Matheus Palhares Viana

***A Metodologia das Redes Complexas para
Caracterização do Sistema de Havers***

São Carlos – SP

Janeiro / 2007

Matheus Palhares Viana

***A Metodologia das Redes Complexas para
Caracterização do Sistema de Havers***

Dissertação apresentada à Banca Examinadora
no Departamento de Física da Universidade de
São Paulo para a obtenção do título de Mestre
em Física Aplicada Computacional.

Orientador:

Prof. Dr. Luciano da Fontoura Costa

MESTRADO EM FÍSICA COMPUTACIONAL
DEPARTAMENTO DE FÍSICA
UNIVERSIDADE DE SÃO PAULO

São Carlos – SP

Março / 2007

Agradecimentos

Agradeço ao meu orientador, Luciano da Fontoura Costa, pela oportunidade de realizar este trabalho, pela orientação competente, pela confiança e amizade.

Aos meus amigos da UFSCAR, Fabio, Antônio Marcos, Jefferson, Helder, Ériton, Wagner e Lara.

Aos amigos da república Asgard, Gerson, Ivan, Lucas e Rafael.

Aos amigos Bruno e Luis, pela ajuda com as questões de programação e pelas discussões sobre Redes Complexas.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brasil e à sociedade brasileira, pelo financiamento deste trabalho.

Aos meus pais e meu irmão, responsáveis diretos por este trabalho.

À Évelyn, pela paciência.

*”É mais importante fazer as coisas que devem
ser feitas do que fazer as coisas como devem ser feitas.”*

P. Drucker

*”O Binômio de Newton é tão belo quanto a Vênus de Milo.
O que há é pouca gente para dar por isso.”*

Álvaro de Campos

Resumo

Esta dissertação apresenta um estudo detalhado do sistema de canais do osso cortical dos animais sob o ponto de vista das redes complexas. Este sistema é composto por canais paralelos ao eixo principal do osso e por canais perpendiculares. Estes canais são chamados de canais de Havers e Volkman, respectivamente. Sua principal função é conduzir os vasos sanguíneos responsáveis pela nutrição das células por toda estrutura óssea.

O conjunto de canais foi mapeado em uma rede na qual consideramos cada vértice como sendo a intersecção de um ou mais canais e as conexões entre os vértices como os próprios canais. Analisamos as características topológicas desta rede utilizando os conceitos clássicos de redes complexas, como conectividade, coeficiente de clusterização, distribuição do comprimento dos menores caminhos e detecção de comunidades. Também utilizamos uma abordagem hierárquica para algumas destas medidas.

Mostramos que a rede do osso cortical é altamente modular, sendo organizada em comunidades bem definidas e espacialmente localizadas, sendo este último fator importante para determinação das características topológicas da comunidade.

Os resultados demonstram que a rede Haversiana é similar a outras redes reais, sugerindo um processo natural de otimização durante sua criação.

Alguns aspectos dinâmicos também foram estudados através do processo de despercolação. Nós mostramos que as comunidades da rede Haversiana não são igualmente resistentes a perda de conexões. Além disso as comunidades mais resistentes estão localizadas na região posterior do osso, onde também estão localizados os prolongamentos ósseos, responsáveis pela sustentação e equilíbrio mecânico da estrutura. Também avaliamos o fluxo entre a medula óssea e o perióstio ósseo durante o processo de despercolação e o comparamos com o fluxo medido quando a rede é submetida a ataques. Nossos resultados indicam que a remoção aleatória de conexões é mais prejudicial às propriedades de transporte desta rede.

Abstract

This work studies in detail the channel system of the cortical bone of animals from the complex networks point of view. This system is composed by channels parallel to the main axis of the bone and by channels perpendicular to it. These channels are called Haversian and Volkmann channels, respectively. Their main function is to lead the blood vessels responsible for cell nourishment through the whole bone structure.

The set of channels was mapped into a network in which we considered each node as the confluence of one or more channels and the edges among the nodes as the channels. We analysed the topological features of the Haversian network using classic concepts of complex networks, such as degree, clustering coefficient, distribution of the shortest path and communities detection. We also used hierarquical approaches for some of these measurements.

We showed that the cortical bone network is highly modular and organized into communities very well defined and spatially localized. The latter feature was important in order to define the topological properties of the communities.

The results indicate that the Haversian network is similar to other man-made networks, suggesting a natural process of optimization during its creation.

Some dynamical aspects were also investigated through despercolation process. We showed that the Haversian communities are not equally resistant to edges removal. Moreover, the more resilient communities were found to be at the posterior portion of the bone, where are also placed the bone protrusions, responsible for support and balance of the structure. We also evaluated the flow between the marrow bone and the periosteum during the despercolation process and compared with the flow while the network was submitted to edges attacks. Our results indicate that the random removal of the edges is more harmful to the transport properties of this network.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	18
1.1	Motivação	20
1.2	Objetivos	20
1.3	Organização da Dissertação	20
2	Descrição Biológica	22
2.1	Introdução	22
2.2	O Sistema de Havers	25
3	Visualização 3D	27
3.1	Introdução e Importância	27
3.2	Procedimentos Utilizados neste Trabalho	28
4	Redes Complexas	32
4.1	Introdução	32
4.2	Nomenclatura e Conceitos Básicos	33
4.3	Os Modelos de Redes	36
4.3.1	Introdução	36
4.3.2	Redes Aleatórias	36
4.3.3	Redes Regulares	36

4.3.4	Redes Mundo Pequeno	37
4.3.5	Redes Livres de Escala	38
4.4	Caracterização Estática	40
4.4.1	Introdução	40
4.4.2	Medidas Clássicas	40
4.4.3	Medidas Hierárquicas	50
4.4.4	Fluxo Máximo	55
4.4.5	Comunidades	56
4.5	Caracterização Dinâmica	58
4.5.1	Introdução	58
4.5.2	Processos de Crescimento	59
4.5.3	Despercolação: A Vulnerabilidade das Redes Complexas	61
5	O Sistema Haversiano pelas Redes Complexas: Resultados e Discussões	66
5.1	O Modelo Ósseo Utilizado	66
5.2	Reconstrução 3D Computadorizada	67
5.3	Mapeamento da Estrutura Biológica em um Grafo	67
5.4	Medidas Clássicas	72
5.5	Medidas Hierárquicas	76
5.6	Comunidades	78
5.7	Robustez da rede Haversiana	85
6	Conclusão	90
	Referências Bibliográficas	93
7	Apêndice	100
7.1	Conectividade Máxima para as Redes Aleatórias	100

7.2	Diâmetro das Redes Aleatórias	101
7.3	O Algoritmo de Prim-Dijkstra	103
7.4	Distribuição de Conectividade Hierárquica no Modelo Barabási-Albert	106

Lista de Figuras

2.1	Representação da estrutura pulmonar humana evidenciando a ramificação na forma de árvore. Adaptado de http://www.bryanchristiedesign.com	23
2.2	Sistema Linfático do corpo humano. Adaptado de http://www.healingdaily.com	24
2.3	Rede neuronal gerada pela conexão de alguns neurônios. Adaptado de http://straddle3.net	24
2.4	Esquema da parede da diáfase dos ossos longos.	25
2.5	Visão superficial de alguns canais de havers e dos osteócitos dispostos na direção radial dos canais. No detalhe vemos um osteócito ampliado.	26
3.1	a) Imagem digitalizada de uma parcela de uma das lâminas histológicas e b) união das quatro imagens que compõem uma lâmina.	28
3.2	Exemplo de segmentação a) da região do canal medular e b) dos canais da matriz óssea.	29
3.3	Software utilizado para reconstrução e visualização dos dados utilizados neste trabalho. 1 - leitura da sequência de imagens, 2 - valor da isosuperfície que deve ser reconstruída, 3 - parâmetros da Suavização gaussiana, 4 - exportação em formato <i>VRML</i> , 5 - visualização do objeto renderizado e 6 - renderização das imagens que foram lidas.	30
3.4	a) exemplo da triangulação gerada pelo algoritmo de reconstrução do VTK, b) e c) representam a superfície (já suavizada) do canal medular.	31
4.1	a) Exemplo de grafo binário não direcionado; b) exemplo de grafo não-binário direcionado.	34

4.2	a) Rede de roteadores da internet (www.isi.edu/scan/scan.html), aqui os vértices representam roteadores e dois vértices estão conectados se existe uma conexão real entre os dois respectivos roteadores, b) rede de um circuito eletrônico (http://www.weizmann.ac.il/mcb/UriAlon/index.html) onde cada vértice representa um componente do circuito e c) rede de interação entre as proteínas da levedura do fermento <i>S. cerevisiae</i> (http://www.nd.edu/alb/). Considerando que os vértices sejam as proteínas, dois vértices estão conectados se as duas proteínas correspondentes apresentam afinidade de ligação	35
4.3	Exemplo de rede aleatória contendo 12 vértices e 15 conexões.	37
4.4	Exemplo de rede regular unidimensional contendo 12 vértices e 36 conexões. Percebemos a existência da condição de contorno periódica, já que o último vértice é vizinho do primeiro.	37
4.5	Exemplo de redes geradas pelo modelo de Watts e Strogatz com a) $p = 0$, b) $p = 0,15$ e c) $p = 1$	38
4.6	Exemplo de rede livre de escala com $N = 12$ e $E = 18$. Pela figura vemos que a maioria dos vértices estabelecem poucas conexões, enquanto poucos são altamente conectados.	39
4.7	Rede genérica com 30 vértices e 38 conexões. Em vermelho destacamos os vértices que pertencem ao menor caminho entre os vértices 17 e 27.	40
4.8	Coeficiente de clusterização médio do modelo mundo pequeno proposto por Watts e Strogatz [1] em função da probabilidade de redirecionamento p . A curva é normalizada pelo valor do coeficiente de clusterização médio da rede regular ($C(p = 0)$).	43
4.9	Distribuições de conectividade dos modelos de rede discutidos neste capítulo. a) rede aleatória de Erdős-Rényi ($N = 1000$ e $\langle k \rangle = 7$), b) rede regular finita ($N = 32 \times 32$ e $\langle k \rangle = 3,875$), c) mundo pequeno ($N = 32 \times 32$, $\langle k \rangle = 3,875$ e $p = 0,03$) e d) livre de escala ($N = 1000$, $\langle k \rangle = 7$ e $\gamma = 2$).	46
4.10	a) rede genérica com o vértice fonte destacado em amarelo, b) respectiva rede estrela e c) respectiva árvore de menor cobertura.	49

4.11	a) Rede genérica onde destacamos as hierarquias do vértice de referência $r = 1$, b) número hierárquico de vértices, c) conectividade hierárquica e d) coeficiente de clusterização hierárquico. Os quadrados vermelhos representam as medidas tomadas considerando o vértice 1 como referência. Para os círculos verdes consideramos o vértice 2 como referência. A faixa azul representa o desvio padrão e a média tomada sobre todos os vértices da rede.	51
4.12	Ilustração do comportamento da curva de conectividade hierárquica para redes de tamanho finito (preto) e infinito (vermelho).	55
4.13	O Algoritmo de fluxo máximo interpreta as conexões não direcionadas do sis- tema haversiano como duas conexões direcionadas em sentidos contrários e pode, possivelmente, realizar operações distintas nas duas direções.	57
4.14	Exemplo de rede com modularidade $Q \approx 0,93$ (equação 4.31). O particiona- mento avaliado contém três comunidades, como é indicado pelas cores dos vértices.	57
4.15	Componentes nos grafos aleatórios de Erdős-Rényi com a) $p = 0.03$, b) $p = 0.1$, c) $p = 0.2$ e d) $p = 0.4$	60
4.16	Dinâmica de anexação preferencial do modelo de Barabási-Albert com parâmetros $m = 1$ e $m = 2$. É interessante notar a ausência de circuitos fechados no caso $m = 1$	61
4.17	Caracterização de alguns modelos de redes complexas utilizando as seguin- tes medidas hierárquicas a) número hierárquico de vértices, b) conectividade hierárquica e c) coeficiente de clusterização hierárquico. A cor de cada curva está associada a cor da respectiva rede apresentada ao lado.	62
4.18	Simulação do processo de falhas em uma rede genérica. As falhas ocorrem pela remoção de a) vértices e b) conexões.	63
4.19	variação do tamanho da maior componente conectada em função de f durante o processo de despercolação. Todas as redes utilizadas possuem $N = 500$ vértices e são apresentadas na figura 4.17.	64
5.1	Ilustração da pata direita de uma galinha. Em destaque apresentamos a falange distal do primeiro dedo, de onde provém o osso utilizado neste trabalho.	66
5.2	Algumas perspectivas da reconstrução tridimensional da estrutura Haversiana.	68

5.3	Secção 205 do osso considerado nesse estudo. Percebemos que esta imagem é formada pela rotação, alinhamento e união de três imagens geradas por microscópio.	69
5.4	Exemplo de bifurcação de um canal. No processo de segmentação, cada trecho é marcado com uma cor.	70
5.5	A primeira seqüência mostra uma região da estrutura cortical onde pode ser visto um canal que estabelece uma conexão com o tecido. Na segunda seqüência mostramos a segmentação deste canal (azul) e do tecido (preto).	70
5.6	Esquema utilizado para mapear a estrutura Haversiana em um grafo. Além de ser utilizada para realizar a reconstrução tridimensional da rede de canais, a seqüência de imagens é processada por uma rotina externa que fornece a posição dos vértices (bifurcações). Em seguida estes vértices são conectados de acordo com as conexões presentes no arquivo de dados.	71
5.7	Representação da Haversiana. As perspectivas a, b, c e d desta imagem acompanham as da figura 5.2	71
5.8	a) Distribuição de conectividade da rede Haversiana e b) representação da correlação $P(k, k')$	73
5.9	a) Distribuição espacial dos valores de conectividade na rede Haversiana, b) distribuição espacial dos vértices com conectividade $k > 4$ e c) distribuição espacial dos valores do coeficiente de clusterização.	74
5.10	Distribuição do comprimento médio dos menores caminhos ao longo da rede haversiana. Também mostramos esta distribuição avaliada sobre a rede haversiana aleatória e sobre a árvore haversiana de menor cobertura. No inset apresentamos o mesmo gráfico em escala $\log \times \log$	74
5.11	O vértice fonte introduzido na rede conecta-se à todos os vértices que estabelecem conexão direta com a medula óssea.	75
5.12	Ilustração do caminho que liga o vértice fonte (0) à um vértice i qualquer. Em a) este caminho é retilíneo, implicando em um fator de rota $q_1 = 1$. Em b) o caminho é mais tortuoso, com fator de rota $q_1 = 2$	76

5.13	Propriedades hierárquicas da rede haversiana: número hierárquico de vértices e conectividade hierárquica. Em cinza mostramos a aproximação por campo médio para uma rede infinita com conectividade média similar a da rede haversiana. Também destacamos o valor de pico da rede haversiana: $h_{pico} = 32$	77
5.14	Coefficiente de clusterização hierárquico da rede haversiana. No inset apresentamos o mesmo resultado em escala linear e um ajuste através de um decaimento exponencial $e^{-\mu h}$, com $\mu \approx 0,15$	78
5.15	Evolução do valor da modularidade da rede Haversiana durante a execução do algoritmo de detecção de comunidades [2]. As linhas vermelhas marcam o valor adotado como modularidade da rede Haversiana neste trabalho.	80
5.16	Disposição espacial das comunidades da rede Haversiana.	81
5.17	Organização topológica das comunidades da rede Haversiana.	82
5.18	a) Curvas de conectividade hierárquica média das comunidades 1 (vermelho) e 10 (preto) e b) coeficiente de clusterização hierárquico médio das mesmas comunidades.	83
5.19	Organização das conexões estabelecidas entre as comunidades da rede Haversiana. Quando consideramos cada comunidade como um único vértice, podemos perceber que existe uma semelhança entre esta imagem e aquela apresentada na figura 4.5b. Este fato indica a existência do efeito mundo pequeno na rede Haversiana a nível das comunidades.	84
5.20	a) Rede obtida quando consideramos cada comunidade da rede Haversiana como um vértice. Dois vértices estão conectados se há ao menos uma conexão entre as comunidades correspondentes e b) distribuição do comprimento dos menores caminhos nesta rede. O valor médio da distribuição é $\langle l \rangle \approx 1,6$	85
5.21	Tamanho da maior componente conectada (escala $\log \times \log$) durante o processo de despercolação: i) na rede Haversiana (em azul) e ii) nas comunidades da rede Haversiana (linhas pretas). As comunidades estão agrupadas na legenda de acordo com a proximidade de suas curvas. Além disso, os índices correspondem aqueles da figura 5.16.	86
5.22	Fluxo máximo que escoar na rede Haversiana durante o processo de despercolação a) aleatório (quadrados vermelhos) e b) coordenado espacialmente (círculos pretos). Φ_m^* representa o máximo valor do fluxo máximo na rede Haversiana.	88

5.23	Evolução da rede Haversiana submetida ao processo de ataque coordenado espacialmente.	88
7.1	Procedimento realizado para obtenção da árvore de menor cobertura através do algoritmo Prim-Dijkstra. A figura A1 apresenta o grafo original com o vértice inicial na execução do algoritmo. A figura D8 mostra a AMC resultante.	105

Lista de Tabelas

- 5.1 Características gerais da rede haversiana e de outras redes biológicas. Na rede de interação proteína-próteína, cada vértice representa uma proteína e dois vértices são conectados se as proteínas correspondentes possuem a capacidade de ligar-se entre si. A rede da cadeia alimentar estudada por Montoya *et al.* foi gerada após estudos das relações alimentares existentes em determinados parques ecológicos. Os dados apresentados na tabela correspondem ao parque Silwood (localizado no *Colégio Imperial de Londres*). As redes metabólicas são redes biológicas onde os vértices representam componentes moleculares e, de acordo com [3], as ligações que apontam para um determinado componente representam as reações em que ele participa como produto. 72
- 5.2 Valores de q_1 (fator de rota) e q_2 (razão entre o comprimento total de conexões de uma rede e de sua respectiva AMC) para a rede Haversiana, bem como para outras rede reais estudadas por Gastner e Newman em [4]. A rede de esgoto é referente a cidade de Bellingham, EUA, a rede de gás WA refere-se a rede de gas natural da cidade de Western, Austrália. A rede de gás IL refere-se a rede de gás natual do estado de Illinois, EUA. A rede de ferrovias refere-se ao sistema de ferrovias da cidade de Boston, EUA. 76
- 5.3 Propriedades topologicas e espaciais das comunidades presentes na rede haver-siana. A segunda linha de cada comunidade corresponde a razão entre o valor medido na comunidade e o valor medida na rede completa. As quantidades tabeladas são: N - número de vértices, $\langle k \rangle$ - conectividade média, $\langle C \rangle$ - coeficiente de clusterização médio, $\langle l \rangle$ - diâmetro, l_{max} - comprimento do maior menor caminho e $\langle D \rangle$ - média da distância entre os vértices e a superfície do canal medular. 79

Notação

$\mathcal{V}$	Conjunto de vértices da rede
N	Número de vértices da rede (número de elementos do conjunto $\mathcal{V}$)
$\mathcal{E}$	Conjunto de conexões da rede
N	Número de conexões da rede (número de elementos do conjunto $\mathcal{E}$)
$u, v, m, \dots$	Vértices específicos da rede
r	Vértice de referência
$(u, v), (m, n), \dots$	Conexões específicas da rede
$\mathcal{A}$	Matriz de adjacência
$a_{u,v}$	Elemento da matriz de adjacência
$\mathcal{W}$	Matriz de pesos
$w_{u,v}$	Elemento da matriz de pesos
k_u	Conectividade do vértice u
$\check{k}$	Conectividade máxima da rede
$\langle k \rangle$	Conectividade média da rede
C_u	Coefficiente de clusterização do vértice u
$\langle C \rangle$	Coefficiente de clusterização médio da rede
e_u	Número de conexões entre os vizinhos do vértice u
$P(k)$	Probabilidade de que um vértice tenha conectividade k
$\mathcal{T}$	Conjunto de vértices que definem uma trajetória
x_i	i -ésimo vértice de uma trajetória
$l_{\mathcal{T}}$	Comprimento da trajetória $\mathcal{T}$
$c_{\mathcal{T}}$	Custo da trajetória $\mathcal{T}$
$\langle l \rangle$	Diâmetro da rede
$\mathcal{L}_{u,v}$	Menor caminho entre os vértices u e v
$\mathcal{C}_{u,v}$	Rota de menor custo entre os vértices u e v
AMC	Árvore de menor cobertura
q_1	Fator de rota
q_2	Razão entre o comprimento médio das conexões de uma rede e de sua respectiva AMC

$\mathcal{H}_r(h)$	Hierarquia h referente ao vértice r
$n_r(h)$	Número hierárquico de vértices referente ao vértice r
$k_r(h)$	Conectividade da hierarquia h referente ao vértice r
$\check{h}_r$	Hierarquia máxima do vértice r
$P_h(k)$	Probabilidade de que um vértice tenha conectividade k na hierarquia h
$C_r(h)$	Coeficiente de clusterização da hierarquia h referente ao vértice r
$e_r(h)$	Número de conexões entre os vértices da hierarquia h referente ao vértice r
$\omega_{u,v}$	Capacidade de transporte da conexão (u, v)
$s_{u,v}$	Fluxo na conexão (u, v)
Φ	Fluxo que escoar na rede
Q	Modularidade
N_C	Número de comunidades da rede
$\beta_{i,j}$	Fração de conexões que ligam vértices da comunidade i com vértices da comunidade j
α_i	Fração de conexões que ligam vértices da comunidade i com os demais vértices das outras comunidades
Z	Tamanho da maior componente
ℓ	Diâmetro da maior componente
$\langle z \rangle$	Tamanho médio das componentes menores

1 *Introdução*

A composição da estrutura esquelética dos seres vivos está em constante alteração ao longo do seu desenvolvimento. Muitas são as variáveis que contribuem para estas alterações, como a variação externa de pressão ao qual os ossos são submetidos e a demanda de cálcio no sangue [5]. As alterações que ocorrem nos ossos são responsáveis pela sua adaptação às mais diversas situações e ambientes. Este processo de adaptação, conhecido como *remodelagem óssea*, consiste na renovação de porções antigas (ou fraturadas) de osso. Células conhecidas como *osteoclastos* são responsáveis por degradar tecidos velhos ou danificados onde, em seguida ocorre uma nova deposição de cálcio pelos *osteoblastos* [6].

