

Introdução às Redes Complexas

Nilton Alves Jr.

10 de julho de 2008

Resumo

O objetivo deste texto¹ é abordar os princípios conceituais básicos da área de Redes complexas, também conhecida como Redes de Conexão ou simplesmente Redes. Este texto será utilizado no curso de mesmo nome da VII Escola do CBPF.

1 Introdução

A cerca de duas décadas, o ambiente acadêmico passou a dar grande importância aos sistemas complexos, em particular às redes de conexões ou redes complexas. Redes, elas estão por todos os lados. Diversas estruturas organizadas em redes estão presentes em nosso cotidiano, e.g., as redes de distribuição elétrica, redes sociais, redes rodoviárias, redes de computadores, redes de neurônios, etc.

Até recentemente os estudos de sistemas em redes eram principalmente descritos por uma área da matemática chamada teoria dos grafos. A falta de dados experimentais de tais sistemas levavam a uma abordagem por meio de uma estrutura aleatória, conhecida por *random-graphs networks*. As primeiras análises de estruturas em redes foram introduzidas por Erdős e Rényi [ER59][ER60] em meados de 1960. O modelo proposto na época consistia de nós interconectados entre si com probabilidade p . Através desse tipo de consideração uma rede aleatória segue uma distribuição de Poisson, fazendo com que seja raro encontrar nós com concentração de conexões ou muito grande ou muito pequena.

Vários trabalhos relativos a redes complexas seguiram esta linha sempre encontrando uma dificuldade comum que era a dificuldade experimental de obter dados e testar suas teorias e seus métodos. Como obter o mapeamento de conexões de nossos neurônios? Como verificar se os modelos se aplicam a rede de conexões aéreas? Como simular e comparar o crescimento de uma colônia de bactérias ou de formigas? Tudo era possível somente com redes pequenas, com poucos nós e conexões.

A Internet, com o advento do ambiente web, com suas páginas com hipertexto², que ligam outras páginas que por sua vez se conectam a outras tantas e assim por diante, revelou-se uma rede complexa, de tamanho grande e crescente e o mais importante, neste caso, passível de gerar imensas quantidades de dados com facilidade.

Sendo assim, a Internet passou a ser uma excelente rede complexa para estudo, pois além de permitir que teorias, modelos e técnicas fossem desenvolvidas e testadas, a extensão disto para outras redes mais inacessíveis, de difícil aquisição de dados tais como rede de neurônios ou rede sociais, é imediata. A Internet se tornou o objeto de estudo de laboratório para pesquisa na sub-área Sistemas Complexos da Física Estatística.

¹É parte integrante da tese de doutoramento do mesmo autor intitulada "Caracterização de Redes Complexas - Aplicação à Modelagem Relacional entre Sistemas Autônomos da Internet", Instituto Politécnico da Universidade do Estado do Rio de Janeiro - IP/UERJ, março de 2007.

²Hipertexto é um texto organizado em forma de rede de itens ou módulos de informação interligados entre si, permitindo ao usuário "navegar" seguindo sua própria seqüência de estudo.

2 Redes de Conexão

Redes complexas, redes de conexão, redes sociais ou simplesmente redes, é uma recente área interdisciplinar que envolve o formalismo matemático da Teoria dos Grafos e a análise baseada em ferramentas da Mecânica Estatística. Uma rede é um conjunto de elementos que são associados a nós ou vértices cuja ligação entre si se dá por meio de uma aresta. Uma maneira intuitiva de trabalhar com redes é utilizar a matriz adjacência $M[i, j]$ característica do sistema, onde os índices i e j representam os nós e os elementos de matriz m_{ij} representam as ligações entre os nós, as arestas. As ligações podem ser unidirecionais, bidirecionais ou sem direção, caso da matriz simétrica. Podem ainda ser simples, $a(i, j) = 1$ ou $a(i, j) = 0$, ou com pesos diferenciados, $a(i, j) \in \mathfrak{R}$.

É comum se considerar o trabalho de Leonhard Euler, em 1735, em St. Peterburg, que resolveu o chamado problema das pontes de Königsberg (Prússia, no séc. XVIII), como sendo o ponto inicial da ciência das redes. Nesta cidade, atualmente Kaliningrado, Rússia, formada por duas grandes ilhas, haviam sete pontes ligando-as. Se discutia na época se seria possível fazer um percurso passando por todas as pontes uma única vez. Apesar da cultura local, Euler provou em seu trabalho [Eul53] a impossibilidade de um caminho fechado e para isto ele elaborou, provavelmente, o primeiro grafo matemático, considerando as pontes como arestas.

O estudo de muitas redes complexas teve como motivação o desejo de entender diversos sistemas reais que vão desde redes de comunicação até sistemas de cadeias ecológicas. Podemos citar como exemplos mais estudados recentemente o ambiente *World Wide Web*, a Internet, a rede de colaboração de atores de cinema, a rede de contatos sexuais humanos, as redes celulares, as redes ecológicas, as redes de telefonia, as redes de citações científicas, as redes lingüísticas, as redes de transmissão elétrica, as redes de neurônios e as redes de interação de proteínas, dentre outras.

De maneira geral, podemos classificar as redes complexas em quatro tipos diferenciados: as redes sociais, as redes tecnológicas, as redes biológicas e as redes de informação.

Uma rede social é a caracterização de um grupo de pessoas ou mais especificamente, um grupo de pessoas com padrão de contato ou relacionamento entre si, [Sco00] e [WF94]. Neste tipo de rede, podemos destacar os trabalhos passados relativos às amizades entre indivíduos [Mor34] e [RH61], às relações comerciais entre empresas [Mar75] e [Miz82] ou ainda às relações familiares [PA93]. Podemos também considerar neste tipo de rede o relacionamento entre super-heróis Marvel Comics [AMJR02] ou entre animais, brilhantemente descrito em [CHB99] onde em uma comunidade de cerca de 400 golfinhos, 14 indivíduos formaram um nível de super-aliança que subjuga o nível comum de aliança entre os demais indivíduos.

Ainda falando em redes sociais, um conjunto de experimentos que teve um grande impacto na sociedade acadêmica mundial é aquele feito por Milgram [Mil67] e [TM69] e que ficou conhecido pelo fenômeno *small world*. Em 1967, Milgram conduziu um experimento seminal para testar a hipótese de que membros de uma grande rede social, no caso a população dos Estados Unidos, estariam ligados entre si por uma pequena cadeia de conhecimentos intermediários. Milgram enviou mensagens para algumas centenas de indivíduos selecionados aleatoriamente para que encaminhassem para alguém próximo, com o objetivo de alcançar um indivíduo alvo em uma região geográfica distante. O resultado médio de seis indivíduos, incluindo ele, necessários para fechar uma cadeia entre ele e o indivíduo destino se transformou em um dogma sociológico e uma verdade popular. Atualmente existe o projeto *Small World* na Universidade de Columbia [Wat07] que tenta, através do correio eletrônico, levantar algumas questões ainda em aberto tais como a generalização do conceito em função de raça, classe social, nacionalidade, ocupação profissional, etc.

