

Universidade Estadual de Ponta Grossa
Setor de Ciências Exatas e Naturais
Programa de Pós-Graduação em Ciências
Área de Concentração: Física

Rafael Ribaski Borges

**Teoria de Valores Extremos Aplicada a Redes
Complexas**

Ponta Grossa
2013

Rafael Ribaski Borges

Teoria de Valores Extremos Aplicada a Redes Complexas

Dissertação submetida ao programa de Pós-Graduação em Ciências - Área de concentração: Física - da Universidade Estadual de Ponta Grossa como requisito parcial para obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Sandro Ely de Souza Pinto

Ponta Grossa
2013

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

B732 Borges, Rafael Ribaski
 Teoria de valores extremos aplicada a
 redes complexas/ Rafael Ribaski Borges.
 Ponta Grossa, 2013.
 112f.

 Dissertação (Mestrado em Ciências -
 Área de Concentração: Física),
 Universidade Estadual de Ponta Grossa.
 Orientador: Prof. Dr. Sandro Ely de
 Souza Pinto.

 1.Eventos extremos. 2.Caminhada
 aleatória. 3.Redes complexas. I.Pinto,
 Sandro Ely de Souza. II. Universidade
 Estadual de Ponta Grossa. Mestrado em
 Ciências. III. T.

CDD: 519.2

TERMO DE APROVAÇÃO

RAFAEL RIBASKI BORGES

“Teoria de Valores Extremos Aplicada a Redes Complexas”

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre 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. Dr. Ricardo Luiz Viana
Departamento de Física, UTFPR/PR



Prof. Dr. Pedro Rodrigues Júnior
Departamento de Física, UEPG/PR

Ponta Grossa, 05 de março de 2013.

Agradecimentos

- À Deus, por estar sempre ao meu lado .
- À minha mãe Olga, pelo amor que recebo;
- Ao meu amor, Jaqueline, pelo carinho, atenção, incentivo e auxílio na realização deste trabalho;
- Ao prof. Dr. Sandro Ely de Souza Pinto, primeiramente pela oportunidade, e por sua orientação na realização deste trabalho;
- Ao prof. Dr. Luiz Antônio Bastos Bernardes, pelo incentivo e indicação;
- À todos do grupo DNLSC, por seu apoio e conhecimento compartilhado;
- Aos colegas doutorandos Rodrigo F. Pereira, Maurício A. Ribeiro e Romeu Miquéias Szmoski pela disposição para auxílio na área computacioanal.
- À CAPES, pelo auxílio financeiro;
- A todos que de alguma forma contribuíram para a realização deste trabalho.

Resumo

A teoria de valores extremos é um ramo da estatística e probabilidade. Ela trata das distribuições assintóticas de valores extremos (máximos ou mínimos) de séries temporais. Os eventos que assumem valores afastados da média são classificados como eventos extremos. Alguns exemplos são desastres naturais, tais como enchentes, terremotos ou um evento que cause um forte impacto na sociedade. Considerando o cenário de redes complexas, alguns exemplos de eventos extremos são congestionamentos em redes de rodovias, quedas de energia em redes de transmissão e servidores de internet congestionados. Assim, a compreensão dos mecanismos que regem tais eventos é de grande interesse, pois com a previsão de ocorrências destes pode-se minimizar seus efeitos ou até mesmo evitá-los. Com isso, os objetivos deste trabalho foram: 1) descrever o comportamento assintótico das excedências de um valor limite especificado por meio da distribuição de valores extremos generalizada; 2) estender o estudo para a probabilidade de eventos extremos em redes complexas com topologia aleatória, mundo pequeno e escala livre. Este trabalho foi realizado por meio de simulações de caminhada aleatória padrão e por menores caminhos. Os resultados obtidos mostram que para os nós, também denominados vértices ou sítios, com menor conectividade (menor grau) nas redes analisadas, a distribuição dos excessos não é do tipo exponencial. Isto implica que esta distribuição é limitada superiormente. Os resultados para os nós com maior grau foram semelhantes, porém, somente para a rede de escala livre este comportamento não ocorre. Isto se deve ao fato de que o número de excedências observadas neste caso são menores do que nos demais. Foi verificado analiticamente e numericamente por meio de simulações de caminhada aleatória padrão, que a probabilidade de evento extremo é maior e que o tempo médio entre eles é menor para os nós com grau menor, quando comparados com nós com grau maior. O espectro de autovalores da matriz adjacência da rede, a qual descreve as ligações entre os nós, fornece condições para uma boa concordância entre os resultados analíticos e das simulações. Para simulações de caminhada aleatória por menores caminhos verificou-se que os nós com menores centralidades de intermediação são mais propensos a ter eventos extremos.

Palavras-chave: Eventos extremos, Caminhada Aleatória, Redes complexas.

Abstract

The extreme value theory is a branch of statistics and probability. It deals with the asymptotic distributions of extreme values (maximum or minimum) temporal series. The events which takes the average values removed are classified as extreme events. Examples include natural disasters such as floods, earthquakes or an event that causes a strong impact on society. Considering the scenario of complex networks, some examples of extreme events are congestion in networks of roads, power outages in power transmission networks and web servers congested. Thus, understanding the mechanisms that occur in such events is of great interest, because the prediction of these occurrences can minimize its effects, or even avoid them. Thus, the objectives of this study were: 1) to describe the asymptotic behavior of exceedances of a threshold specified by the generalized extreme value distribution, 2) extend the study to the probability of extreme events in complex networks with random topology, small world and scale free. This work was carried out by simulations of random walk pattern and shorter paths. The results shows that for the nodes, also called vertices or sites with low connectivity (lesser degree) in the networks analyzed, the distribution of excesses is not of exponential type. This implies that this distribution is bounded above. The results for the nodes with higher degree were similar, but only for the scale-free network this behavior does not occur. This is due to the fact that the number of exceedances observed in this case is significantly smaller than the other. It was checked analytically and numerically simulated by random walk pattern, the probability of extreme event is larger and the average time between them is smaller for nodes with lower degree when compared with nodes with higher degree. The spectrum of eigenvalues of the adjacency matrix of the network, which describes the links between nodes, provides conditions for a good agreement between the analytical results and the simulations. For simulations of random walk for shorter paths it was found that nodes with lower betweenness centralities are more likely to have extreme events.

Key-words: Extreme Events, Random Walk, Complex Networks.

Lista de Figuras

2.1	Série temporal dos máximos anuais da quantidade de precipitação em Fort Collins, Colorado, EUA (1900 a 1999).	21
2.2	Densidade de Gumbel (linha contínua) e duas densidades de Fréchet com $\xi = 0,28$ e $\xi = 0,56$	25
2.3	Densidade de Gumbel (linha contínua) e duas densidades de Weibull com $\xi = -0,28$ e $\xi = -0,56$	25
2.4	Ilustração da determinação das excedências do valor limite q por meio de caminhada aleatória em uma rede complexa.	30
3.1	(a) Ilustração da cidade de Königsberg com as sete pontes. (b) Situação simplificada pelo grafo, onde as arestas são as pontes e os nós são as porções de terra.	33
3.2	Grafo não direcionado com $N = 5$ nós e $E = 6$ arestas.	34
3.3	Grafo direcionado com $N = 5$ nós e $E = 4$ arestas.	35
3.4	(a) Hipergrafo com 5 nós. (b) Respectivo grafo bipartido.	36
3.5	Um caminho de comprimento três em um grafo.	39
3.6	A centralidade de intermediação do vértice A é maior do que a do vértice B, apesar de os dois possuírem o mesmo grau	42
3.7	Distribuição de grau de uma rede aleatória	47

3.8	Rede aleatória no modelo de ER com $p = 0,2$ e $N = 30$ nós.	47
3.9	Procedimento de religação de arestas em uma rede regular	49
3.10	Rede mundo pequeno no modelo de Watts e Strogatz com $p = 0,05$ e $N = 30$ nós.	50
3.11	Rede de Escala Livre no modelo de Barabási e Albert com $N = 30$ nós. . .	51
3.12	Distribuição de grau de redes de escala livre no modelo de Barabási e Albert com $N = 300000$ nós. O coeficiente angular da reta tracejada é $-2,9$	54
3.13	Processo para construção de uma rede hierárquica.	56
3.14	Distribuição de grau de uma rede hierárquica.	56
3.15	Rede determinística.	57
3.16	Densidade espectral para um grafo aleatório com $p = 0,05$. Na figura inserida: A densidade espectral decai exponencialmente, onde F é a função distribuição espectral cumulativa.	59
3.17	Densidade espectral para rede com escala livre no modelo de Barabási e Albert com $m = m_0 = 5$. Na figura inserida: A densidade espectral possui um decaimento do tipo lei de potência.	60
4.1	Caminhada aleatória em um grafo.	62
4.2	Diagrama de fases entre a fase congestionada e a de fluxo livre.	66
5.1	Teste de χ^2 para verificação da aderência de dados amostrais a uma dada distribuição. A área hachurada corresponde à região de não aceitação do teste.	75
5.2	Rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós, $E = 255$ arestas e probabilidade de ligação $p = 0,07$	76

5.3	Rede de escala livre no modelo de Barabási e Albert com $N = 100$ nós e $m = 2$	77
5.4	Rede mundo pequeno no modelo de Watts e Strogatz com $N = 100$ nós, $E = 199$ arestas e probabilidade de religação $p = 0,05$	78
5.5	Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede aleatória no modelo de Erdős e Renyi. Em (a) $N = 100$ nós, $E = 255$ arestas e $W = 510$ caminhantes,(b) $N = 500$ nós, $E = 4978$ arestas e $W = 9956$ caminhantes, (c) $N = 1000$ nós, $E = 6149$ arestas e $W = 12298$ caminhantes.	80
5.6	Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Watts e Strogatz. Em (a) $N = 100$ nós, $E = 199$ arestas e $W = 398$ caminhantes, (b) $N = 500$ nós, $E = 997$ arestas e $W = 1994$ caminhantes, (c) $N = 1000$ nós, $E = 2998$ arestas e $W = 5996$ caminhantes.	82
5.7	Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Newman e Watts. Em (a) $N = 100$ nós, $E = 302$ arestas e $W = 604$ caminhantes, (b) $N = 500$ nós, $E = 1513$ arestas e $W = 3026$ caminhantes, (c) $N = 1000$ nós, $E = 4502$ arestas e $W = 9004$ caminhantes.	83
5.8	Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede hierárquica no modelo de Ravász e Barabási. Em (a) $N = 125$ nós, $E = 410$ arestas e $W = 820$ caminhantes, (b) $N = 625$ nós, $E = 2450$ arestas e $W = 4900$ caminhantes.	84
5.9	Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede de escala livre no modelo de Barabási e Albert com $m = 2$. Em (a) $N = 100$ nós, $E = 189$ arestas e $W = 378$ caminhantes, (b) $N = 500$ nós, $E = 969$ arestas e $W = 1938$ caminhantes, (c) $N = 1000$ nós, $E = 1979$ arestas e $W = 3958$ caminhantes.	85

5.10	Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede de escala livre no modelo de Barabási e Albert com $m = 1$. Em (a) $N = 100$ nós, $E = 99$ arestas e $W = 198$ caminhantes, (b) $N = 500$ nós, $E = 499$ arestas e $W = 998$ caminhantes, (c) $N = 1000$ nós, $E = 999$ arestas e $W = 1998$ caminhantes.	87
5.11	Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós.	88
5.12	Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede mundo pequeno no modelo de Watts e Strogatz com $N = 100$ nós.	88
5.13	Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede mundo pequeno no modelo de Newman e Watts com $N = 100$ nós.	89
5.14	Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede de escala livre não bipartida no modelo de Barabási e Albert com $N = 100$ nós.	89
5.15	Probabilidade de EE em função do grau K com $\Delta = 0$ para algumas redes complexas, $W = 2E$, $W = 4E$ e $W = 6E$. (a) Rede hierárquica no modelo de Ravász e Barabási com $N = 125$ nós.(b) Rede de escala livre não bipartida no modelo de Barabási e Albert com $N = 100$ nós.	90
5.16	Probabilidade de EE em função do grau K com $\Delta = 0$ para algumas redes complexas com $N = 100$ nós, $W = 2E$, $W = 4E$ e $W = 6E$. (a) Rede aleatória no modelo de Erdős e Renyi, (b) Rede mundo pequeno no modelo de Newman e Watts, (c) Rede mundo pequeno no modelo de Watts e Strogatz.	91
5.17	Comparação entre o número de caminhantes para o tráfego médio e o limite q de EE para uma rede com $N = 100$ nós.(a)Nós com grau máximo e (b) Nós com grau mínimo.	93

5.18	Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede aleatória no modelo de Erdős e Renyi. Em (a) com $N = 100$ nós e $W = 255$ caminhantes.(b) $N = 500$ nós e $W = 9956$ caminhantes. (c) $N = 1000$ nós e $W = 12298$ caminhantes.	95
5.19	Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Watts e Strogatz. Em (a) com $N = 100$ nós e $W = 398$ caminhantes.(b) com $N = 500$ nós e $W = 1994$ caminhantes. (c) com $N = 1000$ nós e $W = 5996$ caminhantes.	96
5.20	Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede hierárquica no modelo de Ravász e Barabási. (a) $N = 125$ nós e $W = 820$ caminhantes (b) $N = 625$ nós e $W = 4900$ caminhantes.	97
5.21	Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Newman e Watts. (a) $N = 100$ nós e $W = 604$ caminhantes (b) $N = 500$ nós e $W = 3026$ caminhantes (c) $N = 1000$ nós e $W = 9004$ caminhantes	98
5.22	Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para Rede de escala livre no modelo de Barabási e Albert não bipartida. (a) $N = 100$ nós e $W = 378$ caminhantes (b) $N = 500$ nós e $W = 1938$ caminhantes (c), $N = 1000$ nós e $W = 3958$ caminhantes.	99
5.23	Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para Rede de escala livre no modelo de Barabási e Albert bipartida. (a) $N = 100$ nós e $W = 198$ caminhantes (b) $N = 500$ nós e $W = 998$ caminhantes (c), $N = 1000$ nós e $W = 1998$ caminhantes.	100
5.24	Probabilidade de EE em função da centralidade de intermediação b com $\Delta = 0$ rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós e $W = 9754$ caminhantes.	104

5.25	Probabilidade de EE em função da centralidade de intermediação b com $\Delta = 0$ para rede mundo pequeno no modelo de Newman e Watts com $N = 100$ nós e $W = 9549$ caminhantes.	104
5.26	Probabilidade de EE em função do grau K dos nós com $\Delta = 0$ e com $N = 100$. (a) Rede aleatória no modelo de Erdős e Renyi com $W = 9754$ caminhantes.(b) Rede mundo pequeno no modelo de Newman e Watts com $W = 9549$ caminhantes.	105

Lista de Tabelas

5.1	Teste de hipóteses para distribuição exponencial com $\xi = 0$. Simulação com rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós, $E = 255$ arestas e probabilidade de ligação $p = 0,07$	76
5.2	Teste de hipóteses para distribuição exponencial com $\xi = 0$. Simulação com rede de escala livre no modelo de Barabási e Albert com $N = 100$ nós e $m = 2$	77
5.3	Teste de hipóteses para distribuição exponencial com $\xi = 0$. Simulação com rede mundo pequeno no modelo de Watts e Strogatz com $N = 100$ nós, $E = 199$ arestas e probabilidade de religação $p = 0,05$	78

Sumário

Lista de Figuras	4
Lista de Tabelas	10
Sumário	11
1 Introdução	14
2 Eventos Extremos	19
2.1 Eventos Extremos	19
2.1.1 Teorema dos Tipos Extremos	22
2.1.1.1 Distribuição de Valores Extremos Generalizada	23
2.1.1.2 Modelo Assintótico para Valores Mínimos	26
2.1.2 Modelos para Excedências de um Valor Limite	27
2.1.2.1 Modelo de Caracterização Assintótica	27
3 Redes Complexas	31
3.1 Redes Complexas do Mundo Real	31
3.2 Ferramentas Matemáticas de Redes	32
3.2.1 Redes e suas Representações	32

3.2.2	Matriz Adjacência	33
3.2.3	Redes Direcionadas	34
3.2.4	Hipergrafos e Redes Bipartidas	35
3.2.5	Grau	37
3.2.6	Caminhos	39
3.3	Medidas de Centralidade	40
3.3.1	Centralidade de Grau	40
3.3.2	Centralidade por Autovetor	40
3.3.3	Centralidade de Intermediação	42
3.3.4	Centralidade de Proximidade	43
3.4	Modelos para Redes Complexas	44
3.4.1	Redes Aleatórias: Modelo de Erdős e Renyi	44
3.4.2	Redes <i>Mundo Pequeno</i>	49
3.4.2.1	Modelo de Watts e Strogatz (WS)	49
3.4.3	Redes de Escala Livre	50
3.4.3.1	Modelo de Barabási e Albert	50
3.4.3.2	Redes Hierárquicas	55
3.4.4	Espectro de Autovalores de Redes Complexas	58
3.4.4.1	Matriz Adjacência e o Espectro de um Grafo	58
3.4.4.2	Espectro de Redes Complexas: Lei do Semicírculo	58
4	Caminhada Aleatória em Redes Complexas	61
4.1	Caminhada Aleatória em Grafos	61

4.2	Caminhada Aleatória em Redes Complexas	63
4.3	Estimativa de Probabilidades de EE em Redes Complexas por Meio de Caminhada Aleatória	66
5	Resultados e Discussões	73
5.1	Caminhada Aleatória Padrão	73
5.1.1	Distribuição de Excedências	73
5.1.2	Probabilidade de Evento Extremo	79
5.1.3	Tempos Médios entre Eventos Extremos	94
5.2	Caminhada Aleatória por Menores Caminhos	101
6	Conclusões	106
7	Sugestões de trabalhos futuros	108
	Referências Bibliográficas	109

Capítulo 1

Introdução

Eventos extremos (EE) são tipicamente associados a desastres naturais como terremotos e furacões. A ferramenta matemática para modelar este tipo de evento é a teoria de valores extremos. Este ramo da estatística e probabilidade teve seu início na década de 1920 com trabalhos de Fischer, Fréchet, entre outros, e tem como objetivo principal obter uma caracterização assintótica dos valores máximos ou mínimos de séries temporais (1). Todo evento que toma valores fora do comum ou excede um valor limite, pode ser classificado como EE (2).

Dentre as várias aplicações estão a análise de frequências de ocorrência de enchentes em hidrologia (3); na engenharia, o tempo de vida útil de materiais e, na medicina, pode-se citar a epilepsia, um tipo de EE que ocorre no cérebro humano (4). Tratando-se especificamente de aplicações em hidrologia, são estimados níveis de água em enchentes que podem ser excedidos em T anos (1). Consequências catastróficas podem ocorrer se estes níveis são excedidos. Como exemplo, os níveis estimados para $T = 100$ anos foram excedidos na enchente que ocorreu em 1993, o que causou uma grande devastação nos estados do meio-oeste americano (5). Também, uma enchente que ocorreu na Holanda em 1953, desencadeou um grande investimento na construção de diques. Existem leis nos países baixos (Holanda e Bélgica), que exigem que os diques devam ser construídos para suportar inundações com níveis estimados em $T = 10^4$ anos (5).

Em redes complexas, alguns exemplos de EE são engarrafamentos em redes de rodovias, enchentes em redes de rios, servidores de internet congestionados e quedas de energia em redes de transmissão de energia elétrica (6). Assim, o estudo dos EE é de fundamental importância, devido a estar relacionado a perdas de vidas e/ ou materiais.

Na matemática, uma rede complexa pode ser descrita por um grafo G . Um grafo é um conjunto $G = \{V, E\}$, onde V é o conjunto de vértices ou nós e E é o conjunto das arestas ou ligações que unem pares de vértices (7). Existem vários exemplos de redes complexas no mundo real em diversas áreas, como por exemplo: Redes de cadeia alimentar na biologia, redes de internet e redes sociais. Com o objetivo de descrever de forma satisfatória essas redes reais, foram propostos alguns modelos (7). O modelo de rede aleatória, criado na década de 1950 pelos matemáticos Húngaros Paul Erdos e Alfred Renyi, permaneceu até a poucos anos atrás, como aquele que mais se aproximava de algumas propriedades dessas redes reais.

Porém, no início da década de 1990, com o aumento da capacidade dos computadores e ao acesso à banco de dados de diversas áreas, tornou-se possível analisar o comportamento de algumas propriedades de redes reais para maiores tamanhos de rede. Verificou-se que este modelo não conseguia descrever algumas propriedades, como por exemplo, a propriedade de mundo pequeno e a distribuição de grau. A primeira é caracterizada por apresentar um distância média relativamente curta entre quaisquer dois nós na rede. A segunda refere-se que a maioria das redes reais apresentam distribuições de grau do tipo lei de potência, enquanto que o que foi previsto pelo modelo de Erdos e Renyi era do tipo de Poisson (7). Com isso, foram propostos modelos que descrevem estas duas propriedades.

Em 1998, Watts e Strogatz propuseram um modelo que melhor descreve a propriedade de mundo pequeno (8). Este é intermediário entre um modelo de rede regular e uma aleatória, dependendo da probabilidade p religação entre nós. Este modelo também descreve a transitividade, a qual é a tendência de formação de grupos, e tem uma melhor aproximação com os valores empíricos observados. Em 1999, Barabási e Albert (9) propuseram um modelo que consegue reproduzir uma distribuição de grau lei de potência,

assim como é observado em redes reais.

Os trabalhos que relacionam EE e redes complexas de uma forma ou de outra, utilizam como modelo de transporte de informação a caminhada aleatória (6), (10). Uma caminhada aleatória em um grafo G , com início a partir de um vértice qualquer, é uma sequência de escolhas aleatórias de vértices vizinhos à posição anterior. O problema da caminhada aleatória em grafos foi primeiramente estudado por Pólya (1919). Pólya percebeu a analogia existente entre a caminhada aleatória unidimensional e o jogo de cara ou coroa. Em duas dimensões, com um bêbado vagando pelas ruas de uma cidade e, em três dimensões, com o processo de difusão de uma gás na rede cristalina de um sólido (11). A sua teoria clássica tem o enfoque em grafos regulares infinitos. Os objetivos eram responder à questões tais como: 'Se um caminhante aleatório está em um dado nó em um determinado instante de tempo, ele retornará a este mesmo nó posteriormente? A teoria recente está relacionada com grafos mais gerais, porém com uma quantidade de nós finita. Nela, pode-se obter a probabilidade de um caminhante chegar a um determinado nó, seja qual for o nó de partida. Esta probabilidade é conhecida como estacionária, e está relacionada com o espectro de autovalores da matriz adjacência desta rede, a qual descreve as ligações entre os nós.

Uma classe de EE que foi estudada por meio de caminhada aleatória em redes complexas é o congestionamento (10), (12). Nos trabalhos encontrados na literatura, existe uma limitação na capacidade de armazenagem de caminhantes em um dado nó. Também, é adotada uma taxa de criação de caminhantes, ou pacotes de informação, e uma probabilidade de um determinado nó recusar caminhantes, o que ocorre para o caso da capacidade máxima ter sido atingida. Nesta situação, diz-se que este nó está congestionado. Como resultado são obtidos diagramas de fase que, dependendo dos valores de parâmetros, podem apresentar fases com congestionamento, de fluxo livre ou mesmo de coexistência entre as duas.

Outra abordagem para o estudo de EE em redes complexas também por meio de modelos de caminhada aleatória, padrão ou por menores caminhos, é considerar como

EE toda flutuação grande do tráfego normal de caminhantes nos nós de uma rede (6). Neste caso, não é imposta uma capacidade de armazenagem para os nós. Os EE são caracterizados como aqueles eventos em que um valor limite alto foi ultrapassado. No trabalho de Kishore et. al (2011) (6), o limite adotado é proporcional ao tráfego neste nó, isto é, toda vez que o número de caminhantes em um nó for superior a esse limite, este evento é classificado como extremo. O resultado principal destes pesquisadores foi que os nós com grau menor na rede estão mais propensos a ter EE quando comparados com os nós mais conectados, apesar de que nesses, o tráfego médio de caminhantes é superior. Também, outra característica observada foi que os tempos médios entre EE são menores para os nós com grau menor. Estes resultados permanecem válidos, considerando uma caminhada aleatória por menores caminhos na rede, onde os nós com menor centralidade de intermediação são os mais propensos a ter EE. Estes resultados referem-se a uma rede de escala livre.

O objetivo do presente trabalho é analisar o comportamento assintótico dos valores máximos, mais especificamente da distribuição de excedências, por meio da teoria clássica de valores extremos. Também, reproduzir os resultados encontrados por Kishore et. al (2011) (6) para EE em redes de escala livre, e estender e discutir para outras redes com topologia aleatória, mundo pequeno e hierárquica, bem como para os tempos médios entre eles, considerando a caminhada aleatória padrão e por menores caminhos.

Esta dissertação está estruturada da seguinte maneira: No capítulo 2 é apresentada uma breve descrição dos conceitos básicos da teoria clássica de valores extremos: A distribuição de valores extremos generalizada (VEG) e a distribuição de Pareto generalizada. A VEG contempla para alguns valores de parâmetros as distribuições de Gumbel, Fréchet e Weibull que descrevem o comportamento assintótico dos valores máximos e mínimos. A distribuição de Pareto generalizada descreve este comportamento das excedências de um limite alto.

No capítulo 3 são abordadas algumas ferramentas matemáticas e tipos de redes complexas, juntamente com suas propriedades principais, tais como a distribuição de grau, o

coeficiente de aglomeração e o espectro de autovalores.

No capítulo 4 são apresentados os conceitos básicos sobre caminhada aleatória clássica em grafos e em redes complexas, e a aplicação na determinação das probabilidades de EE. É demonstrada a relação existente entre a caminhada aleatória e o espectro de autovalores do grafo. Nos capítulos de revisão bibliográfica (1 ao 4) foram inseridas algumas figuras elaboradas pelo autor, com o objetivo de tornar mais didática a apresentação dos conceitos.

No capítulo 5 são apresentados os resultados obtidos e as discussões. Nos capítulos 6 e 7 apresentamos as conclusões e sugestões de trabalhos futuros, respectivamente.

Capítulo 2

Eventos Extremos

2.1 Eventos Extremos

Eventos extremos (EE) ocorrem em ambientes naturais, tecnológicos e sociais. Podem ser de origem natural ou causados pela ação do homem, ou mesmo por acaso. Estão frequentemente associados com perdas de vida ou financeiras (4). Os EE são caracterizados por valores extremos, cuja probabilidade de ocorrência é pequena (1). Na linguagem cotidiana, os EE são associados com características como raridade, extremismo e frequentemente com consequências catastróficas. Estes tipos de fenômenos são estudados na matemática, por meio da teoria clássica de valores extremos, que descreve o comportamento assintótico de variáveis aleatórias adequadamente normalizadas (2).

Os EE foram estudados ao longo da trajetória da história humana: Os egípcios, cerca de 5000 anos atrás, iniciaram uma cronologia dos níveis máximo e mínimo do rio Nilo. Havia fome em épocas de seca e desastres devido a grandes inundações. Um fluxo moderado no delta do Nilo era benéfico para a agricultura e a sociedade em geral; Outro exemplo são os terremotos. O terremoto que ocorreu em Lisboa em 1755 deu início às pesquisas nesta área na Europa. Na atualidade, EE atraem a atenção do público em geral e de cientistas devido ao receio de que a magnitude e a frequência de EE sejam aumentadas devido a ação humana (4). A teoria de valores extremos teve seu início na década de 20,

com trabalhos de R. Von Mises (1923), E.L Dodd (1923), L.H.C. Tippett (1925), R.A. Fischer e M. Fréchet (1927), os quais obtiveram distribuições de valores máximos (1).

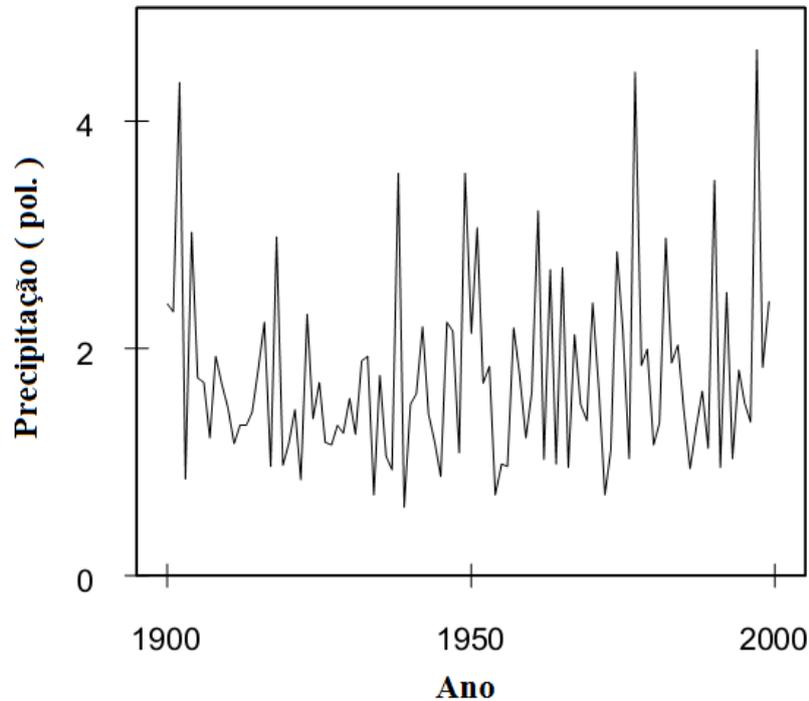
A teoria de valores extremos é um ramo da teoria de probabilidades que já está bem estabelecido, principalmente devido a trabalhos como o de E.J. Gumbel, que trata dos valores extremos (máximo e mínimo) de séries temporais finitas. Estas séries provem de variáveis aleatórias independentes e identicamente distribuídas (i.i.d), aplicadas principalmente nas áreas da climatologia e hidrologia (5).

A principal afirmação desta teoria é que se $(X_n)_{n \in \mathbb{N}}$ é uma sequência de variáveis aleatórias reais i.i.d. de máximos $M_n \equiv \max(X_1, X_2, \dots, X_n)$, adequadamente padronizada, possui uma distribuição limite quando $n \rightarrow \infty$. Esta distribuição pertence a um dos três tipos padrão (Fréchet, Weibull e Gumbel), este resultado conhecido como Teorema dos Tipos Extremos, foi enunciado pela primeira vez por Fisher e Tippet (1928) e depois estendido por Gnedenko (1943)(4). Nesta seção é descrita a base do modelo da Teoria de valores extremos que estuda o comportamento estatístico de:

$$M_n = \max\{X_1, X_2, \dots, X_n\}. \quad (2.1)$$

As variáveis aleatórias X_1, \dots, X_n são independentes e tem em comum entre si, uma função distribuição F. Na prática, X_n geralmente representam medidas em intervalos de tempo regulares, como por exemplo, dados anuais de medidas do valor diário de precipitação, como ilustrado na Fig. 2.1.

