

DISSERTAÇÃO DE MESTRADO

CONTRIBUIÇÕES AO ESTUDO DAS REDES
COMPLEXAS: MODELO DE QUALIDADE

Gabriel Alves Mendes

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE

NATAL, MARÇO DE 2007

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

CONTRIBUIÇÕES AO ESTUDO DAS REDES COMPLEXAS:
MODELO DE QUALIDADE

Dissertação de mestrado submetido à Universidade
Federal do Rio Grande do Norte
sob orientação do Professor **Dr. Luciano Rodrigues da Silva**
para a obtenção do título de
Mestre em Física por **Gabriel Alves Mendes**.

2007

AGRADEÇO

Ao meu orientador, professor **Luciano Rodrigues da Silva**, pela orientação e pelo apoio em todo o período em que estive no programa de pós graduação de física da Universidade Federal do Rio Grande do Norte.

Ao meu pai, **Renio dos Santos Mendes**, pelos conselhos e comentários que possibilitaram o enriquecimento deste trabalho.

Ao professor **Danyel Judson Bezerra Soares**, pelas discussões que contribuíram para minha formação e para a realização deste trabalho.

Ao professor **Adriano de Oliveira Sousa** pelas discussões e sugestões que possibilitaram o enriquecimento desse trabalho.

Aos amigos do DFTE: **Thatyara Freire de Souza**, **Sharon D. da Cunha**, **Marcelo D. S. de Menezes** e **Eliade Ferreira Lima**.

À todos os professores e funcionários do Departamento de Física Teórica.

À CAPES pelo apoio financeiro.

Resumo

Neste trabalho veremos as implicações em utilizar uma distribuição de qualidade obedecendo uma lei de potência na dinâmica de crescimento de uma rede estudada por Bianconi e Barabási. Em particular, começaremos nossos estudos pelas redes aleatórias, que caracterizam ou estão relacionadas com algumas situações reais, como por exemplo o movimento das marés. Neste contexto de redes complexas abordaremos várias redes reais e definiremos alguns conceitos importantes no seu estudo. Na sequência, abordaremos o primeiro modelo de rede de escala livre, o qual foi proposto por Barabási et al., e um modelo modificado por Bianconi e Barabási que incorpora na ligação preferencial as diferentes habilidades (qualidades) dos sítios na competição por ligações. Ao final, apresentaremos nossos resultados, discussões e conclusões.

Abstract

In this work we analyse the implications of using a power law distribution of vertice's quality in the growth dynamics of a network studied by Bianconi and Barabási. In particular, we start studying the random networks which characterize or are related to some real situations, for instance the tide movement. In this context of complex networks, we investigate several real networks, as well as we define some important concepts in the network studies. Furthermore, we present the first scale-free network model, which was proposed by Barabási et al., and a modified model studied by Bianconi and Barabási, where now the preferential attachment incorporates the different ability (fitness) of the nodes to compete for links. At the end, our results, discussions and conclusions are presented.

Conteúdo

| | |
|--|-----------|
| Introdução | 1 |
| 1 Conceitos e topologia das redes reais | 4 |
| 1.1 O que são redes? | 4 |
| 1.2 Distribuição de conectividade | 7 |
| 1.3 Coeficiente de agregação | 10 |
| 1.4 Redes de mundo pequeno | 12 |
| 1.5 Betweenness | 14 |
| 1.6 Redes reais | 16 |
| 1.6.1 6 graus de separação | 16 |
| 1.6.2 World Wide Web | 18 |
| 1.6.3 Internet | 18 |
| 1.6.4 Rede de colaborações de atores de cinema | 20 |
| 1.6.5 Rede de relações sexuais | 21 |
| 1.6.6 Rede celular | 22 |
| 1.6.7 Rede ecológica | 22 |
| 1.6.8 Rede de chamadas telefônicas | 23 |
| 1.6.9 Rede de citações | 23 |
| 1.6.10 Redes linguísticas | 24 |
| 2 Modelos teóricos | 27 |
| 2.1 Modelos clássicos | 27 |
| 2.1.1 Modelo de Erdős-Rényi | 27 |
| 2.1.2 Modelo de Watts e Strogatz | 36 |
| 2.2 Redes livres de escala | 39 |
| 2.2.1 Modelo de Barabási e Albert | 39 |

| | | |
|----------|---|-----------|
| 2.2.2 | Ingredientes para gerar uma rede com uma distribuição de conectividade em lei de potência | 43 |
| 2.2.3 | Propriedades do modelo de Barabási e Albert | 46 |
| 3 | Modelo de qualidade nas redes complexas | 48 |
| 3.1 | Introdução | 48 |
| 3.2 | Resultados | 55 |
| 3.2.1 | Modelo livre de escala | 55 |
| 3.2.2 | Distribuição de qualidade uniforme | 56 |
| 3.2.3 | Distribuição de qualidade em lei de potência | 62 |
| | Conclusão | 72 |
| | Referências bibliográficas | 74 |

Introdução

Redes complexas descrevem uma larga variedade de sistemas ([1],[2],...[34]) de alta tecnologia e importância intelectual. Por exemplo, a célula é melhor descrita como uma rede complexa de conexões químicas conectadas por reações químicas [21]; a Internet é uma rede complexa de roteadores e computadores ligados através de várias ligações físicas ou não físicas (sem fio) ([1],...[11]); idéias espalhadas numa rede social, cujos nós são seres humanos e cujas ligações representam várias relações sociais; a WWW (World Wide Web) é uma enorme rede virtual de páginas conectadas por "hyperlinks". Estes sistemas representam alguns dos muitos exemplos que tem induzido a comunidade científica a investigar os mecanismos que determinam a topologia das redes complexas. O desejo de entender sistemas entrelaçados tem encontrado desafios significantes. Mas há alguns anos tem crescido as ferramentas da Mecânica Estatística oferecendo uma melhor forma para descrever fenômenos como os que foram citados. Com esse desenvolvimento tem sido possível solucionar e criar problemas mais desafiadores.

Tradicionalmente o estudo de redes complexas tem sua origem na teoria de grafos. Desde os anos 50, grandes redes aparentemente sem princípios organizacionais tem sido descritas como grafos aleatórios, proposta como a forma mais simples e direta de rede complexa. Os grafos aleatórios foram estudados primeiramente pelos matemáticos húngaros Paul Erdős e Alfréd Rényi [35]. No modelo de Erdős-Rényi, começamos com N nós e todo par é conectado com probabilidade p , criando uma rede com aproximadamente $pN(N - 1)/2$ ligações distribuídas aleatoriamente. Este modelo tem guiado nossos pensamentos sobre redes complexas por décadas desde sua introdução. Mas o interesse crescente em sistemas complexos tem feito com que alguns cientistas reconsiderassem este modelo paradigmático e refletissem sobre uma questão simples: Redes reais comportam-se como sistemas complexos? A

Internet é fundamentalmente aleatória? Nossa intuição claramente indica que sistemas complexos devem mostrar alguns princípios organizacionais, que devem satisfazer em algum nível sua topologia. Mas se a topologia de suas redes vem de fato dos grafos aleatórios, nós precisamos desenvolver ferramentas de medição para obter em termos quantitativos os princípios organizacionais fundamentais. Antes de continuar nossa introdução, faremos uma breve discussão de três conceitos que ocupam uma posição proeminente no pensamento sobre redes complexas.

Mundo pequeno: O conceito de mundo pequeno descreve simplesmente o fato de que mesmo uma rede possuindo um elevado número de nós, existe um caminho relativamente pequeno entre dois nós quaisquer. A distância entre dois nós é definida como o número de ligações ao longo do menor caminho conectando eles. A manifestação mais popular de pequenos mundos são os "seis graus de separação", encontrada pelo psicólogo Stanley Milgram (1967), que concluiu existir um caminho de conhecimento (pessoas que se conhecem) com um tamanho típico de cerca de seis passos entre um par qualquer de pessoas nos Estados Unidos. As propriedades de mundo pequeno aparecem para caracterizar várias redes complexas: os atores em Hollywood estão em média ligados entre si por três co-estrelas, ou a ligação química numa célula está tipicamente separada por três reações químicas.

Agregação: Uma propriedade comum de uma rede social é a formação de "panelinhas", representando círculos de amigos ou seja todos membros estão ligados entre si. Esta tendência inerente para agrupar é quantificada pelo coeficiente de agregação, um conceito que tem sua raiz na sociologia. O coeficiente de agregação da rede é a média de todos os C_i 's individuais.

Num grafo aleatório, desde que as ligações são distribuídas aleatoriamente, o coeficiente de agregação é constante. Entretanto, na maioria das redes reais o coeficiente de agregação é muito maior que os das redes aleatórias (isto é, tendo o mesmo número de nós e ligações). Esta é uma das razões pela qual é necessário estudar e propor outros modelos de redes complexas.

Distribuição de conectividade: Não são todos nós na rede que possuem o mesmo número de conexões. A distribuição de conectividade dos nós é caracterizada pela função distribuição $P(k)$, que fornece a probabilidade que o nó selecionado aleatoriamente tenha k ligações. Desde que as conexões num grafo aleatório são posicionadas aleatoriamente, a maioria dos nós tem apro-

ximadamente a mesma conectividade, próxima da média $\langle k \rangle$ da rede. A distribuição de conexões de um grafo aleatório é uma distribuição de Poisson com um pico em $k = \langle k \rangle$. Uma das descobertas mais interessantes de redes complexas foi que para grandes redes a distribuição de conectividade desvia significativamente da distribuição de Poisson. Em particular, para um número grande de redes, incluindo a WWW, a Internet e a rede metabólica, a distribuição de conectividade tem uma cauda em lei de potência,

$$P(k) \sim k^{-\gamma}. \quad (1)$$

Redes como essas são ditas **livres de escala**.

Esta descoberta iniciou um restabelecimento do modelamento da rede há alguns anos, resultando na introdução e estudo de três classes principais da modelagem paradigmática. Primeiro, grafos aleatórios, que são variantes do modelo de Erdős-Rényi, são ainda largamente usados em alguns campos e servem como base para algumas modelagens e estudos empíricos. Segundo, motivado pelo coeficiente de agregação, uma classe de modelos, coletivamente chamados de modelos de mundo pequeno, tem sido propostos. Estes modelos interpolam a alta agregação das redes regulares e a topologia dos grafos aleatórios. Por último, a descoberta da distribuição em lei de potência tem estimulado construções e análise de vários modelos livres de escala que focalizam a dinâmica da rede, de modo a oferecer uma teoria universal da evolução de redes.

A dissertação esta organizada em três capítulos como segue: No capítulo 1, **Conceitos e topologia das redes reais**, abordaremos alguns dos conceitos mais usados em redes complexas bem como as propriedades topológicas destes sistemas. No capítulo seguinte, **Modelos teóricos**, discutiremos dois modelos clássicos: o de Erdős-Rényi e o de Watts-Strogatz (estes apresentam distribuição de conectividade com uma escala típica). Além destes modelos, apresentaremos um modelo livre de escala sugerido por Barabási e Albert. Este possui uma distribuição de conectividade em lei de potência. No terceiro capítulo, **Modelo de qualidade nas redes complexas**, mostraremos um modelo que generaliza o modelo de Barabási e Albert, bem como o de Bianconi-Barabási. Neste, mostraremos as motivações e os resultados obtidos. Para finalizar o trabalho, apresentaremos nossas conclusões e perspectivas.

Capítulo 1

Conceitos e topologia das redes reais

O estudo da maioria das redes complexas foi motivado pelo desejo de entender vários sistemas reais que vão das redes de comunicação às redes ecológicas. Assim a base de dados avaliada para estudo são de várias disciplinas. Por hora, revisaremos os conceitos básicos e os sistemas que tem sido estudados por pesquisadores objetivando características gerais das redes complexas como menor caminho médio, coeficiente de agregação e a distribuição de conectividade.

1.1 O que são redes?

Do ponto de vista formal, uma rede (ou grafo) é um conjunto de vértices (sítios ou nós) conectados via linhas (arestas ou arcos). Redes com arestas sem direção são conhecidas como *undirect networks*, redes com arestas com direção são conhecidas como *direct networks*. Um sítio muito conectado é chamado de pólo ou "hub". Nessa dissertação quando falarmos de ligação subentenderemos que esta não apresenta direção (praticamente toda dissertação) caso contrário avisaremos.

O número total de conexões de um vértice é chamado de sua *conectividade* k . Na rede com ligações dirigidas, o número de arestas entrando no vértice é chamado sua conectividade de entrada k_i (i de in), o número de arestas saindo do vértice é chamado sua conectividade de saída k_o (o de out). Assim, $k = k_i + k_o$.

Um caso particularmente importante de grafos são as árvores, que são grafos sem "loops" (conhecidas também como circuitos) podendo ser de qualquer ordem. Um circuito é ilustrado na figura 1.1 (sítios azuis). Se uma árvore não tem partes separadas, é chamada *árvore conectada*. O número total de vértices, N , e arestas, L , numa árvore conectada é relacionado pela seguinte expressão $L = N - 1$. Em geral, o número I de circuitos num grafo conectado arbitrariamente com arestas sem direção é relacionado ao número da suas arestas e vértices, por

$$I = L + 1 - N. \quad (1.1)$$

Esta é uma das fórmulas básicas na teoria dos grafos.

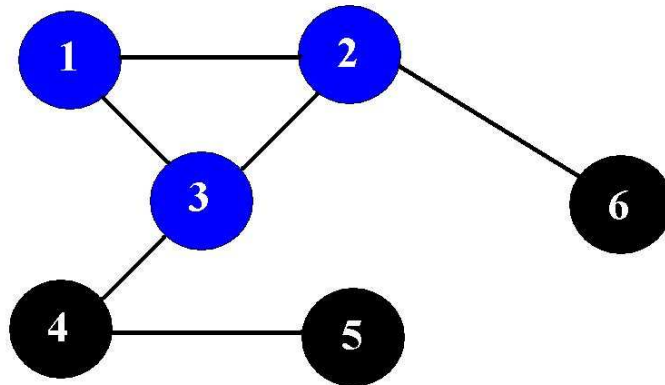


Figura 1.1: A figura ilustra a presença de um circuito formado pelos sítios azuis. Neste caso este é de ordem 3.

Em termos gerais, redes aleatórias são redes com uma arranjo desordenado das arestas. Note que, em geral, na teoria dos grafos o significado de uma rede aleatória é muito mais complicado. Enquanto falamos sobre redes aleatórias, temos que manter na cabeça que é uma rede particular, ou seja é um membro do *ensemble estatístico* de todas as realizações possíveis. Desta forma, quando falamos sobre redes aleatórias, estamos falando sobre ensemble estatístico. A descrição estatística completa de uma rede aleatória sugere a descrição dos ensembles estatísticos correspondentes. Em princípio, redes aleatórias podem conter vértices com conectividade fixa. Em geral, a conectividade dos vértices estão estatisticamente distribuídas.

Do ponto de vista físico, as redes aleatórias podem ser *estáticas* ou *dinâmicas*. Vamos introduzir essas idéias usando alguns exemplos simples.

(1) O exemplo de uma rede aleatória estática (*grafo aleatório clássico*). Este grafo é definido pela seguinte regra:

- (a) O número total de vértices é fixado.
- (b) Escolhe-se aleatoriamente o par de vértices que serão conectados.

Podemos por definição proibir múltiplas conexões mas isso não é necessário para grafos grandes onde conectividades altas estão quase ausentes. Podemos dizer que os vértices dos grafos clássicos aleatórios são estatisticamente equivalentes e independentes. Este modelo simples foi proposto por Erdős e Rényi (1959, 1960) [35]. O procedimento de construção deste grafo pode ser através da adição em sequência de novas arestas entre pares de vértices escolhidos aleatoriamente. Quando o número total de vértices é fixado, este procedimento produz configurações de equilíbrio.

(2) Um exemplo de rede aleatória evoluindo no tempo: é um grafo aleatório simples crescendo através de adições simultâneas de vértices e arestas. O algoritmo deste grafo é:

- (a) Em cada tempo, um novo vértice é adicionado ao grafo.
- (b) Simultaneamente, um par (ou alguns pares) de vértices escolhidos aleatoriamente são conectados por uma aresta.

Notemos que o sistema não está no equilíbrio. As arestas não são distribuídas homogeneamente no grafo. Os nós "mais velhos" são os mais conectados enquanto os sítios novos são pouco conectados. Se em algum momento pararmos de aumentar o número de vértices, mas continuarmos com a adição aleatória de arestas, então a rede tenderá para o "estado de equilíbrio" sem nunca chegar a este estado. De fato, se a conectividade da rede não desaparece, então a não homogeneidade sobrevive. Um "estado de equilíbrio" pode ser aproximado somente se, permitirmos que as arestas mais velhas desapareçam de tempo em tempo.

(3) Outro exemplo de uma rede dinâmica é a *rede de citações*. Vejamos uma versão bem simples dada pelo seguinte algoritmo:

- (a) Em cada tempo, um novo vértice é adicionado ao grafo.
- (b) Este é conectado com algum nó anterior por uma aresta sem direção.

Vejam os que, inicialmente, existirá um sítio sem conexão.

Observemos que se usarmos uma regra diferente para a conexão entre os nós geraremos uma rede de citações diferente.

1.2 Distribuição de conectividade

Como foi explicado anteriormente, a conectividade dos vértices em redes aleatórias estão distribuídas estatisticamente (na dissertação trataremos apenas de ligações sem direção). Seja os vértices distinguíveis, que é a situação padrão no crescimento das redes, introduziremos uma distribuição de conectividade $p(k, s, N)$ para cada vértice. Esta é a probabilidade que o vértice s na rede de tamanho N tenha k conexões (k vizinhos mais próximos). Conhecendo a distribuição de conectividade de cada vértice numa rede, fica fácil encontrar a distribuição de *conectividade total*

$$P(k, N) = \sum_{s=1}^N p(k, s, N). \quad (1.2)$$

Se todos os vértices de uma rede aleatória são estatisticamente equivalentes, como num grafo aleatório clássico, cada um deles terá a mesma distribuição de conectividade $P(k, N)$.

O primeiro momento da distribuição, conectividade média de uma rede, é dado por $\bar{k} = \sum_k kP(k)$. O número total L de arestas numa rede é igual a $\bar{k}N/2$.

Apesar de estudarmos nessa dissertação ligações sem direção, por completeza teórica discutiremos um pouco sobre as ligações que apresentam direções. Analogamente, para vértices individuais numa rede com arestas com direção, introduziremos uma distribuição de conectividade de entrada $p^{(i)}(k_i, s, N)$ e a distribuição de conectividade de saída $p^{(o)}(k_o, s, N)$. Podemos também definir uma distribuição de conectividade total de entrada $P^{(i)}(k_i, N)$ e a distribuição de conectividade total de saída $P^{(o)}(k_o, N)$. Contudo, é muito mais informativo se juntarmos as distribuições de conectividade de entrada e de saída, ficando com $p(k_i, k_o, s, N)$ e $P(k_i, k_o, N)$. Então,

$$P^{(i)}(k_i) = \sum_{k_o} P(k_i, k_o), \quad (1.3)$$

$$P^{(o)}(k_o) = \sum_{k_i} P(k_i, k_o), \quad (1.4)$$

$$P(k) = \sum_{k_i} P(k_i, k - k_i) = \sum_{k_o} P(k - k_o, k_o), \quad (1.5)$$

onde a rede de tamanho N não está explicitamente indicada. Relações semelhantes podem ser escritas para distribuições de nós individuais. Se a rede não tem conexões com o exterior, então a conectividade média entrando e saindo são iguais:

$$\bar{k}_i = \sum_{k_i} k_i P(k_i, k_o) = \bar{k}_o = \sum_{k_o} k_o P(k_i, k_o) = \bar{k}/2. \quad (1.6)$$

A distribuição de conectividade é a característica mais simples e fácil de se obter numa rede. Devido a esta facilidade e a importância dessa quantidade, existe uma grande número de trabalhos que calculam a distribuição de conectividade total (ou conectividade de entrada, ou conectividade de saída). Ela caracteriza somente as propriedades locais de uma rede, mas essa pouca informação sobre a estrutura da rede é suficiente para determinar suas propriedades básicas, as quais serão discutidas durante toda a dissertação. De fato, o problema principal no estudo empírico de redes é a fraca estatística, além do pequeno tamanho das redes. Em geral, temos somente uma única realização de uma rede aleatória real para obter dados a partir dela. Quando fazemos o gráfico da distribuição de conectividade $P(k)$ usando dados empíricos, indicamos o número de vértices com uma conectividade k e então fazemos uma normalização apropriada. Ao unirmos a distribuição de conectividade de entrada e de saída $P(k_i, k_o)$, usamos duas variáveis, k_i e k_o , isso produz flutuações mais pronunciadas.

Quando os vértices de um grafo são estatisticamente independentes (equivalentes), suas conexões são aleatórias e a distribuição de conectividade determina completamente as propriedades estatísticas de uma rede. Impondo que a distribuição de conectividade $P(k)$ está presente (ou seja, apesar de inicialmente todos sítios não possuírem ligações, estamos impondo a condição que cada sítio possuirá um certo número de ligações máxima, que obedecem a distribuição de conectividade escolhida) e se o número total N de nós está fixado, como poderemos construir um grafo que obedeça essa condição?

- (1) Rotulando N vértices.
- (2) Para o vértice j de um grafo atribuímos conectividade k_j a partir de uma distribuição de conectividade $P(k)$. Agora cada vértice será conectado a outro vértice.
- (3) Conectamos aleatoriamente cada par de vértices.

Esta é somente uma realização particular de um grafo, mas dessa forma, é possível construir um ensemble estatístico completo. Semelhantemente, grafos aleatórios no equilíbrio são generalizações naturais de grafos aleatórios clássicos.

Os exemplos seguintes demonstram casos típicos de distribuições de conectividade observadas no estudo de redes complexas.

- (1) *Distribuição de Poisson:*

$$P(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!}. \quad (1.7)$$

Sendo, \bar{k} a conectividade média dada por $\bar{k} = \sum_{k=0}^{\infty} kP(k)$. Um grafo aleatório clássico tem sua distribuição de conectividade quando seu número de vértices tende para o infinito e a conectividade média é fixa (vínculo).

- (2) *Distribuição exponencial:* $P(k) \propto e^{-k/\bar{k}}$. Esta é a distribuição de conectividade de um grafo crescido aleatoriamente descrito na seção anterior, no exemplo (2). Igualmente para redes com tamanhos infinitos, todos os momentos da distribuição (1) e (2) são finitos, $M_m \equiv \sum_{k=0}^{\infty} k^m P(k) < \infty$.

- (3) *Distribuição em lei de potência:* $P(k) \propto k^{-\gamma}$ com $k \neq 0$ e γ é o expoente da distribuição.

A distribuição em lei de potência contrasta com as distribuições de Poisson e exponencial, pois esta não tem nenhuma escala natural (tamanho típico). Redes com essa distribuição são chamadas *livres de escala*.

Em redes com tamanhos infinitos (**não existe fisicamente**), todos momentos de ordem $m \geq \gamma - 1$ da distribuição em lei de potência divergem. A partir disso, podemos ver quais valores do expoente γ são possíveis para redes de escala livre (é claro que γ deve ser maior que 1 para ser normalizável):

- (a) Se a conectividade média da rede de escala livre (primeiro momento da distribuição de conectividade) é finita, então seu expoente γ deve ser maior que 2 ou seja $2 < \gamma < \infty$.

(b) Um exemplo em que a conectividade média diverge é quando o número total de arestas de uma certa rede cresce mais rápido que uma função linear do número total de vértices. Isto ocorre para algumas rede reais se a distribuição de conectividade é estacionária, $1 < \gamma \leq 2$.

