

ESTUDO DA DINÂMICA DE EPIDEMIAS EM REDES COMPLEXAS

Eduardo Ribeiro Pinto

Dissertação apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de Mestre em Biometria.

BOTUCATU
São Paulo - Brasil
Fevereiro – 2018

ESTUDO DA DINÂMICA DE EPIDEMIAS EM REDES COMPLEXAS

Eduardo Ribeiro Pinto

Orientadora: Prof^a. Dr^a. **Andriana S. L. O. Campanharo**

Dissertação apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de Mestre em Biometria.

BOTUCATU
São Paulo - Brasil
Fevereiro – 2018

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Pinto, Eduardo Ribeiro.

Estudo da dinâmica de epidemias em redes complexas /
Eduardo Ribeiro Pinto. - Botucatu, 2018

Dissertação (mestrado) - Universidade Estadual Paulista
"Júlio de Mesquita Filho", Instituto de Biociências de
Botucatu

Orientador: Andriana Susana Lopes de Oliveira
Campanharo

Capes: 900000005

1. Epidemias. 2. Tuberculose - Diagnóstico. 3.
Simulação por Computador. 4. Biometria. 5. Topologia.

Palavras-chave: MBI; Modelos Compartmentais; Redes
Complexas; Tuberculose.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Botucatu



ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE EDUARDO RIBEIRO PINTO, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA, DO INSTITUTO DE BIOCIÊNCIAS - CÂMPUS DE BOTUCATU.

Aos 23 dias do mês de fevereiro do ano de 2018, às 14:00 horas, no(a) Sala de Pós-Graduação do IBB, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Profa. Dra. ANDRIANA SUSANA LOPES DE OLIVEIRA CAMPANHARO - Orientador(a) do(a) Departamento de Bioestatística / Instituto de Biociências de Botucatu - UNESP, Prof. Dr. ERIVELTON GERALDO NEPOMUCENO do(a) Depto. de Engenharia Elétrica / UFSJ, Prof. Dr. FERNANDO LUIZ PIO DOS SANTOS do(a) Departamento de Bioestatística / Instituto de Biociências de Botucatu - UNESP, sob a presidência do primeiro, a fim de proceder a arguição pública da DISSERTAÇÃO DE MESTRADO de EDUARDO RIBEIRO PINTO, intitulada **Estudo da Dinâmica de Epidemias em Redes Complexas**. Após a exposição, o discente foi arguido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: Aprovado . Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.

Profa. Dra. ANDRIANA SUSANA LOPES DE OLIVEIRA CAMPANHARO

Prof. Dr. ERIVELTON GERALDO NEPOMUCENO

Prof. Dr. FERNANDO LUIZ PIO DOS SANTOS

Everything must be made as simple
as possible, but no simpler.

Albert Einstein, 1963

Dedico este trabalho a Deus, a toda minha família e a minha orientadora que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

Agradecimentos

Agradeço a todos que direta ou indiretamente contribuíram para a conclusão deste trabalho, especialmente:

Aos meus pais, Paulo e Luzia, ao meu irmão Leonardo e a todos da minha família por toda força, carinho e incentivo.

A minha orientadora, Dra. Andriana Susana Lopes de Oliveira Campanharo, pela oportunidade de pesquisa, orientação, paciência, apoio e compreensão durante esses anos.

Aos meus amigos, Felipe Teles, Jacqueline Domingues, Marta Helena, Gabriela Colovati, Leandro Marcucci, Wesley Gil, Rodrigo Duron, Rafael Castanha, Luiz Fernando Bomtempo... e me desculpo pelos que possa ter esquecido.

Aos colegas do departamento que me apoiaram e me acompanharam durante esse período.

À banca examinadora deste trabalho, Dr. Erivelton Geraldo Nepomuceno e Dr. Fernando Luiz Pio dos Santos pelas valiosas sugestões e comentários visando o aprimoramento deste trabalho.

Aos professores por todo conhecimento compartilhado.

À Fundação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo apoio financeiro, que foi essencial para o desenvolvimento de toda a pesquisa realizada durante o mestrado.

Sumário

	Página
LISTA DE FIGURAS	viii
LISTA DE TABELAS	xiii
RESUMO	xiv
SUMMARY	xvi
1 INTRODUÇÃO	1
1.1 Relevância	1
1.2 Estrutura do texto	3
2 REDES COMPLEXAS	5
2.1 Introdução histórica	7
2.2 Propriedades topológicas	8
2.2.1 Grau de conectividade	8
2.2.2 Força de conectividade	9
2.2.3 Coeficiente de agrupamento	10
2.2.4 Comprimento do menor caminho	11
2.2.5 Modularidade	12
2.2.6 Comprimento do salto médio	14
2.3 Modelos de redes complexas	14
2.3.1 Modelo de Erdős & Rényi	14
2.3.2 Modelo de Watts & Strogatz	16

	vii
2.3.3 Modelo de Barabási & Albert	19
2.3.4 Modelo modular	21
2.3.5 Modelo hierárquico	22
3 MODELO BASEADO EM INDIVÍDUOS	25
3.1 Formulação matemática	26
3.2 Algoritmo do MBI	30
4 ADEQUAÇÃO DO MBI AO MODELO SIR	32
4.1 Modelo SIR	32
4.2 Comparação entre o MBI e o modelo SIR	34
5 O MBI COM REDES COMPLEXAS	37
5.1 Estimação de Parâmetros	38
5.2 Modelos de redes complexas utilizados	40
6 TUBERCULOSE	51
6.1 Modelo matemático para a tuberculose	53
6.2 Comparação entre o MBI e o modelo SEI	54
6.3 O SEI e redes complexas	57
7 CONCLUSÕES	66
REFERÊNCIAS BIBLIOGRÁFICAS	68
APÊNDICES	74

Lista de Figuras

Página

1	(a) Rede não direcionada e sua respectiva matriz de adjacência A (binária e simétrica). (b) Rede direcionada e sua respectiva matriz de adjacência A (binária e não-simétrica). Fonte: (Rodrigues, 2007).	7
2	O problema das sete pontes de Königsberg. Representação de Euler por meio de um grafo para o problema. Fonte: (Rodrigues, 2007).	7
3	Exemplo de uma rede com 6 vértices (a) e a matriz de menores distâncias associada (b).	12
4	Exemplo de rede modular com 6 vértices, 7 arestas, 2 módulos e $\mathcal{M}_p = 0,357$	13
5	(a) Exemplo de uma rede aleatória de Erdős e Rényi com 20 vértices e $p = 0,1$ e (b) a distribuição do grau de conectividade $P(k)$ para uma rede de Erdős e Rényi com 1.000 vértices e $p = 0,1$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.	16
6	(a) Exemplo de rede mundo pequeno de Watts e Strogatz com 15 vértices, $k = 2$ e $p = 0,1$. (b) Distribuição do grau de conectividade $P(k)$ para uma rede de Watts e Strogatz com 1.000 vértices, $k = 25$ e $p = 0,3$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.	19

7	(a) Exemplo de rede livre de escala de Barabási e Albert com 20 vértices e $N_0 = 1$. (b) Distribuição do grau de conectividade $P(k)$ para uma rede de Barabási e Albert com 1.000 vértices e $N_0 = 1$ e $\gamma = 2,5$ (b). Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.	21
8	(a) Exemplo de uma rede modular com 10 vértices e dois módulos. (b) Distribuição do grau de conectividade $P(k)$ para uma rede modular com $N = 1.000$, $M = 10$, $p = 0,5$ e $r = 0,9$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.	23
9	(a) Exemplo de uma rede hierárquica com $N = 20$, $h = 2$ e $b = 3$. (b) Distribuição do grau de conectividade $P(k)$ para uma rede hierárquica com $N = 1.000$, $h = 2$ e $b = 100$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.	24
10	Representação esquemática do MBI para seis indivíduos. As interações ocorrem em intervalos de um ano. Destacam-se as mudanças de estado em relação à epidemia nos instantes $t = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$	29
11	Fluxograma do MBI. Os estados 0, 1 e 2 correspondem aos indivíduos suscetíveis, infectados e recuperados, respectivamente	31
12	Representação esquemática do modelo compartimental SIR, constituído de indivíduos suscetíveis (S), infectados (I) e recuperados (R).	33
13	Número de indivíduos suscetíveis, infectados e recuperados em função do tempo para os modelos SIR e MBI. Em (a) foram utilizados os parâmetros $N = 1.000$, $\mu_{sir} = \frac{1}{70}$, $\mu_{mbi} = 70$, $\gamma_{sir} = \frac{1}{36}$, $\gamma_{mbi} = 36$, $\beta_{sir} = 0,5$, $\beta_{mbi} = 0,05$, $\Delta t = 0,1$, $S(0) = 700$, $I(0) = 300$, $R(0) = 0$ e $\mathcal{R}_0 \approx 11,89$. Em (b) foram utilizados os parâmetros $N = 1.000$, $\mu_{sir} = \frac{1}{70}$, $\mu_{mbi} = 70$, $\gamma_{sir} = \frac{1}{24}$, $\gamma_{mbi} = 24$, $\beta_{sir} = 0,05$, $\beta_{mbi} = 0,005$, $\Delta t = 0,1$, $S(0) = 700$, $I(0) = 300$, $R(0) = 0$ e $\mathcal{R}_0 \approx 0,89$. As linhas pretas representam a solução do modelo SIR e as demais representam o conjunto de dez soluções do MBI.	36

14	Fluxograma do MBI com redes complexas para o modelo SIR. Os estados 0, 1 e 2 correspondem aos indivíduos suscetíveis, infectados e recuperados, respectivamente.	39
15	Comparação do MBI com redes complexas (linha contínua) com o modelo SIR (linha tracejada).	43
16	Comparação das curvas de indivíduos suscetíveis, onde cada curva é a média aritmética de 1000 simulações do MBI com os modelos de redes complexas em estudo.	44
17	Comparação das curvas de indivíduos infectados, onde cada curva é a média aritmética de 1000 simulações do MBI com os modelos de redes complexas em estudo.	44
18	Comparação das curvas de indivíduos recuperados, onde cada curva é a média aritmética de 1000 simulações do MBI com os modelos de redes complexas em estudo.	45
19	Propagação de uma doença em uma rede aleatória com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5. Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 14 infecta o vértice 17, em $t = 2$, o vértice 14 infecta o vértice 9, em $t = 3$, o vértice 19 infecta o vértice 8, em $t = 4$, o vértice 19 infecta o vértice 6 e em $t = 5$, o vértice 8 infecta o vértice 13. Observa-se, a partir de $t = 3$ o surgimento de vértices em verde (indivíduos recuperados).	46
20	Propagação de uma doença em uma rede mundo pequeno com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5. Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 14 infecta o vértice 16, em $t = 2$, o vértice 9 infecta os vértices 7 e 10, em $t = 3$, o vértice 7 infecta o vértice 5, em $t = 4$, o vértice 14 infecta o vértice 15 e em $t = 5$, o vértice 10 infecta o vértice 12. Observa-se, a partir de $t = 3$ o surgimento de vértices em verde (indivíduos recuperados).	47

- 21 Propagação de uma doença em uma rede livre de escala com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 2 infecta o vértice 9, em $t = 2$, o vértice 2 infecta o vértice 19 e o vértice 1 infecta o vértice 14, em $t = 3$, o vértice 1 infecta o vértice 4, em $t = 4$, o vértice 4 infecta o vértice 13 e o vértice 1 infecta o vértice 5 e em $t = 5$, o vértice 1 infecta o vértice 16. Observa-se, a partir de $t = 1$ o surgimento de vértices em verde (indivíduos recuperados). 48
- 22 Propagação de uma doença em uma rede modular com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 15 infecta o vértice 14, em $t = 2$, o vértice 15 infecta o vértice 13, em $t = 3$, o vértice 13 infecta o vértice 12, em $t = 4$, o vértice 14 infecta o vértice 11 e em $t = 5$, o vértice 9 infecta o vértice 1. Observa-se, a partir de $t = 3$ o surgimento de vértices em verde (indivíduos recuperados). 49
- 23 Propagação de uma doença em uma rede hierárquica com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 1 infecta o vértice 2, em $t = 2$, o vértice 2 infecta o vértice 5, em $t = 3$, o vértice 2 infecta o vértice 4, em $t = 4$, o vértice 5 infecta o vértice 10 e o vértice 6 infecta o vértice 13 e em $t = 5$, o vértice 5 infecta o vértice 11. Observa-se, a partir de $t = 2$ o surgimento de vértices em verde (indivíduos recuperados). 50
- 24 Representação esquemática de um modelo SEI , constituído de indivíduos suscetíveis, expostos e infectados. 53
- 25 Fluxograma do MBI para o modelo SEI . Os estados 0, 1 e 2 correspondem aos indivíduos suscetíveis, expostos e infectados respectivamente. 56

26	Número de indivíduos suscetíveis, expostos e infectados em função do tempo para os modelos SEI e MBI. Em (a) foram utilizados os parâmetros $N = 10.000$, $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,07$, $S(0) = 9000$, $E(0) = 0$, $I(0) = 1000$ e $\mathcal{R}_0 \approx 0,8$. Em (b) foram utilizados os parâmetros $N = 10.000$, $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,13$, $S(0) = 9000$, $E(0) = 0$, $I(0) = 1000$ e $\mathcal{R}_0 \approx 1,48$. As linhas pretas representam a solução do modelo SEI e as demais representam o conjunto de dez soluções do MBI.	58
27	Fluxograma do MBI com redes complexas para o SEI. Os estados 0, 1 e 2 correspondem os indivíduos suscetíveis, expostos e infectados, respectivamente.	61
28	Comparação do modelo MBI com redes complexas (linha contínua) e o modelo SEI (linha tracejada).	63
29	Comparação das curvas de indivíduos suscetíveis, onde cada curva é a média aritmética de 100 simulações do MBI com os modelos de redes complexas em estudo.	64
30	Comparação das curvas de indivíduos expostos, onde cada curva é a média aritmética de 100 simulações do MBI com os modelos de redes complexas em estudo.	64
31	Comparação das curvas de indivíduos infectados, onde cada curva é a média aritmética de 100 simulações do MBI com os modelos de redes complexas em estudo.	65
32	O número básico de reprodução $\mathcal{R}_0 = 2$: um indivíduo infectado introduzido em uma população de indivíduos suscetíveis.	75

Lista de Tabelas

	Página
1	Valores de k_i , E_i e C_i para cada um dos vértices da rede da Figura 1(a). 11
2	Resultados obtidos pelo teste t-Student da comparação das soluções produzidas pelos modelos SIR e MBI. 35
3	Parâmetros para geração das redes aleatória (ER), mundo pequeno (WS), livre de escala (BA), modular (MOD) e hierárquica (HIE) e as características topológicas de tais redes. 42
4	Resultados obtidos pelo teste t-Student da comparação das soluções produzidas pelos modelos SEI e MBI. 57
5	Parâmetros para geração das redes aleatória (ER), mundo pequeno (WS), livre de escala (BA), modular (MOD) e hierárquica (HIE) e suas propriedades topológicas. 62

ESTUDO DA DINÂMICA DE EPIDEMIAS EM REDES COMPLEXAS

Autor: EDUARDO RIBEIRO PINTO

Orientadora: Prof^ª. Dr^ª. ANDRIANA S. L. O. CAMPANHARO

RESUMO

Os Modelos Baseados em Indivíduos (MBI's) têm sido crescentemente empregados na modelagem de processos infecciosos. Um MBI consiste de uma estrutura na qual ocorrem interações entre um certo número de indivíduos, cujo comportamento é determinado por um conjunto de características que evoluem estocasticamente no tempo. Estudos recentes têm mostrado que as redes complexas constituem um suporte natural para o estudo da propagação de uma doença. Redes complexas são descritas por um conjunto de vértices (nós), arestas (conexões, ligações ou *links*) e algum tipo de interação entre os mesmos. Na formulação original do MBI e em modelos SIR (Suscetível, Infectado e Recuperado) e SEI (Suscetível, Exposto e Infectado), as relações entre os indivíduos são representadas por grafos completos, ou seja, todos os indivíduos estão conectados entre si. Como a topologia de uma rede real não pode ser descrita por uma rede puramente aleatória, nesse trabalho o MBI foi implementado de forma a incorporar modelos mais realísticos de redes de contato

na propagação de uma doença infecciosa. De maneira geral, observou-se que redes complexas com diferentes topologias resultam em curvas de indivíduos suscetíveis, infectados e recuperados (ou suscetíveis, expostos e infectados) com diferentes comportamentos. Desta forma, a evolução de uma dada doença, é altamente sensível à topologia de rede utilizada, em particular a tuberculose. Mais especificamente, observou-se que quanto maior o valor do comprimento do salto médio, mais rápida será a propagação da doença e, conseqüentemente, maior será o número de indivíduos infectados.

Palavras-chave: MBI, Redes Complexas, Tuberculose, Modelos Compartimentais.

MATHEMATICAL MODELS FOR EPIDEMIC PROPAGATION BASED ON SOCIAL NETWORKS

Author: EDUARDO RIBEIRO PINTO

Adviser: Prof^a. Dr^a. ANDRIANA S. L. O. CAMPANHARO

SUMMARY

Individual-Based Models have been increasingly employed in the modeling of an infectious process. An IBM consists of a structure in which interactions occur between a certain number of individuals, whose behavior is determined by a set of characteristics that evolve stochastically in time. Recent studies have shown that complex networks are a natural framework for the study of a disease spread. Complex networks are described by a set of vertices (or nodes), edges (connections or links) and some type of interactions between them. In the original IBM approach and in SIR (Susceptible, Infected, Recovered) and SEI (Susceptible, Exposed and Infected) models, the relations between individuals are represented by complete graphs, that is, all individuals are connected to each other. Since the topology of a real network can not be described by a purely random network, in this work an IBM has been implemented in order to incorporate some realistic contact networks

models. In general, it was observed that complex networks with different topologies correspond to curves of susceptible, infected and recovered individuals (or susceptible, exposed and infected) with different behaviors, and thus, that the evolution of a given disease, in particular tuberculosis, is highly sensitive to a network topology. More specifically, it was observed that the higher the value of the mean jump length is, the faster the disease spreads and consequently, the higher is the number of infected individuals.

Keyword: IBM, Complex Networks, Tuberculosis, Compartmental Models.

1 INTRODUÇÃO

1.1 Relevância

As doenças infecciosas são motivo de pânico desde tempos ancestrais, dizimando populações. Entende-se por doença infecciosa, ou doença transmissível, aquela causada por um agente biológico (por exemplo, vírus ou bactéria). Nenhum exemplo sintetiza melhor o efeito desastroso de doença infecciosa do que a *peste negra*, que dizimou um quarto da população europeia durante os anos de 1347 à 1350 (Cox et al., 2003). Outra doença que aterrorizou a Europa, foi a peste bubônica que matou mais de 68 mil pessoas em 1665. O número de mortes causadas pelas grandes epidemias é incomparavelmente maior que o número de mortes provocadas de todas as guerras juntas (Anderson & May, 1991).

