

Elizabete Koehler<sup>1</sup>  
Reginaldo Aparecido Zara<sup>2</sup>

**ALGORITMO PARA CÁLCULO DE DISTÂNCIAS  
MÍNIMAS EM REDES DE ESCALA LIVRE**

RESUMO: Muitos sistemas naturais e tecnológicos podem ser representados matematicamente por uma classe de grafos chamada de redes de escala livre. Exemplos destes sistemas são a Internet e a *World-Wide Web*. Nestas redes, a distância média entre vértices é uma grandeza de fundamental importância em processos de transmissão de dados ou recuperação de informação. Neste trabalho aplicam-se métodos computacionais para estimar as distâncias entre vértices da rede de escala livre resultante do modelo de crescimento de Barabási-Albert. Os resultados mostram que, independentemente da conectividade, as distâncias mínimas são funções logarítmicas crescentes do número de vértices da rede.

PALAVRAS-CHAVE: distribuições de conectividades; grafos; menores caminhos.

ABSTRACT: A variety of natural and technological systems may be mathematically described as a specific kind of graphs classified as scale free networks. The Internet and the World-Wide Web are important examples of systems which exhibit scale free characteristics. In these systems the length of the minimum path between vertices is a fundamental quantity to data recovery and information transmission processes. In this work we applied computational techniques to investigate the length of minimum paths in networks resulting from the growth model proposed by Barabási and Albert. Our results show that, independent of the mean connectivity, the minimum paths are logarithmic functions of the number of vertices in the network.

KEYWORDS: degree distributions; graphs; minimum paths.

Data de recebimento: 03/10/03. Data de aceite para publicação: 26/01/04 .

<sup>1</sup> Bolsista de Iniciação Científica. Programa PIBIC/CNPq/Unioeste. Centro de Ciências Exatas e Tecnológicas. Unioeste – Universidade Estadual do Oeste do Paraná. Endereço eletrônico: elizabetekoehler@hotmail.com.

<sup>2</sup> Professor Adjunto do Centro de Ciências Exatas e Tecnológicas. Unioeste – Universidade Estadual do Oeste do Paraná. Endereço eletrônico: razara@unioeste.br.

## 1. INTRODUÇÃO

Muitos sistemas físicos, químicos, biológicos e sociais podem ser interpretados como grafos onde os indivíduos são representados por vértices e suas interações por arestas (cf. Albert; Barabási, 2002). Tradicionalmente, os grafos são investigados usando a teoria de grafos aleatórios desenvolvida por Erdős e Renyi (ER) (cf. Gersting, 1995). Os resultados de ER levam a distribuições de conectividade ( $P(k)$ ) (que fornecem a probabilidade de que um vértice da rede esteja conectado a  $k$  outros vértices) que obedecem a distribuições de Poisson, o que significa que, sob o ponto de vista de conectividade, todos os vértices do grafo são praticamente equivalentes.

Porém, a análise de dados empíricos extraídos de uma série de redes revela que as distribuições de conectividade seguem de leis de potências (cf. Albert; Barabási, 2002), o que implica na abundância de vértices com poucas ligações e uma pequena, porém significativa minoria, que possui uma grande quantidade de ligações. Devido à ausência de uma escala característica este tipo de grafo é chamado de rede de escala livre.

Estudos recentes têm revelado que as redes de escala livre exibem estrutura topológica rica, com profundas implicações no desenvolvimento de estratégias de segurança e estabilidade das redes de comunicação (cf. Broder *et al*, 2000). Uma quantidade importante é o menor caminho (ou distância mínima) entre dois vértices escolhidos ao acaso,  $D$ , definido como o menor número de conexões que devem ser seguidas para navegar de um vértice para outro. Testes experimentais encontraram que as distâncias mínimas são funções logarítmicas do número de vértices (cf. Albert; Barabási, 2002).

Embora o comportamento das distâncias mínimas esteja empiricamente estabelecido, sob o ponto de vista de modelagem ainda existem pontos a ser explorados. Kim (*et al*, 2002) tem argumentado que o comportamento matemático das distâncias mínimas depende da estratégia de busca utilizada. Neste trabalho adapta-se o algoritmo das queimadas (cf. Herrmann *et al*, 1984), proposto inicialmente para o estudo de sistemas espacialmente desordenados, como estratégia para a determinação das distâncias mínimas entre vértices de uma rede de escala livre e analisa-se como esta distância depende do número de vértices e da conectividade da rede.