Em redes com tamanhos finitos, a distribuição de conectividade apresenta uma cauda larga. Lembrando que todas as redes reais possuem tamanhos finitos. A maior rede de escala livre conhecida é a WWW, contendo somente 10^9 vértices, isto é muito menos do que o número de partículas num sistema macroscópico em sistemas físicos estatísticos e da matéria condensada. De fato, redes reais e simuladas, ambas não são objetos macroscópicas mas *mesoscópicos*, e o efeito do tamanho são importantes à primeira vista.

(4) *Distribuição discreta*: Esse espectro de conectividade é típico em redes crescendo deterministicamente [33], [32], [36],[37] e [38]. Em nossa simulações realizadas nessa dissertação sempre obtemos uma distribuição discreta.

1.3 Coeficiente de agregação

O coeficiente de agregação caracteriza a "densidade" de conexões próximas aos arredores do vértice. Seja uma rede com arestas sem direção, e uma destas arestas tenha z vizinhos próximos. A agregação máxima no aglomerado é obtida, quando todas $z(z - 1)/2$ arestas possíveis estão presentes (todos os sítios vizinhos estão ligados). Convencionalmente ([39],[40] e [41]), o coeficiente de agregação C de um vértice é o quociente entre o número total y de arestas conectando seus vizinhos mais próximos e o número possível de arestas entre todos esses vizinhos,

$$C = \frac{2y}{z(z - 1)}. \quad (1.8)$$

Esta grandeza reflete o grau de conhecimento mútuo de seus "amigos" mais próximos. Por exemplo, observe a figura 1.2.

Podemos introduzir a distribuição de C numa rede, mas em geral somente o valor médio de \bar{C} é considerado, que é o coeficiente de agregação de uma rede. \bar{C} expressa a probabilidade que exista uma aresta entre dois vizinhos de

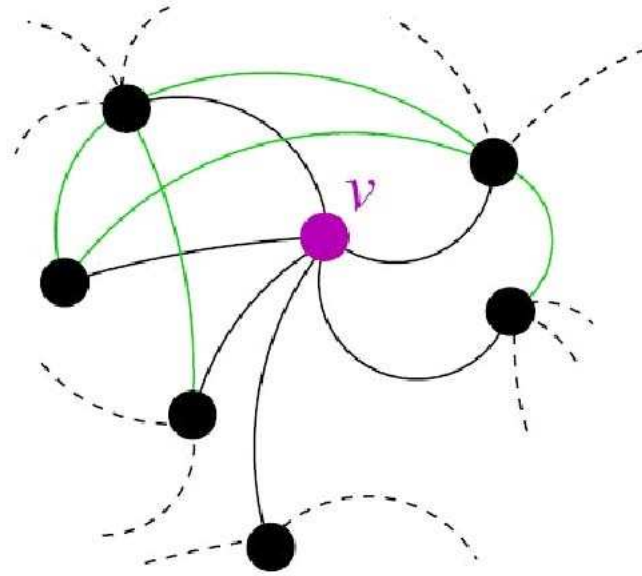


Figura 1.2: Vizinhança de ν com $k_\nu = 6$, resultando $c_\nu = 1/3$.

um vértice escolhido aleatoriamente. Podemos dizer que a agregação média \bar{C} mostra "a densidade" de pequenos circuitos de tamanho 3 numa rede. Então, a noção de agregação é diretamente relacionada aos circuitos. O coeficiente de agregação de uma árvore é zero por definição. A estrutura de uma rede com o valor de \bar{C} alto é muito diferente do tipo árvore. A presença de circuitos é uma forma específica de correlação em redes. Se o coeficiente de agregação de uma rede "infinita" não tende a zero, então a correlação entre vértices de uma rede estão certamente presentes.

Podemos ver que o coeficiente de agregação de uma rede completamente conectada é igual a 1. No grafo aleatório clássico constituído de N vértices aleatoriamente conectados por L arestas, cada par de vértices é conectado com a mesma probabilidade $p \cong z/N$. Aqui $z = \bar{k} = 2L/N$. Então, o coeficiente de agregação de um grafo aleatório clássico é $\bar{C} = z/N$ o que é minúsculo para grandes redes. Grafos aleatórios clássicos tem pouquíssimos circuitos, porém esse não é o caso para algumas redes reais.

1.4 Redes de mundo pequeno

A maioria das redes são compactas. O termo "mundo pequeno" ou "efeito de mundo pequeno" está presente em redes científicas [46]. A compactação, aqui, significa a pequenez do "tamanho linear" de uma rede. Distâncias entre pares de sítios numa rede podem ser medidos usando a seguinte regra: primeiramente, suponha que uma rede possua arestas sem direção e que estas tenham uma unidade de comprimento. Assim, a distância entre dois sítios de uma rede é o comprimento do menor caminho (caminho geodésico) entre eles. As distâncias l entre os pares de vértices de uma rede aleatória são distribuídas com alguma função de distribuição $\mathcal{P}(l)$. Esta é a probabilidade de que o menor caminho entre dois vértices escolhidos aleatoriamente seja igual a l . Assim se quisermos calcular o "tamanho linear" de uma rede, basta usarmos a função $\mathcal{P}(l)$ que é uma das principais características estruturais de uma rede. Para distribuições que diminuem rapidamente, a distância característica é da ordem do menor caminho médio $\bar{l} \equiv \sum_l l\mathcal{P}(l)$. Outra distância característica natural é o tamanho do mais longo caminho existente numa rede, l_{max} .¹ É claro que ambos \bar{l} e l_{max} dependem do número total N de vértices numa rede.

Olhemos o problema por outro ponto de vista. Seja $z \equiv z_1$, o número de vizinho mais próximos, z_2 , o número de segundos vizinhos mais próximos, z_3 , o número de terceiros vizinhos mais próximos, e assim por diante. Em geral, existem z_m vértices numa m -ésima esfera centrada num sítio, $z_0 = 1$. A distância entre um sítio e qualquer um dos seus m -ésimos vizinhos mais próximos é igual a m , e esta é uma forma equivalente para introduzir a noção de distância numa rede. Podemos introduzir uma distribuição de números de m -ésimos vizinhos mais próximos, $P_m(z)$. Esta é a probabilidade que um sítio escolhido aleatoriamente de uma rede tenha m -ésimos vizinhos mais próximos. Obviamente, as distribuições $\mathcal{P}(l)$ e $P_m(z)$ são quantidades relacionadas e como esses valores dependem dos vizinhos mais próximos, logo o menor caminho é uma generalização natural da distribuição de conectividade da rede.

Agora, demonstraremos como estimar o menor caminho médio (tamanho

¹Em diferentes trabalhos, ambos: o comprimento do menor caminho médio e o tamanho do mais longo caminho médio são chamados por *diâmetro* de uma rede.

linear) para uma rede não correlacionada. Se uma rede aleatória é grande, então é habitual que tenha uma *estrutura local tipo árvore*. Isto significa que circuitos (caminho fechados numa rede) estão quase ausentes na vizinhança de um vértice. A estrutura local tipo árvore em redes é um dos pontos-chaves da teoria de redes. Por hora não discutiremos condições para estas realizações, mas chamaremos a atenção que em geral é suficiente que a distribuição de conectividade caia suficientemente rápido. Assim, numa situação contrastante, quando uma fração finita do número total de arestas está anexada a um número finito de vértices, a rede certamente não é do tipo árvore.²

Num outro caso, sem entrar em detalhes, vamos fazer uma simples estimativa para uma rede tipo árvore. Consideremos uma rede aleatória com conectividade média \bar{z} . Ou seja, um sítio qualquer pode visitar em média z outros sítios afastados a uma distância de um passo, veja a figura 1.3. De uma forma geral, podemos falar que um sítio pode visitar \bar{z}^l outros sítios com l passos. Assim, podemos estimar a distância típica da rede, isto é o menor caminho médio \bar{l} , usando a relação $N \sim \bar{z}^{\bar{l}}$.

Obtendo a expressão bem conhecida para \bar{l} ,

$$\bar{l} \approx \frac{\ln N}{\ln \bar{z}}. \quad (1.9)$$

Portanto redes com muitos nós possuem um pequeno valor para o menor caminho médio. Esta notável pequenez é em geral referida como "efeito de mundo pequeno" ([39] e [40]), presente em todas as redes abordadas nesse trabalho.

Para uma rede com arestas que tenham direção, a situação é mais complexa. Neste caso, podem existir dois menores caminhos diferentes entre um par de pontos. O menor caminho de um ponto a outro não coincide com o menor caminho na direção oposta e seus comprimentos não coincidem necessariamente. Podemos ignorar a direção das arestas e introduzir o menor caminho médio \bar{l} para essa rede onde esta sempre terá um valor menor ou igual ao do menor caminho médio de uma rede com arestas dirigidas.

²Notemos que uma estrutura de uma componente gigantesca conectada de uma rede aleatória desvia do tipo árvore. Note também que redes com um alto coeficiente de agregação não são do tipo árvore.

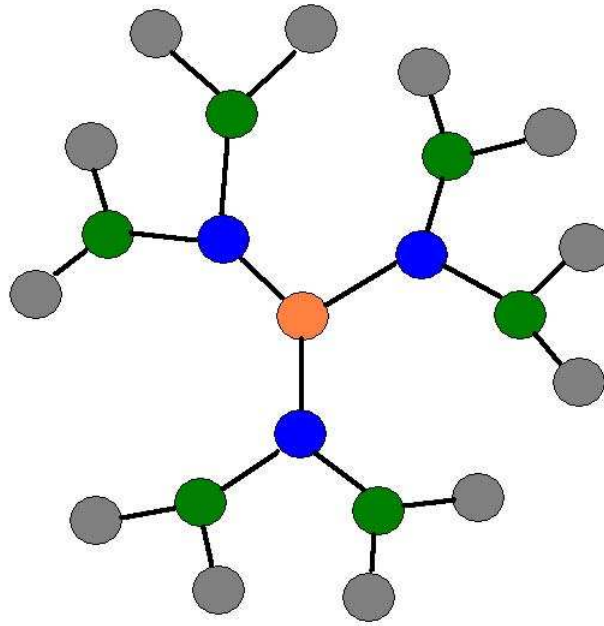


Figura 1.3: Exemplo de uma rede com $\bar{k} = 3$. Pode-se notar que quando $l = 1$, temos $N = 3^1 = 3$ nós visitados. Quando $l = 2$, temos $N = 3^2 = 9$. Quando $l = 3$, temos $N = 3^3 = 27$ e assim por diante ($N \sim \bar{k}^l$).

1.5 Betweenness

O menor caminho é uma das noções centrais na ciência das redes. Mostraremos brevemente outras características estruturais padrões de redes que estão baseadas nesta noção.

Em termos gerais, o "betweenness" $\sigma(m)$ mede a quantidade de menores caminhos que passam através do sítio m . Se o número de menores caminhos entre um par de vértices exceder este valor, então o caminho passando através do vértice m contribui para a "betweenness" com um correspondente peso reduzido. Notemos que todos os pares com o vértice m também entram na conta. Esta quantidade foi introduzida na sociologia [47] para caracterizar a regra "social" de um vértice. Vértices com um grande "betweenness"³ são mais "influentes". De fato, este indica se um vértice é ou não importante no tráfego de uma rede. Por exemplo, se todos pares de vértices de uma rede

³Esta quantidade também é chamada de *load* ([47] e [48]) ou "betweenness centrality" de um vértice.

comunicam com uma mesma taxa e o tráfego vai pelo menor caminho, então o tráfego através de um vértice coincide com o "betweenness" desse vértice. Podemos introduzir outras propriedades a partir do "betweenness" como a distribuição de "betweenness" dos vértices, o "betweenness" médio de um vértice numa rede, e assim por diante.

Se o número total de menores caminhos entre os vértices i e j é $B(i, j) > 0$, e $B(i, m, j)$ deles passam através do vértice m , então a taxa $B(i, m, j)/B(i, j)$ mostra o quão fundamental o vértice m é na regra das conexões entre i e j . Por exemplo, se nenhum dos $B(i, j)$ menores caminhos passam através de m , então pela regra, $\sigma(m)$ para i e j é zero. Contrariamente, se a maioria dos $B(i, j)$ caminhos passam através de m , então m se torna extremamente importante, pois o vértice m controla quase todos os menores caminhos médios da rede. O "betweenness" $\sigma(m)$ do vértice m é definido por:

$$\sigma(m) \equiv \sum_{i \neq j} \frac{B(i, m, j)}{B(i, j)}, \quad (1.10)$$

onde a soma é sobre todos pares de vértices e com a condição que pelo menos um caminho exista, isto é, com $B(i, j) > 0$. Os vértices com alto "betweenness" controlam a rede, como esta ilustrado na figura 1.4. Como é de se esperar, isto sugere que o "betweenness" de um vértice está fortemente correlacionado com sua conectividade.

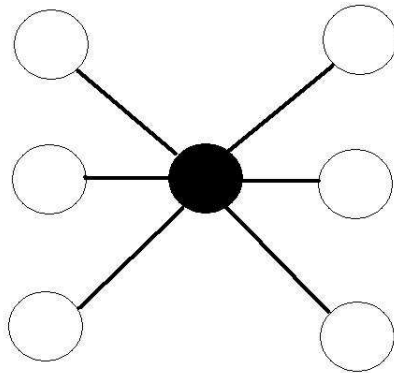


Figura 1.4: Nesta ilustração, podemos perceber que o sítio central é de extrema importância para a rede. Pela definição de "betweenness", esse nó possui o maior "betweenness" desse grafo.

Similarmente, podemos definir "betweenness" para uma aresta e dessa forma detectar e indexar a estrutura hierarquica das redes de forma natural. Uma rede é decomposta em pares de sub redes separadas por sucessivas eliminações de arestas com "betweenness" máximos. Este procedimento é repetido para cada sub rede resultante sucessivamente.

1.6 Redes reais

1.6.1 6 graus de separação

Stanley Milgram [46], psicólogo e sociólogo americano, desenvolveu uma experiência em 1967 com o intuito de saber se era possível que duas pessoas escolhidas aleatoriamente nos Estados Unidos pudessem se conectar através de pessoas conhecidas e se caso fosse possível, encontrar a quantidade de pessoas necessárias para tal feito ("distância").

A pesquisa foi feita da seguinte forma: Stanley Milgram escolheu duas pessoas alvo. Uma dessas duas pessoas era um corretor da bolsa de valores e a outra era a esposa de um estudante de graduação, ambos em Massachusetts. A experiência consistia em enviar envelopes para pessoas aleatoriamente residindo em Nebraska e Kansas (áreas desérticas) que deveriam enviar o envelope à pessoa alvo, ver figura (1.5).

O envelope continha um papel com um pequeno sumário contendo o objetivo do estudo, uma fotografia, nome, endereço, alguns dados pessoais da pessoa alvo e as seguintes instruções:

- (1) Cada pessoa que receber o caderno deve colocar seu nome nele. Isto serve para rastrear a origem do envelope, evitando que este retorne para mesma pessoa.
- (2) Se você não conhece diretamente a pessoa alvo, não tente contactá-la diretamente. Repasse este caderno para algum conhecido que você ache que tenha maior probabilidade de conhecer o destinatário.
- (3) Se você conhece a pessoa alvo, mande este documento diretamente para ele(a).



Figura 1.5: Mapa dos Estados Unidos. Os estados em vermelho (Nebraska e Kansas) representam o ponto de saída das correspondências e o estado em verde (Massachusetts) representa o destino final.

(4) Ao enviar o papel, cada pessoa deve arrancar uma página do caderno, preenchê-la e remetê-la ao cientista que acompanhava o avanço da experiência. Quando o caderno alcançar a pessoa alvo, este deverá enviar o caderno a Milgram para completar a experiência.

O resultado surpreendente foi publicado na revista *Psychology Today* com o nome "problema de mundo pequeno" e mostrou que o número médio de ligações entre a pessoa inicial e o destinatário era aproximadamente de seis pessoas, ou seja seis graus de separação. Como foi visto numa das seções anteriores, este resultado ilustra uma propriedade conhecida no estudo de redes como "mundo pequeno".

1.6.2 World Wide Web

A WWW representa a segunda maior rede conhecida e conta com aproximadamente um bilhão de sítios ([49] e [50]). Os nós da rede são os documentos ("web page") e as conexões são os "hyperlinks" (URL's) que apontam de um documento para outro. O interesse na WWW como uma rede explodiu depois da descoberta que a distribuição de conectividade das "web pages" segue uma lei de potência [1]. As conexões da WWW tem direção e a rede é caracterizada por duas distribuições: a distribuição de conexões saindo, $P_{out}(k)$, significa que o documento tem k "hyperlinks" que fazem sair deste, e a distribuição de conexões entrando, $P_{in}(k)$, é a probabilidade que k "hyperlinks" apontem para o documento. Alguns estudos tem estabelecido que ambos P_{out} e P_{in} tem cauda em lei de potência:

$$P_{out}(k) \sim k^{-\gamma_{out}} \quad e \quad P_{in}(k) \sim k^{-\gamma_{in}}$$

Albert, Jeong, e Barabási (1999) [1] foram os pioneiros no estudo da WWW (contendo 325729 nós) e encontraram $\gamma_{out} = 2,45$ e $\gamma_{in} = 2,1$, ver figura 1.6. Kumar *et al.* (1999) [51], usaram 40 milhões de documentos obtendo $\gamma_{out} = 2,38$ e $\gamma_{in} = 2,1$. Depois Broder *et al.* (2000) [3], usaram 200 milhões de documentos, obtendo $\gamma_{out} = 2,72$ e $\gamma_{in} = 2,1$. Notemos que γ_{in} é o mesmo para todas medidas apesar da quantidade de documentos ter variado nesses dois anos. Neste período a rede cresceu pelo menos cinco vezes. Entretanto, γ_{out} tem aumentado com o tamanho da amostra ou com o tempo. Um outro ponto interessante é a existência do efeito de mundo pequeno. Como foi relatado por Albert, Jeong, e Barabási (1999) que encontraram que o menor caminho médio para uma amostra de 325729 nós é 11,2.

1.6.3 Internet

A Internet é uma rede de ligações físicas entre computadores e outros aparelhos de comunicação. A topologia da Internet é estudada em dois níveis diferentes. Um a nível de roteadores onde os nós são os roteadores, e as ligações são conexões físicas entre eles. O outro a nível de interdomínios (ou sistemas autônomos), onde cada domínio é composto por milhares de roteadores e computadores, sendo representado por um único nó, e a ligação

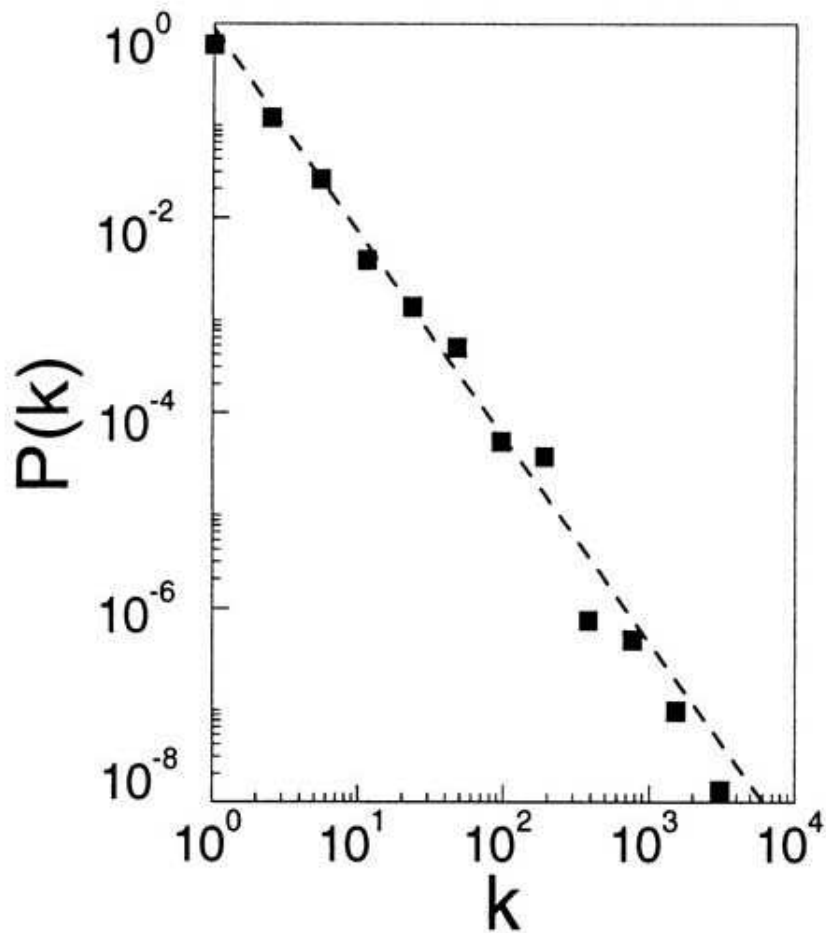


Figura 1.6: Rede WWW com 235.729 sítios, $\langle k \rangle = 5,46$ e $\gamma_{in} = 2,1$. Figura retirada da ref.[1].

é feita entre dois domínios se existir pelo menos um roteador que os ligue. Ambos os níveis estudados pertencem a classe de problemas que seguem uma lei de potência. A Internet como uma rede apresenta o coeficiente de agregação e o menor caminho médio, ambos pequenos. Yook *et al.* (2001) [52], e Pastor-Satorras *et al.* (2001) [10] estudando a Internet a nível de domínio entre 1997 e 1999, encontraram que o coeficiente de agregação está entre 0,18 e 0,3, sendo comparável com $C_{rand} \simeq 0,001$ para redes aleatórias com parâmetros similares (tamanhos comparáveis). O menor caminho médio da Internet a nível de domínio está entre 3,70 e 3,77 e a nível de roteadores é cerca de 9, indicando seu carácter de mundo pequeno.

1.6.4 Rede de colaborações de atores de cinema

Uma base de dados muito estudada é a rede de colaboração de atores de cinema, baseada no banco de dados que tem na Internet sobre cinema, que contem todos filmes e seus elencos desde 1890. Nesta rede, os nós são atores, e dois nós são ligados se os atores correspondentes atuarem no mesmo filme. Esta é uma rede em expansão contínua [16], com 225226 nós em 1998, que cresceu para 449913 nós até maio de 2000. O menor caminho médio de atores da rede é próximo a de um grafo aleatório com o mesmo tamanho e conectividade média, 3,65 comparado com 2,9, mas seu coeficiente de agregação é cem vezes maior que a do grafo aleatório. A distribuição de conectividade da rede de atores de filme é uma lei de potência para k grande, que segue $P(k) \sim k^{-\gamma_{ator}}$, onde $\gamma_{ator} = 2,3 \pm 0,1$. Ver a figura 1.7.

