

# INTRODUÇÃO ÀS REDES COMPLEXAS

Por Aline D. Bessa, Leonardo B. L. Santos\*,  
Lorena P. N. R. Martinez, Mariana C. Costa e Pedro G. S. Cardoso  
Alunos do grupo de Física Estatística e Sistemas Complexos (FESC)  
da Universidade Federal da Bahia

23 de Dezembro de 2010

## 1 Apresentação

Ao se conhecer a Teoria de Redes Complexas (TRC) passamos a notar sua extensa aplicabilidade. As estradas que conectam cidades, as relações de amizade entre pessoas, as interligações entre páginas da *internet*, esses e muitos outros são exemplos do cotidiano em que podemos enxergar a idéia básica que norteia a construção de uma rede complexa.

O conjunto de áreas do conhecimento que se utilizam da Teoria de Redes Complexas é grande e heterogêneo. Há pesquisas envolvendo redes complexas em um leque de campos que vai de Artes a Zoologia, passando por Lingüística e Psicologia. Além de ser particularmente interessante como ciência pura, o conhecimento sobre Redes Complexas proporciona uma excelente metodologia às questões aplicadas.

O curioso é que, apesar de haver uma enorme contribuição dos pesquisadores brasileiros em aplicações de redes complexas, pouca atenção é destinada à teoria, gerando uma lacuna de publicações neste sentido. Diante deste panorama, o presente trabalho se propõe a apresentar a TRC não apenas em nível de divulgação mas também como uma introdução aos leitores interessados em geral.

Este conteúdo sob este formato será disponibilizado apenas pela Internet, para download livre (do arquivo no formato pdf e LaTeX)<sup>1</sup>.

Com votos de proveitosa leitura;

Os autores.

---

\*e-mail: santoslbl@gmail.com

<sup>1</sup>Disponível no repositório 4Shared.

## 2 Conceitos Preliminares

Nesta seção, apresentaremos alguns fundamentos das Teorias de Grafos, Física Estatística e Sistemas Complexos.

### 2.1 Introdução à Teoria dos Grafos

Um **grafo** é um par  $G = (V, E)$  de conjuntos tal que os elementos de  $V$  são seus **vértices** e os elementos de  $E$ , suas **arestas**. Neste livro, consideraremos apenas grafos para os quais a cardinalidade de  $V$  e a de  $E$  são finitas.

Grafos costumam ser representados visualmente da seguinte forma: cada vértice é indicado por um ponto e cada aresta é indicada por uma linha conectando dois pontos.

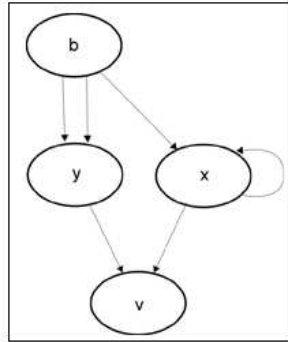


Figura 1: Exemplo de um grafo

Um vértice  $v$  é **incidente** a uma aresta  $e$  se  $v \in e$ . Dois vértices são ditos **adjacentes** se eles são incidentes a uma mesma aresta. Para cada vértice  $i \in V$  é possível, portanto, determinar o conjunto  $A_i \subseteq V$  de seus vértices adjacentes de acordo com a seguinte equação:

$$A_i = \{j \in V | \{i, j\} \in E\} \quad (1)$$

Existem algumas definições simples que são muito importantes para a Teoria dos Grafos. Enumeraremos algumas delas:

- Um grafo é dito **ponderado** quando se associa um valor (normalmente, um número real) a cada uma de suas arestas. Este valor é denominado comumente de **peso da aresta**. O grafo representado na Figura 1, por exemplo, não é ponderado: não há nenhum valor associado a suas arestas<sup>2</sup>.
- Um **laço** é uma aresta que conecta um vértice a si mesmo. Na figura acima, há um laço envolvendo o vértice 11.

<sup>2</sup>Em determinados casos, é conveniente assumir que todas as arestas possuem peso igual a um ou zero.

- Um grafo possui **arestas múltiplas** quando há pelo menos duas arestas incidentes a um mesmo par de vértices. Neste caso, o grafo é considerado um **multigrafo**. As três arestas que incidem nos vértices 6 e 8 da Figura 1 indicam que o grafo em questão é um multigrafo.
- Um grafo é dito ser **desconexo** se há pelo menos um par de vértices para o qual, partindo de um deles e atravessando qualquer sequência finita de arestas, não é possível atingir o outro. Caso esta propriedade não valha para nenhum par de vértices no grafo, ele é **conexo**. Cada subgrafo conexo de um grafo é dito ser uma **componente**. O grafo representado na Figura 1 é desconexo e possui três componentes.

Um grafo é **simples** quando não é um multigrafo, não é ponderado e não possui laços. O grafo representado acima, por exemplo, não é simples. No decorrer deste artigo, por uma questão de simplicidade, manteremos o enfoque apenas em grafos simples.

Por fim, introduziremos o conceito de **isomorfismo entre grafos**; dois grafos  $G(V, E)$  e  $G'(V', E')$  são **isomorfos** se e somente for possível obter um a partir do outro apenas via renumeração de seus vértices. Mais formalmente,  $G(V, E)$  e  $G'(V', E')$  são **isomorfos** se e somente se existir uma função bijetora (isomorfismo) entre  $V$  e  $V'$  que preserve as relações de adjacência de  $G$  e  $G'$ .

Uma rede é um grafo utilizado para a representação de um **sistema complexo**, tipo de sistema que será definido ainda nesta seção. Por este motivo, no corpo deste artigo, todas as definições que se aplicam a grafos serão, com igual naturalidade, aplicadas a redes.

## 2.2 Sobre Física Estatística

Em oposição à visão inerentemente macroscópica da termodinâmica, a Física Estatística trata seus sistemas sob uma perspectiva reducionista, procurando, por equações matemáticas capazes de descrever comportamentos microscópicos, e via operadores de médias (no domínio do tempo, espaço ou frequência), recuperar a fenomenologia macroscópica.

Como exemplo, temos o conceito de temperatura, tomado na Física Estatística como mensuração do grau de agitação (energia cinética) dos constituintes (microscópicos) do sistema (macroscópico).

A Física Estatística está fortemente presente nos fundamentos da Teoria das Redes Complexas uma vez que, por exemplo, diversas características dinâmicas (macroscópicas) de uma rede podem ser descritas tomando como base grandezas básicas, como a sua distribuição de graus (microscópica).