A tuberculose é a doença infecto-contagiosa que mais mata no mundo, superando as mortes causadas pelo HIV e pela malária juntos. Estima-se que um terço da população mundial esteja infectada com o bacilo causador da doença, e que em 2013, ocorreram 9 milhões de novos casos e 1 milhão de óbitos (WHO, 2016). No Brasil foram diagnosticados 68 mil novos casos de tuberculose e cerca de 4,5 mil óbitos em 2015 (WHO, 2016). De acordo com a classificação da OMS 2016-2020, o Brasil ocupa a 20^a posição na lista dos 30 países com maior incidência de tuberculose.

Diante desse cenário, métodos que possam auxiliar à prevenção de doenças infecciosas, de forma a diminuir sua incidência e reduzir custos, tornam-se cada vez mais necessários. Em particular, os modelos matemáticos podem atuar como ferramentas valiosas na investigação de estratégias de controle de doenças infecciosas, podendo inferir informações sobre o impacto epidemiológico de uma vacina,

identificar quais e quantos indivíduos devem ser vacinados, verificar se a doença em estudo é epidêmica ou endêmica e fazer projeções futuras do número de infectados.

Diversas estratégias de modelagem de doenças infecciosas têm sido propostas na literatura através do uso de modelos compartimentais por meio de equações diferenciais ordinárias e parciais, cadeias de Markov, autômatos celulares, métodos de otimização, entre outros (Harris et al., 2016; Cohen et al., 2008; Castillo-Chavez & Feng, 1998; Bhunu et al., 2008; Ditkowsky & Schwartzman, 2014; Channing & Sinanovic, 2014; Tseng et al., 2011; ReVelle et al., 1967). Estudos recentes têm mostrado que as redes complexas constituem um suporte natural para o estudo da propagação de uma doença (Newman, 2002; Almeida, 2011; Liljeros et al., 2001; Barabási, 2009).

Redes complexas são descritas por um conjunto de vértices (nós), arestas (conexões, ligações ou *links*) e algum tipo de interação entre seus vértices (Newman, 2010). Diversos sistemas de importância biológica, tecnológica ou mesmo social podem ser representados por redes complexas. Por exemplo, uma célula pode ser descrita como uma rede de componentes químicos conectados por reações químicas; a Internet é uma rede de roteadores e computadores conectados por conexões físicas ou sem fio; ideias se propagam através de uma rede social cujos vértices são os seres humanos e as arestas representam os diferentes tipos de relações sociais (Barabási, 2009). Na modelagem de uma doença infecciosa por meio de redes complexas, cada um de seus vértices constitui um indivíduo e suas arestas representam contatos entre os mesmos (Liljeros et al., 2001; Almeida, 2011).

Os Modelos Baseados em Indivíduos (MBI's) têm sido crescentemente empregados na modelagem de processos infecciosos (Alvarenga, 2008; Almeida, 2011; Newman, 2003). Um MBI consiste de uma estrutura na qual ocorrem interações entre um certo número de indivíduos, cujo comportamento é determinado por um conjunto de características que evoluem estocasticamente no tempo (Sole et al., 1999; Keeling & Grenfell, 2000). A ideia principal de um modelo MBI é tratar o indivíduo como uma unidade básica, sendo a população o nome dado ao conjunto de entidades dis-

cretas (indivíduos). Na formulação original do MBI e em modelos SIR (Suscetível, Infectado e Recuperado) e SEI (Suscetível, Exposto e Infectado), as relações entre os indivíduos são representadas por grafos completos, onde todos os indivíduos estão conectados entre si. Contudo, a topologia de uma rede real não pode ser descrita por uma rede puramente aleatória. Desta forma, o objetivo deste trabalho foi implementar computacionalmente um MBI e modificá-lo, de forma que o mesmo incorporasse modelos mais realísticos de redes de contato na propagação de uma doença infecciosa, em particular, da tuberculose. Tais modelos compreendem as redes aleatórias de Erdős e Rényi (Erdős & Rényi, 1959); mundo pequeno de Watts e Strogatz (Watts & Strogatz, 1998); livre de escala de Barabási e Albert (Barabási & Albert, 1999); modular de Newman (Newman, 2003) e hierárquica de Dodds *et al.* (Dodds et al., 2003). Em seguida, foi realizado um estudo do efeito da topologia de tais redes nas soluções produzidas pelo MBI e, de como as mesmas se desviam das soluções produzidas pelos modelos SIR e SEI.

Todas as simulações realizadas nesse trabalho foram executadas em um computador Core 2 duo (4Gb de memória) com sistema operacional Linux e distribuição Ubuntu. Convém ressaltar que tanto o sistema operacional quanto os softwares utilizados para a geração dos resultados são gratuitos. Mais especificamente, os códigos foram escritos em linguagem C com o software Code::Blocks (Code::Blocks, 2008) e os gráficos foram gerados com o software XmGrace (Turner, 1995). Por fim, as redes foram geradas com o software Pajek (Batagelj & Mrvar, 1998).

1.2 Estrutura do texto

O texto está organizado da seguinte forma:

- Capítulo 2 apresenta-se um estudo detalhado de redes complexas, incluindo seus aspectos históricos, suas propriedades topológicas e os principais modelos presentes na literatura.
- Capítulo 3 apresenta-se um estudo detalhado de um MBI, incluindo sua for-

mulação matemática e seu algoritmo.

- Capítulo 4 apresenta-se a equivalência dos modelos Suscetível-Infectado-Recuperado (SIR) e MBI. Mais especificamente, mostra-se que as soluções desses modelos apresentam comportamentos similares, independentemente do conjunto de parâmetros utilizado e do número de execuções realizado.
- Capítulo 5 apresentam-se os resultados obtidos do MBI em conjunto com modelos realísticos de redes de contato. Mais especificamente, apresenta-se o efeito da topologia de tais redes nas soluções produzidas pelo MBI e, como tais, se desviam da solução produzida pelo modelo SIR.
- Capítulo 6 apresentam-se os resultados obtidos do MBI em conjunto com modelos realísticos de redes de contato para a modelagem da propagação da tuberculose. Mais especificamente, apresenta-se o efeito da topologia de tais redes nas soluções produzidas pelo MBI e, como tais, se desviam da solução produzida pelo modelo SEI.
- Capítulo 7 apresentam-se algumas conclusões e perspectivas de trabalhos futuros.

2 REDES COMPLEXAS

Uma rede complexa é descrita por um conjunto de vértices (nós), ligações (arestas, links) e algum tipo de interação entre os seus vértices. Em uma rede complexa, os vértices podem ser representados por circunferências e as arestas por arcos ou segmentos de reta (Newman, 2010). Problemas reais podem ser representados por meio de redes complexas, sendo essas divididas em:

- *Redes sociais*, onde cada indivíduo pode ser representado por um vértice e a sua interação com os demais indivíduos representada por arestas;
- *Redes de informação*, onde os vértices representam algum tipo de informação (conhecimento) e as arestas representam relações entre essas informações;
- *Redes tecnológicas*, onde os vértices podem representar, por exemplo, roteadores e as arestas as fibras óticas que os conectam;
- *Redes biológicas*, onde os vértices podem representar, por exemplo, proteínas e as interações entre as mesmas constituem suas arestas.

Matematicamente, uma rede complexa $R = (\mathcal{N}, \mathcal{L})$ é formada por um conjunto de N vértices, $\mathcal{N} = \{n_1, n_2, \dots, n_N\}$ e um conjunto de M arestas, $\mathcal{L} = \{l_1, l_2, \dots, l_M\}$. As arestas podem ser direcionadas quando o sentido da ligação é considerado, e não direcionadas quando o sentido da ligação não é considerado. Se as arestas possuem intensidade, a cada aresta é associado um peso, e neste caso, a rede é representada por $R = (\mathcal{N}, \mathcal{L}, \mathcal{W})$, onde $\mathcal{W} = \{w_1, w_2, \dots, w_M\}$ constitui um conjunto de pesos das M arestas (Newman, 2010).

Em termos computacionais, uma rede complexa pode ser armazenada em dois tipos diferentes de estrutura de dados: *matriz de adjacência* ou *listas*. Na representação em *matrizes de adjacência*, se dois vértices n_i e n_j estiverem conectados, a posição a_{ij} da matriz de adjacência A será igual a um; caso contrário, será igual a zero. Em uma rede não-direcionada, tem-se que $a_{ij} = a_{ji}$ para $i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$ (Fig. 1(a)), ou seja, a matriz de adjacência A é simétrica. No caso de uma rede direcionada, as linhas da matriz A representam as ligações que partem do vértice n_i para n_j , e as colunas de A representam as ligações que chegam de n_j para n_i (Fig. 1(b)). Se a rede não possuir autoconexões, então os elementos da diagonal principal da matriz A assumem o valor zero para todos os a_{ii} , com $i = 1, 2, \dots, N$. Em uma rede com peso, a posição w_{ij} da matriz de adjacência com pesos W será igual ao peso da aresta que conecta n_i e n_j .

Na representação em forma de listas, para uma rede sem peso apenas os pares de vértices (n_i, n_j) conectados por uma aresta são armazenados. Para redes com peso, a lista possui um terceiro elemento relacionado ao peso da aresta, (n_i, n_j, w_{ij}) .

Cada estrutura de armazenamento tem suas vantagens e desvantagens. O uso de listas permite maior economia de memória, principalmente se a rede é do tipo esparsa. Contudo, o acesso as ligações é mais complexo devido a necessidade de buscas. O acesso as ligações em matrizes de adjacências é mais simples, porém, este tipo de estrutura de dado é somente bem aproveitado em redes densas (ou seja, redes que possuem a maior parte de seus vértices conectados entre si) (Rodrigues, 2007).

As redes complexas podem ser estáticas, quando não há variação do número de vértices, de arestas ou na configuração das mesmas; ou dinâmicas, quando há variação de pelo menos uma destas componentes. O tipo mais comum de rede é aquele cujas arestas são direcionadas e com peso (Costa et al., 2007).

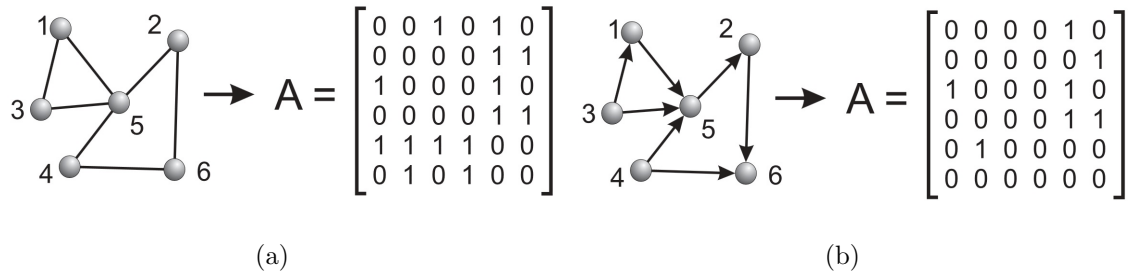


Figura 1 - (a) Rede não direcionada e sua respectiva matriz de adjacência A (binária e simétrica). (b) Rede direcionada e sua respectiva matriz de adjacência A (binária e não-simétrica). Fonte: (Rodrigues, 2007).

2.1 Introdução histórica

O matemático Leonhard Euler em 1736 introduziu a Teoria dos Grafos quando resolveu o famoso problema das Sete Pontes de Königsberg (Prússia no século XVII, atual Kaliningrado, Rússia). O problema colocava em questão a possibilidade de se atravessar as sete pontes que separavam quatro regiões de terra da cidade, passando uma única vez por cada ponte. Euler modelou o problema como um grafo, na qual as regiões de terra eram os vértices e as pontes eram as arestas (Fig. 2). Ele provou que, no caso das Pontes de Königsberg, não existia um caminho que satisfizesse as regras existentes. Euler então concluiu que tal passeio seria possível caso todas as regiões de terra (vértices) tivessem um número par de arestas ou que apenas duas regiões de terra tivessem um número ímpar de pontes (Rodrigues, 2007).

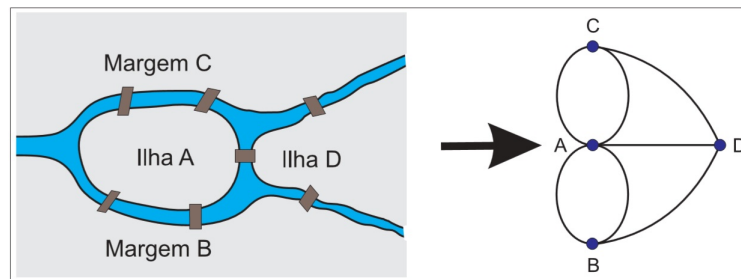


Figura 2 - O problema das sete pontes de Königsberg. Representação de Euler por meio de um grafo para o problema. Fonte: (Rodrigues, 2007).

Desde o estudo inicial de Euler, a teoria de grafos tem sido utilizada para solucionar questões práticas como, por exemplo, distribuir tarefas entre pessoas com a máxima eficiência; colorir regiões de um mapa usando um número mínimo de cores ou ainda preencher n empregos por n pessoas com utilidade máxima. No final do século XX, com o aumento do poder computacional e o surgimento da Internet, houve o surgimento de uma nova teoria, a teoria de redes complexas (Watts & Strogatz, 1998). Apesar de existirem semelhanças entre a teoria de redes complexas e a teoria dos grafos, a primeira está relacionada com modelagem de processos dinâmicos reais, onde as redes em estudo evoluem com o tempo, modificando suas estruturas.

2.2 Propriedades topológicas

As propriedades topológicas de uma rede complexa são definidas como métricas quantitativas capazes de caracterizar sua estrutura. A seguir, são apresentadas algumas dessas propriedades, tais como: o comprimento médio do menor caminho, o grau de conectividade médio, a força de conectividade média, o coeficiente de agrupamento médio e a modularidade.

2.2.1 Grau de conectividade

O grau de conectividade k_i de um vértice n_i é dado pelo número de arestas incidentes desse vértice a outros vértices. Para uma rede não direcionada, tem-se

$$k_i = \sum_{j=1}^N a_{ij} = \sum_{j=1}^N a_{ji}, \quad (1)$$

onde a_{ij} ($i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$) são elementos da matriz de adjacência A associada. No caso de uma rede não direcionada, ambos os somatórios resultam no mesmo valor de k_i , já que A é simétrica. Para uma rede direcionada, existem dois tipos de graus de conectividade: o grau de conectividade de saída k^{out} , igual ao número de arestas que partem do vértice n_i para o vértice n_j , e o grau de conectividade de entrada k^{in} , igual ao número de arestas que chegam ao vértice n_i pelo

vértice n_j , ou seja,

$$k_i^{out} = \sum_{j=1}^N a_{ij}, \quad (2)$$

$$k_i^{in} = \sum_{j=1}^N a_{ji}, \quad (3)$$

$$k_i = k_i^{out} + k_i^{in}, \quad (4)$$

com $i \neq j$. O grau de conectividade médio $\langle k \rangle$ de uma rede é definido como a média aritmética dos graus de conectividade sobre todos os vértices, ou seja,

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i. \quad (5)$$

Por exemplo, o grau de conectividade médio $\langle k \rangle$ da rede da Figura 1(a) é dado por:

$$\langle k \rangle = \frac{1}{6}(2 + 2 + 2 + 2 + 4 + 2) = \frac{14}{6} = 2,33. \quad (6)$$

O grau de conectividade médio $\langle k \rangle$ da rede da Figura 1(b) é dado por:

$$k^{in} = 1 + 1 + 0 + 0 + 3 + 2 = 7 \quad (7)$$

$$k^{out} = 1 + 1 + 2 + 2 + 1 + 0 = 7 \quad (8)$$

$$\langle k \rangle = \frac{1}{6}(k^{in} + k^{out}) = \frac{1}{6}(7 + 7) = \frac{14}{6} = 2,33 \quad (9)$$

2.2.2 Força de conectividade

Assim como o grau de conectividade, a força de conectividade s_i de um vértice n_i , em uma rede com peso, é definida como a soma dos pesos das arestas em n_i incidentes a outros vértices n_j . Para redes com peso e sem direção, tem-se que

$$s_i = \sum_{j=1}^N w_{ij} = \sum_{j=1}^N w_{ji}, \quad (10)$$

onde w_{ij} ($i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$) são elementos da matriz peso W associada. Por outro lado, para redes com peso e com direção, existem dois tipos de força de conectividade: a força de conectividade de saída s^{out} , igual ao peso do arco que

parte do vértice n_i para o vértice n_j , e a força de conectividade entrada s_i^{in} , igual ao peso do arco que chega ao vértice n_i pelo vértice n_j , ou seja,

$$s_i^{out} = \sum_{j=1}^N w_{ij}, \quad (11)$$

$$s_i^{in} = \sum_{j=1}^N w_{ji}, \quad (12)$$

$$s_i = s_i^{out} + s_i^{in}, \quad (13)$$

A força de conectividade média da rede, é definida como a média aritmética das forças de conectividade sobre todos os vértices, ou seja,

$$\langle s \rangle = \frac{1}{N} \sum_{i=1}^N s_i. \quad (14)$$

2.2.3 Coeficiente de agrupamento

Dada uma rede R , suponha que um vértice n_i tenha k_i vizinhos. Logo podem existir $k_i(k_i - 1)/2$ arestas entre tais vizinhos e isto ocorre quando todos os vizinhos de n_i estão conectados entre si. A razão entre o número E_i de arestas existentes entre os vizinhos de n_i pelo número total de arestas possíveis fornece o valor do coeficiente de agrupamento C_i do vértice n_i , ou seja,

$$C_i = \frac{E_i}{(k_i(k_i - 1))/2}. \quad (15)$$

O coeficiente de agrupamento de uma rede sem peso pode ser calculado diretamente por meio de sua matriz de adjacência A associada (Fagiolo, 2007). Ou seja,

$$C_i = \frac{(A^3)_{ii}}{k_i(k_i - 1)}. \quad (16)$$

Por definição, para $k_i = 0$ ou $k_i = 1$, $C_i = 0$. Define-se o coeficiente de agrupamento médio $\langle C \rangle$ da rede como a média aritmética dos C_i , para todos os vértices n_i ($i = 1, 2, \dots, N$):

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i. \quad (17)$$

A Tabela 1 apresenta os valores de k_i , E_i e C_i para cada um dos vértices da rede da Figura 1(a). Assim, com base na Equação (17), o coeficiente de agrupamento médio, é dado por

Tabela 1: Valores de k_i , E_i e C_i para cada um dos vértices da rede da Figura 1(a).

