

Saulo-Davi Soares e Reis

Navegação em redes espacialmente correlacionadas.

Fortaleza – CE

30 de janeiro de 2009

Saulo-Davi Soares e Reis

Navegação em redes espacialmente correlacionadas.

Dissertação apresentada ao Curso de Pós-Graduação em Física da Universidade Federal do Ceará como parte dos requisitos para a obtenção do título de Mestre em Física.

Orientador:

Prof. Dr. José Soares de Andrade Júnior

Co-orientador:

Prof. Dr. André Auto Moreira

DEPARTAMENTO DE FÍSICA
CENTRO DE CIÊNCIAS
UNIVERSIDADE FEDERAL DO CEARÁ

Fortaleza – CE

30 de janeiro de 2009

Dissertação de Mestrado sob o título “*Navegação em redes espacialmente correlacionadas*”, defendida por Saulo-Davi Soares e Reis e aprovada em 30 de janeiro de 2009, em Fortaleza, Ceará, pela banca examinadora constituída pelos doutores:

Prof. Dr. José Soares de Andrade Júnior
Departamento de Física - UFC
(Orientador)

Prof. Dr. André Auto Moreira
Departamento de Física - UFC
(Co-orientador)

Prof. Dr. Roberto Fernandes Silva Andrade
Universidade Federal da Bahia

A minha família e amigos.

Não há vida sem eles.

Agradecimentos

Algumas pessoas merecem ser lembradas aqui, por terem de alguma forma ajudado no trabalho, ou simplesmente por terem sido fonte de inspiração, alegrias e momentos felizes.

- A minha família, antes de mais nada;
- Ao Professor José Soares de Andrade Júnior pela orientação e confiança;
- Ao Professor André Auto Moreira, pelos ensinamentos e discussões sempre precisos e indispensáveis; e pelas conversas sempre descontraídas;
- A todos os professores que de alguma forma contribuíram para o meu crescimento intelectual, em especial aos professores Raimundo Nogueira da Costa Filho, Paulo de Tarso Cavalcante Freire, Nilson de Almeida e José Ramos Gonçalves.
- Aos amigos e companheiros de laboratório, em especial Apiano e Erneson, Pablo e George, Samyr (e seus algoritmos “tampa”), Talita, Lucas, Roberto, Marília, Marco e Petrucio, sem a companhia deles a “vida” no laboratório teria sido no mínimo menos divertida;
- A todos os meus amigos que me deram apoio para o trabalho e que dividiram comigo momentos de lágrimas e de alegrias. Sejam os amigos da Física: Eduardo Girão e Acrísio, Gardênia e Cleânio, Sérgio, Franciné, Roner e Eveline, Igor, Alberto, João Philipe, Namor e Sílvia Helena, Josias e Paolo, Janilson, Roberto Sena, Aldilene, Tayroni, Daniela, Ana Tereza, Paschoal, José Junior, George Frederick, Pablo, Rogelma, Sara Rocha, Marcelo Pires e Alexandre Vieira, Luciana Angélica e Magalhães, Jeanlex, Hudson e Alana, Carlos Alex, Roberto Maluf, Ivan, Wagner, Lino, Pedro, Eduardo Barbosa, Lucas Vieira e Pedro, Klara, Diego, Victor e Alexandre; os da Paper: Maxwell, Tici, Argus, Patrícia e Baila; os da vida e do mundo: Lígia e Carol, George, Alexandre Baratta, Paulo e Armando Sands; e os da Terra Média: Levi, Paulo Júnior, Arthur, Felipe, Tiago Bin, Michael, Rafael Bruno, Lucas, Lúcio e Lígia;
- A todos os escritores que me levaram para mundos e épocas distantes (ou não) da minha realidade e que me apresentaram personagens maravilhosas (ou não), mas que me propiciaram momentos de prazer nos instantes em que queria pensar em tudo;

- Ao meu computador pessoal Salubri e às máquinas do laboratório que tive oportunidade de utilizar, sem eles esse trabalho seria impossível.

*“... a apreciação completa dos fenômenos naturais,
como os vemos,
deve ultrapassar a Física no sentido habitual.
Não pedimos nenhuma desculpa por fazer essas excursões em outros campos,
porque a separação dos campos,
como enfatizamos,
é simplesmente uma conveniência humana,
e uma coisa não natural.
A natureza não está interessada nas nossas divisões,
e muitos dos fenômenos interessantes preenchem as lacunas entre os campos.”*

Richard P. Feynman

Resumo

Um número significativo de redes reais apresentam características espaciais bem definidas. Nós estudamos como topologias de redes espacialmente correlacionadas podem influenciar processos de navegação através das mesmas. Para isso estudamos o comportamento do mínimo caminho médio para redes definidas dentro de modelo de Kleinberg [1, 2] para analisar a navegação ditada por regras de conhecimento global. O modelo que Kleinberg caracteriza-se por permitir conexões de longo alcance entre dois vértices u e v distribuídas por uma distribuição de probabilidade $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$. Para um melhor entendimento das características topológicas apresentadas por essa família de redes, nós aplicamos o modelo epidêmico *suscetível-infectado-suscetível* (**SIS**), e com isso verificamos que o modelo de Kleinberg apresenta fenômeno de mundo pequeno apenas para uma determinada faixa de valores assumidos pelo expoente de agregação α . Em seguida, introduzimos um modelo de redes espacialmente embutidas, conceitualmente inspirado no modelo de Kleinberg. Este traduz-se na introdução de um vínculo para a distribuição das conexões de longo alcance. Associamos este vínculo a um possível custo envolvido no processo de adição de novas conexões de longo alcance à rede. Estudamos como esse vínculo no custo afeta a navegação na rede, tendo como base de comparação os trabalhos de Kleinberg para a navegação com conhecimento local da topologia, e nossos resultados considerando a navegação com conhecimento global.

Abstract

A significant number of real networks have well-defined spatial characteristics. We studied how network with spatially correlated topologies can influence the processes of navigation through them. For this, we study the behavior of the average shortest-path length to networks defined within Kleinberg's model [1, 2] to analyze the navigation dictated by rules of global knowledge. The Kleinberg's model is characterized by allowing long-range connections between two vertices u and v distributed by a probability distribution of $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$. For a better understanding of the topological characteristics presented by this family of networks, we applied the epidemic model *susceptible-infected-susceptible* (**SIS**) and we found that, we see that the Kleinberg's model presents the small-world phenomenon only for a certain range of values of the clustering exponent α . We introduced a model of spatially embedded networks, conceptually based on the Kleinberg's model. This model consist in introduction of a constrain to the distribution of long-range connections. We associate this constrain to a possible cost involved in the process of adding new long-range connections to the network. We studied how this cost constrain affects the navigation through the system, taking as a basis for comparison the work of Kleinberg for navigation with local knowledge, and our results concerning for navigation with global knowledge.

Sumário

Resumo

Abstract

Lista de Figuras

Lista de Tabelas

Introdução	p. 18
1 Redes complexas	p. 21
1.1 Teoria de grafos	p. 22
1.1.1 Definições e notações	p. 22
1.1.2 Conectividade, distribuição de conectividade e correlações de conectividade	p. 23
1.1.3 Mínimo caminho e agregação	p. 24
1.2 Modelo de Erdős-Rényi	p. 25
1.2.1 Evolução estrutural da rede	p. 25
1.2.2 Distribuição de conectividade do modelo de ER	p. 26
1.2.3 Comprimento médio do caminho	p. 26
1.2.4 Coeficiente de agregação	p. 27
1.3 O mundo pequeno	p. 27
1.3.1 O experimento	p. 28
1.3.2 O modelo de Watts-Strogatz	p. 30
1.4 Barabási-Albert: Um modelo de rede livre de escala	p. 33

1.4.1	O modelo do contato preferencial	p. 33
1.4.2	Abordagem através do tratamento contínuo	p. 35
1.4.3	Equação mestra da evolução da rede	p. 37
1.5	O modelo de Kleinberg	p. 40
1.5.1	A rede	p. 40
1.5.2	O algoritmo descentralizado	p. 42
1.5.3	Limites para o valor esperado do tempo de envio da mensagem	p. 42
2	Redes espacialmente correlacionadas: navegação global e efeitos de custo	p. 47
2.1	Mínimo caminho na rede de Kleinberg	p. 48
2.2	Modelo de epidemia SIS aplicado ao modelo de Kleinberg	p. 52
2.2.1	O modelo SIS	p. 52
2.2.2	Resultados e discussões	p. 53
2.3	Navegação em redes espacialmente correlacionadas com custo	p. 57
2.3.1	O modelo de rede com custo limitado	p. 58
2.3.2	Efeitos do custo τ sobre a dinâmica da navegação Kleinberg	p. 59
2.3.3	Navegação com informação global	p. 63
3	Conclusões gerais	p. 68
	Referências Bibliográficas	p. 71

Lista de Figuras

- 1.1 Grafos direcionado e não-direcionado. (a) Um grafo direcionado $G = (V, E)$, onde $V = \{1, 2, 3, 4, 5, 6\}$ e $E = \{(1, 2), (2, 2), (2, 4), (2, 5), (4, 1), (4, 5), (5, 4), (6, 3)\}$. A aresta $(2, 2)$ é um auto-ciclo. (b) Um grafo direcionado $G = (V, E)$, onde $V = \{1, 2, 3, 4, 5, 6\}$ e $E = \{(1, 2), (1, 5), (2, 4), (2, 5), (3, 6)\}$. O vértice 4 está isolado. p. 23
- 1.2 Mapa dos EUA onde estão em destaque os estado do Kansas e Nebraska (*vermelho*), estados onde as pessoas fonte residiam; e Massachusetts (azul), onde estava localizado a pessoa alvo. p. 28
- 1.3 Seis graus de separação. Figura ilustrativa onde para uma informação ser trocada entre dois vértices quaisquer da rede, como por exemplo os vértice A e B , não são necessários mais do que seis intermediários. Figura retirada de *wikipedia.org* p. 29
- 1.4 Processo de troca entre conexões regulares e aleatórias. O número de vértices e arestas permanecem inalterados no grafo. Iniciamos com uma rede circular regular com $N = 20$ vértices e $\langle k \rangle = 4$. Com isso, substituímos uma ligação regular por um atalho com probabilidade p . Para $p = 0$ o círculo original não sofre qualquer modificação; como o aumento de p aumenta a desordem nas ligações do grafo até $p = 1$, quando o grafo torna-se completamente aleatório. Figura retirada de [3]. p. 31
- 1.5 Mínimo caminho média característico $\langle l(p) \rangle$ e o coeficiente de agregação para o modelo WS. Os dados foram normalizados pelos valores de $\langle l(0) \rangle$ e $C(0)$. A comparação entre o comportamento do mínimo caminho médio e do coeficiente de agregação evidencia um intervalo significativo de p para o qual o fenômeno de mundo pequeno está presente, onde temos uma elevado valor para $C(p)$ e valores relativamente pequenos para $\langle l(p) \rangle$. Assim, como no trabalho de watts e Strogatz, utilizamos um tamanho de rede $N = 1000$ e $k = 10$. Foram realizadas 20 execuções do modelo WS. p. 32

- 1.6 Uma rede livre de escala gerada pelo modelo de Barabási-Albert. Estão destacados os três vértices mais conectados da rede. Ilustração retirada de [4] . . . p. 34
- 1.7 Distribuição de conectividade para o modelo BA. Pode-se ver que a regra de contato preferencial utilizado pelo modelo produz o efeito de escala livre na conectividade, evidenciado pelo comportamento e lei de potência da distribuição de conectividade. Na figura, todos os três conjuntos de dados são referentes a um mesmo tamanho de rede $N = m_o + t = 300000$. Foram utilizados os parâmetros $m_o = m = 3$ (*círculos pretos*), $m_o = m = 5$ (*quadrados vermelhos*), $m_o = m = 7$ (*losangos azuis*). A linha tracejada possui inclinação $\gamma = \gamma_{BA} = 3$ p. 35
- 1.8 Distribuição de conectividade do modelo BA reescalado por um fator de $1/2m^2$. Como na Fig. (1.7), $N = 300000$. Foram utilizados os parâmetros $m_o = m = 3$ (*círculos pretos*), $m_o = m = 5$ (*quadrados vermelhos*), $m_o = m = 7$ (*losangos azuis*). A linha tracejada possui inclinação $\gamma = \gamma_{BA} = 3$ p. 36
- 1.9 Evolução temporal de dois vértices da conectividade de dois vértices escolhidos arbitrariamente. O primeiro foi adicionado a rede em um instante $t_1 = 5$ (*círculos pretos*), o segundo no instante $t_2 = 92$ (*quadrados vermelhos*). A linha tracejada possui inclinação $\beta = 0.5$, como previsto pela Eq.(1.19). . . . p. 38
- 1.10 Distribuição de Conectividade para o modelo BA, com $m_o = m = 5$ para vários tamanhos de rede diferentes. Vemos que para o limite assintótico do modelo, $t \rightarrow \infty$, o comportamento em lei de potência da distribuição de conectividade independe do tamanho do sistema. No gráfico $N = 100000$ (*círculos pretos*), $N = 200000$ (*quadrados vermelhos*), $N = 300000$ (*losangos azuis*). . . p. 39
- 1.11 Conexões de um vértice i . Cada vértice i possui ligação com todos os seus vizinhos a um distância $p = 1$ (a , b , c e d), e uma aresta com um vértice escolhido aleatoriamente j . O vértice j é escolhido com probabilidade $P(r_{i,j}) \sim r_{i,j}^{-\alpha}$ p. 41
- 1.12 Expoente η dos limites inferiores do valor esperado para o tempo de envio da mensagem, EX , como função do coeficiente de agregação α . Na figura, assim como em Kleinberg [1], $d = 2$. Podemos verificar claramente que o expoente dos limites de EX possui um mínimo para $\alpha = d$, como previsto analiticamente. Para o valor singular $\alpha = d$, o limite de EX cresce com $(\ln L)^2$, mais lentamente do que uma potência de L p. 45

- 1.13 O mesmo que Kleinberg em [2]. O logaritmo de T , número de passos necessários para que a mensagem seja enviada do vértice fonte f ao vértice alvo a utilizando o algoritmo descentralizado de Kleinberg, é plotado como uma função de α . Assim como Kleinberg encontramos um mínimo que se aproxima de 2 com o aumento do sistema. As diferenças com Fig. 1c. de [2] são devidas ao tamanho de sistemas utilizados. p. 46
- 2.1 Mínimo caminho médio $\langle l \rangle$ como uma função do expoente de agregação α . Não fizemos qualquer restrição aos contatos de longo alcance adicionados a 40% dos vértices da rede. Encontramos que o comprimento do caminho ótimo é alcançado em $\alpha_e = 0$. Com o aumento do valor de α o comprimento das ligações de longo alcance diminuem em média, e como efeito, o mínimo caminho médio $\langle l \rangle$ aumenta. Foram utilizados diversos tamanhos de rede. . . p. 49
- 2.2 $\langle l \rangle$ como função de α . Quando restringimos dois vértices quaisquer na rede, u e v , a possuírem conexões apenas dentro de uma pequena região $r_{u,v} \leq b_l \times L$, encontramos um expoente de agregação efetivo $\alpha_e = 0$. O comportamento de $\langle l \rangle$ é o mesmo para quando não existe qualquer limitação para tamanho das ligações. Aqui utilizamos apenas um tamanho de rede $L = 1000$, e adicionamos contatos de longo alcance a 20% dos vértices presentes na rede. p. 50
- 2.3 $\langle l \rangle$ como função de α . Os vértices da rede estão restritos a uma pequena região da rede $b_l \times L$ novamente. Para um mesmo valor $b_l = 0.04$, encontramos, para três tamanhos de rede diferentes, o mesmo valor $\alpha_e = 0$. Mais uma vez, adicionamos contatos de longo alcance apenas para 20% dos vértices. . . p. 50
- 2.4 Densidade de vértices infectados ρ para os estados estacionários como função de λ , para o modelo de Kleinberg com diferentes valores de α . Podemos verificar facilmente que o limiar epidêmico λ_c é dependente do expoente de agregação α . Nas simulações, foram utilizadas rede com $N = 10^6$ vértices. . . p. 54

- 2.5 Evolução no tempo t para a densidade de vértices infectados $\rho(t)$. Em todos os casos acima apresentados, utilizamos $N = 10^6$ e como condição inicial $\rho(t = 0) = 0.01$. (a) A taxa de espalhamento $\lambda = 0.205$ a 0.213 (de baixo para cima). A infecção torna-se persistente para valores próximos de $\lambda = 0.208$ ou maiores, permitindo que concluamos que $\lambda_c = 0.208$. (b) A taxa de espalhamento $\lambda = 0.205$ a 0.213 (de baixo para cima). $\lambda_c = 0.208$. (c) A taxa de espalhamento $\lambda = 0.205$ a 0.213 (de baixo para cima). Encontramos $\lambda_c = 0.209$. Aqui podemos perceber uma leve mudança no limiar endêmico. (d) A taxa de espalhamento $\lambda = 0.209$ a 0.217 (de baixo para cima). A infecção torna-se persistente para valores próximos de $\lambda_c = 0.216$ ou maiores. (e) A taxa de espalhamento $\lambda = 0.216$ a 0.224 (de baixo para cima). $\lambda_c = 0.224$. p. 55
- 2.6 O mínimo caminho l de cada vértice da rede é calculado em relação ao vértice central da rede para diferentes valores de α . O modelo de rede é construído a partir de uma rede quadrada com L^2 vértices. Aqui $L = 256$, e o custo é fixado em $\tau = L^2$. Facilmente podemos observar que com o aumento de α , a média do mínimo caminho será afetada. p. 59
- 2.7 Densidade de ρ de conexões adicionadas como função do tamanho da rede L . Para valores de $\alpha < 2$, verificamos que $\rho \sim L^{-1}$, comportamento que não se repete acima desse limiar. A linha pontilhada possui inclinação igual a -1 . No *inset* vemos que para $3 < \alpha$ a densidade ρ torna-se praticamente independente do tamanho da rede. Aqui, a linha pontilhada possui inclinação nula. p. 60
- 2.8 Densidade ρ como função do tamanho da rede L . As duas linhas pontilhada delimitam a região onde ρ apresenta escalonamento próximo de $L^{-(3-\alpha)}$, a superior com inclinação nula e a inferior com inclinação -1 . A linha traçada possui inclinação 0.5 e coincide com os valores ρ para $\alpha = 2.5$. No gráfico, $\alpha = 2$ (círculos pretos), 2.2 (estrelas vermelhas), 2.5 (losangos verdes), 2.8 (triângulos azuis), e 3 (quadrados laranjas). p. 61

- 2.9 Tempo médio característico de envio da mensagem $\langle T \rangle$ como função de α . Resultados da simulação do algoritmo descentralizado em redes de tamanho $L = 1024, 2048, 4095$; com ligações de longo alcance distribuídos com probabilidade $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$. Há um custo envolvido na adição dessas conexões de $\tau = \sum r_{u,v} = L \times L$. Esse custo altera a densidade de conexões de longo alcance, causa do mínimo encontrado em α próximo de 3 pela direita. Cada ponto é resultado de pelo menos 4000 realizações. p. 62
- 2.10 $\langle T \rangle$ como função do tamanho da rede L para cinco valores de α diferentes, encontramos um comportamento praticamente idêntico. Aparentemente, $\langle T \rangle \sim L$. A linha pontilhada possui inclinação igual a unidade e serve como comparação do escalonamento. p. 62
- 2.11 $\langle l \rangle$ como uma função de α , para uma rede 1D. Encontramos que o mínimo caminho ótimo é obtido com $\alpha_e = 2$. Com a restrição na soma total do tamanhos dos conexões de longo alcance, o número de conexões de longo alcance não é fixo, por exemplo, para $\alpha = 0$ as conexões de longo alcance conectam vértices distantes maiores tornam-se mais frequentes, o que reduz o número de conexões de longo alcance. Aqui $L = 2^{17}, 2^{18}, 2^{19}$ e 2^{20} (de baixo para cima). p. 64
- 2.12 $\langle l \rangle$ como uma função de α , para uma rede 2D. Encontramos que o mínimo caminho ótimo é obtido com $\alpha_e = 3$ p. 65
- 2.13 Mínimo caminho $\langle l \rangle$ como função do expoente de agregação α . Aqui são permitidas conexões entre vértices apenas na mesma direção (por exemplo, sem relacionarmos cada vértice u a uma posição na rede quadrada (x_u, y_u) , apenas um vértice v com posição (x_u, y_v) , ou (x_v, y_u) , poderá ser sorteado para receber a conexão. Encontramos um $\alpha_e = 2.1$, resultado próximo ao resultado obtido para uma rede unidimensional. Esse resultado não é mera coincidência devido às semelhanças topológicas entre a rede unidimensional e a rede bidimensional com vínculo nas direções. Esse resultado aumenta as possibilidades de estratégias que podem ser utilizadas para otimizar a navegação através da rede. p. 66

2.14 Mínimo caminho $\langle l \rangle$ como função de L para uma rede com o substrato quadrado. Há uma restrição no comprimento total das ligações $\tau = \sum r_{u,v}$. A curva para $\alpha_3 = 3$ cresce mais lentamente com o tamanho da rede do que qualquer outro valor de α . No *inset* vemos que o crescimento de $\langle l \rangle$ na condição ótima $\alpha_e = 3$ é consistente com o comportamento logarítmico. p.67

Lista de Tabelas

- 2.1 Expoente de agregação α e seu respectivo limiar endêmico λ_c . Aqui utilizamos $p = q = 1$, o que representa uma rede com $\langle k_{in} \rangle = 5$. Pela previsão da aproximação de campo médio, $\lambda_c = \langle k_{in} \rangle$ para regimes de mundo pequeno. Como discutido na seção anterior, o modelo de Kleinberg apresenta comportamento próximo ao de mundo pequeno (modelo WS) apenas para a região $0 \leq \alpha \leq 2$. Para valores de $\alpha > 2$ o limiar endêmico distancia-se do previsto para redes que seguem o paradigma de WS. p. 56

Introdução

Em Física, um importante contato entre a Matemática e a Termodinâmica é feito, aproximadamente, na década de 70 do séc. XIX, quando Boltzmann desenvolve seu teorema H e, desse modo, estabelece uma conexão direta entre a entropia e a dinâmica molecular[5]. Contudo, por volta da mesma época, a teoria cinética dos gases indica um novo formalismo baseado em *ensembles* estatísticos. A união dessas poderosas técnicas culminam na criação de uma nova ferramenta para o tratamento matemático de sistemas com um grande número de constituintes, a Mecânica Estatística. Inúmeras aplicações da Mecânica Estatística a problemas físicos, traduzidos na forma de *modelos de rede regular*, ou lattices, foram desenvolvidas com sucesso desde então. Tais modelos tentam descrever propriedades termodinâmicas de sistemas físicos estabelecendo de forma simples as interações entre componentes de uma rede. O modelo de Ising, o modelo de Potts e o modelo de Heisenberg são exemplos de modelos de rede amplamente aplicados a sistemas magnéticos.