## 2.3 Sobre Sistemas Complexos

A Teoria de Sistemas Complexos surge no final do século XX se propondo a tratar de sistemas com algumas características especiais:

- grande número de constituintes que interagem muitas vezes de forma não-linear e se relacionam com o meio - tanto influenciando quanto por ele sendo influenciado;
- invariância por escala, ou seja, a presença de padrões auto-similares (geometria fractal) e/ou auto-afins, e distribuições (de frequência e/ou energia) obedecendo a leis de potência (criticalidade auto-organizada);
- exibição de propriedades coletivas - logo, dividir o sistema em partes menores para tentar analisá-lo melhor, em alguns casos, não é apropriado.

A Teoria de Redes Complexas é, hoje, amplamente aplicada tanto à caracterização quanto à modelagem matemática de Sistemas Complexos. Modelagem, no contexto deste artigo, deve ser entendido como o processo de simplificação de uma realidade física, marcado pelo equilíbrio entre confiabilidade e operacionalidade. Um modelo deve apresentar o máximo de características do sistema original, mas não pode inviabilizar o estudo eficiente de suas propriedades - quer seja analiticamente, quer seja computacionalmente.

Com o desenvolvimento das ferramentas computacionais e a visível ruptura dos limites disciplinares, a Teoria dos Grafos passou a ser cada vez mais utilizada na modelagem de diversos Sistemas Complexos. Isto se deu no início de uma época de síntese do conhecimento científico, ressaltando diversas interconexões entre as mais distintas áreas. À medida que a complexidade dos sistemas modelados crescia (número muito grande de constituintes - vértices, e interações - arestas), tornou-se usual utilizar ferramentas da Física Estatística, como invariância de escala e auto-similaridade, no estudo dos modelos.

### 3 Linha do Tempo

Nesta seção, trataremos de marcos no estudo das Redes Complexas cujos primeiros fundamentos coincidem com as origens da Teoria dos Grafos.

**1736** - O matemático e físico suíço Leonhard Euler, a um só tempo, fundamentou a Teoria dos Grafos e a Topologia quando respondeu o que hoje é conhecido como o **Problema das Pontes de Königsberg**. O problema colocava em questão a possibilidade de atravessar as sete pontes que separavam os quatro bairros da cidade prussiana, sem passar duas vezes pela mesma ponte, como indicado na Figura 2. Euler construiu uma representação simplificada e operacional da cidade, na qual os bairros são vértices e as pontes são arestas - ou seja, ele modelou a cidade através de um grafo. O matemático concluiu que, para qualquer rede de pontes, só é possível fazer um passeio completo, atravessando uma única vez cada ponte, se todos os bairros tiverem um número par de pontes ou se apenas dois bairros tiverem um número ímpar de pontes. No caso de Königsberg, existem quatro bairros e todos são ligados por um número ímpar de pontes. Desta forma, Euler explicou por quê é impossível atravessar cada uma das pontes de Königsberg apenas uma vez, além de ter produzido uma regra que pode ser aplicada a qualquer rede de pontes em qualquer cidade

do mundo.<sup>3</sup>

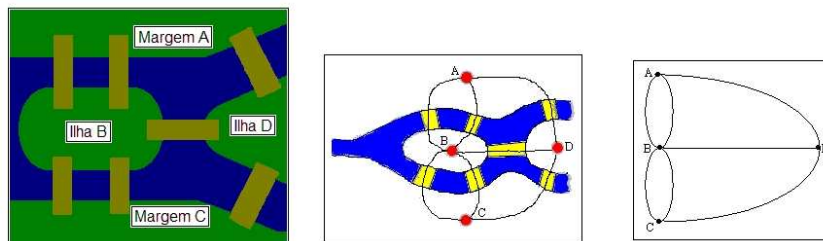


Figura 2: A modelagem matemática de Kaliningrado proposta por Euler

**1847** - Estudando circuitos elétricos, o físico russo Gustav Robert Kirchhoff iniciou o desenvolvimento da Teoria das Árvores, uma particular e importante classe de grafos cuja codificação<sup>4</sup>, feita por Prüfer, data de 1918. A Teoria das Árvores prossegue com Cayley (1857), aplicada à Química Orgânica, e com Jordan (1869), de forma puramente matemática<sup>5</sup>.

**1852** - A conjectura das quatro cores, enunciada em 1852 pelo matemático inglês Francis Guthrie, diz que todo mapa desenhado no plano, dividido em um número qualquer de regiões, pode ser colorido, de forma que regiões fronteiriças não recebam a mesma cor, com um mínimo de quatro cores. Esta conjectura foi demonstrada apenas em 1976, com o auxílio de um computador. O importante conceito de coloração de um grafo surgiu inspirado nessa conjectura. Kempe (1879) e Tait (1880) tentaram, sem sucesso, provar a conjectura; Heawood (1890) não só demonstrou a falsidade da prova de Kempe, como demonstrou a conjectura para cinco cores. Conceitos importantes para a teoria de grafos surgiram inspirados nesse problema, como os polinômios cromáticos de Birkhoff (1912), o grafo dual de Whitney (1931), e o teorema que limita o número cromático (menor número de cores necessárias à coloração de um grafo) de um grafo, devido a Brooks (1941)<sup>6</sup>.

**1859** - O matemático, físico e astrônomo irlandês William Rowan Hamilton inventou um jogo cujo objetivo era percorrer, uma única vez, todos os vértices de um dodecaedro regular. Observe a semelhança entre este jogo e o clássico problema de Euler, que consistia em percorrer, uma única vez, todas as arestas do grafo que representava Königsberg e suas pontes. Inspirados nesses trabalhos, surgiram alguns conceitos simples e importantes para a Teoria dos Grafos: o ciclo e o caminho euleriano e hamiltoniano<sup>7</sup>.

<sup>3</sup>Singh, S.; trad. Calife, J. L.; O Último Teorema de Fermat: a história do enigma que confundiu as maiores mentes do mundo durante 358 anos; Editora Record, Rio de Janeiro, 1998, 2ª Edição

<sup>4</sup>Processo de obtenção da representação única - código - de um grafo em um determinado esquema.