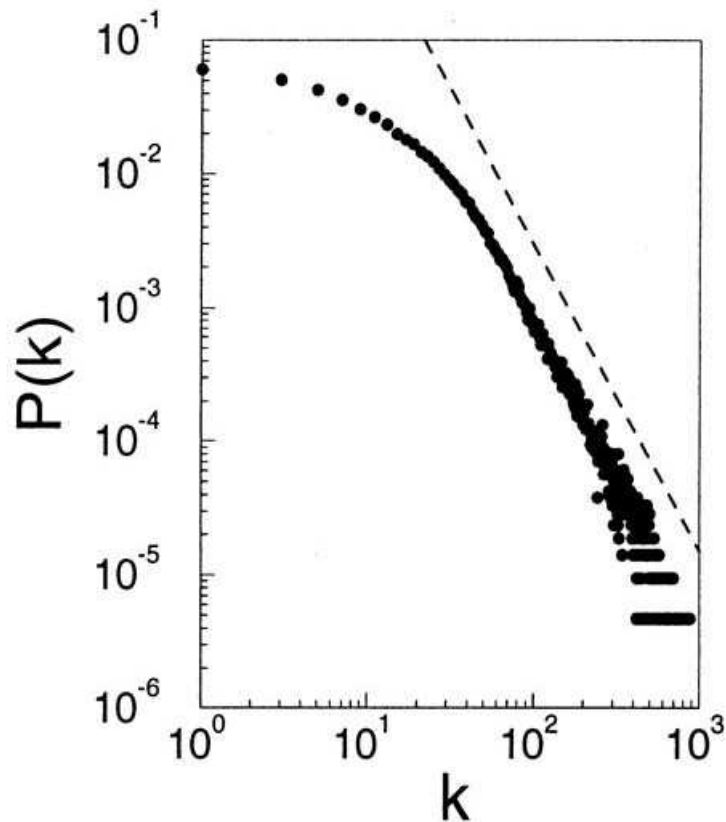


Figura 1.7: Rede de colaborações de atores de cinema com 212.250 sítios, $\langle k \rangle = 28,78$ e $\gamma_{ator} = 2,3$. Figura retirada da ref. [16].

1.6.5 Rede de relações sexuais

Algumas doenças transmitidas sexualmente, incluindo a AIDS, se propagam através da rede de relações sexuais. Liljeros *et al.* (2001) [20], estudaram uma rede formada por 2810 indivíduos, onde foi analisada a distribuição de parceiros sexuais de um grupo em um ano, obtendo uma distribuição de conectividade em lei de potência tanto para homens quanto para mulheres, com o expoente $\gamma_{Males} = 3,3 \pm 0,2$ e $\gamma_{Females} = 3,5 \pm 0,2$ respectivamente. Como esta rede é de escala livre, esperamos que as doenças se propaguem rapidamente o que de fato ocorre. Dessa forma, campanhas para evitar a disseminação das doenças sexualmente transmissíveis devem ser voltadas especialmente para indivíduos com um grande número de parceiros sexuais. Ver a figura 1.8.

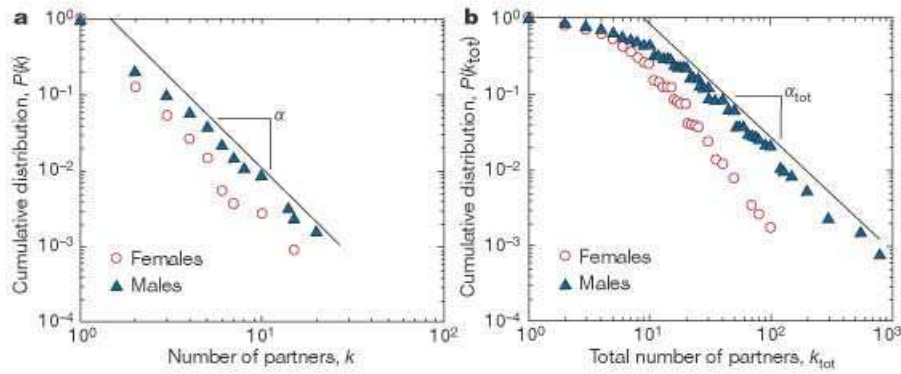


Figura 1.8: Distribuição de conectividade em lei de potência. **a)** Distribuição do número de parceiros durante os últimos 12 meses. **b)** Distribuição do número de parceiros durante toda a vida. Figura retirada da ref. [20].

1.6.6 Rede celular

Jeong *et al.* (2000) [21], estudaram o metabolismo de 43 organismos, representados em três domínios da vida, e organizaram numa rede em que os nós são os substratos (como ATP, ADP, H₂O) e a conexão representa a direção predominante da reação química em que esses substratos podem participar. A distribuição de saída e de entrada é em lei de potência para todos organismos, com o expoente variando entre 2,0 e 2,4. O coeficiente de agregação não foi calculado pois a rede tem direção e quanto ao menor caminho médio foi 3,3 aproximadamente para todos organismos.

Outra rede importante caracterizando a célula, descreve as interações proteína-proteína, onde os nós são proteínas e elas são conectadas se já foi demonstrado experimentalmente que há interações entre elas. O estudo dessas interações físicas mostra que a distribuição de conectividade destas seguem uma lei de potência com uma exponencial cortada $P(k) \sim (k + k_0)^{-\gamma} e^{-(k+k_0)/k_c}$ com $k_0 = 1$, $k_c = 20$, e $\gamma = 2,4$ (Jeong *et al.*, 2001 [23]).

1.6.7 Rede ecológica

Cadeias alimentares são usadas regularmente por ecologistas para quantificar a interação entre várias espécies. Numa cadeia alimentar os nós são espécies e as ligações representam as relações presa-predador. Williams *et al.* (2000) [53], investigaram a topologia das sete maiores cadeias documentadas (Skipwith Pond, Little Rock Lake, Bridge Brook Lake, Chesapeake Bay, Ythan Estuary, Coachella Valley, e St. Martin Island). Estas cadeias são pequenas (a maior delas tem 186 nós) e diferem largamente no número de espécies ou na sua conectividade média, porém elas indicam que espécies em seus habitats interagem com três ou menos espécies. As cadeias alimentares apresentaram um alto coeficiente de agregação, presença de "hubs" e a distribuição de conectividade em lei de potência com o valor de $\gamma = 1,1$.

1.6.8 Rede de chamadas telefônicas

As maiores redes estudadas são: o cérebro humano com seus 10^{11} neurônios (não é livre de escala) e a WWW (10^9 páginas). Existe uma rede com um número menos impressionante que é a rede de chamadas telefônicas que liga duas pessoas em qualquer lugar do mundo. Atualmente o número de telefones no mundo é menor que 10^9 mas esta crescendo rapidamente.

Surpreendentemente, a estrutura global e topológica das chamadas telefônicas praticamente não tem sido estudadas. Uma das principais causas do pouco conhecimento dessa rede é devido às companhias telefônicas não fornecerem muitos dados por razões de privacidade.

Um grafo com ligações dirigidas tem sido construído com chamadas telefônicas padrões de longa distância, onde os nós são números telefônicos e cada chamada telefônica completa é uma ligação, dirigida de quem faz a chamada para quem recebe. Aiello, Chung, e Lu (2000) [29] estudaram o gráfico de chamadas de longa distância feitas durante um único dia, e foi observada uma distribuição de conectividade de saída quanto de entrada seguindo uma lei de potência com o expoente $\gamma_{out} = \gamma_{in} = 2, 1$.

1.6.9 Rede de citações

Uma outra rede pode ser formada pelas citações de artigos científicos. Os nós são os artigos publicados e a direção da ligação representa uma referência para um artigo publicado anteriormente. Redner (1998) [12], estudando a distribuição de citações de 783339 trabalhos catalogados pelo Institute for Scientific Information e 24296 trabalhos publicados na Physical Review D entre 1975 e 1994, encontrou que a probabilidade do trabalho estar citado k vezes segue uma lei de potência com o expoente $\gamma = 3$. Um estudo recente de Vázquez (2001) [15], estendeu esses resultados para a distribuição de conectividade saindo, obtendo uma cauda exponencial.

1.6.10 Redes linguísticas

A complexidade das línguas humanas oferece algumas possibilidades para definir e estudar redes complexas. Recentemente Ferrer i Cancho e Solé (2001) [54], construíram uma rede para a língua inglesa, baseada no British National Corpus, onde os nós são as palavras. Estes são ligados quando uma palavra estiver ao lado ou no máximo a distância de uma palavra na sentença. Para uma rede de 440902 palavras, foi obtido um pequeno valor para o menor caminho médio $l = 2,67$, um alto coeficiente de agregação $C = 0,437$ e um regime duplo da distribuição de conectividade. Palavras com conectividade $k \leq 10^3$ decaem com o expoente $\gamma_{<} = 1,5$, enquanto palavras com $10^3 < k < 10^5$ seguem uma lei de potência com $\gamma_{>} \simeq 2,7$.

Um estudo diferente (Yook, Jeong, e Barabási, 2001b) [52], conecta as palavras baseada nos seus significados, isto é, duas palavras foram conectadas uma com a outra se forem sinônimos de acordo com o dicionário Merriam-Webster. Os resultados indicam a existência de um aglomerado gigante de 22311 palavras de um total de 23279 que são sinônimos, com um menor caminho médio $l = 4,5$, e um alto coeficiente de agregação $C = 0,7$. Além do mais, a distribuição de conectividade segue uma lei de potência com $\gamma = 2,8$. Estes resultados indicam que as línguas formam uma rede complexa com princípios organizacionais não tão diferentes dos exemplos discutidos anteriormente. Outro estudo sobre redes linguísticas foi realizada usando as sílabas da língua portuguesa por M.M. Soares et al. [34]. Nesta, os nós das redes são as sílabas e as ligações são realizadas se duas sílabas pertencem a mesma palavra. O coeficiente de agregação obtido nesta rede foi alto enquanto o menor caminho médio foi pequeno. A distribuição de conectividade da rede silábica segue uma lei de potência com o expoente $\gamma \simeq 1,4$.

Na figura 1.9 estão dispostos os expoentes dinâmicos das redes reais que são livres de escala (dados empíricos). Notemos que a maioria dos expoentes possuem um valor entre 2 e 3. A tabela 1.1 contém alguns dados empíricos das redes apresentadas nesse capítulo.

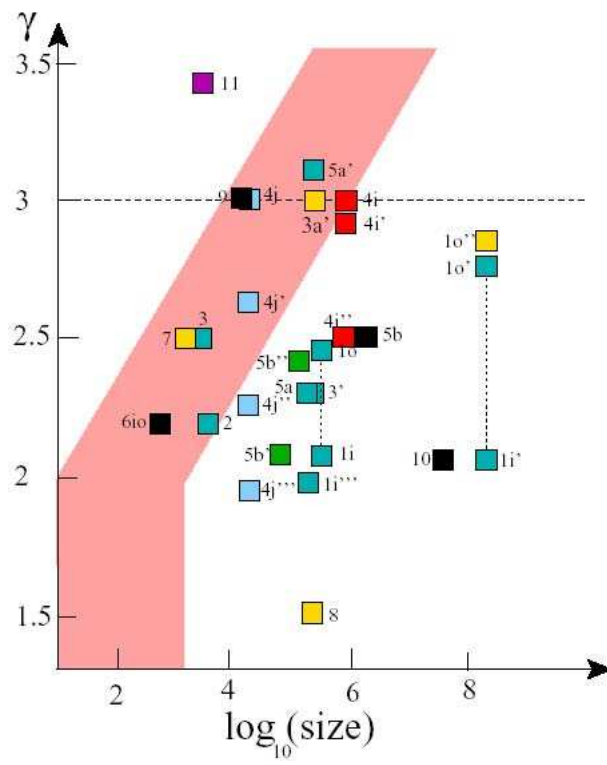


Figura 1.9: Expoente da distribuição de conectividade contendo as redes da tabela 1.1. Figura retirada da ref. [55].

Tabela 1.1: Dados das redes reais

| | Rede ou subgrafo | Número de vértices | Número de arestas | γ | C | C/C_r | \bar{l} | \bar{l}/\bar{l}_r | Refs. |
|----|--|---------------------|---------------------|---|-------------------|--------------------|-----------|---------------------|------------|
| 1 | Mapa completado domínio nd.edu da Web | 325 729 | 1 469 680 | $\gamma_i = 2, 1,$ $\gamma_o = 2, 45$ | - | - | 11,2 | - | [1] |
| 2 | WWW analisado pelo Altvista (1999) | $2, 711 \cdot 10^8$ | $2, 130 \cdot 10^9$ | $\gamma_i = 2, 1,$ $\gamma_o = 2, 7$ | - | - | 16 | 1 | [2],[3] |
| 3 | (outro ajuste) | | | $\gamma_i = 2, 10,$ $\gamma_o = 2, 82$ | - | - | - | - | [4] |
| 4 | Mapa de sites de um domínio da WWW | $2, 60 \cdot 10^5$ | - | $\gamma_i = 1, 94$ | - | - | - | - | [5] |
| 5 | Mapa com ligações sem direções da WWW | 153 127 | $2, 70 \cdot 10^6$ | $\gamma_i = 2, 05$ | 0,108 | $0, 47 \cdot 10^3$ | 3,1 | 0,93 | [6] |
| 6 | Um conjunto de home pages | 4923 | $1, 335 \cdot 10^7$ | $\gamma_i = 2, 05$ | - | - | - | - | [7] |
| 7 | Outro conjunto de home pages | - | - | $\gamma_i = 2, 63$ | - | - | - | - | [7] |
| 8 | Conjunto de home pages de uma universidade | - | - | $\gamma_i = 2, 66,$ $\gamma_o = 2, 82$ | - | - | - | - | [7] |
| 9 | Conjunto de home pages de cientistas | 56880 | - | $\gamma_i = 2, 82$ | - | - | - | - | [7] |
| 10 | Internet (1998) | 4389 | 8256 | $\gamma_i = 2, 2$ | - | - | 4 | 0,6 | [9] |
| 11 | Internet (1999) | 6374 | 13 641 | $\gamma_i = 2, 2$ | 0,24 | $3, 3 \cdot 10^2$ | 3,7 | 0,58 | [10] |
| 12 | Internet a nível de roteadores (1995) | 3888 | 5012 | $\gamma_i = 2, 5$ | - | - | 12,1 | 1,39 | [9] |
| 13 | Internet a nível de roteadores (2000) | 150 000 | 200 000 | $\gamma_i = 2, 3$ | - | - | 10 | 0,8 | [11] |
| 14 | Citações | 783 339 | 6 716 198 | $\gamma_i = 3$ | - | - | - | - | [12] |
| 15 | (outro ajuste) | | | $\gamma_i = 2, 9$ | - | - | - | - | [13] |
| 16 | (outro ajuste) | | | $\gamma_i = 2, 5$ | - | - | - | - | [14] |
| 17 | Citações no Physical Review D | 24 296 | 351 872 | $\gamma_i = 3$ | - | - | - | - | [12] |
| 18 | (outro ajuste) | | | $\gamma_i = 2, 6$ | - | - | - | - | [13] |
| 19 | (outro ajuste) | | | $\gamma_i = 2, 3$ | - | - | - | - | [14] |
| 20 | (outro ajuste) | | | $\gamma_i = 1, 9$ | - | - | - | - | [15] |
| 21 | Rede de colaboração de atores de cinema | 212 250 | 61 085 555 | $\gamma_i = 2, 3$ | - | - | 4,54 | 1,25 | [16] |
| 22 | (outro ajuste) | | | $\gamma_i = 3, 1$ | - | - | - | - | [17] |
| 23 | Rede de colaboração | 1 388 989 | $1, 028 \cdot 10^7$ | $\gamma_i = 2, 5$ | 0,066 | $6 \cdot 10^3$ | 4,6 | 0,9 | [18] |
| 24 | Rede de co-autores | 56 627 | $4, 898 \cdot 10^6$ | $\gamma_i = 1, 2$ | 0,726 | $0, 24 \cdot 10^3$ | 4,0 | 1,88 | [18] |
| 25 | Rede de colaboração | 70 975 | $0, 132 \cdot 10^6$ | $\gamma_i = 2, 1$ | 0,59 | $1, 1 \cdot 10^4$ | 9,5 | 1,16 [19] | [19] |
| 26 | Rede de colaboração | 209 293 | $1, 214 \cdot 10^6$ | $\gamma_i = 2, 4$ | 0,76 | $1, 4 \cdot 10^4$ | 6 | 1,2 | [20] |
| 27 | Rede de relações sexuais | 2810 | - | $\gamma_i = 3, 4$ | - | - | - | - | [20] |
| 28 | Rede de reações metabólicas | $\sim 200 - 800$ | $\sim 600 - 3000$ | $\gamma_i = 2, 2,$ $\gamma_o = 2, 2$ | 0,32 | 12 | 3,2 | 0,95 | [21], [22] |
| 29 | Rede de interações entre proteínas | 1870 | 2240 | $\gamma_i = 2, 5$ | 0,022 | 4,4 | 6,8 | 0,8 | [23], [24] |
| 30 | Cadeia alimentar | 470 000 | $17 000 000$ | $\gamma_i = 2, 7$ | 0,44 | $2, 8 \cdot 10^3$ | 2,65 | 0,87 | [25] |
| 31 | Cadeia alimentar do parque Silwood | 154 | 366 | $\gamma_i = 1$ | 0,15 | 5 | 3,4 | 1,05 | [26] |
| 32 | "Java Development Framework" | 1376 | 2174 | $\gamma_i = 2, 5$ | 0,06 | 25 | 6,39 | 1,02 | [27] |
| 33 | Jogo de computador | 1989 | $4, 78 \cdot 10^3$ | $\gamma_i = 2, 85$ | 0,08 | 35 | 6,2 | 1,28 | [27] |
| 34 | Circuitos eletrônicos | $2 \cdot 10^4$ | $4 \cdot 10^4$ | $\gamma_i = 3$ | $3 \cdot 10^{-2}$ | $1, 5 \cdot 10^2$ | 6 | 1 | [28] |
| 35 | Rede de chamadas telefônicas | $47 \cdot 10^6$ | $8 \cdot 10^7$ | $\gamma_i = 2, 1$ | - | - | - | - | [29] |
| 36 | Rede de e-mail | 5165 | $6, 57 \cdot 10^4$ | $\gamma_i = 1, 5$ | 0,156 | $3, 25 \cdot 10^3$ | 4,95 | 0,48 | [30] |
| 37 | "Energy landscape network for a 14-atom cluster" | 4196 | 87 219 | $\gamma_i = 2, 78$ | 0,073 | 7,4 | 2,32 | 1,04 | [31] |

Capítulo 2

Modelos teóricos

Abordaremos nesse capítulo alguns modelos de redes complexas (clássicos e livres de escala) a fim de mostrar as falhas e os sucessos destes na tentativa de descrever da melhor maneira possível as redes reais. A apresentação desse capítulo respeitará a ordem histórica dos modelos estudados, ou seja: Erdős-Rényi, Watts-Strogatz e Albert-Barabási.

2.1 Modelos clássicos

2.1.1 Modelo de Erdős-Rényi

Em seus primeiros artigos clássicos em redes aleatórias, Erdős e Rényi [35], 1959, definiram um grafo aleatório com N nós que são conectados por n arestas, escolhidas aleatoriamente das $\frac{N(N-1)}{2}$ combinações possíveis. Uma definição alternativa e equivalente de um grafo aleatório é o modelo binomial. Aqui iniciamos com N nós, sendo que cada par de nós é conectado com probabilidade p (ver figura 2.1) e a conectividade média é relacionada com p da seguinte forma

$$\langle k \rangle = \frac{2n}{N} = p(N - 1) \simeq pN. \quad (2.1)$$

Consequentemente o número total de arestas é uma variável aleatória com o valor médio $E(N) = p[N(N - 1)/2]$.

Algumas propriedades dos grafos aleatórios podem ser determinadas usando argumentos probabilísticos. Representaremos uma propriedade qualquer de um grafo pela letra Q . Vejamos algumas questões relevantes sobre redes

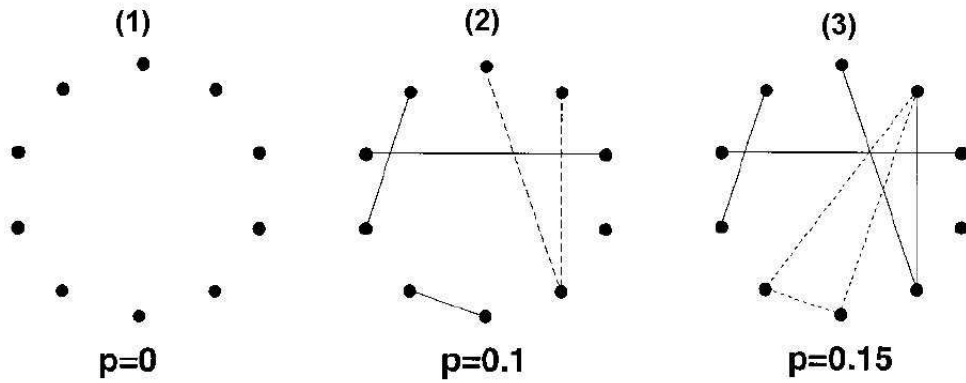


Figura 2.1: Processo de evolução do modelo de Erdős e Rényi para diferentes p 's. (1) Começando com $N = 10$ nós isolados, conectamos cada par de nós com probabilidade p . (2) e (3) mostram dois estágios diferentes no desenvolvimento do grafo, correspondendo a $p = 0,1$ e $p = 0,15$. Podemos perceber o surgimento de árvores (uma árvore de ordem 3, linhas tracejadas) e ciclos (um ciclo de ordem 3, linha tracejada) num grafo, e um aglomerado conectando metade dos nós em $p = 0,15 = 1,5/N$. Figura retirada da ref. [56].

complexas: Um grafo típico é conectado? Ele contém circuitos ("loops")? O diâmetro depende do seu tamanho?

Na literatura matemática, a construção de um grafo aleatório é frequentemente uma evolução de processos: começando com um conjunto de N sítios isolados, o grafo desenvolve pela adição aleatória e sucessiva de arestas aleatórias. Os grafos obtidos em diferentes estágios deste processo correspondem a probabilidades de conexão p cada vez maiores, podendo obter grafos completamente conectados (número máximo de arestas é $n = N(N - 1)/2$ para $p \rightarrow 1$). O principal objetivo da teoria de grafos aleatórios é determinar em qual probabilidade de conexão, p , o sistema apresenta uma transição de fase. A maior descoberta de Erdős e Rényi foi que as redes aleatórias possuem transições de fase. Por exemplo em uma dada probabilidade todos os pares de nós estão quase desconectados um do outro (vários aglomerados pequenos e isolados) e de repente para um certo valor de p isso não ocorre mais (surge um aglomerado gigante).

Observemos que na teoria de grafos aleatórios, a probabilidade de ocupação

é definida como uma função do tamanho do sistema: p representa a fração de arestas que estão presentes das possíveis $N(N - 1)/2$. Deste modo é mais provável que ocorra a presença de circuitos em grafos maiores do que nos menores, se p for fixo em ambos.

A primeira propriedade de grafos aleatórios estudada por Erdős e Rényi foi a presença de sub grafos. Um grafo G_1 consistindo de um conjunto P_1 de nós e um conjunto E_1 de arestas é um sub grafo de um rede $G = \{P, E\}$ se todos nós em P_1 são também nós de P e todas arestas em E_1 são também arestas de E . O exemplo mais simples de sub grafos são circuitos, árvores e sub grafos completos, ver figura 2.1. Um circuito fechado de ordem k possui k arestas em que todos pares de arestas consecutivas tem somente um nó em comum. Ou seja, graficamente um triângulo é um circuito de ordem 3, enquanto um retângulo é um circuito de ordem 4. A conectividade média de um circuito é igual a 2, desde que todos os nós tenham duas arestas. O oposto de circuitos são as árvores, que não formam circuitos fechados. Mais precisamente, um grafo é uma árvore de ordem k se este tem k nós e $k - 1$ arestas, e nenhum de seus sub grafos são circuitos. Sub grafos completos de ordem k contém k nós e todas as possíveis $k(k - 1)/2$ arestas. Em outras palavras, eles estão completamente conectados.

Seja a evolução do processo descrito na figura (2.1) para um grafo $G = G_{N,p}$. Começemos com N nós desconectados, então conectemos cada par de nós com probabilidade p . Para probabilidades de conexão pequenas, as arestas ficam isoladas, e com o aumento do número de arestas, aumenta a probabilidade de duas arestas se ligarem em um mesmo nó, formando uma árvore de ordem 3. Um problema interessante é a determinação da probabilidade crítica $p_c(N)$ em que quase todos grafos G contém uma árvore de ordem 3. Em geral, podemos perguntar se existe uma probabilidade crítica que marca o surgimento de sub grafos arbitrários consistindo de k nós e l arestas.