O osso é composto basicamente por dois tecidos principais conhecidos como, *osso trabecular* e *osso cortical*. O primeiro possui forma irregular, análoga à de uma esponja. No interior desta esponja encontra-se a medula óssea vermelha. A região cortical é compacta e permeada pelos *osteócitos*. Os osteócitos são células cerca de 10 vezes menores que os osteoclastos e são organizadas de maneira radial em torno de canais responsáveis por conduzir capilares sanguíneos provindos da região medular até essas células. Existem basicamente dois tipos de canais: os *canais de Havers* e os *canais de Volkmann*. Os primeiros são paralelos ao eixo principal do osso, e os segundos são perpendiculares ao eixo ósseo, sendo responsáveis pelas interconexões entre os canais de Havers. O conjunto de um canal de Havers e os osteócitos que o cerca é denominado *sistema Haversiano* e é considerado a unidade estrutural do osso cortical.

Como todo tecido ósseo, a região cortical e conseqüentemente, os canais de Havers e Volkmann, sofrem alterações durante o processo de remodelagem. A destruição dos canais próximos aos sítios de fratura ou porções antigas do osso, secundariamente à perda do suprimento vascular local, estimulam uma intensa regeneração de novos sistemas haversianos nestas regiões [7].

Ao longo dos últimos séculos, a estrutura óssea vem sendo exaustivamente estudada. A região cortical foi primeiramente descrita pelo fisiologista Clopton Havers, no final do século XVII [8]. Desde então, inúmeros trabalhos foram dedicados ao estudo desta região, bem como a estrutura dos canais que a irrigam [9, 10], sendo em sua maioria observações qualitativas. Muitos

outros trabalhos buscam explicações definitivas para o processo de remodelagem [6, 11, 12] e o entendimento dos fenômenos de transporte que ocorrem no interior dos ossos [13, 14, 15]. Entretanto, a relação entre a remodelagem óssea e sua influência no transporte de nutrientes através dos canais degredados por ela, tem sido pouco explorada.

O uso de ferramentas alternativas, como simulação por elementos finitos [16, 17] e visualização tridimensional [9, 10, 18], são de grande importância no estudo do sistema esquelético e têm permitido o avanço desta área de pesquisa. A proposta deste trabalho é contribuir com mais uma ferramenta para este estudo. Em particular, iremos investigar a conectividade dos canais responsáveis pela condução de nutrientes através osso. Pretendemos modelar estes canais como um grafo e estudá-lo utilizando conceitos de *Redes Complexas*.

Os trabalhos seminais de Watts et. al. [1] e R. Albert et al. [19] em 1999 abriram as portas para um novo campo da ciência, hoje conhecido como redes complexas. Este campo de pesquisa localiza-se na intersecção entre áreas científicas consolidadas há mais tempo, como a Mecânica Estatística [20, 21] e a Teoria dos Grafos [22]. Porém, muitas outras áreas estão envolvidas nos trabalhos em redes complexas, como técnicas de visualização computacional, estatística, armazenamento e tratamento de dados, entre outras. Devido a esta grande multidisciplinaridade, as redes complexas têm sido vistas por alguns pesquisadores como sendo particularmente importantes para a integração das ciências.

O formalismo das redes complexas é útil para descrever problemas que possuem um grande número de agentes que interagem através de determinadas regras (simples, na maioria das vezes). Nesta descrição, não nos preocupamos com a evolução de um agente específico e sim com a dinâmica que emerge do comportamento coletivo destes agentes. Para representar estes sistemas usamos uma estrutura matemática chamada *Grafo*. Basicamente, um grafo é representado por um conjunto de vértices e um conjunto de conexões que estabelecem relações entre os vértices.

Sistemas como a WWW [23], Internet [24], colaboração entre atores [25], colaboração entre pesquisadores [26], rede de citações científicas [27], rede de contatos sexuais [28], rede de tráfego aéreo [29] e sistemas de ruas urbanas [30], são alguns exemplos de sistemas nos quais esta abordagem tem sido aplicada com sucesso. As redes complexas também estão presentes em vários sistemas biológicos como, interação entre proteínas [31], redes metabólicas [32], redes neurais [33] e cadeias alimentares [34]. Aqui, pretendemos estender esta abordagem ao estudo do sistema de canais interconectados do osso cortical. Para isso, iremos considerar que a confluência de dois canais é mapeada em um vértice e os canais são as próprias conexões entre

os vértices do grafo.

1.1 Motivação

O osso fornece sustentação interna para todo o corpo, além de servir de base para os músculos e tendões. O conhecimento dos processos de crescimento e remodelagem, bem como da organização estrutural é muito importante em aplicações terapêuticas e clínicas. Mesmo sendo tão importante e consequentemente, sendo alvo de intensa pesquisa, muitas questões acerca do sistema esquelético permanecem em aberto como a relação entre a obstrução aleatória de canais durante o processo de remodelagem. Outras questões foram descritas somente de maneira qualitativa, como é o caso da organização topológica dos canais de Havers e de Volkmann.

Acreditamos que informações quantitativas sobre esses aspectos são importantes para aprimorar nossos conhecimentos acerca desta estrutura tão complexa. Essas informações também podem ajudar a melhorar a qualidade das estruturas artificiais criadas para uso em implantes de tecido [35].

Entendemos a área das redes complexas como sendo ideal para a descrição do sistema de canais do osso cortical por tratar-se de um mapeamento direto, favorecido pela estrutura dos canais. Além disso, ela é uma ferramenta poderosa que nos permitirá investigar diferentes aspectos desta estrutura.

1.2 Objetivos

O principal objetivo deste trabalho de mestrado é aplicar os conceitos e métodos das redes complexas na caracterização e análise no sistema de canais interconectados do osso cortical. Com isso, pretendemos responder algumas questões sobre a organização topológica desta estrutura de maneira quantitativa e estabelecer alguma relação dinâmica entre a obstrução aleatória de canais (provocada pelo processo de remodelagem) e a condução de nutrientes através da matriz óssea.

1.3 Organização da Dissertação

No próximo capítulo será apresentada uma breve discussão sobre os aspectos biológicos envolvidos neste trabalho. No capítulo 4 veremos alguns tópicos sobre visualização tridimensional

focado no que foi utilizado neste trabalho. No capítulo 5 introduzimos os conceitos de redes complexas utilizados aqui. No capítulo 6 apresentamos e discutimos os resultados obtidos neste trabalho e por fim, no capítulo 7 descrevemos as conclusões finais, as contribuições e os trabalhos futuros.

2 *Descrição Biológica*

2.1 Introdução

Vivemos em um mundo cercado por redes de diferentes tipos. Desde os circuitos eletrônicos [36] até as relações sociais [27, 28] podem ser descritas através do conceito de rede. O corpo humano não é exceção. Ele possui muitos exemplos de redes de transporte, tanto a nível macroscópico quanto a nível microscópico. Cada um deles possui uma topologia altamente adaptada às suas funções e a região de operação. Por exemplo, o *Sistema Circulatório* sanguíneo, constituído pelo conjunto de veias, artérias e capilares do nosso corpo, é tipicamente estruturado na forma de uma árvore. Os diâmetros típicos do sistema circulatórios vão de $10^4 \mu\text{m}$ para as veias e artérias mais espessas até $10^1 \mu\text{m}$ para os mais finos capilares.

Esta organização topológica também está presente na *Árvore Pulmonar*, cujos dutos conduzem oxigênio no interior dos pulmões. A figura 2.1 mostra uma ilustração deste sistema. O diâmetro dos dutos é da ordem de $10^2 \mu\text{m}$ e tornam-se menores a medida que a distância ao duto central aumenta [37]. Em [38], os autores mostram que esta dependência entre o diâmetro dos dutos e a distância ao o duto central segue uma relação linear.

Estudos recentes mostraram que as bifurcações existentes nos dois sistemas descritos acima são governadas por regras alométricas de escala [39] que também são encontradas na vascularização das plantas [40, 41, 42]. Estas regras vinculam o diâmetro dos canais filhos com o diâmetro do canal pai através da lei de Murray [43]

$$D^3 = d_1^3 + d_1^3 + \dots + d_n^3 \quad (2.1)$$

onde n é o número de ramos originados a partir do canal pai que possui diâmetro D . Em [42] os autores verificaram que as redes que seguem este tipo de regra são otimizadas para maximizar a condutância hidráulica¹.

¹ A condutância hidráulica é definida como a quantidade de material que passa através de um canal por unidade

Figura 2.1: Representação da estrutura pulmonar humana evidenciando a ramificação na forma de árvore. Adaptado de <http://www.bryanchristiedesign.com>.

Outro exemplo de rede de transporte é o *Sistema Linfático*. Ele é responsável por retornar o plasma sanguíneo disperso nos tecidos (linfa) ao sistema circulatório. Este sistema é constituído por capilares e vasos linfáticos que abrangem todo o corpo humano. Como pode ser visto na ilustração da figura 2.2, este sistema difere dos anteriores, pois possui um grande número de ciclos.

Do ponto de vista das conexões, a rede mais complexa do corpo humano talvez seja a aquela pertencente ao *Sistema Nervoso*. A rede neuronal existente no cérebro possui uma topologia completamente não trivial que vem sendo investigada por inúmeros pesquisadores no último século [44, 45, 46]. Na figura 2.3 mostramos uma ilustração da rede formada pelas conexões de alguns neurônios. A região responsável por conduzir os impulsos elétricos aos outros neurônios é chamada axônio e seu diâmetro médio é da ordem de $10^1 \mu\text{m}$.

Como podemos perceber, praticamente todo corpo humano está envolto por redes de transporte que garantem a manutenção da vida. No interior dos ossos não é diferente. Lá existe um sistema de canais interligados responsável pela distribuição do sangue por toda matriz óssea chamado *Sistema de Havers*.

de tempo e por unidade de diferença de pressão.

Figura 2.2: Sistema Linfático do corpo humano. Adaptado de <http://www.healingdaily.com>.

Figura 2.3: Rede neuronal gerada pela conexão de alguns neurônios. Adaptado de <http://straddle3.net>.

Figura 2.4: Esquema da parede da diáfase dos ossos longos.

2.2 O Sistema de Havers

Os ossos não são estruturas completamente sólidas. Sua parte mais externa, denominada *Osso Cortical*, consiste em uma fina camada sólida com pequenos canais. Seu interior, no entanto, apresenta um aspecto esponjoso, onde se localiza a medula óssea vermelha. Denominamos esta região de *Osso Trabecular* (figura 2.4).

A unidade estrutural do sistema ósseo é chamada *Sistema de Havers* (ou ósteon), mas ela não é encontrada em todos os tipos de ossos, ocorrendo apenas na diáfase dos ossos longos. No centro do sistema de Havers encontra-se um canal, designado canal de Havers. Na região do osso cortical, identificam-se facilmente estes canais, que ocorrem paralelamente ao canal medular. Os canais de Havers são unidos por canais radiais, os canais de Volkman 2.4. Esta estrutura de canais não ocorre no osso trabecular, devido a sua forma não regular [7, 47]. O nome dos canais foi dado em homenagem ao anatomista Clopton Havers graças aos seus estudos pioneiros sobre a estrutura do osso cortical por volta de 1691 [8].

As substâncias responsáveis pela nutrição dos osteócitos (células da matriz óssea) passam por uma hierarquia de estruturas anatomicamente muito bem definidas [13]. Os vasos sangüíneos alcançam o canal medular e ramificam-se. Cada nova ramificação penetra na matriz óssea

Figura 2.5: Visão superficial de alguns canais de havers e dos osteócitos dispostos na direção radial dos canais. No detalhe vemos um osteócito ampliado.

através dos canais de Havers e de Volkman. Em seguida, através de difusão molecular, os nutrientes alcançam as células da matriz óssea, veja figura 2.5.

Este processo de difusão que conduz os nutrientes dos canais de Havers e Volkman até os osteócitos ainda não é completamente conhecido e, conseqüentemente, é alvo de discussões entre diversos autores [13, 14, 15]. Além disso, atualmente sabemos que a difusão não é o principal modo pelo qual os nutrientes chegam até as células. Este fato foi observado pela primeira vez em 1967 por Piekarski e Munro [48]. Os autores mostraram que as variações externas de pressão na estrutura óssea alteram o escoamento nos canalículos que ligam os canais de Havers aos osteócitos. Esta alteração é capaz de gerar um fluxo cuja ordem é de 3 a 4 vezes maior do que aquele gerada pelo processo de difusão [49, 15].

Em [50] os autores propuseram um modelo matemático para estudar o fluxo induzido por variações externas de pressão na estrutura óssea. Entre outras coisas, eles investigaram os efeitos do diâmetro dos canalículos na intensidade deste fluxo. Segundo os autores, existe um diâmetro ótimo para produção de um fluxo máximo, capaz de alimentar 4 ou 5 conjuntos circulares de osteócitos (veja figura 2.5). Outro resultado interessante desses autores mostra que a falta de fluxo em porções locais da matriz óssea (devido a obstrução dos canalículos, por exemplo) pode dar início ao processo de osteoporose.

Como todos os tecidos do corpo humano, o sistema ósseo está em constante alteração. Graças aos processos de crescimento e remodelagem do osso, os canais de Havers e Volkman sofrem alterações ocasionais (obstruções temporárias e até mesmo desaparecimento completo) [7, 47]. O crescimento do sistema ósseo vem sendo amplamente estudado, porém explicações definitivas para estes processos ainda não existem. Um dos motivos é a dificuldade de estudar a evolução temporal deste tipo de sistema microscópico rígido no qual experimentos não invasivos são praticamente impossíveis.

3 *Visualização 3D*

Estamos em uma era no qual a aquisição de dados 3D é massiva e continuada. Isso ocorre graças aos avanços tecnológicos que permitem automatizar quase que totalmente a maioria dos processos de aquisição. Porém, para tornar os dados realmente úteis, devemos empregar modelos que os tornem passíveis à análise interativa [51]. Por isso, a visualização interativa tornou-se o principal método de exploração de coleções massivas de dados [51]. Neste capítulo daremos uma breve introdução a visualização científica com enfoque especial nos aspectos utilizados neste trabalho.

3.1 **Introdução e Importância**

Atualmente, existe uma diversidade de métodos para captura dos dados. Na área das ciências biológicas, as técnicas mais comuns são raio-X, tomografia computadorizada, imageamento por ressonância magnética e seccionamento histológico. As três primeiras são técnicas não invasivas, capazes de amostrar as informações anatômicas de uma determinada estrutura interna. Os dados amostrados em geral, consistem em um conjunto de pontos no espaço tridimensional. O próximo passo é realizado por algoritmos especializados capazes de converter este conjunto de pontos em uma estrutura tridimensional.

Já o seccionamento histológico é uma técnica invasiva que consiste em seqüenciar a estrutura de interesse em "fatias". Em seguida, as imagens são digitalizadas e processadas. A fase de processamento é constituída por um conjunto de procedimentos que possuem o objetivo de padronizar as imagens e selecionar a região que deseja-se estudar. Dentre estes procedimentos destacamos, dentre outros, alinhamento, rotação, suavização e segmentação [52]. Em seguida, as imagens digitais são empilhadas por algoritmos especializados e a estrutura tridimensional da região de interesse é recuperada.

As avançadas técnicas de visualização fornecem imagens tridimensionais de alta resolução, representando, com clareza, o formato de uma determinada estrutura. Na área médica esses

Figura 3.1: a) Imagem digitalizada de uma parcela de uma das lâminas histológicas e b) união das quatro imagens que compõem uma lâmina.

procedimentos já permitem o planejamento terapêutico e cirúrgico das mais variadas doenças. Também permitem o desenvolvimento de modelos que descrevem a dinâmica do sistema cardiovascular e do sistema respiratório, o crescimento de tumores e o transporte, difusão e absorção de fármacos. Ultimamente, procedimentos baseados em visualização 3D também têm sido amplamente utilizados no treinamento de especialistas [53].

Na próxima secção daremos uma breve introdução aos procedimentos utilizados para a visualização dos dados presentes neste trabalho.

3.2 Procedimentos Utilizados neste Trabalho

Histologia

Os dados utilizados neste trabalho foram obtidos através de procedimentos histológicos realizados no *Hospital de Veterinária da Universidade Federal de Uberlândia*. A histologia preocupa-se com o estudo de tecidos seccionados por de um micrótomo em fatias finíssimas. Após diversas etapas, o resultado do processo histológico é uma coleção de lâminas histológicas permanentes.

Digitalização

O próximo passo foi digitalizar as lâminas utilizando um microscópio óptico. Um exemplo da imagem resultante é apresentada na figura 3.1-a. Até este ponto, todo o trabalho foi desenvolvido no Setor de Patologia e Veterinária do Hospital da Universidade Federal de Uberlândia, Brasil. Cada lâmina histológica originou quatro imagens digitais.

Figura 3.2: Exemplo de segmentação a) da região do canal medular e b) dos canais da matriz óssea.

Processamento

As quatro imagens de cada lâmina passaram pela fase de processamento, em que foram padronizadas e unidas, formando toda a estrutura presente na lâmina. Um exemplo de imagem após este processo por ser visto na figura 3.1-b. Em seguida realizamos a segmentação das regiões de interesse, que eram duas no nosso caso. A primeira é a região mais interna, que representa o canal medular e o segundo são os canais embebidos na matriz óssea. Exemplos da segmentação dessas estruturas podem ser vistos figura 3.2.

Reconstrução e Visualização

Realizamos a reconstrução 3D por meio do Visualization Toolkit - VTK [54]. O VTK é uma biblioteca implementada em C++ para programação orientada a objetos, que permite trabalhar com computação gráfica 3D e processamento de imagens. O VTK inclui uma grande diversidade de estruturas de dados, como: *unorganized point set*, *polygonal data*, *images*, *volumes* e *unstructured grids*. A biblioteca também inclui filtros para processamento dos dados, que podem ser exportados ou importados com auxílio de suas funções *Writers* e *Readers*, respectivamente. Existem vários softwares para visualização 3D, o AVS (www.avs.com) foi o primeiro software comercial com este propósito. O IBM Data Explorer <http://www.research.ibm.com/dx>, que era originalmente comercial, é mais uma das opções *open-source* hoje em dia. Outros softwares comerciais com esta finalidade são: NAG Explorer (<http://www.nag.com>) e o Template Graphics Amira (<http://www.tgs.com>).

Por ser *open-source* e pela excelente documentação e suporte, o VTK é adotado nas mais diversas áreas de pesquisa. Atualmente instituições como: Rensselaer Polytechnic Institute, State University of New York at Stony Brook, Stanford e o Brigham and Women's Hospital usam

Figura 3.3: Software utilizado para reconstrução e visualização dos dados utilizados neste trabalho. 1 - leitura da sequência de imagens, 2 - valor da isosuperfície que deve ser reconstruída, 3 - parâmetros da Suavização gaussiana, 4 - exportação em formato *VRML*, 5 - visualização do objeto renderizado e 6 - renderização das imagens que foram lidas.

VTK como ferramenta de pesquisa e diagnóstico [55].

Para trabalharmos com o VTK uma interface de acesso às suas funções foi desenvolvida em ambiente C++ Builder e está apresentada na figura 3.3. A interface nos permite ler a sequência de imagens segmentadas em formato *PNG*. Então, algoritmos de reconstrução encarregam-se de criar a triangulação da superfície que desejamos visualizar. Existem basicamente duas técnicas para visualização de dados, sendo que uma delas é exatamente a visualização de superfícies. A outra é a visualização volumétrica, que não será abordada aqui.

Se imaginarmos que todos os pontos do espaço no qual a imagem deverá ser apresentada possuem um valor escalar associado, uma superfície pode ser criada, conectando todos os pontos que possuem o mesmo valor. Este valor é conhecido como *valor de isosuperfície*. Na figura 3.4 mostramos a triangulação gerada para o canal medular e duas perspectivas de reconstrução da superfície deste canal. As superfícies apresentadas foram suavizadas através da rotina *vtkImageGaussianSmooth* [54], que nos permite o controle do índice de suavização e do raio efetivo de suavização. Neste caso utilizamos o valor 3 para o primeiro e 2 para o segundo parâmetro.

Figura 3.4: a) exemplo da triangulação gerada pelo algoritmo de reconstrução do VTK, b) e c) representam a superfície (já suavizada) do canal medular.

4 *Redes Complexas*

4.1 Introdução

Sistemas complexos são estudados há muito tempo por físicos e matemáticos. Em geral, estes sistemas possuem grande número de elementos fundamentais, cuja constituição é simples e o comportamento coletivo é não trivial. Henri Poincaré foi o primeiro a mostrar que sistemas relativamente simples podem ter um comportamento evolutivo indeterminado. Poincaré percebeu este fato após notar que é matematicamente impossível encontrar a trajetória de três planetas interagindo gravitacionalmente entre si. Com o passar do tempo surgiram novas áreas de pesquisa dentro da física dos sistemas complexos: sistemas não lineares, redes neurais, caos e sistemas auto-organizados. A última avalanche de trabalhos ocorreu na década de 90 quando pesquisadores perceberam que muitas estruturas naturais são auto similares como os fractais, que haviam sido descobertos por Benoît Mandelbrot na década de 70.

As atenções voltadas à ciência dos sistemas complexos começaram a mudar de foco a partir de 1999, quando Albert, Jeong e Barábasi realizaram uma descoberta surpreendente, baseada em um trabalho de 1967 do psicólogo Stanley Milgram [56].

O experimento de Milgram consistia em distribuir aleatoriamente cartas entre cidadãos americanos. Cada carta continha o nome de uma pessoa alvo. Se a pessoa que estivesse com a carta não conhecesse o destinatário, ela deveria entregar a carta a um amigo que teria chance maior de conhecê-lo. Ao repassar a carta, a pessoa deveria colocar seu nome na mesma. O destinatário era responsável por reenviar a carta ao responsável pelo estudo. Curiosamente, Milgram notou que, em média, cada carta passava por seis pessoas antes de chegar ao destinatário. Tal resultado ficou conhecido como *seis graus de separação* ou *efeito mundo-pequeno* e mostra que pessoas aparentemente sem nenhuma relação têm uma grande possibilidade de possuírem, em algum grau, amigos em comum que as aproxime.

Este estudo foi aplicado à rede mundial de computadores - WWW por Albert, Jeong e Barábasi [23]. Esta rede é formada por páginas WEB interligadas por hiperlinks que tornam possível a

navegação entre elas. O estudo procurou encontrar o número médio de hiperlinks que separavam duas páginas escolhidas aleatoriamente nesta rede. O resultados encontrado mostra que, em média, quaisquer duas páginas da internet estão separadas por pouco menos de 20 cliques do seu *mouse* e assim o efeito mundo pequeno também está presente neste sistema.

