

**UNIVERSIDADE ESTADUAL DE PONTA GROSSA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS
ÁREA DE CONCENTRAÇÃO - FÍSICA**

ROBERTO ANTONIO VOSGERAU

ANÁLISE DA ENTROPIA EM REDES COMPLEXAS

**PONTA GROSSA
2016**

ROBERTO ANTONIO VOSGERAU

ANÁLISE DA ENTROPIA EM REDES COMPLEXAS

Tese apresentada ao Programa de Pós-Graduação em Ciências, área de concentração Física, da Universidade Estadual de Ponta Grossa, como requisito para obtenção do título de doutor em Ciências.

Orientador: Prof. Dr. Sandro Ely de Souza Pinto

**PONTA GROSSA
2016**

Ficha Catalográfica
Elaborada pelo Setor de Tratamento da Informação BICEN/UEPG

V967 Vosgerau, Roberto Antonio
Análise da entropia em redes complexas/
Roberto Antonio Vosgerau. Ponta Grossa,
2016.
105f.

Tese (Doutorado em Ciências - Área de
Concentração: Física), Universidade
Estadual de Ponta Grossa.

Orientador: Prof. Dr. Sandro Ely de
Souza Pinto.

1.Redes complexas. 2.Robustez.
3.Entropias estruturais. I.Pinto, Sandro
Ely de Souza. II. Universidade Estadual
de Ponta Grossa. Doutorado em Ciências.
III. T.

CDD: 519.2

TERMO DE APROVAÇÃO

ROBERTO ANTONIO VOSGERAU

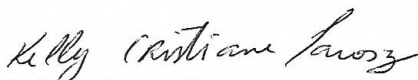
“Análise da Entropia em Redes Complexas”

Tese aprovada como requisito parcial para obtenção do grau de Doutor no Programa de Pós-Graduação em Ciências - Física da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora.

Orientador:



Prof. Dr. Sandro Ely de Souza Pinto
Departamento de Física – UEPG/PR



Prof.ª Dr.ª Kelly Cristiane Iarosz
Universidade de São Paulo – USP



Prof.ª Dr.ª Sani de Carvalho Rutz da Silva
Universidade Federal Tecnológica – UTFPR



Prof. Dr. Adriano Doff Sotta Gomes
Universidade Federal Tecnológica – UTFPR



Prof. Dr. José Danilo Szezech Júnior
Departamento de Matemática e Estatística - UEPG/PR

Ponta Grossa, 25 de abril de 2016

Dedico este trabalho aos meus pais (in memoriam) Fredolim e Iloah e à minha esposa Maria Jurema. A eles meu carinho, gratidão, respeito e profundo reconhecimento.

AGRADECIMENTOS

A Deus pela vida e presença em todos os momentos.

Ao professor e pesquisador Dr. Sandro Ely de Souza Pinto que durante todo o doutorado, colaborou de forma positiva com minha formação científica e intelectual, sempre acreditando em mim.

À minha família pelo apoio, orações e incentivos.

Às minhas filhas Rosana, Lígia e minha neta Maria Eduarda pela compreensão e paciência.

Aos meus filhos Marcos e Flávio, meus amigos de modo especial ao prof. Dr. Romeu M. Szmowski e Dr. Rodrigo F. Pereira que muito ajudaram na realização deste.

Aos professores Dr. Maurício Kaster e Dr. Josmar Ivanqui do DAELE da UTFPR, ao professor Dr. Antonio Marcos, ao professor Dr. Giuliano e aos colegas Rafael R. Borges, Maurício A. Ribeiro, Pedro Jeferson de Miranda e os da sala L-24.

A todos os professores do programa de pós graduação em Ciências e do Departamento de Física da UEPG, pela participação e colaboração nos conhecimentos adquiridos.

A Universidade Tecnológica Federal do Paraná - UTFPR que através do departamento acadêmico de Matemática proporcionou condições para que o curso fosse realizado.

Aos membros da qualificação e da defesa pelas valiosas correções e sugestões.

À Josecler Lepinski, pela competência, eficiente e pronto atendimento às solicitações relacionadas com procedimentos burocráticos.

A todos que direta ou indiretamente contribuíram para realização deste trabalho.

RESUMO

Nesta tese, aplicou-se a entropia de rede para analisar as características de redes complexas. Analisou-se o efeito da remoção de vértices sobre a estrutura e robustez das redes direcionadas e não direcionadas, como resultado de falhas e ataques. A variação na estrutura e robustez das redes devido a esta remoção é quantificada por entropias estruturais com aproximação de ordem zero e um. Estas entropias são calculados para redes reais para verificar a robustez, onde utilizou-se a rede biológica como direcionada e rede literária como não direcionada. As falhas são simuladas através de vértices que são removidos aleatoriamente, enquanto os ataques são realizados intencionalmente nos vértices de maiores graus e de acordo com a centralidade de intermediação. Observa-se que a rede com falhas exibe entropias com valores menores do que a rede original. Quando o ataque é nos vértices de maiores graus ou de acordo com as centralidades de intermediação verifica-se que as entropias diminuem. No entanto, se o ataque não está em vértices de maiores graus ou maiores centralidade de intermediação, é possível verificar valores de entropias maiores do que a rede original. Além disso, a medida do comprimento do menor percurso mostrou uma correlação com a entropia estrutural.

Palavras-chave: Redes complexas, robustez, entropias estruturais.

ABSTRACT

In this thesis, we apply network entropy to analyze characteristics of complex networks. It is analyzed the effect of vertices removal on the structure and robustness of directed and undirected networks, as a result of failures and attacks. The variation in the structure and robustness of networks due to this removal is quantified by structural entropies with zero and one order approximation. These entropies are calculated to real networks to verify the robustness, where we have been using biological network as directed and literary network as undirected. The failures are simulated through vertices that are randomly removal, while the attacks are realized on the hubs and according to betweenness centrality. We observe that the network with failures exhibits entropies with values smaller than the original network. When the attack is on the hubs or according to the betweenness centrality we verify that the entropies decrease. However, if the attack is not on the hubs, it is possible to verify values of entropies larger than the original network. Moreover, the measure of the shortest path length showed a correlation with the structural entropy.

Keywords: complex networks, robustness, structural entropy.

LISTA DE FIGURAS

Figura 1.1 – Representações do problema das setes pontes de Königsberg e seu grafo. . . .	16
Figura 2.1 – Grafo não orientado com 4 nós e 3 arestas.	20
Figura 2.2 – Representação de um grafo direcionado com 4 nós e 3 arestas.	21
Figura 2.3 – Grafo completo K_4 com 4 nós e 6 arestas.	21
Figura 2.4 – Representação de um grafo ponderado com pesos nas arestas.	22
Figura 2.5 – Diâmetro de um grafo direcionado de 5 nós e 4 arestas.	23
Figura 2.6 – Caminho Euleriano de um grafo não direcionado. Cada aresta é utilizada uma única vez.	24
Figura 2.7 – Caminho Hamiltoniano 1-2-3-4-5 num grafo direcionado. Cada vértice é utilizado uma única vez.	24
Figura 2.8 – Grafos com vértices altamente conectados numa rede simples (a) e numa rede ponderada (b), com os pesos das arestas representados por linhas mais espessas.	26
Figura 2.9 – Ilustração de dois diferentes valores para o coeficiente de aglomeração para o vértice em vermelho.	27
Figura 2.10–Distribuição de graus num grafo não direcionado em função do número de conexões de cada vértice.	28
Figura 2.11–Grau médio dos vizinhos do vértice 1, representados pelos vértices 2, 3 e 4. . .	29
Figura 2.12–Gráfico de função assortativa e disassortativa.	30
Figura 2.13–Representação da estrutura de um hidrocarboneto (a) e o seu respectivo grafo (b).	31
Figura 2.14–Grafo simples de cinco nós e seis arestas com energia $\varepsilon = 6,7298$	31
Figura 2.15–Modelos de redes aleatórias com 10 nós e probabilidades $p = 0,1$ e $p = 0,3$. $G(10,0,1)$ na figura (a) e $G(10,0,3)$ na figura (b).	33
Figura 2.16–Exemplo de um grafo aleatório de Erdős e Rényi com $N = 50$ (figura (a)), usando uma probabilidade $p = 0,3$ e sua distribuição de graus (figura (b)).	34
Figura 2.17–Representação do modelo Watts-Strogatz com parâmetros $N = 15$ e $p = 0$ na figura (a) e $N = 15$ e $p = 0,4$ em (b).	36
Figura 2.18–Rede <i>small world</i> com 30 nós com grau 4 e probabilidade de religação	37
Figura 2.19–(a) Modelo de rede do Barabasi com 50 nós e grau médio igual a 4 e (b) distribuição de graus. Em escala (logarítmica) de uma rede livre de escala com 500 vértices.	39
Figura 2.20–Pedacos de arestas denominadas de stubs. Os stubs são ligados entre si até formarem uma rede simples.	41
Figura 3.1 – Rede estelar de sete nós (distinguíveis) com entropia nula.	45
Figura 3.2 – Redes não direcionadas geradas pela sequência de graus $\{3,2,2,2,1\}$, onde o número de redes $\mathcal{N}\{k_i\} = 6$	46
Figura 3.3 – Entropia do modelo Erdős-Rényi calculada para redes (contadas de baixo para cima) com três, quatro, cinco e seis nós respectivamente, calculadas pela equação 3.13.	47
Figura 3.4 – Relação entre a entropia H e γ para o número de nós $N = 100$ (cor vermelha), 1000 (cor verde) e 10000 (cor azul) no modelo livre de escala com grau mínimo $m = 1$	48
Figura 3.5 – Entropia pelo grau médio de uma rede livre de escala com $N = 10000$ nós e grau médio igual a seis.	49
Figura 3.6 – Comportamento da entropia em relação à densidade.	54

Figura 3.7 – Exemplo de rede aleatória com nove nós e treze arestas cuja entropia $S_0 = 2,62$	58
Figura 3.8 – Grafo realizado pela sequência de graus $\{5, 4, 3, 3, 3, 2, 2, 2, 2\}$ com duas comunidades cuja entropia $S_c = 0,79$	65
Figura 3.9 – Grafo direcionado ($N = 4$, $L^{dir} = 5$) com entropia $S_0^{dir} = 2,25$	66
Figura 4.1 – Rede literária da Odisseia.	70
Figura 4.2 – Rede literária social do Senhor dos anéis.....	72
Figura 4.3 – Rede complexa da tolerância oral.	73
Figura 5.1 – Ensemble de 64 grafos com quatro nós e o número de arestas variando de zero a seis (número máximo).	76
Figura 5.2 – Ilustração da rede de amizade do Clube de Karatê constituído por 34 nós e 78 ligações.	78
Figura 5.3 – Entropias S_0 (figuras (a) e (b)) e S_1 (figuras (c) e (d)) relativas aos ataques aos maiores (menores) graus em ordem descendente (ascendente) dos nós da rede Odisseia. O número de ataques é representado por n.....	82
Figura 5.4 – Entropias S_0 (figuras (a) e (b)) e S_1 (figuras (c) e (d)) relativas aos ataques aos nós com maiores (menores) centralidades de intermediação da rede Odisseia. O número de ataques é representado por n.	83
Figura 5.5 – Entropias S_0 (figura (a)) e S_1 (figura (b)) relativas aos ataques aleatórios à rede da Odisseia. O número de ataques é representado por n. Os nós atacados e o índice δ estão representados na figura (c) e os nós atacados e o menor caminho médio l , na figura (d).....	84
Figura 5.6 – Entropias de ordem zero e de ordem um referentes aos ataques nos 100 maiores/menores graus da rede Senhor dos anéis.	85
Figura 5.7 – Entropias dos ensembles microcanônico e canônico dos ataques aleatórios em 100 nós da rede Senhor dos Anéis.	86
Figura 5.8 – Entropias S_0 e S_1 referentes aos ataques aos vértices com maiores(menores) centralidades de intermediação da rede Senhor dos Anéis.	87
Figura 5.9 – Entropias S_0 e S_1 referentes aos ataques nos vértices com maiores centralidades de intermediação.	89
Figura 5.10–Entropias S_0 e S_1 relativas aos ataques dos vértices com maiores graus de entrada.	90
Figura 5.11–Entropias S_0 e S_1 nos ataques aos vértices com maiores graus de saída da rede da Tolerância Oral.	91
Figura 5.12–Entropias S_0 e S_1 da rede Tolerância Oral que sofreu ataques aleatórios nos vértices.	92
Figura 5.13–Rede do Clube Karatê com quatro comunidades construída no igraph.	94
Figura 5.14–Rede da Odisseia com 4 comunidades.	95
Figura 5.15–Rede literária Senhor dos anéis com 5 comunidades.	95
Figura 5.16–Entropia nas cinco comunidades da rede O senhor dos Anéis, destacando-se a comunidade 3 com maior entropia, pois é a mais heterogênea em ligações. .	97

LISTA DE QUADROS

Quadro 4.1 –Componentes imunológicos da rede de tolerância oral.	73
---	----

LISTA DE TABELAS

Tabela 1.1 – Exemplos de redes reais.	17
Tabela 5.1 – Entropias de ordem zero de redes não direcionadas. As redes assinaladas com * foram extraídas das tabelas de (BIANCONI 2008).	79
Tabela 5.2 – Entropias de ordem 1 de redes não direcionadas.	80
Tabela 5.3 – Variação das entropias S_0 e S_1 nos ataques aos maiores/menores graus da rede.	88
Tabela 5.4 – Variação das entropias S_0 e S_1 nos ataques aleatórios aos vértices da rede do Senhor dos anéis e da Odisseia.	88
Tabela 5.5 – Variação das entropias S_0 e S_1 nos ataques aos vértices de maiores/menores centralidades de intermediação na sequência decrescente (crescente).	88
Tabela 5.6 – Entropias S_0^{dir} e S_1^{dir} nos ataques aos vértices com maiores centralidades de intermediação.	89
Tabela 5.7 – Entropias S_0 e S_1 nos ataques aos vértices com maiores graus de entrada. ..	90
Tabela 5.8 – Entropias de redes com ataques nos vértices com maiores graus de saída na rede da tolerância oral	91
Tabela 5.9 – Entropias S_0 e S_1 da rede da Tolerância Oral com ataques aleatórios de dez vértices.	92
Tabela 5.10–Entropias S_c com valores das modularidades para rede Zachary modulada em duas, três e quatro comunidades.	94
Tabela 5.11–Entropias S_c da rede literária Odisseia modulada de duas a quatro comunidades.....	95
Tabela 5.12–Entropia S_c da rede literária O Senhor dos Anéis modulada de duas até cinco comunidades com os respectivos valores das modularidades.	96
Tabela 5.13–Propriedades estruturais das cinco comunidades da rede social Senhor dos anéis.	96

LISTA DE ABREVIATURAS E SIGLAS

MHC II – Complexo de histocompatibilidade do tipo II

RA – Ácido retinóico

iTREGFOXP3+ – Células T regulatórias induzidos do tipo FOXP3 selecionadas positivamente

TGF- β – fator de crescimento e expansão celular do tipo β

IL-2 – Interleucina do tipo 2

IL-10 – Interleucina do tipo 10

IL-27 – Interleucina do tipo 27

CD3+ - Células dendríticas do tipo 3 selecionadas positivamente

TH3 – Células auxiliares do tipo 3

INF-lambda – Interferon lambda

CCR7 – Proteína integral de membrana codificada pelo gene CCR7

FcRn – Receptor de membrana neonatal do tipo Fc

$\alpha 4\beta 7$ integrina – Elemento protein integral de membrane do tipo $\alpha 4\beta 7$

MadCAM-1 – Molécula de adesina musosal vascular de adesão celular do tipo 1

CD11b+ – Célula dendríticas do tipo 11b selecionadas positivamente

CCL25 – Citosina do tipo 25

CD11c+ – Células dentríticas do tipo 11c selecionadas positivamente

LISTA DE SÍMBOLOS

G	Grafo
N	Número de nós, vértices ou sítios
L	Número de arestas, ligações ou conexões
A	Matriz de adjacência
a_{ij}	Elemento da matriz de adjacência
w	Peso das conexões
w_{ij}	Elemento da matriz dos pesos W
k_i	Grau do vértice i
$\langle k \rangle$	Grau médio
$\langle k^2 \rangle$	Grau quadrado médio
$P(k)$	Distribuição de conectividade
ρ	Densidade do grafo
$\langle s \rangle$	Força média da rede
D	Matriz das distâncias
k_i^{in}	Grau de entrada
k_i^{out}	Grau de saída
$k_{nn,i}$	Grau do vizinho mais próximo
$\langle k_{nn,i} \rangle$	Grau médio do vizinho mais próximo
$\varepsilon(G)$	Energia do grafo
$L(G)$	Matriz laplaciana
E	Eficiência do grafo
$\langle L \rangle$	Número médio de arestas
H	Entropia de Shannon
$\langle x_i \rangle$	Valor médio de uma medida
F	Energia livre de Helmholtz
T	Temperatura
$S(p)$	Entropia da distribuição de graus
U	Energia interna
Z	Função de partição
$\mathcal{G}$	Ensemble de grafos G
$\mathcal{N}(G)$	Número de grafos G
$G(N,L)$	Conjunto de grafos com número de nós e de arestas fixados
h_{ij}	Campo auxiliar
$n!$	Fatorial de n
$n!!$	Duplo fatorial de n
S_0	Entropia com aproximação de ordem zero
S_1	Entropia com aproximação de ordem um

S_2	Entropia com aproximação de ordem 2
S_0^{dir}	Entropia com aproximação de ordem zero de rede direcionada
S_1^{dir}	Entropia com aproximação de ordem um de rede direcionada
Z_0	Função de partição de ordem zero
Z_1	Função de partição de ordem um
Z_1^{dir}	Função de partição de ordem um de rede direcionada
Z_2	Função de partição de ordem dois
$\delta(S_i^0)$	Índice de variação das entropias de ordem zero

SUMÁRIO

1	INTRODUÇÃO.....	16
2	CONCEITOS BÁSICOS.....	20
2.1	Grafos	20
2.1.1	Distâncias	23
2.2	Medidas de redes	24
2.2.1	Grau	25
2.2.2	Distribuição das conexões	28
2.2.3	Assortatividade	29
2.3	Energia de um grafo.....	30
2.4	Matriz Laplaciana	32
2.5	Modelos de redes complexas	33
2.5.1	Modelos $G(N,p)$	33
2.5.2	Redes <i>small world</i>	35
2.5.3	Redes livre de escala	37
2.5.3.1	Divergência em momentos de maiores ordens	40
2.6	Redes aleatórias generalizadas (modelo de configuração)	40
2.6.1	Probabilidade de uma aresta	41
2.6.1.1	Distribuição de graus.	42
3	ENTROPIA EM REDES COMPLEXAS.....	43
3.1	Considerações iniciais.....	43
3.1.1	Entropia de Shannon	43
3.1.2	Entropia de Gibbs de uma rede aleatória com uma sequência de graus.....	45
3.2	Entropia de redes modelo $G(N,p)$	45
3.3	Entropia em redes livre de escala	47
3.4	Grafo aleatório exponencial	49
3.5	<i>Ensembles</i>	55
3.5.1	<i>Ensemble</i> microcanônico de redes complexas.....	56
3.5.2	<i>Ensemble</i> canônico.....	56
3.5.3	<i>Ensemble</i> macrocanônico.....	56
3.6	Entropia estrutural de redes complexas.....	57
3.6.1	Entropia em <i>ensembles</i> de redes não direcionadas com aproximação de ordem zero	57
3.6.2	Entropia em <i>ensembles</i> de redes não direcionadas com aproximação de ordem um	60
3.6.3	Entropia em <i>ensembles</i> de redes não direcionadas com estrutura de comunidade (aproximação de ordem 2)	63
3.7	Entropia em <i>ensembles</i> de redes direcionadas.....	66
3.7.1	Entropia de <i>ensembles</i> de redes direcionadas com aproximação de ordem zero....	66
3.7.2	Entropia em <i>ensembles</i> de redes direcionadas com aproximação de ordem um	66
4	REDE LITERÁRIA SOCIAL DA ODISSEIA, SENHOR DOS ANÉIS E REDE IMUNOLÓGICA.....	69
4.1	Rede literária da Odisseia	69
4.2	Rede literária do Senhor dos Anéis.....	70
4.3	Rede direcionada da tolerância oral	72
4.4	Robustez de redes complexas	74

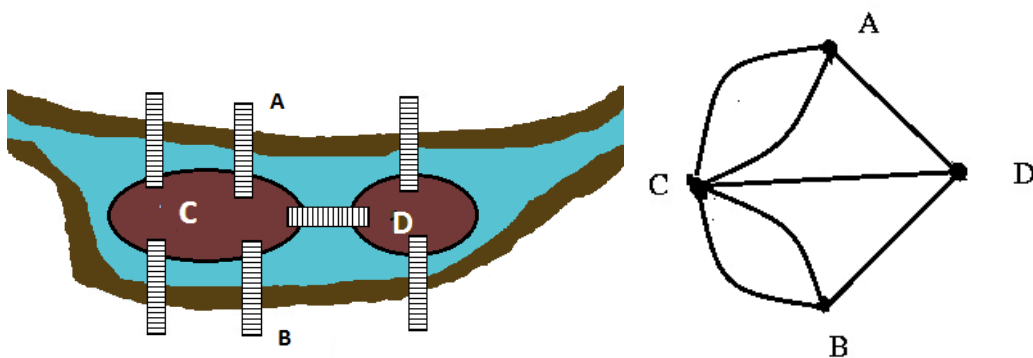
5	ENTROPIA DE REDES - APLICAÇÕES	76
5.1	Ensembles de redes complexas	76
5.2	Exemplos de entropias S_0 em redes não direcionadas	77
5.3	Exemplos de entropias de <i>ensembles</i> com ordem de aproximação 1: S_1	79
5.3.1	Falhas e ataques nas redes complexas	80
5.3.1.1	Rede social literária da obra Odisseia de Homero	81
5.3.1.2	Rede literária Senhor dos anéis	85
5.3.2	Entropia S_0 de redes direcionadas	88
5.4	Entropias S_c de redes com estrutura de comunidade	93
6	CONCLUSÃO	98
	REFERÊNCIAS	100
	APÊNDICE A – MÉTODO DO PONTO DE SELA	103

1 INTRODUÇÃO

A interdisciplinaridade das redes complexas abrange diversas áreas de conhecimento, tais como a Física, Matemática, Ciência da Computação, Engenharia, Biológicas, Sociológica, Linguística e outras. O que é uma rede complexa? Em termos gerais, uma rede é qualquer sistema que admite uma representação por um grafo, cujos vértices identificam os elementos do sistema e o conjunto de ligações representa a interação entre esses elementos. Diferente dos grafos, redes complexas são construídas através de elementos fornecidos pela modelagem de sistemas reais e sua composição estrutural tende a ser não trivial.

Grafo é uma representação das relações existentes entre pares de elementos de um determinado conjunto, os quais são denominados de nós ou vértices e as conexões entre eles, de arestas ou ligações (*links*). No estudo das propriedades topológicas de redes utiliza-se a teoria dos grafos, que historicamente, tem sua concepção inicial atribuída a Leonhard Euler (1707-1783) que analisou o problema das sete pontes da cidade prussiana de Königsberg (atual Kaliningrado na Rússia) em seu trabalho publicado em 1741 (MONTEIRO (2010)). Este problema consistia na possibilidade de alguém atravessar todas as pontes sem repetir nenhuma. Em 1736, Euler provou que não existia tal caminho (euleriano). Inicialmente, Euler percebeu que os comprimentos das pontes não importavam e sim as porções de terra que se conectavam. Uma representação gráfica do local adequada para solucionar esta questão é o grafo mostrado na Figura 1.1, onde as ilhas são os nós C e D e as margens do rio A e B. As pontes são as arestas entre os nós. O problema das sete pontes foi modelado por um grafo, transformando caminhos em arestas e suas interseções em vértices, criando possivelmente o primeiro grafo da História, MONTEIRO (2010) conforme figura abaixo:

Figura 1.1 – Representações do problema das setes pontes de Königsberg e seu grafo.



Fonte: O autor.

Euler foi um matemático suíço que fez carreira em Berlim e São Petersburgo, exerceu extraordinária influência sobre vários campos da matemática, física e engenharia.

Em geral as redes complexas são classificadas em quatro grupos: redes sociais, redes tecnológicas, redes biológicas e redes de informação. A rede social é caracterizada por um

grupo de pessoas com um padrão de contato ou relacionamento entre si. As pessoas são representadas por vértices e as interações entre elas por arestas. As redes tecnológicas são aquelas construídas pelo homem como as redes elétricas, redes de telecomunicações, malhas rodoviárias ou malhas ferroviárias. A internet é uma rede tecnológica muito estudada, em virtude da facilidade na obtenção de dados. As redes complexas biológicas constituem poderosa abstração matemática que permite a representação de grandes volumes de dados e a posterior formulação de hipóteses biológicas. Trabalhos sobre as propriedades estatísticas das redes metabólicas foram estudados por H. Jeong e Barabási (JEONG et al. (2000)). Tem-se como exemplo, as redes de interação Proteína-Proteína, redes de genes, redes do tipo “*Food Web*” e redes Neurais.

As redes de informação ou redes de conhecimento, formam a modalidade de redes complexas descritas por um conjunto de vértices, representando uma entidade que contém informações, e um conjunto de arestas que conectam esses elementos, caso haja alguma relação entre a informação contida nos dois vértices nos quais a aresta incide. Por exemplo, rede de citação (orientada) entre artigos acadêmicos, redes de palavras ou rede de coautoria de artigos. Um outro exemplo clássico de rede complexa que representa o fluxo de informações é o da *World Wide Web*, a rede de páginas acessíveis pela *Internet*. Enquanto as páginas são traduzidas em vértices, as arestas denotam os *hyperlinks* entre duas páginas da *web*.

A Tabela 1.1 apresenta alguns exemplos de redes complexas com seus nós (vértices) e arestas ou conexões:

Tabela 1.1 – Exemplos de redes reais.

Rede complexa	vértices	arestas
cadeia alimentar	espécies	interações presa-predador
metabólica	moléculas	reações químicas
neural	neurônios	sinapses
sistema financeiro	bancos	transações financeiras
citações	artigos	citações
de sinônimos	palavras	significados similares
<i>world wide web</i> (www)	páginas “web”	<i>hyperlinks</i>
árvore genealógica	parentes	relações de parentesco
de amizade	amigos	relacionamento
de contatos	amigos, parentes	mensagens
internet	computadores, roteadores	cabos
de energia elétrica	geradores, transformadores	linhas de transmissão
de telecomunicações	telefones, centrais telefônicas	cabos
de transportes	rodoviárias, aeroportos, estações de trem	rodovias, linhas aéreas, ferrovias

Fonte: MONTEIRO (2010).

Neste trabalho procurou-se utilizar o conceito de redes aleatórias para descrever conjuntos ou *ensembles* de redes não direcionadas e direcionadas, com vínculos estruturais mediante de uma abordagem da mecânica estatística. Estes conjuntos são caracterizados pela sua entropia

que avalia a sua cardinalidade, ou seja, o número de redes do *ensemble*.

No *ensemble* microcanônico, também denotado por (G, N, L) , a rede complexa representada por sua matriz de adjacência A_{ij} de tamanho $N \times N$ tem o número de ligações L fixado. No *ensemble* canônico, cada rede é construída por uma sequência de graus, obedecendo o modelo das configurações, definido no capítulo 2. Quando a rede está modularizada, a entropia é calculada obedecendo os vínculos estruturais da sequência de graus e da estrutura de comunidades, obtida através de diversos métodos, fazendo com que a rede fique dividida em vários módulos.

Este formalismo foi proposto na referência (BIANCONI (2008), BIANCONI (2009)) e fornece as funções de partição Z_0 , com aproximação de ordem zero (microcanônico), Z_1 , de ordem um (canônico) e Z_2 ou Z_c de ordem dois ou com estrutura de comunidade, que são módulos organizacionais das redes, definidas por vértices altamente conectados entre si e pouco conectados com o restante da rede.

Os *ensembles* de redes aleatórias são caracterizados pela sua entropia, ou seja pelo logaritmo normalizado do número de redes que constituem o *ensemble* conforme BIANCONI (2008), que propõe a entropia como um indicador para avaliar o papel de cada característica estrutural numa rede real.

Para caracterizar a complexidade de uma rede real considera-se uma série de redes aleatórias que contém algumas características das redes reais. Particularmente considera-se redes com uma dada sequência de graus, correlação de graus com uma dada estrutura de comunidades.

Apresenta-se neste trabalho, uma análise da relação entre estrutura e robustez de redes complexas mediante remoção de vértices com maiores (menores) graus, maiores (menores) *betweenness centrality* e remoção aleatória de vértices. Para aplicação deste estudo, foram utilizadas bases de dados das redes literárias sociais da Odisseia (MIRANDA (2014)), do livro do Senhor dos anéis (RIBEIRO et al. (2015)) e da rede tolerância oral (MIRANDA et al. (2013)). Através de falhas e ataques intencionais, foi estudada a robustez destas redes segundo a dinâmica de remoção de vértices. Assim a variação da estrutura das redes em virtude deste procedimento foi quantificada pelas entropias estruturais com ordem de aproximação zero e aproximação um.

A tese é composta por seis capítulos. O segundo capítulo abrange conceitos básicos das redes complexas com suas medidas bem como a história do surgimento da teoria dos grafos e tipos de redes aleatórias. O terceiro capítulo apresenta conceitos da entropia da Mecânica Estatística, das redes aleatórias e entropias estruturais dos *ensembles* micro canônico e canônico das redes complexas com aproximação de ordem zero, um e de *ensembles* com estrutura de comunidades, ilustradas com diversos exemplos. O quarto capítulo apresenta a descrição das redes reais literárias sociais da obra da Odisseia de Homero e do Senhor dos anéis de J.R.R. Tolkien e a rede da tolerância oral com suas estruturas.

Na sequência, o capítulo cinco mostra a aplicação e análise da entropia em *ensembles* de

ordem zero, um e com estrutura de comunidades, bem como ataques às redes da Odisseia, do Senhor dos anéis e da rede biológica da tolerância oral. Para isso, analisou-se a robustez dessas redes com auxílio da entropia estrutural de ordem zero e de ordem um. Foi criado um índice representativo da variação das entropias em cada sequência de ataques aleatórios na rede da Odisseia possibilitando estabelecer uma correlação com o menor caminho médio da rede. O último capítulo (capítulo seis) apresenta as considerações finais e as possibilidades de trabalhos futuros.

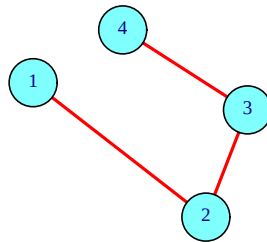
Esta tese resultou no artigo publicado pela Revista Brasileira de Ensino de Física em abril do corrente ano, denominado de Rede social complexa de O Senhor dos Anéis (referência RIBEIRO et al. (2015)).

2 CONCEITOS BÁSICOS

2.1 Grafos

Visando facilitar a compreensão do texto e das seções seguintes, são colocadas algumas definições sobre a teoria dos grafos. Um grafo G é definido por um par ordenado de conjuntos V e E , ou seja $G = (V, E)$, onde $V \neq \emptyset$ é um conjunto finito representando o conjunto de vértices ou nós e $E = \{l_{i,j}, i, j \in V\}$, o conjunto das arestas, conexões ou ligações que relacionam dois vértices do grafo, podendo possuir ou não um custo (peso) associado. O número de vértices N é denominado de ordem do grafo, $|V| = N$. O tamanho L do grafo é representado pelo número de arestas: $|E| = L$. Por exemplo, $V = \{1, 2, 3, 4\}$, $E = \{(1, 2), (2, 3), (3, 4)\}$. O grafo correspondente está representado na Figura 2.1:

Figura 2.1 – Grafo não orientado com 4 nós e 3 arestas.



Fonte: O autor.

O tamanho deste grafo não orientado é $L = 3$ e a ordem é $N = 4$. Um grafo é não orientado quando a presença da aresta que parte do vértice i e chega ao vértice j implica na existência da mesma aresta que parte de j e chega em i e portanto o par (i, j) é não ordenado. Num grafo direcionado ou orientado, essa implicação não existe, isto é, o par (i, j) é ordenado e as arestas são representadas por uma seta. Os nós i e j são denominados adjacentes ou vizinhos, se existe uma aresta entre eles. Este mesmo exemplo pode ser utilizado para representar um grafo direcionado conforme a Figura 2.2.

Para um um grafo não direcionado com N nós, o número máximo de arestas é

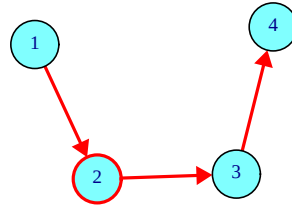
$$\binom{N}{2} = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2}. \quad (2.1)$$

Enquanto que para grafos direcionados, o número máximo é o dobro desse, ou seja,

$$2 \times \binom{N}{2} = N(N-1) \quad (2.2)$$

arestas. Rede de participação em filmes é um exemplo de grafo não direcionado, enquanto que rede de citação de textos científicos é uma rede direcionada.

Figura 2.2 – Representação de um grafo direcionado com 4 nós e 3 arestas.



Fonte: O autor.