As redes tecnológicas são aquelas feitas ou construídas pelo homem diretamente tais como as redes

de distribuição de eletricidade [ASBS00], as rotas das linhas férreas [SDC⁺02] ou a rede que retrata o movimento de pedestres em uma cidade [CHE02].

Outra rede tecnológica muito estudada é a Internet, principalmente pela facilidade na obtenção dos dados. A Internet é uma rede de conexões físicas entre computadores que possuem endereçamento IP³ chamados *hosts*. Como a taxa de crescimento do número de *hosts* é muito grande, normalmente os estudos são feitos usando um outro sistema de maior granulometria como por exemplo roteadores, servidores de e-mail ou de ambiente Web ou através de Sistemas Autônomos - AS (sigla do nome em inglês).

Outro tipo de rede, as redes biológicas, por possuírem intrinsecamente uma grande dificuldade na obtenção de dados, se utilizam também das teorias e modelos desenvolvidos para os outros tipos de redes. Estudos sobre as propriedades estatísticas das redes metabólicas, por exemplo, foram bem estudadas por Jeong *et al.* [JTA⁺00], Stelling *et al.* [SKB⁺02] e Wagner e Fell [WF01].

O quarto tipo de rede considerado, redes de informação ou redes de conhecimento, tem como exemplo clássico a rede de citações entre artigos acadêmicos [ER90]. Por exemplo, os artigos citados neste trabalho formam uma rede complexa de tópicos relativos. Cada artigo citado é um vértice e uma conexão, neste caso unidirecional, entre o artigo *A* e o artigo *B*, significa que o artigo *B* cita o artigo *A* em seu conteúdo. Devido a propriedade intrínseca de que cada artigo cita aqueles outros já existentes, esta rede basicamente aponta os relacionamentos do passado. Um dos primeiros estudos neste tipo de rede foi realizado por Alfred Lotka [Lot26], em 1926, quando desenvolveu a Lei de Produtividade Científica, que basicamente afirma que a distribuição do número de trabalhos publicados *k* por cientistas segue uma lei de potência do tipo $k^{-\gamma}$ onde γ é o expoente da lei a ser determinado.

Apesar das primeiras bases de dados da área de bibliometria⁴ terem sido integradas em 1960 por Eugene Garfield e outros pioneiros, neste trabalho vamos citar os estudos mais recentes de Seglen [Seg92], Redner [Red98] e Tsallis e de Albuquerque [TdA00].

Um outro importante exemplo de uma rede de informação é o ambiente *World Wide Web* que é composto de servidores Web e o emaranhado de conexões entre suas páginas de hipertexto. Esta rede, diferentemente das redes de citações científicas, apresenta uma característica cíclica. Este ambiente foi amplamente estudado desde o seu início no começo dos anos 90 com destaque para os trabalhos de Albert *et al.* [AJB99], Barabási *et al.* [BAJ00], Kleinberg [Kle00] e Broder *et al.* [BKM⁺00]. É importante destacar que o ambiente WWW não é a Internet propriamente dita, que é o objeto deste trabalho e que é uma rede física de computadores, dispositivos computacionais (impressoras, *data storages*, *clusters*, etc) e equipamentos de rede (roteadores, comutadores, etc).

3 Sistemas Complexos

Sistemas constituídos por muitos corpos, em escalas diferentes, que interagem entre si e com o próprio meio de forma não linear, são considerados sistemas complexos. Estes sistemas podem ser naturais como um formigueiro, o conjunto de neurônios em cérebros humanos, ou não naturais como as linhas de transmissões de energia elétrica, a economia e a Internet.

No final do século XX, parte da comunidade de físicos passou a dar importância ao estudo da dinâmica dos sistemas complexos que tem como principal característica a interação não linear entre seus elementos. A tentativa de construir teorias e modelos que contemplassem estes sistemas originou a teoria do caos e a física dos sistemas complexos.

³IP é um acrônimo para a expressão inglesa *Internet Protocol*, que é um protocolo usado entre duas máquinas em rede do ethernet, para encaminhamento dos dados.

⁴Bibliometria é um campo da ciência da informação, que infere sobre a produção bibliográfica de um determinado autor, tentando medir a produtividade de cada autor e assim criar métodos de comparação entre vários. O termo foi primeiramente usado por Pritchard e por Nalimov e Mulchenko em 1969. <http://www.steunpuntoos.be/bibliometrics.html> em 31/01/07.

Em sistemas lineares, quando o elemento A interage com o elemento B de forma linear, uma mudança em A implica em uma mudança proporcional em B . Quando existem muitas não-linearidades nas interações em um sistema de muitos componentes, o sistema não pode ser analisado como sendo a soma de suas partes constituintes. Nestes casos, a predição de um comportamento não é possível e, assim, teorias e métodos desenvolvidos na área de sistemas complexos são aplicáveis.

Os conceitos de criticalidade auto-organizada [BTW87][BTW88], auto-similaridade [LTWW94], fractais e leis de potência [Mit03] fazem parte da física estatística mais recente e em particular, da física de sistemas complexos.

4 Redes Aleatórias

Redes aleatórias é um campo da Ciência da Computação que teve um desenvolvimento mais intenso nos anos recentes. Está baseada no estudo de grafos aleatórios que está em uma interseção entre a teoria dos grafos e a teoria de probabilidades.

Em termos matemáticos, a representação de um grafo é feita através da notação $G = \{P, E\}$, onde P é um conjunto de N nós ou vértices ou pontos P_1, P_2, \dots, P_N e E é um conjunto de arestas ou conexões ou linhas entre dois elementos de P . Na figura 1 observa-se um grafo simples bidirecional com o conjunto de vértices ou nós $V = \{1, 2, 3, 4, 5, 6\}$ e com um conjunto de arestas ou conexões $E = \{\{1, 2\}, \{1, 5\}, \{2, 3\}, \{2, 5\}, \{3, 4\}, \{4, 5\}, \{4, 6\}\}$, com o mapeamento dos pesos das conexões w sendo a identidade.

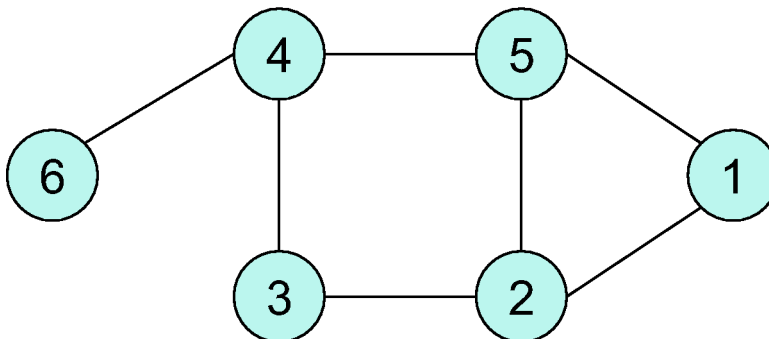


Figura 1: Exemplo de um grafo com seis vértices ou nós. Figura retirada de http://pt.wikipedia.org/wiki/Teoria_dos_grafos em fevereiro de 2007.