Os trabalhos seguintes representam as primeiras tentativas de modelagem da internet. Para isso os pesquisadores passaram a utilizar como ferramenta os grafos. Mesmo sendo conhecida desde Leonard Euler no início do século XVIII, a teoria dos grafos só foi desenvolvida mais profundamente no século XX.

4.2 Nomenclatura e Conceitos Básicos

Um grafo $G(\mathcal{V}, \mathcal{E})$ pode ser definido como um conjunto de vértices $\mathcal{V}$ e um conjunto de conexões $\mathcal{E}$. Em alguns casos, cada vértice do conjunto $\mathcal{V}$ está associado a uma posição no espaço euclidiano, nestes casos, dizemos que o grafo é *geográfico*. Cada elemento do conjunto $\mathcal{E}$ associa dois elementos do conjunto $\mathcal{V}$, assim, se $(u, v) \in \mathcal{E}$ então existe uma conexão entre o vértice u e o vértice v do grafo. Chamamos de grafos *direcionados* aqueles que fazem distinção entre os elementos (u, v) e (v, u) , isto é, a conexão que vai de u para v não é a mesma que vai no sentido contrário. Quando esta restrição não é imposta dizemos que os grafos são *não-direcionados*. Conexões direcionadas são geralmente representadas por setas indicando a direção da ligação, enquanto conexões não direcionadas são representadas por uma linha simples.

A maioria dos grafos reais não são binários, ou seja, as conexões não podem ser representados simplesmente por 1 (existe) ou 0 (não-existe), nestes casos devemos introduzir uma nova quantidade que informa a intensidade da interação entre dois vértices. Com esse objetivo, definimos o peso da conexão entre o vértice u e o vértice v como $w_{u,v}$. Assumimos que não existe conexão entre u e v quando $w_{u,v} = 0$.

Por vezes, é muito útil representar um grafo por uma matriz, que recebe o nome de *matriz de adjacência*. No caso dos grafos binários, iremos denotar a matriz de adjacência por $\mathcal{A}$ (com elementos $a_{u,v}$), enquanto a notação usada para os grafos não binários será $\mathcal{W}$ (com elementos $w_{u,v}$). O elemento $a_{u,v}$ é igual a unidade se existe um conexão entre os vértices u e v , caso contrário seu valor é nulo. O mesmo ocorre para os grafos não-binários porém, neste caso o valor de cada elemento da matriz de adjacência é igual a intensidade da conexão entre u e v , ou seja $w_{u,v}$.

Por exemplo, o grafo binário e não direcionado apresentado na figura figura 4.1-a pode ser representado pela seguinte matriz de adjacência:

Figura 4.1: a) Exemplo de grafo binário não direcionado; b) exemplo de grafo não-binário direcionado.

$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad (4.1)$$

É interessante notar que, pelo fato do grafo ser não direcionado, sua matriz de adjacência é simétrica, ou seja $a_{u,v} = a_{v,u}$. A representação matricial do grafo apresentado na figura 4.1-b é

$$\mathcal{W} = \begin{pmatrix} 0 & 0,8 & 0 & 0 \\ 0 & 0 & 1,5 & 0,3 \\ 0 & 0 & 0 & 0,4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.2)$$

Vemos que neste caso a simetria não está presente na matriz de adjacência, já que o grafo é do tipo direcionado.

O formalismo dos grafos e a utilização da sua notação matricial são ferramentas poderosas para o estudo de sistemas complexos como a internet, interação entre as proteínas [31, 57], relacionamento humano [58, 28, 59, 60], sistemas de transporte aéreo [61, 29, 62], entre outros. Por exemplo, a rede WWW pode ser facilmente escrita em termos de grafos supondo-se que cada vértice representa uma página e que os hiperlinks representam as conexões entre os vértices.

No decorrer deste trabalho, vamos nos referir aos grafos por redes, já que esta terminologia é mais comum entre os físicos.

Figura 4.2: a) Rede de roteadores da internet (www.isi.edu/scan/scan.html), aqui os vértices representam roteadores e dois vértices estão conectados se existe uma conexão real entre os dois respectivos roteadores, b) rede de um circuito eletrônico (<http://www.weizmann.ac.il/mcb/UriAlon/index.html>) onde cada vértice representa um componente do circuito e c) rede de interação entre as proteínas da levedura do fermento *S. cerevisiae* (<http://www.nd.edu/alb/>). Considerando que os vértices sejam as proteínas, dois vértices estão conectados se as duas proteínas correspondentes apresentam afinidade de ligação

4.3 Os Modelos de Redes

4.3.1 Introdução

Utilizamos o termo *redes complexas* [19, 63, 64] para nos referir aos grafos que possuem organizações topológicas específicas e estão relacionados a um sistema físico real. Alguns exemplos de redes complexas estão apresentados na figura 4.2, bem como suas descrições e os sistemas físicos que representam. Também devemos destacar a importância das redes complexas provindas de modelos matemáticos que buscam reproduzir os aspectos quantitativos que emergem das redes complexas reais. De fato, ao longo da última década os novos modelos propostos nos tornaram capazes de modelar os mecanismos de crescimento de muitos sistemas complexos, como a Internet [65], por exemplo.

As próximas seções deste capítulo serão devotadas a apresentação dos modelos clássicos de redes, bem como sua origem. Em seguida, apresentaremos os conceitos utilizados para caracterizar cada tipo de rede tanto do ponto de vista estático, quanto dinâmico.

4.3.2 Redes Aleatórias

Os trabalhos de P. Erdős e A. Rényi [66] em meados de 1950 constituem uma das bases da teoria moderna das redes complexas. Nesses trabalhos, os autores estudaram, do ponto de vista matemático, um sistema formado por E conexões distribuídas aleatoriamente entre N vértices. Mesmo sendo o primeiro modelo amplamente estudado, o mesmo não possui muitas aplicações relevantes já que poucos sistemas possuem tal grau de aleatoriedade. Anatol Rapoport propôs [67] a utilização das redes aleatórias para explicar o efeito de mundo pequeno encontrado no experimento de Milgram. Com este modelo ele conseguiu obter o valor 6 para o número de pessoas que separam dois indivíduos escolhidos ao acaso, entretanto o modelo não leva em conta o fato de que as pessoas possuem um núcleo local de amizades e não distribuídas de maneira aleatória por toda população. Na figura 4.3 mostramos um exemplo de rede aleatória com $E = 20$ e $N = 10$.

4.3.3 Redes Regulares

Este é o modelo conhecido pelos físicos há mais tempo. Geralmente, essas redes são geográficas, ou seja, os vértices possuem uma posição bem definida no espaço Euclidiano. Além disso, as conexões são distribuídas somente entre os vizinhos topológicos de cada vértice. Na Física do estado sólido, a rede regular é uma ferramenta indispensável para representação dos

Figura 4.3: Exemplo de rede aleatória contendo 12 vértices e 15 conexões.

Figura 4.4: Exemplo de rede regular unidimensional contendo 12 vértices e 36 conexões. Percebemos a existência da condição de contorno periódica, já que o último vértice é vizinho do primeiro.

átomos em uma estrutura e suas interações locais [68]. Na figura 4.4 temos um exemplo de rede regular unidimensional com condições de contorno periódica onde cada um dos $N = 12$ vértices está conectado aos seus primeiros vizinhos topológicos.

4.3.4 Redes Mundo Pequeno

Este foi o modelo de rede proposto por D. J. Watts e S. Strogatz [1] para explicar o efeito mundo pequeno emergente no experimento de Milgram. Ele incorpora a aleatoriedade e a localidade dos dois modelos discutidos anteriormente de uma maneira que o torna completamente diferente de ambos. O modelo consiste basicamente em reorientar todas as conexões de uma rede regular com uma probabilidade p . Quando $p = 0$ nada se altera e mantemos a topologia regular. Quando $p = 1$, todas as conexões são alteradas e a rede torna-se aleatória. Para os valores de p contidos nesse intervalo a rede assume a topologia mundo pequeno¹.

Este modelo é capaz de capturar as principais características do sistema estudado por Milgram,

¹Na verdade, a topologia mundo pequeno é alcançada para valores muito pequenos de p . Quando 20% ($p = 0.2$) das conexões são reorientadas a rede já pode ser considerada aleatória [1].

Figura 4.5: Exemplo de redes geradas pelo modelo de Watts e Strogatz com a) $p = 0$, b) $p = 0,15$ e c) $p = 1$.

já que possui alta performance de comunicação entre pessoas distantes, graças as poucas conexões aleatórias e ainda mantém certo grau de regularidade, representando os núcleos locais de amizade.

4.3.5 Redes Livres de Escala

Graças aos avanços computacionais nas técnicas de aquisição e armazenamento de dados de redes reais, os pesquisadores perceberam que grande parte dos sistemas físicos não se adequava a nenhum dos modelos estudados anteriormente. Do ponto de vista topológico, dois fatores tornavam as redes reais incapazes de serem investigadas através dos modelos existentes. São eles: i) as redes reais estão em constante evolução, com vértices e conexões entrando e saindo do sistema a todo instante; ii) os novos vértices tendem a se conectar com vértices que possuem grande número de conexões. Este segundo fator ficou conhecido na literatura como *dinâmica de anexação preferencial*.

Em 1998 Watts e Strogatz mostraram que a rede neural do verme *Caenorhabditis elegans*, a rede de energia elétrica do oeste do Estados Unidos e a rede de colaboração de atores exibiam o efeito de mundo pequeno [1]. Curiosamente este efeito também foi verificado em 1999 na rede mundial de computadores (WWW) por R. Albert *et. al.* [23]. Neste mesmo ano A.-L. Barabási e R. Albert reportaram um novo trabalho [19] onde exploravam uma nova característica topológica das redes reais estudadas por Watts e Strogatz e da rede WWW. Eles perceberam que a grande maioria dos vértices dessas redes realizava muito poucas conexões, enquanto um número muito pequeno de vértices conectava-se a muitos outros vértices. Este fato já havia sido observado por Redner [27] ao estudar uma rede de citações científicas, em que cada vértice representava um artigo científico e o autor considerava que dois artigos estavam conectados se um deles fizesse referência ao outro.

Matematicamente falando, o que esses autores perceberam é que a distribuição de conectividade

Figura 4.6: Exemplo de rede livre de escala com $N = 12$ e $E = 18$. Pela figura vemos que a maioria dos vértices estabelecem poucas conexões, enquanto poucos são altamente conectados.

das redes reais obedece a uma lei de potência, caracterizada por um decaimento exponencial. Ao contrário, a distribuição de conectividade das redes aleatórias segue a distribuição de Poisson, caracterizada por um pico bem definido. Nas redes regulares infinitas a distribuição de conectividade é dada pela função delta de Dirac, pois todos os vértices realizam o mesmo número de conexões. Isso demonstra que a presença de vértices com intensa conectividade (*hubs*) nas redes aleatórias, regulares e mundo pequeno (já que esta é a interpolação das duas primeiras) é praticamente nula e que, como foi comprovado, este não é o caso da maior parte das redes reais.

As redes cuja distribuição de conectividade obedece a uma lei de potência, ficaram conhecidas como *redes livres de escala*. Este nome faz referência a ausência de um pico na distribuição de conectividade, indicando uma escala característica ao sistema (como ocorre nos grafos aleatórios). Na figura 4.6 mostramos um exemplo de rede livre de escala com $N = 12$ e $E = 18$.

No mesmo trabalho em que divulgam sua descoberta [19], A.-L. Barabási e R. Albert propõem um modelo para a evolução de redes livres de escala. O modelo, hoje conhecido como *Barabási-Albert* concorda muito bem com os dados empíricos apresentados no trabalho. Em [69] os mesmos autores, com a colaboração de H. Jeong, estudaram as principais características do modelo do ponto de vista da teoria de campo médio. A solução exata do modelo Barabási-Albert foi apresentada pela primeira vez em 2000 por Dorogovtsev, Mendes e Samukhin [70]. Desde então, formas não lineares [71, 72, 73] e muitas outras variações deste modelo vêm sendo propostas, estudadas e aplicadas com grande sucesso na descrição de redes complexas [74, 75].

O restante deste capítulo será devotado à introdução das medidas clássicas usadas em redes complexas, bem como à caracterização dos diferentes modelos de redes, tanto do ponto de vista estático quanto dinâmico.

Figura 4.7: Rede genérica com 30 vértices e 38 conexões. Em vermelho destacamos os vértices que pertencem ao menor caminho entre os vértices 17 e 27.

4.4 Caracterização Estática

4.4.1 Introdução

As medidas que não incluem nenhum tipo de evolução temporal são as mais comuns na área das redes complexas e fornecem informações valiosas a respeito de suas características topológicas. Abaixo discutimos o comportamento dessas grandezas nos modelos apresentados anteriormente.

4.4.2 Medidas Clássicas

Conectividade

A característica mais fundamental de cada vértice é sua conectividade. Denotada por k_i , ela indica o número de conexões estabelecidas entre o vértice i e os demais. Na rede da figura 4.7, por exemplo, podemos observar que $k_1 = 1$, $k_5 = 2$ e $k_{22} = 5$. Os k_i vértices que conectam-se ao vértice i , são comumente chamados de *vizinhos* do vértice i . Quando estamos trabalhando com redes não direcionadas a conectividade está vinculada à matriz de adjacência da seguinte maneira:

$$k_u = \sum_{i=1}^N a_{u,i} \quad (4.3)$$

Denotamos por $\langle k \rangle$ a média da conectividade sobre todos os vértices da rede. Nas redes não direcionadas, onde cada conexão contribui com o grau de dois vértices, o valor desta quantidade é

$$\langle k \rangle = \frac{2E}{N} \quad (4.4)$$

Quando desconsideramos a possibilidade de um vértice conectar-se a ele mesmo (auto-loop) e de não conectar-se à qualquer outro (vértice isolado), sua conectividade fica limitada ao intervalo $k_i \in [1, N - 1]$. Além disso, a maioria das redes que já foram estudadas têm em comum o fato de que $\langle k \rangle \ll \check{k}$ (onde $\check{k} = N - 1$). Estas redes são ditas esparsas. Por exemplo, a rede de energia elétrica do oeste norte-americano estudada por Watts e Strogatz [1] possui $N = 4941$ vértices e conectividade média $\langle k \rangle = 2,67$, mostrando que, em média, cada estação de distribuição de energia elétrica conecta-se diretamente com 2,67 outras estações.

Existem dois valores de conectividade que são particularmente importantes na teoria das redes complexas. O primeiro deles é denotado por $\check{k}$ e chamado de *conectividade máxima possível* [76]. Esta quantidade nos informa o máximo valor de conectividade que um vértice pode ter em uma rede com N vértices. Se consideramos que não existem auto-conexões (conexão de um vértice com ele mesmo) e múltiplas conexões (mais de uma conexão entre dois vértices), então o valor de $\check{k}$ é limitado, segundo a equação:

$$\check{k} = N - 1 \quad (4.5)$$

A segunda quantidade é chamada de *conectividade de corte*, k_c . Para uma determinada rede, a probabilidade de que um vértice tenha conectividade k , tal que $k > k_c$, é praticamente nula. Como veremos, esta quantidade depende das particularidades de cada modelo.

Coefficiente de Clusterização

Cada vértice possui uma quantidade associada que mede o nível de conectividade entre seus vizinhos. O coeficiente de clusterização é definido como segue:

$$C_u = \frac{2e_u}{k_u(k_u - 1)} \quad (4.6)$$

aqui, e_u indica o número de conexões compartilhadas entre os vizinhos do vértice u . O valor do

coeficiente de clusterização varia sempre entre 0 e 1. $C_u = 0$ indica que os vizinhos do vértice u "não se conhecem" e portanto, não estabelecem conexões entre si. Ao contrário, quando $C_u = 1$ existem conexões entre todos os vizinhos de u . É interessante notar que o coeficiente de clusterização também indica a densidade de triângulos² que envolvem o vértice u . Na rede da figura 4.7, o coeficiente de clusterização do vértice 22 é: $C_{22} = 2(3/20) = 0,3$. Muitas vezes estaremos interessados somente na média dos coeficientes de clusterização, denotada por:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i \quad (4.7)$$

Nos grafos aleatórios, o coeficiente de clusterização médio, $\langle C \rangle_{al}$ é proporcional à conectividade média e inversamente proporcional ao tamanho da rede [22]:

$$\langle C \rangle_{al} = \frac{\langle k \rangle}{N}. \quad (4.8)$$

Como foi discutido, $\langle k \rangle \ll N - 1$ na maioria das redes reais de interesse científico (redes esparsas). Sendo assim, fosse a topologia destas redes completamente aleatória, esperaríamos coeficientes de clusterização muito baixos, o que não se verifica nos dados empíricos.

Para as redes regulares temos $e_u = 3/2 \langle k \rangle (\langle k \rangle - 1)$, para todo u . Assim, usando a equação 4.6 o coeficiente de clusterização médio, $\langle C \rangle_{reg}$ é independente de N e toma a forma:

$$\langle C \rangle_{reg} = \frac{3 \langle k \rangle - 6}{4 \langle k \rangle - 4}. \quad (4.9)$$

O coeficiente de clusterização das redes mundo pequeno, $\langle C \rangle_{mp}$ foram estimados pela primeira vez por Barrat e Weigt em [77]. Neste trabalho eles mostram que esta quantidade é dada por

$$\langle C \rangle_{mp} = \left[\frac{3 \langle k \rangle - 6}{4 \langle k \rangle - 4} \right] (1 - p)^3 \quad (4.10)$$

No gráfico da figura 4.8 mostramos o comportamento do coeficiente de clusterização para a rede mundo pequeno em função da probabilidade de redirecionamento p . Destacamos três regiões com características topológicas diferentes. A primeira, corresponde aos baixos valores de p onde o coeficiente de clusterização é praticamente invariante. Esta região corresponde à rede regular. Para grandes valores de p , temos o coeficiente de clusterização indo à zero, logo associamos esta região à completa desorganização das redes aleatórias. Entre as duas regiões

²Denotamos por triângulos a estrutura de conexão do tipo: m liga-se com n e com v e n liga-se com v .

Figura 4.8: Coeficiente de clusterização médio do modelo mundo pequeno proposto por Watts e Strogatz [1] em função da probabilidade de redirecionamento p . A curva é normalizada pelo valor do coeficiente de clusterização médio da rede regular ($C(p = 0)$).

identificamos as redes de mundo pequeno. É importante ressaltar que o limite entre essas regiões não é bem definido na literatura.

O coeficiente de clusterização médio do modelo de Barabási-Albert para as redes livres de escala, $\langle C \rangle_{BA}$ foi encontrado por Klemm e Eguíluz em [78]. O resultado obtido por eles foi:

$$\langle C \rangle_{BA} = \frac{5}{6} - \frac{7}{30m} + \mathcal{O}(m^{-2}) \quad (4.11)$$

onde $m \geq 1$ é um parâmetro do modelo. Vemos que o modelo de Barabási-Albert fornece redes livres de escala com alto grau de clusterização, já que seu valor mínimo é aproximadamente 0,72 quando $m = 2$ ³.

O coeficiente de clusterização possui ampla diversidade de valores nas redes reais, como por exemplo: 0,27 para a rede WWW [23], 0,34 rede de circuitos eletrônicos [36] e 0,79 para a rede de atores de filmes [19].

³Quando $m = 1$ a rede resultante não possui circuitos fechados, logo o coeficiente de clusterização é nulo para todos os vértices.

Distribuição de Conectividade

Sabemos que a maioria das redes reais são esparsas, pois o número de conexões é da ordem do número de vértices, ou seja, pela equação 4.4 temos: $\langle k \rangle \sim \mathcal{O}(1)$ [76]. Para sabermos mais sobre a conectividade da rede de maneira global definimos uma quantidade chamada distribuição de conectividade (DC) $P(k)$, que nos informa a fração de vértices que possuem conectividade k [79, 80]. Agora, podemos definir a conectividade média como o primeiro momento da DC, como mostra a equação 4.12:

$$\langle k \rangle = \frac{1}{N} \sum_{k=0}^{\infty} kP(k) \quad (4.12)$$

Para efeitos de cálculo, é comum aproximarmos a soma por uma integral, assumindo que a distribuição de conectividade é contínua. Porém, somente os valores inteiros da distribuição possuem significado físico.

É interessante notar que a DC é capaz de diferenciar alguns modelos de rede que possuem mesmo valor de $\langle k \rangle$. Porém, não é capaz de diferenciar todos os modelos de redes existentes na literatura. Para isso, precisamos de ferramentas adicionais como as medidas hierárquicas [81, 82, 83, 84, 85] e outras que veremos à frente [24, 81].

A DC das redes aleatórias é conhecida desde os primeiros trabalhos de P. Erdős e A. Rényi [66]. Neste caso ela toma a forma da lei binomial, dada por:

$$P(k) = \frac{(N-1)!}{k!(N-1-k)!} p^k (1-p)^{N-1-k} \quad (4.13)$$

onde p é um parâmetro do modelo. Como veremos mais adiante, p indica a probabilidade de que dois vértices quaisquer estejam ligados. Usando a equação 4.12 vemos que a conectividade média deste modelo é:

$$\langle k \rangle_{al} = p(N-1) \quad (4.14)$$

No limite em que $N \gg \langle k \rangle$, $P(k)$ converge para a distribuição de Poisson muito bem conhecida da teoria dos processos estocásticos [76].

$$P(k)_{al} = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!} \quad (4.15)$$

No Apêndice 1 mostramos que, graças a forma da distribuição de conectividade dos grafos aleatórios, existe um limite superior para os valores de k . Este limite está vinculado pela expressão $k_c \left[\ln \left(\frac{k_c}{\langle k \rangle} \right) - 1 \right] = \ln N - \langle k \rangle$, sendo que praticamente não existem vértices na rede com conectividade superior a k_c . Por exemplo, para uma rede aleatória com $N = 10^6$, temos $k_c = 20$, que é muito menor do que o máximo possível $\tilde{k} = 10^6 - 1$. Um exemplo da distribuição de conectividade de uma rede aleatória pode ser vista na figura 4.9-a.

As redes regulares infinitas ou periódicas de contorno possuem DC dada pela delta de Dirac, como mostra a equação 4.16. Isso ocorre porque todo vértice realiza o mesmo número de conexões.

$$P(k)_{reg} = \delta_{k, \langle k \rangle} \quad (4.16)$$

A equação 4.16 nos informa que $P(k) = 1$ se $k = \langle k \rangle$ e $P(k) = 0$ se $k \neq \langle k \rangle$.

