, implica que não haja *múltiplas conexões* entre os nós da rede. Além disso, os sítios da rede não conterão *auto-conexões*, implicando que os elementos da diagonal principal da matriz de adjacências sejam iguais a zero, ou seja, $A_{ii} = 0$. Uma rede que não apresenta múltiplas conexões é reconhecida na literatura como uma *rede* ou *grafo simples* [16].

2.2 Medidas em redes

Sem qualquer exagero, pode-se dizer que toda informação possível de se obter da rede está contida na matriz de adjacência. Por exemplo, a medida da *conectividade* ou *grau* de um nó i pertencente a uma rede indireta pode ser obtida de A_{ij} somando-se todos os elementos da linha i da matriz

$$k_i = \sum_{j=1}^N A_{ij}. \quad (2.2)$$

Esta medida *local* permite-nos estabelecer o tamanho da vizinhança de um sítio de nosso interesse. Apesar de simples de se obter, o grau é uma medida muito útil na caracterização *local* da estrutura da rede e, em muitos casos, da *centralidade* de um sítio.

O conhecimento geral do grau de todos os sítios da rede pode ser resumido através da *distribuição de graus*, definida através da equação

$$P(k) = \frac{N_k}{N}, \quad (2.3)$$

em que N_k representa o número de nós com grau k e N representa o número de nós da rede.

Notemos que, devido à sua natureza probabilística, $P(k)$ deve satisfazer a condição de normalização

$$\sum_{k=k_{\min}}^{k_{\max}} P(k) = 1. \quad (2.4)$$

Por meio da distribuição de graus, podemos facilmente obter o n -ésimo momento $\langle k^n \rangle$ da distribuição como

$$\langle k^n \rangle = \sum_{k=k_{\min}}^{k_{\max}} k^n P(k). \quad (2.5)$$

Dentre os momentos da distribuição de graus, devemos destacar os primeiro e segundo momentos, $\langle k \rangle$ e $\langle k^2 \rangle$, respectivamente, pela sua importância no cálculo de parâmetros estruturais, e mesmo dinâmicos em teorias de campo médio heterogênea dos modelos epidêmicos que nos interessarão no Capítulo IV.

2.3 Caminho e menor distância

A noção de caminho em redes complexas está implícita em muitas redes que usamos no nosso dia-à-dia - a *web*, ao navegar-se na internet; a *rede de aeroportos*, ao fazer-se conexões aéreas ou mesmo a *rede ferroviária*, ao fazer-se baldeação entre estações de trem, para nomear apenas alguns exemplos [17].

Definamos, portanto, o que é *caminho* em uma rede. Haverá um caminho de comprimento 1 entre dois nós i e j , caso, haja um link entre eles, ou seja, caso $A_{ij} = 1$. Um caminho de comprimento 2, caso haja um caminho de dois links conectando os dois nós, ou seja, caso $A_{im}A_{mj} = 1$, para algum m [16]. Podemos generalizar a ideia para um caminho l' , ou seja, haverá um caminho de comprimento l' entre dois nós i e j , caso

$$\underbrace{A_{im}A_{mn} \cdots A_{qj}}_{l' \text{ vezes}} = 1. \quad (2.6)$$

O número total de caminhos de comprimento 2 entre i e j é, portanto, dado pela expressão

$$N_{ij}^{(2)} = \sum_k A_{ik}A_{kj} = [A^2]_{ij}. \quad (2.7)$$

De forma similar, o número de de caminhos de comprimento 3 entre os mesmos dois nós é dado pela expressão

$$N_{ij}^{(3)} = \sum_{k,m} A_{ik}A_{km}A_{mj} = [A^3]_{ij}. \quad (2.8)$$

As expressões acima podem ser generalizadas para um caminho de comprimento qualquer, cujo número total de caminhos de comprimento l' , entre dois nós quaisquer, será dado por $N_{ij}^{(l')}$.

Como podemos ver até aqui, entre dois nós quaisquer da rede pode haver muitos caminhos. No entanto, muitas vezes interessa-nos o *menor caminho* ou *distância* entre dois nós, definido como

$$d_{ij} = \min(l'_{ij}). \quad (2.9)$$

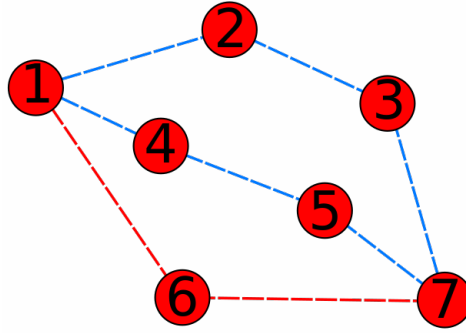


Figura 2.2 – Menor caminho, em vermelho.

A *distância média* $\langle l \rangle$ da rede pode ser então definida como

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{ij} d_{ij}. \quad (2.10)$$

Vamos agora olhar para uma rede com grau médio $\langle k \rangle$ e tamanho N . Olhando a partir de um nó qualquer, tem-se que ele tem distância 1 de seus primeiros $\langle k \rangle$ vizinhos, distância 2 de seus $\langle k \rangle^2$ segundos vizinhos e assim por diante, até totalizarmos $N(l)$ nós dentro de uma distância l [2]. Assim,

$$N_l \approx 1 + \sum_{j=1}^{j=l} \langle k \rangle^j = \frac{\langle k \rangle^{l+1} - 1}{\langle k \rangle - 1}. \quad (2.11)$$

Assumindo-se que $\langle k \rangle \gg 1$, o último termo da equação acima domina e temos, para $l = d_{max}$, que $N_{d_{max}} \approx N$

$$N \approx \langle k \rangle^{d_{max}}. \quad (2.12)$$

Tomando-se o logaritmo de ambos os lados da equação acima, temos

$$d_{max} \approx \frac{\ln N}{\ln \langle k \rangle}. \quad (2.13)$$

Apesar do cálculo estimar a distância máxima, também chamada de *diâmetro da rede*, esta estimativa mostra bastante bem o comportamento da distância média $\langle d \rangle$

em função do tamanho da rede, visto que a distância máxima d_{max} é dominada por alguns caminhos extremamente grandes, enquanto que $\langle d \rangle$ origina-se da média de todos os caminhos.

Podemos perceber da equação (2.13) que ela cresce com o logaritmo do tamanho da rede, diferentemente do que se espera de uma rede regular com dimensão n , cuja distância média se comporta como $N^{\frac{1}{n}}$. Para $N = 10^6$ e $\langle k \rangle = 3$, a distância média em uma rede aleatória é estimada em $\langle d \rangle \approx 4$, enquanto que para uma rede tridimensional, $\langle d \rangle = 100$. Este reduzido número de passos conectando quaisquer dois sítios de uma rede complexa carrega o nome de *efeito de mundo pequeno* e tem um papel importante na transmissão de epidemias na rede.

2.4 Coeficiente de aglomeração

É muito comum em redes de relacionamento interpessoal que uma pessoa que acabamos de conhecer seja também amiga ou conhecida de algum amigo bastante popular. Como veremos, este fato corriqueiro no dia-à-dia é o reflexo de uma medida estrutural de redes complexas: o *coeficiente de aglomeração* associado a um sítio da rede. Trata-se de uma medida do percentual dos primeiros vizinhos de um sítio que também são vizinhos entre si. Dados dois sítios vizinhos da rede, i e j , é possível que eles possuam um vizinho k em comum. Se um nó i e um de seus primeiros vizinhos j possuem um vizinho k em comum, isto significa que $A_{ij}A_{jk}A_{ki} = 1$.

A medida apresentada acima conta se há ou não há um 'triângulo' formado entre os sítios i , j e k , como visto na figura 2.14 a seguir. Assim, o número de triângulos formados em torno de i é dado por

$$\epsilon_i = \frac{1}{2} \sum_{jk} A_{ij}A_{jk}A_{ki}, \quad (2.14)$$

onde o fator $\frac{1}{2}$ aparece para evitar contar-se duas vezes o número de triângulos, devido ao somatório duplo.

Caso todos os vizinhos de um nó i estejam conectados entre si, i e sua vizinhança formam uma rede completamente conectada, também conhecida como grafo completo. Portanto, para medir o quão próximo a vizinhança de um nó está de ser um grafo completo, temos o coeficiente de aglomeração

$$C_i = \frac{2\epsilon_i}{k_i(k_i - 1)}, \quad (2.15)$$

definido na referência [16].

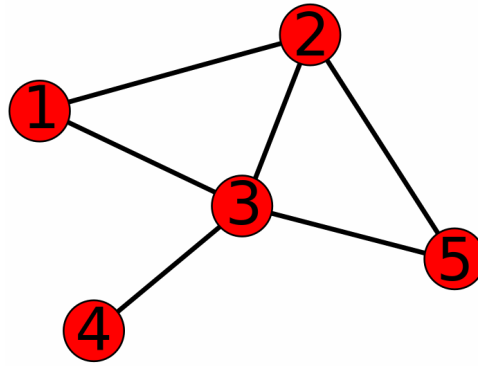


Figura 2.3 – Rede exemplificando cálculo de coeficiente de aglomeração.

Na figura 2.3, vemos que os sítios de índices 1, 2 e 5 possuem apenas dois vizinhos, que são também vizinhos entre si, o que lhes confere automaticamente $C_i = 1$. O sítio central de índice 3 possui 4 vizinhos, cujo grafo completo associado renderia um total de 6 triângulos. No entanto, dentre seus vizinhos, há a formação de apenas 2 triângulos entre ele e seus vizinhos, portanto, $C_3 = \frac{1}{3}$.

2.4.1 Correlação de grau

A afinidade de ligação entre sítios de graus diferentes é uma medida que pode trazer informações importantes sobre a estrutura da rede. Uma forma de expressar essa afinidade dá-se por meio da probabilidade de que, dado que um determinado nó tenha grau k , ele esteja conectado a um nó de grau k' , sendo fornecida pela equação

$$P(k'|k) = \frac{E_{kk'}}{kN_k}, \quad (2.16)$$

onde $E_{kk'}$ é o número de links conectando nós de grau k e k' e kN_k é o número de links saindo de todos os nós de grau k . Uma medida mais direta dessa afinidade é dada pelo grau médio dos vizinhos mais próximos κ_{NN}

$$\kappa_{NN}(k) = \sum_{k'} k' P(k'|k). \quad (2.17)$$

Dos comportamentos possíveis de κ_{NN} , podemos destacar a *correlação desassociativa*, na qual κ_{NN} decresce monotonicamente em função de k ; a *correlação associativa*, em que o κ_{NN} cresce monotonicamente em função de k e a *correlação neutra*, em que κ_{NN} é constante.

Em redes com **correlação desassociativa**, nós de grau pequeno tendem a ter vizinhos de grau elevado e vice-versa. O resultado disso é que o *grau médio dos vizinhos mais próximos* κ_{NN} tenderá a ser elevado para nós de grau pequeno e baixo para nós de grau elevado.

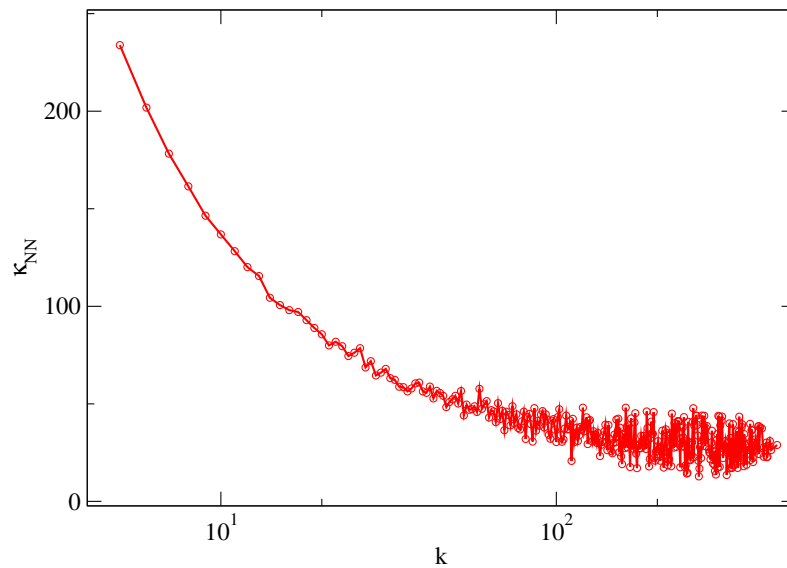


Figura 2.4 – Grau médio dos vizinhos mais próximos κ_{NN} com comportamento desassociativo. Modelo de Barabási-Albert generalizado, com os parâmetros $A = -2.4$ e $k_{min} = 5$

A figura (2.4) ilustra como o padrão de ligação desassociativa se manifesta em redes complexas, exibindo neste exemplo a inclinação de sítios de grau baixo de estarem conectados a sítios de grau elevado e vice-versa.

No caso de **correlação associativa**, nós tendem a ter vizinhos de grau semelhante aos seus. Isto quer dizer que nós de grau pequeno tem tendência a se ligar com nós de grau pequeno ou comparável, ocorrendo o mesmo com nós de grau elevado.

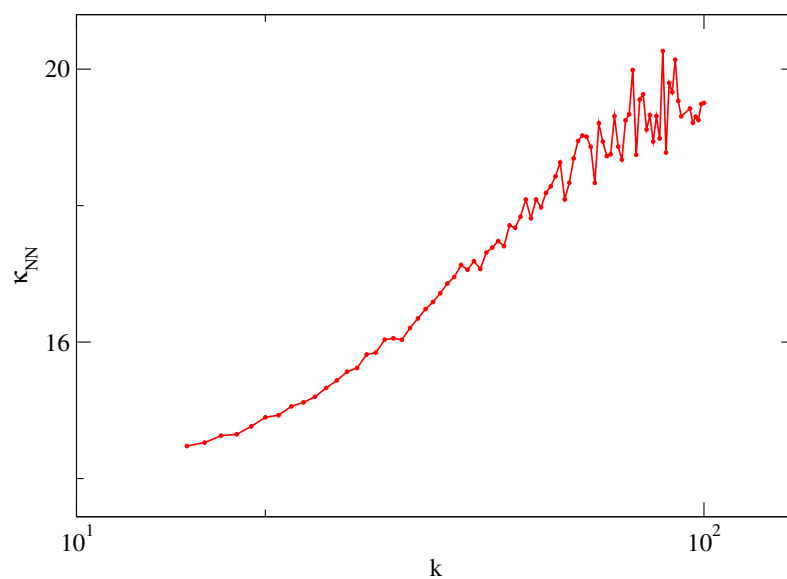


Figura 2.5 – Grau médio dos vizinhos mais próximos κ_{NN} com comportamento associativo. Modelo de Barabási-Albert generalizado com os parâmetros $A = 5.0$ e $k_{min} = 5$.

Padrões de correlação de graus associativos podem ser bem ilustrados através da figura (2.5). Nela, fica bastante claro que, em média, sítios de grau baixo tendem a estar ligados a sítios de grau comparavelmente baixo. Analogamente, nós de grau elevado tendem a estar conectados a nós de grau também elevado.

Em redes **descorrelacionadas** não ocorrem as tendências anteriores e o grau médio do vizinho mais próximo é distribuído homoganeamente.

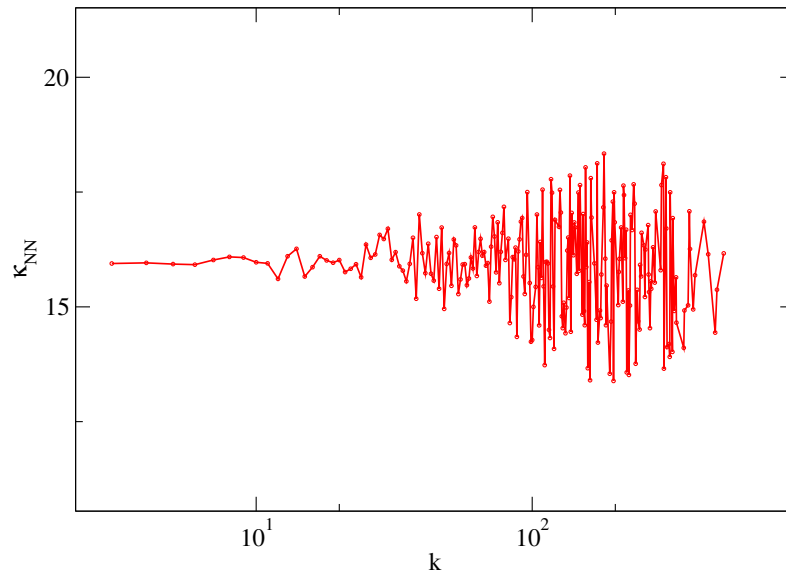


Figura 2.6 – Grau médio dos vizinhos mais próximos κ_{NN} em rede sem correlação de grau. Modelo de configurações não-correlacionado, $\gamma = 3.0$.

Na figura (2.6), é apresentada um exemplo de distribuição de grau dos primeiros vizinhos κ_{NN} descorrelacionada. Neste padrão de comportamento, espera-se que, em média, a vizinhança de sítios de qualquer grau seja homogênea.

Este padrão de correlação permite-nos fazer algumas previsões. Sem correlação de grau, podemos esperar que $P(k'|k)$ não dependa de k , sendo dado por

$$P(k'|k) = \frac{k'N}{\sum_{k''} k''N_{k''}} = \frac{k'N}{N\langle k \rangle}, \quad (2.18)$$

portanto

$$P(k'|k) = \frac{k'P(k')}{\langle k \rangle}. \quad (2.19)$$

Neste caso, usando-se (2.19) na equação (2.17), temos

$$\kappa_{NN} = \frac{\langle k^2 \rangle}{\langle k \rangle}. \quad (2.20)$$

2.4.2 Modelo de Erdős-Rényi

O modelo de Erdős-Rényi se baseia na construção de uma rede aleatória de tamanho N de uma forma simples. Tomando-se dois sítios da rede aleatoriamente, estes são conectados com uma probabilidade p . Sendo assim, a distribuição de probabilidades de encontrarmos na rede um sítio com k vizinhos é $p^k(1-p)^{N-k-1}$. Em outras palavras, esta é a probabilidade de que um sítio esteja conectado com k outros sítios e desconectado dos demais nós da rede. Ora, k sítios vizinhos podem ser tomados de $\binom{N-1}{k}$ formas [2]. Portanto,

$$P_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \quad (2.21)$$

que é uma distribuição *binomial*.