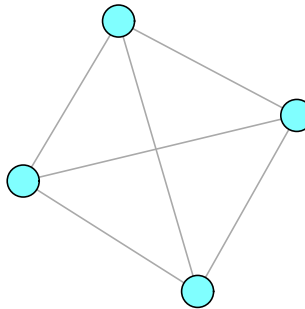
A densidade ρ de um grafo é a razão entre o número L de arestas existentes e o número máximo de arestas:

$$\rho = \frac{2L}{N(N-1)} \quad (\text{grafos não direcionados}), \quad (2.3)$$

$$\rho = \frac{L}{N(N-1)} \quad (\text{grafos direcionados}). \quad (2.4)$$

Um grafo é considerado denso se $\rho \simeq 1$ e esparso se $0 < \rho \ll 1$. O grafo da Figura 2.2 tem densidade igual a 0,25 (esparso) e o da Figura 2.3 tem $\rho = 1$ (denso). Um grafo completo K_N é aquele em que existe uma aresta conectando cada par de vértices (Figura 2.3), ou seja é um grafo constituído do número máximo de arestas ou de todas as arestas possíveis.

Figura 2.3 – Grafo completo K_4 com 4 nós e 6 arestas.



Fonte: O autor.

Grafo simples (Figura 2.3) é aquele que não possui laços (auto conexões) nem arestas múltiplas. Um grafo $G(V, E)$ pode ser representado por sua matriz de adjacência $A = N \times N$ de N nós, cujos elementos a_{ij} são definidos por:

$$a_{ij} = \begin{cases} 1, & \text{se existe uma aresta ligando vértices } i \text{ a } j, \\ 0, & \text{caso contrário.} \end{cases}$$

Se o grafo é não direcionado, então $a_{ij} = a_{ji}$ e a matriz A é simétrica, indicando que não se diferencia o sentido da ligação. Para a rede direcionada tem-se que $a_{ij} \neq a_{ji}$, informando

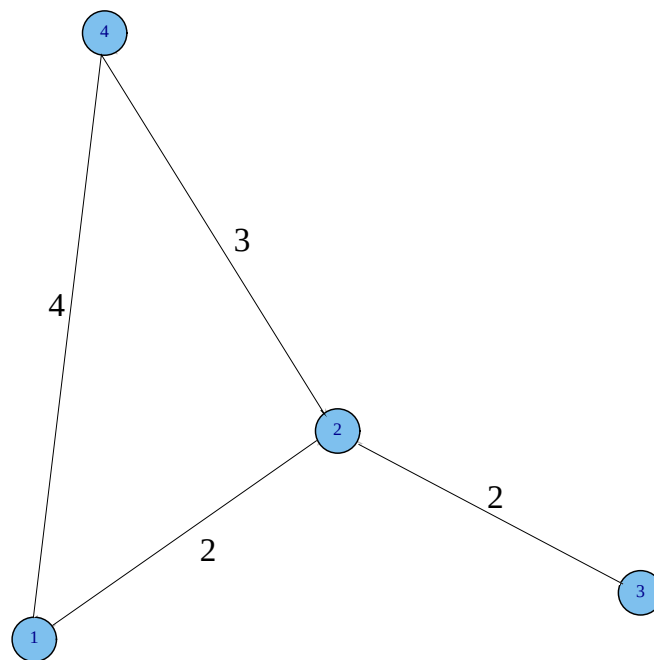
que o sentido da ligação é importante, além disso a matriz de adjacência A é não simétrica (assimétrica). A matriz de adjacência do grafo dado na Figura 2.1 é dada por:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Nota-se que esta matriz apresenta sua diagonal principal com elementos nulos, indicando que não existe autoconexão dos vértices.

Quando atribuem-se pesos $w(a_{ij})$ às arestas, a rede é denominada de rede ponderada ou valorada (Figura 2.4), sua matriz associada é a matriz de valores ou dos pesos W , matriz que não é constituída somente por zeros ou uns. Como exemplo, o grafo representado pela Figura 2.4 é um grafo ponderado.

Figura 2.4 – Representação de um grafo ponderado com pesos nas arestas.



Fonte: O autor.

A matriz dos pesos W deste grafo é dada por:

$$W = \begin{pmatrix} 0 & 2 & 0 & 4 \\ 2 & 0 & 2 & 3 \\ 0 & 2 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{pmatrix}.$$

Os elementos de W representam os pesos das ligações e podem ser entendidos como a proporção de força da ligação, por exemplo, a conexão do vértice 1 ao vértice 4 tem peso 4 e está representada por um peso duas vezes maior que a ligação do vértice 2 ao vértice 3.

2.1.1 Distâncias

Uma questão importante na estrutura de grafos é a acessibilidade dos vértices, que é definida como a possibilidade de ir de um vértice a outro por meio das conexões da rede. Uma família de ligações sucessivamente adjacentes é denominada caminho ou percurso. Num percurso fechado, também denominado de ciclo ou *loop*, a última ligação da sucessão é adjacente à primeira.

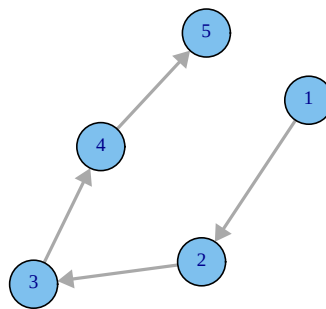
A definição de caminho leva ao conceito de distância entre dois vértices. O número de arestas l_{ij} que compõem o caminho de menor comprimento entre os nós i e j é denominado geodésica ou distância (d_{ij}) entre esses nós. Caso não exista tal caminho, convencionou-se que $d_{ij} = \infty$, sendo que esta expressão não é computada no cálculo do comprimento médio $\langle d \rangle$ para grafos não direcionados bem como para direcionados.

Os valores de d_{ij} formam a matriz de distâncias D que é simétrica para redes não direcionadas. No grafo direcionado da Figura 2.2, tem-se

$$D = \begin{pmatrix} 0 & 1 & 2 & 3 \\ \infty & 0 & 1 & 2 \\ \infty & \infty & 0 & 1 \\ \infty & \infty & \infty & 0 \end{pmatrix}.$$

O valor máximo de d_{ij} recebe a denominação de diâmetro do grafo (é o mais longo caminho mínimo). Para o grafo da Figura 2.5 o diâmetro é $d = 4$.

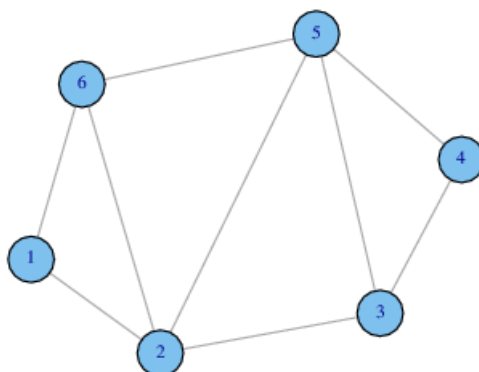
Figura 2.5 – Diâmetro de um grafo direcionado de 5 nós e 4 arestas.



Fonte: O autor.

O percurso aberto ou fechado que utiliza cada ligação do grafo uma única vez, é denominado caminho Euleriano. Na Figura 2.6 é representado por 3-4-5-3-2-5-6-2-1-6.

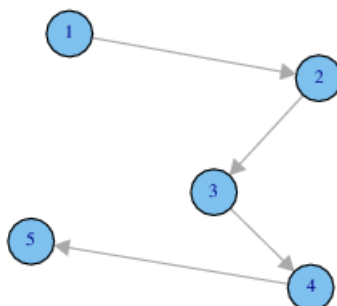
Figura 2.6 – Caminho Euleriano de um grafo não direcionado. Cada aresta é utilizada uma única vez.



Fonte: O autor.

Enquanto que o percurso que utiliza cada vértice uma única vez, denomina-se Hamiltoniano, representado na Figura 2.7 pelo percurso 1-2-3-4-5.

Figura 2.7 – Caminho Hamiltoniano 1-2-3-4-5 num grafo direcionado. Cada vértice é utilizado uma única vez.



Fonte: O autor.

2.2 Medidas de redes

As redes podem ser classificadas numa mesma categoria, se suas medidas estruturais fornecem valores estatisticamente semelhantes. Assim a escolha das medidas torna-se muito importante, uma vez que a utilização de um número reduzido de medidas podem fornecer classificações errôneas. Deste modo, quanto mais características topológicas um dado modelo é capaz de satisfazer, mais completo este se torna. A cardinalidade do conjunto N de nós e a cardinalidade do conjunto E de arestas constituem dois exemplos elementares de medidas. Devido à diversidade de aspectos locais e globais as medidas se tornam úteis e necessárias para a caracterização topológica de redes complexas. Algumas medidas de redes importantes serão consideradas neste texto e serão relacionadas a seguir.

2.2.1 Grau

O grau de um nó i é o número de arestas que concorrem em i . É representado por k_i e calculado a partir dos elementos da matriz de adjacência A de um grafo não direcionado, por:

$$k_i = \sum_{j=1}^N a_{ij} = \sum_{j=1}^N a_{ji}. \quad (2.5)$$

Propriedade (segundo Euler):

$$\sum_{i=1}^N k_i = 2L. \quad (2.6)$$

Em palavras, a soma dos graus de todos os nós de uma rede é o dobro do número de arestas L . Como o grau de um nó é uma medida local, caracteriza-se a estrutura da rede não direcionada mediante o grau médio ou conectividade média $\langle k \rangle$ que é uma medida global, dado por:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2L}{N}. \quad (2.7)$$

Se o grafo for regular, todos os vértices tem o mesmo grau e portanto, o grau médio $\langle k \rangle = k$.

Para um grafo direcionado tem-se que o grau de entrada k_i^{in} recebe arestas convergentes e o grau de saída k_i^{out} é formada por arestas divergentes e são obtidos por:

$$k_i^{in} = \sum_{j=1}^N a_{ij}, \quad k_j^{out} = \sum_{i=1}^N a_{ij}. \quad (2.8)$$

Denomina-se sumidouro, o vértice que possui somente grau de entrada e fonte ou origem, o nó que possui somente grau de saída. Para evitar nós isolados, assume-se que $k_i^{in} + k_i^{out} > 0$, para todo $i = 1, 2, \dots, N$. O grau de um nó i é $k_i = k_i^{in} + k_i^{out}$, onde $k_i^{in} = \sum_{j=1}^N a_{ji}$ e $k_i^{out} = \sum_{j=1}^N a_{ij}$.

A densidade da rede em função do grau médio, é dada por

$$\rho = \frac{\langle k \rangle}{N - 1}. \quad (2.9)$$

No caso de redes ponderadas, o grau do vértice s_i é dado pela soma dos pesos das arestas e é denominado intensidade ou força do vértice.

Assim,

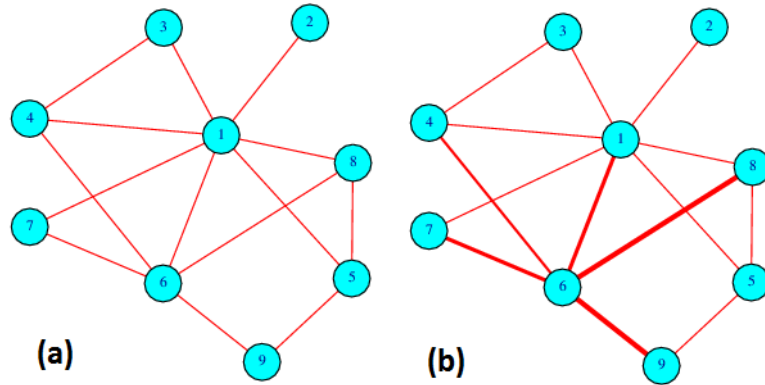
$$s_i = \sum_{j=1}^N w_{ij}, \quad (2.10)$$

onde w_{ij} é um elemento da matriz dos W e a respectiva medida global é a força média da rede:

$$\langle s \rangle = \frac{1}{N} \sum_{i=1}^N s_i. \quad (2.11)$$

Estas medidas podem ser utilizadas para identificar *hubs* que são vértices altamente conectados ou que possuem maior intensidade (ALBERT; BARABÁSI, 2002).

Figura 2.8 – Grafos com vértices altamente conectados numa rede simples (a) e numa rede ponderada (b), com os pesos das arestas representados por linhas mais espessas.



Fonte: O autor.

Na Figura 2.8 (a), o *hub* é o vértice 1 que possui o maior grau da rede, $k = 7$ enquanto que o vértice 6 é o *hub* da Figura 2.8 (b), pois é o vértice com maior intensidade (o peso das arestas está representado por uma aresta mais espessa).

Os *hubs* são muito importantes na formação da estrutura de redes complexas, pois sua retirada pode causar fragmentação da rede, resultando em componentes não conectados ou vértices isolados. Este procedimento está relacionado com falhas aleatórias e ataques em redes que dependendo da aplicação pode-se obter redes robustas ou vulneráveis. A internet é um exemplo de robustez pois interrompendo um ou mais roteadores ela ainda continua funcionando. Como exemplo de vulnerabilidade pode-se ter uma rede formada por pessoas de uma comunidade em que há propagação de epidemia, sendo desejável que a parte infectada seja isolada (vértices removidos) do restante da rede ((BARBIERI, 2010)).

O coeficiente de agregação (ou clusterização) C_i do vértice i é a razão entre o número e_i de ligações que existem entre os vizinhos de i e o número máximo de arestas possíveis entre esses vizinhos. Para um grafo não direcionado tem-se que:

$$C_i = \frac{e_i}{\binom{k_i}{2}} = \frac{2e_i}{k_i(k_i - 1)}, \quad \text{para } k_i \geq 2. \quad (2.12)$$

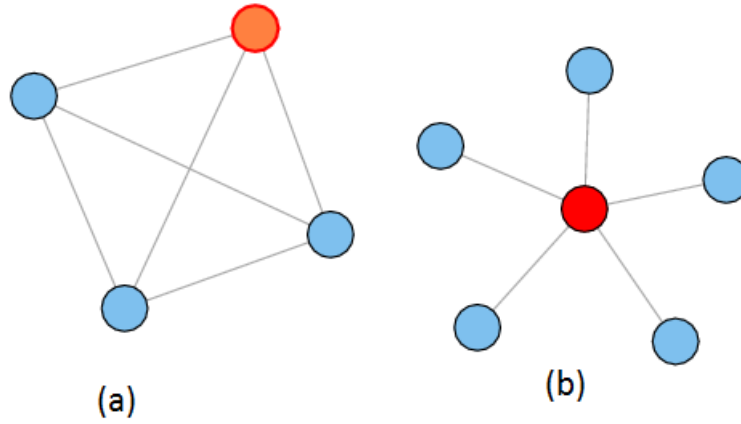
O valor de e_i é obtido da matriz de adjacência A , através da expressão:

$$e_i = \frac{1}{2} \sum_{j=1}^N \sum_{l=1}^N a_{ij} a_{jl} a_{li}. \quad (2.13)$$

Para $k_i = 0$ ou 1, define-se $C_i = 0$. Nas redes sociais, esta propriedade é denominada de aglomeração ou transitividade, indicando a probabilidade de dois amigos quaisquer terem um outro amigo em comum.

Na Figura 2.9 (a) o coeficiente de aglomeração para o vértice em vermelho é $C_i = 1$ e para a Figura 2.9 (b), o coeficiente é $C_i = 0$. Para quantificar a ocorrência de triângulos na rede,

Figura 2.9 – Ilustração de dois diferentes valores para o coeficiente de aglomeração para o vértice em vermelho.



Fonte: O autor.

utiliza-se outra medida global, fornecida pela média do coeficiente de aglomeração $\langle C \rangle$ para todos os vértices, calculado pela expressão:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i. \quad (2.14)$$

Outra medida global de grande importância é o comprimento do menor caminho médio $\langle l \rangle$ representado pelo valor médio de l_{ij} , levando em conta todos os possíveis pares de vértices i e j , sendo calculado por

$$\langle l \rangle = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N l_{ij}, \quad (2.15)$$

para redes não direcionadas. E para grafos direcionados, por:

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N l_{ij}. \quad (2.16)$$

A medida local, denominada centralidade de proximidade (*closeness centrality*) $C_p(i)$ do nó i é inversamente proporcional ao somatório do comprimento dos caminhos mais curtos entre esse nó e os outros $N - 1$ nós que constituem a rede. Logo,

$$C_p(i) = \frac{N-1}{\sum_{j=1}^N l_{ij}}. \quad (2.17)$$

Assim, quanto menores forem as distâncias entre o nó i e os demais, mais central será esse nó.

A centralidade de intermediação (*betweenness centrality*) $C_{int}(i)$, do nó i é o número $\sigma_{jl}(i)$ de caminhos mais curtos entre os nós j e l que passam por i , dividido pelo número σ_{jl} de

caminhos mais curtos entre os nós j e l , considerando todos os possíveis pares de nós j e l , com $j \neq i$ e $l \neq i$. Para normalizar, este coeficiente deve-se dividi-lo pelo número máximo de vezes que o nó i pode aparecer no caminho mais curto entre dois nós quaisquer (excluindo o nó i). Para um grafo não direcionado, o número máximo vale $\binom{N-1}{2}$, portanto:

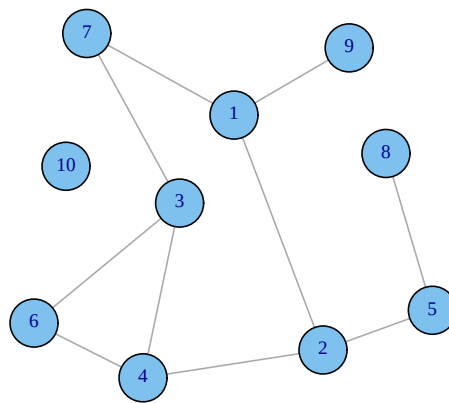
$$C_{int}(i) = \frac{2}{(N-1)(N-2)} \sum_{j=1}^{N-1} \sum_{l=j+1}^N \frac{\sigma_{jl}(i)}{\sigma_{jl}}, \quad j \neq l \neq i. \quad (2.18)$$

Em 1972, Bonacichi((MONTEIRO, 2010)), sugeriu que a centralidade de um nó fosse proporcional à soma das centralidades dos nós aos quais ele está conectado. Desse modo, se v_i é centralidade do nó i , então $v_i = \frac{1}{\lambda} \sum_{j=1}^N a_{ij} v_j \Rightarrow \sum_{j=1}^N a_{ij} v_j = \lambda v_i$ sendo $\frac{1}{\lambda}$, a constante de proporcionalidade. Portanto λ é o autovalor e v é o autovetor da matriz de adjacência A .

2.2.2 Distribuição das conexões

A distribuição de graus $P(k) = p_k = \frac{N_k}{N}$ indica a fração de nós com grau k , em que N_k representa o número de nós com grau k . É interpretada como sendo a probabilidade de que um nó selecionado aleatoriamente possua grau k . As quantidades p_k representam a distribuição de graus da rede. Por exemplo, um grafo com 10 nós conforme Figura 2.10 em que um vértice tem grau zero, dois tem grau um, quatro com grau dois, dois tem grau três e um tem grau quatro. Então os valores de p_k para $k = 0, 1, 2, 3, 4$, são: $p_0 = \frac{1}{10}$, $p_1 = \frac{2}{10}$, $p_2 = \frac{3}{10}$, $p_3 = \frac{2}{10}$, $p_4 = \frac{1}{10}$.

Figura 2.10 – Distribuição de graus num grafo não direcionado em função do número de conexões de cada vértice.



Fonte: O autor.

De maneira análoga define-se a distribuição de força $P(s) = p_s$ onde são consideradas as forças dos vértices. O grau médio pode ser obtido a partir de $P(k)$:

$$\langle k \rangle = \sum_{k_{\min}}^{k_{\max}} k P(k), \quad (2.19)$$

onde $k_{\min}$ e $k_{\max}$ são os valores mínimo e máximo de k encontrado no grafo.

2.2.3 Assortatividade

Grande parte das redes reais são correlacionadas no sentido de que a probabilidade de um vértice com grau k esteja conectado com outro vértice de grau k' depende de k (probabilidade condicional $P(k'|k)$). Esta probabilidade é dada por:

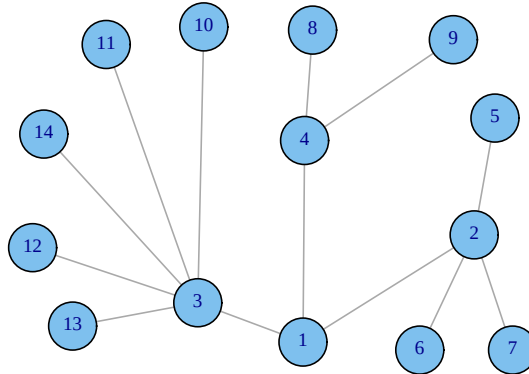
$$P(k' | k) = \frac{k' P(k')}{\langle k \rangle}. \quad (2.20)$$

Observa-se que $\sum_{k'} P(k' | k) = 1$ e para redes não direcionadas, $P(k' | k) = P(k | k')$ e $k' P(k | k') P(k') = k P(k' | k) P(k)$. O grau médio dos vizinhos mais próximos do nó i é dado por:

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in V(i)} k_j \quad (2.21)$$

Exemplo: na Figura 2.11, $k_1 = 3 \Rightarrow k_{nn,1} = (3 + 4 + 6)/3 = 4,3$

Figura 2.11 – Grau médio dos vizinhos do vértice 1, representados pelos vértices 2, 3 e 4.



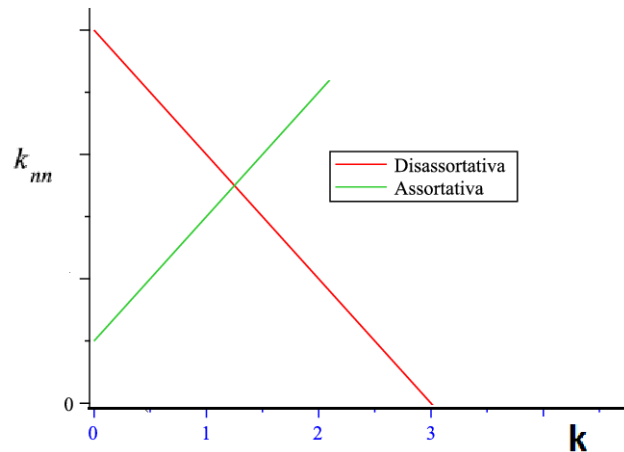
Fonte: O autor.

Usando a média dos vizinhos mais próximos de vértices com grau k , tem-se

$$\langle k_{nn} \rangle = \sum_{k'} k' P(k' | k). \quad (2.22)$$

Quando $\langle k_{nn} \rangle$ é uma função crescente em k , vértices com alto grau tendem a se conectarem entre si. Estas redes são classificadas como assortativas. Caso $\langle k_{nn} \rangle$ for decrescente em k , vértices com alto grau tendem a se conectarem com vértices de graus menores, as redes são classificadas como disassortativas.

Figura 2.12 – Gráfico de função assortativa e disassortativa.



Fonte: O autor.

Outra maneira de determinar a correlação de grau é considerando o coeficiente de correlação de Pearson nos extremos de uma aresta (NEWMAN, 2010)

$$s = \frac{\sum_{i,j} (a_{ij} - k_i k_j / 2L) k_i k_j}{\sum_{i,j} (k_i \delta_{ij} - k_i k_j / 2L) k_i k_j}, \quad (2.23)$$

onde a_{ij} é elemento da matriz de adjacência A , k_i e k_j são as conectividades dos vértices i e j , L , o número de arestas e δ_{ij} , é a função delta de Kronecker. Se $s > 0$ a rede é assortativa, se $s < 0$, disassortativa e para $s = 0$ não há correlação entre os graus dos vértices.

2.3 Energia de um grafo

Segundo a prof.^a Nair M. Abreu (ABREU et al., 2012), o físico-químico Erich Hückel estudou os hidrocarbonetos insaturados, onde a estrutura da molécula foi representada por um grafo, como mostra a Figura 2.13.

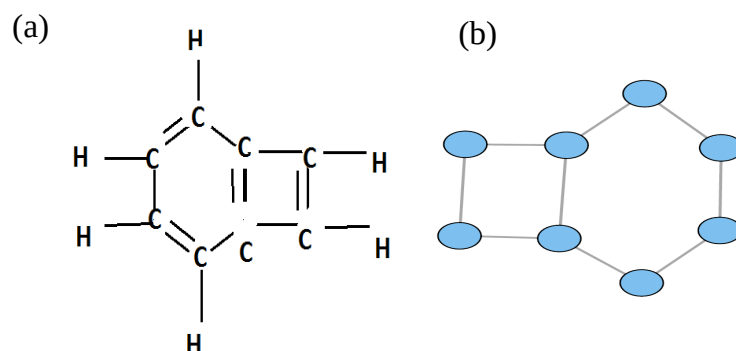
Os vértices do grafo da figura 2.13(b) correspondem aos átomos de carbono do hidrocarboneto representado na 2.13(a). Dois vértices são adjacentes se, e somente se, existe uma ligação entre os átomos de carbono correspondentes.

Em 1978, o professor de físico-química, Ivan Gutman (FRITSCHER, 2011) introduziu o conceito de energia de um grafo na representação de hidrocarbonetos. Porém essa definição não ficou subordinada a grafos moleculares, tendo sido estendida para qualquer grafo. Assim, a energia de um grafo G com N vértices é definida por:

$$\varepsilon(G) = \sum_{i=1}^N |\lambda_i|, \quad (2.24)$$

onde os λ_i são os autovalores da matriz de adjacência do grafo G . O cálculo da energia de um grafo depende da determinação do seu espectro. Desde o início do estudo deste conceito,

Figura 2.13 – Representação da estrutura de um hidrocarboneto ((a)) e o seu respectivo grafo ((b)).



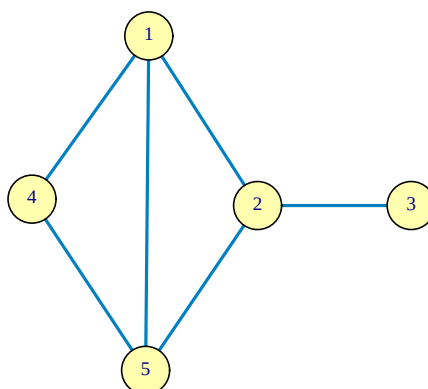
Fonte: Adaptada de (ABREU et al., 2012).

estimativas foram feitas para se determinar a magnitude desta grandeza, que resultaram no estabelecimento de limites superiores e inferiores para a mesma, de tal forma que uma das proposições para estes limitantes é a seguinte (ABREU et al., 2012): Se G é um grafo com n vértices e m arestas, então

$$\varepsilon(G) \leq \sqrt{2mn}. \quad (2.25)$$

Sendo A uma matriz simétrica, todos seus autovalores são reais e somam zero porque o $\text{tr}(A) = 0$. O polinômio característico de A do grafo G é calculado por $\text{pol}(\lambda) = \det(A - \lambda I)$. O espectro do grafo G é constituído por todos os autovalores com suas multiplicidades. O maior autovalor de G é denominado índice de G e denotado por $\text{ind}(G)$. Exemplo: calcular a energia de um grafo G com cinco vértices representado na Figura 2.14:

Figura 2.14 – Grafo simples de cinco nós e seis arestas com energia $\varepsilon = 6,7298$.



Fonte: O autor.

Matriz de adjacência do grafo da figura 2.14:

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

Seu polinômio característico é do quinto grau na variável λ : $\text{pol}(\lambda) = \lambda^5 - 6\lambda^3 - 4\lambda^2 + 3\lambda + 2$.

Seu espectro é $\text{spect}(G) = \begin{bmatrix} 2,6412 & 0,7237 & -0,5892 & -1 & -1,7757 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$. Nesta matriz os autovalores estão na primeira linha (cuja soma é igual ao traço da matriz de adjacência) e suas multiplicidades na segunda linha.

Logo $\text{ind}(G) = 2,6412$. A energia do grafo é $\varepsilon(G) = 6,7298$.

Se G é um grafo com N vértices e L arestas e seu polinômio característico $\text{pol}(\lambda) = \lambda^N + a_1\lambda^{N-1} + a_2\lambda^{N-2} + a_3\lambda^{N-3} + \dots + a_{N-1}\lambda + a_N$, então (FRITSCHER, 2011):

1. $a_1 = 0$
2. $a_2 = -L$ (número de arestas com sinal negativo)
3. $a_3 = -2t$, onde t é o número de triângulos do grafo.
4. Um grafo é bipartido se e somente se, possui espectro simétrico em relação a zero.

No exemplo acima, o grafo tem $-L = -6 \Rightarrow L = 6$ arestas e $-2t = -4 \Rightarrow t = 2$ triângulos.

2.4 Matriz Laplaciana

Seja D a matriz diagonal consistindo dos graus dos vértices de um grafo G , também conhecida como matriz de valência, com entradas $d_{ii} = k_i$ e A a matriz de adjacência. Então a matriz $L = D - A$ é definida como a matriz Laplaciana do grafo G . Por exemplo, a matriz Laplaciana do grafo da Figura 2.14, é:

$$L(G) = \begin{bmatrix} 3 & -1 & 0 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 2 & -1 \\ -1 & -1 & 0 & -1 & 3 \end{bmatrix}.$$

O polinômio característico Laplaciano de G é $P(\lambda) = \det(L(G) - \lambda I)$. O espectro laplaciano é o conjunto das raízes do polinômio Laplaciano, com suas multiplicidades. Usualmente são enumerados em ordem decrescente $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \lambda_N$. A matriz Laplaciana $L(G)$ é simétrica e por conseguinte (da Álgebra Linear) possui seus autovalores reais não negativos.

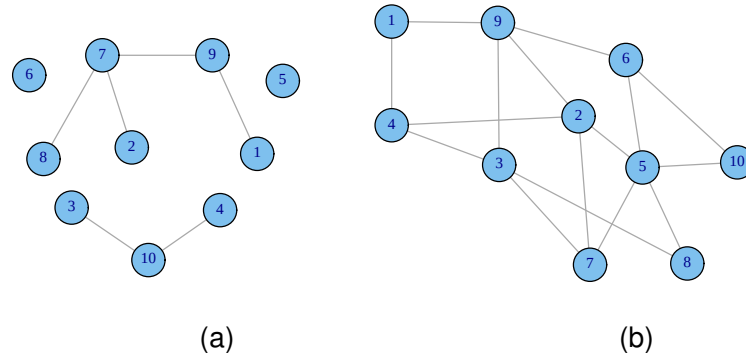
2.5 Modelos de redes complexas

2.5.1 Modelos $G(N,p)$

Em 1959, os matemáticos húngaros, Paul Erdős e Alfréd Rényi (BARBIERI, 2010), começaram a escrever artigos sobre grafos aleatórios, provando diversas características fundamentais da estrutura de um grafo gerado por estes modelos. O modelo $G(N, p)$, também é conhecido por modelo binomial ou por “modelo de rede aleatória”. Possui apenas dois parâmetros, N e p , em que N representa o número de nós do grafo rotulados de 1 a N e o parâmetro p determina a probabilidade de uma determinada aresta ser incluída na rede, independentemente das outras.

Vê-se que o modelo não determina a estrutura da rede, e sim o processo aleatório que irá gerar a estrutura da rede. Além disso, não define uma única rede, podendo levar a diferentes realizações (conjunto de redes possíveis com diferentes probabilidades). Tem-se como exemplo uma rede com 10 nós ($N = 10$) e probabilidades $p = 0,1$ e $p = 0,3$, respectivamente, conforme Figura 2.15.

Figura 2.15 – Modelos de redes aleatórias com 10 nós e probabilidades $p = 0,1$ e $p = 0,3$. $G(10,0,1)$ na figura (a) e $G(10,0,3)$ na figura (b).



Fonte: O autor.

O espaço amostral S deste modelo é dado pelo conjunto de redes que pode ser gerado. Cada aresta pode pertencer ou não à rede e o total de arestas possíveis da rede é $\binom{N}{2}$ que é utilizado no cálculo do tamanho do espaço amostral:

$$|S| = 2^{\binom{N}{2}}. \quad (2.26)$$

Assim se $N = 4 \Rightarrow |S| = 2^6 = 64$ redes do ensemble. A probabilidade p de uma aresta figurar na rede gerada pelo modelo é independente da aresta e a probabilidade de outra aresta não aparecer na rede é $1 - p$. Logo, a probabilidade de ter um grafo particular $G(N,p)$ é:

$$P(G(N, p)) = p^L (1 - p)^{\binom{N}{2} - L}, \quad (2.27)$$

ou seja, cada grafo $G(N,p)$ aparece com probabilidade $P(G(N,p))$. Se por exemplo, $N = 5$, $p = 0,25$, tem-se que a probabilidade do grafo $G(5,0,25)$ com cinco arestas é aproximadamente $2 \cdot 10^{-4}$. O número de grafos com exatamente L arestas é dado pelo número de combinações das L arestas de $\binom{N}{2}$ possíveis arestas do grafo. Desse modo, obtém-se:

$$P(L) = \binom{\binom{N}{2}}{L} p^L (1-p)^{\binom{N}{2}-L}. \quad (2.28)$$

Observa-se que é uma distribuição binomial com parâmetros $\binom{N}{2}$ e p , consequentemente o valor esperado do número de ligações $\langle L \rangle$ no grafo aleatório é:

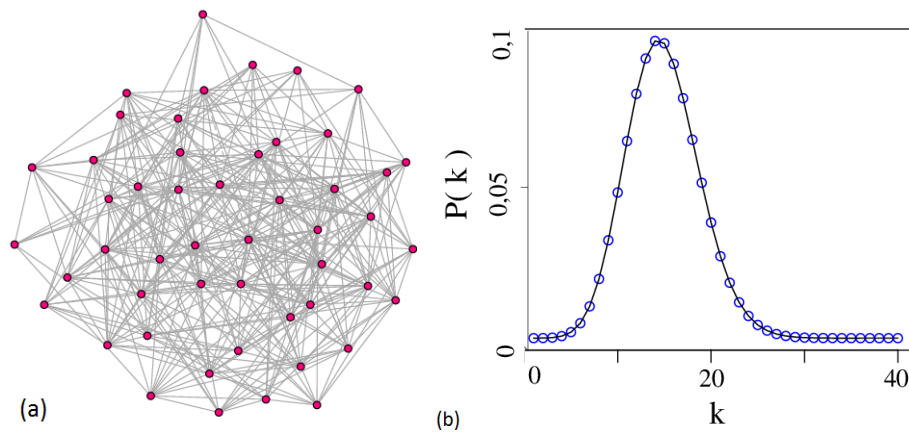
$$\langle L \rangle = \sum_{L=0}^{\binom{N}{2}} L p(L) = p \frac{N(N-1)}{2}, \quad (2.29)$$

e grau médio $\langle k \rangle$ é obtido por

$$\langle k \rangle = \frac{2L}{N} = p(N-1) \simeq pN. \quad (2.30)$$

Exemplo: Rede aleatória de Erdős-Rényi, com 50 nós, Figura 2.16(a), usando a probabilidade $p = 0,3$, tendo grau médio $\langle k \rangle = 15,28$.