A Teoria dos Grafos teve sua origem no século XVIII com o trabalho de Leonhard Euler que se concentrava em grafos pequenos e com alto grau de regularidade. Já no século XX, a Teoria dos Grafos passou a ser utilizada na análise de sistemas grandes e em conjunto com conceitos estatísticos e através de algoritmos computacionais. Muitos conceitos foram desenvolvidos a partir de grafos aleatórios, onde as conexões são aleatórias. Inicialmente passaram a ser estudados também por esta teoria as redes com topologia complexa e princípios de conectividade desconhecidos.

A Teoria de Grafos Aleatórios foi inicialmente introduzida por Paul Erdős e Alfréd Rényi [ER59], [ER60] e [ER61]. A título de referência deve-se considerar o artigo de Cohen [Coh88] e o livro de Bollobás [Bol01]. O modelo de Erdős-Rényi começa pela definição de um grafo aleatório com N nós e com n conexões escolhidas de maneira aleatória entre as $N(N-1)/2$ conexões possíveis. Existem $C_{N(N-1)/2}^n$ grafos possíveis que formam um espaço de probabilidade em que cada grafo é igualmente provável. Conseqüentemente, o número total de conexões é uma variável aleatória proporcional a

probabilidade p e com valor esperado

$$E(n) = p[N(N-1)/2] \quad (1)$$

Considerando que G_0 é um grafo com N nós e n conexões, a probabilidade de obtê-lo por esta construção é dada por

$$P(G_0) = p^n(1-p)^{N(N-1)/2-n} \quad (2)$$

Este modelo possui algumas características que não se aplicam às redes complexas. Podemos destacar dois aspectos:

- O número fixo de elementos em contraposição com o número crescente encontrado em muitas redes
- O grau de conectividade de muitas redes não está em concordância com a função de Poisson

Uma propriedade marcante na teoria dos grafos, no caso de grafos aleatórios (*random graphs*), é a existência de uma probabilidade crítica acima da qual emerge a existência de um aglomerado gigante envolvendo na maioria das vezes todos os elementos do grafo. Abaixo desta probabilidade de conexão, o grafo é composto de vários aglomerados menores e isolados. Este fenômeno é similar à teoria da percolação amplamente estudada em matemática e mecânica estatística [bAH00].

O modelo de Erdős-Rényi, apesar de ser robusto para muitas redes, em particular para as redes aleatórias, não satisfaz a descrição daquelas redes que apresentam crescimento contínuo e baseado em uma lei de potência. Na próxima seção abordaremos este tipo de rede.

5 Redes Sem Escala

Muitas das redes reais diferem das redes aleatórias pois apresentam uma distribuição de graus de conectividade que segue a lei de potência (3). Como as leis de potência são livres de qualquer escala característica, estas redes são chamadas de redes livres de escala, redes de escala livre ou redes sem escala [Ada07].

Redes sem escala, são um tipo de rede complexa que atraiu a atenção dos pesquisadores porque várias redes reais caem nesta categoria. Diferentemente das redes aleatórias onde a distribuição de conectividade segue a distribuição de Poisson, no caso de redes livres de escala, alguns poucos elementos são muito conectados enquanto a maioria dos demais possuem baixo índice de conectividade. Este tipo de rede é independente do número N de elementos. Sua principal característica, que a diferencia da rede aleatória, é a probabilidade de conexão que é dada por [BA99] [SFFF03]

$$P(k) \sim k^{-\gamma} \quad (3)$$

onde k é o coeficiente de conectividade ou número de conexões e o expoente γ varia aproximadamente entre 2 e 3 para a maioria das redes reais [BA02].

Uma rede livre de escala pode ser construída adicionando-se elementos progressivamente à rede existente através de conexões com os elementos já participantes da rede seguindo o princípio de conexão preferencial com a probabilidade sendo dada por

$$P(k_i) = \frac{k_i}{\sum_{j=1}^N k_j} \quad (4)$$

onde k_i é o número de conexões do i ésimo elemento ou nó e N é o número total destes elementos [BA99].

Aqueles elementos que possuem um elevado valor do coeficiente de conectividade k são considerados *hubs*.

A figura 2 (pág. 7) apresenta um diagrama comparativo entre redes aleatórias que seguem a distribuição de Poisson, coluna esquerda e as redes em lei de potência que não possuem escala, coluna direita.

Na figura 2a, uma rede baseada no modelo de grafo aleatório de Erdős-Rényi é construída com $N = 10$ nós conectados aos pares com probabilidade $p = 0,2$. Considerando que as conexões possuem pesos iguais e unitários, $\{i, j\} = \{j, i\}$, então de acordo com a equação (1), página 5, o valor esperado do número de conexões para este caso é $p[N(N - 1)/2] = 9$.

Na figura 2b é utilizado o modelo sem escala que considera que a rede cresce continuamente e que a cada instante um novo nó se conecta a dois outros nós já existentes, preferencialmente aqueles com maior número de conexões. Este mecanismo é chamado de conexão preferencial.

A rede de conexão pode ser caracterizada pela probabilidade de conexão $P(k)$ que o i ésimo nó tenha k_i conexões. Na figura 2c, observa-se que para rede aleatória, $P(k)$ segue a distribuição de Poisson que está centrada no valor médio $\langle k \rangle$. Basicamente esta figura mostra que a maioria dos nós possuem $k = \langle k \rangle$ e que a probabilidade decai exponencialmente a medida que o número de conexões se afasta deste valor.

Já na figura 2d, é visível que a probabilidade segue uma lei de potência, dada pela equação (4) onde é muito provável nós com poucas conexões e ao contrário, pouco provável nós muito conectados.

Nas figuras 2e e 2f, observam-se redes de 130 nós e 430 conexões e onde é possível constatar que a rede aleatória é mais homogênea enquanto que na rede sem escala, a maioria dos nós estão conectados a alguns poucos nós, chamados de *hubs*. A importância destes nós *hubs* para este tipo de rede pode ser avaliada pelo fato de, estatisticamente, 60% dos nós serem alcançados através de apenas 5 deles em comparação com 27% no caso de redes aleatórias.

Observa-se que a estabilidade da rede de conexões apesar de ser maior se ela for do tipo sem escala, os nós *hubs* não podem ser afetados pois, se isto ocorrer, existe a possibilidade de criação de aglomerados de nós e conseqüente isolamento.

Na natureza estes dois principais tipos de redes de conexões existem, assim como existem também aquelas que são uma mistura delas. As redes aleatórias são mais simples e foram primeiramente bem estudadas. As redes sem escala tiveram sua natureza e propriedades conhecidas mais recentemente, principalmente nas últimas duas décadas (extensa documentação no recente livro de Newman et al. [NBW06]).

Na próxima seção, são descritos os principais modelos genéricos de crescimento ordenado de redes de conexões. É também apresentado um modelo específico e original para a rede Internet, proposto neste trabalho.