<sup>5</sup>Boaventura Netto, P. O.; Grafos: Teoria, Modelos, Algoritmos; Editora Edgard Blücher, São Paulo, 2001, 2ª Edição Revista e Ampliada

<sup>6</sup>*idem*

<sup>7</sup>*ibid.*

**1959** - Erdős e Rényi iniciaram o estudo sistemático de grafos aleatórios, com o propósito de, através de métodos probabilísticos, estudar propriedades dos grafos em função do crescimento de conexões aleatórias entre vértices. Um grafo é dito **aleatório** quando existe uma disposição desordenada de arestas conectando os seus vértices<sup>8</sup>.

**1967** - Stanley Milgran, psicólogo social da Universidade de Harvard, promoveu um experimento para estudar o chamado **Problema do Pequeno Mundo**, para avaliar o grau de ligação entre pessoas. O experimento consistia em pedir a indivíduos de algumas cidades que enviassem cartas a conhecidos, com o objetivo de chegar a determinados residentes em Boston. A partir deste experimento, surgiu o conceito de seis graus de separação entre pessoas, mostrando que há uma probabilidade alta de que indivíduos desconhecidos possuam amigos em comum<sup>9</sup>.

**1998** - Duncan J. Watts e Steven Strogatz desenvolveram um algoritmo baseado em grafos aleatórios para estudar o Problema do Pequeno Mundo de maneira mais geral, procurando assemelhar a rede criada à estrutura encontrada em interações sociais. Neste modelo, é possível obter redes de pequeno mundo ajustando o nível de aleatoriedade com que são reconectados os vértices da rede regular original. Com o máximo de aleatoriedade, obtém-se uma rede aleatória<sup>10</sup>.

**1999** - Albert-László Barabási e Réka Albert publicaram um artigo propondo um modelo genérico para construir redes que se assemelham à estrutura encontrada em redes genéticas, ou redes de *internet*. Tais redes foram chamadas de redes livres de escala. Neste modelo, uma rede original aleatória, com poucos vértices, é expandida continuamente pela adição de novos vértices, que são ligados preferencialmente a vértices com maiores números de conexões<sup>11</sup>.

## 4 Representando Redes através de Matrizes

Para representar uma rede matematicamente, utilizamos formalismos aplicados à Teoria dos Grafos. Nesta seção, discutiremos dois tipos de matrizes utilizados nesta tarefa: a **matriz de adjacência** e a **matriz de vizinhança**. A primeira delas é bastante comum na literatura de Teoria dos Grafos e a segunda, apresentada pela primeira vez em um artigo<sup>12</sup> de 2008, possui algumas características

---

<sup>8</sup>Erdős, P., Rényi, A.; On Random Graphs. I.; Publicationes Mathematicae 6: 290–297, 1959

<sup>9</sup>Travers, J., Milgram, S.; An Experimental Study of the Small World Problem; Sociometry, Vol. 32, No. 4, pp. 425–443, 1969

<sup>10</sup>Watts, D.J., Strogatz, S.H.; Collective dynamics of 'small-world' networks; Nature 393 (6684): 409–10, 1998, doi:10.1038/30918

<sup>11</sup>Barabási, A. L., Albert, R; Emergence of scaling in random networks; Science, 286:509–512, October 15, 1999

<sup>12</sup>Andrade, R. F. S., Miranda, J. G. V., Pinho, S. T. R., Lobão, T. P.; Characterization of complex networks by higher order neighborhood properties; European Physical Journal B, v. 61, p. 247–256, 2008

que fazem dela uma representação bastante útil para redes.

## 4.1 A Matriz de Adjacência

A matriz de adjacência guarda informação sobre todas as relações de adjacência de uma rede. Se ela possui  $n$  vértices, sua matriz de adjacência  $M_{n,n}$  é construída da seguinte forma:

$$M_{i,j} = \begin{cases} 1, & \text{se há uma aresta entre os vértices } i \text{ e } j \text{ do grafo} \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

	1	2	3	4	5
1	0	1	1	0	0
2	1	0	1	0	0
3	1	1	0	1	0
4	0	0	1	0	1
5	0	0	0	1	0

Figura 3: A matriz de adjacência de um grafo  $G$

## 4.2 A Matriz de Vizinhaça

Para definir a matriz de vizinhaça, é preciso compreender a construção das **matrizes de adjacência de ordem superior** ou, simplesmente, **MLsreferencia fesc**. Para isto, é importante que o leitor domine o conceito de distância, oriundo da Teoria dos Grafos<sup>13</sup>

Quando dois vértices  $i$  e  $j$  são adjacentes, temos que sua distância  $d$  obedece à seguinte equação:

$$d(i, j) = 1 \quad (3)$$

Esta característica pode ser visualizada através da matriz de adjacência. Mas, para esta mesma matriz, não é tão fácil visualizar a distância entre vértices não-adjacentes. Uma solução para este problema pode ser apresentada através do conceito de adjacência de ordem superior<sup>14</sup>. Este conceito estende a idéia de adjacência para quaisquer dois vértices  $i, j$  cuja distância esteja entre zero e o diâmetro da rede. Dada uma rede de  $n$  vértices e diâmetro  $m$ , a construção de qualquer  $ML_{n,n}$ ,  $0 \leq L \leq m$ , obedece à seguinte equação:

<sup>13</sup>Apresentamos este conceito na seção **Conceitos Preliminares**, subseção **Introdução à Teoria dos Grafos**.

<sup>14</sup>Andrade, R. F. S. *et al*; *op. cit.*; Eur. Phys. J. B, v. 61, p. 247-256, 2008

$$ML_{i,j} = \begin{cases} 1, & \text{se } d(i,j) = L \\ 0, & \text{caso contrário} \end{cases} \quad (4)$$

Se a distância entre dois vértices  $i,j$  for infinita, temos que:

$$ML_{i,j} = 0 \quad \forall L \quad (5)$$

Com este conceito bem estabelecido, é trivial definir a matriz de vizinhança  $V_{n,n}$  de uma rede com  $n$  vértices e diâmetro  $m$ :

$$V_{i,j} = \sum_{L=0}^m LML_{i,j}. \quad (6)$$

A matriz de vizinhança condensa a informação contida nas  $MLs$ , ao tempo em que o fator  $L$  explicita a ordem de vizinhança em  $V_{n,n}$ , facilitando a visualização e estudo de propriedades associadas à distância entre vértices em uma rede. Esta matriz não aumenta o poder de representação da matriz de adjacência, mas simplifica consideravelmente o cálculo e a interpretação de diversos índices caracterizadores das Redes Complexas.

