



Ministério da
Ciência e Tecnologia



aa/bb/cc/dd-TDI

REDES COMPLEXAS DE BASE TERRITORIALIZADA

Leonardo Bacelar Lima Santos

Monografia de Qualificação para
o Doutorado do Curso de Pós-
Graduação em Computação Apli-
cada, orientada pelo Dr. Antônio
Miguel Vieira Monteiro.

URL do documento original:

[<http://urlib.net/xx/yy>](http://urlib.net/xx/yy)

INPE

São José dos Campos

2012

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6923/6921

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE (RE/DIR-204):

Presidente:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Membros:

Dr^a Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr^a Regina Célia dos Santos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

Dr. Horácio Hideki Yanasse - Centro de Tecnologias Especiais (CTE)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Vivéca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



Ministério da
Ciência e Tecnologia



aa/bb/cc/dd-TDI

REDES COMPLEXAS DE BASE TERRITORIALIZADA

Leonardo Bacelar Lima Santos

Monografia de Qualificação para
o Doutorado do Curso de Pós-
Graduação em Computação Apli-
cada, orientada pelo Dr. Antônio
Miguel Vieira Monteiro.

URL do documento original:

<<http://urlib.net/xx/yy>>

INPE

São José dos Campos

2012

Dados Internacionais de Catalogação na Publicação (CIP)

Santos, Leonardo Bacelar Lima.
Cutter Redes Complexas de Base Territorializada / Leonardo Bacelar
Lima Santos. – São José dos Campos : INPE, 2012.
xi + 29 p. ; (aa/bb/cc/dd-TDI)

Monografia de Qualificação (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2012.

Orientador : Dr. Antônio Miguel Vieira Monteiro.

1.Redes Complexas 2. Modelagem computacional 3.Interações
Clima-Ambiente-Saúde

CDU 000.000

Copyright © 2012 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, reprográfico, de microfilmagem ou outros, sem a permissão escrita do INPE, com exceção de qualquer material fornecido especificamente com o propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2012 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, microfilming, or otherwise, without written permission from INPE, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

RESUMO

Redes Complexas são grafos, conjuntos de vértices possivelmente conectados uns aos outros por arestas, cujas propriedades topológicas não são necessariamente nem regulares, como em autômatos celulares regulares, nem completamente aleatórias, como nos bem conhecidos grafos aleatórios. Podem ser consideradas como grafos não triviais com um grande número de vértices, muito apropriados para representação de conhecimento, tratamento de dados e modelagem de sistemas complexos. Fenômenos de interação Clima-Ambiente-Saúde são exemplos nos quais Redes Complexas podem representar uma poderosa ferramenta. A incorporação de elementos espacialmente explícitos em tais dinâmicas implica na necessidade de revisitar as relação vértices-arestas, novas interpretações para índices topológicos tradicionais e o desenvolvimento de índices locais e globais para as chamadas Redes Complexas de Base Territorializada. Este trabalho traz um panorama histórico sobre as contribuições à área desde a teoria dos grafos até os artigos do final do século XX que inauguraram a moderna Teoria de Redes Complexas; os principais artigos publicados no século XXI com enfoque em Redes Complexas de Base Territorializada e algumas das lacunas presentes na literatura que podem ser tratadas no doutoramento ao qual este texto se reporta.

SPATIALLY EMBEDDED COMPLEX NETWORKS

ABSTRACT

Complex Networks are graphs, sets of nodes possibly connected to each other by edges, whose topological properties are not necessarily regular, as in cellular automata, neither completely random, like the well known random graphs. Can be considered as non-trivial graphs with a large number of nodes, very suitable for knowledge representation, data processing and modeling of complex systems. Climate-Environment-Health interaction phenomena are examples in which complex networks can be a powerful tool. The incorporation of spatially explicit elements in such dynamics implies the needing to revisit the nodes-edges relationship, new interpretations to traditional topological indexes and the development of local and global indexes for Geographical Embedded Complex Networks. This work presents a historical overview on the contributions to the area from graph theory to the articles of the late twentieth century that ushered in the modern theory of complex networks, the major articles published in the twenty-first century with a focus on Geographical Embedded Complex Networks and some gaps in the literature that can be treated in the PhD to which this text refers.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Kaliningrado no século XIX (GRIBKOVSKAIA, 2007) e atualmente (BESSA et al., 2009)	3
2.2 Evolução temporal do número de artigos publicados na área de Redes Complexas.	5
3.1 Grafo G	15
3.2 Grafo G com vértices localizados de acordo com uma configuração pré-estabelecida.	17
3.3 Grafo G com vértices localizados de acordo com uma configuração pré-estabelecida e área de cobertura de cada vértice para cálculo do coeficiente de aglomeração e do caminho mínimo médio.	17
1.1 Grafo G	23
1.2 Matriz de adjacência relativa ao grafo G	24
1.3 Matriz de vizinhança relativa ao grafo G	25

LISTA DE ABREVIATURAS E SIGLAS

RC	–	Redes Complexas
TG	–	Teoria dos Grafos
TRC	–	Teoria de Redes Complexas
RCBT	–	Redes Complexas de Base Territorializada

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 REVISÃO BIBLIOGRÁFICA	3
2.1 CONTEXTUALIZAÇÃO HISTÓRICA	3
2.2 REDES COMPLEXAS DE BASE TERRITORIALIZADA	6
2.3 QUESTÕES DE ESCALA E MODELAGEM	10
2.4 LACUNAS NA LITERATURA	13
3 ESBOÇO DE CARACTERIZAÇÃO GEOGRÁFICA	15
3.1 CARACTERIZAÇÃO TOPOLÓGICA	15
3.2 CARACTERIZAÇÃO GEOGRÁFICA	16
REFERÊNCIAS BIBLIOGRÁFICAS	19
ANEXO 1 - CONCEITOS BÁSICOS DE REDES COMPLEXAS .	23
1.1 DEFINIÇÃO E REPRESENTAÇÃO	23
1.2 ÍNDICES BÁSICOS	24
1.2.1 Tamanho e Ordem de um grafo	25
1.2.2 Grau	25
1.2.3 Coeficiente de aglomeração	25
1.2.4 Caminho mínimo médio	25
1.3 TOPOLOGIAS E MODELOS DE GERAÇÃO	26
1.3.1 Redes Aleatórias	26
1.3.2 Redes Livres de Escala	26
1.3.3 Redes de Pequeno Mundo	27

1 INTRODUÇÃO

A área de Redes Complexas (RC) tem sua origem na tradicional Teoria dos Grafos (TG). Um grafo (\mathbf{G}) é um conjunto de vértices (\mathbf{V}) e arestas (\mathbf{E}), denotado por $\mathbf{G} = \mathbf{G}(\mathbf{V}, \mathbf{E})$, no qual os vértices podem estar conectados uns aos outros pelas arestas (NETTO, 2001). Uma RC é um grafo não trivial, como um grande número de vértices e propriedades topológicas não necessariamente típicas nem de grafos regulares, como em autômatos celulares regulares, nem de grafos completamente aleatórios (NEWMAN et al., 2006; RÉKA; BARABASI, 2002; NEWMAN, 2003).

A abordagem de RC é vastamente utilizada hoje em dia para representação de conhecimento, tratamento de dados e modelagem de sistemas complexos – sistemas formados por muitas partes, estas interrelacionadas possivelmente de forma não linear, apresentando comportamentos emergentes e em multiescalas. Exemplos de sistemas complexos podem ser encontrados especialmente em domínios biológicos, sociais e climáticos (NEWMAN, 2011; RÉKA; BARABASI, 2002; NEWMAN, 2003).