Na teoria de redes aleatórias, existe uma prova rigorosa feita por Bollobás em 1985 [57] sobre a quantidade de sub grafos existentes numa rede. Consideremos um grafo aleatório $G = G_{N,p}$ e um pequeno grafo F consistindo de k nós e l arestas. Em princípio, o grafo aleatório G pode conter alguns sub grafos F . Nosso primeiro objetivo é determinar quantos sub grafos existem. Os k nós podem ser escolhidos do número total de nós N em C_N^k formas e

l arestas são formadas com probabilidade p^l . Assim, podemos permutar os k nós e obter $k!$ redes novas (o valor correto é $k!/a$, onde a é o número de grafos que são **isomórficos**). Deste modo o número esperado de sub grafos F contidos em G é

$$E(X) = C_n^k \frac{k!}{a} p^l \simeq \frac{N^k p^l}{a}. \quad (2.2)$$

onde X é o número atual de sub grafos. Notemos que as sub redes não tem que estar isoladas.

A equação 2.2 indica que se $p(N)$ é de tal forma que $p(N)N^{k/l} \rightarrow 0$ com $N \rightarrow 0$, o número esperado de sub grafos $E(X) \rightarrow 0$, ou seja, quase nenhum dos grafos aleatórios contem um grafo F . Entretanto, se $p(N) = cN^{-k/l}$, o número médio de sub grafos é um número finito, denotado por $\lambda = c^l/a$, indicando que esta função pode ser a probabilidade crítica. A validade disso, pode ser testada calculando a distribuição de números de sub grafos, $P_p(X = r)$, obtendo-se

$$\lim_{N \rightarrow \infty} P_p(X = r) = e^{-\lambda} \frac{\lambda^r}{r!}. \quad (2.3)$$

A probabilidade que G contenha pelo menos um sub grafo F é

$$P_p(G \supset F) = \sum_{r=1}^{\infty} P_p(X = r) = 1 - e^{-\lambda}, \quad (2.4)$$

que converge para 1 com c crescendo. Para valores de p satisfazendo $pN^{k/l} \rightarrow \infty$ a probabilidade $P_p(G \supset F)$ converge para 1. Assim a probabilidade crítica em que quase todos grafos contem um sub grafo com k nós e l arestas é $p_c(N) = cN^{-k/l}$.

Alguns casos importantes seguem da equação 2.4:

- (a) A probabilidade crítica de ter uma árvore de ordem k é $p_c(N) = cN^{-k/(k-1)}$.
- (b) A probabilidade crítica de ter um circuito de ordem k é $p_c(N) = cN^{-1}$.

É instrutivo olhar os resultados acima de um ponto de vista diferente. Consideremos um grafo aleatório com N nós e que a probabilidade de conexão $p(N)$ varie com N^z , onde z é um parâmetro que pode tomar qualquer valor entre $-\infty$ e 0.

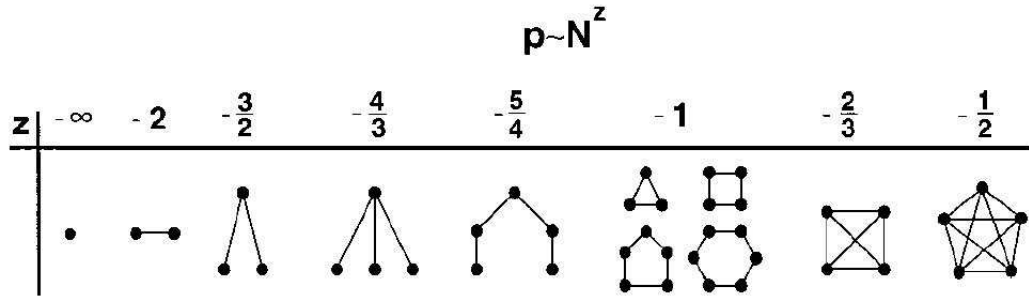


Figura 2.2: As probabilidades limites em que diferentes sub grafos aparecem num grafo aleatório. Para $pN^{-3/2} \rightarrow 0$ árvores de ordem três aparecem. Em $p \sim N^{-4/3}$ surgem árvores de ordem 4. Em $p \sim N^{-1}$ estão presentes árvores de todas ordens, e ao mesmo tempo aparecem circuitos de todas ordens. A probabilidade $p \sim N^{-2/3}$ marca a aparição de sub grafos completos de ordem 4 e $p \sim N^{-1/2}$ correspondendo aos sub grafos completos de ordem 5. Com z tendendo a 0, o grafo contém sub grafos completos de ordens crescentes. Figura retirada da ref. [56].

Para z menor que $-3/2$, quase todos grafos contem somente nós e arestas isoladas. Quando z passa por $-3/2$, árvores de ordem 3 aparecem de repente. Quando z alcança $-4/3$, árvores de ordem 4 aparecem, e com z aproximando de -1 , o grafo contém árvores de ordens cada vez maiores. Entretanto, com $z < -1$, a conectividade média do grafo $\langle k \rangle = pN \rightarrow 0$ com $N \rightarrow \infty$, o grafo é uma união das árvores não conectadas e não apresenta ciclos. Quando z passa exatamente por -1 , correspondendo a $\bar{k} = \text{constante}$, a probabilidade assintótica de ciclos de todas ordens passam de 0 para 1. Os ciclos de ordem 3 podem ser observados como sub grafos completos de ordem 3. Sub grafos completos de ordem 4 aparecem em $z = -2/3$, e com z crescendo, sub grafos completos de ordens cada vez maiores vão aparecendo. Finalmente, com z tendendo a 0, o grafo contem sub grafos completos de todas ordens finitas.

Quando temos $p \propto N^{-1}$ e a conectividade média dos nós é dada por $\langle k \rangle = \text{constante}$, ou seja, $z = -1$ podemos obter alguns resultados: neste caso, um grafo aleatório contém árvores e ciclos de todas ordens. Até agora não discutimos o tamanho e a estrutura de uma componente de um grafo típico. Uma componente de um grafo é por definição um sub grafo conectado

e isolado. Erdős e Rényi num de seus trabalhos mostraram existir uma mudança abrupta na estrutura do aglomerado de um grafo aleatório com $\langle k \rangle$ próximo de 1 como será discutida a seguir.

Se $0 < \bar{k} < 1$, a maioria dos aglomerados são árvores. Apesar dos circuitos estarem presentes, quase todos nós pertencem a estruturas tipo árvore. O número médio de aglomerados é da ordem de $N - n$, onde n é o número de arestas, ou seja, neste intervalo quando uma nova aresta é adicionada, o número de aglomerados diminui de um.

Para $\bar{k} < 1$ o maior aglomerado é uma árvore e ao ultrapassar o valor crítico $\langle k \rangle_c = 1$, a estrutura do grafo varia abruptamente. Percebemos que se \bar{k} cresce, os menores aglomerados colapsarão num aglomerado gigante. Assim os aglomerados menores tem maiores chances de sobreviverem, ou seja não ser "engolidos" pelo aglomerado gigante. Um exemplo prático é: imagine uma rede formada por pessoas representando os sítios e elas estarão ligadas se uma conhece a outra. Assim se $\langle k \rangle = 1$ na rede, não existirá segredo entre ninguém, pois qualquer segredo que você conte a uma dessas pessoas que participam da rede, será passada a todas as outras.

A. Distribuição de conectividade

Erdős e Rényi em 1959 [35] foram os primeiros a estudar a distribuição de conectividade máxima e mínima de um grafo aleatório. Numa rede aleatória com probabilidade de conexão p , a conectividade k_i de um nó i segue uma distribuição binomial:

$$P(k_i = k) = p^k (1 - p)^{N-1-k} C_{N-1}^k. \quad (2.5)$$

onde o primeiro termo é a probabilidade do sítio ter k_i ligações, o segundo termo é a probabilidade de que as outras probabilidades estejam ausentes e o último termo é o número de diferentes combinações em que as ligações podem estar distribuídas. No limite $N \rightarrow \infty$, essa distribuição tende a uma distribuição de Poisson, dada por

$$P(k) \approx e^{-pN} \frac{(pN)^k}{k!} = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}, \quad (2.6)$$

onde $\langle k \rangle$ é a conectividade média e é regida pela seguinte expressão: $\langle k \rangle = p(N - 1)$. Assim, a distribuição decresce rapidamente para valo-

res afastados da conectividade média. O comportamento da distribuição de conectividade para o modelo de Erdős e Rényi esta ilustrado na figura 2.3.

Os resultados indicam que para um grande intervalo de valores de p , as conectividades máximas e mínimas são determináveis e finitas. Outro ponto interessante é que apesar das arestas serem aleatórias a maioria dos nós possuem o mesmo número de arestas.

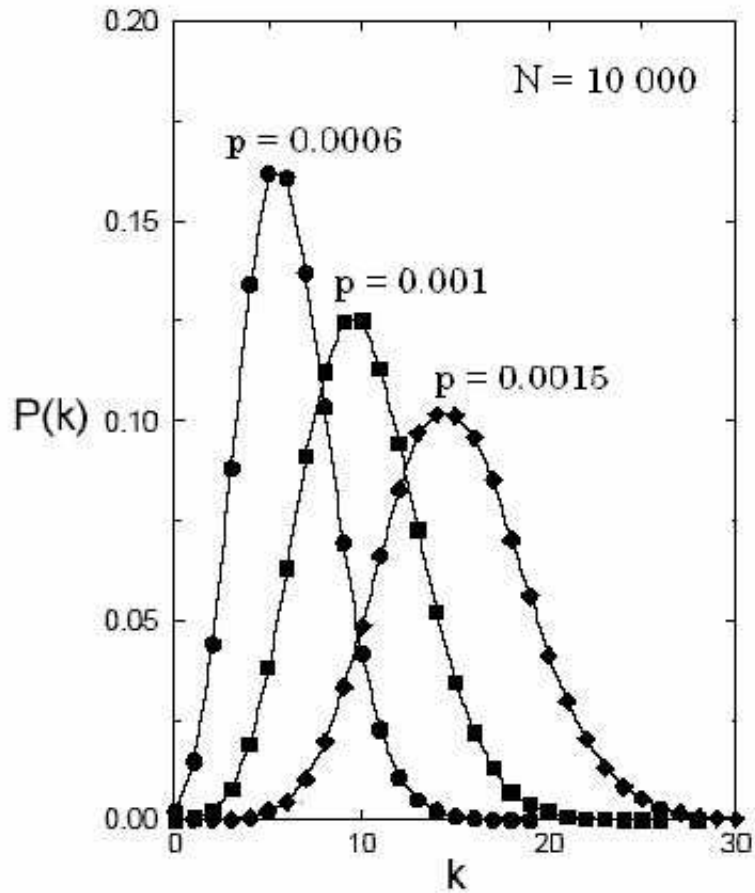


Figura 2.3: Distribuição de conectividade do modelo de Erdős e Rényi, para redes com $N = 10000$ e $p = 0,0006$ (círculos), $p = 0,001$ (quadrados), $p = 0,0015$ (diamantes). Notemos que os sítios possuem aproximadamente o mesmo número de conexões. Figura retirada da ref. [56].

B. Diâmetro da rede

O diâmetro de um grafo é a distância máxima entre qualquer par de nós da rede. Porém, o diâmetro de um grafo desconectado (ou seja, grafos que possuem aglomerados isolados) é infinito, pois não é possível conectar todos pares de nós nessa rede. Por esta razão é mais apropriado definir o diâmetro da rede como o diâmetro máximo dos seus aglomerados. Grafos aleatórios tendem a ter pequenos diâmetros, para p não muito pequeno. A razão disso, é que um grafo é igualmente distribuído com uma probabilidade já definida. Seja uma rede aleatória onde os sítios tenham em média k ligações, $\langle k \rangle$, como na figura (1.3). Isso significa que de um sítio qualquer, podemos visitar em média k outros sítios afastados a distância de um passo. De cada um desses k sítios, podemos visitar k outros sítios. Deste modo podemos visitar $\langle k \rangle^n$ sítios com n passos. Se a rede tiver N sítios, $\langle k \rangle^n$ não pode exceder o tamanho N , ou seja $N = \langle k \rangle^l$. Assim, para alcançar os N sítios da rede são necessários em média,

$$l \simeq \frac{\log N}{\log \langle k \rangle}, \quad (2.7)$$

passos. Isso explica o motivo pelo qual partindo-se de um sítio inicial qualquer, será possível com poucos passos encontrar qualquer sítio na rede e esta foi a primeira tentativa de explicar o problema dos seis graus de separação de Milgram.

Vejam alguns resultados importantes:

- (a) Se $\bar{k} = pN < 1$, uma rede típica é composta de árvores isoladas e seu diâmetro é igual ao diâmetro de uma árvore.
- (b) Se $\bar{k} > 1$, surge um aglomerado gigante. O diâmetro da rede é igual ao diâmetro do aglomerado gigante se $\bar{k} \geq 3, 5$.
- (c) Se $\bar{k} \geq \log(N)$, quase todo grafo esta conectado.

Outra forma de caracterizar o espalhamento de um grafo aleatório é calculando a distância média entre todos pares de nós, ou o tamanho do menor caminho médio. Esperamos que o tamanho do menor caminho médio varie com o número de nós da mesma forma como o diâmetro,

$$l_{rand} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad (2.8)$$

Ao comparar o menor caminho médio das redes reais e o da teoria de grafos aleatórios obtemos uma estimativa razoável (ver figura 2.4).

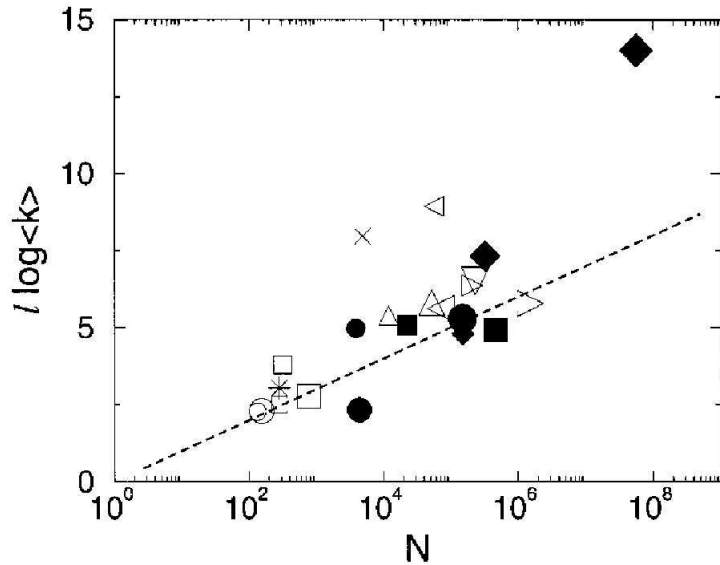


Figura 2.4: Comparação entre o caminho médio de uma rede real (símbolos) e a predição da equação (2.8) da teoria de redes aleatórias (linha pontilhada). Figura retirada da ref. [56].

C. Coeficiente de agregação

Se considerarmos um nó do grafo aleatório e seus primeiros vizinhos, a probabilidade que dois destes vizinhos estejam conectados é igual a probabilidade que dois nós selecionados aleatoriamente estejam conectados. Consequentemente, o coeficiente de agregação de um grafo aleatório é

$$C_{rand} = p = \frac{\langle k \rangle}{N}. \quad (2.9)$$

Na figura (2.5), observamos o gráfico de $C_{real}/\langle k \rangle$ em função dos seus tamanhos e comparamos este com a predição da equação (2.9). O gráfico indica convincentemente que redes reais não seguem a predição dos grafos aleatórios. A razão $C/\langle k \rangle$ não diminui com N^{-1} , ao invés disso, parece ser independente de N .

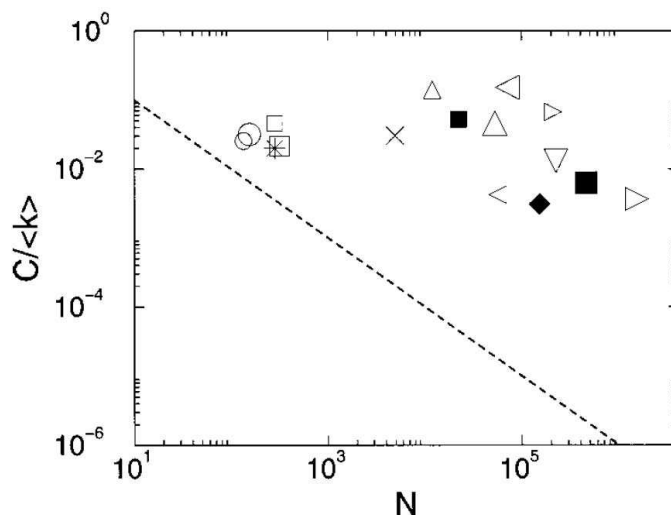


Figura 2.5: Comparação entre o coeficiente de agregação das redes reais e dos grafos aleatórios. A linha tracejada corresponde a equação (2.9). Figura retirada da ref. [56].

2.1.2 Modelo de Watts e Strogatz

Watts e Strogatz (1998) [40] propuseram um modelo que interpola uma rede de dimensão finita ordenada e uma rede aleatória. A razão para isso, é que uma rede ordenada possui um coeficiente de agregação que independe do tamanho da rede (ver figura 2.5) mas depende do número de coordenação. O algoritmo deste modelo é descrito a seguir (ver figura 2.6):

- (1) Inicia-se com um círculo que possui N nós (rede regular) onde cada nó esta conectado aos seus primeiros k vizinhos. Seja $N \gg k \ln(N) \gg 1$ para que todos nós não estejam conectados.
- (2) Reescrevemos aleatoriamente cada aresta da rede com uma probabilidade p (proibimos qualquer auto conexão e conexões duplas).

Este processo introduz ligações de longo alcance que conectam diferentes vizinhanças. Ao variarmos p a partir de zero, rede regular, até $p = 1$ onde a rede é aleatória, podemos observar a transição que ocorre com essa variação.

Esse modelo tem sua origem nos sistemas sociais em que a maioria das pessoas são amigas de seus vizinhos (por exemplo, da mesma rua). Entretanto, todos tem amigos distantes, ou seja de outro bairro, cidade ou país, podemos

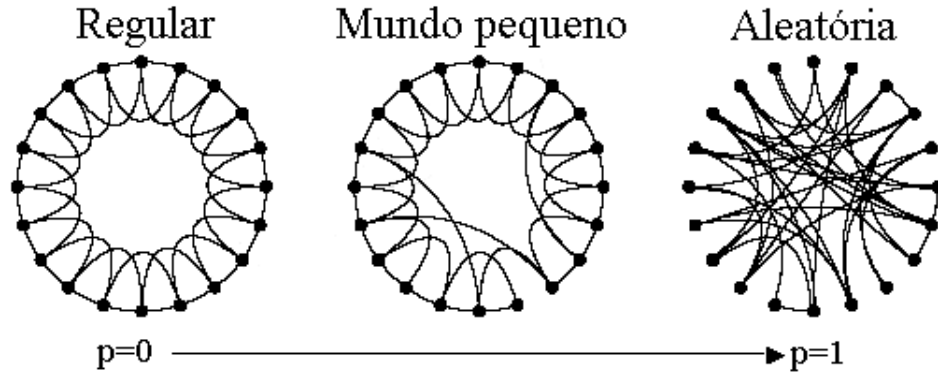


Figura 2.6: A figura mostra três valores de p no modelo de Watts e Strogatz. Inicialmente temos uma rede regular, $p = 0$, com um alto coeficiente de agregação. A seguir com um p intermediário existe o efeito de mundo pequeno, devido as distâncias pequenas entre os sítios e ao alto valor do coeficiente de agregação. Na rede mais a direita, temos uma rede aleatória com $p = 1$. Nessa ilustração fica claro que o aumento do valor de p é acompanhado pelo aumento da desorganização na rede. Figura retirada da ref.[40].

representar esta característica da rede usando o modelo de Watts-Strogatz. Para isto basta reescrever aleatoriamente as ligações da rede regular ($p = 0$) o que gera ligações de longo alcance.

Agora, analisaremos o gráfico 2.7 que ilustra o comportamento do coeficiente de agregação, $C(p)$, e o tamanho do menor caminho médio, $l(p)$. Para uma rede regular, $p = 0$, $l(0) \simeq N/2k \gg 1$ e $C(0) \simeq 3/4$. Notemos que l varia linearmente com o tamanho do sistema, e que o coeficiente de agregação é grande. Por outro lado, para $p \rightarrow 1$ o modelo converge para um grafo aleatório em que $l(1) \sim \ln(N)/\ln(k)$ e $C(1) \sim k/N$ (l varia logicamente com N , e o coeficiente de agregação decresce com N). Estes casos limites sugerem que altos valores de C são sempre associados com altos valores de l e pequenos valores de C com pequenos valores de l . Porém Watts e Strogatz (1998) notaram em seu modelo que para um certo intervalo de p a situação explicada acima não ocorre. Ou seja, existe um regime em que coexiste valores pequenos de l e valores altos de C , esta região apresenta uma boa concordância com as características das redes reais discutidas no capítulo anterior. Além do mais, nesta região a rede é considerada de mundo pequeno.

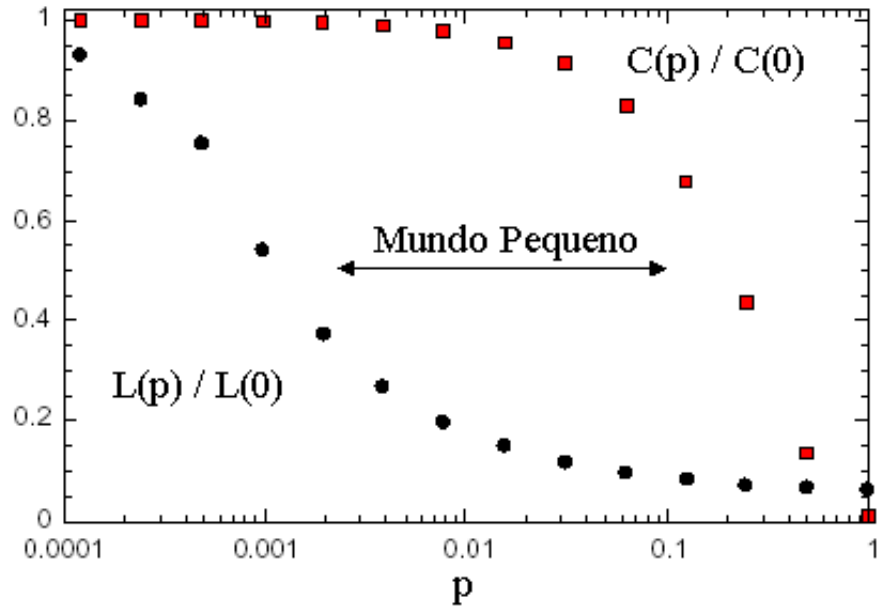


Figura 2.7: Comparação entre o menor caminho médio e o coeficiente de agregação para o modelo de Watts-Strogatz com $N = 10000$. Notemos a rápida queda do menor caminho médio e a constância do coeficiente de agregação para pequenos valores de p , responsável pelo efeito de mundo pequeno na rede. Figura retirada da ref. [40].

2.2 Redes livres de escala

Os resultados empíricos sobre redes reais discutido no capítulo anterior demonstram que várias redes são livres de escala, isto é, sua distribuição de conectividade segue uma lei de potência. Na seção sobre redes clássicas, vimos que a teoria de grafos aleatórios e o modelo de Watts e Strogatz não apresentam esta característica e então surge uma questão: Qual o mecanismo responsável pela característica de redes livres de escala?