Decorre desta distribuição de probabilidades que o grau médio da rede é dado pela expressão

$$\langle k \rangle = \sum_{k=0}^{N-1} k \binom{N-1}{k} p^k (1-p)^{N-1-k} = p(N-1), \quad (2.22)$$

Da expressão para o *n-ésimo momento*, dado pela equação (2.5), podemos, de maneira similar, obter

$$\langle k^2 \rangle = p(1-p)N + p^2 N^2. \quad (2.23)$$

Munidos da expressão (2.23) e levando em conta que a variância obedece à relação $\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2$, podemos obter o desvio padrão através da equação

$$\sigma_k = (p(1-p)N)^{\frac{1}{2}} = (\langle k \rangle (1-p))^{\frac{1}{2}}. \quad (2.24)$$

Fazendo-se $N \rightarrow \infty$ e $p \rightarrow 0$, mas mantendo $\langle k \rangle$ fixo, podemos demonstrar que a distribuição binomial (2.21) se reduz à distribuição de Poisson

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}, \quad (2.25)$$

tendo desvio padrão

$$\sigma_k = \langle k \rangle^{\frac{1}{2}}. \quad (2.26)$$

Como podemos ver, o desvio padrão de k não depende do número de sítios, ficando a escala da rede dependente apenas de $\langle k \rangle$.

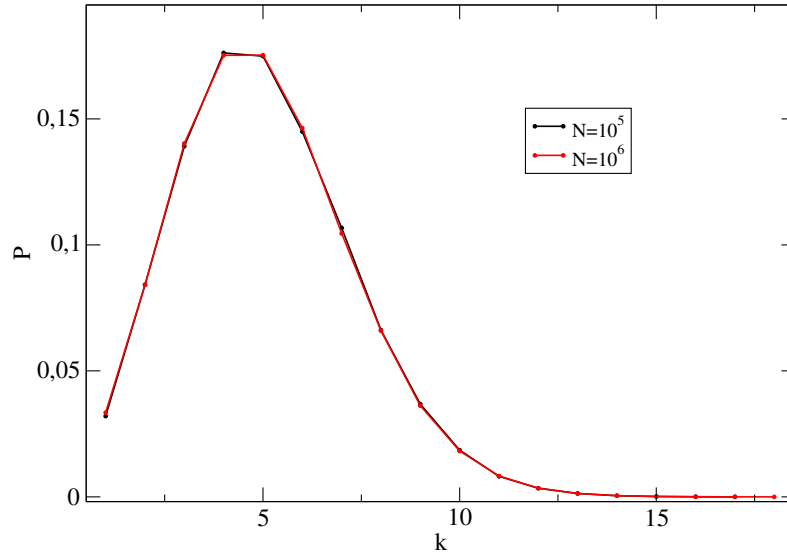


Figura 2.7 – Distribuições de grau para o modelo de Erdős-Rényni de tamanhos $N = 10^5$ e $N = 10^6$ e grau médio $\langle k \rangle = 5$.

A figura (2.7) mostra a distribuição de graus para duas redes de Erdős-Rényni com mesmo grau médio $\langle k \rangle$. Apesar dos tamanhos distintos, as duas distribuições de graus coincidem.

2.4.3 Grau médio e a componente gigante

Uma rede pode ser composta de diferentes fragmentos, ou seja, subgrafos desconectados uns dos outros. Em outras palavras, não há caminhos possíveis entre quaisquer duas componentes desconectadas, ficando impossível que um vértice de uma componente seja alcançado - via uma sequência de vértices - por um vértice de outra componente.

Para que surja um subgrafo formado por uma fração finita da rede, é necessário que, em média, os primeiros vizinhos de um sítio tenham grau igual ou maior que 2. Em outras palavras, a estrutura de uma rede com este atributo favorece que, dado que dois sítios quaisquer estejam conectados, estes estarão também conectados a outros sítios. Portanto, munidos da equação (2.20), podemos expressar este critério como

$$\frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2. \quad (2.27)$$

A equação (2.27) representa o *critério de Molloy-Reed* [18]. Comparando-se o critério de Molloy-Reed com a equação (2.20), válida para uma rede não-correlacionada, podemos dar à equação (2.27) uma interpretação muito simples: para o aparecimento da componente gigante, é necessário que o número médio de vizinhos de um vizinho da rede seja igual ou maior do que 2, a fim de que um sítio da rede esteja conectado aos demais.

Levando-se em consideração o segundo momento da distribuição de Poisson e o critério dado pela equação (2.27), vemos que, para uma rede de Erdős-Rényni, o surgimento de uma componente contendo uma fração finita da rede passa a se dar para $\langle k \rangle = 1$.

A figura 2.8 ilustra a fração da componente altamente conectada em função de $\langle k \rangle$ para uma rede de Erdős-Rényi, ocorrendo no ponto previsto pela equação (2.27). A transição arredondada é devida ao tamanho finito da rede.

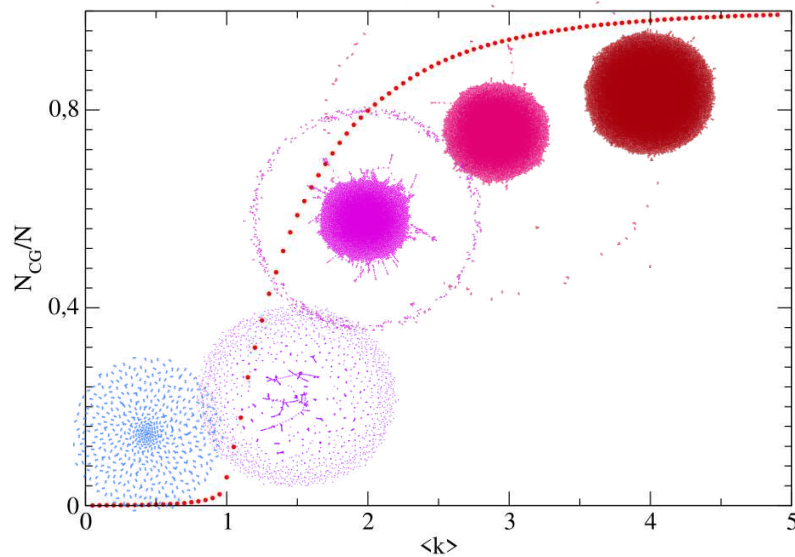


Figura 2.8 – Fração da componente gigante vs grau médio. Figuras ilustram amostras de rede com o respectivo grau médio

2.4.4 Modelo de Barabási-Albert

O modelo de Barabási-Albert apresenta uma combinação entre *crescimento* e *ligação preferencial* para o surgimento de redes *sem escala* [2], em contraposição ao modelo de Erdős-Rényi que, como vimos anteriormente, apresenta escala. O modelo mostra de forma simples como se dá o surgimento de redes com distribuição de grau em lei de potência e leva em conta que a rede é também uma entidade dinâmica, crescendo ao longo do tempo.

Neste modelo, começa-se com uma rede possuindo um número N_0 de sítios conectados inicialmente por m_0 arestas. A cada instante de tempo t um novo sítio é adicionado e se conecta com m sítios já existentes na rede. A probabilidade de que um sítio j vá se conectar com o novo sítio é dada por

$$\Pi_j = \frac{k_j}{\sum_{i=1}^N k_i}. \quad (2.28)$$

Visto que um sítio de grau k_j tem m oportunidades de mudar seu grau num instante de tempo t , com probabilidade de ser escolhido dado pela equação acima, sua taxa de variação temporal é dada por

$$\frac{dk_j}{dt} = \frac{mk_j}{\sum_{i=1}^N k_i}. \quad (2.29)$$

Consideremos que no instante de tempo $t = 0$ houvesse N_0 sítios, interligados por m_0 links, contanto que nenhum sítio fique sem ligações. Após t passos de tempo, haverá $m_0 + mt$ links na rede, agora formada por $N = N_0 + t$. No instante t , um sítio chega e se liga a m sítios da rede. Para ele, há $N = N_0 + t - 1$ sítios na rede e um total de $m_0 + m(t-1)$ links, portanto, $\sum_{i=1}^{N-1} k_i = 2(m_0 + m(t-1))$. Para $t \gg 1$, a equação (2.29) fica

$$\frac{dk_j}{dt} = \frac{k_j}{2t}. \quad (2.30)$$

Integrando-se esta equação entre $t' = t_j$ e $t' = t$ e levando-se em conta que, em $t' = t_j$ o sítio j chega à rede com m ligações, temos

$$k_j(t) = m \left(\frac{t}{t_j} \right)^{\frac{1}{2}}. \quad (2.31)$$

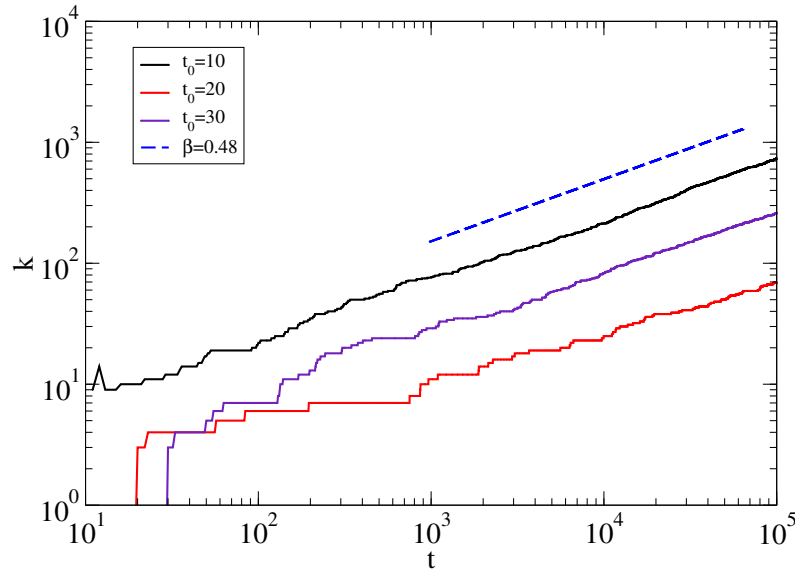


Figura 2.9 – Evolução temporal do grau de 3 sítios da rede. Parâmetros: $N_0 = 5$, $m_0 = 4$ e $m = 3$

Nesta equação para a evolução temporal do grau de um sítio que chega no instante de tempo t_j , vemos que, quanto menor o seu tempo de chegada, maior será seu grau num instante de tempo t . Isto significa que os sítios com maior conectividade serão aqueles que chegam mais cedo à rede ou que já estavam nela no seu início [2].

Para um instante de tempo t_j há um número $t_j - 1$ de sítios na rede com grau maior do que o grau do sítio que chega. Portanto, há um número $t_j = t \left(\frac{m}{k} \right)^2$ de sítios com grau maior do que k . Visto que, para tempos já longos, $N \approx t$, tempos que, a probabilidade de se encontrar um sítio com grau maior do que k é $\left(\frac{m}{k} \right)^2$. Portanto, a

probabilidade acumulada de se encontrar um sítio de grau menor ou igual a k é

$$P(k) = 1 - \left(\frac{m}{k}\right)^2. \quad (2.32)$$

Visto que esta é a probabilidade acumulada, devemos derivar esta equação no intuito de se achar a distribuição de probabilidades de se encontrar um sítio de grau k .

$$p(k) = \frac{dP}{dk} = 2mk^{-3}. \quad (2.33)$$

Notemos que este é um cálculo que aproxima o grau dos sítios da rede como uma variável contínua, sendo o cálculo exato feito levando-se em conta a natureza discreta da distribuição de graus [19].

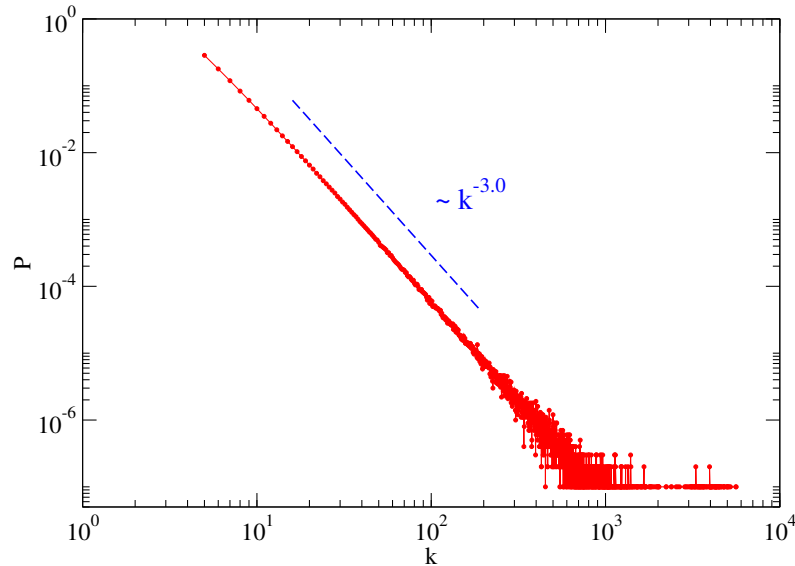


Figura 2.10 – Distribuição de graus para rede de Barabási-Albert. Parâmetros: $N = 10^6$, $N_0 = 10$, $m_0 = 5$ e $m = 5$

2.4.5 Corte natural de uma rede sem escala

Redes com distribuição de graus da forma $P(k) \propto k^{-\gamma}$ aparecem em uma gama enorme de sistemas, indo desde redes de aeroportos a redes de relacionamento humano e, devido à sua constante presença, redes dessa natureza tem sido estudadas extensivamente.

Definindo-se a probabilidade de obter um sítio de grau k a menos de uma constante de proporcionalidade e usando a condição de normalização, temos

$$\sum_k Ak^{-\gamma} = 1, \quad (2.34)$$

o que nos leva a

$$A = (\gamma - 1) \left(\frac{1}{k_{min}^{1-\gamma}} - \frac{1}{k_{max}^{1-\gamma}} \right), \quad (2.35)$$

que, para $k_{max} \gg 1$, reduz-se a

$$A = k_{min}^{\gamma-1}. \quad (2.36)$$

A fim de estimar o grau máximo da rede, devemos esperar encontrar apenas um sítio com grau máximo ou maior.

Assim,

$$N \int_{k_{max}}^{\infty} p(k) dk = 1. \quad (2.37)$$

Resolvendo-se a equação para k_{max} , temos

$$k_{max} \simeq k_{min} N^{\frac{1}{\gamma-1}}. \quad (2.38)$$

Este é um resultado importante que estima o grau máximo de uma distribuição em lei de potência e é conhecido como *corte natural da rede*. Uma consequência interessante do comportamento do corte natural da rede é que ele determina um expoente mínimo para redes complexas sem múltiplas ligações. Visto que $\frac{1}{\gamma-1} > 1$ para $\gamma < 2$, o grau máximo da rede se torna maior do que o número de sítios da rede neste caso. Assim, isto será apenas realizável com a presença de múltiplas ligações [20].

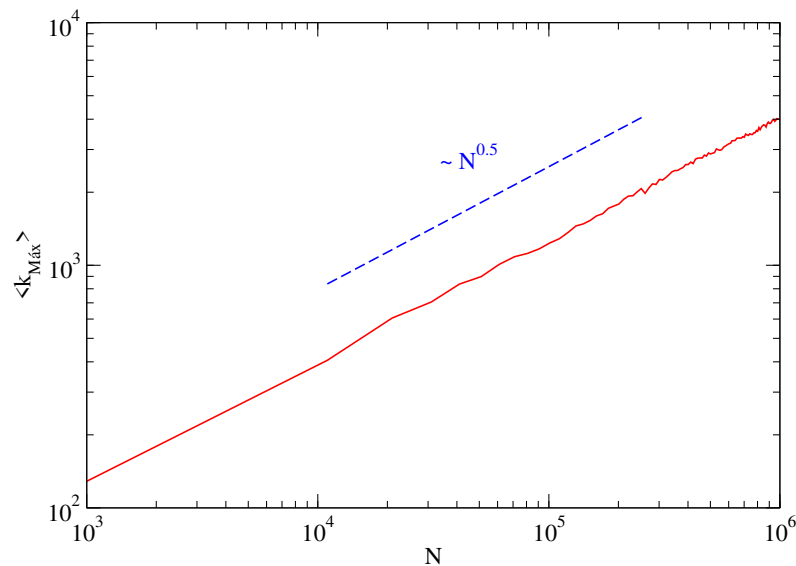


Figura 2.11 – Grau máximo médio para rede de Barabási-Albert. Parâmetros: $N_0 = 10^3$, $N = 10^6$, $m_0 = 5$ e $m = 5$

Vale notar que o grau máximo definido acima só é válido para a média de uma coleção de redes definidas com o mesmo parâmetro, havendo na prática flutuações grandes em torno deste valor. A figura (2.11) mostra como escala o grau máximo de redes de Barabási-Albert com o tamanho, concordando com a previsão dada pela equação (2.38). Vemos, no entanto, que o mesmo havendo um grau máximo esperado para uma rede de

Barabási-Albert finita, ele está muito acima do que se espera de uma rede de Erdős-Rényi, como mostra a figura 2.12.