Fenômenos de interação Clima-Ambiente-Saúde são exemplos nos quais a Teoria de Redes Complexas (TRC) pode representar uma poderosa ferramenta, particularmente para modelagem dinâmica. A incorporação de elementos espacialmente explícitos em tais dinâmicas implica na necessidade de um novo aparato conceitual - relação vértices-arestas, nova interpretação para índices topológicos tradicionais e desenvolvimento de novos índices locais e globais para as chamadas Redes Complexas de Base Territorializada (RCBT) - redes nas quais os vértices tem localização geográfica conhecida e a lei de criação de arestas apresenta dependência funcional frente a variáveis espaciais.

Alguns trabalhos foram já publicados em RCBT, entretanto diversas lacunas ainda persistem, especialmente no contexto de modelagem dinâmica espacialmente explícita. Preencher algumas destas lacunas é o objetivo central do projeto de doutorado relativo a esta monografia.

A presente monografia está assim organizada:

- a) na seção Contextualização Histórica é traçado um panorama histórico sobre as contribuições à área desde a TG até os artigos do final do século XX que inauguraram a moderna TRC;
- b) na seção Redes Complexas de Base Territorializada é discutida - do ponto de vista histórico e conceitual - a incorporação de elementos espacialmente

explícitos às redes, com a apresentação de alguns modelos clássicos de construção de RCBT;

- c) na seção Questões de Escala e Modelagem há uma breve análise sobre as relações entre diferentes escalas e diferentes técnicas de modelagem, apresentando o papel das RCBT nesse contexto;
- d) na seção Lacunas na Literatura são listados alguns pontos ainda não explorados nas publicações da área de RCBT;
- e) na seção Esboço de Caracterização Geográfica são apresentados alguns resultados preliminares de desenvolvimento de índices espaciais para caracterização de redes complexas.

A monografia conta ainda com um anexo que traz alguns conceitos introdutórios sobre Redes Complexas ¹.

¹Sugere-se a leitura deste anexo, antes de prosseguir com o presente texto, para aqueles que não têm familiaridade com os termos e conceitos introdutórios da TRC.

2 REVISÃO BIBLIOGRÁFICA

2.1 CONTEXTUALIZAÇÃO HISTÓRICA

A área de Redes Complexas tem sua origem na tradicional Teoria dos Grafos (TG). Publicado em 1736, o artigo do matemático suíço Leonhard Euler (1707-1783), sobre o problema das sete pontes de Königsberg, é considerado o primeiro resultado da TG (NETTO, 2001).

Conforme (NETTO, 2001; BESSA et al., 2009), Königsberg, antiga cidade prussiana (hoje Kaliningrado), foi erguida às margens do rio Preguel e consistia em quatro bairros separados, ligados por sete pontes (hoje as pontes estão parcialmente destruídas – Figura 2.1). Alguns dos moradores se perguntavam se seria possível fazer um passeio completo, atravessando as sete pontes, sem ter de atravessar duas vezes uma mesma ponte. Os cidadãos de Königsberg tentaram várias rotas mas todas terminaram em fracasso. Euler também não conseguiu encontrar tal rota, porém conseguiu explicar por que a jornada era impossível.

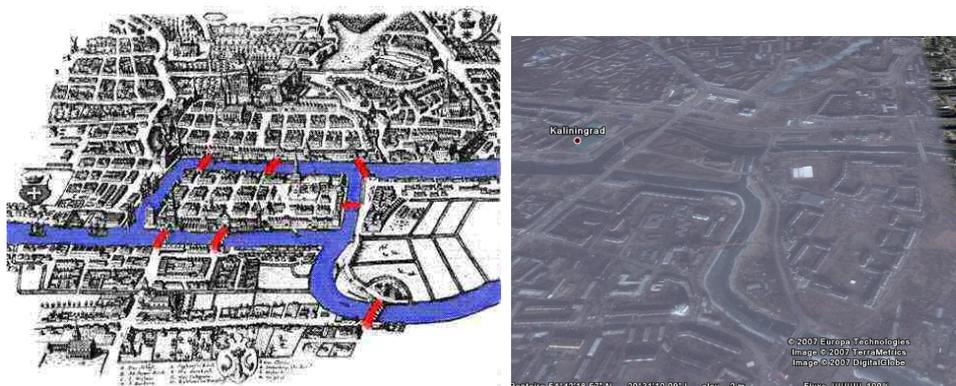


Figura 2.1 - Kaliningrado no século XIX (GRIBKOVSKAIA, 2007) e atualmente (BESSA et al., 2009)

Euler começou com uma planta da cidade e a partir dela produziu uma representação simplificada, na qual os trechos de terra são reduzidos a pontos e as pontes a linhas. Então argumentou que, de modo a fazer uma jornada bem-sucedida (ou seja, cruzando todas as pontes só uma vez), um ponto deve ser ligado por um número par de linhas. Isto acontece porque quando o viajante passa por uma massa de terra, ele deve chegar por uma ponte e sair por outra. Só existem duas exceções a esta regra: quando o viajante começa e quando termina sua jornada. Euler concluiu que, para qualquer rede de pontes, só é possível fazer um passeio completo atravessando uma

única vez cada ponte se todas as massas de terra tiverem um número par de pontes ou exatamente duas massas de terra tiverem um número ímpar de pontes. No caso de Königsberg, existem quatro massas de terra no total e todas são ligadas por um número ímpar de pontes. Euler tinha sido capaz de explicar por que seria impossível atravessar cada uma das pontes de Königsberg somente uma vez e, além disso, produziu uma regra que pode ser aplicada a qualquer rede de pontes em qualquer cidade¹.

Após o trabalho original de Euler diversos outros estudos foram feitos na área, com importantes contribuições em 1847 do físico alemão Gustav Kirchhoff (1824-1887), estudando circuitos elétricos, em 1857 com o químico inglês Arthur Cayley (1821-1895) e suas aplicações em química orgânica, em 1859 com o matemático irlandês William Hamilton (1805-1865) e seus jogos de caminhos em grafos (generalizando o problema original de Euler) e em 1869 com o matemático francês Camille Jordan (1838-1922) e suas teorias de árvores (grafos sem caminhos fechados).

Nos meados do século XX a TG ganha duas fundamentais contribuições, cujos resultados influenciam fortemente a moderna TRC:

i) Em 1959 os matemáticos húngaros Paul Erdős (1913-1996) e Alfréd Rényi (1921-1970) publicam um estudo sistemático de grafos aleatórios, com importantes resultados analíticos com enfoque probabilístico (ERDÖS; RÉNYI, 1959);

ii) Em 1967 o psicólogo social americano Stanley Milgran (1933-1984) promoveu um experimento para estudar o assim batizado “Problema do Pequeno Mundo”, para avaliar o grau de ligação entre indivíduos, obtendo um resultado que gerou o conceito dos seis graus de separação entre pessoas, mostrando que havia uma probabilidade alta de que indivíduos desconhecidos possuam amigos em comum (MILGRAN, 1967).

Uma busca na literatura², conforme a Figura 2.2, antes de 1990 eram publicados anualmente em periódicos menos de 20 artigos completos relacionados à palavra de busca “complex networks”. Tais artigos eram sobre redes de força, quase todos publicados em revistas de Engenharia Elétrica/Telecomunicações ou de Pesquisa Operacional. Em 1990 o total de papers, nunca maior que 20, passou para 85, e em

¹Tal texto é baseado na explicação compilada em (BESSA et al., 2009), que por sua vez se reporta a: Singh, S., Calife, J. L. T. (1998). O Último Teorema de Fermat: a história do enigma que confundiu as maiores mentes do mundo durante 358 anos; Editora Record, Rio de Janeiro, 2a Edição.