Figura 2.16 – Exemplo de um grafo aleatório de Erdős e Rényi com $N = 50$ (figura (a)), usando uma probabilidade $p = 0,3$ e sua distribuição de graus (figura (b)).



Fonte: O autor.

Para N com alto valor e p com valor baixo pode-se utilizar a aproximação dada pela Equação 2.32 para obter a distribuição de graus $P(k)$. A distribuição $P(k)$ que corresponde à probabilidade de que um nó selecionado ao acaso tenha grau k é dada por:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \quad (2.31)$$

onde $\binom{N-1}{k}$ representa o número de maneiras para selecionar k vértices dentre os $N-1$, de modo que se realizem k ligações; p^k representa a probabilidade de que o vértice escolhido esteja conectado a k outros, e $(1-p)^{N-1-k}$ é a probabilidade de que tal vértice não esteja conectado aos $(N-1-k)$ nós restantes. Assim:

$$\binom{N-1}{k} = \frac{(N-1)!}{k!(N-1-k)!} = \frac{(N-1)^k}{k!}.$$

Utilizando a série de Taylor para $x = 0$, $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$, tem-se que:

$$\begin{aligned} \ln \left[(1-p)^{N-1-k} \right] &= \left((N-1-k) \ln \left(1 - \frac{\langle k \rangle}{N-1} \right) \right) = -\langle k \rangle \implies \\ \implies (1-p)^{N-1-k} &= e^{-\langle k \rangle}. \end{aligned}$$

Portanto

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}, \quad (2.32)$$

que é a distribuição de Poisson com valor esperado $\langle k \rangle = p(N-1)$.

Erdős e Rényi mostraram que quando $p < \frac{1}{N} \Rightarrow \langle k \rangle < 1$, isto é, há uma coleção de nós isolados, para $p > \frac{1}{N}$ existe uma componente gigante que abrange a maioria dos vértices e para $p > \frac{\ln(N)}{N}$, todos os vértices estão conectados (NEWMAN, 2010). Para estas redes, $\langle k \rangle$ é o número médio de vértices vizinhos, com $l = 1$ (distância unitária). O número médio de vértices com $l = 2$ é $\langle k \rangle^2$. Para a distância $\langle l \rangle$, o número médio de vizinhos é aproximadamente $\langle k \rangle^{\langle l \rangle}$. Assim, é estimado que o comprimento médio do caminho mais curto $\langle l \rangle$, considerando que

$$\begin{aligned} N = 1 + \langle k \rangle + \langle k \rangle^2 + \dots + \langle k \rangle^d &= \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1} = \langle k \rangle^{\langle l \rangle} \implies \\ \langle l \rangle &= \frac{\ln(N)}{\ln(\langle k \rangle)}. \end{aligned} \quad (2.33)$$

2.5.2 Redes *small world*

Ao estudarem redes reais (NEWMAN, 2010), Duncan Watts e Steven Strogatz (1998) (WATTS; STROGATZ, 1998), observaram que algumas redes possuíam as seguintes características estruturais em comum:

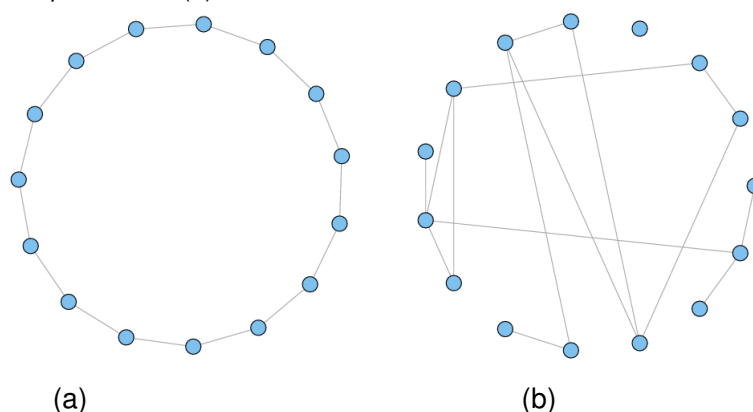
- alta clusterização (presença de grande número de triângulos ou ciclos fechados de ordem três),
- baixa distância entre os vértices e
- muito esparsas.

Eles denominaram estas redes de “rede mundo pequeno”, do inglês *small world network* (WS) e propuseram um modelo matemático representando estas três características. O modelo é

baseado numa configuração inicial de rede regular, esparsa e com alta clusterização, seguida de mudanças aleatórias para induzir baixas distâncias entre os vértices.

O modelo WS (Watts-Strogatz) não representa nenhum aspecto específico do processo de criação de redes reais. Para obtenção deste modelo, inicia-se com uma rede regular formada de um círculo com N vértices (rede em forma de um anel), todos com grau dois. Este processo inicial dá origem a uma grade regular onde o grau de cada vértice é $k = 2m$, onde m representa o número de vizinhos do vértice. Reescreve-se aleatoriamente cada aresta da rede com uma probabilidade p , não sendo permitido autoconexões e nem arestas múltiplas; deste modo são criados atalhos entre partes distantes do grafo. Em geral cada nó i está conectado aos $2m$ vizinhos mais próximos (m nós à direita e m nós à esquerda). Cada vértice tem grau $2m$ e estão dispostos no círculo. Na Figura 2.17, tem-se a configuração inicial de uma rede regular com $N = 15$, $p = 0$ e $m = 1$, religando as arestas com probabilidade $p = 0,4$.

Figura 2.17 – Representação do modelo Watts-Strogatz com parâmetros $N = 15$ e $p = 0$ na figura (a) e $N = 15$ e $p = 0,4$ em (b).

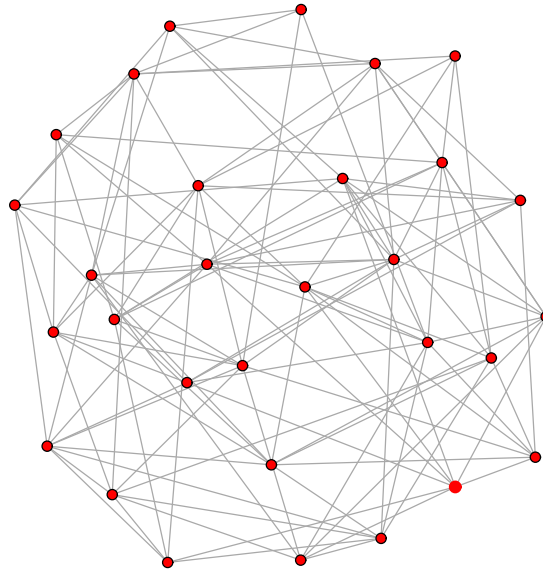


Fonte: O autor.

Observa-se que o modelo WS possui três parâmetros: N, m, p . O parâmetro p , determina a probabilidade de reposicionamento das arestas. Quando $p = 0$, tem-se a grade regular inicial e quando $p = 1$ todas as arestas são reposicionadas aleatoriamente e tem-se um modelo parecido com o $G(N, p)$.

Na Figura 2.18 a rede “mundo pequeno” foi gerada a partir de uma estrutura regular com 30 nós, $m = 4$ e $p = 0,4$, notando-se elevado número de ciclos fechados de ordem três (triângulos).

Figura 2.18 – Rede *small world* com 30 nós com grau 4 e probabilidade de religação .



Fonte: O autor.

A rede regular inicial ($p = 0$) possui o comprimento médio do caminho mais curto entre os vértices, dadas por:

$$\langle l \rangle_{reg} = \frac{N}{4m}, \quad (2.34)$$

e coeficiente de agrupamento médio $\langle C \rangle_{reg} = \frac{3m - 3}{4m - 2}$.

A rede complexa mundo pequeno (WS) por ser um caso intermediário da rede regular e da rede completamente aleatória, é tal que:

$$\frac{\log(N)}{\log(\langle k \rangle)} < \langle l \rangle_{WS} < \frac{N}{4m}, \quad (2.35)$$

e que

$$\frac{\langle k \rangle}{N} < \langle C \rangle_{WS} < \frac{3(m - 1)}{2(2m - 1)}, \quad (2.36)$$

onde $\langle k \rangle = k' = 2m$.

2.5.3 Redes livre de escala

Diferente de outros modelos, na rede livre de escala, o número N de nós não é fixado, onde inicialmente tem-se uma rede simples com um número qualquer de nós no instante ($t_0 = 0$) e a seguir, em cada instante de tempo, novo nó é inserido na rede e conectado a outros nós da rede inicial, seguindo princípio da conexão preferencial, em que a probabilidade de um novo nó conectar-se a outro é proporcional ao grau deste. Neste modelo, observa-se que a distribuição de graus obedece a uma lei de potência (NEWMAN, 2010).

Uma distribuição é uma lei de potência se sua função de probabilidade possui a seguinte forma:

$$P(x) = cx^{-\gamma}, \quad (2.37)$$

onde $P(x)$ é a probabilidade da variável discreta aleatória X assumir o valor x , c é a constante de normalização e γ uma constante e parâmetro da distribuição com $\gamma > 1$. Esta distribuição possui uma “cauda pesada” pois valores arbitrariamente grandes de x possuem volume maior de probabilidades não desprezíveis.

Em 1999, Albert R. Barabasi e Reka Albert (BARABÁSI; ALBERT, 1999) analisando a estrutura da rede mundial de computadores (do inglês “*world wide web*”) descobriram que além dela apresentar o fenômeno *small world*, sua distribuição de graus não era aleatória e sim do tipo $P(k) = k^{-\gamma}$, modelo *scale free* ou *livre de escala*. A distribuição $P(k) = ck^{-\gamma}$, com c e γ constantes é chamada livre de escala, pois trocando-se k por ck não se altera a maneira pela qual essa distribuição depende da variável k . Assim, $P(ck) = ca^{-\gamma}k^{-\gamma} = c'k^{-\gamma}$. Uma observação importante deste modelo é que seu gráfico traçado na escala log-log é uma reta: $\log(P(k)) = \log(ck^{-\gamma}) \implies \log(P(k)) = \log(c) - \gamma \log(k)$, para k suficientemente grande. A inclinação da reta é $-\gamma$.

Leis de potência também são conhecidas como lei de Zipf ou distribuições de Pareto. Analisando textos escritos em inglês, conforme (MONTEIRO, 2010), em 1949, Zipf propôs que a frequência relativa $P(k)$ com que aparece uma palavra tem classificação k , na lista de ocorrência, em ordem decrescente, é aproximadamente inversamente proporcional a k , ou seja $P(k) = k^{-\gamma}$ com $\gamma \simeq 1$. Assim, Zipf, fez $k = 1$ corresponder à palavra “*the*” (a palavra mais frequente), “*of*” a $k = 2$, “*to*” a $k = 3$ e assim por diante. Isto significa que a palavra “*the*” é utilizada duas vezes mais que “*of*” e três vezes mais que “*to*”.

Diferente do modelo de Erdős-Rényi, onde o número de vértices N é fixo e são aleatoriamente conectados ou reordenados conforme uma distribuição Poissoniana de graus com um pico em $\langle k \rangle$, caindo exponencialmente para graus muito elevados. Porém, o mesmo não ocorre nas redes livre de escala, onde o grau médio $\langle k \rangle$ não corresponde ao ponto máximo de $P(k)$ pois não é um grau típico da rede, que possui muitos nós com poucas conexões e poucos nós com muitas conexões. Redes reais são formadas pela adição contínua de novos vértices, assim, o número de vértices aumenta com o passar do tempo. Por exemplo, a rede mundial de computadores (*World Wide Web*), a rede de citações científicas, redes de interação social, entre outras. Então, uma característica comum destes sistemas é que a rede cresce continuamente pela adição de novos vértices que se conectam a outros que já estão no sistema.

Também é possível verificar que, em algumas redes, para cada nova relação criada há uma probabilidade maior de que o nó que a receba seja um nó altamente conectado. Tal característica pode ser vista na rede de citações científicas, onde artigos já bastante citados apresentam maior probabilidade de receber novas citações frente a artigos desconhecidos. Assim, Barabási e Albert introduziram o modelo livre de escala (sigla SF), que se mostra mais consistente com redes reais e que pode ser definido em duas etapas ((COHEN; HAVLIN, 2010)):

- Crescimento: a rede deve ser iniciada com uma quantidade N_0 de nós, sendo adicionado um novo nó a cada passo. Cada nó adicionado possui, a princípio, uma quantidade m ($m \leq N_0$) de arestas.

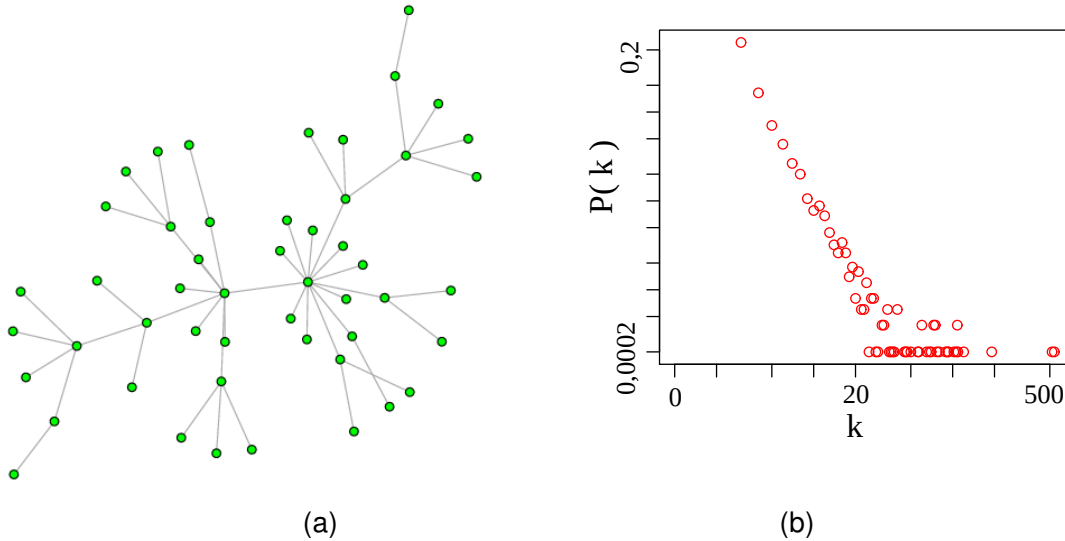
• Conexão preferencial (*preferential attachment*): cada aresta adicionada deve ligar o novo nó a um nó i já existente na rede, escolhido com probabilidade proporcional à sua quantidade de conexões:

$$p(i) = \frac{k_i}{\sum_j k_j}, \quad (2.38)$$

onde k_i é o grau do nó i .

Repetindo-se estas operações após t passos de tempo, a rede terá $N = N_0 + t$ nós e $L = L_0 + mt$ ligações. Exemplo: Uma rede livre de escala com 50 nós e $\langle k \rangle \simeq 4$, Figura 2.19 (a), onde se destacam os *hubs*, característica do modelo. Na Figura 2.19 (b), a distribuição de graus de uma rede livre de escala com $N = 500$ nós, $\gamma = 1$ e $m = 2$.

Figura 2.19 – (a) Modelo de rede do Barabasi com 50 nós e grau médio igual a 4 e (b) distribuição de graus. Em escala (logarítmica) de uma rede livre de escala com 500 vértices.



Fonte: O autor.

A distribuição de graus, neste modelo, no formalismo discreto é dada por:

$$P(k) = p_k = ck^{-\gamma}, \quad \gamma > 1 \quad (2.39)$$

onde C é a constante de normalização e γ é um parâmetro da distribuição. Como $\sum_{k=1}^{\infty} p_k = 1$, implica em que $1 = c \sum_{k=1}^{\infty} k^{-\gamma}$. Portanto:

$$c = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta(\gamma)}, \quad (2.40)$$

é a constante de normalização. Tendo em vista que:

$$\zeta(\gamma) = \sum_{k=1}^{\infty} k^{-\gamma} = 1 + \frac{1}{2^\gamma} + \frac{1}{3^\gamma} + \dots, \quad (2.41)$$

é a função zeta de Riemann, cujo domínio é $\mathbb{C} - \{1\}$, porque a série harmônica de ordem 1 não converge e a parte real de γ , $\Re(\gamma) > 1$. Logo

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}. \quad (2.42)$$

Para $k > 0$ assume-se que não existe nó desconectado na rede. A distribuição de graus no formalismo contínuo pode ser calculada por:

$$\int_{K_{min}}^{\infty} P(k) dk = 1 \implies C = \frac{1}{\int_{K_{min}}^{\infty} k^{-\gamma} dk} = (\gamma - 1) K_{min}^{\gamma-1}. \quad (2.43)$$

Deste modo, tem-se

$$P(k) = (\gamma - 1)(K_{min})^{\gamma-1} k^{-\gamma}. \quad (2.44)$$

A conectividade média $\langle k \rangle$ pode ser obtida, assumindo que a variável k é contínua, calculando a integral entre valor do grau mínimo m e o valor do grau *cutoff* :

$$\langle k \rangle = \int_m^K k \cdot p(k) dk = \int_m^K k c k^{-\gamma} dk = \frac{c}{2-\gamma} (K^{2-\gamma} - m^{2-\gamma}), \quad (2.45)$$

ou ainda:

$$\langle k \rangle = \frac{\gamma - 1}{2 - \gamma} m \left(N^{\frac{2-\gamma}{\gamma-1}} - 1 \right), \quad (2.46)$$

onde $K = m N^{\frac{1}{\gamma-1}}$.

2.5.3.1 Divergência em momentos de maiores ordens

O enésimo momento da distribuição de graus é dado por: $\langle k^n \rangle = \int_{K_{min}}^{\infty} k^n P(k) dk$. Tomando o valor de $P(k)$ na Equação 2.44, vem:

$$\langle k^n \rangle = (\gamma - 1)(K_{min})^{\gamma-1} \int_{K_{min}}^{\infty} k^{n-\gamma} dk = \frac{\gamma - 1}{n - \gamma + 1} (K_{min})^{\gamma-1} [k^{n-\gamma+1}]_{K_{min}}^{\infty}.$$

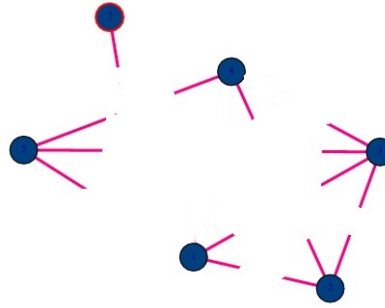
Se $n - \gamma + 1 < 0 \implies \langle k^n \rangle = -\frac{\gamma - 1}{n - \gamma + 1} (K_{min})^n$.

Se $n - \gamma + 1 > 0$, a integral diverge. Para um γ fixado, todos os momentos com $n > \gamma - 1$, diverge. A maioria dos expoentes são menores que três, então $\langle k^2 \rangle$ diverge.

2.6 Redes aleatórias generalizadas (modelo de configuração)

O modelo de configurações ((NEWMAN, 2010)) é um grafo aleatório cuja construção obedece a uma determinada sequência de graus: $D = \{k_1, k_2, \dots, k_N\}$, tal que cada vértice i possui um grau k_i , com $i = 1, 2, 3, \dots, N$. Cada vértice tem um número de “stubs” (pedaços) de arestas igual ao seu grau. A seguir os “stubs” são ligados aleatoriamente para formar arestas. Observa-se que $\sum_i k_i = 2m$.

Figura 2.20 – Pedacos de arestas denominadas de stubs. Os stubs são ligados entre si até formarem uma rede simples.



Fonte: O autor

2.6.1 Probabilidade de uma aresta

A propriedade principal no modelo das configurações é a probabilidade p_{ij} da ocorrência de uma aresta entre dois vértices i e j , onde supõe-se que $k_i, k_j > 0$. No total, existem $2m$ de *stubs* ou $2m - 1$, excluindo aquele que é gerado por i . Qualquer “*stub*” é suscetível de ser ligado a outro, então a probabilidade de um “*stub*” do vértice i ser ligado a qualquer outro em torno do vértice j é $\frac{k_j}{2m - 1}$. Mas existem k_i “*stubs*” no vértice i , então a probabilidade total de uma ligação entre i e j é

$$p_{ij} = \frac{k_i k_j}{2m - 1}. \quad (2.47)$$

No limite para um valor muito alto de m , este número torna-se muito pequeno, podendo-se ignorar o subtraendo 1 do denominador, escrevendo-se:

$$p_{ij} = \frac{k_i k_j}{2m}. \quad (2.48)$$

Nota-se que, mesmo tendo assumido que $k_i, k_j > 0$, essa expressão também fornece o resultado se qualquer grau for zero e portanto a probabilidade de ligação é nula. A probabilidade p_{ij} de ter uma aresta entre os vértices i e j é obtida da Equação 2.48. A probabilidade do vértice i estar conectado ao vértice l é $p_{il} = \frac{k_i k_l}{2m}$ e assim o número de *stubs* do vértice l é reduzido de uma unidade. Portanto, a probabilidade de existência de uma aresta entre j e l é igual a $p_{jl} = \frac{k_j (k_l - 1)}{2m}$. Multiplicando-se estas probabilidades e somando sobre l , encontra-se o valor esperado do número de vizinhos comuns de i e j :

$$N_{ij} = \sum_{l=1}^N \frac{k_i k_l}{2m} \frac{k_j (k_l - 1)}{2m} = \frac{k_i k_j}{2m} \frac{1}{N \langle k \rangle} \sum_{l=1}^N k_l (k_l - 1) = p_{ij} \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}. \quad (2.49)$$

Assim, a probabilidade de compartilhar um vizinho comum é igual a probabilidade dada pela Equação 2.47 multiplicada pelo fator que depende somente da variância e da média da distribuição de graus.

2.6.1.1 Distribuição de graus.

Supondo que um vértice escolhido aleatoriamente (ou não) seja ligado a outro nó deste modelo, qual a probabilidade de que este nó tenha grau k ? Existem $\sum_{i=1}^N k_i = 2m$ stubs no total ou $2m - 1$, excluindo um que tem início nesta aresta. Desde que k seja ligado a um vértice de grau k , a aresta terá probabilidade $\frac{k}{2m - 1}$ ou $\frac{k}{2m}$ no limite $N \rightarrow \infty$. Sabendo-se que p_k é a fração dos vértices com grau k , então o número total de tais vértices é: $N_k = Np_k$. Combinando estes resultados, encontra-se a probabilidade de uma aresta conectada a um vértice qualquer de grau k (NEWMAN, 2010):

$$\frac{k}{2m} Np_k = \frac{kp_k}{\langle k \rangle}. \quad (2.50)$$

Vê-se que a probabilidade de encontrar um vértice com grau k de uma aresta, nestas condições é proporcional a kp_k e não p_k .

O grau médio $\langle k_{NN} \rangle$ de um vizinho de um vértice é definido por:

$$\langle k_{NN} \rangle = \sum_k k \frac{kp_k}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle}. \quad (2.51)$$

Além disso,

$$\langle k_{NN} \rangle - \langle k \rangle = \frac{\langle k^2 \rangle}{\langle k \rangle} - \langle k \rangle = \frac{\langle k^2 \rangle - \langle k \rangle^2}{\langle k \rangle} = \frac{\sigma_k^2}{\langle k \rangle} \geq 0, \quad (2.52)$$

onde σ_k^2 é a variância da distribuição de graus. Da inequação $\frac{\langle k^2 \rangle}{\langle k \rangle} - \langle k \rangle > 0 \implies \frac{\langle k^2 \rangle}{\langle k \rangle} > \langle k \rangle$. O grau médio de um vizinho de um vértice é maior que o grau médio do vértice.

3 ENTROPIA EM REDES COMPLEXAS

3.1 Considerações iniciais

O conceito de entropia tem sua origem na área da termodinâmica clássica criada por Rudolph Clausius no estudo do engenho a vapor, desenvolvido inicialmente por Carnot. Nessa época, concluiu-se que uma parte da energia gerada pela máquina a vapor era desperdiçada em forma de calor residual. Sendo assim, foi proposta uma equação que relacionava calor com a perda de energia (LOPES, 2011):

$$\Delta S = \frac{\Delta Q}{T},$$

sendo que a entropia era medida na mudança de estado. Considerando-se uma transformação reversível infinitesimal, durante a qual a entropia varia de uma quantidade dS e o sistema recebe uma quantidade de calor dQ a uma temperatura T , tem-se que

$$dS = \frac{dQ}{T}. \quad (3.1)$$

A entropia de um sistema constituído de várias partes é igual a soma das entropias de todas as partes, desde que a energia seja a soma das energias de todas as partes e que o trabalho realizado pelo sistema seja a soma dos trabalhos realizados pelas partes.

3.1.1 Entropia de Shannon

Na teoria da informação, a entropia corresponde à incerteza probabilística associada a uma distribuição de probabilidade. Cada distribuição reflete um certo grau de incerteza e diferentes graus de incerteza estão associados a diferentes distribuições (embora diferentes distribuições possam refletir o mesmo grau de incerteza). De um modo geral, quanto mais “espalhada” a distribuição de probabilidade, maior incerteza ela irá propagar. Por exemplo, se alguém lança um dado de seis faces, sem saber se ele é viciado ou não, a probabilidade mais razoável a ser atribuída a cada resultado possível é $1/6$, ou seja, a incerteza é representada usando a distribuição uniforme. Uma importante característica da entropia na teoria da informação, é que ela está diretamente associada ao grau de similaridade entre as probabilidades de uma distribuição. Este aspecto confere uma importante versatilidade à essa noção de entropia que lhe permite ser estendida e adaptada, enquanto conceito, à várias outras disciplinas.

Em 1948, Shannon, propõe uma medida para quantificar o grau de incerteza de uma distribuição de probabilidade. Na teoria da informação, a entropia é normalmente representada por H e sua expressão formal para distribuições discretas de probabilidades é fornecida por (JAYNES, 1957):

$$H(P) = -\sum_{i=1}^N p_i \log(p_i), \quad (3.2)$$

em que $P = \{p_1, p_2, p_3, \dots, p_N\}$ é a distribuição de probabilidade. A entropia H é não negativa e assume seu valor máximo $H(P) = \ln(N)$, quando $P = \left\{\frac{1}{N}, \frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right\}$ (distribuição uniforme). Qualquer outra distribuição faz com que H seja menor que $\ln(N)$. Seu mínimo ocorre em zero, ou seja, $H(P) = 0$ caso em que $P = \{0, 0, 0, \dots, 1, 0, 0, 0, \dots, 0\}$, onde uma das p_i s é um e as demais são nulas. As propriedades mais relevantes da entropia H , são:

1. $H(P)$ é duas vezes diferenciável em P .
2. $H(p_1, p_2, \dots, p_N)$ é simétrica em relação à permutação dos p_i s, com $i = 1, 2, \dots, N$.
3. $H\left(\frac{1}{N}, \frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right)$ deve ser monotonicamente crescente de N .
4. $H(P)$ é uma função estritamente côncava de P .

A primeira propriedade permite a aplicação de técnicas para maximização de funções diferenciáveis. A segunda propriedade afirma que as p_i s podem ter sua ordem invertida que no cálculo de H seu valor não se altera. Na terceira propriedade, a entropia será máxima quando a distribuição for uniforme, ou seja, ela cresce quanto maior for o número de resultados possíveis N . Por último, a quarta propriedade significa que H terá um máximo global, mesmo quando sujeita a restrições lineares.

A forma usual da entropia de Shannon para uma variável aleatória X é:

$$H(X) = -\sum_{x \in X} P(x) \log_2(P(x)), \quad (3.3)$$

de forma que $\sum_{x \in X} P(x) = 1$ e a base 2 indica que a unidade de medida da entropia é o *bit*, que representa a quantidade de informação do sistema, quando é necessário efetuar uma escolha entre dois eventos possíveis equiprováveis. Se a base de logaritmos for a natural, então a unidade de medida é o *nats*. A entropia H da distribuição de conectividade $P(k)$ quantifica a heterogeneidade na organização das conexões de uma rede complexa.

Em redes complexas, a entropia de Shannon é utilizada geralmente na análise de organização das conectividades, levando-se em conta a distribuição dos graus dos nós. A entropia H da distribuição de graus $p_k = P(k)$ é dada por

$$H = -\sum_k P(k) \log(P(k)), \quad (3.4)$$

medida esta que quantifica a heterogeneidade na organização das ligações de uma rede complexa que por sua vez está relacionada com a robustez da rede. Quanto mais homogênea for a distribuição de graus, maior será a entropia. O valor máximo da entropia é $H_{\max} = \log(N - 1)$ para $p_k = \frac{1}{N - 1}$, $\forall k = 1, 2, \dots, N$, e o valor mínimo $H = 0$ ocorre quando $p_k = \{0, 0, 0, \dots, 1, \dots, 0, 0\}$.

3.1.2 Entropia de Gibbs de uma rede aleatória com uma sequência de graus

Seja a sequência gráfica de graus $\{k_1, k_2, k_3, \dots, k_N\}$ de um *ensemble* de redes. Então, a entropia de Gibbs (ZHAO et al., 2011) é dada por:

$$H(\{k_i\}) = \frac{1}{N} \log(\mathcal{N}(\{k_i\})), \quad (3.5)$$

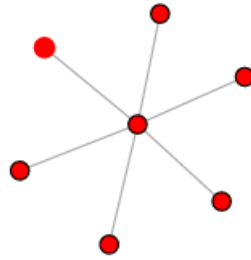
em que $\mathcal{N}(\{k_i\})$ representa o número de redes com a sequência de graus especificada e N é o número de nós rotulados da rede. A entropia de Gibbs depende do número de arestas sobre a sequência de graus específica:

$$\sum_{i=1}^N k_i = 2m \quad (\text{número de stubs de arestas}). \quad (3.6)$$

Exemplos:

1) Dada a sequência de graus $\{k_i\} = \{1, 1, 1, 1, 1, 1, 6\} \implies \mathcal{N}(\{k_i\}) = 1$.

Figura 3.1 – Rede estelar de sete nós (distinguíveis) com entropia nula.



Fonte: O autor.

Neste caso, $\mathcal{N}\{k_i\} = 1$, $H[\{k_i\}] = \frac{1}{7} \log(1) = 0$.

2) Com a sequência de graus $\{k_i\} = \{1, 2, 2, 2, 3\} = \{3, 2, 2, 2, 1\}$, tem-se que $\mathcal{N}(\{k_i\}) = 6$. Pela figura 3.2, tem-se que $\mathcal{N}\{k_i\} = 6 \implies H[\{k_i\}] = \frac{1}{5} \log(6) \simeq 0,358$.

3.2 Entropia de redes modelo G(N,p)

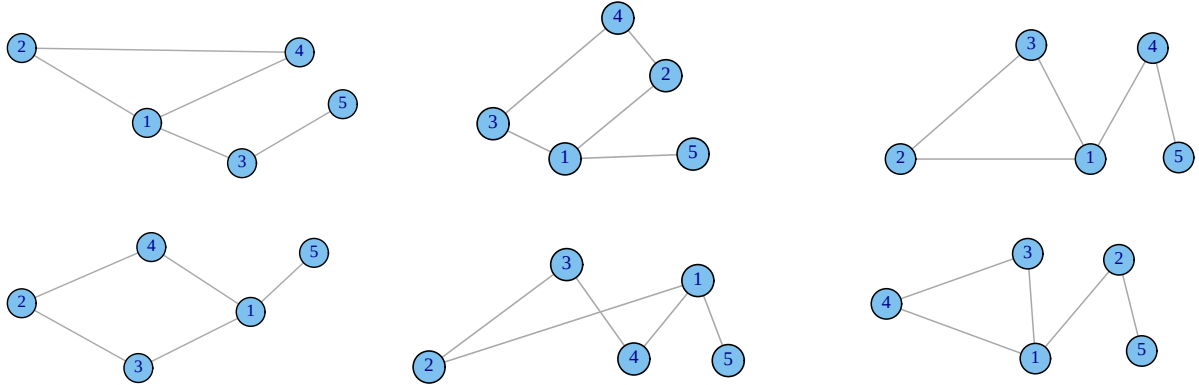
Pela teoria da entropia de Shannon, tem-se que as probabilidades são diferentes para diferentes estados, ou seja (Jl et al., 2008):

$$S(\Omega, P) = -K \sum_{i=1}^{\Omega} P_i \log(P_i), \quad (3.7)$$

onde Ω é o número de possíveis estados do sistema, P é a probabilidade do i -ésimo estado e K é uma constante. Tomando-se um valor apropriado para K , $K = \log_2 10$, tem-se que a entropia passa a ser calculada por

$$S(\Omega, P) = - \sum_{i=1}^{\Omega} P_i \log_2(P_i). \quad (3.8)$$

Figura 3.2 – Redes não direcionadas geradas pela sequência de graus $\{3,2,2,2,1\}$, onde o número de redes $\mathcal{N}\{k_i\} = 6$.



Fonte: Adaptada de (ZHAO et al., 2011).