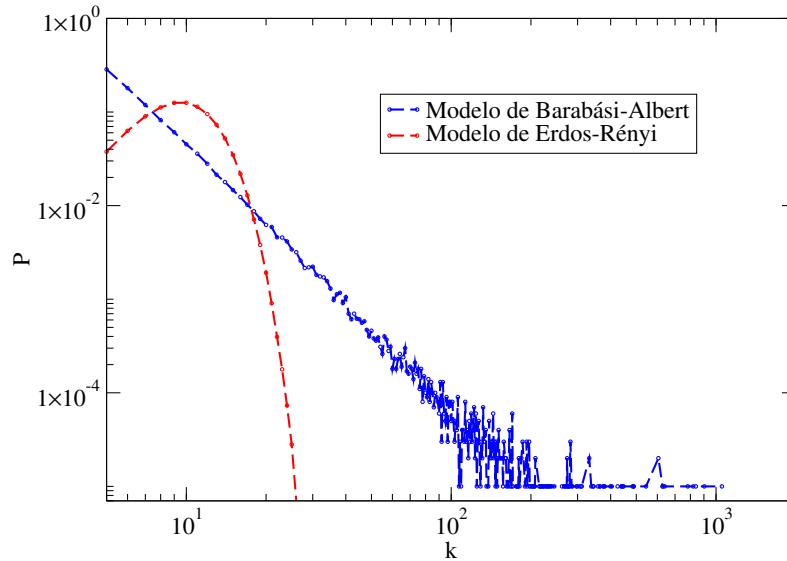


Figura 2.12 – Distribuições de grau de redes de Barabási-Albert e Erdős-Rényi. Ambas com grau médio $\langle k \rangle = 10$

2.4.6 Conflito entre falta de escala e correlação neutra

À priori, redes aleatórias não possuem correlação de graus quando múltiplas conexões são permitidas [2]. No entanto, ao evitar-se múltiplas conexões em simulações de redes, não é isto o que se observa. Acontece que é muito comum obter-se um padrão de correlação *disassociativo* e veremos o porquê.

Para tanto, calculemos o número esperado de ligações entre um sítio de grau k e grau k' , relaxando-se a condição de que não haja múltiplas conexões e auto-conexões na rede, a fim de se estimar como o valor esperado de ligações entre os sítios emerge. A probabilidade conjunta $P(k, k')$ de que, sendo tomado um link da rede ao acaso, em uma ponta haja um sítio de grau k e na outra um sítio de grau k' é dada por

$$P(k, k') = \frac{E_{kk'}}{\langle k \rangle N}. \quad (2.39)$$

Dada esta expressão, podemos estimar o número de ligações entre sítios de graus diferentes

$$E_{kk'} = \langle k \rangle N P(k, k'). \quad (2.40)$$

No caso de uma distribuição independente, ou seja, sem correlação, esta deve satisfazer

$$P(k, k') = \frac{k p(k) k' p(k')}{\langle k \rangle^2}.$$

Com isto, a equação (2.40) torna-se

$$E_{kk'} = N \frac{kp(k)k'p(k')}{\langle k \rangle}. \quad (2.41)$$

Vemos, portanto, que o número de ligações entre sítios de diferentes graus é proporcional ao seu grau em redes neutras. Isto pode levar a conflitos quando se deseja criar redes descorrelacionadas e que sejam simples, isto é, que não possuam múltiplas ou auto-conexões. Portanto, deve haver um valor máximo de grau, abaixo do qual o número de ligações entre quaisquer sítios seja menor ou igual a 1. Lembrando que a probabilidade de se encontrar o sítio mais conectado da rede deve ser $p(k_{max}) = \frac{1}{N}$ e substituindo-se este valor na equação (2.41), temos

$$\frac{k^2}{N\langle k \rangle} \leq 1, \quad (2.42)$$

levando-nos a

$$k_C \leq (\langle k \rangle N)^{\frac{1}{2}}. \quad (2.43)$$

Abaixo deste valor, podemos esperar que a rede seja neutra, sendo conhecido na literatura como *corte estrutural da rede*. Agora podemos entender o mecanismo que introduz correlação em redes sem escala, feitas via simulação. Ao se conectar pontas ao acaso, é esperado que sítios de grau elevado se conectem, devido ao seu alto número de pontas disponíveis. Quando se trata da primeira ligação, não há quaisquer conflitos, no entanto, ao evitar-se múltiplas conexões, estas ligações serão rejeitadas e, na busca de novas conexões, sítios de grau pequeno serão escolhidos, devido à sua abundância na rede, gerando assim correlação desassociativa. Porém, este conflito entre neutralidade e falta de escala fica resolvido ao se introduzir o corte estrutural k_C dado pela equação (2.43).

2.5 Decomposição k-core

Uma ferramenta importante na investigação da estrutura da rede, baseada no posicionamento dos sítios relativamente a sua organização, é a decomposição *k-core*, permitindo decompor a rede em *cascas*, indo dos sítios mais externos aos sítios mais centrais.

A decomposição *k-core* ou *k-shell* é realizada como segue: retira-se primeiramente os sítios de grau mais baixo k_{min} da rede; estes recebendo índice $k_S = k_{min}$. Após a remoção, seus vizinhos tem então seu grau atualizado. Seguida desta primeira poda, é possível que ainda haja sítios com grau $k \leq k_S$, resultantes da modificação do grau dos vizinhos dos nós removidos. Neste caso, procede-se removendo estes os sítios com grau $k \leq k_S$ restantes e, por sua vez, atualiza-se o grau de seus vizinhos. O procedimento deve ser realizado até que não haja mais quaisquer sítios de grau $k \leq k_S$.

Os sítios removidos recebem o índice k_S atual [21]. Agora, aumenta-se em uma unidade o índice k_S ($k_S = k_S + 1$) e remove-se os sítios de grau $k \leq k_S$ caso eles existam na rede. Caso não existam, aumenta-se em mais uma unidade o índice k_S . A poda de sítios é realizada de maneira idêntica à realizada anteriormente, até que não haja mais sítios de grau $k \leq k_S$. Os sítios removidos nesta etapa recebem índice k_S . Procede-se desta maneira até que todos os sítios da rede sejam removidos. Os sítios possuidores do índice k_S máximo pertencem ao k -core, que é comumente referido como o *núcleo denso* [22]. A figura 2.13 exemplifica a decomposição k -core.

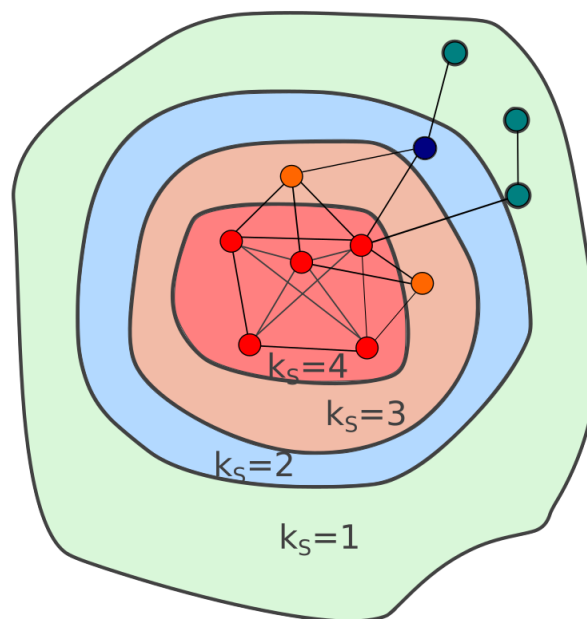


Figura 2.13 – Ilustração da decomposição k -core para rede com $N = 11$ sítios.

Como veremos em capítulos subsequentes, o núcleo mais denso da rede pode vir a atuar como importante agente em mecanismos de ativação da fase endêmica do modelo SIS.

III Redes sintéticas e redes reais

Uma rede pode ser considerada uma abstração de um sistema complexo, como apresentado no Capítulo II. Ocorre que, geralmente, uma rede é uma entidade dinâmica em si, ou seja, esta evolui ao longo do tempo, mudando sua estrutura. Portanto, sua natureza mutável pode impor dificuldades na elaboração de algoritmos que simulem processos dinâmicos.

No entanto, verifica-se em muitos casos que a escala de tempo τ_{din} de processos dinâmicos é muito menor que a escala de evolução τ_{rede} da rede. A consequência disto é que, para fins práticos, o estado da rede pode ser considerado *congelado*. A fim de se descrever a dinâmica de uma rede, torna-se, pois, conveniente utilizar uma *coleção* ou *ensemble* de redes como substrato - todas possuindo estatisticamente as mesmas características métricas, como distribuição e correlação de graus, para citar exemplos mais típicos.

Obter um número satisfatório de estados de uma rede real ou mesmo extrapolar seu tamanho para o limite termodinâmico, $N \rightarrow \infty$, possui, portanto, dificuldades atreladas ao seu tempo de evolução. No intuito de se contornar esta dificuldade, costuma-se gerar uma *coleção* de *redes sintéticas* possuindo características métricas estatísticas de nosso interesse.

3.1 O modelo de configurações e o modelo de configurações des-correlacionado

Na simulação de processos dinâmicos em redes complexas é muitas vezes conveniente fixar uma distribuição de graus de interesse a fim de se obter uma coleção de redes com características métricas semelhantes. Para alcançar este objetivo, atribui-se a cada sítio um grau proveniente da distribuição de probabilidades desejada. A soma dos graus deve ser um número par, de forma que não sobrem pontas soltas. Além do mais, devemos lembrar que, para que a rede gerada tenha uma *componente gigante*, a condição dada pela equação (2.27) deve ser respeitada. Na figura (3.1), exemplificamos a criação de uma coleção de três redes geradas usando-se a mesma distribuição de graus.

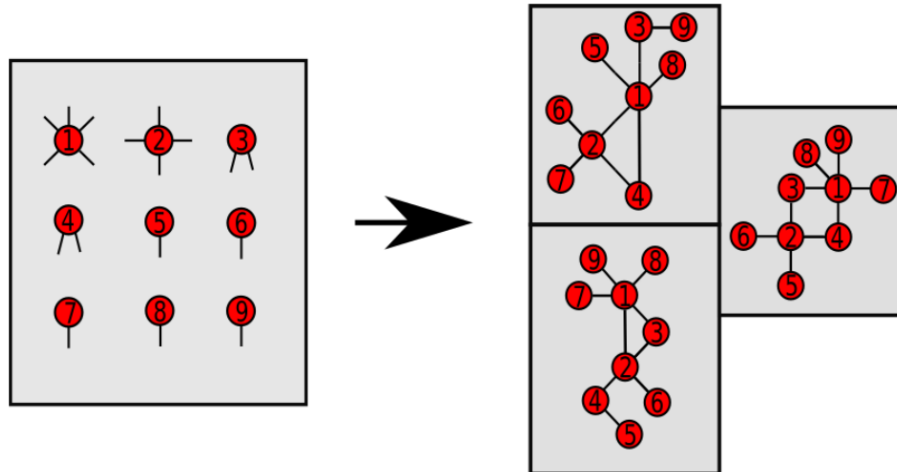


Figura 3.1 – Distribuição de graus fixa e três exemplos de realizações de redes com a mesma distribuição

De uma forma prática, lançando-se mão da distribuição desejada, cria-se uma lista possuindo como tamanho a soma dos graus. Após sua criação, ela deve ser preenchida com os índices de cada sítio i , repetidos na lista k_i vezes, como ilustra a figura (3.2).

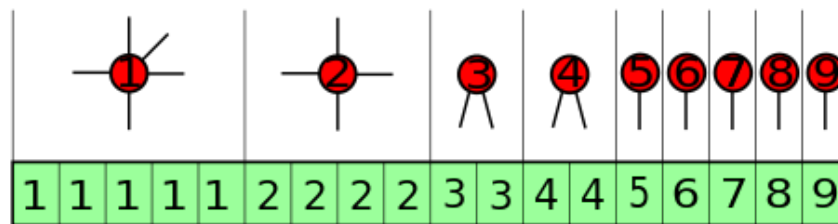


Figura 3.2 – Distribuição de graus fixa e lista de pontas.

Com a intenção de se conectar dois sítios, um par de pontas é tomado ao acaso por meio de um gerador de números pseudo-aleatórios. Caso a escolha não se trate de múltipla conexão, tampouco de auto-conexão, a escolha é aceita e os sítios são então conectados; do contrário, a ligação é rejeitada e a escolha de um novo par é feita. Caso a ligação seja feita com sucesso, desloca-se os últimos valores da lista para o local das pontas anteriormente aceitas e encurta-se a lista em duas unidades, a fim de que ela contenha apenas pontas que não foram conectadas. Eis a nossa implementação do *Modelo de Configurações*. Para que a rede gerada tenha correlação de graus neutra, como a do exemplo ilustrado na figura (2.6), deve-se impor que o grau máximo da distribuição gerada seja menor ou igual ao corte estrutural da rede [23], resultado obtido na equação (2.43). Este é o *Modelo de Configurações Descorrelacionado* (UCM - *Uncorrelated Configuration Model*), amplamente utilizado na criação de redes sem correlação de grau, condição necessária para a discussão de certos modelos analíticos.

3.2 Redes com nós periféricos

Nesta seção descrevemos os modelos utilizados para a geração de redes sintéticas, formadas por núcleo com distribuição de grau em lei de potência $P(k) \propto k^{-\gamma}$ e a adição de nós periféricos de forma controlada. Consideramos um modelo com a inclusão somente de vértices de grau $k = 1$, referidos também como *folhas*, em analogia ao grafo estrela, e a adição de sítios de grau $k = 2$, referidos neste trabalho como *tubos*. Os últimos podem desempenhar papéis diferentes na propagação de uma epidemia. Um deles é reduzir a distância entre hubs, formando atalhos e caminhos redundantes, ao gerarem rotas alternativas entre os principais propagadores da infecção.

Ainda nesta seção, discutimos brevemente algumas propriedades métricas em redes reais. Nelas, ocorre ser mais conveniente controlar o número de sítios periféricos não pela medida baseada diretamente no grau do sítio, mas em seu índice *k-shell*. A escolha desta medida de centralidade foi feita motivada pelo fato de haver em redes reais estruturas tubulares, e muitas vezes ramificadas, que não fazem conexão entre sítios do núcleo. Com efeito, baseado na decomposição *k-core*, estas estruturas pertencem também à camada $k_S = 1$. Estruturalmente elas se assemelham mais a folhas do que a tubos. As metodologias de criação de redes apresentada nas subseções 3.2.2 e 3.2.3 justificam ainda esta escolha pois sítios de grau $k = 1$ e 2 pertencem às camadas $k_S = 1$ e 2 , respectivamente, neste contexto. Portanto, o índice k_S foi a escolha mais natural de centralidade. As estruturas mencionadas são exibidas na figura 3.3. Nela, a rede *Air Traffic* apresenta algumas dessas estruturas destacadas em azul. No processo dinâmico devemos, quando conveniente, *imunizar* os sítios periféricos, tornando-os indiferentes à dinâmica.

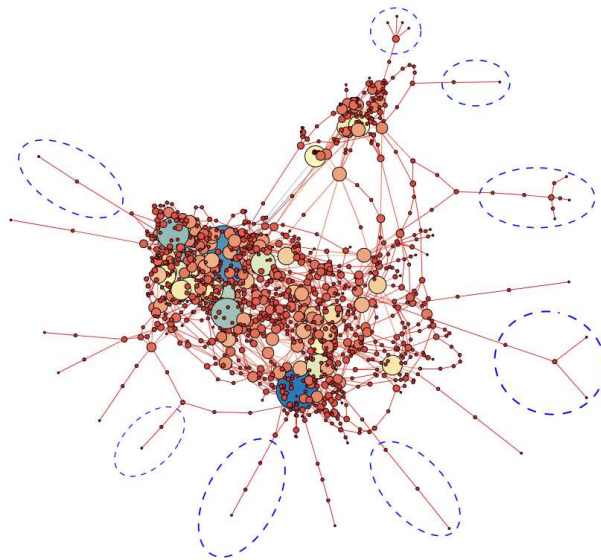


Figura 3.3 – Rede *Air Traffic* com estruturas da camada $k_S = 1$ ressaltadas em azul.

3.2.1 Redes reais

A figura 3.4 mostra algumas propriedades da rede *AS Caida*. Esta é rica em ambos os tipos de sítios periféricos (veja tabela 3.1). A distribuição de grau da rede original é comparada antes e depois da remoção das 1- e 2-*shell* na figura 3.4(a). Nela podemos ver redução progressiva do grau máximo k_{\max} com a remoção das shells. Definimos a seguir a medida

$$P_{NN}(k|k_S) = \frac{E_{kk_S}}{E_{k_S}}, \quad (3.1)$$

em que E_{kk_S} é a quantidade de arestas que ligam a camada k_S a sítios de grau k fora da camada e E_{k_S} é o número de arestas que saem da camada k_S . Ela mede a distribuição de grau dos vizinhos da camada k_S . A figura 3.4(c) demonstra predominante correlação de graus desassociativa. Qualitativamente, este caráter não muda com a remoção dos vértices periféricos. Enquanto isso, distribuições de graus dos vizinhos mais próximos dos sítios componentes da 1- e da 2-*shell*, exibidas na figura 3.4(b), demonstram a afinidade dessas camadas com sítios de grau elevado.

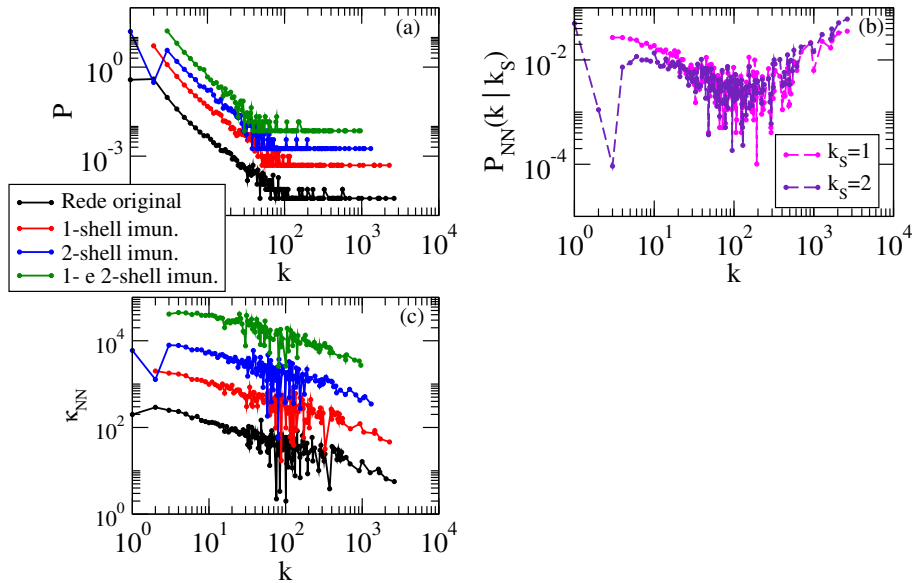


Figura 3.4 – Propriedades estruturais da rede AS Caida. Métricas da rede original antes e após imunização das 1-*shell* e da 2-*shell* (independente e concomitantemente). (a) Distribuição de grau $P(k)$; (b) distribuição de grau dos vizinhos mais próximos $P_{NN}(k|k_S)$ da 1- e da 2-*shell*; (c) grau médio de vizinhos mais próximos κ_{NN} .