²Pesquisa realizada em 14 de novembro de 2011, no portal ISI - Web of Knowledge e Web of Science.

1991 para 679 - nenhum dos título dos artigos, entretanto, com o termo “complex networks” - o número é em grande parte justificado por contribuições às áreas da Inteligência Artificial (especialmente via Redes Neurais Artificiais) e trabalhos de Química Orgânica e Psicologia Social.

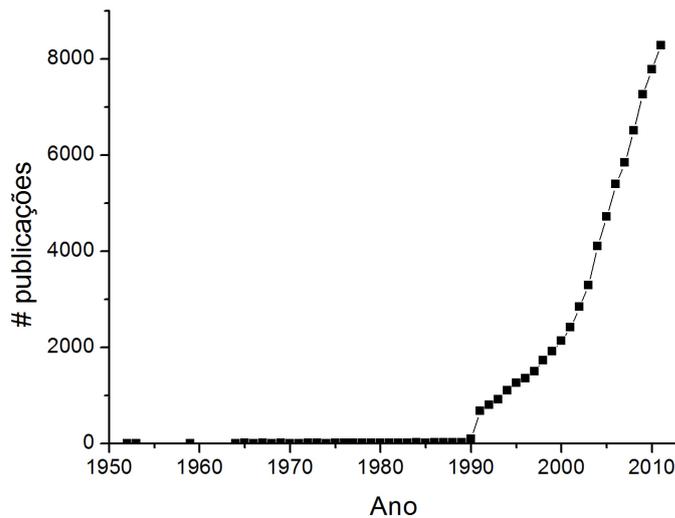


Figura 2.2 - Evolução temporal do número de artigos publicados na área de Redes Complexas.

Os trabalhos geralmente citados como os pioneiros da moderna TRC são:

- i) (1998) “Collective dynamics of ‘small-world’ networks” (WATTS; STROGATZ, 1998), do físico australiano Duncan J. Watts (1971) e do matemático americano Steven H. Strogatz (1959).
- ii) (1999) “Emergence of scaling in random networks” (BARABASI; RÉKA, 1999), do físico húngaro Albert-László Barabasi (1967) e da física também húngara Réka Albert³ - artigo mais citado da área, com 6716 citações.

Vale a pena destacar que estes dois artigos foram publicados nas duas revistas científicas internacionais possivelmente de maior fator de impacto desde aquela época até os dias atuais: Nature e Science, respectivamente.

Dois artigos de revisão merecem destaque pela grande quantidade de conceitos e aplicações discutidos e seu alto número de citações: o “Statistical mechanics of complex

³Ano de nascimento não encontrado.

networks” (Albert e Barabasi, 2002) (RÉKA; BARABASI, 2002) (5371 citações) e “The structure and function of complex networks” (Newman, 2003)(NEWMAN, 2003) (3731 citações). O livro “The Structure and Dynamics of Networks” (Newman, Barabasi e Watts, 2006)(NEWMAN et al., 2006) também foi bem recebido pela comunidade.

2.2 REDES COMPLEXAS DE BASE TERRITORIALIZADA

Conforme (Kuby et al., 2005) (KUBY, 2005), a representação do espaço via teoria de Redes Complexas é topológica mas não geográfica. A localização dos vértices, a proximidade entre eles e o comprimento das arestas são tratados como irrelevantes.

A despeito da relevância da geografia tanto na dinâmica de formação (conexão) quanto na dinâmica de processos nas redes, a imensa maioria dos trabalhos versam sobre a topologia e não a geografia das redes (HAYASHI, 2006).

Ao prescindir de questões geográficas a TRC focou-se nas relações, ganhou poder de abstração e generalização para operar em contexto topológico, no qual não a posição dos vértices mas sim a relação entre eles era importante naquele momento. Alguns anos depois destes primeiros passos, representações e análises geográficas mostraram relevância. O foco destes primeiros trabalhos foi em modelos para geração de redes sobre reticulados, com critérios espaciais: envolvendo a distância entre os pontos do reticulado sob os quais foram localizados os vértices.

A abordagem geográfica para Redes Complexas é, sobre certo ponto de vista, uma reciprocidade ao uso de grafos como estruturas de dados em geografia (e posteriormente via Sistemas de Informações Geográficas), especialmente no sub-domínio da Geografia dos Transportes. Conforme Sousa (2010) (SOUSA, 2010), a análise de redes de transporte e comunicação ganhou relevância na década de 60 (GARRISON, 1960; NYSTUEN; DACEY, 1961; KANSKY, 1963) e teve interesse renovado já no século XXI, graças aos SIG (Gastner e Newman, 2006). Ainda segundo Sousa (2010), no domínio das ciências geográficas, os SIGs vieram contornar um dos problemas apontados à TG, mais precisamente a sua abstração espacial e a sua incapacidade de incorporar os atributos de localização (coordenadas), “excluindo uma variável fundamental na explicação dos processos inter-ativos e intra-ativos entre os biótopos naturais e os biótopos artificiais de que resulta a configuração de padrões espaciais e a paisagem: o território”.

Três são os modelos mais recorrentes da literatura para construção de RCBT. Todos estes são modelos probabilísticos. Na descrição abaixo, para cada caso é fornecida

a função distribuição de probabilidade que norteia a conexão entre os vértices da rede. Cabe ressaltar que o enfoque probabilístico dado aos modelos de geração destas redes é recorrente na literatura da simulação de processos dinâmicos sobre redes, especialmente no contexto de propagação de epidemias (GOMEZ et al., 2011; LIN et al., 2005; PASTOR-SATORRAS; VESPIGNANI, 2001). Os modelos para construção de RCBT mais citados são:

- a) *Spatial Embedded Random Network (SERN)* (DALL; CHRISTENSEN, 2002) – Segundo *Barnet (2007)*, o modelo *SERN* de N vértices é especificado da seguinte forma:
- Uma métrica m sobre o espaço S onde o grafo será alocado,
 $m : S \times S \longrightarrow \mathfrak{R}^+$;
 - Uma variável aleatória, X , para a distribuição de vértices que toma valores em S ;
 - Uma função de probabilidade de conexão, não-crescente, entre dois vértices i e j , $\gamma(r(x_i, x_j)) : \mathfrak{R}^+ \longrightarrow [0, 1]$, onde r é a distância entre tais vértices.

A dinâmica de construção do grafo acontece associando a cada aresta uma probabilidade $\gamma(r(x_i, x_j)) = c(x, y)$, com y *iid* a x – ou seja, a probabilidade de conexão depende da posição relativa entre os vértices. O primeiro momento dessa distribuição é: $\mu \equiv E(c(x, y))$. O grau médio desta rede é: $\langle K \rangle = (N - 1)\mu$. Para o modelo de Erdős e Rényi⁴ $c(x, y) = p$, uma constante, independente da distância entre os vértices (ERDÖS; RÉNYI, 1959; BOLLOBÁS, 1998).

Vale ressaltar que este modelo não gera nem redes livres de escala nem redes com efeito de pequeno mundo.

- b) *General Spatial Embeeded COmplex Network (GSECON)* (KOSMIDIS et al., 2008) – O objetivo neste caso é dispor vértices em um reticulado de forma a minimizar o comprimento das arestas que conectam tais vértices.