No modelo de redes aleatórias de Erdos Rényi, tem-se que o número total de possíveis arestas é dada por $M = \frac{N(N-1)}{2}$, o número atual ou efetivo de arestas representado por L obedece à distribuição binomial:

$$P(L) = \binom{M}{L} p^L (1-p)^{M-L}, \quad L = 0, 1, 2, \dots, M. \quad (3.9)$$

Para M possíveis arestas e L efetivas ligações, tem-se que

$$C_M^L = \frac{M!}{L!(M-L)!}, \quad (3.10)$$

desde que os nós sejam rotulados (distinguíveis). O número de configurações possíveis $\Omega(N, P)$ para o número total de nós N e probabilidade P é dado por

$$\Omega(N, P) = \sum_{L=0}^M C_M^L = \sum_{L=0}^M \frac{M!}{L!(M-L)!}. \quad (3.11)$$

Para cada L , existem C_M^L configurações possíveis correspondentes. Assim, P_i ocorre com a probabilidade:

$$P_i = \frac{P(L)}{C_M^L}, \quad i = 1, 2, \dots, \Omega. \quad (3.12)$$

Como as probabilidades de cada configuração C_M^L para um certo L são as mesmas, o valor da entropia do total de configurações $\sum_{L=0}^M C_M^L$ pode ser calculada para diferentes valores de L . Então a entropia S pode ser escrita como:

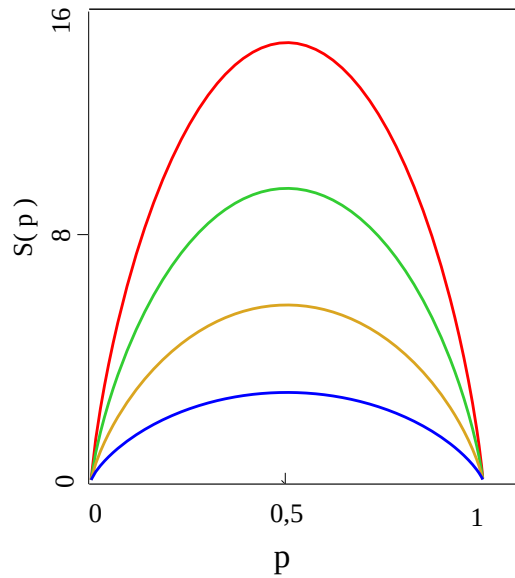
$$S(N, P) = \sum_{L=0}^M C_M^L [p^L (1-p)^{M-L}] \log_2 [p^L (1-p)^{M-L}].$$

Simplificando-se esta equação, obtém-se a entropia da rede referente à configuração da topologia na forma analítica:

$$S(N, P) = -\frac{1}{2}N(N-1) [P\log_2(P) + (1-P)\log_2(1-P)] . \quad (3.13)$$

Na Figura 3.3, exibe-se os casos em que $N = 3, 4, 5, 6$ e as probabilidades de ligação variam de 0,05 a 0,95, corroborando os resultados obtidos em (JI et al., 2008) :

Figura 3.3 – Entropia do modelo Erdős-Rényi calculada para redes (contadas de baixo para cima) com três, quatro, cinco e seis nós respectivamente, calculadas pela equação 3.13.



Fonte: O autor.

3.3 Entropia em redes livre de escala

Nas redes livre de escala, a distribuição dos graus dos vértices é $p(k) = ck^{-\gamma}$, em que o grau k varia entre $k = k_{min}, \dots, K$ e $K =$ conectividade *cutoff* (grau de corte) da rede. Nas redes complexas, a entropia da distribuição de graus (WANG et al., 2006) dada por

$$H = -\sum_{k=1}^{N-1} p(k)\log(p(k)), \quad (3.14)$$

fornece uma medida média da heterogeneidade da rede, uma vez que mede a diversidade ou variedade das distribuições das ligações. Pode ocorrer dois casos extremos:

i) O valor máximo da entropia $H_{\max} = \log(N-1)$ obtido de $p(k) = \frac{1}{N-1}$,
 $\forall k = 1, 2, 3, \dots, N-1$.

ii) O valor mínimo da entropia $H = 0$, ocorrendo quando $p(k) = \{0, 0, \dots, 1, 0, 0, \dots, 0\}$.

Para a distribuição de graus da lei de potência de uma rede livre de escala, o impacto da diversidade (cauda longa) é óbvio: aumenta a incerteza. Aumentando o expoente ou diminuindo o *cutoff*, a rede se torna menos heterogênea e, como resultado, gera uma entropia menor.

Considerando-se a aproximação contínua para a distribuição de graus para a rede livre de escala, pode-se expressar a entropia da seguinte forma:

$$H = \int_{m=k_{min}}^K p(k) \log(p(k)) dk = - \int_m^K c k^{-\gamma} \log(c k^{-\gamma}) dk,$$

$$H = -c \log(c) \int_m^K k^{-\gamma} dk + c \int_m^K k^{-\gamma} \log(k^{-\gamma}) dk,$$

$$H = -\frac{c \log(c)}{1-\gamma} [k^{1-\gamma}]_m^K + c \left[\log(k^{-\gamma}) \frac{k^{1-\gamma}}{1-\gamma} + \frac{\gamma}{(1-\gamma)^2} k^{-\gamma} \right]_m^K,$$

$$H = -\frac{c \log(c)}{1-\gamma} (K^{1-\gamma} - m^{1-\gamma}) + \frac{c\gamma}{1-\gamma} \left[K^{1-\gamma} \log(K) - m^{1-\gamma} \log(m) - \frac{1}{1-\gamma} (K^{1-\gamma} - m^{1-\gamma}) \right]. \quad (3.15)$$

Numa rede finita, a maior conectividade pode ser estimada da integral

$$\int_K^\infty p(k) dk = \frac{1}{N} \Rightarrow -c \frac{K^{1-\gamma}}{1-\gamma} = \frac{1}{N} \Rightarrow K = m N^{\frac{1}{\gamma-1}}, \text{ onde a constante } c \text{ foi obtida de}$$

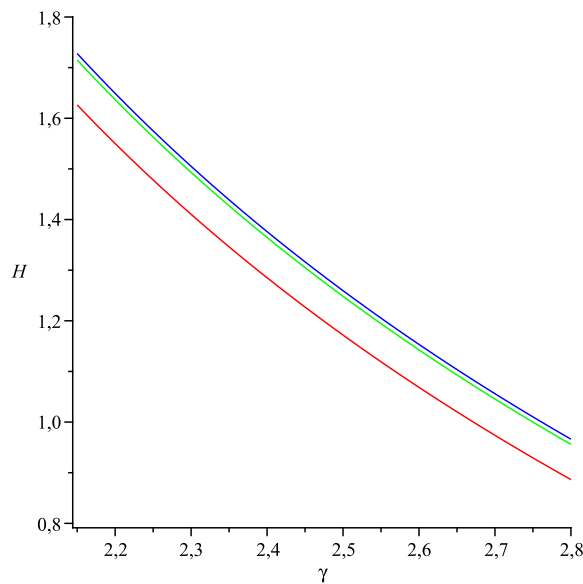
$$1 = \int_m^\infty p(k) dk = c \int_m^\infty k^{-\gamma} dk \Rightarrow c = -(1-\gamma)m^{\gamma-1} = (\gamma-1)m^{\gamma-1}$$

Assim pode-se escrever a entropia como função da conectividade mínima m , do expoente γ e do número de nós da rede N :

$$H(\gamma, m, N) = \left[\log\left(\frac{\gamma-1}{m}\right) + \frac{\gamma}{1-\gamma} \right] \left(\frac{1}{N} - 1 \right) - \frac{\gamma}{\gamma-1} \frac{\log(N)}{N}. \quad (3.16)$$

A Figura 3.4 mostra a relação entre a entropia H e o expoente γ para diferentes valores de N ($N = 100, 1000$ e 10000) quando a conectividade mínima m é definida ($m = 1$).

Figura 3.4 – Relação entre a entropia H e γ para o número de nós $N = 100$ (cor vermelha), 1000 (cor verde) e 10000 (cor azul) no modelo livre de escala com grau mínimo $m = 1$.



Fonte : O autor.

Vê-se que com o crescimento do expoente γ , a conectividade *cutoff* K decresce, o que torna a rede menos heterogênea, diminuindo a entropia e a conectividade média $\langle k \rangle$ pode ser obtida através da aproximação contínua dada por:

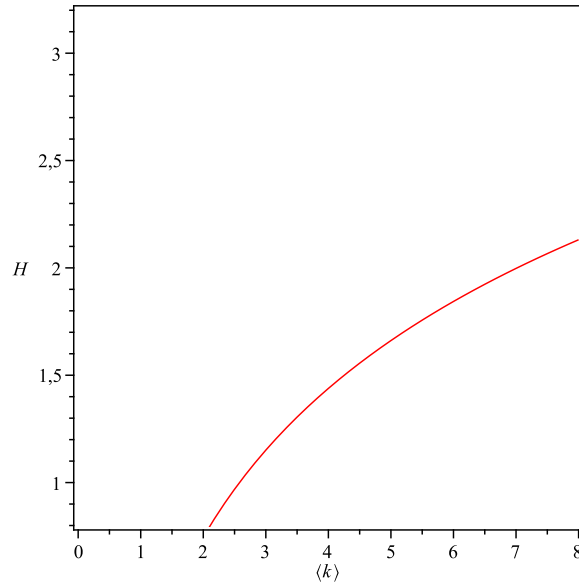
$$\langle k \rangle = \int_m^K k \cdot p(k) dk = \int_m^K k c k^{-\gamma} dk = \frac{c}{2-\gamma} (K^{2-\gamma} - m^{2-\gamma}), \quad (3.17)$$

ou ainda:

$$\langle k \rangle = \frac{\gamma-1}{2-\gamma} m \left(N^{\frac{2-\gamma}{\gamma-1}} - 1 \right). \quad (3.18)$$

Na Figura 3.5, tem-se a entropia numa rede *scale free* de 10^4 nós com grau médio $\langle k \rangle = 6$.

Figura 3.5 – Entropia pelo grau médio de uma rede livre de escala com $N = 10000$ nós e grau médio igual a seis.



Fonte : O autor.

3.4 Grafo aleatório exponencial

Considere um conjunto $\mathcal{G} = \{G_1, G_2, G_3, \dots, G_n\}$ de grafos simples (sem auto conexões e sem arestas múltiplas) com N vértices. A cada grafo $G_i \in \mathcal{G}$ associa-se uma probabilidade $P(G_i)$, tal que $\sum_{G \in \mathcal{G}} P(G_i) = 1$. O valor médio ou esperado de uma medida x_i no ensemble é dado por

$$\langle x_i \rangle = \sum_{G \in \mathcal{G}} P(G) x_i(G), \quad (3.19)$$

no qual $x_i(G)$ é o valor medido no grafo (número de arestas, número de triângulos, coeficientes de aglomeração, etc). A Equação (3.19) torna-se uma restrição sobre a distribuição de probabilidades dos grafos. Pode-se mostrar que a melhor escolha da distribuição de probabilidades é

aquela que maximiza a entropia de Gibbs para um ensemble de grafos:

$$S = - \sum_{G \in \mathcal{G}} P(G) \ln(P(G)), \quad (3.20)$$

sujeita a restrições conhecidas e à condição de normalização

$$\sum_{G \in \mathcal{G}} P(G) = 1. \quad (3.21)$$

A máxima entropia sujeita às condições das Equações (3.21) e (3.19) pode ser encontrada pelo método dos multiplicadores de Lagrange α e θ_i . O valor ótimo é um conjunto de valores que maximiza a quantidade:

$$\frac{\partial}{\partial P(G)} \left[S - \alpha \left(\sum_G P(G) - 1 \right) - \sum_i \theta_i \left(\sum_G x_i(G) P(G) - \langle x_i \rangle \right) \right], \quad (3.22)$$

para todo grafo G .

O valor de x_i no grafo G é representado por $x_i(G)$. Derivando a Equação (3.22) em relação à probabilidade $P(G)$, vem:

$$-\ln(P(G)) - 1 - \alpha - \sum_i \theta_i x_i(G) = 0.$$

Logo:

$$P(G) = \exp(-\alpha - 1 - \sum_i \theta_i x_i(G)), \quad (3.23)$$

ou:

$$P(G) = \frac{e^{-H(G)}}{Z}, \quad (3.24)$$

em que

$$Z = e^{1+\alpha} = \sum_G e^{-H(G)}, \quad (3.25)$$

é a função de partição e

$$H(G) = \sum_i \theta_i x_i(G), \quad (3.26)$$

é denominado Hamiltoniano do grafo.

Do valor esperado de $\langle x_i \rangle$ no *ensemble* conclui-se que:

$$\langle x_i \rangle = \frac{1}{Z} \sum_{G \in \mathcal{G}} e^{-\sum_i \theta_i x_i} x_i = -\frac{1}{Z} \frac{\partial}{\partial \theta_i} \sum_G e^{-\sum_i \theta_i x_i(G)} = -\frac{1}{Z} \frac{\partial}{\partial \theta_i} Z = \frac{\partial}{\partial \theta_i} (-\ln(Z)). \quad (3.27)$$

A energia livre de Helmholtz do *ensemble* é dada por

$$F = -\ln(Z). \quad (3.28)$$

Assim, a Equação (3.27) pode ser escrita como:

$$\langle x_i \rangle = \frac{\partial}{\partial \theta_i} F. \quad (3.29)$$

Para calcular o valor esperado de um observável é necessário calcular a função de partição, daí, calcula-se a energia livre F e diferencia-se em relação aos multiplicadores de Lagrange. Conforme (NEWMAN, 2010), as Equações (3.24), (3.25) e (3.26) definem o modelo de grafo aleatório exponencial, que é definido sobre um conjunto de grafos sujeito a certas restrições que maximizam a entropia (NEWMAN, 2010). A distribuição de probabilidade sobre os microestados deve reproduzir um certo valor da energia interna que pode ser identificada com o valor esperado da energia do sistema:

$$U = \langle E \rangle = \sum_i p_i E_i, \quad (3.30)$$

em que a função de partição é dada por

$$Z = \sum_i e^{-\theta E_i}, \quad (3.31)$$

e, portanto, a entropia máxima é

$$S(p_i) = \ln(Z) + \theta U, \quad (3.32)$$

$$S(p_i) = -\sum_i p_i (-\theta E_i - \ln(\sum_i e^{-\theta E_i})), \text{ com } p_i = \frac{e^{-\theta E_i}}{\sum_i e^{-\theta E_i}}.$$

Em geral, pode-se tomar o parâmetro θ como o inverso da temperatura T , tendo em vista a relação da termodinâmica

$$\frac{\partial S}{\partial U} = \frac{1}{T}, \quad (3.33)$$

onde assume-se que a constante de Boltzmann é igual a unidade. Assim,

$$\frac{\partial S}{\partial U} = \frac{\partial}{\partial U} (\ln(Z) + \theta U) = \theta = \frac{1}{T}. \quad (3.34)$$

E a Equação (3.32) pode ser escrita como:

$$S = \ln(Z) + \frac{U}{T}. \quad (3.35)$$

Exemplo 1: O modelo mais simples de grafo aleatório exponencial é aquele em que se fixa o número esperado de arestas L de uma rede não direcionada.

Neste caso o Hamiltoniano toma a forma

$$H(G) = \theta L(G) \quad (3.36)$$

onde L é o número de arestas e θ é o parâmetro de acoplamento das arestas (ou pode ser pensado como o inverso da temperatura). A função de partição para este Hamiltoniano é calculada

para *ensemble* de redes simples. Os grafos individuais aparecem no conjunto com probabilidades

$$P(G) = \frac{e^{-\theta L}}{Z} = \frac{e^{-\theta L}}{\sum_G e^{-\theta L}}. \quad (3.37)$$

Assim,

$$\langle x \rangle = \frac{1}{Z} \sum_G x(G) e^{H(G)}. \quad (3.38)$$

A matriz de adjacência A do grafo é simétrica e de ordem N (número de nós) tal que $a_{ij} = 1$ se o nó i estiver conectado com o nó j e $a_{ij} = 0$, caso contrário. O número de arestas L é dado por:

$$L = \sum_{i < j} a_{ij}. \quad (3.39)$$

A densidade de ligações no grafo é

$$\rho = \frac{L}{\binom{N}{2}} = \frac{2L}{N(N-1)}, \quad (3.40)$$

e a função de partição:

$$Z = \sum_{\{a_{ij}\}} e^{-\theta \sum_{i < j} a_{ij}} = \sum_{\{a_{ij}\}} \prod_{i < j} e^{-\theta a_{ij}} = \prod_{i < j} \sum_{\{a_{ij}\}} e^{-\theta a_{ij}} = \prod_{i < j} (1 + e^{-\theta}) = (1 + e^{-\theta})^{\binom{N}{2}}. \quad (3.41)$$

Neste caso, a energia livre $F = -\ln(Z)$ será obtida por:

$$F = -\binom{N}{2} \ln(1 + e^{-\theta}). \quad (3.42)$$

Usando a Equação (3.38), pode-se calcular o valor esperado do número de arestas:

$$\langle L \rangle = \frac{1}{Z} \sum_G L e^{-\theta L} = -\frac{1}{Z} \frac{\partial Z}{\partial \theta} = -\frac{\partial \ln(Z)}{\partial \theta} = \frac{\partial F}{\partial \theta} = \binom{N}{2} \frac{1}{1 + e^{\theta}}. \quad (3.43)$$

Como $\langle L \rangle = \binom{N}{2} p$, pode-se escrever $p = \frac{1}{1 + e^{\theta}}$, e a probabilidade de ocorrência de um grafo G do ensemble é dada por:

$$P(G) = \frac{e^{-H}}{Z} = \frac{e^{-\theta L}}{(1 + e^{-\theta})^{\binom{N}{2}}} = p^L (1 - p)^{\binom{N}{2} - L}, \quad (3.44)$$

ou seja, $P(G)$ é a probabilidade de ocorrência de um grafo G em que cada uma das possíveis $\binom{N}{2}$ arestas aparecem com probabilidade independente p .

Da Equação (3.43) pode-se tirar o valor da expressão do multiplicador de Lagrange θ :

$$\theta = \ln \left(\frac{\binom{N}{2} - \langle L \rangle}{\langle L \rangle} \right). \quad (3.45)$$

Para uma rede com dez nós ($N = 10$) e probabilidade $p = 1/3$ tem-se $\langle L \rangle = \frac{1}{3} \cdot 45 = 15$, isto é, $\theta = \ln(2)$. Calculando-se a probabilidade p_{uv} de que existe uma aresta entre os vértices u e v em que é dada a média do elemento a_{uv} da matriz de adjacência. Da Equação (3.27) tem-se que

$$p_{uv} = \langle a_{uv} \rangle = \frac{1}{Z} \sum_{\{a_{ij}\}} a_{uv} e^{-\theta \sum_{i < j} a_{ij}} = \frac{\sum_{a_{uv}=0,1} a_{uv} e^{-\theta a_{uv}}}{\sum_{a_{uv}=0,1} e^{-\theta a_{uv}}} = \frac{1}{1 + e^\theta} = \frac{\langle L \rangle}{\binom{N}{2}} = \langle \rho \rangle. \quad (3.46)$$

A probabilidade de ocorrência de uma aresta entre um dado par de vértices é a mesma neste modelo para cada par e é idêntica à densidade esperada de ligações do *ensemble*.

A entropia deste grafo é obtida mediante a Equação (3.32) e é dada por

$$S = - \binom{N}{2} \frac{1}{1 + e^\theta} + \theta L. \quad (3.47)$$

Tendo em vista que o parâmetro θ é função de L , pode-se escrever a entropia em termos da densidade ρ :

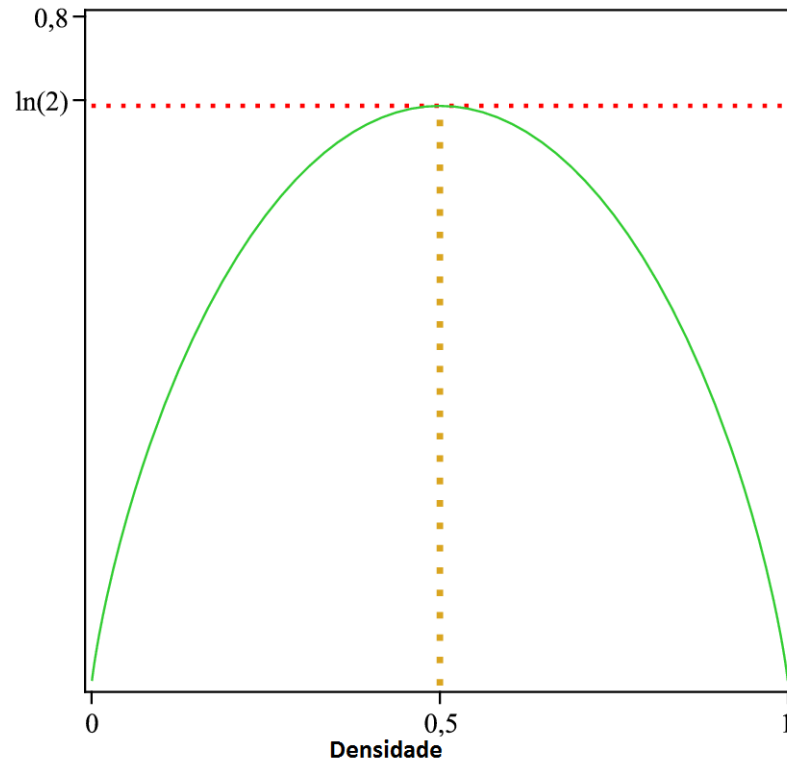
$$S'(\rho) = \frac{S}{\binom{N}{2}} = \ln \left[\frac{1}{\rho} \left(\frac{\rho}{1 - \rho} \right)^{1 - \rho} \right]. \quad (3.48)$$

A Figura (3.6) mostra o comportamento da entropia S' com respeito à densidade ρ . Anula-se em zero e um, atingindo um valor máximo para uma densidade de 0,5, com $S'(0,5) = \ln(2) \simeq 0,7$, que é a quantidade típica da entropia associada a um bit de informação.

Exemplo 2: Grafos aleatórios generalizados, que são realizados por uma sequência de graus $\{k_i\}$ e o modelo de grafo aleatório exponencial correspondente possui o Hamiltoniano:

$$H = \sum_i \theta_i k_i. \quad (3.49)$$

Figura 3.6 – Comportamento da entropia em relação à densidade.



Fonte : O autor.

Observando que $k_i = \sum_j a_{ij}$ e $\sum_i k_i = 2L = \sum_{i,j} a_{ij} = \sum_{i,i} (A^2)_{ii} = \text{tr}(A^2)$ para um grafo não direcionado, tem-se que

$$H = \sum_{i,j} \theta_i a_{ij} = \sum_{i < j} \theta_i a_{ij} + \sum_{i > j} \theta_i a_{ij} = \sum_{i < j} \theta_i a_{ij} + \sum_{i < j} \theta_j a_{ji} = \sum_{i < j} (\theta_i + \theta_j) a_{ij}. \quad (3.50)$$

Então, a função de partição é dada por:

$$Z = \sum_{\{a_{ij}\}} e^{-\sum_{i < j} (\theta_i + \theta_j) a_{ij}} = \prod_{i < j} (1 + e^{-(\theta_i + \theta_j)}). \quad (3.51)$$

E a energia livre:

$$F = -\sum_{i < j} \ln(1 + e^{-(\theta_i + \theta_j)}). \quad (3.52)$$

A probabilidade de conexão entre dois vértices quaisquer u e v é fornecida por:

$$p_{uv} = \langle a_{uv} \rangle = \frac{1}{Z} \sum_{\{a_{ij}\}} a_{uv} e^{-\sum_{i < j} (\theta_i + \theta_j) a_{ij}} = \frac{1}{1 + e^{\theta_u + \theta_v}} \Rightarrow$$

$$p_{ij} = \frac{1}{1 + e^{\theta_i + \theta_j}}. \quad (3.53)$$

A conectividade média de um nó é caracterizada pelo campo local θ_i :

$$\langle k_i \rangle (\theta_i) = \frac{\partial F}{\partial \theta_i} = \sum_{j=1}^N \frac{1}{1 + e^{\theta_i + \theta_j}} = \sum_{j=1}^N p_{ij}. \quad (3.54)$$

Estas duas últimas expressões mostram a relação entre o grau médio e os multiplicadores de Lagrange. Assim, tem-se que graus de valores menores correspondem a multiplicadores fixos com grandes valores e vice versa, e vértices com grande número de conexões correspondem a multiplicadores fixos pequenos.

O grau médio $\langle k \rangle = \frac{\sum_i k_i}{N} = \frac{2L}{N} \implies 2L = \langle k \rangle N$. Como $p_{ij} = \frac{k_i k_j}{2L}$ e pela Equação (3.53)

$$p_{ij} \simeq e^{-(\theta_i + \theta_j)} = \frac{k_i k_j}{\langle k \rangle N} \implies e^{-\theta_i} = \frac{k_i}{\sqrt{\langle k \rangle N}}. \quad (3.55)$$

Assim, cada multiplicador fixo poderá ser obtido aproximadamente pelo logaritmo neperiano do segundo membro da Equação (3.55).

3.5 Ensembles

A Mecânica Estatística é uma teoria probabilística que preocupa-se em estabelecer uma conexão entre dois níveis de descrição: o macroscópico (Termodinâmica) e o microscópico (Mecânica). Um ensemble estatístico é uma distribuição de probabilidade para os microestados do sistema. Gibbs observou que a escolha do conjunto depende das restrições macroscópicas utilizadas. Em 1902, Gibbs, apresentou três ensembles para sistemas físicos (FARKAS et al., 2004):

- **Microcanônico**

O sistema microcanônico é uma expressão estatística da conservação de energia para um sistema fechado. Atribui uma probabilidade igual a todos os microestados com energia especificada. Os outros microestados tem probabilidade zero.

- **Canônico**

O ensemble canônico governa o comportamento de equilíbrio de um sistema a uma temperatura fixa. Um sistema pode trocar energia com um reservatório de calor a uma temperatura constante.

$$\text{A probabilidade de ocorrência de um estado } s \text{ com energia } E_s \text{ é } p(s) = \frac{\exp(-\beta E_s)}{Z}.$$

A termodinâmica do ensemble canônico baseia-se na energia livre de Helmholtz:

$$F(T, V, N) = E - TS.$$

- **Macrocanônico**

O ensemble macrocanônico descreve sistemas em equilíbrio que podem trocar partículas e energias com um reservatório de calor com temperatura e potencial químico constantes. O potencial químico é um regulador do número de partículas e a temperatura, da energia. A probabilidade de ocorrência de um estado s com N_s partículas e energia E_s é $p(s) = \frac{\exp(-\beta(E_s - \mu N_s))}{Z}$. A termodinâmica do ensemble macrocanônico baseia-se no grande potencial $\Phi(T, V, \mu) = -PV$.

Observação: num *ensemble* microcanônico de redes, as arestas (*links*) representam partículas, um grafo corresponde a um estado do sistema e o número de vértices é análogo ao volume da Mecânica Estatística.

3.5.1 *Ensemble* microcanônico de redes complexas

Nas redes complexas, o *ensemble* microcanônico, é constituído pelo número N de nós e o número fixo L de arestas, onde todos os grafos são equiprováveis. Seja $\mathcal{G}$ o conjunto dos grafos G e $\mathcal{N}(G)$ o número de grafos pertencentes a $\mathcal{G}$, então a probabilidade de cada grafo, que tem L arestas e energia E é:

$$P(G) = \frac{1}{\mathcal{N}(G)}, \quad (3.56)$$

e zero para os outros grafos.

3.5.2 *Ensemble* canônico

É constituído por grafos com um número fixo de arestas e cada grafo tem a probabilidade:

$$P(G) = \frac{e^{-\frac{E_G}{T}}}{Z}, \quad (3.57)$$

onde T é a temperatura, E_G é a energia do grafo e $Z = \sum_G e^{-\frac{E_G}{T}}$ é a função de partição do *ensemble* canônico. Da Equação (3.57) pode-se calcular o valor da energia do grafo:

$$E_G = -T(\ln(P(G)) + \ln(Z)).$$

3.5.3 *Ensemble* macrocanônico

É caracterizado por uma temperatura T e um potencial químico μ constantes (fixos). A energia e o número de arestas podem variar no sistema e a probabilidade do grafo G é:

$$P(G) = \frac{e^{-\frac{(E_G - \mu L_G)}{T}}}{Z}, \quad (3.58)$$

em que E_G e L_G representam a energia e o número de arestas do grafo G respectivamente, enquanto que

$$Z = \sum_g e^{-\frac{(E_g - \mu L_g)}{T}}, \quad (3.59)$$

é a função de partição do sistema macrocanônico.

3.6 Entropia estrutural de redes complexas

A restrição estrutural sobre uma rede é formulada como sendo um vínculo sobre a matriz de adjacência A do grafo:(BIANCONI, 2009)

$$\vec{F}(A) = \vec{C}, \quad (3.60)$$

onde a matriz $\vec{C}$ é uma matriz de adjacência com $c_{ij} = 1$ se o vértice i estiver conectado com o vértice j e $c_{ij} = 0$, caso contrário. A função de partição Z do *ensemble* é definida por:(BIANCONI, 2009)

$$Z = \sum_A \delta(\vec{F}(A) - \vec{C}) \exp\left(\sum_{ij} h_{ij} \Theta(a_{ij}) + r_{ij} a_{ij}\right), \quad (3.61)$$

onde h_{ij} e r_{ij} são campos auxiliares, $\Theta(x)$ é a função degrau unitário ($\Theta(x) = 1$ se $x > 0$ e $\Theta(x) = 0$, se $x \leq 0$) e $\delta(\cdot)$ é o delta de Kronecker.

A entropia de Gibbs por nó do *ensemble* de redes e a probabilidade de uma aresta existir entre os nós i e j são definidas por:

$$NS = \ln(Z) \big|_{h_{ij}=r_{ij}=0} \forall(i, j), \quad (3.62)$$

$$p_{ij} = \frac{\partial \ln(Z)}{\partial h_{ij}} \big|_{h_{ij}=r_{ij}=0} \forall(i, j), \quad (3.63)$$

e para um *ensemble* de redes ponderadas:

$$w_{ij} = \frac{\partial \ln(Z)}{\partial r_{ij}} \big|_{h_{ij}=r_{ij}=0} \forall(i, j). \quad (3.64)$$

3.6.1 Entropia em *ensembles* de redes não direcionadas com aproximação de ordem zero

O primeiro *ensemble* de redes estudado é o *ensemble* microcanônico (ordem de aproximação zero) tendo o número de nós N e o número de arestas L fixos. O modelo é representado por $G(N, L)$. Seja A a matriz de adjacência do grafo G , então $L = \sum_{i < j} a_{ij}$ e o número total $\mathcal{N}_0$

de redes que satisfazem estas condições é: $\mathcal{N}_0 = \binom{\binom{N}{2}}{L}$. Então a entropia por nó, S_0 no nível zero de aproximação deste conjunto de redes é dada por:

$$NS_0 = \ln \left(\frac{N(N-1)}{2} \right), \quad (3.65)$$

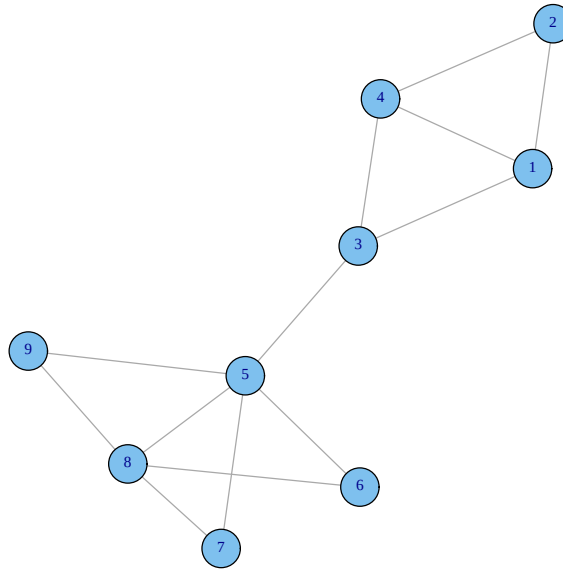
onde a binomial representa o número de redes $\mathcal{N}_0$ do *ensemble*.

Utilizando a aproximação de Stirling ($\ln(n!) = n\ln(n) - n$) tem-se que a entropia por nó é dada por:

$$NS_0 = \frac{N(N-1)}{2} \ln\left(\frac{N(N-1)}{2}\right) - L\ln(L) - \left(\frac{N(N-1)}{2} - L\right) \ln\left(\frac{N(N-1)}{2} - L\right). \quad (3.66)$$

Por exemplo: a entropia do grafo com nove nós ($N = 9$) e treze arestas ($L = 13$) gerado aleatoriamente e representado na Figura (3.7), é $S_0 = 2,62$.

Figura 3.7 – Exemplo de rede aleatória com nove nós e treze arestas cuja entropia $S_0 = 2,62$.



Fonte : O autor.

A probabilidade p_{ij} de uma aresta (i, j) para cada par de nós i, j é dada por $p_{ij}^0 = \frac{L}{\frac{N(N-1)}{2}}$, em que o denominador representa o número de todas as possíveis arestas da rede. Assim $p_{ij}^0 = \frac{13}{36}$ (todas as ligações tem a mesma probabilidade).

A probabilidade de ocorrência de um grafo G no *ensemble* microcanônico é $P(G) = \frac{1}{n_G}$, em que n_G representando o número de grafos com L arestas e energia ε .