Para motivar esse estudo, consideramos algumas redes reais frequentemente usadas para análise de processos dinâmicos [24, 25, 26]. Analisamos a quantidade de sítios periféricos ($k_S = 1$ e 2) nestas redes e os resultados são resumidos na tabela 3.1. Podemos ver que as frações destes vértices podem ser significativamente grandes (até 55%), justificando a importância do seu estudo.

Em particular, podemos notar a importância dos vértices da camada $k_S = 1$ na composição da vizinhança de sítios de grau elevado da rede original através do valor de k_{\max} antes e depois da remoção destes sítios periféricos, como mostrado na tabela 3.1. Além disso, a remoção dos vértices da camada $k_S = 2$ pode reduzir drasticamente o grau máximo k_{\max} da rede, mostrando que eles também estão ligados a hubs e podem ser atores centrais na propagação da epidemia.

Tabela 3.1 – Análise de vértices periféricos em um conjunto de redes reais. Mostramos o tamanho (N) da rede, a fração de vértices de índice $k_S = 1$, $P(k_S = 1)$, e $k_S = 2$, $P(k_S = 2)$, respectivamente, o grau mais elevado da rede original (k_{\max}) e depois da remoção dos vértices de índice $k_S = 1$ (k_{\max}^*), $k_S = 2$ (k_{\max}^{**}) e remoção de ambos (k_{\max}^{***}).

Rede	N	$P(k_S = 1)$	$P(k_S = 2)$	k_{\max}	k_{\max}^*	k_{\max}^{**}	k_{\max}^{***}
Air Traffic	1230	0.120	0.555	34	33	28	28
AS Caida	26500	0.384	0.430	2628	2277	1311	960
AS Oregon	6500	0.379	0.420	1458	1181	785	508
Astrophysics	14800	0.067	0.1108	360	360	360	360
Astrophysics(1993)	17900	0.055	0.089	504	504	504	504
Cite Seer	365100	0.149	0.147	1739	1687	1643	1591
CondMat(1993)	21400	0.082	0.147	279	279	275	275
CondMat(1995)	13900	0.101	0.186	107	107	107	107
Cora	23200	0.148	0.164	377	343	331	297
DBLP(citation)	12500	0.389	0.140	709	675	673	639
Digg	29700	0.393	0.188	283	268	265	250
English	7400	0.131	0.328	2568	2457	2060	2050
Facebook(links)	63400	0.133	0.090	1098	1095	1092	1089
Facebook(wall)	44000	0.242	0.175	223	218	219	214

3.2.2 Criação de redes com folhas

Sítios de grau 1 (ou folhas), apesar de sua baixa conectividade, são capazes de alimentar atividade infecciosa em torno dos sítios mais conectados [10]. Portanto, sua presença é capaz de aumentar o tempo de vida da infecção na vizinhança do hub, favorecendo assim o estabelecimento da epidemia. O algoritmo para criação de redes com acréscimo controlado de folhas na rede é explicado a seguir

1. Uma rede de tamanho N_c (núcleo da rede) é gerada usando o algoritmo UCM, descrito na seção 3.1. O expoente γ da rede é escolhido dentre os valores 2.5, 2.7 e 3.5;
2. definimos a quantidade N_1 de sítios de grau 1 que será adicionada ao núcleo da rede. Neste trabalho, a quantidade usada foi $N_1 = 0.5 \times N_c$;
3. os N_1 sítios são ligados ao núcleo da rede, com um sítio i escolhido ao acaso e aceito com probabilidade $\Pi(k_i) = \left(\frac{k_i}{k_{max}}\right)^\alpha$. Consideramos valores de $\alpha = 1$ (conexão preferencial linear) e $\alpha = 2$ (ligação preferencial quadrática).

O algoritmo utilizado na criação desta rede pode ser encontrado em <https://github.com/JCMSchott/complexTube>.

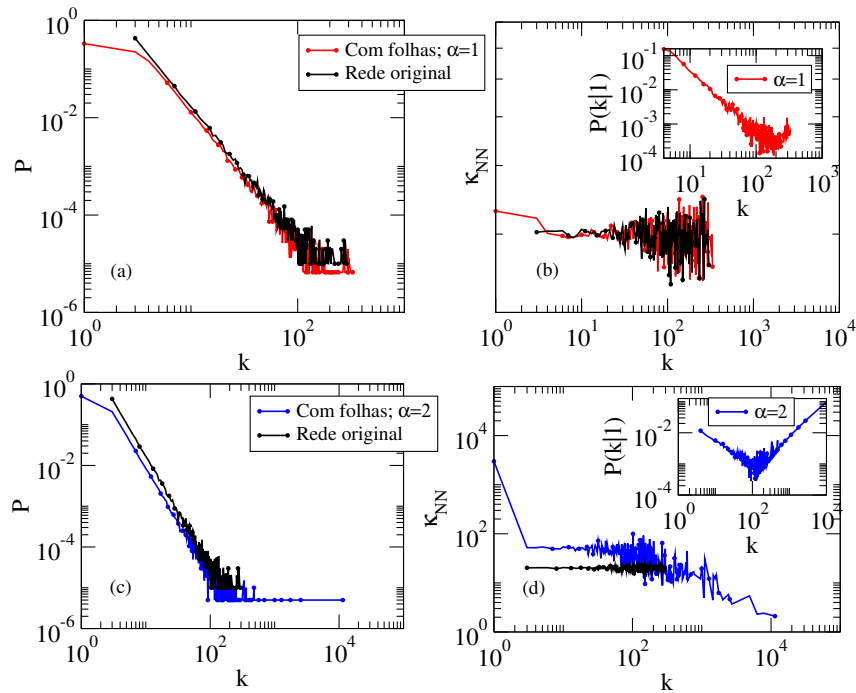


Figura 3.5 – Redes UCM com folhas usando $\gamma = 2.7$ e $N_c = 10^5$ sítios ligados a $N_1 = 0.5 \times N_c$ folhas. (a): Distribuição de graus: rede original e rede com adição de folhas via $\alpha = 1$. (b): Grau médio de primeiros vizinhos κ_{NN} das referidas redes. (c): Distribuição de graus: rede original e rede com adição de folhas, $\alpha = 2$. (d): Grau médio de primeiros κ_{NN} vizinhos das referidas redes. As inserções nas figuras (b) e (d) mostram distribuição de graus da vizinhança dos sítios de grau 1.

As medidas de κ_{NN} exibidas na figura 3.5(b) mostram correlação de grau neutra, como apresentados na figura (2.6). Por outro lado, a mesma medida, desta vez com $\alpha = 2$, representada em azul na figura 3.5(d), apresenta correlação de graus desassociativa, como tipificado na figura (2.4). Via ligação preferencial linear nenhuma mudança significativa foi gerada nos hubs, como mostrado nas figuras 3.5(a) e (b), onde mostramos a distribuição de grau $P(k)$ e o grau médio de vizinhos mais próximos κ_{NN} .

Exceto pela inclusão do grau $k = 1$, ambas as distribuições são muito parecidas, mantendo a ausência de correlação de graus. Na inserção da figura 3.5(b), mostramos a distribuição de graus dos vizinhos das folhas. Nós observamos um padrão muito similar ao da distribuição global da rede, confirmando que os efeitos são muito pequenos. Para $\gamma = 2.3$, os efeitos são ainda menores e tornam-se observáveis para $\gamma = 3.5$, embora ainda sejam bastante pequenos.

Os dados exibidos nas figuras 3.5(c) e (d) mostram os efeitos da ligação quadrática de folhas ao núcleo. Um aumento sensível no grau dos sítios mais conectados da rede pode ser observado. Além disso, a introdução de correlação na rede também pode ser visto, manifestando caráter desassociativo para graus elevados. O mesmo efeito pode ser visto na distribuição de graus dos vizinhos das folhas $P(k|1)$, mostrado na inserção da figura 3.5(d).

3.2.3 Criação de redes com tubos

Dentre os efeitos da adição de sítios de grau $k = 2$ estão o aumento do tempo de atividade na vizinhança dos hubs e a adição de atalhos e caminhos redundantes entre hubs e sítios do núcleo k -core. Abaixo é apresentado o algoritmo utilizado para a criação de redes com tubos neste trabalho

1. Criamos uma rede de tamanho N_c . Sua distribuição de graus é definida como $P(k) \propto k^{-\gamma}$, com $3 \leq k \leq N_c^{\frac{1}{2}}$. Rede gerada via algoritmo UCM, descrito na seção 3.1;
2. expoente γ com os valores 2.5, 2.7 e 3.5 são considerados;
3. define-se a quantidade N_2 de sítios de grau 2 que será adicionada ao núcleo. Neste trabalho, a quantidade usada foi $N_2 = 0.25 \times N_c$;
4. sítios de grau 2 são ligados ao núcleo da rede original apenas, com um sítio i escolhido ao acaso e aceito com probabilidade $\Pi(k_i) = \left(\frac{k_i}{k_{max}}\right)^\alpha$. Neste trabalho $\alpha = 1$ (conexão preferencial linear) e $\alpha = 2$ (ligação preferencial quadrática).

O algoritmo utilizado na criação desta rede pode ser encontrado em <https://github.com/JCMSchott/complexTube>.

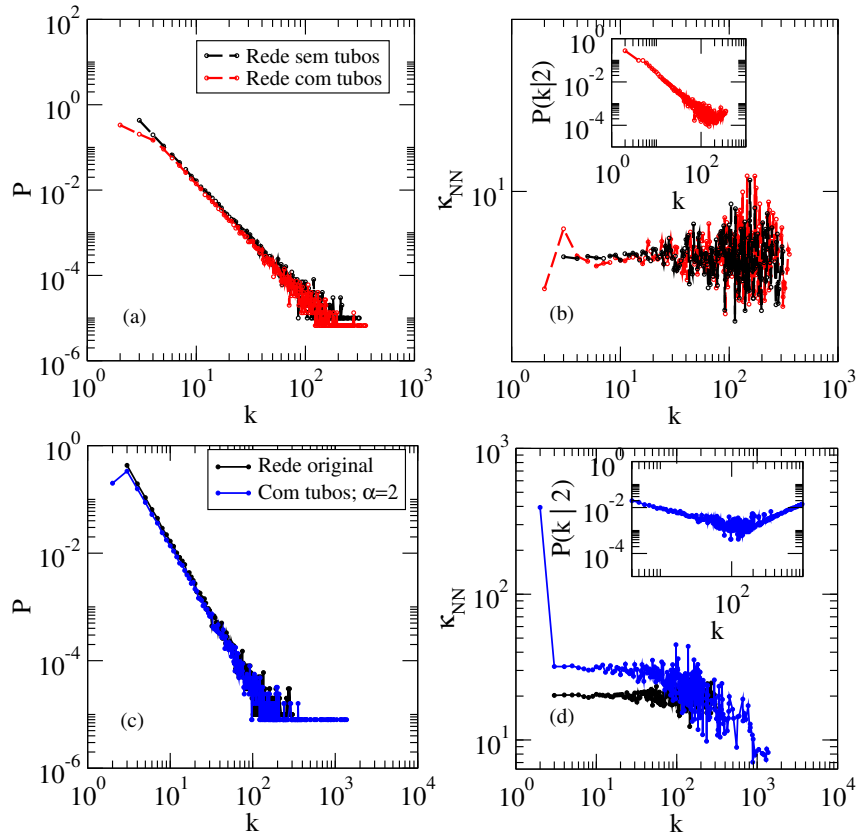


Figura 3.6 – Redes UCM usando $\gamma = 2.7$ e $N_c = 10^5$ sítios ligados a $N_2 = 0.5 \times N_c$ tubos. (a): Distribuição de graus: rede original e rede com adição de tubos ($\alpha = 1$). (b): κ_{NN} das referidas redes. (c): Distribuição de graus para rede original e rede com tubos ($\alpha = 2$). (d): κ_{NN} para as redes associadas à figura (c). As inserções nas figuras (b) e (d) mostram distribuição de grau dos vizinhos dos sítios de grau 2.

IV Processos epidêmicos em redes complexas

A infraestrutura formada pelas mais variadas redes de transportes e informação que representam nosso mundo globalizado é certamente a principal responsável pela sensação de encurtamento de distâncias que se tem atualmente. Redes de aeroportos espalhados por todo o globo, além das malhas de transporte ferroviário e rodoviário, permitem diariamente um fluxo elevado de seres-humanos todos os dias. A facilidade hoje obtida no transporte humano permite também o espalhamento de agentes infecciosos em uma velocidade sem precedentes. O surto do vírus ebola que ameaçou a humanidade em 2014 foi acompanhado em tempo real pela comunidade científica [27] e é um bom exemplo dos efeitos colaterais da mesma infraestrutura que traz a sensação de encurtamento de distâncias. Poder de previsão, planejamento de estratégias de contenção agentes patogênicos tornam-se então imprescindíveis [28] num mundo em que distâncias já não significam tanto para possíveis ameaças à saúde e vida humanas. Redes de relacionamento humano representam uma rede de contato direto para a propagação de doenças infectocontagiosas, como o vírus *influenza* e o vírus HIV, trazendo à tona a necessidade de estratégias otimizadas de imunização [29]. Em paralelo às redes de relacionamento, redes de computadores e smartphones desempenham um papel semelhante para programas maliciosos, enquanto desinformação e crenças se propagam, de modo semelhante a agentes patogênicos, na rede de relações humanas da qual fazemos parte e em redes sociais internet afora [30], mudando inclusive o modo como agentes políticos tem feito suas campanhas eleitorais [31].

4.1 Modelos epidêmicos

Na literatura há uma variedade de modelos epidêmicos extensamente estudados sobre redes complexas. Uma abordagem bastante comum no estudo desses modelos é a compartimentação de sítios de acordo com seu estado de saúde.

O estudo do **processo de contato** pode ser bastante útil na modelagem de doenças sem imunidade de longo prazo transmitidas em *redes de contato* [32]. Neste processo dinâmico, cada sítio pode estar nos estados *suscetível*, representado por $\sigma = 0$, e *infectado*, representado por $\sigma = 1$. Um sítio no estado suscetível não é portador da infecção, mas é capaz de contraí-la, enquanto um sítio infectado de grau k pode contagiar cada um de seus vizinhos suscetíveis com taxa $\lambda_i = \frac{\lambda}{k_i}$. Sítios infectados são curados espontaneamente a uma taxa μ . Teorias de campo médio de primeira ordem, em redes não correlacionadas, apontam neste modelo para o mesmo limiar epidêmico $\lambda_C = 1$, separando a fase inativa (sem sítios infectados após um intervalo de tempo transiente) da fase ativa (com uma quantidade finita de sítios infectados após um intervalo de tempo transiente).

Em estudos epidemiológicos, o modelo **SIS - suscetível-infectado-suscetível**, assim como o processo de contato, está associado processos epidêmicos envolvendo doenças sem imunidade de longo prazo. No entanto, no modelo SIS, um sítio *suscetível* é infectado por meio do contato com sítios *infectados* em sua vizinhança a uma taxa $n\lambda$, onde λ é a taxa de infecção da epidemia e n é o número de primeiros vizinhos infectados. Por sua vez, um sítio infectado é curado a uma taxa μ . Em *redes regulares*, o modelo SIS e o processo de contato são equivalentes, sendo a relação entre os limiares dos dois modelos $\lambda_{SIS} = \frac{\lambda_{CP}}{k}$, onde k é o grau dos sítios da rede regular. No entanto, os dois modelos entram em discordância em redes complexas. Neste trabalho nos ocuparemos exclusivamente do modelo SIS.

Em contraposição ao modelo SIS e o processo de contato, o **Modelo SIRS (suscetível-infectado-recuperado-suscetível)** está associado a processos epidêmicos envolvendo doenças com imunidade de longo prazo. Neste modelo, um sítio tem agora um novo estado que pode ser acessado - o estado *recuperado*, representado por $\sigma = 2$. Este modelo se assemelha ao modelo SIS, com a diferença de que um sítio infectado se torna recuperado com taxa μ , não podendo infectar e tampouco ser infectado. Com uma taxa α , um sítio recuperado torna-se suscetível, podendo contrair novamente a infecção. Do ponto de vista de teorias de campo médio padrão, há resultados presentes na literatura que apontam para a equivalência entre os modelos SIRS e SIS [11].

4.1.1 A equação mestra

Nesta dissertação trataremos de processos Markovianos em que as taxas de transição não dependem da história que levou ao estado específico. Estes processos são descritos pela equação mestra sobre a qual discorreremos a seguir. Em um processo dinâmico em que os sítios da rede tem acesso a um número finito de estados $\sigma = 0, 1, 2, \dots$, define-se a probabilidade de o sistema se encontrar em um estado microscópico $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ em um instante de tempo t como $P(\sigma, t)$. Sujeito às taxas de transição $W_{\sigma \rightarrow \sigma'}$, do estado σ para o estado σ' , estamos em posição de definir a *equação mestra* que rege a dinâmica do processo

$$\frac{dP(\sigma, t)}{dt} = \sum_{\sigma'} [P(\sigma', t)W_{\sigma' \rightarrow \sigma} - P(\sigma, t)W_{\sigma \rightarrow \sigma'}]. \quad (4.1)$$

O primeiro termo do somatório representa as taxas de transição do sistema para o estado σ , enquanto o segundo termo representa as taxas de transição para fora do estado σ . A probabilidade $P(\sigma, t)$ deve estar normalizada em qualquer instante de tempo

$$\sum_{\sigma} P(\sigma, t) = 1. \quad (4.2)$$

Além disso, a média de qualquer variável dinâmica pode ser calculada para um instante

$$\langle B(t) \rangle = \sum_{\sigma} B(\sigma) P(\sigma, t). \quad (4.3)$$

Um olhar para a equação mestra (4.1) nos mostra que uma condição suficiente para a existência de um estado estacionário é a condição de que

$$P(\sigma', t) W_{\sigma' \rightarrow \sigma} = P(\sigma, t) W_{\sigma \rightarrow \sigma'}. \quad (4.4)$$

Conhecida como *balanço detalhado*, esta é uma condição necessária para a existência de *equilíbrio microscópico* ou *reversibilidade microscópica* em um processo dinâmico ou, em outras palavras, expressando que os fluxos de probabilidade entre quaisquer dois estados microscópicos se contrabalanceiem. Há ainda a possibilidade de um processo não possuir equilíbrio microscópico, não sendo satisfeita a condição dada pela equação (4.4) em cujo caso, podendo ainda, no entanto, haver equilíbrio global - condição que garante a existência de estado estacionário.