	1	2	3	4	5
1	0	1	1	2	3
2	1	0	1	3	3
3	1	1	0	1	2
4	2	2	1	0	1
5	3	3	2	1	0

Figura 4: A matriz de vizinhança de um grafo  $G$

Dada uma rede com seus vértices já enumerados, há uma e somente uma matriz de adjacência que a representa. Em contrapartida, dada uma matriz de adjacência, há uma e somente uma família de grafos isomorfos que traduzem



fielmente sua informação, uma vez que podemos enumerar os  $n$  vértices de um grafo de  $n!$  formas distintas.

## 5 Índices Básicos

Mensurar grandezas é tarefa fundamental a um sistema formal de estudo. Nesta seção, portanto, discutiremos alguns dos índices básicos à análise de uma rede.

### 5.1 Tamanho e Ordem

Seja  $G(V, E)$  um grafo. A **ordem** de  $G$  é a cardinalidade de seu conjunto de vértices  $V$  e seu **tamanho** é a cardinalidade de seu conjunto de arestas  $E$ .

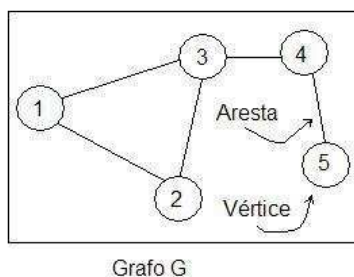


Figura 5: Grafo  $G$  com ordem e tamanho igual a 5

### 5.2 Grau de um Vértice, Grau Médio de um Grafo e Distribuição de Graus

Seja  $i$  um vértice qualquer de um grafo  $G$ . O **grau** de  $i$ ,  $k_i$ , é definido como a cardinalidade do conjunto de todos os vértices adjacentes a  $i$ . Tomemos, por exemplo, o vértice 3 do grafo da Figura 5. Seu grau é 3, pois ele está ligado a 3 vértices: 1, 2 e 4. O **grau médio**  $\langle k \rangle$  de  $G$  é a média aritmética dos graus de cada vértice. O grau médio de  $G$  é  $\langle k \rangle = 10/5 = 2$ . Fazendo um histograma dos possíveis graus  $k$  encontrados em um grafo, temos a sua **distribuição de graus**. Além disso, podemos definir o índice  $p_k$  como a probabilidade de um vértice, escolhido aleatoriamente em um grafo, ter grau igual a  $k$ . A distribuição de graus é uma propriedade estatística, utilizada para caracterizar Redes Complexas.

### 5.3 Coeficiente de Aglomeração de um Vértice, Coeficiente de Aglomeração Médio

Seja  $i$  um vértice qualquer de uma rede  $A$ . O **coeficiente de aglomeração** de  $i$  é a probabilidade de que os vértices pertencentes a  $A_i$  sejam adjacentes entre

si. Podemos calcular seu coeficiente de aglomeração a partir da igualdade

$$C_i = \frac{2n_i}{k_i(k_i - 1)} \quad (7)$$

sendo  $n_i$  o número de arestas entre seus adjacentes e  $k_i$  o grau do vértice.

O coeficiente de aglomeração médio de  $A$  é a média aritmética dos coeficientes de aglomeração de cada vértice. O coeficiente de aglomeração do grafo  $G$  (Figura 5), por exemplo, é 0.467.

#### 5.4 Caminho Mínimo Médio de um Vértice, Caminho Mínimo Médio e Diâmetro de um Grafo

Um **caminho** entre dois vértices  $i$  e  $j$  de um grafo  $G$  é uma sequência de  $k$  vértices  $v_1, \dots, v_k$  tal que

- $v_1 = i$  e  $v_k = j$ ;
- existe uma aresta entre o vértice  $v_l$  e o vértice  $v_{l+1}$ ,  $1 \leq l \leq k$ ;
- não há repetição de vértices nesta sequência, nem de arestas entre estes vértices.

O **comprimento de um caminho** entre dois vértices  $i$  e  $j$  equivale ao número de arestas que conectam todos vértices pertencentes a este caminho. A **distância** entre dois vértices  $i$  e  $j$  é o comprimento do menor caminho entre estes vértices - ou seja, é o comprimento associado à menor sequência de vértices entre  $i$  e  $j$ .

Neste artigo, toda vez que discutirmos alguma noção relacionada a caminho entre vértices, estaremos nos referindo a seu comprimento - um número - e não a uma sequência bem definida de vértices.

Seja  $i$  um vértice qualquer de um grafo  $G$ . O **caminho mínimo médio de  $i$**  é a razão entre sua distância em relação a todos os outros vértices de  $G$ <sup>15</sup> e seu grau. O **caminho mínimo médio de  $G$** , por sua vez, é a média aritmética dos caminhos mínimos médios de cada um de seus vértices. Para o grafo  $G$  da figura acima, o caminho mínimo médio é 4.13.

## 6 Classificação e Modelos de Geração

As Redes Complexas podem ser classificadas tomando como objeto de observação as suas propriedades estatísticas, destacando-se entre elas a distribuição de graus e o coeficiente de aglomeração. Os tipos de redes mais encontrados em trabalhos sobre Redes Complexas são:

- redes regulares;
- redes aleatórias;

---

<sup>15</sup>Estamos considerando a distância entre dois vértices pertencentes a componentes distintas como zero.

- redes livres de escala;
- redes de pequeno mundo;
- redes hierárquicas e modulares.

## 6.1 Redes Regulares

Em redes regulares, todos os vértices apresentam o mesmo grau. Devido a facilidades operacionais, redes regulares são fundamentais ao estudo de diversos sistemas. No contexto da Física, por exemplo, modelos atômicos são estudados através de redes regulares.

## 6.2 Redes Aleatórias

As redes aleatórias são geradas a partir de ligações aleatórias entre os vértices de um conjunto. Em outras palavras, dado um conjunto de vértices, é atribuída, para cada um de seus elementos, igual probabilidade de que ele se conecte com outro elemento qualquer deste conjunto.

Podemos construir uma rede aleatória com  $n$  vértices e  $m$  arestas a partir da seguinte relação entre estes valores:

$$m = p \frac{n(n-1)}{2} \quad (8)$$

Na equação acima,  $p$  é a probabilidade de haver uma aresta entre quaisquer  $n$  vértices da rede. A fração  $\frac{n(n-1)}{2}$  indica o número máximo de arestas em uma rede simples com  $n$  vértices.