Parte-se de um grafo cuja distribuição de graus, $P(k)$, é conhecida e do tipo *poisson*. Para cada vértice i , para cada um dos $k(i)$ vértices, de acordo com a distribuição $P(k)$, escolhe-se uma distância r em relação ao vértice i com probabilidade $\gamma(r) = cr^{d-1}r^{-\delta}$, onde δ controla a força do decaimento

⁴Vide anexo 1

de γ frente à distância, d é a dimensão do espaço e c é um termo de normalização: tal que $\int_1^L \gamma(r)dr = 1$, com $L = N^{\frac{1}{d}}$. Toma-se um vértice que diste r do vértice i e conecta-se a i . Remove-se conexões múltiplas quando preciso for.⁵

Este modelo já foi testado em espaços uni e bidimensionais e o seguinte resultado foi obtido (KOSMIDIS et al., 2008):

- $\delta < d$: o grafo formado é semelhante a um grafo aleatório;
- $d < \delta < 2d$: há características de redes livres de escala;
- $\delta > 2d$: o comportamento é parecido ao de grafos regulares.

c) *Lattice Embedded Scale-Free Network (LESFN)* (ROZENFELD et al., 2002) – De acordo com Hayashi(2006) e Yang et al. (2008), este modelo gera redes livres de escala em um reticulado de tamanho R com condições periódicas de contorno.

Para cada nó do reticulado associa-se um grau k , de acordo com uma distribuição de probabilidade $P(k)$ tal que $P(k)k^{-\lambda}$. Seleciona-se um vértice j qualquer. Conecta-se, probabilisticamente, tal vértice aos vizinhos mais próximos até que o seu grau seja k_j , ou que todos os vértices até uma distância limite, $Rmax_j$, tenham sido analisados, com $Rmax_j = Ak_j^{\frac{1}{d}}$, onde A é uma constante positiva previamente determinada e d é a dimensão do reticulado⁶. A probabilidade de conexão para um vértice j é dada por: $p_j(r) = D \exp(-(r/Rmax_j)^2)$ (YANG et al., 2008), uma distribuição gaussiana de média nula e desvio padrão $R\sqrt{2}/2$. Em (LIN et al., 2005), foi discutida a influência do parâmetro A em índices topográficos de redes criadas com tal algoritmo. Entretanto, o foco do trabalho foi em aspectos teóricos do processo de sincronização – questões de deformações dos índices por parâmetros espaciais foram tratadas apenas superficialmente, e índices propriamente espaciais não foram discutidos.

As RCBT do tipo livre de escala tem atraído forte atenção dos pesquisadores, com interesses não apenas na estrutura topológica estática da rede (HERRMANN et al., 2003) e sua robustez (HUANG et al., 2005), mas também na sua dinâmica, como passeios aleatórios (HUANG, 2006), sincronização (LIND, 2004), e os clássicos modelos de física estatística para magnetismo (ANDRADE; HERRMANN, 2004).

⁵Para os últimos vértices o grau pode não ser respeitado.

⁶Quando $A \rightarrow \infty$ tem-se um grafo aleatório (YANG et al., 2008).

Alguns artigos da literatura de RCBT merecem atenção especial (seguem em ordem cronológica crescente de publicação):

- *Kuby et al., 2005* (KUBY, 2005) – Traz um tipo de especificação para metodologias de representação de conhecimento, compara “estruturas” tipo Sistemas de Informações Geográficas, Redes Complexas, Problemas de Programação Linear, Modelagem Baseada em Agentes. No tocante a interfaces entre SIG e RC discute as RCBT de um ponto de vista teórico-conceitual.
- *Zhang et al., 2006* (ZHANG et al., 2006) – Discute que o processo de crescimento de redes, que utiliza como mecanismo preferencial de criação de arestas critérios de proximidade espacial, tenta capturar as dinâmicas reais de crescimento de redes por interações locais; sendo, portanto, uma opção natural para modelagem de processos onde tal tipo de vizinhança desempenha papel relevante.
- *Xu et al., 2006* e *Xu et al., 2007* (XU et al., 2006; XU et al., 2007) – São os únicos artigos com um modelo epidemiológico, mas tratam-se dos modelos Susceptível-Infetado (SI) e Susceptível-Infetado-Susceptível (SIS): os mais elementares dos modelos epidemiológicos. Não são exibidos os padrões espaciais das simulações, apenas a evolução temporal da prevalência, incidência e suas velocidades. Além disso, a hipótese de transmissão usada nestes trabalhos é que a probabilidade de um vértice i ser infectado por um vértice j é: $P(i, j) = \frac{1}{d(i, j)^\alpha}$, onde $d(i, j)$ é a distância euclidiana entre os vértices i e j , e α é um parâmetro a ser ajustado. Tal campo de probabilidade é, portanto, não homogêneo, mas ainda é isotrópico: possivelmente incoerente em situações reais.
- *Kosmidis et al., 2008* (KOSMIDIS et al., 2008) – Faz uma boa revisão dos artigos da área, discute as RCBT tanto de um ponto de vista de alto nível (esquema teórico-conceitual e possibilidades de aplicações) quanto baixo nível (modelos de construção e resultados matemáticos).
- *Bullock et al., 2010* (BULLOCK et al., 2010) – Sem entrar em questões matemáticas, aborda diversas questões conceituais e faz um vasto leque de citações a artigos correlacionados. É um dos poucos artigos a abordar o efeito de pequeno mundo em redes espaciais, discutindo tal possibilidade incorporando não uma restrição determinística de “raio máximo” para haver conexão mas com um termo de probabilidade:

- alta o suficiente para ligações de “longo alcance” – responsáveis pela queda do mínimo caminho médio da rede
- não tão alta, para não afetar o alto coeficiente de aglomeração: conseguido pela “redundância” das conexões “locais”, de “curto alcance”.

2.3 QUESTÕES DE ESCALA E MODELAGEM

Segundo *Santos e Barcellos (2006)* (SANTOS; BARCELLOS, 2006), a escala é uma relação entre o mapa e o mundo real. O termo escala, no contexto de modelagem que será abordado deste texto, é dado por Gibson et al. (2000) (GIBSON, 2000), e se refere à dimensão espacial, temporal, quantitativa ou analítica usada para medir e estudar objetos e processos. Já o conceito de nível, reporta à localizações ao longo da escala. Para uma dada escala, diversos níveis podem ser considerados. Toda escala tem dois constituintes essenciais: extensão (“A”) e resolução (“n”). A resolução (n) para observar uma região depende da extensão envolvida: para uma região muito extensa, normalmente as medidas devem usar uma grande agregação de unidades individuais, para que o número total de unidades (n) não seja grande demais a ponto de inviabilizar a operacionalização da análise. Segundo Gibson et al. (2000), tais conceitos na área de modelagem apresentam, portanto, diferenças em relação ao seu emprego em áreas como cartografia (LEVIN, 1992; GIBSON, 2000; TURNER, 2001).

Pode-se organizar dois grandes grupos de escalas para trabalhos de modelagem:

- Escala local: Se faz necessária uma grande quantidade de informação específicas para as unidade – muitos detalhes.
- Escala global: A área de cobertura do fenômeno é vasta, mas o tamanho de cada unidade de análise pode ser grande demais para capturar algumas propriedades, escapando do analista a percepção de alguns comportamentos.

Relacionar as diferentes escalas com diferentes técnicas de modelagem é tarefa não trivial (LEVINS, 1966) e, a depender do fenômeno que se quer modelar, diferentes visões podem ser igualmente apropriadas.

Pode-se dizer que para escalas mais finas, como a intra-urbana, modelos baseados em agentes são candidatos “naturais”, devido ao fato de que ambos os enfoques

necessitam de informações detalhadas sobre a área de estudo – muitos parâmetros. A modelagem baseada em agentes trata de processos de interação indivíduo-indivíduo (NIAZI; HUSSAIN, 2011; MARETTO et al., 2011; ANDRADE et al., 2011).

Do outro lado, ao aumentar a área de abrangência de cada unidade de estudo, os comportamentos médios⁷ ganham força dentro de cada unidade. Neste casos, abordagens baseadas em equações diferenciais podem parecer mais apropriadas.