6 Modelos de Crescimento Organizado de Redes Complexas

A Internet pode ser vista como sendo baseada nas interconexões de Sistemas Autônomos que apresentam uma aparente natureza aleatória porém, na realidade, são descritas por uma lei de potência, sendo por isto considerada uma rede de topologia do tipo sem escala (*scale-free network*).

Muitos modelos foram desenvolvidos e testados em redes de diversos tipos. De maneira resumida podemos dividir os modelos em dois tipos: os básicos e os específicos. Os modelos básicos são aqueles utilizados para o desenvolvimento e testes de teorias e conceitos necessários para o desenvolvimento dos modelos específicos que incluem ingredientes dinâmicos existentes nas redes a que serão submetidos e

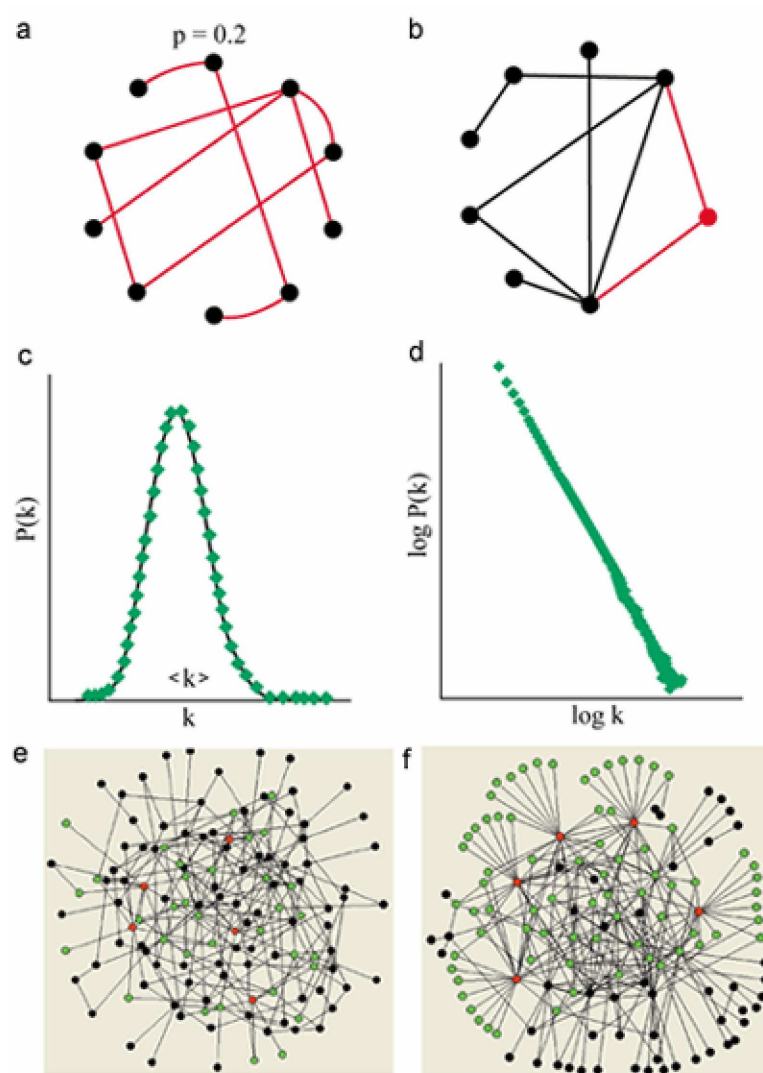


Figura 2: Esquema comparativo entre redes de conexões que possuem a distribuição do coeficiente de conectividade k_i do tipo aleatório ou Poisson e do tipo lei de potência, nas colunas esquerda e direita, respectivamente. Figura extraída de [Bar01].

avaliados.

Nesta seção, são apresentados os principais modelos que caracterizam as redes livres de escala que serviram de base para o desenvolvimento de um modelo mais abrangente e específico para Internet. Os modelos base são os modelos Barabási-Albert e a sua extensão, o modelo Dorogovtsev-Mendes e o modelo Zhou-Mondragón. Estes modelos serviram para o desenvolvimento de um modelo específico, original e aplicado aos dados experimentais considerados neste trabalho.

Nas primeiras sub-seções deste capítulo, 6.1 a 6.3, são apresentados os modelos base. Estes modelos tratam dos seguintes conceitos básicos: crescimento contínuo, conexão preferencial, atração inicial, novas conexões e exclusão de conexões. Na sub-seção 6.4 é abordado um modelo que acrescenta um ingrediente na dinâmica de conexões entre sistemas autônomos da Internet, a probabilidade não linear. Na sub-seção 6.5, é apresentado o modelo proposto neste trabalho, que contém ingredientes dos modelos descritos anteriormente. Posteriormente, em outro trabalho, são apresentados os resultados das modelagens computacionais destes modelos, comparativamente com os resultados dos dados experimentais.

6.1 Modelo Barabási-Albert

Baseado nos dois princípios fundamentais: crescimento contínuo e conexão preferencial, Barabási e Albert [BA99] propuseram o seguinte modelo:

- **Crescimento contínuo:** o modelo começa com um pequeno número de nós sem conexões n_0 e a cada instante é adicionado um novo nó que faz m novas conexões a diferentes nós já presentes na rede.
- **Conexão preferencial:** as conexões iniciadas pelo novo nó são realizadas de acordo com a probabilidade dada pela equação (4) que será repetida aqui na forma

$$P(k_i) = \frac{k_i}{\sum_{j=1}^N k_j} \quad (5)$$

onde $P(k_i)$ e k_i são a probabilidade e o grau de conectividade do i ésimo nó, respectivamente, e N é o número de nós a qualquer instante da evolução da rede.

Neste modelo, a rede resultante a cada instante terá N nós e mt conexões depois de t passos, onde $N = t + n_0$. Através de simulações numéricas, é possível demonstrar que a rede resultante segue uma lei de potência cujo expoente γ é aproximadamente igual a 3, independentemente do valor de m .

Voltando um pouco no tempo, em 1999, Barabási e Albert [BA99] e Albert, Jeong e Barabási [AJB99] calcularam a dependência temporal do grau de conectividade do i ésimo nó k_i utilizando aproximações contínuas. Este grau, também chamado de coeficiente de conectividade k_i , aumenta toda vez que um novo nó entra no sistema e se conecta ao nó i com uma probabilidade dada pela equação (5).