2.2.1 Modelo de Barabási e Albert

Os modelos discutidos até agora possuem um número fixo N de vértices que são aleatoriamente conectados ou reodernados. Já no modelo de Barabási e Albert [16], a rede cresce com a adição contínua de novos nós. Começam com um pequeno número de nós e com o passar do tempo, o número de nós aumenta durante a idade da rede por adições sucessivas de sítios novos. Por exemplo: a WWW (World Wide Web), a rede de citações científicas, etc.

Notemos que várias redes reais exibem uma *ligação preferencial*, ou seja, a conexão à um nó depende da conectividade do nó. Por exemplo, uma página da Internet provavelmente terá "hyperlinks" para páginas mais populares (documentos com alta conectividade).

Esses dois ingredientes, adição de sítios novos e ligação preferencial, é a base do modelo de Barabási e Albert que é um dos primeiros modelos teórico a gerar uma distribuição de conectividade em lei de potência. O algoritmo que descreve o modelo de Barabási e Albert é o seguinte:

- (1) Inicia-se a rede com m_0 sítios.
- (2) A cada passo de tempo é adicionado um sítio novo. Este sítio é conectado com outros m ($\leq m_0$) sítios do aglomerado da rede pré-existente.
- (3) A probabilidade de um conexão ser realizada com o sítio i é proporcional a k_i e é dada por

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \quad (2.10)$$

(4) Repete-se as operações (2) e (3) até o tamanho desejado e após t passos de tempo, a rede terá $N = t + m_0$ sítios, para $t \rightarrow \infty$, teremos mt ligações.

É interessante estudarmos a evolução da conectividade dos sítios e a distribuição de conectividade da rede. Estas duas propriedades podem ser obtidas no modelo de Barabási e Albert usando um tratamento contínuo o qual será exposto a seguir.

Tratamento contínuo: Foi introduzido por Barabási, Albert e Jeong [58]. Notemos que no modelo em questão a conectividade k_i de um dado nó i cresce com o tempo (notemos que o tempo passa a medida que novos sítios são adicionados). Seja k_i uma variável contínua e real, esperamos que a taxa com que k_i varie seja proporcional a $\Pi(k_i)$. Assim k_i satisfaz a seguinte equação

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_j k_j}. \quad (2.11)$$

Notemos que a cada passo de tempo, a soma das conectividades aumenta $2mt$, assim $\sum_j k_j = 2mt$. Deste modo a equação (2.11) pode ser escrita como

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t}. \quad (2.12)$$

A solução dessa equação com a condição inicial $k_i(t_i) = m$ (todo sítio novo sempre possui m conexões), é

$$k_i(t) = m \left(\frac{t}{t_i} \right)^\beta, \quad (2.13)$$

onde $\beta = 1/2$. O gráfico (2.8) mostra o resultado da simulação numérica, que está de acordo com a teoria.

A equação (2.13) mostra que a conectividade de todos os nós evolui da mesma forma e segue uma lei de potência.

Usando a equação (2.13), podemos escrever a probabilidade que um nó tenha uma conectividade $k_i(t)$ menor que k , $P[k_i(t) < k]$, como

$$P \left[m \frac{t^\beta}{t_i^\beta} < k \right] = P \left[t_i^\beta > \frac{mt^\beta}{k} \right] = P \left[t_i > \frac{m^{\frac{1}{\beta}} t}{k^{\frac{1}{\beta}}} \right]. \quad (2.14)$$

Como adicionamos os nós em intervalos de tempo iguais na rede, os valores t_i obedecem uma densidade de probabilidade constante dada por

$$P(t_i) = \frac{1}{m_0 + t}. \quad (2.15)$$

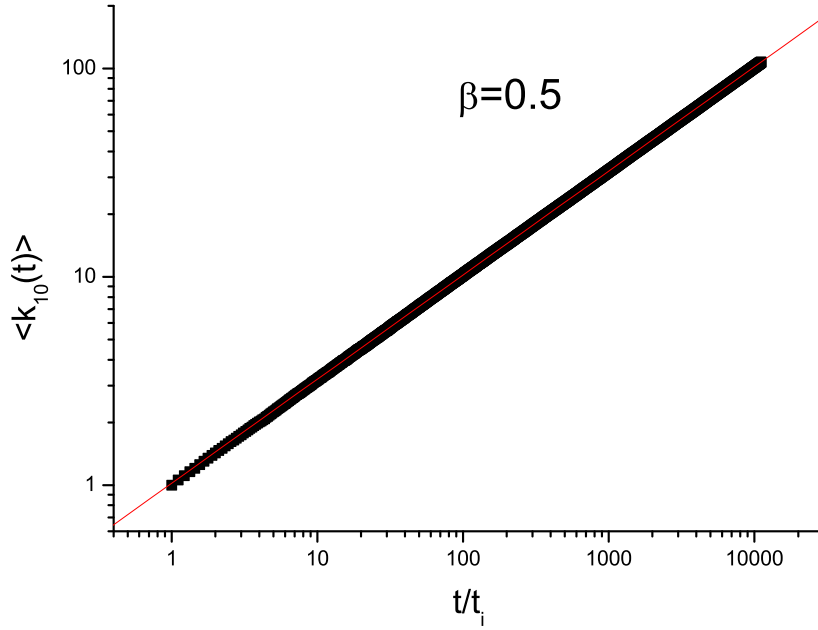


Figura 2.8: Evolução temporal da conectividade. Resultado obtido através de simulação numérica obtendo-se $\beta = 0,5$ ($N = 10^5$).

Substituindo esta equação na equação (2.14), obtemos

$$P \left[t_i > \frac{m^{\frac{1}{\beta}} t}{k^{\frac{1}{\beta}}} \right] = \int_{\frac{m^{\frac{1}{\beta}} t}{k^{\frac{1}{\beta}}}}^{m_0+t} \frac{1}{m_0 + t} dt = 1 - \frac{m^{\frac{1}{\beta}} t}{k^{\frac{1}{\beta}}(t + m_0)}, \quad (2.16)$$

e a distribuição de conectividade $P(k)$ será obtida como segue

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{2m^{\frac{1}{\beta}} t}{m_0 + t} \frac{1}{k^{\frac{1}{\beta}+1}}, \quad (2.17)$$

No limite de $t \rightarrow \infty$, encontramos o coeficiente γ :

$$P(k) \sim 2m^{\frac{1}{\beta}} k^{-\gamma}, \quad (2.18)$$

com $\gamma = 1/\beta + 1 = 3$, concordando com os resultados numéricos (ver a figura 2.9). Notemos que a distribuição de conectividade independe de m .

Em redes reais, observamos leis de potência em sistemas de diferentes tamanhos e que independem do tempo. Assim é esperado que um modelo que descreva esses sistemas, tenha $P(k)$ independente do tempo ou do tamanho da rede ($N = m_0 + t$), mostrando que apesar do crescimento contínuo, a rede atinge um estado estacionário livre de escala como na equação (2.17).

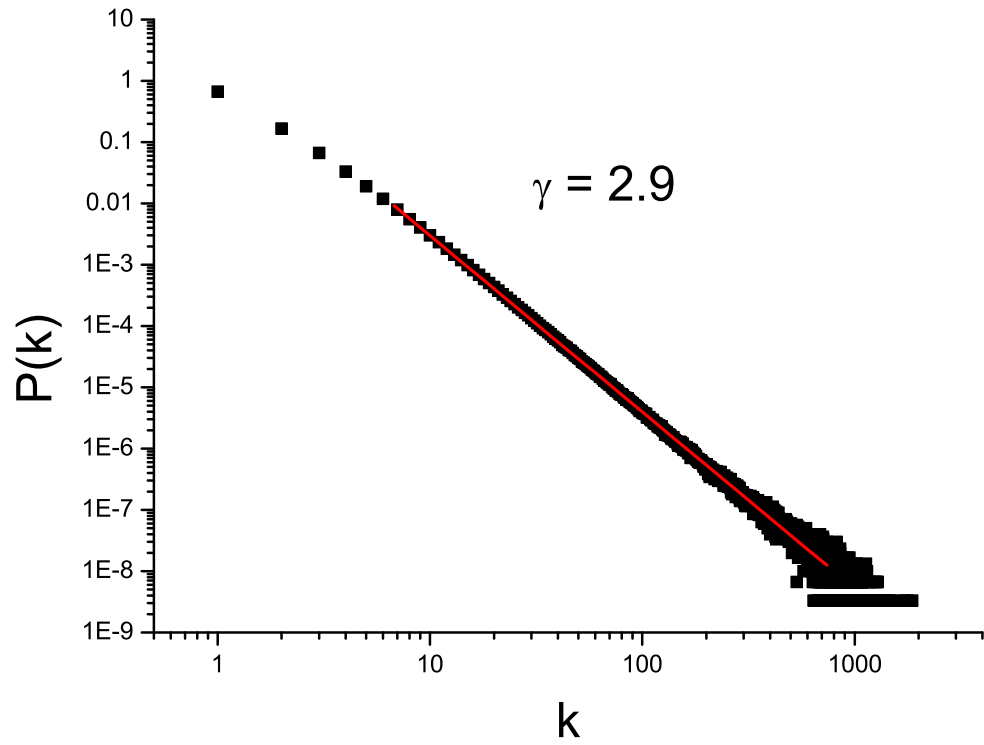


Figura 2.9: Simulação numérica do modelo de Barabási e Albert, $N = 10^5$ e $m = 1$, distribuição de conectividade em escala log-log. O valor de γ concorda com o resultado teórico obtido na equação (2.18).

2.2.2 Ingredientes para gerar uma rede com uma distribuição de conectividade em lei de potência

O modelo de Barabási e Albert possui dois ingredientes importantes para obter uma distribuição de conectividade em lei de potência: **crescimento e ligação preferencial**. Mostraremos de que forma esses ingredientes influenciam a distribuição de conectividade. Lembrando que tal distribuição tem sido verificada em diferentes áreas do conhecimento. Primeiramente, veremos o que ocorre quando mantemos a característica do crescimento da rede sem a ligação preferencial. Iniciamos com um pequeno número de nós (m_0) e em cada passo de tempo adicionamos um novo nó com $m(\leq m_0)$ arestas. Os novos sítios serão conectados com igual probabilidade aos j nós existentes no sistema, ou seja, $\Pi(k_i) = 1/(m_0 + t - 1)$, independe de k_i . Com isso podemos escrever a variação da conectividade da seguinte forma:

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = \frac{m}{(m_0 + t - 1)} \quad (2.19)$$

e obtemos a conectividade de um determinado sítio, dada por

$$k_i(t) = m \ln(m_0 + t - 1) + mC' = \ln(m_0 + t - 1)^m + C \quad (2.20)$$

Da condição inicial, $k_i(t_i) = m$, obtemos $C = m - \ln(m_0 + t - 1)$. Deste modo a evolução temporal de um determinado sítio i pode ser expressa, da seguinte forma:

$$k_i(t_i) = \ln \left[\frac{(m_0 + t - 1)}{m_0 + t_i - 1} \right]^m + m, \quad (2.21)$$

ou seja, $k_i(t)$ segue uma dependência temporal logaritmica, o que não ocorre no modelo de Barabási (modelo livre de escala).

Usando a equação (2.21), podemos escrever a probabilidade que um nó tenha uma conectividade $k_i(t)$ menor que k , $P[k_i(t) < k]$, como

$$P[k_i(t) < k] = P \left[\ln \left[\frac{m_0 + t - 1}{m_0 + t_i - 1} \right]^m + m < k \right], \quad (2.22)$$

ou escrevendo em termos da função exponencial, temos

$$P[k_i(t) < k] = P \left[t_i > \frac{e}{e^{\frac{k}{m}}} (m_0 + t - 1) - m_0 + 1 \right]. \quad (2.23)$$

Notemos que a função densidade de probabilidade é constante, pois adicionamos os nós em intervalos de tempos iguais na rede e dada por

$$P(t_i) = \frac{1}{m_0 + t}. \quad (2.24)$$

Substituindo esta na equação (2.23), ficamos com

$$\begin{aligned} P(k_i(t) < k) &= \int_{\frac{e^{-\frac{k}{m}}}{e^{\frac{k}{m}}(m_0+t-1)-m_0+1}}^{m_0+t} P(t_i) dt_i \\ &= 1 - \frac{e^{-\frac{k}{m}}}{e^{\frac{k}{m}}} \frac{m_0+t-1}{m_0+t} + \frac{m_0-1}{m_0+t}, \end{aligned} \quad (2.25)$$

e a distribuição de conectividade $P(k)$ será obtida como segue

$$P(k) = \frac{\partial P(k_i(t) < k)}{k} = \frac{e^{-\frac{k}{m}}}{m} \frac{m_0+t-1}{m_0+t}. \quad (2.26)$$

No limite de $t \rightarrow \infty$, encontramos a distribuição de conectividade

$$P(k) = \frac{e}{m} \exp\left(-\frac{k}{m}\right). \quad (2.27)$$

Este resultado para a distribuição de conectividade tem um caráter exponencial, mostrando que a ausência da ligação preferencial elimina o caráter de uma rede livre de escala.

Agora veremos o que acontece se a rede não cresce (número de sítios constante durante a evolução da rede) e existe ligação preferencial no sistema. Para isso faremos uma rede com N sítios isolados (nenhuma ligação existe no sistema). A cada passo de tempo um nó é selecionado aleatoriamente e conectado com probabilidade $\Pi(k_i) = k_i / \sum_j k_j$ a um nó i do sistema. Simulações numéricas indicam que em tempos iniciais, o modelo exhibe escala em lei de potência, figura 2.10. Desde que N seja constante e o número de arestas aumente com o tempo, teremos após $T \simeq N^2$ passos de tempo o sistema num estado em que todos os nós estão conectados. Vejamos como é a evolução da conectividade dos sítios. Para isso é preciso usar novamente a teoria contínua como esta a seguir,

$$\frac{\partial k_i(t)}{\partial t} = m \Pi(k_i), \quad (2.28)$$

onde

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} = \frac{k_i}{2t}. \quad (2.29)$$

Notemos que $m = 2$, pois o sítio recém incorporado a rede conecta-se a outros dois nós a cada passo de tempo. Substituindo a equação (2.29) em (2.28) e usando a condição inicial, $k_i(t=1) = 2/N$, obtemos a solução final, ou seja,

$$k_i(t) = \frac{2t}{N}. \quad (2.30)$$

A partir do gráfico 2.10, podemos perceber que a distribuição de conectividade é uma gaussiana em torno do seu valor médio, pois a partir de um determinado tempo os nós devem ter o mesmo valor e é dado pela equação (2.30). Como foi observado, os modelos estudados nessa seção não fornecem uma distribuição livre de escala, indicando que o crescimento e a ligação preferencial são simultaneamente necessários para produzir uma distribuição estacionária em lei de potência, observada em redes reais.

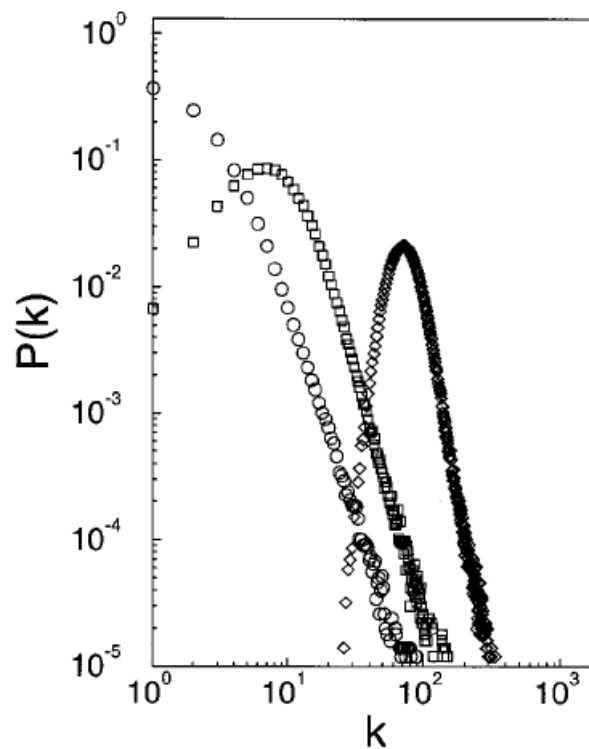


Figura 2.10: Distribuição de conectividade para um modelo em que a quantidade de sítios é fixa ($N=800\ 000$), porém o número de ligações varia. Legenda: círculo ($m_0 = m = 1$), quadrado ($m_0 = m = 3$), diamante ($m_0 = m = 5$) e triângulo ($m_0 = m = 7$). Figura retirada da ref. [58].

2.2.3 Propriedades do modelo de Barabási e Albert

A. Tamanho do menor caminho médio

O gráfico 2.11 mostra a comparação entre o tamanho do menor caminho médio da rede de Barabási-Albert e de uma rede aleatória, para serem comparáveis, tomemos a conectividade média $\langle k \rangle = 4$ e o tamanho da rede, N . O gráfico ilustra que o tamanho do menor caminho médio do modelo de Barabási é menor que a de um grafo aleatório para qualquer N . Ou seja neste modelo a rede é mais coesa e o tamanho do menor caminho da rede de Barabási-Albert cresce logaritmicamente com N ,

$$l = A \ln(N - B) + C. \quad (2.31)$$

O fato de crescer logaritmicamente com N , expressa que este sistema possui efeito de mundo pequeno.

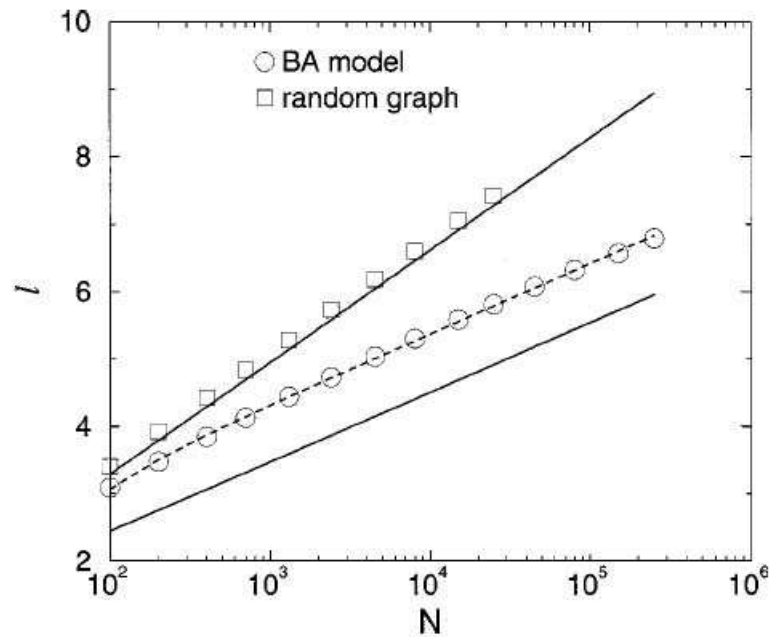


Figura 2.11: Gráfico do menor caminho médio, l , pelo tamanho da rede, N , no modelo de Barabási e Albert com $\langle k \rangle = 4$, comparado com um grafo aleatório de igual tamanho e conectividade média. Figura retirada da ref. [56].

B. Coeficiente de agregação

A figura 2.12 mostra o coeficiente de agregação para a rede de Barabási e Albert e para um grafo aleatório (que possui um coeficiente de agregação $C_{rand} = \langle k \rangle / N$), ambos com conectividade média $\langle k \rangle = 4$ e tamanhos diferentes. O coeficiente de agregação do primeiro modelo é cerca de cinco vezes maior que a de um grafo aleatório e diminui lentamente com o crescimento da rede. Além do mais o coeficiente de agregação da rede livre de escala segue uma lei de potência $C \sim N^{-0.75}$ enquanto o outro é dado por $C = \langle k \rangle N^{-1}$.

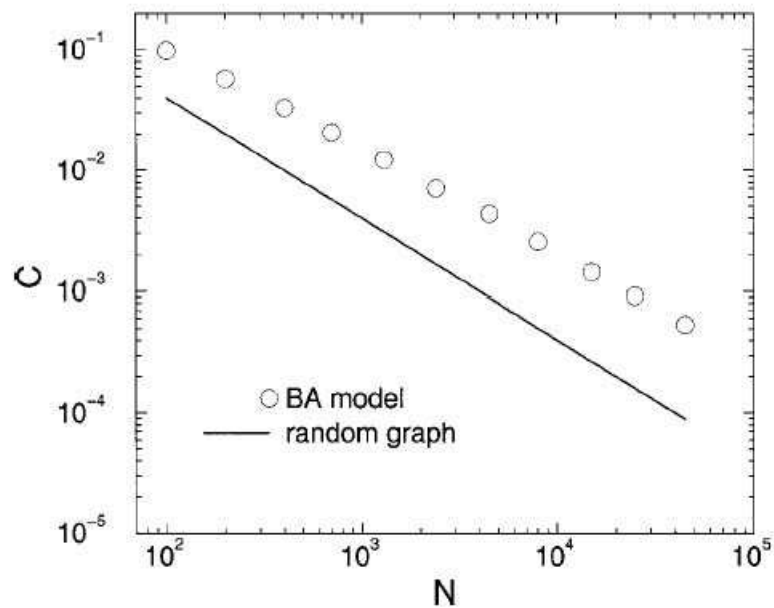


Figura 2.12: Gráfico do coeficiente de agregação versus tamanho da rede. Neste gráfico comparamos o modelo de Barabási e Albert com $\langle k \rangle = 4$ e uma rede aleatório com $C_{rand} \simeq \langle k \rangle / N$. Figura retirada da ref. [56].

Capítulo 3

Modelo de qualidade nas redes complexas

Atualmente os modelos de redes livres de escala estão em evidência na literatura científica, com inúmeras aplicações em diversas áreas do conhecimento ([1],[2],...[31]). Uma das características destes modelos é a descrição das grandezas físicas através de distribuições em lei de potência (por exemplo o modelo de Barabási-Albert (BA) [16]). Antes da proposta de BA era comum se admitir que as distribuições de conectividade obedeciam leis exponenciais. No modelo BA, cada sítio novo que se incorpora à rede se conecta aleatoriamente aos sítios já existentes de acordo com a conectividade destes sítios. Quanto mais conectado maior a chance de receber a ligação do novo sítio. Bianconi juntamente com Barabási modificaram a forma como a rede é construída no modelo BA, introduzindo um fator de qualidade [59]. Nesse capítulo faremos uma generalização deste modelo e veremos suas características topológicas bem como suas aplicações.

3.1 Introdução

Nesta etapa da dissertação já sabemos que as interações entre os constituintes das redes (os nós) determinam a complexidade da rede, ou seja suas, propriedades topológicas. Como por exemplo, a sociedade pode ser imaginada como uma rede complexa cujo os nós são indivíduos e as ligações representam o grau de proximidade (amizade no caso de uma rede social ser entre amigos) entre estes, ou a WWW forma uma rede complexa cujos os

sítios são os documentos e as ligações são as URLs. Uma propriedade genérica destes sistemas complexos é que nestes a quantidade de sítios e ligações variam constantemente através da adição ou remoção de nós e arestas. A fim de abordarmos estes sistemas, temos que estudar as forças dinâmicas que atuam em nível de nós individuais, cujo efeito acumulativo determina o sistema topológico em larga escala. Um dos primeiros modelos teóricos que conseguiu tratar relativamente bem esta propriedade genérica citada a pouco foi sugerida por Barabási e Albert. Neste modelo eles obtiveram uma distribuição de conectividade em lei de potência e resultados que melhor se relacionam com os dados empíricos das redes reais. Assim esperamos que modelos parecidos com este (ligação preferencial e crescimento) seja capaz de descrever melhor os sistemas reais, ou seja, ser mais realístico obtendo propriedades topológicas parecidas com as redes reais. Neste capítulo analisaremos as algumas propriedades (evolução da conectividade dos sítios, distribuição de conectividade, menor caminho médio e coeficiente de agregação) para uma variação do modelo BA.