Redes são sistemas compostos por *vértices*, por vezes chamados de nós ou sítios, os quais possuem algum tipo de conexão entre os mesmos que chamaremos de *aresta*. Portanto, é natural desejar representar *sistemas complexos* constituídos por um grande número de agentes, e que de alguma forma interagem entre si, por uma rede; nesse caso os vértices desempenham o papel dos agentes e as arestas expressam as interações entre os mesmos. A complexidade de tais sistemas está nas suas características emergentes devido a interação desses constituintes, ou seja, os sistemas como um todo possuem propriedades que não podem ser verificadas quando analisamos o comportamento de um constituinte isoladamente.

Alguns exemplos de *redes reais* são encontrados em diversos ramos da ciência, como na Informática, com a World Wide Web (WWW) [6, 7], na Tecnologia com a Internet [8], na Biologia, como por exemplo em redes celulares de processos metabólicos [9], e nas Ciências Sociais em redes de relações de indivíduos em grupos sociais [10, 11]. Isso levou conceitos de Mecânica Estatística a diferentes campos de conhecimento, visando o melhor tratamento de sistemas complexos. Nestes sistemas, é comum encontrarmos uma rede de interações não trivial, impossível de ser representada por uma rede regular de vizinhos como no modelo de Ising; no entanto, o levantamento de dados dessas redes reais evidenciaram a existência de propriedades comuns a redes inicialmente tão distintas. Assim, estudar características topológicas dessas re-

des são fundamentais para a compreensão de sistemas complexos [4].

A grande parte dos modelos de redes propostos para modelar redes reais são comumente desprovidos de características espaciais. Usualmente, considera-se que as possíveis interações entre componentes distantes espacialmente podem ser reduzidas a um *campo médio* de tal forma que uma verdadeira correlação espacial entre as interações dos vértices é ignorada. Recentemente, Kleinberg [1, 2] estudou a dinâmica de navegação do envio de mensagens em um modelo de rede que possui ligações de longo alcance distribuídas com probabilidade na forma de uma lei de potência $\sim r_{u,v}^{-\alpha}$, onde u e v são vértices da rede. Neste trabalho, abordaremos o problema da navegação em redes complexas espacialmente correlacionadas através de uma distribuição de probabilidade do tipo introduzida por Kleinberg. Esta distribuição possui uma interpretação bastante simples, ela nos diz que a probabilidade de dois vértices estarem conectados dependerá da distância que os separa. Com isso, é mais provável que dois vértices estejam conectados quanto menor for a distância entre eles. Podemos encontrar outros modelos de rede onde a probabilidade de dois vértices serem conectados possui um decaimento em lei de potência com uma distância como por exemplo em [12], onde a distância r em questão é a distância de um vértice adicionado a rede e o baricentro da mesma.

Esta dissertação está dividida da seguinte forma. No Capítulo 1, apresentaremos uma breve exposição da teoria de redes complexas, para tanto, introduziremos conceitos fundamentais da teoria dos grafos, importantes para o tratamento e interpretação dos fenômenos presentes nas redes reais. Em seguida, apresentaremos os três principais características que redes complexas podem possuir. São eles: A aleatoriedade, apresentada aqui através do modelo dos matemáticos P. Erdős e A. Rényi [13]; o fenômeno de mundo pequeno, introduzido no meio acadêmico pelo sociólogo americano S. Milgram [14], e posteriormente melhor explicado através do modelo de D. Watts e S. Strogatz [15, 16], que hoje serve de paradigma para a elaboração de modelos de rede que busquem reproduzir o fenômeno de mundo pequeno; e a liberdade de escala para a distribuição de conectividade, primeiramente estudado por D. J. de S. Price [17, 18], e posteriormente extensivamente estudado e esclarecido por A.-L. Barabási e R. Albert [19, 20]. Depois, introduziremos o modelo de Kleinberg para redes de mundo pequeno, e apresentaremos suas conseqüências sob um processo de navegação que se utiliza apenas da informação local que a rede fornece.

No Capítulo dois, apresentaremos os estudos realizados por nós sobre a navegação com informação global no modelo de Kleinberg, isso é realizado com o pressuposto de que, se o vértice possui o conhecimento global da posição de todos os demais vértices da rede, e de todas as conexões de longo alcance presentes na rede, o tempo de navegação pode ser caracterizado através do cálculo do mínimo caminho médio da rede. Como veremos, estes resultados in-

dicam que existem algumas classes de redes definidas pelo modelo de Kleinberg que não se comportam como redes de mundo pequeno. Para verificar este fato, aplicaremos uma dinâmica de processos epidêmicos no modelo de Kleinberg e discutiremos seus resultados. Introduziremos também um modelo de rede, inspirado conceitualmente no modelo de Kleinberg, que possui um vínculo para a ligações de longo alcance. Este vínculo está representado na forma de um custo para a adição de novas ligações. Veremos como esse custo altera os processos de navegação no modelo, tanto para o processo de navegação local proposto por Kleinberg, quanto para o processo de navegação que se utiliza de todo o conhecimento da estrutura da rede. Por fim, apresentaremos um último, em que estão nossas conclusões sobre os resultados obtidos, acompanhadas de perspectivas para trabalhos futuros.

1 *Redes complexas*

Redes complexas têm sido o foco de muitos trabalhos na comunidade científica como um todo. Em especial, a abordagem de sistemas complexos onde tem-se mostrado um tema produtivo na conexão de campos tradicionalmente distintos como a Física, a Química, a Biologia, a Ecologia, e os campos das ciências sociais como a Sociologia e a Economia. Sistemas complexos podem ser definidos como um grande número de componentes que interagem entre si de alguma forma tal que produz características como a emergência de fenômenos devidos ao conjunto, ausentes quando estudamos um componente individualmente. Uma das idéias mais fundamentais na área de redes complexas, inspirado no trabalho seminal de D. Watts e S. Strogatz [3], tem sido o estudo comparativo de redes presentes em diferentes ramos da ciência; com isto, grande parte das redes complexas construídas nasce das observações de propriedades de sistemas reais. Podemos dividir os dados obtidos em quatro classes de redes (não muito rígidas em sua definição): redes sociais, redes de informação, redes tecnológicas e redes biológicas. Estatisticamente, estas redes compartilham características matemáticas similares [4, 19, 21]. Academicamente, as ciências sociais possuem uma longa história no estudo de *redes reais*, particularmente, os trabalhos de J. Moreno durante a década de 20 e 30 que estudaram padrões de amizade dentro de pequenos grupos de pessoas. [22]. Mais recentemente, estudos de associações de negócios financeiros [23], de padrões em contatos sexuais [24–26], têm atraído muita atenção da comunidade científica.

No presente capítulo, apresentaremos uma breve revisão sobre redes complexas. Para tanto, introduziremos em princípio, conceitos da teoria de grafos, como a conectividade, o mínimo caminho e a agregação. Em seguida, apresentamos uma breve discussão a respeito de redes reais, para então introduzirmos as principais topologias de redes complexas, acompanhadas sempre dos modelos que apresentaram os mecanismos responsáveis por suas formação. Os modelos escolhidos são o modelo Watts-Strogatz [3] para mundo pequeno, que define um paradigma para modelos de mundo pequeno; o modelo Barabási-Albert [20], que mostra que o princípio de contato preferencial é o mecanismo responsável pela liberdade de escala na distribuição de conectividade; e por fim, apresentamos o modelo de Kleinberg [1, 2] para redes de mundo

pequeno espacialmente embutidas. Este iniciou uma interessante abordagem da topologia de mundo pequeno através de um algoritmo fortemente inspirado no famoso experimento de Milgram [14].

1.1 Teoria de grafos

O primeiro problema real a ser resolvido pela teoria de redes, na forma matemática da teoria de grafos, é freqüentemente citado como sendo a celebrada solução do matemático suíço Leonhard Euler (1707-1783), no ano de 1736, para o problema das *sete pontes de Königsberg* (hoje, cidade conhecida como Kaliningrado, exclave da Federação Russa situado entre a Polônia, a Lituânia e o mar Báltico), e, ao longo do último século, a teoria de grafos desenvolveu-se como um importante ramo do conhecimento científico.

Nesta seção, apresentaremos algumas propriedades que caracterizam a estrutura de grafos, e as definições da teoria de grafos. Ambas, são ferramentas importantes para o entendimento e para o estudo de redes complexas.

1.1.1 Definições e notações

Pode-se representar matematicamente uma rede, seja complexa ou não, por meio de um *grafo*. Um grafo é um par de conjuntos $G = (V, E)$ [27], onde V é um conjunto finito de N vértices, e E é uma relação binária de V , que representa as arestas entre dois vértices de V . Cada aresta é representada pelo par de dois vértices (i, j) , e pode ser do tipo *direta*, ou seja, a aresta possui um sentido determinado, como uma rua de mão única; ou *indireta*, onde ela não determina um único sentido.

Um grafo composto por arestas direcionadas é chamado de grafo direto ou *dígrafo*. Redes que representem alguma estrutura de envio de mensagens, seja de mensagens eletrônicas ou de cartas, ou uma rede de ligações telefônicas, devem ser representadas por um grafo direto, visto que cada mensagem possui uma única direção. A Fig. 1.1(a) é uma representação pictórica de um grafo direto G , com o conjunto de vértices $V = \{1, 2, 3, 4, 5, 6\}$. Os círculos representam os vértices, e as setas representam as ligações. Grafos diretos aceitam, ou não, a existência de auto-ciclos, ou seja, vértices possuem arestas direcionadas a eles mesmos.

Um grafo *não-direcionado* é representado pelo conjunto G da mesma forma que o grafo direcionado, porém o conjunto de arestas E é um conjunto não ordenado, ou seja, $(i, j) = (j, i)$, como mostrado na Fig. 1.1(b) que apresenta um exemplo deste tipo de grafo. Em grafos não-direcionados, $i \neq j$, proibindo com isto a presença de auto-ciclos, e implicando que, uma aresta

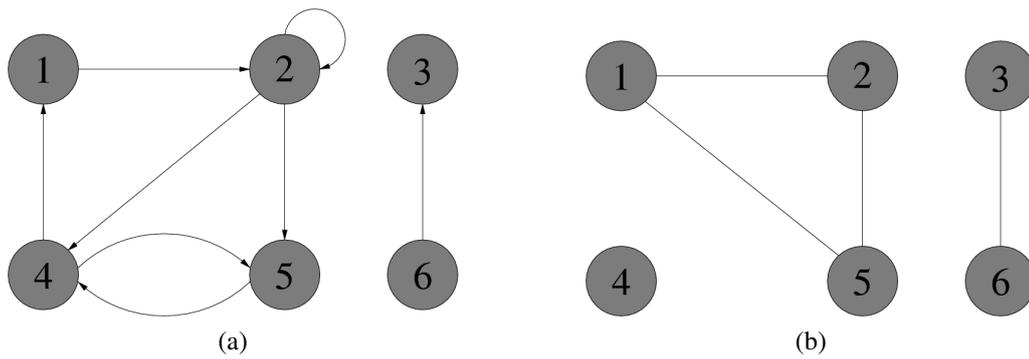


Figura 1.1: Grafos direcionado e não-direcionado. (a) Um grafo direcionado $G = (V, E)$, onde $V = \{1, 2, 3, 4, 5, 6\}$ e $E = \{(1, 2), (2, 2), (2, 4), (2, 5), (4, 1), (4, 5), (5, 4), (6, 3)\}$. A aresta $(2, 2)$ é um auto-ciclo. (b) Um grafo direcionado $G = (V, E)$, onde $V = \{1, 2, 3, 4, 5, 6\}$ e $E = \{(1, 2), (1, 5), (2, 4), (2, 5), (3, 6)\}$. O vértice 4 está isolado.

(i, j) consiste em um par de dois vértices distintos.

1.1.2 Conectividade, distribuição de conectividade e correlações de conectividade

A *conectividade*, ou o grau de um vértice i em uma rede não-direcionada, é definido como o número de vértices *adjacentes* ao vértice i [21], ou seja, a conectividade k_i de i é igual ao seu número de vizinhos. No entanto, se estivermos tratando de um grafo direcionado, a conectividade de um vértice é constituído por duas parcelas distintas: k_i^{out} que contabiliza o número de arestas que saem do vértice i e incidem em outros vértices, e a conectividade k_i^{in} que, por sua vez, representa o número de ligações incidentes a i , de forma que,

$$k_i = k_i^{in} + k_i^{out}. \quad (1.1)$$

Se definirmos uma fração p_k de vértices com conectividade k , teremos com isto, de forma equivalente, que p_k é a probabilidade de escolhermos, de forma aleatória, um vértice com conectividade k . Uma representação gráfica dessas probabilidade p_k pode ser feita através de um histograma o qual é conhecido como distribuição de conectividade $P(k)$. Esta distribuição de conectividade é um dos principais meios de caracterização de uma rede complexa. No entanto, existem redes que não são caracterizadas de forma completa pela sua distribuição de conectividade. Algumas redes reais podem apresentar correlações em sua conectividade [28], de modo que, é interessante introduzir uma *probabilidade condicional* $P(k'|k)$, ao explicitar que dado um vértice com conectividade k , a probabilidade dele ser conectado a um vértice com conectividade k' depende de sua própria conectividade k .

O estudo da probabilidade condicional $P(k'|k)$ em redes reais produz resultados muitas

vezes difíceis de interpretar. Com o objetivo de se caracterizar correlações de conectividade, torna-se muito útil trabalhar com a média da conectividade dos vizinhos mais próximos, \bar{k}_{nn} , dos vértices com conectividade k , dado por:

$$\bar{k}_{nn}(k) = \sum_{k'} k' P(k'|k). \quad (1.2)$$

Em muitas redes sociais, é verificada a preferência de vértices com alta conectividade conectarem-se a outros vértices também altamente conectados, neste caso, $\bar{k}_{nn}(k)$ é uma função crescente de k . No entanto, em muitas redes tecnológicas e redes biológicas, verifica-se que vértices muito conectados possuem preferência de se conectarem a vértices de baixa conectividade, com isso, $\bar{k}_{nn}(k)$ apresenta-se como uma função decrescente de k [29]. Portanto, o estudo do comportamento de \bar{k}_{nn} fornece uma forma simples de caracterizar correlações de conectividade por conter parte das informações contidas na função probabilidade condicional $P(k'|k)$.

1.1.3 Mínimo caminho e agregação

O caminho ℓ de um vértice i a um outro vértice i' em uma rede representada por um grafo $G = (V, E)$ é uma seqüência $\langle i_0, i_1, i_2, \dots, i_\ell \rangle$ onde $i = i_0$, $i' = i_\ell$, e $(i_{m-1}, i_m) \in E$ para $m = 0, 1, 2, \dots, \ell$ [27]. Assim, o comprimento do caminho ℓ é o número de arestas que compõe o caminho.

Definindo o mínimo caminho l_{ij} como a seqüência de arestas que minimiza o caminho ℓ entre os vértice i e j , podemos, com isto, definir o *mínimo caminho médio* $\langle l \rangle$ [30] de uma rede como

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{i,j \in G; i \neq j} l_{ij}. \quad (1.3)$$

A definição acima pode ser problemática em redes que sejam constituídas por mais de um componente onde existem pares de vértices que não possuem um caminho ℓ que os conecte. Usualmente, considera-se o caminho mínimo de um par de vértices desses tipo como sendo infinito, o que resulta em um valor também infinito para $\langle l \rangle$. Para contornar esse problema, é comum desconsiderar pares de vértices que pertençam a diferentes componentes, ou seja, apenas pares que possuam um caminho que os conecte são considerados no cálculo de $\langle l \rangle$.

Outro importante parâmetro topológico de caracterização de uma rede complexa é a *agregação* ou transitividade. Por agregação, entenda-se a probabilidade de que dado dois vértices i' e i'' , ambos conectados a um vértice i , estes possuam uma aresta conectando-os. No jargão das redes sociais, podemos dizer que agregação é a probabilidade de um amigo de um amigo seu seja, também, seu amigo. Traduzindo para termos da teoria de grafos, a agregação mede a probabilidade da existência de triângulos na rede, ou seja, a existência de conjuntos de três

vértices onde cada um destes é conectado aos demais. A agregação pode ser quantificado através do coeficiente de agregação C que pode ser calculado pelo método introduzido por Watts e Strogatz [3], utilizando o *coeficiente de agregação local* definido como sendo a fração de arestas existentes entre os primeiros vizinhos de i , portanto,

$$c_i = \frac{2e_i}{k_i(k_i - 1)}. \quad (1.4)$$

Em que e_i é o número de arestas existentes entre os vértices conectados a i , e o termo $k_i(k_i)/2$ representa todas as possíveis arestas que podem existir conectando os vizinhos de i . O coeficiente de agregação é a média de c_i sobre todos os vértices presentes na rede, ou seja,

$$C = \frac{1}{N} \sum_{i \in G} c_i. \quad (1.5)$$

1.2 Modelo de Erdős-Rényi

O artigo inicial de Paul Erdős e Alfréd Rényi [13] sobre redes aleatórias inicia, aparentemente, o interesse da comunidade científica em grafos com topologia complexa. Neste trabalho, Erdős e Rényi (**ER**) propõem um modelo de redes com N vértices conectados por n arestas escolhidas aleatoriamente com probabilidade p . O modelo ER define um *ensemble*, $G_{N,p}$, de todas as redes tal que uma rede com m arestas aparece com probabilidade $p^m(1-p)^{M-n}$, onde $M = \frac{1}{2}N(N-1)$ é o número máximo de arestas possível. Muitos resultados exatos podem ser extraídos do modelo ER para redes dentro do limite $N \rightarrow \infty$. Neste limite, a rede de ER possui um valor constate de conectividade média dada por:

$$\langle k \rangle = p(N-1) \simeq pN. \quad (1.6)$$

Nesta seção temos por objetivo apresentar algumas características presentes no modelo ER.

1.2.1 Evolução estrutural da rede

Muitas propriedades estruturais do modelo de ER apresentam comportamento dependente de p . Podemos assumir que $p(N) \sim N^{-z}$, onde z é um parâmetro ajustável e varia dentro do intervalo $[0, \infty)$. Diversos valores de z apresentam-se como limiar para diferentes arranjos estruturais da rede de ER. Para valores de z maiores que $3/2$, a presença de vértices isolados é dominante, porém à medidas que z se aproxima de 1, a rede apresenta componentes de ordem cada vez maior. Esta análise do modelo revela a presença de uma transição de fase para $z = 1$, correspondendo a uma probabilidade crítica $p_c = 1/N$ e uma média de conectividade crítica

$\langle k \rangle_c = 1$, de modo que:

- Para valores de $p < p_c$, a rede não possui componentes maiores que $O(\ln N)$, e nenhum componente tem mais que um ciclo.
- Para $p = p_c$, emerge o maior componente da rede de tamanho $O(N^{2/3})$; na transição, o tamanho dos componentes estão distribuídos em lei de potência com expoente $\tau = 5/2$ (ou $3/2$ se estivermos analisando o componente a que pertence um vértice escolhido aleatoriamente).
- Para valores de $p > p_c$, a rede tem um componente de tamanho $O(N)$.

Esta é uma transição de fase de segunda ordem e pertence à classe de universalidade da percolação de dimensão infinita.

1.2.2 Distribuição de conectividade do modelo de ER

A distribuição de conectividade $P(k_i)$ de uma rede de ER com probabilidade de conexão p para um vértice i de conectividade k_i segue a distribuição binomial [21]:

$$P(k_i = k) = C_{N-1}^k p^k (1-p)^{N-1-k}. \quad (1.7)$$

Visto que o vértice i pode possuir até $N - 1$ arestas, temos que a probabilidade de que este tenha k_i conexões é dada por p^{k_i} e a probabilidade de que as demais $(N - 1 - k_i)$ conexões não existam é $(1 - p)^{N-1-k_i}$. Temos também um fator de $C_{N-1}^{k_i}$, que é o número de formas que as k_i ligações podem ser combinadas dentro das $N - 1$ possibilidades. Para o limite de $N \rightarrow \infty$, podemos aproximar a distribuição binomial por uma distribuição de Poisson:

$$P(k) \simeq e^{-pN} \frac{(pN)^k}{k!} = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (1.8)$$

Portanto, apesar de termos aleatoriedade na posição das arestas, a rede de ER é aproximadamente homogênea, pois a grande maioria dos vértices possui o mesmo número de arestas.

1.2.3 Comprimento médio do caminho

Uma forma de caracterizar difusão, espalhamento e navegação de algum tipo de propriedade ou informação em uma rede aleatória, como o modelo de ER, é calculando-se o *comprimento médio do caminho* $\langle \ell \rangle$ entre qualquer par de vértices presentes na rede; ou seja, o comprimento médio do caminho entre dois vértices em uma rede de ER, $\langle \ell \rangle_{ER}$, é a média do comprimento do

caminho ℓ_{ER} entre todos os pares de vértices da rede.

Como podemos dizer que cada vértice da rede possui em média $\langle k \rangle$ vizinhos e que cada um destes resulta, em média, em $\langle k \rangle^2$ segundos vizinhos que, por sua vez, resultam em $\langle k \rangle^3$ terceiros vizinhos, e assim por diante, teremos que se o número N de vértices for suficientemente grande, e com isso $\ln(N) \gg \langle k \rangle$, o número de vértices a um caminho de comprimento ℓ_{ER} continuará crescendo exponencialmente até que o número de vértices distantes entre si por um caminho de comprimento ℓ_{ER} é $O(N)$. Equacionando $\langle k \rangle^{\ell_{ER}}$ com N por meio da função logarítmica, teremos:

$$\ell_{ER} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad (1.9)$$

Como $\langle \ell \rangle_{ER}$ não é nada além do que a média de ℓ_{ER} , teremos que:

$$\langle \ell \rangle_{ER} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad (1.10)$$

Portanto, temos que o comprimento do caminho médio do modelo de ER depende de forma logarítmica de N . Na próxima seção veremos que esse comportamento caracteriza fenômeno de mundo pequeno. Ao utilizarmos o resultado da Eq. (1.6) podemos reescrever $\langle \ell \rangle$ como sendo:

$$\langle \ell \rangle \sim \frac{1}{1-z}. \quad (1.11)$$

1.2.4 Coeficiente de agregação

Considerando um vértice em uma rede aleatória do tipo ER e seus primeiros vizinhos, a probabilidade de dois destes estarem conectados é igual à probabilidade de dois vértices escolhidos aleatoriamente estarem conectados. Portanto, o coeficiente de agregação de uma rede de ER é

$$C_{ER} = p = \frac{\langle k \rangle}{N}. \quad (1.12)$$

Portanto, pela Eq. (1.12), a fração $C_{ER}/\langle k \rangle$ decresce com N^{-1} . Este comportamento do modelo ER faz com que ele não seja ideal na modelagem de redes reais onde é verificado que o coeficiente de agregação é independente do tamanho da rede.