Assumindo que k_i é uma variável real e contínua, a taxa de variação temporal com que muda, deve ser proporcional à probabilidade $P(k_i)$. Lembrando que m é o número inicial de conexões, conseqüentemente, k_i deve satisfazer a equação dinâmica

$$\frac{\partial k_i}{\partial t} = mP(k_i) = m \frac{k_i}{\sum_{j=1}^{N-1} k_j} \quad (6)$$

Observando que agora o somatório do denominador não considera o novo nó e que cada conexão é simétrica e por isso contada duas vezes então, no limite ($t \gg m$), esta soma é dada por

$$\sum_{j=1}^{N-1} k_j = 2(mt - m) \Rightarrow 2mt \quad (7)$$

A simples substituição da equação (7) em (6), leva a

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t} \quad (8)$$

Observando que o i ésimo nó é acrescentado na rede no instante t_i com o número inicial de conexões $k_i = m$, a solução da equação (8) com a condição de contorno inicial $k_i(t_i) = m$ é dada por

$$k_i(t) = m \left(\frac{t}{t_i} \right)^\beta, \quad \beta = 1/2 \quad (9)$$

Observe que a equação (9) atesta que o coeficiente de conectividade de qualquer um dos nós, é uma lei de potência diferenciada pelo valor de t_i .

$$k_i(t) = m \frac{t^\beta}{t_i^\beta} \Rightarrow t_i^\beta = \frac{mt^\beta}{k_i(t)} \Rightarrow t_i = \frac{m^{1/\beta} t}{(k_i(t))^{1/\beta}} \quad (10)$$

Dessa forma, a probabilidade de um nó possuir grau $k_i(t)$ menor que k é dada por

$$P[k_i(t) < k] = P \left[t_i > \frac{m^{1/\beta} t}{k_i^{1/\beta}} \right] \quad (11)$$

Assumindo que o modelo de crescimento considera intervalos de tempo iguais, os valores de t_i possuem uma densidade de probabilidade igual a

$$P(t_i) = \frac{1}{n_0 + t} = \frac{1}{N} \quad (12)$$

Substituindo (12) em (11), obtêm-se

$$P \left[t_i > \frac{m^{1/\beta} t}{k_i^{1/\beta}} \right] = 1 - \frac{m^{1/\beta} t}{k_i^{1/\beta} N} \quad (13)$$

A distribuição do coeficiente de conectividade $P(k)$ é obtida calculando sua derivada parcial em relação a k . Para obter a forma algébrica das equações (14), (15) e (16), o valor de N foi substituído na equação (13) e o valor numérico de $\beta = 1/2$ foi usado quando conveniente.

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{2m^{1/\beta} t}{(n_0 + t)} \frac{1}{k^{1/\beta+1}} \quad (14)$$

Asimptoticamente quando ($t \rightarrow \infty$) em (14), as contribuições de t do numerador e denominador se cancelam e obtêm-se

$$P(k) \approx 2m^{1/\beta} k^{-\gamma} \quad (15)$$

onde o expoente da lei de potência é dado por

$$\gamma = \frac{1}{\beta} + 1 \quad (16)$$

Para $\beta = 1/2$, é imediato chegar ao valor $\gamma = 3$ comentado anteriormente. Este resultado de independência do valor de γ em relação à m obtido analiticamente por Barabási e Albert em [BA02] também foi obtido numericamente através de simulações computacionais do modelo e são apresentadas na seção (4.5.1).

Assim, a equação (15) mostra que a distribuição de probabilidades do coeficiente de conectividade k_i é descrita por uma lei de potência de expoente γ independente do número de conexões iniciais a que todo novo nó está sujeito.

A equação (15) também revela que o modelo Barabási-Albert é independente do tempo e por consequência, independente do tamanho pois $N = n_0 + t$, indicando que apesar do crescimento contínuo, a rede de conexões chega ao seu estado estacionário, que é também livre de escala ou sem escala.

Este processo analítico foi baseado nas premissas do crescimento contínuo e da conexão preferencial descritas na pág. 8. Nos artigos citados, [BA99] e [AJB99], os autores também provam que a liberdade de escala necessariamente depende destas duas condições básicas do modelo. O procedimento é simples e apoiado em uma análise de duas situações diferentes onde somente uma das premissas é considerada de cada vez.

Conexão Preferencial

Nesta primeira análise não é considerada a premissa do crescimento contínuo e portanto não há aumento do número de nós N que permanece constante ao longo do tempo. A cada instante um nó é escolhido aleatoriamente e realiza uma conexão preferencial dada com o nó i com probabilidade $P(k_i)$ dada pela equação (5).

As simulações feitas por Barabási e Albert [BA99] revelaram que no início o modelo apresenta uma lei de potência porém, com o passar das iterações, é alcançado um estado onde todos os nós estão conectados entre si, estado conhecido como *full meshed network*, o que elimina a possibilidade de uma rede livre de escala

Crescimento Contínuo

Nesta análise a premissa do crescimento contínuo é testada com a probabilidade de conexão de um novo nó sendo agora igual para qualquer um dos outros nós já existentes. Se a rede começa com n_0 nós e a cada novo nó são realizadas $m < n_0$ conexões, então a equiprobabilidade é dada por

$$P(k_i) = \frac{1}{n_0 + mt - 1} \quad (17)$$

Neste caso, Barabási e seus colaboradores nos artigos já citados, demonstram que existe uma dependência logarítmica de coeficiente de conectividade $k_i(t)$ com o tempo que no limite $t \rightarrow \infty$ leva a distribuição de probabilidades à seguinte expressão:

$$P(k, t \rightarrow \infty) = \frac{e}{m} \exp\left(-\frac{k}{m}\right) \quad (18)$$

o que também elimina a propriedade de rede livre de escala.

Em cada uma das duas situações consideradas, somente uma das premissas fundamentais está presente e a condição de liberdade de escala não é alcançada, portanto é possível concluir que para o crescimento

de uma rede livre de escala são necessários simultaneamente os mecanismos de crescimento contínuo e conexão preferencial dada pela equação (5).

Resumidamente, é possível dizer que este modelo reproduz as distribuições de probabilidades estacionárias livres de escala observadas, por exemplo, ou em redes genéticas ou mesmo no ambiente Web. Isto é um indicativo de que o desenvolvimento de grandes redes de conexões é governado por um fenômeno robusto de auto-organização que predomina sobre os princípios dos sistemas individuais, participantes da rede.

6.2 Modelo Barabási-Albert Estendido

Barabási e Albert observaram que algumas outras características deveriam ser acrescentadas ao seu modelo básico principalmente porque em algumas redes de conexões, o valor experimental do expoente γ diferia daquele obtido na equação (16).

A primeira mudança no modelo básico está no conceito de atração inicial, que é uma necessidade diante dos demais mecanismos acrescentados ao modelo, novas conexões e rearranjo, e que serão descritos na seqüência [BAJ00].

De acordo com a equação (5), se um nó tiver zero ligações, então a probabilidade deste nó receber uma nova ligação também será zero e isto na maioria das redes reais não é a realidade, pois sempre existe alguma possibilidade de, por exemplo na rede de citações, um artigo que ainda não foi citado, receber citações.

Para solucionar esta questão pode-se adicionar uma constante A , usualmente $A = 1$, garantindo a possibilidade do nó i ter chance de receber novas ligações e assim a equação (5) fica:

$$P(k_i) = \frac{k_i + A}{\sum_j (k_j + A)} \quad (19)$$

Esta equação permite que o modelo considere a atração inicial que é intrínseca a cada elemento da rede independente de estar já conectado ou ainda isolado.