A DC das redes mundo pequeno é praticamente idêntica a das redes regulares que as originaram, já que a perturbação (valor de p) é muito pequena. Esse fato pode ser visualizado nos gráficos b e c da figura 4.9. Vemos que a DC da rede regular não possui o comportamento esperado quando comparamos com a equação 4.16. Isso ocorre porque o gráfico representa a DC de uma rede regular finita e não periódica. Assim, os vértices que estão na extremidade da rede não possuem 4 vizinhos, como os vértices centrais.

Pesquisas recentes [19, 25, 57, 86] mostraram que a DC da maioria das redes reais não se assemelha àquelas citadas acima, principalmente pelo fato de não apresentarem um pico característico como na distribuição de Poisson (veja a figura 4.9-d). Essa ausência de escala característica deu nome as redes livres de escala. Essas redes possuem DC na forma de uma lei de potência:

$$P(k)_{le} = \alpha k^{-\gamma} \quad (4.17)$$

Dorogovtsev e Samukhin mostraram [87] que as flutuações encontradas na cauda das distribuições que seguem leis de potência são muito maiores do que quando comparadas as flutuações da

Figura 4.9: Distribuições de conectividade dos modelos de rede discutidos neste capítulo. a) rede aleatória de Erdős-Rényi ($N = 1000$ e $\langle k \rangle = 7$), b) rede regular finita ($N = 32 \times 32$ e $\langle k \rangle = 3,875$), c) mundo pequeno ($N = 32 \times 32$, $\langle k \rangle = 3,875$ e $p = 0,03$) e d) livre de escala ($N = 1000$, $\langle k \rangle = 7$ e $\gamma = 2$).

distribuição de Poisson. Conseqüentemente, a ocorrência de *hubs* (vértices com alta conectividade) é muito maior nas redes livres de escala em relação as redes aleatórias. De fato, o valor limite de conectividade para as redes livres de escala é da ordem $k_c \approx N^{1/(\gamma-1)}$. Para o caso de uma rede com $N = 10^6$ e um valor típico de $\gamma = 2$, temos $k_c \propto \check{k}$.

Caminhos

Antes de avançarmos, iremos introduzir o conceito de caminho que será necessário nas sessões seguintes. Um caminho é definido como um subconjunto ordenado de $\mathcal{V}$, denotado por $\mathcal{T} = \{x_i\}$, onde $i = 1, 2, \dots$. Os elementos deste conjunto devem satisfazer a seguinte propriedade de adjacência: $(x_i, x_{i+1}) \in \mathcal{E}$. Com isso, garantimos que existe um caminho do primeiro ao último elemento de $\mathcal{T}$. Chamamos de comprimento do caminho a cardinalidade (número de elementos) de $\mathcal{T}$, ou seja:

$$l_{\mathcal{T}} = ||\mathcal{T}|| \quad (4.18)$$

Menor Caminho e Diâmetro da Rede

O caminho entre u e v para o qual o valor de $l_{\mathcal{T}}$ é mínimo, é denotado por $\mathcal{L}_{u,v}$. Dizemos que $\mathcal{L}_{u,v}$ é o menor caminho entre u e v pois representa o número mínimo de vértices necessários para ir de u até v ⁴. Por exemplo, para a rede genérica mostrada na figura 4.7, o menor caminho entre os vértices 17 e 27 é dado por $\mathcal{L}_{17,27} = \{17, 16, 13, 7, 10, 22, 24, 26, 27\}$. Na figura, este caminho é representado pelos vértices destacados em amarelo e seu comprimento é $l_{\mathcal{L}_{17,27}} = 9$.

Muitas vezes, estamos interessados apenas no valor médio de $l_{\mathcal{L}_{u,v}}$ (a média é realizada sobre todos os pares de vértices u e v), denotado por $\langle l \rangle$ e comumente chamado de *diâmetro da rede*. Para a rede da figura 4.7 temos $\langle l \rangle \approx 4$.

Quantificar esta característica é muito importante em redes complexas, pois as informações trafegam baseadas na lei de minimização da energia. Sendo assim, caminhos mais curtos são preferencialmente adotados para a transferência de informação e consequentemente, redes mais compactas (baixos valores de $\langle l \rangle$) desempenham melhor este papel.

É fácil mostrar (Apêndice 2) que o diâmetro das redes aleatórias é da forma:

$$\langle l \rangle \propto \frac{\ln N}{\ln \langle k \rangle} \quad (4.19)$$

Este resultado também é válido para as redes mundo pequeno. A dependência de $\langle l \rangle$ com o $\ln N$ é responsável pelo efeito mundo pequeno observado nesses dois modelos. Através da equação 4.19 vemos que uma rede com $N = 10^6$ e $\langle k \rangle = 4$ possui diâmetro da ordem $\langle l \rangle \approx 10$.

O fato de não existir efeito mundo pequeno nas redes regulares implica em valores de $\langle l \rangle$ superiores aos casos anteriores. A ausência do efeito mundo pequeno ocorre porque as conexões respeitam as restrições espaciais vinculadas a posição dos vértices. Assim, conexões de longo alcance são ausentes⁵ e as quantidades $\langle l \rangle$ e N estão relacionados pela dimensionalidade espacial (D) através da equação⁶:

⁴ Alguns autores adotam o menor caminho como sendo o número mínimo de conexões necessárias para ir de u até v e $\mathcal{L}_{u,v}$ representa o conjunto destas conexões [64].

⁵ A rede de aeroportos [61] é uma exceção. Apesar dos aeroportos (vértices) possuírem coordenadas espaciais associadas, dizemos que existe uma conexão entre dois aeroportos se existe um voo de um a outro. Obviamente, não existe restrição espacial para estas conexões.

⁶ Válido para redes regulares finitas.

$$\langle l \rangle = N^{1/D} \quad (4.20)$$

Recentemente R. Cohen e S. Havlin [88] mostraram que as redes livres de escala são ainda mais compactas do que as redes em que o efeito mundo pequeno está presente. De fato, a dependência de $\langle l \rangle$ e N para as redes livres de escala é da forma⁷:

$$\langle l \rangle \propto \ln(\ln N) \quad (4.21)$$

Para a rede do exemplo anterior, onde $N = 10^6$, encontramos o valor $\langle l \rangle \approx 3$. O fato do valor inteiro da função $\ln(\ln N)$ ser constante para muitas escalas de N explica por exemplo, porque o diâmetro das redes metabólicas é praticamente independente de N [88, 32].

Rota de Menor Custo

Dado o caminho $\mathcal{T}_{u,v}$ entre os vértices u e v , definimos seu custo como:

$$c_{\mathcal{T}_{u,v}} = \sum_{i=1}^{l_{\mathcal{T}_{u,v}}-1} w_{x_i, x_{i+1}} \quad (4.22)$$

onde $w_{x_i, x_{i+1}}$ representa o peso da conexão existente entre o elemento x_i e o elemento x_{i+1} do conjunto $\mathcal{T}_{u,v}$. A adição de pesos é uma maneira de quantificar a interação entre os vértices. Neste caso, a transferência de informação pode não seguir necessariamente pelo caminho que passa pelo menor número de vértices e sim através do conjunto de conexões que oferece o menor custo ao transporte.

O caminho que minimiza o valor de $c_{\mathcal{T}_{u,v}}$ é chamado *rota de menor custo* e é denotado pelo conjunto $\mathcal{C}_{u,v}$. É interessante notar que, para os casos onde $w_{u,v} = w$ para qualquer valor de u e v , ou seja, quando todas as conexões da rede possuem o mesmo peso, a rota de menor custo coincide com o menor caminho, ou seja $\mathcal{L}_{u,v} = \mathcal{C}_{u,v} \quad \forall (u, v)$.

⁷Este resultado é válido para $2 \leq \gamma \leq 3$ (onde enquadra-se a maioria das redes reais). Quando $\gamma = 3$ a relação é $\langle l \rangle \propto \ln N / \ln \ln N$ e para $\gamma > 3$ o comportamento é similar ao das redes aleatórias, $\langle l \rangle \propto \ln N$.

Figura 4.10: a) rede genérica com o vértice fonte destacado em amarelo, b) respectiva rede estrela e c) respectiva árvore de menor cobertura.

Fator de Rota

Segundo [4] existem basicamente dois critérios para quantificar a eficiência das redes geográficas de transporte que possuem um vértice fonte. O primeiro critério diz respeito ao comprimento da rota de menor custo $l_{\mathcal{C}_{u,0}}$, estabelecida entre cada vértice u e o vértice fonte, denotado pelo índice 0 (aqui, o peso das conexões diz respeito a distância entre os vértices que ela conecta).

Sob este aspecto, a rede que apresenta maior eficiência é aquela onde todos os vértices são ligados diretamente ao vértice fonte por uma linha reta, essas redes são conhecidas como *Redes Estrela* (veja um exemplo na 4.10b).

O fator de rota, q_1 é uma medida que nos auxilia a quantificar o quanto as redes reais divergem de suas versões estrela.

$$q_1 = \frac{1}{N} \sum_{i=1}^N \frac{l_{\mathcal{C}_{i,0}}}{d_{i,0}} \quad (4.23)$$

onde $d_{i,0}$ representa a distância Euclidiana entre os vértices i e o vértice fonte 0. Para exemplificar o uso desta medida, vejamos o exemplo apresentado na figura 4.10-a, onde apresentamos um exemplo de rede geográfica com $N = 8$ vértices e um vértice fonte. Em 4.10b mostramos sua versão estrela. O fator de rota associado a esta rede é estimado da seguinte maneira:

$$q_1 = \frac{1}{8} \left(\frac{16}{4} + \frac{14}{6} + \frac{10}{4} + \frac{8}{6} + \frac{4}{4} + \frac{8}{6} + \frac{10}{4} + \frac{6}{6} \right) = 2$$

Portanto, deslocar-se sobre a rede a partir da fonte até qualquer outro vértice, equivale em média, a percorrer o dobro da distância que separa estes vértices. Obviamente, quanto mais próximo de 1 for o valor de q_1 , maior a eficiência da rede [4].

Árvore de Menor Cobertura

Todo subconjunto de $\mathcal{E}$ com tamanho $N - 1$ capaz de manter os N vértices conectados em apenas uma componente, é chamado de árvore [22]. A principal característica das árvores é a ausência de ciclos. Consequentemente existe apenas um caminho que liga dois vértices quaisquer e o coeficiente de clusterização é sempre nulo para todos os vértices ($C_i = 0$ para qualquer i). Dada uma rede qualquer, o número estimado de árvores que ela possui é da ordem $n_{ar} = \mathcal{O}(3^\Delta)$, onde Δ representa o número de vezes que ocorrem conexões na forma de triângulo, ou seja, a liga com b e com c e b liga com c .

Dentre todas essas possibilidades, existe uma que é particularmente importante para o segundo critério de eficiência [4]. Se estamos interessados em redes econômicas e de fácil manutenção, o modelo estrela deveria ser o último a ser adotado. Isso ocorre pois nele as conexões são longas e suas construções, ou reparos demandam muito material. Sob este aspecto, uma rede ideal é aquela onde as conexões possuem o menor comprimento possível. A árvore que obedece este requisito é denominada *Árvore de Menor Cobertura - AMC*. Neste trabalho, a árvore de menor cobertura foi obtida pelo algoritmo *Prim-Dijkstra* [89] cujo funcionamento é discutido no Apêndice 3.

Para compararmos uma rede real e sua respectiva AMC sob este segundo critério de eficiência, podemos mensurar o comprimento médio de cada conexão nos dois casos e tomar a razão entre eles:

$$q_2 = \frac{\langle e \rangle}{\langle e \rangle_{AMC}}$$

Quanto mais próximo de 1 for o valor de q_2 , mais eficiente é a rede sob o ponto de vista do segundo critério de eficiência.

Em [4], os autores mostram que redes reais como as redes de ferrovias, redes de esgoto, e redes de gás possuem valores de q_1 e q_2 muito satisfatórios, indicando que deve existir um processo natural de otimização durante a construção desses sistemas.

4.4.3 Medidas Hierárquicas

Até pouco tempo, a caracterização estática das redes complexas baseava-se apenas em medidas locais (e em seus valores médios), que não forneciam uma visão ampla da organização das

Figura 4.11: a) Rede genérica onde destacamos as hierarquias do vértice de referência $r = 1$, b) número hierárquico de vértices, c) conectividade hierárquica e d) coeficiente de clusterização hierárquico. Os quadrados vermelhos representam as medidas tomadas considerando o vértice 1 como referência. Para os círculos verdes consideramos o vértice 2 como referência. A faixa azul representa o desvio padrão e a média tomada sobre todos os vértices da rede.

redes como um todo [83]. Ao longo do tempo, medidas mais elaboradas foram desenvolvidas, porém ainda eram incapazes de gerar um mapa completo das redes. Exemplos destas medidas são *correlação de conectividade* [24] e *betweenness centrality* [81]. Há pouco tempo, uma nova classe de medidas foi proposta [81, 82, 83, 84, 85] como alternativa para a investigação global das propriedades das redes complexas. As medidas que enquadram-se nesta nova classe são ditas *medidas hierárquicas*.

As medidas hierárquicas são generalizações e complementações das medidas clássicas capazes de nos informar mais a respeito das progressivas vizinhanças de cada vértice. Até mesmo o conceito de vizinhança pode ser generalizado e estendido, englobando toda rede.

O conceito de hierarquia está intimamente relacionado ao conceito de vértice de referência [90, 91], que iremos denotar por r . Dizemos que o conjunto $\mathcal{H}_r(h)$ de todos os vértices u que satisfazem a relação $l_{\mathcal{L}_{u,r}} = h - 1$ representa a hierarquia h do vértice de referência r . De maneira menos formal, dizemos que $\mathcal{H}_r(h)$ é o conjunto de todos os vértices que estão à distância $h + 1$ do vértice r (a distância é estimada pelo menor caminho que conecta u e r).

Para exemplificar este conceito, apresentamos na figura 4.11-a as hierarquias relativas ao vértice de referência 1.

Número Hierárquico de Vértices

A extensão da propriedade de vizinhança, através do conceito de hierarquia, nos traz a necessidade de quantificar o "tamanho da vizinhança" a medida que nos afastamos do vértice de referência. Para isso, definimos a quantidade $n_r(h)$ que nos informa quantos vértices existem na hierarquia h referente ao vértice r . Formalmente escrevemos:

$$n_r(h) = ||\mathcal{H}_r(h)|| \quad (4.24)$$

Por exemplo, para a rede mostrada na figura 4.11-a os valores de $n_r(h)$ são apresentados no gráfico 4.11-b, quando tomamos como referência os vértices 1 e 2 destacados na rede. Também apresentamos o valor médio e o desvio padrão de $n_r(h)$ tomado sobre todos os vértices da rede.

Conectividade Hierárquica

Definimos como conectividade hierárquica do vértice de referência r a quantidade $k_r(h)$ [83, 84, 85]. Ela nos informa o número de conexões estabelecidas entre os vértices do conjunto $\mathcal{H}_r(h-1)$ e os do conjunto $\mathcal{H}_r(h)$. É interessante notar que esta medida é uma generalização da conectividade clássica, já que para $h = 1$ temos: $k_r(1) = k_r$.

Através desta nova formulação, a propriedade de conectividade de cada vértice passa a ser representada por uma função $f(h) = k_r(h)$, ao invés de um único valor como proposto inicialmente. Este fato permite distinguir com maior fidelidade cada vértice, pois é muito pequena a probabilidade de que dois vértices tenham exatamente a mesma curva de conectividade. Na figura 4.11-c mostramos a curva de conectividade para o vértice de referência 1 (quadrados) destacado na rede da figura 4.11-a. Nesta figura, também apresentamos a curva de conectividade hierárquica quando consideramos o vértice 2 como referência (círculos). A curva de conectividade média para todos os vértices da rede também é apresentada, bem como seu desvio padrão.

Na figura, vemos claramente que o vértice 1 apresenta um comportamento que diverge do valor médio, e que isto não ocorre para 2. No primeiro caso, o pico da curva está localizado em $h = 4$ enquanto que no segundo o pico está em $h = 6$. Outro ponto importante capaz de diferenciar os vértices 1 e 2 quando eles são adotados como vértices de referência é o número de hierarquias

que cada um deles possui. Pela figura, vemos que a hierarquia máxima ($\check{h}_r$) para o primeiro vértice é $\check{h}_1 = 6$, enquanto que para o segundo temos $\check{h}_2 = 10$. Além disso, o valor de pico de $k_1(h)$ é significativamente maior que o de $k_2(h)$.

Podemos entender essas diferenças olhando simplesmente para a posição topológica de cada um dos vértices na rede. De fato, vemos que o vértice 2 está localizado na periferia da rede, enquanto 1 é um vértice central. Conforme dilatamos as hierarquias, é provável que nós de referência centralizados alcancem as extremidades da rede após poucas hierarquias, enquanto vértices periféricos necessitam um maior número de hierarquias para chegar as extremidades opostas.

As diferenças mencionadas acima estão intimamente relacionadas com a importância do vértice na rede. Vértices mais centralizados, ou seja, que possuem menor valor de $\check{h}_r$ são, em geral, mais importantes do que aqueles para o qual $\check{h}_r$ é grande. Assim, vemos que a conectividade hierárquica é um artefato útil para quantificar a importância dos vértices, bem como estabelecer diferenças entre eles. Em [84] os autores mostram como o conceito de conectividade hierárquica, dentre outros, é capaz de distinguir vértices em redes que sofreram redirecionamento de conexões.

Denotamos por $P_h(k)$ a fração de vértices com conectividade hierárquica $k(h)$. A formulação matemática para a distribuição de conectividade hierárquica dos modelos básicos de redes complexas, ainda é um problema em aberto. Os primeiros passos nesse sentido foram dados recentemente [92], quando conseguimos calcular analiticamente o valor exato de $P_2(k)$ para o modelo livre de escala de Barabási-Albert [69] com parâmetro $m = 1$ (veja o Apêndice 4). A equação 4.25 apresenta a distribuição encontrada por Barabási-Albert [69] e em seguida (equação 4.26) apresentamos nosso resultado para a distribuição de $P_2(k)$.

$$P_1(k) = \frac{4}{k(k+1)(k+2)} \quad (4.25)$$

$$P_2(k) = \sum_{r=1}^{\check{k}} \left\{ \Delta_{1,k} + \frac{r-1}{2r+k+2} \sum_{f=0}^{k-1} \left[\frac{\Gamma(2r+f+2)}{\Gamma(r+f)} \right] P(r-1, f) \right\} \quad (4.26)$$

onde $\Delta_{1,q}$ é a correlação entre os vértices de conectividade unitária e os vértices com conectividade k - veja o Apêndice 4. Aqui vemos que a distribuição $P_1(k)$ é proporcional a k^{-3} quando $k \gg 2$, assim, pela equação 4.17 obtemos o expoente de escala $\gamma = 3$ para a conectividade da primeira hierarquia no modelo de Barabási-Albert [69].

Coeficiente de Clusterização Hierárquico

A mesma generalização realizada para a conectividade, pode ser estendida ao coeficiente de clusterização [83, 84, 85]. O coeficiente de clusterização hierárquico do vértice de referência r é denotado por $C_r(h)$:

$$C_r(h) = \frac{2e_r(h)}{n_r(h)(n_r(h) - 1)} \quad (4.27)$$

onde $e_r(h)$ representa o número de conexões existentes entre os vértices do conjunto $\mathcal{H}_r(h)$. Na figura 4.11-d apresentamos as curvas de $C_r(h)$ para os vértices 1 e 2 da rede mostrada em 4.11-a. Fica claro que o comportamento de $C_1(h)$ assemelha-se a média da rede enquanto que, para o vértice 2 a curva apresenta grandes flutuações em torno média.

Generalização de Outras Medidas

As caracterizações hierárquicas podem ser estendidas a muitas outras medidas clássicas a fim de enriquecer a descrição das redes complexas. Nesta tese ficaremos restritos as medidas hierárquicas aqui apresentadas. Para uma leitura complementar veja [81, 82, 83, 84, 85].

Efeito do Tamanho Finito

É importante ressaltar o fato de que toda medida hierárquica sofre influência do efeito de tamanho finito das redes. O valor de $\check{h}$ é um reflexo direto desta influência. De fato, se imaginarmos uma rede aleatória infinita onde todos os vértices possuem conectividade $\langle k \rangle$ (aproximação por campo médio), é claro que não existe limite para o valor de $\check{h}$. Além disso, o valor da conectividade hierárquica seria dado por uma lei de potência do tipo $k_r(h) \propto \langle k \rangle^h$. No entanto, esse não é o comportamento observado na natureza, exatamente pelo fato das redes reais possuírem um número limitado de vértices. Em geral, o que ocorre com a conectividade hierárquica é que ela cresce (de acordo com a potência $\langle k \rangle^h$) até alcançar um pico de máximo. Em seguida começa a cair, indo a zero para $h > \check{h}$. O gráfico da figura 4.12 ilustra este fenômeno.

Como exemplo, podemos observar o comportamento da curva de conectividade hierárquica média da figura 4.11-b, referente a rede apresentada na figura 4.11-a. Aqui vemos claramente a presença de um pico quando $h \approx 6$.

Figura 4.12: Ilustração do comportamento da curva de conectividade hierárquica para redes de tamanho finito (preto) e infinito (vermelho).

4.4.4 Fluxo Máximo

Imaginemos um vértice (fonte) que produz um determinado material a uma taxa constante no tempo e um vértice (destino) que consome este material também a uma taxa fixa. Agora supomos que estes vértices estão conectados através de um grafo e que cada conexão (u, v) deste grafo possui uma capacidade máxima de transporte $\omega_{u,v}$. O problema do fluxo máximo consiste em saber qual a quantidade máxima de material que pode ser conduzida da fonte até o destino [93]. Sua solução é de grande interesse em diversas áreas da ciência, como na hidrodinâmica, para tratar o escoamento de fluidos em canais com propriedades distintas, na eletricidade, para o estudo do transporte elétrico em redes com componentes que possuem características elétricas diferentes, no planejamento rodoviário, para otimização de rotas, entre outras. Em geral o fluxo nas redes é dado por uma taxa, nos casos citados acima, algumas das unidades possíveis são, respectivamente: *volume/tempo*, *carga/tempo* e *veículos/tempo*.