A solução exata da equação mestra pode trazer o entendimento de muitos aspectos envolvidos na dinâmica [16], sendo raramente possível, no entanto, devido ao enorme número de variáveis estocásticas envolvidas em processos associados a sistemas muito grandes. Métodos de aproximação conhecidos como *teorias de campo médio* desempenham um papel importante na tarefa de buscar uma solução aproximada, tomando como objeto de observação os valores médios do estado da dinâmica, ao invés de seu estado microscópico. Além de métodos de aproximação, métodos computacionais de simulação representam um poderoso arsenal para se testar os resultados obtidos com teorias de campo médio e mesmo propôr novas teorias [33].

4.1.2 O modelo SIS

Algumas doenças, como a gripe comum e o vírus influenza, não apresentam imunização de longo prazo após o contágio, devido à mutação constante dos agentes virais, tornando-se o indivíduo suscetível à doença novamente após a cura. Doenças como esta podem ser melhor entendidas teoricamente sob o ponto de vista do modelo SIS. Neste modelo, um sítio pode encontrar-se nos estados $\sigma = 0$ e $\sigma = 1$, associados ao estado *suscetível* e *infectado*, respectivamente, como explicado brevemente no início desta seção. Um sítio infectado transmite a infecção a uma taxa λ para cada um de seus vizinhos e se cura espontaneamente a uma taxa μ . Dadas as regras dinâmicas do modelo SIS, é facilmente feita a constatação de que, para $\sigma = \{0, 0, \dots, 0\}$, não há mais qualquer dinâmica, ficando o processo preso para sempre nesta configuração, conhecida na literatura como *estado absorvente*.

4.1.3 Teoria HMF

Ao invés de olhar para a dinâmica do ponto de vista microscópico, uma abordagem mais simples seria encarar sítios de mesmo grau como estatisticamente equivalentes, além da divisão dos sítios em compartimentos de acordo com seu estado. Seja S_k o número de sítios de grau k suscetíveis e I_k o número de sítios de grau k infectados da rede num instante de tempo t . Denotaremos a densidade de sítios de grau k infectados por $\rho_k = \frac{I_k}{N}$, onde N é o número de sítios da rede.

O número de arestas que saem dos S_k sítios suscetíveis do compartimento k é kS_k , dos quais, podemos esperar que, dentre seus vizinhos, uma fração $\rho_{k'}$ seja de sítios infectados de grau k' , conectados a sítios de grau k com probabilidade $P(k'|k)$, que transmitem a infecção aos S_k sítios a uma taxa λ , tornando-os infectados à taxa $\lambda k S_k \rho_{k'} P(k'|k)$. Esta é a contribuição proveniente apenas dos sítios de grau k' para a ocorrência da infecção dos sítios de grau k , a taxa total sendo $\sum_{k'} \lambda k S_k \rho_{k'} P(k'|k)$. Esta taxa total de infecção funciona como uma taxa de entrada de sítios de grau k suscetíveis para o compartimento dos sítios infectados. No entanto, sítios infectados se curam espontaneamente a uma taxa μ , sendo a taxa total de cura de sítios de grau k infectados igual a μI_k , funcionando como uma taxa de saída do compartimento dos infectados. Vamos fixar a taxa de cura $\mu = 1$, sem perda de generalidade, pois isso muda apenas a unidade de tempo em que estudamos o problema. Com estas taxas de entrada e saída, podemos estabelecer agora a equação dinâmica da teoria HMF

$$\frac{dI_k}{dt} = -I_k + \sum_{k'} \lambda k S_k \rho_{k'} P(k'|k). \quad (4.5)$$

Dividindo-se a equação (4.5) por N , o termo S_k dá lugar ao termo $s_k = \frac{S_k}{N}$ que, junto com ρ_k satisfazem a equação $\rho_k + s_k = 1$, o que nos permite representar $s_k = 1 - \rho_k$. Com isto, a equação dinâmica toma sua forma final

$$\frac{d\rho_k}{dt} = -\rho_k + \lambda k (1 - \rho_k) \sum_{k'} \rho_{k'} P(k'|k). \quad (4.6)$$

Um primeiro olhar para a equação (4.6) já nos sugere que a correlação de graus desempenhe um papel importante na dinâmica. Visto que uma solução para um caso geral não é possível sem conhecermos a forma explícita de $P(k'|k)$, vamos olhar um caso simples, no qual $P(k'|k)$ não dependa de k , o que corresponde a redes de correlação neutra, também chamadas de não correlacionadas. Neste caso, $P(k'|k) = \frac{k' P(k')}{\langle k \rangle}$. Vamos proceder fazendo uma análise de estabilidade linear em torno do ponto fixo $\rho_k = 0$ e fazer uso da substituição $\rho_k = \rho_{k'} \delta_{kk'}$, para que a equação (4.6) tome a forma

$$\frac{d\rho_k}{dt} = \sum_{k'} B_{kk'} \rho_{k'}, \quad (4.7)$$

onde $B_{kk'} = -\delta_{kk'} + \lambda \frac{kk'P(k')}{\langle k \rangle}$ é a matriz jacobiana.

As soluções da equação (4.7) são dadas por

$$\vec{\rho}(t) = \sum_i e^{\beta_i t} a^{(i)} \vec{b}_i, \quad (4.8)$$

onde \vec{b}_i e β_i são o i -ésimo autovetor e respectivo autovalor da matriz $B_{kk'}$.

Visto que os autovalores de $B_{kk'}$ em geral são complexos, as exponenciais podem ser representadas como

$$e^{\beta_i t} = e^{\Re(\beta_i)t} e^{i\Im(\beta_i)t}. \quad (4.9)$$

Enquanto a parte imaginária corresponde a um termo oscilatório, a parte puramente real corresponde à sua amplitude, acarretando em um termo que não tem fronteiras quando $t \rightarrow \infty$, para $\Re(\beta_i) > 0$ (divergindo do ponto fixo) e um termo cuja amplitude vai a 0 com $t \rightarrow \infty$ (convergindo para o ponto fixo), para $\Re(\beta_i) < 0$. As soluções começam, portanto, a divergir do ponto fixo $\rho_k = 0$, quando o maior valor da parte real entre todos os autovalores da Jacobiana torna-se nulo, ou seja, $-1 + \lambda\Lambda_m = 0$, onde Λ_m é maior autovalor da matriz $\frac{kk'P(k')}{\langle k \rangle}$, que é simétrica e não negativa. O teorema de Perron-Frobenius nos garante, portanto, a existência de um autovalor real, sendo este seu maior autovalor, associado ao único autovetor com todos os elementos positivos [16]. No caso da matriz $B_{kk'}$, o maior auto-valor corresponde a

$$\Lambda_m = \frac{\langle k^2 \rangle}{\langle k \rangle}, \quad (4.10)$$

levando o limiar previsto pela teoria a

$$\lambda_C = \frac{\langle k \rangle}{\langle k^2 \rangle}. \quad (4.11)$$

Este resultado, apesar de simples, prevê que, para redes com distribuição de graus em lei de potência com $\gamma \leq 3$, o limiar se anula no limite termodinâmico, visto que o segundo momento da distribuição diverge neste regime. Para $\gamma > 3$, o segundo momento $\langle k^2 \rangle$ não diverge no limite termodinâmico, sendo portanto previsto um limiar finito pela teoria HMF.

4.1.4 Teoria QMF

Alternativamente a considerar sítios de mesmo grau como estatisticamente equivalente, a teoria *QMF* (*Quenched Mean-Field*) aborda o problema de outro ponto de vista, que consiste na dinâmica da probabilidade de um sítio i encontrar-se infectado num instante de tempo t . No entanto, diferentemente da teoria HMF, a matriz de adjacências entra em jogo explicitamente, não apenas refletida na distribuição de graus da rede.

Assim como na teoria HMF, correlações dinâmicas não são levadas em conta nesta teoria de campo médio. De forma análoga à dedução da equação dinâmica da teoria HMF, a equação dinâmica da teoria QMF toma a forma

$$\frac{d\rho_i}{dt} = -\rho_i + \lambda(1 - \rho_i) \sum_j A_{ij} \rho_j. \quad (4.12)$$

Considerando-se $\rho_i \ll 1$ (a fim de se linearizar a equação) e utilizando-se os mesmos argumentos de análise de estabilidade linear da seção anterior, com $B_{ij} = -\delta_{ij} + \lambda A_{ij}$, temos que o limiar crítico corresponde a

$$\lambda_C = \frac{1}{\Lambda_m}. \quad (4.13)$$

onde Λ_m é o maior auto-valor da matriz A_{ij} , que é real, como garante o teorema de Perron-Frobenius.

O maior auto-valor da matriz A_{ij} para redes em lei de potência é fornecido na referência [34] e dado por

$$\Lambda_m \approx \begin{cases} \sqrt{k_{max}}, & \gamma > \frac{5}{2} \\ \frac{\langle k^2 \rangle}{\langle k \rangle}, & 2 < \gamma < \frac{5}{2}, \end{cases} \quad (4.14)$$

cujos resultado transforma a equação (4.13) em

$$\lambda_C \approx \begin{cases} \frac{1}{\sqrt{k_{max}}}, & \gamma > \frac{5}{2} \\ \frac{\langle k \rangle}{\langle k^2 \rangle}, & 2 < \gamma < \frac{5}{2}. \end{cases} \quad (4.15)$$

Este resultado fica em discordância com a teoria HMF para $\gamma > 3$. Um fato já conhecido é o de que o grau máximo da rede escala com o tamanho da rede, como visto no primeiro capítulo. A teoria QMF prevê, portanto, que o limiar vá para 0 no limite termodinâmico, independente do valor de γ .

4.2 O método quase-estacionário

O modelo SIS possui um *estado estacionário* ativo no limite termodinâmico. No entanto, sistemas com tamanho finito invariavelmente cairão no estado absorvente, ou seja, no estado em que todos os componentes estão no estado suscetível. Isto ocorre pois, no regime de tamanhos finitos, o único *estado estacionário* do sistema é o estado absorvente [12].

A fim de se contornar esta dificuldade em simulações, implementa-se o *método quase-estacionário*, no qual, ao alcançar-se o estado absorvente, a dinâmica é devolvida a alguma configuração visitada anteriormente.

Seja uma variável estocástica X_t , associada a um processo Markoviano, tomando valores de 0 a n num instante de tempo t com probabilidade $P_n(t)$. Definamos também a *probabilidade de sobrevivência* P_s , definida como a probabilidade de que a dinâmica não alcance o estado absorvente até o instante de tempo t , dada por

$$P_s(t) = \sum_{n \geq 1}^N P_n(t). \quad (4.16)$$

Por meio da probabilidade de sobrevivência, estamos agora em posição de definir a probabilidade quase-estacionária [35]

$$P_{QS} = \lim_{t \rightarrow \infty} \frac{P_n(t)}{P_s(t)}. \quad (4.17)$$

Podemos também caracterizar quaisquer momentos da distribuição quase-estacionária

$$\langle \rho^S \rangle = \frac{1}{N^S} \sum_{n=1}^N n^S P_{QS}(n). \quad (4.18)$$

O parâmetro de ordem é dado por

$$\langle \rho \rangle = \frac{1}{N} \sum_{n=1}^N n P_{QS}(n), \quad (4.19)$$

que é a densidade de sítios infectados. O limiar epidêmico pode ser estimado através da função resposta representada pela suscetibilidade dinâmica [36]

$$\chi = N \frac{\langle \rho^2 \rangle - \langle \rho \rangle^2}{\langle \rho \rangle}. \quad (4.20)$$

A partir da probabilidade quase-estacionária, pode-se obter o tempo de vida média de uma epidemia, definido como o intervalo de tempo médio em que a dinâmica visita o estado absorvente

$$\tau = \frac{1}{P_{QS}(1)}. \quad (4.21)$$

Para melhor entender o método quase-estacionário, vale olhar a equação mestra associada à dinâmica [37]

$$\frac{dP_n(t)}{dt} = \sum_m (W_{nm} P_m(t) - W_{mn} P_n(t)), \quad (4.22)$$

onde W_{nm} é a taxa de transição do estado $X_t = n$ para $X_t = m$.

Com a probabilidade quase-estacionária dada pela equação (4.17), a equação mestra deve ser modificada, obtendo a forma

$$\frac{dP_n}{dt} = \sum_{m > 0} [W_{nm} P_m - W_{mn} P_n + W_0 P_m P_n], \quad (4.23)$$

onde o termo adicional redistribui a probabilidade de a dinâmica cair no estado absorvente às demais configurações. Por não ser linear, a equação obtida não pode ser considerada uma equação mestra [38]. No entanto, ela nos sugere um método para obter a distribuição de probabilidades quase-estacionária por meios de simulação. Guardando-se estados da dinâmica X_t até o instante de tempo t , a dinâmica deve ser devolvida a algum desses estados após o estado absorvente ser visitado. De forma prática, armazena-se estados da dinâmica em uma lista contendo m vetores contendo os estados microscópicos da dinâmica

$$M = (\vec{\sigma}_1, \vec{\sigma}_2, \dots, \vec{\sigma}_j, \dots, \vec{\sigma}_m). \quad (4.24)$$

A fim de se descorrelacionar a dinâmica de sua própria história e garantir a manutenção da característica Markoviana da dinâmica, designa-se uma probabilidade δ e, para cada instante de tempo t , sorteia-se um número aleatório p . Caso $p \leq \delta$, um estado da lista M é sorteado e trocado pelo estado no instante de tempo atual.

4.3 Mecanismos de ativação no modelo SIS

Dois mecanismos importantes competem na ativação da epidemia. No primeiro, a atividade endêmica é fomentada no núcleo k -core da rede [39]. Como descrito na seção 2.5, o núcleo k -core forma uma estrutura densamente conectada, onde todos os vértices tem grau maior do que o índice k_S máximo [9, 39]. Neste mecanismo, sítios do núcleo k -core mantêm-se mutuamente ativos e, por sua vez, espalham a infecção para o restante da rede. Na figura 4.1, apresentamos uma ilustração da ativação epidêmica no núcleo da rede.

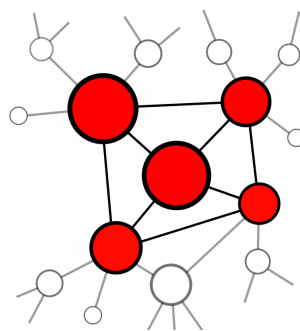


Figura 4.1 – Ativação da epidemia no núcleo k -core. Neste caso, o k -core corresponde a $k_S = 3$

O segundo mecanismo envolve a ativação local de hubs, vértices de elevada conectividade (que diverge no limite termodinâmico), e o espalhamento subsequente da atividade infecciosa a uma parte extensiva da rede. Vale ressaltar que este mecanismo de ativação ocorre entre os hubs a distâncias que, devido à propriedade de mundo pequeno [16], são suficientemente curtas para permitir a ativação mútua de hubs mesmo quando eles não estão diretamente conectados [10, 11].

Em outras palavras, um hub de grau k , uma vez inativo, pode ser reinfectado por seus vizinhos. Isto lhe confere a capacidade de manter a atividade infecciosa em torno de si por um longo tempo τ_k^{rec} . Ocasionalmente, o hub e sua vizinhança se tornam inativos após este intervalo. No entanto, outro hub de grau k' ainda ativo pode, à distância, enviar uma infecção para o hub inativo, reanimando a atividade neste dentro de um intervalo $\tau_{k',k}^{inf} < \tau_{k'}^{rec}$ [11]. A figura 4.2 exemplifica o mecanismo de reinfeção existente entre hubs. Este mecanismo de interação à distância torna-se preponderante no estabelecimento da epidemia quando a rede não possui um núcleo densamente conectado.

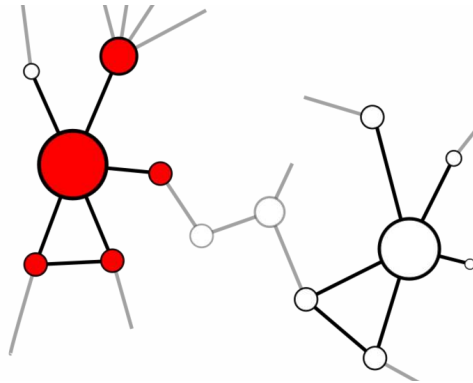


Figura 4.2 – Ilustração interação hub-hub. A cor vermelha representa sítios infectados, enquanto a cor branca representa sítios suscetíveis.

Em redes não correlacionadas com distribuição de graus em lei de potência, o k -core máximo governa a ativação para expoentes $\gamma < 2.5$. Por outro lado, a ativação via hubs domina o processo para $\gamma > 2.5$ [9]. Isto inclui mesmo o caso de redes em lei de potência não scale-free, quando $\gamma > 3$, onde a distribuição de graus possui uma variância finita [2, 40].

V Resultados

Neste capítulo apresentamos medidas baseadas no *método quase-estacionário padrão*, descrito na seção 4.2. Como substratos para a evolução da dinâmica, usamos redes reais e redes sintéticas, estas últimas geradas por meio do modelo UCM e a inserção de folhas (grau $k = 1$) ou tubos (grau $k = 2$), como descrito na seção 3.2.