Uma outra questão é que em muitas redes reais o expoente γ é diferente de 3, que é o valor esperado no modelo original, sugerindo a existência de mecanismos não considerados até então. Esses mecanismos são os chamados eventos locais, tais como: adição ou remoção de conexões entre os nós já existentes, exclusão de nós e rearranjo de conexões. Para tornar seu modelo mais realista, Barabási e Albert propuseram uma extensão ao seu modelo básico. Começar com n_0 nós isolados e a cada instante de tempo uma das três operações a seguir é realizada:

- Com probabilidade p são adicionadas m novas conexões. Um dos nós da nova conexão é escolhido aleatoriamente e outro com probabilidade descrita pela equação (19). Este procedimento é repetido m vezes.
- Com probabilidade q são rearranjadas m conexões. Para isso, é selecionado um nó aleatoriamente e é removida uma de suas conexões que é substituída por outra com um nó escolhido com probabilidade descrita pela equação (19). Este procedimento também é repetido m vezes.
- Com probabilidade $r = (1 - p - q)$ é adicionado um novo nó. Este novo nó realiza m conexões com probabilidade descrita pela equação (19).

Nesta extensão do modelo básico, as probabilidades p e q de novas conexões e de rearranjos entre as conexões existentes, podem variar no intervalo $0 \leq p \leq 1$ e $0 \leq q \leq 1$, respectivamente, porém,

satisfazendo a $p + q + r = 1$, onde r é a probabilidade de somente adicionar um novo nó. No caso de ambas as probabilidades serem nulas, o modelo estendido reduz-se ao modelo básico.

Barabási e colaboradores nos artigos [BA99] e [BA02] também utilizaram as mesmas aproximações contínuas introduzidas no modelo básico, para descrever a taxa de variação temporal do coeficiente de conectividade k_i do i -ésimo nó. A contribuição de cada novo mecanismo é descrita a seguir.

Novas Conexões

Nesta operação de nova conexão entre os nós existentes, um dos nós é escolhido aleatoriamente, portanto, com probabilidade $1/N$. O outro nó da nova conexão é selecionado preferencialmente com probabilidade dada pela equação (19), com $A = 1$. Esta operação é repetida m vezes com probabilidade p . A taxa de variação do coeficiente de conectividade nesta operação de adição de conexões preferenciais entre os nós já existentes, é dada por:

$$\left(\frac{\partial k_i}{\partial t}\right)_a = mp \frac{1}{N} + mp \frac{k_i + 1}{\sum_j (k_j + 1)} \quad (20)$$

Rearranjo de Conexões

O mesmo raciocínio empregado na operação anterior é utilizado no caso da operação de rearranjo de conexão. O primeiro termo da equação (21) é negativo porque nesta operação ocorre primeiramente a remoção com probabilidade q de uma conexão do nó sorteado aleatoriamente, portanto um decréscimo no coeficiente de conectividade. O segundo termo continua positivo porque representa o acréscimo no coeficiente de conectividade com probabilidade q quando a conexão é feita com um novo nó, agora com probabilidade preferencial.

$$\left(\frac{\partial k_i}{\partial t}\right)_b = -mq \frac{1}{N} + mq \frac{k_i + 1}{\sum_j (k_j + 1)} \quad (21)$$

Adição de um novo nó

Esta é a operação definida no modelo básico, agora realizada com probabilidade $r = 1 - p - q$. Neste caso ocorre somente o acréscimo no coeficiente de conectividade do nó existente que irá receber a conexão. De novo a operação é realizada m vezes agora com probabilidade r . A taxa de variação é dada por:

$$\left(\frac{\partial k_i}{\partial t}\right)_c = mr \frac{k_i + 1}{\sum_j (k_j + 1)} \quad (22)$$

O valor total da taxa de variação do coeficiente de conectividade do i -ésimo nó, k_i , é dada pela soma das equações (20), (21) e (22) relativas a cada uma das três operações descritas.

$$\frac{\partial k_i}{\partial t} = \left(\frac{\partial k_i}{\partial t}\right)_a + \left(\frac{\partial k_i}{\partial t}\right)_b + \left(\frac{\partial k_i}{\partial t}\right)_c \quad (23)$$

$$\frac{\partial k_i}{\partial t} = m(p - q) \frac{1}{N} + m \frac{k_i + 1}{\sum_j (k_j + 1)} \quad (24)$$

Considerando $N = n_0 + (1 - p - q)t$, $\sum_j (k_j) = 2mt(1 - q) - m$ e a condição de conectividade inicial

do i ésimo nó $k_i(t) = m$, no limite quando $t \rightarrow \infty$, a solução da equação (24) é dada por

$$k_i(t) = [A(p, q, m) + m + 1] \left(\frac{t}{t_i} \right)^{\frac{1}{B(p, q, m)}} - A(p, q, m) - 1 \quad (25)$$

onde

$$A(p, q, m) = (p - q) \left(\frac{2m(1 - q)}{1 - p - q} + 1 \right) \quad (26)$$

$$B(p, q, m) = \frac{2m(1 - q) + 1 - p - q}{m} \quad (27)$$

Finalmente, o modelo Barabási-Albert estendido leva a uma distribuição de probabilidade do coeficiente de conectividade k_i tendo uma lei de potência generalizada com a seguinte forma:

$$P(k) \propto [k + A(p, q, m) + 1]^{-\gamma} \quad (28)$$

onde os parâmetros $A(p, q, m)$ e γ são definidos pelas equações (26) e (29), respectivamente.

$$\gamma = \frac{2m(1 - q) + 1 - p - q}{m} + 1 \quad (29)$$

A equação (28) somente é válida para $A(p, q, m) + m + 1 > 0$ e a estrutura livre de escala só é observada para $q < q_{max}$, onde

$$q_{max} = \min \left\{ (1 - p), \frac{(1 - p + m)}{(1 + 2m)} \right\} \quad (30)$$

De maneira resumida, o modelo Barabási-Albert estendido inclui os mecanismos intrínseco de atração inicial e probabilísticos de novas conexões e de rearranjos entre nós já existentes. A distribuição de probabilidades de conexão é descrita por uma lei de potência dada pela equação (28).

6.3 Modelo Dorogovtsev-Mendes

Partindo do modelo básico Barabási-Albert, Dorogovtsev e Mendes [DM00] propuseram dois novos mecanismos de crescimento: i. desenvolvimento de redes e ii. estrutura de decaimento.

O mecanismo “desenvolvimento de redes” considera a possibilidade do surgimento de novas conexões entre os nós já existentes. O mecanismo “estrutura de decaimento”, ao contrário, admite a exclusão de conexões entre os nós já existentes. Estes mecanismos levam à seguinte expressão para o expoente γ :

$$\gamma = 2 + \frac{1}{1 + 2C} \quad (31)$$

onde C é o número de conexões incluídas (> 0) ou removidas (< 0). Quando $C = 0$ obtemos o valor esperado do modelo original de Barabási-Albert. Em sua forma original o modelo Dorogovtsev-Mendes considera um ou outro mecanismo.

6.4 Modelo Zhou-Mondragón

Estes modelos apresentados nas seções anteriores, o modelo Barabási-Albert, sua extensão e o modelo Dorogovtsev-Mendes, são modelos gerais utilizados no desenvolvimento de teorias e modelagem de redes complexas de uma maneira geral.