Reis et al (2010) (REIS et al., 2009) discutem o problema de escala em modelagem tomando como base o crescimento de tumores. Em sua fase inicial, com poucas células e ausência de estruturas de transporte de nutrientes não difusivas (ou seja, sem angiogênese) modelos do tipo autômatos celulares se mostram apropriados. Para poder capturar comportamentos típicos de tumores maiores, já com angiogênese e até mesmo metástase, os modelos contínuos, de equações diferenciais ordinárias ou parciais, já ganham maior pertinência.

Uma “regra de ouro” para modelagem é que um bom modelo tem que ter o máximo de detalhes possíveis sobre o fenômeno tratado mas de forma a ainda ser operacional, tanto do ponto de vista fenomenológico (envolvendo questões de escala) quanto computacionais: relação com complexidade dos algoritmos de solução/simulação numérica e arquitetura e eficiência computacional.

Modelos do tipo “Toy models”⁸ se destinam a captar comportamentos mais gerais, pautados em mecanismos, são modelos mecanicistas, tradicionais da Física Estatística e Sistemas Dinâmicos – que acabam por, muitas vezes, deixar de lado certo grau de precisão.

Já modelos mais específicos, considerados certas vezes como apropriados apenas para estudos de caso, se adequam muito bem a condições muito particulares, e para tal precisam de um grande número de parâmetros a serem devidamente ajustados.

A busca pelo meio termo entre generalidade e precisão é bastante complexa. Segundo Colizza et al. (2007) (COLIZZA, 2007) o debate entre realismo, precisão e generalidade ainda está bastante vivo na comunidade de modelagem (LEVINS, 1966; MAY, 1972; KEELING, 2005).

Usar modelos genéricos o bastante para captar comportamentos globais e ter bons

⁷Na comunidade de Física Estatística e de Sistemas Dinâmicos é muito usado o termo comportamento de campo médio

⁸A tradução para o português seria como “Modelos de brinquedo”.

mecanismos para o ajuste dos parâmetros, visando ganhar especificidade e precisão, é um caminho promissor. O uso de formas de representação do conhecimento que permitam tal abordagem é fundamental. RCBT se enquadram neste cenário. A representação de processos via redes complexas representa a passagem de uma visão baseada no espaço euclidiano (geográfico) para uma visão topológica. Entretanto, tal transição não deve desconsiderar as propriedades geográficas inerentes às dinâmicas. As RCBT tentam, simultaneamente, possibilitar análises topológicas e análises geográficas.

Em escala local, no nível urbano, um exemplo de investigação foi apresentado por (CHOWELL et al., 2005). Tal estudo constrói uma rede de locais em uma cidade, incluindo residências, escritórios, centros comerciais e áreas de lazer. As ligações entre os locais, ou seja, as arestas entre os vértices, representam o fluxo de pessoas indo em um dado momento de um lugar a outro. As topologias destas redes, em geral, são heterogêneas e variam caso a caso.

Procedimentos similares tem sido empregados tomando como vértices cidades, estados ou países, e como arestas a movimentação das pessoas seja por rodovias estaduais e federais. Nestes casos, as redes mostram-se, geralmente, mais homogêneas, especialmente devido à fortes restrições espaciais existentes (MONTIS et al., 2007) – tais redes tem peso, porém, os pesos são relacionados à intensidade do fluxo e não a diretamente propriedades geográficas. Neste trabalho, um resultado com potencial geográfico é apresentado: são apresentados mapas com legenda de acordo com propriedades topológicas, e a distância geográfica entre *hubs*, vértices de alto grau, são analisadas.

Chegando à escalas globais, a caracterização é feita tomando-se como vértices a infraestrutura de transporte aéreo, conectando aeroportos quando havendo voos diretos entre eles. (COLIZZA, 2006)

Em todos estes casos, entretanto, e em diversos outros trabalhos na literatura, os efeitos combinados de topologia e geografia não foram explorados.

2.4 LACUNAS NA LITERATURA

Quatro oportunidades de investigação podem ser listadas com base na avaliação da literatura:

- a) Interpretar geograficamente índices de caracterização topológica tanto do

ponto de vista local (valor individual para cada vértice) quanto global (média para todo o grafo);

- b) Criar índices para caracterização geográfica das RCBT;
- c) Construir RCBT com vértices e arestas inspirados diretamente em problemas reais. Por exemplo: conectar duas séries temporais de um mesmo processo (como epidemias em diferentes cidades) caso a distância geográfica entre as cidades seja menor que um valor máximo ou a correlação entre suas séries temporais maior que um valor mínimo – e então caracterizar topologicamente e geograficamente tais grafos;
- d) Estudar processos epidêmicos em RCBT com base em modelos mais sofisticados que o modelo Susceptível-Infetado (SI), em especial modelos que retratem epidemias de transmissão indireta.

Os únicos trabalhos da literatura usando RCBT em modelagem de processos epidêmicos, até onde esta foi analisada⁹ são o *Xu et. al., 2006* e *Xu et. al., 2007*, que trazem os mais simples modelos epidemiológicos, o SI e o SIS, e sem nenhuma comparação com dados reais ou análises espaciais com mapas e recursos de geoprocessamento. A investigação da dinâmica espacialmente explícita de epidemias de forte ligação com questões espaciais (especialmente meteorológicas e ambientais), via modelos não tão simplistas (como o SI e o SIS), é, até onde foi constatado, inédita na literatura.

⁹Todos os artigos completos da base do ISI e Web of Science, suas referências e diversos artigos encontrados em buscas em ambientes como o google, e suas referências.

3 ESBOÇO DE CARACTERIZAÇÃO GEOGRÁFICA

A seguir será apresentado um conjunto de resultados preliminares, pertinentes à criação de índices para caracterização geográfica de redes complexas.

3.1 CARACTERIZAÇÃO TOPOLÓGICA

Dado o grafo G a seguir, figura 3.1, a caracterização topológica confere os seguintes resultados:

- Vértice 1: grau=2, coef. de aglom.=1, min. cam. médio= $\frac{7}{4}$;
- Vértice 2: grau=2, coef. de aglom.=1, min. cam. médio= $\frac{7}{4}$;
- Vértice 3: grau=3, coef. de aglom.= $\frac{1}{3}$, min. cam. médio= $\frac{5}{4}$;
- Vértice 4: grau=2, coef. de aglom.=0, min. cam. médio= $\frac{3}{2}$;
- Vértice 5: grau=1, coef. de aglom.=0, min. cam. médio= $\frac{9}{4}$.

O grau médio do grafo é, portanto, $\langle k \rangle = 2$, o coeficiente médio de aglomeração $\langle c \rangle = 0.47$, o caminho mínimo médio $\langle l \rangle = 1.7$ e o diâmetro 3. Estes são resultados da caracterização clássica de redes, desprezando questões espaciais.

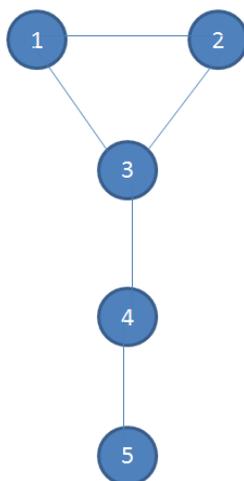


Figura 3.1 - Grafo G .