Obedecendo a restrição do número de arestas $L = \sum_{i < j} a_{ij}$, a função de partição Z_0 (BIANCONI, 2008) pode ser determinada por:

$$Z_0 = \sum_{\{a_{ij}\}} \delta(L - \sum_{i < j} a_{ij}) \cdot \exp\left(\sum_{i < j} h_{ij} a_{ij}\right), \quad (3.67)$$

onde h_{ij} representa um campo auxiliar. Como exemplo, utilizou-se a definição da função de partição Z_0 para o cálculo da entropia do grafo da Figura (3.7), ou seja:

$$Z_0 = \sum_{\{a_{ij}\}} \delta(13 - (a_{12} + a_{13} + \dots + a_{89})) \cdot \exp\left(\sum_{i < j} h_{ij} a_{ij}\right).$$

O campo auxiliar dado por $h_{ij} = -\theta L$ conforme a Equação (3.36) faz com que a função de partição tome o valor:

$$Z_0 = \delta(13 - 13) \sum_{\{a_{ij}\}} \prod_{i < j} \exp(-\theta a_{ij}) = \prod_{i < j} (1 + \exp(-\theta)),$$

$$Z_0 = (1 + \exp(-\theta))^{\binom{9}{2}}.$$

Logo, a entropia por nó é:

$$NS_0 = \ln(Z_0) = \binom{9}{2} \ln(1 + \exp(-\theta)) = 36 \ln \left(1 + \exp\left(-\ln\left(\frac{23}{13}\right)\right) \right) = 1,792,$$

onde o multiplicador de Lagrange θ , é obtido da Equação (3.45).

Verifica-se uma diferença neste resultado com aquele encontrado pela Equação 3.66, em virtude de não colocar a função delta de Kronecker na forma integral, quando utiliza-se a definição.

Por outro lado, considerando-se a probabilidade de ocorrência de um estado s com energia que o campo auxiliar $h_{ij} = -\beta \varepsilon_{ij} = -\frac{1}{T} \varepsilon_{ij}$ como na Mecânica Estatística (com $k_B = 1$), T a temperatura, tem-se para a função de partição:

$$Z_0 = \delta(13 - 13) \sum_{\{a_{ij}\}} \prod_{i < j} \exp\left(-\frac{1}{T} \varepsilon_{ij} a_{ij}\right) = \prod_{i < j} (1 + \exp\left(-\frac{1}{T} \varepsilon_{ij}\right))$$

Para o caso específico em que as ligações tem a mesma energia $\varepsilon_{ij} = \varepsilon$, a função de partição será dada por:

$$Z_0 = (1 + \exp\left(-\frac{1}{T} \varepsilon\right))^{36}.$$

Assim, a entropia é obtida como:

$$S_0 = 36 \ln(1 + \exp\left(-\frac{1}{T} \varepsilon\right)) = \binom{9}{2} \ln(1 + \exp\left(-\frac{1}{T} \varepsilon\right)).$$

Analisando esta expressão alguns valores da temperatura T tem-se:

i) Para $T = 1$, $S_0 = 36 \ln(1 + e^{-\varepsilon})$ resultado coincidente com o anterior, onde

$$\varepsilon = \theta = \ln\left(\frac{23}{13}\right) \simeq 0,57, \text{ a energia tem o mesmo valor do multiplicador de Lagrange } \theta.$$

ii) Para $T = 0 \Rightarrow S_0 = 0$.

iii) Para $T \rightarrow \infty \Rightarrow S_0 = 36 \ln(2) = \binom{9}{2} \ln(2)$ que é o valor máximo da entropia neste ensemble.

3.6.2 Entropia em *ensembles* de redes não direcionadas com aproximação de ordem um

Nestes *ensembles* tem-se um número fixo de nós N e a restrição estrutural é a sequência de graus $k = \{k_1, k_2, \dots, k_N\}$ com a propriedade $\sum_i k_i = 2L$. O *ensemble* constituído é denominado canônico com a função de partição Z_1 definida por (BIANCONI, 2008):

$$Z_1 = \sum_{\{a_{ij}\}} \prod_i \delta\left(\sum_j a_{ij} - k_i\right) \exp\left(\sum_{i < j} h_{ij} a_{ij}\right). \quad (3.68)$$

Substituindo δ por sua forma integral juntamente com os multiplicadores de Lagrange w_i , com $i = 1, 2, \dots, N$, encontra-se:

$$Z_1 = \int Dw e^{-\sum_i w_i k_i} \prod_{i < j} (1 + e^{w_i + w_j + h_{ij}}), \quad (3.69)$$

onde $Dw = \prod_i \frac{dw_i}{2\pi}$. Resolve-se esta integral pelo método do ponto de sela (apêndice) e a entropia é aproximada (no limite $N \gg 1$), por:

$$NS_1 = \ln(Z_1) \simeq -\sum_i w_i^* k_i + \sum_{i < j} \ln(1 + e^{w_i^* + w_j^*}) - \frac{1}{2} \sum_i \ln(2\pi \alpha_i). \quad (3.70)$$

Derivando esta expressão em relação w_i^* e igualando a zero, obtém-se o valor de k_i , ou seja:

$$k_i = \sum_{j \neq i} \frac{e^{w_i^* + w_j^*}}{1 + e^{w_i^* + w_j^*}}. \quad (3.71)$$

Os coeficientes α_i , que representam a derivada de k_i em relação a w_i são definidos por :

$$\alpha_i = \sum_{j \neq i} \frac{e^{w_i^* + w_j^*}}{(1 + e^{w_i^* + w_j^*})^2}. \quad (3.72)$$

A probabilidade de uma aresta neste ensemble é dada por ((BIANCONI, 2008)):

$$p_{ij} = \frac{e^{w_i^* + w_j^*}}{1 + e^{w_i^* + w_j^*}}. \quad (3.73)$$

Considerando, que o *cutoff* estrutural na rede, satisfaz $k_{\max} = k_i < \sqrt{\langle k \rangle N}$, pode-se aproximar na Equação (3.71), $e^{w_i^*} \simeq \frac{k_i}{\langle k \rangle N}$ e, $\alpha_i \simeq k_i$. Neste limite, a rede é não correlacionada e as probabilidades de uma aresta são dadas por $p_{ij} = \frac{k_i k_j}{\langle k \rangle N}$, desde que $w_i^* < 0$ (BIANCONI, 2008).

A entropia destes ensembles não correlacionados é denominada entropia estrutural S_1 e pode ser calculada da expressão

$$NS_1 \simeq -\sum_i \ln\left(\frac{k_i}{\sqrt{\langle k \rangle N}}\right) \cdot k_i + \sum_{i < j} \ln\left(1 + \frac{k_i k_j}{\langle k \rangle N}\right) - \frac{1}{2} \sum_i \ln(2\pi k_i).$$

Utilizando $\ln(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1} = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 - \dots$, $|x| < 1$, para desenvolver a segunda parcela da expressão acima, tem-se:

$$NS_1 \simeq -\sum_i \ln\left(\frac{k_i}{\sqrt{\langle k \rangle N}}\right) \cdot k_i + \frac{1}{2} \sum_{ij} \frac{k_i k_j}{\langle k \rangle N} - \frac{1}{4} \sum_{ij} \frac{k_i^2 k_j^2}{(\langle k \rangle N)^2} + \frac{1}{6} \sum_{ij} \frac{k_i^3 k_j^3}{(\langle k \rangle N)^3} - \frac{1}{2} \sum_i \ln(2\pi k_i),$$

$$NS_1 = -\sum_i (\ln(k_i) - 1)k_i + \frac{1}{2} \langle k \rangle N (\ln(\langle k \rangle N) - 1) - \frac{1}{4} \left(\frac{\langle k^2 \rangle}{\langle k \rangle}\right)^2 + \frac{1}{6N \langle k \rangle} \left(\frac{\langle k^3 \rangle}{\langle k \rangle}\right)^2 - \frac{1}{2} \sum_i \ln(2\pi k_i), \quad (3.74)$$

que fornece a entropia por nó, S_1 . Por exemplo, a entropia S_1 da rede realizada pela sequência de graus $\{5, 4, 3, 3, 3, 2, 2, 2, 2\}$ é 1,309.

A Equação (3.74), fornece uma estimativa para determinar o número de redes $\mathcal{N}$ do *ensemble*:

$$\mathcal{N} \simeq \frac{(\langle k \rangle N)!!}{\prod_i k_i!} e^{-\frac{1}{4} \left(\frac{\langle k^2 \rangle}{\langle k \rangle}\right)^2 + O(\ln(N))}, \quad (3.75)$$

onde $\langle k \rangle N = 2L$, L é o número total de arestas da rede e $()!!$ representa o duplo fatorial.

O termo $(\langle k \rangle N)!! \simeq (\langle k \rangle N - 1)!!$ fornece o número de redes para N muito grande, estimado por:

$$\mathcal{N} \propto \frac{(2L - 1)!! e^{-\frac{1}{4} \left(\frac{\langle k^2 \rangle}{\langle k \rangle}\right)^2}}{\prod_i k_i!}. \quad (3.76)$$

O número de redes para a sequência $\{k_i\} = \{3, 2, 2, 2, 1\}$ dada no exemplo 2 da secção 2.7, é $\mathcal{N} = 5,87 \simeq 6$ redes.

Calculando, pela definição, a entropia S_1 do grafo com a sequência de graus $\{5, 4, 3, 3, 3, 2, 2, 2, 2\}$, em que $N = 9$, $L = 13$ e grau médio $\langle k \rangle = \frac{26}{9}$, onde

$$\langle k \rangle = \frac{\sum_i k_i}{N} = \frac{2L}{N}.$$

O campo auxiliar é dado por $h = -\sum_i \theta_i k_i$ e a função de partição:

$$Z_1 = \sum_{\{a_{ij}\}} \delta(k_1 - (a_{12} + a_{13} + \dots + a_{19})) \delta(k_2 - (a_{23} + a_{24} + \dots + a_{29})) \dots \exp\left(\sum_{i,j} -\theta_i a_{ij}\right),$$

$$Z_1 = \sum_{\{a_{ij}\}} \delta(k_1 - (a_{12} + a_{13} + \dots + a_{19})) \delta(k_2 - (a_{23} + a_{24} + \dots + a_{29})) \dots \exp\left(\sum_{i < j} -(\theta_i + \theta_j) a_{ij}\right),$$

$$Z_1 = 1.1..1 \sum_{\{a_{ij}\}} \prod_{i < j} \exp(-(\theta_i + \theta_j) a_{ij}) = \prod_{i < j} ((1 + \exp(-\theta_i - \theta_j))). \quad (3.77)$$

Como $e^{-\theta_i} = \frac{k_i}{\sqrt{\langle k \rangle N}}$, tem-se que $Z_1 = \prod_{i < j} \left(1 + \frac{k_i k_j}{\langle k \rangle N}\right)$.

Com os elementos da sequência de graus, calcula-se a função de partição:

$$Z_1 = \left(1 + \frac{5 \times 4}{26}\right) \left(1 + \frac{5 \times 3}{26}\right) \left(1 + \frac{5 \times 3}{26}\right) \cdot \left(1 + \frac{5 \times 3}{26}\right) \dots \left(1 + \frac{2 \times 2}{26}\right),$$

$$Z_1 = 8.445,37.$$

A entropia por nó é dada por:

$$N.S_1 = \ln(Z_1) = 9,0413 \implies S_1 = 1,0046. \text{ O valor da entropia calculado pela Equação (3.74): } S_1 = 1,309$$

Observações:

- A energia livre de Helmholtz $F = -\ln(Z)$ permite escrever que

$$F = -\sum_{i < j} \ln(1 + e^{-\theta_i - \theta_j}) = -\sum_{i < j} \ln(1 + e^{-\Theta_{ij}}).$$

- Assim, pode-se calcular a probabilidade de ocorrência p_{ij} de uma aresta entre os nós i e j , pois:

$$p_{ij} = \langle a_{ij} \rangle = \frac{\partial F}{\partial \Theta_{ij}} = \frac{1}{1 + e^{\Theta_{ij}}},$$

e também o valor esperado do grau k_i de um nó i , ou seja,

$$\langle k_i \rangle = \frac{\partial F}{\partial \theta_i}$$

- Colocando o campo auxiliar como $h_{ij} = -\frac{1}{T} \varepsilon_{ij}$, onde T representa a temperatura e ε_{ij} a energia de uma ligação dos nós i e j , tem-se:

$$Z_1 = \sum_{\{a_{ij}\}} \delta(k_1 - (a_{12} + a_{13} + \dots + a_{19})) \delta(k_2 - (a_{23} + a_{24} + \dots + a_{29})) \dots \exp\left(\sum_{i < j} -\frac{1}{T} \varepsilon_{ij} a_{ij}\right).$$

A energia de uma ligação a_{ij} é dada pela soma das energias de seus vértices: $\varepsilon_{ij} = \varepsilon_i + \varepsilon_j$.

$$\text{Então: } Z_1 = 1 \cdot \prod_{i < j} \left(1 + \exp\left(\frac{-\varepsilon_i - \varepsilon_j}{T}\right)\right).$$

A probabilidade de um grafo é

$$P(G) = \prod_{i < j} p_{ij}^{a_{ij}} (1 - p_{ij})^{1-a_{ij}}$$

em que $p_{ij}(T) = \frac{1}{\frac{e^{\frac{\varepsilon_i + \varepsilon_j}{T}}}{\varepsilon_i} + 1} = \frac{x_i x_j}{1 + x_i x_j}$ é a probabilidade de existência de uma ligação entre i e j ; $x_i = e^{-\frac{\varepsilon_i}{T}}$ recebe a denominação de nó *fitness* ou variável oculta e representa uma outra maneira de escrever o multiplicador de Lagrange associado a k_i .

Observa-se que $p_{ij} > \frac{1}{2}$ se $(\varepsilon_i + \varepsilon_j) < 0$ e $p_{ij} < \frac{1}{2}$ se $(\varepsilon_i + \varepsilon_j) > 0$ independentemente de T . Considerando-se os casos em que a temperatura T tende para infinito e para zero, vem $p_{ij}(\infty) = \frac{1}{2}$. Como a probabilidade é constante, a rede se torna um grafo aleatório com $p = \frac{1}{2}$. Neste caso dois grafos G_1 e G_2 são equiprováveis $P(G_1) = P(G_2)$.

Quando $T = 0$, tem-se $p_{ij}(0) = \Theta(-\varepsilon_{ij}) = \begin{cases} 1, & \text{se } -\varepsilon_{ij} > 0 \\ 0 & \text{se } -\varepsilon_{ij} < 0. \end{cases}$

Esta equação significa que somente os pares de vértices em que $\varepsilon_{ij} < 0 \implies p_{ij} = 1$ estão conectados (grafo denso e completo) ou representa um grafo vazio (esparso) se $\varepsilon_{ij} > 0 \implies p_{ij}(0) = 0$.

No caso de $T = 1$, recai-se no caso inicial em que as energias representam os multiplicadores de Lagrange: $\varepsilon_i = \theta_i$ e $\varepsilon_j = \theta_j$.

3.6.3 Entropia em ensembles de redes não direcionadas com estrutura de comunidade (aproximação de ordem 2)

Estruturas modulares denominadas de comunidades são muito comuns em redes complexas. As comunidades são subgrafos que são densamente conectados internamente com poucas conexões externas. Detectar comunidades consiste em buscar conjuntos de vértices densamente conectados entre si dentro da rede. Descobrir a melhor divisão em comunidades de uma rede é um problema NP-completo, exigindo que o desenvolvimento de soluções viáveis baseiem-se em heurísticas como por exemplo, medidas de qualidade. Os métodos de detecção de comunidades em redes complexas baseiam-se na adição ou remoção de arestas sendo classificados como métodos aglomerativos ou divisivos, respectivamente.

Considerando-se uma rede com estrutura de Q comunidades, tal que cada nó $i = 1, 2, \dots, N$ pertence a comunidade $q_i = 1, 2, \dots, Q$ finita, com $Q \ll \sqrt{N}$. A comunidade do nó i representada por q_i , o número fixo de ligações entre comunidades, por $A(q, q')$, com $q_i = q$ e $q_j = q'$, respectivamente. Então para uma rede não direcionada, $A(q, q')$ é dada pela seguinte expressão: (BIANCONI, 2009)

$$A(q, q') = \sum_{i < j} \delta(\underline{q_{ij}} - q) \delta(\overline{q_{ij}} - q') a_{ij}, \quad (3.78)$$

onde $\underline{q_{ij}} = \min(q_i, q_j)$ e $\overline{q_{ij}} = \max(q_i, q_j)$.

A função de partição para este ensemble de redes com uma sequência de graus $\{k_i\}$ e um

numero $A(q, q')$ de ligações fixas entre as comunidades é:(BIANCONI, 2009)

$$Z_2 = \sum_{[a_{ij}]} \prod_i \delta(\sum_j a_{ij} - k_i) \prod_{q \neq q'} \delta(A(q, q') - \sum_{i < j} \delta(\underline{q_{ij}} - q) \delta(\overline{q_{ij}} - q') a_{ij}) \exp(\sum_{i < j} h_{ij} a_{ij}). \quad (3.79)$$

Na Equação (3.79), utilizando a forma integral da função δ e os multiplicadores de Lagrange w_i , $i = 1, 2, \dots, N$, tem-se que:

$$Z_2 = \int Dw Dw_{qq'} \exp(\sum_i w_i (\sum_j a_{ij} - k_i)) \cdot \exp\left(\sum_{q \leq q'} w_{qq'} \left(\sum_{i < j} \delta(\underline{q_{ij}} - q) \delta(\overline{q_{ij}} - q') a_{ij} - A(q, q')\right)\right) \times \exp(\sum_{i < j} h_{ij} a_{ij}),$$

em que $Dw = \prod_i \frac{dw_i}{2\pi}$ e $Dw_{qq'} = \frac{dw_{qq'}}{2\pi}$.

A entropia S_2 deste ensemble é dada por:

$$NS_2 = \ln(Z_2) \simeq -\sum_i k_i w_i^* - \sum_{q \leq q'} A(q, q') w_{q, q'}^* + \sum_{i < j} \ln(1 + \exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*)) - \frac{1}{2} \sum_i \ln(2\pi \alpha_i) - \frac{1}{2} \sum_{q, q'} \ln(2\pi \alpha_{q, q'}), \quad (3.80)$$

com os multiplicadores de Lagrange $\{w_i\}$ e $\{w_{q, q'}\}$ satisfazendo as equações de ponto de sela (apêndice A):

$$k_i = \sum_{i \neq j} \frac{\exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*)}{1 + \exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*)}, \quad (3.81)$$

e

$$A(q, q') = \sum_{i < j} \delta(\underline{q_{ij}} - q) \delta(\overline{q_{ij}} - q') \frac{\exp(w_i^* + w_j^* + w_{q, q'}^*)}{1 + \exp(w_i^* + w_j^* + w_{q, q'}^*)}, \quad (3.82)$$

em que α_i e $\alpha_{q, q'}$ são definidos por:

$$\alpha_i = \sum_j \frac{\exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*)}{(1 + \exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*))^2}, \quad (3.83)$$

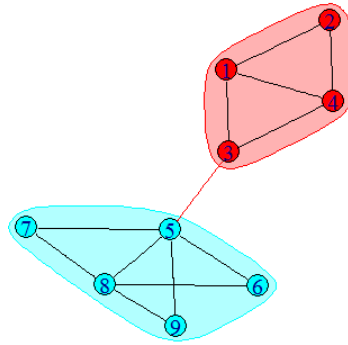
$$\alpha_{q, q'} = \sum_{i < j} \delta(\underline{q_{ij}} - q) \delta(\overline{q_{ij}} - q') \frac{\exp(w_i^* + w_j^* + w_{q, q'}^*)}{(1 + \exp(w_i^* + w_j^* + w_{q, q'}^*))^2}. \quad (3.84)$$

A probabilidade de conexão entre os nós i e j é igual a :

$$p_{ij}^{(2)} = \frac{\exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*)}{1 + \exp(w_i^* + w_j^* + w_{\underline{q_{ij}}, \overline{q_{ij}}}^*)}. \quad (3.85)$$

Calculando-se a entropia do grafo realizado pela sequência de graus $\{5, 4, 3, 3, 3, 2, 2, 2, 2\}$ do exemplo anterior e considerando o grafo modulado em duas comunidades conforme Figura (3.8), com uma aresta fixa ligando as comunidades 1 e 2.

Figura 3.8 – Grafo realizado pela sequência de graus $\{5, 4, 3, 3, 3, 2, 2, 2, 2\}$ com duas comunidades cuja entropia $S_c = 0,79$.



Fonte : O autor.

Calculando o número de ligações fixas entre as comunidades:

$$A(q, q') = \delta(\underline{q}_{12} - q)\delta(\overline{q}_{12} - q')a_{12} + \dots + \delta(\underline{q}_{35} - q)\delta(\overline{q}_{35} - q')a_{35} = 1$$

onde $q = 1$ (comunidade 1) e $q' = 2$ (comunidade 2).

O campo auxiliar é representado por $h_{ij} = \theta_q L_q + \theta_i k_i \Rightarrow$

$$Z_c = 1. \sum_{\{a_{ij}\}} \exp \left(- \left(\sum_{i < j} \theta_q L_q + \sum_{i < j} \theta_i k_i \right) a_{ij} \right).$$

Então:

$$Z_c = \prod_{i < j} (1 + \exp(-\theta_{L_q} - \theta_i - \theta_j)).$$

O multiplicador de Lagrange θ_{L_q} é obtido da expressão: $\theta_{L_q} = \ln \left(\frac{\binom{N}{2} - L}{L} \right)$.

Portanto:

$$Z_c = \prod_{i < j} (1 + \exp(-\ln(35)) \cdot \exp(-\theta_i - \theta_j)) \Rightarrow Z_c = \prod_{i < j} \left(1 + \frac{1}{35} \cdot \frac{k_i k_j}{\langle k \rangle N} \right) \Rightarrow$$

$$\Rightarrow Z_c = \left(1 + \frac{1}{35} \frac{3 \times 3}{10} \right) \left(1 + \frac{1}{35} \frac{3 \times 2}{10} \right) \dots \left(1 + \frac{1}{35} \frac{2 \times 2}{14} \right) = 1,364.$$

Então, a entropia por nó $S_c = \ln(Z_c) = 0,3111$. A comunidade q (comunidade 1) em cor vermelha na Figura (3.8) é constituída pelos vértices $\{1, 2, 3, 4\}$ cuja sequência de graus é $\{3, 3, 2, 2\}$ e a comunidade q' (comunidade 2) em cor azul, pelos vértices $\{5, 6, 7, 8, 9\}$ com a sequência de graus: $\{4, 4, 2, 2, 2\}$

Calculando pela fórmula, a entropia por nó, encontra-se: $S_c = 0,79$.

3.7 Entropia em ensembles de redes direcionadas

3.7.1 Entropia de ensembles de redes direcionadas com aproximação de ordem zero

Considerando-se $\mathcal{N}_0^{dir} = \binom{L^{max}}{L^{dir}} = \binom{N(N-1)}{L^{dir}}$ como o número de redes direcionadas com o número de nós N e o número de arestas L^{dir} fixados. Então a entropia será dada por

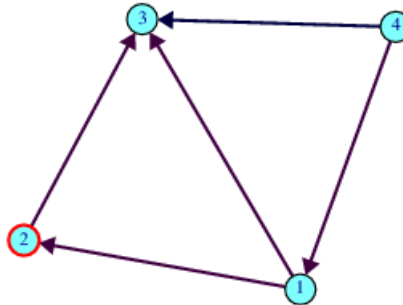
$$NS_0^{dir} = \ln(\mathcal{N}_0^{dir}) = \ln \left(\frac{N(N-1)}{L^{dir}} \right). \quad (3.86)$$

Esta expressão desenvolvida com a utilização da fórmula de Stirling fornece:

$$S_0^{dir} = N(N-1) \ln \left(\frac{N(N-1)}{N(N-1) - L^{dir}} \right) + L^{dir} \ln \left(\frac{N(N-1) - L^{dir}}{L^{dir}} \right). \quad (3.87)$$

Considerando-se uma rede direcionada de quatro nós e cinco arestas (Figura 3.9), tem-se que $S_0^{dir} = 2,25$.

Figura 3.9 – Grafo direcionado ($N = 4$, $L^{dir} = 5$) com entropia $S_0^{dir} = 2,25$.



Fonte : O autor

3.7.2 Entropia em *ensembles* de redes direcionadas com aproximação de ordem um

Neste *ensemble*, o grau de cada vértice é constituído por um par (k_i^{in}, k_i^{out}) representando, respectivamente, o grau de entrada e saída das arestas adjacentes ao nó k_i . Dada uma sequência $D = \{(k_1^{in}, k_1^{out}), (k_2^{in}, k_2^{out}), \dots, (k_N^{in}, k_N^{out})\}$ de pares ordenados, de inteiros não negativos, pode-se construir um grafo direcionado simples $G(V, E)$ para todo $i = 1, 2, \dots, N$. Um dígrafo simples não possui nó com auto conexão nem arestas múltiplas na mesma direção entre dois nós quaisquer.

A sequência D também é denominada sequência bi grau (*bi-degree sequence - bds*). Diz-se que a *bds* é gráfica quando existe um dígrafo simples que realiza D . Para evitar nós isolados, assume-se que $k_i^{in} + k_i^{out} > 0$, para todo $i = 1, 2, \dots, N$. O grau de um nó i é $k_i = k_i^{in} + k_i^{out}$, onde $k_i^{in} = \sum_{j=1}^N a_{ji}$ e $k_i^{out} = \sum_{j=1}^N a_{ij}$.

A função de partição deste *ensemble* é constituída pelas restrições nas conectividades de entrada e de saída:(BIANCONI, 2009)

$$Z_1^{dir} = \sum_{\{a_{ij}\}} \prod_i \delta(k_i^{out} - \sum_j a_{ij}) \prod_i \delta(k_i^{in} - \sum_j a_{ji}) \exp(\sum_{i,j} h_{ij} a_{ij}),$$

$$Z_1^{dir} = \sum_{\{a_{ij}\}} e^{-\sum_i w_i^{int} k_i^{in} - \sum_i w_i^{out} k_i^{out}} \prod_{i \neq j} e^{(w_i^{in} + w_i^{out} + h_{ij}) a_{ij}}.$$

Procedendo de modo análogo ao cálculo da entropia por nó para redes não direcionadas, tem-se que

$$NS_1^{dir} = \ln(Z_1^{dir}) = -\sum_i w_i^{out} k_i^{out} - \sum_i w_j^{in} k_j^{in} + \ln(1 + e^{w_j^{in} + w_i^{out}}) - \frac{1}{2} \ln((2\pi)^2 \alpha_j^{in} \alpha_i^{out}), \quad (3.88)$$

equação da entropia de uma rede direcionada com ordem de aproximação 1, com os multiplicadores de Lagrange satisfazendo as equações do ponto de sela:

$$k_i^{out} = \sum_{i \neq j} \frac{e^{w_j^{in} + w_i^{out}}}{1 + e^{w_j^{in} + w_i^{out}}}, \quad (3.89)$$

$$k_j^{in} = \sum_{i \neq j} \frac{e^{w_j^{in} + w_i^{out}}}{1 + e^{w_j^{in} + w_i^{out}}}, \quad (3.90)$$

$$\alpha_i^{out} = \sum_{i \neq j} \frac{e^{w_j^{in} + w_i^{out}}}{(1 + e^{w_j^{in} + w_i^{out}})^2}, \quad (3.91)$$

$$\alpha_i^{in} = \sum_{i \neq j} \frac{e^{w_j^{in} + w_i^{out}}}{(1 + e^{w_j^{in} + w_i^{out}})^2}. \quad (3.92)$$

A probabilidade de uma ligação direcionada do vértice i ao vértice j é $p_{ij}^{dir,1} = \frac{e^{w_j^{in} + w_i^{out}}}{1 + e^{w_j^{in} + w_i^{out}}}$.

Se $w_i^{out} + w_j^{in} < 0, \forall i, j = 1, 2, \dots, N$, a rede direcionada torna-se não correlacionada e a probabilidade de ligação ente os vértices i e j , é dada por:

$p_{ij}^{dir,1} = \frac{k_i^{out} k_j^{in}}{\sqrt{\langle k^{in} \rangle N}}$. Com esta condição, tem-se a entropia da rede direcionada com ordem de aproximação 1 dada pela expressão:

$$NS_1^{dir} \simeq -\sum_i \ln[(k_i^{in})!(k_i^{out})!] + \ln[(N \langle k^{in} \rangle)!] - \frac{1}{2} \frac{\langle (k^{in})^2 \rangle \langle (k^{out})^2 \rangle}{\langle k^{in} \rangle \langle k^{out} \rangle} - \frac{1}{2} \ln((2\pi)^2 \cdot k_j^{in} k_i^{out}), \quad (3.93)$$

entropia por nó deste ensemble para $k_j^{in}, k_i^{out} \geq 0$ (BIANCONI, 2009).

Utilizando-se a aproximação de Stirling, a entropia por nó pode ser calculada por:

$$\begin{aligned}
NS_1^{dir} = (N \langle k^{in} \rangle) \ln[(N \langle k^{in} \rangle)] - (N \langle k^{in} \rangle) - \left[\sum_i (k_i^{in} \ln(k_i^{in}) + (k_i^{out}) \ln(k_i^{out})) - \right. \\
\left. - (k_i^{in} + k_i^{out}) \right] - \frac{1}{2} \frac{\langle (k^{in})^2 \rangle \langle (k^{out})^2 \rangle}{\langle k^{in} \rangle \langle k^{out} \rangle} - \frac{1}{2} \ln[(2\pi)^2 k_i^{in} k_i^{out}] \quad (3.94)
\end{aligned}$$

Exemplo: No grafo da Figura (3.9), o grau médio de entrada $\langle k^{in} \rangle = \langle k^{out} \rangle = 1,5$ e a entropia do grafo é $S_1^{dir} = 0,1054$.

4 REDE LITERÁRIA SOCIAL DA ODISSEIA, SENHOR DOS ANÉIS E REDE IMUNOLÓGICA

4.1 Rede literária da Odisseia

Na forma de poema, a estória narra o retorno de Ulisses à Grécia após a guerra ocorrida em Ílion, conhecida também por Troia. Esta guerra foi descrita em *Ilíada*, obra do mesmo autor criada anos antes. Apesar de ser considerada uma sequência de *Ilíada*, a *Odisseia* difere da primeira em alguns aspectos. Primeiro, por ter sido escrita nos últimos anos de vida do poeta, (a crítica considera que a narrativa de *Odisseia* não apresenta o mesmo vigor criativo da antecessora). Além disso, a primeira narra os feitos bélicos e sacrifícios realizados por grandes heróis, feitos estes tão espetaculares que os próprios deuses participaram ativamente da narrativa. Nesta obra, o próprio Olimpo está mais próximo dos humanos, para influenciar suas ações e decisões. Na *Odisseia* ocorre exatamente o oposto: é a descrição de tempos de paz, um momento de reconstrução do povo heleno, dedicados à vida em sociedade e ao ambiente doméstico. Ela narra a longa viagem de retorno dos heróis de guerra, que em suas reuniões com chefes dos povos, em palácios, relembram em um ambiente tranquilo as intempéries e dificuldades vencidas, os diferentes povos e costumes que conheceram. Mesmo os feitos bélicos são lembrados de forma romântica. O Olimpo se afasta dos humanos para ocupar o Céu e adquire aspectos espirituais, de forma muito parecida com o entendimento moderno. Finalmente, o alcance geográfico de *Ilíada* é muito menor que de *Odisseia*. Enquanto a primeira se passa quase que exclusivamente em Ílion, a segunda alcança o Norte da África e boa parte do Mediterrâneo.

Por ser o militarismo uma característica essencial dos povos da antiguidade, a *Ilíada* teve importância muito maior nesta época, pois simbolizava a força, o poder e o feito dos seus heróis. Porém, para os povos modernos, civilizados pelo cristianismo, a *Odisseia* teve sua importância elevada por descrever tempos de paz e de reconstrução social e, desta forma, se aproximar dos ideais destes povos.

Atribui-se aos sábios de Alexandria a divisão e classificação tanto de *Ilíada* quanto de *Odisseia*. Ambos foram divididos em vinte e quatro livros, só que aos livros da *Ilíada* apuseram letras gregas maiúsculas e aos da *Odisseia*, minúsculas. Originalmente em hexâmetro grego, ao ser traduzido para o português, a *Odisseia* foi adaptado para a forma de decassílabo. A métrica utilizada manteve o ritmo da narrativa, essencial à poesia, porém os versos ficaram destituídos de rima.

Tanto a *Ilíada* quanto a *Odisseia* tem valor inestimável na literatura clássica. Influenciaram não só a literatura como também outras áreas do conhecimento, como por exemplo a historiografia (HOMERO, 1960). A rede foi construída baseando-se na interpretação textual do poema, onde foram definidos alguns critérios normativos na extração textual.

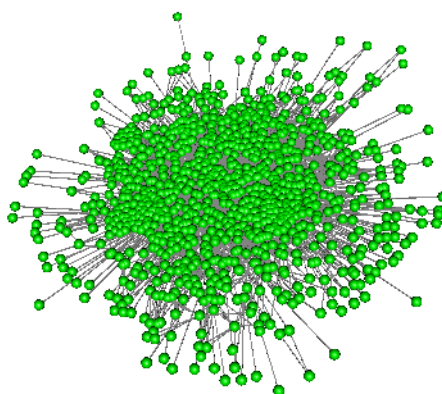
Nas redes sociais derivadas de narrativas, os vértices representam os personagens e as

arestas representam uma forma de interação entre eles. Segundo Miranda (MIRANDA; BAPTISTA; PINTO, 2013)), tais interações foram definidas pelos seguintes critérios:

1. Personagens que estiverem no mesmo local no mesmo instante, demonstrando participar da ação: critério de compartilhamento de evento.
2. Personagens que interajam diretamente: critério de interação direta.
3. Personagens que citam outro à terceiros. Esse critério deve prevalecer se, no ato da citação, o interlocutor especificar um mínimo de conhecimento acerca de seu subjecto: critério de interação indireta.