Figura 2.1: Série temporal dos máximos anuais da quantidade de precipitação em Fort Collins, Colorado, EUA (1900 a 1999).



Fonte: Adaptado de Reiss et. al (2007)(3).

Assim, M_n representa o valor máximo do processo ao longo do tempo depois de n observações (2). Nesta teoria, a distribuição dos valores de M_n pode ser obtida para todos os valores de n :

$$Pr\{M_n \leq z\} = Pr\{X_1 \leq z, X_2 \leq z, \dots, X_n \leq z\} = Pr\{X_1 \leq z\} \dots Pr\{X_n \leq z\} = \{F(z)\}^n. \quad (2.2)$$

No entanto, na prática isto não é muito útil, devido à função F ser desconhecida. Uma forma para contornar este problema é usar as técnicas padrão da estatística para estimar F a partir de dados coletados, porém, uma pequena discrepância na estimativa de F pode gerar uma diferença considerável em F^n . Outra alternativa é ainda considerar F como desconhecida e procurar aproximações para famílias de F^n , as quais podem ser estimadas com base somente nos dados extremos. Este procedimento é análogo à prática usual da obtenção da distribuição de médias de uma amostra por uma distribuição normal, usando

o teorema do limite central (1).

Na teoria de valores extremos a distribuição assintótica de M_n não degenerada deve pertencer a uma das três possíveis famílias, independentemente da função distribuição original. Além disso, não é necessário o conhecimento detalhado da natureza de F para se determinar a forma limite, isto é, o 'domínio de atração' a que ela pertence (13).

A distribuição de M_n pode ser degenerada quando $n \rightarrow \infty$, ou seja, os valores máximos podem se repetir para diferentes instantes de tempo. Para evitar esse problema realiza-se a padronização nesta variável, como segue:

$$M_n^* = \frac{M_n - b_n}{a_n}. \quad (2.3)$$

Escolhas adequadas de sequência de constantes $\{a_n\}$ e $\{b_n\}$ estabilizam a localização e a escala de M_n^* quando n aumenta, evitando os problemas que M_n apresenta. Assim observa-se a distribuição limite de M_n^* com $\{a_n\}$ e $\{b_n\}$ adequadas, em vez de M_n .

2.1.1 Teorema dos Tipos Extremos

As possíveis distribuições limite para M_n^* são dadas pelo teorema abaixo (2):

Teorema 1: Se existem sequências de constantes $a_n > 0$ e b_n , tal que $Pr \left\{ \frac{M_n - b_n}{a_n} \leq z \right\} \rightarrow G(z)$ quando $n \rightarrow \infty$, onde G é uma função distribuição não degenerada, então G pertence a uma das seguintes famílias:

$$(Tipo I) \quad G(z) = exp \left\{ -exp \left[- \left(\frac{z - b}{a} \right) \right] \right\}, \quad -\infty \leq z \leq \infty; \quad (2.4)$$

$$(Tipo II) \quad G(z) = \begin{cases} 0 & \text{se } z \leq b, \\ exp \left\{ - \left(\frac{z - b}{a} \right)^{-\alpha} \right\}, & \text{se } z > b. \end{cases} \quad (2.5)$$

$$(Tipo III) \quad G(z) = \begin{cases} \exp \left\{ - \left[- \left(\frac{z-b}{a} \right)^\alpha \right] \right\}, & \text{se } z < b. \\ 1, & \text{se } z \geq b; \end{cases} \quad (2.6)$$

para parâmetros $a > 0$, b e para os caso das famílias (II) e (III), $\alpha \geq 0$.

O Teorema 1 garante que a distribuição de $(M_n - b_n)/a_n$ converge para uma das três famílias rotuladas por (I), (II) e (III), que são denominadas distribuições de valores extremos. Estas são conhecidas como famílias de Gumbel, Fréchet e Weibull, respectivamente, onde cada uma possui uma posição e um parâmetro de escala b e a . As famílias de Fréchet e Weibull possuem um parâmetro de contorno α . Este teorema pode ser considerado como análogo ao teorema do limite central, pois quando M_n é normalizado com as sequências adequadas a_n e b_n , a sua distribuição limite converge para um dos três tipos padrão de distribuição de valores extremos (2).

2.1.1.1 Distribuição de Valores Extremos Generalizada

Os três tipos de limites no Teorema 1 possuem diferentes formas e comportamentos, correspondendo ao formato da cauda da função distribuição F (13). Em aplicações iniciais da teoria de valores extremos, é comum adotar uma das três famílias que mais se aproxime dos dados coletados, e então estimar os parâmetros da distribuição. A inferência para estes parâmetros presume que esta escolha é a correta. A melhor forma de análise é a reformulação do modelo no Teorema 1, combinando as três famílias em uma forma generalizada, da seguinte maneira (2):

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z-\mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (2.7)$$

definido no conjunto $\left\{ z : 1 + \xi \left(\frac{z-\mu}{\sigma} \right) > 0 \right\}$, onde os parâmetros satisfazem $-\infty \leq z \leq \infty$, $\sigma > 0$ e $-\infty \leq \xi \leq \infty$.

Esta é a família de distribuições de valores extremos generalizada (VEG) que possui três parâmetros: Um parâmetro de posição, μ ; um de formato, ξ e um de escala, σ . Os tipos II e III das classes de distribuição de valores extremos correspondem respectivamente aos casos $\xi > 0$ e $\xi < 0$, e o subconjunto de VEG, com $\xi = 0$, pode ser interpretado como o limite de $\xi \rightarrow 0$, que corresponde a função distribuição de Gumbel.

Esta unificação das três famílias originais de distribuição de valores extremos em uma única família simplifica a implementação estatística, por meio da inferência de ξ , e com os dados coletados pode ser determinado qual família é mais apropriada. Assim, pode-se reescrever o Teorema 1 da seguinte forma (2):

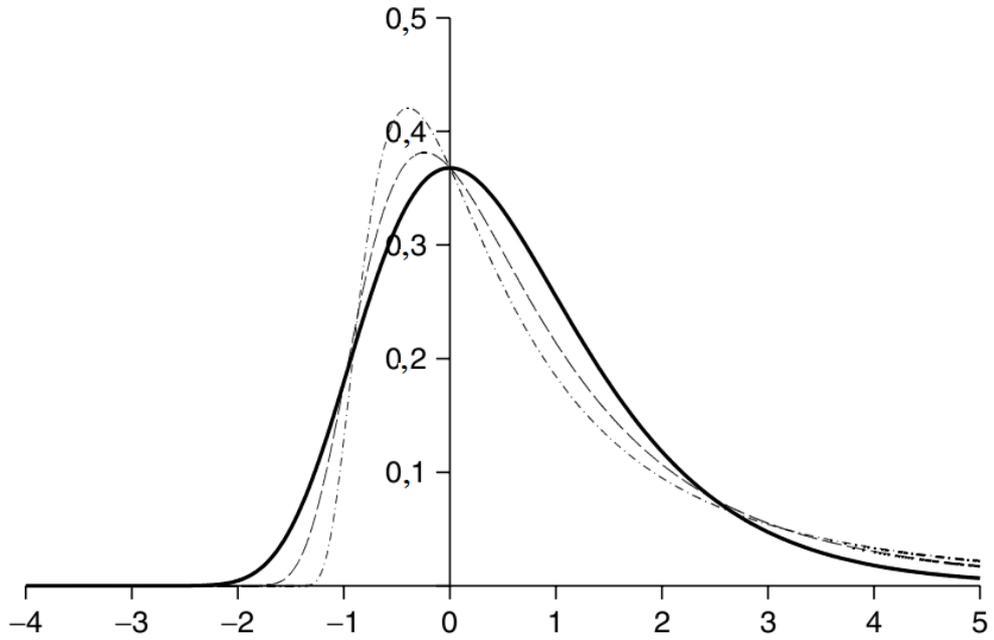
Teorema 2: Se existem sequências de constantes $a_n > 0$ e b_n tal que $Pr \left\{ \frac{M_n - b_n}{a_n} \leq z \right\} \rightarrow G(z)$ quando $n \rightarrow \infty$, onde G é uma função distribuição não degenerada, então G pertence a família VEG:

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (2.8)$$

definido em $\left\{ z : 1 + \xi \left(\frac{z - \mu}{\sigma} \right) > 0 \right\}$, onde os parâmetros satisfazem $-\infty \leq z \leq \infty$, $\sigma > 0$ e $-\infty \leq \xi \leq \infty$.

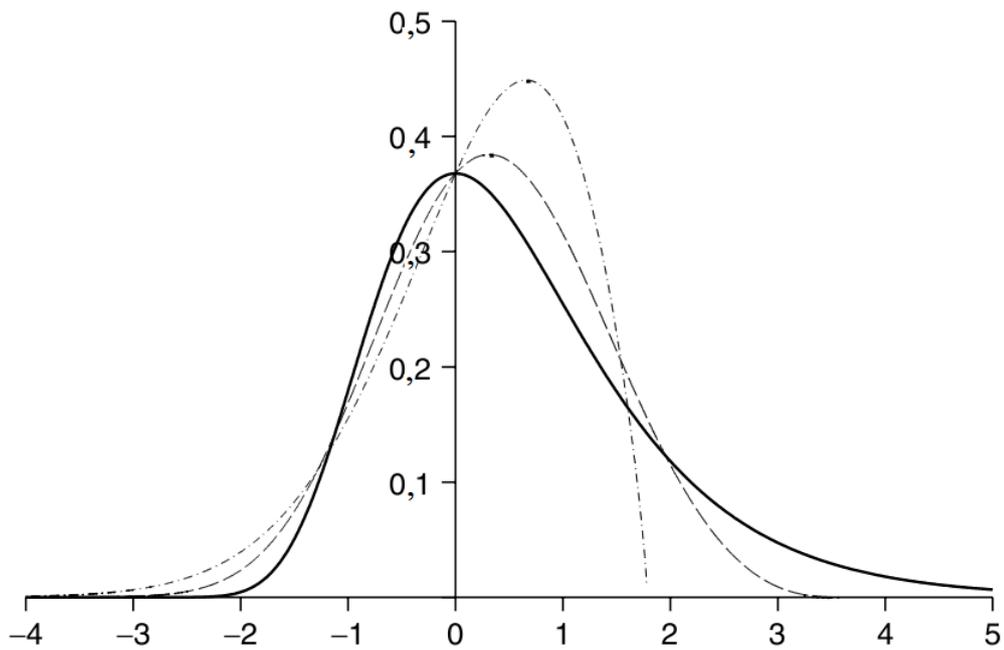
Uma forma para aproximar a distribuição VEG é a partir de uma série de observações independentes X_1, X_2, \dots , onde os dados são alocados em blocos de tamanho n suficientemente grande, gerando uma série máximos em cada bloco, $M_{n,1}, \dots, M_{n,m}$, em que a distribuição VEG pode ser ajustada. Uma ilustração das famílias VEG são apresentadas nas Figs. 2.2 e 2.3 (5).

Figura 2.2: Densidade de Gumbel (linha contínua) e duas densidades de Fréchet com $\xi = 0,28$ e $\xi = 0,56$.



Fonte: Adaptado de Beirlant et. al (2004) (5)

Figura 2.3: Densidade de Gumbel (linha contínua) e duas densidades de Weibull com $\xi = -0,28$ e $\xi = -0,56$.



Fonte: Adaptado de Beirlant et.al (2004)(5)

2.1.1.2 Modelo Assintótico para Valores Mínimos

Para algumas aplicações, existe a necessidade de modelos para descrever observações extremamente pequenas, ao invés de extremamente grandes. Como exemplo, para modelos de falhas em sistemas estes podem ser construídos tal que o tempo médio de vida útil do sistema seja igual ao tempo médio de um número n de componentes individuais. Assim, o tempo médio de vida dos sistemas como um todo é então $\widehat{M}_n = \min \{X_1, X_2, \dots, X_n\}$, onde X_i denota o tempo médio das componentes individuais. Assumindo que a variável aleatória X_i é independente e identicamente distribuída, os mesmos argumentos aplicados para M_n permanecem válidos para \widehat{M}_n , levando a uma distribuição limite para a distribuição de \widehat{M}_n adequadamente normalizada (2).

Pode-se também obter os mesmos resultados a partir dos correspondentes da distribuição de M_n . Seja $Y_i = -X_i$, para $i = 1, 2, \dots, n$, o sinal (-) indica que pequenos valores de X_i correspondem a valores grandes de Y_i . Portanto, se $\widehat{M}_n = \min \{X_1, X_2, \dots, X_n\}$ e $M_n = \max \{Y_1, Y_2, \dots, Y_n\}$, então $\widehat{M}_n = -M_n$, e desta forma, para um n suficientemente grande, tem-se:

$$Pr\{\widehat{M}_n \leq z\} = Pr\{-M_n \leq z\} = Pr\{M_n \geq -z\} = 1 - Pr\{M_n \leq -z\}, \quad (2.9)$$

$$\begin{aligned} Pr\{\widehat{M}_n \leq z\} &\approx 1 - \exp \left\{ - \left[1 + \xi \left(\frac{-z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} = \\ &= 1 - \exp \left\{ - \left[1 - \xi \left(\frac{z - \widehat{\mu}}{\sigma} \right) \right]^{-1/\xi} \right\}, \end{aligned} \quad (2.10)$$

definido em $\left\{ z : 1 - \xi \left(\frac{z - \widehat{\mu}}{\sigma} \right) > 0 \right\}$, onde $\widehat{\mu} = -\mu$. Esta é a distribuição de VEG para os valores mínimos. Isto é formalizado no seguinte teorema (2):

Teorema 3: Se existem sequências de constantes $a_n > 0$ e b_n tal que $Pr \left\{ \frac{\widehat{M}_n - b_n}{a_n} \leq z \right\} \rightarrow \widehat{G}(z)$ quando $n \rightarrow \infty$, onde \widehat{G} é uma função distribuição não degenerada, então \widehat{G} pertence

a família de distribuição VEG para valores mínimos:

$$\widehat{G}(z) = 1 - \exp \left\{ - \left[1 - \xi \left(\frac{z - \widehat{\mu}}{\sigma} \right) \right]^{-1/\xi} \right\}. \quad (2.11)$$

definido em $\left\{ z : 1 - \xi \left(\frac{z - \widehat{\mu}}{\sigma} \right) > 0 \right\}$, onde os parâmetros satisfazem $-\infty \leq \mu \leq \infty$, $\sigma > 0$ e $-\infty \leq \xi \leq \infty$.

2.1.2 Modelos para Excedências de um Valor Limite

Seja X_1, \dots, X_n uma sequência de variáveis aleatórias independentes e identicamente distribuídas que possuem uma função distribuição F . Pode-se considerar X_i como um evento extremo se este exceder um alto valor limite q . Denotando um termo arbitrário na sequência X_i por X , a descrição do comportamento estocástico de eventos extremos será dada pela probabilidade condicional (2):

$$Pr\{X > q + y | X > q\} = \frac{1 - F(q + y)}{1 - F(q)} \quad y > 0. \quad (2.12)$$

Se a distribuição F é conhecida, então a distribuição dos limites de excessos também será conhecida. Como em situações práticas não conhecemos F , devemos considerar aproximações para a distribuição dos máximos.

2.1.2.1 Modelo de Caracterização Assintótica

Teorema 4: Seja X_1, \dots, X_n uma sequência de variáveis aleatórias independentes que possui uma função distribuição F , e seja $M_n = \max\{X_1, \dots, X_n\}$. Denotando um termo arbitrário na sequência X_i por X e supondo que F satisfaz o Teorema 2, então para n suficientemente grande temos (2):

$$Pr\{M_n \leq z\} \approx G(z), \quad (2.13)$$

onde

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}. \quad (2.14)$$

para algum $\mu, \sigma > 0$ e ξ . Então, para q suficientemente grande a função distribuição de $(X - q)$, condicionada ao fato que $X > q$ é aproximadamente:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\hat{\sigma}} \right)^{-1/\xi}. \quad (2.15)$$

definida em $\left\{ y : y > 0 \text{ e } \left(1 + \frac{\xi y}{\hat{\sigma}} \right) > 0 \right\}$ onde $\hat{\sigma} = \sigma + \xi(q - \mu)$.

A família de distribuições definida na Eq. (2.15) é denominada Família de Pareto Generalizada. O parâmetro ξ é dominante no comportamento qualitativo da distribuição, pois se $\xi < 0$, a distribuição de excessos possui um limitante superior $q - \hat{\sigma}/\xi$; se $\xi \geq 0$ não possui um limitante superior e considerando o limite $\xi \rightarrow 0$, tem-se (2):

$$H(y) = 1 - \exp \left(-\frac{y}{\hat{\sigma}} \right). \quad (2.16)$$

correspondendo a uma distribuição exponencial com parâmetro $1/\hat{\sigma}$.

Prova: Seja X uma função distribuição de F . A partir da hipótese do Teorema 2, para um n suficientemente grande temos (2):

$$F^n(z) \approx \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (2.17)$$

para alguns parâmetros $\mu, \sigma > 0$ e ξ . Portanto:

$$n \log F(z) \approx - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi}, \quad (2.18)$$

Mas para z suficientemente grande, a expansão em série de Taylor fornece:

$$\log F(z) = -\{1 - F(z)\}, \quad (2.19)$$

Substituindo na equação anterior, e fazendo $z = q$ tem-se:

$$1 - \log F(q) \approx \frac{1}{n} \left[1 + \xi \left(\frac{q - \mu}{\sigma} \right) \right]^{-1/\xi}, \quad (2.20)$$

para q suficientemente grande. De forma análoga, para $y > 0$ tem-se:

$$1 - \log F(q + y) \approx \frac{1}{n} \left[1 + \xi \left(\frac{q + y - \mu}{\sigma} \right) \right]^{-1/\xi}, \quad (2.21)$$

Substituindo as equações (2.20) e (2.21) em (2.12) tem-se:

$$Pr \{X > q + y | X > q\} \approx \frac{\frac{1}{n} \left[1 + \xi \left(\frac{q + y - \mu}{\sigma} \right) \right]^{-1/\xi}}{\frac{1}{n} \left[1 + \xi \left(\frac{q - \mu}{\sigma} \right) \right]^{-1/\xi}} = \left[1 + \frac{\xi y}{\hat{\sigma}} \right]^{-1/\xi}. \quad (2.22)$$

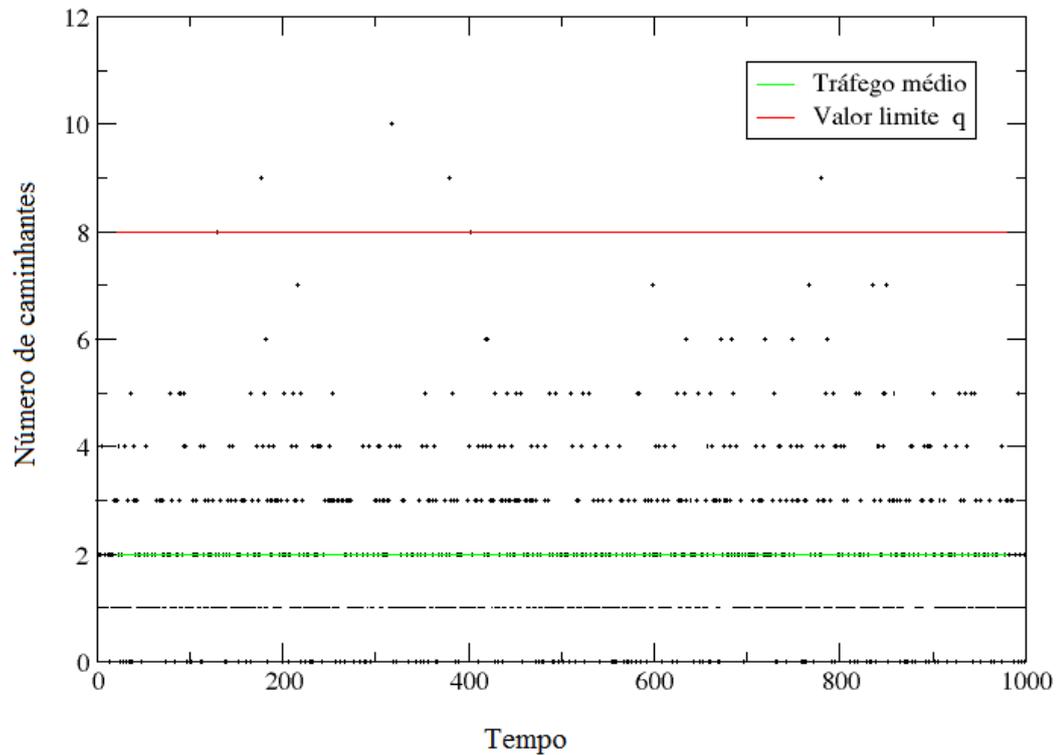
onde $\hat{\sigma} = \sigma + \xi(q - \mu)$.

O Teorema 4 indica uma forma de construir um modelo para valores extremos. Os dados brutos consistem em uma sequência de medidas independentes e identicamente distribuídas. Os EE são identificados definindo um alto valor limite q , em que as excedências são $\{x_i : x_i > q\}$. Indicando essas excedências por $x_{(1)}, \dots, x_{(k)}$ e definindo as alturas $y_j = x_j - q$, para $j = 1, \dots, k$, os valores de y_j podem ser considerados como variáveis aleatórias independentes cuja distribuição pode ser aproximada por um membro da família generalizada de Pareto. A inferência consiste em ajustar os valores das excedências observadas à essa família.

No capítulo 4, descrevemos o trabalho de Kishore et. al (6), que estudaram as probabilidades de EE em redes complexas, onde foi adotado o valor limite q como sendo

proporcional ao tráfego de caminhantes aleatórios nos vértices desta rede. Na Fig 2.4 é ilustrado por meio de uma simulação de caminhada aleatória em uma rede complexa, o processo de classificação de eventos como sendo extremos em um determinado vértice, conforme o trabalho de Kishore et. al (6). Para este vértice os valores determinados por meio da simulação para o tráfego médio e o valor limite q é igual a 2 e 8 caminhantes, respectivamente. Assim, se em um instante de tempo o número de caminhantes neste vértice excede o limite $q = 8$ caminhantes, o evento correspondente é classificado como extremo (conforme itens 4.3 e 5.1.1).

Figura 2.4: Ilustração da determinação das excedências do valor limite q por meio de caminhada aleatória em uma rede complexa.



Fonte: O autor

Capítulo 3

Redes Complexas

3.1 Redes Complexas do Mundo Real

Uma ampla variedade de sistemas com alta importância tecnológica e intelectual podem ser descritos por estruturas complexas com o formato de teia (7). Com o aumento da capacidade e popularidade dos computadores, principalmente durante a década de 90, tornou-se possível o acesso e a análise de várias bases de dados de diversas áreas. Isso permitiu a comparação de redes do mundo real com os modelos de rede já existentes, em particular com o modelo de Erdős e Renyi (ER) (14) .

Para tentar descrever de uma forma satisfatória uma rede de mundo real por meio de um modelo, várias quantidades e medidas foram propostas. Contudo, três conceitos ainda ocupam um lugar privilegiado na ideia de redes complexas: as propriedades de mundo pequeno, aglomeração e distribuição de grau (7). A propriedade de mundo pequeno descreve o fato de que a maioria das redes apresentam uma distância média pequena entre quaisquer dois nós na rede, sendo a distância entre eles definida como o número de arestas ao longo do menor caminho que os conectam. A propriedade de aglomeração, que é quantificada pelo coeficiente de aglomeração, é comum em redes sociais. Esta, é a tendência de formação de cliques, isto é, círculos de amizades e conhecidos, em que todos os membros se conhecem. A terceira propriedade, a distribuição de grau, descreve

o fato que os nós em uma rede, em geral não possuem o mesmo grau. O espalhamento do número de arestas por nó é caracterizado por uma função distribuição $P(K)$, que fornece a probabilidade de um nó selecionado aleatoriamente possuir exatamente K arestas.

Alguns exemplos de redes de mundo real são: redes de computadores onde as conexões são físicas, uma aresta existe entre dois nós (computadores) na rede se eles são fisicamente ligados por um cabo; redes tecnológicas que são as redes de transmissão de energia elétrica; redes de telefonia; redes de transportes; redes tecnológicas virtuais e WWW de páginas HTML, onde não existem conexões físicas; redes sociais que consistem em interações sociais entre indivíduos (amizades, relações de trabalho) e redes biológicas, como por exemplo, as que representam cadeias alimentares etc (15). Muitas destas redes apresentam distribuição de grau do tipo lei de potência (7).

3.2 Ferramentas Matemáticas de Redes

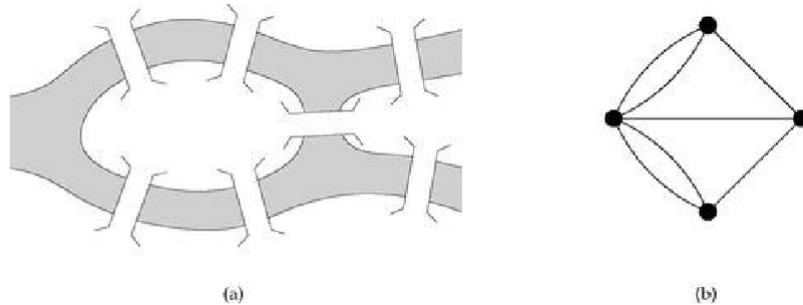
3.2.1 Redes e suas Representações

Uma rede, também denominada grafo na literatura matemática, é uma coleção de vértices ligados por arestas. Um grafo é formado por um par de conjuntos $G = \{V, E\}$, onde V é o conjunto dos N nós (ou vértices, ou pontos) e E é o conjunto das arestas que conectam dois elementos de V (7). Vértices e arestas são também chamados de nós e links na ciência da computação, sítios e conexões na física e atores e relações na sociologia (15). Quando dois nós estão conectados por mais do que uma aresta, dizemos que existem arestas múltiplas, já quando um nó está conectado com ele mesmo, dizemos que existe um laço. Uma rede que não possui arestas múltiplas e laços é chamada de rede simples ou grafo simples. Caso existam arestas múltiplas, o grafo é um multigrafo.

O Estudo formal da teoria de grafos tem sua origem no século 18, com Leonard Euler (1707-1783), cujos trabalhos iniciais concentravam-se em grafos pequenos e com alto grau de regularidade. Euler, em 1735, resolveu o problema das sete pontes de Königsberg (atual Kaliningrado, na Rússia). Nesta cidade existiam sete pontes que conectavam duas

ilhas e as margens do rio Pregel (atual Rio Pregolia) conforme a Fig. 3.1 (16):

Figura 3.1: (a) Ilustração da cidade de Königsberg com as sete pontes. (b) Situação simplificada pelo grafo, onde as arestas são as pontes e os nós são as porções de terra.



Fonte: Adaptado de Newman (2010)(15).

O problema consistia na seguinte questão: é possível andar por toda a cidade de modo a passar por todas as pontes uma única vez? Euler percebeu que o comprimento das pontes não importava, mas sim, quais porções de terra elas ligavam. Com isso ele conseguiu simplificar o problema, como pode ser visto na Fig. 3.1(a). Assim, Euler mostrou que o problema não tinha solução, verificando o número de conexões de cada nó (grau), pois neste grafo todos os nós possuíam grau ímpar. Um problema deste tipo tem solução somente se o grafo possuir zero ou dois nós com grau ímpar (para o primeiro caso o ponto de partida é o mesmo de chegada).

3.2.2 Matriz Adjacência

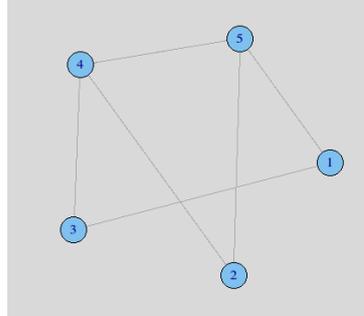
Existem várias formas de representar uma rede matematicamente, uma delas é através da sua matriz adjacência. Para um grafo simples com N vértices, a matriz adjacência é uma matriz quadrada de ordem N , definida da seguinte maneira:

$$A_{ij} = \begin{cases} 1 & \text{se existe uma aresta entre } i \text{ e } j. \\ 0, & \text{caso contrário;} \end{cases} \quad (3.1)$$

Como, por exemplo, temos o grafo gerado no software R (17) com o pacote igraph

(18) na Fig. 3.2:

Figura 3.2: Grafo não direcionado com $N = 5$ nós e $E = 6$ arestas.



Fonte: O autor.

A sua respectiva matriz adjacência A é dada por:

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix} \quad (3.2)$$

Para um grafo não direcionado a matriz adjacência é simétrica, e para um grafo sem laços os elementos da diagonal principal são nulos. Também é possível representar um grafo com arestas múltiplas e com loops usando a matriz adjacência. Uma aresta múltipla é representada na matriz pela sua multiplicidade, por exemplo, uma aresta dupla entre os vértices i e j é representada $A_{ij} = 2$, e arestas que formam laços, o elemento correspondente na diagonal da matriz é $A_{ii} = 2$, pois esta aresta possui duas extremidades (15).

3.2.3 Redes Direcionadas

Uma rede direcionada ou um grafo direcionado, também denominado dígrafo, é um grafo no qual existe uma direção para as arestas, apontando de um vértice para outro. Dentre as redes de mundo real, algumas são direcionadas, como por exemplo, redes de

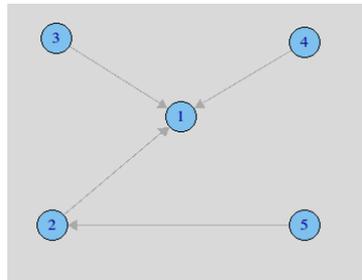
cadeia alimentar predador-presa e redes de citações.

A matriz adjacência neste caso, é definida por:

$$A_{ij} = \begin{cases} 1 & \text{se existe uma aresta a partir de } j \text{ para } i. \\ 0, & \text{caso contrário;} \end{cases} \quad (3.3)$$

Notamos que a direção da aresta é a partir do segundo índice para o primeiro, essa convenção é mais conveniente matematicamente (15). Em geral a matriz adjacência de um grafo direcionado não é simétrica.

Figura 3.3: Grafo direcionado com $N = 5$ nós e $E = 4$ arestas.



Fonte: O autor.

A sua respectiva matriz adjacência A é dada por:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.4)$$

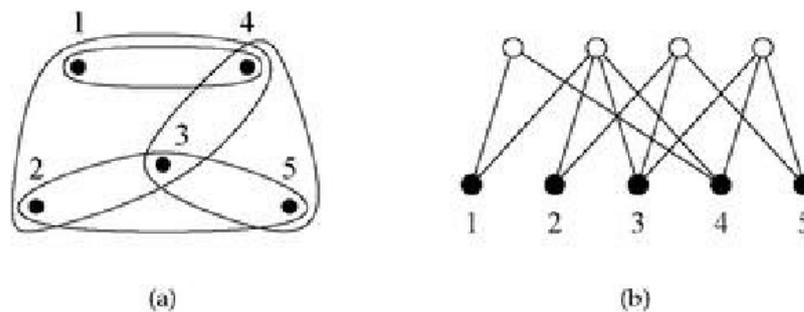
3.2.4 Hipergrafos e Redes Bipartidas

Em alguns tipos de redes as arestas podem ligar mais que dois vértices ao mesmo tempo. Em redes sociais, como por exemplo, as famílias em uma grande comunidade podem ter elementos em comum e a melhor forma de representar esta situação é através da utilização de arestas que ligam mais que dois nós. Tais arestas são denominadas

hiperarestas e a rede correspondente são os hipergrafos. Em sociologia estas redes são chamadas redes de afiliação (15).