5.1 Dinâmica SIS em redes reais

Dos resultados apresentados nesta seção, focamos em redes com alta densidade de sítios periféricos. Por meio das curvas de suscetibilidade χ e densidade de sítios infectados ρ , comparamos seus efeitos sobre a dinâmica.

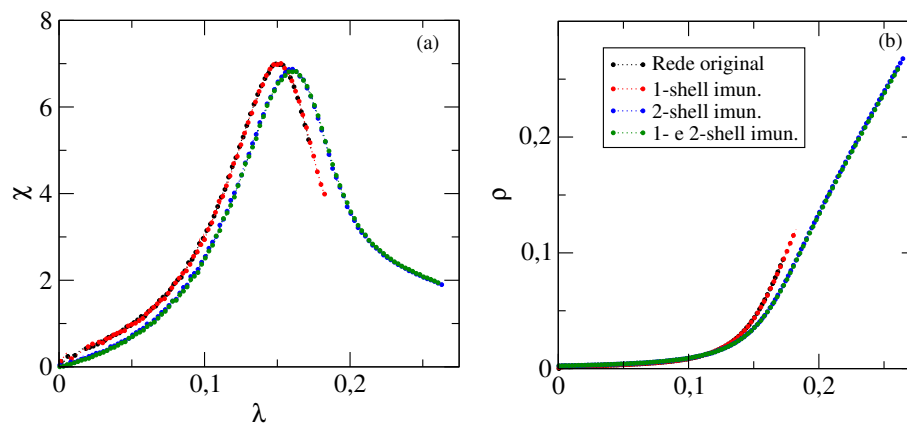


Figura 5.1 – Medidas quase-estacionárias na rede *Air traffic*. (a) Suscetibilidade χ e (b) densidade de sítios infectados ρ . Resultados comparam efeitos da imunização independente e conjunta de sítios com índice $k_S = 1$ e 2.

Imunizar a camada $k_S = 1$ na rede *AirTraffic* não gera efeitos apreciáveis no limiar, como mostra a curva de suscetibilidade χ , na Figura 5.1(a). Isto ocorre a despeito da proporção de folhas (principais componentes da camada $k_S = 1$) na rede ser de 7.9%. Além do mais, o tamanho da rede, $N = 1230$ sítios, não favorece o aparecimento de hubs com tempo de vida maior que o tempo de vida do núcleo denso da rede.

Ainda na figura 5.1(b) há um efeito sensível da imunização da camada $k_S = 2$. Como citado na seção 4.3, no regime de ativação k -core, a distância muito curta entre sítios de mesmo índice k_S determina sua eficiência em espalhar a epidemia. Devido à preponderância de tubos de comprimento baixo na rede *Air Traffic*, um possível efeito da remoção dos referidos tubos é o de remover caminhos redundantes para que a epidemia se espalhe no núcleo da rede ou mesmo de aumentar as distâncias entre sítios do núcleo, deslocando o limiar para direita.

As redes *AS CAIDA*, de tamanho $N = 26475$ sítios, e *AS Oregon*, de tamanho $N = 6474$ sítios, apresentam ainda uma quantidade considerável de sítios de grau 1 e 2 (37.5% e 36.8%, respectivamente). Estes sítios são os principais componentes das shells $k_S = 1$ e 2, cuja afinidade de ligação com sítios de grau elevado é considerável. No entanto, a remoção da primeira camada da rede ainda produz efeitos pequenos na dinâmica. Os valores baixos dos expoentes da distribuição de grau, $\gamma = 2.03$ e 1.8, é uma possível explicação para esse efeito pequeno, que é associado à ativação via *k-core*. De maneira similar à rede *Air Traffic*, a remoção de sítios da segunda camada ($k_S = 2$) diminui o número de caminhos redundantes entre sítios do núcleo, bem como eliminando possíveis atalhos.

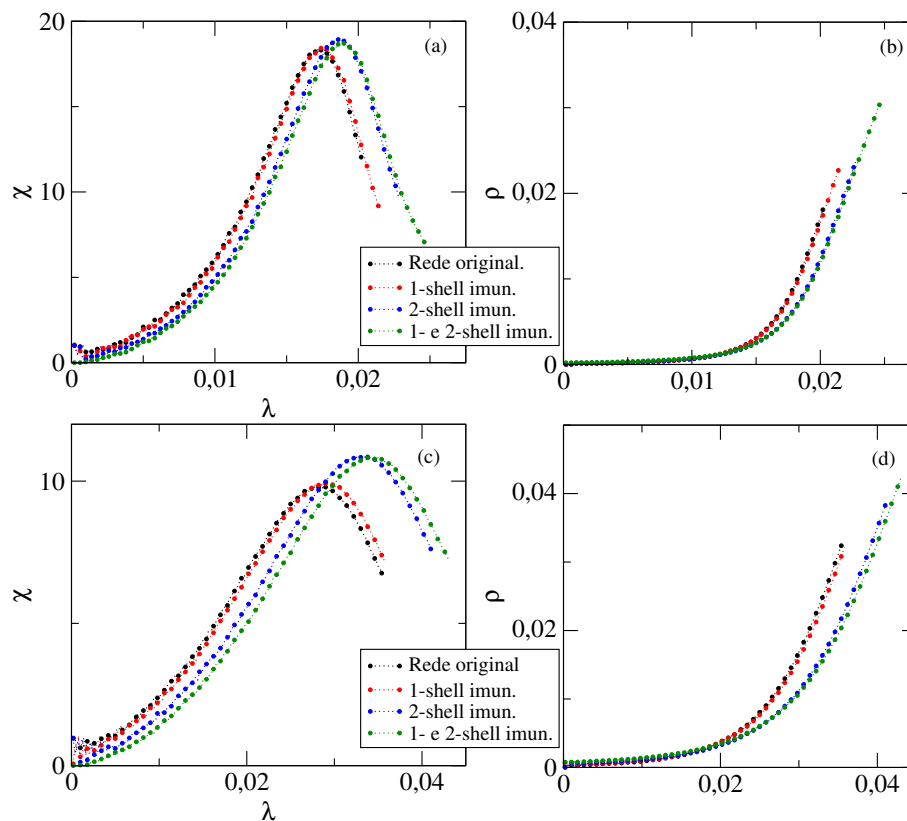


Figura 5.2 – (a) Suscetibilidade χ e (b) densidade de sítios ρ na rede *AS CAIDA*. (c) Suscetibilidade χ e (d) densidade ρ na rede *AS Oregon*. Efeitos da remoção conjunta ou independente das camadas $k_S = 1$ e 2 são mostrados em cada figura.

5.2 Resultados para redes sintéticas

O presente capítulo exhibe nossos principais resultados em redes sintéticas, comparando os efeitos da adição de folhas e tubos na dinâmica do núcleo da rede. Dentre os resultados estão medidas quase-estacionárias em amostras típicas e escalas de tamanho finito.

5.2.1 Dinâmica em redes com folhas

Nesta subseção estão disponíveis resultados dinâmicos em redes sintéticas após adição de folhas, segundo o modelo descrito na subseção 3.2.2. As próximas figuras apresentam curvas de suscetibilidade χ e densidade de sítios infectados ρ em amostras de redes típicas. Nelas são investigados os efeitos causados por folhas sobre a dinâmica do modelo SIS no núcleo de redes com distribuição de grau em lei de potência, com expoentes $\gamma = 2.3$ e 2.7 .

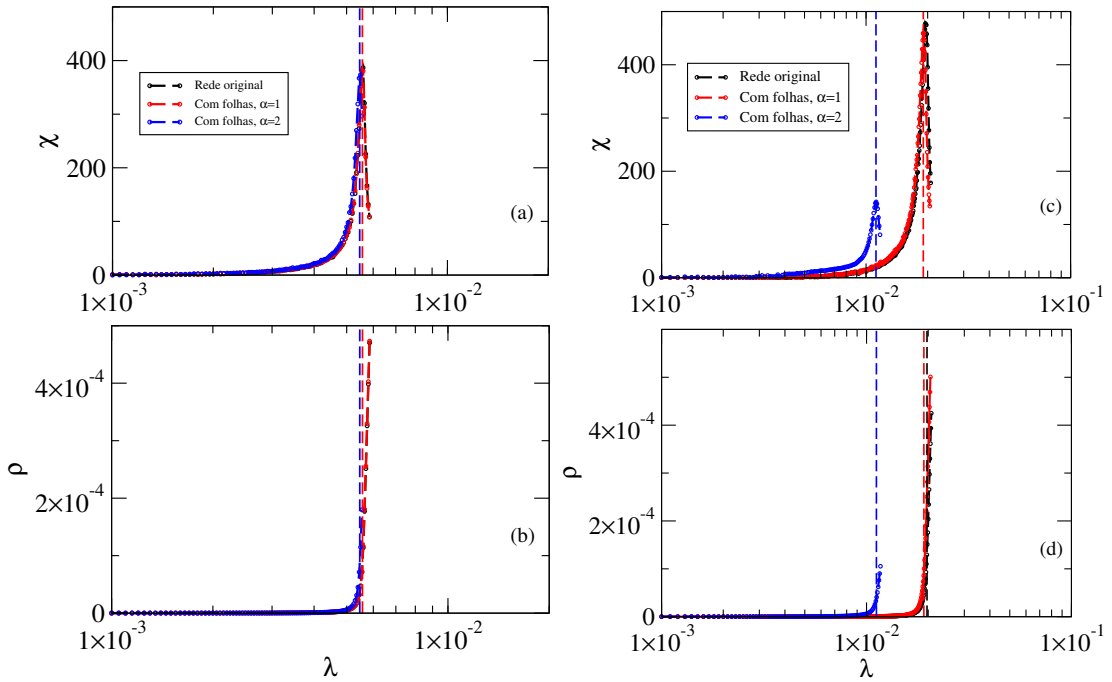


Figura 5.3 – (a) Suscetibilidade χ e (b) densidade de sítios infectados ρ em redes com $\gamma = 2.3$. (c) Suscetibilidade e (d) densidade de sítios infectados, em redes com $\gamma = 2.7$. Redes com distribuição $P(k) \propto k^{-\gamma}$ de tamanho $N_c = 10^7$ antes e após a adição de $N_1 = 0.5 \times N_c$ folhas ($\alpha = 1$ e 2).

As figuras 5.3(a) e (b) sugerem efeitos fracos da inserção de folhas em redes com $\gamma = 2.3$ através dos mecanismos de conexão de sítios periféricos considerados. Pequenos desvios sugerem que a ativação no núcleo k -core seja robusta mesmo sob a inserção massiva e altamente orientada aos sítios mais centrais da rede, representada pela ligação preferencial quadrática.

Por sua vez, as curvas mostradas nas figuras 5.3(c) e (d), referentes a redes com $\gamma = 2.7$, mostram uma pequena mudança no limiar epidêmico, ocorrendo após ligação preferencial linear ($\alpha = 1$) de folhas aos sítios do núcleo $P(k) \propto k^{-\gamma}$. Isto se deve à fraca ligação das folhas aos hubs e, conseqüentemente, do reduzido efeito do aumento do grau dos referidos hubs, como mostram os resultados do Capítulo III. No entanto, um deslocamento significativo é observado com a ligação preferencial quadrática ($\alpha = 2$). Os resultados do Capítulo III mostram que esta forma de inserção de sítios periféricos tem

a capacidade de promover super-hubs. Como discutido na seção 4.3, isto ocasiona um aumento no tempo de vida da infecção na vizinhança dos hubs assim surgidos.

Na figura 5.4 são apresentadas escalas de tamanho finito para o limiar epidêmico λ_p , pico de suscetibilidade χ_p e densidade de sítios infectado ρ_p no limiar. As medidas se referem a redes com expoente $\gamma = 2.3$.

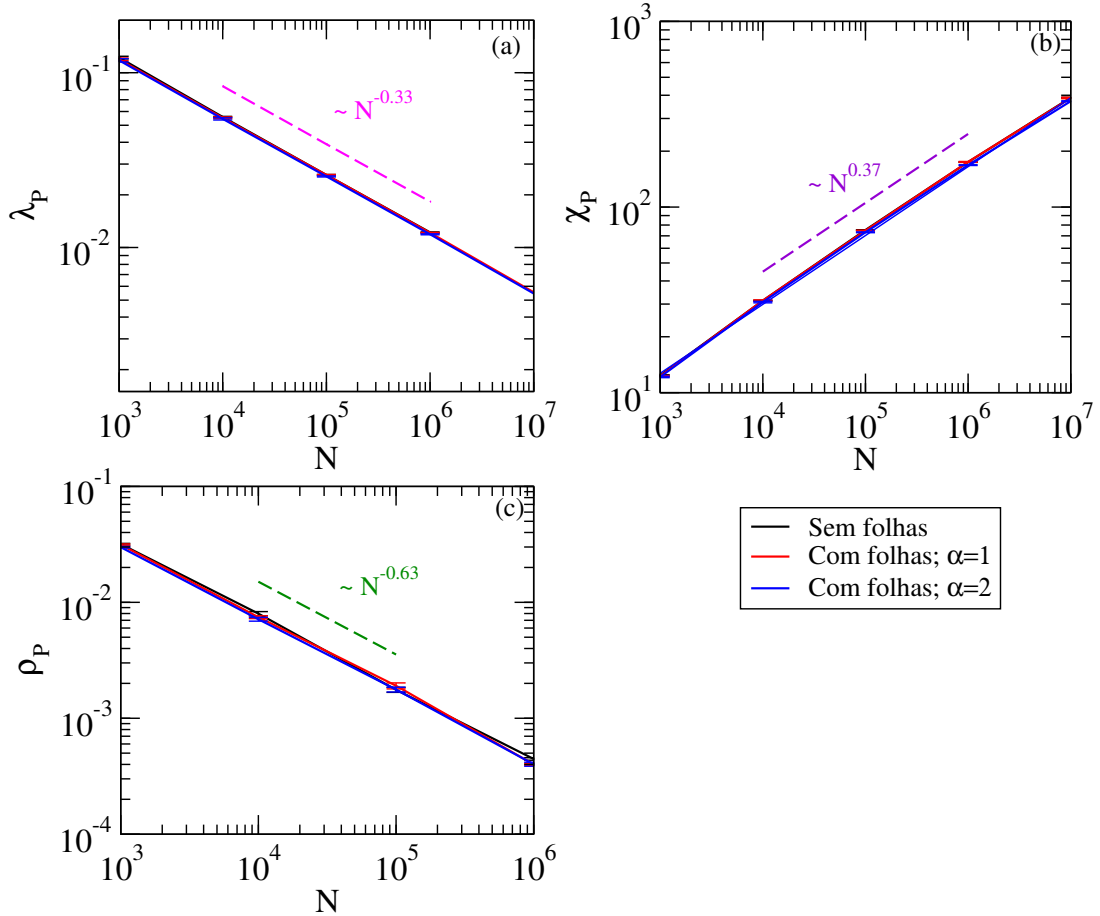


Figura 5.4 – Análise de tamanho finito de variáveis dinâmicas sobre redes em lei de potência com expoente $\gamma = 2.3$. Comparamos efeitos na rede original e nos núcleos após adição de folhas, via ligação preferencial linear ($\alpha = 1$) e ligação preferencial quadrática ($\alpha = 2$). (a) Limiar epidêmico, (b) suscetibilidade no pico, (c) densidade de sítios infectados no pico. Média sobre 5 amostras.

Os resultados da Figura 5.5 mostram como escalam as medidas quase-estacionárias de λ_p , ρ_p e χ_p com o tamanho do núcleo de redes com $\gamma = 2.7$, após a inserção de folhas. Como já sugeriam os gráficos das figuras 5.3(c) e (d), a ligação preferencial linear apresenta mudanças pequenas nos limiares, se comparadas àquelas provocadas pela ligação quadrática. Isto se deve ao fato de a adição de folhas via ligação preferencial linear não aumentar significativamente o tempo de vida dos hubs. No entanto, a ligação quadrática ($\alpha = 2$) apresenta efeitos apreciáveis no regime de ativação via hub-hub. É curioso notar que as curvas associadas à ligação quadrática desviam sistematicamente do comportamento usual, em lei de potência, conforme o tamanho da rede é aumentado.

A propósito, vale lembrar que a ativação via k -core não é inexistente para $\gamma > 2.5$ [9], mas antes, compete com a ativação hub-hub no estabelecimento da fase ativa. No entanto, para tamanhos crescentes da rede, os hubs que emergem da inserção de folhas se tornam tão grandes que passam a ter um tempo de vida que sobrepuja os efeitos da ativação k -core.