Na construção da rede foi possível representar 342 personagens ligados socialmente por 1747 interações (Figura 4.1).

Figura 4.1 – Rede literária da Odisseia.



Fonte: O autor.

É importante enfatizar que a rede é limitada em termos da sociedade da época em que se passou a estória. Portanto, é uma representação de parte da sociedade (MIRANDA; BAPTISTA; PINTO, 2013)). Um sumário das propriedades estruturais encontradas na rede da Odisseia e de outras redes mitológicas podem ser encontradas na tabela I do artigo de Miranda (MIRANDA; BAPTISTA; PINTO, 2013)) e do artigo de Carron (CARRON; KENNA, 2012).

Esta rede social apresentou-se disassortativa com índice de $-0,15$, mostrando que vértices com alta conexão se relacionam com vértices de baixa conectividade. Efetuou-se o cálculo das entropias estruturais de ordem zero, um e dois (com estrutura de comunidade). Foi verificado a variação da entropia com estrutura de comunidade para a rede modularizada em até 4 comunidades, constatando-se que quanto maior o número de comunidades da rede menor é sua entropia, ou seja, quanto mais vínculos a rede tiver menor será sua entropia.

4.2 Rede literária do Senhor dos Anéis

Os livros de "O Senhor dos Anéis" constituem a continuação da obra de ficção iniciada pelo autor John Ronald Reuel Tolkien no livro "O Hobbit". A ligação entre estas duas obras se dá

pela continuidade da história do Anel do Poder. Apesar de ter sido escrito entre os anos de 1937 e 1949, e o autor ter vivido na Inglaterra durante a Primeira Guerra Mundial, ele informa que a obra não contém alegorias ou referências aos ocorridos durante a Segunda Guerra Mundial (TOLKIEN, 2000a).

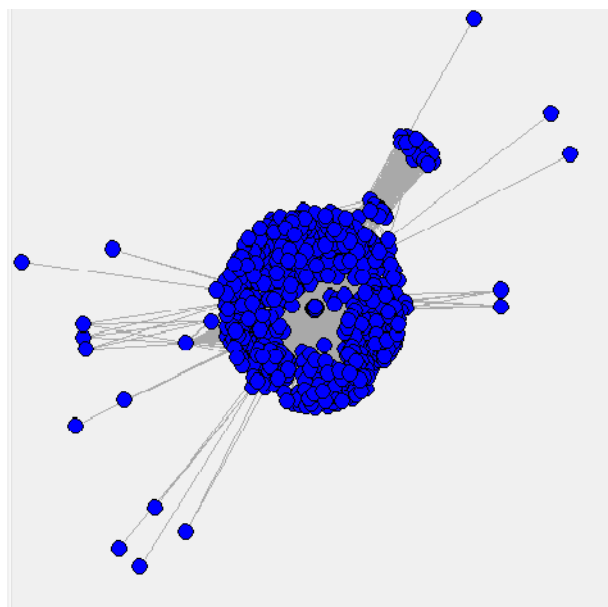
Os livros constituem a narrativa da Grande Guerra do Anel ocorrida na Terra Média, do ponto de vista dos Hobbits. O primeiro livro, chamado *A Sociedade do Anel*, descreve como o anel do poder herdado por Frodo foi assim identificado, a criação e dissolução da Sociedade do Anel. Esta era composta por criaturas de diferentes povos e que têm a missão de ajudar Frodo a destruir o Anel, impedindo que ele retorne ao Senhor do Escuro e este subjogue todos os outros povos (TOLKIEN, 2000a). O segundo livro, chamado *As Duas Torres*, narra a aventura de Frodo e Sam em direção à Mordor e seu encontro com Gollum, além de relatar a vitória dos humanos e elfos em Rohan contra as tropas de Saruman (TOLKIEN, 2000b). O terceiro livro, *O Retorno do Rei*, descreve a união dos humanos, elfos e anões para lutar contra as forças de Sauron e defender Minas Tirith, além da saga de Frodo e Sam rumo a Mordor e a destruição do Anel (TOLKIEN, 2000c).

O Senhor dos Anéis é uma obra de grande complexidade devido aos inúmeros cenários e povos criados, somando-se o alto nível de detalhamento utilizado para descrever estes elementos da narrativa. Cada povo tem características peculiares e diferentes interesses em relação à Guerra do Anel. Além disso, cada criatura carrega consigo as características de seu povo, mas também apresentam ambições e motivações individuais que ficam evidentes no decorrer dos acontecimentos narrados. Em resumo, trata-se de uma longa história de ficção, cujo autor teve por objetivo principal o entretenimento dos leitores. Foi traduzida para vários idiomas e adaptada para o cinema (TOLKIEN, 2000a).

A rede literária social do Senhor dos anéis (Figura 4.2) foi construída com a ligação de seus personagens mediante 368 nós e 19.462 arestas (ligações entre os personagens), tendo grau médio $\langle k \rangle = 63$ arestas.

Para a formação desta rede foi utilizada a obra do escritor J.R.R. Tolkien, onde na mineração de personagens das obras considerou-se as restrições sociais de alusão, compartilhamento de evento e interação direta, formando a conexão entre os personagens. Esta rede social apresentou-se disassortativa com índice de -0,367, mostrando que vértices com alta conexão se relacionam com vértices de baixa conectividade. Efetuou-se no capítulo cinco, o cálculo das entropias estruturais de ordem zero, um e dois (com estrutura de comunidade). Foi verificada a variação da entropia com estrutura de comunidade para a rede modularizada em até 4 comunidades, constatando-se também que quanto maior o número de comunidades da rede menor é sua entropia.

Figura 4.2 – Rede literária social do Senhor dos anéis.



Fonte: O autor.

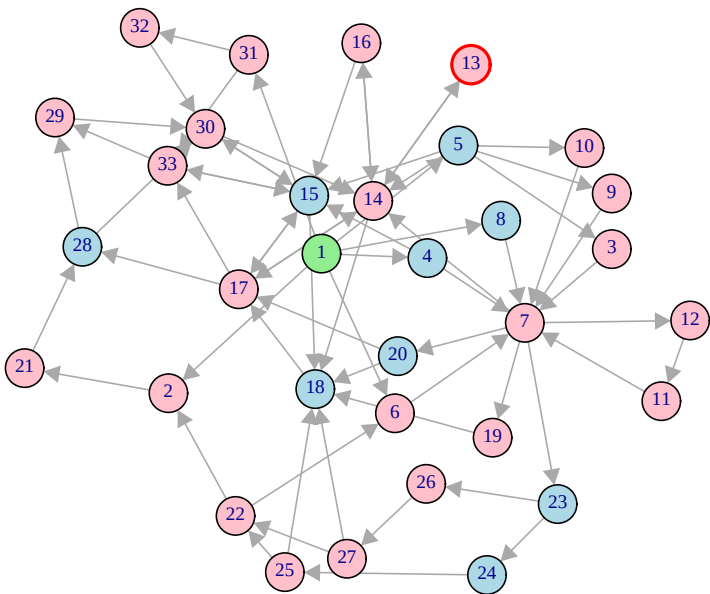
4.3 Rede direcionada da tolerância oral

A tolerância oral pode ser descrita como a supressão de resposta imunológica humoral e/ou celular do organismo a um antígeno, após a administração prévia deste pela via oral. Está relacionada com a prevenção de hipersensibilidade a alimentos e bactérias que fazem parte da flora intestinal, bem como à supressão de resposta imunológica contra o próprio organismo. Por isto, desempenha um papel central na resposta imunológica (WEINER, 2000).

Embora não totalmente compreendida, sabe-se que a tolerância oral se desenvolve em uma complexa atividade de diferentes órgãos, tecidos, células e moléculas presentes no organismo, além de ser dose dependente. O intestino tem papel fundamental neste processo, por estar exposto a inúmeros antígenos provenientes da alimentação. Apesar desta ação local, verifica-se que a tolerância para antígenos solúveis é percebida em todo o animal exposto, o que sugere a participação do sangue ou linfa na disseminação destes antígenos. Enfim, sugere-se que este tipo de tolerância imunológica é um processo multifacetado que envolve processos de indução/manutenção da resposta aos antígenos e migração de células de defesa por todo o organismo, garantindo a regulação da resposta imune em todos os tecidos (PABST; MOWAT, 2012).

Na Figura (4.3), as relações imunológicas associadas ao fenômeno da tolerância oral apresentam 23 componentes celulares (cor de rosa) e 9 componentes humorais (azuis), enquanto que um antígeno genérico é representado pela cor verde.

Figura 4.3 – Rede complexa da tolerância oral.



Fonte: O autor.

A lista dos elementos associados aos números dos vértices da rede está descrita no Quadro 4.1 (MIRANDA et al., 2013)).

Quadro 4.1 – Componentes imunológicos da rede de tolerância oral.

1.Antígeno	12.INF-Lambda	23.RA
2.slgC	13.Th3	24. $\alpha 4\beta 7$ integrina
3.Células calciformes	14.TGE-B	25.MadCAM-1
4.Macrófagos CX3CR1	15.II-10	26.CCR9
5.Enterocito	16.Natural TregFoxP3	27.CCL25
6.slgA	17.iTregFoxP3	28.CD11c ⁺
7.CD103 ⁺	18.NaiveCD4FoxP3 ⁻	29.IL-27
8.Junções oclusivas	19.CCR7	30.Trl
9.Via transcelular	20.IL-2	31.Célula M
10.MHC II	21.FcRn	32.pCDs
11.CD3 ⁺	22.Célula B	33.CD11b ⁺

Fonte: Extraída de MIRANDA et al. (2013).

Conforme Miranda MIRANDA (2014),

“o fenômeno da tolerância oral se refere à tolerância local e sistêmica induzida no intestino delgado após a exposição do tecido mucoso a antígenos inócuos. Foi determinado recentemente que as componentes celulares e moleculares envolvidos nesse processo são de natureza complexa e de interdependência mútua, de modo que a rede de interações é concebível, mas a forma desta fenomenologia complexa ainda é desconhecida. ”

Para a rede de interações da tolerância oral são propostos ataques nos vértices, de modo que cada ataque representa a ausência de um elemento celular ou humoral relacionado no fenômeno. Com isso é possível saber a importância do elemento retirado, utilizando-se o valor da entropia da rede de interações “saudáveis” e o grafo sem o vértice em questão. É importante salientar que, neste contexto, o vértice fonte é o antígeno, pois o processo de estímulo que irá se desenvolver na rede de interação é iniciado por ele. O fluxo de informações, para este contexto, é o quanto um elemento na rede é ativado. Essa aproximação permite o estudo de cada componente imunológico em termos de sua parcela de importância na tolerância oral, que é concebida por uma rede complexa onde ocorre um processo dinâmico de estímulos mútuos, de modo que haja um estado ideal: o vetor de fluxo da rede em seu estado completo ou saudável.

A obtenção dos valores das entropias para cada ataque foram obtidos de duas maneiras: utilizando as entropias estruturais de ordem zero para redes direcionadas S_0^{dir} e as de ordem um S_1^{dir} para ataques obedecendo a sequência de graus e a centralidade de intermediação dos vértices.

4.4 Robustez de redes complexas

Nestes últimos anos, pesquisadores passaram a dar grande importância aos sistemas complexos, uma vez que tais sistemas podem ser modelados como redes de conexões ou redes complexas. Percebeu-se que estruturas complexas dos mais diversos campos da ciência podem ser organizadas e modeladas como redes. Estas estruturas estão presentes em nosso cotidiano, a exemplo das redes de distribuição elétrica, redes sociais, redes rodoviárias, redes de computadores, redes de neurônios, redes literárias dentre outras. Tais redes são passíveis de modelagem, o que possibilita o estudo e análise de estratégias que possam aperfeiçoar tanto seu projeto quanto sua dinâmica.

Para analisar diversos tipos de informação que se propagam por uma rede complexa é necessário conhecer-se a robustez da rede sob o enfoque de falhas aleatórias e ataques intencionais, pois a propagação poderá ser interrompida dependendo de quais vértices sejam atacados. Os ataques simulados devem ser de dois tipos:

- Aleatórios, onde investiga-se o resultado de uma ação casual.

- Intencionais segundo as características topológicas da rede, onde analisa-se o resultado de uma ação planejada sobre alvos potenciais.

Robustez é definida como sendo a capacidade do sistema de manter suas propriedades de comunicação após a remoção de um ou mais vértices. Deste modo, a robustez de uma rede complexa pode ser quantificada pela análise da variação de sua organização quando submetida a falhas e ataques. Barabási (ALBERT; JEONG; BARABÁSI, 2000)) foi um dos primeiros pesquisadores a publicar um trabalho no ano de 2000 sobre a dinâmica de falhas e ataques a redes complexas, analisando como a estrutura de uma rede está relacionada com a robustez, verificando a variação do diâmetro e da componente gigante da Internet quando uma fração de nós é removida.

Sendo a entropia uma das mais importantes medidas de redes, fez-se um estudo da variação das entropias de ordem zero e de ordem um da rede da Odisseia antes e após realizarem-se ataques aleatórios e intencionais na rede com a retirada de 100 nós com maiores (menores) graus e maiores (menores) *betweenness*. A extração destes nós pode desorganizar a estrutura da rede. Também foi calculado a entropia da rede com estrutura de comunidades para fins de comparação com as entropias de ordem zero e de ordem um. Vários tipos de informação podem se propagar numa rede complexa, tal como doenças sendo disseminadas entre indivíduos, troca de pacotes entre computadores na internet.

Para investigar a propagação da informação, é necessário conhecer a robustez da rede sob o ponto de vista de falhas aleatórias e ataques intencionais, pois a propagação pode ser interrompida dependendo dos vértices atacados. Uma tolerância a falhas é definida como a capacidade do sistema manter sua propriedade de comunicação após a remoção de um ou de uma fração de vértices. Sistemas complexos são altamente robustos às falhas aleatórias, como observa-se na internet, onde os roteadores são desligados a todo instante, mas a transmissão de dados a nível mundial é mantida. Para se estudar a tolerância de redes complexas, deve-se analisar sua organização quando os vértices ou arestas são retirados. Como exemplo de robustez de rede, tem-se a internet que possui uma diversidade de caminhos, pelos quais a informação pode ir de um computador a outro, com uma alta taxa de sucesso na entrega de pacotes entre os roteadores. Se ocorrer a falha num roteador, o sistema da internet é capaz de reorientar o caminho pelo qual a informação irá percorrer até alcançar seu destino.

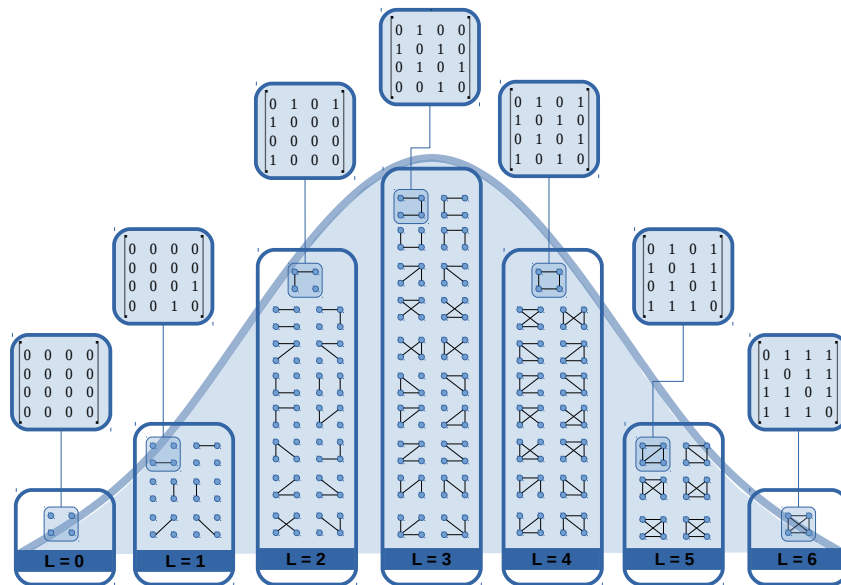
5 ENTROPIA DE REDES - APLICAÇÕES

5.1 Ensembles de redes complexas

Considerando-se o *ensemble* de redes constituído por $N = 4$ e o número de ligações variando de $L = 0$ a $L = 6$, ou seja de um grafo vazio até um grafo completo para análise das entropias de ordem zero e de ordem um. O espaço amostral deste *ensemble* é constituído por

64 grafos (ou seja $2^{\binom{4}{2}} = 2^6 = 64$), conforme Figura 5.1 .

Figura 5.1 – Ensemble de 64 grafos com quatro nós e o número de arestas variando de zero a seis (número máximo).



Fonte: O autor.

Neste conjunto, o número de grafos com $L = 0$ é um (grafo vazio) cuja sequência de graus é $\{0, 0, 0, 0\}$, para $L = 1$ tem-se seis grafos simples distintos com a sequência $\{1, 1, 0, 0\}$. Para $L = 2$, quinze grafos com 12 grafos simples realizando a sequência $\{2, 1, 1, 0\}$ e mais 3 grafos com a sequência $\{1, 1, 1, 1\}$. Para $L = 3$, 20 grafos simples com 3 sequência de graus: 4 representantes da sequência $\{3, 1, 1, 1\}$, 12 da sequência $\{2, 2, 1, 1\}$ e 4 da sequência $\{2, 2, 2, 0\}$. Para $L = 4$, tem-se 15 grafos simples (3 com a sequência de graus $\{2, 2, 2, 2\}$ e 12 realizado pela sequência $\{3, 2, 2, 1\}$); para $L = 5$, tem-se 6 grafos simples realizado pela sequência $\{3, 3, 2, 2\}$ e, para $L = 6$, tem-se um grafo simples regular realizado pela sequência $\{3, 3, 3, 3\}$.

Verifica-se que neste ensemble tem grafos com sequência de graus complementares ou

conjunto de arestas complementares, como no caso da sequência $\{1, 1, 0, 0, \}$ para $L = 1$ com a sequência $\{3, 3, 2, 2\}$ para $L = 5$, o que acarreta a entropia $S_0 = 2,703$ de mesmo valor para os dois conjuntos de redes, ocorrendo o mesmo com os outros grafos complementares. O maior grau desta rede é $k_{max} = N - 1 = 3$. Assim o valor de cada grau complementar é o que falta para completar 3, quando $N = 4$. Pode-se afirmar que a entropia S_0 apresenta-se equivalente em relação aos grafos complementares, ou seja, apresentam o mesmo valor para cada par de grafos complementares.

Para redes k -regulares, pode-se determinar o número de grafos de cada ensemble utilizando a Equação assintótica (SQUARTINI et al., 2015)

$$\Omega_N(k) = \frac{(Nk)!}{\left(\frac{Nk}{2}\right)! 2^{\frac{Nk}{2}} (k!)^N} \exp\left(-\frac{k^2 - 1}{4} - \frac{k^3}{12N} + O\left(\frac{k^2}{N}\right)\right). \quad (5.1)$$

Aplicando-se a fórmula de Stirling e propriedades do cálculo combinatório pode-se reescrever a Equação (5.1):

$$\Omega_N(k) = \frac{\sqrt{2} \left(\frac{Nk}{e}\right)^{\frac{Nk}{2}}}{(k!)^N} \exp\left(-\frac{k^2 - 1}{4} - \frac{k^3}{12N} + O\left(\frac{k^2}{N}\right)\right). \quad (5.2)$$

Exemplos:

No ensemble da Figura 5.1 tem-se para as sequências homogêneas, o número de grafos de:

- $\{1, 1, 1, 1\} \implies \Omega_4(1) = 3,06 \simeq 3$ grafos
- $\{2, 2, 2, 2\} \implies \Omega_4(2) = 2,66 \simeq 3$ grafos
- $\{3, 3, 3, 3\} \implies \Omega_4(3) = 0,622 \simeq 1$ grafo

Para uma rede 3-regular com $N = 6$ nós, realizada pela sequência $\{3, 3, 3, 3, 3, 3\}$, o número de grafos é $\Omega_6(3) = 69,017 \simeq 69 \implies S = \ln(69) = 4,23$, a entropia é 4,23.

5.2 Exemplos de entropias S_0 em redes não direcionadas

Inicialmente será calculada a entropia dada pela Equação 3.65 de outra forma, usando a forma integral da função delta de Kronecker.

Da função partição $Z_0 = \sum_{\{a_{ij}\}} \delta(L - \sum_{i < j} a_{ij}) \cdot \exp\left(\sum_{i < j} h_{ij} a_{ij}\right)$ tem-se que:

$$Z_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{\{a_{ij}\}} e^{ik((L - \sum_{i < j} a_{ij}))} e^{\sum_{i < j} h_{ij} a_{ij}} dk = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikL} \sum_{\{a_{ij}\}} \prod_{i < j} e^{(-ik + h_{ij})a_{ij}} dk,$$

$$Z_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikL} \prod_{i < j} \sum_{\{a_{ij}\}} e^{(-ik + h_{ij})a_{ij}} dk = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikL} \prod_{i < j} (1 + e^{-ik}) dk,$$

$$Z_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikL} (1 + e^{-ik}) \binom{N}{2} dk,$$

$$Z_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikL} \sum_{r=0}^{\binom{N}{2}} \binom{\binom{N}{2}}{r} e^{-ikr}.$$

Fazendo $r = L$, vem:

$$Z_0 = \binom{\binom{N}{2}}{L}$$

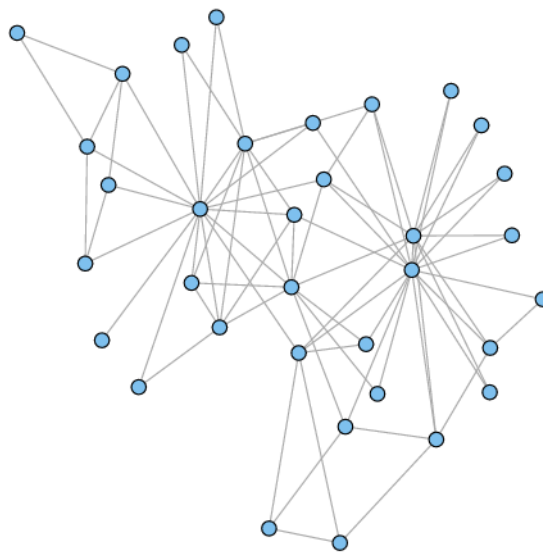
Então a entropia por nó neste ensemble é dada por:

$$NS_0 = \ln(Z_0) = \ln \left[\binom{\binom{N}{2}}{L} \right]$$

que é justamente o valor definido na Equação (3.65).

A rede de amizade de um clube de karatê foi estudada por Wayne W. Zachary de 1970 a 1972. A rede possuía 34 membros, apresentando 78 ligações de pares entre os membros. Durante o estudo surgiu um conflito entre o administrador e o instrutor, o que levou à divisão do clube em dois. Metade dos membros formaram um novo clube em torno do administrador, os membros da outra parte encontrou um novo instrutor ou desistiu de karatê. Rede do Clube de Karatê Zachary com $N = 34$ nós e $L = 78$ arestas, Figura 5.2, cuja entropia com ordem de aproximação zero $S_0 = 6,653003$.

Figura 5.2 – Ilustração da rede de amizade do Clube de Karatê constituído por 34 nós e 78 ligações.



Fonte : O autor.

Para comprovação da Equação (3.65) foram verificados os valores da entropia S_0 das redes das Tabelas 1 e 2 do artigo (BIANCONI, 2008), cujos valores foram confirmados, ou seja:

- Rede da cadeia alimentar Little Rock FW com $N = 183$ nós e $L = 2.494$ arestas. Entropia $S_0 = 38,42$.
- Rede da cadeia alimentar marinha Seagrass com $N = 48$ nós e $L = 226$ arestas. Entropia $S_0 = 11,771$.
- Rede Neural (C. elegans) com $N = 306$ e $L = 2.359$ arestas. Entropia $S_0 = 30,52086$.
- Rede Metabólica com $N = 896$ e $L = 964$ arestas. Entropia $S_0 = 7,562798$.

Também foram calculadas as entropias S_0 das redes literárias (conforme Tabela 5.1):

- Odisseia com $N = 342$ nós e $L = 1.747$ arestas.
- Hobbit com $N = 52$ nós, $L = 547$ arestas.
- Senhor dos anéis com $N = 618$ nós e $L = 19462$ arestas.

Tabela 5.1 – Entropias de ordem zero de redes não direcionadas. As redes assinaladas com * foram extraídas das tabelas de (BIANCONI 2008).

Rede	N	L	S_0
Zachary	34	78	6,65
Little Rock FW*	183	2494	38,42
Seagrass FW*	48	226	11,77
Metabólica*	896	964	7,56
Neural net (C. elegans)*	306	2359	30,52
Odisseia	342	1747	22,94
Senhor dos anéis	618	19462	101,69
Hobbit	52	547	17,28

Fonte: Autor

5.3 Exemplos de entropias de *ensembles* com ordem de aproximação 1: S_1

As entropias de *ensembles* com ordem de aproximação 1, de algumas redes da Figura 5.1, obedecendo a sequência de graus de cada uma, foram determinadas, utilizando a Equação 3.74, obtendo-se os seguintes resultados:

- $N = 4$ e $L = 2 \Rightarrow \{1, 1, 1, 1\}$, a entropia $S_1 = 0,222$.
- $N = 4$ e $L = 3 \Rightarrow \{3, 1, 1, 1\}$, $S_1 = 0,137$
- $N = 4$ e $L = 4 \Rightarrow \{3, 2, 2, 1\}$, $S_1 = 0,4348$
- $N = 4$ e $L = 5 \Rightarrow \{3, 3, 2, 2\}$, $S_1 = 0,201954$
- $N = 4$ e $L = 6 \Rightarrow \{3, 3, 3, 3\}$, $S_1 = 0,18202$

Outros exemplos:

- A rede social do Clube de Karatê Zachary (ZACHARY, 1977) com $N = 34$ nós e a sequência de graus $x = [17 \ 16 \ 12 \ 10 \ 9 \ 6^{(2)} \ 5^{(3)} \ 4^{(6)} \ 3^{(6)} \ 2^{(11)} \ 1]$ arestas. A entropia estrutural de ordem 1, é dada por $S_1 = 3,917485$.
- Rede Littlerock FW com $N = 183$ nós. A sequência de graus e grau médio $\langle k \rangle = 26,79781$. Entropia $S_1 = 27,27$.
- Rede Seagrass FW com $N = 48$ nós. A sequência de graus e grau médio $\langle k \rangle = 9,142857$. Entropia $S_1 = 8,536811$.
- Rede Neural net. (C. Élegans) com $N = 306$. A sequência de grau e grau médio $\langle k \rangle = 14,43$. Entropia $S_1 = 20,45$.
- Rede metabólica com $N = 896$. A sequência de graus e grau médio $\langle k \rangle = 2,0785$. Entropia $S_1 = 4,498$.
- Redes literárias: Odisseia com $N = 342$ nós, sequência de graus e grau médio $\langle k \rangle = 10,216$. Entropia $S_1 = 14,71231$; Hobbit com $N = 52$ nós, sequência de graus e grau médio $\langle k \rangle = 21,03846$. Entropia $S_1 = 9,225819$.

Tabela 5.2 – Entropias de ordem 1 de redes não direcionadas.

Rede	N	$\langle k \rangle$	S_1
Zachary	34	4,59	3,92
Little Rock FW	183	26,80	27,27
Seagrass FW	48	9,14	8,54
Odisseia	342	10,22	14,71
Hobbit	52	21,04	9,23

Fonte : O autor.

Verifica-se que as entropias de ordem 1 são menores que as entropias de ordem zero em virtude de estar condicionada a um maior número de restrições.

5.3.1 Falhas e ataques nas redes complexas

A robustez das redes complexas pode ser analisada mediante dois modos: falhas aleatórias ou ataque. No primeiro caso, os vértices são retirados sem qualquer preferência, ou seja, todos são removidos com a mesma probabilidade. No caso dos ataques, os vértices são removidos de forma intencional, direcionada, selecionando-se desde os vértices mais conectados aos menos conectados. Para cada caso de ataques em vértices, adotou-se as seguintes estratégias:

remoções por ordem decrescente(ascendente) do grau e da centralidade de intermediação, calculado sobre a rede inicial. Na literatura, apresentam-se vários trabalhos com simulação de tais remoções a fim de medirem a variação do diâmetro e do tamanho da maior componente (maior subgrafo conectado) para observar os efeitos ocorridos na estrutura da rede. Neste trabalho, procurou-se analisar a variação da entropia estrutural que é uma medida global da rede.

O primeiro trabalho publicado sobre a dinâmica de falhas e ataques em redes complexas foi no ano de 2000 por Réka Albert (ALBERT; JEONG; BARABÁSI, 2000), em que os pesquisadores estudaram o relacionamento da robustez com a estrutura da rede. Eles analisaram a variação do diâmetro e do maior componente da Internet quando uma fração f de vértices é retirada, concluindo que a Internet é altamente robusta a falhas aleatórias e vulnerável a ataques intencionais.

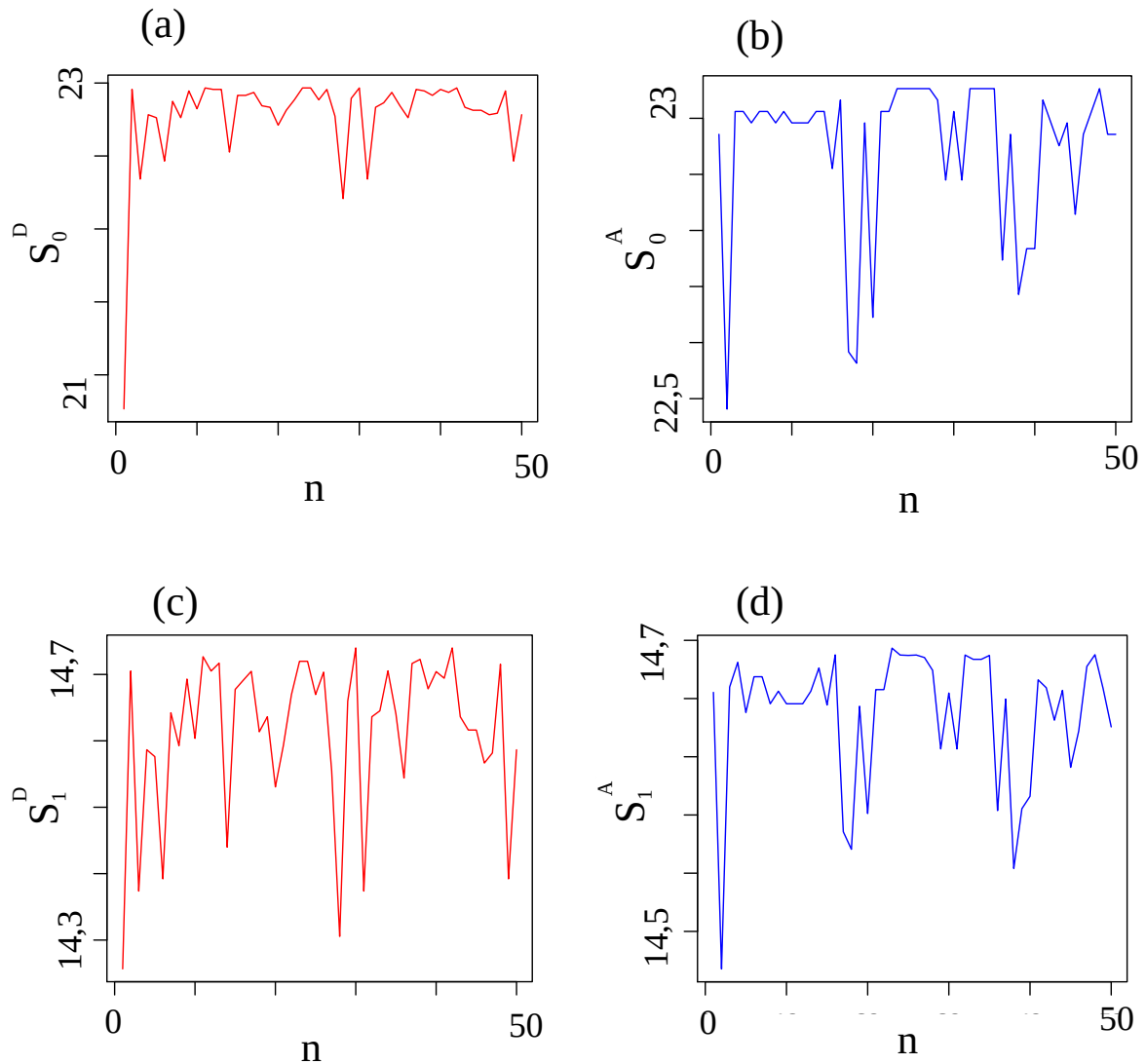
Neste trabalho, a estratégia de ataque adotada foi a de retirar os vértices de maiores graus e centralidades de intermediação, um por um, de forma decrescente e com a reposição do nó retirado para o ataque seguinte. O mesmo procedimento foi adotado na ordem inversa, isto é efetuou-se ataques aos menores graus e centralidades de intermediação na forma crescente. Notou-se que as remoções dos graus são estratégias locais, enquanto as outras duas com base na centralidade de intermediação são globais.

A outra diferença importante entre as estratégias de ataques aos graus e nas centralidades de intermediação é que o primeiro se concentra na redução do número total de arestas da rede enquanto o segundo destina-se a destruir possíveis geodésicas da rede. As redes reais podem sofrer falhas e ataques em sua estrutura que podem danificá-la, dependendo de sua capacidade de recuperação. Assim, analisou-se a robustez das redes reais literárias da Odisseia e Senhor dos Anéis e a rede biológica da tolerância oral sob o enfoque da entropia estrutural, após serem submetidas às estratégias de ataques.