O conjunto de vértices conectados por hiperarestas em um hipergrafo também pode ser representado de forma mais conveniente por um grafo bipartido, também chamado de rede de dois modos na sociologia. Em tais redes existem dois conjuntos de vértices, um representando os vértices originais e outro representando os grupos nos quais eles estão inseridos. Por exemplo, uma rede de atores e filmes em que eles atuaram pode ser representada por uma rede bipartida, considerando o conjunto de atores e um conjunto de filmes, onde cada ator está conectado por uma aresta a um filme em que atuou. Também podemos construir uma rede, a partir da rede bipartida, considerando as ligações entre atores, onde existirá uma aresta entre dois atores se eles atuaram juntos em um filme (19).

Figura 3.4: (a) Hipergrafo com 5 nós.(b) Respectivo grafo bipartido.



Fonte: Adaptado de Newman (2010) (15).

A equivalente matriz adjacência para grafos bipartidos é a matriz retangular, denominada matriz incidência. Se n é o número de pessoas ou outros participantes da rede, e g é o número de grupos, então a matriz incidência B é de ordem g por n , definida da seguinte maneira:

$$B_{ij} = \begin{cases} 1 & \text{se o vértice } j \text{ pertence ao grupo } i. \\ 0, & \text{caso contrário;} \end{cases} \quad (3.5)$$

3.2.5 Grau

O grau de um vértice i , denotado por K_i em um grafo, é o número de arestas conectadas a ele. Para uma rede não direcionada com N vértices, o grau pode ser escrito em termos da sua matriz adjacência (16):

$$K_i = \sum_{j=1}^N A_{ij}. \quad (3.6)$$

Cada aresta não direcionada em um grafo possui duas extremidades, e se existem E arestas, então no total existem $2E$ extremidades de arestas, porém, o número de extremidades é igual ao somatório dos graus de todos os nós na rede:

$$2E = \sum_{i=1}^N K_i, \quad (3.7)$$

O grau médio $\langle K \rangle$ de um grafo não direcionado é dado por:

$$\langle K \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2E}{N}. \quad (3.8)$$

Em um grafo não direcionado simples (que não possui arestas múltiplas ou laços), o número máximo de arestas é dado por $\binom{N}{2} = N(N-1)/2$. Assim, podemos definir a densidade ρ de um grafo, como a fração de arestas realmente existentes pelo número total de possíveis arestas:

$$\rho = \frac{E}{\binom{N}{2}} = \frac{2E}{N(N-1)} = \frac{\langle K \rangle}{N-1}. \quad (3.9)$$

Onde $0 \leq \rho \leq 1$. Uma rede para a qual ρ tende a uma constante quando $N \rightarrow \infty$, é dita ser densa. Nestas redes, a fração de elementos não nulos na matriz adjacência

permanece constante quando a rede aumenta. A rede em que $\rho \rightarrow 0$ quando $N \rightarrow \infty$, é dita esparsa (15).

As redes em que todos os vértices apresentam o mesmo grau são denominadas regulares. Em uma rede regular em que todos os vértices possuem mesmo grau K , algumas vezes são denominadas K - regular. Como exemplo de redes regulares, temos as redes triangulares ($K=3$) e quadradas ($K=4$).

Para redes direcionadas, cada vértice possui dois tipos de graus, o de entrada K^{ent} e o de saída K^{sai} . Como para um grafo direcionado, o elemento na matriz adjacência é $A_{ij} = 1$, se existe uma aresta a partir de j até i , os graus de entrada e saída são:

$$K_i^{ent} = \sum_{j=1}^N A_{ij}, \quad (3.10)$$

$$K_j^{sai} = \sum_{i=1}^N A_{ij}, \quad (3.11)$$

O número de arestas E em um grafo direcionado é igual ao número total de arestas de entrada de todos os vértices e também igual ao número total de arestas de saída de todos os vértices, assim:

$$E = \sum_{i=1}^N K_i^{ent} = \sum_{j=1}^N K_j^{sai} = \sum_i \sum_j A_{ij}. \quad (3.12)$$

O grau médio de entrada é igual ao grau médio de saída em um rede direcionada:

$$\langle K_{ent} \rangle = \frac{1}{N} \sum_{i=1}^N K_i^{ent} = \frac{1}{N} \sum_{j=1}^N K_j^{sai} = \langle K_{sai} \rangle. \quad (3.13)$$

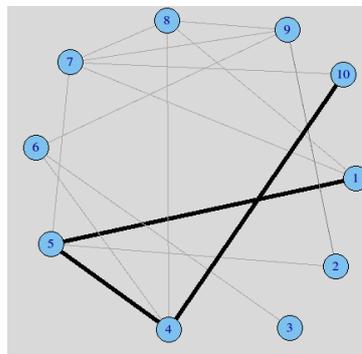
Assim, denotando o grau médio por $\langle K \rangle$ obtém-se:

$$\langle K \rangle = \frac{E}{N}. \quad (3.14)$$

3.2.6 Caminhos

Um caminho em uma rede é uma sequência de vértices tal que cada par de vértices consecutivos são conectados por uma aresta. Em uma rede não direcionada, as arestas podem ser percorridas em qualquer direção, porém para redes direcionadas, estas devem ser percorridas na direção correta. O comprimento de um caminho é o número de arestas que devem ser percorridas, sendo que estas podem ser percorridas mais de uma vez (20).

Figura 3.5: Um caminho de comprimento três em um grafo.



Fonte: O autor.

Pode-se calcular o número de caminhos de comprimento l na rede: para um grafo simples, o elemento da matriz adjacência é $A_{ij} = 1$ se existe uma aresta ligando os vértices i e j , caso contrário $A_{ij} = 0$, portanto o produto $A_{ik}A_{kj} = 1$ se existe um caminho de comprimento 2 ligando os vértices i a j passando por k , e é igual a zero, caso contrário. Assim, o número de caminhos de comprimento 2 a partir de i para j passando por algum outro vértice será dado por:

$$N_{ij}^{(2)} = \sum_{k=1}^N A_{ik}A_{kj} = A^2. \quad (3.15)$$

De modo geral, o número de caminhos de comprimento l é então (16);

$$N_{ij}^{(l)} = A^l. \quad (3.16)$$

Podemos classificar os caminhos em: geodésicos, Eulerianos ou Hamiltonianos. Ca-

minhos geodésicos ou menores caminhos são aqueles com menor comprimento (distância geodésica) que conectam dois vértices na rede, isto é, o menor valor de l tal que $A^l > 0$ (15). Podemos definir o diâmetro de um grafo através da maior distância geodésica que conectam dois vértices quaisquer e se existirem vértices isolados o diâmetro é infinito. Um caminho Euleriano é um caminho que percorre cada aresta apenas uma vez, ao passo que, um caminho Hamiltoniano é um caminho que visita cada vértice apenas uma vez (14).

3.3 Medidas de Centralidade

Uma grande quantidade de trabalhos relacionados a redes se concentram no tema centralidade, isto é, em como quantificar o quão central é um determinado vértice, no sentido de importância nesta rede (21). Nesta seção são descritas algumas das medidas mais utilizadas na literatura de redes complexas.

3.3.1 Centralidade de Grau

A medida mais simples de centralidade de um vértice na rede é simplesmente o seu grau, ou seja, a quantidade de arestas desse vértice. Em redes direcionadas, os vértices possuem um determinado grau de entrada e um grau de saída. Embora o grau de um vértice seja uma medida simples de centralidade, ele pode fornecer um bom indicador desta, como por exemplo, em redes sociais é razoável supor que os indivíduos que sejam mais conectados sejam os mais influentes e que tenham um maior acesso à informações (14).

3.3.2 Centralidade por Autovetor

A centralidade por autovetor, proposta primeiramente por Bonacich em 1987, é uma extensão da centralidade pontual determinada pelo grau de um vértice. Este tipo de centralidade leva em conta também a importância dos vizinhos deste vértice, fornecendo

um índice que é proporcional aos índices de seus vizinhos (15).

Para deduzir uma expressão para a centralidade por autovetor x_i de um vértice i , supõe-se que $x_i = 1$ para todo i , e assim pode-se determinar x'_i , que é definida como a soma das centralidades de seus vizinhos. Então:

$$x'_i = \sum_j A_{ij}x_j. \quad (3.17)$$

onde A_{ij} é um elemento da matriz adjacência. Podemos escrever esta expressão na forma matricial $\mathbf{x}' = A\mathbf{x}$, onde \mathbf{x} é um vetor cujas componentes são as centralidades dos vértices vizinhos. Repetindo o processo, depois de t iteradas, o vetor de centralidade $x(t)$ será dado por (15):

$$\mathbf{x}(t) = A^t\mathbf{x}(0), \quad (3.18)$$

Pode-se escrever $\mathbf{x}(0)$ como sendo a combinação linear dos autovetores da matriz adjacência:

$$\mathbf{x}(0) = \sum_i c_i \mathbf{v}_i, \quad (3.19)$$

Portanto:

$$\mathbf{x}(t) = A^t \sum_i c_i \mathbf{v}_i = \sum_i c_i \lambda_i^t \mathbf{v}_i = \lambda_1^t \sum_i c_i \left(\frac{\lambda_i}{\lambda_1}\right)^t \mathbf{v}_i. \quad (3.20)$$

onde λ_i são os autovalores da matriz adjacência A , e λ_1 sendo o maior deles. Como $\lambda_i/\lambda_1 \leq 1$, todos os termos no somatório tendem a zero quando $t \rightarrow \infty$ e assim $\mathbf{x}(t) \rightarrow c_1 \lambda_1 \mathbf{v}_1$. Assim, o vetor de centralidades é proporcional ao autovetor principal da matriz adjacência e pode-se escrever:

$$A\mathbf{x} = \lambda_1 \mathbf{x}, \quad (3.21)$$

Assim tem-se que a centralidade de autovetor de um vértice i será dada por:

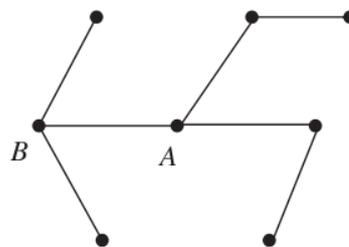
$$x_i = \lambda_1^{-1} \sum_j A_{ij} x_j. \quad (3.22)$$

Este tipo de centralidade pode assumir valores altos para alguns vértices devido a presença de muitos vizinhos, ou devido a vizinhos importantes ou ambos.

3.3.3 Centralidade de Intermediação

A centralidade de intermediação é ferramenta essencial para a análise de redes sociais, pois é a medida de centralidade mais adotada para mensurar a importância de um determinado vértice na rede (21). Esta medida de centralidade foi proposta por Freeman (1977), e fornece a importância de um vértice baseado na sua posição na rede. Como visto no item anterior, o grau de um vértice é uma medida de centralidade, porém, nem sempre ele determina a importância de um determinado vértice. Como exemplo, podemos considerar uma pequena rede destacando dois vértices A e B com o mesmo grau, como na Fig. (3.6). Se removermos o vértice B, somente dois vértices serão desconectados, porém, removendo o vértice A, uma grande quantidade de vértices serão desconectados, indicando que o vértice A é o mais importante do que o vértice B (14).

Figura 3.6: A centralidade de intermediação do vértice A é maior do que a do vértice B, apesar de os dois possuírem o mesmo grau



Fonte: Adaptado de Cohen et.al (2010) (14).

A centralidade de intermediação de um determinado vértice $v \in G$ é baseada na contagem dos menores caminhos através deste, da seguinte maneira (21):

$$b(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}. \quad (3.23)$$

onde σ_{st} é a quantidade de menores caminhos que ligam todos os pares de vértices, e $\sigma_{st}(v)$ é a quantidade de menores caminhos que ligam todos os pares de vértices que passam por v .

Os vértices com uma alta centralidade de intermediação podem ter uma considerável influência dentro da rede, pois por eles passam uma maior quantidade de informação se comparados com os demais vértices. Removendo alguns desses vértices pode-se também interromper a comunicação entre alguns vértices, devido ao alto número de caminhos que passam por eles (22).

3.3.4 Centralidade de Proximidade

A centralidade de proximidade é uma medida de centralidade que leva em conta a distância média de um vértice a todos os outros da rede. Supondo que d_{ij} é o comprimento do menor caminho entre i e j , ou seja, o número de arestas ao longo do caminho. Então a distância média entre i e j , calculada para todos outros j vértices da rede será dada por (15):

$$l_{ij} = \frac{1}{n} \sum_j d_{ij}. \quad (3.24)$$

Os vértices que possuem uma distância média pequena aos outros vértices na rede, podem ter um acesso maior às informações e influenciar diretamente outros vértices. Em uma rede social, por exemplo, uma pessoa com uma distância média pequena em relação as outras, pode levar suas opiniões para uma quantidade maior de pessoas da comunidade mais facilmente que outras pessoas com distância média maior. Desde que baixos valores de l_i indicam uma maior centralidade, pode-se definir a centralidade de proximidade como o inverso da distância média (21):

$$C_i = \frac{1}{l_i} = \frac{n}{\sum_j d_{ij}}. \quad (3.25)$$

3.4 Modelos para Redes Complexas

A teoria de grafos tem sido a principal ferramenta para o estudo de redes complexas. Até a metade do século passado, esta teoria se concentrava em grafos regulares, porém, no final da década de 50, com trabalhos dos matemáticos Húngaros Paul Erdős e Alfred Renyi, a teoria dos grafos tornou-se mais probabilística e algorítmica (7). O modelo de grafos aleatórios proposto por Erdős e Renyi para descrever redes grandes de mundo real guiou as pesquisas sobre redes complexas por algumas décadas.

3.4.1 Redes Aleatórias: Modelo de Erdős e Renyi

Em geral um grafo aleatório é um modelo de rede em que algum conjunto específico de parâmetros toma valores fixos, mas a rede é aleatória em outros aspectos. Um exemplo simples de um grafo aleatório é considerar uma rede em que fixamos o número de vértices N e de arestas E , e assim conectamos pares de vértices de forma aleatória, não permitindo arestas múltiplas e laços. Este modelo é conhecido como modelo $G(N,E)$ (15).

No primeiro artigo fundamental sobre grafos aleatórios, Erdős e Rényi definem um grafo aleatório como N nós rotulados conectados por E arestas, que são escolhidas aleatoriamente a partir de $N(N-1)/2$ possíveis arestas. No total existem $C_{N(N-1)/2}^E$ grafos com N nós e E arestas formando o espaço amostral. Uma definição alternativa para grafo aleatório é a partir do modelo binomial: inicia-se com N nós, onde cada par de nós é conectado com uma probabilidade p . Assim, o número total de arestas é uma variável aleatória com o valor esperado $E(E) = p \frac{N(N-1)}{2}$ (7).

Uma das propriedades estudadas por Erdős e Renyi foi o surgimento de subgrafos. Um grafo G_1 que possui um conjunto P_1 de nós e um conjunto E_1 de arestas é um subgrafo do grafo $G = \{P, E\}$ se todos os nós de P_1 são também nós de P e todas as arestas de E_1

são também arestas de E . Exemplos simples de subgrafos são ciclos, árvores e subgrafos completos. Um ciclo de ordem k é um percurso fechado com k arestas tais que cada duas arestas consecutivas possuem um nó comum. Graficamente um triângulo é um ciclo de ordem 3 e um retângulo é um ciclo de ordem 4. O grau médio de um ciclo é igual a 2, pois cada nó possui duas arestas. Grafos com característica de árvores, são opostos aos ciclos, pois não possuem percursos fechados. Um grafo é uma árvore de ordem k se possui k nós e $k - 1$ arestas. O grau médio de uma árvore de ordem k é $\langle k \rangle = 2 - 2/k$, que é aproximadamente 2 para árvores grandes. Um subgrafo completo de ordem k contém k nós e todas as possíveis $k(k - 1)/2$ arestas, isto é, um grafo totalmente conectado (7).

A distribuição dos graus máximo e mínimo de grafos aleatórios foi obtida a Erdős e Rényi e a distribuição completa foi obtida em 1981 por Bollobás (7). Para redes aleatórias ela segue uma distribuição de Poisson para N grande, fornecendo a probabilidade de ao ser sorteado aleatoriamente um nó no grafo, este possua grau K .

Para a determinação da distribuição de grau de uma rede aleatória, notamos que um vértice pode ser conectado com probabilidade p a no máximo $N - 1$ vértices. Assim, a probabilidade de um dado vértice ser conectado com outros K vértices é dada por (15):

$$p_K = \binom{N-1}{K} p^K (1-p)^{N-1-K}. \quad (3.26)$$

Assim $G(N, p)$ possui uma distribuição de grau binomial. Muitas redes tendem a ter um grau médio aproximadamente constante quando N é muito grande, por exemplo, a quantidade de amigos de uma pessoa não depende do número de pessoas no mundo. Considerando o fato de que o grau médio de uma rede aleatória é $\langle K \rangle = (N - 1)p$, isto é, $p = \frac{\langle K \rangle}{N - 1}$, então tem-se a seguinte aproximação para $N \rightarrow \infty$:

$$\ln(1-p)^{N-1-K} = (N-1-K) \ln(1-p) = (N-1-K) \ln\left(1 - \frac{\langle K \rangle}{N-1}\right) \quad (3.27)$$

Expandindo o logaritmo em série de Taylor tem-se:

$$\ln(1-p)^{N-1-K} \approx -(N-1-K) \frac{\langle K \rangle}{N-1} \approx -\langle K \rangle, \quad (3.28)$$

pois $N-1 \gg K$, portanto:

$$\ln(1-p)^{N-1-K} = -\langle K \rangle \quad (1-p)^{N-1-K} = e^{-\langle K \rangle} \quad N \rightarrow \infty, \quad (3.29)$$

Considerando novamente $N \rightarrow \infty$, tem-se:

$$\binom{N-1}{K} = \frac{(N-1)!}{K!(N-1-K)!} = \frac{(N-1) \cdot (N-1-1) \dots (N-1-K+1)}{K!}, \quad (3.30)$$

Como $N \gg K$:

$$\binom{N-1}{K} = \frac{(N-1)^K}{K!}, \quad (3.31)$$

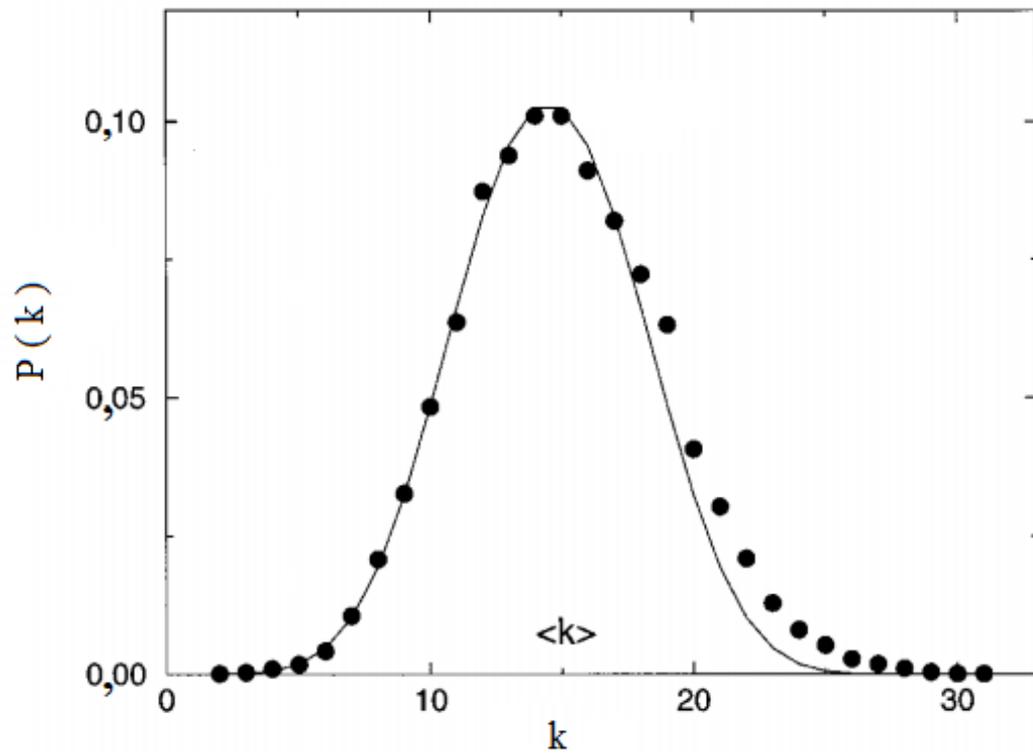
Assim, substituindo (3.29) e (3.31) em (3.26) teremos:

$$p_K = \frac{(N-1)^K}{K!} p^K e^{-\langle K \rangle} = \frac{(N-1)^K}{K!} \left(\frac{\langle K \rangle}{N-1} \right)^K e^{-\langle K \rangle} = e^{-\langle K \rangle} \frac{\langle K \rangle^K}{K!}. \quad (3.32)$$

No limite quando N é grande, a distribuição de grau do modelo $G(N, p)$ tende a uma distribuição de Poisson (7) como se pode verificar na Fig. 3.7.

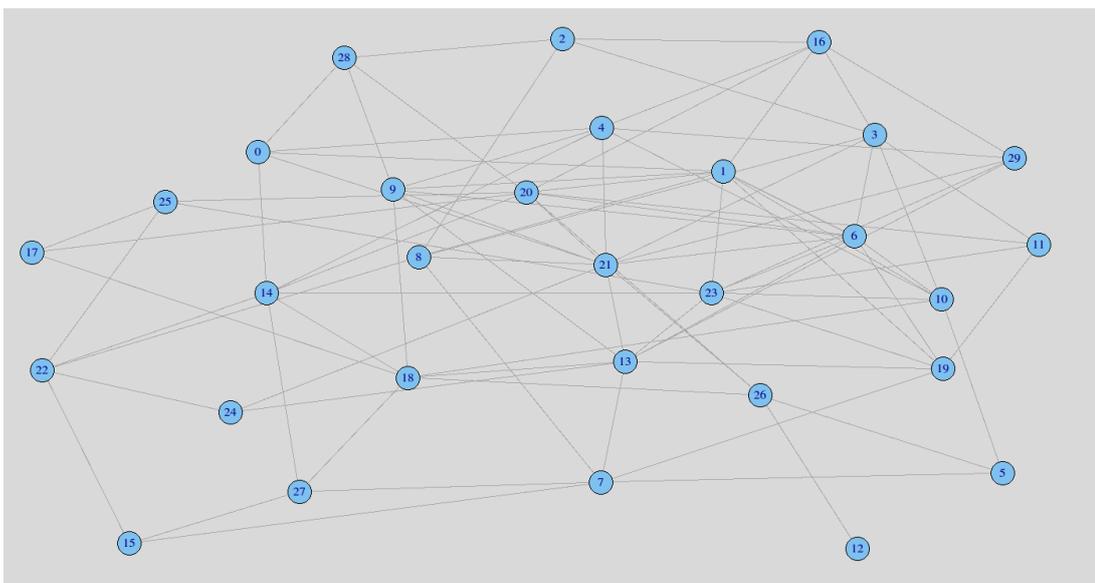
Portanto, nesse tipo de rede uma grande parte dos nós, apresentam conectividade próxima ao valor do grau médio. A diferença entre os valores do grau máximo e mínimo é relativamente pequena. Estes fatos são ilustrados na Fig. 3.8.

Figura 3.7: Distribuição de grau de uma rede aleatória .



Fonte: Adaptado de Barabási e Albert (2002) (7).

Figura 3.8: Rede aleatória no modelo de ER com $p = 0,2$ e $N = 30$ nós.



Fonte: O autor.

Conectividade e Diâmetro

O diâmetro de um grafo é a máxima distância entre quaisquer pares de vértices. Para um grafo que possui vértices isolados o diâmetro pode ser considerado infinito. O diâmetro d de grafos aleatórios é proporcional a $\ln(N)/\ln(\langle K \rangle)$, isto é, depende do logaritmo do número de nós (7):

$$d = \frac{\ln(N)}{\ln pN} = \frac{\ln(N)}{\ln(\langle K \rangle)}. \quad (3.33)$$

Também podemos calcular a distância média entre os pares de vértices, ou o comprimento de caminho médio l . Como esperado, este é proporcional ao número de nós, da mesma maneira que o diâmetro (7):

$$l \approx \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad (3.34)$$

Coefficiente de Aglomeração

O coeficiente de aglomeração C é a medida da transitividade da rede. É definido como a probabilidade de dois vizinhos de um dado vértice também serem vizinhos um do outro (23). Redes complexas reais exibem um alto coeficiente de aglomeração (7). Se considerarmos um vértice de uma rede aleatória e seus primeiros vizinhos, a probabilidade de que dois desses vizinhos também sejam conectados é igual a probabilidade de que dois vértices selecionados aleatoriamente sejam conectados. Portanto, o coeficiente de aglomeração de um grafo aleatório é (15):

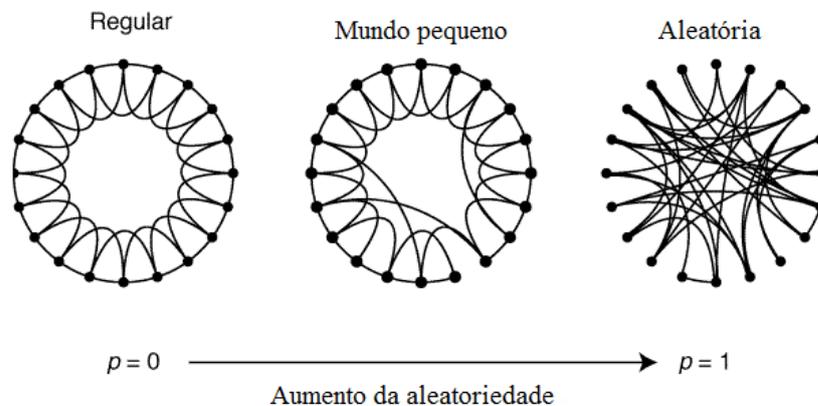
$$C = p = \frac{\langle k \rangle}{N}. \quad (3.35)$$

3.4.2 Redes *Mundo Pequeno*

3.4.2.1 Modelo de Watts e Strogatz (WS)

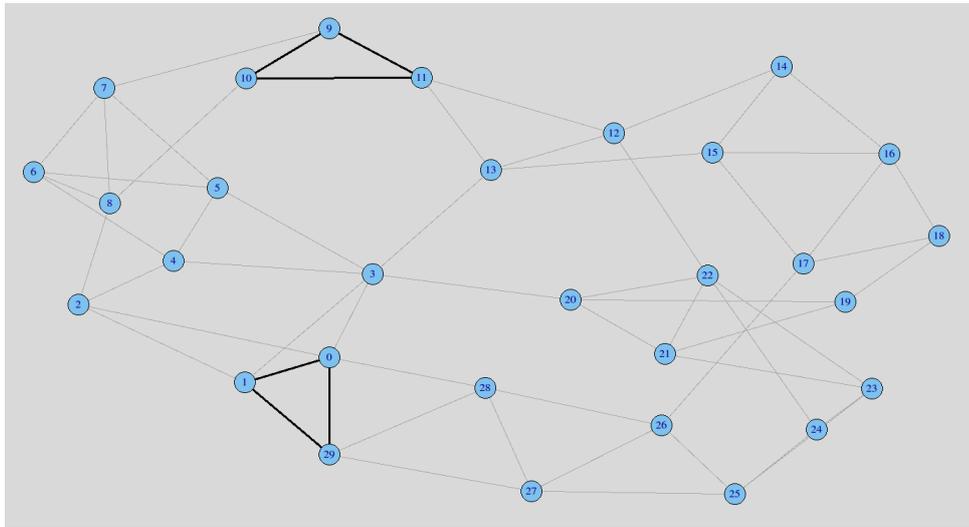
Uma das mais conhecidas características de redes de mundo real é a transitividade ou aglomeração, a qual é mensurada pelo coeficiente de aglomeração C , e está relacionada com a probabilidade de dois vizinhos de um vértice serem também vizinhos um do outro. Para uma rede aleatória, quando $N \rightarrow \infty$, o coeficiente de aglomeração tende a zero. Porém, este resultado é muito menor que o observado em redes reais (23). Uma outra característica é que em média a distância entre quaisquer dois nós na rede é relativamente pequena. Esta quantidade é mensurada pelo comprimento de caminho médio l . Em 1998, Watts e Strogatz propuseram um modelo de rede que consegue reproduzir estas duas propriedades (8). O algoritmo para construção dessa classe de redes é o seguinte: inicia-se com uma rede regular, como um anel por exemplo, e com uma probabilidade p , um nó escolhido aleatoriamente na rede é religado com algum outro nó. Esse processo insere ligações de longo alcance, o que implica em um baixo valor de l . No entanto, a estrutura regular da rede original permanece, e assim a transitividade é alta. Este modelo possui características intermediárias de uma rede regular e uma aleatória, como é ilustrado na Fig. 3.9. O alto coeficiente de aglomeração pode ser verificado pela presença de diversos triângulos na rede (Fig. 3.10).

Figura 3.9: Procedimento de religação de arestas em uma rede regular



Fonte: Adaptado de Watts e Strogatz (1998)(8).

Figura 3.10: Rede mundo pequeno no modelo de Watts e Strogatz com $p = 0,05$ e $N = 30$ nós.



Fonte: O autor.

Este modelo possui suas raízes em sistemas sociais, onde a maioria das pessoas são amigas de seus vizinhos imediatos, vizinhos na mesma rua, escola etc. Em outro sentido, cada pessoa tem um amigo em algum lugar distante, em outras cidades ou países, que são representados pelas arestas de longo alcance obtidas pelo religamento no modelo de WS.

Uma variante do modelo original de WS foi proposto em 1999 por Newman e Watts (24). A diferença básica neste modelo é que com uma probabilidade p adicionamos arestas ("atalhos") entre dois vértices escolhidos de forma aleatória na rede, mantendo a rede regular original.