Redes aleatórias apresentam uma distribuição de graus característica: a distribuição de Poisson (normal - gaussiana), com grau médio dado por  $p(n-1)$ . O coeficiente de aglomeração médio é dado por  $p$  e, portanto, independe de  $n$ .

Dois dos mais importantes pesquisadores que trabalharam com redes aleatórias foram Erdős e Rényi.

## 6.3 Redes Livres de Escala

Na geração de uma rede livre de escala, a cada passo, surgem arestas conectando os vértices que, no momento, apresentam os maiores graus. Por este motivo, a distribuição de graus de uma rede livre de escala segue uma lei de potência na qual poucos vértices possuem altos graus<sup>16</sup> e a maioria dos vértices apresentam graus baixos. O modelo de Barabási e Albert é um dos mais usados para gerar redes livres de escala. A idéia fundamental é o crescimento da rede via o princípio da ligação preferencial: quão maior o grau de um vértice  $i$ , maior a probabilidade de que a cardinalidade de  $A_i$  aumente na próxima iteração do

<sup>16</sup>Em uma rede livre de escala, poucos vértices possuem altos graus, mas em maior quantidade do que em redes aleatórias.



Figura 6: Erdős (1913 - 1996) e Rényi (1921 - 1970)

processo de crescimento. A rede é gerada a partir de um número pequeno de vértices iniciais e, a cada passo de tempo, são acrescentados novos vértices e arestas, definidos a partir de regras como:

- $p(k) = \frac{k}{\sum k}$ , sendo  $p$  a probabilidade de surgir uma aresta e  $k$  o grau de determinado vértice;
- $n = t + mt$ , sendo  $n$  o número de vértices após um número  $t$  de passos de tempo e  $mt$  o atual número de arestas.

Um processo de **ataque** em uma rede consiste na remoção de um vértice de alto grau. Em contrapartida, uma **falha** é uma remoção indiscriminada de um vértice da rede. Uma noção muito discutida, e importante, sobre redes livre de escala é sua **tolerância a falhas**. Isto implica que remoções aleatórias de vértices da rede atingem, em grande parte das vezes, vértices de baixo grau, pois eles são maioria. Com isto, considerando o grau de um vértice como uma medida da sua informação, há baixa probabilidade de uma grande perda de informação em processos de falha em uma rede livre de escala. Por outro lado, este tipo de rede é extremamente sensível a ataques<sup>17</sup>.

## 6.4 Redes de Pequeno Mundo

Como indicado na Linha do Tempo, Stanley Milgran realizou um experimento no qual aproximadamente cento e cinquenta famílias em Omaha, Nebraska e Wichita, Kansas deveriam entregar correspondências a pessoas alvo em Boston, utilizando apenas seus amigos. Cada envelope continha:

- Nome, endereço e alguns dados pessoais da pessoa alvo;

<sup>17</sup>Redes aleatórias, por sua vez, são sensíveis a falhas e tolerantes a ataques.



Figura 7: Barabási e Albert

- um conjunto de regras instruindo a pessoa intermediária que o recebeu sobre como proceder, podendo ser assim resumido: se você não conhece diretamente a pessoa alvo, repasse este envelope para um amigo que você conhece pessoalmente e que, provavelmente, a conheça;
- cada pessoa que recebesse este envelope deveria escrever seu nome nele, pois desta forma estaria evitando que uma mesma pessoa o recebesse novamente.

Inicialmente, Milgran e seus colegas acreditavam que as correspondências chegariam ao seu destino em torno de cem passos aproximadamente. Ao término do experimento, no entanto, os resultados mostravam que o trânsito das correspondências levavam entre cinco e seis passos em média e, assim, surgiu o conceito de **seis graus de separação**, que comprova que pessoas aparentemente sem relação alguma têm uma grande probabilidade de possuírem, em algum grau, amigos em comum que as aproximem.

Watts e Strogatz, em 1998, propuseram um algoritmo baseado em redes aleatórias, no qual buscavam mimetizar a topologia de interações sociais em um modelo abstrato para tentar estudar este mesmo problema de uma maneira mais geral. O fruto desse algoritmo são as redes de pequeno mundo, que possuem comportamento e características entre os apresentados por uma rede aleatória e uma rede regular. Redes de pequeno mundo apresentam coeficiente de aglomeração médio maior e mínimo caminho médio menor que uma rede aleatória de mesmo número de vértices e arestas.

Redes de pequeno mundo podem ser geradas por uma aplicação que retire conexões de uma rede regular e promova reconexões ou, simplesmente, acrescente ligações entre os vértices. No caso da promoção de reconexões, se reconecta com probabilidade  $p$ , aleatoriamente, uma rede com  $n$  vértices de grau  $k$ <sup>18</sup>. Assim, temos uma rede regular para  $p = 0$  e uma aleatória para  $p = 1$ . Para valores intermediários de  $p$ , obtemos redes de pequeno mundo.

Pode-se tomar o conceito de rede de pequeno mundo como passível de generalização a um **efeito dinâmico de pequeno mundo**: situação de crescimento

<sup>18</sup>Neste caso, o número de reconexões é igual a  $\frac{p \times n \times k}{2}$ .



Figura 8: Watts e Strogatz

do coeficiente de aglomeração médio e decréscimo do mínimo caminho médio em uma rede.

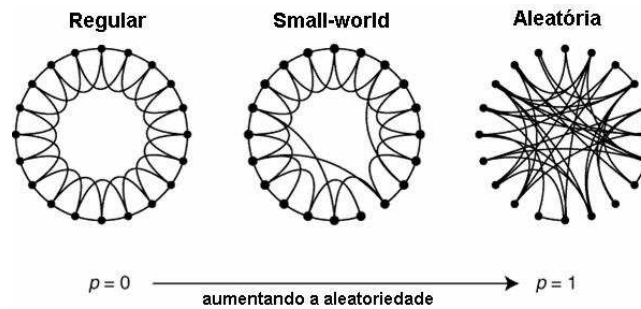


Figura 9: Gerando uma rede de pequeno mundo a partir de uma rede regular

## 6.5 Redes Modulares e Hierárquicas

A característica mais importante de uma rede hierárquica é a relação de lei de potência entre o coeficiente de aglomeração de um vértice e seu grau. A arquitetura hierárquica implica que vértices distantes são partes de áreas de alta aglomeração e que a comunicação<sup>19</sup> entre estas áreas, as quais serão chamadas de agora em diante de módulos, é feita por um pequeno número de vértices.