Vértice	k_i	E_i	C_i
1	2	1	1
2	2	0	0
3	2	1	1
4	2	0	0
5	4	1	$\frac{1}{6}$
6	2	0	0

$$\langle C \rangle = \frac{(1 + 0 + 1 + 0 + \frac{1}{6} + 0)}{6} \cong 0,361. \quad (18)$$

2.2.4 Comprimento do menor caminho

O comprimento do menor caminho entre dois vértices n_i e n_j é um caminho que os conecta cujo comprimento é mínimo (Costa et al., 2007). Em uma rede sem peso, o comprimento do menor caminho entre dois vértices n_i e n_j é o menor número de arestas que os conecta. Os menores caminhos entre todos os pares de vértices de uma rede podem ser representados através de uma matriz de menores distâncias L , cujos elementos l_{ij} representam o valor do menor caminho entre os vértices n_i e n_j . A Figura 3(a) apresenta uma rede e a sua matriz de menores distâncias.

A maior distância entre dois vértices n_i e n_j é chamada de *diâmetro* da rede. Matematicamente, o mesmo pode ser representado como $d_{\max} = \max_{i,j} \{d_{ij}\}$ com $(i = 1, 2, \dots, N \text{ e } j = 1, 2, \dots, N)$ (Newman, 2010). O comprimento médio do menor caminho $\langle L \rangle$ da rede é definido como a média dos elementos não-nulos de L , ou seja,

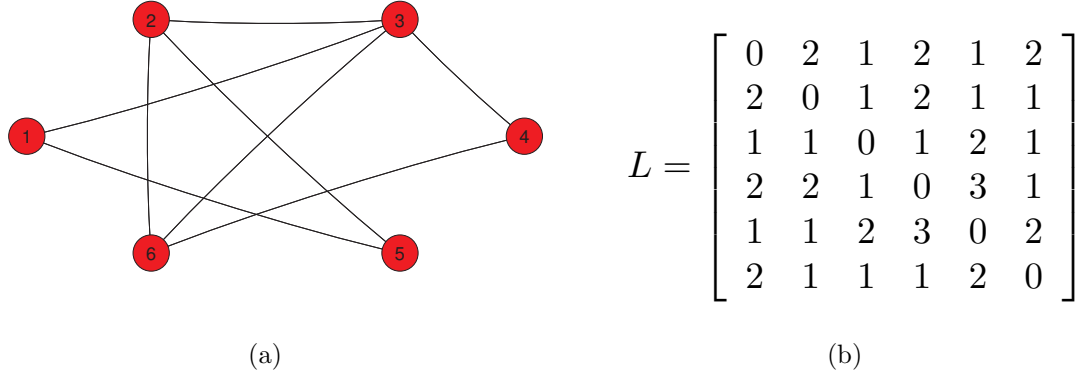


Figura 3 - Exemplo de uma rede com 6 vértices (a) e a matriz de menores distâncias associada (b).

$$\langle L \rangle = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N l_{ij}. \quad (19)$$

Desta forma, o comprimento médio do menor caminho da rede da Figura 3(a) é dado por:

$$\langle L \rangle = \frac{1}{6 \cdot 5} (2 + 1 + 2 + 1 + \dots + 1 + 1 + 1 + 2) = \frac{38}{6 \cdot 5} = 1,26. \quad (20)$$

Existem diversos métodos numéricos que calculam o menor caminho entre quaisquer pares de vértices n_i e n_j de uma rede complexa, tais como, os algoritmos de Dijkstra, Floyd-Warshall, Bellman-Ford, dentre outros. Nesse trabalho, o valor de $\langle L \rangle$ foi obtido com base no algoritmo de Floyd-Warshall (Floyd, 1962) por se tratar de um algoritmo que utiliza matrizes de adjacência como estrutura de dados, de fácil implementação e complexidade de ordem cúbica, ou seja, N^3 .

2.2.5 Modularidade

Uma grande parte das redes reais, incluindo redes sociais, biológicas e tecnológicas, possui uma estrutura modular, ou seja, arestas densamente distribuídas entre vértices pertencentes a um mesmo grupo e esparsamente distribuídas entre vértices pertencentes a grupos distintos (Newman, 2003). Um exemplo típico de

uma rede modular consiste de uma rede de contato entre as pessoas de uma pequena cidade, onde indivíduos de uma mesma família apresentam-se altamente conectados entre si quando comparados com indivíduos de famílias distintas.

A modularidade é uma medida que quantifica a possibilidade de se dividir uma rede em módulos (Guimerá et al., 2004). Para uma partição p , a modularidade $\mathcal{M}_p$ pode ser definida por

$$\mathcal{M}_p = \sum_{s=1}^r \left[\frac{M_s}{M} - \left(\frac{d_s}{2M} \right)^2 \right], \quad (21)$$

onde r é o número de módulos da rede, M é o total de arestas na rede, M_s é o número de arestas entre os vértices pertencentes ao s -ésimo módulo e d_s é a soma dos graus de conectividade dos vértices pertencentes ao s -ésimo módulo. A modularidade $\mathcal{M}$ de uma rede é definida como o máximo valor de $\mathcal{M}_p$ dentre todas as possíveis partições. Quando a rede é composta por um único módulo, então $\mathcal{M}_p = 0$. Como toda rede pode ser particionada dessa forma, conclui-se que $\mathcal{M} \geq 0$. A modularidade de uma rede é mais elevada quando a mesma é composta por módulos desconexos ou conectados por um número pequeno de arestas. A Figura 4 mostra uma rede particionada em dois módulos. Nela, os vértices 1, 2 e 3 pertencem ao módulo 1 e os vértices 4, 5 e 6 pertencem ao módulo 2. Utilizando-se a Equação (21), com $r = 2$, $M = 7$, $M_1 = 3$, $d_1 = 7$, $M_2 = 3$, $d_2 = 7$, obtém-se $\mathcal{M}_p = 0,357$.

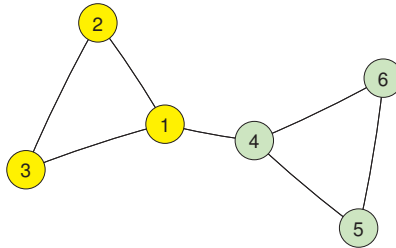


Figura 4 - Exemplo de rede modular com 6 vértices, 7 arestas, 2 módulos e $\mathcal{M}_p = 0,357$.

2.2.6 Comprimento do salto médio

A partir da matriz de adjacência A , é possível obter a matriz de transição de Markov P , de tal forma que $\sum_j p_{ij} = 1$. Desta forma, cada elemento $p_{ij} = a_{ij} / \sum_{j=1}^N a_{ij}$ representa a probabilidade de transição do vértice n_i para o vértice n_j . Assim, é possível realizar um *caminho aleatório* sobre a rede R e calcular o comprimento do salto médio Δ (Campanharo & Ramos, 2016), definido como:

$$\Delta = \frac{1}{S} \sum_{s=1}^S \delta_s(i, j), \quad (22)$$

onde $s = S$ são os saltos de comprimentos $\delta_l(i, j) = |i - j|$, onde $i, j = 1, \dots, N$ são os índices dos vértices definidos por P . O cálculo do comprimento do salto médio, pode ser calculado com base na expressão:

$$\Delta = \frac{1}{N} \text{tr}(GP^T), \quad (23)$$

onde P^T é a transposta de P , G é a matriz $N \times N$ com os elementos $g_{i,j} = |i - j|$ e tr é o traço da matriz resultante.

2.3 Modelos de redes complexas

2.3.1 Modelo de Erdős & Rényi

Em 1959, dois matemáticos húngaros, Paul Erdős e Alfred Rényi criaram um modelo de rede baseado em ligações aleatórias, que ficou conhecido como grafos aleatórios de Erdős e Rényi (Erdős & Rényi, 1959). As redes complexas geradas por este modelo possuem uma estrutura altamente homogênea. De acordo com tal modelo, uma rede aleatória é construída da seguinte forma:

- Inicialmente são considerados N vértices totalmente desconectados e uma probabilidade $0 \leq p \leq 1$;
- Um número aleatório $0 \leq w \leq 1$ é sorteado para cada par de vértices da rede;

- Cada valor de w é comparado à probabilidade p de uma aresta existir. Se $w < p$, a aresta é adicionada à rede conectando o par de vértices associado a ela. Ou seja, pares de vértices são conectados com probabilidade p .

Analisando-se o algoritmo descrito anteriormente, é possível concluir que quando $p = 0$, os vértices da rede estarão todos desconectados. Por outro lado, para $p = 1$, todos os vértices estarão conectados entre si. Para uma rede com N vértices, construída com probabilidade p , é esperado que o grau de conectividade médio seja dado por:

$$\langle \bar{k} \rangle \approx pN, \quad (24)$$

para $N \gg 1$ (Camargo, 2016). Além disso, o comprimento médio do menor caminho é pequeno nessas redes, podendo ser estimado como a razão entre o logaritmo do tamanho da rede e o logaritmo do grau de conectividade médio, ou seja,

$$\langle \bar{L} \rangle \approx \frac{\log N}{\log \langle \bar{k} \rangle}. \quad (25)$$

Por fim, o coeficiente de agrupamento médio é dado por:

$$\langle \bar{C} \rangle \approx p. \quad (26)$$

A Figura 5(a) mostra um exemplo de uma rede aleatória com 20 vértices e a Figura 5(b) mostra o gráfico da distribuição do grau de conectividade para uma rede aleatória com 1.000 vértices. Em ambos os casos, as redes foram geradas pelo modelo de Erdős e Rényi com $p = 0,1$. Observa-se que a distribuição do grau de conectividade pode ser aproximada por uma distribuição de Poisson¹ (Rodrigues, 2007). Além disso, observa-se que o valor do grau de conectividade médio esperado para a rede da Figura 5(b) (Eq. (24)), dado por $\langle \bar{k} \rangle = 100$, se aproxima do valor obtido numericamente (Eq. (5)), dado por $\langle k \rangle = 99,8$. De maneira análoga, o valor estimado para o comprimento médio do menor caminho (Eq. (25)), dado por $\langle \bar{L} \rangle = 1,5$, se aproxima do valor numérico (Eq. (19)), dado por $\langle L \rangle = 1,90$.

¹A distribuição de Poisson é um tipo de distribuição discreta de probabilidade, cuja forma analítica é dada por $P(k) = \frac{\lambda^k e^{-\lambda}}{k!}$ onde λ é o parâmetro da distribuição.

O mesmo pode ser observado para o valor estimado do coeficiente de agrupamento médio (Eq. (26)), dado por $\langle \bar{C} \rangle = 0,1$, que se aproxima do valor obtido numericamente (Eq. (17)), dado por $\langle C \rangle = 0,099$.

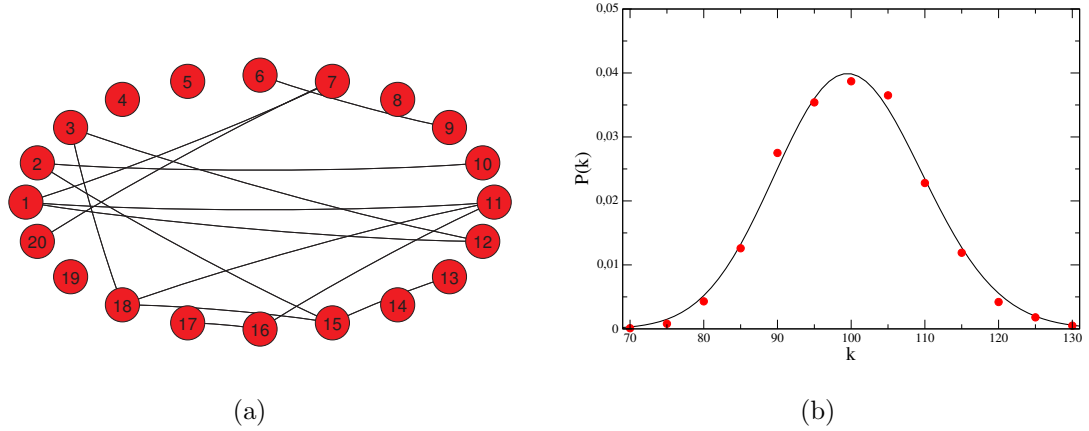


Figura 5 - (a) Exemplo de uma rede aleatória de Erdős e Rényi com 20 vértices e $p = 0,1$ e (b) a distribuição do grau de conectividade $P(k)$ para uma rede de Erdős e Rényi com 1.000 vértices e $p = 0,1$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.

Diversos sistemas na natureza e na sociedade, tais como redes de reações químicas, redes de roteadores e a de Internet foram inicialmente representados como redes aleatórias. Contudo, estudos posteriores concluíram que a topologia e a evolução de redes reais eram governadas por princípios de organização complexas e que as mesmas não poderiam ser representadas por um modelo de rede puramente aleatório (Albert & Barabási, 2002; Newman, 2002; Watts & Strogatz, 1998; Dodds et al., 2003).

2.3.2 Modelo de Watts & Strogatz

Duncan Watts e Steven Strogatz (Watts & Strogatz, 1998) observaram que a conexão topológica de diversas redes reais, incluindo redes biológicas, sociais e tecnológicas, não era completamente regular ou completamente aleatória, todavia se situava entre esses dois extremos. Ou seja, que redes reais poderiam ser

altamente agrupadas, como as redes regulares, mas com características de pequeno caminho, como os grafos aleatórios. Esta característica é conhecida como a propriedade “mundo pequeno” (em inglês, “small world”). O conceito de redes mundo pequeno, em termos simples, descreve o fato de que apesar de seu tamanho amplo, na maior parte das redes há um menor caminho entre quaisquer dois vértices. A manifestação mais popular das redes mundo pequeno é a “separação de grau seis”, descoberta pelo psicólogo Stanley Milgran (Milgran, 1967), que concluiu que há em média um caminho de comprimento de seis pessoas entre a maioria de pares de pessoas nos Estados Unidos (Newman, 2010).

Baseados nessa descoberta, Watts e Strogatz propuseram um modelo alternativo ao modelo de grafos aleatórios, conhecido como o *modelo mundo pequeno* de Watts-Strogatz. A construção desse modelo é feita através de um processo de interpolação entre uma rede regular e uma rede aleatória, a partir do seguinte processo:

- Inicia-se com uma rede circular e regular com N vértices, cada qual conectado aos seus k vizinhos à direita mais próximos;
- Escolhe-se aleatoriamente um vértice e uma aresta que o conecta a seu vizinho mais próximo no sentido horário. Com probabilidade p , reconecta-se essa aresta a um vértice escolhido aleatoriamente (de maneira uniforme) ao longo da rede, com proibição de duplicações ou auto-conexões. Em outras palavras, muda-se a aresta de lugar;
- Repete-se este processo sobre todos os vértices da rede até que uma volta seja completada;
- Consideram-se em seguida as arestas que conectam vértices a seus segundos vizinhos mais próximos. Como descrito anteriormente, reescreve-se aleatoriamente cada uma dessas arestas com probabilidade p , e repete-se este processo, circulando sobre a rede e indo em direção aos vizinhos mais distantes após cada

volta, até que cada ligação na rede original tenha sido considerada uma vez. Como há Nk arestas em toda a rede, o processo de reescrita para após k voltas.

Analisando-se o algoritmo descrito anteriormente, é possível concluir que quando $p = 0$ as conexões da rede em questão não mudam de lugar, ou seja, a rede permanece regular. Por outro lado, para $p = 1$ a rede resultante é aleatória. Portanto, para $0 < p < 1$ tal modelo produz redes que se situam entre a completa regularidade e aleatoriedade. O grau de conectividade de uma rede mundo pequeno, com N vértices e k vizinhos à direita, é dado por:

$$\langle \bar{k} \rangle = 2k. \quad (27)$$

O comprimento médio do menor caminho de uma rede regular, $\langle \bar{L} \rangle$, é dado por

$$\langle \bar{L} \rangle \approx \frac{N}{4k}. \quad (28)$$

Em um modelo mundo pequeno espera-se que o comprimento do menor caminho possua como limitantes:

$$\frac{\log N}{\log \langle \bar{k} \rangle} < \langle \bar{L} \rangle_{\text{mundopequeno}} < \frac{N}{4k}. \quad (29)$$

Portanto, $\langle \bar{L} \rangle_{\text{mundopequeno}}$ deve ser maior que $\langle \bar{L} \rangle_{\text{aleatoria}}$ e menor que $\langle \bar{L} \rangle_{\text{regular}}$. Entretanto, $\langle \bar{L} \rangle_{\text{mundopequeno}}$ deve se aproximar mais de $\langle \bar{L} \rangle_{\text{aleatoria}}$ por possuir arestas de longo alcance (Camargo, 2016).

O coeficiente de agrupamento médio de uma rede regular é dado por:

$$\langle \bar{C} \rangle = \frac{3k - 3}{4k - 2}. \quad (30)$$

De maneira similar ao menor caminho médio, tem-se que $\langle \bar{C} \rangle_{\text{mundopequeno}}$ está entre $\langle \bar{C} \rangle_{\text{aleatoria}}$ e $\langle \bar{C} \rangle_{\text{regular}}$, ou seja,

$$p < \langle \bar{C} \rangle_{\text{mundopequeno}} < \frac{3k - 3}{4k - 2}. \quad (31)$$

Para uma rede com N vértices, construída com probabilidade p , é esperado que o grau de conectividade médio seja dado por:

$$\langle \bar{C} \rangle \approx \frac{3(k - 1)}{2(2k - 1)}(1 - p)^3. \quad (32)$$

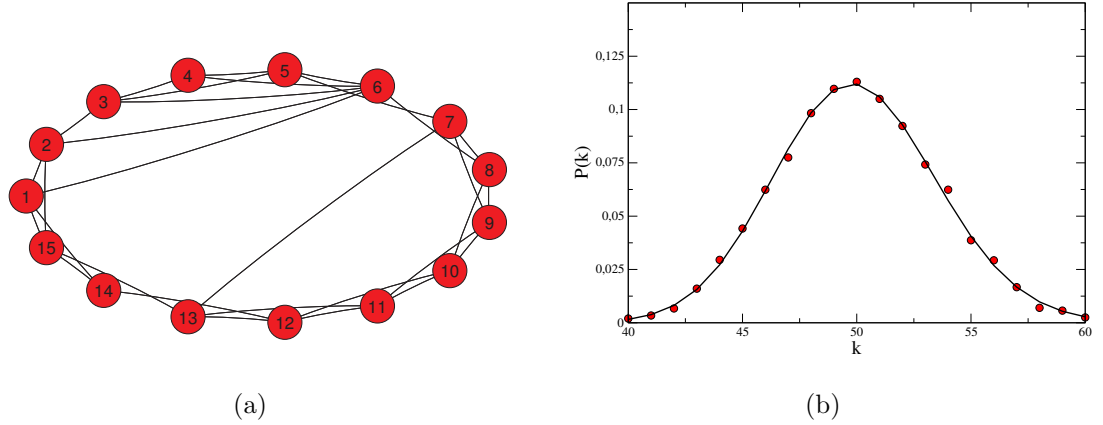


Figura 6 - (a) Exemplo de rede mundo pequeno de Watts e Strogatz com 15 vértices, $k = 2$ e $p = 0,1$. (b) Distribuição do grau de conectividade $P(k)$ para uma rede de Watts e Strogatz com 1.000 vértices, $k = 25$ e $p = 0,3$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.