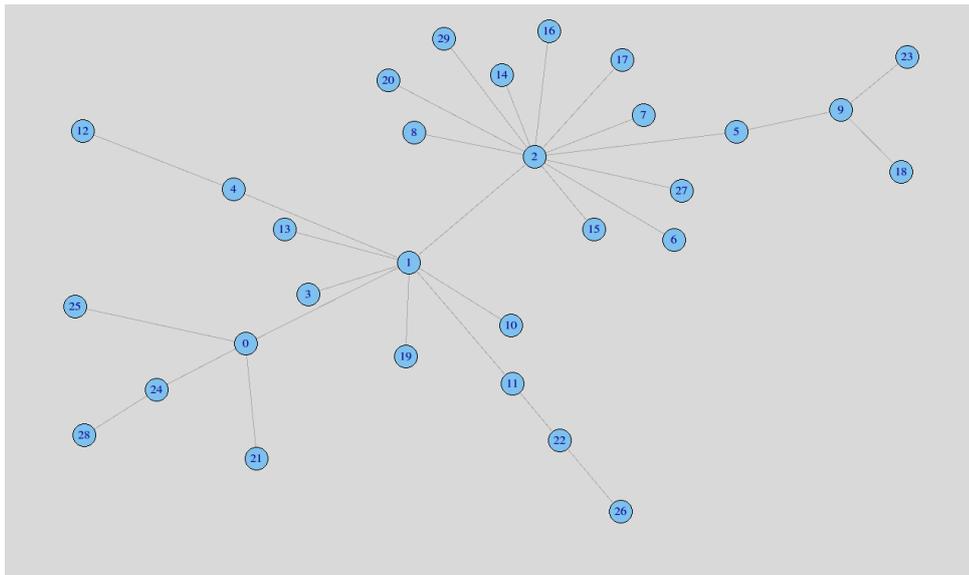
3.4.3 Redes de Escala Livre

3.4.3.1 Modelo de Barabási e Albert

Resultados empíricos demonstram que muitas redes grandes são de escala livre, isto é, elas seguem uma distribuição de grau do tipo lei de potência para um grande K . O surgimento de distribuições de grau do tipo lei de potência foi rastreada por dois mecanismos que faltam no modelo clássico de redes aleatórias. Primeiro, a maioria das redes crescem pela adição de novos nós que se ligam aos já presentes no sistema, e continuamente elas

se expandem pela adição de novos nós. Segundo, a grande parte de redes reais apresentam um comportamento de ligação preferencial, isto é, os novos nós adicionados irão conectar-se com maior probabilidade aos nós com maior conectividade na rede. Assim, estas redes possuem poucos nós com grau alto e a maioria apresentam grau baixo. Este fato é ilustrado na Fig. 3.11.

Figura 3.11: Rede de Escala Livre no modelo de Barabási e Albert com $N = 30$ nós.



Fonte: O autor.

Crescimento e ligação preferencial são os ingredientes fundamentais deste modelo, o qual possui uma distribuição de grau do tipo lei de potência. O algoritmo para gerar uma rede de escala livre é o seguinte (9):

Crescimento: Iniciando com um pequeno conjunto (m_0) de nós, em cada instante de tempo adicionamos um novo nó com $m \leq m_0$ arestas que ligam o novo nó a m nós já presentes no sistema;

Ligação preferencial: A probabilidade de que o novo nó conecte-se a um nó i é proporcional ao grau k_i tal que:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \quad (3.36)$$

As propriedades dinâmicas do modelo de escala livre podem ser obtidas através de

algumas aproximações analíticas. A teoria do *continuum* proposta por Barabási e Albert (1999) (9) foca na dinâmica do grau dos nós.

Teoria do *continuum*: Esta aproximação calcula a dependência temporal do grau k_i de um dado nó i . Este grau irá aumentar a cada instante de tempo devido a entrada de novos nós no sistema que se ligam ao nó i , sendo a probabilidade deste processo $\Pi(k_i)$. Assumindo que k_i é uma variável real contínua, é esperado que a taxa com que k_i varia seja proporcional a $\Pi(k_i)$. Assim, k_i satisfaz a equação:

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m \frac{k_i}{\sum_{j=1}^{N-1} k_j}. \quad (3.37)$$

A soma no denominador é sobre todos os nós no sistema, exceto o novo nó introduzido, e seu valor é $\sum_j k_j = 2mt - m$, assim temos:

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t}. \quad (3.38)$$

A solução desta equação, com a condição inicial que cada nó i introduzido possua $k_i(t_i) = m$, é:

$$k_i(t) = m \left(\frac{t}{t_i} \right)^\beta, \quad (3.39)$$

com $\beta = \frac{1}{2}$.

Esta equação indica que o grau de todos os nós evoluem da mesma maneira, seguindo uma lei de potência.

A probabilidade de que um nó possua grau $k_i(t)$ menor que k , $P(k_i(t) < k)$ pode ser escrita como

$$P(k_i(t) < k) = P(t_i > \frac{m^{1/\beta}t}{k^{1/\beta}}). \quad (3.40)$$

Assumindo que os nós são adicionados em intervalos de tempo iguais na rede, os

valores de t_i possuem uma densidade de probabilidade constante.

$$P(t_i) = \frac{1}{m_0 + t}, \quad (3.41)$$

Substituindo na equação (3.40), obtemos:

$$P(t_i > \frac{m^{1/\beta}t}{k^{1/\beta}}) = 1 - \frac{m^{1/\beta}t}{k^{1/\beta}(m_0 + t)}. \quad (3.42)$$

A distribuição de grau pode ser obtida como:

$$P(k) = \frac{\partial}{\partial k} P[k_i(t) < k] = \frac{1}{\beta} \frac{m^{1/\beta}t}{m_0 + t} \frac{1}{k^{1/\beta+1}}, \quad (3.43)$$

Quando $t \rightarrow \infty$, temos:

$$P(k) \sim 2m^2 k^{-3}. \quad (3.44)$$

Na Fig. 3.12 é apresentada a distribuição de grau obtida para uma rede com $N = 300000$ nós.

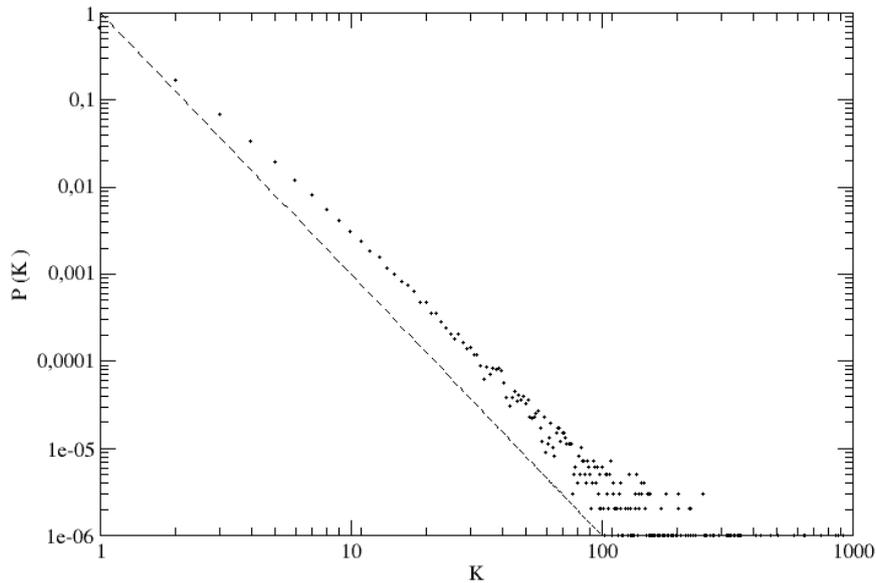
Assim, neste modelo a distribuição de grau é do tipo lei de potência $P(k) \sim k^{-\gamma}$ com $\gamma \rightarrow 3$. No entanto, para redes de mundo real, os valores para este expoente estão entre $2 \leq \gamma \leq 3$ (7). Existem algumas variações para este modelo. Uma delas adota a probabilidade de ligação preferencial ainda linear, mas com um deslocamento A , ou seja, (25):

$$\Pi(i) = \frac{k_i + A}{\sum_j (k_j + A)}. \quad (3.45)$$

Com isso, a distribuição de grau é do tipo lei de potência $P(k) \sim k^{-\gamma}$, com o expoente γ dado por (14):

$$\gamma = 3 + A/m \quad A > -m. \quad (3.46)$$

Figura 3.12: Distribuição de grau de redes de escala livre no modelo de Barabási e Albert com $N = 300000$ nós. O coeficiente angular da reta tracejada é $-2,9$.



Fonte: O autor.

Com essa probabilidade de ligação preferencial, a distribuição de grau é uma lei de potência com γ dado em função do parâmetro A com valores $2 \leq \gamma \leq \infty$. A outra variante adota a probabilidade de ligação preferencial não linear dada por Krapivsky et al. (2001) (26):

$$\Pi(i) = \frac{k_i^\alpha}{\sum_j k_j^\alpha}. \quad (3.47)$$

Para este caso existem 3 regimes possíveis (14): O sublinear, para $\alpha < 1$, $\alpha = 1$ e o superlinear com $\alpha > 1$. No primeiro caso, a distribuição de grau possui um corte exponencial. O segundo recupera o modelo original de Barabási e Albert, e para o terceiro tem-se uma situação em que a maioria dos nós estão conectados a um único nó (14).

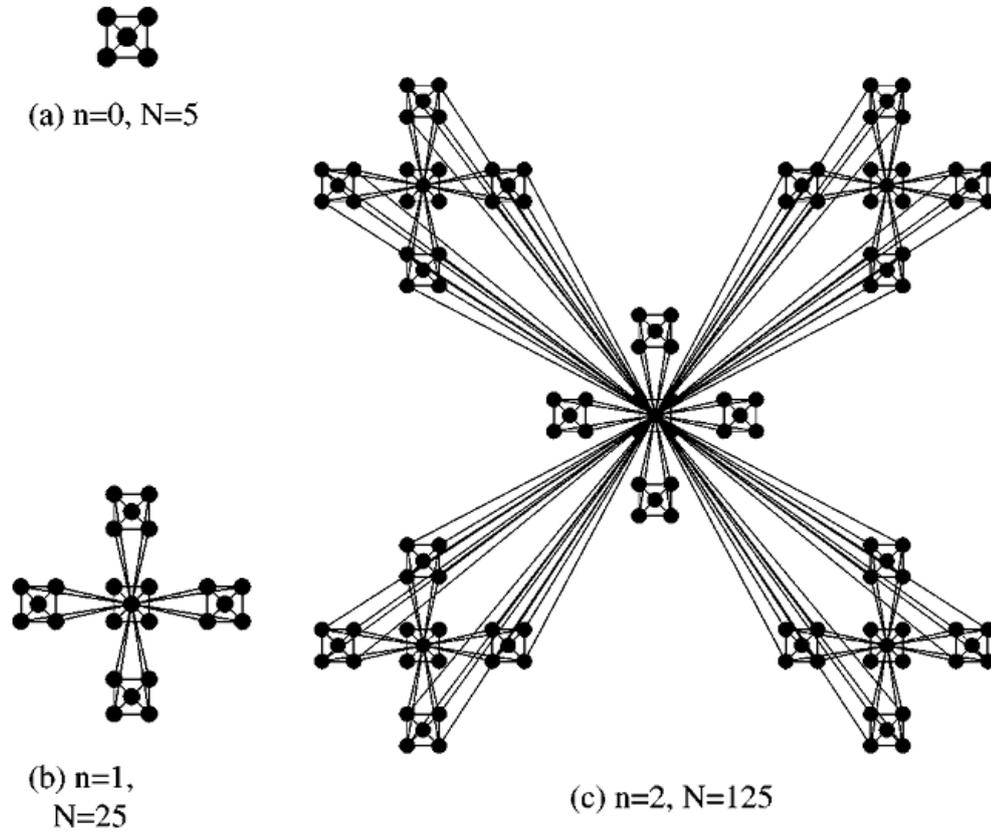
3.4.3.2 Redes Hierárquicas

Muitas redes de mundo real compartilham as propriedades de serem redes de escala livre e um alto coeficiente de aglomeração. Foi demonstrado que as duas características são consequências da organização hierárquica, em que pequenos grupos de nós se organizam de forma hierárquica formando redes maiores (27). Estes autores propuseram modelos para construção de tais redes de forma hierárquica (28) ou determinística (29), e mostraram analiticamente que a distribuição de grau segue uma lei de potência $P(k) \sim k^{-\gamma}$, com o expoente $\gamma = 1 + \frac{\ln 5}{\ln 4} \approx 2,16$ e $\gamma = \frac{\ln 3}{\ln 2} \approx 1,58$ respectivamente.

A construção das redes nos dois modelos apresentados se dá de forma iterativa. Para o modelo proposto em (28), inicia-se em $t = 0$ com um grafo completo com $N = 5$ nós, sendo o nó central denominado nó raiz. Em $t = 1$ são construídas 4 réplicas e então são conectados todos os nós periféricos com o nó raiz, gerando uma rede com $N = 25$ nós. Em $t = 2$ novamente são construídas 4 réplicas da rede do instante anterior, e todos os nós periféricos são conectados com o nó raiz gerando uma rede com $N = 125$ nós. Esse procedimento pode ser repetido para um número qualquer de passos, e assim a rede é construída de forma iterativa como é mostrado na Fig. 3.13:

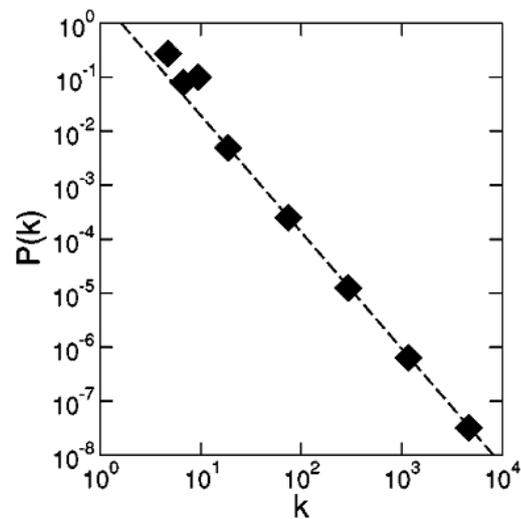
A rede gerada possui um alto coeficiente de aglomeração $C \approx 0,743$ e a distribuição de grau é do tipo lei de potência com $\gamma = 1 + \frac{\ln 5}{\ln 4} \approx 2,16$ obtido numericamente, como pode ser verificado na Fig. 3.14 (28).

Figura 3.13: Processo para construção de uma rede hierárquica.



Fonte: Adaptado de Ravasz e Barabási (2003) (28).

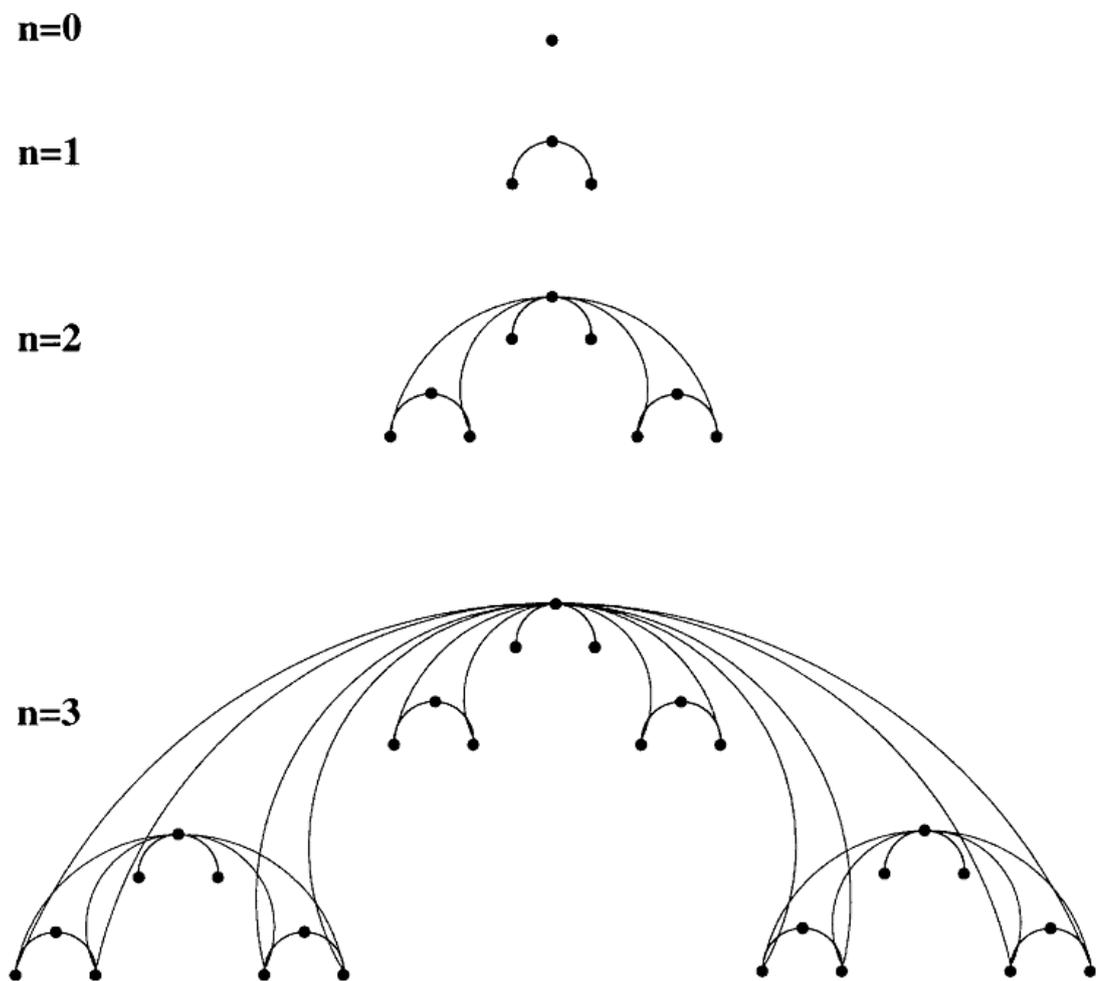
Figura 3.14: Distribuição de grau de uma rede hierárquica.



Fonte: Adaptado de Ravasz e Barabási (2003) (28).

Para o modelo proposto em (29), a construção da rede é feita de modo análogo, porém a rede é iniciada em $t = 0$ com apenas um nó (raiz), e a cada iterada adicionamos duas réplicas da rede na iterada anterior. Em geral, na n -ésima iterada, adicionamos duas réplicas da rede da iterada anterior, com 3^{n-1} nós cada uma, e conectamos os 2^n nós inferiores com o nó raiz, conforme Fig. 3.15. A distribuição de grau para esta rede segue uma lei de potência com $\gamma = \frac{\ln 3}{\ln 2} \approx 1,58$.

Figura 3.15: Rede determinística.



Fonte: Adaptado de Barabási et. al (2001) (29).

3.4.4 Espectro de Autovalores de Redes Complexas

3.4.4.1 Matriz Adjacência e o Espectro de um Grafo

Um grafo G pode ser representado pela sua matriz adjacência $A(G)$ onde $A_{ij} = 1$ se os vértices i e j são conectados e, caso contrário, $A_{ij} = 0$. O espectro de autovalores de um grafo é o conjunto de autovalores de sua matriz adjacência e ele fornece informações sobre as suas propriedades estruturais (14).

A densidade espectral é a densidade de autovalores $\rho(\lambda)$ de sua matriz adjacência, e pode ser escrita como (30):

$$\rho(\lambda) = \frac{1}{N} \sum_{j=1}^N \delta(\lambda - \lambda_j). \quad (3.48)$$

que converge para uma função contínua quando $N \rightarrow \infty$. λ_j é o j -ésimo maior autovalor de sua matriz adjacência.

A densidade espectral pode ser relacionada diretamente com as propriedades topológicas do grafo, onde o k -ésimo momento pode ser escrito como:

$$M_k = \frac{1}{N} \sum_{j=1}^N (\lambda_j)^k = \frac{1}{N} \text{Tr}(A^k) = \frac{1}{N} \sum_{i_1, i_2, \dots, i_k} A_{i_1, i_2} A_{i_2, i_3} \dots A_{i_k, i_1}. \quad (3.49)$$

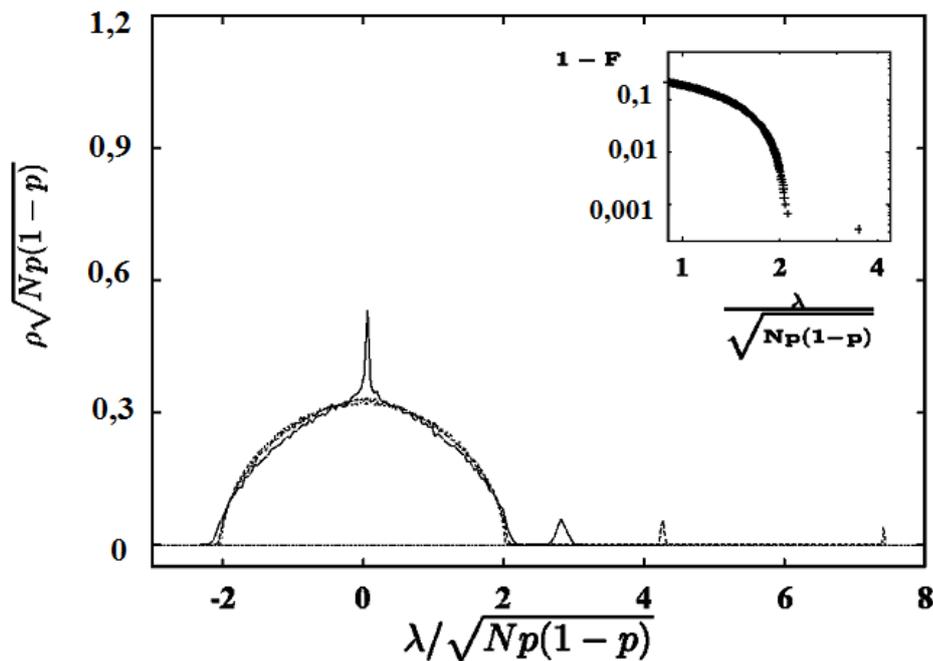
3.4.4.2 Espectro de Redes Complexas: Lei do Semicírculo

A forma geral da lei do semicírculo para matrizes simétricas é a seguinte: Se $A_{N \times N}$ é uma matriz real, simétrica e não correlacionada com $\langle A_{ij} \rangle = 0$ e $\langle A_{ij}^2 \rangle = \sigma^2$ para cada $i \neq j$, e se com o aumento de N os demais momentos permanecem finitos, então, quando $N \rightarrow \infty$, a densidade espectral limite ou a densidade e autovalores de A/\sqrt{N} converge para a distribuição semicircular (14) e (30):

$$\rho(\lambda) = (2\pi\sigma^2)^{-1} \sqrt{4\sigma^2 - \lambda^2} \quad \text{se } |\lambda| < 2\sigma. \quad (3.50)$$

A matriz adjacência de grafos não segue as condições da lei de Wigner, pois o valor esperado de seus elementos é uma constante diferente de zero. Porém, para grafos aleatórios no modelo de ER, a distribuição espectral normalizada converge para a distribuição semicircular, conforme pode ser visualizado na Fig. 3.16 (30).

Figura 3.16: Densidade espectral para um grafo aleatório com $p = 0,05$. Na figura inserida: A densidade espectral decai exponencialmente, onde F é a função distribuição espectral cumulativa.



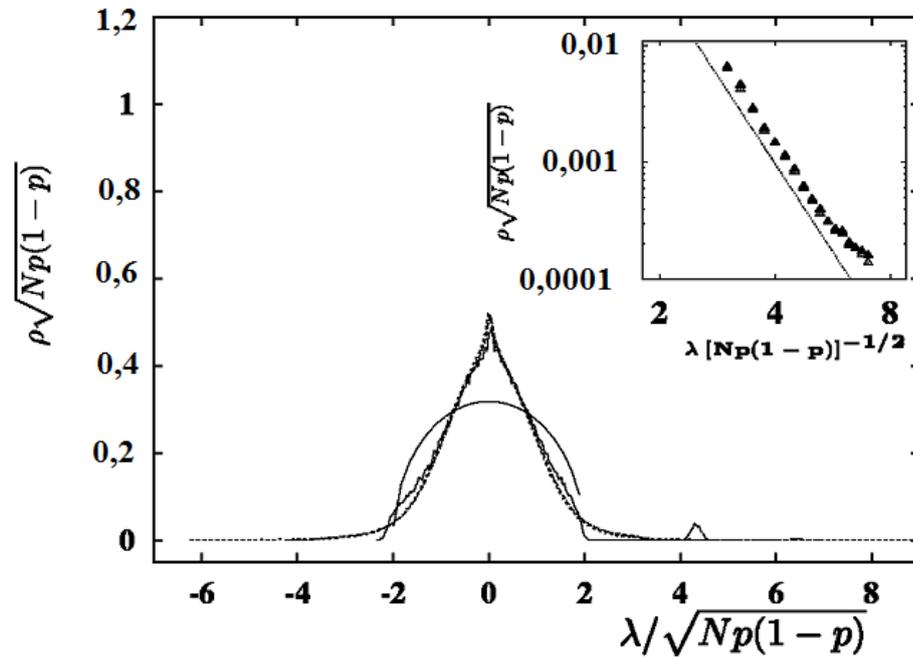
Fonte: Adaptado de Farkas et. al (2001) (30).

Vemos que o autovalor principal (maior autovalor λ_1), está afastado dos demais. Ou seja, ele cresce mais rapidamente que $\{\lambda_2, \dots, \lambda_N\}$ pois λ_1 cresce com N enquanto os demais são proporcionais a \sqrt{N} . Para as regiões onde $\lambda \approx \pm 2\sqrt{Np(1-p)}$, a densidade espectral decai exponencialmente quando $N \rightarrow \infty$. Este fato é mostrado na Fig. 3.16 inserida. Para redes de mundo pequeno, temos uma situação análoga, pois no limite $p = 1$ temos uma rede aleatória de ER e, portanto, a convergência para a lei de Wigner (30).

Considerando redes de escala livre, se $m = m_0 = 1$ a rede é, por definição, uma árvore. Assim, o seu espectro é simétrico e, portanto $\lambda_1 = -\lambda_N$ (7). Para $m > 1$, o autovalor principal λ_1 é deslocado dos demais autovalores. O conjunto $\{\lambda_2, \dots, \lambda_N\}$ converge para uma função contínua no formato triangular para valores normalizados de λ e possui cauda tipo

lei de potência, o que pode ser verificado na Fig. (3.17). Assim o espectro de autovalores de uma rede com escala livre não segue a lei do semicírculo de Wigner (30).

Figura 3.17: Densidade espectral para rede com escala livre no modelo de Barabási e Albert com $m = m_0 = 5$. Na figura inserida: A densidade espectral possui um decaimento do tipo lei de potência.



Fonte: Adaptado de Farkas et. al (2001) (30).

Assim, alguns tipos de redes complexas apresentam o espectro de autovalores simétrico e outras assimétrico. Esta propriedade tem grande influência no problema de caminhada aleatória nessas redes, conforme será abordado no capítulo seguinte.

Capítulo 4

Caminhada Aleatória em Redes Complexas

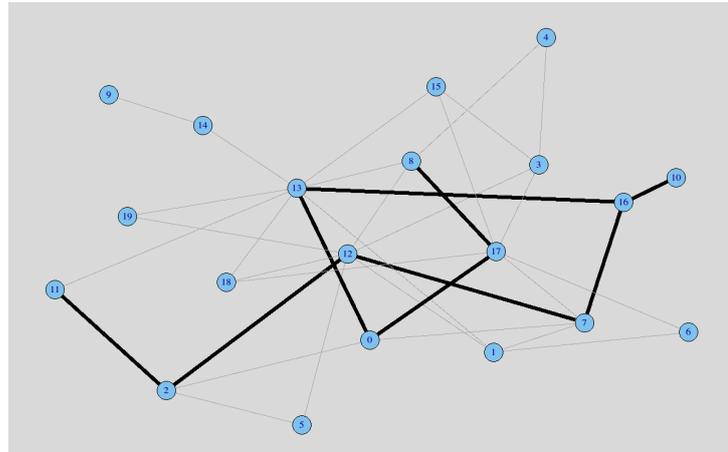
4.1 Caminhada Aleatória em Grafos

Um passo de uma caminhada aleatória em um grafo G , iniciado em um vértice i , é realizado escolhendo algum vértice j do conjunto de vizinhos de i de forma aleatória. O passo seguinte se dá pela escolha aleatória de algum vizinho de j , e assim sucessivamente. A repetição deste processo indefinidamente caracteriza uma caminhada aleatória no grafo (Fig. 4.1).

O problema clássico de um caminhante errante nas ruas de uma cidade foi primeiramente estudado por Pólya (1919) (31). Ele percebeu a analogia entre o problema da caminhada aleatória em uma dimensão com o jogo de cara e coroa; Em duas dimensões, com uma pessoa caminhando sem rumo nas ruas de uma cidade e, em três dimensões, com o processo de difusão de um gás na rede cristalina de um sólido (11).

A teoria clássica está concentrada em grafos simples, porém infinitos. Estuda o comportamento qualitativo da caminhada, onde as principais questões propostas são: pode-se afirmar que um caminhante aleatório retornará com certeza absoluta ao ponto de partida, isto é, a probabilidade de retorno é igual a 1? Se isto é verdade, ele retorna infinitas

Figura 4.1: Caminhada aleatória em um grafo.



Fonte: O autor.

vezes? G. Pólya (1921) provou que se a caminhada aleatória for realizada em uma rede d -dimensional, então o caminhante retornará com certeza ao ponto de início, e infinitas vezes, se $d = 1$ ou $d = 2$. Esta caminhada é denominada recorrente ou persistente. No entanto, se $d \geq 3$ o retorno acontecerá um número finito de vezes. Esta caminhada é dita transiente (31).

Uma caminhada aleatória em um grafo não direcionado pode ser vista como uma cadeia de Markov finita e reversível no tempo. Já quando a caminhada ocorre em um grafo regular, está relacionada com uma cadeia de Markov simétrica (32).

A teoria recente de caminhada aleatória concentra-se em grafos mais gerais, porém finitos, e estuda o seu comportamento de forma quantitativa (11). Nesta teoria algumas questões de interesse são: quanto tempo um caminhante leva para retornar ao ponto de partida? E qual é o tempo necessário para ele chegar a outro nó qualquer? Quão rápido a distribuição dos caminhantes tende a uma distribuição limite? Algumas destas questões serão tratadas a seguir.

4.2 Caminhada Aleatória em Redes Complexas

Considere uma rede finita, conectada e não direcionada com N nós e E arestas. As ligações entre os nós são descritas pela matriz adjacência A , cujos elementos A_{ij} são 1 ou 0, dependendo se os nós i e j forem conectados por uma aresta ou não.

A principal equação para a caminhada aleatória em tempo discreto em um grafo é dada por (14):

$$P_{t+1}(i) = \sum_j \frac{P_t(j)}{k_j}. \quad (4.1)$$

Esta representa a probabilidade do caminhante estar no instante $t + 1$ no nó i , em função da probabilidade de ele estar no instante anterior em algum nó j vizinho de i .