1.3 O mundo pequeno

Todos nós já ouvimos e, muito provavelmente, utilizamos o termo “mundo pequeno”. Em 1967, Stanley Milgram, psicólogo e sociólogo, professor em Harvard, introduziu o termo no vocabulário científico [14]. Em seu trabalho, Milgram propôs um experimento cujo objetivo

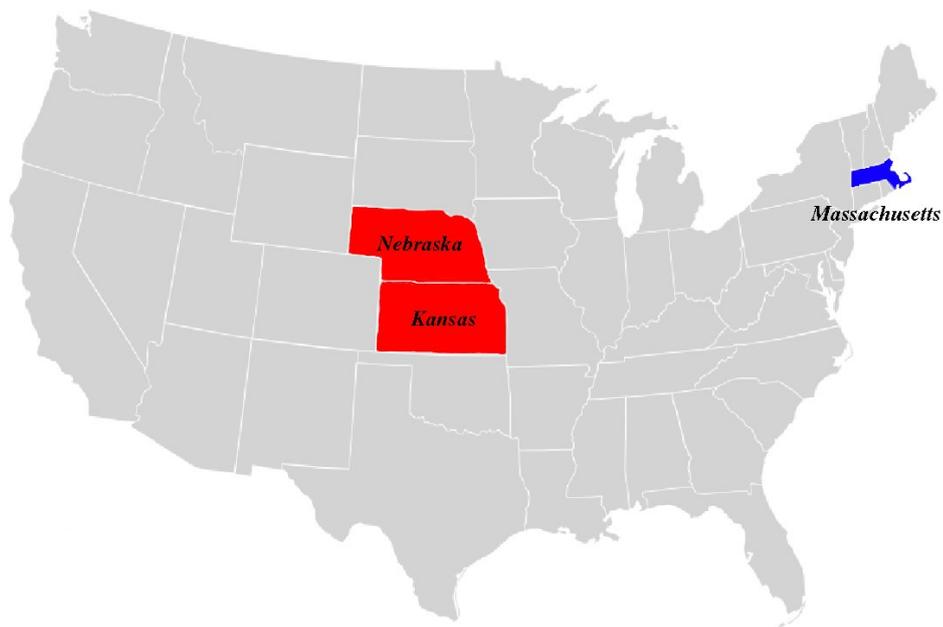


Figura 1.2: Mapa dos EUA onde estão em destaque os estado do Kansas e Nebraska (*vermelho*), estados onde as pessoas fonte residiam; e Massachusetts (*azul*), onde estava localizado a pessoa alvo.

era medir o comprimento dos caminhos entre duas pessoas quaisquer no Estados Unidos da América.

1.3.1 O experimento

O experimento desenvolvido por Milgram consistia basicamente no envio de correspondência entre duas pessoas nos Estados Unidos da América (**EUA**). Dentro do vocabulário empregado na teoria de redes complexas, essas pessoas podem ser vistas como vértices presentes em uma rede social real, onde as ligações são algum tipo de vínculo. No experimento, foi escolhido uma pessoa “alvo” que era um corretor em Boston, Massachusetts; e, aleatoriamente, pessoas “fontes” em outras duas cidades estadunidenses, Wichita, no estado do Kansas, e Omaha, no Nebraska (Fig. 1.3).

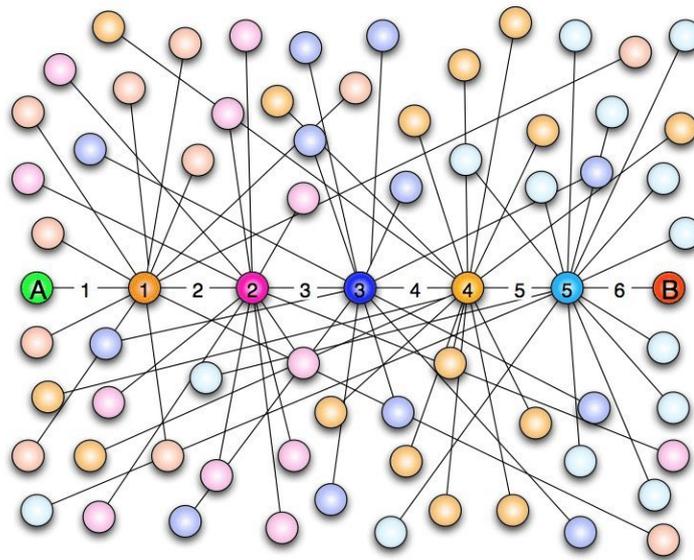


Figura 1.3: Seis graus de separação. Figura ilustrativa onde para uma informação ser trocada entre dois vértices quaisquer da rede, como por exemplo os vértice *A* e *B*, não são necessários mais do que seis intermediários. Figura retirada de *wikipedia.org*

Os indivíduos fonte receberam uma caixa que continha um breve resumo do experimento, juntamente com informações sobre o alvo, como uma foto, nome e endereço, além de uma lista e um conjunto de cartões postais.

Procedimento Básico

1. O participante deveria adicionar seu nome e o nome da próxima pessoa que receberia a caixa na lista.
2. Deveria pegar um dos cartões, existentes em grande quantidade na caixa, e enviar de volta a Harvard. Isso permitiu que Milgram acompanhasse o “caminho” do pacote.
3. Se ela conhecesse¹ a pessoa alvo, a caixa deveria ser enviada diretamente a ela.
4. Caso o portador da caixa não conhecesse o alvo, ele deveria enviá-la para alguém conhecesse e que, de uma forma ou de outra, acreditasse que agilizaria o processo de envio da correspondência à pessoa alvo.

Muitas cartas não chegaram ao alvo, porém, as cartas que finalizaram seu caminho, indicaram, para a surpresa de Milgram, que as caixas precisaram passar por apenas seis intermediários

¹No experimento, por motivos culturais, *conhecer* significava tratar a pessoa pelo primeiro nome, o que nos EUA representa, por exemplo, a existência de alguma relação de amizade, coleguismo ou vínculo familiar.

em média. Esse resultado, conhecido hoje como “seis graus de separação”, apesar de receber inúmeras críticas, foi, de certo modo, reproduzido por Duncan Watts e seus colaboradores[15] em um experimento utilizando mensagens de *e-mail* que contou com 18 pessoas alvo escolhidas em 13 países, um total de 61,163 indivíduos espalhados por países 166 que constituíram 24,163 cadeias de mensagens distintas.

Os seis graus de separação não nos dizem apenas o quão fortemente conectados estamos, mas também nos indica que estamos a poucos “passos” uns dos outros, mostram nos que o mundo é pequeno.

1.3.2 O modelo de Watts-Strogatz

O primeiro modelo de redes de mundo pequeno foi inicialmente proposto por D. Watts e S. Strogatz (SW) [3], baseando-se no fato de que um possível modelo de rede de mundo pequeno deveria permitir a coexistência de duas importantes características, evidenciadas pelo experimento de Milgram. Primeiro, um alto valor para o coeficiente de agregação da rede C , o que no experimento social significa que vértices com vizinhos comuns possuem alta probabilidade de estarem conectados. Segundo um valor pequeno para o mínimo caminho médio da rede $\langle l \rangle$ mesmo para grandes tamanhos de rede. Produzindo assim, o *fenômeno do mundo pequeno* [3, 16] no sentido de que o diâmetro topológico cresce lentamente (*logaritimicamente*) com o aumento do tamanho da rede.

Para redes aleatórias, o coeficiente de agregação (Eq. 1.12) é muito baixo para modelar redes reais de mundo pequeno. Enquanto isso, as redes regulares (redes onde a média da conectividade $\langle k \rangle$ é uma constante) apresentam altos valores para $\langle l \rangle$ quando lidamos com tamanhos de rede muito grandes², o que as tornam incapazes de modelar, por exemplo, os seis graus de separação. No entanto, devido sua aleatoriedade, redes aleatórias possuem altos valores de $\langle l \rangle$; em contra partida em muitas redes regulares, dependendo de sua estrutura local, temos alta probabilidade de formação de ciclos, ou seja, dependendo da topologia a rede, C possui valores compatíveis com os encontrados em redes reais.

Com isso Watts e Strogatz conjecturaram, e provaram, que um modelo capaz de reproduzir, de alguma forma, o fenômeno do mundo pequeno, deve possuir uma topologia situada entre uma topologia de uma rede regular e de uma rede completamente aleatória.

O modelo de WS consiste basicamente na adição de ligações aleatórias a uma rede regular de dimensão e topologia quaisquer. Seguindo o modelo WS, a adição desses *atalhos* pode ser realizada de duas formas, uma através da simples inserção de ligações aleatórias distri-

²Um exemplo de rede regular é a rede quadrada presente em modelos de sistemas magnéticos. Em uma rede desse tipo $\langle l \rangle$ deve ser $O(N)$.

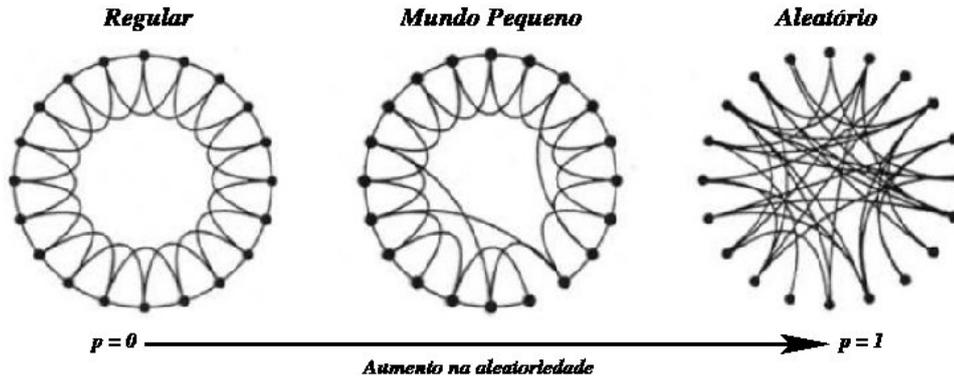


Figura 1.4: Processo de troca entre conexões regulares e aleatórias. O número de vértices e arestas permanecem inalterados no grafo. Iniciamos com uma rede circular regular com $N = 20$ vértices e $\langle k \rangle = 4$. Com isso, substituímos uma ligação regular por um atalho com probabilidade p . Para $p = 0$ o círculo original não sofre qualquer modificação; como o aumento de p aumenta a desordem nas ligações do grafo até $p = 1$, quando o grafo torna-se completamente aleatório. Figura retirada de [3].

buídas uniformemente entre pares de vértices com probabilidade p , ou por meio da troca de uma ligação presente na rede regular por um atalho, também com probabilidade p . Em ambos os procedimentos, quando $p = 0$ é sorteado, alteração alguma na rede deve ser realizada. A forma mais extensamente estudada para o modelo WS consiste em uma rede unidimensional com condições de contorno periódicas, em outras palavras, com N vértices conectados a seus K primeiros vizinhos a direita e a esquerda, uma rede circular regular, onde $\langle k \rangle = 2K$. A partir dessa rede regular, inicia-se o processo de substituição de arestas por atalhos, como mostrado na Fig. (1.4). No modelo WS, auto-ligações e ligações redundantes são proibidos.

Para melhor entender o efeito da adição desses atalhos, devemos analisar o comportamento do coeficiente de agregação $C(p)$ e do mínimo caminho médio $\langle l(p) \rangle$ em função da probabilidade p . É esperado que para $p \rightarrow 0$, $\langle l(0) \rangle \simeq N/2k$, e $C(0) \simeq 3/4$, como deve ser uma rede regular, onde

$$C = \frac{3k - 3}{4k - 2}, \quad (1.13)$$

que para o limite de $k \rightarrow \infty$ nos leva a $C \rightarrow 3/4$. O que se traduz em uma rede altamente agregada e um mínimo caminho que depende linearmente de N . No limite oposto, $p \rightarrow 1$, a rede transforma-se em uma rede aleatória, onde $\langle l \rangle \simeq \ln(N)/\ln(k)$, e $C \simeq 2k/N$.

No entanto, a Fig. (1.5) mostra resultados numéricos para o cálculo de $C(p)$ e $\langle l(p) \rangle$, onde se nota que existe um intervalo razoável de p no qual o mínimo caminho é quase tão pequeno quanto uma rede aleatória, porém, $C(p) \gg C_{ER}$, por exemplo.

Pode-se notar também que para valores muito pequenos de p a presença de poucos “defeitos”, oriundos da troca de conexões locais por atalhos, reflete-se de forma drástica em um

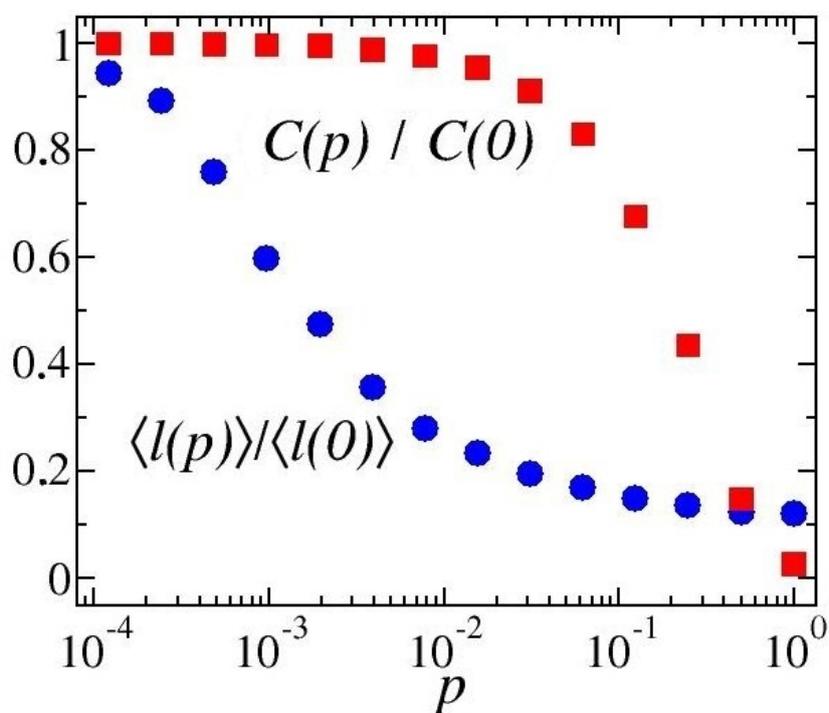


Figura 1.5: Mínimo caminho média característico $\langle l(p) \rangle$ e o coeficiente de agregação para o modelo WS. Os dados foram normalizados pelos valores de $\langle l(0) \rangle$ e $C(0)$. A comparação entre o comportamento do mínimo caminho médio e do coeficiente de agregação evidencia um intervalo significativo de p para o qual o fenômeno de mundo pequeno está presente, onde temos um elevado valor para $C(p)$ e valores relativamente pequenos para $\langle l(p) \rangle$. Assim, como no trabalho de Watts e Strogatz, utilizamos um tamanho de rede $N = 1000$ e $k = 10$. Foram realizadas 20 execuções do modelo WS.

comportamento extremamente não-linear de $\langle l(p) \rangle$. Em contrapartida, apenas quando a aleatoriedade da rede começa a ser dominante, temos alguma alteração considerável no valor do coeficiente de agregação, o que faz com que $C(p)$ seja praticamente constante para pequenos valores de p . A apresentação dos dados em escala logarítmica na Fig. (1.5) mostra o quão rápido $\langle l(p) \rangle$ decresce, e que durante essa variação intensa, $C(p)$ permanece praticamente inalterado, indicando que a transição de uma topologia regular para uma topologia cujo o fenômeno de mundo pequeno está presente é praticamente imperceptível em um nível local.

1.4 Barabási-Albert: Um modelo de rede livre de escala

O princípio de interpolação entre topologias regulares e aleatórias proposto por Watts e Strogatz no intuito de modelar redes reais sociais que apresentem o fenômeno do mundo pequeno mostrou-se bastante eficaz. No entanto, um número significativo de redes reais apresentam uma característica topológica importante que não está presente no modelo ER ou WS. Diversos sistemas parecem possuir uma distribuição de conectividade que decai na forma de uma lei de potência $P(k) \sim k^{-\gamma}$. Apesar de tal característica evidenciar que apenas a conectividade da rede é livre de escala, visto que é possível encontrar efeitos de escala em outras características presentes nessas redes, estas redes popularizaram-se como *redes livre de escala*.

O físico Derek de Solla Price descreveu em 1965 o que hoje conhecemos como a primeira rede livre de escala. Price estudou uma rede de citações de artigos científicos [17] e constatou que, para sua surpresa, tanto a conectividade de entrada como a conectividade de saída (número de vezes que o artigo foi citado e o número de demais publicações que o artigo cita) possui uma distribuição que decai em lei de potência. Anos depois, Price [18] forneceu o que hoje aceitamos como a explicação para leis de potência. Seus argumentos baseavam-se na idéia de que leis de potência surgem quando os “ricos ficam mais ricos”, o que ele chamou de *vantagem cumulativa*. Hoje, esse efeito é conhecido como *contato preferencial*, termo cunhado por R. Albert e A. L. Barabási[20] que propuseram um modelo capaz de reproduzir o efeito. Nessa seção focaremos nossa atenção no modelo de Barabási-Albert (**BA**).

1.4.1 O modelo do contato preferencial

Muitas redes reais crescem com o tempo, no sentido de que, o número de vértice N presente na rede não é um parâmetro fixo no processo de sua formação. Barabási e Albert propuseram que esta característica aliada a um possível contato preferencial durante o processo de crescimento seriam capazes de emergir a propriedade de invariância de escala na conectividade. O estudo de redes reais indica que a probabilidade de um vértice novo formar uma nova aresta com um vértice já existente na rede não é, de fato, aleatório, por exemplo, na WWW, quando uma página “entra no ar”, é muito provável que ela tente se ligar a uma outra página que seja bastante conhecida. O modelo BA é constituído por estas duas características básicas:

- (i) *Crescimento*: A rede deve ser construída a partir de uma quantidade inicial de vértices m_0 . Esses compõem um caroço inicial ao qual, a cada passo de tempo, é adicionado um novo vértice que acrescenta $m(\leq m_0)$ arestas que conectam o novo vértice aos demais presentes na rede.

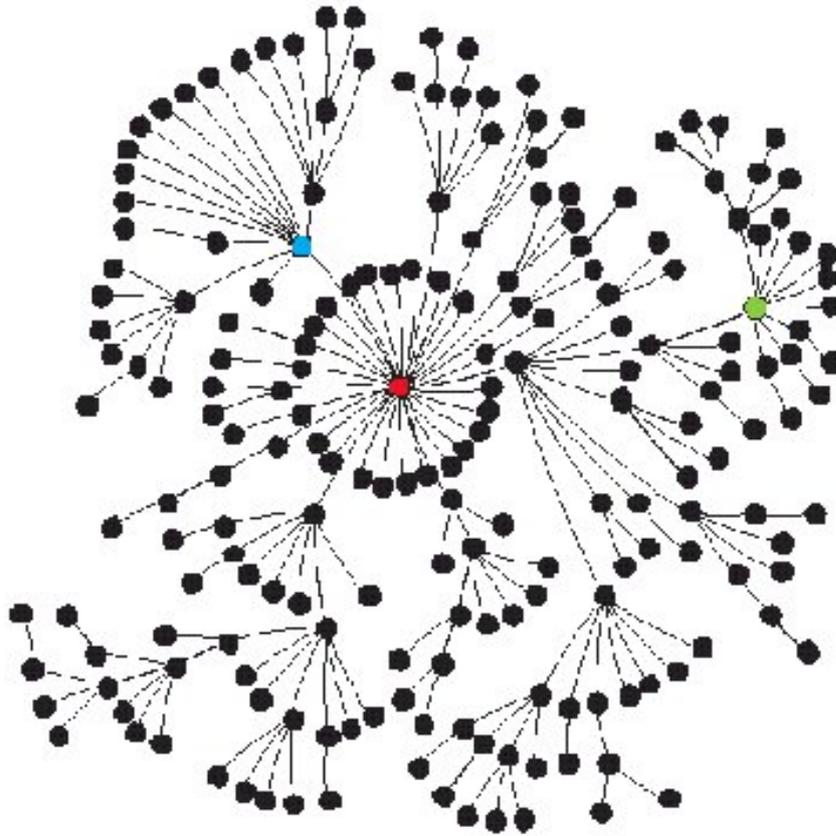


Figura 1.6: Uma rede livre de escala gerada pelo modelo de Barabási-Albert. Estão destacados os três vértices mais conectados da rede. Ilustração retirada de [4]

- (ii) *Contato preferencial*: Quando o novo vértice é adicionado, são escolhidos m vértices com probabilidade Π que depende da conectividade do vértice escolhido, ou seja,

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}, \quad (1.14)$$

onde i é o vértice escolhido para se receber a conexão com o novo vértice e k_i o valor de sua conectividade. A soma no denominador é efetuada sobre todos os vértices presentes na rede até o passo de tempo anterior.

Após t passos de tempo, este procedimento resulta em uma rede com $N = t + m_o$ vértices e mt arestas³. Este modelo alcança um estado estacionário onde a probabilidade de que um vértice i escolhido aleatoriamente possua conectividade k_i decai em lei de potência com um expoente $\gamma_{BA} = 3$. Uma característica interessante desse modelo é que este expoente de escala é independente de m , o único parâmetro, de fato, do modelo BA.

³A rede inicial formada por m_o vértices pode possuir uma quantidade qualquer de arestas dentro do intervalo $[0, m_o(m_o - 1)]$. No entanto podemos ignorar essa variável adicional sem falta de generalidade.

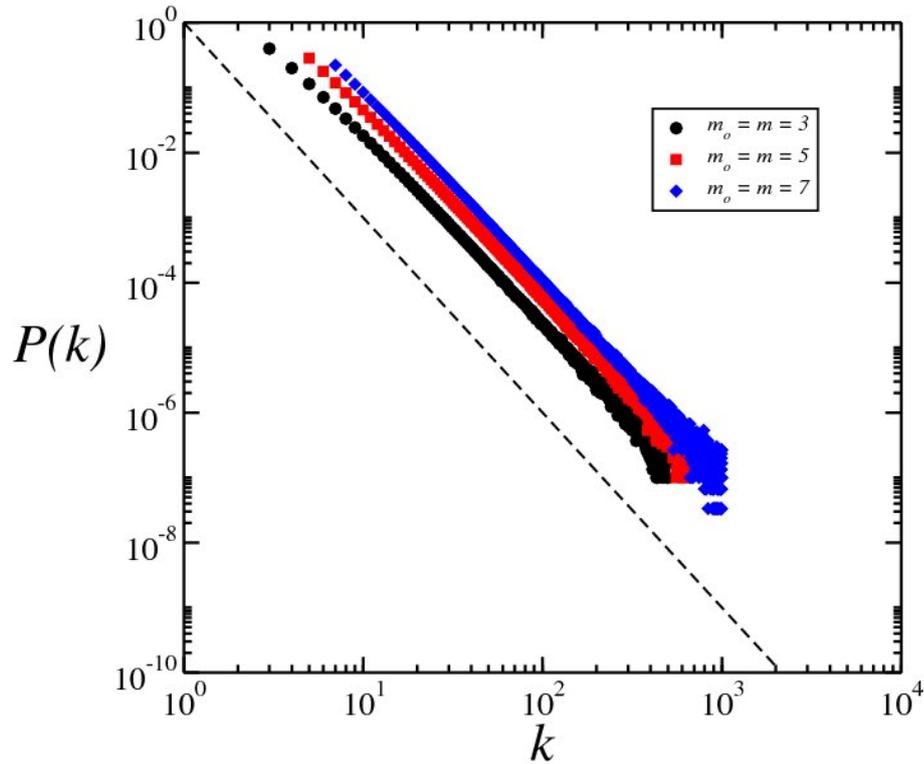


Figura 1.7: Distribuição de conectividade para o modelo BA. Pode-se ver que a regra de contato preferencial utilizado pelo modelo produz o efeito de escala livre na conectividade, evidenciado pelo comportamento e lei de potência da distribuição de conectividade. Na figura, todos os três conjuntos de dados são referentes a um mesmo tamanho de rede $N = m_o + t = 300000$. Foram utilizados os parâmetros $m_o = m = 3$ (círculos pretos), $m_o = m = 5$ (quadrados vermelhos), $m_o = m = 7$ (losangos azuis). A linha tracejada possui inclinação $\gamma = \gamma_{BA} = 3$.