5.3.1.1 Rede social literária da obra Odisseia de Homero

Esta rede possui 342 vértices e 1747 ligações. A entropia calculada da rede inicial, de ordem zero foi $S_0 = 22,94985$ e a de ordem 1, $S_1 = 14,71231$. Efetuando-se ataques intencionais nos cinquenta primeiros nós de maior grau, a entropia S_0 ficou entre os valores $20,76 \leq S_0 \leq 22,96$ enquanto que a entropia por nó S_1 permaneceu no intervalo $14,25 \leq S_1 \leq 14,73$, sobre os cinquenta vértices com menores graus, a entropia S_0 teve a seguinte variação: $22,69 \leq S_0 \leq 22,97$ e S_1 : $14,46 \leq S_1 \leq 14,74$ conforme gráfico da Figura 5.3, onde n representa o número de ataques, S_0^D , a entropia da rede, de ordem zero considerando-se os maiores graus em ordem decrescente; S_0^A , a entropia de ordem zero sobre os menores graus dos vértices em ordem ascendente (crescente). Analogamente, S_1^D é a entropia de ordem 1 sobre os maiores graus dos vértices em ordem decrescente e S_1^A , sobre os menores graus em ordem crescente. Nos ataques aos vértices de maiores graus, as entropias S_0 tiveram uma variação significativa, indicando a vulnerabilidade da rede para estes ataques intencionais.

Figura 5.3 – Entropias S_0 (figuras (a) e (b)) e S_1 (figuras (c) e (d)) relativas aos ataques aos maiores (menores) graus em ordem decrescente (ascendente) dos nós da rede Odisseia. O número de ataques é representado por n .

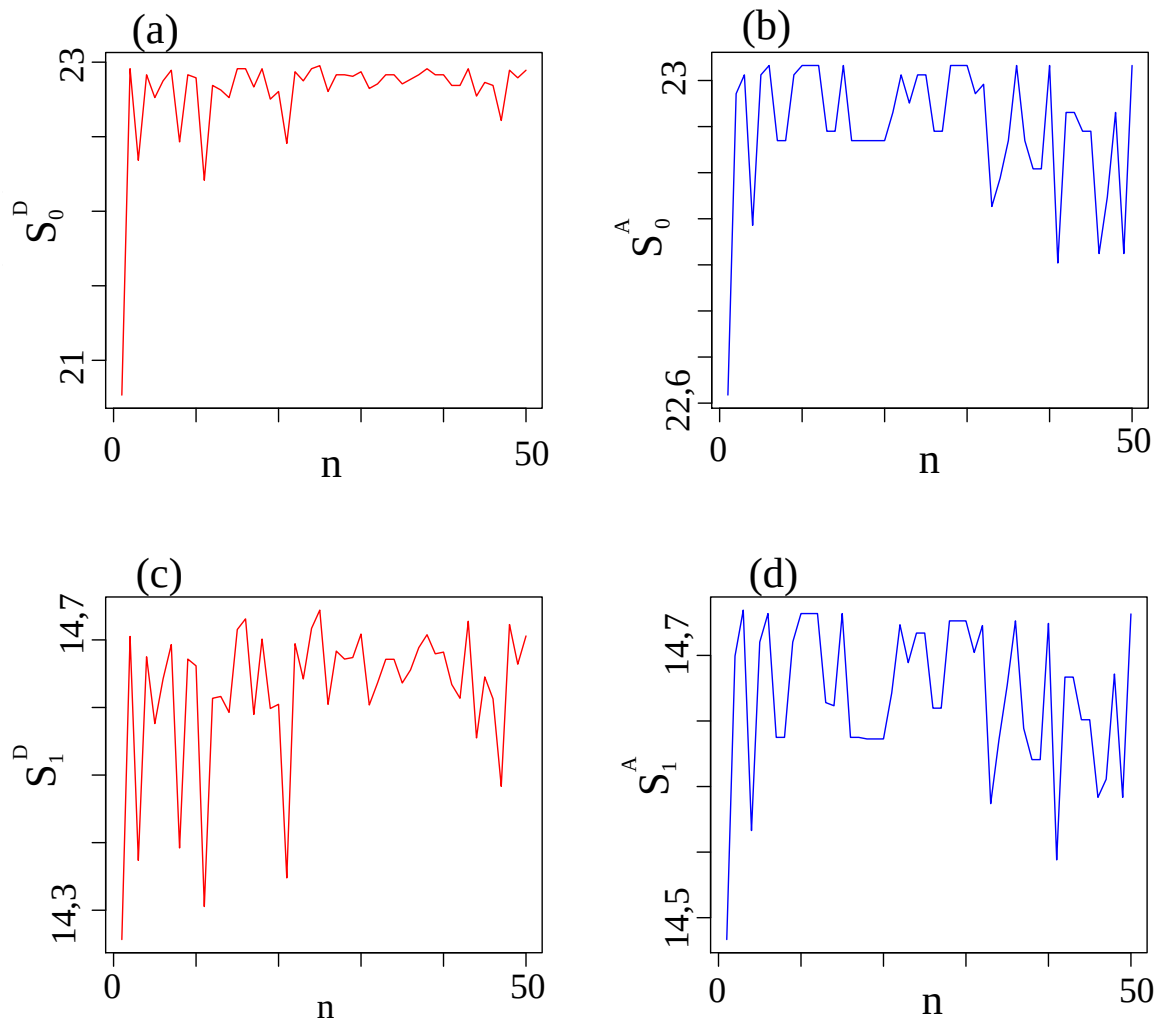


Fonte : O autor.

Nos ataques direcionados aos nós com maiores/menores *betweenness*, tomados em ordem decrescente/crescente (Figura 5.4) obteve-se as seguintes variações das entropias:

- $20,76 \leq S_0 \leq 22,97$ e $14,25 \leq S_1 \leq 14,74$ para as cinquenta maiores.
- $22,69 \leq S_0 \leq 22,96$ e $14,48 \leq S_1 \leq 14,73$ para as cinquenta menores.

Figura 5.4 – Entropias S_0 (figuras (a) e (b)) e S_1 (figuras (c) e (d)) relativas aos ataques aos nós com maiores (menores) centralidades de intermediação da rede Odisseia. O número de ataques é representado por n .



Fonte : O autor.

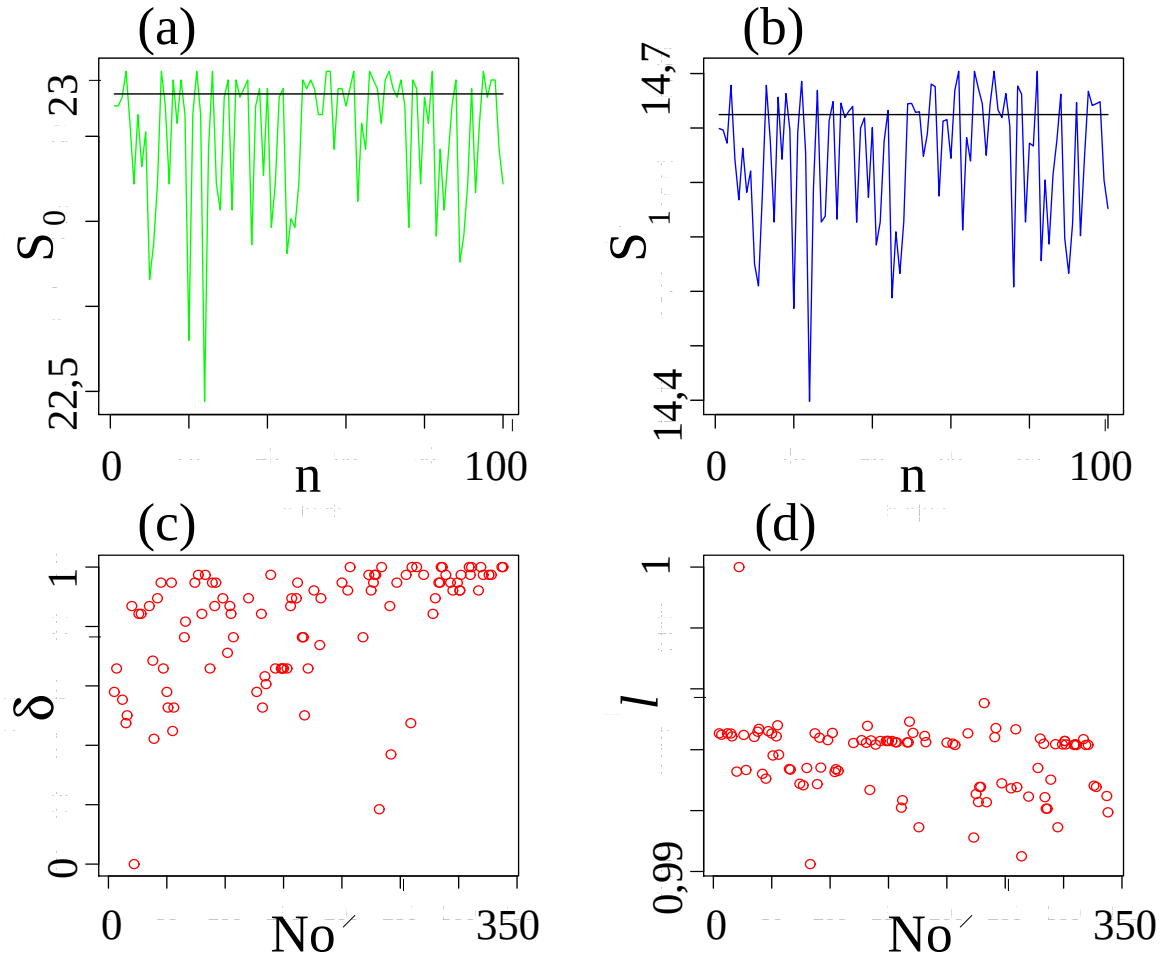
Em seguida, atacando-se aleatoriamente com vértices da rede da Odisseia, obteve-se que as seguintes variações das entropias de ordem zero e de ordem 1:

$22,45 \leq S_0 \leq 22,97$ e $14,34 \leq S_1 \leq 14,74$ conforme o gráfico da Figura 5.5 cujas representações na parte inferior se referem aos nós que foram atacados.

De acordo com a Figura 5.5, nos ataques aleatórios (falhas), cada nó é escolhido ao acaso, gerando assim um conjunto de n de novas redes com entropias de ordem zero:

$S_{min}^0 = S_1^0, \dots, S_n^0 = S_{max}^0$ cujos valores podem ser normalizados pela Equação (DEME-

Figura 5.5 – Entropias S_0 (figura (a)) e S_1 (figura (b)) relativas aos ataques aleatórios à rede da Odisseia. O número de ataques é representado por n . Os nós atacados e o índice δ estão representados na figura (c) e os nós atacados e o menor caminho médio l , na figura (d).



Fonte : O autor.

TRIUS; MANKE, 2005):

$$\delta(S_i^0) = \frac{S_i^0 - S_{\min}^0}{S_{\max}^0 - S_{\min}^0}, \quad (5.3)$$

onde $0 \leq \delta(S_i^0) \leq 1$. Observou-se que:

- Quando a entropia é máxima no intervalo, então $\delta(S_i^0) = \delta(S_{\max}^0) = 1$.
- Quando a entropia é mínima no intervalo, então $\delta(S_{\min}^0) = 0$.
- Para a entropia de ordem 1, valem as mesmas considerações.

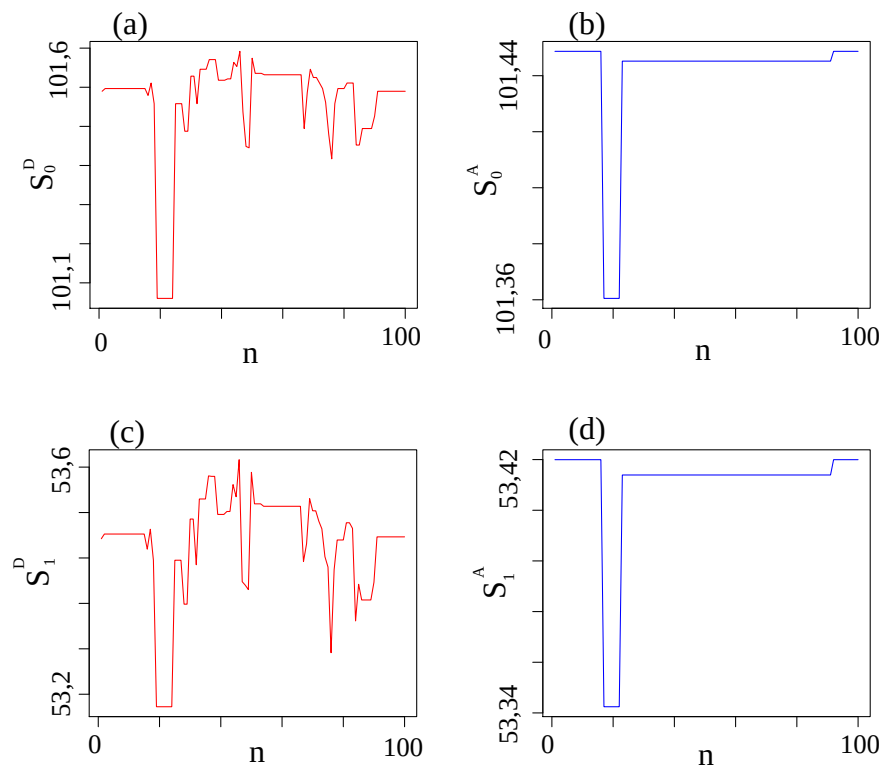
Verificou-se que, nos ataques aleatórios, as entropias apresentaram nos extremos de suas variações, na maioria dos casos, um relacionamento com o caminho médio l entre dois vértices, ou seja quando o ataque maximiza a entropia, l é minimizado. No exemplo dado pela Figura

5.5, o ataque número 53 relativo ao vértice 31, apresenta uma entropia mínima $S_0 = 22,45$ para um caminho médio máximo $l = 2,62$, enquanto que o ataque número 88 referente ao nó 290, obteve-se uma entropia máxima $S_0 = 22,97$, para o caminho médio mínimo $l = 2,57$. Para os ataques da rede através das sequências de graus e das centralidades de intermediação não verificou-se esta relação.

5.3.1.2 Rede literária Senhor dos anéis

Descrita no capítulo 3, é uma rede constituída por 618 sítios (personagens) e 19462 conexões (relações entre os personagens). Suas entropias de ordem zero e de ordem 1 são respectivamente: 101,69 e 53,64. Atacando-se intencionalmente os cem nós de maiores graus da rede obteve-se as seguintes variações das entropias: $101,06 \leq S_0 \leq 101,69$ e $53,17 \leq S_1 \leq 53,71$. O ataque nos cem nós com menores graus apresentou $101,36 \leq S_0 \leq 101,44$ e $53,34 \leq S_1 \leq 53,43$, conforme gráfico da Figura 5.6

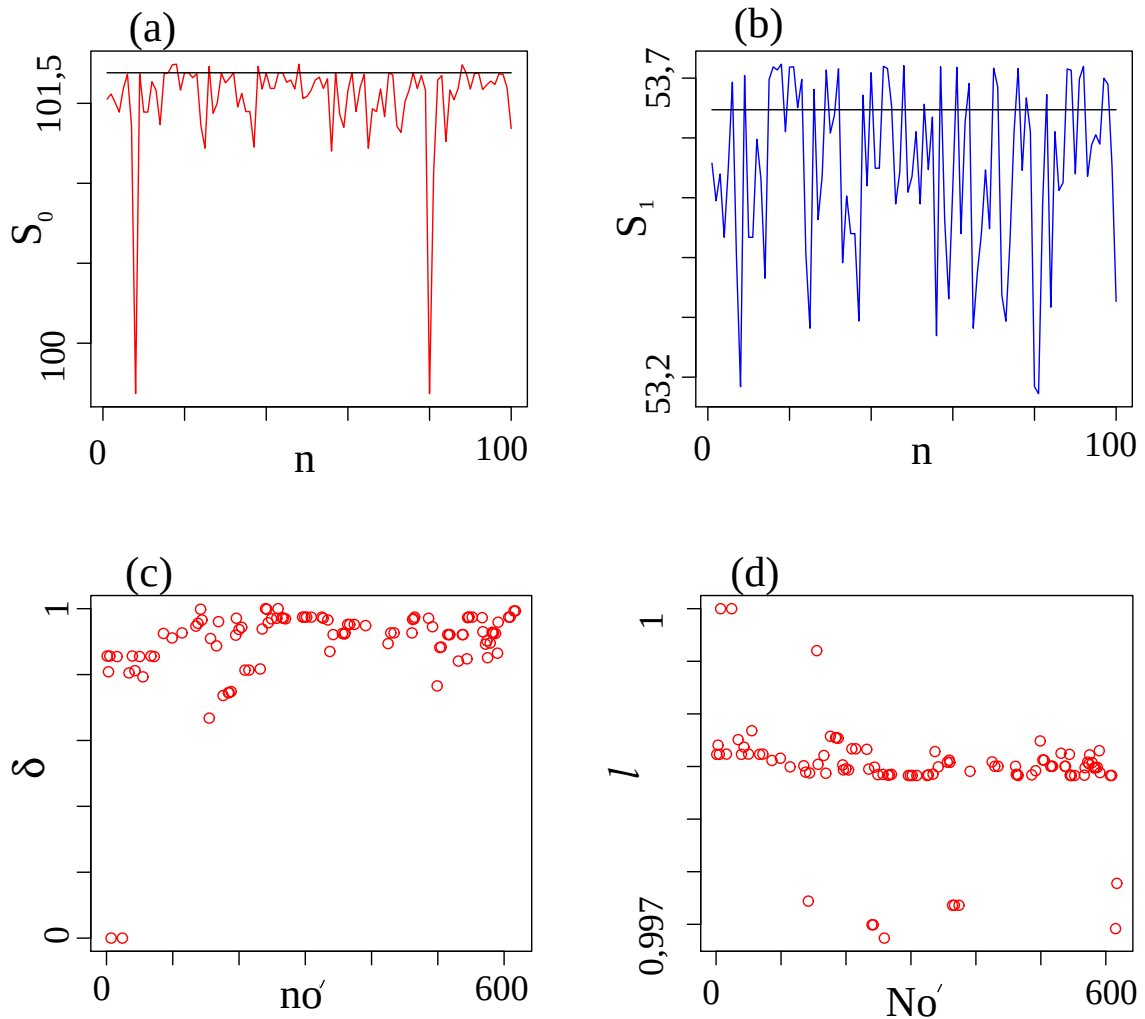
Figura 5.6 – Entropias de ordem zero e de ordem um referentes aos ataques nos 100 maiores/menores graus da rede Senhor dos anéis.



Fonte : O autor.

Os ataques aleatórios em 100 nós, apresentou as seguintes variações das entropias: $99,68 \leq S_0 \leq 101,67$ e $53,17 \leq S_1 \leq 53,68$ conforme Figura 5.7. Nota-se uma variação maior da entropia com ordem de aproximação zero.

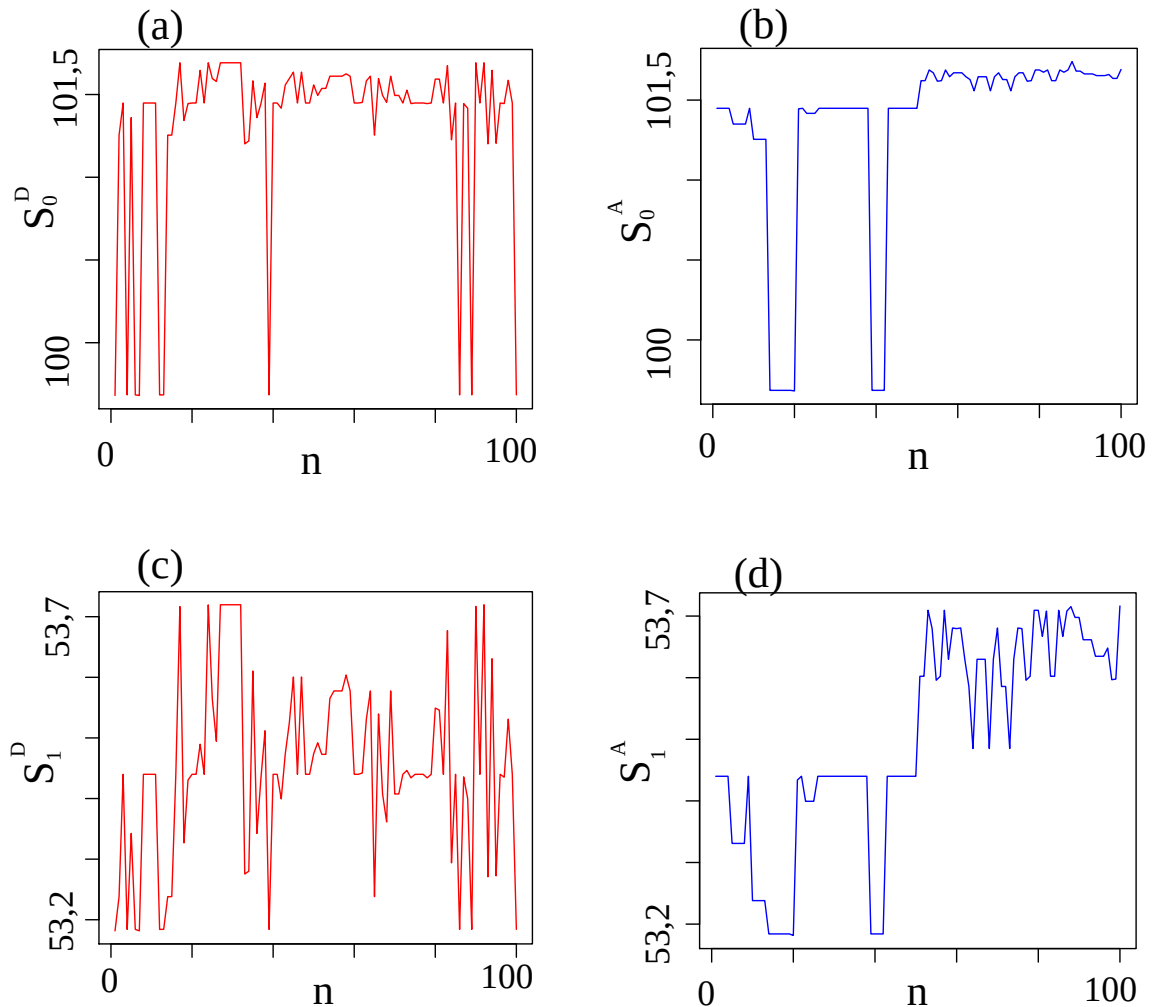
Figura 5.7 – Entropias dos ensembles microcanônico e canônico dos ataques aleatórios em 100 nós da rede Senhor dos Anéis.



Fonte : O autor.

O ataque intencional relativo às 100 maiores centralidades de intermediação (*betweenness centrality*) dos nós, apresentou a seguinte variação das entropias: $99,68 \leq S_0 \leq 101,69$ e $53,18 \leq S_1 \leq 53,71$; com as 100 menores *betweenness*: $99,68 \leq S_0 \leq 101,74$ e $53,18 \leq S_1 \leq 53,71$, conforme Figura 5.8.

Figura 5.8 – Entropias S_0 e S_1 referentes aos ataques aos vértices com maiores(menores) centralidades de intermediação da rede Senhor dos Anéis.



Fonte : O autor.

Observou-se que nos resultados dos ataques, houve uma variação mais significativa da entropia S_0 na rede literária da Odisseia, devido a extração dos vértices de maiores graus, de maiores centralidades de intermediação e também no ataque aleatório, enquanto que nas outras modalidades de ataques, a variação das entropias de ordem zero, foi menor, quando comparadas com a entropia da rede inicial (sem ataque). Verificou-se menor variação nas entropias de ordem um, nas diferentes modalidades de ataques quando comparadas com a entropia S_1 da rede inicial.

Na Tabela 5.3 encontra-se os valores das entropias das redes referentes aos ataques direcionados aos maiores (menores) graus.

Tabela 5.3 – Variação das entropias S_0 e S_1 nos ataques aos maiores/menores graus da rede.

Rede	Graus decrescentes	Graus crescentes
Senhor dos anéis	$101,06 \leq S_0 \leq 101,69$	$101,36 \leq S_0 \leq 101,44$
	$53,17 \leq S_1 \leq 53,71$	$53,34 \leq S_1 \leq 53,43$
Odisseia	$20,76 \leq S_0 \leq 22,96$	$22,69 \leq S_0 \leq 22,97$
	$14,25 \leq S_1 \leq 14,73$	$14,46 \leq S_1 \leq 14,74$

Fonte : O autor.

A variação das entropias relativas aos ataques aleatórios, encontra-se na Tabela 5.4

Tabela 5.4 – Variação das entropias S_0 e S_1 nos ataques aleatórios aos vértices da rede do Senhor dos anéis e da Odisseia.

Rede	Entropias
Senhor dos anéis	$99,68 \leq S_0 \leq 101,63$
	$53,17 \leq S_1 \leq 53,68$
Odisseia	$22,45 \leq S_0 \leq 22,97$
	$14,34 \leq S_1 \leq 14,74$

Fonte : O autor.

A Tabela 5.5 apresenta a variação das entropias das redes após sofrerem ataques às maiores (menores) centralidades de intermediação, seguindo a ordem decrescente (crescente).

Tabela 5.5 – Variação das entropias S_0 e S_1 nos ataques aos vértices de maiores/menores centralidades de intermediação na sequência decrescente (crescente).

Rede	<i>Betweenness</i> decrescentes	<i>Betweenness</i> crescentes
Senhor dos anéis	$99,68 \leq S_0 \leq 101,69$	$99,68 \leq S_0 \leq 101,74$
	$53,18 \leq S_1 \leq 53,71$	$53,18 \leq S_1 \leq 53,71$
Odisseia	$20,76 \leq S_0 \leq 22,97$	$22,69 \leq S_0 \leq 22,96$
	$14,25 \leq S_1 \leq 14,74$	$14,48 \leq S_1 \leq 14,73$

Fonte : O autor.

5.3.2 Entropia S_0 de redes direcionadas

Foi analisada na rede da tolerância oral (MIRANDA et al., 2013) os componentes imunes e suas respectivas interações que geram, mantém e regulam a tolerância oral. A função de partição Z está relacionada com as restrições estruturais das redes, como por exemplo, o número de nós N , o número de ligações L , sequência de graus, estrutura de comunidades. Estes vínculos trazem informações da rede, através dos *hubs*, do número de ligações de cada vértice com seus vizinhos, mediante vértices que são intermediadores do tráfego de informações, dado os caminhos da rede (medida conhecida como centralidade de intermediação). Vértices com estas características formam um tráfego importante das informações da estrutura e portanto são suscetíveis a ataques e falhas. Estes ataques podem desorganizar a rede, alterando sua complexidade.

Baseando-se nestas medidas do sistema microcanônico e canônico respectivamente efetua-se o ataque aos vértices, com exceção do vértice 1 que é o antígeno da rede, com a retirada de um vértice (elemento celular ou humoral) para verificar sua importância, comparando o valor da entropia da rede original com a rede sem o vértice em questão. Estes ataques aos nós foram realizados adotando-se a sequência decrescente de graus e a centralidade de intermediação. A retirada de um vértice da rede permite avaliar sua importância representada no fenômeno da tolerância oral. Assim, o ataque ao vértice 7 (células de CD103+) causou o maior impacto no valor da entropia da rede, estando de acordo com (MIRANDA et al., 2013).

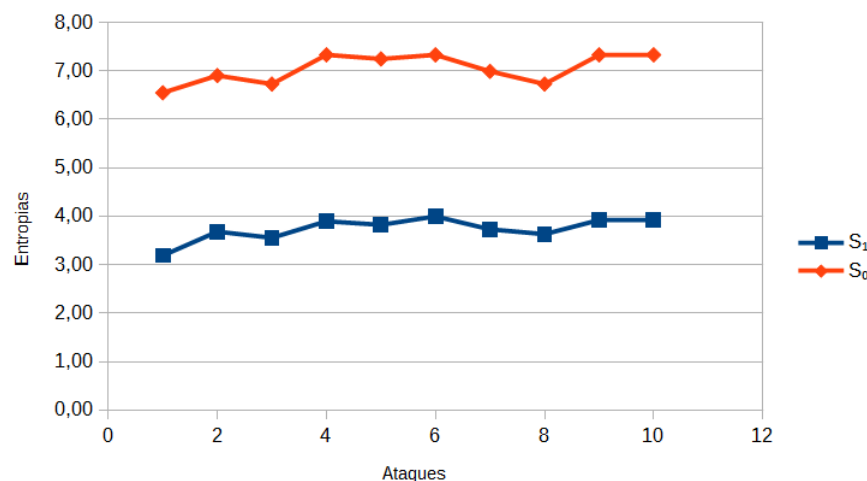
1) Analisou-se a rede direcionada real que descreve o fenômeno da tolerância oral com ataques em 10 vértices com maiores *betweenness* (Tabela 5.6), apresentando a variação das entropias de ordem zero S_0^{dir} e de ordem um S_1^{dir} conforme a Figura 5.9.

Tabela 5.6 – Entropias S_0^{dir} e S_1^{dir} nos ataques aos vértices com maiores centralidades de intermediação.

Vértice KO	<i>Betweenness</i>	Entropia S_1^{dir}	Entropia S_0^{dir}
7 (CD103)	245,75	3,19	6,55
17 (iTregFoxP3)	121,06	3,68	6,90
14 (TGF- β)	111,3	3,56	6,72
23 (RA)	98	3,89	7,33
22 (B-Cell)	83	3,82	7,24
6 (slgA)	59,41	3,99	7,33
18 (Naive CD4FoxP3)	54	3,73	6,99
15 (IL-10)	52,13	3,63	6,72
27 (CCL25)	51,5	3,92	7,33
25 (MadCAM-1)	51,5	3,92	7,33

Fonte : O autor.

Figura 5.9 – Entropias S_0 e S_1 referentes aos ataques nos vértices com maiores centralidades de intermediação.



Fonte : O autor.

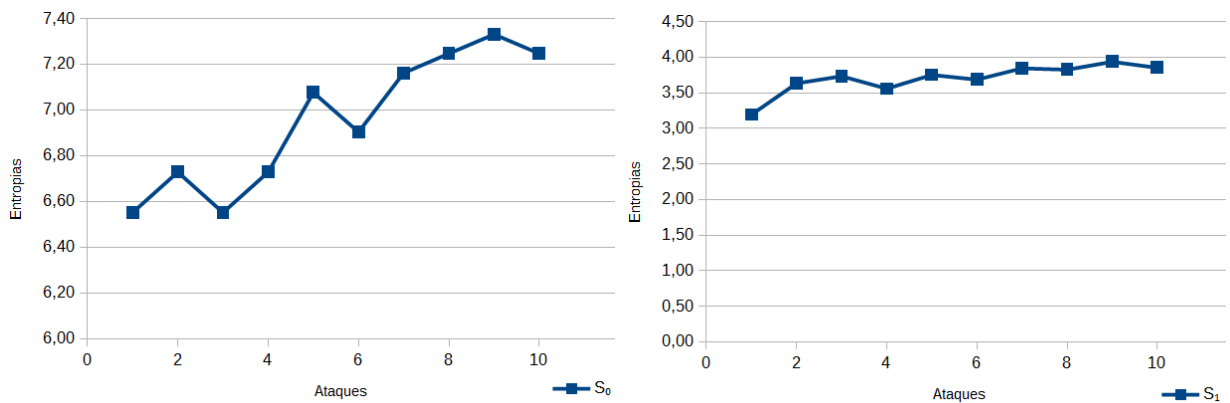
2) Ataques e variação das entropias S_0^{dir} e S_1^{dir} para vértices com maiores graus de entrada (Tabela 5.7) estão representadas na Figura 5.10.

Tabela 5.7 – Entropias S_0 e S_1 nos ataques aos vértices com maiores graus de entrada.

Vértices	Graus de entrada	S_0^{dir}	S_1^{dir}
7	7	6,55	3,19
14	6	6,73	3,56
15	6	6,73	3,63
17	4	6,90	3,68
18	6	6,55	3,73
22	2	7,24	3,82
28	2	7,25	3,85
29	2	7,33	3,93
30	4	7,08	3,75
33	3	7,16	3,84

Fonte: O autor.

Figura 5.10 – Entropias S_0 e S_1 relativas aos ataques dos vértices com maiores graus de entrada.



Fonte : O autor.

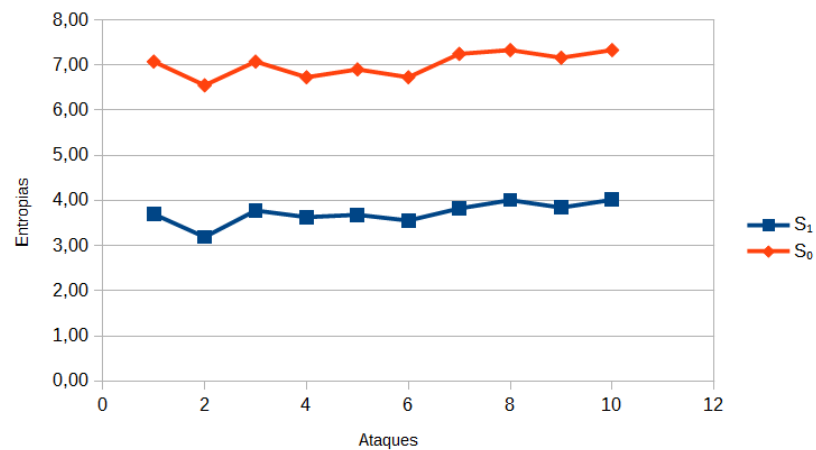
3) As entropias referentes aos ataques nos vértices com maiores graus de saída e variação das entropias de ordem zero e de ordem um, S_0^{dir} e S_1^{dir} respectivamente, estão representadas na Figura 5.11 e descritas na Tabela 5.8.