Pode-se escrever esta equação em termos da matriz adjacência do grafo, considerando que a probabilidade do caminhante saltar do nó i para o nó j é A_{ij}/K_i . Assim, a Eq. (4.1) para a probabilidade do n -ésimo salto de um caminhante, iniciando a partir de um nó i no tempo $t = 0$, até um tempo $t + 1$, é escrita como segue:

$$P_{ij}(t + 1) = \sum_k \frac{A_{kj}}{K_k} P_{ik}(t). \quad (4.2)$$

Considerando o vetor $\pi_t = (P_t(1), P_t(2), \dots, P_t(N))^t$ o qual fornece as probabilidades de um caminhante estar em cada um dos nós no instante t , pode-se escrever a Eq. (4.2) na forma matricial:

$$\pi_{t+1} = M\pi_t, \quad (4.3)$$

onde $M = DA$, A é a matriz adjacência do grafo e D é uma matriz diagonal com o i -ésimo elemento igual a $1/K_i$.

A configuração no instante $t = \tau$ pode ser obtida a partir do seu estado inicial:

$$\pi_\tau = M^\tau \pi_0. \quad (4.4)$$

Se o grafo for conectado, isto é, não existirem componentes isoladas, a solução da Eq. (4.4) será (14):

$$\pi_j = \frac{K_j}{2E}. \quad (4.5)$$

Este vetor é denominado vetor de probabilidade estacionário e corresponde ao autovetor de M com autovalor 1. A dedução do resultado apresentado na Eq. (4.5) está relacionada ao fato da matriz M ser assimétrica, a menos que o grafo G seja regular. Isto implica que os autovetores à esquerda e à direita não são iguais. Para contornar este problema, considera-se a matriz $P = D^{1/2}AD^{1/2} = D^{-1/2}MD^{1/2}$, onde P é simétrica e, portanto, com autovetores à esquerda e à direita iguais. Assim esta pode ser escrita na forma espectral (33):

$$P = \sum_{k=1}^N \lambda_k \mathbf{v}_k \mathbf{v}_k^T, \quad (4.6)$$

onde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ são os autovalores de P e $\mathbf{v}_1, \mathbf{v}_2, \dots$, são os correspondentes autovetores. O vetor $\mathbf{w}_i = \sqrt{K_i}$ é um autovetor de P com o autovalor 1, pois $P\mathbf{w}_i = 1\mathbf{w}_i$, e como ele é positivo, pelo Teorema de Frobenius-Perron temos que $\lambda_1 = 1 > \lambda_2 \geq \dots \geq \lambda_N \geq -1$ e que $\mathbf{v}_1 = \left(1/\sqrt{2E}\right) \mathbf{w} = \sqrt{K_i/2E} = \sqrt{\pi(i)}$, pois:

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|} = \frac{\mathbf{w}_i}{\sqrt{\sum_i^N (\sqrt{K_i})^2}} = \frac{\mathbf{w}_i}{\sqrt{\sum_i^N K_i}} = \frac{1}{\sqrt{2E}} \mathbf{w}_i, \quad (4.7)$$

Assim, a matriz M pode ser escrita da seguinte forma:

$$M^t = (D^{1/2}PD^{-1/2})^t = D^{1/2}P^tD^{-1/2} = \sum_{k=1}^N \lambda_k^t D^{1/2} \mathbf{v}_k \mathbf{v}_k^T D^{-1/2}. \quad (4.8)$$

Portanto,

$$M^t = \lambda_1^t D^{1/2} \mathbf{v}_1 \mathbf{v}_1^T D^{-1/2} + \sum_{k=2}^N \lambda_k^t D^{1/2} \mathbf{v}_k \mathbf{v}_k^T D^{-1/2} = Q + \sum_{k=2}^N \lambda_k^t D^{1/2} \mathbf{v}_k \mathbf{v}_k^T D^{-1/2}. \quad (4.9)$$

onde $Q_{ij} = \pi(j)$. Assim:

$$p_{ij}^t = \pi(j) + \sum_{k=2}^N \lambda_k^t v_{ki} v_{kj} \sqrt{\frac{K_j}{K_i}}. \quad (4.10)$$

Para um grafo G não bipartido tem-se que $\lambda_N > -1$, e então $|\lambda_k| < 1$ para $k = 2, 3, \dots, N$. Isto implica que (33):

$$p_{ij}^t \rightarrow \pi(j) = \frac{K_j}{2E} \quad (t \rightarrow \infty). \quad (4.11)$$

Outra maneira de verificar a Eq. (4.5) é escrevendo a expressão para a probabilidade de transição $P_{ij}(t)$, de forma iterativa a partir da equação (4.2), do seguinte modo (34):

$$P_{ij}(t) = \sum_{j_1, j_2, \dots, j_{t-1}} \frac{A_{ij_1}}{K_i} \cdot \frac{A_{j_1 j_2}}{K_{j_1}} \cdots \frac{A_{j_{t-1} j}}{K_{j_{t-1}}}, \quad (4.12)$$

Comparando as expressões para P_{ij} e P_{ji} nota-se que:

$$K_i P_{ij} = K_j P_{ji}, \quad (4.13)$$

Isto se deve à rede ser não direcionada. Para a solução estacionária, esta equação implica que $K_i P_j^\infty = K_j P_i^\infty$, e portanto tem-se que:

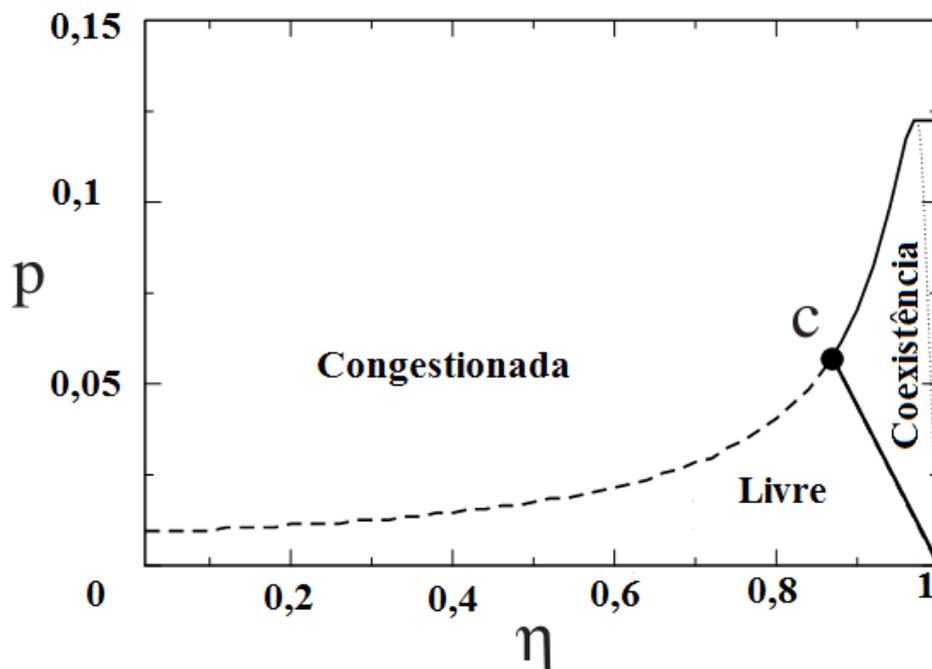
$$P_j^\infty = \lim_{n \rightarrow \infty} P_{ij}(n) = p_j = \frac{K_j}{2E}. \quad (4.14)$$

Logo, a distribuição estacionária é diretamente proporcional ao grau de um nó, e isso implica que quanto mais conectado é um determinado nó, mais caminhantes passarão por ele.

4.3 Estimativa de Probabilidades de EE em Redes Complexas por Meio de Caminhada Aleatória

Eventos que assumem valores extremos, ou seja, que se afastam da média, são classificados como EE. Outra forma de caracterização de EE é por meio de excedências de um valor limite alto. Em redes complexas, um tipo de EE que tem sido estudado é na forma de congestionamentos (10), (35), (12), (36). Estes trabalhos são baseados em caminhadas aleatórias nas redes. Em De Martino et. al (2009)(10), foi obtido um diagrama de fases comparando as fases congestionada e a de fluxo livre. Neste trabalho, existe uma taxa de criação de caminhantes p e uma probabilidade η de um nó recusar caminhantes, quando a sua capacidade de armazenamento foi atingida. É possível observar pela Fig. 4.2 que, dependendo dos valores destes parâmetros, tem-se uma região congestionada, uma de fluxo livre e outra de coexistência, onde pode-se encontrar uma fração de nós congestionados e os demais com fluxo livre coexistindo. A linha tracejada da Fig. 4.2 representa uma transição de fase de segunda ordem, e a linha contínua, uma de primeira ordem.

Figura 4.2: Diagrama de fases entre a fase congestionada e a de fluxo livre.



Fonte: Adaptado de De Martino et. al (2009)(10).

Outra linha de trabalhos referente à EE foi proposta em 2011 por Kishore et. al (6). Nestes, são considerados EE todos os valores que se afastem demasiadamente da média. O modelo de transporte adotado é o da caminhada aleatória. A diferença entre este modelo e os anteriores, é que o número total de caminhantes é fixo, ou seja, não há adição de novos caminhantes na rede. Também não há restrição quanto à capacidade de armazenamento para os nós. Neste caso, em um instante de tempo existiria uma probabilidade não nula de que todos os caminhantes estivessem em um único nó.

Kishore et. al (2011) (6) consideraram que em um nó de grau K ocorreu um EE, se neste passou uma elevada quantidade maior que q caminhantes em um instante de tempo, isto é, se houve uma flutuação muito alta do tráfego normal de caminhantes. O modelo proposto por estes autores é descrito a seguir:

Seja $f(w)$ a probabilidade de que um nó de grau K possua w caminhantes. Considerando que os caminhantes são independentes e não interagentes, a probabilidade de se encontrar w caminhantes em um dado nó é p^w , enquanto que o resto $W - w$ dos caminhantes são distribuídos em todos os outros nós. Este fato pode ser modelado por uma distribuição binomial, dada por:

$$f(w) = \binom{W}{w} p^w (1-p)^{W-w}. \quad (4.15)$$

Como a média e a variância para uma distribuição binomial são dadas por $\bar{x} = np$ e $S^2 = np(1-p)$, respectivamente (37), o número médio de caminhantes e a sua variância para um dado nó com grau K podem ser escritos como:

$$\langle f \rangle = \frac{WK}{2E}, \quad (4.16)$$

$$S^2 = \frac{WK}{2E} \left(1 - \frac{K}{2E}\right), \quad (4.17)$$

Também, este fato pode ser mostrado da seguinte maneira:

$$\langle f \rangle = \sum_{w=0}^W f(w)w = \sum_{w=0}^W \binom{W}{w} p^w (1-p)^{W-w} w. \quad (4.18)$$

Mas,

$$wp^w = p \frac{\partial}{\partial p} p^w. \quad (4.19)$$

Portanto,

$$\langle f \rangle = \sum_{w=0}^W f(w)w = \sum_{w=0}^W \binom{W}{w} p \frac{\partial}{\partial p} p^w (1-p)^{W-w}, \quad (4.20)$$

$$\langle f \rangle = p \frac{\partial}{\partial p} \sum_{w=0}^W \binom{W}{w} p^w (1-p)^{W-w} = p \frac{\partial}{\partial p} (p+1-p)^W = pW(p+1-p)^{W-1}, \quad (4.21)$$

Logo,

$$\langle f \rangle = Wp = W \frac{K}{2E}. \quad (4.22)$$

Para a variância:

$$S^2 = \overline{(w - \bar{w})^2} = \overline{w^2} - \bar{w}^2. \quad (4.23)$$

Tem-se que:

$$\bar{w}^2 = \langle f \rangle^2 = (Wp)^2, \quad (4.24)$$

e

$$\overline{w^2} = \sum_{w=0}^W f(w)w^2 = \sum_{w=0}^W \binom{W}{w} p^w (1-p)^{W-w} w^2, \quad (4.25)$$

Mas

$$w^2 p^w = p \frac{\partial}{\partial p} \left[p \frac{\partial}{\partial p} p^w \right], \quad (4.26)$$

Portanto,

$$\overline{w^2} = p \frac{\partial}{\partial p} \left[p \frac{\partial}{\partial p} \right] \sum_{w=0}^W \binom{W}{w} p^w (1-p)^{W-w} = p \frac{\partial}{\partial p} \left[p \frac{\partial}{\partial p} \right] [p+1-p]^W. \quad (4.27)$$

Assim,

$$\overline{w^2} = p \frac{\partial}{\partial p} \left\{ p \left[W(p+1-p)^{W-1} \right] \right\} = (pW)^2 + Wp(1-p) = \overline{w^2} + Wp(1-p). \quad (4.28)$$

Substituindo as Eqs. (4.24) e (4.28) na Eq. (4.23), tem-se:

$$S^2 = Wp(1-p) = W \frac{K}{2E} \left(1 - \frac{K}{2E} \right). \quad (4.29)$$

Uma extensão para a Eq. (4.15) pode ser generalizada para levar em conta as flutuações na quantidade de caminhantes. Suponha que o número total de caminhantes seja uma variável aleatória uniformemente distribuída no intervalo $[W - \Delta, W + \Delta]$. Então, a probabilidade de se encontrar w caminhantes se torna:

$$f^\Delta(w) = \sum_{j=0}^{2\Delta} \frac{1}{2\Delta+1} \binom{W'+j}{w} p^w (1-p)^{W'+j-w}. \quad (4.30)$$

onde $W' = W - \Delta$. A média e a variância para esta distribuição serão:

$$\langle f^\Delta \rangle = \langle f \rangle, \quad (4.31)$$

$$S_{\Delta}^2 = \langle f^{\Delta} \rangle \left[1 + \langle f^{\Delta} \rangle \left\{ \frac{\Delta^2}{3W^2} + \frac{\Delta}{3W^2} - \frac{1}{W} \right\} \right], \quad (4.32)$$

Com base na teoria de valores extremos, as probabilidades de ocorrência de um EE são muito pequenas, e estão tipicamente associadas com a cauda da distribuição de probabilidades (1),(2). Contabilizando todas as excedências em um intervalo de tempo suficientemente grande, pode-se inferir a probabilidade de EE, denotada por $F(K)$, em um nó com grau K :

$$F(K) = f^{\Delta}(k) + f^{\Delta}(k+1) + \dots + f^{\Delta}(W' + j), \quad (4.33)$$

onde $k = \lfloor q \rfloor + 1$ e $W' + j$ é o número total de caminhantes. Portanto, tem-se:

$$F(K) = \sum_{j=0}^{2\Delta} \frac{1}{2\Delta + 1} \sum_{k=\lfloor q \rfloor + 1}^{W' + j} \binom{W' + j}{k} p^k (1-p)^{W' + j - k}. \quad (4.34)$$

Kishore et. al (2011)(6) definiram também o valor limite q para cada nó dependendo do tráfego de caminhantes que passa por este nó da seguinte forma:

$$q = \langle f \rangle + cS, \quad (4.35)$$

onde c é um número real.

A Eq. (4.34) não foi escrita de uma forma fechada, porém, para o caso particular com $\Delta = 0$, ela pode ser reescrita da seguinte forma:

$$F(K) = \sum_{k=\lfloor q \rfloor + 1}^W f(k) = I_p(\lfloor q \rfloor + 1, w - \lfloor q \rfloor), \quad (4.36)$$

onde I_p é a função beta incompleta regularizada (38).

Kishore et. al (2011)(6) também investigaram os intervalos de retorno de EE. Isto é, tendo ocorrido um EE, daqui a quando tempo em média ocorrerá um novo EE. Para

determinar esta quantidade foi considerado que os eventos em cada nó não estavam correlacionados, pois os caminhantes não interagiram entre si. Foram definidas as seguintes probabilidades $P(t)$: probabilidade de um nó ficar um tempo t sem a ocorrência de EE e, $F(K)dt$, a probabilidade de que em um nó de grau K ocorra um EE entre t e $t + dt$. Assim, a probabilidade de um nó ficar um tempo $t + dt$ sem EE será dado por:

$$P(t + dt) = P(t) (1 - F(K)dt), \quad (4.37)$$

A solução geral desta equação é:

$$P(t) = C_0 e^{-F(K)t}. \quad (4.38)$$

onde C_0 é uma constante. Mas, considerando $P(0) = 1$, tem-se que $C_0 = 1$. Logo:

$$P(t) = e^{-F(K)t}. \quad (4.39)$$

Definindo $P(t)dt$ como a probabilidade de que um nó, depois de ficar um tempo t sem EE, sofra um EE no intervalo de tempo entre t e $t + dt$, vem:

$$P(t) = e^{-F(K)t} F(K)dt, \quad (4.40)$$

O tempo médio entre EE será dado por $\langle \tau \rangle = \int_0^\infty tP(t)dt$, assim :

$$\langle \tau \rangle = \int_0^\infty t e^{-F(K)t} F(K)dt \quad (4.41)$$

Cuja solução é:

$$\langle \tau \rangle = \frac{1}{F(K)}. \quad (4.42)$$

A distribuição de recorrência temporal será:

$$P(t) = e^{-t/\langle\tau\rangle}. \quad (4.43)$$

Portanto, na medida em que o tempo passa, a probabilidade de um determinado nó ficar sem sofrer algum EE diminui.

Capítulo 5

Resultados e Discussões

Para a construção de algumas redes complexas e análise de suas propriedades, foi utilizado o software R (17) e o pacote igraph (18). Somente as redes hierárquicas foram construídas conforme o algoritmo de Ravasz e Barabási (28).

5.1 Caminhada Aleatória Padrão

Para todas as simulações de caminhada aleatória, a posição inicial de cada caminhante foi escolhida aleatoriamente e um tempo transiente de 10^3 passos foi permitido para cada caminhante, para a estabilização da caminhada. Em seguida, observou-se por um intervalo de 10^6 passos a quantidade de caminhantes a cada instante em cada um dos nós para, determinar o número de caminhantes médio, $\langle f \rangle$, e o respectivo desvio padrão S .

5.1.1 Distribuição de Excedências

Foi implementada uma caminhada aleatória em redes complexas por $t = 3,65 \times 10^6$ passos, para a determinação da média e o desvio padrão do número de caminhantes em um nó com grau K . Assim, foi adotado para o número total de caminhantes $W = 2E$, onde E é o número total de arestas, e o valor limite, $q = \langle f \rangle + 4S$ foi adotado para a classificação de eventos como extremos, de forma análoga ao trabalho de Kishore et al. (2011) (6).

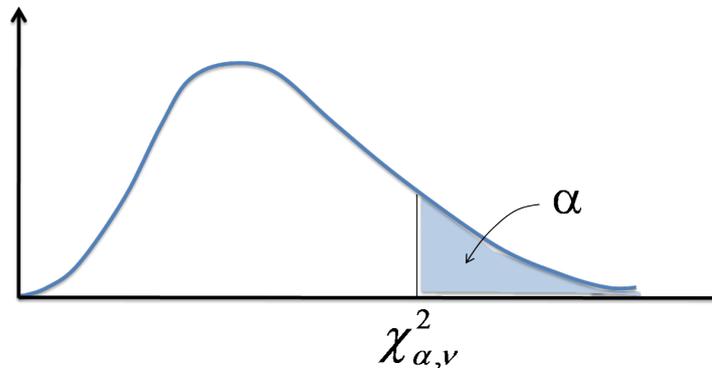
Foram investigadas redes com topologia aleatória, mundo pequeno e escala livre, onde foi selecionado em cada uma destas um subconjunto de nós com grau mínimo e outro com grau máximo. Em cada um desses subconjuntos foi escolhido um nó com centralidade de intermediação máxima e mínima. Foi observada a quantidade de caminhantes a cada instante somente neste nó escolhido, e toda vez que esta quantidade excedeu o limite especificado, este evento foi classificado como extremo. O valor da excedência também foi determinado. Este corresponde a diferença entre o valor observado e o valor limite.

Para a análise dos dados obtidos foi utilizado o pacote "extRemes" do software R (39), (40). Esta ferramenta permite estimar os parâmetros ξ e σ da distribuição de excedências (distribuição de Pareto generalizada), dada pela Eq.(2.15), bem como verificar se existe o regime exponencial dado pela Eq.(2.16). A verificação foi feita por meio do teste de hipóteses de χ^2 ao nível de significância de $\alpha = 5\%$, com o objetivo de analisar a qualidade do ajuste. Este teste requer uma amostra aleatória de tamanho n , obtida a partir de uma população com distribuição de probabilidade desconhecida. Estas n observações são dispostas em um histograma que possui k intervalos de classes. Seja O_i a frequência observada no i -ésimo intervalo e a partir da distribuição em teste, a frequência esperada neste intervalo foi calculada e denotada por E_i . Com isso, tem-se uma estatística de teste calculada do seguinte modo (41):

$$X_0^2 = \sum_1^k \frac{(O_i - E_i)^2}{E_i}. \quad (5.1)$$

Se a população segue a distribuição em teste, então X_0^2 segue uma distribuição de χ^2 com $\nu = k - p - 1$ graus de liberdade, onde p representa o número de parâmetros da distribuição em teste. Não se aceita a hipótese de que a distribuição dos dados segue a distribuição em teste se for obtido $X_0^2 > \chi_{\alpha, \nu}^2$, onde $\chi_{\alpha, \nu}^2$ é denominado valor crítico, que delimita as regiões de aceitação e não aceitação do teste, conforme pode ser visualizado na Fig.(5.1).

Figura 5.1: Teste de χ^2 para verificação da aderência de dados amostrais a uma dada distribuição. A área hachurada corresponde à região de não aceitação do teste.

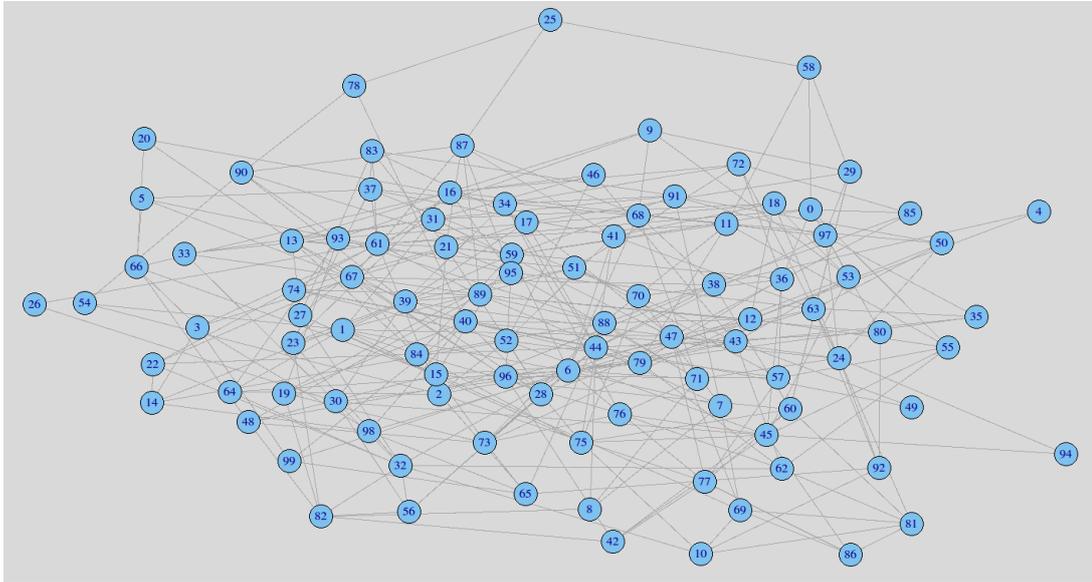


Fonte: Adaptado de Montgomery et. al (41).

Para o nível de significância e o grau de liberdade utilizados, o valor encontrado para $\chi_{\alpha, \nu}^2$ foi 3,84. A seguir são apresentadas ilustrações das redes que foram analisadas (Figs.5.2 a 5.4), bem como os resultados obtidos (Tabelas 5.1 a 5.3). Os valores das excedências, da estatística de teste X_0^2 , e dos parâmetros σ e ξ correspondem à média aritmética de três simulações.

A primeira rede analisada foi uma no modelo de Erdős e Renyi (7) com $N = 100$ nós, $E = 255$ arestas e probabilidade de ligação $p = 0.07$ (Fig 5.2). O número total de caminhantes utilizado foi $W = 510$, distribuídos de forma aleatória nesta rede. O número de excedências foi maior para os nós com grau menor (Tabela 5.1). Para os nós com mesmo grau porém com centralidade de intermediação b distintas, esta quantidade continuou superior para os nós com valores de b maiores. Para todos os casos, o valor da estatística de teste X_0^2 obtido foi sempre superior ao valor de $\chi^2 = 3,84$ e, portanto, como conclusão do teste, não se aceita a hipótese exponencial com $\xi = 0$. Como o valor de ξ foi negativo, isto implica que a distribuição de excedências possui um limitante superior dado por $q - \sigma/\xi$, conforme discutido no item (2.1.2.1). Como por exemplo, para os valores contidos na primeira linha da Tabela 5.1, temos que $q - \sigma/\xi = 11,4$, e assim, na respectiva série temporal, foi possível observar no máximo 11 caminhantes neste nó.

Figura 5.2: Rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós, $E = 255$ arestas e probabilidade de ligação $p = 0,07$.



Fonte: O autor

Tabela 5.1: Teste de hipóteses para distribuição exponencial com $\xi = 0$. Simulação com rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós, $E = 255$ arestas e probabilidade de ligação $p = 0,07$.

K	b	q	Excedências	X_0^2	σ	ξ	Conclusão
2	3,0	8	72	23,26	1,680	-0,490	N.A.
2	4,7	8	84	22,31	1,717	-0,440	N.A.
13	249,5	28	30	5,08	2,385	-0,434	N.A.
13	251,1	28	32	12,12	2,733	-0,745	N.A.

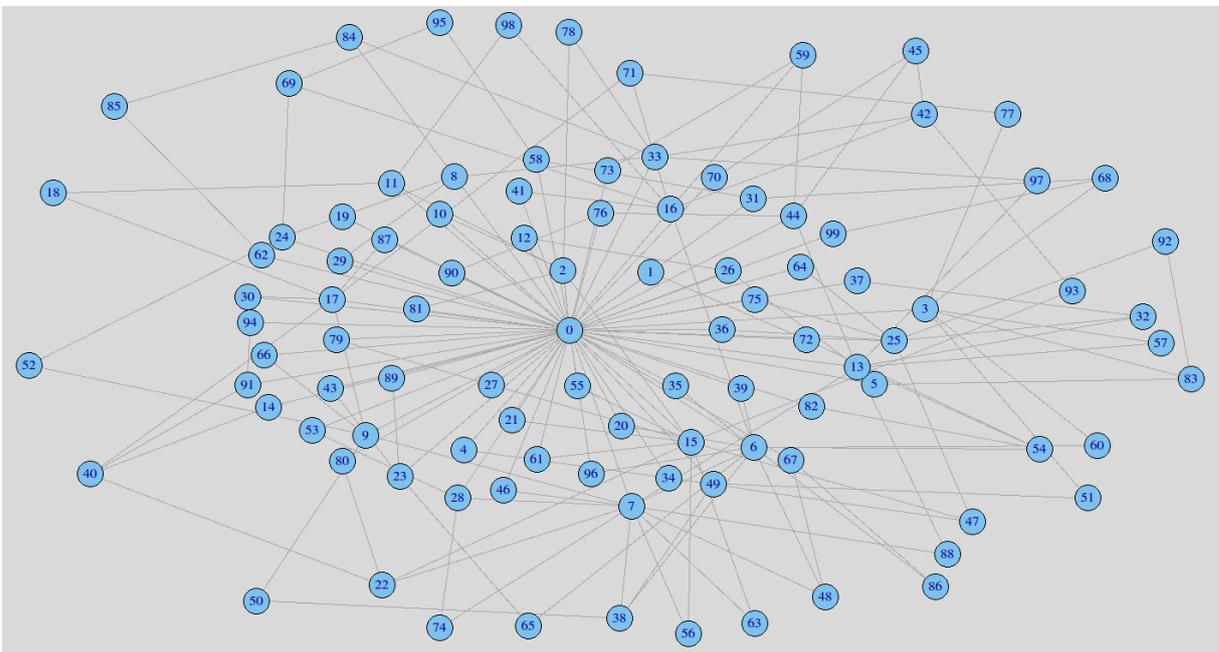
N.A.: não aceita a hipótese exponencial. Fonte: O autor

A segunda rede analisada foi uma de escala livre no modelo de Barabási e Albert (9) com $N = 100$ nós, $E = 189$ arestas (Fig. 5.3) e com $W = 378$ caminhantes. No processo de construção desta rede, cada novo nó adicionado a cada instante de tempo foi conectado a dois nós já presentes nesta rede, isto é, $m = 2$. Observou-se somente um nó com grau máximo ($K=64$) e uma grande quantidade de nós com grau mínimo. Esta é a característica principal deste tipo de rede. Assim, verifica-se a partir da Tabela 5.2 que o número de excedências foi novamente maior para os nós com grau menor.

Considerando os nós com menor grau, aquele com maior valor de b apresentou nova-

mente o número de excedências superior. O valor de X_0^2 obtido para estes nós foi superior ao valor de $\chi^2 = 3,84$ e, portanto, como conclusão do teste não se aceita a hipótese exponencial com $\xi = 0$. No entanto, para o nó com grau máximo, o valor de X_0^2 obtido foi menor que $\chi^2 = 3,84$. Assim, não se rejeita a hipótese exponencial e a distribuição não é limitada. Porém, foram somente 15 excedências para realizar o ajuste, pois o limite $q = 94$ é um valor que é muito maior que os demais limites.

Figura 5.3: Rede de escala livre no modelo de Barabási e Albert com $N = 100$ nós e $m = 2$.



Fonte:O autor

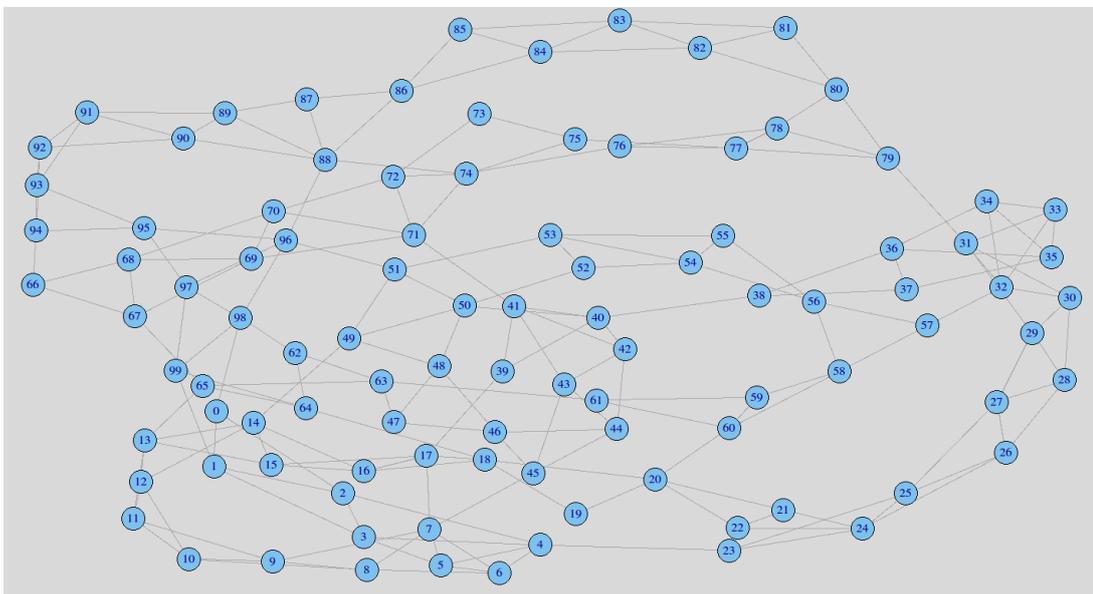
Tabela 5.2: Teste de hipóteses para distribuição exponencial com $\xi = 0$. Simulação com rede de escala livre no modelo de Barabási e Albert com $N = 100$ nós e $m = 2$.