Com o intuito de motivar o estudo da generalização do modelo de qualidade, daremos alguns exemplos que mostram propriedades que não esta presente no modelo BA mas que existe em sistemas reais como a taxa de crescimento da conectividade dos sítios que depende não somente da idade dos sítios isoladamente (conectividade). Por exemplo, num sistema social não é todo mundo que faz amigos a mesma taxa: alguns indivíduos são melhores que outros em tornar um encontro aleatório numa amizade (ligação). Na rede WWW, alguns documentos através de uma combinação de bons conteúdos e "marketing" adquirem um maior número de ligações num pequeno intervalo de tempo, podendo alcançar "websites" que estejam no ar por mais tempo. Também em Hollywood alguns atores num pequeno intervalo de tempo participam de filmes e colecionam ligações com seus colegas de profissão que pode ser maior que alguns atores que estão no negócio há muito mais tempo. Finalmente, alguns trabalhos científicos conseguem num pequeno intervalo de tempo adquirir um grande número de citações e até ultrapassar seus coteporâneos publicados. Em todos estes exemplos, vemos o seguinte padrão similar: alguns nós adquirem ligações numa taxa muito maior que outros sítios quando comparados ao modelo de Barabási; e existe uma maior quantidade de pólos na rede. Pretendemos associar estas dife-

renças a algumas qualidades intrínsecas dos nós, tal como a personalidade de um indivíduo, o conteúdo de uma "webpage", o talento de um ator ou o conteúdo de um artigo científico. No estudo de redes, chamaremos este fator de *qualidade do nó*. A partir destes exemplos vemos que o modelo livre de escala (BA) não incorpora muito bem o aspecto competitivo dos sistemas vistos na natureza apesar do sucesso na predição da topologia das redes reais. Este modelo prediz que todos os nós aumentam suas conectividades no tempo com $k_i(t) = (t/t_i)^\beta$, onde $\beta = 1/2$ e t_i é o tempo que o nó i tem sido adicionado no sistema. Conseqüentemente, os nós mais velhos terão maior número de ligações, se a estrutura temporal para adquirir ligações for longa. Desde que novos nós se anexem preferencialmente em nós mais conectados, então os sítios mais antigos continuarão facilmente adquirindo mais ligações numa taxa maior do que os sítios novos. Este fenômeno é responsável pela cauda em lei de potência da distribuição de conectividade. Assim, se dois nós surgem ao mesmo tempo, eles terão aproximadamente em qualquer tempo o mesmo número de ligações, a parte das flutuações estatísticas.

Na direção de propor um modelo simples, que nos permita investigar este aspecto competitivo da rede real em termos quantitativos, Bianconi e Barabási adicionaram na ligação preferencial um fator de qualidade (depende da qualidade do nó). Para facilitar o entendimento do leitor, apresentaremos o nosso modelo de qualidade generalizado com o parâmetro $\alpha = 0$ que é justamente o modelo de Bianconi-Barabási (será explicado mais adiante o significado de α). Por hora veremos que a conectividade dos nós individuais para o modelo de qualidade segue uma lei de potência no tempo, $k_i(t) \sim t^{\beta_i}$, porém o expoente dinâmico β_i dependerá da qualidade do nó. Desenvolveremos um modelo contínuo para esta competitividade, de modo a calcular analiticamente β e derivar uma expressão para a distribuição de conectividade.

O modelo de qualidade ou modelo Bianconi - Barabási - Os exemplos discutidos nos capítulos anteriores indicam que os nós possuem diferentes habilidades (qualidades) em competir por ligações. Para explicar estas diferenças, introduziremos o algoritmo do modelo de Bianconi-Barabási a seguir:

- (1) Inicia-se a rede com m_0 sítios, onde cada um deles possui um parâmetro de qualidade η . A qualidade é escolhida aleatoriamente respeitando uma distribuição de qualidade $\rho(\eta)$.
- (2) A cada passo de tempo é adicionado um novo sítio que já possui um parâmetro de qualidade η escolhido a partir da distribuição de qualidade (passo 1). Então este sítio é conectado com outros m ($\leq m_0$) sítios do aglomerado da rede pré-existente.
- (3) A probabilidade de uma conexão ser realizada com um sítio i é proporcional a k_i e η_i , e é dada por

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}. \quad (3.1)$$

- (4) Repete-se as operações (2) e (3) até o tamanho desejado. No limite termodinâmico, tempo suficientemente grande, a rede terá $N = t + m_0$ sítios e mt ligações.

Esta generalização da ligação preferencial incorpora a combinação mais simples possível que a qualidade e a conectividade presentes determinem a taxa com que novas ligações são adicionadas a um dado nó, isto é, possibilita o sítio relativamente mais jovem que possui algumas ligações obter uma alta taxa de conexões, para isso basta ter um grande parâmetro de qualidade. A fim de discutir as propriedades de escala desse modelo, primeiramente utilizamos uma teoria contínua de modo a obter a distribuição de conectividade. Lembrando que esta teoria possui aproximações que foram tratadas em mais detalhes na seção sobre o modelo BA. Um nó i irá aumentar sua conectividade k_i numa taxa que é proporcional a probabilidade dada pela equação (3.1), da seguinte forma

$$\frac{\partial k_i}{\partial t} = m \frac{\eta_i k_i}{\sum_j k_j \eta_j}. \quad (3.2)$$

O fator m ilustra o fato que cada sítio novo incrementa m ligações ao sistema. Um caso particular é $\rho(\eta) = \delta(\eta - 1)$, ou seja toda qualidade é igual a um. Substituindo a distribuição de qualidade na equação (3.2), obtem-se o modelo livre de escala, que prediz que $k_i(t) \sim t^{1/2}$. Para resolver a equação (3.2), Bianconi e Barabási sugeriram a seguinte solução [59], inspirados no modelo livre de escala,

$$k_{\eta_i}(t, t_0) = m \left(\frac{t}{t_0} \right)^{\beta(\eta_i)}, \quad (3.3)$$

onde t_0 é o tempo em que o nó foi introduzido no sistema. Notemos que eles introduziram no ansatz uma multi escala no expoente dinâmico. Outra coisa que é verificada nesse expoente, $\beta(\eta)$, é que ele é limitado no intervalo $0 < \beta(\eta) < 1$ pois um nó sempre aumenta o número de ligações no tempo, ($\beta(\eta) > 0$) e $k_i(t)$ não pode aumentar mais rápido que t , ($\beta(\eta) < 1$). Agora calcularemos a média da soma $\sum_j \eta_j k_j$. Desde que cada nó seja produzido num tempo diferente de t_0 , a soma sobre j pode ser escrita como uma integral sobre t_0

$$\begin{aligned} \langle \sum_j \eta_j k_j \rangle &= \int d\eta \rho(\eta) \eta \int_1^t dt_0 k_\eta(t, t_0) \\ &= \int d\eta \rho(\eta) \eta \int_1^t \frac{dt_0}{t_0^{\beta(\eta)}} \\ &= \int d\eta \eta \rho(\eta) m \frac{(t - t^{\beta(\eta)})}{1 - \eta(\eta)}. \end{aligned} \quad (3.4)$$

Desde que $\beta(\eta) < 1$, no limite em que $t \rightarrow \infty$, $t^{\beta(\eta)}$ pode ser desprezado comparado com t , obtendo

$$\langle \sum_j \eta_j k_j \rangle = C m t (1 + O(t^{-\epsilon})), \quad (3.5)$$

onde

$$\begin{aligned} \epsilon &= (1 - \max \beta(\eta)) > 0, \\ C &= \int d\eta \rho(\eta) \frac{\eta}{1 - \beta(\eta)}. \end{aligned} \quad (3.6)$$

Usando a equação (3.5), e a notação $k_\eta = k_{\eta_i}(t, t_0)$, a equação dinâmica (3.2) pode ser escrita como

$$\frac{\partial k_\eta}{\partial \eta} = \frac{\eta k_\eta}{Ct}, \quad (3.7)$$

que tem uma solução da seguinte forma,

$$k_\eta(t, t_0) = At^{\frac{\eta}{C}}, \quad (3.8)$$

e impondo a condição inicial, $k_\eta(t_0, t_0) = m$, encontramos o expoente dinâmico dado por

$$\beta(\eta) = \frac{\eta}{C}, \quad (3.9)$$

Notemos que $\beta(\eta)$ depende de C e logicamente dependerá de $\rho(\eta)$. Deste modo temos que obter o valor de C . Para isso substituiremos $\beta(\eta)$ na equação (3.6), ficando com

$$1 = \int_0^{\eta_{max}} d\eta \rho(\eta) \frac{1}{\frac{C}{\eta} - 1}, \quad (3.10)$$

onde η_{max} é a máxima qualidade possível no sistema. Aparentemente a equação (3.10) é uma integral com uma singularidade. Entretanto, desde que $\beta(\eta) = \eta/C < 1$ para qualquer valor de η , temos $C > \eta_{max}$, assim o limite de integração nunca terá uma singularidade. Verificamos também que, desde que $\sum_j \eta_j k_j \leq \eta_{max} \sum_j k_j = 2mt\eta_{max}$, e usando a equação (3.5) que $C \leq 2\eta_{max}$.

Finalmente, podemos calcular a distribuição da conectividade $P(k)$, que fornece a probabilidade que o nó tenha k ligações. Mas antes de fazer este cálculo, discutiremos as implicações obtidas em ter o expoente dinâmico como na equação 3.9. Este expoente obtido, ilustra como é o comportamento da evolução da conectividade dos constituintes da rede (sítios) e como o expoente β depende do parâmetro η (qualidade intrínseca do nó). Outro ponto que devemos prestar atenção é a distribuição de conectividade que é dada pela seguinte expressão, $P(k) \sim k^{-\gamma}$ onde $\gamma = 1/\beta + 1$. Assim $P(k)$ será obtido por uma média sobre diferentes leis de potência. Para achar $P(k)$, precisamos calcular a probabilidade acumulada para um nó fixo $k_\eta > k$, como é feito a seguir

$$P(k_\eta(t) > k) = P \left[m \left(\frac{t}{t_0} \right)^{\beta(\eta)} > k \right] = P \left[t_0 < t \left(\frac{m}{k} \right)^{\frac{\eta}{C}} \right]. \quad (3.11)$$

Como adicionamos os nós em intervalos de tempo iguais na rede, os valores t_i tem uma densidade de probabilidade constante dada por

$$P(t_i) = \frac{1}{m_0 + t} \quad (3.12)$$

com isso encontramos o valor de $P(k_\eta > k)$, dado por

$$P(k_\eta(t) > k) = \int_0^{t\left(\frac{m}{k}\right)^{\frac{C}{\eta}}} \frac{1}{m_0 + t} dt_0 = \frac{1}{m_0 + t} t \left(\frac{m}{k}\right)^{\frac{C}{\eta}} \quad (3.13)$$

Assim a distribuição de conectividade, ou seja a probabilidade que o nó tenha k ligações, é dada pela integral

$$P(k) = \int_0^{\eta_{max}} d\eta \frac{\partial P(k_\eta(t) > k)}{\partial k} \propto \int_0^{\eta_{max}} d\eta \rho(\eta) \frac{C}{\eta} \left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}. \quad (3.14)$$

3.2 Resultados

Nessa etapa, abordaremos algumas distribuições de qualidade no modelo de Bianconi e Barabási, e veremos o que ocorre com suas características topológicas: distribuição de conectividade, menor caminho médio, coeficiente de agregação e evolução da conectividade dos nós. A análise da distribuição de conectividade foi realizada com $m = 1$, pois usamos o fato da rede ser livre de escala, isto nos poupa tempo de processamento (quanto maior m mais trabalhoso se torna o algoritmo). Por outro lado, o problema torna-se mais interessante se usarmos valores de m maiores que 1 para o menor caminho médio e o coeficiente de agregação, pois, a maioria das redes reais possui valores maiores que a unidade.

3.2.1 Modelo livre de escala

Dada uma distribuição de qualidade $\rho(\eta)$, a teoria contínua permite a predição dinâmica da rede, descrita pelo expoente dinâmico $\beta(\eta)$ (equações (3.9) e (3.10)), e a topologia, caracterizada pela distribuição da conectividade $P(k)$ (equação (3.14)). O modelo livre de escala (BA) é um dos mais simples (ligação preferencial depende apenas da conectividade dos nós) e apresenta todos os sítios com qualidades iguais (notemos que se $\eta_i = \text{constante}$, da equação (3.7) obtemos a equação dinâmica do modelo de Barabási-Albert). Assim podemos representar a distribuição de qualidade deste modelo da seguinte forma, $\rho(\eta) = \delta(\eta - 1)$, e inserindo na Eq.(3.10), encontraremos $C = 2$, que representa o maior valor possível de C . Usando a Eq.(3.9), obtemos $\beta = 1/2$ e da Eq.(3.14), chegamos a $P(k) \propto k^{-3}$, a conhecida relação do modelo livre de escala que foi explicada e ilustrada (gráficos da distribuição de conectividade (gráfico 2.9) e evolução temporal da conectividade dos sítios (gráfico 2.8)) no capítulo anterior. Assim, o modelo livre de escala representa um caso limite do modelo de qualidade considerado neste capítulo, com o expoente da conectividade possuindo o maior valor possível.

3.2.2 Distribuição de qualidade uniforme

A distribuição de qualidade uniforme é o caso mais simples que possui diferentes valores de qualidades competindo por ligações que oferecem escalas múltiplas não triviais (diferentes valores de β). Este é obtido quando $\rho(\eta)$ é escolhido uniformemente num intervalo $[0,1]$, ou seja, $\rho(\eta) = \text{constante}$. A constante C pode ser determinada novamente pela equação (3.10). Fazendo uma mudança de variável, $y = C - \eta$ e assim $dy = -d\eta$, teremos

$$1 = \int_{C-1}^C dy \frac{(C-y)}{y}. \quad (3.15)$$

Essa integral é facilmente resolvida e obtemos dela a expressão para encontrar C , dada por

$$\exp(-2/C) = 1 - 1/C, \quad (3.16)$$

cuja solução é $C = 1,255$. Assim, de acordo com a Eq.(3.9), cada nó terá um diferente expoente dinâmico, dado por $\beta(\eta) \sim \frac{\eta}{C^*}$. Usando a equação (3.14), obtemos

$$P(k) \propto \int_0^1 d\eta \frac{C^*}{\eta} \frac{1}{k^{1+C^*/\eta}} \sim \frac{k^{-(1+C^*)}}{\log(k)}, \quad (3.17)$$

ou seja, a distribuição de conectividade segue uma lei de potência generalizada, com um logaritmo inverso.

A fim de verificar as predições da teoria contínua, bem como testar a validade da solução teste (equação 3.3), realizamos algumas simulações numéricas do modelo de qualidade usando uma distribuição constante para o fator de qualidade, $\rho(\eta) = \text{constante}$. A solução teste pode ser comprovada através dos resultados expressos nos gráficos 3.1, 3.2, 3.3 e 3.4. Neste gráfico usamos $m = 1$, pois a evolução da conectividade do sítio independe do valor escolhido para m , além do mais a escolha de um m cada vez maior, aumenta o tempo de execução da simulação.

Notemos no gráfico 3.1, que $k_i(t)$ segue uma lei de potência para todos η . Já no gráfico 3.4, vemos que o expoente de escala, $\beta(\eta)$, cresce com η . Outro ponto interessante neste gráfico é que o comportamento do sistema é dado por $\beta(\eta) = \eta/C$, justamente o que foi obtido pelo cálculo analítico na equação 3.9. Para fazermos este gráfico, usamos o seguinte algoritmo:

- (1) Inicia-se a rede com m_0 sítios, onde cada um deles possui um parâmetro de qualidade η . A qualidade é escolhida aleatoriamente respeitando uma distribuição de qualidade $\rho(\eta)$. Para o sítio que desejamos observar a evolução de conectividade, escolhamos um parâmetro de qualidade no intervalo entre 0 e 1.
- (2) A cada passo de tempo é adicionado um novo sítio que já possui um parâmetro de qualidade η escolhido a partir da distribuição de qualidade (passo 1). Então este sítio é conectado com outros m ($\leq m_0$) sítios do aglomerado da rede pré-existente.
- (3) A probabilidade de uma conexão ser realizada com um sítio i é proporcional a k_i e η_i , e é dada por

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}. \quad (3.18)$$

- (4) Repete-se as operações (2) e (3) até o tamanho desejado. No limite termodinâmico, tempo suficientemente grande, a rede terá $N = t + m_0$ sítios e mt ligações.

Após este algoritmo ser realizado, é necessário realizar o gráfico da evolução da conectividade do sítio (que havíamos fixado seu parâmetro de qualidade, no passo 1) pela razão de t/t_0 , onde t_0 é o tempo em que o sítio nasceu.

Vimos que a equação 3.5 prediz que a soma $\langle \sum_i \eta_i k_i \rangle / mt \rightarrow C$ no limite de $t \rightarrow \infty$, onde C é dada pela equação (3.16) com $C = 1.255$, como indicado no gráfico 3.2. Por fim, no gráfico 3.3, mostramos a concordância entre a predição da equação 3.17 e os resultados obtidos numericamente para a distribuição de conectividade $P(k)$.

O gráfico 3.3 possui uma característica muito interessante quando olhamos o comportamento dos sítios à nível de ligações realizadas durante a evolução da rede. Observamos neste, a presença de alguns nós que possuem uma alta conectividade ("hubs ou pólos") que aparecem como uma longa linha horizontal num gráfico log-log, presente em vários sistemas, incluindo a rede WWW [1, 2, 3, 4, 5, 6, 7] e a de metabolismo das células [21, 22]. Isto indica que "super pólos" é uma característica de sistemas competitivos.

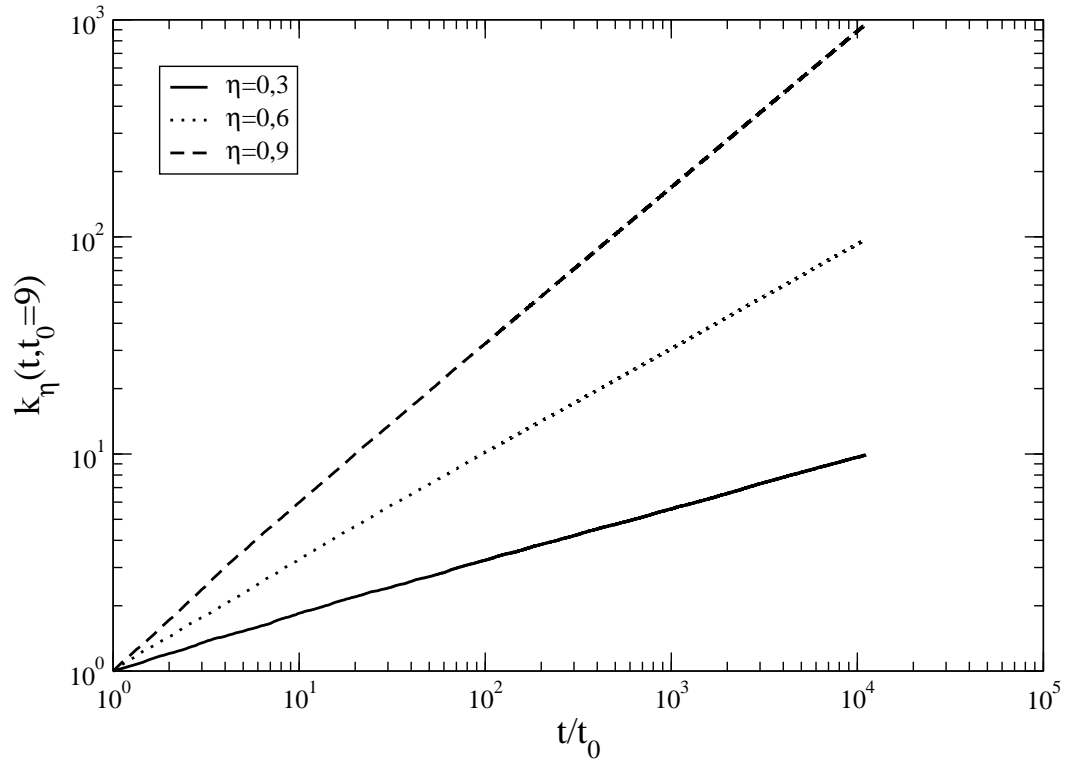


Figura 3.1: Simulação numérica para uma rede com $m = 1$ e $N = 10^5$ mostrando a dependência temporal da conectividade, $k_\eta(t)$, para sítios com qualidade $\eta = 0,3, 0,6$ e $0,9$. Notemos que com o aumento do fator de qualidade, a conectividade do sítio cresce mais rapidamente durante a evolução da rede.

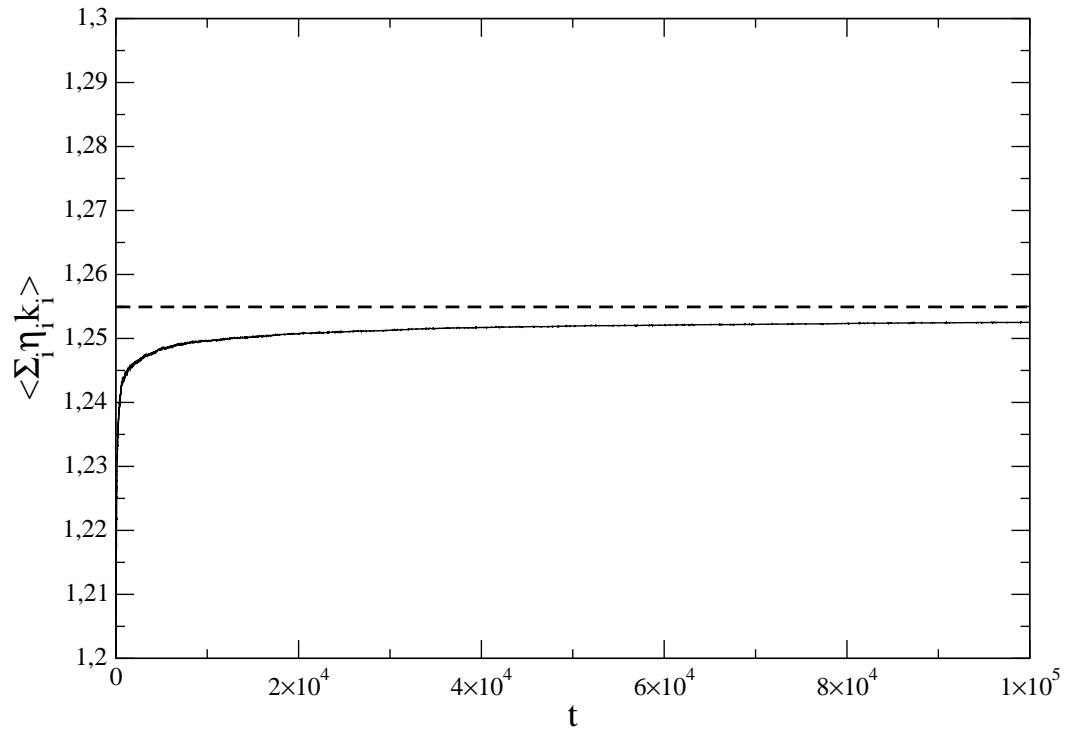


Figura 3.2: Comportamento assintótico de C com $t \rightarrow \infty$, notemos a concordância com a predição analítica, linha tracejada, (equação 3.5).