3.2 CARACTERIZAÇÃO GEOGRÁFICA

Para cada um dos índices topológicos básicos, k, c, l , foi construído um índice geográfico associado: k_r, c_r, l_r . O k_r de um vértice mede o número de vértices que distam até r deste¹. O c_r de um vértice sugere qual a probabilidade de que os vértices que a este distem até r estejam entre eles a uma distância até r . Ou seja, se os vértices dentro da área de cobertura de um vértice estão dentro da área de cobertura um do outro – matematicamente o valor de c_r é dado pela razão do número de vezes que a situação anteriormente descrita ocorre sobre o número de vezes que ela poderia ocorrer: $\frac{[k_r \times (k_r - 1)]}{2}$ ². O l_r mede o número médio de áreas de cobertura que é preciso percorrer para ir de um vértice a todos os outros, com o denominador para a média sendo o número de vértices “especialmente conexos”, ou seja, que podem ser alcançados a um número finito de “passos” sobre as áreas de cobertura.

Para o grafo G , figura 3.1, pode-se localizar os vértices de distintas formas, cada uma gerando um resultado possivelmente diferente de caracterização geográfica. Para exemplificar os índices escolheu-se a disponibilidade dos vértices conforme a figura 3.2. A figura 3.3 plota o grafo espacializado com a área de cobertura de cada vértice. A caracterização geográfica aqui proposta dá os seguintes resultados:

- Vértice 1: grau=1, coef. de aglom.=0, min. cam. médio= $\frac{5}{3}$;
- Vértice 2: grau=2, coef. de aglom.=1, min. cam. médio= $\frac{4}{3}$;
- Vértice 3: grau=0, coef. de aglom.=0, min. cam. médio= 0;
- Vértice 4: grau=2, coef. de aglom.=1, min. cam. médio= $\frac{4}{3}$;
- Vértice 5: grau=3, coef. de aglom.= $\frac{1}{3}$, min. cam. médio= 1.

O grau médio do grafo é, portanto, $\langle k \rangle = 1.6$, o coeficiente médio de aglomeração $\langle c \rangle = 0.47$, o caminho mínimo médio $\langle l \rangle = 5.3$ e o diâmetro 2 – mas com a existência de um vértice “geograficamente desconexo”: o vértice 3 não pode ser alcançado por um número finito de passos sobre a área de cobertura partindo de nenhum dos outros vértices.

¹Excluí-se o próprio vértice da contabilidade.

²Fórmula semelhante à usada para cálculo de c , no caso sem abordagem espacial.

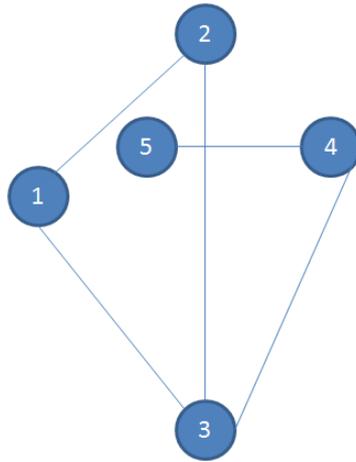


Figura 3.2 - Grafo G com vértices localizados de acordo com uma configuração pré-estabelecida.

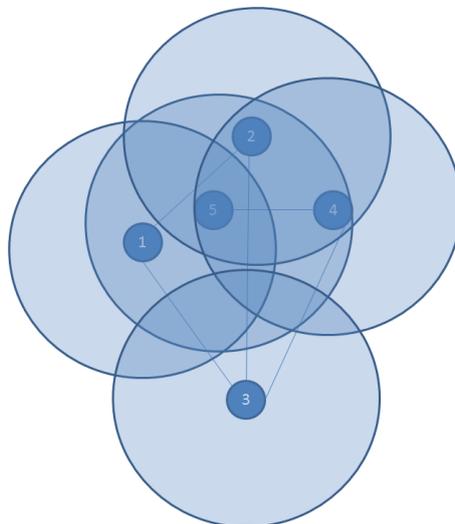


Figura 3.3 - Grafo G com vértices localizados de acordo com uma configuração pré-estabelecida e área de cobertura de cada vértice para cálculo do coeficiente de aglomeração e do caminho mínimo médio.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, P. R.; MONTEIRO, A. M. V.; CÂMARA, G. Entities and relations for agent-based modelling of complex spatial systems. **Proceedings of Second Brazilian Workshop on Social Simulation**, v. 1, p. 111–118, 2011. 11
- ANDRADE, R. F. S.; HERRMANN, H. J. Magnetic models on apollonian networks. **Phys. Rev. E**, v. 71, p. 056131, 2004. 9
- ANDRADE, R. F. S.; MIRANDA, J. G. V.; PINHO, S. T. R.; LOBAO, T. P. Characterization of complex networks by higher order neighborhood properties. **Eur. Phys. J. B.**, v. 61, p. 247–256, 2008. 24
- BARABASI, A. L.; RéKA, A. Emergence of scaling in random networks. **Science**, v. 286, p. 509–512, 1999. 5
- BESSA, A. D.; SANTOS, L. B. L.; MARTINEZ, P. N. R.; COSTA, M. C.; CARDOSO, P. G. S. Introdução às redes complexas. p. 21, 2009. Material livre, disponível em:
http://wiki.dpi.inpe.br/doku.php?id=ser301-2011:wiki_eonardosantos.vii, 3, 4, 23
- BOLLOBÁS, B. **Modern Graph Theory**. [S.l.]: Graduate Texts in Mathematics. Springer, 1998. V. 184. 7
- BULLOCK, S.; BARNETT, L.; PAOLO, E. A. D. Spatial embedding and the structure of complex networks. **Wiley Periodicals**, v. 16, n. 2, 2010. DOI 10.1002/cplx.20338. 9
- CHOWELL, G.; HYMAN, J. M.; EUBANK, S.; CASTILLO-CHAVEZ, C. Scaling laws for the movement of people between locations in a large city. **Phys. Rev. E**, v. 68, p. 066102, 2005. 12
- COLIZZA, V. *et al.*. The modeling of global epidemics: Stochastic dynamics and predictability. **Bulletin of Mathematical Biology**, v. 68, p. 1893–1921, 2006. 12
- _____. Epidemic modeling in complex realities. **C. R. Biologies**, v. 330, p. 364374, 2007. 12
- DALL, J.; CHRISTENSEN, M. Random geometric graphs. **Physical Review E**, v. 66, p. 016121, 2002. 7
- ERDÖS, P.; RÉNYI, A. On random graphs. **Publicationes Mathematicae**, v. 6, p. 290–297, 1959. 4, 7

- GARRISON, W. Connectivity of the interstate highway system. **Papers of the Regional Science Association**, 1960. 6
- GIBSON, C. C. *et al.*. The concept of scale and the human dimensions of global change: a survey. **Ecological Economics**, v. 32, n. 2, p. 217–239, 2000. 10
- GOMEZ, S.; ARENAS, A.; HOLTHOEFER, J. B.; MELONI, S.; MORENO, Y. Probabilistic framework for epidemic spreading in complex networks. **Int. J. Complex Systems in Science**, v. 1, p. 47–54, 2011. 7
- GRIBKOVSKAIA, I. *et al.*. The bridges of konisberg – a historical perspective. **Special Issue: Dedicated to Leonhard Euler**, v. 49, n. 3, p. 199–203, 2007. vii, 3
- HAYASHI, Y. A review of recent studies of geographical scale-free networks. **IP SJ Digital Courier**, v. 2, p. 155–164, 2006. 6
- HERRMANN, C.; BARTHÉLEMY, M.; PROVERO, P. Connectivity distribution of spatial networks. **Physical Review E**, v. 68, p. 026128, 2003. 9
- HUANG, L.; YANG, L.; YANG, K. Hollowing strategies for enhancing robustness of geographical networks. **Europhys. Lett.**, v. 72, p. 144, 2005. 9
- HUANG, Z. G. *et al.*. Walks on apollonian networks. 2006. ArXiv:cond-mat/0601357v1. 9
- KANSKY, K. Structure of transportation networks. **Department of Geography Research**, v. 84, 1963. 6
- KEELING, M. J. Models of foot-and-mouth disease. **Proc. R. Soc. B**, v. 272, p. 11951202, 2005. 12
- KOSMIDIS, K.; HAVLIN, S.; BUNDE, A. Structural properties of spatially embedded networks. **EPL**, v. 82, p. 48005, 2008. DOI: 10.1209/0295-5075/82/48005. 7, 8, 9
- KUBY, M. e. A comparison of geographic information systems, complex networks, and other models for analyzing transportation network topologies. **NASA/CR**, v. 213522, 2005. 6, 9
- LEVIN, S. A. The problem of pattern and scale in ecology. **Ecology**, v. 73, p. 1943–1967, 1992. 10