Mais formalmente, um **módulo** pode ser assim definido: dado um grafo  $G(V, E)$  e um conjunto  $S \subseteq V$ , dizemos que  $S$  se trata de um módulo se e somente se o número de arestas entre os vértices de  $S$  for muito maior do que o número de arestas que ligam os vértices de  $S$  a vértices de  $V \setminus S$ . Não há, entretanto, uma razão crítica entre estes números de arestas que determine se um conjunto  $S$  de vértices é ou não um módulo; logo, definir se um rede é ou não modular depende muito dos objetivos da modelagem matemática em questão.

Dado um sistema complexo em que uma grandeza  $y$  varie em lei de potência frente a uma grandeza  $x$

<sup>19</sup>A comunicação entre dois vértices, neste caso, deve ser compreendida como um caminho que os conecta.

$$y \sim x^a \quad (9)$$

temos que  $a$  é o expoente característico desta relação. Como estamos diante de uma lei de potência,  $a$  é uma informação relevante para compreender como  $y$  depende da ordem do sistema<sup>20</sup>.

Uma rede é dita hierárquica se e somente se seu coeficiente de aglomeração  $\langle C \rangle$  variar em lei de potência com o grau  $\langle K \rangle$ . Portanto, modularidade não implica em hierarquia. Em uma rede modular não hierárquica, os módulos estão relativamente isolados do resto da rede.

Atente ao fato de que redes hierárquicas e redes livres de escala são conceitos independentes: a primeira apresenta relação de lei de potência quanto ao coeficiente de aglomeração; a segunda, quanto ao grau.

### 6.5.1 Outra Maneira de Definir Hierarquia: a Rede Hierárquica do Diamante (RHD)

A RHD é um tipo de rede bastante utilizado na Física Estatística, especialmente como substrato à construção de modelos de *spins*. É fundamental ressaltar que o conceito de hierarquia na RHD é distinto do adotado anteriormente. As regras de formação da RHD podem ser assim resumidas:

- A princípio, a rede possui dois vértices (sítio raiz) ligados por uma aresta;
- substitui-se a aresta original por quatro novas, formando um losango com os dois vértices originais e dois vértices novos;
- repete-se o passo anterior um determinado número de vezes.

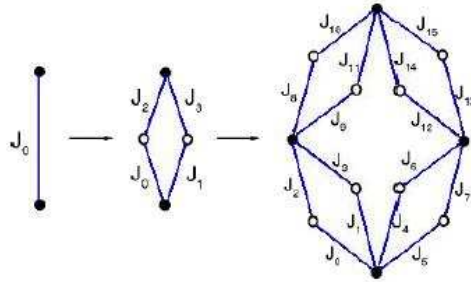


Figura 10: Construção iterativa da RHD

Após  $t$  iterações, a RHD possuirá  $n = \frac{2}{3} \times (2 + 4^t)$  vértices e  $m = 4^t$  arestas. A maior parte dos vértices, gerados na última iteração, possuirá apenas dois vértices adjacentes; alguns poucos, os iniciais, possuirão grau  $2^t$ ; os demais vértices apresentarão um grau entre estes dois extremos.

<sup>20</sup>A ordem do sistema será denotada pelo número de elementos que o constituem. Esta definição é análoga à apresentada na seção 5.1 deste texto.

Na figura abaixo, mostramos três tipos de redes: aleatória (A), livre de escala não hierárquica (B), livre de escala hierárquica (C); bem como uma configuração padrão da rede (a), a distribuição de graus (b), e a distribuição do coeficiente de aglomeração (c).

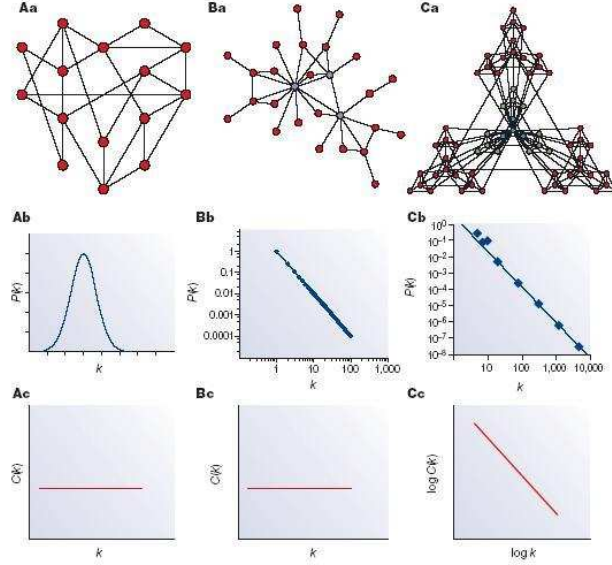


Figura 11: Comparação entre as redes aleatórias, livre de escala não hierárquica, e livre de escala hierárquica

Muitas dessas classificações não são excludentes. Como exemplo, temos as redes Apolonianas, que apresentam distribuição de graus em lei de potência, baixo mínimo caminho médio e alto coeficiente de aglomeração quando comparadas com redes aleatórias de mesma ordem, e relação de escala entre o coeficiente de aglomeração e o grau de cada vértice. Logo, estas redes são, simultaneamente, livres de escala, de pequeno mundo e hierárquicas.

## 7 Aspectos Computacionais e Algoritmos

Nesta seção, trataremos de algoritmos para armazenar redes complexas e calcular os índices básicos apresentados anteriormente.

### 7.1 Representando um grafo em uma matriz de adjacência

Para representar um grafo através de uma matriz de adjacência<sup>21</sup>, é preciso conhecer de antemão a cardinalidade do conjunto de vértices do grafo, denotada por  $n$ . Uma matriz  $A_{n,n}$ , em seguida, deve ser alocada na memória principal. A alocação desta matriz pode ser estática, sendo pré-definida por  $n$ . Caso

<sup>21</sup>Markenzon, Lilian; Representações Computacionais de Grafos; São Carlos, SP; SBMAC, 2006.