Observa-se que o valor do grau de conectividade médio esperado (Eq. (27)), seja dado por $\langle \bar{k} \rangle = 50$, que é exatamente o mesmo valor obtido numericamente (Eq. (5)), dado por $\langle k \rangle = 50$. De maneira análoga, o intervalo estimado para o comprimento médio do menor caminho (Eq. (29)) é dado por $1,765 < \langle \bar{L} \rangle < 10$, e o valor numérico (Eq. (19)), $\langle L \rangle = 2,49$, se encontra nesse intervalo. O mesmo pode ser observado para o valor estimado do coeficiente de agrupamento médio (Eq. (32)), dado por $\langle \bar{C} \rangle = 0,535$ que se aproxima do valor obtido numericamente (Eq. (17)), dado por $\langle C \rangle = 0,541$.

2.3.3 Modelo de Barabási & Albert

Albert Barabási e Réka Albert (Albert & Barabási, 2002) mapearam a topologia das conexões entre as páginas da WWW (World-Wide Web) e descobriram que além da mesma apresentar o fenômeno mundo pequeno, sua distribuição de conexões não era aleatória, mas do tipo livre de escala, ou seja, da forma $P(k) \approx k^{-\gamma}$. A distribuição livre de escala é um tipo de distribuição de probabilidade que reflete a

invariância de escala. Este tipo de distribuição possui uma lei de potência associada, que é uma curva continuamente decrescente sem um pico característico sendo descrita por um único expoente. As quantidades geradas aleatoriamente possuem uma escala típica, sendo descritas por curvas características definidas por uma média e um desvio padrão. Diferentemente da uniformidade, as leis de potência sugerem que muitos eventos pequenos podem coexistir com poucos eventos grandes (Campanharo, 2011). Desta forma, uma rede livre de escala é caracterizada pela presença de poucos vértices com elevado grau de conectividade (denominados “hubs”) e grande parte dos vértices com baixo grau de conectividade.

Barabási e Albert propuseram um modelo capaz de gerar redes livres de escala, baseado nos seguintes passos:

- *Crescimento*: Inicia-se uma rede com um número reduzido de vértices desconexos N_0 . A cada passo é adicionado um novo vértice com N_0 arestas que se conectam com os vértices já presentes na rede;
- *Ligação preferencial*: O novo vértice a ser adicionado na rede tende a se conectar com os vértices com maior grau de conectividade, ou seja, a probabilidade de um vértice n_j ser escolhido é proporcional ao seu grau de conectividade.

Na distribuição livre de escala, o grau de conectividade médio $\langle k \rangle$ não é o grau de conectividade frequentemente encontrado nas distribuições de Erdős & Rényi e Watts & Strogatz (Camargo, 2016). Para uma rede livre de escala, o grau de conectividade médio $\langle \bar{k} \rangle$ pode ser estimado por:

$$\langle \bar{k} \rangle \approx 2N_0, \quad (33)$$

O comprimento médio do menor caminho, $\langle \bar{L} \rangle$ é dado, aproximadamente, por:

$$\langle \bar{L} \rangle \approx \frac{\log N}{\log(\log N)}. \quad (34)$$

Observa-se que o valor do grau de conectividade médio esperado (Eq. (33)), dado por $\langle \bar{k} \rangle = 2$, se aproxima do valor obtido numericamente (Eq. (5)),

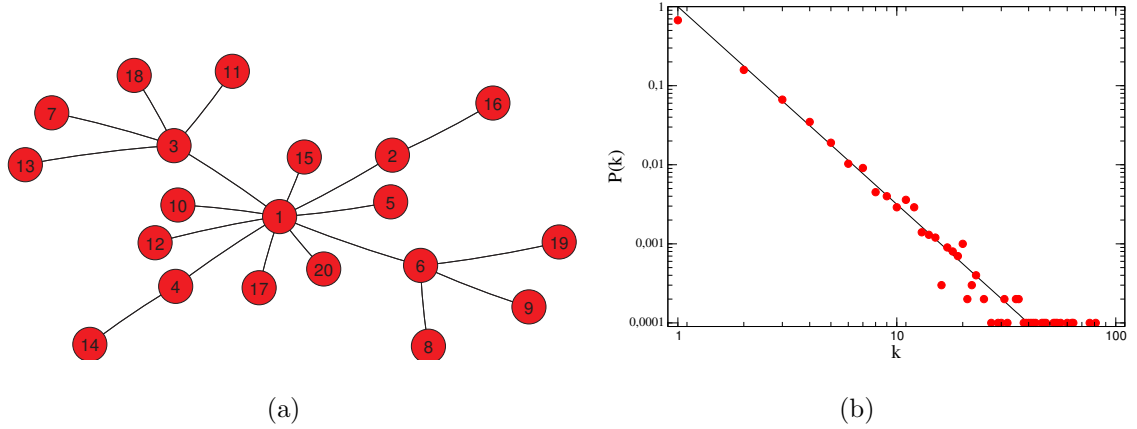


Figura 7 - (a) Exemplo de rede livre de escala de Barabási e Albert com 20 vértices e $N_0 = 1$. (b) Distribuição do grau de conectividade $P(k)$ para uma rede de Barabási e Albert com 1.000 vértices e $N_0 = 1$ e $\gamma = 2,5$ (b). Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.

dado por $\langle k \rangle = 1,99$. De maneira análoga, o valor estimado para o comprimento médio do menor caminho (Eq. (34)), dado por $\langle \bar{L} \rangle = 6,28$, é próximo do valor numérico obtido (Eq. (19)), dado por $\langle L \rangle = 6,90$.

2.3.4 Modelo modular

Um modelo de rede modular é aquele capaz de produzir redes com uma estrutura modular, ou seja, com arestas densamente distribuídas entre vértices pertencentes a um mesmo grupo (módulo) e esparsamente distribuídas entre vértices pertencentes a grupos distintos. Nesse trabalho, utiliza-se o algoritmo proposto por Mark Newman (Newman, 2002) conhecido como *modelo modular aleatório*². Mais especificamente, dados N vértices e M módulos, pares de vértices são conectados com probabilidade $0 \leq p \leq 1$ dentro de cada módulo e com probabilidade $0 \leq r \leq 1$ em módulos distintos. A distribuição de arestas dentro de cada módulo obedece o

²A implementação computacional desse algoritmo pode ser encontrada http://strategic.mit.edu/downloads.php?page=matlab_networks.

modelo de Erdős-Rényi (Subseção 2.3.1). Uma vez executado tal algoritmo, o mesmo produz uma matriz de adjacências da rede associada e um vetor indicando o módulo que cada vértice pertence.

Considerando $N = 10$, $M = 2$, $p = 0,5$ e $r = 0,9$, a matriz de adjacência produzida pelo algoritmo descrito anteriormente, é dada por:

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \quad (35)$$

Como $M = 2$, cada um dos vértices foi distribuído em um dos dois módulos, ou seja:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \end{bmatrix} \quad (36)$$

Os vértices 1, 2, 3, 4 e 5 pertencem ao módulo 1 e os vértices 6, 7, 8, 9 e 10 pertencem ao módulo 2. A Figura 8(a) apresenta a rede modular associada a matriz de adjacência em questão. A Figura 8(b) apresenta a distribuição do grau de conectividade de uma rede modular com $N = 1.000$, $M = 10$, $p = 0,5$ e $r = 0,9$.

2.3.5 Modelo hierárquico

Diversas redes reais, tais como, redes sociais, metabólicas e de interação de proteínas, além de possuírem um número reduzido de *hubs* e grande parte de seus vértices com baixo grau de conectividade, possuem um coeficiente de agrupamento elevado. Como os modelos de geração de redes livres de escala produzem redes com

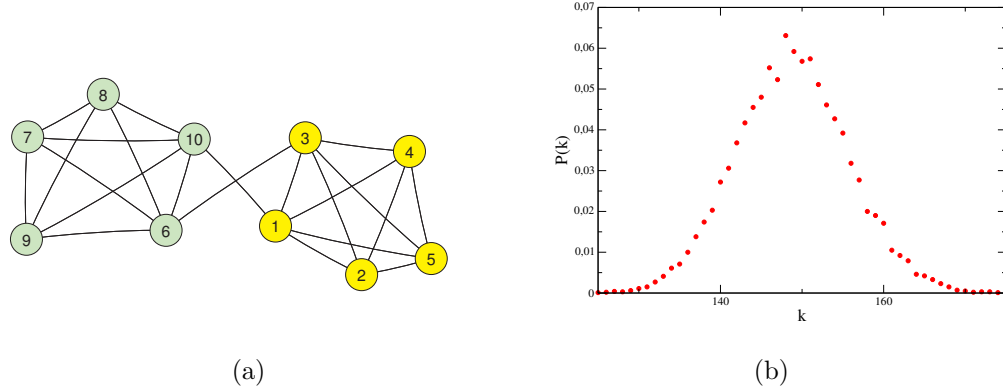
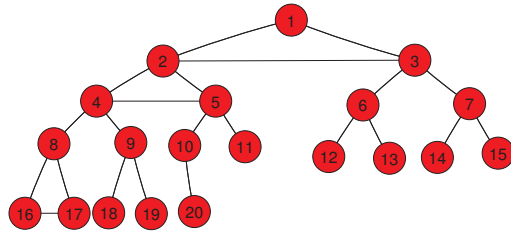


Figura 8 - (a) Exemplo de uma rede modular com 10 vértices e dois módulos. (b) Distribuição do grau de conectividade $P(k)$ para uma rede modular com $N = 1.000$, $M = 10$, $p = 0,5$ e $r = 0,9$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.

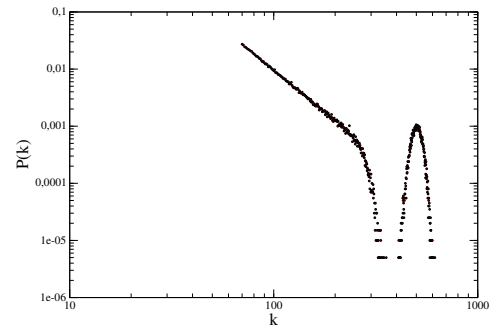
baixo coeficiente de agrupamento, os modelos hierárquicos foram propostos com o intuito sanar essa deficiência, ou seja, de produzirem redes livres de escala altamente agrupadas. Uma rede produzida por um modelo hierárquico possui uma estrutura em formato de árvore, ou seja, seus vértices são organizados hierarquicamente de cima para baixo, ou seja, de vértices com maior grau de conectividade para vértices com menor grau de conectividade. (Ravasz & Barabási, 2003).

Nesse trabalho, utiliza-se o algoritmo proposto por Peter Dodds, Duncan Watts e Charles Sabel (Dodds et al., 2003) conhecido como *modelo hierárquico aleatório*³. Com base nesse modelo, dados N vértices, o fator hierárquico h e b arestas a serem adicionadas à rede, cada vértice n_i conecta-se com h vértices até atingir os N vértices, formando assim uma rede hierárquica (Fig. 9(a)). Em seguida, b arestas são distribuídas de maneira aleatória em tal rede.

³A implementação computacional desse algoritmo pode ser encontrada http://strategic.mit.edu/downloads.php?page=matlab_networks.



(a)



(b)

Figura 9 - (a) Exemplo de uma rede hierárquica com $N = 20$, $h = 2$ e $b = 3$. (b) Distribuição do grau de conectividade $P(k)$ para uma rede hierárquica com $N = 1.000$, $h = 2$ e $b = 100$. Cada ponto no gráfico em (b) é a média aritmética dos valores de $P(k)$ sobre 10 redes.

3 MODELO BASEADO EM INDIVÍDUOS

Um MBI consiste de uma estrutura na qual ocorrem relacionamentos entre um certo número de indivíduos, cujo comportamento é determinado por um conjunto de características (Sole et al., 1999; Keeling & Grenfell, 2000). À medida que os relacionamentos ocorrem, as características de cada indivíduo podem mudar ou não.

O MBI proposto nesse trabalho baseia-se no MBI proposto por Nepomuceno (2005), e o mesmo foi desenvolvido a partir das seguintes premissas epidemiológicas:

1. S, I e R : número de indivíduos suscetíveis, infectados e recuperados, respectivamente;
2. β : *taxa de contato* - também chamado de coeficiente de transmissão, é o número médio de contatos de um indivíduo por unidade de tempo, onde há a transmissão da doença;
3. γ : *taxa de recuperação* - γ^{-1} representa o tempo médio que um indivíduo permanece na classe “infectado”;
4. μ : *taxa de renovação* ou *mortalidade* - número de indivíduos por unidade de tempo que morrem e, em mesma quantidade, nascem outros suscetíveis;
5. *Tamanho da população*: a população tem tamanho constante N ;
6. *Característica do indivíduo*: cada indivíduo possui como única característica - seu estado em relação à epidemia, e assume um e somente um dos possíveis estados;

7. *Distribuição estatística*: a mortalidade e recuperação dos indivíduos seguem uma distribuição não necessariamente uniforme em cada intervalo de tempo.

Além disso, a probabilidade de contato entre os indivíduos é proporcional ao grau de relacionamento entre os mesmos;

8. *Regras de mudança de estado*: em cada instante de tempo cada indivíduo pode alterar de estado, conforme as seguintes regras:

- $S, I, R \rightarrow S$: ocorre a renovação dos indivíduos, ou seja, os mesmos morrem e, para que a população permaneça com tamanho constante, um novo indivíduo suscetível nasce;
- $S \rightarrow I$: ocorre quando um indivíduo suscetível tem contato com um indivíduo infectado, adquirindo a doença e passando para o estado I ;
- $I \rightarrow R$: ocorre a recuperação de um indivíduo infectado, tornando-se imune a uma nova infecção. Essa premissa é válida na modelagem de doenças virais, que conferem imuidade ao indivíduo, como o sarampo e a rubéola.

3.1 Formulação matemática

Um MBI pode ser expresso por uma matriz $P_{m \times n}$, onde cada uma de suas linhas representa um indivíduo e cada uma de suas colunas representa uma dada característica, como por exemplo, sexo, idade ou localização espacial. Cada indivíduo pode ser representado por:

$$I_{m,t} = [C_1 \quad C_2 \quad \dots \quad C_n], \quad (37)$$

onde m é o número de indivíduos de uma população, t é o instante em que um indivíduo apresenta um conjunto específico de características e C_n é uma característica do indivíduo. Portanto, a população pode ser representada por:

$$P_t = [I_{1,t} \quad I_{2,t} \quad \dots \quad I_{m,t}]^T. \quad (38)$$

A formulação matemática do MBI é bastante genérica e permite incorporar várias características aos indivíduos, tornando o modelo cada vez mais realístico. Com base no MBI proposto por Nepomuceno (2005), nesse trabalho o MBI utiliza as seguintes características:

- C_1 é o estado de um indivíduo em relação à epidemia (suscetível, infectado ou recuperado). C_1 assumirá zero, caso o indivíduo esteja no estado suscetível, um no estado infectado ou dois no estado recuperado.
- C_2 é a idade de um indivíduo em anos, sendo que a cada iteração é adicionado Δt na sua idade. Em $t = 0$, a idade é zero, ou seja, $C_2 = 0$.
- C_3 é a idade máxima em que um indivíduo poderá viver. No momento do seu nascimento, essa idade é obtida por:

$$C_3 = -\mu_{mbi} \log(a_u), \quad (39)$$

onde μ_{mbi} é a expectativa de vida da população e a_u é uma variável aleatória uniformemente distribuída entre 0 e 1.

- C_4 é o tempo em anos que um indivíduo se encontra infectado. Em $t = 0$, esse tempo é zero, ou seja, $C_4 = 0$.
- C_5 é o tempo máximo em que um indivíduo permanece infectado. No momento da infecção, esse tempo é obtido por:

$$C_5 = -\gamma_{mbi} \log(a_u), \quad (40)$$

onde γ_{mbi} é o período da infecção.

Por definição, para indivíduos suscetíveis e recuperados $C_4 = 0$ e $C_5 = 0$, para $\Delta t \geq 0$. O número de indivíduos suscetíveis, infectados e recuperados em cada instante t da população é denotado por $S(t)$, $I(t)$ e $R(t)$, respectivamente.

A Figura 10 apresenta a representação esquemática do MBI para seis indivíduos que apresentam ou não mudanças de estado em relação à epidemia nos instantes $t = 0, 1, 2, 3, 4, 5, 6, 7, 8$. Mais especificamente,

$$t = 1 \quad \bullet \quad I_{2,0}(0; 0; 15; 0; 0) \rightarrow I_{2,1}(1; 1; 15; 0; 3).$$

Esse indivíduo foi infectado e o período de infecção será de 3 anos.

$$t = 2 \quad \bullet \quad \text{Não houve alteração no estado epidemiológico dos indivíduos.}$$

$$t = 3 \quad \bullet \quad I_{1,2}(0; 2; 6; 0; 0) \rightarrow I_{1,3}(1; 3; 6; 0; 2).$$

Esse indivíduo foi infectado e o período de infecção será de 2 anos.

$$\bullet \quad I_{4,2}(1; 2; 20; 2; 2) \rightarrow I_{4,3}(2; 3; 20; 0; 0).$$

Esse indivíduo atingiu o tempo máximo da infecção, então o mesmo tornou-se recuperado.

$$t = 4 \quad \bullet \quad I_{5,3}(1; 3; 25; 3; 3) \rightarrow I_{5,4}(2; 4; 25; 0; 0).$$

Esse indivíduo atingiu o tempo máximo da infecção, então o mesmo tornou-se recuperado.

$$t = 5 \quad \bullet \quad I_{2,4}(1; 4; 15; 3; 3) \rightarrow I_{2,5}(2; 5; 15; 0; 0).$$

Esse indivíduo atingiu o tempo máximo da infecção, então o mesmo tornou-se recuperado.

$$\bullet \quad I_{3,4}(0; 4; 7; 0; 0) \rightarrow I_{3,5}(1; 5; 7; 0; 3).$$

Esse indivíduo foi infectado e o período de infecção será de 3 anos.

$$t = 6 \quad \bullet \quad I_{1,5}(1; 5; 6; 2; 2) \rightarrow I_{1,6}(2; 6; 6; 0; 0).$$

Esse indivíduo atingiu o tempo máximo da infecção, então o mesmo tornou-se recuperado.

$$t = 7 \quad \bullet \quad I_{1,6}(2; 6; 6; 0; 0) \rightarrow I_{1,7}(0; 0; 15; 0; 0).$$

Esse indivíduo recuperado atingiu a idade máxima, então o mesmo foi substituído por um novo indivíduo suscetível cuja a idade máxima é de 15 anos, ou seja, $C_1 = 0$, $C_2 = 0$ e $C_3 = 15$.

$$t = 8 \quad \bullet \quad I_{3,7}(0; 7; 7; 2; 3) \rightarrow I_{1,7}(0; 0; 23; 0; 0).$$

Esse indivíduo infectado atingiu a idade máxima, então o mesmo foi substituído por um novo indivíduo suscetível cuja a idade máxima é de 23 anos, ou seja, $C_1 = 0$, $C_2 = 0$ e $C_3 = 23$.