São muito úteis no entendimento dos mecanismos existentes nos estudos relacionados com o crescimento das redes complexas em geral. Certamente servem de base para o desenvolvimento de modelos

mais realistas e aplicáveis a redes complexas específicas.

Para a rede de conexões da Internet, Zhou e Mondragón [ZM03][ZM04] introduziram o conceito de conexão preferencial não-linear onde a probabilidade de conexão agora apresenta um expoente que tem a principal característica de amplificar o efeito de conectividade preferencial. A probabilidade de conexão de cada nó agora é dada por

$$\prod(k_i) = \frac{k_i^\alpha}{\sum_j k_j^\alpha} \quad (32)$$

O que está por trás desta probabilidade, que passaremos a chamar probabilidade alfa ou probabilidade não-linear, é a possibilidade de aumentar a importância de grandes *hubs*⁵ para a rede toda.

Além deste tipo não-linear de probabilidade, este modelo considera também a possibilidade do surgimento de novas conexões entre nós já existentes.

O modelo Zhou-Mondragón, apesar de considerar a rede de conexões entre ASs, não foi submetido a dados retirados da tabela *full routing* BGP. Os autores utilizaram dados obtidos através da utilização da ferramenta computacional *traceroute*. Esta ferramenta basicamente é um programa que determina a rota por onde passam os pacotes de informação em uma rede de computadores.

6.5 Modelo Proposto

Havia nesta fase do trabalho, as alternativas de desenvolvimento de um modelo baseado em uma estatística não-extensiva, como aquele feito por Soares et al. em 2005 [STMdS05], que em outros sistemas também em escala livre apresentou ótima concordância com os dados experimentais, ou um modelo baseado na dinâmica das redes (Chen e Shi, 2004) [CS04] ou ainda aquele fenomenológico proposto por Oliveira (2005) [Oli05].

Porém, o modelo Zhou-Mondragón chamou atenção pelo fato de estar bem próximo dos modelos já estudados com simulações e portanto não havendo necessidade de introduzir conceitos oriundos de outras teorias e também por apresentar facilidade de implementação.

Tomando como base os modelos apresentados nas seções anteriores, neste trabalho é proposto um modelo original que considera alguns dos mecanismos mencionados. O modelo proposto considera o crescimento contínuo, a probabilidade alfa de conexão dada pela equação (32), pág. 14, e as possibilidades de inclusão e de rearranjo de novas conexões entre os nós já existentes.

Os dois princípios básicos, crescimento contínuo e conexão preferencial, são mecanismos de longo alcance, genéricos e de interferência nas conexões da rede de modo coletivo, como um sistema amplo. Existem outros mecanismos que caracterizam eventos locais e que contribuem de maneira importante nas propriedades coletivas. Dois novos mecanismos locais não específicos da rede são aqueles que representam a possibilidade de adição de novas conexões entre os nós já existentes e de remoção de conexões já existentes.

A remoção simples de conexões no sistema Internet é muito raro, representando aqueles poucos casos em que o nó, sistema autônomo - AS no caso da Internet, deixa de existir sendo normalmente absorvido por um outro, caso de fusão entre dois ou mais *Internet Services Providers* - ISPs. É bem mais comum a operação chamada de rearranjo que é a remoção seguida de uma nova conexão, representando a troca de vizinhança devido principalmente ao custo/benefício.

A simples inclusão do expoente na probabilidade alfa de conexão faz com que aqueles nós, aqueles ASs, considerados *hubs* passem a ser ainda mais preferidos nas conexões sorteadas se compararmos com

⁵Neste trabalho, *hub* significa um nó que possui alto grau de conectividade k_i , não havendo um limite inferior deste valor para que o nó seja assim considerado.

a expressão inicial.

O modelo aqui proposto considera a inclusão de um novo AS, a unidade básica do ciclo temporal, a cada iteração na implementação. Considera também, com probabilidade p e q , as possibilidades de novas conexões e de rearranjo entre ASs já existentes, respectivamente.

Assim, a cada ciclo ou iteração, além da inclusão do AS e suas m novas conexões, é sorteada uma probabilidade, $0 < prob < 1$. Se $prob$ for menor ou igual a p , o modelo adiciona uma nova conexão entre ASs introduzidos em um ciclo anterior. Se a probabilidade estiver entre p e $p + q$ inclusive, uma operação de rearranjo é efetuada. E por fim, se a probabilidade for maior que $p + q$, nada ocorre além da inclusão deste novo AS e suas conexões.

Os parâmetros p e q tem seus valores ajustados nas simulações de cada ano para que a rede de conexões gerada seja compatível com os números de nós e de conexões, obtidos dos arquivos de dados experimentais.

Este modelo proposto está baseado em três funções principais descritas a seguir.

SortProbAlfa: nesta função primeiramente é construído um vetor va de números reais onde cada elemento representa a probabilidade alfa dada pela equação (32) e posteriormente é sorteado um número entre 0 e 1 que é procurado em cada um dos elementos de va definindo assim o nó sorteado.

IncluiConexão: esta função utiliza a função *SortProbAlfa* para escolher um nó de acordo com a probabilidade alfa.

ExcluiConexão: esta função sorteia uma das conexões do nó previamente escolhido e a exclui verificando se existem outras, evitando assim o isolamento do nó.

Com estas três funções principais e outras mais, a implementação do modelo proposto de crescimento da rede complexa Internet permite a utilização dos diversos mecanismos considerados na dinâmica da evolução temporal. A seguir, uma breve descrição da implementação em cinco etapas.

1. Definir os valores do número de nós total N e inicial n_0 , do número inicial de conexões m , das propabilidades de novas conexões p e de rearranjo q entre os nós já existentes.
2. Iniciar a rede com dois nós conectados, que é a condição básica de uma rede.
3. A cada iteração, incluir um novo nó com m conexões definidas pela probabilidade alfa dada pela equação (32).
4. A cada iteração, também sortear uma probabilidade $prob$ e compará-la com as probabilidades das dinâmicas de novas conexões ou de rearranjos entre nós já existentes, de acordo com:
 - $prob \leq p$ realizar uma nova conexão;
 - $p < prob \leq (p + q)$ realizar rearranjo;
 - $(p + q) < prob$ nada a fazer.
5. Repetir o processo até que o número de nós seja N .

A implementação deste algoritmo deve ser otimizada, visto que o número de iterações é bastante grande no caso da Internet. Além disto, a quantidade de memória deve ser apropriado pois o tamanho global da grande rede está atualmente perto de 257.000 redes distribuídas em cerca de 22.000 sistemas autônomos

7 Conclusão

Neste breve artigo, os conceitos básicos do tema Redes Complexas foram abordados, assim como os modelos de crescimento da rede do tipo sem escala. Foram comentados os principais modelos básicos e específicos para Internet e também um outro proposto pelo autor.

Em um trabalho posterior, são definidos os objetos de estudo da Internet, as fontes de dados experimentais e o pré-tratamento dos dados. Também são apresentadas as primeiras análises dos dados obtidos em janeiro de 2004 e as modelagens computacionais dos modelos básicos submetidas a comparações com estas análises iniciais.