$n$  se altere durante o processamento do grafo, é possível utilizar uma matriz alocada dinamicamente, para que não haja perda de informação sobre os vértices ou mesmo desperdício de espaço. Basicamente, cada campo  $A_{i,j}$  dessa matriz é assumido inicialmente como nulo, caso não haja uma aresta conectando os vértices  $i$  e  $j$ , ou com um, se essa aresta existe. Como os possíveis valores iniciais de  $A_{n,n}$  são apenas dois, utilizar um único *bit* para cada campo é suficiente. Isso, inclusive, proporciona uma grande economia de espaço em memória. O acesso à memória *bit a bit*, entretanto, ainda não é suportado de maneira muito confortável pela maioria das linguagens de programação. Por este motivo, esta representação não é tão comum e, na prática,  $A_{i,j}$  é normalmente representada como uma matriz de inteiros ou *booleans*.

## 7.2 Matriz de Vizinhaça

A matriz de vizinhaça  $V_{n,n}$  pode ser alocada em memória principal da mesma forma que a matriz de adjacência. Como ela é uma combinação linear de várias *MLs*, a sua inicialização é menos trivial. Apresentaremos um algoritmo que inicializa  $V_{n,n}$  a partir de  $M_1$  - ou seja, a partir da matriz de adjacência  $A_{n,n}$ .

Neste algoritmo, geramos indutivamente  $M_p$  a partir de  $M_{p-1}$ ,  $2 \leq p \leq d$  e  $d$  o diâmetro da rede, e de uma matriz que armazena a soma de todas as  $M_k$ ,  $1 \leq k \leq p-1$ . A esta última matriz, daremos o nome de *MS*. *MS* deve ser inicializada também com a matriz de adjacência.

Para gerar  $M_p$ , fazemos:

```

1  for  $i \leftarrow 1$  to  $n$  do
2      for  $j \leftarrow 1$  to  $n$  do
3          if  $MS_{i,j} = 1$  then
4               $Mp_{i,j} \leftarrow 0$ ;
5          else
6              if  $MS_{i,j} = 0$  then
7                  for  $k \leftarrow 1$  to  $n$  do
8                      if  $M_{p-1,i,k} = 1$  e  $M_{p-1,k,j} = 1$  then
9                           $Mp_{i,j} \leftarrow 1$ ;
10                     end
11                 end
12             end
13         end
14     end
15 end

```

**Algoritmo 1:** Geração indutiva da matriz de vizinhaça.

Se os vértices  $i$  e  $j$  são adjacentes para alguma  $M_{p-1}$ , eles não podem ser adjacentes para a  $M_p$ , pois há um número de arestas menor que  $p$  os conectando. Isto justifica as linhas 3 e 4 do algoritmo, que utilizam a informação armazenada pela *MS*. A linha 5 indica que um outro procedimento deve ser tomado, uma vez que ainda não se determinou a ordem da adjacência entre  $i$  e  $j$ . O restante das linhas determina se os vértices adjacentes a  $i$  em  $M_{p-1}$  são adjacentes a  $j$  nesta

mesma matriz, uma vez que  $i$  e  $j$ , de acordo com  $MS$ , não foram adjacentes até a ordem avaliada.

Finalmente, com todas as  $ML_s$  determinadas até o diâmetro, podemos gerar  $V_{n,n}$  através de uma combinação linear entre elas.

### 7.3 Grau e Grau Médio de um Grafo

Para determinar o grau de cada um dos  $n$  vértices de uma matriz de adjacência  $A_{n,n}$  armazenada na memória, devemos consultar cada uma de suas linhas, somando o conteúdo de cada um de seus campos. Para o processamento de cada linha, obteremos o grau de um único vértice - ou seja, para armazenar a informação de  $n$  graus, devemos utilizar um vetor de tamanho  $n$ ,  $GR$ . Abaixo, um algoritmo simples para obtenção dos graus de um grafo.

```

1 for  $i \leftarrow 1$  to  $n$  do
2    $GR_i \leftarrow 0$ ;
3   for  $j \leftarrow 1$  to  $n$  do
4      $GR_i \leftarrow GR_i + A_{i,j}$ ;
5   end
6 end

```

**Algoritmo 2:** Obtenção dos graus de todos os vértices de um grafo.

Este algoritmo funciona para grafos conexos e desconexos.

Para calcular o grau médio  $\langle k \rangle$  do grafo, a fim de criar uma visão global da estrutura, devemos fazer uma média aritmética de todos os valores do vetor.

### 7.4 Coeficiente de Aglomeração

O coeficiente de aglomeração mede a probabilidade de que, dados três vértices de um grafo  $G$ ,  $u$ ,  $v$  e  $w$ , tais que haja uma aresta entre  $u$  e  $v$  e uma aresta entre  $u$  e  $w$ , haja uma aresta entre  $v$  e  $w$ . Ou seja, este coeficiente analisa a probabilidade de que, dado um vértice  $u$  de  $G$ , vértices distintos adjacentes a  $u$  sejam adjacentes entre si. Daremos a esta situação o nome de **situação triângulo**.

Para definir computacionalmente o coeficiente de aglomeração de um vértice  $i$ , visitamos a  $i$ -ésima linha da matriz de adjacência  $A_{n,n}$ <sup>22</sup>, calcular seu grau e armazenar a informação de quais são seus vértices adjacentes. Em seguida, calcula-se o número máximo de situações triângulo nos quais  $i$  pode estar envolvido. Este número, que será chamado de  $CT_i$ , pode ser obtido da seguinte forma:

$$CT_i \leftarrow \frac{GR_i \times (GR_i - 1)}{2}$$

<sup>22</sup>Para todos os algoritmos, consideramos que os vértices estão enumerados de modo que a  $i$ -ésima linha das matrizes de adjacência e vizinhança armazenem informação sobre o  $i$ -ésimo vértice

A divisão por dois é feita para que se elimine a contagem repetida de arestas.

Em seguida, caso  $CT_i$  seja diferente de zero, ou seja,  $i$  pode participar de pelo menos uma situação triângulo, fazemos:

```

1  $CL_i \leftarrow 0$ ;
2 for  $j \leftarrow 1$  to  $n$  do
3   if  $A_{i,j} = 1$  then
4     for  $k \leftarrow 1$  to  $n$  do
5       if  $A_{i,k} = 1$  e  $A_{j,k} = 1$  then
6          $CL_i \leftarrow CL_i + 1$ ;
7       end
8     end
9   end
10 end
11  $CL_i \leftarrow \frac{CL_i}{2}$ 

```

**Algoritmo 3:** Contabilizando situações triângulo envolvendo todos os vértices de um grafo.

As linhas de 2 a 6 do algoritmo acima acumulam em  $CL_i$  o número de vezes em que, para todo  $j$  adjacente a  $i$ ,  $j$  é adjacente a um  $k$  que também é adjacente a  $i$ . Este número, entretanto, precisa ser dividido por dois por contabilizar situações equivalentes duas vezes (linha 11). Com isto,  $CL_i$  passa a armazenar a quantidade de situações triângulo das quais  $i$  realmente faz parte.