O fluxo $s_{u,v}$ na conexão (u, v) pode ser aumentado pela quantidade: $g_{u,v} = \omega_{u,v} - s_{u,v}$, até que atinja seu valor máximo. Além disso, assumimos que existe conservação de fluxo em todos os vértices do grafo, exceto na fonte e no destino. O fluxo no sistema é descrito pelo conjunto dos valores $\{s_{u,v}\}$ em todas as conexões e dizemos que um estado de fluxo é possível se $s_{u,v} < \omega_{u,v}$ para todas as conexões (u, v) . Qualquer estado de fluxo possível determina uma quantidade Φ que escoar da fonte para o destino. Dizemos que o fluxo é máximo se Φ for finito e se não existe outro estado de fluxo para o qual o escoamento é maior do que Φ . As restrições deste problema são formalmente escritas como:

$$\sum_{m=1}^N s_{m,u} a(m,u) = \sum_n s_{u,n} a(u,n) \quad u \in [1, N-1] \quad (4.28)$$

$$0 \leq s_{u,v} \leq \omega_{u,v} \quad \forall (u,v) \quad (4.29)$$

onde $\mathcal{A}$ é a matriz de adjacência. Além disso, consideramos que o vértice fonte e o vértice de destino são representados pelos índices 0 e N , respectivamente. A primeira expressão nos informa que tudo que chega a um vértice deve sair. Na segunda está explícito o fato de que o fluxo em um canal não deve ultrapassar sua capacidade máxima. O fluxo total que deve ser maximizado é dado por:

$$\Phi = \sum_{m=1}^N s_{0,m} a(0,m) = \sum_{m=1}^N s_{m,N} a(m,N) \quad (4.30)$$

Aqui supomos que os vértices fonte e destino possuem índices 0 e N , respectivamente. Teoricamente, o fluxo máximo pode ser encontrado analiticamente por tratar-se de um problema de programação linear. Entretanto durante os últimos anos, vários algoritmos computacionais foram desenvolvidos, tornando a tarefa de encontrar numericamente o fluxo máximo bem mais simples e eficiente. O primeiro algoritmo foi proposto por Ford e Fulkerson em 1956 [94]. Este método possui complexidade da ordem $\mathcal{O}(E\Phi)$ (onde E é o número de conexões da rede) quando os valores de $\omega_{i,j}$ são inteiros. Outros métodos com tempo de cálculo independente do valor de Φ foram desenvolvidos a partir de então [95].

Neste trabalho, vamos utilizar uma implementação do algoritmo proposto por Ford e Fulkerson denotada na literatura como *Matrix Labeling Process* (detalhes desta técnica podem ser encontrados em [93]).

O algoritmo citado acima é aplicável a redes direcionadas. No entanto quando não existe informações acerca da direção real do fluxo no sistema, o algoritmo entende que cada uma das conexões não direcionadas pode ser representada por duas conexões direcionadas em sentidos contrários, como na figura 4.13.

4.4.5 Comunidades

Muitas redes complexas reais exibem uma estrutura topológica organizada em comunidades, ou seja apresentam um particionamento natural dos vértices em grupos bem conectados. Várias técnicas para a detecção destas comunidades foram propostas nos últimos anos. Os primeiros

Figura 4.13: O Algoritmo de fluxo máximo interpreta as conexões não direcionadas do sistema haversiano como duas conexões direcionadas em sentidos contrários e pode, possivelmente, realizar operações distintas nas duas direções.

Figura 4.14: Exemplo de rede com modularidade $Q \approx 0,93$ (equação 4.31). O particionamento avaliado contém três comunidades, como é indicado pelas cores dos vértices.

baseavam-se em métodos espectrais [96, 97] e hierárquicos [98], porém possuíam uma funcionalidade restrita a somente alguns casos. Em 2002, Girvan e Newman [99] propuseram um método geral para detecção de comunidades utilizando a medida de *edge betweeness*. A grosso modo, podemos dizer que esta quantidade mede o número de menores caminhos que passam por uma determinada conexão. O método proposto foi aplicado com muito sucesso a um grande número de redes complexas, porém a alta demanda computacional o tornava viável apenas para redes com poucos milhares de vértices. No início de 2004, Newman [100] propôs um novo algoritmo de detecção baseado na otimização da medida de modularidade [101]. Usando este método o tempo de processamento caiu para $\mathcal{O}(n^2)$ em redes esparsas, contra $\mathcal{O}(n^3)$ do primeiro método. Neste trabalho utilizamos uma versão deste algoritmo, proposta em [2]. Utilizando uma estrutura de dados diferente da proposta originalmente por Newman, esta versão é capaz de realizar a detecção com um tempo de processamento da ordem de $\mathcal{O}(n \log^2 n)$.

A modularidade Q é uma quantidade que qualifica um dado particionamento da rede:

$$Q = \sum_{i=1}^{N_C} (\beta_{ii} - \alpha_i^2) \quad (4.31)$$

onde N_C é o número de comunidades do particionamento que deseja-se avaliar, β_{ij} corresponde

a fração de conexões que ligam vértices da comunidade i com vértices da comunidade j . A quantidade α_i representa a fração de conexões que ligam vértices de i aos demais vértices que não estão contidos em i . Por exemplo, a rede mostrada na figura 4.14 está particionada em $N_C = 3$ comunidades. Seu valor de modularidade é estimado como segue:

$$Q = \left[\frac{28}{64} - \left(\frac{3}{64} \right)^2 \right] + \left[\frac{9}{64} - \left(\frac{2}{64} \right)^2 \right] + \left[\frac{23}{64} - \left(\frac{3}{64} \right)^2 \right] \approx 0,93$$

Segundo [2] valores de Q maiores que 0.3 são bons indicativos da presença de comunidades. Na literatura, geralmente encontramos valores para Q entre 0,3 e 0,75. Como veremos, a rede do sistema Haversiano apresenta valor de modularidade próximo de 0,95.

4.5 Caracterização Dinâmica

4.5.1 Introdução

Os processos dinâmicos em redes complexas podem ser classificados em três tipos: i) dinâmica da rede, ii) dinâmica na rede e iii) interação entre a dinâmica da rede e as dinâmicas na rede [102, 103, 104]. O primeiro refere-se a maneira como a rede é construída e como sua topologia se altera ao longo do tempo. Como exemplo, temos o modelo de Barabási-Albert para geração de redes livres de escala através do mecanismo de anexação preferencial [69]. Em segundo lugar temos os processos dinâmicos que ocorrem na rede. Aqui enquadram-se processos de disseminação de doenças [105, 106, 107, 108], tráfego de informações [109, 110], fenômenos de sincronização [111, 112], percolação [113, 114], tolerância a ataques [115, 116, 29], dentre outros.

Enquadram-se no terceiro tipo as dinâmicas que acoplam processos da primeira e da segunda classe. Estes processos ainda são pouco estudados na literatura porque, em geral, as escalas temporais que envolvem processos do tipo i) e do tipo ii) não são compatíveis. Um exemplo que permite estabelecer a conexão entre diferentes classes de processos foi reportado recentemente por Rocha e Costa em [117]. Aqui os autores propõem uma maneira de estudar a eficiência de redes com diferentes topologias que combatem a propagação de doenças, sendo que as doenças propagam-se no espaço bidimensional através de reações de difusão.

Neste trabalho estaremos interessados apenas nos modelos de evolução e na aplicação do processo de despercolação. Assim, iremos devotar esta seção a uma breve introdução aos modelos

de crescimento de redes mais utilizados na literatura. Em seguida apresentaremos o modelo de despercolação, suas características e aplicações.

4.5.2 Processos de Crescimento

Geração de Grafos Aleatórios

Em 1967 os matemáticos Erdős e Rényi reportaram um trabalho [118] onde propunham um modelo estocástico para geração de grafos aleatórios. O procedimento é bastante simples e resume-se em: dado um conjunto de N vértices, devemos conectar cada par de vértices com probabilidade p . Muitas propriedades deste modelo são solúveis analiticamente, e praticamente todas foram exploradas por Erdős e Rényi em seus trabalhos. Por exemplo, eles mostraram que seu modelo possui conectividade média $\langle k \rangle = p(N - 1)$. Assim, pela equação 4.4 vemos que o número médio de conexões adicionadas ao sistema é: $E = pN(N - 1)/2$ ou seja, a fração p de todas as conexões possíveis. Os autores também mostram que a distribuição de conectividade deste modelo é dada pela equação 4.13.

Uma das características mais importantes deste modelo é que ele sofre uma transição de fase [103] a medida que o valor de p aumenta. A grosso modo, dizemos que um sistema sofre uma transição de fase quando há uma mudança abrupta em alguma de suas características. Neste caso há o surgimento de uma *componente gigante* na rede.

Chamamos de componente o conjunto de vértices unidos por pelo menos um caminho (por exemplo, o grafo da figura 4.15-a possui 7 componentes, enquanto o da figura 4.15-b possui 4 componentes). Quando o valor de p é pequeno, a rede é composta por componentes com poucos vértices. Acima de um valor crítico p_c , uma componente se destaca e emerge englobando praticamente todos os vértices da rede. Dizemos que esta é uma componente gigante. O valor de p_c para este modelo é

$$p_c = \frac{1}{N - 1} \quad (4.32)$$

Na figura 4.15 mostramos a evolução de um grafo e o surgimento de uma componente gigante. O valor crítico de p deste exemplo é $p_c = 1/11 \approx 0.1$. Vemos que na figura 4.15-c-d, onde $p > p_c$ existe uma componente gigante, enquanto em 4.15-a-b isso não se verifica.

Figura 4.15: Componentes nos grafos aleatórios de Erdős-Rényi com a) $p = 0.03$, b) $p = 0.1$, c) $p = 0.2$ e d) $p = 0.4$.

Anexação Preferencial e as Redes Livres de Escala

Enquanto no modelo de Erdős-Rényi o valor de N é fixo, na maioria das redes reais isso não se verifica. Elas estão em constante evolução, perdendo e ganhando vértices e conexões. Barabási *et al.* [19, 69] perceberam esse fato ao estudar a rede WWW formada por páginas HTML interligadas através de hyperlinks. Depois de observarem que a distribuição de conectividade dessa e de outras redes não assemelhava-se a distribuição de Poisson dos grafos aleatórios eles propuseram um modelo para a construção deste novo tipo de rede. As redes ficaram conhecidas como livres de escala (devido à ausência de escala da distribuição de conectividade) e o modelo de crescimento ficou conhecido como anexação preferencial.

O modelo começa com uma pequena componente de m vértices. Inicialmente esta componente está completamente conectada ou seja, todos os vértices conectam-se a todos os outros. Em seguida, um novo vértice é adicionado à rede. Este novo vértice estabelece conexões com m vértices escolhidos de acordo com a probabilidade:

$$\pi(k) = Ak^\alpha \quad (4.33)$$

onde A é um fator de normalização. Esta equação nos dá a probabilidade de que um vértice com conectividade k receba uma conexão do novo vértice que está entrando na rede. O esquema da dinâmica de anexação preferencial está ilustrado na figura 4.16.

Em [69] os autores utilizam teoria de campo médio para estimar os coeficientes de escala deste modelo. Eles encontraram que a conectividade de cada vértice escala com o tempo de acordo com $k_i \propto mt^{0.5}$. Além disso, mostraram que o coeficiente de escala da distribuição de conectividade é $\gamma = 3$. A solução exata deste modelo foi apresentada por Dorogovtsev *et al.* [70] e revela que na verdade, o expoente de escala depende do parâmetro m da seguinte maneira:

Figura 4.16: Dinâmica de anexação preferencial do modelo de Barabási-Albert com parâmetros $m = 1$ e $m = 2$. É interessante notar a ausência de circuitos fechados no caso $m = 1$.

$$\gamma = 2 + 1/m.$$

Exemplos

Na figura 4.17-a-c apresentamos algumas características hierárquicas para três exemplos de rede, a saber: Rede aleatória de Erdős-Rényi (4.17-f) e dois exemplos da rede livre de escala Barabási-Albert, um com parâmetro $m = 1$ (figura 4.17-d) e outro com parâmetro $m = 2$ (figura 4.17-e).

O efeito da ausência de ciclos na rede de Barabási-Albert com $m = 1$ torna-se visível nos gráficos (4.17-a-b). Vemos que a falta de caminhos alternativos torna a rede dispersa, induzindo um alto valor de $\check{h}$. Ao contrário, para as outras duas redes analisadas, os valores de $\check{h}$ são baixos, indicando que as redes são topologicamente compactas. Essas afirmações ficam evidentes quando olhamos para a representação das redes nas figuras 4.17-d-f.

Outra característica interessante pode ser observada para os valores de $C(h)$ apresentados na 4.17-c para rede de Barabási-Albert com $m = 2$ e para a rede de Erdős-Rényi com $\langle k \rangle = 4$. A primeira possui valores de $C(h)$ relativamente altos para as primeiras hierarquias, tendendo a zero em seguida. Na segunda rede o valor de $C(h)$ é praticamente constante para todas as hierarquias. Em [85] os autores demonstram como as medidas hierárquicas dependem dos parâmetros de cada modelo.

4.5.3 Despercolação: A Vulnerabilidade das Redes Complexas

Praticamente todos os sistemas físicos e biológicos estão sujeitos a falhas e essa característica é herdada pela redes complexas nas quais estes sistemas são mapeados. Por exemplo, o mal

Figura 4.17: Caracterização de alguns modelos de redes complexas utilizando as seguintes medidas hierárquicas a) número hierárquico de vértices, b) conectividade hierárquica e c) coeficiente de clusterização hierárquico. A cor de cada curva está associada a cor da respectiva rede apresentada ao lado.

funcionamento de um roteador (falha no vértice), ou a má comunicação entre dois roteadores (falha na conexão), pode levar a uma falha na rede da internet. Essa falha pode ter ou não um reflexo global, dependendo da importância do roteador [115].

As falhas são simuladas nas redes por processos de remoção de vértices ou conexões. A robustez com que a rede supera estes processos de remoção está intimamente relacionada à estabilidade do sistema que ela representa [116]. Em geral, os processos de remoção são classificados em dois tipos, de acordo com a maneira como ocorrem. Chamamos de *erros* as falhas cuja ocorrência é aleatória. Quando existe algum mecanismo de seleção governando as falhas, estas são denominadas de *ataques*.

A simulação de um processo de falha consiste, simplesmente, na remoção continuada de vértices (ou conexões), onde denotamos por f a fração de vértices (ou conexões) removidos [119]. Para avaliar a vulnerabilidade das redes, algumas quantidades são monitoradas durante o processo de remoção. Entre elas destacam-se: i) tamanho da maior componente, ii) diâmetro da maior

Figura 4.18: Simulação do processo de falhas em uma rede genérica. As falhas ocorrem pela remoção de a) vértices e b) conexões.

componente e iii) tamanho médio das componentes menores.

i) *Tamanho da maior componente* - Z - É uma medida direta da capacidade da rede em manter-se conectada quando a fração f de vértices (ou conexões) foram removidos. Variações abruptas desta quantidade indicam a fragmentação e a consequente falta de funcionalidade da rede. É importante ressaltar que esta medida é apresentada como a fração dos N vértices contidos na componente gigante.

ii) *Diâmetro da maior componente* - ℓ - Quantifica o menor caminho médio entre todos os vértices da maior componente quando uma fração f de vértices (ou conexões) foram removidos. Para $f = 0$ temos $\ell = \langle l \rangle$, já que a maior componente corresponde a toda a rede.

iii) *Tamanho médio das componentes menores* - $\langle z \rangle$ - A média que o valor de f aumenta, algumas componentes são desagrupadas da maior componente. Esta medida corresponde ao valor médio do tamanho dessas novas componentes que surgem durante o processo de remoção.

Figura 4.19: variação do tamanho da maior componente conectada em função de f durante o processo de despercolação. Todas as redes utilizadas possuem $N = 500$ vértices e são apresentadas na figura 4.17.

Na figura 4.18 ilustramos um processo de falha em uma rede genérica com $N = 12$ e $E = 28$. Levamos em consideração tanto a remoção de vértices (4.18-a) quanto a remoção de conexões (4.18-b). É interessante notar o comportamento de Z e $\langle z \rangle$ nos casos a) e b). Em ambos, as quantidades inicialmente possuem valor $Z = 1$ e $\langle z \rangle = 0$ (rede completamente conectada com uma só componente) e a medida que o processo se desenvolve, os valores convergem para $Z = 0$ e $\langle z \rangle = 1$ (rede formada por vértices isolados). Ao contrário, a quantidade ℓ possui comportamento diferenciado para a remoção de vértices e de conexões. No primeiro, o valor de ℓ é sempre decrescente e no segundo ele aumenta e depois diminui. Na prática, o que ocorre em redes com muitos vértices é que o valor de ℓ sempre aumenta. Esse aumento ocorre até que a influência do efeito de tamanho finito torne-se forte o suficiente para deter o crescimento [115, 116].

Neste trabalho estaremos particularmente interessados no processo de falha de conexões, ou seja, nos processos em que conexões são removidas aleatoriamente da rede [115, 116, 120, 121]. Vamos denotar este processo por "despercolação". Este nome deve-se ao fato deste processo ser o inverso da percolação onde conexões são adicionadas aleatoriamente à rede enquanto algumas quantidades de interesse (como Z , $\langle z \rangle$ e ℓ , por exemplo) são monitoradas. Além disso, iremos

dar preferência à grandeza Z para quantificar a vulnerabilidade da rede de interesse pois, como veremos adiante, ela é a que possui mais significado biológico.

Na figura 4.19 apresentamos o comportamento de Z para os três modelos de rede apresentados na figura 4.17 quando submetidos à despercolação. Pelo gráfico vemos que a rede de Barabási-Albert com parâmetro $m = 1$ é muito frágil e após cerca de 30% das conexões removidas, o tamanho da maior componente cai pela metade. Ao contrário quando o parâmetro m possui valor 2, o comportamento de Z assemelha-se ao da rede aleatória de Erdős-Rényi considerado. Ambos mostram-se menos vulneráveis à remoção de conexões e, após cerca de 30% das conexões serem removidas, o tamanho da maior componente é praticamente inalterado. Neste ponto, podemos ver como a presença de ciclos interfere de maneira positiva para a robustez das redes complexas, já que a primeira rede não os possui. Os caminhos alternativos criados pelos ciclos garantem que a rede irá se manter conectada por mais tempo, tornando-se menos susceptível a falhas aleatórias das conexões.

5 *O Sistema Haversiano pelas Redes Complexas: Resultados e Discussões*

Neste trabalho, nos restringimos ao estudo da rede estática formada pelos canais de Havers e pelos canais de Volkman na região cortical da falange distal do primeiro dedo da pata direita de uma galinha jovem - figura 5.1. Mapeamos esta estrutura em um grafo e em seguida estudamos suas propriedades topológicas, bem como algumas propriedades dinâmicas, de acordo com a metodologia proposta no capítulo anterior. Abaixo, apresentemos e discutimos os resultados obtidos.

5.1 O Modelo Ósseo Utilizado

O osso utilizado como modelo neste estudo foi a falange distal do primeiro dedo da pata direita de uma galinha (figura 5.1) necropsada no Setor de Patologia e Veterinária do Hospital da Universidade Federal de Uberlândia, Brasil. Após os procedimentos histológicos clássicos, o osso foi seccionado em 800 seções seriais com $5\mu m$ de espessura. Em seguida, cada seção serial foi digitalizada seguindo o procedimento Schmorl [7]. A digitalização foi realizada em um microscópio *Olympus Triocular BX40*, conectado a uma *Camêra Oly-200* interligada a um

Figura 5.1: Ilustração da pata direita de uma galinha. Em destaque apresentamos a falange distal do primeiro dedo, de onde provém o osso utilizado neste trabalho.

computador PC por uma placa digitalizadora *Datatranslation 3135*.

5.2 Reconstrução 3D Computadorizada

A estrutura tridimensional do sistema de canais do osso cortical foi reconstruída a partir da sequência de imagens utilizando o software e os procedimentos descritos no segundo capítulo desta dissertação. A figura 5.2 apresenta as seguintes perspectivas da estrutura: a) visão frontal da parte anterior, b) visão frontal da parte posterior, c) visão lateral e d) visão em declive. A imagem 5.2 apresenta as reconstruções para diferentes valores de iso-superfície. Percebemos que, a medida que os valores de iso-superfície tornam-se menores, somente os canais mais largos podem ser visualizados.

A qualidade da reconstrução é fortemente limitada pela qualidade das secções histológicas. Estas por sua vez, são comprometidas pelos artefatos da técnica histológica empregada e por problemas de manuseamento durante a geração das lâminas histológicas. Nos casos em que houve comprometimento parcial ou completo da secção histológica, a imagem foi substituída por sua antecessora. Acreditamos que este procedimento não afetou a fidelidade da reconstrução 3D perante a estrutura real, já que a espessura de cada seção corresponde a cerca de 1/3 do diâmetro médio dos canais. Mesmo assim, é evidente que reconstruções baseadas em dados tomográficos ou de raio-X são mais fiéis e de melhor qualidade [9, 10, 18].

Para nossos propósitos, a reconstrução baseada nos dados histológicos foi suficiente e nos ajudou a perceber algumas peculiaridades da estrutura de canais do osso cortical. A primeira delas é a clara preferência dos canais a se alinharem paralelamente ao canal medular (veja figura 5.2c). Os canais que possuem esta tendência são denominados canais de Havers. Aqueles perpendiculares ao canal medular e que conectam os canais de Havers são denominados canais de Volkman. Em segundo lugar, destacamos a existência de prolongamentos ósseos na extremidade posterior, como pode ser visto na figura 5.2b. Analisando a geometria desta estrutura e relacionando-a com sua posição no dedo da galinha, percebemos que os prolongamentos servem de apoio e tornam o conjunto mecânicamente estável.

5.3 Mapeamento da Estrutura Biológica em um Grafo

As imagens digitalizadas foram submetidas a um pré-processamento visando a padronização e correção de defeitos. As principais etapas deste pré-processamento foram: i) rotação e ii) alinhamento. Na figura 5.3 mostramos uma das imagens da porção posterior do osso após ser

Figura 5.2: Algumas perspectivas da reconstrução tridimensional da estrutura Haversiana.

Figura 5.3: Secção 205 do osso considerado nesse estudo. Percebemos que esta imagem é formada pela rotação, alinhamento e união de três imagens geradas por microscópio.

submetida à fase de pré-processamento.

Em seguida, realizamos a segmentação dos canais de Havers e Volkman manualmente. De forma a diferenciar cada trecho dos canais, eles foram visualizados com cores distintas, como é mostrado na figura 5.4. Algumas perspectivas da estrutura tridimensional reconstruídas a partir das imagens segmentadas podem ser vistas na figura 5.2. É interessante perceber que existem duas saliências na região posterior do osso. Elas são chamadas de projeções ósseas e têm a finalidade de dar sustentação e manter o equilíbrio mecânico da estrutura.