- LEVINS, R. The strategy of model building in population biology. **Am. Sci.**, v. 54, p. 421–431, 1966. [11](#), [12](#)
- LIN, Z. C.; YANG, L.; Q., Y. K. Geographical effects on complex networks. **Chin. Phys. Lett.**, v. 22, n. 12, p. 3214, 2005. [7](#), [8](#)
- LIND, P. G. *et al.*. Coherence in scale-free networks of chaotic maps. **Phys. Rev. E**, v. 70, p. 056207, 2004. [9](#)
- MARETTO, R. V.; ASSIS, T. O.; GAVLAK, A. A. Simulating urban growth and residential segregation through agent-based modeling. **Proceedings of Second Brazilian Workshop on Social Simulation**, v. 1, p. 52–57, 2011. [11](#)
- MAY, R. M. **Stability and Complexity in Model Ecosystems**. Princeton, NJ, USA: Princeton University Press, 1972. [12](#)
- MILGRAN, S. The small world problem. **Psychology today**, v. 1, p. 61–67, 1967. [4](#)
- MONTIS, A.; BARTHÉLEMY, M.; CHESSA, A.; VESPIGNANI, A. The structure of inter-urban traffic: a weighted network analysis. **Env. Plann. J. B.**, v. 34, p. 905–924, 2007. [12](#)
- NETTO, P. O. B. **Grafos: Teoria, Modelos, Algoritmos**. São Paulo: Editora Edgard Blücher, 2001. 402 p. [1](#), [3](#)
- NEWMAN, M. E. J. The structure and function of complex networks. **SIAM Review**, v. 45, p. 167–256, 2003. [1](#), [6](#)
- _____. Complex systems: A survey. **Am. J. Phys.**, p. 800–810, 2011. DOI 10.1119/1.3590372. [1](#)
- NEWMAN, M. E. J.; BARABASI, A. L.; WATTS, D. J. **The Structure and Dynamics of Networks**. [S.l.]: Princeton Studies in Complexity, 2006. ISBN-10: 0691113572. [1](#), [6](#)
- NIAZI, M.; HUSSAIN, A. Agent-based computing from multi-agent systems to agent-based models: A visual survey. **Springer Scientometrics**, v. 89, n. 2, p. 479–499, 2011. DOI:10.1007/s11192-011-0468-9. [11](#)
- NYSTUEN, J.; DACEY, M. A graph theory interpretation of nodal regions. **Papers and Proceedings of the Regional Science Association**, v. 7, p. 29–42, 1961. [6](#)

- PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic spreading in scale-free networks. **Physical Review Letters**, v. 86, n. 14, 2001. 7
- REIS, E. A.; SANTOS, L. B. L.; PINHO, S. T. R. A cellular automata model for avascular solid tumor growth under the effect of therapy. **Physica A**, v. 388, n. 7, p. 1303–1314, 2009. 11
- RÉKA, A.; BARABASI, A. L. Statistical mechanics of complex networks. **Rev. Mod. Phys.**, v. 74, p. 47–97, 2002. 1, 6
- ROZENFELD, A. F.; COHEN, R.; AVRAHAM, D. ben; HAVLIN, S. Scale-free networks on lattices. **Phys. Rev. Lett.**, v. 89, p. 218701, 2002. 8
- SANTOS, S. M.; BARCELLOS, C. **Abordagens espaciais na saúde pública. Série Capacitação e Atualização em Geoprocessamento em Saúde.** [S.l.]: Graduate Texts in Mathematics. Springer, 2006. 136 p. Serie B. 10
- SOUSA, P. A. M. **Efeito Estruturante das Redes de Transporte no Território.** 313 p. Tese em Geografia — Universidade de Lisboa, Lisboa, 2010. 6
- TURNER, M. G. *et al.* **Landscape ecology in theory and practice: pattern and process.** New York: Springer-Verlag, 2001. 10
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of “small-world” networks. **Nature**, v. 393, p. 440–442, 1998. 5
- XU, X.; WANG, W.; ZHOU, T.; CHEN, G. Geographical effects on epidemic spreading in scale-free networks. **Int. J. Mod. Phys. C**, v. 17, p. 1815, 2006. 9
- XU, X.; WUB, Z.; CHEN, G. Epidemic spreading in lattice-embedded scale-free networks. **Physica A**, v. 377, p. 125–130, 2007. 9
- YANG, K. Q.; L., Y.; GONG, B. H.; LIN, Z. C.; HE, H. S.; HUANG, L. Geographical networks: geographical effects on network properties. **Front. Phys. China**, v. 3, n. 1, p. 105–111, 2008. 8
- ZHANG, Z.; RONG, Z.; COMELLAS, Z. Evolving small-world networks with geographical attachment preference. **J. Phys. A: Math. Gen.**, v. 39, p. 3253–3261, 2006. 9

ANEXO 1 - CONCEITOS BÁSICOS DE REDES COMPLEXAS

A seguir são apresentados alguns conceitos básicos da Teoria de Redes Complexas, com ênfase nos tópicos que tem relação direta com os assuntos discutidos no corpo do texto da presente monografia de qualificação. O texto usado como base para este é o *Bessa et al., 2009* (BESSA et al., 2009).

1.1 DEFINIÇÃO E REPRESENTAÇÃO

Para uma definição formal, pode-se dizer que um grafo é um par $G = (V, E)$, sendo V um conjunto de vértices e E um conjunto de arestas, que equivalem cada uma a um subconjunto de V com dois vértices ¹.

Para o grafo G , grafo de exemplo da Figura 1.1, tem-se a seguinte representação: $G = (V, E)$, sendo $V = \{1, 2, 3, 4, 5\}$ e $E = \{1, 2, 1, 3, 2, 3, 3, 4, 4, 5\}$.

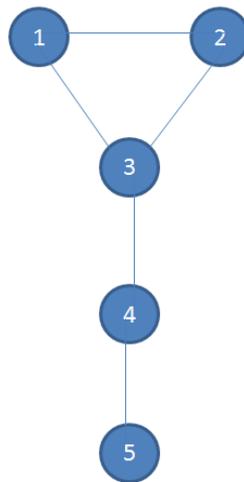


Figura 1.1 - Grafo G.

Para uma representação matemática/computacional de uma rede, pode-se empreender a construção de uma matriz de adjacência. Se o grafo possui N vértices, constrói-se a matriz de adjacência $M(N \times N)$. Para cada elemento $M(i, j)$ da matriz:

¹Considera-se aqui grafos não ponderados e não direcionados.

- $M(i, j) = 1$, se há uma aresta entre os vértices i e j do grafo;
- $M(i, j) = 0$, se não há essa ligação.

A matriz de adjacência do grafo G representado na Figura 1.1 é a seguinte:

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

Figura 1.2 - Matriz de adjacência relativa ao grafo G .