Tabela 5.8 – Entropias de redes com ataques nos vértices com maiores graus de saída na rede da tolerância oral .

Vértices KO	Grau de saída	Entropias S_1^{dir}	Entropias S_0^{dir}
1	6	3,70	7,07
5	5	3,77	7,07
7	5	3,19	6,55
15	6	3,63	6,72
17	4	3,68	6,90
14	4	3,55	6,73
22	2	3,82	7,24
28	2	3,85	7,24
29	2	3,93	7,33
33	2	3,84	7,16

Fonte : O autor.

Figura 5.11 – Entropias S_0 e S_1 nos ataques aos vértices com maiores graus de saída da rede da Tolerância Oral.



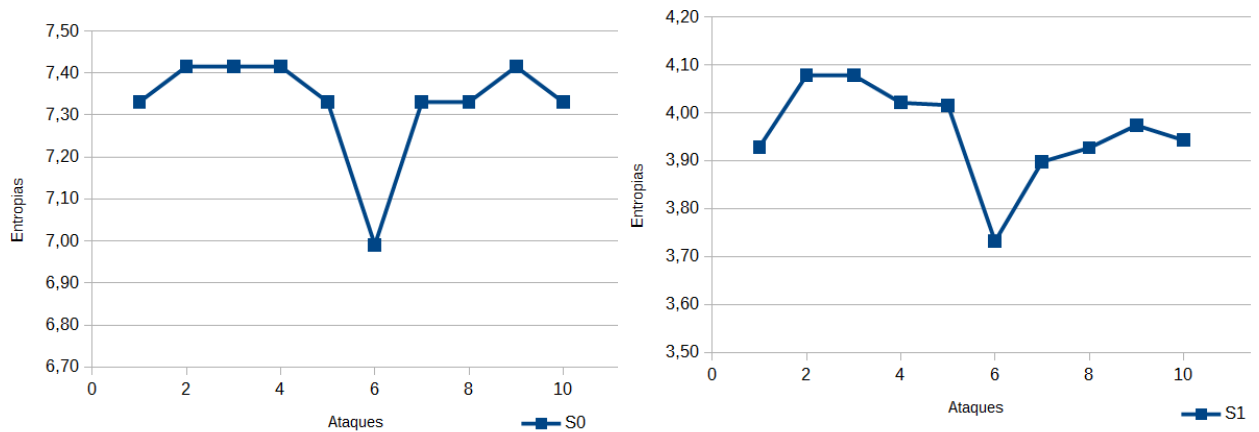
Fonte : O autor.

4) Outra maneira de efetuar os ataques aos vértices é de maneira aleatória, onde retira-se um nó ao acaso e calculam-se as entropias estruturais S_0^{dir} e S_1^{dir} . Os resultados estão apresentados na Tabela 5.9. As falhas (ataques aleatórios) e a variação das entropias S_0^{dir} e S_1^{dir} estão representadas na Figura 5.12.

Tabela 5.9 – Entropias S_0 e S_1 da rede da Tolerância Oral com ataques aleatórios de dez vértices.

Vértices KO	Entropias S_1^{dir}	Entropias S_0^{dir}
2	3,93	7,33
3	4,08	7,41
10	4,08	7,41
11	4,02	7,41
16	4,02	7,33
18	3,73	6,99
23	3,89	7,33
25	3,93	7,33
26	3,97	7,41
31	3,94	7,33

Fonte : O autor.

Figura 5.12 – Entropias S_0 e S_1 da rede Tolerância Oral que sofreu ataques aleatórios nos vértices.

Fonte : O autor.

Os ataques efetuados nesta rede apresentaram uma variação de entropia que depende do vértice a ser retirado. Deste modo, em ataques aleatórios, a estrutura não sofreu mudanças significativas sob o ponto de vista da entropia, enquanto que o ataque seguindo as sequências de graus de entrada e de saída e da centralidade de intermediação, apresentaram um decaimento significativo da entropia. Isto ocorre face a importância dos vértices atacados, diminuindo a entropia e rearranjando a rede, pois ao se retirar os vértices nas sequências de graus de forma decrescente, tiram-se caminhos que transportam informação fazendo com que a entropia seja modificada. As Figuras 5.9, 5.10, 5.11 e 5.12 representam o comportamento das entropias S_0 e S_1 mediante os ataques efetuados na rede direcionada da tolerância oral. O vértice 7, correspondente à célula CD103, foi o que apresentou menor entropia, tanto no ensemble micro-canônico como no canônico ao sofrer ataque aos vértices de maiores *betweenness*, causando grande impacto no processamento de informação, de acordo com (MIRANDA et al., 2013).

5.4 Entropias S_c de redes com estrutura de comunidade

Outra entropia estrutural apresentada em (BIANCONI, 2008) é sobre ensembles de redes com estrutura de comunidades e realizando uma sequência de graus. Supõe-se que a rede possui Q comunidades, tal que cada nó $i, j, \dots, N$ pertence a comunidade $q_i = 1, 2, \dots, Q$, Q finito. Existem vários algoritmos de detecção de comunidades em redes complexas na literatura. No entanto, novas pesquisas tem surgido com o objetivo de detectar comunidades de forma eficiente e com menor custo computacional. Destacam-se cinco algoritmos de detecção de comunidades (*betweenness* (GIRVAN; NEWMAN, 2002), caminhada aleatória (CSARDI; NEPUSZ, 2006), *fastgreedy* (CLAUSET; NEWMAN; MOORE, 2004), autovalores e autovetores e *springlass* (REICHARDT; BORNHOLDT, 2006)) por meio da medida de modularidade, tempo de execução e número de comunidades encontradas.

Newman publicou um trabalho (NEWMAN, 2006) sobre a função de modularidade Q , que mede a qualidade de uma possível divisão da rede em comunidades, ou seja, de uma determinada divisão do grafo ser ou não significativa. Tal função é dada por:

$$Q = \sum_i (e_{ii} - a_i^2), \quad (5.4)$$

em que $a_i = \sum_j e_{ij}$, corresponde à fração de arestas conectadas na comunidade i , o número esperado de arestas intra comunidade é a_i^2 e os elementos e_{ij} representam a fração do total de ligações entre duas comunidades diferentes; a fração de ligações exclusivamente dentro de uma comunidade é denotada por e_{ii} .

Todos os algoritmos escolhidos estão presentes na biblioteca *igraph* do programa R. Os algoritmos tem diferentes abordagens: divisivos, aglomerativos, espectrais e otimização da modularidade. Os métodos selecionados foram:

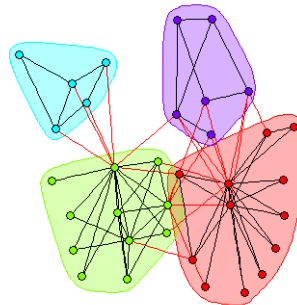
- *Betweenness*: algoritmo divisivo que se baseia na contagem de caminhos mínimos entre vértices. Dessa forma, definem-se quais arestas com valor baixo de *betweenness* pertencem à mesma comunidade; arestas com alto valor de *betweenness* separam diferentes comunidades. Baseia-se na remoção das arestas entre as comunidades.
- Caminhada aleatória (*walktrap*): algoritmo em que caminhos aleatórios curtos tendem a estar na mesma comunidade.
- *Fastgreedy*: algoritmo que otimiza a medida de modularidade original de Newman utilizando uma busca “gulosa”.
- Autovetores: algoritmo que reformula o conceito de modularidade em termos dos autovetores e autovalores de uma nova matriz, a matriz de modularidade. É denominado método espectral.

- *Spinglass*: algoritmo que otimiza a modularidade de Newman usando *Simulated Annealing*. É um método heurístico eficiente utilizado para localizar o mínimo absoluto de funções com muitos mínimos locais.

Para calcular a entropia de conjuntos de redes randomizados com dada estrutura de comunidade, impõe-se que os vértices tenham uma sequência de graus fixada e que se tenha um número fixo de arestas entre as comunidades. Foram determinadas as entropias com ordem de aproximação $2(S_2)$ ou com estrutura de comunidade (S_c) utilizando-se as expressões do Capítulo 3 para os seguintes exemplos:

1. Rede do clube karate (Zachary club) (Figura 5.13)

Figura 5.13 – Rede do Clube Karatê com quatro comunidades construída no igraph.



Fonte : O autor.

Tabela 5.10 – Entropias S_c com valores das modularidades para rede Zachary modulada em duas, três e quatro comunidades.

Comunidades	Modularidade	Entropia S_c
2	0,39	3,27
3	0,40	2,96
4	0,42	2,87

Fonte : O autor.

O valor da entropia encontrada para duas comunidades foi aproximadamente igual ao encontrado em (BIANCONI, 2008). Verifica-se que a menor entropia é referente à divisão da rede em 4 comunidades, indicando melhor organização entre seus elementos.

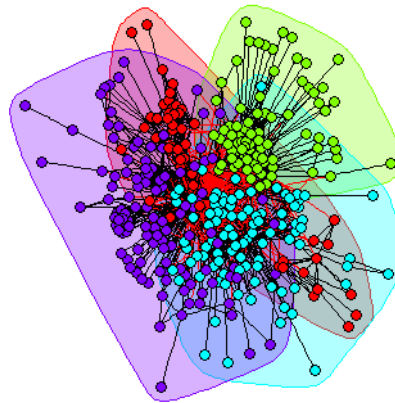
2. Rede literária da Odisseia (Figura 5.14)

Os resultados encontrados para as entropias S_c estão apresentados na Tabela 5.11. A rede foi dividida em 2, 3 e posteriormente em 4 comunidades. Também verificou-se que a menor entropia ocorre na rede com quatro comunidades, em virtude de ter mais vínculos, indicando melhor organização de seus elementos.

3. Rede literária Senhor dos Anéis (Figura 5.15)

Esta rede foi construída com a ligação de seus personagens representados por 368 nós com 19462 ligações, tendo grau médio $\langle k \rangle = 63$, assortatividade de -0,367 e centralidade de intermediação média de 304,894.

Figura 5.14 – Rede da Odisseia com 4 comunidades.



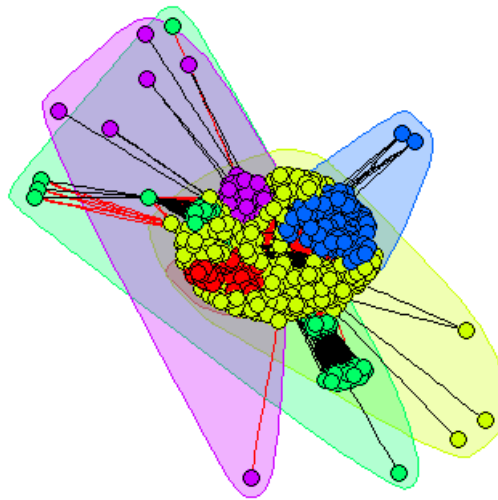
Fonte : O autor.

Tabela 5.11 – Entropias S_c da rede literária Odisseia modulada de duas a quatro comunidades.

Comunidades	Modularidade	Entropia S_c
2	0,36	14,06
3	0,49	12,75
4	0,53	11,79

Fonte : O autor.

Figura 5.15 – Rede literária Senhor dos anéis com 5 comunidades.



Fonte : O autor.

A entropia S_c da rede do Senhor dos Anéis, foi determinada para a rede modulada em 2, 3, 4 e 5 comunidades, cujos valores são apresentados na Tabela 5.12.

Tabela 5.12 – Entropia S_c da rede literária O Senhor dos Anéis modulada de duas até cinco comunidades com os respectivos valores das modularidades.

Comunidades	Modularidade	Entropia S_c
2	0,31	52,82
3	0,41	44,71
4	0,427	40,50
5	0,426	37,11

Fonte : O autor.

Verifica-se que para cinco comunidades, a rede social literária apresenta uma entropia menor, indicando que houve melhor organização dos elementos em relação às modulações anteriores, além de ser imposta maior número de vínculos. Estas comunidades da estrutura social foram identificadas (RIBEIRO et al., 2015), calculando-se as medidas em cada comunidade, conforme Tabela 5.13 onde $\langle k \rangle$ representa o grau médio dos vértices, T a transitividade, A a assortatividade, $\langle B \rangle$ a centralidade de intermediação média, d a distância de caminho médio, N o número de vértices, E o número de arestas e S_1 a entropia estrutural de ordem 1, calculada em cada comunidade.

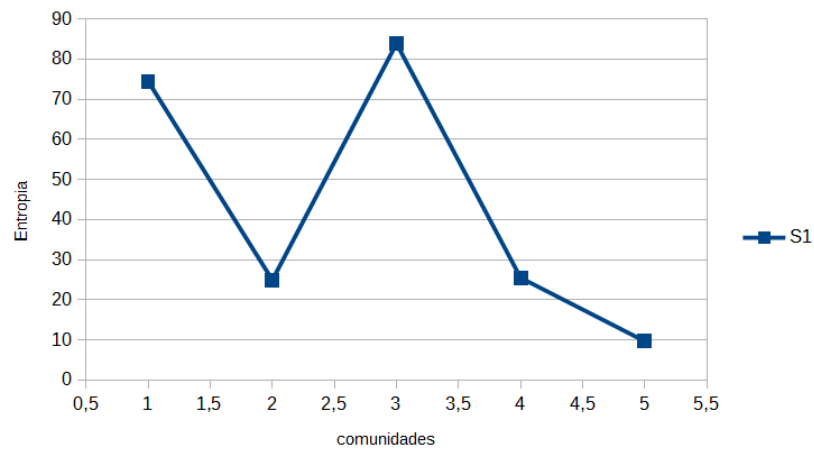
Tabela 5.13 – Propriedades estruturais das cinco comunidades da rede social Senhor dos anéis.

Comunidades	$\langle k \rangle$	T	A	$\langle B \rangle$	d	N	E	S_1
1	39	0,894	-0,5567	67,47	1,896	331	6595	74,43
2	64	0,958	-0,5355	61,82	0,997	77	2494	24,94
3	67	0,829	-0,496	502,13	1,574	151	5105	83,85
4	21	0,989	-0,4294	22,27	1,201	28	305	25,49
5	28	0,966	-0,3671	362,93	1,062	31	436	9,68

Fonte: (RIBEIRO et al., 2015).

As comunidades apresentaram valores de transitividade altos, indicando alta densidade em ligações, caracterizada pelo número de triângulos formados pela relação social. Dentre as comunidades, destaca-se a comunidade 3 (Figura 5.16) como a mais heterogênea em ligações, pois sua entropia estrutural recebe a restrição da sequência de graus. Assim, a estrutura das obras de (TOLKIEN, 2000a) é densa em ligações e suas comunidades são heterogêneas com o vínculo da sequência de graus.

Figura 5.16 – Entropia nas cinco comunidades da rede O senhor dos Anéis, destacando-se a comunidade 3 com maior entropia, pois é a mais heterogênea em ligações.



Fonte (RIBEIRO et al., 2015).

Foram calculadas com auxílio do programa R, utilizando o pacote *igraph*, *matlab* e *Matrix*, as entropias de ensembles de redes complexas com estrutura de comunidade e com uma sequência de graus. Para as modularidades foi adotado o processo *spinglass* do programa R, o qual permite que se estipule o número de comunidades a priori, enquanto que nos outros métodos, a modularidade é determinada pelo programa.

6 CONCLUSÃO

Considerada como uma área relativamente nova da Ciência, a teoria das redes complexas, baseia-se em dados empíricos tais como aqueles que são obtidos das interações biológicas, sociais e tecnológicas. A complexidade de uma estrutura de rede depende da organização global que as ligações exercem com seus elementos. As redes reais apresentam diferentes níveis de organização em virtude da heterogeneidade das conexões entre os vértices. Para mensurar o grau de organização ou complexidade de uma rede, utilizou-se a de entropia da Mecânica Estatística, entropia esta que é calculada com vínculos estruturais, como a sequência de graus, número de ligações fixas, número de vértices fixo, densidade de ligações e estrutura de comunidades (BIANCONI (2008),BIANCONI (2009)).

Neste trabalho, utilizando a entropia estrutural, analisou-se a robustez das redes complexas não direcionadas de caráter literário social (da Odisseia e do Senhor dos anéis) e de redes direcionadas de caráter biológico (Tolerância oral), com a utilização das entropias estruturais de ordem de aproximação zero e um, ao serem submetidas a ataques aleatórios e intencionais mediante sequências de grau e da centralidade de intermediação.

O índice $\delta(S_0^i)$ função das variações das entropias de ordem zero S_0 , verificadas em uma série de ataques (100) na rede da Odisseia, permitiu verificar uma relação existente entre as entropias extremas destes ataques com o comprimento do menor caminho médio. Quando a entropia de uma rede que sofreu o ataque a um vértice for máxima, o menor caminho médio é mínimo e quando a entropia for mínima, o comprimento do menor caminho médio tem valor máximo durante o ataque, na maioria dos casos, sendo que o índice vale um, quando a entropia de um ataque i for máxima e o comprimento do caminho médio é mínimo, ao passo que quando o índice for zero, a variação é nula e o comprimento do caminho médio é máximo. Observou-se que a rede é menos robusta a ataques intencionais, onde ocorreu uma variação significativa da entropia de ordem zero S_0 e é resiliente a ataques aleatórios, onde a variação das entropias foi menor.

Na rede do Senhor dos Anéis, verificou-se uma significativa mudança nas entropias de ordem zero e de ordem um quando submetida aos ataques intencionais, seguindo a sequência de centralidade de intermediação decrescente e ataques aleatórios, onde a rede mostrou-se vulnerável, enquanto que mostrou-se mais robusta para ataques direcionados seguindo a sequência de graus em ordem decrescente. Modulou-se esta rede em cinco comunidades, calculando-se as entropias S_1 de cada uma, verificando-se que a comunidade três apresentou maior valor da entropia, devido a característica da heterogeneidade das ligações.

A rede literária em questão foi modulada em comunidades aplicando-se o método *walkstrap* e do *spinglass*, onde as entropias estruturais tinham mais um vínculo: o da estrutura de comunidades. Em virtude das redes com estrutura de comunidades terem mais vínculos estruturais que as anteriores, as entropias S_c dessas redes foram menores que as entropias S_0 e S_1 .

Na rede biológica da tolerância oral, os ataques ocorreram de forma aleatória e de forma

intencional, atacando os vértices na sequência decrescente dos graus de entrada k_{in} , dos graus de saída k_{out} e dos vértices com maiores centralidades de intermediação. Verificou-se uma variação significativa das entropias estruturais nos ataques aos vértices com maiores centralidades de intermediação e com maiores graus de entrada, indicando que a rede é vulnerável para estes ataques, enquanto que nos outros ataques, a variação entrópica foi menor, indicando que a rede é robusta para esta estratégia de ataques. O ataque do vértice 7 (célula CD103) provocou o menor valor das entropias indicando que a rede diminui a robustez para a remoção deste vértice.

Nestes estudos, verificou-se a importância da entropia, tida como a principal medida global de redes e sua aplicabilidade na teoria das redes complexas, apresentando algumas propriedades, das quais destaca-se que a entropia da rede possui valores que dependem da dinâmica de remoção de vértices. A entropia está correlacionada com a robustez. Valores máximos da entropia surgem quando aumenta a robustez, enquanto que nos valores mínimos, a robustez é reduzida.

Como trabalho futuro, diversas investigações podem ser realizadas considerando outras medidas topológicas das redes relacionadas com a entropia. Remoção de uma fração de vértices ou de arestas poderão ser investigadas através de abordagem semelhante. Analisar os ataques às redes sob o enfoque da entropia estrutural de comunidades S_c ou de ordem dois S_2 . A verificação da robustez de uma rede complexa poderá ser feita com a adição de nós à rede, analisando-se as entropias das redes atacadas, utilizando como exemplo uma rede real de “*power grid*” (redes elétricas).

REFERÊNCIAS

- ABREU, N. et al. **Teoria Espectral de Grafos-Uma Introdução**. [S.l.]: SBMAC, 2012.
- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. **Reviews Of Modern Physics**, APS, v. 74, n. 1, p. 47, 2002.
- ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Error and attack tolerance of complex networks. **Nature**, Nature Publishing Group, v. 406, n. 6794, p. 378–382, 2000.
- AMARAL, L. A.; OTTINO, J. M. Complex networks. **The European Physical Journal B-Condensed Matter and Complex Systems**, Springer, v. 38, n. 2, p. 147–162, 2004.
- ANAND, K.; BIANCONI, G. Entropy measures for networks: Toward an information theory of complex topologies. **Physical Review E**, APS, v. 80, n. 4, p. 045102, 2009.
- ANTIQUERA, L. et al. Strong correlations between text quality and complex networks features. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 373, p. 811–820, 2007.
- BARABÁSI, A. **Linked how Everything is Connected to Everything Else and what it Means for Business, Science, and Everyday Life**. [S.l.]: A Plume Book, 2002.
- BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, American Association for the Advancement of, v. 286, n. 5439, p. 509–512, 1999.
- BARBIERI, A. L. **Análise de robustez em redes complexas**. Tese (Doutorado) — Universidade de São Paulo, 2010.
- BAUER, M.; BERNARD, D. Maximal entropy random networks with given degree distribution. **arXiv preprint cond-mat/0206150**, 2002.
- BERG, J.; LÄSSIG, M. Correlated random networks. **Physical Review Letters**, APS, v. 89, n. 22, p. 228701, 2002.
- BIANCONI, G. The entropy of randomized network ensembles. **EPL (Europhysics Letters)**, IOP Publishing, v. 81, n. 2, p. 28005, 2008.
- BIANCONI, G. Entropy of network ensembles. **Physical Review E**, APS, v. 79, n. 3, p. 036114, 2009.
- BIANCONI, G.; GULBAHCE, N.; MOTTER, A. E. Local structure of directed networks. **Physical Review Letters**, APS, v. 100, n. 11, p. 118701, 2008.
- BOCCALETTI, S. et al. Complex networks: Structure and dynamics. **Physics Reports**, v. 424, p. 175–308, fev. 2006.
- BOGACZ, L.; BURDA, Z.; WACŁAW, B. Homogeneous complex networks. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 366, p. 587–607, 2006.
- CALLAWAY, D. S. et al. Network robustness and fragility: Percolation on random graphs. **Physical Review Letters**, APS, v. 85, n. 25, p. 5468, 2000.
- CANCHO, R. F. I.; SOLÉ, R. V. Optimization in complex networks. In: **Statistical mechanics of complex networks**. [S.l.]: Springer, 2003. p. 114–126.

CARRON, P. M.; KENNA, R. Universal properties of mythological networks. **EPL (Europhysics Letters)**, IOP Publishing, v. 99, n. 2, p. 28002, 2012.

CLAUSET, A.; NEWMAN, M. E.; MOORE, C. Finding community structure in very large networks. **Physical review E**, APS, v. 70, n. 6, p. 066111, 2004.

COHEN, R.; HAVLIN, S. **Complex networks: structure, robustness and function**. Cambridge, New York: Cambridge University Press, 2010.

CSARDI, G.; NEPUSZ, T. The igraph software package for complex network research. **InterJournal, Complex Systems**, v. 1695, n. 5, p. 1–9, 2006.

DEMETRIUS, L.; MANKE, T. Robustness and network evolution-an entropic principle. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 346, n. 3, p. 682–696, 2005.

DOROGOVTSSEV, S.; MENDES, J.; SAMUKHIN, A. Principles of statistical mechanics of uncorrelated random networks. **Nuclear Physics B**, v. 666, n. 3, p. 396 – 416, 2003. ISSN 0550-3213.

DOROGOVTSSEV, S. N.; GOLTSEV, A. V.; MENDES, J. F. F. Critical phenomena in complex networks. **Rev. Mod. Phys.**, American Physical Society, v. 80, p. 1275–1335, 2008.

FARKAS, I. et al. Equilibrium statistical mechanics of network structures. In: **Complex networks**. [S.l.]: Springer, 2004. p. 163–187.

FRITSCHER, E. **Propriedades espectrais de um grafo**. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, 2011.

GIRVAN, M.; NEWMAN, M. E. Community structure in social and biological networks. **Proceedings of the national academy of sciences**, National Acad Sciences, v. 99, n. 12, p. 7821–7826, 2002.

HOMERO. **Odisseia**. São Paulo: Atena, 1960.

JAYNES, E. T. Information theory and statistical mechanics. **Physical review**, APS, v. 106, n. 4, p. 620, 1957.

JEONG, H. et al. The large-scale organization of metabolic networks. **Nature**, Nature Publishing Group, v. 407, n. 6804, p. 651–654, 2000.

JI, L. et al. Network entropy based on topology configuration and its computation to random networks. **Chinese Physics Letters**, IOP Publishing, v. 25, n. 11, p. 4177, 2008.

LOPES, G. A. W. **Um Modelo de Rede Complexa para Análise de Informações Textuais**. Dissertação (Mestrado) — Dissertação apresentada ao Curso de Mestrado em Inteligência Artificial Aplicada à Automação Industrial do Centro Universitário da FEI, São Paulo, 2011.

MIRANDA, P. J. **Emergência e Fluxo de Informação em Redes Complexas**. Dissertação (Mestrado) — Universidade Estadual de Ponta Grossa, 2014.

MIRANDA, P. J.; BAPTISTA, M. S.; PINTO, S. E. D. S. Analysis of communities in a mythological social network. **arXiv preprint arXiv:1306.2537**, 2013.

MIRANDA, P. J. et al. The oral tolerance as a complex network phenomenon. **arXiv preprint arXiv:1311.3237**, 2013.

- MONTEIRO, L. H. A. **Sistemas dinâmicos complexos**. São Paulo: Editora Livraria da Física, 2010.
- NEWMAN, M. **Networks: An Introduction**. [S.l.]: OUP Oxford, 2010. ISBN 9780199206650.
- NEWMAN, M.; BARABASI, A.-L.; WATTS, D. J. **The structure and dynamics of networks**. [S.l.]: Princeton University Press, 2006.
- NEWMAN, M. E. Modularity and community structure in networks. **Proceedings of the national academy of sciences**, National Acad Sciences, v. 103, n. 23, p. 8577–8582, 2006.
- PABST, O.; MOWAT, A. Oral tolerance to food protein. **Mucosal immunology**, Nature Publishing Group, v. 5, n. 3, p. 232–239, 2012.
- REICHARDT, J.; BORNHOLDT, S. Statistical mechanics of community detection. **Physical Review E**, APS, v. 74, n. 1, p. 016110, 2006.
- RIBEIRO, M. A. et al. A rede social complexa de o senhor dos anéis. **Revista Brasileira Ensino de Física**, 2015.
- SOLÉ, R. V.; VALVERDE, S. Information theory of complex networks: On evolution and architectural constraints. In: **Complex networks**. [S.l.]: Springer, 2004. p. 189–207.
- SQUARTINI, T. et al. Breaking of ensemble equivalence in networks. **arXiv preprint arXiv:1501.00388**, 2015.
- TOLKIEN, J. R. R. **O Senhor dos Aneis - Vol. 1 - A Sociedade do Anel**. [S.l.]: Martins Fontes, 2000.
- TOLKIEN, J. R. R. **O Senhor dos Aneis - Vol. 2 - As Duas Torres**. [S.l.]: Martins Fontes, 2000.
- TOLKIEN, J. R. R. **O Senhor dos Aneis - Vol. 3 - O Retorno do Rei**. [S.l.]: Martins Fontes, 2000.
- WANG, B. et al. Entropy optimization of scale-free networks robustness to random failures. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 363, n. 2, p. 591–596, 2006.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of small-world networks. **nature**, Nature Publishing Group, v. 393, n. 6684, p. 440–442, 1998.
- WEINER, H. L. Oral tolerance, an active immunologic process mediated by multiple mechanisms. **The Journal of Clinical Investigation**, The American Society for Clinical Investigation, v. 106, n. 8, p. 935–937, 10 2000. Disponível em: <<http://www.jci.org/articles/view/11348>>.
- ZACHARY, W. W. An information flow model for conflict and fission in small groups. **Journal of Anthropological Research**, JSTOR, p. 452–473, 1977.
- ZHAO, K. et al. Entropy rate of nonequilibrium growing networks. **Physical Review E**, APS, v. 84, n. 6, p. 066113, 2011.

APÊNDICE A – MÉTODO DO PONTO DE SELA

MÉTODO DO PONTO DE SELA

Uma ferramenta útil para cálculo aproximado de integrais é o método do ponto de sela, limitado aqui ao eixo real. Neste caso, o método está baseado no desenvolvimento de Taylor

$$I = \int_a^b \exp(f(x))dx \simeq \int_a^b \exp(f(x_0)) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$$

Supondo, que $f(x)$ tenha um único máximo em $x = x_0$, com $x_0 \in [a, b]$, e que $f''(x_0) \neq 0$, tem-se que: $f'(x_0) = 0$ e $f''(x_0) < 0$. Considerando que pode-se estender o intervalo $[a, b]$ até o intervalo $(-\infty, +\infty)$ com $f(\pm\infty) = 0$, a integral I torna-se uma gaussiana, ou seja:

$$I \simeq \int_{-\infty}^{+\infty} \exp(f(x_0)) \cdot \exp\left(-\frac{1}{2}|f''(x_0)|(x - x_0)^2\right) dx = \sqrt{\frac{2\pi}{|f''(x_0)|}} \exp(f(x_0)).$$

O cálculo da integral fica reduzido a encontrar o valor crítico da função $f(x)$.

Além disso, para calcular a integral aproximada de uma função positiva e não nula $g(x)$, quando é possível aplicar o método, basta reescrevê-la como

$$\int_a^b g(x)dx = \int_a^b \exp(\ln(g(x)))dx$$

e aplicar o procedimento acima, fazendo $f(x) = \ln(g(x))$.

Por exemplo, o comportamento assintótico da função gama é:

$$\Gamma(n+1) = \int_0^\infty x^n \exp(-x)dx = \int_a^b \exp(n \ln(x) - x)dx$$

$$\text{Então } f(x) = n \ln(x) - x \implies f'(x) = \frac{n}{x} - 1 = 0 \implies x_0 = n \implies f''(x_0) = -\frac{1}{n}$$

Substituindo esses valores na integral I , vem:

$$\Gamma(n+1) \simeq n^n \ln(n) \sqrt{2\pi n} \exp(-n)$$

Que é a fórmula de Stirling, extremamente útil para $n \gg 1$.

Método do ponto de sela para calcular integrais é de grande utilidade em Física Teórica e em particular na Física Estatística. Aplica-se a integrais do tipo:

$$g(x) = \int_C \exp(x \cdot f(z))dz$$

onde C é um contorno aberto com a propriedade de que $\text{Re}(f(z)) \rightarrow -\infty$ em ambas as suas extremidades. A partir de agora, o número complexo $f(z)$ será decomposto em sua parte real e imaginária:

$$f(z) = f_R(z) + i f_I(z)$$

Consideremos valores positivos e grandes de x . Como

$$e^{xf(z)} = e^{xf_R(z)} e^{ixf_I(z)}$$

e $|e^{ixf_I(z)}| = 1$, o módulo do integrando na equação é dado por $e^{xf_R(z)}$. Esta função, para um dado x , varia de um valor máximo, atingido quando $f_R(z)$ é máximo, até zero, pelo menos

nos extremos. Para $x > 0$ e muito grande, temos um “pico” muito elevado, de onde o valor da integral cai rapidamente para o “vale” (região de baixos valores). Além disso, podemos utilizar a possibilidade de deformar o contorno, para fazer com que ele fique “a maior parte do tempo” nos vales, subindo ao pico pelo caminho mais íngreme. Desta maneira, apenas uma pequena parte do contorno contribuirá efetivamente para a integral. O método do ponto sela é isto: achar o contorno mais íngreme, passando pelo pico. Note que são os valores muito grandes de x que acentuam essas propriedades extremas. Logo, o método se presta para calcular valores assintóticos. A determinação do caminho mais íngreme passando pelo pico pode ser feita assim: considere as curvas de nível de $f_R(z)$, ou seja, as curvas ao longo das quais $f_R(z)$ é constante. O que procura-se são as curvas que cortem essas curvas de nível ortogonalmente: são estas as que “sobem mais rapidamente”. Ora, essas curvas são, como se sabe da teoria de funções analíticas de uma variável complexa, as curvas ao longo das quais $f_I(z)$ é constante. Logo, tem-se que achar a curva dessa família que passa pelo “pico”. No “pico” (que é o ponto sela) temos $\frac{d}{dz}f_R(z) = 0$. Vimos agora que, pelo caminho escolhido, $f_I(z)$ é constante, e, portanto, $\frac{d}{dz}f_I(z) = 0$. Logo, o ponto de sela satisfaz a equação complexa

$$\frac{df(z)}{dz} = 0$$

Seja z_0 o ponto em que essa equação é satisfeita (pode haver vários). Expandindo a função em torno desse ponto, tem-se

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \frac{f''(z_0)}{2!}(z - z_0)^2$$

mais termos de ordem superior. A derivada primeira é nula, por definição de ponto sela. Logo, temos, para a parte real do integrando,

$$e^{xf(z)} = e^{xf(z_0)} e^{\frac{f''(z_0)}{2}(z - z_0)^2}$$

com $\left(\frac{d^2f}{dz^2}\right)_{z_0} < 0$, ao longo do contorno, por ser um máximo de $f_R(z)$. Logo,

$$\int_C \exp(xf(z))dz = \exp(xf(z_0)) \int_C \exp\left(-|f''(z_0)|\frac{(z - z_0)^2}{2}\right) dz$$

que, em geral, por ser a integral de uma gaussiana, pode ser calculada facilmente.