No decorrer do processo de segmentação dos canais, mantivemos um arquivo texto sempre atualizado com as conexões entre as cores. Por exemplo, para o canal mostrado na figura 5.4, teríamos a seguinte anotação: vermelho - azul / verde - azul (as anotações foram feitas em termos das componentes *RGB* de cada cor). Este arquivo de dados foi utilizado em seguida para criação do grafo do sistema Haversiano, como mostramos na figura 5.6.

Ao fazer este mapeamento, estamos assumindo que a intersecção dos canais representam os vértices do grafo e as conexões são representadas pelos canais. Esta hipótese é muito útil para estudar sistemas rígidos em termos de grafos. Em [122] este método, aliado aos conceitos de redes complexas, foi utilizado pelos autores para caracterizar as propriedades topológicas das galerias subterrâneas construídas por formigas. Ele também foi, recentemente, empregado por Buhl *et al.* [30] e por Laemer *et al.* [123] para estudar redes de ruas urbanas de acordo com a teoria das redes complexas.

Durante o processo de segmentação os canais que realizavam conexões com o canal medular ou

Figura 5.4: Exemplo de bifurcação de um canal. No processo de segmentação, cada trecho é marcado com uma cor.

Figura 5.5: A primeira sequência mostra uma região da estrutura cortical onde pode ser visto um canal que estabelece uma conexão com o tecido. Na segunda sequência mostramos a segmentação deste canal (azul) e do tecido (preto).

com o tecido também foram destacados. Este procedimento está exemplificado na figura 5.5.

O passo seguinte foi escrever uma rotina para capturar a posição espacial dos canais. Como entrada para esta rotina foram passadas as componentes *RGB* da cor correspondente ao canal desejado. Para um dado valor, a rotina realizava uma busca (sempre no sentido positivo do eixo $\hat{z}$) por todo volume da estrutura haversiana e verificava a última posição em que essa cor ocorria. Então esta posição era assimilada a um novo vértice. Em seguida, estabelecemos as conexões entre os vértices utilizando o arquivo contendo a lista de conexões. Um esquema deste processo é ilustrado na figura 5.6.

A rede haversiana resultante deste procedimento é apresentada na figura 5.7. Na tabela 5.1 apresentamos algumas características da rede haversiana e as comparamos com algumas redes biológicas reais.

Figura 5.6: Esquema utilizado para mapear a estrutura Haversiana em um grafo. Além de ser utilizada para realizar a reconstrução tridimensional da rede de canais, a sequência de imagens é processada por uma rotina externa que fornece a posição dos vértices (bifurcações). Em seguida estes vértices são conectados de acordo com as conexões presentes no arquivo de dados.

Figura 5.7: Representação da Haversiana. As perspectivas a, b, c e d desta imagem acompanham as da figura 5.2

Rede	N	E	$\langle k \rangle$	$\check{k}$	$\langle C \rangle$	γ	Ref.
haversiana	5592	6643	2,37	12	0,028	6,0	—
haversiana de gato	988	1120	2,26	9	0,009	3,8	[124]
proteína-proteína (<i>S. cerevisiae</i>)	1870	4469	2,39	56	0,07	2,4	[3]
cadeia alimentar	178	801	4,5	27	0,15	1,13	[125]
metabólica (<i>E. coli</i>)	778	5757	7,4	110	—	2,2	[32]

Tabela 5.1: Características gerais da rede haversiana e de outras redes biológicas. Na rede de interação proteína-próteína, cada vértice representa uma proteína e dois vértices são conectados se as proteínas correspondentes possuem a capacidade de ligar-se entre si. A rede da cadeia alimentar estudada por Montoya *et al.* foi gerada após estudos das relações alimentares existentes em determinados parques ecológicos. Os dados apresentados na tabela correspondem ao parque Silwood (localizado no *Colégio Imperial de Londres*). As redes metabólicas são redes biológicas onde os vértices representam componentes moleculares e, de acordo com [3], as ligações que apontam para um determinado componente representam as reações em que ele participa como produto.

5.4 Medidas Clássicas

Como vimos no capítulo sobre redes complexas, as árvores não possuem ciclos e o número de conexões nelas existente é $E = N - 1$. Assim, pela equação 4.4 vemos que a conectividade média destas redes é aproximadamente $\langle k \rangle \approx 2$. Este valor é próximo aquele apresentado na tabela 5.1 para a rede haversiana. Isso explica o baixo número de ciclos que observamos nesta rede. O número estimado de conexões responsáveis por ciclos é $E - (N - 1)$ que corresponde a aproximadamente $\approx 15\%$ da conexões desta rede. A baixa densidade de ciclos explica o baixo valor encontrado para o coeficiente de clusterização médio, também apresentado na tabela 5.1.

A fim de complementar os dados de conectividade apresentados na tabela 5.1 apresentamos na figura 5.8a o gráfico da distribuição de conectividade da rede haversiana. É evidente a tendência linear dos valores de $P(k)$ em função de k para $k > 2$. Isso mostra que nesta região a conectividade da rede é governada por uma lei de potência como aquela representada pela equação 4.17. De fato, o ajuste destes pontos por esta lei de potência nos fornece o expoente de escala $\gamma = 6,0$. Este alto valor para o expoente de escala reflete o fato da rede estar imersa em um espaço físico real. Assim, suas conexões respeitam regras de adjacência espacial que tornam muito improvável a ocorrência de vértices com conectividade maior que uma dezena. Na figura 5.8b apresentamos o gráfico da correlação $P(k, k')$, ou seja, a probabilidade de que um vértice com conectividade k conecte-se à outro com conectividade k' . Vemos que o valor máximo desta quantidade ocorre para $k = k' = 3$. Este resultado é compatível com o que foi obtido em nosso trabalho anterior [124].

Além disso, percebemos que os valores de $P(k)$ para $k = 1$ fogem a aproximação pela lei de

Figura 5.8: a) Distribuição de conectividade da rede Haversiana e b) representação da correlação $P(k, k')$.

potência. Uma possível explicação para este fato é que os vértices com conectividade unitária representam terminações de um canal e, como já foi observado, estas regiões terminais não possuem somente a função de dar suporte ao transporte de nutrientes. Elas estão intimamente relacionadas aos locais de desenvolvimento da rede, porque são nestes locais que ocorrem o surgimento e a bifurcação de novos canais.

Na figura 5.9a associamos o valor da conectividade de cada vértice à uma cor e transparência. Tons azulados e com alta transparência indicam valores baixos de conectividade, enquanto tons vermelhos e sem transparência indicam alta conectividade. Podemos perceber que existe uma distribuição espacial homogênea dos valores de k e uma dominância do tom verde, correspondente ao valor de $k = 3$. Na figura 5.9b apresentamos a disposição espacial dos *hubs* da rede haversiana (consideramos todos os vértices com conectividade $k > 4$). Aqui percebemos o surgimento de algumas regiões concentradoras de hubs. Estas regiões são, em geral, dispostas paralelamente ao canal medular.

A distribuição dos valores do coeficiente de clusterização para cada vértice é apresentada na figura 5.9c. Nela vemos que existe uma dominância de altos valores na região posterior e o surgimento de duas faixas paralelas ao canal medular com valores próximos a 0,5. As faixas localizam-se em média próximas ao canal medular, evidenciando a presença de ciclos do tipo triângulo nestas regiões.

Na próxima figura, figura 5.10, apresentamos a distribuição do comprimento dos menores caminhos (l) da rede haversiana. Além disso, realizamos duas análises adicionais para estudar a importância dos ciclos e o efeito das restrições espaciais nesta medida.

Para isso, consideramos as seguintes redes: i) rede haversiana, ii) árvore de menor cobertura da rede haversiana e iii) rede haversiana aleatória. A segunda foi obtida pelo algoritmo *Prim-*

Figura 5.9: a) Distribuição espacial dos valores de conectividade na rede Haversiana, b) distribuição espacial dos vértices com conectividade $k > 4$ e c) distribuição espacial dos valores do coeficiente de clusterização.

Dijkstra [89]. O funcionamento deste algoritmo é discutido no Apêndice 3. A terceira rede foi obtida reorientando aleatoriamente todas as conexões da rede haversiana. É interessante notar que essa reorientação aleatória quebra a restrição espacial imposta às conexões pela matriz óssea. Isso faz com que a rede haversiana passe por uma transição similar àquela sofrida pela redes regulares que dão origem as redes mundo pequeno. Portanto, esperamos que a versão aleatória da rede haversiana apresente efeito de mundo pequeno.

Figura 5.10: Distribuição do comprimento médio dos menores caminhos ao longo da rede haversiana. Também mostramos esta distribuição avaliada sobre a rede haversiana aleatória e sobre a árvore haversiana de menor cobertura. No inset apresentamos o mesmo gráfico em escala $\log \times \log$.

Seguindo a terminologia desta dissertação, o valor médio das distribuições apresentadas na figura 5.10 representa o diâmetro da rede em questão. Assim, vemos que a rede haversiana possui diâmetro em torno de $\langle l \rangle \approx 32$. Removendo os ciclos desta rede, vemos que o valor

Figura 5.11: O vértice fonte introduzido na rede conecta-se à todos os vértices que estabelecem conexão direta com a medula óssea.

de $\langle l \rangle$ aumenta para $\langle l \rangle \approx 60$. Isso mostra que os ciclos, além de criarem caminhos alternativos para a passagem de nutrientes, são organizados de modo a tornar significativamente menor o caminho entre duas confluências quaisquer e conseqüentemente entre locais diferentes da matriz óssea.

Os efeitos da reorientação aleatória ficam claros quando observamos a curva de cor preta na figura 5.10. Vemos que ela é descrita por uma distribuição normal (típica para redes aleatórias) com baixa variância e valor médio da ordem de $\langle l \rangle \approx 12$. Este diâmetro é significativamente menor do que aquele da rede haversiana e é bem aproximado pela expressão 4.19 que resulta no valor $\langle l \rangle \approx 10$. Assim, confirmamos a ocorrência de uma transição mundo pequeno na rede haversiana decorrente da reorientação aleatória de suas conexões.

Para investigarmos a eficiência da rede haversiana com relação ao transporte de nutrientes, utilizamos dois critérios propostos em [4]. O primeiro, denominado fator de rota (q_1), estabelece a eficiência da rede para ligar um vértice fonte aos demais vértices. O segundo critério indica o quanto a estrutura da rede difere de sua árvore de menor cobertura.

O primeiro passo para investigação destas características foi estabelecer um vértice fonte para a rede haversiana. Para isso, unimos em um único vértice (o vértice fonte), todos aqueles que estabeleciam uma ou mais conexões com a medula. Este procedimento está ilustrado na figura 5.11. Em seguida as medidas de eficiência foram estimadas para esta nova rede. Os resultados estão sumarizados na tabela 5.2, bem como os valores destas quantidades para outras redes reais da literatura.

Pela tabela, vemos que a rede haversiana possui alto fator de rota, indicando que em média, o menor caminho que liga a fonte ao vértice i (ao longo das conexões da rede) é 2,4 vezes maior do que a distância euclidiana que separa estes vértices. O alto fator de rota da rede haversiana indica que a estrutura de canais é tortuosa e conseqüentemente, que ela abrange melhor a matriz óssea. Assim, o material difundido através das superfícies dos canais alcança um número maior

Figura 5.12: Ilustração do caminho que liga o vértice fonte (0) à um vértice i qualquer. Em a) este caminho é retilíneo, implicando em um fator de rota $q_1 = 1$. Em b) o caminho é mais tortuoso, com fator de rota $q_1 = 2$.

Rede	q_1	q_2
haversiana	2,4	1,38
rede de esgoto	1,59	1,18
rede de gás WA	1,13	1,27
rede de gás IL	1,48	1,63
ferrovia	1,14	1,12

Tabela 5.2: Valores de q_1 (fator de rota) e q_2 (razão entre o comprimento total de conexões de uma rede e de sua respectiva AMC) para a rede Haversiana, bem como para outras rede reais estudadas por Gastner e Newman em [4]. A rede de esgoto é referente a cidade de Bellingham, EUA, a rede de gás WA refere-se a rede de gas natural da cidade de Western, Austrália. A rede de gás IL refere-se a rede de gás natural do estado de Illinois, EUA. A rede de ferrovias refere-se ao sistema de ferrovias da cidade de Boston, EUA.

de osteócitos (figura 5.12).

Ainda na tabela 5.2 apresentamos o valor de q_2 , referente a razão entre a soma do comprimento de todas as conexões da rede haversiana e a soma do comprimento de todas as conexões da árvore haversiana de menor cobertura. Vemos que o valor $q_2 = 1,38$ é bem próximo dos demais apresentados para outras redes reais, mostrando que deve existir um processo natural de otimização na construção dos canais do osso cortical.

5.5 Medidas Hierárquicas

Utilizamos os conceitos de medidas hierárquicas para completar a caracterização topológica da rede haversiana. Na figura 5.13 apresentamos as curvas das funções $n(h) \times h$ e $k(h) \times (h)$. Estas curvas são muito similares, refletindo a baixa densidade de circuitos fechados na rede. Vale lembrar que para uma árvore (rede sem circuitos fechados) temos $n(h) = k(h)$. Além disso, fica

Figura 5.13: Propriedades hierárquicas da rede haversiana: número hierárquico de vértices e conectividade hierárquica. Em cinza mostramos a aproximação por campo médio para uma rede infinita com conectividade média similar a da rede haversiana. Também destacamos o valor de pico da rede haversiana: $h_{pico} = 32$.

claro o valor de pico para o qual as medidas começam a sofrer influências do efeito de tamanho finito da rede. Estimamos este valor como $h_{pico} = 32$. Também apresentamos em cinza a curva esperada pela teoria de campo médio para os valores de $n(h)$ em uma rede infinita com mesma conectividade média que a rede haversiana.

Outra propriedade hierárquica que investigamos neste trabalho foi o coeficiente de clusterização hierárquico. O resultado desta análise está sumarizado na figura 5.14. Nesta figura apresentamos a curva $C(h) \times h$ em escala $\log \times \log$ bem como em escala linear no inset do gráfico. Observamos um decaimento exponencial do coeficiente de clusterização hierárquico em função da hierarquia. De fato, os dados são bem ajustados por uma curva do tipo $C(h) \propto e^{-\mu h}$, com expoente de decaimento $\mu = 0.15$.

O caráter assintoticamente decrescente do coeficiente de clusterização hierárquico observado aqui contrasta com os resultados obtidos em nosso trabalho prévio sobre a estrutura de canais do osso cortical [124]. Neste trabalho os resultados sugerem que a curva de $C(h)$ possui um pico para baixas hierarquias ($h \approx 3$) e decresce em seguida.

A explicação para esta discrepância é obtida levando-se em conta que os trabalhos utilizam diferentes espécies animais como objeto de estudo. Em nosso trabalho prévio usamos o fêmur de um gato adulto, enquanto nesta dissertação de mestrado o osso utilizado foi a falange distal de uma galinha jovem. Diferenças entre a estrutura de canais haversianos em diferentes espécies animais já haviam sido observadas por R. Georgia *et al.* [126, 127] por volta de 1982. Além

Figura 5.14: Coeficiente de clusterização hierárquico da rede haversiana. No inset apresentamos o mesmo resultado em escala linear e um ajuste através de um decaimento exponencial $e^{-\mu h}$, com $\mu \approx 0,15$.

disso, há indícios de que existem diferenças entre estas estruturas até mesmo em animais de mesma espécie porém, com idades diferentes.

5.6 Comunidades

Para estudar a organização interna da rede haversiana utilizamos o conceito de comunidades. Uma comunidade é formada por um subconjunto de vértices que estabelecem mais conexões entre si do que com o restante da rede. O algoritmo utilizado para detectar a presença de comunidades na rede haversiana foi proposto por Aaron Clauset *et al.* [2] e pode ser obtido no endereço eletrônico <http://www.cs.unm.edu/aaron/>. Basicamente, o algoritmo busca otimizar uma quantidade denominada modularidade. A modularidade é uma medida que avalia uma dada configuração de comunidades. O valor máximo da modularidade é 1, indicando uma rede altamente modular e 0 para uma rede onde não existem comunidades (rede aleatória).

Os valores da modularidade da rede haversiana em função do passo do algoritmo são apresentados na figura 5.15. O pico de modularidade ocorre com valor $Q = 0.95$ indicando a forte presença de comunidades nesta rede. Este valor de modularidade corresponde a uma configuração com 65 comunidades com tamanhos médios da ordem de algumas dezenas de vértices. Entretanto, no nosso caso não é interessante uma configuração como esta, já que não teria significado qualquer propriedade estatística sobre esta quantidade de vértices.

comunidade	N	$\langle k \rangle$	$\langle C \rangle$	$\langle l \rangle$	l_{max}	$\langle D \rangle$
1	473 0.08	2.75 ± 1.16 1.16	0.03 ± 0.12 0.90	9 ± 12 0.29	24 0.29	68 ± 38 0.98
2	680 0.12	2.75 ± 1.23 1.16	0.04 ± 0.13 1.21	10 ± 16 0.32	28 0.34	63 ± 49 0.91
3	440 0.07	2.56 ± 1.16 1.08	0.04 ± 0.14 1.21	11 ± 15 0.35	27 0.32	56 ± 35 0.81
4	323 0.05	2.21 ± 1.26 0.93	0.02 ± 0.08 0.60	16 ± 22 0.51	43 0.51	51 ± 24 0.74
5	439 0.08	2.47 ± 1.20 1.04	0.05 ± 0.15 1.51	15 ± 26 0.48	40 0.48	76 ± 41 1.10
6	701 0.13	2.16 ± 1.25 0.91	0.02 ± 0.08 0.60	32 ± 80 1.03	81 0.97	61 ± 27 0.88
7	154 0.03	2.18 ± 1.23 0.92	0.03 ± 0.10 0.9	13 ± 13 0.42	36 0.43	74 ± 28 1.07
8	231 0.04	2.20 ± 1.25 0.93	0.01 ± 0.06 0.30	13 ± 15 0.42	31 0.37	50 ± 25 0.72
9	568 0.10	2.20 ± 1.27 0.93	0.03 ± 0.09 0.90	20 ± 36 0.64	51 0.61	84 ± 29 1.22
10	276 0.05	2.19 ± 1.24 0.92	0.02 ± 0.09 0.60	23 ± 36 0.74	59 0.71	71 ± 24 1.03
11	576 0.10	2.15 ± 1.27 0.90	0.02 ± 0.09 0.60	32 ± 68 1.03	81 0.97	84 ± 30 1.22
12	172 0.03	2.19 ± 1.20 0.92	0.03 ± 0.11 0.9	14 ± 16 0.45	38 0.46	65 ± 26 0.94
13	121 0.02	2.05 ± 1.25 0.86	0.02 ± 0.11 0.60	10 ± 9 0.32	26 0.31	72 ± 33 1.04
14	155 0.03	2.18 ± 1.26 0.92	0.03 ± 0.10 0.90	14 ± 16 0.45	36 0.43	84 ± 31 1.21
15	283 0.05	2.12 ± 1.24 0.89	0.01 ± 0.07 0.30	20 ± 30 0.64	54 0.65	86 ± 34 1.24

Tabela 5.3: Propriedades topológicas e espaciais das comunidades presentes na rede haversiana. A segunda linha de cada comunidade corresponde a razão entre o valor medido na comunidade e o valor medido na rede completa. As quantidades tabeladas são: N - número de vértices, $\langle k \rangle$ - conectividade média, $\langle C \rangle$ - coeficiente de clusterização médio, $\langle l \rangle$ - diâmetro, l_{max} - comprimento do maior menor caminho e $\langle D \rangle$ - média da distância entre os vértices e a superfície do canal medular.

Figura 5.15: Evolução do valor da modularidade da rede Haversiana durante a execução do algoritmo de detecção de comunidades [2]. As linhas vermelhas marcam o valor adotado como modularidade da rede Haversiana neste trabalho.

Para contornar esta situação executamos novamente o algoritmo e deixamos que ele avançasse até que a modularidade atingisse o valor $Q = 0.85$ depois do pico. Usando este procedimento o algoritmo é capaz de agrupar comunidades que contribuem pouco para a modularidade. Como resultado, obtivemos uma configuração final constituída por 15 comunidades. As principais propriedades de cada uma são apresentadas na tabela 5.3.

Na figura 5.16 apresentamos a disposição espacial das comunidades na rede haversiana, pela qual vemos que as comunidades topológicas também são comunidades espaciais, ou seja, são bem concentradas em uma região do espaço. Esse fenômeno se deve a ausência de conexões de longo alcance (efeito mundo pequeno) que por sua vez é reflexo das restrições espaciais impostas pela matriz ósea. Além disso, percebemos que existem formas espaciais distintas, que possivelmente poderiam ser utilizadas para categorização destas comunidades. É interessante notar que as comunidades 1 e 2, mais densamente conectadas segundo a tabela 5.3, estão localizadas exatamente sobre a região das projeções ósseas.

A organização topológica das comunidades é apresentada na figura 5.17. Aqui fica visível a maior densidade de conexões nas comunidades 1 e 2. Além disso, percebemos que existe um gradiente decrescente de ciclos a medida que avançamos no índice das comunidades. É interessante perceber que, percorrer estes índices corresponde, a grosso modo, a caminhar no sentido *posterior* $\rightarrow$ *anterior* do osso.

Para compreendermos melhor diferença entre as comunidades da região posterior e da região anterior, olhamos para as características hierárquicas ($k(h)$ e $C(h)$) das comunidades 1 (região

Figura 5.16: Disposição espacial das comunidades da rede Haversiana.

Figura 5.17: Organização topológica das comunidades da rede Havrasi.

Figura 5.18: a) Curvas de conectividade hierárquica média das comunidades 1 (vermelho) e 10 (preto) e b) coeficiente de clusterização hierárquico médio das mesmas comunidades.

posterior) e 10 (região anterior). Na figura 5.18 apresentamos as curvas de conectividade hierárquica e do coeficiente de clusterização hierárquico destas comunidades (as curvas representam os valores médios tomados sobre todos os vértices da comunidade). Na figura, percebemos que as comunidades realmente possuem características muito distintas. Pela figura 5.18a vemos que a comunidade 1 é mais bem conectada e mais compacta (baixo número de hierárquicas), enquanto a comunidade 10 é bem menos compacta e possui baixa conectividade para todas as hierarquias. Através da figura 5.18b vemos que o valor do coeficiente de clusterização das comunidades 1 e 10 é praticamente o mesmo que o da rede haversiana completa quando $h = 1$ e decrescem em seguida, como a rede haversiana (figura 5.14). Porém, ao contrário do comportamento assintoticamente decrescente de $C(h)$ para a rede completa, aproximadamente em $h = 10$ os valores de $C(h)$ para ambas comunidades começam a aumentar até atingirem um pico em $h = 15$ para a comunidade 1 e $h = 27$ para a comunidade 10.