Além disto, os resultados das simulações do modelo aqui proposto são comparados aos valores experimentais na análise da evolução da Internet sobre os aspectos da topologia e do parâmetro menor caminho médio ao longo da década de 1998 à 2007.

Referências

- [Ada07] L. A. Adamic. Zipf, power-laws, and pareto - a ranking tutorial. <http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html>, acessado em março de 2007. Information Dynamics Lab, HP Labs, Palo Alto, CA 94304.
- [AJB99] R. Albert, H. Jeong, and A. L. Barabási. Diameter of the world-wide web. *Nature*, 401:130–131, 1999.
- [AMJR02] R. Alberich, J. Miro-Julia, and F. Rossello. Marvel universe looks almost like a real social network. *Cornell University Library arXiv.org*, cond-mat/0202174, 11feb 2002.
- [ASBS00] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *Proc. Natl. Acad. Sci. USA*, 97:11.149–11.152, 2000.
- [BA99] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [BA02] A. L. Barabási and R. Albert. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002.
- [bAH00] D. ben Avraham and S. Havlin. Diffusion and reactions in fractals and disordered systems. *Cambridge University Press*, 2000.
- [BAJ00] A.-L. Barabási, R. Albert, and H. Jeong. Scale-free characteristics of random networks: The topology of the world wide web. *Physica A*, 281:69–77, 2000.
- [Bar01] A. L. Barabási. The physics of the web. *Physics World*, pages 33–38, jul. 2001.
- [BKM⁺00] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. *Computer Networks*, 33:309–320, 2000.
- [Bol01] B. Bollobás. Random graphs. *Cambridge University Press, 2nd Edition*, 2001.
- [BTW87] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: an explanation of the $1/f$ noise. *Physical Review Letter*, 59:381–384, 1987.
- [BTW88] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Physical Review A*, 38:364–374, 1988.

- [CHB99] R. C. Connor, M. R. Heithaus, and L. M. Barre. Superalliance of bottlenose dolphins. *Nature*, 397:571–572, 1999.
- [CHE02] G. Chowell, J. M. Hyman, and S. Eubank. Analysis of a real world network: The city of portland. *Department of Biological Statistics and Computational Biology*, Technical Report Cornell University:BU-1604-M, 2002.
- [Coh88] J. E. Cohen. *Discrete Appi. Math.*, 19:113, 1988.
- [CS04] Q. Chen and D. Shi. The modeling of scale-free networks. *Physica A*, 335:240–248, 2004.
- [DM00] S. N. Dorogovtsev and J. F. F. Mendes. Scaling behavior of developing and decaying networks. *Europhysics Letters*, 52:33-39, 2000.
- [ER59] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae (Debrecem)*, 6:290–297, 1959.
- [ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- [ER61] P. Erdős and A. Rényi. *Bull. Inst. Int. Stat.*, 38:343, 1961.
- [ER90] L. Egghe and R. Rousseau. Introduction to informetrics. Elsevier, Amsterdam, 1990.
- [Eul53] L. Euler. Dissertatio de principio mininiaie actionis, ' una cum examine objectionum cl. prof. koenigii. *ibid.*, 1753. in octavo.
- [JTA⁺00] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási. The large-scale organization of metabolic networks. *Nature*, 407:651–654, 2000.
- [Kle00] J. M. Kleinberg. The small-world phenomenon: An algorithmic perspective. *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, Association of Computing Machinery, New York*, pages 163–170, 2000.
- [Lot26] A. J. Lotka. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12):317–323, June 1926.
- [LTWW94] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1994.
- [Mar75] P. Mariolis. Interlocking directorates and control of corporations: The theory of bank control. *Social Science Quarterly*, 56:425–439, 1975.
- [Mil67] S. Milgram. The small world problem. *Sychology Today*, P2:60–67, 1967.
- [Mit03] M. Mitzenmacher. A brief history of generative models for power law and longnormal distributions. *Internet Mathematics*, 1(2):226–251, 2003.
- [Miz82] M. S. Mizruchi. The american corporate network, 1904-1974. *SAGE Library of Social Research - SAGE Publications Inc.*, 1982. ISBN-13: 978-0803917781.
- [Mor34] J. L. Moreno. Who shall survive? *Beacon House, NY*, 1934.

- [NBW06] M. Newman, A. L. Barabási, and D. J. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, 2006. ISBN-13: 978-0-691-11357-9.
- [Oli05] J. C. C. Oliveira. Simulação computacional de redes sem escala: Uma aplicação à modelagem relacional da internet. Master's thesis, Instituto Politécnico do Estado do Rio de Janeiro - IPRJ/UERJ, Nova Friburgo, RJ, março 2005.
- [PA93] J. F. Padgett and C. K. Ansell. Robust action and the rise of the medici, 1400-1434. *Am. J. Sociol.*, 98:1.259–1.319, 1993.
- [Red98] S. Redner. How popular is your paper? an empirical study of the citation distribution. *Eur. Phys. J. B*, 4:131–134, 1998.
- [RH61] A. Rapoport and W. J. Horvath. A study of a large sociogram. *Behavioral Science*, 6:279–291, 1961.
- [Sco00] J. Scott. *Social network analysis: A handbook*. Sage Publications, London, 2000.
- [SDC⁺02] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna. Small-world properties of the indian railway network. *Preprint arXiv*, cond-mat/0208535, 2002.
- [Seg92] P. O. Seglen. The skewness of science. *J. Amer. Soc. Inform. Sci.*, 43:628–638, 1992.
- [SFFF03] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos. Power-laws and the as-level internet topology. *IEEE/ACM Transactions on Networking (TON)*, 2003.
- [SKB⁺02] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420:190–193, 2002.
- [STMdS05] D. J. B. Soares, C. Tsallis, A. M. Mariz, and L. R. da Silva. Preferential attachment growth model and nonextensive statistical mechanics. *Europhysics Letters*, 1:70–76, 2005.
- [TdA00] C. Tsallis and M. P. de Albuquerque. Are citations of scientific papers a case of nonextensivity. *European Physical Journal B, Società Italiana di Fisica SV*, 13:777–780, 2000.
- [TM69] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32:425–443, 1969.
- [Wat07] D. J. Watts. Small world project. *Columbia University*, acessado em fevereiro de 2007. <http://smallworld.columbia.edu/index.html>.
- [WF94] S. Wasserman and K. Faust. *Social network analysis*. Cambridge University Press, 1994.
- [WF01] A. Wagner and D. Fell. The small world inside large metabolic networks. *Proc. Royal Society London B*, 268(1478):1803–1810, 7 september 2001.
- [ZM03] S. Zhou and R. J. Mondragón. In J. Charzinski, editor, *Proceedings of the 18th International Teletraffic Congress (ITC18)*, volume 5a of *Teletraffic Science and Engineering*, pages 121–130. Elsevier (Berlin), 2003.
- [ZM04] S. Zhou and R. J. Mondragón. Accurately modeling the internet topology. *Physical Review E*, 70:066108–1–8, 2004.