K	b	q	Excedências	X_0^2	σ	ξ	Conclusão
2	0,0	8	88	24,68	1,765	-0,474	N.A.
2	89,0	8	93	26,45	1,662	-0,463	N.A.
64	4040,8	94	15	1,53	2,429	-0,347	N.R.

N.A.: não aceita a hipótese exponencial. N.R.: não rejeita a hipótese exponencial.

A terceira rede analisada foi uma com topologia mundo pequeno no modelo de Watts e Strogatz (8) com $N = 100$ nós, $E = 199$ nós e $W = 398$ caminhantes (Fig. 5.4). A probabilidade de religação empregada foi $p = 0,05$. Os resultados para esta rede são apresentados na Tabela 5.3. Os valores X_0^2 obtidos foram novamente superiores ao valor de $\chi^2 = 3,84$ para todos os casos. Analogamente, como para as demais redes, os nós com menor conectividade apresentaram uma quantidade superior de excedências.

Figura 5.4: Rede mundo pequeno no modelo de Watts e Strogatz com $N = 100$ nós, $E = 199$ arestas e probabilidade de religação $p = 0,05$.



Fonte: O autor

Tabela 5.3: Teste de hipóteses para distribuição exponencial com $\xi = 0$. Simulação com rede mundo pequeno no modelo de Watts e Strogatz com $N = 100$ nós, $E = 199$ arestas e probabilidade de religação $p = 0,05$.

K	b	q	Excedências	X_0^2	σ	ξ	Conclusão
2	0	8	85	21,48	1,685	-0,395	N.A.
2	7,2	8	80	25,91	1,692	-0,474	N.A.
6	401,4	16	50	15,14	2,315	-0,542	N.A.
6	905,5	16	56	17,67	1,911	-0,532	N.A.

N.A.: não aceita a hipótese exponencial. Fonte: O autor

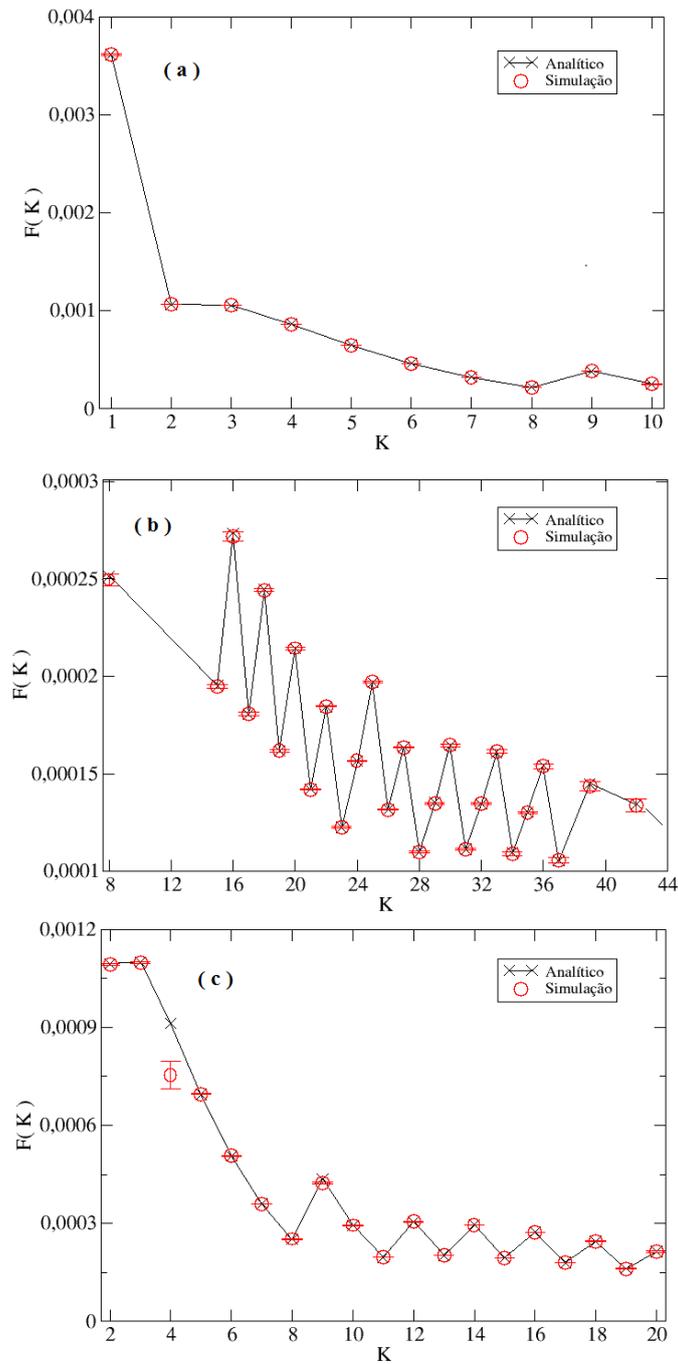
5.1.2 Probabilidade de Evento Extremo

Foi implementada uma caminhada aleatória padrão nas redes complexas analisadas para determinar a frequência de EE em nós com grau K . Para isto, inicialmente os caminhantes foram aleatoriamente distribuídos na rede e um tempo transiente de 10^3 passos foi permitido para cada um deles, para a estabilização da caminhada. Em seguida, foi observado a caminhada por 10^7 passos para a determinação da média e do desvio padrão do tráfego em cada nó, permitindo assim a determinação do limite para EE q da mesma maneira como em Kishore et al. (2011) (6). Com isso, novamente foi observada a caminhada por mais 10^7 passos contabilizando somente as excedências do limite q em cada nó e, ao final, foi calculada uma média para todos os nós com o mesmo grau. A probabilidade de EE em função do grau K , $F(K)$, foi obtida através da média aritmética entre dez simulações para cada rede.

Na Figura (5.5) são apresentados os resultados das simulações juntamente com os resultados analíticos para EE para redes com topologia aleatória no modelo de Erdős e Renyi com $N = 100$, $N = 500$ e $N = 1000$ nós. Neste caso, o número total de caminhantes adotado foi $W = 2E$ e com flutuação $\Delta = 0$. Foi observada uma boa concordância entre os resultados obtidos analiticamente e por meio das simulações.

A barra de erros indica que o afastamento dos valores obtidos para cada uma das dez simulações é muito pequeno. Este fato se deve ao espectro de autovalores das matrizes adjacências dessas redes. Para este tipo de rede o espectro não é simétrico, ou seja, o autovalor principal λ_1 se encontra afastado dos demais (30). Assim, o espectro da matriz M , definida na seção (4.2), apresenta os autovalores $\lambda_1 = 1 > \lambda_2 > \dots > \lambda_N > -1$. Portanto, esta rede não é bipartida, o que implica na convergência da probabilidade estacionária dada pela Eq. (4.5). Os resultados obtidos mostram que os nós com grau menor na rede são mais propensos a ter EE, quando comparados com os nós mais conectados.

Figura 5.5: Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede aleatória no modelo de Erdős e Renyi. Em (a) $N = 100$ nós, $E = 255$ arestas e $W = 510$ caminhantes, (b) $N = 500$ nós, $E = 4978$ arestas e $W = 9956$ caminhantes, (c) $N = 1000$ nós, $E = 6149$ arestas e $W = 12298$ caminhantes.



Fonte: O autor

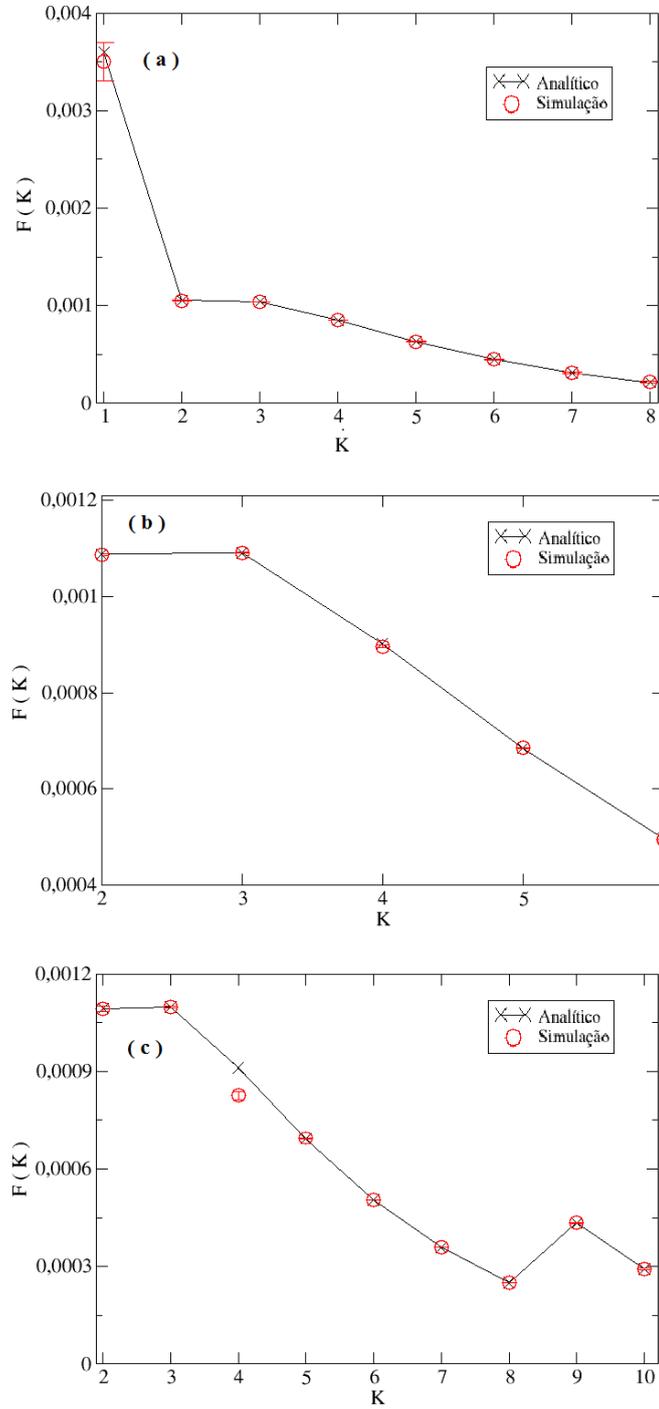
A distribuição de grau para este tipo de rede é Poissoniana, ou seja, existem poucos nós com grau pequeno e com grau alto. Este fato explica a diferença observada na figura (5.5 c) para os nós com grau $K = 4$, entre o resultado analítico e o das simulações. Nesta rede, o número de nós com grau $K = 4$ é pequeno, o que interfere no cálculo do tráfego médio de caminhantes nestes nós. Mas, com o aumento do tempo das simulações, há uma melhor concordância.

Resultados semelhantes foram obtidos considerando redes com topologia mundo pequeno, como no modelo de Watts e Strogatz e Newman e Watts, hierárquica no modelo de Ravász e Barabási e escala livre no modelo de Barabási e Albert. Para a construção deste última rede, cada novo nó adicionado a cada instante de tempo, foi conectado a dois nós já presentes na rede, isto é, $m = 2$. Assim, todas estas redes e incluindo a de escala livre também possuem os autovalores $\lambda_1 = 1 > \lambda_2 > \dots > \lambda_N > -1$, ou seja, o espectro de autovalores da matriz adjacência destas redes não é simétrico, o que indica que estas redes não são bipartidas.

No entanto, assim como para as redes aleatórias, as redes do tipo mundo pequeno apresentam uma diferença relativamente baixa na amplitude de grau, quando comparada com redes de escala livre e hierárquica. Para aquelas redes os valores de $F(K)$ entre os nós com grau menor e maior apresentam uma diferença pequena se comparados com os resultados para redes de escala livre e hierárquica. Estes resultados podem ser visualizados nas Figs. (5.6) a (5.9).

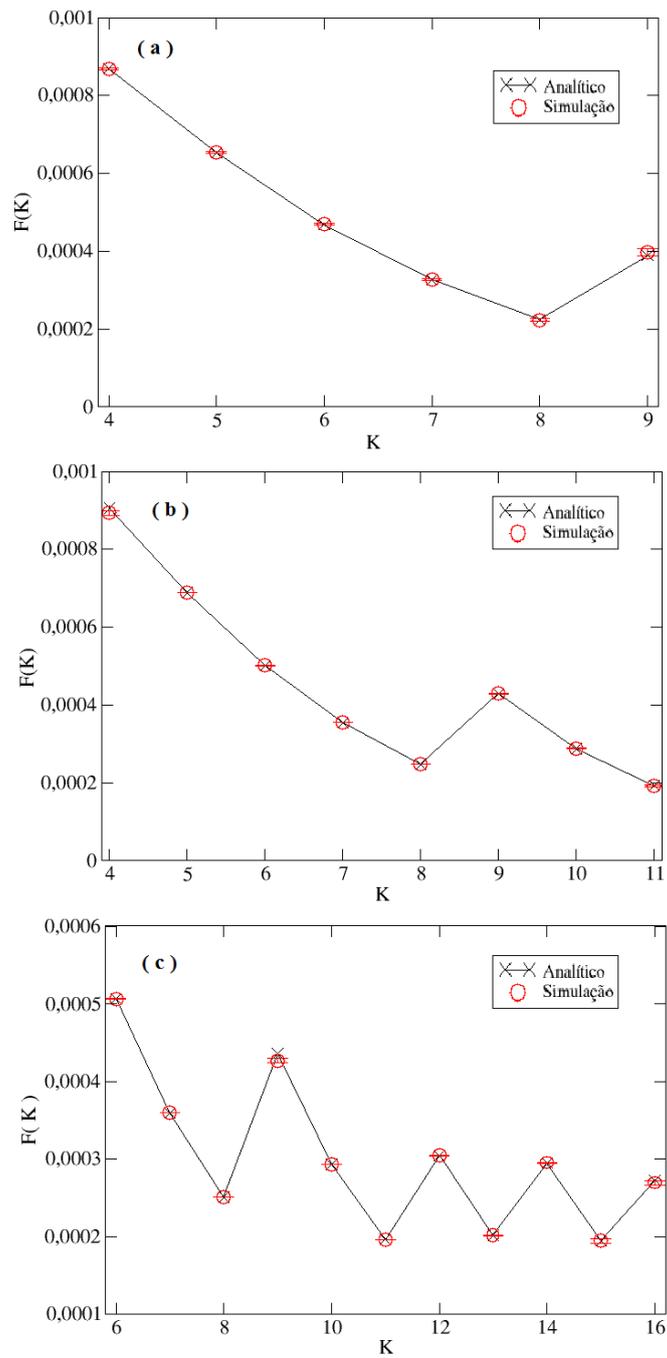
Para a Fig.(5.6a), referente a uma rede no modelo de Watts e Strogatz com 100 nós, há uma boa concordância entre o resultado analítico e a média entre as dez simulações. Porém, nota-se por meio da barra de erros, uma diferença entre os resultados das simulações para os nós com grau $K = 1$. Isto se deve ao fato de que existem poucos nós com este grau na rede, influenciando assim na determinação do tráfego médio de caminhantes. Novamente, aumentando o tempo das simulações, verifica-se a convergência dos resultados das simulações para o resultado analítico.

Figura 5.6: Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Watts e Strogatz. Em (a) $N = 100$ nós, $E = 199$ arestas e $W = 398$ caminhantes, (b) $N = 500$ nós, $E = 997$ arestas e $W = 1994$ caminhantes, (c) $N = 1000$ nós, $E = 2998$ arestas e $W = 5996$ caminhantes.



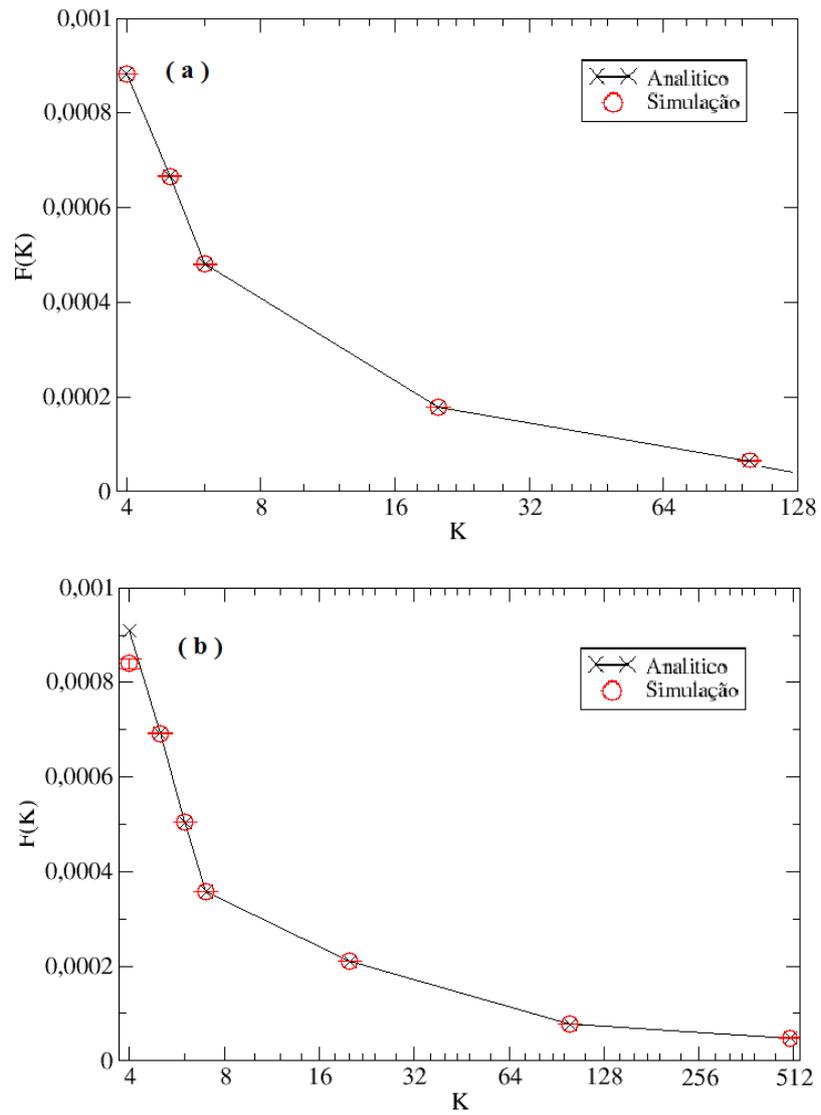
Fonte: O autor

Figura 5.7: Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Newman e Watts. Em (a) $N = 100$ nós, $E = 302$ arestas e $W = 604$ caminhantes, (b) $N = 500$ nós, $E = 1513$ arestas e $W = 3026$ caminhantes, (c) $N = 1000$ nós, $E = 4502$ arestas e $W = 9004$ caminhantes.



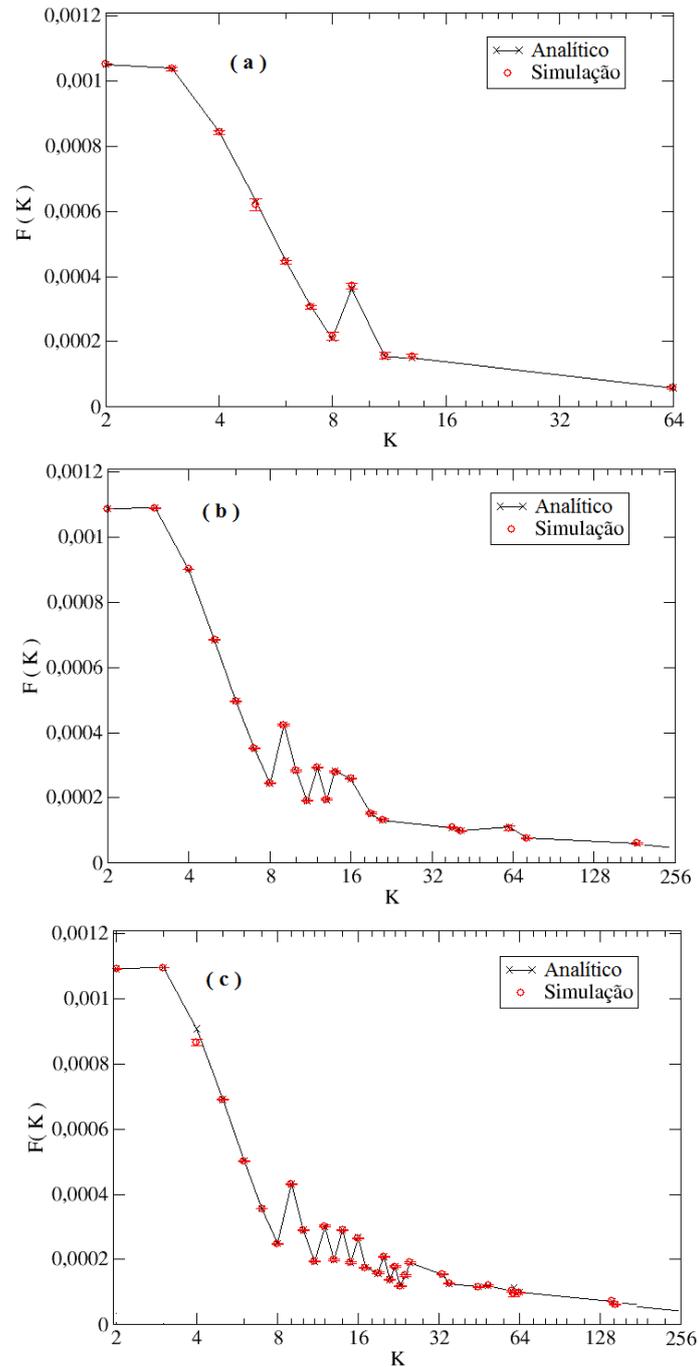
Fonte: O autor

Figura 5.8: Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede hierárquica no modelo de Ravász e Barabási. Em (a) $N = 125$ nós, $E = 410$ arestas e $W = 820$ caminhantes, (b) $N = 625$ nós, $E = 2450$ arestas e $W = 4900$ caminhantes.



Fonte: O autor

Figura 5.9: Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede de escala livre no modelo de Barabási e Albert com $m = 2$. Em (a) $N = 100$ nós, $E = 189$ arestas e $W = 378$ caminhantes, (b) $N = 500$ nós, $E = 969$ arestas e $W = 1938$ caminhantes, (c) $N = 1000$ nós, $E = 1979$ arestas e $W = 3958$ caminhantes.



Fonte: O autor

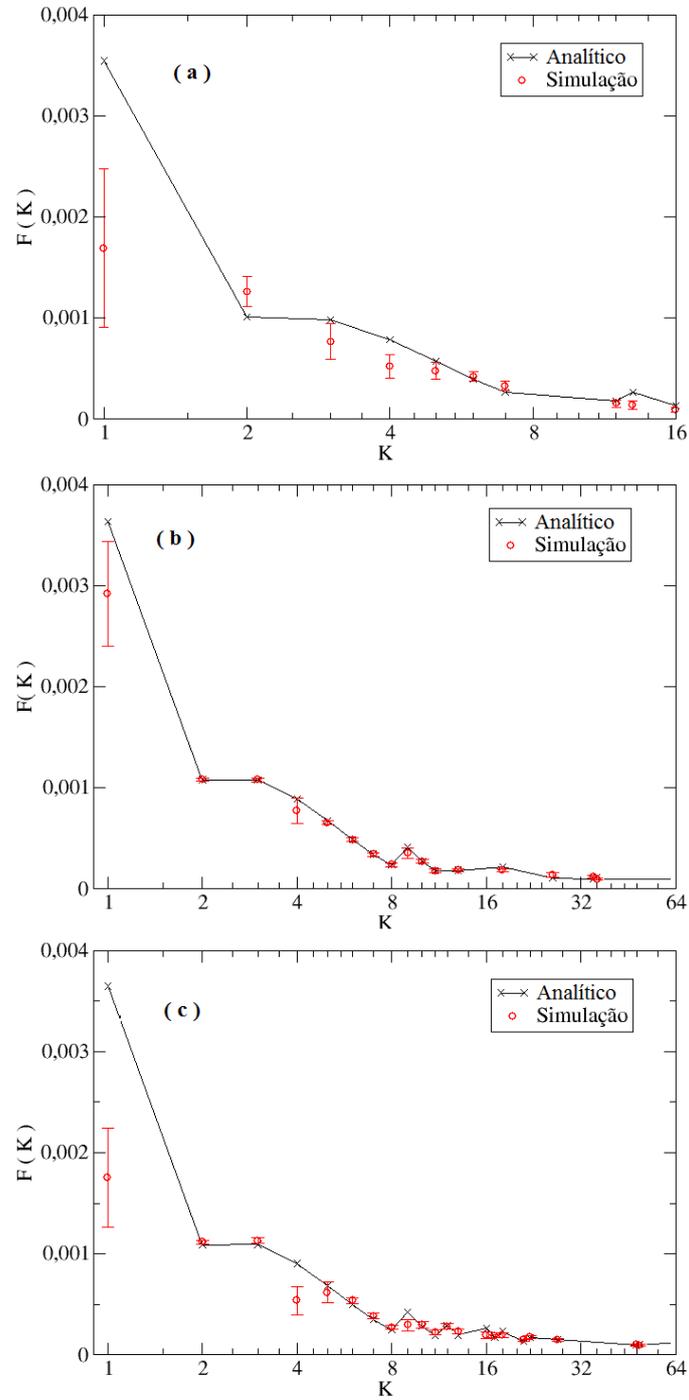
Na Fig.(5.10) é apresentado o resultado analítico e das simulações para redes de escala livre no modelo de Barabási e Albert com $N = 100$, $N = 500$ e $N = 1000$ nós. Porém, no processo de construção destas redes, cada novo nó adicionado a cada instante de tempo, foi conectado a apenas um nó já presente nesta rede, isto é, $m = 1$. Assim, estas redes são por definição do tipo árvores e possuem espectro de autovalores simétrico (7).

Ainda, a partir da Fig. (5.10) observa-se que o resultado principal permanece válido: os nós na rede com grau K menor possuem uma probabilidade de EE maior quando comparados com os nós mais conectados, mesmo considerando uma rede que apresente os autovalores $\lambda_1 = 1 > \lambda_2 > \dots > \lambda_N \geq -1$ na sua matriz M . No entanto, existe uma divergência entre o resultado das simulações e o analítico (Fig. 5.10). Neste caso, como há também a presença do autovalor -1 no espectro de M , esta rede é bipartida (33).

Os nós de uma rede bipartida podem ser separados em dois subconjuntos, onde não existem ligações entre nós no mesmo conjunto, e sim, somente entre nós de conjuntos distintos. Assim, um caminhante aleatório que se encontra em um determinado conjunto no instante t , no instante posterior $t + 1$ ele necessariamente estará no outro conjunto, e assim ficará alternando entre esses dois conjuntos (14). Já para os nós maior grau na rede, a diferença entre o resultado analítico e as simulações diminui. Nestes nós, o tráfego médio de caminhantes é muito superior do que nos demais nós. Por isso, os resultados das simulações aproximam-se do resultado analítico, mesmo considerando que, para estas redes a quantidade nós com grau alto é pequena.

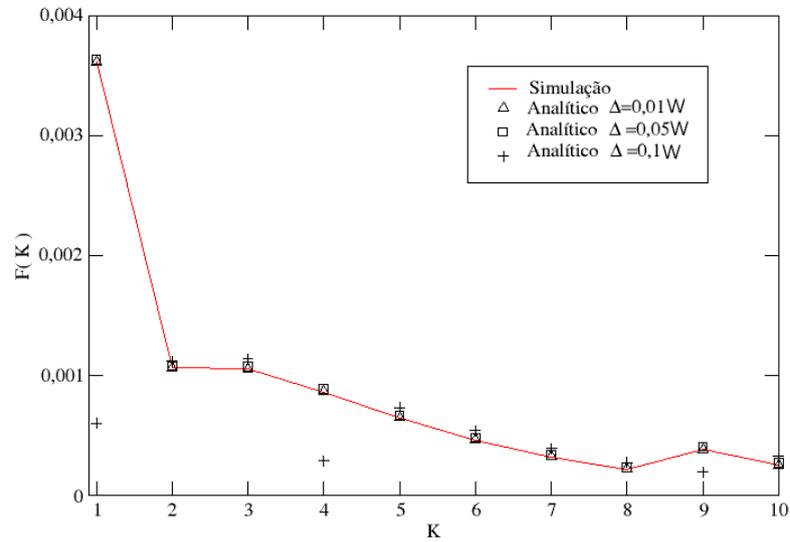
Foi implementada também a probabilidade de evento extremo dada pela Eq. (4.34) para algumas redes complexas com $\Delta \neq 0$, para verificar se este resultado é robusto mesmo quando existe uma flutuação no número total de caminhantes. Novamente, mesmo com uma flutuação no número de caminhantes, os nós com grau menor são mais propensos a ter EE. Nota-se também que para Δ pequeno há uma boa concordância entre a simulação de caminhada aleatória e o resultado analítico. Porém, à medida que esta flutuação aumenta, os resultados divergem (Figs. 5.11 a 5.14).

Figura 5.10: Probabilidade de EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede de escala livre no modelo de Barabási e Albert com $m = 1$. Em (a) $N = 100$ nós, $E = 99$ arestas e $W = 198$ caminhantes, (b) $N = 500$ nós, $E = 499$ arestas e $W = 998$ caminhantes, (c) $N = 1000$ nós, $E = 999$ arestas e $W = 1998$ caminhantes.



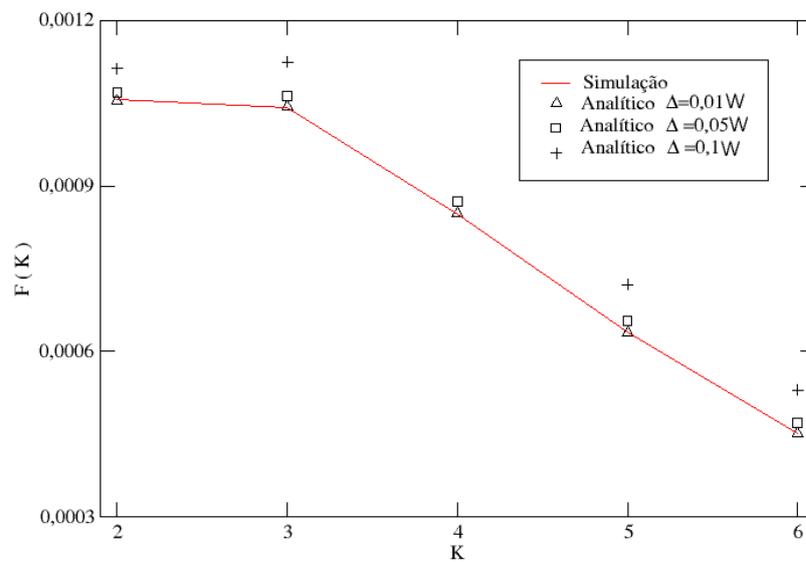
Fonte: O autor

Figura 5.11: Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós.