Outra forma de representar uma rede é via matriz de vizinhança. Há na literatura diversos tipos de matriz de vizinhança. Neste texto será abordada a apresentada por (Andrade et. al., 2008) (ANDRADE et al., 2008). Se um grafo possui N vértices, constrói-se a matriz de vizinhança $V(N \times N)$ tal que:

- $V(i, j) = l$, onde l é o número mínimo de arestas para, partindo do vértice i , alcançar o vértice j ;
- $V(i, j) = 0$, se os vértices i e j forem desconexos: não for possível sair do vértice i e alcançar o vértice j percorrendo exclusivamente arestas do grafo.

A partir da matriz de vizinhança é possível visualizar mais claramente não apenas a relação de vizinhança do tipo primeiros vizinhos, mas todas as ordens de vizinhança entre todos os vértices da rede. Ou seja, ela consegue condensar a relação de distância entre todos os vértices da rede em questão. A figura 1.3 traz a matriz de vizinhança relativa ao grafo G .

1.2 ÍNDICES BÁSICOS

Os índices de caracterização de redes complexas mais usados na literatura estão descritos nas subseções seguintes.

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

Figura 1.3 - Matriz de vizinhança relativa ao grafo G .

1.2.1 Tamanho e Ordem de um grafo

O número representativo da ordem de um grafo é o correspondente ao número de vértices que o compõem, e o seu tamanho é quantificado a partir do número de arestas que ligam seus vértices.

1.2.2 Grau

O grau k de um vértice i é determinado pela quantidade de vértices aos quais o vértice i está ligado, ou seja, o número de vizinhos do vértice i . Formalmente, define-se grau de um vértice como a cardinalidade do seu conjunto de adjacência: $|AdjG(i)|$. O grau médio $\langle k \rangle$ de um grafo é calculado pela média aritmética dos graus de cada vértice. Fazendo um histograma dos graus k encontrados em uma rede, temos a distribuição (frequência absoluta) de graus dessa rede. Define-se o índice $p(k)$ como a probabilidade (frequência relativa) de um vértice escolhido aleatoriamente em uma rede ter grau igual a k .

1.2.3 Coeficiente de aglomeração

O coeficiente de aglomeração de um vértice é a probabilidade de seus vizinhos serem vizinhos entre eles. Pode-se calcular o coeficiente de aglomeração de um vértice i de uma rede a partir da igualdade $C_i = 2E_i/k_i(k_i - 1)$, sendo E_i o número de arestas entre seus vizinhos e k_i o grau do vértice. O coeficiente de aglomeração médio de uma rede é a média dos coeficientes de aglomeração de cada vértice.

1.2.4 Caminho mínimo médio

Define-se um caminho como o conjunto de vértices e arestas, alternadamente dispostos, compondo um "percurso" que vai de um vértice a outro da rede. O comprimento

representativo desse caminho é o número de arestas que liga o primeiro ao último vértice do conjunto. O menor caminho que liga um vértice i a um vértice j da rede é chamado caminho mínimo entre os vértices i e j , e a média dos caminhos mínimos entre o vértice i e os demais vértices do grafo é o caminho mínimo médio – l – do vértice i . O caminho mínimo médio do grafo, por sua vez, é a média dos caminhos mínimos médios de cada vértice. O diâmetro de uma rede é definido como o maior dos caminhos mínimos entre quaisquer dois vértices.

1.3 TOPOLOGIAS E MODELOS DE GERAÇÃO

As Redes Complexas podem ser classificadas tomando como objeto de observação as suas propriedades estatísticas e índices topológicos. Os principais tipos de redes, ou os mais encontrados em trabalhos sobre Redes Complexas, são as redes aleatórias, rede livres de escala e redes de pequeno mundo.

1.3.1 Redes Aleatórias

As redes aleatórias são geradas a partir de ligações aleatórias entre os vértices de um conjunto, ou seja, são formadas a partir da atribuição de igual probabilidade de cada vértice de um conjunto se ligar a outro qualquer desse mesmo conjunto de vértices. Pode-se construir uma rede aleatória a partir de um conjunto de N vértices e L arestas, com um número $L = p \times N(N - 1)/2$ de arestas, sendo p a probabilidade de haver uma ligação entre vértices e $N(N - 1)/2$ o número máximo de possíveis arestas.

Redes desse tipo apresentam uma distribuição de graus característica: a distribuição gaussiana, com grau médio dado por $p(N - 1)$. O coeficiente de aglomeração médio é dado por p , e independe de N .

1.3.2 Redes Livres de Escala

A distribuição de graus de uma rede livre de escala segue uma lei de potência na qual poucos vértices possuem altos graus e a maioria dos vértices apresentam graus baixos. O modelo de *Barabási e Albert (2009)* é um dos mais usados para gerar redes livre de escala. A ideia fundamental é o crescimento do grafo via o princípio da ligação preferencial: quanto maior o grau de um vértice, maior a probabilidade de que ele receba mais vizinhos na próxima iteração do processo de crescimento².

Uma noção muito importante acerca da rede livre de escala é sua tolerância a falhas:

²Partindo de um pequeno grafo aleatório inicial.

remoções aleatórias de vértices atingem, em grande parte das vezes, vértices de baixo grau, pois eles são maioria; de forma que, considerando o grau de um vértice como uma medida da sua informação, há baixa probabilidade de uma grande perda de informação em processos de falha numa rede livre de escala. Por outro lado, redes desse tipo são vulneráveis a ataques: remoção de vértices por um processo que, apesar de apresentar baixa frequência, tem como alvo vértices de alto grau.

1.3.3 Redes de Pequeno Mundo

Stanley Milgran, um psicólogo social da Universidade de Harvard, em 1967 decidiu realizar um experimento no qual aproximadamente 150 famílias em Omaha, Nebraska e Wichita, Kansas deveriam entregar correspondências a pessoas alvo em Boston, utilizando apenas seus amigos (uma rede de conhecidos). Cada envelope continha:

- Nome, endereço e alguns dados pessoais do indivíduo alvo;
- Um conjunto de regras instruindo a pessoa intermediária que o recebeu sobre como proceder, podendo ser resumidas como: se você não conhece diretamente a pessoa alvo, repasse este envelope para um amigo que provavelmente a conheça. Você deve conhecer este amigo pessoalmente;
- Cada pessoa que recebesse este envelope deveria escrever seu nome nele, pois desta forma estaria evitando que uma mesma pessoa o recebesse novamente.

Inicialmente, Milgran e seus colegas acreditavam que as correspondências chegariam ao seu destino em torno de 100 passos aproximadamente. Ao término do experimento, no entanto, os resultados mostravam que estas haviam levado entre cinco a seis passos apenas e, assim, surgiu o conceito de seis graus de separação, que comprova que pessoas aparentemente sem relação alguma terão uma grande probabilidade de possuírem, em algum grau, amigos em comum que as aproximem.

Watts e Strogatz em 1998 propuseram um algoritmo que buscava mimetizar a topologia de interações sociais em um modelo abstrato, para tentar estudar este mesmo problema de uma maneira mais geral. Os frutos desse algoritmo são as redes de pequeno mundo, que possuem comportamento e características entre os apresentados por uma rede aleatória e uma rede regular. Redes de pequeno mundo apresentam coeficiente de aglomeração médio maior e mínimo caminho médio menor que uma rede aleatória de mesmo número de vértices e arestas.

Redes de pequeno mundo podem ser geradas por uma aplicação que retire conexões de uma rede regular (rede na qual todos os vértices têm o mesmo número de arestas) e promova reconexões (com probabilidade p). Assim, tem-se uma rede regular para $p = 0$ e uma aleatória para $p = 1$.

Vale ressaltar que o conceito de rede de pequeno mundo é passível de generalização a um efeito de pequeno mundo: situação de crescimento do coeficiente de aglomeração médio e decréscimo do mínimo caminho médio em uma rede.

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.