1.4.2 Abordagem através do tratamento contínuo

Muitas propriedades dinâmicas do modelo BA podem ser obtidas através de um tratamento analítico. Barabási e Albert propuseram uma abordagem contínua para determinar a forma da distribuição de conectividade do modelo de contato preferencial proposto por eles. Nessa abordagem, assumimos que a conectividade de um vértice i pode ser considerada uma variável contínua no tempo, assim,

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_j k_j}. \quad (1.15)$$

Ou seja, a cada passo de tempo a conectividade k_i pode ser alterada caso o novo vértice seja conectado a i , portanto podemos tomar sua taxa de variação como sendo igual a probabilidade de que essa conexão ocorra vezes o número de arestas que o novo vértice está adicionando. Como soma no denominador e efetuada sobre todos os vértices presentes na rede, excetuando-

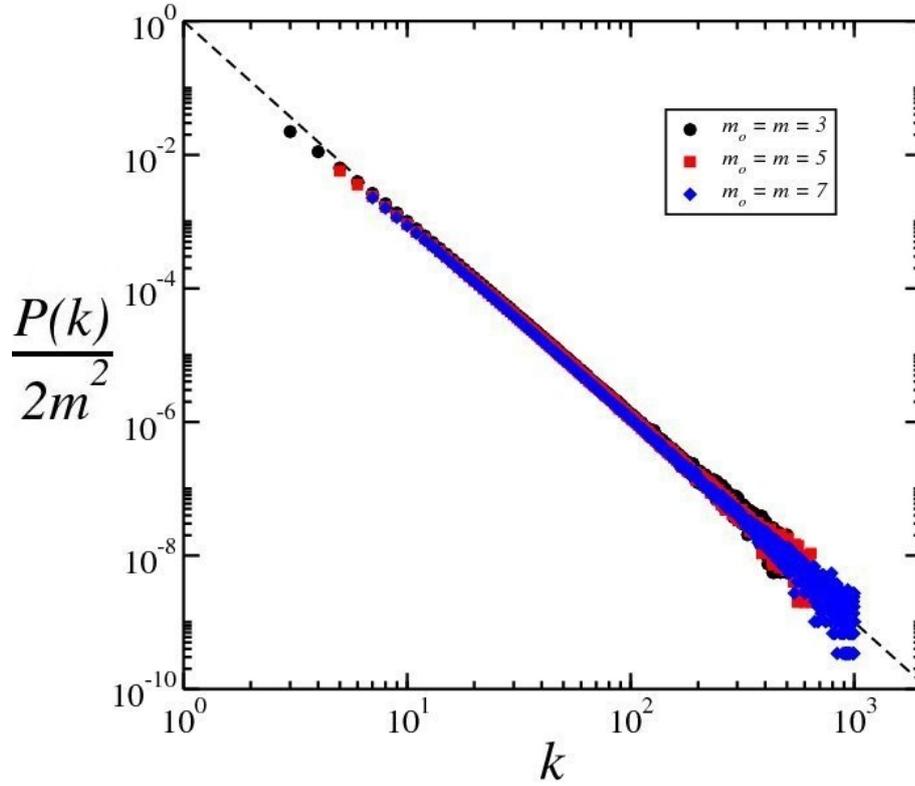


Figura 1.8: Distribuição de conectividade do modelo BA reescalado por um fator de $1/2m^2$. Como na Fig. (1.7), $N = 300000$. Foram utilizados os parâmetros $m_o = m = 3$ (círculos pretos), $m_o = m = 5$ (quadrados vermelhos), $m_o = m = 7$ (losangos azuis). A linha tracejada possui inclinação $\gamma = \gamma_{BA} = 3$.

se o que está sendo adicionado no instante t , sendo assim, temos

$$\sum_{j=1}^{N-1} k_{j=1} = 2m(t-1), \quad (1.16)$$

que no limite assintótico é simplesmente:

$$\sum_{j=1}^{N-1} k_{j=1} \simeq 2mt. \quad (1.17)$$

Portanto,

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t}. \quad (1.18)$$

Como, por condição inicial, temos que a conectividade de um vértice ao ser adicionado é $k_i(t_i) = m$, temos como solução da Eq. (1.18):

$$k_i(t_i) = m \left(\frac{t}{t_i} \right)^\beta, \quad (1.19)$$

onde $\beta = 1/2$.

A Eq. (1.19) diz-nos que a probabilidade de um vértice i possuir conectividade k_i menor do que um valor k é

$$p[k_i < k] = P \left[t_i > \frac{m^{1/\beta} t}{k^{1/\beta}} \right]. \quad (1.20)$$

Como a adição de vértices ocorre em intervalos iguais, temos que t_i possui densidade de probabilidade constante

$$\rho(t_i) = \frac{1}{m_o + t}. \quad (1.21)$$

A Eq. (1.20) diz-nos que um vértice i de conectividade k_i deve ter sido adicionado à rede em um instante t_i maior que $m^{1/\beta} t / k^{1/\beta}$. De posse da Eq. (1.21), temos que a probabilidade de um instante t_i ser maior do que $m^{1/\beta} t / k^{1/\beta}$ é dada por:

$$p \left[t_i > \frac{m^{1/\beta} t}{k^{1/\beta}} \right] = 1 - \frac{m^{1/\beta} t}{k^{1/\beta} (m_o + t)},$$

o que nos diz que

$$p[k_i < k] = 1 - \frac{m^{1/\beta} t}{k^{1/\beta} (m_o + t)}. \quad (1.22)$$

A distribuição de conectividade $P(k)$ pode ser obtida usando

$$P(k) = \frac{\partial p[k_i < k]}{\partial k},$$

resultando em

$$p(k) = \frac{2m^{1/\beta} t}{m_o + t} k^{-(1+1/\beta)}, \quad (1.23)$$

que para o limite assintótico $t \rightarrow \infty$ nos leva a

$$P(k) \sim 2m^{1-\beta} k^{-\gamma}. \quad (1.24)$$

Onde

$$\gamma = 1 + \frac{1}{\beta} = 3.$$

Assintoticamente, Eq. (1.23) diz-nos que a distribuição de conectividade é independente do tempo e conseqüentemente do tamanho da rede. Diz-nos também que $P(k)$ é proporcional a m^2 , como mostrado na Fig. (1.8).

1.4.3 Equação mestra da evolução da rede

Uma forma alternativa para demonstrar que, apesar de simples, o modelo BA leva a uma rede cuja a distribuição de conectividade decai em lei de potência, é a escrever a equação mestra

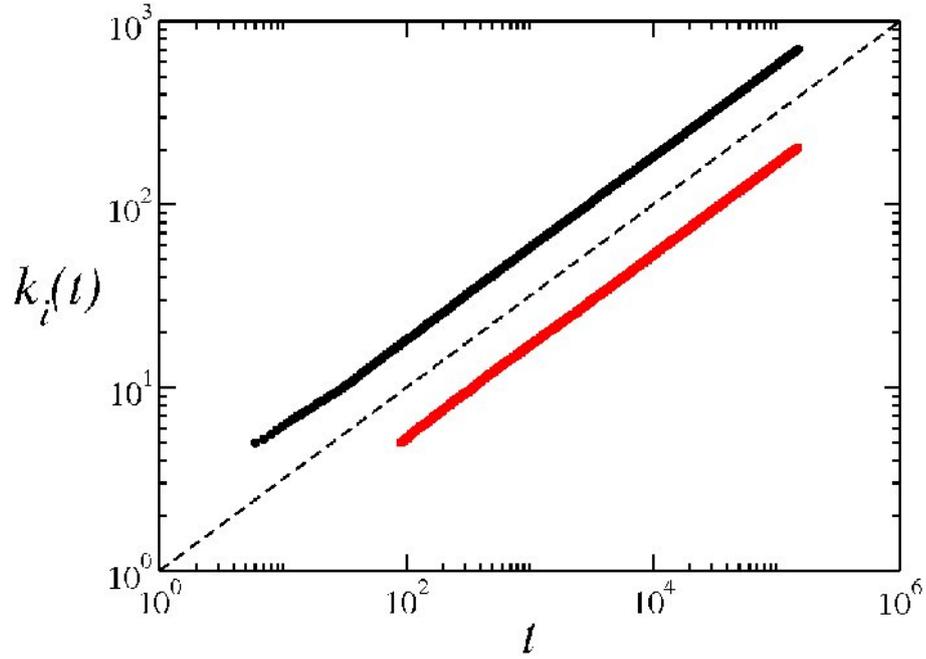


Figura 1.9: Evolução temporal de dois vértices da conectividade de dois vértices escolhidos arbitrariamente. O primeiro foi adicionado a rede em um instante $t_1 = 5$ (círculos pretos), o segundo no instante $t_2 = 92$ (quadrados vermelhos). A linha tracejada possui inclinação $\beta = 0.5$, como previsto pela Eq.(1.19).

para a evolução do número de vértices cuja a conectividade é k [31]. Assim, devemos estudar a probabilidade $P(k, N)$ de em um dado passo de tempo onde o número de vértices é N , um dado vértice de conectividade k pode ter o valor de sua conectividade alterada com a inserção de um novo vértice. Podemos escrever a probabilidade de uma das novas arestas serem conectadas a um vértice de conectividade k no instante em que a rede possui N vértices como sendo

$$\frac{kP(k, N)}{\sum_k P(k, N)} = \frac{kP(k, N)}{2m}. \quad (1.25)$$

Como o novo vértice adiciona m novas arestas, a probabilidade de um vértice de conectividade k receber uma dessas é $m \times kP(k, N)/2m = \frac{1}{2}kP(k, N)$. Assim, a equação mestra da mudança do número de vértices presentes na rede para o modelo BA possui a forma:

$$(N + 1)P(k, N + 1) - NP(k, N) = \frac{1}{2}(k - 1)P(k - 1, N) - \frac{1}{2}kP(k, N), \quad (1.26)$$

para $k \geq m + 1$, ou

$$(N + 1)P(m, N + 1) - NP(m, N) = 1 - \frac{1}{2}kP(k, N), \quad (1.27)$$

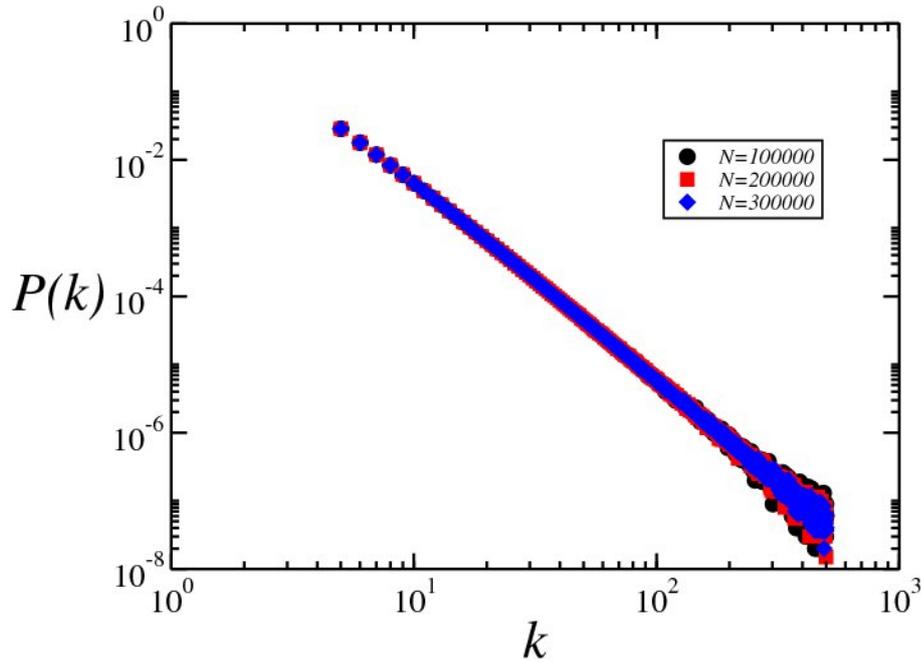


Figura 1.10: Distribuição de Conectividade para o modelo BA, com $m_o = m = 5$ para vários tamanhos de rede diferentes. Vemos que para o limite assintótico do modelo, $t \rightarrow \infty$, o comportamento em lei de potência da distribuição de conectividade independe do tamanho do sistema. No gráfico $N = 100000$ (círculos pretos), $N = 200000$ (quadrados vermelhos), $N = 300000$ (losangos azuis).

para $k = m$. Não há nenhum vértice com conectividade $k < m$.

Para o regime estacionário de $t \rightarrow \infty$, $P(k, N+1) = P(k, N) = P(k)$, assim

$$P(k) = \frac{(k-1)}{(k+2)} P(k-1), \quad (1.28)$$

para $k \geq m+1$, e

$$P(m) = \frac{2}{(m+2)}, \quad (1.29)$$

para $k=m$.

Através da relação de recursividade entre Eq. (1.28) e Eq. (1.29), temos

$$P(k) = \frac{(k-1)(k-2)\dots m}{(k+2)(k+1)\dots(m+3)} P(m) = \frac{2m(m+1)}{k(k+1)(k+2)}. \quad (1.30)$$

No limite de k muito grande, teremos uma distribuição de conectividade em lei de potência $P(k) \sim k^{-3}$.

Recentemente, M. V. Simkin e V. P. Roychowdhury utilizaram uma abordagem semelhante para calcular um valor estimado para o número de pesquisadores que realmente lêem os artigos que eles citam em seus trabalhos [32]. É bem conhecido que a distribuição de conectividade para uma rede de citação em artigos científicos segue uma lei de potência com expoente

~ 3.0 [33]. Eles estimaram que apenas 20% dos citantes realmente leram o artigo científico original, concluindo que erros nas citações não devem ser consideradas como apenas um mero acontecimento aleatório, e sim devido a possível cópia da citação de demais artigos.

1.5 O modelo de Kleinberg

Como vimos anteriormente, redes sociais exibem o fenômeno de mundo pequeno. Desde o resultado apresentados por Milgram, diversos modelos foram propostos no intuito de reproduzir o fenômeno. O modelo WS foi o primeiro a ser bem sucedido, servindo de paradigma para novos modelos bem sucedidos. No entanto, nenhum dos modelos foi capaz de indicar o motivo de por que dois vértices arbitrários, em redes sociais reais, são capazes de encontrar caminhos curtos que os conecte. Kleinberg mostrou que os modelos propostos até então eram insuficientes para explicar porque pessoas em redes sociais acham esses caminhos.

Kleinberg [1, 2] estudou *algoritmos descentralizados* no intuito de reproduzir o experimento de Milgram. Dentro dessa classe de algoritmos, os vértices presentes na rede possuem apenas conhecimento local, ou seja, eles conhecem apenas a posição na rede daqueles vértices aos quais estão conectados. Assim, Kleinberg propõe um família de redes aleatórias que naturalmente generalizam o modelo de WS, mostrando que apenas um modelo dentro dessa família é realmente eficiente.

1.5.1 A rede

O modelo de Kleinberg é inspirado no paradigma apresentado por Watts e Strogatz, ou seja, uma rede rica em conexões locais, e com a presença de algumas poucas ligações de longo alcance. No entanto, ao invés de utilizar uma rede circular como substrato para a rede, Kleinberg optou por uma rede regular quadrada. Temos com isso um conjunto de vértices (que são os indivíduos das redes sociais), que são identificados por um conjunto de pontos em quadrado $L \times L$, $\{u = (i, j) : i \in \{1, 2, \dots, L\}, j \in \{1, 2, \dots, L\}\}$. Assim, a distância entre dois vértices $u = (i, j)$ e $v = (m, n)$ será:

$$r_{u,v} = |m - i| + |n - j|. \quad (1.31)$$

Podemos notar que, geograficamente, a distância aqui é medida como a “distância Manhattan”, conhecida assim devido a disposição dos quarteirões da ilha Manhattan, na cidade Nova Iorque, EUA. No modelo de Kleinberg são definidos três parâmetros universais:

- $p \geq 1$ - O vértice u possui aresta direcionada com todos os vértices presentes na rede a

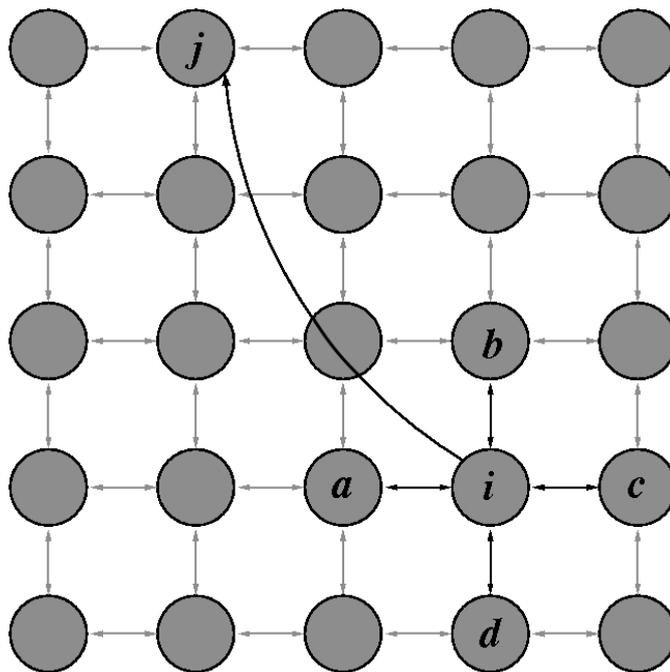


Figura 1.11: Conexões de um vértice i . Cada vértice i possui ligação com todos os seus vizinhos a uma distância $p = 1$ (a , b , c e d), e uma aresta com um vértice escolhido aleatoriamente j . O vértice j é escolhido com probabilidade $P(r_{i,j}) \sim r_{i,j}^{-\alpha}$.

uma distância Manhattan p . Estes são os contatos locais de u .

- $q \geq 0$ e $\alpha \geq 0$ - Também são adicionados, de forma aleatória, q arestas direcionadas entre u e v (contato de longo alcance). A ligação aleatória entre u e v é adicionada com probabilidade $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$.

Este modelo possui uma interpretação geográfica bastante simples: indivíduos presentes em uma rede com correlações espaciais conhece todos os seus vizinhos até uma certa distância p , e alguma pessoas distribuídas através da rede. A distribuição de probabilidade em lei de potência $P(r) \sim r^{-\alpha}$, diz-nos que é mais fácil conhecer pessoas mais próximas geograficamente do que pessoas mais distantes na rede. Se olharmos para p e q como constantes, o modelo de Kleinberg possui um parâmetro α que define toda uma família de redes aleatórias. Quando $\alpha = 0$, o modelo de Kleinberg é equivalente ao modelo WS, onde a escolha da nova aresta é independente da posição na rede do novo contato de u . À medida que α aumenta, os contatos de longo alcance de u tornam-se cada vez mais agregados em sua vizinhança.

1.5.2 O algoritmo descentralizado

O algoritmo que compõe o modelo de Kleinberg é fortemente inspirado pelo experimento idealizado por Milgram. O algoritmo descentralizado proposto por Kleinberg tem por objetivo calcular quantos passos são necessários para que uma mensagem enviada por vértice fonte, que chamaremos de f , leva para chegar a um vértice alvo a . Ambos os vértices são escolhidos de forma arbitrária. A cada passo, o portador da mensagem u deve escolher entre um de seus contatos, sejam eles os contatos locais ou de longo alcance, um vértice v que será o novo portador da mensagem. Essa escolha deve levar em consideração a distância $r_{v,a}$ entre o vértice v e o vértice alvo a . Esse ponto do algoritmo reproduz um dos principais ingredientes do experimento de Milgram: o portador da mensagem deve escolher um contato o qual ele acredita irá otimizar o envio da mensagem.

Durante esse processo portador u possui apenas conhecimento local da rede, ou seja:

- (i) O substrato geográfico da rede (a rede quadrado).
- (ii) A sua posição na rede de a .
- (iii) A posição na rede de seus contatos locais e longo alcance, assim como a posição na rede e os contatos de longo alcance de todos os vértices que tocaram na mensagem.

Uma característica importante do algoritmo é que u não conhece os contatos de longo alcance dos vértices que não tocaram na mensagem. Assim, o *valor esperado do tempo de envio* da mensagem é o valor esperado do número de passos necessários para que o algoritmo descentralizado envie a mensagem através da rede gerada pela distribuição $P(r)$ de um vértice fonte f a um vértice alvo a escolhidos uniformemente dentro do conjunto de vértices da rede.

1.5.3 Limites para o valor esperado do tempo de envio da mensagem

Kleinberg estudou a forma de como a estrutura da rede pode interferir em um algoritmo descentralizado que tem por objetivo encontrar caminhos curtos entre dois vértices arbitrários. Quando o parâmetro $\alpha = 0$, os contatos de longo alcance são gerados de forma independente da geometria da rede, podemos usar argumentos da teoria de grafos aleatórios para provar que existem caminhos ligando todos os pares de vértices da rede e que esses caminhos estão limitados por um polinômio $O(\log L)$, ou seja, os mínimos caminhos entre os pares de vértices crescem com o tamanho da rede de forma logarítmica, assim como o modelo WS. Essa característica é esperada, visto que Kleinberg formulou sua rede inspirado pelo paradigma proposto

por Watts e Strogatz. No entanto, como será mostrado a seguir, vértices possuidores apenas de conhecimento local, operando segundo o algoritmo descrito na seção anterior, são incapazes de encontrar tais caminhos. Para analisarmos o problema de forma mais geral, consideraremos uma rede quadrada d -dimensional.

Pelo modelo proposto, sabemos que a probabilidade que o vértice u escolha v como seu contato de longo alcance é dada por:

$$P(r_{u,v}) \sim [r_{u,v}]^{-\alpha}, \quad (1.32)$$

que deve ser normalizada por um fator de $\sum_{v \neq u} r_{u,v}^{-\alpha}$, que escala na forma:

$$\begin{aligned} \sum_{v \neq u}^{\infty} r_{u,v}^{-\alpha} &\sim \int_1^L x^{-\alpha} x^{(d-1)} dx \\ &= \begin{cases} (\alpha - d)^{-1} & , \text{ se } \alpha > d, \\ \ln L & , \text{ se } \alpha = d, \\ L^{d-\alpha} & , \text{ se } \alpha < d. \end{cases} \end{aligned} \quad (1.33)$$

Para $0 < \alpha < d$, vamos considerar uma região de raio R arbitrário em torno do vértice alvo a , tal que $R = L^\delta$. A Eq. (1.33) nos diz que a probabilidade de um vértice u possuir contato de longo alcance com um vértice v é

$$P(r_{u,v}) \sim \frac{r_{u,v}^{-\alpha}}{L^{d-\alpha}} \leq \frac{1}{L^{d-\alpha}}. \quad (1.34)$$

Assim, a probabilidade de um vértice escolhido aleatoriamente conectar qualquer um dos vértices contidos dentro da casca definida pelo raio R deve respeitar a relação

$$\begin{aligned} P(r_{u,v} \mid r_{a,v} < R) &\leq \frac{R^d}{L^{d-\alpha}} \\ &\leq L^{\delta d - d + \alpha}. \end{aligned} \quad (1.35)$$

Como $\delta < 1$, e $R = L^\delta$, quando temos o limite de L muito grande, R crescerá muito mais lentamente que L , o que indica que o vértice fonte f muito provavelmente estará fora da casca. Esse fato leva a duas conclusões, primeiro, qualquer caminho entre f e a deverá ter pelo menos uma ligação de longo alcance para dentro da casca. Segundo, O valor esperado para o número de passos necessários para a mensagem ser entregue é limitado a ser, no mínimo, igual a R , pois a ligação de longo alcance pode levar a mensagem para dentro da casca em uma região próxima a borda da casca de raio R , assim no caso onde a mensagem não seja enviada através de nenhuma ligação de longo alcance, ela deverá percorrer toda a distância R da casca.