A distinção entre as comunidades da região posterior e anterior, talvez possa ser explicada pela adaptação óssea. Isso porque a região posterior é responsável pela sustentação e pelo equilíbrio mecânico da estrutura e está sujeita a forças externas diferentes do restante do osso. Como foi dito no capítulo introdutório, muitos estudos comprovaram que as variações externas de pressão são responsáveis por criar fluxo de nutrientes nos canais haversianos. Aqui, inferimos que elas alteram não só o fluxo, mas atuam de maneira mais geral, interferindo na própria construção e organização dos canais.

Apesar de ocorrerem com menor intensidade, conexões entre comunidades também são observadas no sistema. A figura 5.19 mostra (em cinza) como estão organizadas as conexões entre as comunidades da rede haversiana. Percebemos que as conexões distribuem-se uniformemente ao longo de cada comunidade, demonstrando que não há regiões preferenciais. Entretanto, o

Figura 5.19: Organização das conexões estabelecidas entre as comunidades da rede Haversiana. Quando consideramos cada comunidade como um único vértice, podemos perceber que existe uma semelhança entre esta imagem e aquela apresentada na figura 4.5b. Este fato indica a existência do efeito mundo pequeno na rede Haversiana a nível das comunidades.

número de conexões realizadas por uma determinada comunidade é proporcional ao número de vértices que ela possui.

Um fato interessante surge quando consideramos uma nova rede, em que cada vértice corresponde a uma comunidade. Obviamente, dois vértices estão conectados se existem conexões que ligam as respectivas comunidades. Esta nova rede é ilustrada na figura 5.20a. Na figura 5.20b mostramos a distribuição do comprimento dos menores caminhos nesta rede. A presença de um alto coeficiente de clusterização ($C \approx 0.48$) e diâmetro dado pela equação 4.19 ($\langle l \rangle \approx 1,6$ e $\log 15 / \log 4,4 \approx 1,82$) indicam a presença de efeito mundo pequeno nesta rede. De maneira surpreendente, este efeito emerge quando olhamos para a rede haversiana sob a escala das comunidades. A ocorrência do efeito mundo pequeno nesta rede evidencia o fato de que a rede haversiana é organizada de maneira diferente em escalas diferentes.

Figura 5.20: a) Rede obtida quando consideramos cada comunidade da rede Haversiana como um vértice. Dois vértices estão conectados se há ao menos uma conexão entre as comunidades correspondentes e b) distribuição do comprimento dos menores caminhos nesta rede. O valor médio da distribuição é $\langle l \rangle \approx 1,6$.

5.7 Robustez da rede Haversiana

Ao longo do desenvolvimento ósseo, os canais de Havers e de Volkman estão sujeitos a alterações estruturais. Estas alterações conduzem a obstruções temporárias, novas bifurcações ou até mesmo na extinção de alguns canais. Para modelar os efeitos das obstruções aleatórias na rede Haversiana utilizamos o conceito de despercolação. Durante este processo, conexões são retiradas da rede aleatoriamente enquanto monitoramos o tamanho da sua maior componente. Esta medida indica o quão robusta é a rede sob o ponto de vista da perda de canais. Nos casos em que a resistência é baixa, a rede se fragmenta rapidamente e torna-se inútil para exercer funções de transporte, por exemplo.

Na figura 5.21 mostramos os resultados deste processo para a rede Haversiana (em azul) bem como para cada uma de suas comunidades. Comparando a curva obtida para a rede Haversiana com a figura 4.19, percebemos que sua resistência encontra-se entre a resistência do modelo Barabási-Albert com $m = 1$ e deste modelo com $m = 2$. Isso é mais ou menos compatível com a variação do grau médio nessas redes: $\langle k \rangle = 2$ ($m = 1$), $\langle k \rangle = 2,37$ (Haversiana) e $\langle k \rangle = 4$ ($m = 2$).

Um fato interessante é que algumas comunidades da rede Haversiana são mais resistentes do que a rede completa, enquanto a maioria das comunidades possui resistência similar ou inferior a da rede completa. Outra curiosidade é que as quatro comunidades mais resistentes do que a rede Haversiana são exatamente aquelas localizadas na região posterior, a saber: 1,2,3 e 5. Mais um indício de que a posição geográfica exerce influência na construção dos canais.

Figura 5.21: Tamanho da maior componente conectada (escala $\log \times \log$) durante o processo de despercolação: i) na rede Haversiana (em azul) e ii) nas comunidades da rede Haversiana (linhas pretas). As comunidades estão agrupadas na legenda de acordo com a proximidade de suas curvas. Além disso, os índices correspondem aqueles da figura 5.16.

Se as comunidades mais resistentes estão na região posterior, estarão as menos resistentes na região anterior? Na verdade isso não ocorre. Podemos perceber pela figura que as comunidades menos resistentes são a 6 e a 11. Pela figura 5.16, vemos que a primeira está localizada exatamente na região central e a segunda na região anterior. Curiosamente, as duas comunidades possuem o mesmo aspecto espacial na forma de anel. Isso sugere a existência de uma relação forma-função [52] para as comunidades da rede haversiana. Além disso, a estratégia de distribuir as comunidades menos resistentes ao longo da matrix óssea garante que nenhuma região seja fragilizada em demasia, o que ocorreria por exemplo, se as comunidades 6 e 11 estivessem na região anterior.

O fluxo máximo que escoar na rede Haversiana durante o processo de despercolação também foi monitorado. Porém, para utilizar a teoria desenvolvida no capítulo anterior, precisamos definir um vértice de destino para o material que está sendo transportado. Para isso, procedemos como na definição do vértice fonte, ou seja, introduzimos um novo vértice à rede conectado a todos os vértices que estabelecem conexão direta com o perióstio ósseo. Assim, a rede haversiana pode ser esquematizada da seguinte maneira:

O objetivo deste estudo é verificar as influências da remoção aleatória de conexões no escoamento de material na rede Haversiana, ainda que a interpretação biológica do vértice de destino não seja direta, já que o perióstio também possui vasos sanguíneos que irrigam a matriz óssea.

O resultado da medida de fluxo máximo durante a despercolação está apresentado em vermelho na figura 5.22, onde vemos que a rede Haversiana é muito pouco robusta sob este aspecto, pois o fluxo máximo torna-se praticamente zero aproximadamente após 50% das conexões serem removidas aleatoriamente.

Este resultado demonstra que a rede Haversiana não é otimizada para suportar tais danos, e que uma disfunção causadora de obstruções aleatórias em um número razoável de canais pode danificar todo tecido ósseo de maneira irreversível. Obviamente as obstruções causadas pelo processo de remodelagem, apesar de serem aleatórias, não ocorrem com tanta intensidade (a ponto de obstruir 50% dos canais) e conseqüentemente este estado crítico não é observado, em geral.

Acreditamos os danos causados por fraturas e lesões comprometam um número maior de canais do que o processo de remodelagem. Para estudar seus efeitos sobre o escoamento de nutrientes através do osso utilizamos um processo de ataque, no qual o mecanismo de seleção baseia-se na posição espacial dos vértices. O funcionamento deste mecanismo é bastante simples e consiste

Figura 5.22: Fluxo máximo que escoar na rede Haversiana durante o processo de despercolação a) aleatório (quadrados vermelhos) e b) coordenado espacialmente (círculos pretos). Φ_m^* representa o máximo valor do fluxo máximo na rede Haversiana.

Figura 5.23: Evolução da rede Haversiana submetida ao processo de ataque coordenado espacialmente.

em remover todos os vértices que estejam no interior de um cilindro com raio r . Na figura 5.23 mostramos a rede Haversiana submetida a este processo de ataque com diferentes valor de r .

O resultado obtido é apresentado na figura 5.21 em preto, onde vemos uma considerável diferença do processo de despercolação. Através desta figura, percebemos que o fluxo na rede Haversiana é bem menos sensível a este tipo de dano, espacialmente localizado. Quando cerca de 50% das conexões são removidas através deste processo, o fluxo máximo cai para cerca de 60% do seu valor máximo. Isso garante que lesões de grandes proporções não causem a morte do tecido ósseo.

6 Conclusão

Analizamos, nesta dissertação, o sistema de canais contidos no interior do osso cortical dos animais sob o ponto de vista das redes complexas. Para isso, utilizamos como modelo de osso a falange distal do primeiro dedo da pata direita de uma galinha jovem. O osso foi histologicamente processado, resultando em uma sequência de imagens digitais. As imagens foram segmentadas e utilizadas para reconstrução do sistema de canais. Para isso, desenvolvemos softwares e ferramentas específicas em ambiente C++ integrado à biblioteca VTK (*Visualization ToolKit*). A reconstrução tridimensional do sistema de canais nos permitiu visualizar as diferenças estruturais entre as regiões do osso cortical, sendo a principal delas a presença de prolongamentos ósseos na porção anterior, que não se verifica na região oposta.

Mapeamos a estrutura de canais em um grafo, considerando as intersecções de um ou mais canais como seus vértices. Neste grafo aplicamos os conceitos desenvolvidos recentemente na área das redes complexas para extrair informações sobre a organização topológica da estrutura Haversiana. Através de medidas como conectividade, coeficiente de clusterização e distribuição do comprimento dos menores caminhos, mostramos que a rede Haversiana é altamente esparsa, com conectividade média $k = 2,37$ e apresenta uma distribuição de conectividade que obedece uma lei de potência com expoente de escala $\gamma = 6,0$. Também mostramos que a rede Haversiana não apresenta efeito mundo pequeno, fato este que reflete o caráter geográfico do sistema.

Complementamos estas medidas com suas versões hierárquicas. Elas tornaram possíveis a caracterização da rede de maneira mais global. Mostramos que o coeficiente de clusterização hierárquico é aproximado por $C(h) \propto e^{-\mu h}$, com $\mu \approx 0,15$. Este resultado não é compatível com aquele apresentado em nosso primeiro trabalho [124] no qual estudamos o sistema de canais presentes na porção cortical do fêmur de um gato adulto. Este fato evidencia a diferença entre a estrutura de canais em diferentes espécies, como já havia sido observado em trabalhos anteriores [126, 127].

Alguns aspectos de eficiência também foram investigados, segundo os critérios estabelecidos em [4]. O valor do fator de rota para a rede Haversiana, $q_1 = 2,4$, mostra que para alcançar dois

vértices, caminhando pela rede, percorrer aproximadamente 2,4 vezes a distância que separa os dois vértices. Este valor é elevado comparado com outras redes reais porém, demonstra que os canais são tortuosos e isso aumenta a eficiência com que os canais cobrem a matriz óssea. A razão entre o comprimento total das conexões da rede Haversiana e o comprimento das conexões de sua AMC, $q_2 = 1,38$ é compatível com outras redes reais e indica um possível processo de otimização durante a criação desta rede.

Mostramos também que a rede Haversiana é organizada de tal maneira que existem regiões mais bem conectadas, denominadas comunidades, de tal maneira que conexões ocorrem com mais frequência entre vértices da mesma comunidade. No total, detectamos 15 comunidades distribuídas ao longo da rede Haversiana, sendo cada uma espacialmente bem localizada. Nossos resultados indicaram diferenças relevantes entre as estruturas de cada comunidade, sendo que aquelas localizadas na extremidade posterior apresentam maior conectividade média e são mais compactas. Esta região se diferencia das demais pois sofre variações de pressão externa ativamente, já que ali estão localizados os prolongamentos ósseos. Como já é bem sabido na literatura, as variações de pressão são responsáveis por causar fluxo nos canalículos que ligam os canais de Havers e Volkmann até as células ósseas. Aqui inferimos que a influência da pressão externa é mais geral, interferindo na própria organização do sistema de canais.

Para estudar as conexões entre comunidades, nós consideramos uma nova rede, no qual os vértices representavam as comunidades e que dois vértices estavam conectados se existisse ao menos uma conexão entre as comunidades correspondentes. Nesta nova rede, observamos a presença do efeito mundo pequeno, devido ao valor médio de $\langle l \rangle$ ser da ordem do logaritmo do número de vértices e ao alto coeficiente de modularidade. Este resultado demonstra que a rede Haversiana é organizada de maneira diferente em escala distintas.

Sob o ponto de vista dinâmico, estudamos a rede Haversiana usando um processo de falhas aleatórias. Este processo (aqui chamado de despercolação) simula a remoção aleatória de canais, em analogia as obstruções reais que ocorrem no osso cortical devido à seu desenvolvimento. Durante este processo monitoramos o tamanho da maior componente conectada da rede, bem como de suas comunidades. Os resultados indicam que as comunidades localizadas na região posterior são significativamente mais resistentes e que as comunidades menos resistentes estão espalhadas pela estrutura óssea.

Durante o processo de despercolação também monitoramos o fluxo máximo que escoava entre a medula óssea e o periósteo ósseo. Verificamos que o processo de remoção aleatória é altamente prejudicial à este fluxo e que ele torna-se praticamente nulo quando 50% das conexões são removidas. Nós comparamos este resultado com aquele obtido quando as conexões não

são removidas aleatoriamente, mas sim através de um critério de seleção baseado na posição espacial do vértices. Chamamos este caso de ataque espacialmente coordenado. Este processo simula fraturas ou lesões na estrutura óssea onde canais adjacentes são removidos. Verificamos que a rede Haversiana é altamente otimizada para suportar este tipo de dano, sendo que após 50% das conexões serem removidas o fluxo máximo ainda é cerca de 60% do inicial.

É interessante ressaltar que no decorrer deste trabalho, muitas ferramentas para geração, análise e visualização de redes complexas foram desenvolvidas e que, possivelmente serão disponibilizadas na internet, bem como os dados aqui utilizados. Por fim, acreditamos que este tenha sido um trabalho altamente original no que diz respeito a aplicação das técnicas das redes complexas, evidenciando o caráter multidisciplinar deste novo ramo da ciência.

Nossas propostas para trabalhos futuros incluem o estudo da rede Haversiana em diferentes estágios do desenvolvimento animal e a caracterização destas redes em diferentes espécies animais. Além disso, novas ferramentas devem ser empregadas, bem como uma técnica para automatização do processo de obtenção da rede, tornando assim a segmentação manual desnecessária. Outra possibilidade é a criação de um modelo matemático para geração de redes que resgatem as propriedades observadas neste sistema.

Referências Bibliográficas

- [1] WATTS, D. J.; STROGATZ, S. Collective dynamics of "small-world" networks. *Nature*, v. 393, p. 440, 1998.
- [2] CLAUSET, A.; NEWMAN, M. E. J.; MOORE, C. Finding community structure in very large networks. *Phys. Rev. E*, v. 70, p. 066111, 2004.
- [3] DEZSO, Z.; OLTVAI, Z. N.; BARABÁSI, A.-L. Bioinformatics analysis of experimentally determined protein complexes in the yeast *saccharomyces cerevisiae*. *Genomic Research*, v. 13, p. 2450, 2003.
- [4] GASTNER, M. T.; NEWMAN, M. E. J. Shape and efficiency in spatial distribution networks. *J. of Stat. Mech.*, v. 2006, p. P01015+, 2006.
- [5] COWIN, S. C. Mechanosensation in bone under microgravity. *Bone*, v. 22, p. 119, 1998.
- [6] DEMPSTER, D. W. *Bone Remodeling*. [S.l.]: Lippincott-Raven Publishers, 1995.
- [7] JUNQUEIRA, L. C.; CARNEIRO, J. *Basic Histology*. [S.l.]: McGraw-Hill/Appleton & Lange, New York, 2002.
- [8] HAVERS, C. Some new observations of the bones and the parts belonging to them, with the manner of their accretion and nutrition,. *Communicated to the Royal Society in Several Discourses*, no-tag, p. no-tag, 1961.
- [9] COOPER, D. M. L.; HALLGRÍMSSON, B. Application of 3d μ CT to the analysis of cortical bone porosity. *Am. J. Phys. Anthropol.*, v. 117, p. 56, 2002.
- [10] COOPER, D. M. L. et al. Quantitative 3d analysis of the canal network in cortical bone by micro-computed tomography. *Anat. Rec.*, v. 274, p. 169, 2003.
- [11] WANG, L. et al. In situ measurement of solute transport in the bone lacunar-canalicular system. *PNAS*, v. 102, p. 11911, 2005.
- [12] GOULET, G. C. et al. Modeling fluid flow and tracer transport in haversian bone. *J. Bone Min. Res.*, v. 20, p. S210, 2005.
- [13] GURUJARA, S. et al. Modeling deformation-induced fluid flow in cortical bone's canalicular-lacunar system. *Ann. Bio. Eng.*, v. 33, p. 7, 2005.
- [14] MCDONALD, F.; FORD, T. R. P. The effect of blood on the material properties of bone. *Ann. Bio. Eng.*, v. 22, p. 194, 1994.
- [15] WANG, L. et al. Modeling tracer transport in an osteon under cyclic loading. *Ann. Bio. Eng.*, v. 28, p. 1200, 2000.

- [16] JACOBS, C. R. et al. Numerical instabilities in bone remodelling simulations: the advantages of a node-based finite element approach. *J. Biomech.*, v. 28, p. 449, 1995.
- [17] CHENA, G. et al. Comparison of two numerical approaches for bone remodelling. *Med. Eng. Phys.*, v. 29, p. 134, 2007.
- [18] TANCK, E. et al. Cortical bone development under the growth plate is regulated by mechanical load transfer. *J. Anat.*, v. 208, p. 73, 2006.
- [19] BARABÁSI, A.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, p. 509, 1999.
- [20] CHANDLER, D. *Introduction to Modern Statistical Mechanics*. [S.l.]: Oxford University Press, 1987.
- [21] HUANG, K. *Statistical Mechanics*. [S.l.]: Wiley, John and Sons, 1990.
- [22] BALLOBÁS, B. *Random Graphs*. [S.l.]: Academic Press, London, 1985.
- [23] ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world wide web. *Nature*, v. 401, p. 130, 1999.
- [24] PASTOR-SATORRAS, R.; VASQUEZ, A.; VESPIAGNANI, A. Dynamical and correlation properties of the internet. *Phys. Rev. Lett.*, v. 87, p. 258701, 2001.
- [25] JIANG, Y.; XU, T.; HE, D.-R. A self-adaptive model of actor-collaboration network. *Int. J. Mod. Phys. B*, v. 18, p. 17–19, 2004.
- [26] BARABASI, A. L. et al. Evolution of the social network of scientific collaborations. *Physica A*, v. 311, p. 590, 2002.
- [27] REDNER, S. How popular is your paper? an empirical study of the citation distribution. *Eur. Phys. J. B*, v. 4, p. 131, 1998.
- [28] LILJEROS, F. et al. The web of human sexual contacts. *NATURE*, v. 411, p. 907, 2001.
- [29] CHI, L. P.; CAI, X. Structural changes caused by error and attack tolerance in us airport network. *Int. J. Mod. Phys. B*, v. 18, p. 17, 2004.
- [30] BUHL, J. et al. Topological patterns in street networks of self organized urban settlements. *Eur. Phys. J. B*, v. 49, p. 513, 2006.
- [31] VLAZQUEZ, A. et al. Modeling of protein interaction networks. *Complexus*, v. 1, p. 38–44, 2003.
- [32] JEONG, H. et al. The large-scale organization of metabolic networks.
- [33] HILGETAG, C. C.; KAISER, M. Clustered organization of cortical connectivity. *Neuroinformatics*, v. 2, p. 353, 2004.
- [34] CAMACHO, J.; GUIMERA, R.; AMARAL, L. A. N. Robust patterns in food web structure. *Phys. Rev. Lett.*, v. 88, p. 228102, 2002.
- [35] INTERNATIONAL, I. C. <http://www.interpore.com/>.

- [36] CANCHO, R. F. i; JANSEEN, C.; SOLE, R. V. Topology of technology graphs: Small world patterns in eletronic circuits. *Phys. Rev. E*, v. 64, p. 046119, 2001.
- [37] DAWSON, C. A. et al. Structure-function relationships in the pulmonary arterial tree. *J. Appl. Physiol.*, v. 86, p. 569, 1999.
- [38] KARAU, K. L. et al. Pulmonary arterial morphometry from microfocal x-ray computed tomography. *Am. J. Physiol. - Heart Circ. Physiol.*, v. 281, p. H2747, 2001.
- [39] PAINTER, P. R. Allometric scaling of the maximum metabolic rate of mammals: Oxygen transport from the lungs to the heart is a limiting step. *Theor. Biol. Med. Model.*, v. 2, p. 31, 2005.
- [40] STEVENS, P. S. *Patterns in Nature*. [S.l.]: Little Brown, Boston, 1974.
- [41] BANAVAR, J. R.; MARITAN, A.; RINALDO, A. Size and form in efficient transportation networks. *Nature*, v. 399, p. 130, 1999.
- [42] MCCULLOH, K. A.; SPERRY, J. S.; ADLER, F. R. Water transport in plants obeys murray's law. *Nature*, v. 421, p. 939, 2003.
- [43] MURRAY, C. D. The physiological principle of minimum work. i. the vascular system and the cost of blood volume. *Proc. Natl Acad. Sci. USA*, v. 12, p. 207, 1926.
- [44] SPORNS, O. et al. Organization, development and function of complex brain networks. *TRENDS in Cognitive Sciences*, v. 8, p. 418, 2004.
- [45] COSTA, L. F.; DIAMBRA, L. Topographical maps as complex networks. *Phys. Rev. E*, p. 021901, 2005.
- [46] OYAMA, Y.; YANAGITA, T.; ICHINOMIYA, T. Numerical analysis of fitzhugh-nagumo neurons on random networks. *Prog. Theor. Phys. Suppl.*, v. 161, p. 389, 2006.
- [47] MCLEAN, F. C.; IRIST, M. R. *Fundamentals of the Physiology of Skeletal Tissue*. [S.l.]: The University of Chicago Press, Chicago, 1968.
- [48] MUNRO, K. P. e M. Transport mechanism operating between blood supply and osteocytes in long bones. *Nature*, v. 269, p. 80, 1977.
- [49] ZENG, Y.; COWIN, S. C.; WEINBAUM, S. A fiber matrix model for fluid flow and streaming potentials in the canaliculi of an osteon. *Ann. Bio. Eng.*, v. 22, p. 280, 1994.
- [50] KUFAHL, R. H.; SAHA, S. A theoretical model for stress-generated fluid flow in the canaliculi-lacunae network in bone tissue. *J. Biomech.*, v. 23, p. 171, 1990.
- [51] RIBARSKY, B.; RUSHMEIER, H. 3d reconstruction and visualization. *IEEE Comp. Graph. Appl.*, v. 23, p. 20, 2003.
- [52] COSTA, L. F.; JR., R. M. C. *Shape analysis and classification: theory and practice*. [S.l.]: CRC Press, Boca Raton, 2001.
- [53] HOO, M. C. *Physiological control system analysis, Simulation and Estimation*. [S.l.]: IEEE Press, New York, 2000.