<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>0</td><td>0</td><td>6</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>15</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>7</td><td>0</td><td>0</td></tr><tr><td>1</td><td>0</td><td>20</td><td>0</td><td>2</td></tr><tr><td>1</td><td>0</td><td>25</td><td>0</td><td>3</td></tr><tr><td>2</td><td>0</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 0$	C_1	C_2	C_3	C_4	C_5	0	0	6	0	0	0	0	15	0	0	0	0	7	0	0	1	0	20	0	2	1	0	25	0	3	2	0	10	0	0	<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>0</td><td>1</td><td>6</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>15</td><td>0</td><td>3</td></tr><tr><td>0</td><td>1</td><td>7</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>20</td><td>1</td><td>2</td></tr><tr><td>1</td><td>1</td><td>25</td><td>1</td><td>3</td></tr><tr><td>2</td><td>1</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 1$	C_1	C_2	C_3	C_4	C_5	0	1	6	0	0	1	1	15	0	3	0	1	7	0	0	1	1	20	1	2	1	1	25	1	3	2	1	10	0	0	<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>0</td><td>2</td><td>6</td><td>0</td><td>0</td></tr><tr><td>1</td><td>2</td><td>15</td><td>1</td><td>3</td></tr><tr><td>0</td><td>2</td><td>7</td><td>0</td><td>0</td></tr><tr><td>1</td><td>2</td><td>20</td><td>2</td><td>2</td></tr><tr><td>1</td><td>2</td><td>25</td><td>2</td><td>3</td></tr><tr><td>2</td><td>2</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 2$	C_1	C_2	C_3	C_4	C_5	0	2	6	0	0	1	2	15	1	3	0	2	7	0	0	1	2	20	2	2	1	2	25	2	3	2	2	10	0	0
C_1	C_2	C_3	C_4	C_5																																																																																																							
0	0	6	0	0																																																																																																							
0	0	15	0	0																																																																																																							
0	0	7	0	0																																																																																																							
1	0	20	0	2																																																																																																							
1	0	25	0	3																																																																																																							
2	0	10	0	0																																																																																																							
C_1	C_2	C_3	C_4	C_5																																																																																																							
0	1	6	0	0																																																																																																							
1	1	15	0	3																																																																																																							
0	1	7	0	0																																																																																																							
1	1	20	1	2																																																																																																							
1	1	25	1	3																																																																																																							
2	1	10	0	0																																																																																																							
C_1	C_2	C_3	C_4	C_5																																																																																																							
0	2	6	0	0																																																																																																							
1	2	15	1	3																																																																																																							
0	2	7	0	0																																																																																																							
1	2	20	2	2																																																																																																							
1	2	25	2	3																																																																																																							
2	2	10	0	0																																																																																																							
<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>1</td><td>3</td><td>6</td><td>0</td><td>2</td></tr><tr><td>1</td><td>3</td><td>15</td><td>2</td><td>3</td></tr><tr><td>0</td><td>3</td><td>7</td><td>0</td><td>0</td></tr><tr><td>2</td><td>3</td><td>20</td><td>0</td><td>0</td></tr><tr><td>1</td><td>3</td><td>25</td><td>3</td><td>3</td></tr><tr><td>2</td><td>3</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 3$	C_1	C_2	C_3	C_4	C_5	1	3	6	0	2	1	3	15	2	3	0	3	7	0	0	2	3	20	0	0	1	3	25	3	3	2	3	10	0	0	<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>1</td><td>4</td><td>6</td><td>1</td><td>2</td></tr><tr><td>1</td><td>4</td><td>15</td><td>3</td><td>3</td></tr><tr><td>0</td><td>4</td><td>7</td><td>0</td><td>0</td></tr><tr><td>2</td><td>4</td><td>20</td><td>0</td><td>0</td></tr><tr><td>2</td><td>4</td><td>25</td><td>0</td><td>0</td></tr><tr><td>2</td><td>4</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 4$	C_1	C_2	C_3	C_4	C_5	1	4	6	1	2	1	4	15	3	3	0	4	7	0	0	2	4	20	0	0	2	4	25	0	0	2	4	10	0	0	<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>1</td><td>5</td><td>6</td><td>2</td><td>2</td></tr><tr><td>2</td><td>5</td><td>15</td><td>0</td><td>0</td></tr><tr><td>1</td><td>5</td><td>7</td><td>0</td><td>3</td></tr><tr><td>2</td><td>5</td><td>20</td><td>0</td><td>0</td></tr><tr><td>2</td><td>5</td><td>25</td><td>0</td><td>0</td></tr><tr><td>2</td><td>5</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 5$	C_1	C_2	C_3	C_4	C_5	1	5	6	2	2	2	5	15	0	0	1	5	7	0	3	2	5	20	0	0	2	5	25	0	0	2	5	10	0	0
C_1	C_2	C_3	C_4	C_5																																																																																																							
1	3	6	0	2																																																																																																							
1	3	15	2	3																																																																																																							
0	3	7	0	0																																																																																																							
2	3	20	0	0																																																																																																							
1	3	25	3	3																																																																																																							
2	3	10	0	0																																																																																																							
C_1	C_2	C_3	C_4	C_5																																																																																																							
1	4	6	1	2																																																																																																							
1	4	15	3	3																																																																																																							
0	4	7	0	0																																																																																																							
2	4	20	0	0																																																																																																							
2	4	25	0	0																																																																																																							
2	4	10	0	0																																																																																																							
C_1	C_2	C_3	C_4	C_5																																																																																																							
1	5	6	2	2																																																																																																							
2	5	15	0	0																																																																																																							
1	5	7	0	3																																																																																																							
2	5	20	0	0																																																																																																							
2	5	25	0	0																																																																																																							
2	5	10	0	0																																																																																																							
<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>2</td><td>6</td><td>6</td><td>0</td><td>0</td></tr><tr><td>2</td><td>6</td><td>15</td><td>0</td><td>0</td></tr><tr><td>1</td><td>6</td><td>7</td><td>1</td><td>3</td></tr><tr><td>2</td><td>6</td><td>20</td><td>0</td><td>0</td></tr><tr><td>2</td><td>6</td><td>25</td><td>0</td><td>0</td></tr><tr><td>2</td><td>6</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 6$	C_1	C_2	C_3	C_4	C_5	2	6	6	0	0	2	6	15	0	0	1	6	7	1	3	2	6	20	0	0	2	6	25	0	0	2	6	10	0	0	<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>0</td><td>0</td><td>15</td><td>0</td><td>0</td></tr><tr><td>2</td><td>7</td><td>15</td><td>0</td><td>2</td></tr><tr><td>1</td><td>7</td><td>7</td><td>2</td><td>3</td></tr><tr><td>2</td><td>7</td><td>20</td><td>0</td><td>0</td></tr><tr><td>2</td><td>7</td><td>25</td><td>0</td><td>0</td></tr><tr><td>2</td><td>7</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 7$	C_1	C_2	C_3	C_4	C_5	0	0	15	0	0	2	7	15	0	2	1	7	7	2	3	2	7	20	0	0	2	7	25	0	0	2	7	10	0	0	<table><tr><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th></tr><tr><td>1</td><td>1</td><td>15</td><td>1</td><td>1</td></tr><tr><td>2</td><td>8</td><td>15</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>23</td><td>0</td><td>0</td></tr><tr><td>2</td><td>8</td><td>20</td><td>0</td><td>0</td></tr><tr><td>2</td><td>8</td><td>25</td><td>0</td><td>0</td></tr><tr><td>2</td><td>8</td><td>10</td><td>0</td><td>0</td></tr></table> $t = 8$	C_1	C_2	C_3	C_4	C_5	1	1	15	1	1	2	8	15	0	0	0	0	23	0	0	2	8	20	0	0	2	8	25	0	0	2	8	10	0	0
C_1	C_2	C_3	C_4	C_5																																																																																																							
2	6	6	0	0																																																																																																							
2	6	15	0	0																																																																																																							
1	6	7	1	3																																																																																																							
2	6	20	0	0																																																																																																							
2	6	25	0	0																																																																																																							
2	6	10	0	0																																																																																																							
C_1	C_2	C_3	C_4	C_5																																																																																																							
0	0	15	0	0																																																																																																							
2	7	15	0	2																																																																																																							
1	7	7	2	3																																																																																																							
2	7	20	0	0																																																																																																							
2	7	25	0	0																																																																																																							
2	7	10	0	0																																																																																																							
C_1	C_2	C_3	C_4	C_5																																																																																																							
1	1	15	1	1																																																																																																							
2	8	15	0	0																																																																																																							
0	0	23	0	0																																																																																																							
2	8	20	0	0																																																																																																							
2	8	25	0	0																																																																																																							
2	8	10	0	0																																																																																																							

Figura 10 - Representação esquemática do MBI para seis indivíduos. As interações ocorrem em intervalos de um ano. Destacam-se as mudanças de estado em relação à epidemia nos instantes $t = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.

3.2 Algoritmo do MBI

A Figura 11 mostra o fluxograma do MBI. A população inicial é determinada de modo aleatório. Em cada instante de tempo, cada indivíduo é considerado e verifica-se por meio de distribuições probabilísticas, qual transição ocorrerá. Um indivíduo infectado poderá infectar qualquer outro indivíduo da população. Após os N indivíduos serem avaliados, o tempo de simulação é incrementado em Δt unidades. O algoritmo termina quando o tempo de simulação atinge o valor final t_f .

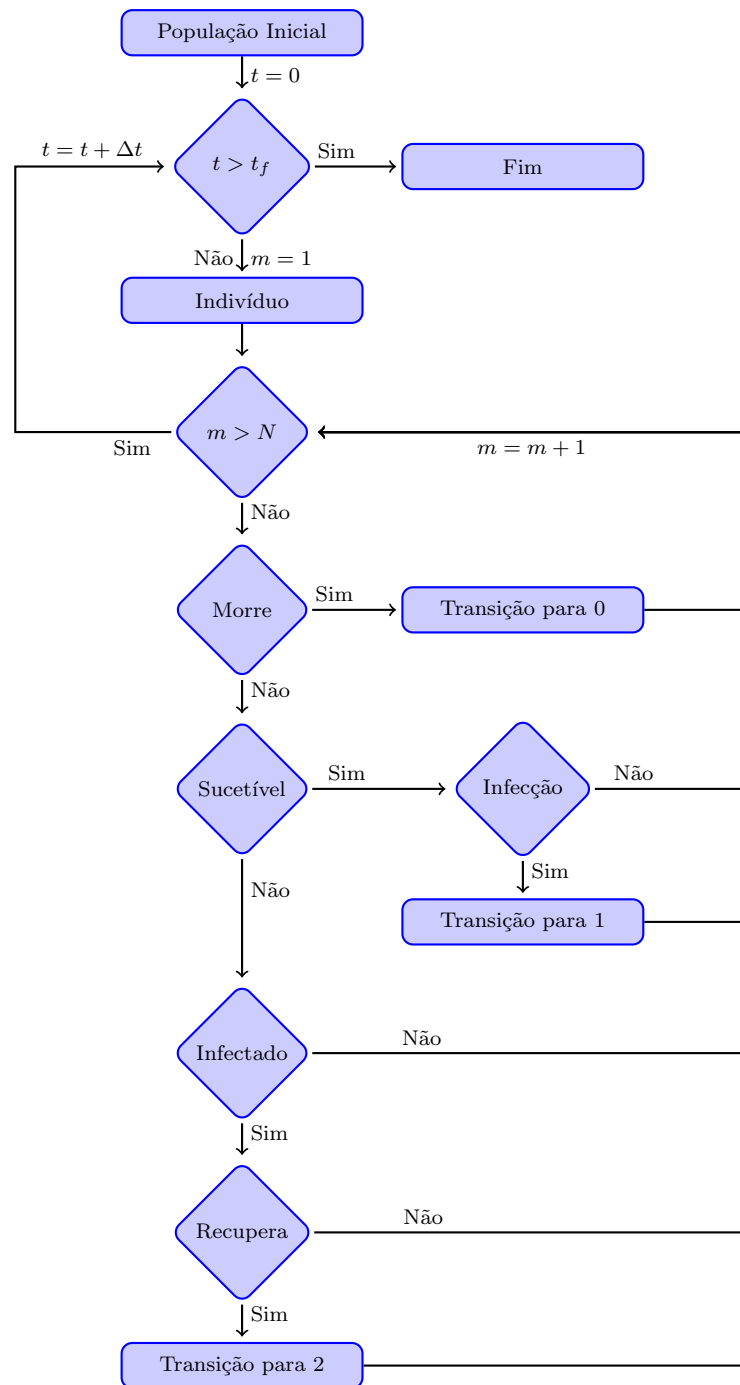


Figura 11 - Fluxograma do MBI. Os estados 0, 1 e 2 correspondem aos indivíduos suscetíveis, infectados e recuperados, respectivamente

4 ADEQUAÇÃO DO MBI AO MODELO SIR

4.1 Modelo SIR

O modelo SIR foi proposto em 1927 por Kermack e McKendrick, sendo também encontrado na literatura como K-M em homenagem a seus primeiros proponentes (Kermack & McKendrick, 1927). Tal modelo divide a população em três classes disjuntas: (S) suscetíveis, (I) infectados e (R) recuperados. O modelo clássico SIR considera que a distribuição de indivíduos é espacial e temporalmente homogênea e o mesmo descreve a evolução temporal do tamanho de cada classe epidemiológica.

A Figura 12 ilustra um fluxograma do modelo compartimental SIR. Nesse, o número total de indivíduos que morre, $\mu S + \mu I + \mu R$ (setas saindo dos compartimentos S , I e R), é igual ao número de indivíduos que é acrescido, μN , na classe de suscetíveis (seta que chega em S), mantendo a população com tamanho constante N . Com base na definição de β , $\beta I/N$ é o número médio de contatos de indivíduos infectados com suscetíveis por unidade de tempo. Consequentemente, $\beta IS/N$ é o número de novas infecções por unidade de tempo (seta que chega em I). Após a duração da infecção, γ^{-1} , um indivíduo infectado se recupera e passa para a classe recuperado (seta que chega em R).

A partir do fluxograma apresentado na Figura 12, pode-se construir

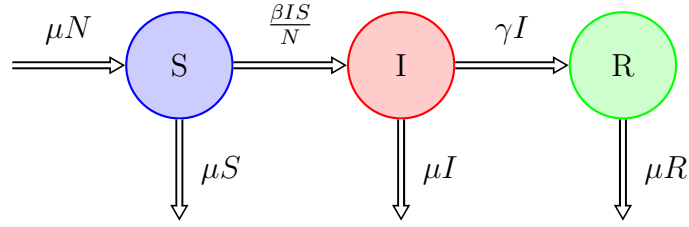


Figura 12 - Representação esquemática do modelo compartmental SIR, constituído de indivíduos suscetíveis (S), infectados (I) e recuperados (R).

um sistema de equações diferenciais ordinárias, dado por:

$$\begin{cases} \frac{dS}{dt} = \mu N - \frac{\beta IS}{N} - \mu S, & S(0) = S_0 \geq 0, \\ \frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I - \mu I, & I(0) = I_0 \geq 0, \\ \frac{dR}{dt} = \gamma I - \mu R, & R(0) = R_0 \geq 0, \end{cases} \quad (41)$$

onde:

S, I, R : número de indivíduos suscetíveis, infectados e recuperados, respectivamente;

N : número total de indivíduos da população de tamanho constante, $N = S(t) + I(t) + R(t)$ para todo $t \geq 0$;

β : *taxa de contato*: também chamado de coeficiente de transmissão, é o número médio de contatos de um indivíduo por unidade de tempo, ocasionando a transmissão da doença;

γ : *taxa de recuperação*: representa a taxa de indivíduos infectados, por unidade de tempo, que passa para a classe de recuperados;

μ : *taxa de renovação*: número de indivíduos que morrem, por unidade de tempo e, neste mesmo número, nascem outros suscetíveis.

A análise da estabilidade do Sistema (41) pode ser encontrada no Apêndice A.

4.2 Comparação entre o MBI e o modelo SIR

Ao se levar em conta que o MBI é discreto e com base nos parâmetros do modelo SIR, pode-se atribuir as seguintes relações (Nepomuceno, 2005):

$$\begin{cases} \beta_{mbi} = \beta_{sir}\Delta t, \\ \mu_{mbi} = \frac{1}{\mu_{sir}}, \\ \gamma_{mbi} = \frac{1}{\gamma_{sir}}. \end{cases} \quad (42)$$

Desta forma, existe uma equivalência entre o modelo SIR e o MBI, e conseqüentemente, as soluções dos mesmos podem apresentar comportamentos similares (Almeida, 2011). O Apêndice C apresenta uma descrição detalhada da equivalência entre ambos os modelos.

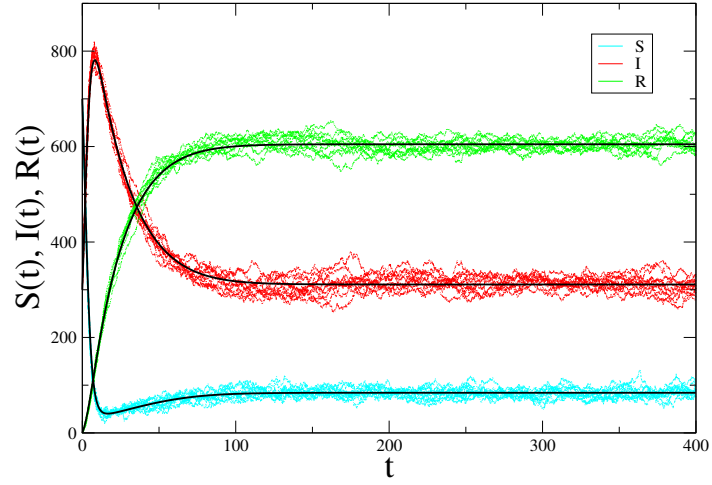
A Figura 13 ilustra a equivalência entre as soluções dos modelos SIR e MBI. Mais especificamente, a Figura 13(a) apresenta a solução do modelo SIR (linhas em preto) para o estado epidêmico (número reprodutivo básico $\mathcal{R}_0 > 1$), $\mu_{sir} = \frac{1}{70}$, $\gamma_{sir} = \frac{1}{36}$ e $\beta_{sir} = 0,5$. São ainda apresentadas as soluções do MBI para $\mu_{mbi} = 70$, $\gamma_{mbi} = 36$, $\beta_{mbi} = 0,05$ e $\Delta t = 0,1$, obtidas sobre 10 diferentes realizações. Em ambos os modelos foram considerados $N = 1.000$, $S(0) = 700$, $I(0) = 300$ e $R(0) = 0$. Pode-se observar ainda que, para todas as realizações efetuadas, as soluções do MBI se aproximam da solução do modelo SIR. A proximidade entre as soluções de tais modelos ocorre independentemente do conjunto de parâmetros utilizado, conforme mostra a Figura 13(b). Nesse caso, foram considerados para o modelo SIR os valores $\mu_{sir} = \frac{1}{70}$, $\gamma_{sir} = \frac{1}{24}$ e $\beta_{sir} = 0,05$ ($\mathcal{R}_0 < 1$) e, para o MBI, os valores $\mu_{mbi} = 70$, $\gamma_{mbi} = 24$ e $\beta_{mbi} = 0,005$.