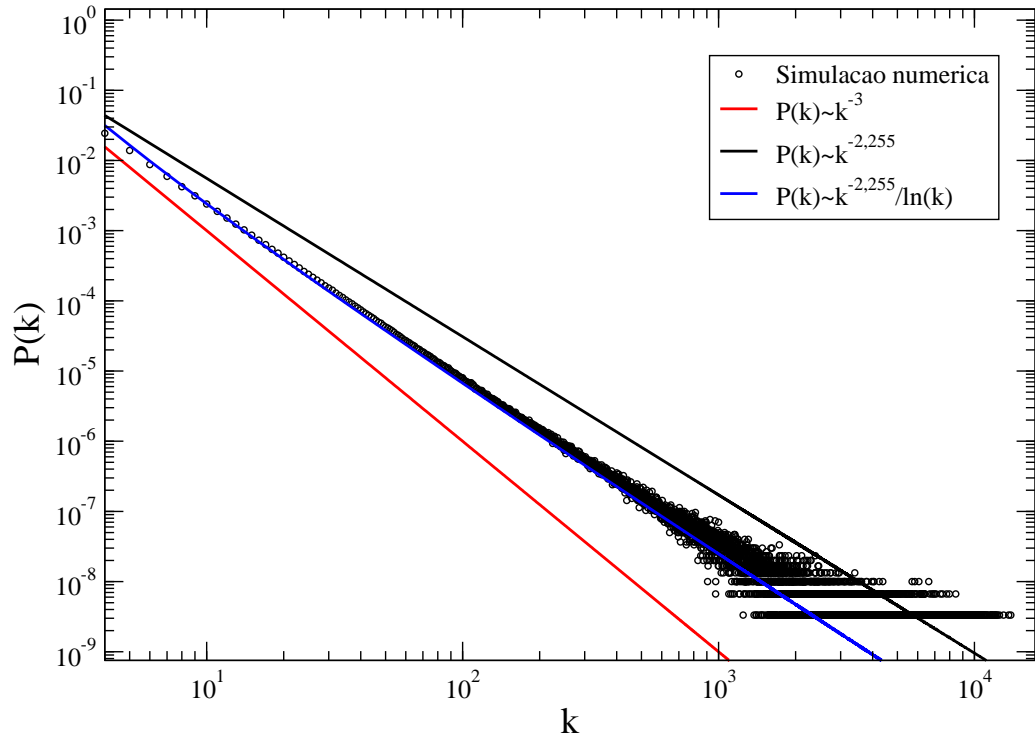


Figura 3.3: Distribuição de conectividade no modelo de qualidade, para uma rede com $m = 1$ e $N = 10^5$. Os círculos sólidos representam a simulação numérica que está em concordância com a predição teórica dada pela equação 3.17 (linha azul), com $\gamma = 2,255$. A linha preta representa um ajuste com $P(k) \sim k^{-2,255}$ sem a correção logarítmica. A linha vermelha corresponde a $P(k) \sim k^{-3}$ como prediz o modelo livre de escala, ou seja, todas as qualidades são iguais.

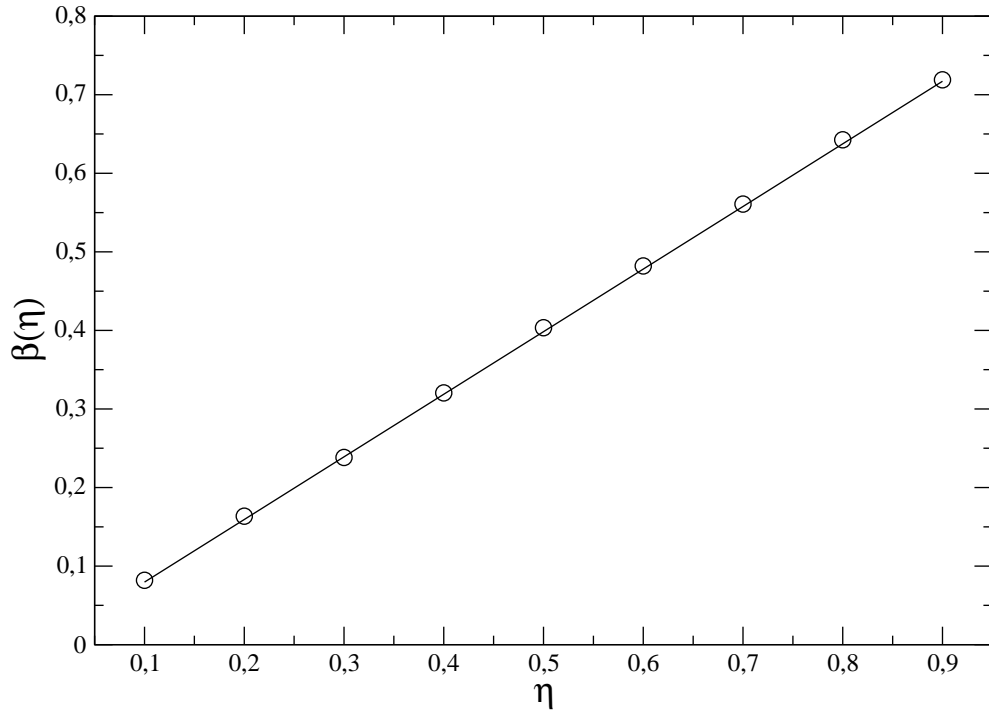


Figura 3.4: Resultado da simulação numérica do modelo de Bianconi com qualidade uniforme (rede com $N = 10^5$ e $m = 1$): Dependência do expoente dinâmico $\beta(\eta)$ com o parâmetro da qualidade η para uma distribuição uniforme $\rho(\eta) = \text{constante}$. Os círculos foram obtidos através da simulação numérica enquanto a linha foi feita com base na predição analítica, $\beta(\eta) = \eta/1,255$. A simulação numérica foi feita, fixando a qualidade do nó nascido em $t_0 = 9$.

3.2.3 Distribuição de qualidade em lei de potência

Nesta seção generalizaremos o modelo de Bianconi e Barabási de forma a incorporar uma interessante característica das distribuições em lei de potência que ilustra bem a ocorrência dos eventos raros e dos não raros. Para deixar claro, observemos o gráfico da distribuição de conectividade: existem poucos sítios com alta conectividade (eventos raros) enquanto os com baixa conectividade ocorrem com frequência (eventos não raros). Se imaginarmos que a distribuição de qualidade pode ser escrita em lei de potência, seria dizer que existem poucos nós com uma alta qualidade e muitos nós com uma baixa qualidade. Isto parece muito sugestivo para vários sistemas físicos, como exemplo pensemos na rede WWW onde a maioria dos nós (páginas) devem apresentar um conteúdo ruim (baixa qualidade), enquanto poucos possuem um conteúdo bom (alta qualidade). Assim usaremos uma distribuição de qualidade em lei de potência no modelo de qualidade afim de observar as implicações obtidas. Neste novo modelo (generalizado) mostraremos que ele apresenta como casos limites:

(1) $\rho(\eta) = \delta(\eta - 1)$, modelo de Barabási-Albert.

(2) $\rho(\eta) = \text{constante}$, modelo de Bianconi e Barabási (seção anterior).

Iniciaremos nosso estudo, calculando o valor da constante C que além de mostrar como é a competição dos sítios por ligações, fornece também o comportamento da distribuição de conectividade do sistema. Para obter essa constante, usaremos a equação 3.10, escrita logo abaixo:

$$1 = \int_0^{\eta_{max}} d\eta \rho(\eta) \frac{1}{\frac{C}{\eta} - 1}.$$

Onde a nossa distribuição de qualidade é expressa em forma de **lei de potência**, da seguinte forma:

$$\rho(\eta) = A\eta^\alpha \tag{3.19}$$

e sua normalização é

$$\int_0^1 \rho(\eta) d\eta = 1. \tag{3.20}$$

Assim, obtemos o valor de $A = \alpha + 1$. Notemos que se α for um número negativo, então a distribuição de qualidade diverge assim estudaremos somente o efeito da variação dos valores de α positivos. Agora com o valor de ρ já obtido, substituiremos ele na equação 3.10 e faremos a seguinte mudança de variável, $y = C - \eta$, obtendo a seguinte equação integral para C:

$$\frac{1}{A} = \int_{C-\eta_{max}}^C dy \frac{(C-y)^{\alpha+1}}{y}. \quad (3.21)$$

Assim, de acordo com a equação 3.9, cada sítio terá um expoente dinâmico diferente, dado por $\beta \sim \frac{\eta}{C}$, onde o valor de C dependerá do α escolhido para o sistema. O caso em que $\alpha = 0$, temos a distribuição de qualidade constante. Os gráficos 3.5 e 3.6 mostram respectivamente: o comportamento do expoente dinâmico (de um sítio com qualidade igual a 1) versus o parâmetro α do sistema e o segundo ilustra como o valor de C varia com α .

É interessante notar no gráfico interior 3.5, o fato da qualidade média tender ao valor um, juntamente com o comportamento de β tender ao valor 0,5 faz com que nesse limite a rede se comporte como o modelo proposto por Barabási. Também podemos verificar que um valor α grande, diminui a taxa com que o sítio obtem ligações bem como a competitividade do sistema por ligações, devido ao expoente β estar intimamente relacionado a evolução da conectividade dos sítios, que é justamente a interpretação da equação 3.3.

Outro ponto interessante a ser estudado é o valor de C , visto que a partir da relação $\beta \sim \eta/C(\alpha)$, podemos entender como ocorre a competição numa rede que possui a mesma estrutura sugerida inicialmente pelo modelo de qualidade. Para isto analisamos o gráfico 3.6 que fornece diferentes valores de C^* para α 's diferentes. Com este gráfico é possível obter a evolução da conectividade dos nós. É interessante notar que quando $\alpha \rightarrow \infty$, o valor de C^* tende para 2, que é o mesmo valor obtido para o modelo livre de escala. Esses dados poderiam ser obtidos através de simulação pelo cálculo da $\langle \sum_i \eta_i k_i \rangle / mt \rightarrow C^*$ no limite de $t \rightarrow \infty$, obtendo os mesmos resultados como no gráfico 3.2.

Os gráficos vistos nesta seção dão grandes indícios que o modelo de Bianconi-Barabási com qualidade em lei de potência possui dois casos limites: o modelo com qualidade constante (isso é fácil de perceber a partir de $\rho(\eta) = A\eta^\alpha$) e o modelo de Barabási quando $\alpha \rightarrow \infty$.

Um ponto interessante a se observar (ver figura 3.7) é a dependência

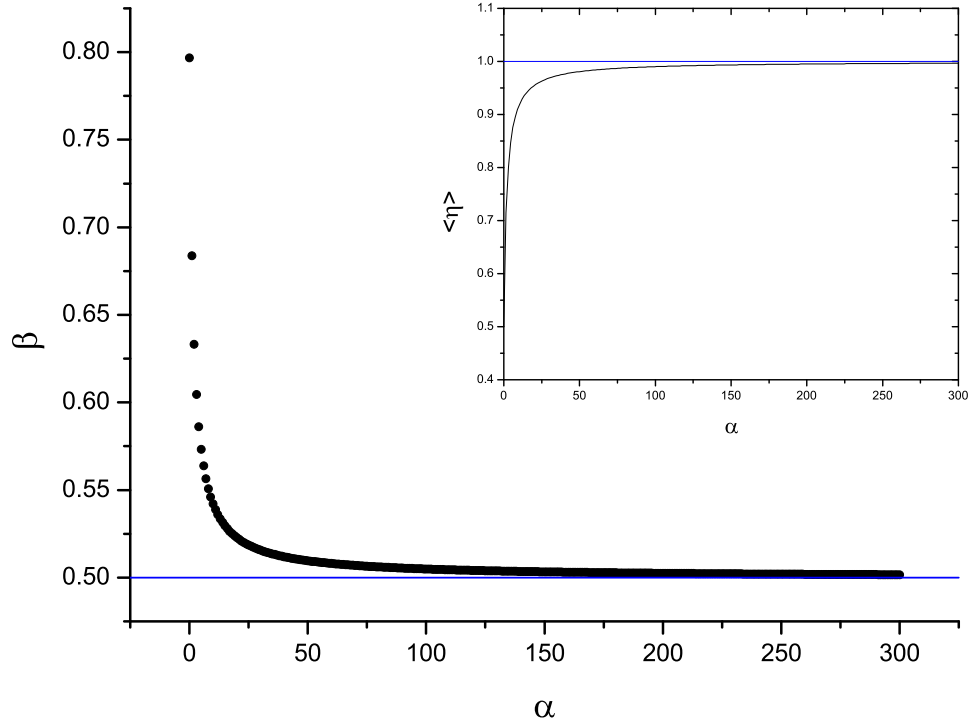


Figura 3.5: Modelo de Bianconi-Barabási: O gráfico mostra o comportamento de um sítio com qualidade igual a 1 ao variarmos o valor de α . O gráfico interno esboça o comportamento assintótico do valor médio de uma rede com $\rho = A\eta^\alpha$. Resultados obtidos a partir da equação 3.21.

temporal da conectividade, $k_\eta(t, t_0)$, que mostra o seguinte: sítios com altas qualidades possuem uma maior competitividade por ligação e incrementando o valor de α , estaremos minimizando a importância de um sítio possuir uma alta qualidade (lembrando que a medida que o valor de α cresce, as qualidades dos sítios tendem a serem semelhantes até o limite em que $\alpha \rightarrow \infty$ e todos os nós possuem qualidade igual a 1). Assim, vemos que estes dois fatores são de extrema importância para as propriedades topológicas da rede estudada por nós. Este comportamento está muito bem ilustrado pelos gráficos 3.7 e 3.8 (este último ilustra a dependência de β com a qualidade e α). Além do mais, podemos a partir do gráfico 3.8 obter os valores de C para diferentes sistemas (redes com diferentes valores de α), lembrando que $\beta = \eta/C$, assim basta usar os valores de η e β obtidos no gráfico para calcular C .

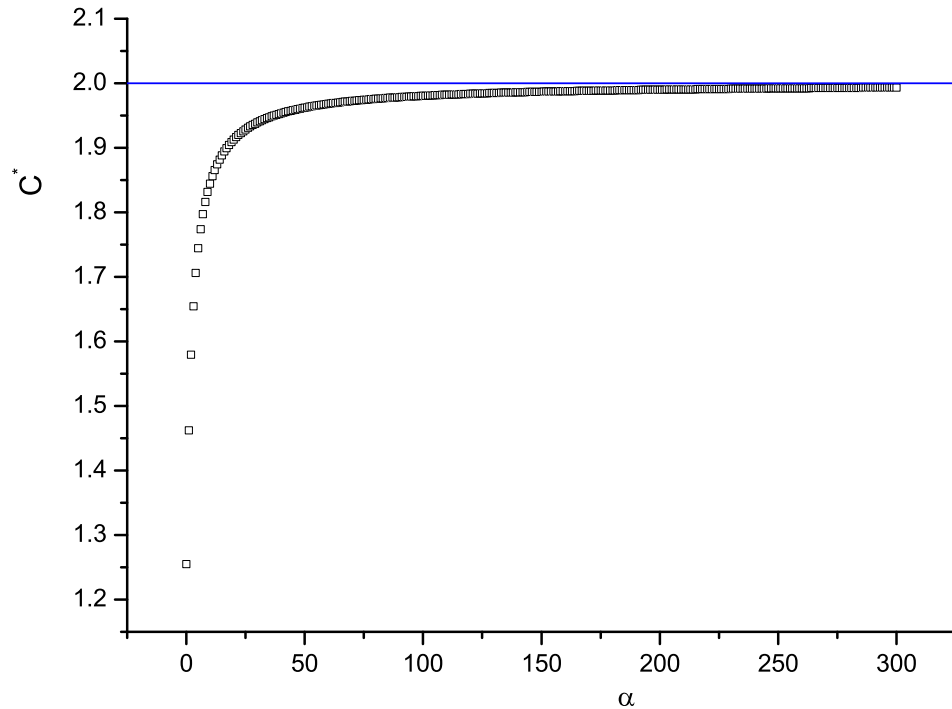


Figura 3.6: Modelo de Bianconi-Barabási: Comportamento assintótico de C^* obtido através da equação 3.21. Para isso basta fixar o valor de α e resolver a equação.

Ainda podemos obter mais informações a respeito do comportamento topológico das redes em estudo, apenas com os resultados analíticos. Como é o caso da distribuição de conectividade, onde vimos da equação 3.14 que se diferentes valores de C são usados nesta equação, obtemos diferentes distribuições, exemplificadas no gráfico 3.9 (simulação do modelo de qualidade generalizado). Neste gráfico, notemos que o nosso modelo interpola desde o modelo de Bianconi-Barabási ($\gamma = 2, 255$) até o modelo de Barabási-Albert ($\gamma = 3$). Podemos perceber também que a variação de $P(k)$ é pequena ao variarmos α porém este intervalo de valores de γ , $[2, 25; 3]$, é obtido em várias redes reais estudadas que são livres de escala (ver figura 1.9). O gráfico oferece um interessante sinal da evolução dos nós num ambiente competitivo, por exemplo: o modelo de Bianconi com $\alpha = 0$ possui nós com "hubs" muito mais conectados do que os do modelo de Barabási ou com $\alpha > 0$. Este fato

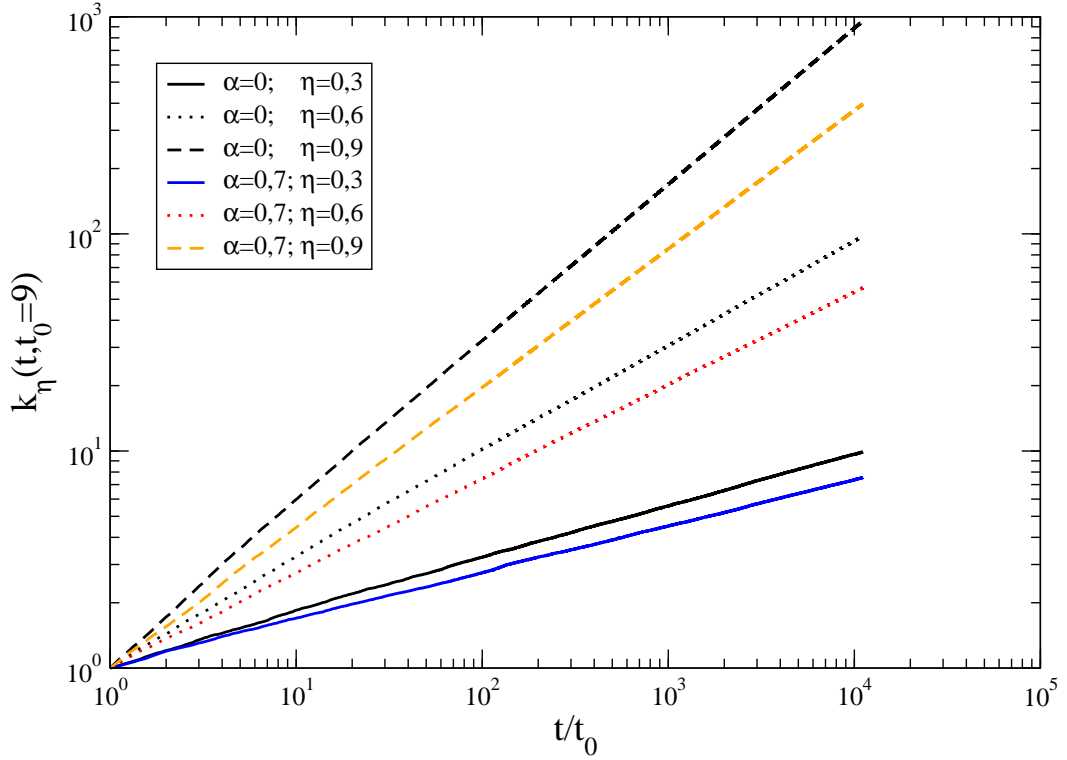


Figura 3.7: Resultado da simulação do modelo de Bianconi com qualidade em lei de potência (rede com $N = 10^5$ e $m = 1$): Observemos que a evolução da conectividade depende de um sítio α e η .

é explicado da seguinte forma: se um sítio nasce no tempo em que a rede foi criada e possui uma alta qualidade, este terá provavelmente uma alta conectividade. Agora, se o sítio tiver só uma alta qualidade ou se for velho na rede, será menos provável que este tenha uma alta conectividade. Por isso, vemos no gráfico 3.9 que o modelo de Barabási ou de Bianconi com α grande possui sítios com menor número de conexões comparado ao modelo de qualidade com α pequeno.

Como vimos nos capítulos introdutórios, o menor caminho médio e o coeficiente de agregação estão relacionados com a conectividade dos sítios ou seja, com a distribuição de conectividade. Desta forma, esperamos que estas duas quantidades não sejam iguais para diferentes valores de α , devido às diferentes curvas observadas de $P(k)$ no gráfico 3.9. O estudo dessas duas quantidades foi realizada com simulações numéricas devido a dificuldade em realizar cálculos analíticos em sistemas como os que estudamos.

O gráfico 3.10 ilustra o comportamento de mundo pequeno, pois o me-

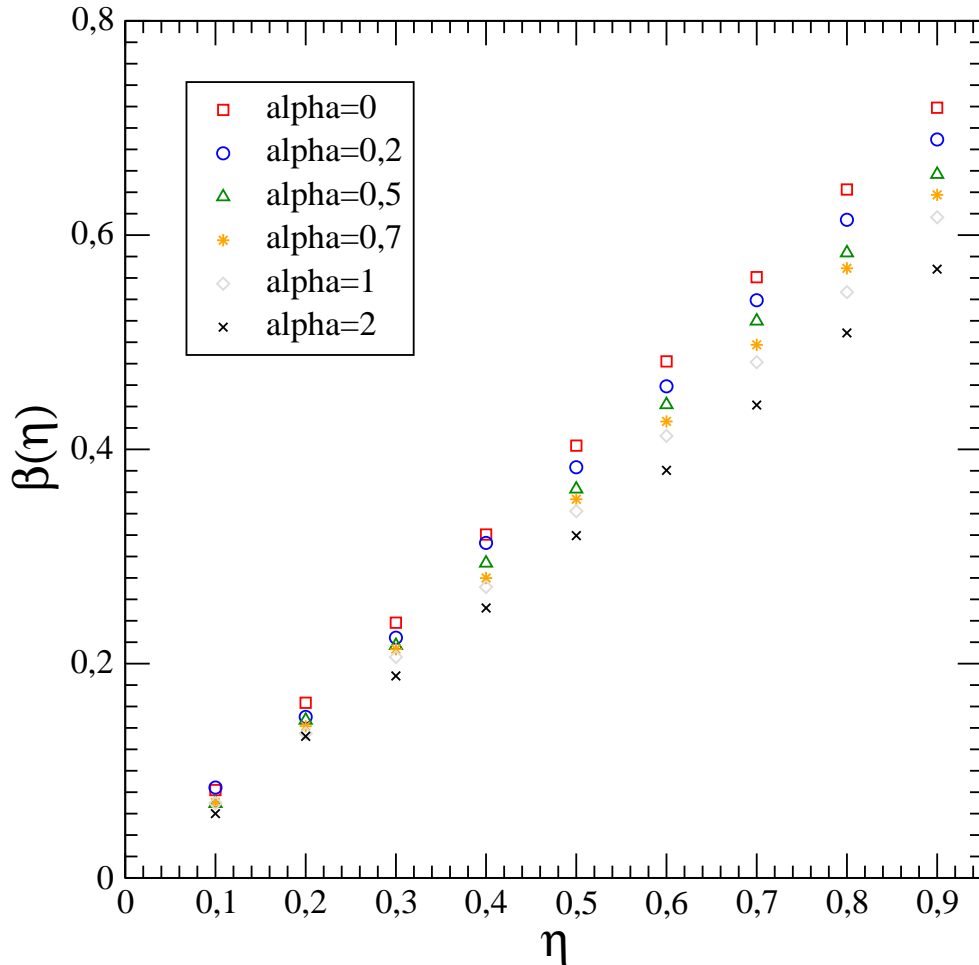


Figura 3.8: Dependência temporal da conectividade, $\beta(\eta)$, para sítios com diferentes qualidades e com diferentes α . Notemos que com o aumento do fator de qualidade, a conectividade do sítio cresce muito mais rapidamente durante a evolução da rede.

nor caminho cresce com $\log(N)$. Outro fato observado é que o modelo de Barabási apresenta uma distância média entre os sítios maior quando comparado com o modelo de Bianconi-Barabási para qualquer valor de α . Por fim, podemos através dos gráficos 3.10 e 2.11 dizer que as redes aleatórias possuem um menor caminho médio maior que o modelo de Barabási e do que o modelo de Bianconi-Barabási com distribuição de qualidade em lei de potência. Podemos entender este resultado da seguinte forma (sugestão): os pólos (sítios com muitas ligações) fazem diminuir a distância entre os nós e como o modelo de Bianconi possui mais pólos que o modelo de Barabási ou do que as redes aleatórias (que não possuem pólos), então o modelo de qua-

lidade possui um valor menor para o menor caminho médio. Queremos enfatizar que esta quantidade (menor caminho médio) é de extrema importância para vários sistemas na natureza e na tecnologia, por exemplo se a distância entre os primeiros vizinhos de uma rede apresentam um tamanho único e estas transportarem informações, então as redes com uma quantidade maior de "hubs" conduzirão informações mais rapidamente do que as que não possuem tantos pólos. Outro exemplo é a rede da cadeia alimentar, onde o menor caminho médio é a distância na cadeia alimentar entre a presa e o predador, se $\langle l \rangle = 1$ todos são predadores ou presas um do outro, por outro lado se $\langle l \rangle = 2,65$ (dado da tabela 1.1)[25] existe uma determinada hierarquia entre os animais na cadeia alimentar.