## 2. METODOLOGIA E RESULTADOS

A maioria dos modelos de redes evolutivas é baseada em dois ingredientes: crescimento e vinculação preferencial (cf. Albert; Barabási, 2002). Implementamos um código computacional que simula o modelo de crescimento de redes de escala livre formulado por Barabási-Albert (BA) (cf. Albert, 2002). Neste modelo, o crescimento da rede começa com grafo aleatório de  $m_0$  nós. A hipótese de crescimento é incorporada admitindo-se que em cada unidade de tempo, um novo vértice seja adicionado ao grafo, conectando-se a  $m$  diferentes vértices e seguindo uma hipótese de vinculação preferencial. Em cada instante de tempo cada vértice possui uma probabilidade  $p$  de adquirir uma nova conexão.

No modelo BA de  $N$  vértices, a probabilidade de que o novo vértice introduzido se vincule ao vértice  $i$  é expressa como a razão entre o número atual de conexões do vértice  $i$  e o número total de conexões na rede, expressando a vinculação preferencial aos vértices de maior conectividade. Ao final do processo de crescimento de uma rede de  $N$  vértices, a conectividade média é  $2m$  (cf. Albert; Barabási, 2002).

Uma vez construída a rede com o valor de conectividade escolhido *a priori*, estima-se a menor distância entre todos os possíveis pares de vértices e efetua-se a média sobre os pares. Repete-se o procedimento para redes de diferentes tamanhos, para investigar as variações das distâncias mínimas como função do número de vértices da rede.

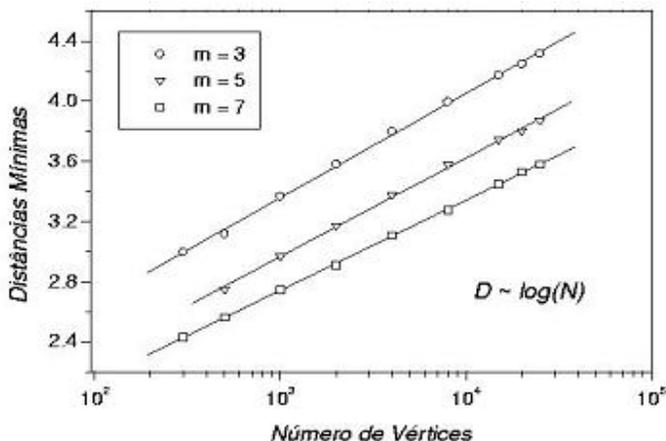
Para estimar as distâncias mínimas entre vértices, adaptamos o algoritmo da queimada (cf. Herrmann *et al*, 1984), utilizando as idéias do algoritmo de busca de extensão por fronteira (cf. Cormen *et al*, 2002). A idéia central do algoritmo está na atribuição de rótulos aos vértices indicando suas distâncias em relação a um vértice de origem escolhido de antemão. Considere um grafo formado por  $N$  vértices. Inicialmente selecionam-se, ao acaso, dois vértices como origem e destino, respectivamente, e inicia-se um processo de busca a partir da origem. Os vértices diretamente conectados com a origem recebem um rótulo de índice 1, indicando que estão à distância de uma aresta em relação à origem. Em uma segunda etapa, todos os vizinhos dos vértices de índice 1 são rotulados como o índice 2. Em seguida os vizinhos do vértice de rótulo 2 são consultados e aqueles não rotulados recebem o índice 3

e, assim, sucessivamente, até que cada vértice receba um rótulo correspondente à sua distância em relação ao vértice de origem. Este procedimento é interrompido quando o vértice de destino for encontrado. Desta forma, o rótulo do vértice de destino corresponde à sua distância em relação à origem. Assim como no algoritmo das queimadas, não é necessário que todos os vértices do grafo sejam consultados para que a distância entre os dois pontos seja encontrada, o que contribui para que grafos de milhares de vértices possam ser investigados.