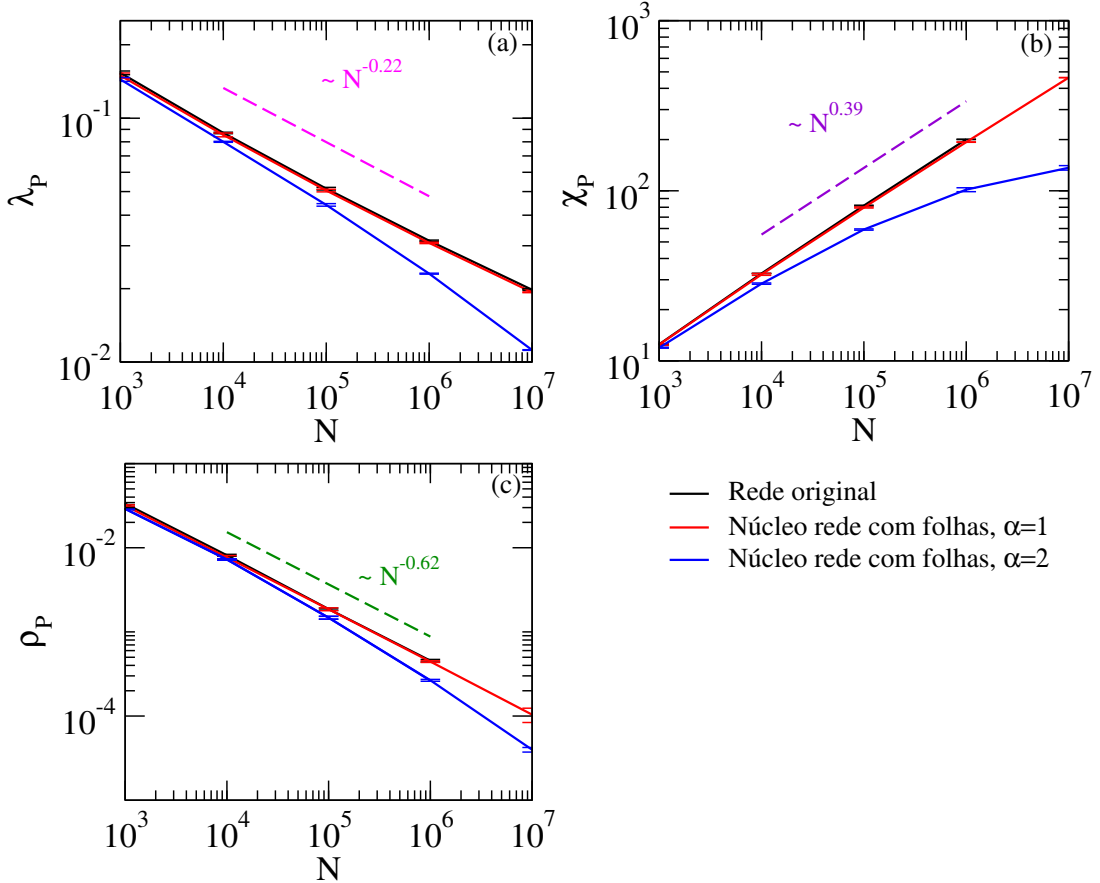


Figura 5.5 – Análise de tamanho finito de variáveis dinâmicas em redes com $\gamma = 2.7$. Efeitos da inserção de folhas são comparados em escalas de tamanho finito para o limiar λ_P (a), para o pico de suscetibilidade χ_P (b) e para a densidade de sítios infectados no limiar ρ_P (c). As linhas tracejadas mostram o comportamento em lei de potência da dinâmica da rede original. Média sobre 5 amostras.

Os gráficos a seguir apresentam amostras típicas para $\gamma = 3.5$. Para tamanhos pequenos de rede ainda é possível ver uma fase ativa bem definida, sem a presença de efeitos fortes de localização e, como no caso apresentado anteriormente, o limiar apresenta mudança significativa apenas para ligação quadrática ($\alpha = 2$). As figuras 5.6(a) e (b) apresentam resultados que comparam a dinâmica em redes com porção scale-free de tamanho $N_c = 10^4$ com a dinâmica nas mesmas redes, após a adição de $N_1 = 0.5 \times N_c$ folhas, com $\alpha = 1$ e 2.

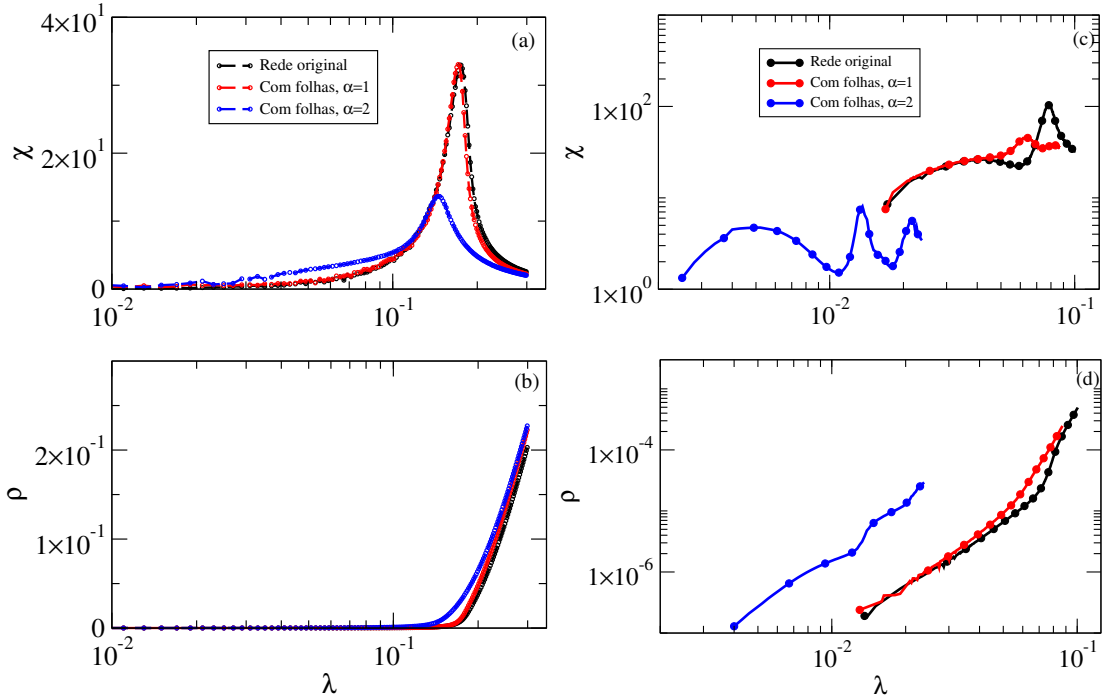


Figura 5.6 – Suscetibilidade χ e densidade de sítios infectados ρ em redes com $\gamma = 3.5$. (a) e (b): rede original, de tamanho $N_c = 10^4$ e redes após inserção de $N_1 = 0.5 \times N_c$ folhas, via ligação com $\alpha = 1$ e 2. (c) e (d): redes de tamanho $N_c = 10^7$, na ausência e na presença de $N_1 = 0.5 \times N_c$ folhas, inseridas via ligação linear e quadrática.

Conforme o tamanho das redes se torna maior, no entanto, efeitos de localização em hubs se tornam preponderantes, mesmo com ligação preferencial linear. Isto torna difícil estabelecer o limiar da fase global, levando a múltiplos picos na curva de suscetibilidade, como mostram as figuras 5.6(c) e (d). Nelas são apresentados resultados em redes com tamanho $N_c = 10^7$ na ausência e na presença de sítios periféricos, em cujo caso foi feita a adição de $N_1 = 0.5 \times N_c$ folhas através de ligação preferencial linear ($\alpha = 1$) e quadrática ($\alpha = 2$).

Um aspecto interessante a se observar com a adição das folhas é que os valores máximos das suscetibilidades são bastante reduzidos e a transição para o estado absorvente é menos aguda. Alguns trabalhos recentes [41, 42] sugerem que, quanto mais localizada for a atividade epidêmica, mais suave será a divergência da suscetibilidade nas transições. Essa conjectura está em consonância com os resultados apresentados nesta seção.

5.2.2 Dinâmica em redes com tubos

As curvas exibidas na figura 5.7 apresentam medidas quase-estacionárias utilizadas na investigação dos efeitos de sítios de grau $k = 2$ na dinâmica do modelo SIS em redes sintéticas com distribuição de grau em lei de potência. Para obtê-las, adicionamos a redes de tamanho $N_c = 10^5$ uma quantidade $N_2 = 0.25 \times N_c$ de sítios de grau $k = 2$. As medidas apresentadas na figura 5.7 tipificam os efeitos dos tubos em redes com $\gamma = 2.3$ e 2.7.

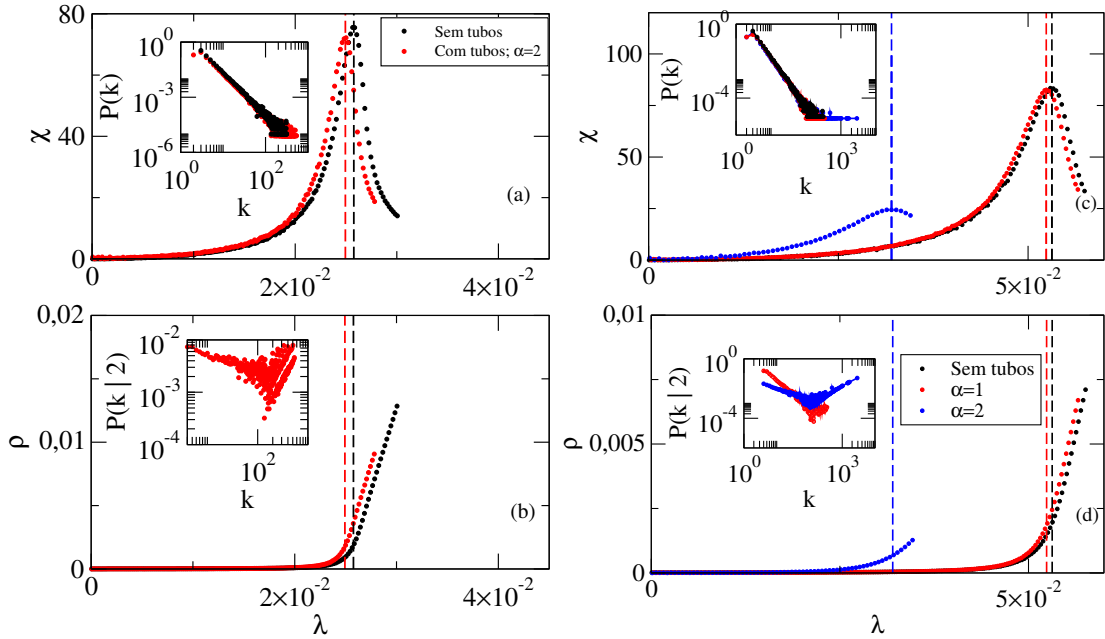


Figura 5.7 – (a) Suscetibilidade χ e (b) densidade de sítios infectados ρ em redes com expoente $\gamma = 2.3$. (c) Suscetibilidade χ e (d) densidade ρ em redes com $\gamma = 2.7$. Redes de tamanho $N_c = 10^5$ antes e após a adição de $N_2 = 0.25 \times N_c$, via $\alpha = 1$ e 2. Inserções mostram distribuições de graus da rede total e distribuição de graus de vizinhos de sítios de grau $k = 2$, compartilhando da mesma legenda que a sub-figura na qual estão inseridas.

A dinâmica em redes acopladas a tubos inseridos via ligação preferencial linear aparenta pequenos desvios da fenomenologia usual tanto em redes com $\gamma = 2.3$, quanto em redes com $\gamma = 2.7$, como mostra a figura 5.7. A inserção da figura 5.7(d), em vermelho, mostra a baixa afinidade de tubos, ligados à rede via mecanismo linear, com os hubs da rede.

No entanto, a ligação quadrática provocou forte desvio, mostrando que a presença dos tubos é relevante. O abaixamento do pico de suscetibilidade sugere forte correlação da atividade na vizinhança dos hubs, principais agentes de transmissão em redes com $\gamma > 2.5$ [9]. Além disso, a redução acentuada do máximo de suscetibilidade novamente sugere uma forte localização na epidemia, associada a desordens ou inomogeneidades estendidas fortemente correlacionadas, eventualmente ocasionada pela aproximação dos hubs provocada pelos tubos.

A seguir, na figura 5.8, são mostrados comportamentos de escala de tamanho finito em redes com $\gamma = 2.3$ após adição de tubos, via $\alpha = 2$. Os desvios poucos acentuados do limiar, vistos na figura 5.7(a), confirmam-se nos comportamentos de escala da figura 5.8(a). Estes resultados encontram mais uma vez explicação na robustez da ativação no núcleo *k-core*. Devido à densa conexão entre os sítios do núcleo, a adição de tubos não é capaz de acrescentar atalhos e a presença de novos caminhos, ou mesmo caminhos redundantes, não traz qualquer ganho na transmissão da infecção entre os sítios do núcleo *k-core*. Para completar, estimamos os expoentes obtidos nesta análise de tamanho finito. Os valores de $\lambda_P \sim N^{-0.33}$, $\chi_P \sim N^{0.38}$ e $\rho_P \sim N^{-0.62}$ estão em acordo com os encontrados para o modelo SIS em outros trabalhos [36, 43].

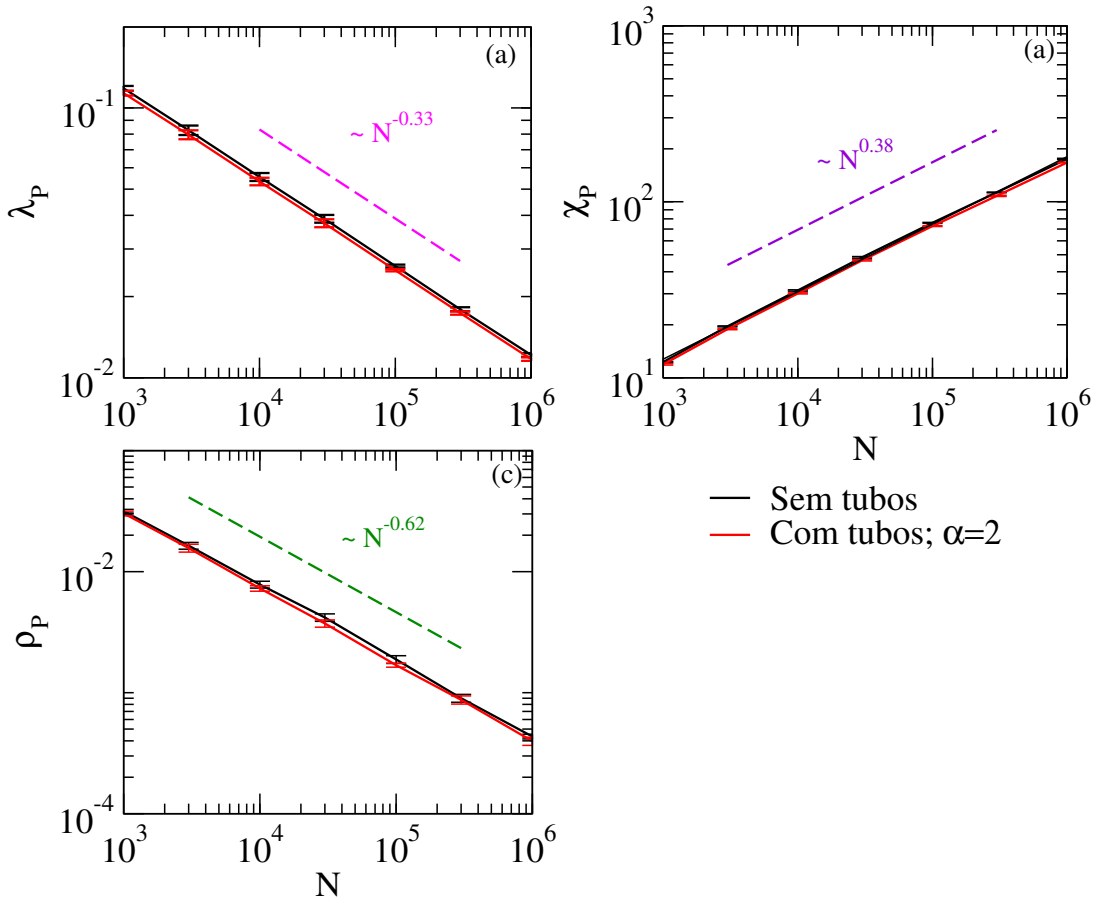


Figura 5.8 – Escalas de tamanho finito. (a) Limiar de transição λ_P , (b) pico de suscetibilidade χ_P e (c) densidade de sítios infectados no limiar ρ_P . Efeitos da adição de tubos (ligados com $\alpha = 2$) são comparados àqueles presentes em redes com $\gamma = 2.3$ sem tubos. Linhas tracejadas servem como guia para os olhos. Média sobre 25 amostras

Os comportamentos de escala apresentados na figura 5.9, corroboram a hipótese de uma transição smeared devido aos fortes desvios dos comportamentos usuais de escala, caracterizadas pelos expoentes críticos.

É conhecido na literatura que a classe de *Percolação Direcionada* não possui resiliência a desordens estendidas fortemente correlacionadas [44]. Fortuitamente, a presença de correlações fortes na dinâmica entre os hubs, causada não apenas pelo aumento do tempo de vida da atividade em suas vizinhanças, mas também pelo acréscimo de atalhos, foi o fator dominante dos desvios observados. Para o caso $\alpha = 1$, obtivemos novamente expoentes em acordo com a literatura [36, 43].