Observou-se, ainda, que as médias amostrais dos indivíduos infectados da Figura 13(a), $\bar{x}_{mbi} = 568,3$ e $\bar{x}_{sir} = 568,7$ e da Figura 13(b), $\bar{x}_{mbi} = 100,7$ e $\bar{x}_{sir} = 101,9$, não possuem diferenças significativas. A fim de quantificar com maior precisão tais diferenças, calculou-se o intervalo de confiança para a comparação das suas médias populacionais, ϕ_{mbi} e ϕ_{sir} , ao nível de significância de 5%, ($IC_{\phi_{mbi}-\phi_{sir}}$).

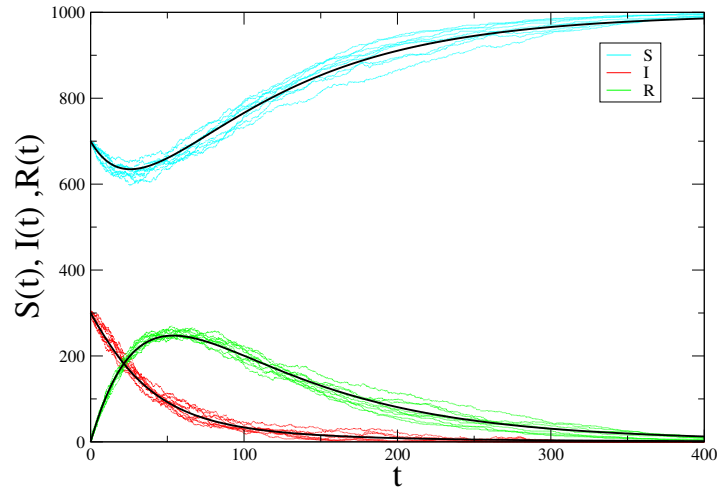
Supondo-se que a hipótese nula seja a igualdade das médias dos dois modelos e, conseqüentemente, a alternativa como a diferença entre elas, através do resultado do teste t, determinou-se o valor- p associado a esta estatística. A Tabela 2 exhibe os resultados obtidos da comparação das soluções produzidas pelos modelos SIR e MBI. Observa-se que não há diferenças estatisticamente significativas entre as médias populacionais dos dois modelos, o que implica na não rejeição da hipótese nula, o que pode ser comprovado pelo valor- p , sendo valor- $p \in [0, 1]$. Da tabela observa-se que esta estatística aponta mais evidências a favor da hipótese de nulidade, corroborando com o resultado do IC. Portanto, à 95% de confiança $\phi_{mbi} = \phi_{sir}$.

Tabela 2: Resultados obtidos pelo teste t-Student da comparação das soluções produzidas pelos modelos SIR e MBI.

	$IC(\phi_{mbi} - \phi_{sei}; 0, 95)$	valor- p
Figura 13(a)	$(-3,7520; 4,5452)$	0,8514
Figura 13(b)	$(-3,2027; 2,1063)$	0,6856



(a)



(b)

Figura 13 - Número de indivíduos suscetíveis, infectados e recuperados em função do tempo para os modelos SIR e MBI. Em (a) foram utilizados os parâmetros $N = 1.000$, $\mu_{sir} = \frac{1}{70}$, $\mu_{mbi} = 70$, $\gamma_{sir} = \frac{1}{36}$, $\gamma_{mbi} = 36$, $\beta_{sir} = 0,5$, $\beta_{mbi} = 0,05$, $\Delta t = 0,1$, $S(0) = 700$, $I(0) = 300$, $R(0) = 0$ e $\mathcal{R}_0 \approx 11,89$. Em (b) foram utilizados os parâmetros $N = 1.000$, $\mu_{sir} = \frac{1}{70}$, $\mu_{mbi} = 70$, $\gamma_{sir} = \frac{1}{24}$, $\gamma_{mbi} = 24$, $\beta_{sir} = 0,05$, $\beta_{mbi} = 0,005$, $\Delta t = 0,1$, $S(0) = 700$, $I(0) = 300$, $R(0) = 0$ e $\mathcal{R}_0 \approx 0,89$. As linhas pretas representam a solução do modelo SIR e as demais representam o conjunto de dez soluções do MBI.

5 O MBI COM REDES COMPLEXAS

Na formulação original do MBI e no modelo SIR, as relações entre os indivíduos são representadas por grafos completos, onde todos os indivíduos estão conectados entre si. Contudo, a topologia de uma rede real não pode ser descrita por uma rede puramente aleatória. Desta forma, nesse trabalho o MBI foi adaptado de forma a incorporar modelos mais realísticos de redes de contato na propagação de uma doença infecciosa entre um conjunto de indivíduos. Nesse sentido, o objetivo foi analisar o efeito da topologia de tais redes nas soluções produzidas pelo MBI e, como tais, se desviam da solução produzida pelo modelo SIR. Mais especificamente, para cada tipo de simulação na qual o MBI foi executado, utilizou-se os seguintes modelos de redes complexas:

1. Rede complexa aleatória de Erdős e Rényi (Erdős & Rényi, 1959);
2. Rede complexa mundo pequeno de Watts e Strogatz (Watts & Strogatz, 1998);
3. Rede complexa livre de escala de Barabási e Albert (Barabási & Albert, 1999);
4. Rede complexa modular de Newman (Newman, 2003);
5. Rede complexa hierárquica de Dodds *et al.* (Dodds et al., 2003).

A Figura 14 mostra o fluxograma do MBI com redes complexas. A população inicial é determinada de modo aleatório. Em cada instante de tempo, cada indivíduo é considerado e verifica-se, por meio de distribuições probabilísticas, qual transição ocorrerá. Diferentemente do MBI sem redes complexas, um indivíduo infectado poderá infectar outro indivíduo se os mesmos estiverem conectados. Após

os N indivíduos serem avaliados, o tempo de simulação é incrementado em Δt unidades. O algoritmo termina quando o tempo de simulação atinge o valor final t_f .

5.1 Estimação de Parâmetros

Uma vez conhecida a solução produzida pelo MBI para uma dada rede complexa, estima-se os parâmetros μ , β e γ (Eq. (41)) que melhor adequam a solução do modelo SIR à tal solução.

Dada uma série temporal de indivíduos infectados I_{obs} e um conjunto de parâmetros $\wp = \{\mu, \beta, \gamma\}$, define-se como função custo $Custo(I_{obs}, \wp)$ o erro quadrático entre I_{obs} e a curva de infectados obtida pelo modelo SIR com os parâmetros $\wp$, denotada por $I_\wp$. Assim:

$$Custo(I_{obs}, \wp) = \sum_t (I_{obs}(t) - I_\wp(t))^2, \quad (43)$$

onde t representa o instante de tempo. O conjunto de parâmetros $\hat{\wp}$, cujo modelo SIR melhor representa I_{obs} , é dado por:

$$\hat{\wp} = \arg \min_{\wp} Custo(I_{obs}, \wp). \quad (44)$$

Na Equação (44), $\hat{\wp}$ pode ser obtido por vários métodos de otimização. Nesse trabalho, utilizou-se o método de Levenberg-Marquardt (Marquardt, 1963). O método de Levenberg-Marquardt é frequentemente utilizado na resolução de problemas de ajustes de curvas ou modelos de mínimos quadrados não lineares (dos Santos Benedito & Pio dos Santos, 2017). Trata-se de um método muito eficiente, que combina as principais qualidades de dois métodos clássicos de otimização, a saber, a convergência do método de *Gauss-Newton* e a estabilidade do método do *Gradiente* (de máxima descida) na busca de uma dada solução, nesse caso, do conjunto de parâmetros ótimo. Uma descrição mais detalhada do método de Levenberg-Marquardt pode ser encontrada no Apêndice B.

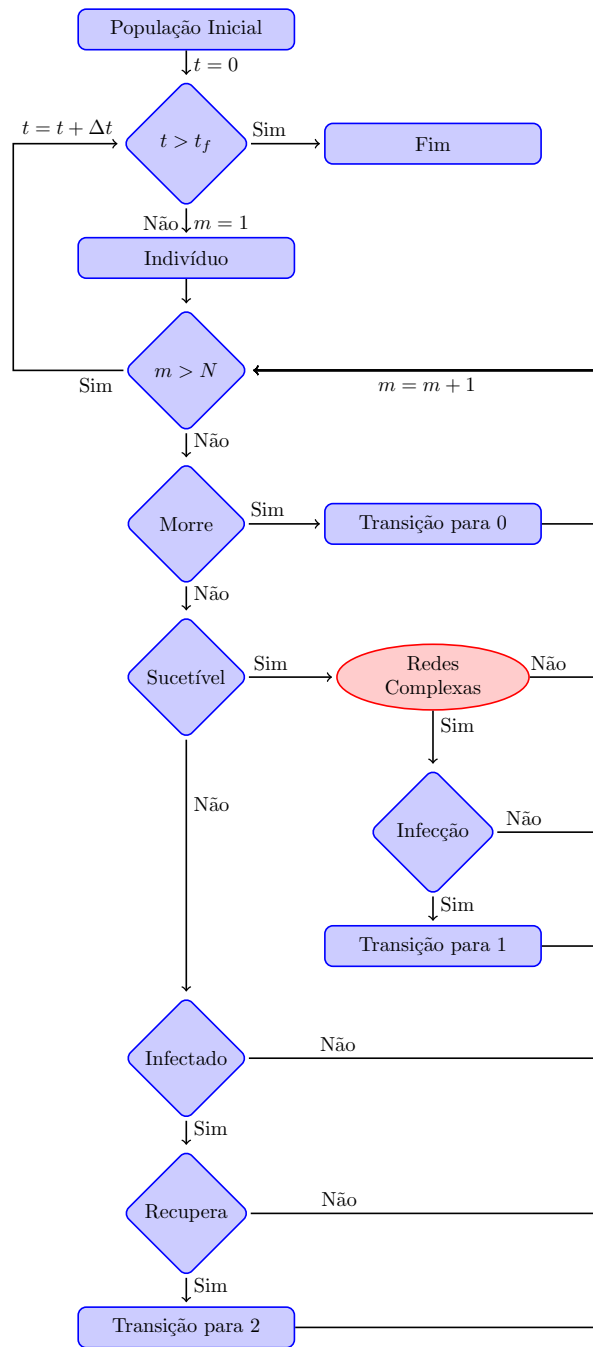


Figura 14 - Fluxograma do MBI com redes complexas para o modelo SIR. Os estados 0, 1 e 2 correspondem aos indivíduos suscetíveis, infectados e recuperados, respectivamente.

5.2 Modelos de redes complexas utilizados

Redes complexas com topologias aleatória, mundo pequeno, livre de escala, modular e hierárquica foram geradas utilizando-se diversos parâmetros, conforme mostra a Tabela 3. Tal tabela apresenta, ainda, as medidas topológicas utilizadas para a caracterização das redes em estudo, tais como, o grau de conectividade médio, o comprimento médio do menor caminho, o coeficiente de agrupamento médio, o comprimento do salto médio e modularidade. Para efeito de comparação dos resultados obtidos, a densidade de conexões dessas redes foi fixada em $d = 0,07$, valor frequentemente encontrado em redes de contatos reais (Fournet & Barrat, 2016).

Para cada tipo de rede em estudo, foram obtidas as curvas médias de indivíduos suscetíveis, infectados e recuperados, com base em 1.000 diferentes realizações do MBI correspondente. Os parâmetros utilizados em cada realização foram $N = 2000$, $S_0 = 1990$, $I_0 = 10$, $R_0 = N - S_0 - I_0$, $\gamma_{mbi} = 20$, $\beta_{mbi} = 0,15$ e $\mu_{mbi} = 100$.

Dada uma rede complexa com densidade d , $0 < d \leq 1$ e taxa de contato entre indivíduos β_{sir} , é esperado que as simulações do MBI sobre tal rede forneça uma nova taxa de contato $\bar{\beta}_{sir} = d\beta_{sir}$ e os parâmetros originais γ_{sir} e μ_{sir} se mantenham os mesmos, uma vez que eles não têm relação com a estrutura social na qual a doença está inserida (Almeida, 2011). Além disso, com base no método de estimação de parâmetros descrito na Seção 5.1, foram estimados os valores μ_{sir} , γ_{sir} e β_{sir} (Tabela 3), de tal forma que a solução produzida pelo modelo SIR se tornasse mais próxima possível à da solução do MBI correspondente.

A Figura 15 ilustra as curvas médias de indivíduos suscetíveis e infectados, solução do MBI sobre cada uma das redes em estudo, e a solução do modelo SIR correspondente. Conforme esperado, a solução do MBI com uma rede aleatória e a do modelo SIR correspondente coincidem. Contudo, para as demais redes há uma diferença entre as soluções dos modelos SIR e MBI. Isso se deve ao fato das topologias dessas redes se desviarem da topologia de uma rede puramente aleatória.

As Figuras 16, 17 e 18 ilustram as soluções do modelo MBI, mais

especificamente, os números de indivíduos suscetíveis, infectados e recuperados para os cinco modelos de redes complexas em estudo. Os parâmetros utilizados foram $N = 2000$, $S_0 = 1990$, $I_0 = 10$, $R_0 = N - S_0 - I_0$, $\gamma_{mbi} = 20$, $\beta_{mbi} = 0,15$ e $\mu_{mbi} = 100$. Observa-se, de modo geral, que redes complexas com diferentes topologias resultam em curvas de indivíduos suscetíveis, infectados e recuperados com diferentes comportamentos. Desta forma, a evolução de uma dada doença é altamente sensível à topologia de rede utilizada.

Mais especificamente, observa-se que quanto menor o valor de $\langle L \rangle$ e maior o valor de Δ , mais rápida será a propagação da doença e, conseqüentemente, maior será o número de indivíduos infectados (Tabela 3)

As Figuras 19, 20, 21, 22 e 23 mostram exemplos de redes com 20 vértices, obtidas a partir dos modelos em estudo para $0 \leq t \leq 5$. Em todos os casos, vértices em azul, vermelho e verde representam indivíduos suscetíveis, infectados e recuperados, respectivamente. Observa-se que na rede aleatória, as conexões de longo alcance fazem com que um dado vértice infecte, preferencialmente, seus vizinhos mais distantes. Em uma rede mundo pequeno, tal infecção ocorre, preferencialmente, entre vizinhos mais próximos de um dado vértice. Em uma rede livre de escala, observa-se que a propagação da doença ocorre do centro da rede para a periferia da mesma, já que os vértices centrais possuem graus de conectividade superiores aos vértices periféricos. Em uma rede modular, a doença se propaga, preferencialmente, entre vértices pertencentes a um mesmo módulo. Por fim, em uma rede hierárquica, existe um sentido preferencial de propagação da doença, ou seja, do topo da rede (vértices com maior hierarquia) para a base (vértices com menor hierarquia).

Tabela 3: Parâmetros para geração das redes aleatória (ER), mundo pequeno (WS), livre de escala (BA), modular (MOD) e hierárquica (HIE) e as características topológicas de tais redes.

	p	k	n_0	m	r	h	b	$\hat{\mu}$	$\hat{\beta}$	$\hat{\gamma}$	$\langle k \rangle$	$\langle L \rangle$	$\langle C \rangle$	Δ	$\mathcal{M}$
ER	0,07	-	-	-	-	-	-	0,0107	0,1033	0,0520	139,9	1,93	0,07	667,83	0,07
WS	0,01	70	-	-	-	-	-	0,0192	0,1112	0,0613	140	2,628	0,723	75,68	0,717
BA	-	-	70	-	-	-	-	0,0132	0,1909	0,0789	139,9	1,93	0,183	702,96	0,065
MOD	0,07	-	-	10	0,95	-	-	0,0134	0,0944	0,0518	140,1	2,239	0,605	100,32	0,849
HIE	-	-	-	-	-	2	137931	0,0140	0,2061	0,0873	139,9	1,94	0,393	757,07	0,192

p probabilidade de conexão (ER e MOD), e probabilidade de reconexão (WS)

k número de vizinhos à direita

n_0 número inicial de vértices da rede

m número de módulos

r razão de ligações dentro e fora do módulo

h fator hierárquico

b número de ligações adicionais na rede

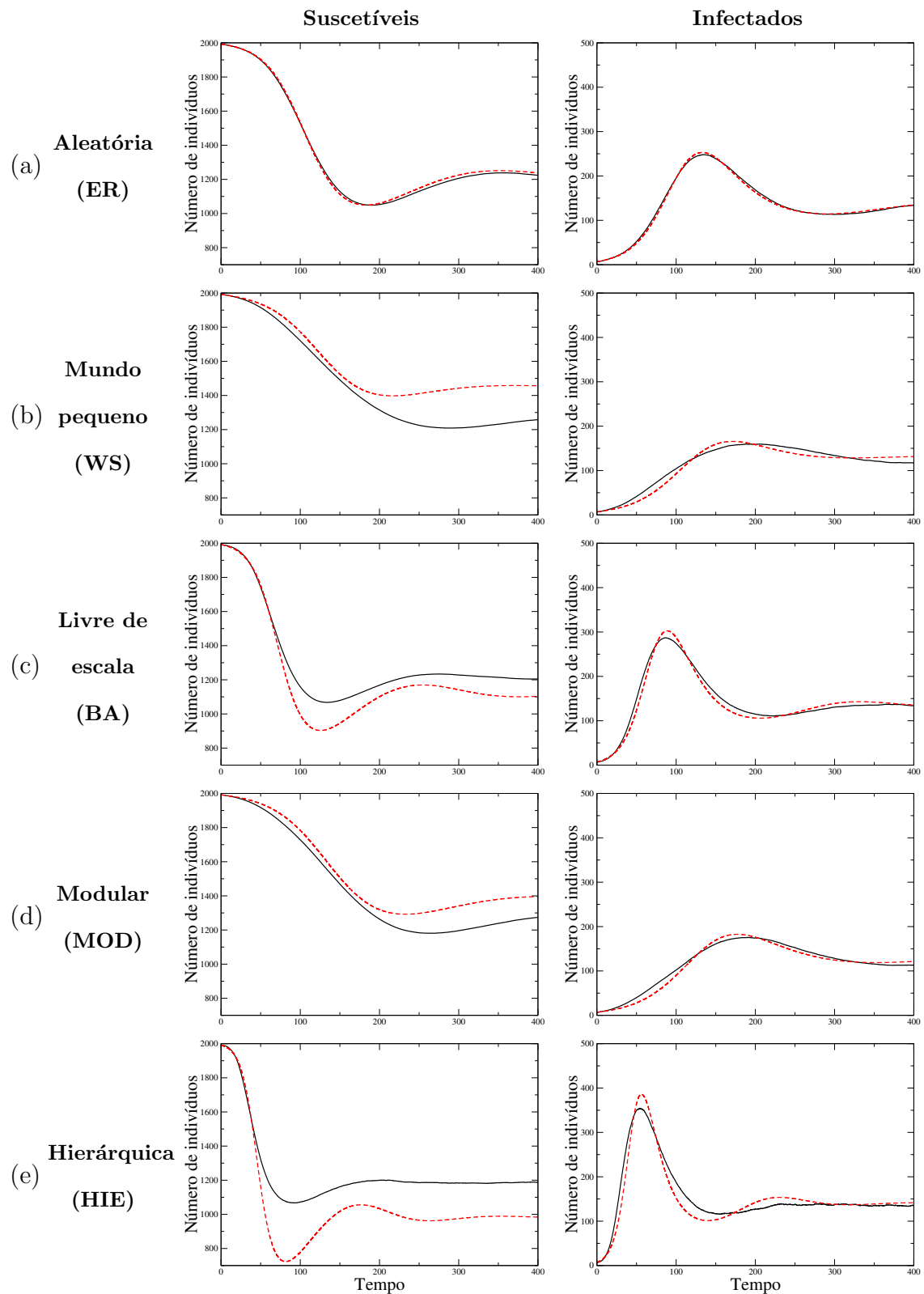


Figura 15 - Comparação do MBI com redes complexas (linha contínua) com o modelo SIR (linha tracejada).