A probabilidade de um vértice com uma conexão desse tipo ser encontrado dentro de pelo

menos R passos é menor que $R \times L^{\delta d - d + \alpha}$. Se essa probabilidade anular-se no limite $L \rightarrow \infty$ então o valor esperado do número de passos também será limitado a ser no mínimo R . Para que essa probabilidade anule-se é necessário que:

$$\begin{aligned} 0 &< R \times L^{\delta d - d + \alpha} < 1, \\ 0 &< L^{\delta(d+1) - d + \alpha} < 1, \end{aligned}$$

que implica em:

$$\begin{aligned} \delta(d+1) - d + \alpha &< 0, \\ \delta &< \frac{(d-\alpha)}{(1+d)}. \end{aligned} \quad (1.36)$$

Portanto, o número esperado de passos não pode ser menor que $L^{(d-\alpha)/(1+d)}$ para o regime de $0 < \alpha < d$.

Para $\alpha > d$, a probabilidade de termos uma ligação maior que L^γ ($0 < \gamma < 1$), de acordo com Eq. (1.33) deve escalonar com o tamanho do sistema na forma:

$$\int_{L^\gamma}^{\infty} \frac{x^{-(\alpha-d+1)}}{\alpha-d} dx \sim L^{\gamma(d-\alpha)}. \quad (1.37)$$

De modo que, a probabilidade de que a mensagem dê um passo maior do que L^γ , em um número de L^β ($0 < \beta < 1$) é menor do que $L^\beta L^{\gamma(d-\alpha)}$. No entanto, essa probabilidade anula-se no limite $L \rightarrow \infty$, o que faz com que a distância total percorrida pela mensagem durante o processo de envio não supere o valor de $L^{\beta+\gamma}$. Porém, o valor da distância que separa o vértice fonte e o vértice alvo deve ser proporcional a L , assim precisamos que

$$\beta + \gamma = 1. \quad (1.38)$$

Por outro lado, temos que a probabilidade de conexões maiores que L^γ existam desaparecem quando $\beta + \gamma(d - \alpha) < 0$. Estas duas condições nos fornecem:

$$\beta < \frac{(\alpha - d)}{(\alpha - d + 1)}, \quad (1.39)$$

de modo que, o valor esperado do tempo de envio da mensagem não deve ser menor que $L^\beta \sim L^{(\alpha-d)/(\alpha-d+1)}$.

Para o valor singular de $\alpha = d$, podemos elaborar a seguinte análise: o vértice alvo a está rodeado por um número de m cascas de raio $e^{m-1} < R < e^m$, $m = 1, 2, \dots$. Agora, suponha que o portador u da mensagem esteja na casa m , a probabilidade de que u possua uma conexão com

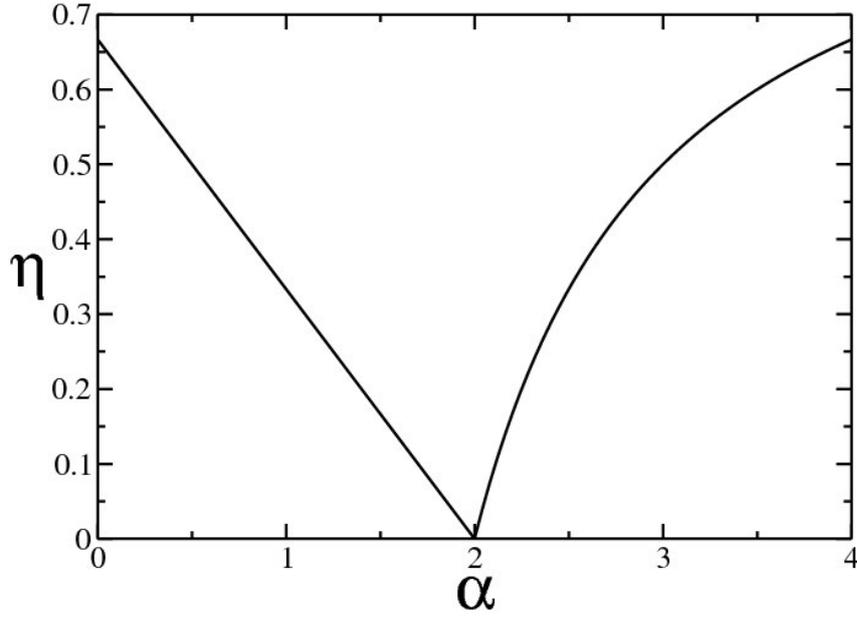


Figura 1.12: Expoente η dos limites inferiores do valor esperado para o tempo de envio da mensagem, EX , como função do coeficiente de agregação α . Na figura, assim como em Kleinberg [1], $d = 2$. Podemos verificar claramente que o expoente dos limites de EX possui um mínimo para $\alpha = d$, como previsto analiticamente. Para o valor singular $\alpha = d$, o limite de EX cresce com $(\ln L)^2$, mais lentamente do que uma potência de L .

algum vértice v na casca $m - 1$ é, de acordo com a Eq. (1.33),

$$P(r_{u,v}) \sim \int_{e^{m-1}}^{e^m} \frac{y^{-1}}{\ln L} dy = \frac{1}{\ln L}. \quad (1.40)$$

A probabilidade de se alcançar a próxima casca $m - 1$ em mais do que x passos é $p(x) = (1 - 1/\ln L)^x$, e o número médio de passos necessários para fazê-lo é

$$\langle x \rangle = \int_0^\infty p(x) dx \sim \frac{1}{\ln L}. \quad (1.41)$$

Como a maior casca é aquela para qual $e^m = L$, o número de cascas entre o vértices fonte e alvo deve ser $O(\ln L)$, portanto o valor esperado para o tempo de envio da mensagem não deve superar um valor $O((\ln L)^2)$.

Analicamente, verificamos que o valor esperado para o tempo de envio da mensagem EX de um vértice fonte f a um vértice alvo a , arbitrários, possuem limites na forma L^η , excluindo um valor singular $\alpha = d$ para o qual teremos um comportamento limitado superiormente por $(\ln L)^2$, que cresce mais lentamente que as potências de L encontradas acima. A Fig. (1.12) mostra o comportamento do expoente η como função de α para $d = 2$.

A Fig. (1.13) apresenta a realização do algoritmo descentralizado para três diferentes tamanhos de rede com $d = 2$. Podemos notar que, com o aumento da rede, existe um mínimo

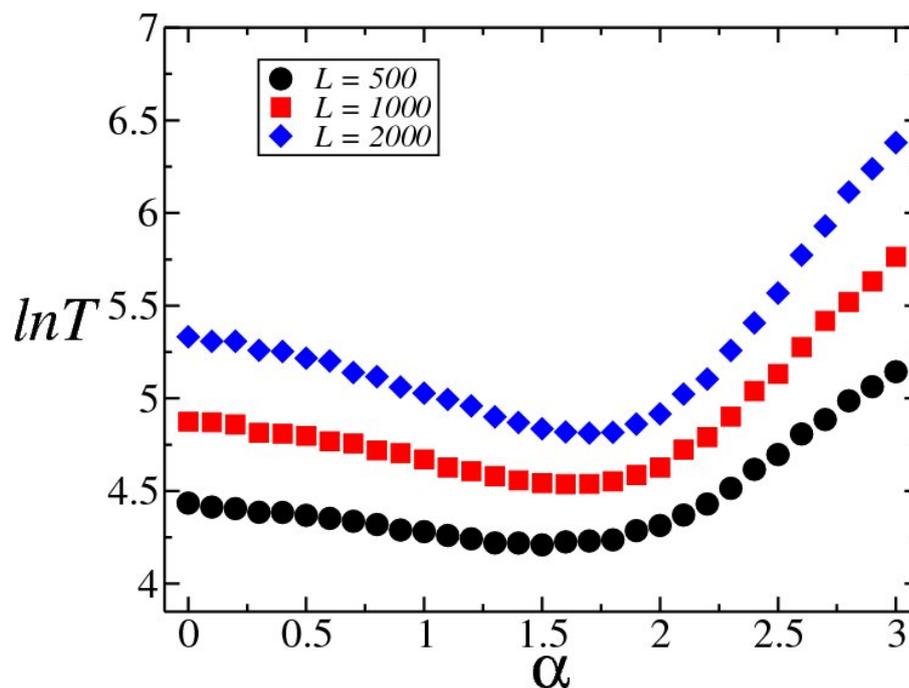


Figura 1.13: O mesmo que Kleinberg em [2]. O logaritmo de T , número de passos necessários para que a mensagem seja enviada do vértice fonte f ao vértice alvo a utilizando o algoritmo descentralizado de Kleinberg, é plotado como uma função de α . Assim como Kleinberg encontramos um mínimo que se aproxima de 2 com o aumento do sistema. As diferenças com Fig. 1c. de [2] são devidas ao tamanho de sistemas utilizados.

aproximando-se do valor previsto por Kleinberg de $\alpha = 2$. Esse comportamento pode ser visto em Kleinberg [2]. M. R. Roberson e D. ben-Avrahan [34] mostraram que este expoente *efetivo* é menor do que o limite idealizado por uma quantidade $O(1/(\ln L)^2)$, o que concorda com o resultado apresentado na Fig. (1.13).

O modelo de Kleinberg leva-nos a acreditar que as correlações entre a estrutura local e as conexões de longo alcance em redes de mundo pequeno fornecem “pistas” fundamentais para que os vértices da rede sejam capazes de encontrar caminhos mais eficazes em conectá-los. Essa correlação possui um limite em $\alpha = d$ onde a estrutura dos contatos de longo alcance forma um tipo de “gradiente” que permite com que os vértices encontrem tais caminhos. Para valores de α acima desse regime crítico, a rede torna-se mais homogênea o que faz com que as pistas oriundas da correlação desapareçam. Percebe-se também que, para uma distribuição homogênea, esses *atalhos* não deixam de existir, porém, o algoritmo descentralizado, baseado apenas em informação local, torna os vértices incapazes de encontrá-los com eficiência.

2 *Redes espacialmente correlacionadas: navegação global e efeitos de custo*

Como vimos anteriormente, existem várias características importantes na teoria de redes complexas e também diversos tipos diferentes de redes. No entanto, a comunidade científica tem dedicado sua atenção principalmente a quatro tipos de redes: Redes regulares, redes aleatórias, redes de “mundo pequeno” e redes livres de escala. Redes regulares são aquelas constituídas por vértices idênticos de conectividade constante, como a rede quadrada e a rede de Bethe, ambas amplamente conhecidas e utilizadas na Física computacional e teórica. No capítulo anterior, abordamos o modelo de Erdős-Rényi para redes aleatórias o qual possibilita vários resultados analíticos devido sua distribuição de conectividade poissoniana. Vimos também o modelo de Watts-Strogatz para redes de mundo pequeno que determina um paradigma para modelos de redes de mundo pequeno como sendo uma interpolação entre redes regulares e redes aleatórias. Como exemplo de redes livres de escala, apresentamos uma pequena revisão sobre o modelo de Barabási-Albert e suas propriedades emergentes do crescimento regido pela regra do contato preferencial.

Em nenhum desses casos, a configuração espacial da rede é considerada. Usualmente, assume-se que vínculos espaciais podem ser ignorados e que abordagens de *campo médio* podem ser utilizadas. Contudo, diversas redes como a Internet, redes de rotas aéreas, de transporte e de comunicação sem fios estão embutidas em um espaço $2D$, ou até mesmo $3D$. Tais exemplos de redes podem possuir vínculos espaciais importantes para a modelagem do problema real, de modo que uma aproximação de campo médio poderá por ser descartada.

O modelo proposto por Kleinberg para navegação de informação apresentado no capítulo anterior, além de abrir a discussão de modelos de rede de mundo pequeno através de uma abordagem algorítmica fornece o que talvez seja o primeiro modelo de rede em que os vínculos espaciais são considerados como um componente fundamental do problema. Permitir vínculos espaciais, como os estabelecidos no modelo de Kleinberg, tornam a rede resultante difíceis de serem estudadas, porém, permitem-nos trabalhar com modelos mais próximos das redes reais.

Neste capítulo abordaremos o problema de navegação com conhecimento global em redes

construídas a partir do modelo de Kleinberg. Para realizarmos essa tarefa, faremos uma analogia entre o tempo de viagem de uma mensagem na rede como o caminho por ela percorrido. Com isso, para o processo de navegação com conhecimento global, temos que o viajante, ou o portador da mensagem, conhece toda a configuração da rede, a posição de todos os vértices da rede e a estrutura de todas as conexões de longo alcance. Assim, podemos dizer que o caminho percorrido pela mensagem é equivalente ao mínimo caminho entre o vértice fonte f e o vértice alvo a . Essa análise nos permitirá dizer que, para determinados valores do expoente de agregação, o modelo de Kleinberg não apresenta o fenômeno do mundo pequeno. Constataremos esse fato através da aplicação de modelo epidêmico *suscetível-infectado-suscetível*. Em seguida apresentaremos um modelo cuja adição de contatos de longo alcance na rede, é limitada por um vínculo τ de custo, ou seja, iremos introduzir um limite total para a soma dos comprimentos das conexões de longo alcance adicionadas. Estudaremos o efeito desse custo τ na estrutura da rede, através da análise da dinâmica do processo de navegação de Kleinberg e de navegação global, a segunda novamente por meio do cálculo do mínimo caminho $\langle l \rangle$.

2.1 Mínimo caminho na rede de Kleinberg

O algoritmo descentralizado apresentado por Kleinberg aborda problema de navegação de informação em um modelo de rede de mundo pequeno [1, 2] em um processo que considera apenas informações locais. Como vimos no capítulo anterior, podemos considerar uma rede quadrada d -dimensional com $N = L^d$ vértices. Kleinberg considerou que cada vértice u da rede estaria conectado com todos os seus vizinhos em um perímetro $p \geq 0$ e poderia ainda receber um número $q \geq 0$ de contatos de longo alcance. A probabilidade de que u possua um contato de longo alcance com um vértice v , escolhido aleatoriamente, é proporcional a $r_{u,v}^{-\alpha}$, onde $r_{u,v}$ é a distancia na rede entre os vértices, e α é o expoente de agregação. Podemos ver que para $\alpha = 0$, temos uma distribuição uniforme, assim como no modelo WS, onde as conexões de longo alcance são geradas, de forma independente das características “geográficas” da rede. Quando $\alpha > 0$, os contatos de longo alcance de u tornam-se cada vez mais agregados em torno de sua vizinhança.

Nessa seção, estudaremos o comportamento do mínimo caminho médio $\langle l \rangle$ da rede de Kleinberg, com isso estaremos analisando um problema de navegação de informação através da rede onde o conhecimento global da rede é considerado. Para isso, abordaremos o problema em redes similares às que Kleinberg estudou com dimensão $d = 2$. No primeiro modelo, consideraremos que cada vértice da rede está conectado aos seus quatro vizinhos no perímetro de $p = 1$, e adicionaremos a apenas 40% dos vértices um contato de longo alcance direcionado. O

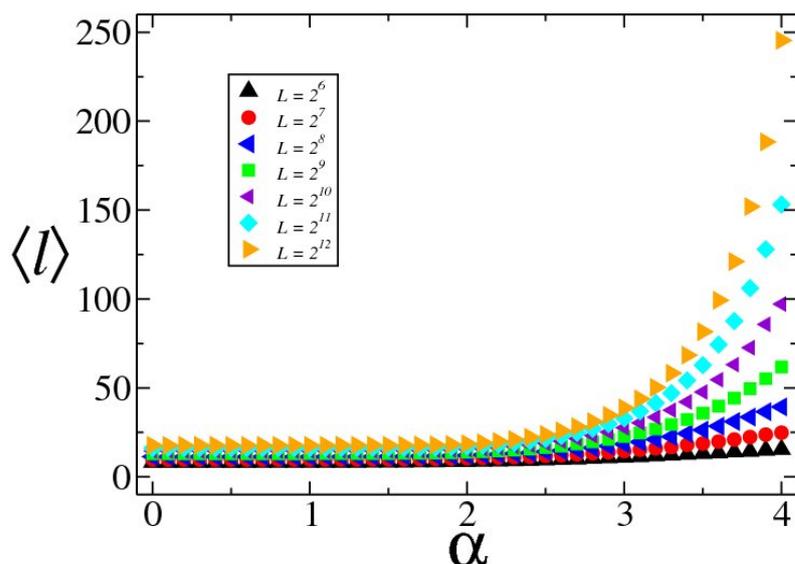


Figura 2.1: Mínimo caminho médio $\langle l \rangle$ como uma função do expoente de agregação α . Não fizemos qualquer restrição aos contatos de longo alcance adicionados a 40% dos vértices da rede. Encontramos que o comprimento do caminho ótimo é alcançado em $\alpha_e = 0$. Com o aumento do valor de α o comprimento das ligações de longo alcance diminuem em média, e como efeito, o mínimo caminho médio $\langle l \rangle$ aumenta. Foram utilizados diversos tamanhos de rede.

cálculo de $\langle l \rangle$ é efetuado seguindo a relação:

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{u \neq v} l_{u,v}, \quad (2.1)$$

onde $l_{u,v}$ é o comprimento do menor caminho que conecta u a v .

A relação entre uma navegação com conhecimento global da informação presente na rede e $\langle l \rangle$, justifica-se a partir do momento que com conhecimento global, todos os vértices “sabem” qual o menor caminho entre ele e os demais vértices presentes na rede. A Fig. (2.1) nos mostra o cálculo de $\langle l \rangle$ para diversos tamanhos de rede, de acordo com a Eq. (2.1). Diferentemente do que é observado para o algoritmo descentralizado de Kleinberg, onde o portador da mensagem possui apenas conhecimento local, e tínhamos um *expoente de agregação efetivo* $\alpha_e = 2$, em um sistema bidimensional encontramos um valor diferente de expoente de agregação efetivo $\alpha_e = 0$ quando calculamos $\langle l \rangle$. A presença de uma distribuição uniforme para $\alpha = 0$ possibilita uma maior frequência de contatos de longo alcance v mais distantes de u , ou seja, podemos pensar que os atalhos na rede serão maiores com maior frequência para $\alpha = 0$ do que para $\alpha > 0$. Com isso, sendo o processo de envio da mensagem baseado em um algoritmo de informação global, o vértice portador da mensagem tirará maior proveito da distribuição $P(r_{u,v}, \alpha = 0) \sim [r_{u,v}]^{-\alpha}$.

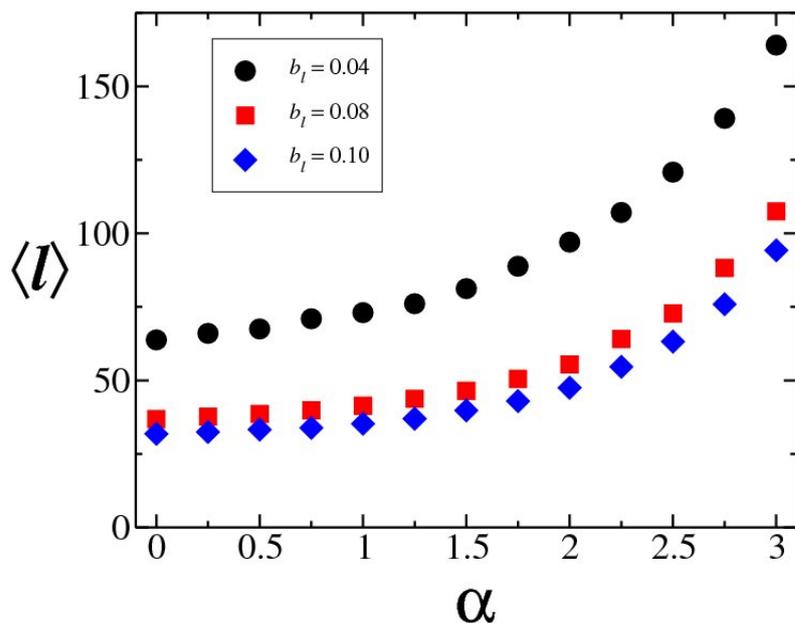


Figura 2.2: $\langle l \rangle$ como função de α . Quando restringimos dois vértices quaisquer na rede, u e v , a possuírem conexões apenas dentro de uma pequena região $r_{u,v} \leq b_l \times L$, encontramos um expoente de agregação efetivo $\alpha_e = 0$. O comportamento de $\langle l \rangle$ é o mesmo para quando não existe qualquer limitação para tamanho das ligações. Aqui utilizamos apenas um tamanho de rede $L = 1000$, e adicionamos contatos de longo alcance a 20% dos vértices presentes na rede.

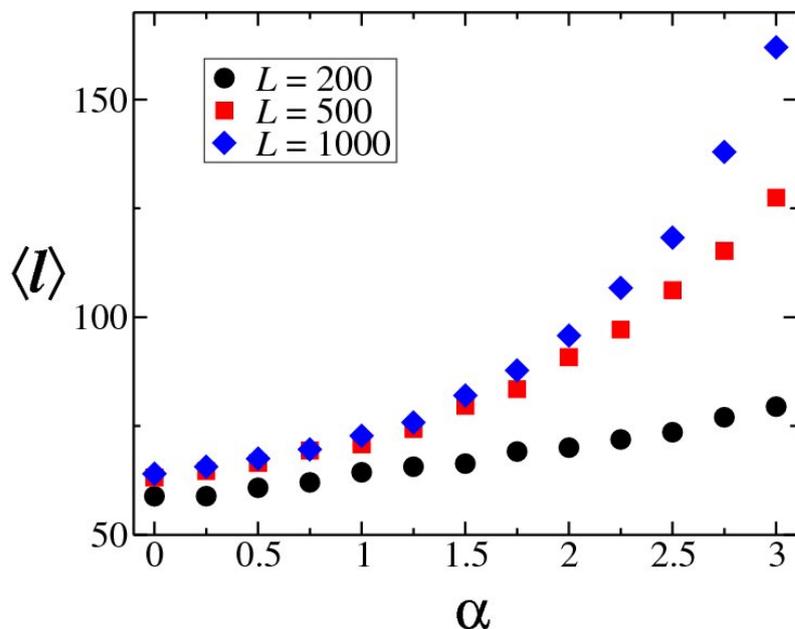


Figura 2.3: $\langle l \rangle$ como função de α . Os vértices da rede estão restritos a uma pequena região da rede $b_l \times L$ novamente. Para um mesmo valor $b_l = 0.04$, encontramos, para três tamanhos de rede diferentes, o mesmo valor $\alpha_e = 0$. Mais uma vez, adicionamos contatos de longo alcance apenas para 20% dos vértices.