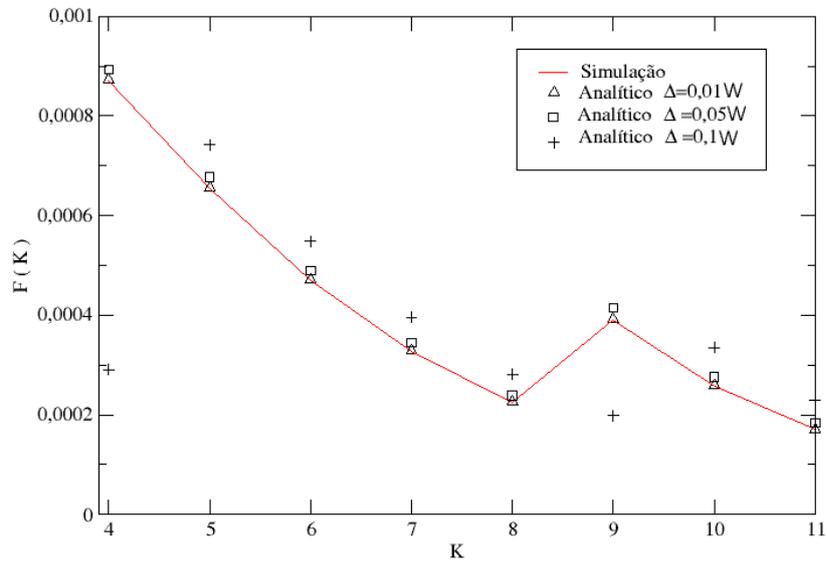
Fonte: O autor

Figura 5.12: Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede mundo pequeno no modelo de Watts e Strogatz com $N = 100$ nós.



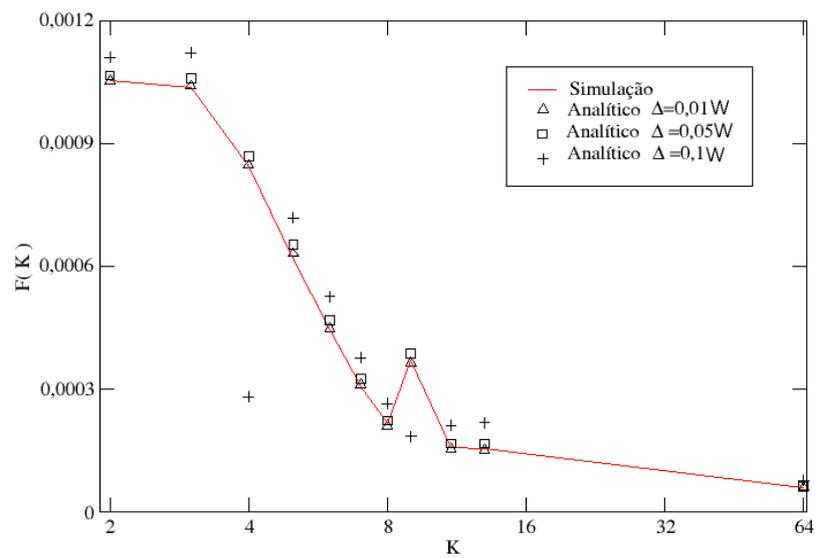
Fonte: O autor

Figura 5.13: Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede mundo pequeno no modelo de Newman e Watts com $N = 100$ nós.



Fonte: O autor

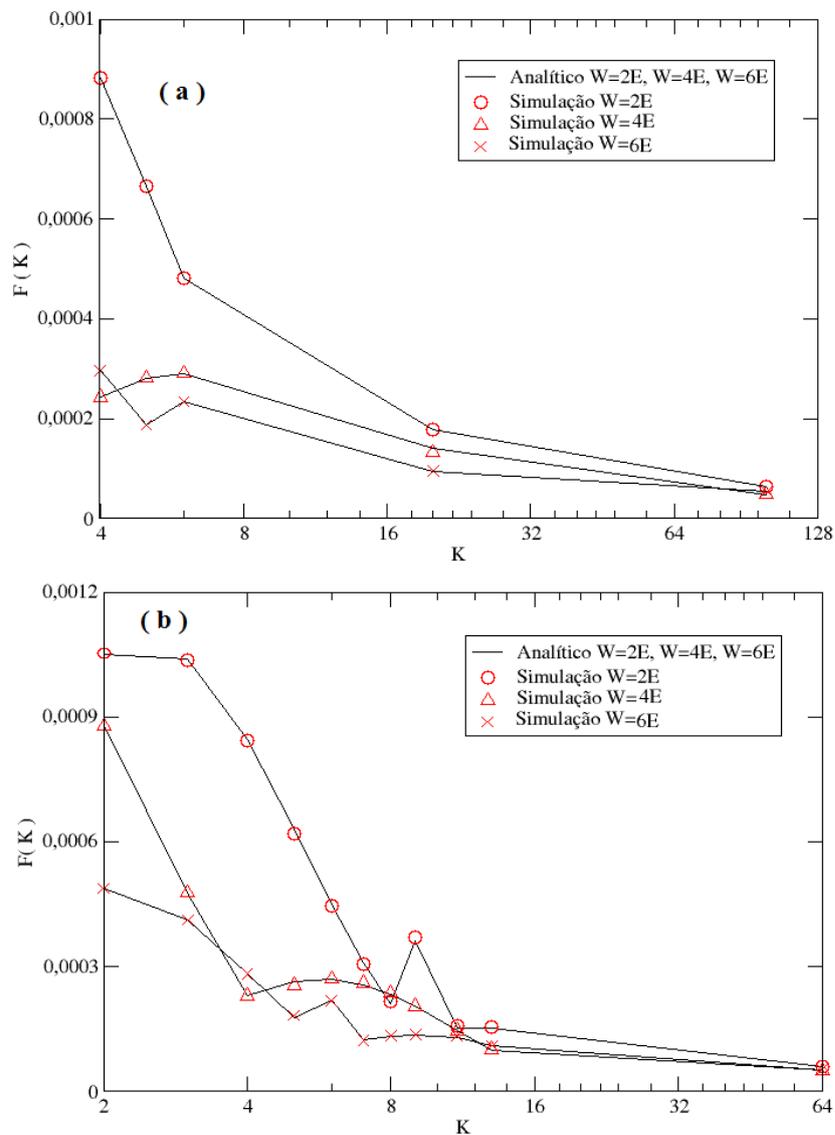
Figura 5.14: Probabilidade de EE em função do grau K com $\Delta \neq 0$ e $W = 2E$ para rede de escala livre não bipartida no modelo de Barabási e Albert com $N = 100$ nós.



Fonte: O autor

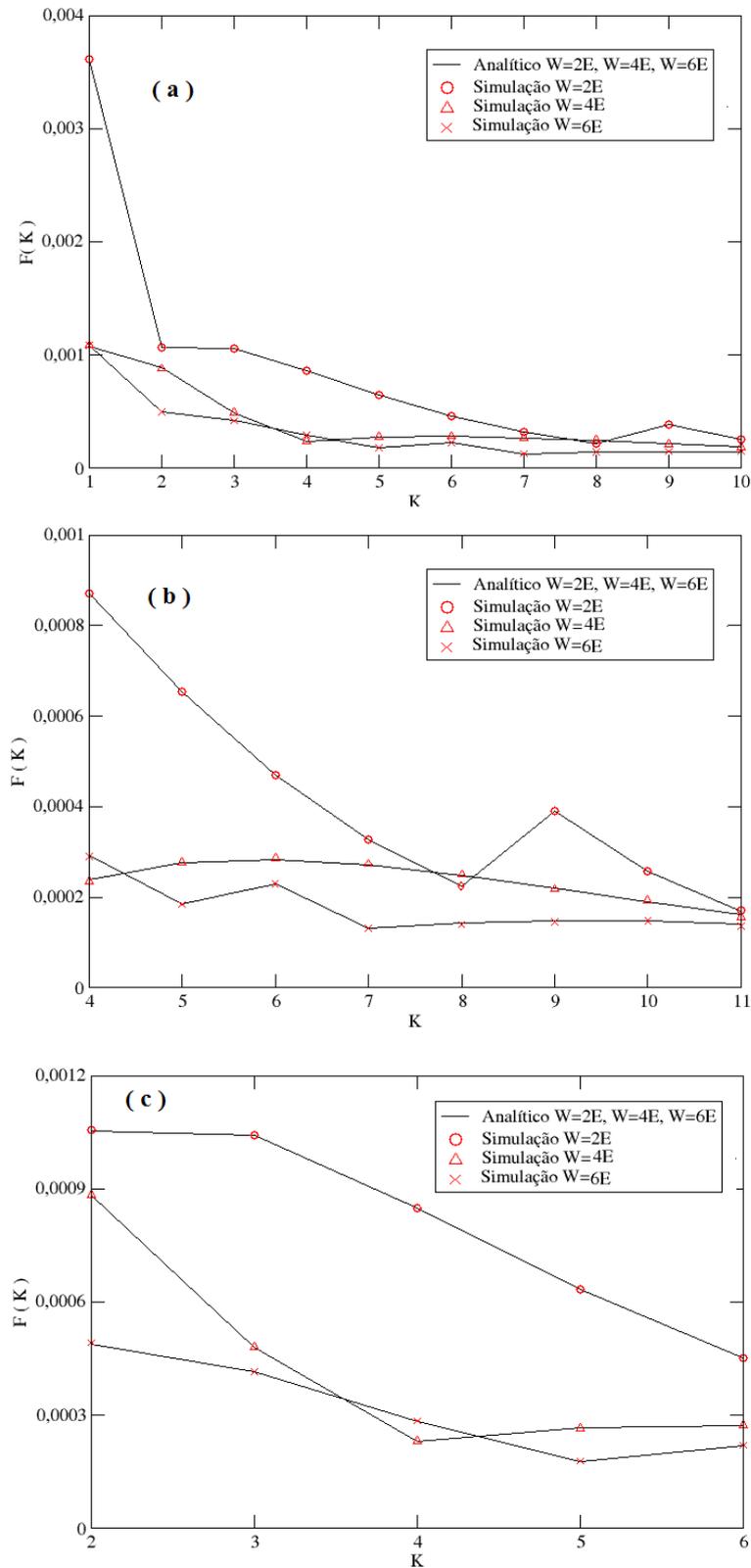
Também foram realizadas simulações variando o número total de caminhantes para verificar o comportamento das probabilidades de EE $F(K)$. Os números de caminhantes utilizados, além de $W = 2E$, foram com $W = 4E$ e $W = 6E$. Notou-se que ao aumentar a quantidade de caminhantes, o valor de $F(K)$ diminuiu, principalmente para os nós com grau menor na rede, (Figs. 5.15 e 5.16).

Figura 5.15: Probabilidade de EE em função do grau K com $\Delta = 0$ para algumas redes complexas, $W = 2E$, $W = 4E$ e $W = 6E$. (a) Rede hierárquica no modelo de Ravász e Barabási com $N = 125$ nós. (b) Rede de escala livre não bipartida no modelo de Barabási e Albert com $N = 100$ nós.



Fonte: O autor

Figura 5.16: Probabilidade de EE em função do grau K com $\Delta = 0$ para algumas redes complexas com $N = 100$ nós, $W = 2E$, $W = 4E$ e $W = 6E$. (a) Rede aleatória no modelo de Erdős e Renyi, (b) Rede mundo pequeno no modelo de Newman e Watts, (c) Rede mundo pequeno no modelo de Watts e Strogatz.

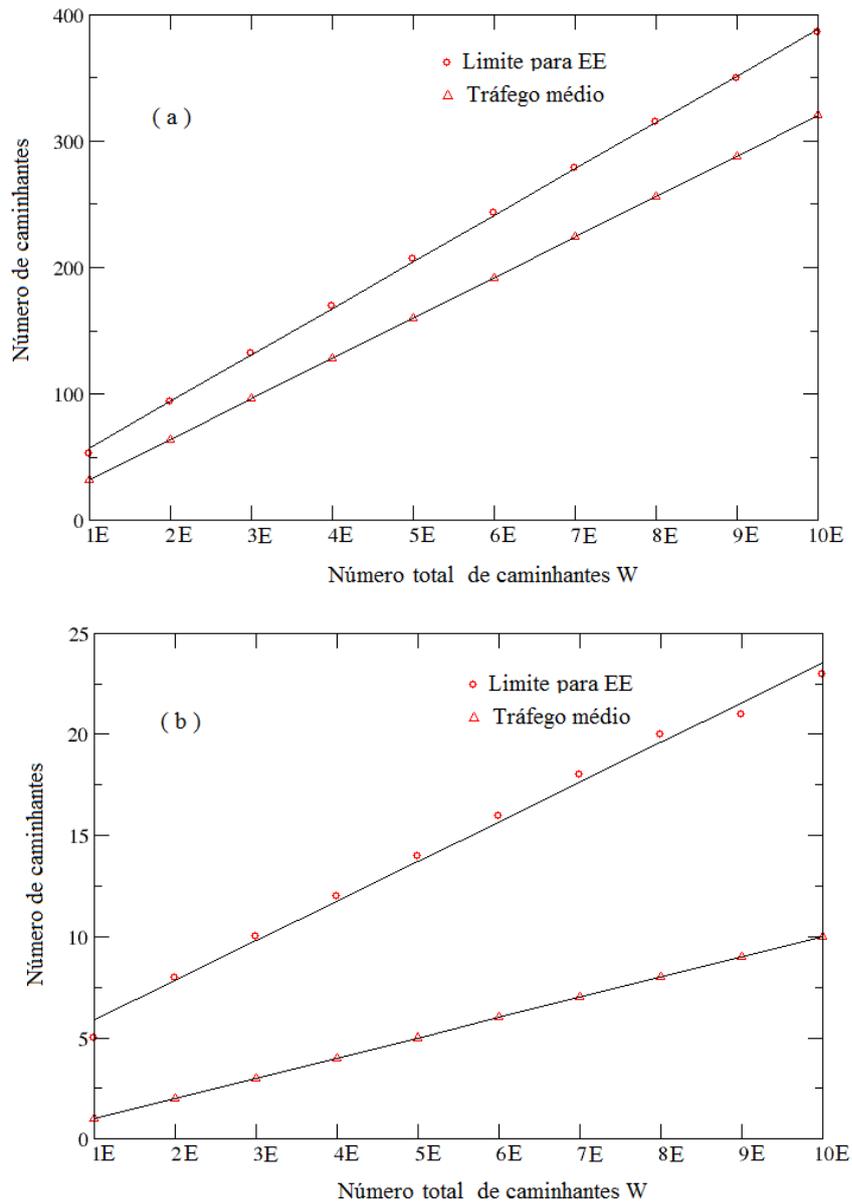


Fonte: O autor

Nas Figs. (5.17a) e (5.17b), são comparados o crescimento do valor limite e do tráfego médio para nós com grau máximo e mínimo, à medida que se aumenta a quantidade total de caminhantes W na rede. Para a Fig. (5.17a), os coeficientes angulares das retas de ajuste foram 36,77 e 32 para o valor limite e tráfego médio, respectivamente. Ao passo que, para na Fig. (5.17b), estes foram 1,96 e 1. Para ambos os casos o valor limite q aumenta mais rapidamente do que o tráfego médio, quando a quantidade total de caminhantes na rede é aumentada, pois seus coeficientes angulares são maiores.

Ao comparar as Figs. (5.17a) e (5.17b), observa-se que o aumento do valor limite q ocorre de forma mais acentuada na Fig.(5.17b). Isto foi verificado comparando os coeficientes angulares das retas de ajuste para cada caso. Para a Fig. (5.17a) a diferença percentual foi de 12,9% em relação ao coeficiente angular da reta de ajuste do valor limite q . Enquanto que para na Fig. (5.17b) a diferença percentual de 48,9%. Portanto, para os nós com grau menor na rede, à medida que W aumenta, o valor limite cresce mais rapidamente do que o tráfego médio. Assim, alguns eventos que seriam anteriormente classificados como extremos, não ultrapassam o limite, reduzindo a frequência de eventos extremos.

Figura 5.17: Comparação entre o número de caminhantes para o tráfego médio e o limite q de EE para uma rede com $N = 100$ nós. (a) Nós com grau máximo e (b) Nós com grau mínimo.



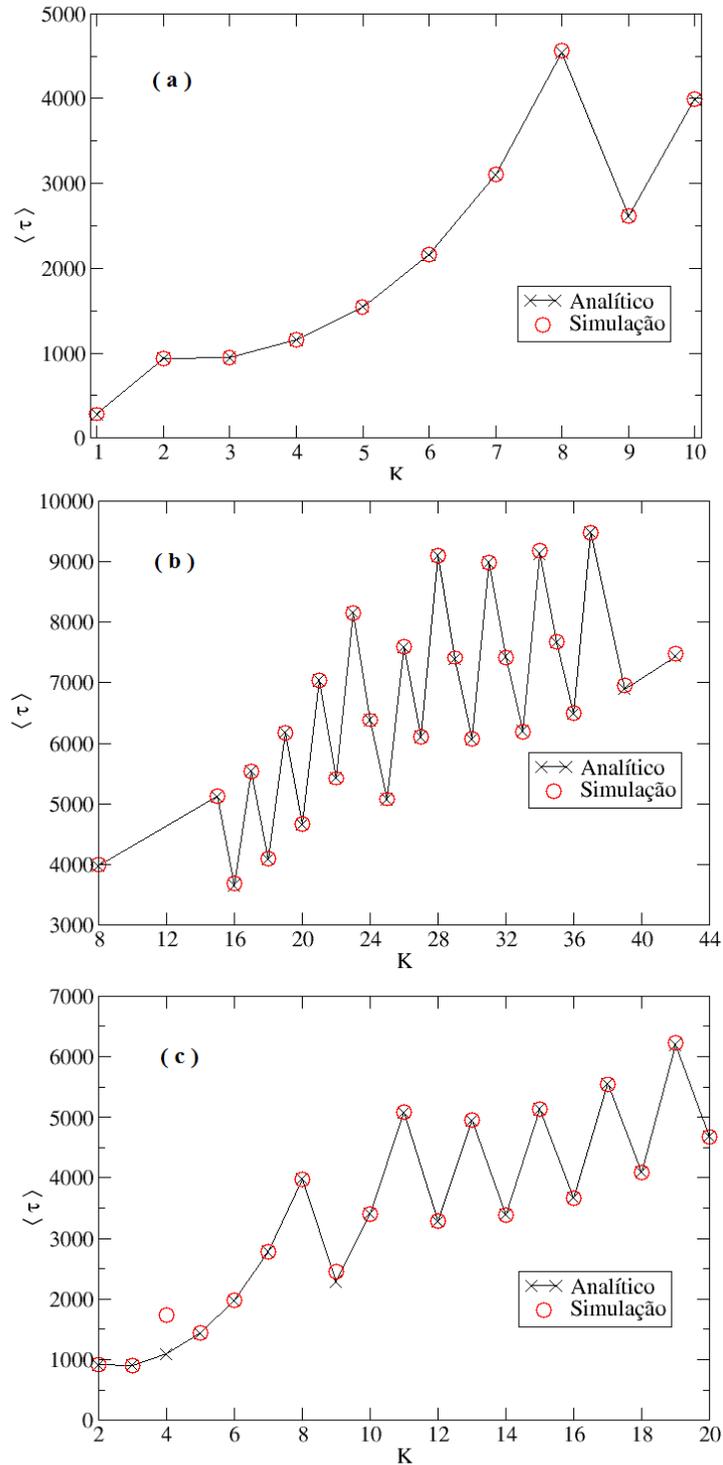
Fonte: O autor

5.1.3 Tempos Médios entre Eventos Extremos

Para as simulações de tempo médio entre EE, foram determinados os EE da mesma forma como descrito anteriormente (item 5.1.2), e também o tempo entre um EE e o posterior. Com isso, foi feita uma média entre esses intervalos para cada nó e em seguida foi determinada a média aritmética para todos os nós com o mesmo grau. Novamente, foi considerada a média entre 10 simulações e esta foi comparada com o resultado analítico dado pela Eq. (4.42).

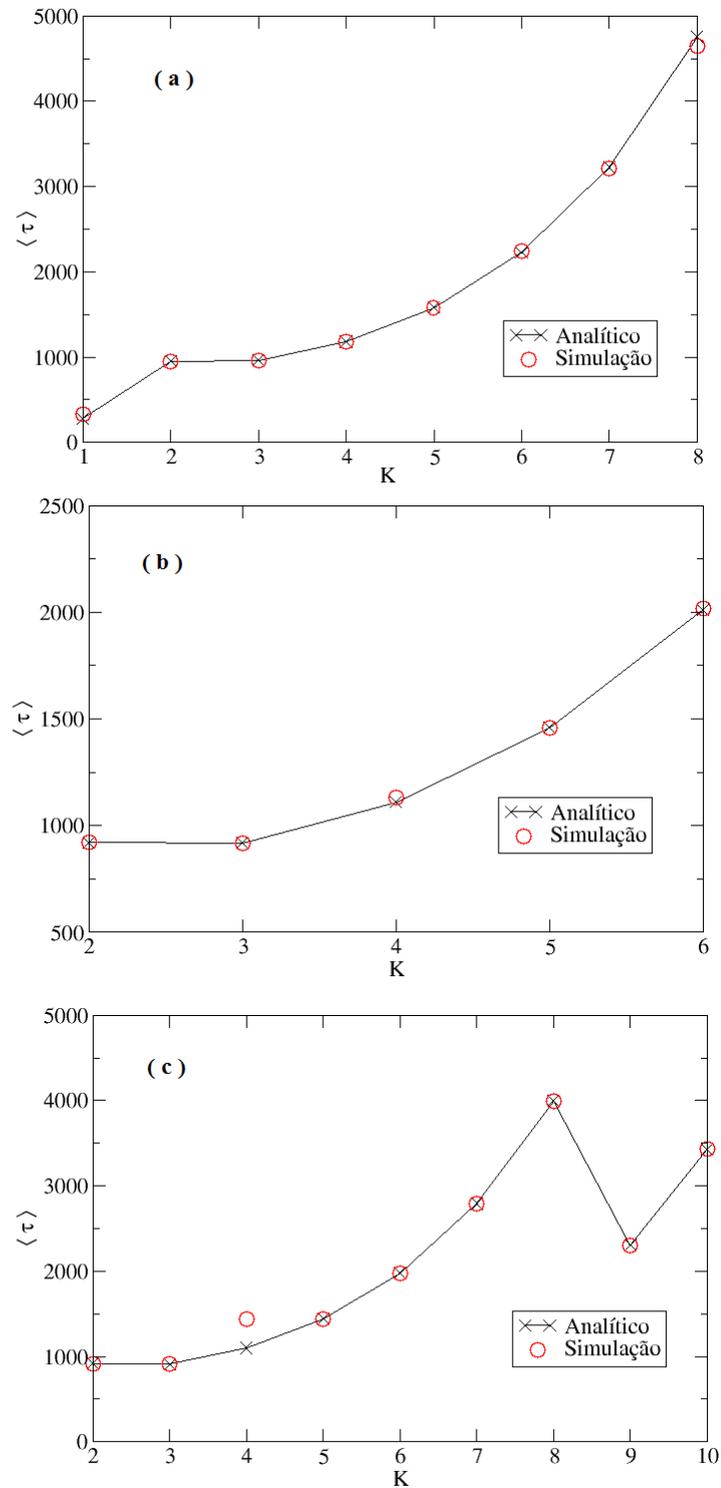
Nas Figs.(5.18) a (5.23) são apresentados os resultados obtidos das simulações e a comparação com os resultados analíticos para os tempos médios entre EE. Novamente, observou-se uma boa concordância entre estes resultados, e somente ocorreram diferenças no caso das redes bipartidas. De uma forma geral os nós com grau baixo apresentam tempos médio de retorno menor quando comparados com os nós mais conectados na rede. Assim, os nós com grau menor na rede são mais propensos a ter EE e também a ter EE sucessivos em um curto intervalo de tempo. Também, foi verificado que, para as redes com topologia aleatória e mundo pequeno, os tempos de retorno são no máximo em torno de 10000 passos para os nós com maior conectividade. Porém, para as redes hierárquicas e livres de escala, estes chegam a alcançar valores próximos a 25000, pois o grau máximo para estas redes é muito superior.

Figura 5.18: Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede aleatória no modelo de Erdős e Renyi. Em (a) com $N = 100$ nós e $W = 255$ caminhantes. (b) $N = 500$ nós e $W = 9956$ caminhantes. (c) $N = 1000$ nós e $W = 12298$ caminhantes.



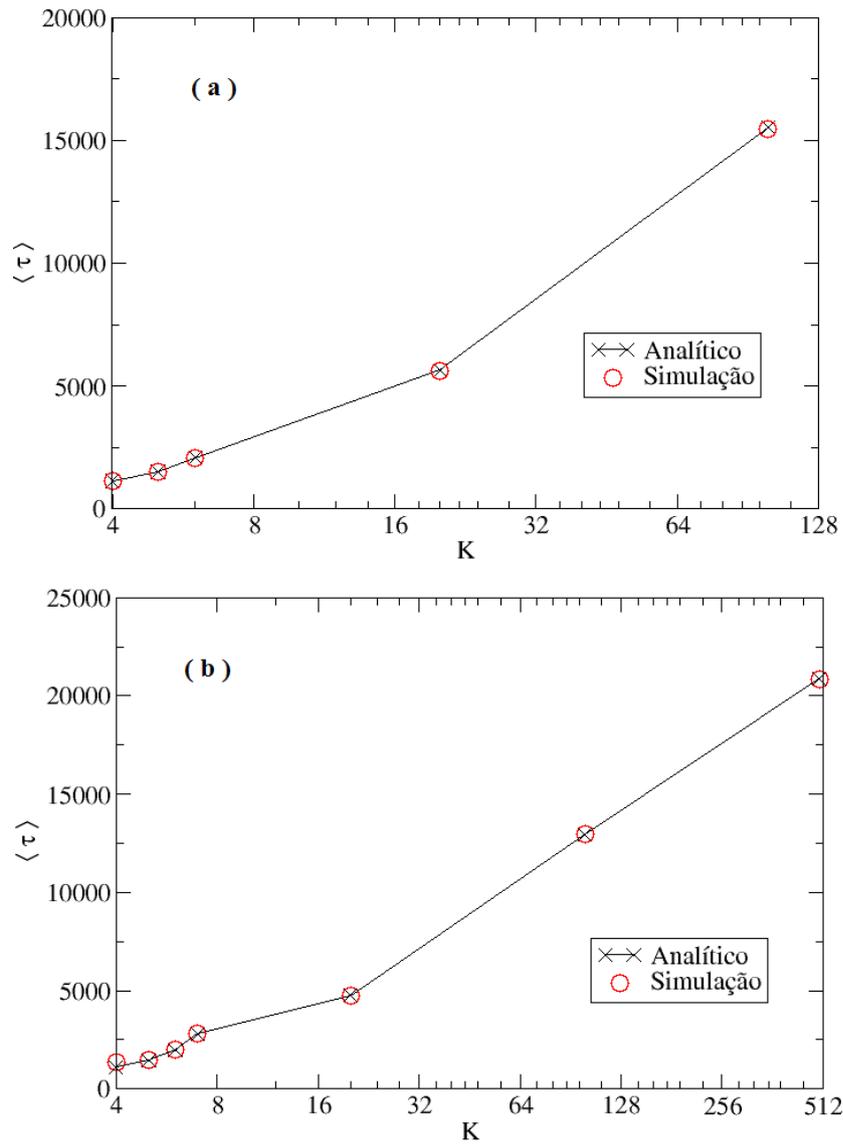
Fonte: O autor

Figura 5.19: Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Watts e Strogatz. Em (a) com $N = 100$ nós e $W = 398$ caminhantes. (b) com $N = 500$ nós e $W = 1994$ caminhantes. (c) com $N = 1000$ nós e $W = 5996$ caminhantes.



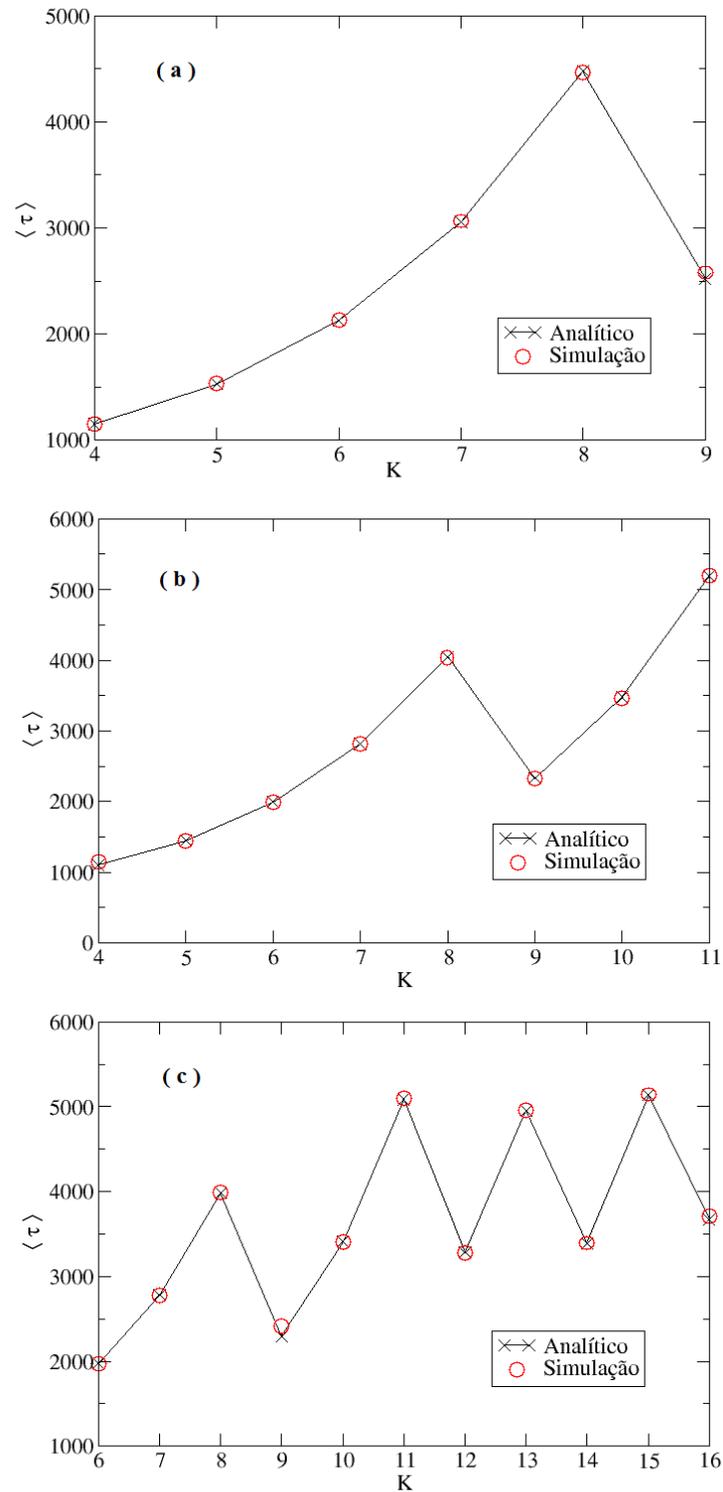
Fonte: O autor

Figura 5.20: Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede hierárquica no modelo de Ravász e Barabási. (a) $N = 125$ nós e $W = 820$ caminhantes (b) $N = 625$ nós e $W = 4900$ caminhantes.