Fizemos o estudo do coeficiente de agregação da rede com diferentes valores de α e os resultados são apresentados no gráfico 3.11. Analisando e comparando os gráficos 3.11 e 2.12, vemos que o modelo de qualidade com $\alpha = 0$ (competitividade alta) possui um coeficiente de agregação maior do que aqueles encontrados em todos os outros modelos citados nessa dissertação. Ao aumentarmos o valor de α fazemos com que o coeficiente de agregação diminua até chegar no valor limite que é o do modelo de Barabási ($\alpha = 50$ já encontramos resultados similares aos do modelo BA). Outro fato observado, é que ao aumentarmos o número de ligações, m , adicionadas a cada passo de tempo, aumentamos o valor deste coeficiente como observamos no gráfico interior 3.11. Lembrado que grafos que não possuem circuitos, possuem coeficiente de agregação nulo. Devido a este fato não foi colocado dados sobre este coeficiente em redes com $m = 1$. Esta quantidade localmente esta associada com o número relativo das conexões entre os vizinhos mais próximos de um determinado nó, assim entendemos que uma rede com elevado coeficiente de agregação possuirá um pequeno valor para o caminho médio. Este fato é verificado ao analisarmos os gráficos 3.10 e 3.11.

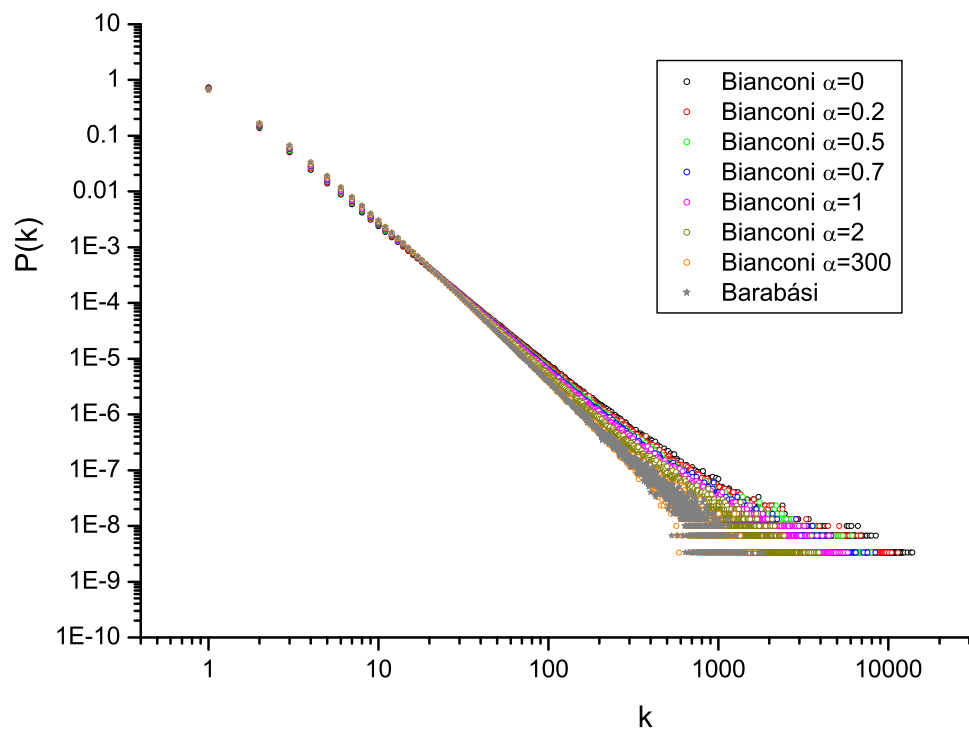


Figura 3.9: Resultado da simulação do modelo de Bianconi com qualidade em lei de potência (rede com $N = 10^5$ e $m = 1$): Observamos que a distribuição de conectividade é modificada ao variarmos α .

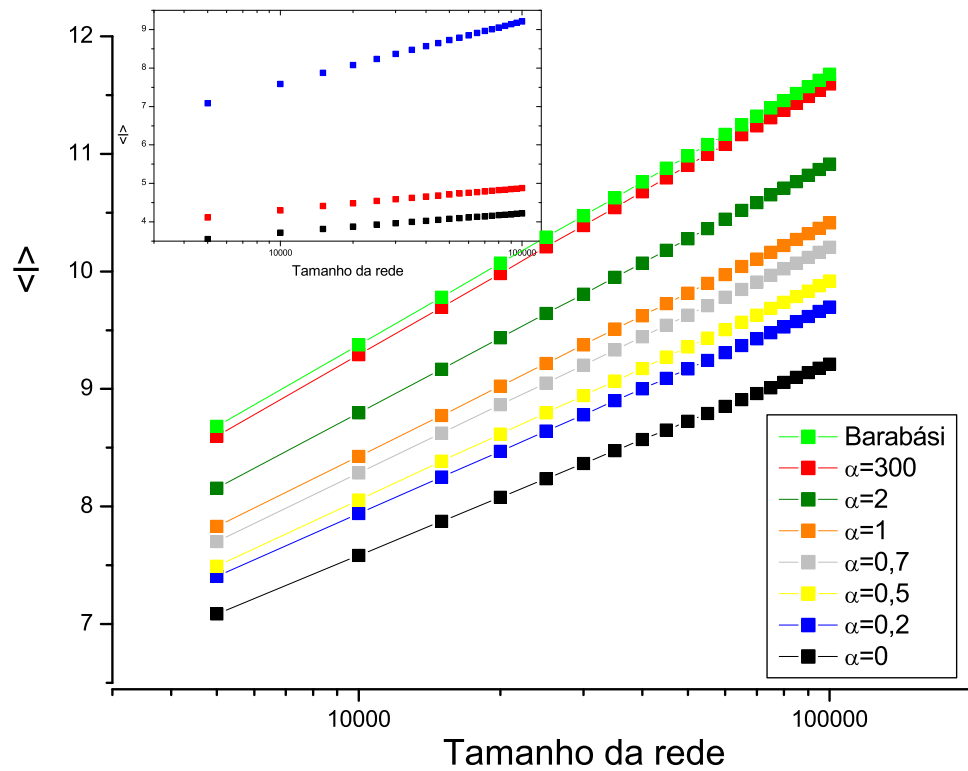


Figura 3.10: Resultado da simulação numérica do modelo de Bianconi com qualidade em lei de potência: Menor caminho médio versus o tamanho da rede em escala linear-log para um rede com $m = 1$. O gráfico interno compara o menor caminho médio entre uma rede com $m = 1$ (pontos azuis), $m = 2$ (pontos vermelhos), $m = 3$ (pontos pretos).

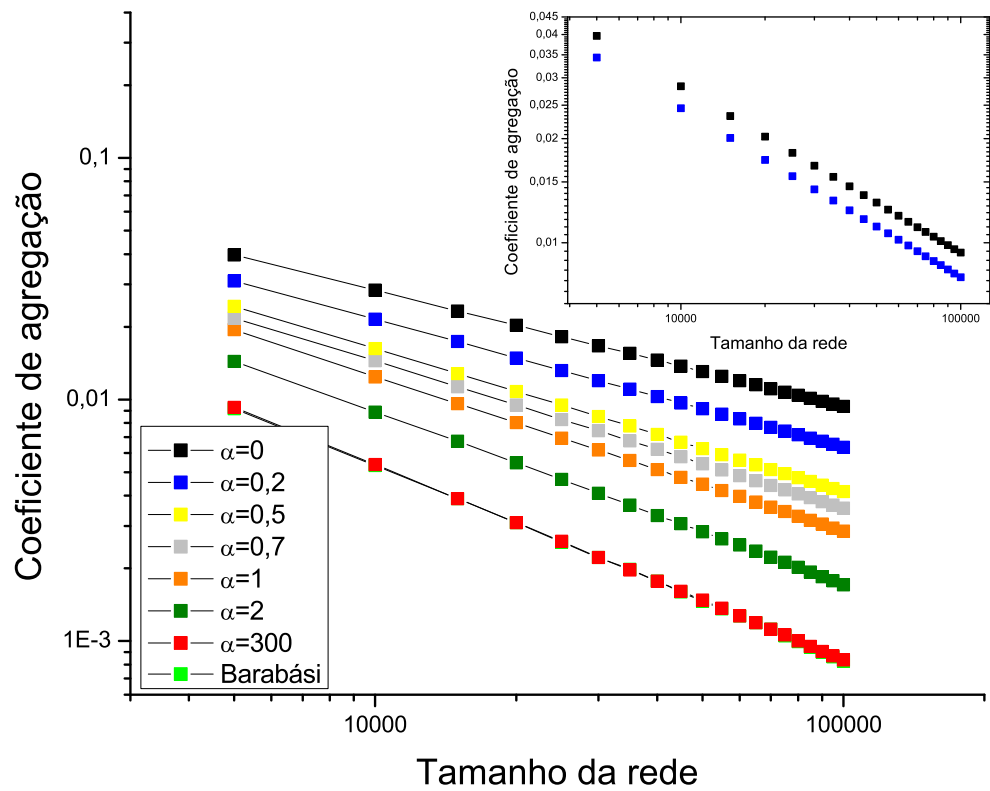


Figura 3.11: Resultado da simulação numérica do modelo de Bianconi com distribuição de qualidade em lei de potência: Coeficiente de agregação total versus o tamanho da rede no gráfico log-log para um rede com $m = 3$. O gráfico interno compara o coeficiente de agregação entre uma rede com $m = 2$ (pontos azuis) e $m = 3$ (pontos pretos).

Conclusão

Muitos são os problemas interessantes e fascinantes na natureza a serem estudados. A teoria dos sistemas complexos tem ajudado para que muitos destes problemas se tornem compreensíveis para o homem. Uma das características dos sistemas complexos é que são formados por um grande número de unidades simples, porém conectadas entre si (uma unidade influencia a outra) resultando num sistema macroscópico surpreendente. A evolução do sistema se torna complexa devido às várias interações entre as unidades. A complexidade surge do fato que de vizinho em vizinho surge correlação de longo alcance, assim não podemos estudar partes do sistema senão o todo. Um tratamento analítico destes sistemas constitui uma dificuldade aos cientistas, devido a complexidade e a dinâmica envolvidas. Nesse trabalho conseguimos entender parte do sistema estudado através de técnicas analíticas e de forma complementar fizemos uso da simulação numérica. Estudamos a evolução temporal de redes complexas generalizando o modelo de Bianconi-Barabási, onde incluímos uma distribuição de qualidade obedecendo uma lei de potência ($\rho \propto \eta^\alpha$). Os resultados obtidos oferecem interessantes sinais da evolução dos constituintes (nós) num ambiente competitivo. O nosso modelo reflete as propriedades básicas de vários sistemas reais em que os nós competem entre si para adquirir ligações (os sítios já existentes disputam para ver quem se liga aos sítios recém incorporados a rede) semelhante ao modelo estudado por Bianconi-Barabási. No modelo padrão livre de escala, onde cada nó tem a mesma qualidade, todos os nós aumentam sua conectividade seguindo o mesmo expoente de escala $\beta = 1/2$. Contrariamente, no modelo de qualidade com uma distribuição constante verificamos sítios com diferentes qualidades, resultando numa escala múltipla (cada sítio tem um expoente dinâmico) e o expoente dinâmico dependente do parâmetro de qualidade do sistema, $\beta = \eta/C$. Isto permite que nós com uma alta qualidade

entrem no sistema num tempo posterior a nós que estejam no sistema há mais tempo e tornam-se muito conectados. Por exemplo, nós com um parâmetro de qualidade alto correspondem, por exemplo, a pessoas com uma alta qualidade social (extrovertido, amigável, etc); websites com melhores conteúdos ou serviços; ou artigos que tratam de algumas descobertas importantes. O que é mais interessante é que apesar destas diferenças significativas em suas qualidades, *todos os nós continuam aumentando sua conectividade seguindo uma lei de potência no tempo.*

Todos estes resultados e esta familiaridade com os sistemas reais, nos motivaram a estudar o modelo de qualidade de forma mais detalhada e nada melhor que verificar o que ocorre com este modelo ao mudar sua distribuição de qualidade, particularmente, estudamos a distribuição de qualidade em lei de potência. Vimos que a distribuição de qualidade em lei de potência ($\rho(\eta) \propto \eta^\alpha$) varia o valor de C de 1,255 ($\alpha = 0$) até 2 ($\alpha \rightarrow \infty$) onde todos os valores de C puderam ser calculados analiticamente, em particular foi mostrado os valores de C para sistemas com α variando de 0 até 300 (gráfico 3.6). Assim obtivemos o expoente dinâmico, $\beta = \eta/C$, para todos estes sistemas. Desta expressão vimos que os sítios com maiores qualidades possuem uma capacidade maior em adquirir novas conexões do que sítios com baixas qualidades.

O nosso modelo apresenta dois limites, o primeiro com $\alpha = 0$ (caso em que a distribuição de qualidade é constante) e o segundo com $\alpha \rightarrow \infty$ (modelo de Barabási-Albert). Assim nosso modelo apresenta várias distribuições de conectividade que estão entre estes dois limites (gráfico 3.3) porém a variação de α entre estes extremos ocorre de forma suave. Outras propriedades da rede apresentaram diferenças relevantes com a variação de α , como o menor caminho médio e o coeficiente de agregação. A partir de C obtemos a distribuição de conectividade $P(k)$.

Observamos que a rede em estudo tem um comportamento de mundo pequeno, pois o menor caminho médio cresce com $\log(N)$. Outro ponto interessante é que quanto maior for o valor de α em nossa rede, maior será o valor do menor caminho médio, porém sempre terá um valor menor que o do modelo de redes aleatórias independentemente do valor de m . Nos modelos trabalhados nesta dissertação, estas características parecem estar relacionadas com o coeficiente de agregação, devido ao seguinte comportamento: uma

rede tendo um grande menor caminho médio, possui um pequeno coeficiente de agregação e vice versa. Para o coeficiente de agregação, se aumentamos o valor de m , aumentaremos o valor deste coeficiente.

Outras propriedades para redes podem ser estudadas, aqui trabalhamos com as quantidades mais conhecidas e estabelecidas a fim de compreender e verificar o quão satisfatório é o nosso modelo em relação aos sistemas reais, que tem sido cada vez mais estudados não só na física como por várias outras áreas da ciência. Danyel et al. [44] e Marcelo et al. [45] estudaram o crescimento das redes complexas levando em conta uma origem para o sistema, sendo esta o seu centro de massa. No trabalho deles a distância Euclidiana foi incorporada a ligação preferencial. Em nosso trabalho a rede estudada é puramente topológica. Neste sentido, podemos refazer o nosso estudo incluindo estas características que permitem levar em conta a distância Euclidiana entre os sítios. Outros trabalhos que podemos fazer no futuro é verificar a robustez das nossas redes. Também pretendemos estudar outros modelos, em particular o que leva em conta a semelhança entre as qualidade dos sítios. Este modelo propicia o estudo de diferentes propriedades, uma delas é a presença de comunidades. Consideramos também interessante estudar nas nossas redes propagação de informação, difusão de epidemias, etc.

Bibliografia

- [1] R. Albert, H. Jeong and A.-L. Barabási, *Nature* **401**, 130 (1999).
- [2] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal, *Proceedings of the 19th ACM Symposium on Principles of Database Systems*, 1-10 (2000a)
- [3] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, *Comput. Netw.* **33**, 309 (2000).
- [4] M.E.J. Newman, S.H. Strogatz and D.J. Watts, *Phys. Rev. E* **64**, 026118 (2001).
- [5] L.A. Adamic and B.A. Huberman, *Science* **287**, 2115a (2000).
- [6] L.A. Adamic, *Proceedings of ECDL'99, LNCS 1696*, 443-452 (1999).
- [7] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, *Comput. Netw.* **33**, 309 (2000).
- [8] AD.M. Pennock, G.W. Flake, S. Lawrence, E.J. Glover and C.L. Giles, *Proc. Natl Acad. Sci. USA* **99**, 5207 (2002).
- [9] M. Faloutsos, P. Faloutsos and C. Faloutsos, *Comp. Commun. Rev.* **29**, 251 (1999).
- [10] R. Pastor-Satorrs, A. Vázquez and A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).
- [11] R. Govindan and H. Tangmunarunkit, *Proceedings of the 2000 IEEE INFOCOM Conference*, 1371-1380 (2000).
- [12] S. Redner, *Eur. Phys. J. B* **23**, 267 (1998).

- [13] C. Tsallis and M.P. de Albuquerque, *Eur. Phys. J. B* **13**, 777 (2000).
- [14] P.L Krapivsky, S. Redner and F. Leyvraz, *Phys. Rev. Lett.* **85**, 4629 (2000).
- [15] A. Vásquez, Statistics of citation networks, cond-mat/0105031 (2001b).
- [16] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [17] R. Albert and A.-L. Barabási, *Phys. Rev. Lett.* **85**, 5234 (2000).
- [18] M.E.J. Newman, *Phys. Rev. E* **64**, 016131 (2001e).
- [19] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert and T. Vicsek, *Physica A* **311**, 590 (2002).
- [20] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley and Y. Aberg, *Nature* **411**, 907 (2001).
- [21] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai and A.-L. Barabási, *Nature* **407**, 651 (2000).
- [22] A. Wagner and D.A Fell, *Proc. R. Soc. London B* **268**, 1803 (2001).
- [23] H. Jeong, S.P. Mason, A.-L. Barabási and Z.N. Oltvai, *Nature* **411**, 41 (2001).
- [24] A. Wagner, *Mol. Biol. Evol.* **18**, 1283 (2001a).
- [25] R. Ferrer i Cancho and R.V. Solé, *Proc. R. Soc. B* **268**, 2261 (2001a).
- [26] J.M. Montoya and R.V. Solé, Working Papers of Santa Fe Institute, 01-11-069 (2001).
- [27] S. Valverde and R.V. Solé, *Physica A* **312**, 636 (2002).
- [28] R. Ferrer i Cancho and R. Solé, cond-mat/0111222.
- [29] W. Aiello, F. Chung and L. Lu, Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, 171-180 (2000).
- [30] S. Abe and N. Suzuki, *Phys. Rev. E* **67**, 016106 (2003).
- [31] J.P.K. Doye, *Phys. Rev. Lett.* **88**, 238701 (2002).

- [32] D.J.B. Soares, J.S. Andrade Jr., H.J. Herrmann and L.R. da Silva, International Journal of Modern Physics C **17**, 1219 (2006).
- [33] J.S. Andrade Jr., H.J. Herrmann, R.F. Andrade and L.R. da Silva, Physical Review Letters **94**, 01870 (2005).
- [34] M.M. Soares, G. Corso and L.S. Lucena, Physica A **355**, 678-684 (2005).
- [35] P. Erdős and A. Rényi, Publ. Math. **6** (1959).
- [36] D.J.B. Soares, J.S. Andrade JR, H.J. Herrmann and L.R. da Silva, International Journal of Modern Physics C, **17**, 1219 (2006).
- [37] J.S. Andrade JR, H.J. Herrmann, R.F. Andrade and L.R. da Silva, Physical Review Letters **94**, 018702 (2005).
- [38] S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes, Phys. Rev. E **65**, 066122 (2002).
- [39] D.J. Watts, Princeton University Press (1999).
- [40] D.J. Watts and S.H. Strogatz, Nature **393**, 440 (1998).
- [41] S.H. Strogatz, Nature **410**, 268 (2001).
- [42] D.J.B. Soares, Tese de doutorado, Departamento de Física-UFRN, (2004).
- [43] A.A. Moreira, Tese de doutorado, Departamento de Física-UFC, (2002).
- [44] D.J.B. Soares, C. Tsallis, A.M. Mariz and L.R. da Silva, Europhysics Letters, **70**, 70-71 (2005).
- [45] M.D. de Meneses, D.J.B. Soares, S.D. da Cunha and L.R. da Silva, Progress of Theoretical Physics Supplement, **162**, 131-137 (2006).
- [46] S. Milgram, Psychol. today **2**, 60 (1967).
- [47] L.C. Freeman, Sociometry **40**, 35 (1977).
- [48] K.-I. Goh, B.Kahng and D.Kim, Phys. Rev. E **64**, 051903 (2001).
- [49] S. Lawrence and C.L. Giles, Science **280**, 98 (1998).

- [50] S. Lawrence and C.L. Giles, *Nature* **400**, 107 (1999).
- [51] R. Kumar, P.Raghavan, S. Rajalopagan and A. Tomkins, *Proceedings of the 9th ACM Symposium on Principles of Database Systems*, 1 (1999).
- [52] S.H. Yook, H. Jeong and A.-L. Barabási, *PNAS*, **99**, 13382 (2002).
- [53] R.J. Williams, N.D. Martinez, E.L. Berlow, J.A. Dunne and A.-L. Barabási, *Santa Fe Institute Working Paper*, 01-07-036 (2000).
- [54] R. Ferrer i Cancho and R. Solé, *Santa Fe Institute Working Paper*, 01-03-016 (2001).
- [55] S.N. Dorogovtsev and J.F.F. Mendes, *Adv. Phys* **51**, 1079 (2002).
- [56] R. Albert and A.-L. Barabási, *Rev, Mod Phys.*, **74**, 1 (2002).
- [57] B. Bollobás, *Random Graphs*, (1985).
- [58] A.-L. Barabási, R. Albert and H. Jeong, *Physica A*, **272**, 173 (1999).
- [59] G. Bianconi and A.-L. Barabási, *Europhys. Lett.*, **54**, 436 (2001).

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)