Esse resultado não é alterado quando impomos uma restrição no tamanho dos atalhos. Por exemplo, podemos restringir os vértices u e v a estarem restritos a uma pequena região de raio $b_l \times L$ ($0 \leq b_l \leq 1$); em outras palavras, $r_{u,v} \leq b_l \times L$. A Fig. (2.2) mostra o cálculo de $\langle l \rangle$ para diferentes valores de b_l para redes com $L = 1000$, enquanto que na Fig. (2.3) fixamos o valor de $b_l = 0.04$ e realizamos o cálculo de $\langle l \rangle$ para três tamanhos de rede diferentes. Em todos os casos encontramos $\alpha_e = 0$. Assim podemos concluir que a navegação de mensagens baseadas em processos nos quais os vértices da rede são permitidos a conhecerem toda a configuração da rede, ou seja, não apenas as posições de cada vértice da rede mas também a posição dos contatos de longo alcance de cada um deles. Este resultado independe do percentual de vértices que recebem contatos de longo alcance (os resultados apresentados Fig. (2.1) foram obtidos para um caso onde 40% dos vértices recebem tais contatos, no entanto, a Fig. (2.2) e a Fig. (2.3) apresentam realizações onde apenas 20% dos vértices são permitidos a possuírem contatos de longo alcance), independem de qualquer restrição feita ao tamanho dos atalhos adicionados a rede (Fig. (2.2)) e independem do tamanho do sistema (Fig. (2.3)).

Kleinberg mostrou que, para redes sociais, indivíduos são incapazes de encontrar caminhos eficazes entre eles quando os contatos de longo alcance são gerados de forma uniformemente aleatória ($\alpha = 0$), que representa ambiente onde atalhos de longo alcance existem. No entanto, os resultados aqui obtidos nos permitem concluir que a distribuição uniforme, assim como proposto no modelo WS para redes de mundo pequeno, produz uma rede onde os vértices, desde que permitidos a possuírem conhecimento global da rede, são capazes de enviar mensagens com eficácia maior do que em qualquer outro modelo de rede que pertença à família de redes definida por Kleinberg. A Fig. (2.1) mostra também que para valores de $\alpha \leq 2$, teremos $\langle l \rangle$ praticamente constante independentemente do tamanho da rede. Os valores apresentados para $\langle l \rangle$ são pequenos comparados ao tamanho das redes. No entanto, podemos perceber claramente que para valores de $\alpha > 2$ o comportamento de $\langle l \rangle$ torna-se extremamente dependente do tamanho da rede. Podemos perceber que para o limite termodinâmico $L \rightarrow \infty$ obteremos valores extremamente altos para $\langle l \rangle$ de modo que a rede perde o caráter de mundo pequeno. Recentemente, K. Komidis, S. Havlin e A. Bunde [35], mostraram para um modelo de rede conceitualmente similar ao modelo de Kleinberg e ao modelo de percolação de longo alcance [36], resultados similares aos encontrados no presente trabalho. Na análise realizada por eles, as redes são capazes de manter suas características de mundo pequeno apenas para $\alpha < d$.

2.2 Modelo de epidemia SIS aplicado ao modelo de Kleinberg

Vimos na seção anterior que o modelo de Kleinberg perde suas características de mundo pequeno para valores elevados de α . Com o intuito de verificar as propriedades de mundo pequeno do modelo de Kleinberg, propomos estudar o efeito da topologia do modelo na dinâmica de um processo de espalhamento de epidemia. Para isso, estudaremos o modelo de epidemia *suscetível-infectado-suscetível* (SIS) e sua aplicação ao modelo de Kleinberg.

2.2.1 O modelo SIS

No modelo epidêmico SIS, cada vértice em uma rede complexa representa um indivíduo e cada conexão representa uma ligação pela qual uma infecção pode espalhar-se para outro indivíduo. Neste modelo simples, cada vértice pode apresentar apenas dois estados discretos e distintos. Um estado *suscetível a infecção*, ou estado “saudável”, e um estado infectado. A cada passo de tempo, cada vértice suscetível é infectado com probabilidade ν se estiver conectado com algum vértice infectado. Durante o mesmo passo de tempo, um vértice infectado pode curar-se com uma probabilidade δ , e tornar-se suscetível novamente. Podemos definir uma taxa de espalhamento efetiva

$$\lambda = \frac{\nu}{\delta},$$

e, sem perda de generalidade, considerar $\delta = 1$ [28] o que nos fornece

$$\lambda = \nu. \tag{2.2}$$

Para topologias padrões, o resultado mais importante é a predição de um limite não nulo λ_c [37]. Se o valor de λ estiver acima do valor limite λ_c ($\lambda \geq \lambda_c$), a infecção torna-se persistente no tempo, ou seja, não se extinguirá e entrará em um estado de equilíbrio. Caso $\lambda < \lambda_c$ a infecção dever desaparecer da rede de forma exponencialmente rápida. É natural pensar que o limiar epidêmico λ_c seja equivalente a um ponto crítico em uma transição de fase fora do equilíbrio. Neste caso, o ponto crítico separa uma fase ativa com uma densidade ρ de vértices infectados (estado endêmico) de uma fase absorvente onde apenas vértices saudáveis estão presentes e temos atividade nula.

2.2.2 Resultados e discussões

No modelo WS para redes de mundo pequeno, podemos abordar analiticamente o modelo SIS considerando uma simples equação de reação para uma aproximação de campo médio, aplicada à densidade de vértices infectados como uma função do tempo, ou seja, $\rho(t)$ [28]. Como o modelo de WS possui uma distribuição exponencial para a conectividade dos vértices, podemos escrever a equação de reação na forma:

$$\partial_t \rho(t) = -\rho(t) + \lambda \langle k \rangle \rho(t) [1 - \rho(t)]. \quad (2.3)$$

Na Eq. (2.3) foram desconsideradas os termos de ordem superiores, visto que estamos interessados no período inicial da infecção, perto da transição de fase, isto é, $\rho(t) \ll 1$. O primeiro termo do lado direito da Eq. (2.3) considera que os vértices infectados tornam-se saudáveis por unidade de tempo. O segundo termo representa a media de vértices saudáveis que são infectados por cada vértice infectado, a qual é proporcional a taxa efetiva de infecção λ , o número médio de contatos que emana de cada vértice $\langle k \rangle$, a probabilidade de que uma dada ligação aponte para um vértice saudável $[1 - \rho(t)]$. Agora, se impormos a condição para o estado estacionário

$$\partial_t \rho(t) = 0,$$

obteremos a equação

$$\rho_e [-1 + \lambda \langle k \rangle (1 - \rho_e)] = 0 \quad (2.4)$$

para o estado estacionário ρ_e . A Eq. (2.4) resulta em

$$\rho_e = 0, \text{ se } \lambda < \lambda_c, \quad (2.5)$$

$$\rho_e = \frac{\lambda - \lambda_c}{\lambda}, \text{ se } \lambda > \lambda_c; \quad (2.6)$$

com um limiar endêmico $\lambda_c = \langle k \rangle^{-1}$.

A aproximação de campo médio para o modelo SIS, bem conhecida para redes aleatórias, é válida para o modelo WS independentemente das diferentes estruturas locais que esse possa assumir [28]. Sendo o modelo de Kleinberg construído em cima do paradigma definido pelo modelo WS, esperamos que o limiar endêmico do modelo não seja menor do que $\langle k \rangle^{-1}$ para um equivalente WS. A verificação do valor de λ_c associado ao modelo de Kleinberg é realizada através de simulações da dinâmica regida pelo modelo SIS em redes de Kleinberg para diferentes valores de α . No modelo de Kleinberg dois parâmetros são importantes para o cálculo da conectividade média, p e q , que, na forma como são definidos, equivale a dizer que todos os vértices da rede possuem não menos que $k_{out} = (4 \sum_{i=1}^p i) + q$ contatos, $(4 \sum_{i=1}^p i)$ locais e q contato de longo alcance saindo. Como cada contato deve apontar para outro vértice, podemos

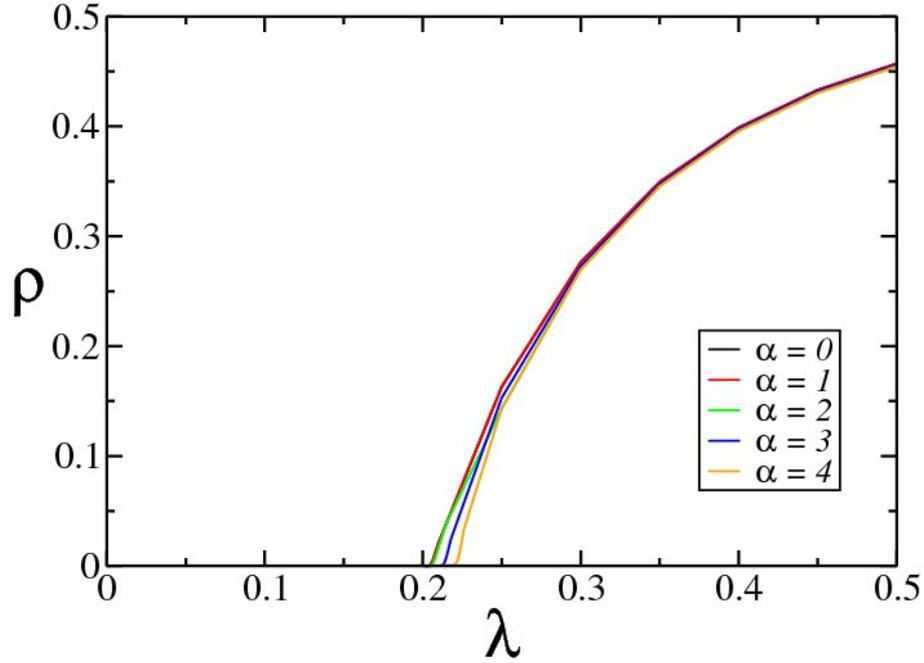


Figura 2.4: Densidade de vértices infectados ρ para os estados estacionários como função de λ , para o modelo de Kleinberg com diferentes valores de α . Podemos verificar facilmente que o limiar epidêmico λ_c é dependente do expoente de agregação α . Nas simulações, foram utilizadas rede com $N = 10^6$ vértices.

dizer que a conectividade média para este modelo é

$$\langle k \rangle = k_{out} + q = \left(\sum_{i=1}^p 4i \right) + 2q. \quad (2.7)$$

Para o modelo SIS, a cura de um vértice é independente da conectividade do mesmo, porém sua infecção depende (segundo termo na Eq. (2.3)). A rede definida pelo modelo de Kleinberg é não-direcionada, visto que foi proposta para o modelar experimento de Milgram; como em uma dinâmica epidêmica, a probabilidade de ser infectado depende das ligações que chegam em um vértice, podemos reescrever a Eq. (2.3) na forma:

$$\partial_t \rho(t) = -\rho(t) + \lambda \langle k_{in} \rangle [1 - \rho(t)]. \quad (2.8)$$

Seguindo a mesma análise, podemos facilmente concluir que $\lambda_c \sim \langle k_{in} \rangle$ em um regime de mundo pequeno. Sem problemas podemos concluir que $\langle k_{in} \rangle = \langle k \rangle - q = \langle k_{out} \rangle$, ou seja,

$$\lambda_c \sim \langle k_{in} \rangle = \left(\sum_{i=1}^p 4i \right) + q. \quad (2.9)$$

Em nossas simulações, utilizamos os parâmetros $N = 10^6$ e $p = q = 1$, o que pela Eq. (2.9)

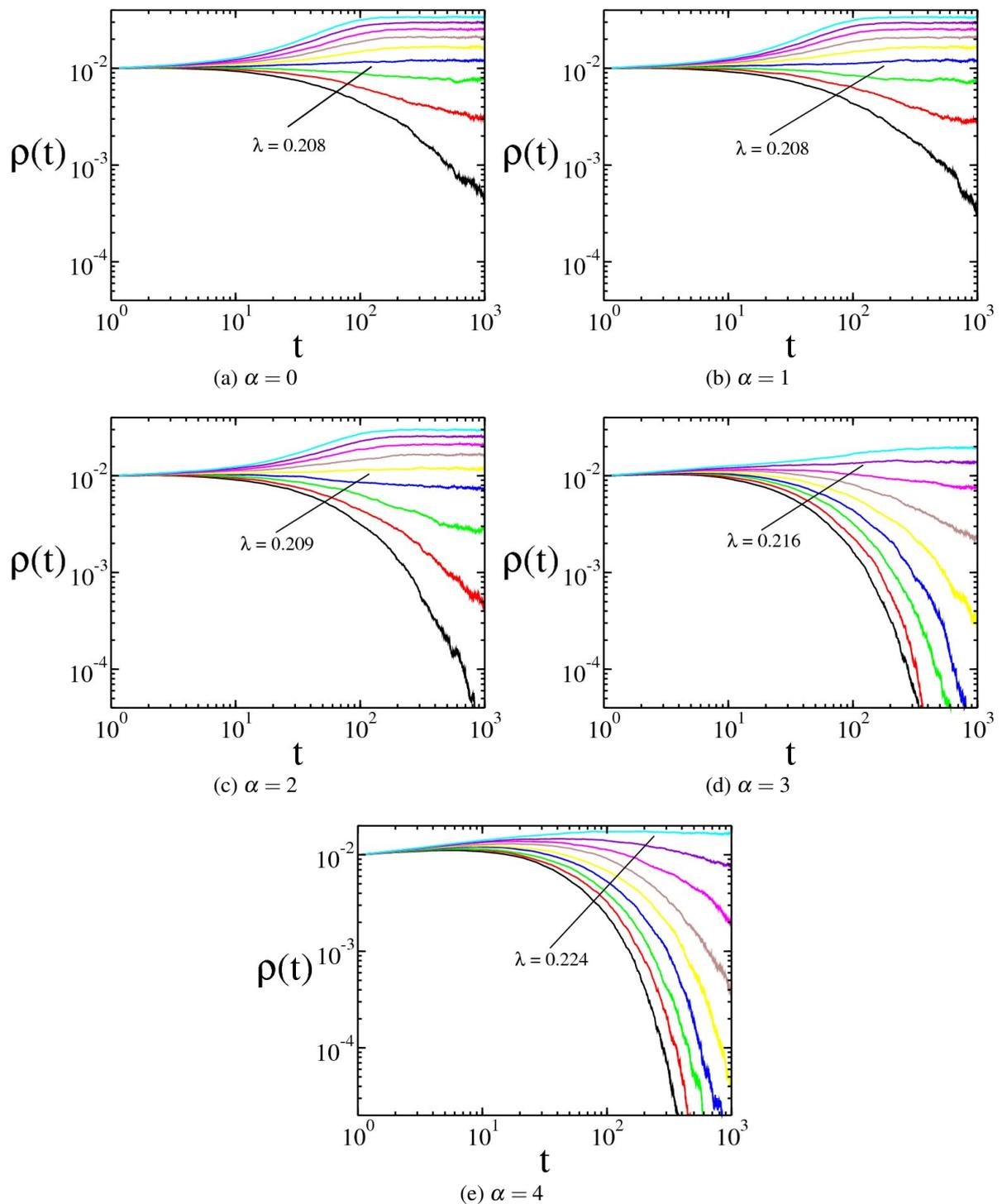


Figura 2.5: Evolução no tempo t para a densidade de vértices infectados $\rho(t)$. Em todos os casos acima apresentados, utilizamos $N = 10^6$ e como condição inicial $\rho(t = 0) = 0.01$. (a) A taxa de espalhamento $\lambda = 0.205$ a 0.213 (de baixo para cima). A infecção torna-se persistente para valores próximos de $\lambda = 0.208$ ou maiores, permitindo que concluamos que $\lambda_c = 0.208$. (b) A taxa de espalhamento $\lambda = 0.205$ a 0.213 (de baixo para cima). $\lambda_c = 0.208$. (c) A taxa de espalhamento $\lambda = 0.205$ a 0.213 (de baixo para cima). Encontramos $\lambda_c = 0.209$. Aqui podemos perceber uma leve mudança no limiar endêmico. (d) A taxa de espalhamento $\lambda = 0.209$ a 0.217 (de baixo para cima). A infecção torna-se persistente para valores próximos de $\lambda_c = 0.216$ ou maiores. (e) A taxa de espalhamento $\lambda = 0.216$ a 0.224 (de baixo para cima). $\lambda_c = 0.224$.

nos leva a $\lambda_c \sim 0.2$. A Fig. (2.4) mostra que o limiar endêmico do modelo de Kleinberg está próximo do previsto pela aproximação de campo médio; no entanto, com o aumento do expoente de agregação α , podemos ver claramente que esse limiar se afasta do valor idealizado $\lambda_c \sim 0.2$. Quando analisamos a evolução temporal de $\rho(t)$ (Fig. (2.5)), verificamos que para os valores de α estudados, a epidemia atinge um estado endêmico com valores de $\lambda > 0.2$. No entanto, apenas para valores de $0 \leq \alpha \leq 2$ encontramos valores para o limiar endêmico tal que este esteja de acordo com o previsto pela aproximação de campo médio.

Devemos acreditar que a aproximação de campo médio, válida para o modelo WS, também seja válida como exemplo de comparação para o modelo de Kleinberg. A topologia espacial do modelo de Kleinberg deve diferenciar-se do modelo WS na forma como influencia o tempo de relaxação da epidemia, ou seja, o tempo que epidemia leva para entrar em seu estado estacionário. Contudo, dentro deste estado estacionário, ou estado endêmico, a topologia espacialmente embutida do modelo de Kleinberg não deve influenciar na taxa de variação de $\rho(t)$. A Tab. (2.1), onde estão listados os valores de λ_c encontrados em Fig. (2.5), mostra que essa mesma topologia influencia nos valores do limiar endêmico.

Estes resultados ajudam a esclarecer, conceitualmente, o valor mínimo de $\alpha = d$, presente no trabalho de Kleinberg. Quando a distribuição de tamanhos dos contatos de longo alcance é gerada de forma próxima da uniforme, a maior frequência de contatos mais distantes dificulta com que o indivíduo tire proveito das características de mundo pequeno da rede de Kleinberg. À medida que o expoente de agregação afasta-se de $\alpha = 0$, contatos de longo alcance tornam-se “menos” distantes, e dessa forma a correlação entre a estrutura local e estrutura de contatos de longo alcance da rede permitem com que o indivíduo inserido na própria comece a ver o mundo pequeno a sua volta e possa aproveitar sua estrutura para achar caminhos curtos entre ele e os demais indivíduos da rede. No entanto, com o aumento do expoente de agregação, as conexões

α	λ_c
0.0	~ 0.208
1.0	~ 0.208
2.0	~ 0.209
3.0	~ 0.216
4.0	~ 0.224

Tabela 2.1: Expoente de agregação α e seu respectivo limiar endêmico λ_c . Aqui utilizamos $p = q = 1$, o que representa uma rede com $\langle k_{in} \rangle = 5$. Pela previsão da aproximação de campo médio, $\lambda_c = \langle k_{in} \rangle$ para regimes de mundo pequeno. Como discutido na seção anterior, o modelo de Kleinberg apresenta comportamento próximo ao de mundo pequeno (modelo WS) apenas para a região $0 \leq \alpha \leq 2$. Para valores de $\alpha > 2$ o limiar endêmico distancia-se do previsto para redes que seguem o paradigma de WS.

de longo alcance tornam-se freqüentemente curtas, por mais que isso facilite a identificação, por parte de indivíduo, da estrutura de contatos de longo alcance do mundo a sua volta, a nova topologia do mundo não permite com que ele ache caminhos eficientes, visto que o mundo já não é mais pequeno.

2.3 Navegação em redes espacialmente correlacionadas com custo

Muitos sistemas estudados em ciências e em áreas tecnológicas possuem as características de navegação de informação que utiliza a organização local da rede. No entanto, outros sistemas podem ser construídos de forma que a navegação em sua rede representativa não seja dada através do conhecimento local, mas sim, por meio da estrutura global da rede, ou seja, o portador da informação no vértice fonte f , seja ele algum mecanismo de navegação, ou mesmo o próprio vértice, pode ser permitido a conhecer todas as conexões presentes na rede. Considere, por exemplo, uma rede de estações de metrô; as rotas fornecidas pela companhia de transporte responsável são planejadas no intuito de otimizar, o tempo médio das viagens através da rede de estações. Sem sombra de dúvidas, tal tarefa é realizada levando em conta toda a estrutura da rede: a posição de cada uma das estações de metrô, que são os vértices da rede; e seus túneis, que conectam determinados pares de estações. Se quisermos minimizar a distância média entre as estações de metrô, e com isso, reduzir o tempo médio de viagem, podemos utilizar a estratégia de Kleinberg e adicionar conexões de longo alcance entre os vértices da rede. Perceba que essa relação entre o tempo médio da navegação da informação na rede, não necessariamente é feita apenas como redes de estações de metrô, ela pode ser feita com qualquer rede real feita pelo homem para troca ou transporte de informação.

Quando consideramos a estrutura global da rede, o mínimo caminho médio, $\langle l \rangle$, não apresenta o comportamento especial para a navegação baseada apenas na estrutura local [1, 2]. A Fig. (2.1) mostra esse comportamento de $\langle l \rangle$ para sete diferentes tamanhos de rede, e, como discutido na seção 2.1, o expoente de agregação efetivo $\alpha_e = 0$, independentemente do tamanho da rede e de sua dimensão. Infelizmente, para redes de estações de metrô, por exemplo, o custo financeiro envolvido para a construção de uma grande extensão de novos túneis torna a adição de túneis construídos com uma distribuição uniforme de tamanho impossível. Habitualmente, existe uma quantidade limitada de dinheiro disponível para a construção de uma ampliação na malha de túneis da rede de metrô. Portanto, se a estratégia escolhida para otimizar o tempo de viagem na rede envolver a adição de novas conexões, o que para rede de estações de metrô significa a construção de novos túneis, é necessário trabalhar em uma estratégia que envolva um

vínculo que limite o comprimento total das conexões de longo alcance

$$\tau = \sum r_{u,v}. \quad (2.10)$$

Em outras palavras, pode-se fixar a soma total de todas as novas conexões de longo alcance, no intuito de representar este custo limitado.

2.3.1 O modelo de rede com custo limitado

Nesta seção, propomos um modelo para redes embutidas espacialmente, são adicionadas conexões de longo alcance na tentativa de minimizar o tempo de navegação através da rede. Neste modelo, consideramos um custo limitado τ dado pela Eq. (2.10). Nosso modelo é construído de forma bastante simples. Considere uma rede d -dimensional quadrada regular, com $N = L^d$ vértices. Cada vértice u é conectado aos seus vizinhos mais próximos, este é o substrato regular da rede (a rede quadrada). Na analogia feita anteriormente com a rede de estações de metrô, este substrato representa os túneis já existentes na rede de estações sobre a qual serão adicionados novas conexões. A partir desse substrato, adicionamos conexões de longo alcance, os novos túneis, a alguns vértices u , sorteados de forma uniformemente aleatória, a um outro vértice v , o qual é aleatoriamente escolhido com probabilidade proporcional a $[r_{u,v}]^{-\alpha}$, da mesma forma como é feito no modelo de Kleinberg. A grandeza $r_{u,v}$ é a distância “Manhattan” entre os vértice u e v na rede d -dimensional, e α é o expoente de agregação.