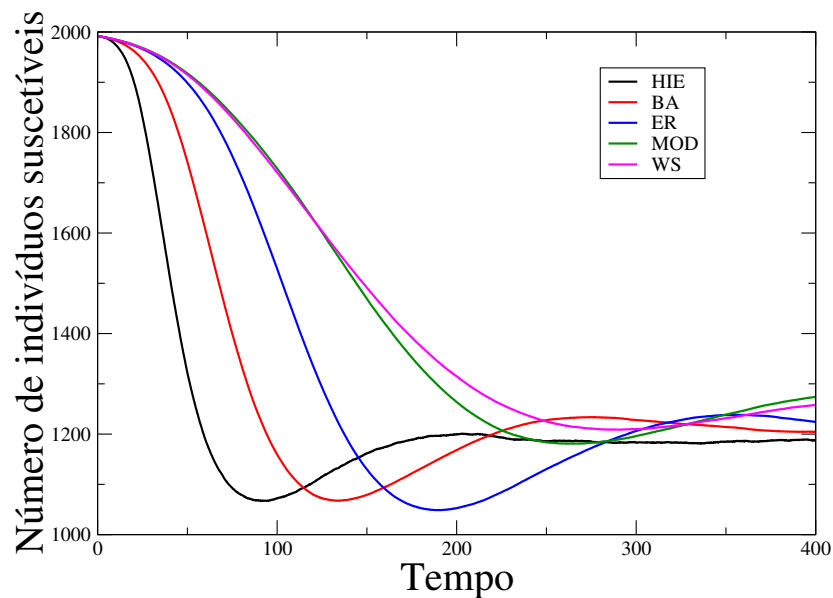


Figura 16 - Comparação das curvas de indivíduos suscetíveis, onde cada curva é a média aritmética de 1000 simulações do MBI com os modelos de redes complexas em estudo.

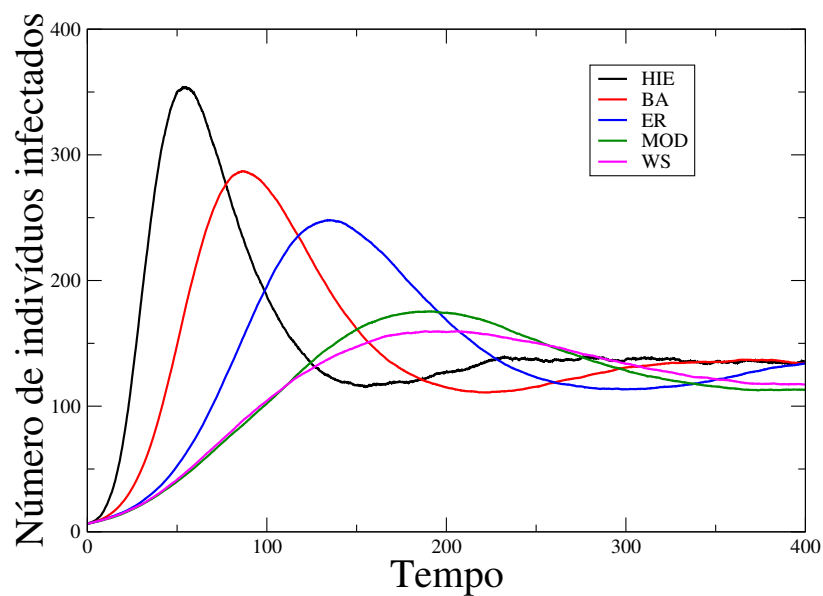


Figura 17 - Comparação das curvas de indivíduos infectados, onde cada curva é a média aritmética de 1000 simulações do MBI com os modelos de redes complexas em estudo.

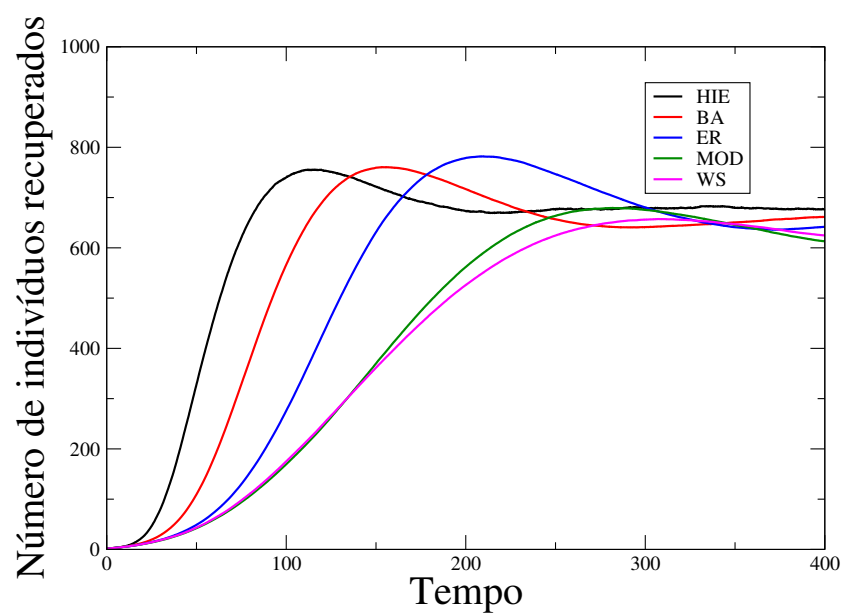


Figura 18 - Comparação das curvas de indivíduos recuperados, onde cada curva é a média aritmética de 1000 simulações do MBI com os modelos de redes complexas em estudo.

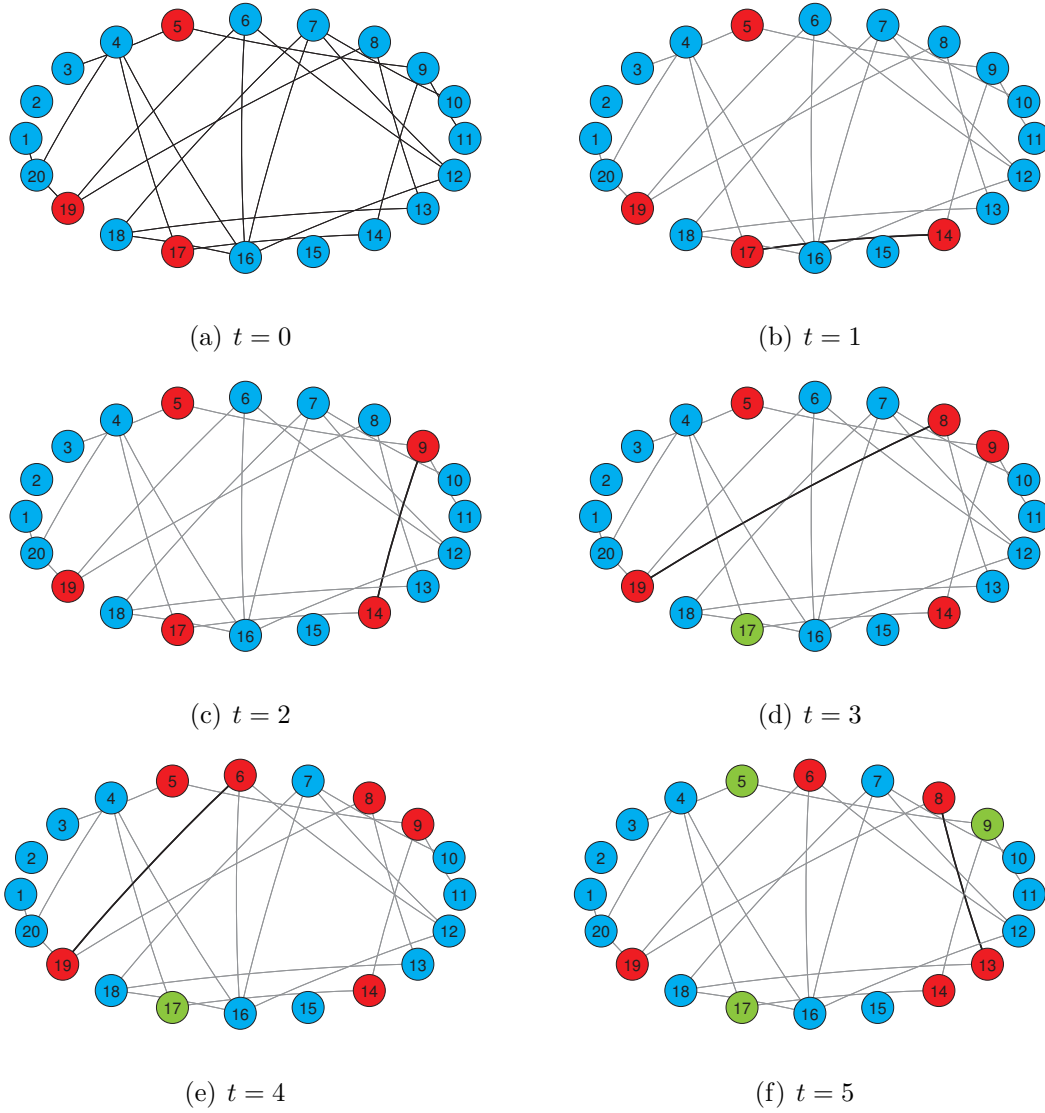


Figura 19 - Propagação de uma doença em uma rede aleatória com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 14 infecta o vértice 17, em $t = 2$, o vértice 14 infecta o vértice 9, em $t = 3$, o vértice 19 infecta o vértice 8, em $t = 4$, o vértice 19 infecta o vértice 6 e em $t = 5$, o vértice 8 infecta o vértice 13. Observa-se, a partir de $t = 3$ o surgimento de vértices em verde (indivíduos recuperados).

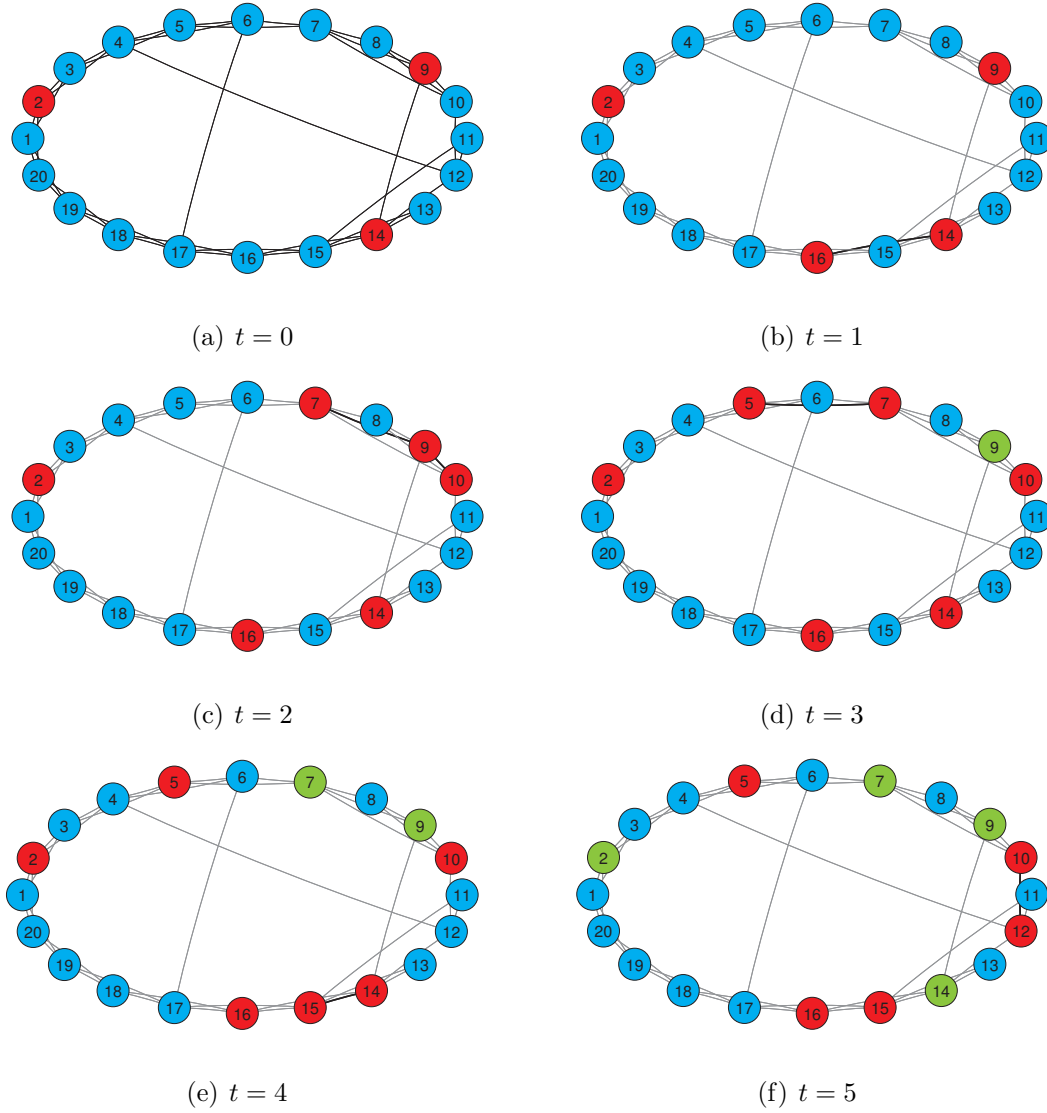


Figura 20 - Propagação de uma doença em uma rede mundo pequeno com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 14 infecta o vértice 16, em $t = 2$, o vértice 9 infecta os vértices 7 e 10, em $t = 3$, o vértice 7 infecta o vértice 5, em $t = 4$, o vértice 14 infecta o vértice 15 e em $t = 5$, o vértice 10 infecta o vértice 12. Observa-se, a partir de $t = 3$ o surgimento de vértices em verde (indivíduos recuperados).

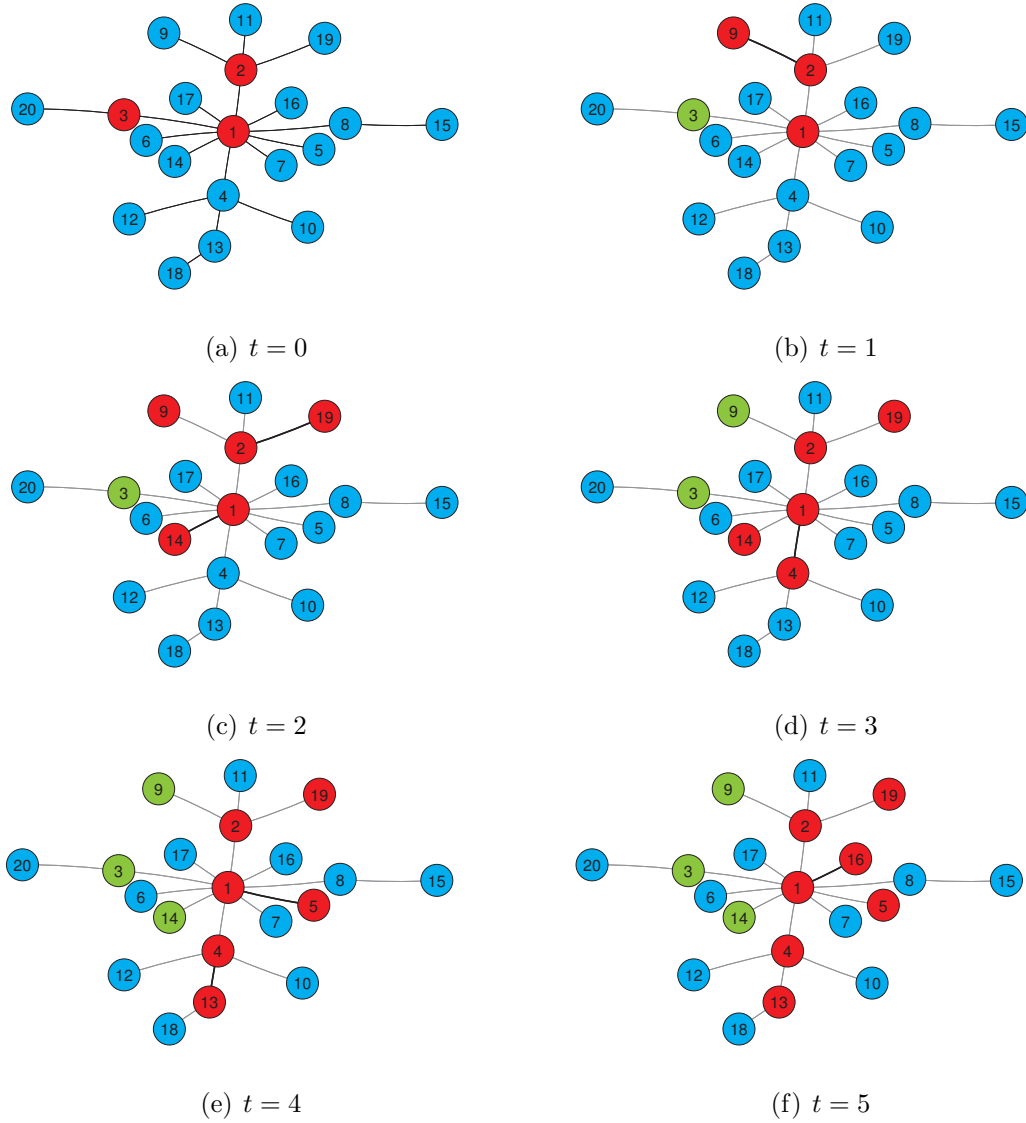


Figura 21 - Propagação de uma doença em uma rede livre de escala com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 2 infecta o vértice 9, em $t = 2$, o vértice 2 infecta o vértice 19 e o vértice 1 infecta o vértice 14, em $t = 3$, o vértice 1 infecta o vértice 4, em $t = 4$, o vértice 4 infecta o vértice 13 e o vértice 1 infecta o vértice 5 e em $t = 5$, o vértice 1 infecta o vértice 16. Observa-se, a partir de $t = 1$ o surgimento de vértices em verde (indivíduos recuperados).