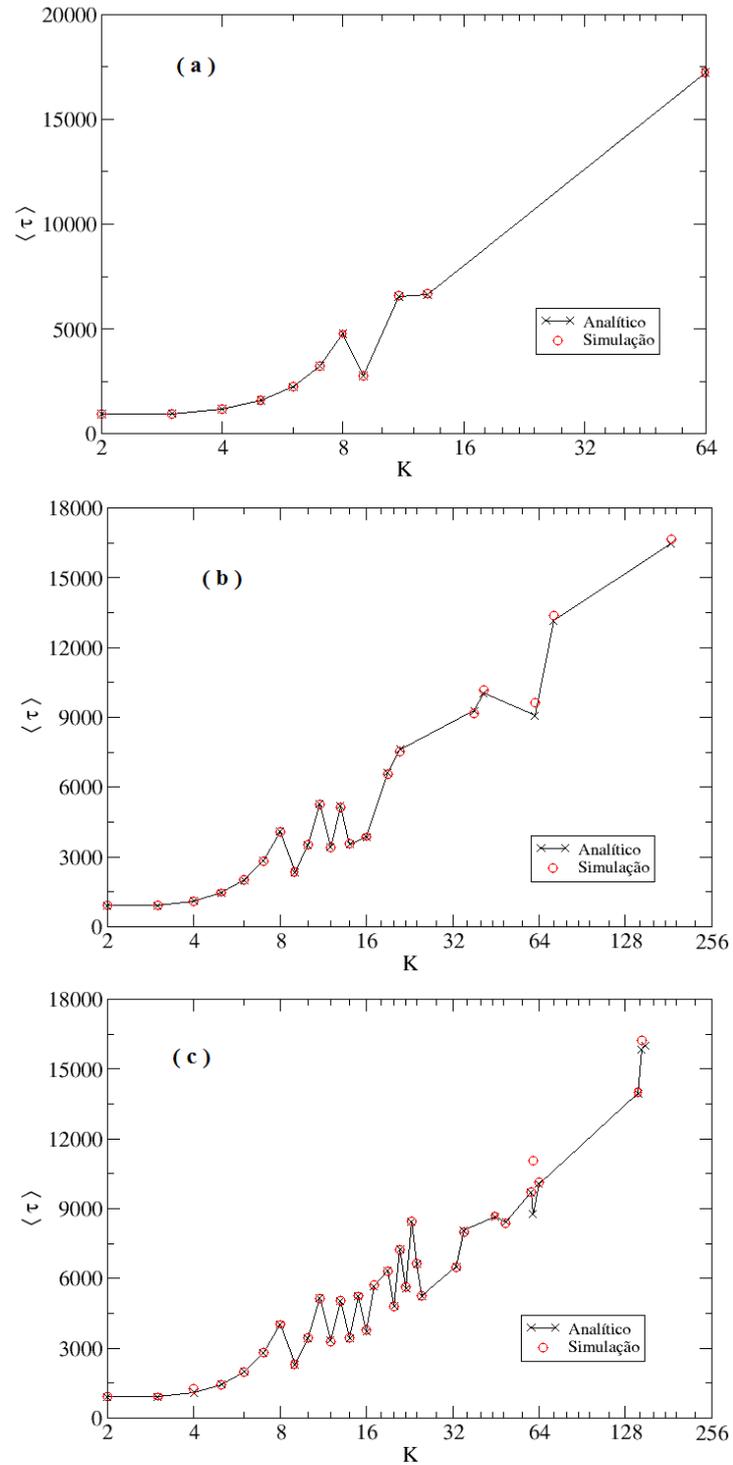
Fonte: O autor

Figura 5.21: Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para rede mundo pequeno no modelo de Newman e Watts. (a) $N = 100$ nós e $W = 604$ caminhantes (b) $N = 500$ nós e $W = 3026$ caminhantes (c) $N = 1000$ nós e $W = 9004$ caminhantes



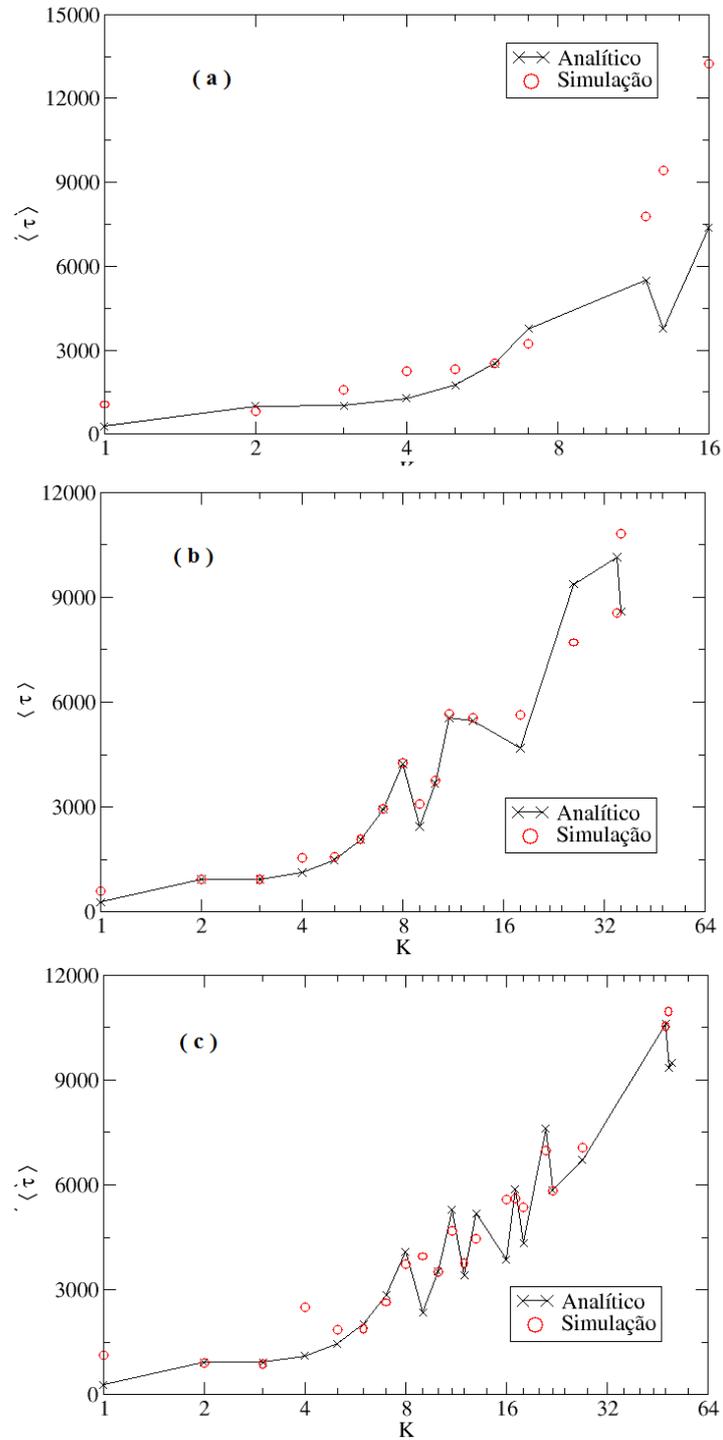
Fonte: O autor

Figura 5.22: Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para Rede de escala livre no modelo de Barabási e Albert não bipartida. (a) $N = 100$ nós e $W = 378$ caminhantes (b) $N = 500$ nós e $W = 1938$ caminhantes (c), $N = 1000$ nós e $W = 3958$ caminhantes.



Fonte: O autor

Figura 5.23: Tempo médio entre EE em função do grau K com $\Delta = 0$ e $W = 2E$ para Rede de escala livre no modelo de Barabási e Albert bipartida. (a) $N = 100$ nós e $W = 198$ caminhantes (b) $N = 500$ nós e $W = 998$ caminhantes (c), $N = 1000$ nós e $W = 1998$ caminhantes.



Fonte: O autor

5.2 Caminhada Aleatória por Menores Caminhos

Para a implementação de uma caminhada aleatória com a restrição de que esta fosse realizada apenas pelos menores caminhos na rede complexa, foram utilizados algoritmos de caminho mínimo em grafos (42). O algoritmo escolhido para este trabalho de dissertação foi o de Floyd-Warshall (20), (43), o qual é matricial e trabalha unindo todos os pares de vértices. Este utiliza um vértice base k a partir do qual são obtidos os menores caminhos por meio de desigualdades triangulares. Da forma como o processo é conduzido, se os vértices forem rotulados de 1 a n , o índice do vértice base usado em uma iteração corresponderá ao valor do contador de iterações, e as desigualdades serão da forma:

$$d_{ij}^k = \min(d_{ij}^{k-1}, (d_{ik}^{k-1} + d_{kj}^{k-1})). \quad (5.2)$$

As modificações nos valores são inscritas na própria matriz vigente D^{k-1} , que será convertida em D^k ao final da iteração. Para registrar as modificações de itinerário correspondentes foi usada uma matriz de roteamento auxiliar, $R = [r_{ij}]$, a qual é uma matriz de índices. Esta é inicializada com:

$$r_{ii} = i; r_{ij} = j \text{ se } v_{ij} < \infty; r_{ij} = 0; \text{ caso contrário.} \quad (5.3)$$

Os elementos desta matriz são os rótulos dos vértices. A operação indicada no algoritmo a seguir ($r_{ij} \leftarrow r_{ik}$), executada sempre que um caminho mais curto é encontrado, indica um vértice pelo qual este caminho passa (20).

Algoritmo

Início: Dados $G = (P, E)$; matriz de valores $V(G)$

$r_{ij} \leftarrow j \forall i; D^0 \leftarrow V(G);$

para $k = 1, \dots, n$ **fazer**

.....**Início**

.....**para** $k = 1, \dots, n$ **fazer**

```

.....se  $d_{ik} + d_{kj} < d_{ij}$  então
..... início
.....  $d_{ij} \leftarrow d_{ik} + d_{kj}$ ;
.....  $r_{ij} \leftarrow r_{ik}$ 
.....fim;
..... fim;
fim;
```

Foi denotado o conjunto dos nós vizinhos de um determinado nó i por $Viz(i) = \{j \in G; A_{i,j} = 1\}$. Como as redes complexas analisadas não são direcionadas e nem ponderadas, por meio deste algoritmo foi calculado um peso w_{ij} para cada um dos nós j vizinhos de um determinado nó i . Desse modo, foi obtida uma matriz ponderada W , por meio da determinação da quantidade de menores caminhos $\sigma_{iv}(j)$ que conectam todos os pares de nós i e v , passando por um nó $j \in Viz(i)$. Assim, partindo de um nó i e passando por $j \in Viz(i)$, pode-se atingir uma maior quantidade de nós na rede por menores caminhos. Portanto, um nó $j \in Viz(i)$ terá uma maior probabilidade de ser escolhido por um caminhante que esteja em i , se este tiver um maior valor de $\sigma_{iv}(j)$. Logo, os elementos da matriz ponderada W são determinados por:

$$w_{i,j} = \frac{\sigma_{iv}(j)}{\sum_j \sigma_{iv}(j)} \quad j \in Viz(i) \quad (5.4)$$

Foi determinada a probabilidade de transição de um caminhante de um nó i para um nó j com a restrição de menor caminho da seguinte forma:

$$P_{ij}^{mc} = p_1 \cdot P_{ij} + p_2 \cdot W_{ij} \quad (5.5)$$

onde P_{ij} é a probabilidade de transição considerando a escolha do nó vizinho de forma equiprovável (caminhada aleatória padrão).

Foram adotados os valores de parâmetros $p_1 = 0,1$ e $p_2 = 0,9$ para todas as redes, para que todos os nós vizinhos tivessem uma probabilidade finita de ser escolhido. Porém, os

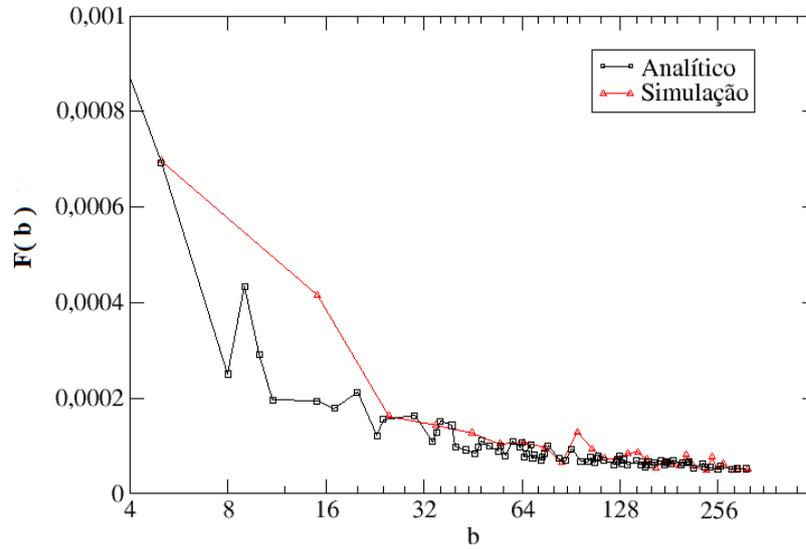
nós com uma quantidade maior de menores caminhos continuam sendo preferencialmente escolhidos para o salto de um caminhante que se encontra no nó i .

Kishore et. al (6) obtiveram um resultado analítico para as probabilidades de EE em função da centralidade de intermediação b dos nós na rede, análoga à equação (4.34), porém considerando uma caminhada aleatória por menores caminhos. Estes pesquisadores adotaram como probabilidade estacionária $p = \beta b/B$, onde $B = \sum_{i=1}^N b_i$ e $\beta \approx 0,94$, obtido numericamente. Este resultado analítico, bem como este parâmetro $\beta \approx 0,94$ foram utilizados neste trabalho de dissertação.

Nas Figs. (5.24) e (5.25) é apresentado este resultado analítico, comparando com o das simulações de caminhada aleatória por menores caminhos. Nas simulações, o valor limite q para determinação de EE foi obtido de forma análoga ao item 5.1.2. Porém, a centralidade de intermediação b não corresponde necessariamente à um número inteiro. Neste caso, utilizamos o ponto médio de um intervalo de amplitude dez, que englobasse um valor de b , para representar este intervalo.

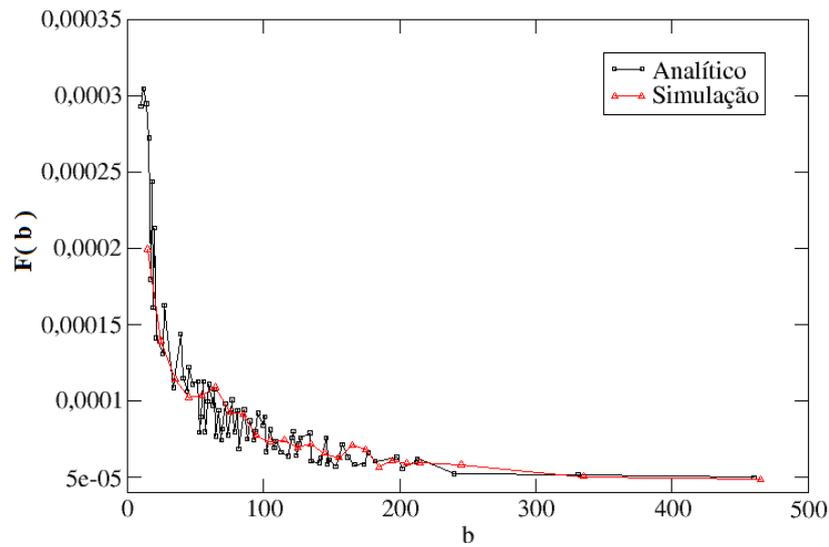
De uma forma qualitativa para todas as redes analisadas os resultados analíticos e as simulações demonstraram que as probabilidades de EE são maiores para os nós com menor centralidade de intermediação b , quando comparados com os nós com maiores valores deste, apesar de que estes apresentam um tráfego médio de caminhantes maior. Existe uma diferença entre estes resultados, como pode ser verificado na Fig. (5.25), para os nós com valores de b menores na rede. Isto se deve ao fato de existirem neste caso, poucos nós com estes valores de b para o cálculo dos valores médios.

Figura 5.24: Probabilidade de EE em função da centralidade de intermediação b com $\Delta = 0$ rede aleatória no modelo de Erdős e Renyi com $N = 100$ nós e $W = 9754$ caminhantes.



Fonte: O autor

Figura 5.25: Probabilidade de EE em função da centralidade de intermediação b com $\Delta = 0$ para rede mundo pequeno no modelo de Newman e Watts com $N = 100$ nós e $W = 9549$ caminhantes.

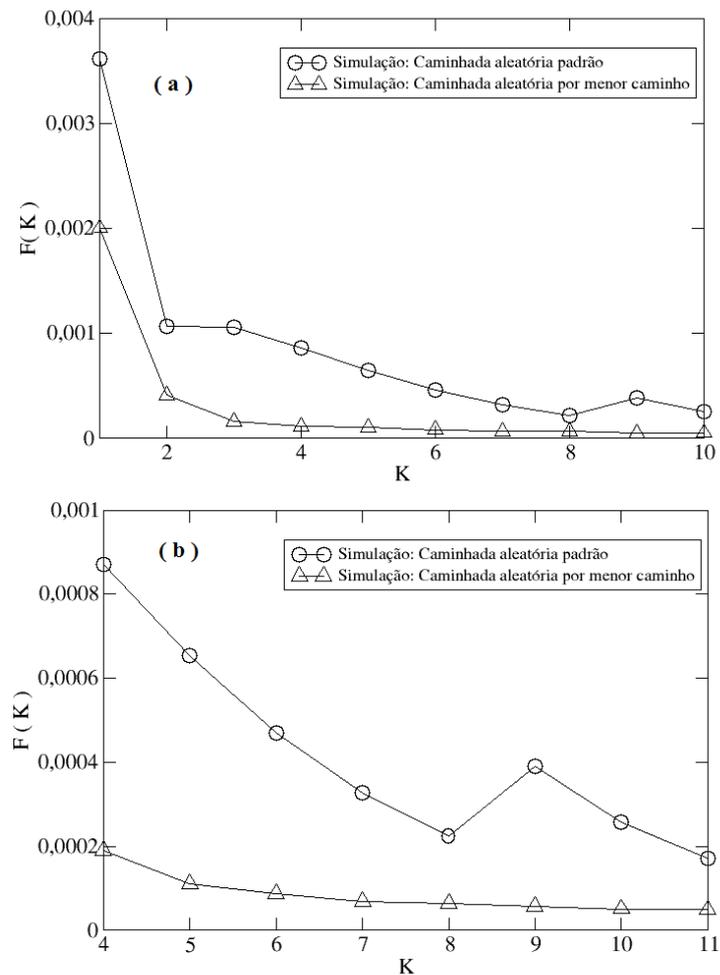


Fonte: O autor

Foram também comparados os resultados das simulações de caminhada aleatória padrão e por menores caminhos em função do grau K dos nós conforme mostrado na Fig.

(5.26). Os resultados permanecem os mesmos de forma qualitativa, ou seja, os nós com menor grau são aqueles mais propensos a ter eventos extremos. As simulações de caminhada aleatória por menores caminhos exibem menores probabilidades se comparadas com os resultados da caminhada aleatória padrão. Isto pode ser explicado pelo direcionamento de caminhantes para os nós com maiores valores de $\sigma_{iv}(j)$, que na maioria das vezes são os mais conectados ou com maiores centralidades de intermediação, diminuindo assim as probabilidades de EE.

Figura 5.26: Probabilidade de EE em função do grau K dos nós com $\Delta = 0$ e com $N = 100$. (a) Rede aleatória no modelo de Erdős e Renyi com $W = 9754$ caminhantes. (b) Rede mundo pequeno no modelo de Newman e Watts com $W = 9549$ caminhantes.



Fonte: O autor

Capítulo 6

Conclusões

Foram obtidos os parâmetros ξ e σ da distribuição de excedências generalizada de Pareto, para os valores que ultrapassaram o limite q para nós com grau mínimo e máximo de cada rede complexa analisada. Com isso, verificou-se que as excedências, para todas as redes com nós com grau mínimo, não importando o valor da sua centralidade de intermediação, não seguem a distribuição exponencial. Isto implica que a distribuição de excessos possui um limitante superior. Os valores encontrados para a estatística de teste X_0^2 no teste de hipóteses, foram superiores ao valor crítico de $\chi_{\alpha,\nu}^2 = 3,84$.

Para os nós mais conectados, resultados semelhantes foram obtidos. No entanto, para o nó com grau máximo na rede de escala livre, foi obtido um resultado diferente dos demais. Neste caso, o valor calculado para a estatística de teste X_0^2 foi 1,53, e, portanto, menor que $\chi_{\alpha,\nu}^2 = 3,84$. Assim a conclusão do teste de hipóteses realizado, não rejeita a hipótese exponencial. Porém, o valor limite $q = 94$ é muito superior aos correspondentes nas outras redes. Como consequência desse fato, foram observadas somente 15 excedências.

Para todas as redes complexas analisadas, verificou-se que os nós com menor grau possuem uma maior probabilidade de EE, quando comparados com nós com grau maior. Os tempos médios entre EE também são menores para estes nós. Há uma boa concordância entre os resultados analíticos e as simulações de caminhada aleatória, a não ser para redes bipartidas, onde o autovalor -1 pertence ao espectro de autovalores da matriz

adjacência. Isto implica que a rede pode ser subdividida em dois subconjuntos de nós onde não existem ligações entre nós no mesmo conjunto e sim, somente entre um conjunto e outro. Assim um caminhante aleatório que se encontra em um nó de um dado conjunto no instante seguinte de tempo, estará necessariamente em um nó do outro conjunto. Com isso a posição deste caminhante se alterna entre estes dois conjuntos e não se tem convergência da probabilidade estacionária.

Os resultados permanecem válidos, mesmo considerando uma flutuação no número total de caminhantes. Também verificou-se que aumentando a quantidade total de caminhantes, as probabilidades de EE sofrem um decréscimo, principalmente para os nós com menor grau na rede. Isto deve-se ao fato de que ao aumentarmos a quantidade total de caminhantes na rede, o valor limite q cresce mais rapidamente do o valor do tráfego médio para o mesmo número de caminhantes. Demonstramos que isto ocorre de forma mais acentuada para os nós com menor grau na rede.

Os resultados para simulações de caminhada aleatória por menores caminhos indicam que os nós na rede com menores centralidades de intermediação, são mais propensos a ter EE. Comparando os resultados das probabilidades de EE em função do grau dos nós, para ambas as simulações, verificou-se que por menores caminhos estas probabilidades são menores. Porém, mantém-se a mesma característica: Os nós mais propensos a ter EE são aqueles com menor conectividade na rede.

Capítulo 7

Sugestões de trabalhos futuros

Os próximos passos neste trabalho de pesquisa são: 1) A obtenção e a caracterização (Fréchet, Gumbel ou Weibull) da distribuição de máximos generalizada; 2) estender o modelo de Kishore et. al (2011)(6) para probabilidades de EE, considerando redes direcionadas e 3) obter um resultado analítico para as probabilidades de EE em função do grau dos nós, considerando a caminhada aleatória por menores caminhos para a comparação com os resultados das simulações.

Referências Bibliográficas

1. GUMBEL, E. J. **Statistics of Extremes**. New York: Dover Publ., 2004. 371 p.
2. COLES, S. **An Introduction to Statistical Modeling of Extreme Values**. London: Springer, 2001. 228 p.
3. REISS, R. D.; THOMAS, M. **Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields**. 3. ed. Berlin: Birkhäuser, 2007. 511 p.
4. ALBEVERIO, S.; JENTSCH, V.; KANTZ, H. **Extreme Events in Nature and Society**. Berlin: Springer, 2006. 352 p.
5. BEIRLANT, J. et al. **Statistics of Extremes: Theory and Applications**. England: John Wiley & Sons, 2004. 522 p.
6. KISHORE, V.; SANTHANAM, M.; AMRITKAR, R. Extreme Events on Complex Networks. **Physical Review Letters**, v. 106, n. 18, 4 p., 2011.
7. ALBERT, R.; BARABÁSI, A. L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, v. 74, n. 1, p. 47-97, 2002.
8. WATTS, D. J.; STROGATZ, S. H. Collective dynamics of "small-world" networks. **Nature**, v. 393, n. 6684, p. 440-442, 1998.
9. BARABÁSI, A. L.; ALBERT, R. Emergence of Scaling in Random Networks. **Science**, v. 286, n. p. 509-512, 1999.

10. MARTINO, D. de.; DALL'ASTA, L.; BIANCONI, G.; MARSILI, M. Congestion phenomena on complex networks. **Physical Review E**, v. 79, n. 1, 4 p., 2009.
11. HUGHES, B. D. **Random Walks and Random Environments**: Random walks. Oxford: Clarendon Press, 1995. 631 p.
12. ECHENIQUE, P.; GÓMEZ-GARDEÑES, J.; MORENO, Y. Dynamics of jamming transitions in complex networks, **Europhysics Letters**, v. 71, n. 2, p. 325-331, 2007.
13. LEADBETTER, M.R.; LINDGREN, G.; ROOTZÉN, H. **Extremes and Related Properties of Random Sequences and Processes**. Springer: London, 2012. 352 p.
14. COHEN, R.; Havlin, S. **Complex Networks: Structure, Robustness and Function**. Cambridge: Cambridge University Press, 2010. 248 p.
15. NEWMAN, M. E. J. **Networks: An Introduction**. USA: Oxford University Press, 2010. 720 p.
16. CHARTRAND, G. **Introductory Graph Theory**. New York: Courier Dover Publications, 1977. 294 p.
17. R Development Core Team. **R: A Language and Environment for statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2009. 409 p.
18. CSARDI, G.; NEPUSZ, T. The igraph software package for complex network research. **InterJournal**, Complex Systems, 2006. 1695 p.
19. NEWMAN, M.E.J.; BARABÁSI, A.L.; WATTS, D.J. **The Structure and Dynamics of Networks**. Princeton: Princeton University Press, 2011. 624 p.
20. NETTO, P. O. B. **Grafos: teoria, modelos, algoritmos**. São Paulo: Edgard Blücher, 2003. 314 p.
21. BRANDES, U. A faster algorithm for betweenness centrality. **Journal of Mathematical Sociology**, v. 25, n. 2, p. 163-177, 2001.

22. NEWMAN, M. E. J. A measure of betweenness centrality based on random walks. **Social Networks**, v. 27, p. 39-54, 2005.
23. NEWMAN, M.; WATTS, D. Scaling and percolation in the small-world network model. **Physical Review E**, v. 60, n. 6, p. 7332-7342, 1999.
24. NEWMAN, M.; STROGATZ, S.; WATTS, D. Random graphs with arbitrary degree distributions and their applications. **Physical Review E**, v. 64, n. 2, 18 p., 2001.
25. DOROGOVTSSEV, S.; MENDES, J.; SAMUKHIN, A. Size-dependent degree distribution of a scale-free growing network. **Physical Review E**, v. 63, n. 6, 4 p., 2001.
26. KRAPIVSKY, P.; REDNER, S. Organization of growing random networks. **Physical Review E**, v. 63, n. 6, 15 p., 2001.
27. BORNHOLDT, S.; Schuster, H. G. **Handbook of Graphs and Networks**. Darmstadt: Wiley, 2006. 417 p.
28. RAVASZ, E.; BARABASI, A. Hierarchical organization in complex networks. **Physical Review E**, v. 67, n. 2, 7 p., 2003.
29. BARABASI, A. L.; RAVASZ, E.; VICSEK, T. Deterministic scale-free networks. **Physica A-Statistical Mechanics and Its Applications**, v. 299, n. 3-4, p. 559-564, 2001.
30. FARKAS, I. et al. Spectra of "real-world" graphs: Beyond the semicircle law. **Physical Review E**, v. 64, n. 2, 12 p., 2001.
31. WOESS, W. **Random Walks on Infinite Graphs and Groups**. Cambridge: Cambridge University Press, 2000. 348 p.
32. BOLLOBÁS, B. **Modern Graph Theory**. New York: Springer Verlag, 1998. 411 p.

33. LOVÁSZ, L. Random Walks on Graphs: A Survey. Bolyai Society Mathematical Studies, v. 2, 46 p., 1993.
34. NOH, J.; RIEGER, H. Random walks on complex networks. **Physical Review Letters**, v. 92, n. 11, 4 p., 2004.
35. WANG, W. et al. Abrupt transition to complete congestion on complex networks and control. **Chaos**, v. 19, n. 3, 7 p., 2009.
36. ZHAO, L. et al. Onset of traffic congestion in complex networks. **Physical Review E**, v. 71, n. 2, 8 p., 2005.
37. REIF, F. **Fundamentals of Statistical and Thermal Physics**. USA: McGraw-Hill Book Company, Inc., 1965. 651 p.
38. ABRAMOWITZ, M., STEGUN, I. A. **Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables**. USA: Dover Publ., 1965. 1046 p.
39. GILLELAND, E.; KATZ, R. W. New Software to Analyze How Extremes Change Over Time. **Eos, Transactions American Geophysical Union**, v. 92, n. 2, p. 13-14, 2011.
40. HEFFERNAN, J. E.; STEPHENSON, R. **Ismev: An Introduction to Statistical Modeling of Extreme Values**, 2012.
41. MONTGOMERY, D. C., RUNGER, G. C. **Applied Statistics and Probability for Engineers**. 3. ed. USA: John Wiley & Sons, 2003. 822 p.
42. BALAKRISHNAN, V. **Schaum's Outline of Graph Theory: Including Hundreds of Solved Problems**. USA: McGraw-Hill Companies, Inc., 1997. 288 p.
43. CORMEN, T. H. et al. **Introduction To Algorithms**. 2. ed. Cambridge: MIT Press, 2001. 1180 p.