- [54] SCHROEDER, W. J.; MARTIN, K.; LORENSEN, B. *Vizualization Toolkit, An Object-Oriented Approach to 3d Graphics*. [S.l.]: Prentice Hall PTR, New Jersey, 1998.
- [55] SCHROEDER, W. J.; AVILA, L. S.; HOFFMAN, W. Vizualization with vtk: A tutorial. *IEEE Computers Graphycs and Aplications*, v. 20, p. 20, 2000.
- [56] MILGRAM, S. The small world problem. *Psychology Today*, v. 2, p. 60, 1967.
- [57] BAGLER, G.; SINHA, S. Network properties of protein structures. *Physica A*, v. 346, p. 27, 2005.
- [58] GUPTA, S.; ANDERSON, R. M.; MAY, R. M. Networks of sexual contacts: Implications for the pattern of spread of hiv. *AIDS*, v. 3, p. 807, 1989 1989.
- [59] MOODY, J. Race, school integration, and friendship segregation in america. *Am. J. Soc.*, v. 107, p. 679, 2001.
- [60] LILJEROS, F. et al. Social networks (communication arising): Sexual contacts and epidemic thresholds. *NATURE*, v. 423, p. 606, 2003.
- [61] GUIMERÁ, R.; AMARAL, L. A. N. Modeling the world-wide airport network. *Eur. Phys. J. B*, v. 38, p. 381, 2004.
- [62] HE, Y.; ZHU, X.; HEY, D.-R. Statistics and developing model of chinese skyway network. *Int. J. Mod. Phys. B*, v. 18, p. 17, 2004.
- [63] ALBERT; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, v. 74, p. 47, 2002.
- [64] DOROGOVTSSEV, S. N.; MENDES, J. F. F. Evolution of networks. *Adv. Phys.*, v. 51, p. 1079, 2002.
- [65] COHEN, R. et al. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, v. 86, p. 3682, 2001.
- [66] ERDÖS, P.; RÉNYI, A. *Random Graphs*. [S.l.]: Academic, London, 1985.
- [67] RAPOPORT, A. Contributions to the theory of random and biased nets. *Bull. of Math. Biol.*, v. 19, p. 257, 1957.
- [68] KITTEL, C. *Introduction to Solid State Physics*. [S.l.]: Wiley, New York, 1968.
- [69] BARABÁSI, A.-L.; ALBERT, R.; JEONG, H. Mean-field theory for scale-free random networks. *Physica A*, v. 272, p. 173, 1999.
- [70] DOROGOVTSSEV, S. N.; MENDES, J. F. F.; SAMUKHIN, A. N. Structure of growing networks with preferential linking. *Phys. Rev. Lett.*, v. 85, p. 4633, 2000.
- [71] KRAPIVSKY, P. L.; REDNER, S.; LEYVRAZ, F. Connectivity of growing random networks. *Phys. Rev. Lett.*, v. 85, p. 4629, 2000.
- [72] KRAPIVSKY, P. L.; REDNER, S. Organization of growing random networks. *Phys. Rev. E*, v. 63, p. 066123, 2001.

- [73] ONODY, R. N.; CASTRO, P. A. de. Nonlinear barabási-albert network. *Physica A*, v. 336, p. 491, 2004.
- [74] DOROGOVTSSEV, S. N.; MENDES, J. F. F. Evolution of networks with aging of sites. *Phys. Rev. E*, v. 62, p. 1842, 2000.
- [75] ONODY, R. N.; CASTRO, P. A. de. Complex network study of brazilian soccer players. *Phys. Rev. E*, v. 70, p. 037103, 2004.
- [76] EVANS, T. Complex networks. *Contemporary Physics*, v. 45, p. 455, 2004.
- [77] BARRAT, A.; WEIGT, M. On the properties of small-world network models. *Eur. Phys. J. B*, v. 13, p. 547, 2000.
- [78] KLEMM, K.; EGUÍLUZ, V. M. Growing scale-free networks with small-world behavior. *Phys. Rev. E*, v. 65, p. 057102, 2002.
- [79] ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, v. 74, p. 47, 2002.
- [80] DOROGOVTSSEV, S. N. et al. Principles of statistical mechanics of random networks. *Nucl. Phys. B*, v. 666, p. 396, 2003.
- [81] GOH, K.-I.; KAHNG, B.; KIM, D. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.*, v. 87, p. 278701, 2001.
- [82] TRUSINA, A. et al. Hierarchy measures in complex networks. *Phys. Rev. Lett.*, v. 92, p. 178702, 2004.
- [83] COSTA, L. F. The hierarchical backbone of complex networks. *Phys. Rev. Lett.*, v. 93, p. 098702, 2004.
- [84] COSTA, L. F.; ROCHA, L. da. A generalized approach to complex networks. *Eur. Phys. J. B*, v. 50, p. 237, 2006.
- [85] COSTA, L. F.; SILVA, F. N. Hierarchical characterization of complex networks. *arXiv:cond-mat/0412761*, 2006.
- [86] KALAPALA, V. et al. Scale invariance in road networks. *Phys. Rev. E*, v. 73, p. 026130, 2006.
- [87] DOROGOVTSSEV, S. N.; SAMUKHIN, A. N. Mesoscopes and fluctuations in networks. *Phys. Rev. E*, v. 67, p. 037103, 2003.
- [88] COHEN, R.; HAVLIN, S. Scale-free networks are ultrasmall. *cond-mat/0205476*.
- [89] PRIM, R. C. Shortest connecting networks and some generalizations. *Bell System Tech. J.*, v. 36, p. 1389, 1957.
- [90] FALOUTSOS, M.; FALOUTSOS, P.; FALOUTSOS, C. On power-law relationships of the internet topology. *SIAM Review*, v. 29, p. 251, 1999.
- [91] NEWMAN, M. E. J. Who is the best connected scientist? a study of scientific coauthorship networks. *Phys. Rev. E*, v. 64, p. 016131, 2001.

- [92] VIANA, M. P.; COSTA, L. F. The dynamics of hierarchical evolution of complex networks. *cond-mat/0504346*.
- [93] EVEN, S. *Graph Algorithms*. [S.l.]: W. H. Freeman and Co., New York - USA, 1979.
- [94] FORD, J. L. R.; FULKERSON, D. R. Maximal flow through a network. *Can. J. Math.*, v. 8, p. 399, 1956.
- [95] GOLDBERG, A. V.; TARJAN, R. E. A new approach to the maximum flow problem. *J. Assoc. Comput. Mach.*, v. 35, p. 921, 1988.
- [96] FIEDLER, M. Algebraic connectivity of graphs. *Czech. Math. J.*, v. 23, p. 208, 1970.
- [97] POTHEN, A.; SIMON, H.; LIOU, K.-P. Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. Matrix Anal. Appl.*, v. 11, p. 430, 1990.
- [98] SCOTT, J. *Social Network Analysis: A Handbook*. [S.l.]: Sage, London, 2000.
- [99] GIRVAN, M.; NEWMAN, M. E. J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*, v. 99, p. 7821, 2002.
- [100] NEWMAN, M. E. J. Fast algorithm for detecting community structure in networks. *Phys. Rev. E*, v. 69, p. 066133, 2004.
- [101] NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. *Phys. Rev. E*, v. 69, p. 026113, 2004.
- [102] DALL'ASTA, L. *Phénomènes Dynamiques sur des Réseaux Complexes*. Tese (Doutorado) — Université de Paris, 2006.
- [103] NEWMAN, M. E. J. The structure and function of complex networks. *SIAM Review*, v. 45, p. 167, 2003.
- [104] BOCCALETTI V. LATORA, Y. M. C. S.; HWANG, D.-U. Complex networks: Structure and dynamics. *Phys. Rep.*, v. 424, p. 175, 2006.
- [105] FERGUSON, N. M.; GARNETT, G. P. More realistic models of sexually transmitted disease transmission dynamics: Sexual partnership networks, pair models and moment closure. *Sex. Transm. Dis.*, v. 27, p. 600, 2000.
- [106] MEYERS, L. A. et al. Applying network theory to epidemics: Control measures for outbreaks of mycoplasma pneumoniae. *Emerg. Inf. Dis.*, v. 9, p. 204, 2001.
- [107] PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E*, v. 63, p. 2001, 2001.
- [108] NEWMAN, M. E. J. Spread of epidemic disease on networks. *Phys. Rev. E*, v. 66, p. 016128, 2002.
- [109] BARTHELEMY, M.; GONDRAN, B.; GUICHARD, E. Large scale cross-correlations in internet traffic. *Phys. Rev. E*, v. 66, p. 056110, 2002.
- [110] VALVERDE, S.; SOLE, R. Self-organized critical traffic in parallel computer networks. *Physica A*, v. 312, p. 636, 2002.

- [111] LI, X.; JIN, Y. Y.; CHEN, G. Complexity and synchronization of the world trade web. *Physica A*, v. 328, p. 287, 2003.
- [112] ARENAS, A.; DIAZ-GUILERA, A.; PEREZ-VICENTE, C. J. Synchronization reveals topological scales in complex networks. *Phys. Rev. Lett.*, v. 96, p. 114102, 2006.
- [113] CALLWAY, D. S.; STROGATZ, M. E. J. N. S. H.; WATTS, D. J. Network robustness and fragility: percolation on random graphs. *Phys. Rev. Lett.*, v. 18, p. 5468, 2000.
- [114] MOORE, C.; NEWMAN, M. E. J. Exact solution of site and bond percolation on small-world networks. *Phys. Rev. E*, v. 62, p. 7059, 2000.
- [115] ALBERT, R. et al. Error and attack tolerance of complex networks. *Nature*, n. 406, p. 378, 2000.
- [116] COHEN1, R. et al. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, v. 86, p. 3682, 2001.
- [117] ROCHA, L. E. C. da; COSTA, L. F. 2d pattern evolution constrained by complex network dynamics. *arXiv:physics/0610001*, 2006.
- [118] ERDÖS, P.; RÉNYI, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, v. 5, p. 17, 1967.
- [119] HOLME, P.; KIM, B. J. Attack vulnerability of complex networks. *Phys. Rev. E*, v. 65, p. 056109, 2002.
- [120] ZHAO, L. et al. Tolerance of scale-free networks against attack-induced cascades. *Phys. Rev. E*, v. 72, p. 025104, 2005.
- [121] COSTA, L. F. Reinforcing the resilience of complex networks. *Phys. Rev. E*, v. 69, p. 066127, 2004.
- [122] BUHL, J. et al. Efficiency and robustness in ant networks of galleries. *Eur. Phys. J. B*, v. 42, p. 123, 2004.
- [123] LAEMMER, S.; GEHLSSEN, B.; HELBING, D. Scaling laws in the spatial structure of urban road networks. *Physica A*, v. 363, p. 89, 2006.
- [124] COSTA, L. F.; VIANA, M. P.; BELETTI, M. E. The complex channel networks of bone structure. *Appl. Phys. Lett.*, v. 88, p. 033903, 2006.
- [125] MONTROYA, J. M.; SOLÉ, R. V. Small world patterns in food webs. *J. Theor. Biol.*, v. 214, p. 405, 2002.
- [126] GEORGIA, R. et al. Comparative aspects of the density and diameter of haversian canals in the diaphyseal compact bone of man and dog. *Morphol. Embryol.*, v. 28, p. 11, 1982.
- [127] GEORGIA, R.; ALBU, I. . The haversian canal network in the femoral compact bone in some vertebrates. *Morphol. Embryol.*, v. 34, p. 155, 1988.

7 *Apêndice*

7.1 Conectividade Máxima para as Redes Aleatórias

Pretendemos mostrar como a distribuição de Poisson impõe um limite natural para os valores de k . Para isso, vamos supor que a maior conectividade presente na rede seja k_c . Assim, é razoável supor que a fração de vértices com esta conectividade é de ordem $\mathcal{O}(1)$. Sem perda de generalidade, iremos assumir que $P(k_c) = 1/N$ (ou seja, um vértice possui conectividade k_c). Tomando o logaritmo em ambos os lados da equação 4.15 quando $k = k_c$, temos:

$$\ln\left(\frac{1}{N}\right) = k_c \ln \langle k \rangle - \langle k \rangle - \ln k_c! \quad (7.1)$$

Agora, usamos a aproximação de Stirling: $\ln k_c! = k_c \ln k_c - k_c$, isso nos dá:

$$k_c \left[\ln\left(\frac{k_c}{\langle k \rangle}\right) - 1 \right] = \ln N - \langle k \rangle \quad (7.2)$$

Através desta expressão, encontramos para um grafo com $N = 10^6$ e $\langle k \rangle = 4$ o valor de $k_c = 20$, que é muito menor do que a maior conectividade possível $\check{k} = 10^6 - 1$ (uma discussão mais detalhada pode ser encontrada em [76]).

7.2 Diâmetro das Redes Aleatórias

Usando a teoria do campo médio, podemos assumir que todos os vértices de uma rede aleatória infinita realizam $\langle k \rangle$ conexões. Tomando o vértice r como referência, podemos inferir que o número hierárquico de vértices, $n_r(h)$, segue a lei de potência:

$$n_r(h) = \langle k \rangle^h \quad (7.3)$$

Agora, vamos assumir que esta lei seja válida para redes finitas. Sendo assim, deve ser verdade a seguinte equação de normalização:

$$\frac{1}{N} \sum_{h=1}^{\check{h}_r} n_r(h) = 1 \quad (7.4)$$

Aproximando a soma por uma integral e utilizando a equação 2.3, temos:

$$\int_{h=1}^{\check{h}_r} \langle k \rangle^h \approx N \quad (7.5)$$

Efetuando a integração e tomando o logaritmo em ambos os lados, vem que

$$\check{h}_r \approx \frac{\ln[N \ln \langle k \rangle]}{\ln \langle k \rangle} \quad (7.6)$$

onde usamos o fato de que $\langle k \rangle^{\check{h}_r} \gg \langle k \rangle$. Além disso, sabemos que para as redes esparsas $\langle k \rangle \ll N$, logo podemos usar a aproximação $\mathcal{O}(N \log \langle k \rangle) = \mathcal{O}(N)$ e assim:

$$\check{h}_r \approx \frac{\ln N}{\ln \langle k \rangle} \quad (7.7)$$

Considerando que $\langle l \rangle \propto 1/N \sum_{r=1}^N \check{h}_r$, temos:

$$\langle l \rangle \approx \frac{\ln N}{\ln \langle k \rangle} \quad (7.8)$$

A dedução rigorosa deste resultado por ser encontrada em [22].

7.3 O Algoritmo de Prim-Dijkstra

Este apêndice tem o objetivo de descrever o funcionamento do algoritmo Prim-Dijkstra utilizado para encontrar a árvore de menor cobertura (AMC) de um grafo no qual as conexões possuem pesos distintos. O grafo que será utilizado como exemplo possui 32 vértices e 47 conexões, organizadas como mostra a figura 7.1-A1. O peso das conexões são representados por sua espessura (conexões mais espessas possuem peso maior e portanto devem ser evitadas na AMC). O algoritmo pode ser dividido em duas etapas:

1 - Buscamos pela conexão de menor peso do grafo. Quando existem mais de uma, a escolha é feita aleatoriamente. Em seguida, destacamos um dos vértices desta conexão. Este vértice está colorido em branco na imagem 7.1-1A.

2 - Buscamos pela conexão de menor peso que conecta um dos vértices brancos a um dos vértices pretos. Em seguida, colorimos esta conexão em vermelho e o novo vértice em branco.

O passo 2 é repetido até que não haja mais vértices pretos. O resultado final do algoritmo está na figura 7.1-D8 e a implementação (sumarizada) utilizada neste trabalho é mostrada abaixo:

```
nodes=SEARCH_NODE_ZERO(); /*executa o passo 1*/
while(blacks>0){
|...w_min=w_max; /*limite superior*/
|...node_start=nodes;
|...while(nodes!=NULL){ /*percorre a lista de nós brancos*/
|...|...n1=nodes->node;
|...|...VZ=(int*)malloc(DEGREES[n1]*sizeof(int));
|...|...NEIGHBOORS(n1,VZ); /*encontra os vizinhos de n1*/
|...|...for(i=0;i<DEGREES[n1];i++){ /*percorre os vizinhos*/
|...|...|...n2=VZ[i];
|...|...|...if(USED[n2]==0&&w(n1,n2)<w_min){ /*se n2 não é branco*/
|...|...|...|...w_min=w(n1,n2); /*e se w(n1,n2)<w_min*/
|...|...|...|...prev=n1; /*então n1 e n2 são possíveis candidatos*/
|...|...|...|...next=n2;
|...|...|...}
|...|...}
|...|...free(VZ);
|...|...nodes=nodes->next;
|...}
```



```
|...n1=prev;n2=next;
|...WRITE_CONNECTION(n1,n2); /*grava a conexão n1-n2*/
|...nodes=node_start; /*a lista volta ao começo*/
|...list*new_node=(list*)malloc(sizeof(list));
|...new_node->node=next; /*n1 e n2 são adicionados a lista dos nós brancos*/
|...new_node->next=nodes;
|...nodes=new_node;
|...USED[next]=1;
|...used++;
}
```

Figura 7.1: Procedimento realizado para obtenção da árvore de menor cobertura através do algoritmo Prim-Dijkstra. A figura A1 apresenta o grafo original com o vértice inicial na execução do algoritmo. A figura D8 mostra a AMC resultante.

7.4 Distribuição de Conectividade Hierárquica no Modelo Barabási-Albert

Krapivsky [72] foi um dos pesquisadores que estudou de maneira rigorosa o comportamento da distribuição de conectividade no modelo evolutivo de Barabási-Albert [69]. Para isso, ele partiu da quantidade $N_k(t)$ que fornece o número de vértices com conectividade k no instante t . Segundo este autor, a equação que governa a evolução desta quantidade é:

$$\frac{\partial N_k}{\partial t} = \frac{1}{M(t)} [(k-1)N_{k-1} - kN_k] + \delta_{1,k} \quad (7.9)$$

onde $M(t)$ é um fator de normalização dado por $M(t) = 2mt$. No caso em que $m = 1$, a solução desta equação é dada por:

$$N_k(t) = \frac{4t}{k(k+1)(k+2)} \quad (7.10)$$

A relação entre a distribuição de conectividade e a quantidade $N_k(t)$ é: $P_1(k) = N_k(t)/N$, mas neste modelo, $N \approx t$, logo:

$$P_1(k) = \frac{4}{k(k+1)(k+2)} \quad (7.11)$$

Nosso objetivo foi utilizar o conceito de hierarquia para estudar a conectividade no modelo Barabási-Albert de maneira um pouco mais geral, pois estávamos interessados na evolução da distribuição de conectividade na segunda hierarquia de cada vértice [92]. Para isso, partimos da seguinte quantidade: $N_{k,q}$. Este valor informa o número de vértices com a seguinte curva de conectividade: $k(h) = \{k(1) = k, k(2) = q, k(3) = \forall, k(4) = \forall, \dots\}$, ou seja, o número de vértice cujo valor da conectividade na primeira hierarquia é k e na segunda hierarquia é q , sendo que os valores de conectividade nas demais hierarquias podem ser quaisquer.

Em [92] descrevemos detalhadamente os procedimentos adotados para chegar à equação mestra que governa a evolução de $N_{k,q}$ apresentada abaixo:

$$\frac{\partial N_{k,q}(t)}{\partial t} = \frac{k-1}{2k+q+2} N_{k-1,q} + \frac{k+q-1}{2k+q+2} N_{k,q-1} + \frac{q}{4+q} N_q \delta_{1,k} \quad (7.12)$$

Aqui supomos uma solução linear do tipo:

$$N_{k,q}(t) = P(k, q)t \quad (7.13)$$

onde $P(k, q)$ é a fração de vértices com conectividade k na primeira hierarquia e q na segunda hierarquia. Utilizando esta relação na equação mestra, obtemos a seguinte relação de recorrência:

$$P(k, q) = \frac{k-1}{2k+q+2}P(k-1, q) + \frac{k+q-1}{2k+q+2}P(k, q-1) + \frac{q}{4+q}P_1(q)\delta_{1,k} \quad (7.14)$$

Por simplicidade, fazemos:

$$A_q = n_{k,q} \quad (7.15)$$

$$F_q = \frac{k-1}{2k+q+2}P(k-1, q) + \frac{q}{4+q}P_1(q)\delta_{1,k} \quad (7.16)$$

$$G_q = \frac{k+q-1}{2k+q+2} \quad (7.17)$$

Assim, podemos escrever a equação de recorrência como:

$$A_q = F_q + G_q A_{q-1} \quad (7.18)$$

Com a condição de que existe somente uma componente na rede, $P(k, 0) = 0$, a solução desta equação é dada por:

$$A_q = \frac{1}{G_{q+1}} \sum_{f=1}^q F_f \prod_{g=f}^q G_{g+1} \quad (7.19)$$

que posta em termos das variáveis originais torna-se:

$$\begin{aligned}
P(k, q) &= \left(\frac{2k+q+3}{k+q} \right) \sum_{f=1}^q \left[\frac{(k-1)P(k-1, f)}{2k+f+2} \right] \prod_{g=f}^q \left(\frac{k+g}{2k+g+3} \right) \\
&+ \frac{\Gamma(q+1)}{\Gamma(q+5)} \sum_{f=1}^{k^{(2)}} f(f+1)(f+2)(f+3)P_1(f)
\end{aligned}$$

Substituindo a equação 7.11 na expressão acima, vem que:

$$P(k, q) = \Delta_{1,q} + \frac{k-1}{2k+q+2} \sum_{f=0}^{q-1} \left[\frac{\Gamma(2k+f+2)}{\Gamma(k+f)} \right] P(k-1, f) \quad (7.20)$$

onde o termo $\Delta_{1,q}$ representa a correlação entre os vértices com conectividade unitária e os vértices com conectividade q [72]. Este termo é escrito como:

$$\Delta_{1,q} = \frac{2q(q+7)}{(q+1)(q+2)(q+3)(q+4)} \quad (7.21)$$

A distribuição de conectividade para a segunda hierarquia é obtida efetuando uma soma sobre todos os valores de k , como segue:

$$P_2(q) = \sum_{k=1}^{\check{k}} P(k, q) \quad (7.22)$$