Se  $N_p$  é o número de pares possíveis, a distância mínima entre vértices da rede ( $D(N)$ ) deve ser calculada como a média dos menores caminhos  $d_{ij}$  entre todos os  $N_p$  pares, isto é

$$D(N) = \frac{1}{N_p} \sum_{i,j} d_{ij}$$

Em um grafo conectado formado por  $N$  vértices são possíveis  $N_p = N(N-1)$  pares. Como se trata de sistemas formados por milhares de vértices, são milhões de pares possíveis e, com isso, nas maiores redes, a análise de todos os pares torna-se computacionalmente proibitiva. Para contornar esta dificuldade, utiliza-se, como espaço amostral, um subconjunto de  $M$  pares ( $M \leq N_p$ ), escolhidos aleatoriamente e toma-se a média entre os  $M$  pares escolhidos.



**Figura 1** - Distâncias mínimas contra o número de vértices para valores de conectividade escolhidos arbitrariamente. Observa-se que  $D(N) \propto \log(N)$  concordando com as estimativas experimentais.

Na Figura 1 tem-se o gráfico das distâncias mínimas entre dois vértices da rede de escala livre como função do tamanho da rede para diferentes valores de conectividade. Para valores fixos de  $N$  verificamos que a distância entre vértices diminui quando a conectividade da rede aumenta. Esta observação concorda com a expectativa de que, quanto maior o número de caminhos entre dois vértices, maior será a probabilidade de encontrar um caminho mais curto.

Observamos ainda que a estratégia de busca leva à conclusão de que o comprimento do menor caminho entre dois vértices é função logarítmica do número de vértices, independentemente de sua conectividade, em concordância com as observações experimentais efetuadas na rede formada pela *World-Wide Web* (cf. Broder *et al*, 2000 e Albert; Barabási, 2002 ). Sob o ponto de vista de navegação na *Web*, esta observação significa que apenas algumas conexões separam um documento (ou página da *Web*) de qualquer outro, sugerindo que a navegação entre dois documentos pode ser feita passando através de poucas conexões, e que o aumento do número de documentos e/ou da conectividade da rede interfere pouco nas distâncias entre os diferentes documentos.

Os resultados também sugerem que arquivos podem ser espalhados por toda a rede de computadores de uma maneira muito eficiente. Se tais arquivos contiverem vírus, temos uma “epidemia” (cf. Lloyd; May, 2001), caracterizada por uma alta taxa de infecção, implicando no rápido espalhamento da infecção.

### 3. CONSIDERAÇÕES FINAIS

Neste trabalho adapta-se o algoritmo das queimadas utilizando idéias de busca em grafos para estimar as distâncias mínimas entre vértices das redes de escala livre resultantes do modelo de Barabási-Albert. Investiga-se a dependência entre as distâncias mínimas entre vértices da rede e o número de vértices e a conectividade média.

Verifica-se que, em concordância com as observações experimentais, as distâncias mínimas são funções logarítmicas do número de vértices do sistema independentemente da conectividade média da rede. Esta observação pode ser útil no desenvolvimento de estratégias de busca e recuperação de informações e na aplicação de métodos de proteção da *World-Wide Web* e da Internet contra pragas virtuais, como vírus e *worms*.

#### 4. REFERÊNCIAS

- ALBERT, R.; BARABÁSI, A-L. “Statistical Mechanics of Complex Networks”. *Rev. Mod. Phys.* 74, p. 47, 2002.
- BRODER, A.; KUMAR, R. [et al.]. “Graph structure in the Web”. *Computer Networks.* 33, p. 309, 2000.
- CORMEN, T. H.; LEISERSON, C. E. [et al.]. *Algoritmos: Teoria e Prática.* 2. ed. Rio de Janeiro: Campus, 2002.
- GERSTING, J. L. *Fundamentos matemáticos para a ciência da computação.* 3. Ed. Rio de Janeiro: LTC, 1995.
- HERRMANN, H. J.; HONG, D. C.; STANLEY, H. E. “Backbone and elastic backbone obtained by the new method of ‘burning’”. *J. Phys A. Math.Gen.*, 17, p. L261, 1984.
- KIM, B. J.; YOON, C. N. [et al.]. “Path finding strategies in scale-free networks”. *Phys. Rev. E*, 65, 027103, 2002.
- LLOYD, A. L.; MAY, R. M. “How viruses spread among computers and people”. *Science* 292, p.1316, 2001.