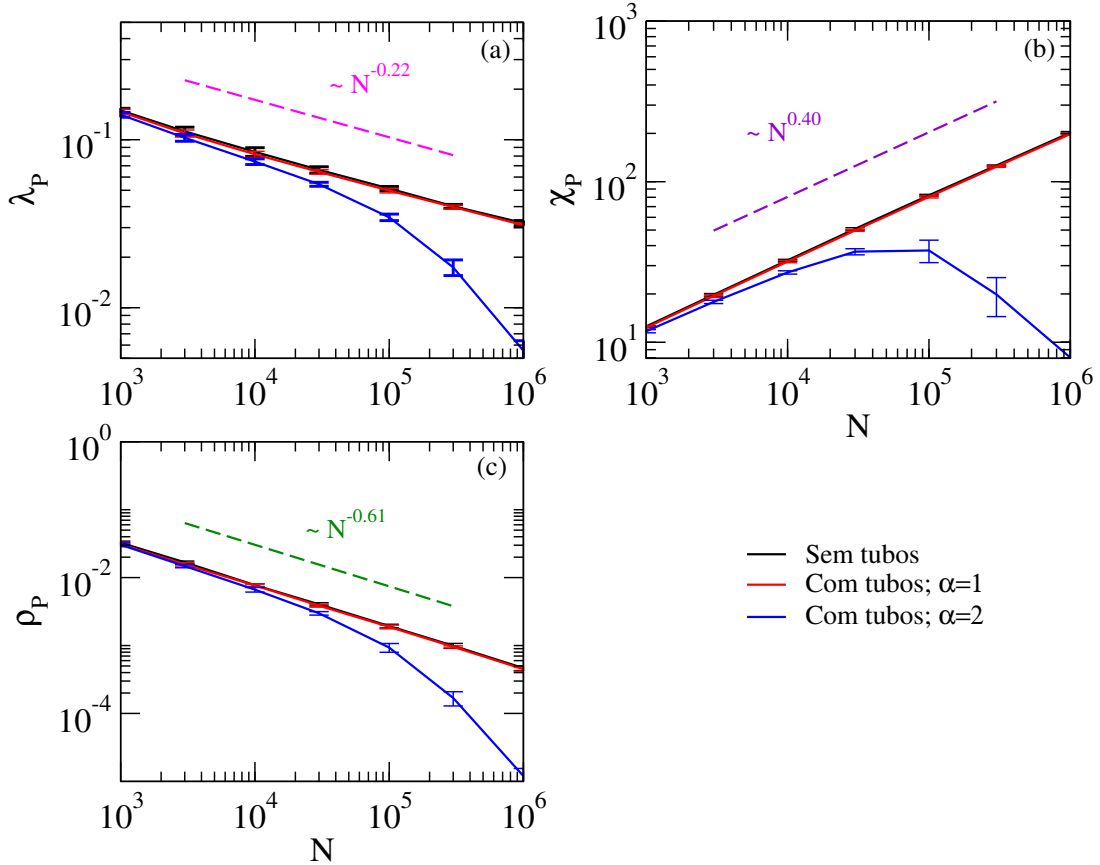


Figura 5.9 – Escala de tamanho finito para (a) limiar de transição λ_P , (b) pico de suscetibilidade χ_P e (c) densidade de sítios infectados no limiar ρ_P . Efeitos de sítios de grau 2 (ligados com $\alpha = 1$ e 2) são comparados em redes com $\gamma = 2.7$. Linhas tracejadas servem como guia para os olhos. Média sobre 25 amostras

A figura 5.10 apresenta resultados da inserção de tubos em redes com $\gamma = 3.5$. Neste regime, é preponderante a ação dos hubs [9]. No entanto, diferentemente dos resultados apresentados até então, um deslocamento do limiar epidêmico já é apreciável via ligação preferencial linear, como sugere figura 5.10(a).

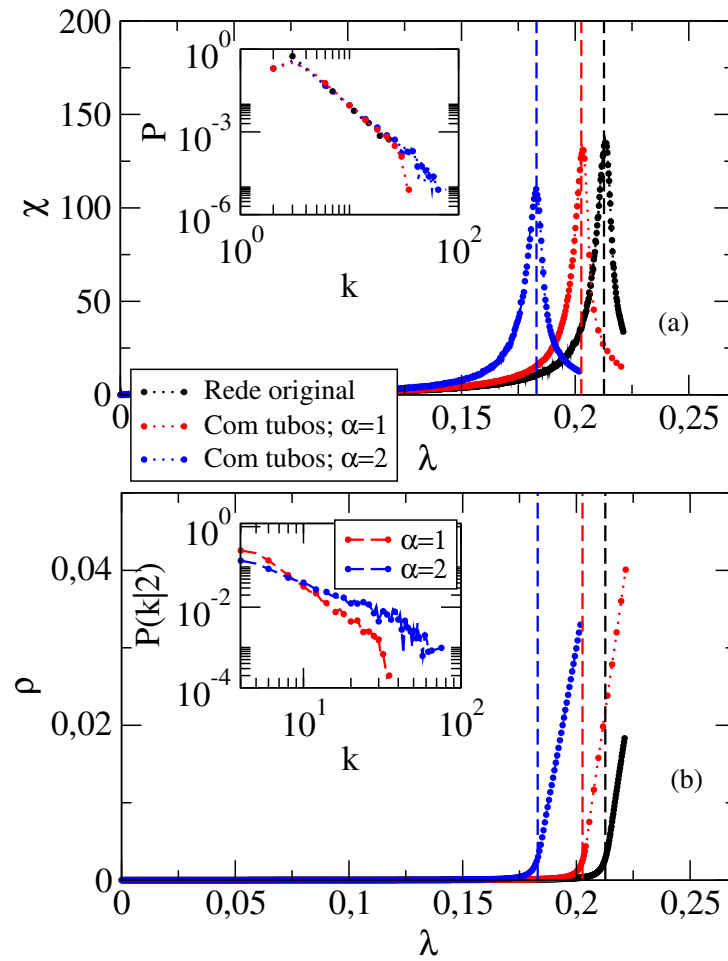


Figura 5.10 – Medidas quase-estacionárias em redes com lei de potência, de tamanho $N = 10^5$ com $\gamma = 3.5$. (a): Suscetibilidade χ . (b): densidade de sítios infectados ρ . Curvas comparam resultados em: rede com distribuição puramente em lei de potência e redes após adição de $N_2 = 0.25 \times N_c$ sítios de grau $k = 2$, via $\alpha = 1$ e 2. Inserções: (a) exhibe distribuições de graus e (b) exhibe distribuições de graus de vizinhos de sítios de grau $k = 2$. Linhas tracejadas mostram o ponto de transição. Todas as figuras compartilham da mesma legenda de cores.

VI Considerações finais

Nos primeiros capítulos deste trabalho apresentamos conceitos fundamentais à teoria de redes complexas, importantes na obtenção das principais métricas. Além disso, o método empregado na criação de redes descorrelacionadas, importantes no estudo de modelos teóricos, foi apresentado através do modelo UCM. Nosso intuito foi aproximar o leitor da metodologia aplicada na criação das redes utilizadas como substrato neste trabalho. Nosso algoritmo de criação de redes ricas em folhas, bem como tubos, foi também apresentado e discutido.

Ainda em capítulos subsequentes, fizemos a apresentação do modelo associado ao processo dinâmico investigado neste trabalho, o modelo SIS. Fizemos também uma breve apresentação e discussão de conceitos fundamentais à teoria de transições de fase fora do equilíbrio, como o de estado microscópico da dinâmica e a probabilidade de encontrar o sistema em algum destes estados, cuja evolução é descrita pela equação mestra. Apresentamos e discutimos as principais teorias de campo médio, utilizadas como meio de simplificar a equação mestra na busca de soluções aproximadas, devido às importantes informações qualitativas trazidas por estas teorias sobre o processo dinâmico. Por fim, o método quase-estacionário é apresentado, fazendo uma conexão entre abordagem teórica e métodos computacionais, motivando os métodos empregados neste trabalho.

Nossos resultados, apresentados no capítulo V, sugerem que o papel dos sítios periféricos dependerá fortemente do regime de ativação predominante na epidemia (*k-core* ou *interação de hubs*) e seu regime de inserção na rede (preferencial linear ou quadrática). O esquema de ligação quadrática, na inserção de folhas, foi capaz de facilitar a interação entre hubs, via aumento do tempo de vida da atividade em suas vizinhanças, em redes com $\gamma = 2.7$ e 3.5 . Como discutido, este mecanismo é fundamental na ativação da epidemia em redes com $\gamma > 2.5$ [9]. Em parte, a inserção de tubos, via $\alpha = 2$, pôde realizar o mesmo efeito. Inclusive, a escolha da quantidade de sítios de grau $k = 2$ inseridos na rede foi realizada no intuito de comparar seus efeitos ao das folhas no que tange o aumento do grau dos hubs, proporcionado por ambos os tipos de sítios periféricos. Com isto, pudemos discriminar entre o efeito do aumento do tempo de atividade na vizinhança dos hubs e o efeito da formação de atalhos e caminhos redundantes.

Em geral, a ligação preferencial linear ($\alpha = 1$) de folhas e tubos ao núcleo da rede não foi capaz de gerar efeitos significativos, a não ser para $\gamma = 3.5$. O acoplamento fraco de folhas e tubos aos hubs da rede neste regime de ligação não foi suficiente para elevar o tempo de atividade na vizinhança próxima aos hubs. Ainda, no caso dos tubos, sua inserção neste regime foi insuficiente para aproximar os principais agentes do estabelecimento da epidemia e, assim, fortalecer a correlação de suas atividades. Em contraposição, em redes com $\gamma = 3.5$, mais homogêneas, os sítios periféricos apresentaram maior afinidade

de ligação com os hubs.

Enquanto os resultados da ligação quadrática para $\gamma = 2.7$ mostram efeitos decorrentes dos sítios periféricos, para $\gamma = 2.3$ os métodos de ligação considerados neste trabalho não foram capazes de trazer qualquer amplificação na ativação via *k-core*. Neste regime, o núcleo densamente conectado da rede, associado a sítios com índice k_S máximo, é por si só responsável por sustentar a epidemia de maneira coletiva, suplantando as características individuais de cada um dos sítios [21]. Neste regime de ativação, a inserção de folhas e tubos não apresentou efeitos via quaisquer método de ligação.

Ainda, com $\gamma = 3.5$, efeitos interessantes e diversos foram observados. Tanto o método de ligação linear, quanto o método de ligação quadrática foram capazes de alterar drasticamente a dinâmica. Para tamanhos pequenos de rede, apenas a ligação via $\alpha = 2$ foi capaz de alterar significativamente o limiar epidêmico, sugerindo por meio deste resultado que nossa interpretação dos resultados para $\gamma = 2.7$ estão corretos. No entanto, com o aumento do tamanho da rede, mesmo a ligação preferencial linear foi capaz de promover o surgimento de hubs acima do grau máximo da rede original de tal forma que sua interação foi capaz de deslocar o limiar para a esquerda. A ligação quadrática, por sua vez, teve os maiores efeitos, em consonância com os resultados para $\gamma = 2.7$, apresentando, no entanto, efeitos de localização da atividade em torno dos hubs. Este efeito, no entanto, torna difícil distinguir entre a fase global e a fase localizada, tornando a análise de tamanho finito ambígua.

Em redes com tubos, assim como no caso das folhas, a dinâmica em redes com $\gamma = 2.3$ não foi sensível à sua inserção em nenhum regime de conexão. Os resultados estão de acordo com a conclusão de que a ativação no *k-core* não é sensível ao grau dos sítios, devido à interação já forte dos envolvidos na ativação [21].

Os resultados obtidos neste trabalho nos permitem delinear algumas estratégias de otimização de algoritmos de simulação para o modelo SIS. Estas envolvem principalmente a imunização de sítios de grau $k = 1$ e 2 , anterior à simulação do processo dinâmico, baseada na análise da distribuição de graus, visto que o comportamento quase-estacionário foi resiliente à ligação linear em redes com $\gamma = 2.3$ e 2.7 .

Para trabalhos futuros, visamos investigar de forma mais aprofundada a ligação de folhas e tubos ao núcleo $P(k) \propto k^{-\gamma}$, para $\gamma > 3$, utilizando ainda outros mecanismos de ligação ao núcleo.

Referências

- [1] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks*. Cambridge University Press, 2008.
- [2] A.-L. Barabási *et al.*, *Network science*. Cambridge university press, 2016.
- [3] C. Viboud, K. Sun, R. Gaffey, M. Ajelli, L. Fumanelli, S. Merler, Q. Zhang, G. Chowell, L. Simonsen, A. Vespignani, *et al.*, “The rapid ebola forecasting challenge: Synthesis and lessons learnt,” *Epidemics*, vol. 22, pp. 13–21, 2018.
- [4] Y. Lu and J. Liu, “The impact of information dissemination strategies to epidemic spreading on complex networks,” *Physica A: Statistical Mechanics and its Applications*, p. 120920, 2019.
- [5] W. Liu and S. Zhong, “Web malware spread modelling and optimal control strategies,” *Scientific reports*, vol. 7, p. 42308, 2017.
- [6] D. Balcan, B. Gonçalves, H. Hu, J. J. Ramasco, V. Colizza, and A. Vespignani, “Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model,” *Journal of computational science*, vol. 1, no. 3, pp. 132–145, 2010.
- [7] G. F. De Arruda, A. L. Barbieri, P. M. Rodríguez, F. A. Rodrigues, Y. Moreno, and L. da Fontoura Costa, “Role of centrality for the identification of influential spreaders in complex networks,” *Physical Review E*, vol. 90, no. 3, p. 032812, 2014.
- [8] R. Ghosh and K. Lerman, “Rethinking centrality: the role of dynamical processes in social network analysis,” *arXiv preprint arXiv:1209.4616*, 2012.
- [9] C. Castellano and R. Pastor-Satorras, “Competing activation mechanisms in epidemics on networks,” *Scientific reports*, vol. 2, p. 371, 2012.
- [10] M. Boguná, C. Castellano, and R. Pastor-Satorras, “Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks,” *Physical review letters*, vol. 111, no. 6, p. 068701, 2013.
- [11] S. C. Ferreira, R. S. Sander, and R. Pastor-Satorras, “Collective versus hub activation of epidemic phases on networks,” *Physical Review E*, vol. 93, no. 3, p. 032314, 2016.
- [12] R. Dickman and M. M. de Oliveira, “Quasi-stationary simulation of the contact process,” *Physica A: Statistical Mechanics and its Applications*, vol. 357, no. 1, pp. 134–141, 2005.

-
- [13] L. Antigueira, M. d. G. V. Nunes, O. Oliveira Jr, and L. d. F. Costa, “Strong correlations between text quality and complex networks features,” *Physica A: Statistical Mechanics and its Applications*, vol. 373, pp. 811–820, 2007.
- [14] I. Dobson, “Synchronization and your morning coffee,” *Nature Physics*, vol. 9, pp. 133–134, 2013.
- [15] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley, “Finding a better immunization strategy,” *Physical review letters*, vol. 101, no. 5, p. 058701, 2008.
- [16] M. Newman, *Networks: an introduction*. Oxford university press, 2010.
- [17] M. Bockholt and K. A. Zweig, *Paths in Complex Networks*, pp. 1–11. New York, NY: Springer New York, 2017.
- [18] M. Molloy and B. Reed, “A critical point for random graphs with a given degree sequence,” *Random structures & algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995.
- [19] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, “Structure of growing networks with preferential linking,” *Physical review letters*, vol. 85, no. 21, p. 4633, 2000.
- [20] M. Boguná, R. Pastor-Satorras, and A. Vespignani, “Cut-offs and finite size effects in scale-free networks,” *The European Physical Journal B*, vol. 38, no. 2, pp. 205–209, 2004.
- [21] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, “Identification of influential spreaders in complex networks,” *Nature physics*, vol. 6, no. 11, p. 888, 2010.
- [22] C. Giatsidis, F. D. Malliaros, N. Tziortziotis, C. Dhanjal, E. Kiagias, D. M. Thilikos, and M. Vazirgiannis, “A k-core decomposition framework for graph clustering,” *arXiv preprint arXiv:1607.02096*, 2016.
- [23] M. Catanzaro, M. Boguná, and R. Pastor-Satorras, “Generation of uncorrelated random scale-free networks,” *Physical review e*, vol. 71, no. 2, p. 027103, 2005.
- [24] F. Radicchi and C. Castellano, “Breaking of the site-bond percolation universality in networks,” *Nat. Commun.*, vol. 6, p. 10196, dec 2015.
- [25] F. Radicchi, “Predicting percolation thresholds in networks,” *Phys. Rev. E*, vol. 91, p. 010801, jan 2015.
- [26] C. Castellano and R. Pastor-Satorras, “Relating topological determinants of complex networks to their spectral properties: Structural and dynamical effects,” *Phys. Rev. X*, vol. 7, p. 041024, oct 2017.

- [27] M. F. Gomes, A. P. y Piontti, L. Rossi, D. Chao, I. Longini, M. E. Halloran, and A. Vespignani, “Assessing the international spreading risk associated with the 2014 west african ebola outbreak,” *PLoS currents*, vol. 6, 2014.
- [28] P. Kim and C. H. Lee, “Epidemic spreading in complex networks with resilient nodes: Applications to fmd,” *Complexity*, vol. 2018, 2018.
- [29] F. Nian and X. Wang, “Efficient immunization strategies on complex networks,” *Journal of theoretical biology*, vol. 264, no. 1, pp. 77–83, 2010.
- [30] W. Quattrociocchi, M. Del Vicario, A. Bessi, F. Zollo, *et al.*, “The spreading of misinformation online,” 2016.
- [31] J. Ratkiewicz, M. D. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. M. Menczer, “Detecting and tracking political abuse in social media,” 2011.
- [32] J. S. Koopman, S. E. Chick, C. S. Riolo, A. L. Adams, M. L. Wilson, and M. P. Becker, “Modeling contact networks and infection transmission in geographic and social space using germs,” *Sexually transmitted diseases*, vol. 27, no. 10, pp. 617–626, 2000.
- [33] 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.
- [34] F. Chung, L. Lu, and V. Vu, “The spectra of random graphs with given expected degrees,” *Internet Mathematics*, vol. 1, no. 3, pp. 257–275, 2004.
- [35] J. Marro and R. Dickman, *Nonequilibrium phase transitions in lattice models*. Cambridge University Press, 2005.
- [36] 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.
- [37] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. Elsevier, 1992.
- [38] R. Dickman and R. Vidigal, “Quasi-stationary distributions for stochastic processes with an absorbing state,” *Journal of Physics A: Mathematical and General*, vol. 35, no. 5, p. 1147, 2002.
- [39] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, “K-core organization of complex networks,” *Physical review letters*, vol. 96, no. 4, p. 040601, 2006.

-
- [40] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek, “Evolution of the social network of scientific collaborations,” *Physica A: Statistical mechanics and its applications*, vol. 311, no. 3-4, pp. 590–614, 2002.
- [41] W. Cota, S. C. Ferreira, and G. Odor, “Griffiths effects of the susceptible-infected-susceptible epidemic model on random power-law networks,” *Physical review E*, vol. 93, no. 3, p. 032322, 2016.
- [42] D. H. Silva and S. C. Ferreira, “Activation thresholds in epidemic spreading with motile infectious agents on scale-free networks,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 12, p. 123112, 2018.
- [43] W. Cota, A. S. Mata, and S. C. Ferreira, “Robustness and fragility of the susceptible-infected-susceptible epidemic models on complex networks,” *Phys. Rev. E*, vol. 98, p. 012310, Jul 2018.
- [44] M. Dickison and T. Vojta, “Monte carlo simulations of the smeared phase transition in a contact process with extended defects,” *Journal of Physics A: Mathematical and General*, vol. 38, no. 6, p. 1199, 2005.