Uma característica especial dessa restrição em τ é a completa ausência de controle no número de conexões de longo alcance, ou seja, a fração de vértices com conexões de longo alcance não é fixa. Para termos uma representação pictórica do nosso problema, tomamos uma simples realização para o cálculo do mínimo caminho l de cada um dos vértices da rede com respeito ao vértice mais central da rede para diferentes valores de α . Para executarmos esse cálculo, primeiramente escolhemos um vértice *raiz*, na Fig. (2.6) este é o vértice central. Uma vez escolhido o vértice central, visitamos todos os seus contatos, incluindo, caso este exista, o contato de longo alcance. Estes vértices são então classificados como distantes um passo de tempo do vértice raiz. Depois, visitamos os contatos classificados anteriormente como “distantes um passo de tempo”, e se estes não forem outros vértices distantes um passo de tempo, os classificamos como vértices distantes dois passos de tempo do vértice raiz. Seguindo esse procedimento para toda a rede, obtemos o valor de l para cada vértice da rede com respeito ao vértice raiz. A Fig. (2.6) é uma representação em código de cores para este cálculo. Com $\alpha = 1$ temos um pequeno número de conexões de longo alcance, com isto as regiões coloridas que representam distâncias a poucos passos de tempo do vértice raiz apresentam-se bastante espar-

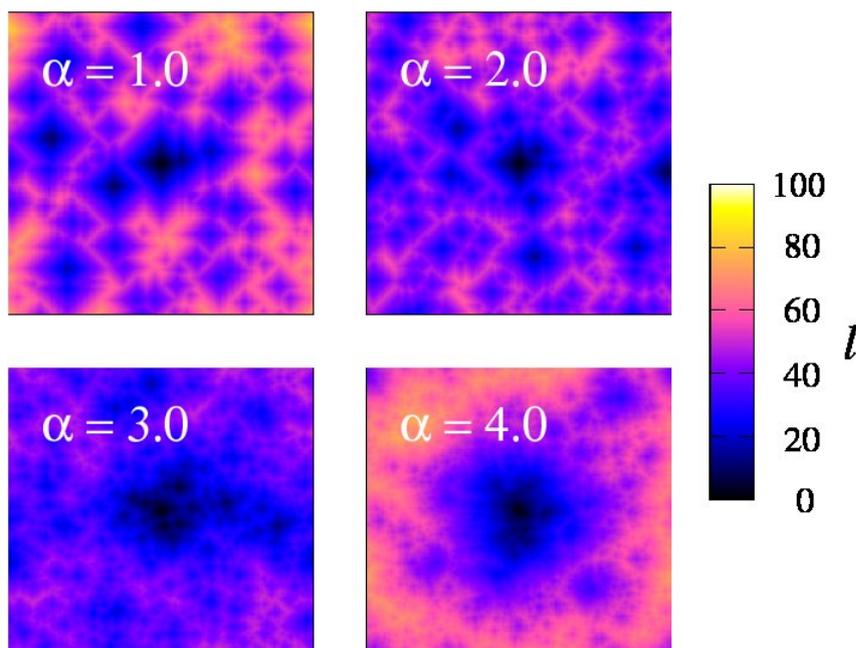


Figura 2.6: O mínimo caminho l de cada vértice da rede é calculado em relação ao vértice central da rede para diferentes valores de α . O modelo de rede é construído a partir de uma rede quadrada com L^2 vértices. Aqui $L = 256$, e o custo é fixado em $\tau = L^2$. Facilmente podemos observar que com o aumento de α , a média do mínimo caminho será afetada.

Para $\alpha = 3$, o número de conexões de longo alcance aumenta, produzindo um efeito obvio em $\langle l \rangle$. Contudo, para $\alpha = 4$, devido a distribuição em lei de potência, as conexões de longo alcance conectam vértices não muito distantes entre si, como efeito disso, as regiões coloridas que representam vértices a poucos passos de tempo do vértice raiz tornam-se densas em torno da vizinhança de mesmo.

2.3.2 Efeitos do custo τ sobre a dinâmica da navegação Kleinberg

Primeiramente, podemos analisar o efeito que o custo τ pode causar à dinâmica regida pela navegação de Kleinberg. Isto se torna importante devido a origem do modelo, derivado do modelo de Kleinberg. A diferença crucial entre o modelo de Kleinberg e o nosso modelo, reside na ausência de controle na densidade de conexões de longo alcance adicionadas ao substrato regular. Para obter resultados consistentes com os tamanhos diferentes de rede, fixamos o custo $\tau = N = L^d$. Podemos escrever essa densidade de conexões de longo alcance como sendo $\rho = n/N = n/L^d$, onde n é o número de conexões de longo alcance dado por:

$$n = \tau / \langle r \rangle, \quad (2.11)$$

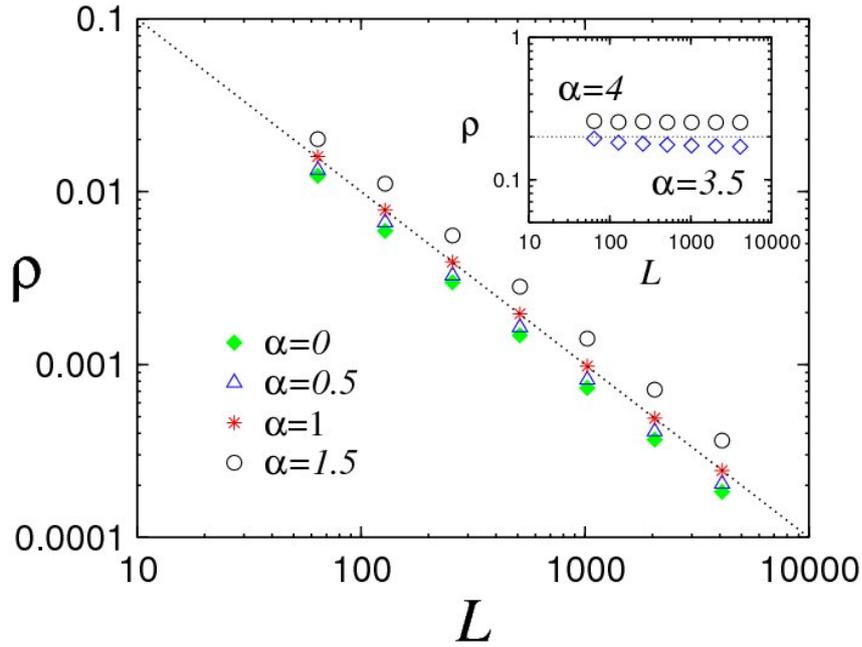


Figura 2.7: Densidade de ρ de conexões adicionadas como função do tamanho da rede L . Para valores de $\alpha < 2$, verificamos que $\rho \sim L^{-1}$, comportamento que não se repete acima desse limiar. A linha pontilhada possui inclinação igual a -1 . No *inset* vemos que para $3 < \alpha$ a densidade ρ torna-se praticamente independente do tamanho da rede. Aqui, a linha pontilhada possui inclinação nula.

onde $\langle r \rangle$ representa o valor médio das conexões de longo alcance, com isto teremos que

$$\rho \propto \frac{1}{\langle r \rangle}. \quad (2.12)$$

De forma natural, podemos fixar uma menor distância para as conexões de longo alcance como sendo r_{min} , assim

$$\langle r \rangle \propto \int_{r_{min}}^L r \times r^{-(\alpha-1)} \times r^d dr, \quad (2.13)$$

que resulta em

$$\langle r \rangle \propto L^{d-\alpha+1}. \quad (2.14)$$

No entanto, para $\alpha > d + 1$, teremos $L^{d-\alpha+1} \ll r_{min}^{d-\alpha+1}$ de modo que $\langle r \rangle$ deve ser independente do tamanho da rede para este intervalo, com isso teremos $\rho \sim L^0$.

No intervalo $0 \leq \alpha < d$, o tamanho médio das conexões de longo alcance deve divergir no limite termodinâmico pois crescerá mais rápido que o tamanho da rede. No entanto isto não acontece, visto que seu valor está sempre limitado pelo próprio tamanho da rede. Dessa forma podemos concluir que para $0 \leq \alpha < d$ teremos $\rho \sim 1/\langle r \rangle = L^{-1}$. Podemos resumir o

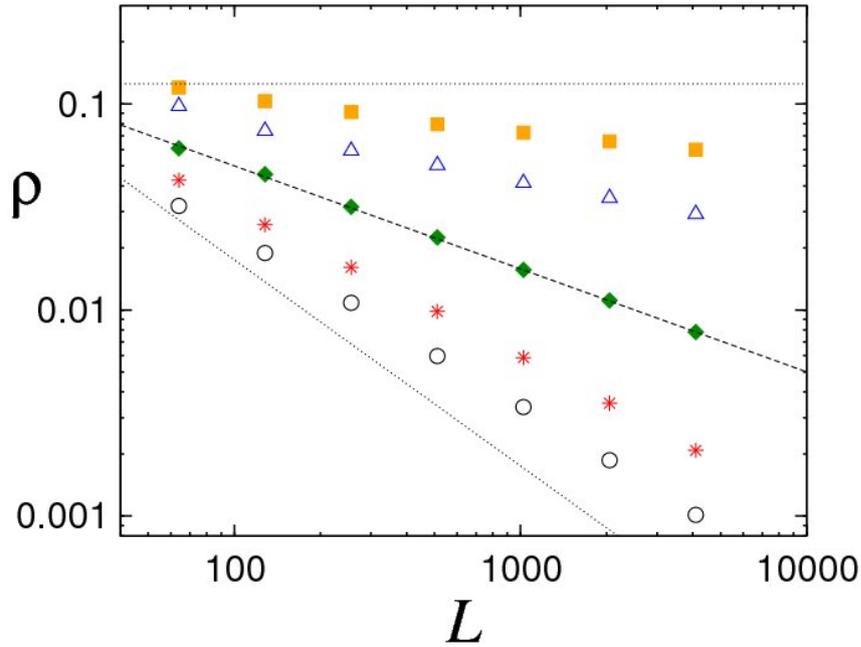


Figura 2.8: Densidade ρ como função do tamanho da rede L . As duas linhas pontilhada delimitam a região onde ρ apresenta escalonamento próximo de $L^{-(3-\alpha)}$, a superior com inclinação nula e a inferior com inclinação -1 . A linha tracejada possui inclinação 0.5 e coincide com os valores ρ para $\alpha = 2.5$. No gráfico, $\alpha = 2$ (círculos pretos), 2.2 (estrelas vermelhas), 2.5 (losangos verdes), 2.8 (triângulos azuis), e 3 (quadrados laranjas).

comportamento da densidade de conexões de longo alcance da seguinte forma:

$$\rho \sim \begin{cases} L^{-1} & , \text{ para } 0 \leq \alpha \leq d; \\ L^{-(d-\alpha+1)} & , \text{ para } d \leq \alpha \leq d+1; \\ L^0 & , \text{ para } d+1 \leq \alpha. \end{cases} \quad (2.15)$$

A Fig. (2.7) podemos perceber uma clara dependência da densidade ρ de novas conexões com relação ao tamanho da rede. Para $\alpha < 2$, verificamos que $\rho \sim L^{-1}$, a linha pontilhada é simplesmente comparativa, e possui inclinação igual a -1 . No entanto, a figura inserida de Fig. (2.7) mostra, através da comparação com um linha pontilhada de inclinação nula, que para valores maiores de α , precisamente para $\alpha > 3$, a densidade ρ torna-se independente do tamanho da rede.

A Fig. (2.8) mostra valores intermediários de α , inseridos no intervalo $2 \leq \alpha \leq 3$. Percebe-se que esta fase da densidade apresenta uma maior variação no seu comportamento associado ao tamanho de rede L . Nesse regime, a densidade passa de uma fase onde $\rho(\alpha < 2) \sim L^{-1}$ para um regime independente de L , ou seja, $\rho(3 < \alpha) \sim L^0$. Nossos resultados sugerem que para esse regime intermediário, $\rho(2 < \alpha < 3) \sim L^{-(3-\alpha)}$. Na Fig. (2.8) estão representadas três linha, duas pontilhadas e um tracejada. A linha inferior, pontilhada possui inclinação igual a

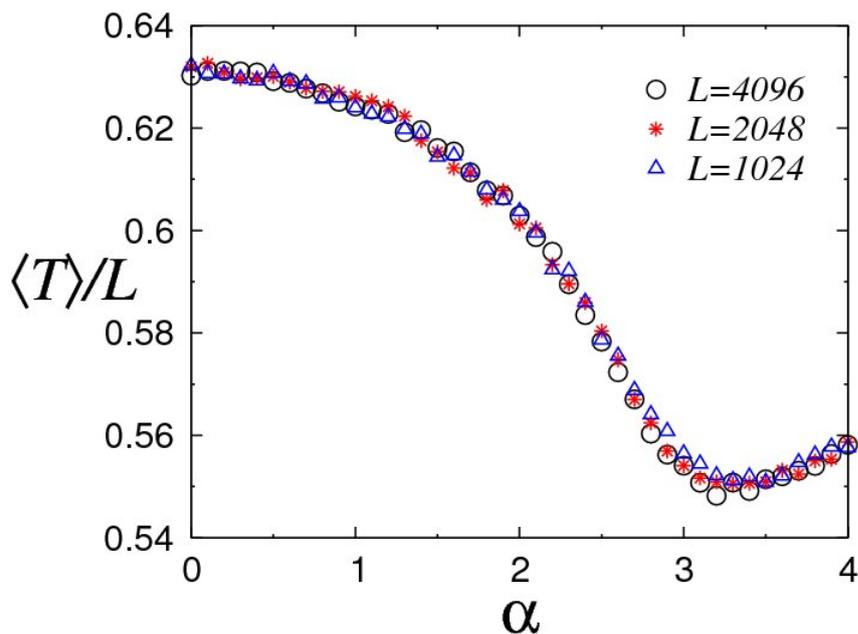


Figura 2.9: Tempo médio característico de envio da mensagem $\langle T \rangle$ como função de α . Resultados da simulação do algoritmo descentralizado em redes de tamanho $L = 1024, 2048, 4095$; com ligações de longo alcance distribuídos com probabilidade $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$. Há um custo envolvido na adição dessas conexões de $\tau = \sum r_{u,v} = L \times L$. Esse custo altera a densidade de conexões de longo alcance, causa do mínimo encontrado em α próximo de 3 pela direita. Cada ponto é resultado de pelo menos 4000 realizações.

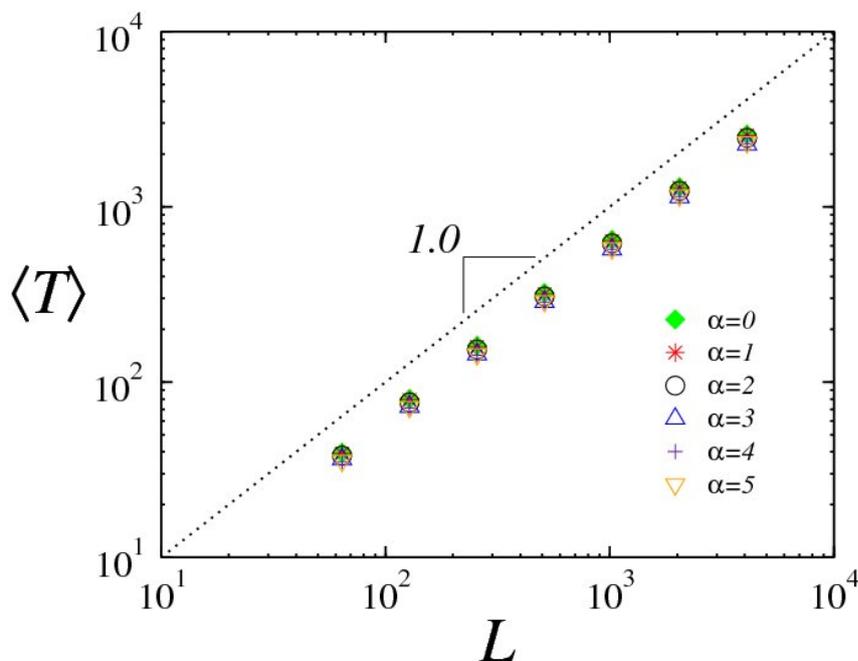


Figura 2.10: $\langle T \rangle$ como função do tamanho da rede L para cinco valores de α diferentes, encontramos um comportamento praticamente idêntico. Aparentemente, $\langle T \rangle \sim L$. A linha pontilhada possui inclinação igual a unidade e serve como comparação do escalonamento.

unidade, enquanto que a linha pontilhada inferior possui inclinação nula. De forma Pictórica, elas separam as fases para os três regimes de comportamento da densidade. A linha tracejada ao centro possui inclinação igual a 0.5, e coincide com precisão, com a inclinação do melhor ajuste para $\alpha = 2.5$. Isso nos deixa confortáveis em dizer que $\rho(L, 2 < \alpha < 3) \sim L^{(3-\alpha)}$.

Para determinar a influência da densidade na navegação de Kleinberg, realizamos uma dinâmica regida pelo algoritmo descentralizado de Kleinberg no modelo com custo nas conexões de longo alcance e medimos com isso, o tempo médio $\langle T \rangle$ necessário para uma mensagem partir de um vértice fonte f e chegar em um vértice alvo a , ambos arbitrariamente escolhidos. A fig (2.9) apresenta o resultado dessas realizações. Podemos constatar uma queda suave no valor de $\langle T \rangle$ até α atingir valor $\alpha = 2$. Após esse regime, essa se acentua, estabilizando-se em um valor de $\alpha \simeq 3$. À medida que α distancia-se de $\alpha = 3$, um leve aumento no valor de $\langle T \rangle$ pode ser notado.

Quando verificamos o comportamento de $\langle T \rangle$ com relação aos tamanhos da rede para valores fixos de α , temos que $\langle T \rangle \sim L$ (Fig. (2.10)), ou seja, para nenhum valor de α teremos um tempo médio de envio que cresça mais devagar (de forma logarítmica) que o crescimento da rede. Com isto podemos concluir que, para a dinâmica de navegação de Kleinberg, o modelo de redes com custo para a adição de conexões geograficamente correlacionadas por uma distribuição de probabilidade $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$ não apresenta características de mundo pequeno quando existe um vínculo de custo na adição de conexões de longo alcance. Este vínculo não permite que as “pistas” fornecidas pela correlação entre a estrutura local, o substrato quadrado da rede, e a estrutura de conexões de longo alcance apareçam da mesma forma que apareciam quando tínhamos uma densidade constante de contatos de longo alcance. Com isso, a mensagem não consegue tirar proveito da estrutura de longo alcance da mesma forma que para o caso estudado por Kleinberg, pois dependerá sempre de estar em mão de um portador que, além de possuir uma conexão de longo alcance, que está seja um salto maior em direção ao alvo do que as conexões oriundas da estrutura local. Para o caso de Kleinberg, a probabilidade que a mensagem possuía de ser enviada a um novo portador com contatos de longo alcance, era independente do expoente de agregação, com isso, como vimos anteriormente, a navegação ótima era atingida para o valor de α que preservava as características de mundo pequeno da rede. Agora, temos um α ótimo para o valor de α para o qual a densidade de conexões de longo alcance é independente do tamanho da rede, $\alpha > 3$.

2.3.3 Navegação com informação global

A introdução de um custo τ na adição de contatos de longo alcance, remove as características de mundo pequeno para navegação em redes espacialmente embutidas cuja distribuição de

tamanhos siga uma lei de potência. Isso nos leva a crer que caso se busque otimizar a navegação de informação em redes reais através da construção de novas conexões, utilizando-se um processo que cada nova conexão possui um preço agregado por sua introdução na rede, existe um único modelo dentro da família de redes definida pela associação da condição

$$P(r_{u,v}) \sim r_{u,v}^{-\alpha}, \quad (2.16)$$

com o vínculo

$$\tau = \sum r_{u,v} = \text{constante}. \quad (2.17)$$

que otimiza essa navegação. Essa otimização, como vimos, não leva a uma topologia que apresente o fenômeno de mundo pequeno.

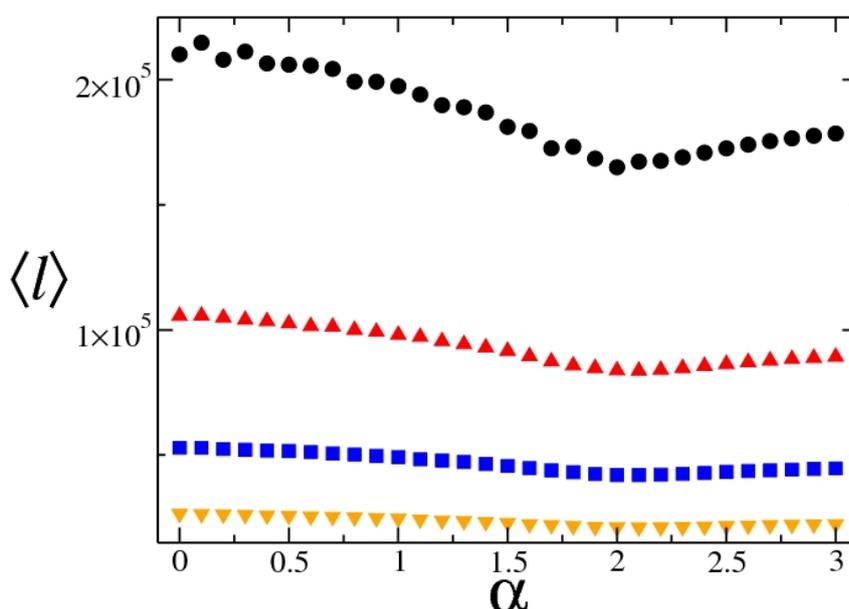


Figura 2.11: $\langle l \rangle$ como uma função de α , para uma rede 1D. Encontramos que o mínimo caminho ótimo é obtido com $\alpha_e = 2$. Com a restrição na soma total do tamanhos dos conexões de longo alcance, o número de conexões de longo alcance não é fixo, por exemplo, para $\alpha = 0$ as conexões de longo alcance conectam vértices distantes maiores tornam-se mais freqüentes, o que reduz o número de conexões de longo alcance. Aqui $L = 2^{17}, 2^{18}, 2^{19}$ e 2^{20} (de baixo para cima).