Finalmente, o coeficiente de aglomeração  $CA_i$  de  $i$  é calculado como a relação entre as situações triângulo efetivamente encontradas e a totalidade de situações triângulo possíveis para  $i$ :

$$CA_i \leftarrow \frac{CL_i}{CT_i}$$

Para cada vértice  $i$  de  $G$ , calculamos  $CA_i$ . Caso  $CT_i$  seja igual a zero,  $CA_i$  deve ser definido como zero. Estas informações são particulares, de modo que, para representar o grafo  $G$  globalmente, calculamos o coeficiente de aglomeração  $CA$  de  $G$ :

$$CA \leftarrow \frac{\sum CA_i}{n}$$

## 7.5 Caminho Mínimo Médio de um Vértice

A idéia de caminho mínimo médio está intimamente relacionada com os conceitos de matriz de vizinhança e de grau. Para a especificação do algoritmo, utilizaremos a matriz de vizinhança  $V_{n,n}$  e o vetor de graus  $GR$  definidos nas subseções anteriores. Armazenaremos o caminho mínimo médio de cada vértice  $i$  de um grafo  $G$  em um vetor de tamanho  $n$ , ao qual daremos o nome de  $CMM$ . O algoritmo abaixo utiliza estas estruturas para calcular o caminho mínimo médio de todos os  $n$  vértices de  $G$ .

Para o algoritmo acima, estamos considerando que o o grafo  $G$  é conexo. Isto implica, em particular, que  $GR_i \neq 0$  para qualquer vértice  $i$  de  $G$ . Se  $G$  for

```

1 for  $i \leftarrow 1$  to  $n$  do
2    $CMM_i \leftarrow 0$ ;
3   for  $j \leftarrow 1$  to  $n$  do
4      $CMM_i \leftarrow CMM_i + V_{i,j}$ ;
5   end
6    $CMM_i \leftarrow \frac{CMM_i}{GR_i}$ ;
7 end

```

**Algoritmo 4:** Cálculo do mínimo caminho médio de todos os vértices de um grafo.

desconexo, podemos aplicar este mesmo algoritmo para cada uma de suas componentes conexas com mais de um vértice separadamente. Se a componente possuir apenas um vértice, ela é constituída de um único vértice **isolado** e, portanto, não faz sentido calcular o caminho mínimo médio deste vértice.

## 8 Algumas Ferramentas

Discutiremos brevemente nesta seção algumas ferramentas úteis ao estudo de Redes Complexas. Trataremos então, essencialmente, de

- um poderoso *software* de visualização de redes - o **Pajek**<sup>23</sup> ;
- uma forma alternativa de visualização da matriz de vizinhança - a **matriz de cores** .

### 8.1 O Pajek

O *Pajek*<sup>24</sup> é um programa destinado à análise e visualização de redes, especialmente útil quando ela possui um grande número de vértices.

O *Pajek* é capaz de gerar a representação gráfica de uma rede a partir de um arquivo que contenha sua matriz de adjacência ou sua relação de arestas. Inicialmente, ele apresenta os vértices na tela, mantendo no centro os de maiores graus e, na periferia, os de menores graus. Todavia, é possível movê-los na tela, ou mesmo atribuir uma localização fixa para um determinado número de vértices, a depender da necessidade do usuário. Com este *software*, também é possível colorir vértices e arestas da rede e exportar a imagem para diversos formatos de arquivo.

### 8.2 A Matriz de Cores

De posse da matriz de vizinhança de uma rede, podemos utilizar um *software* de plotagem para visualizá-la a partir de uma representação via cores ou tons

<sup>23</sup>Batagelj, Vladimir e Mrvar, Anderj; Universidade de Ljubljana, Eslovênia; <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.

<sup>24</sup>A palavra *Pajek*, em Esloveno, significa aranha.

de cinza. Cada número, ou faixa de números, é representado com uma cor ou tom diferente, com frequência correlacionada ao número. Exibimos na figura abaixo uma rede e sua respectiva matriz de cores.

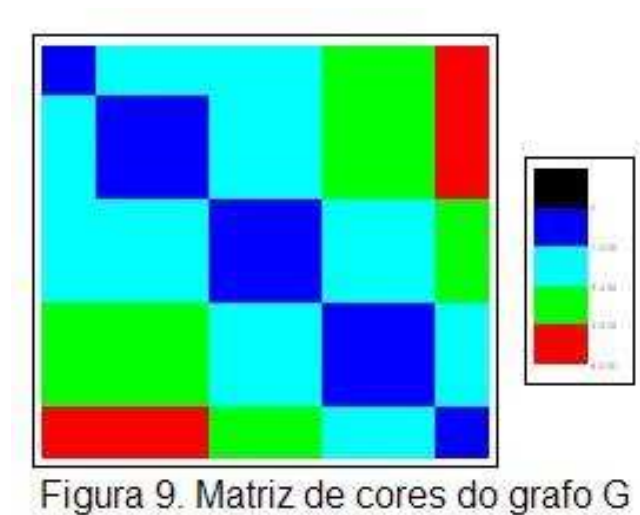


Figura 12: A matriz de cores de uma rede complexa de pequeno porte

## 9 Comentários Finais

Todo autor espera que sua obra deixe algo para a reflexão em seus leitores e conosco não é diferente. Este material possui o intuito de instigar - não de satisfazer. Estamos abertos a sugestões, críticas e dúvidas para que, a cada nova edição, este material seja aperfeiçoado e aumente seu alcance de público.

## 10 Agradecimentos

O apoio dos professores pesquisadores do grupo FESC foi fundamental para a escrita e revisão deste trabalho. Agradecemos, portanto, aos professores Ernesto P. Borges, José G. V. Miranda, Roberto F. S. Andrade, Suani T. R. Pinho e Thierry P. C. Lobão pelo tempo dedicado à construção e ao aperfeiçoamento deste projeto. Somos gratos também a todos os membros do grupo FESC, especialmente a Chares N. de Santana, Daniel Cason, Ivan C. da Rocha Neto e Viviane M. Galvão.