Para extrair mais informações, utilizaremos mais uma vez o paralelo entre o tempo médio de envio de uma mensagem através da rede entre dois vértices quaisquer arbitrários $\langle T \rangle$, e o mínimo caminho médio da rede $\langle l \rangle$. Calculamos $\langle l \rangle$ para diferentes valores de α e diferentes tamanhos de rede, para sistemas unidimensionais (1D) e sistemas bidimensionais (2D). Na Fig. (2.11) e Fig. (2.12) há mínimos óbvios presentes indicando que a navegação ótima é obtida para $\alpha = 2$, para o modelo 1D; e $\alpha = 3$, para o caso 2D, quando um custo é envolvido no

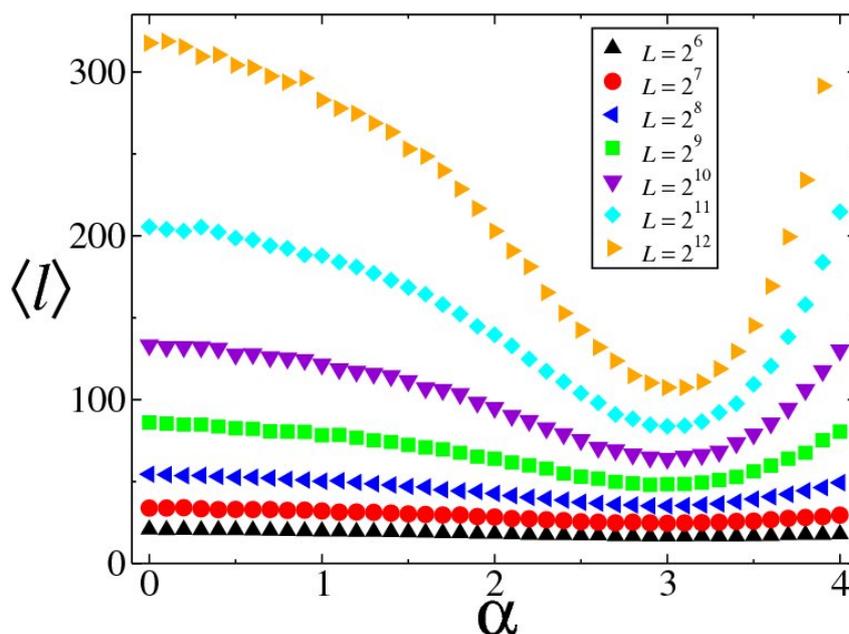


Figura 2.12: $\langle l \rangle$ como uma função de α , para uma rede 2D. Encontramos que o mínimo caminho ótimo é obtido com $\alpha_e = 3$.

processo de adição de novas conexões, com dependência geográfica, a redes já existentes. Este comportamento especial indica uma relação muito próxima entre a distribuição de tamanhos das novas conexões e o custo envolvido para adicioná-las. Mais que isso, mostra a existência de uma combinação especial destes fatores na otimização da navegação através da rede.

Quando permitimos que as conexões sejam apenas adicionadas em uma direção na rede quadrada, ou seja, se definirmos uma direção norte e uma direção sul, e, à medida que elas fossem introduzidas, conectassem vértices apenas distantes entre si na direção norte, ou na direção sul, encontramos que o mínimo é obtido em um valor menor de $\alpha = 2.1$, aproximadamente igual a $d - 1$ (Fig. (2.13)). Isto é justificado devido ao fato de que, com esse vínculo, a uma mesma distância r de um vértice não temos mais $4r$ possíveis novos contatos. Agora teremos sempre quatro candidatos, fazendo com que os processos de escolha dos novos contatos seja equivalente ao processo realizado em um sistema unidimensional, quando também temos um número fixo de possíveis novos contatos. Para o caso unidimensional, é fácil perceber que sempre teremos 2 candidatos. Note que outra semelhança entre o sistema bidimensional com vínculo nas direções e o sistema unidimensional sem este vínculo, é a conexão com vértices na mesma direção (afinal, todos os vértices em uma rede unidimensional estão dispostos na mesma direção). Esse resultado aumenta o leque de opções disponíveis para a tarefa de otimização da navegação da informação na rede. Apesar de ligeiramente diferentes (o valor de $\langle l \rangle$ mínimo para o sistema com o vínculo adicional na direção das conexões de longo alcance é ligeiramente menor do que o encontrado no sistema que possui apenas vínculo associado ao custo), os

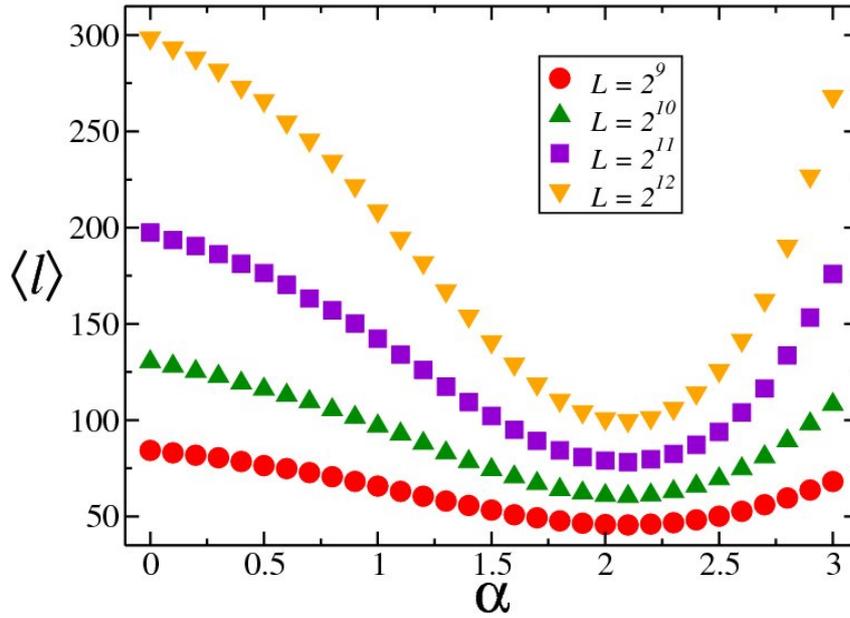


Figura 2.13: Mínimo caminho $\langle l \rangle$ como função do expoente de agregação α . Aqui são permitidas conexões entre vértices apenas na mesma direção (por exemplo, sem relacionarmos cada vértice u a uma posição na rede quadrada (x_u, y_u) , apenas um vértice v com posição (x_u, y_v) , ou (x_v, y_u) , poderá ser sorteado para receber a conexão). Encontramos um $\alpha_e = 2.1$, resultado próximo ao resultado obtido para uma rede unidimensional. Esse resultado não é mera coincidência devido às semelhanças topológicas entre a rede unidimensional e a rede bidimensional com vínculo nas direções. Esse resultado aumenta as possibilidades de estratégias que podem ser utilizadas para otimizar a navegação através da rede.

valores do mínimo caminho $\langle l \rangle$ ótimos pertencem aproximadamente a mesma faixa. Com isso, análises detalhadas do sistema real, e dos custos envolvidos, podem ser realizadas no intuito de escolher a melhor opção entre os casos apresentados.

Focando nossa atenção no comportamento do mínimo caminho em relação ao tamanho da rede (Fig. (2.14)), Apesar de o comportamento de sistema é bastante similar ao encontrado com a navegação de Kleinberg, encontramos comportamentos diferentes para $\langle l \rangle$ como função de L . Para diferentes valores de α o comportamento de $\langle L \rangle$ aparenta ser $O(L^\varepsilon)$ com $0 < \varepsilon < 1$. No entanto, apenas para o expoente efetivo $\alpha_e = 3$, $\langle l \rangle \sim (\log(L))^{2.7}$; o que evidencia a presença do fenômeno de mundo pequeno.

Os valores apresentados na Fig. (2.11) e Fig. (2.12) podem ser interpretados como um limite inferior para a navegação de Kleinberg para o caso com custo. Essa semelhança nos valores de α_e , tanto para a navegação de Kleinberg, quanto para a navegação com conhecimento global, mostra que, para realizarmos a tarefa de otimizar a navegação estaremos resolvendo dois problema simultaneamente. Além disso, mostra que ao inserirmos um custo para a adição de conexões de longo alcance, a densidade inconstante das conexões adicionais faz com que a

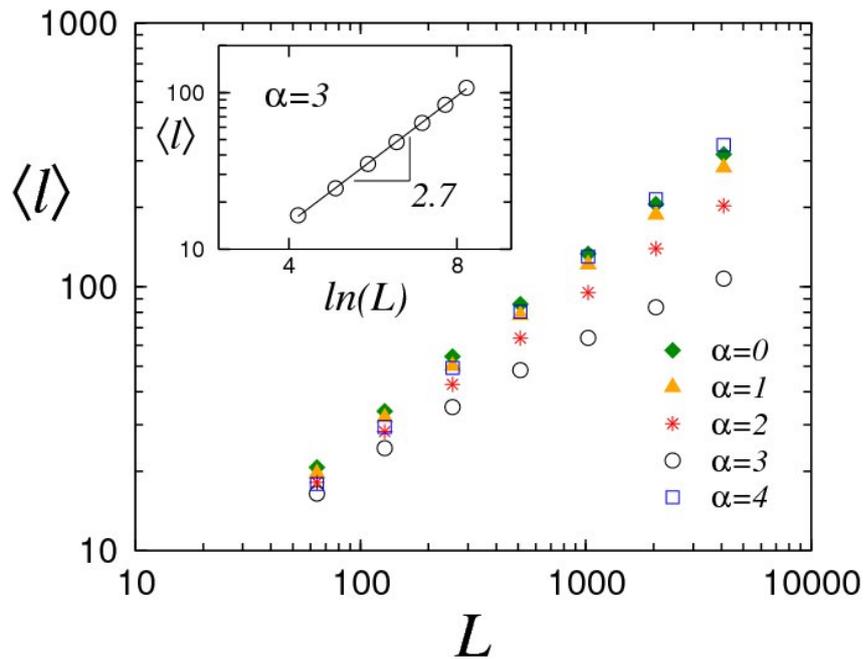


Figura 2.14: Mínimo caminho $\langle l \rangle$ como função de L para uma rede com o substrato quadrado. Há uma restrição no comprimento total das ligações $\tau = \sum r_{u,v}$. A curva para $\alpha_3 = 3$ cresce mais lentamente com o tamanho da rede do que qualquer outro valor de α . No *inset* vemos que o crescimento de $\langle l \rangle$ na condição ótima $\alpha_e = 3$ é consistente com o comportamento logarítmico.

rede presente o fenômeno de mundo pequeno apenas para os valores próximos do valor ótimo $\alpha_e = 3$. Este resultado responde a alteração no valor de α_e da navegação de Kleinberg, onde a correlação entre a estrutura local da rede quadrada com a estrutura das conexões de longo alcance fornecerá um cenário favorável para a navegação da rede.

Neste capítulo estudamos o comportamento do mínimo caminho médio $\langle l \rangle$ do modelo de Kleinberg, e verificamos que não apenas ele apresenta o fenômeno de mundo pequeno para um intervalo limitado de valores do expoente de agregação α , como verificamos, através do modelo epidêmico SIS, que para este intervalo a aproximação de campo médio é aplicável. Por último, introduzimos um modelo, derivado do modelo de Kleinberg, onde um custo τ limitava a adição de novas conexões de longo alcance. Estudamos o efeito que esse custo produz na dinâmica de navegação de Kleinberg, utilizando o algoritmo não-global proposto pelo mesmo, e sobre uma navegação global através do estudo do mínimo caminho médio para esse tipo de rede.

3 *Conclusões gerais*

Os efeitos de vínculos espaciais em redes complexas pode ser uma fonte muito rica em temas de estudo. Nesta dissertação apresentamos primeiramente um estudo sobre a navegação em redes pertencentes à classe de redes definida pelo modelo de Kleinberg [1, 2]. Este consistiu em analisar a navegação através da rede para o caso onde os constituintes possuem informação global da mesma, ou seja, a cada vértice é permitido saber o posicionamento dos demais vértices da rede e também todas as conexões presentes no sistema sem restrições de custo. Constatamos que a navegação ótima é obtida para os menores valores do expoente de agregação α , no entanto, o valor obtido para o mínimo caminho $\langle l \rangle$ dentro do intervalo $0 \leq \alpha \leq 2$ aparecia praticamente inalterados. Verificamos também que este aumento se tornava muito dependente do aumento do tamanho da rede, indicando um possível comportamento sem a presença do efeito de mundo pequeno para valores $\alpha > 2$. Esse comportamento pode ser verificado com a aplicação da dinâmica de epidemia suscetível-infectado-suscetível (SIS), onde o valor do limiar endêmico λ_c para o intervalo $\alpha > 2$ distanciava-se do previsto pela aproximação de campo médio, válido para redes de mundo pequeno [28]. Isto pode ser explicado devido à distribuição de tamanhos das conexões de longo alcance seguir uma lei de potência do tipo $P(r_{u,v}) \sim r_{u,v}^{-\alpha}$, portanto para valores suficientemente grandes do expoente de agregação, as conexões de longo alcance de um vértice arbitrário u acabam por se tornar muito agregadas em torno de sua vizinhança, fazendo com isso, que a informação em fluxo através da rede, seja ela um navegante ou uma mensagem sendo encaminhada entre componentes da rede, pois não há da estrutura de longo alcance presente visto que as conexões adicionadas são, efetivamente, conexões de curto alcance.

Os resultados obtidos em nosso trabalho mostram que interessantes comportamentos podem surgir quando consideramos redes espacialmente correlacionadas. Na primeira parte do trabalho, identificamos um intervalo do parâmetro α para o qual o modelo de Kleinberg se comporta como uma rede de mundo pequeno; no entanto, para a mesma, pode-se perceber facilmente que a rede deve apresentar coeficiente de agregação baixo, que deve anular-se para o limite termodinâmico [35]. No entanto, o modelo de Kleinberg é proposto em cima do pa-

radigma definido por Watts e Strogatz de que efeitos de mundo pequenos aparecem quando estamos em um mundo que interpola características regulares com características aleatórias [3]. Para Strogatz e Watts, o aparecimento do fenômeno é devido a coexistência de um auto valor para o coeficiente de agregação com baixos valores para o mínimo caminho médio da rede. No entanto, outros pesquisadores preferem considerar o efeito de mundo pequeno caracterizado apenas pelo comportamento $\langle l \rangle \sim \log(L)$ do mínimo caminho [21, 35]. Porém, vimos que, mesmo para a rede de Kleinberg, onde o coeficiente de agregação anula-se para $L \rightarrow \infty$ dentro do intervalo $0 \leq \alpha \leq d$, o mínimo caminho da rede é bastante pequeno dentro desse intervalo e o algoritmo não global de Kleinberg permite que os vértices encontrem caminhos curtos para enviar mensagens entre si. Note que isso é possível apenas devido o substrato regular da rede ser uma rede quadrada. Caso utilizássemos uma rede triangular como substrato, o expoente de agregação não se anularia no limite termodinâmico. Ao mesmo tempo, a previsão de campo médio, válida para o modelo de Watts e Strogatz, é aplicável para esse intervalo. Isto reabre dois questionamentos. Primeiro: o que realmente caracteriza o mundo pequeno? Segundo: sabendo-se o que o caracteriza, o mundo pequeno é equivalente a uma aproximação de campo médio? Nossos resultados sugerem que apenas o comportamento logarítmico de $\log(L)$ é suficiente para caracterizar o efeito de mundo pequeno e que neste mundo é equivalente, para o modelo de Kleinberg com expoente de agregação $\alpha \leq 2$, a uma aproximação de campo médio.

Modelos de redes correlacionadas espacialmente, ou seja, a probabilidade de que dois vértices da rede estejam conectados dependa de seus posicionamentos na rede, podem servir de referência para diversos sistemas reais construídos pelo homem. De forma geral, estamos sempre em busca de produzir mecanismos que realizem tarefas com maior eficácia. Redes de estações de metrô e a internet são dois exemplos de redes reais em que o aumento da conectividades da rede implica na construção de novas conexões, seja novos túneis conectando estações de metrô, ou a construção de uma rede de fibras óticas para melhorar o serviço de internet banda larga. Infelizmente, essas realizações possuem um valor financeiro agregado, limitando a construção de novas conexões. Neste trabalho apresentamos um estudo com redes espacialmente correlacionadas onde um custo τ para a “construção” (adição) de novas conexões está associado. Analisamos o efeito desse custo para dois processos de navegação, o primeiro o processo de Kleinberg, onde um algoritmo não-global restringe o conhecimento dos vértices, e o segundo um processo de navegação global onde permitimos os vértices a terem conhecimento absoluto da estrutura da rede, e com isso, fazendo com que o valor do tempo médio de viagem da mensagem (ou viajante) seja equivalente ao mínimo caminho médio da rede. Verificamos que para ambos os processos, a navegação ótima é obtida para valores do expoente de agregação muito próximos um do outro. Para o algoritmo não global, encontramos $\alpha_e \gtrsim 3$, enquanto que para o

mínimo caminho da rede, que representa a navegação global, encontramos $\alpha = 3$. Esse resultado indica que a introdução do custo τ para adição de novas conexões, modifica sobremaneira o comportamentos de ambas as navegações. Verificamos que tal comportamento está relacionado com a densidade de conexões de longo alcance não ser uma constante no problema. Para valores de $0 \leq \alpha \leq 2$, obtemos que $\rho \sim L^{-1}$, e para $2 \leq \alpha \leq 3$ obtivemos $\rho \sim L^{-(3-\alpha)}$ e, por último, nossos resultados mostram que ρ é praticamente independente do tamanho do sistema quando estamos trabalhando com $\alpha > 3$. Esse comportamento da densidade ρ faz com que um navegante com conhecimento global ou local em uma rede bidimensional, só consiga tirar proveito da estruturas de conexões de longo alcance apenas nas proximidades $\alpha_e = 3$, valor para o qual a configuração da densidade começa a equivaler a configuração do problema sem um vínculo no custo. Os resultados obtidos para redes unidimensional, e para redes bidimensionais onde conexões diagonais são proibidas indicam que esse expoente de agregação efetivo $\alpha = d + 1$, onde d é a dimensão do sistema. Futuramente, podemos estudar o efeito do custo τ em rede onde as conexões de longo alcance possuem um peso relacionado ao seu comprimento, com isso teríamos uma analogia perfeita para redes reais de transporte, como redes representativas de ruas e avenidas em uma cidade.

Referências Bibliográficas

- 1 KLEINBERG, J. M. The Small-World Phenomenon: An Algorithm Perspective. *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, p. 163–170, 2000.
- 2 KLEINBERG, J. M. Navigation on Small-World. *NATURE*, 406, p. 845, 2000.
- 3 WATTS, D.; STROGATZ, S. Collective dynamics of ‘small-world’ networks. *NATURE*, 393, p. 440–442, 1998.
- 4 STROGATZ, S. Exploring complex networks. *NATURE*, 410, p. 268–276, 2001.
- 5 K.PATHRIA, R. *Statistical Mechanics, 2nd ed.* Oxford, UK: Elsevier, 1996.
- 6 LAWRENCE, S.; GILES, C. L. Accessibility of information on the web. *NATURE*, 400, p. 107–109, 1999.
- 7 GILES, C. L.; LAWRENCE, S.; KROVETZ, B. Access to information on the Web. *SCIENCE*, 280, p. 1815, 1998.
- 8 LAOUTSOS, M.; FALOUTSOS, P.; FALOUTSOS, C. On Power Law Relationships of the Internet Topology. *Computer Communications Review*, 29, 1999.
- 9 JEONG, H. et al. The large-scale organization of metabolic networks. *NATURE*, 407, p. 651–654, 2000.
- 10 SCOTT, J. *Social Network Analysis: A Handbook*. London: Sage Publications, 2000.
- 11 WASSERMANN, S.; FAUST, K. *Social Network Analysis*. Cambridge: Cambridge University Press, 1994.
- 12 SOARES, D. J. B. et al. Preferential Attachment growth and nonextensive statistical mechanics. *EUROPHYSICS LETTERS*, 70, p. 70–76, 2005.
- 13 ERDŐS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae*, 6, p. 290–297, 1959.
- 14 MILGRAM, S. The Small-World Problem. *Psychology Today*, 1, 1967.
- 15 DODDS, P. S.; MUHAMAD, R.; WATTS, D. J. An Experimental Study of Search in Global Social Networks. *SCIENCE*, 301, p. 827–829, 2003.
- 16 WATTS, D. J. *Small Worlds*. Princeton, NJ: Princeton Press, 1999.
- 17 PRICE, D. J. S. Networks of Scientific Papers. *SCIENCE*, 149, p. 510–515, 1965.
- 18 PRICE, D. J. S. A general Theory of Bibliometric and Other Cumulative Advantage Processes. *J. Amer. Soc. Inform. Sci.*, 27, p. 292–306, 1976.

- 19 ALBERT, R.; BARABASI, A. Statistical mechanics of complex networks. *REVIEWS OF MODERN PHYSICS*, 74, p. 47–97, 2002.
- 20 ALBERT, R.; BARABASI, A. Emergence of Scaling in Random Networks. *SCIENCE*, 286, p. 509–512, 1999.
- 21 NEWMAN, M. The structure and function of complex networks. *SIAM REVIEW*, 45, p. 167–256, 2003.
- 22 MORENO, J. L. *Who Shall Survive?* Beacon, NY: Beacon House.
- 23 GALASKIEWICS, J. *Social Organization of an Human Grants Economy*. New York: Academic Press, 1985.
- 24 KLOVDAHL, A. S. et al. Social Networks and Infectious Disease: The Colorado Springs Study. *SOC. SCI. MED.*, 38, p. 79–88, 1994.
- 25 LILJEROS, F. et al. The Web of Human Sexual Contacts. *NATURE*, 411, p. 907–908, 2001.
- 26 MORRIS, M. Sexual Networks and HIV. *AIDS 97: YWEAR IN REVIEW*, 11, p. 209–216.
- 27 CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L. Cambridge & New York: McGraw-Hill Book Company & The MIT press, 1990.
- 28 PASTOR-SATORRAS, R.; VAZQUEZ, A.; VESPIGNANI, A. Dynamical and correlation properties of the Internet. *PHYSICAL REVIEW LETTERS*, 87, p. 258701,, 2001.
- 29 NEWMAN, M. E. J. Assortative Mixing in Networks. *PHYSICAL REVIEW LETTERS*, 89, p. 208701, 2002.
- 30 BOCCALETTI, S. et al. Complex networks: Structure and dynamics. *PHYSICS REPORTS-REVIEW SECTION OF PHYSICS LETTERS*, 424, p. 175–308, 2006.
- 31 DOROGOVTSSEV, S. N.; MENDES, J. F. F.; SAMUKHIN, A. N. Structure of Growing Networks with Preferential Linking. *PHYSICAL REVIEW LETTERS*, 85, p. 4633–4636, 2000.
- 32 SIMKIN, M. V.; ROYCHOWDHURY, V. P. Read Before you Cite. *COND-MAT/0212043*, 2002.
- 33 REDNER, S. How Popular is your Paper? An Empirical Study of Citation Distribution. *EUR. PHYS. J. B*, 4, p. 131–134, 1998.
- 34 ROBERSON, M. R.; BEN-AVRAHAN, D. Kleinberg Navigation in Fractal Small-World Networks. *PHYSICAL REVIEW E*, 74, p. 017101, 2006.
- 35 KOMIDIS, K.; HAVLIN, S.; BUNDE, A. Structural properties of Spatially embedded Networks. *EUROPHYSICS LETTERS*, 82, p. 48005, 2008.
- 36 MOUKARZEL, C. F.; MENEZES, M. Argollo de. Shortest Paths on Systems with Power-Law Distributed Long-Range Connections. *PHYSICAL REVIEW E*, 65, p. 056709, 2002.
- 37 MURRAY, J. D. *Mathematical Biology*. Berlin: Springer Verlag, 1993.