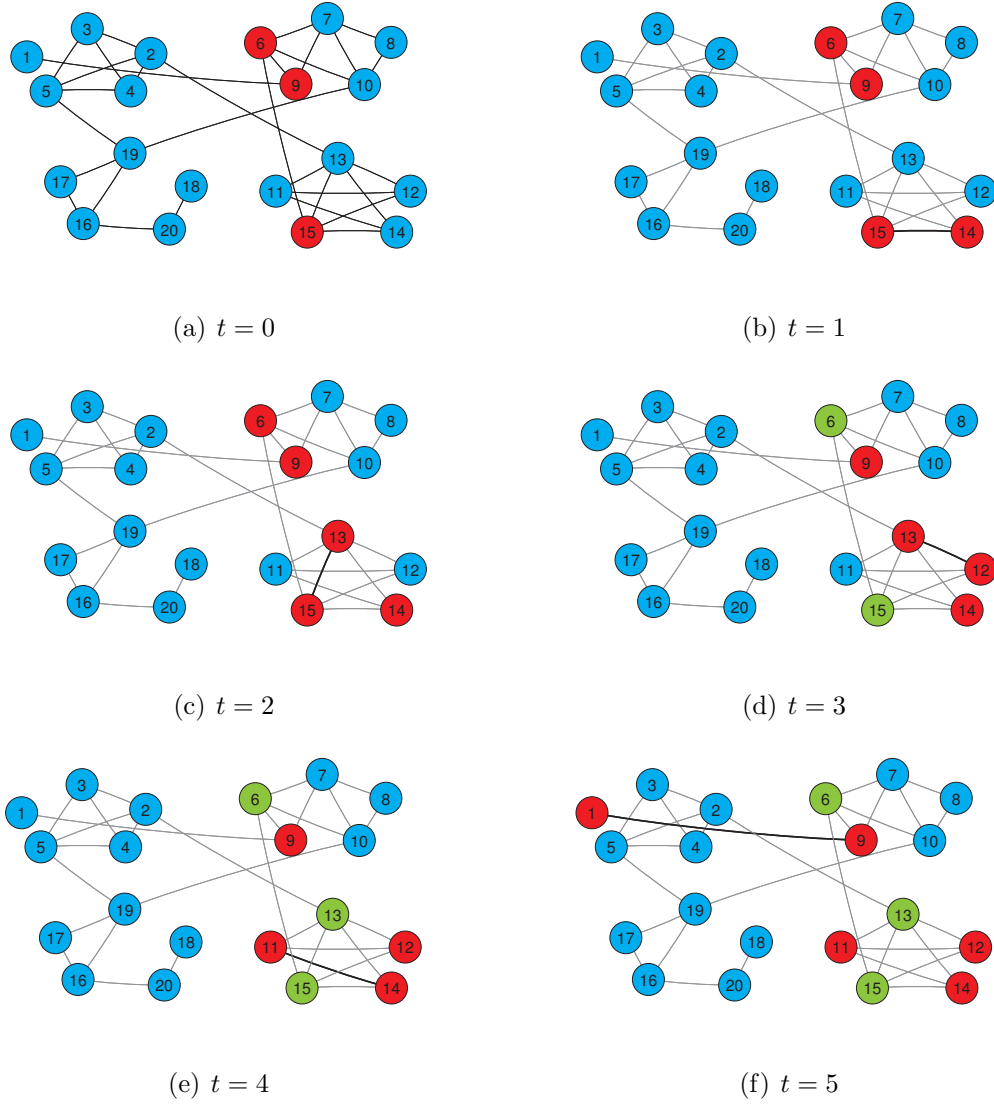


Figura 22 - Propagação de uma doença em uma rede modular com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 15 infecta o vértice 14, em $t = 2$, o vértice 15 infecta o vértice 13, em $t = 3$, o vértice 13 infecta o vértice 12, em $t = 4$, o vértice 14 infecta o vértice 11 e em $t = 5$, o vértice 9 infecta o vértice 1. Observa-se, a partir de $t = 3$ o surgimento de vértices em verde (indivíduos recuperados).

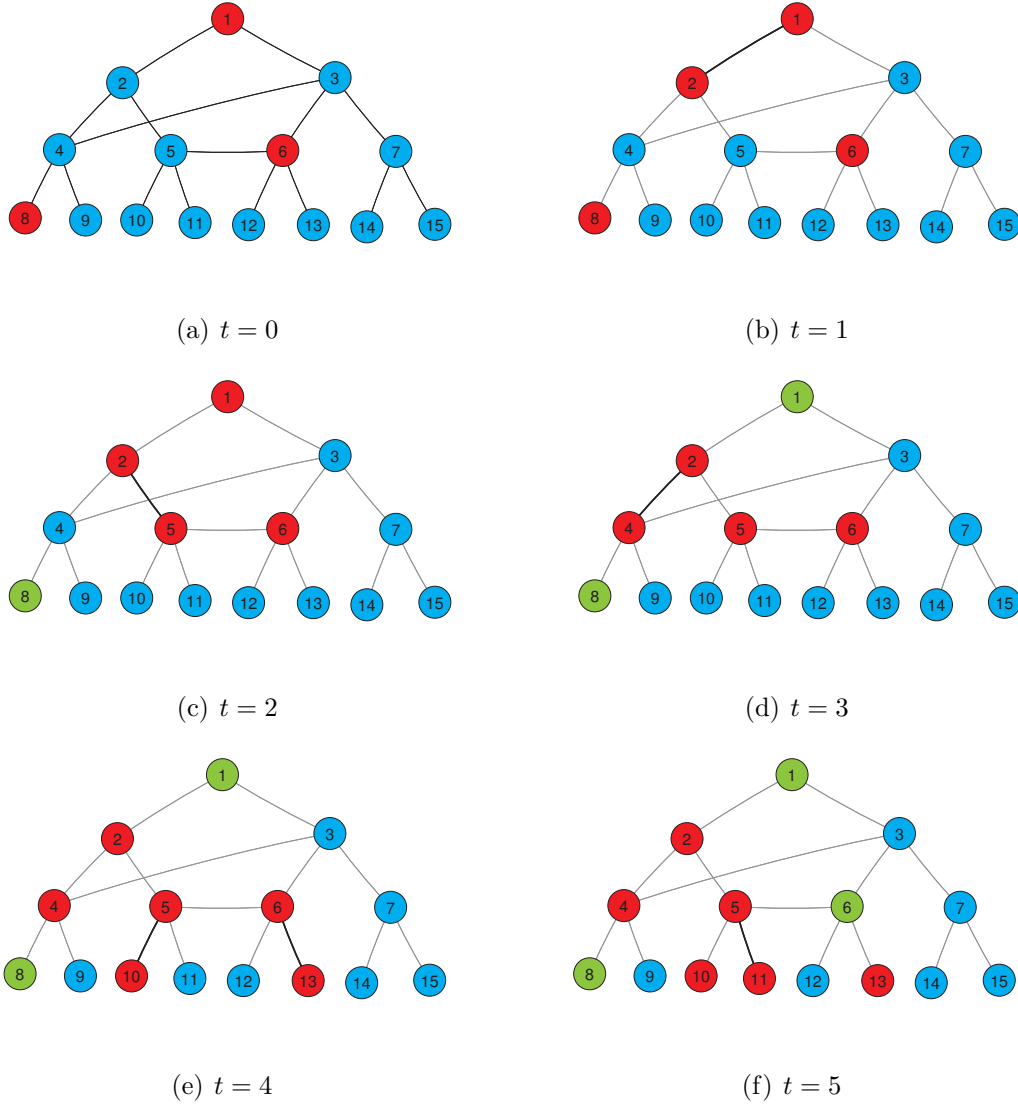


Figura 23 - Propagação de uma doença em uma rede hierárquica com $N = 20$ vértices e $t = 0, 1, 2, 3, 4$ e 5 . Inicialmente, 3 vértices são infectados aleatoriamente ($t = 0$). Em $t = 1$, o vértice 1 infecta o vértice 2, em $t = 2$, o vértice 2 infecta o vértice 5, em $t = 3$, o vértice 2 infecta o vértice 4, em $t = 4$, o vértice 5 infecta o vértice 10 e o vértice 6 infecta o vértice 13 e em $t = 5$, o vértice 5 infecta o vértice 11. Observa-se, a partir de $t = 2$ o surgimento de vértices em verde (indivíduos recuperados).

6 TUBERCULOSE

A tuberculose é uma doença infecto-contagiosa causada por uma bactéria que afeta principalmente os pulmões, ela tem como agente etiológico o *Mycobacterium tuberculosis* que foi identificado em 1882 por Robert Koch, e tem como principal hospedeiro e transmissor o homem. A transmissão da tuberculose ocorre de forma aérea e direta, ou seja, de pessoa a pessoa. A pessoa portadora de tuberculose ativa (tuberculose pulmonar) expõe, ao falar, espirrar ou tossir, pequenas gotas de saliva que contêm o agente infeccioso e podem ser aspiradas por outro indivíduo, contaminando-o. Má alimentação, falta de higiene, tabagismo, alcoolismo ou qualquer outro fator que gere baixa resistência orgânica, também favorece o estabelecimento da tuberculose (Marcolino, 2016).

A expansão da tuberculose tem como aliada a falsa ideia de que a doença está erradicada devido o advento do tratamento antimicrobiano no final da década de 60. Ela é frequentemente confundida com outras doenças, o que faz com que o indivíduo não tenha o tratamento necessário, promovendo não somente um maior comprometimento da saúde do mesmo, como também a disseminação da doença (Marcolino, 2016). A quimioterapia tem duração mínima de 6 meses, e após um mês do início do tratamento o indivíduo tem a sensação de já estar curado e acaba abandonando o tratamento, continuando a ser o transmissor da tuberculose (Gomes, 2004).

Alguns pacientes que se encontram no estado exposto não exibem nenhum indício da tuberculose, outros apresentam sintomas aparentemente simples que são ignorados durante alguns meses (ou anos). Contudo, a maioria dos infectados com tuberculose apresenta sintomas como cansaço excessivo, falta de apetite e

tosse. Nos casos mais graves, o mesmo pode apresentar dificuldades na respiração, eliminação de grande quantidade de sangue e colapso do pulmão (Marcolino, 2016).

Cerca de 10% dos indivíduos que tem contato com o agente infeccioso terão tuberculose. Desses, em torno de 5% desenvolverão a doença nos primeiros anos, ou seja, por progressão direta e os outros 5% poderão desenvolver ao longo de sua vida de forma lenta, ou seja, por reativação endógena (Gomes, 2004).

No Brasil foram diagnosticados 68 mil novos casos de tuberculose e cerca de 4,5 mil óbitos em 2015 (WHO, 2016). De acordo com a classificação da OMS 2016-2020, o Brasil ocupa a 20^a posição na lista dos 30 países com maior incidência de tuberculose. Em consequência, a OMS propôs diversas estratégias para a erradicação da tuberculose, estabelecendo metas para os países no programa “*Estratégia pós-2015 - Um mundo livre da tuberculose*”. Espera-se que até o ano de 2035, reduza-se em 95% o número de mortes causada pela tuberculose e em 90% a taxa de incidência. Tais estratégias incluem o diagnóstico precoce e tratamento da doença, além de campanhas de vacinação.

Modelos matemáticos podem atuar como ferramentas valiosas para a elaboração de estratégias no controle da tuberculose, podendo inferir informações sobre o impacto epidemiológico das vacinas, identificar quais e quantos indivíduos devem-se vacinar, verificar se a doença em estudo é epidêmica ou endêmica e fazer projeções futuras do número de infectados. Diversas técnicas de modelagem da tuberculose têm sido propostas na literatura através do uso de modelos compartimentais em conjunto com equações diferenciais ordinárias e parciais, cadeias de Markov, autômatos celulares, métodos de otimização, entre outros (Harris et al., 2016; Cohen et al., 2008; Castillo-Chavez & Feng, 1998; Bhunu et al., 2008; Ditkowsky & Schwartzman, 2014; Channing & Sinanovic, 2014; Tseng et al., 2011; ReVelle et al., 1967). Nesse trabalho, o MBI em conjunto com a teoria de redes complexas foram utilizados na modelagem da propagação da tuberculose.

6.1 Modelo matemático para a tuberculose

Na modelagem da propagação da tuberculose, uma dada população pode ser dividida em três classes compartimentais, ou seja, indivíduos que nunca tiveram contato com o bacilo, denominados suscetíveis (S); indivíduos que possuem o bacilo mas não desenvolvem a doença, denominados expostos ou latentes (E) e, por fim, indivíduos que desenvolvem a tuberculose pulmonar, responsáveis pela transmissão da doença, denominados infectado (I).

A Figura 24 apresenta um fluxograma de interação entre os indivíduos das classes S , E e I . Os parâmetros μ , β , δ e ρ representam as taxas de mortalidade, infecção, infecção exógena e recaída, respectivamente. Observa-se que uma parcela q dos indivíduos suscetíveis que contraem o agente bacilo, não desenvolvem a tuberculose em um período γ^{-1} após a infecção, enquanto $(1 - q)$ desenvolvem a tuberculose pulmonar.

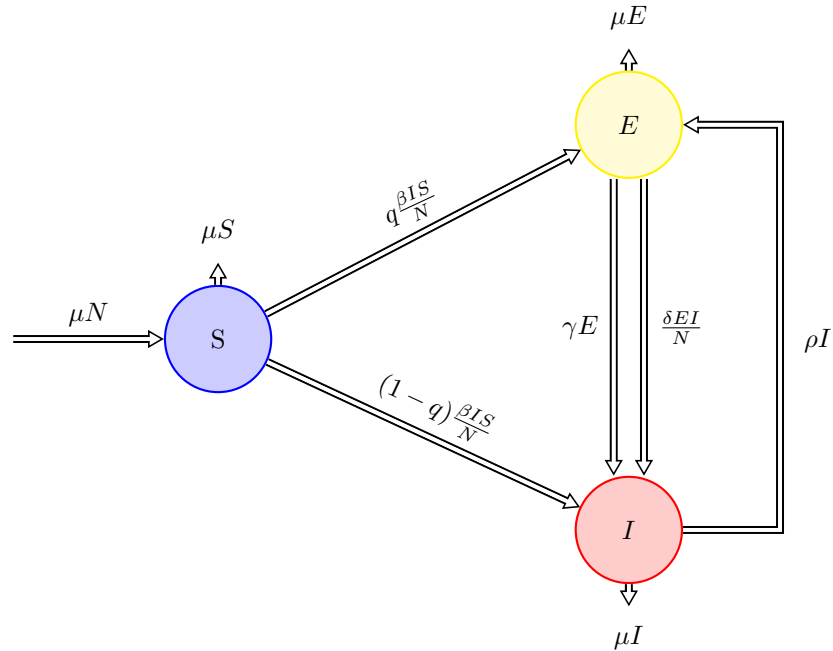


Figura 24 - Representação esquemática de um modelo SEI , constituído de indivíduos suscetíveis, expostos e infectados.

A partir do fluxograma apresentado na Figura 24, pode-se construir um sistema de equações diferenciais ordinárias para a modelagem da tuberculose (Moreno et al., 2017), dado por:

$$\begin{cases} \frac{dS}{dt} = \mu N - \frac{\beta I S}{N} - \mu S \\ \frac{dE}{dt} = \rho I + q \frac{\beta I S}{N} - \left(\mu + \gamma + \frac{\delta I}{N} \right) E \\ \frac{dI}{dt} = (1 - q) \frac{\beta I S}{N} + \left(\gamma + \frac{\delta I}{N} \right) E - (\rho + \mu) I \end{cases} \quad (45)$$

A análise da estabilidade do Sistema (45) pode ser encontrada no Apêndice A.

6.2 Comparação entre o MBI e o modelo SEI

Analogamente ao modelo SIR, ao se levar em conta que o MBI é discreto e com base nos parâmetros do modelo SEI, pode-se atribuir as seguintes relações:

$$\begin{cases} \mu_{mbi} = \frac{1}{\mu_{sei}}, \\ \beta_{mbi} = \beta_{sei} \Delta_t, \\ \gamma_{mbi} = \frac{1}{\gamma_{sei}}, \\ \rho_{mbi} = \frac{1}{\rho_{sei}}, \\ \delta_{mbi} = \delta_{sei} \Delta_t. \end{cases} \quad (46)$$

Nesse trabalho, além das características C_1 , C_2 , C_3 , C_4 e C_5 descritas na Seção 3.1, utiliza-se duas características adicionais para a modelagem da tuberculose com o MBI. São elas:

- C_6 é o tempo em anos que um indivíduo se encontra no estado de latência. Em $t = 0$, esse tempo é zero, ou seja $C_6 = 0$.

- C_7 é o tempo máximo em que um indivíduo permanece no estado de latência. No momento que o indivíduo entra no estado de latência esse tempo é obtido por:

$$C_7 = -\rho \log(a_u), \quad (47)$$

onde ρ é o período de latência.

A Figura 25 mostra o fluxograma de um MBI. A população inicial é determinada de modo aleatório. Em cada instante de tempo, cada indivíduo é considerado e verifica-se, por meio de distribuições probabilísticas, qual transição ocorrerá. A infecção ocorre de forma aleatória, ou seja, caso um indivíduo esteja infectado, o mesmo poderá infectar aleatoriamente qualquer outro indivíduo da população. Após os N indivíduos serem avaliados, o tempo de simulação é incrementado em Δt . O algoritmo termina quando o tempo de simulação atinge o valor final t_f .

Analogamente ao modelo SIR, existe uma equivalência entre o modelo SEI e o MBI, e consequentemente, as soluções dos mesmos podem apresentar comportamentos similares. O Apêndice D apresenta uma descrição detalhada da equivalência entre ambos os modelos. A Figura 26 ilustra a equivalência entre as soluções dos modelos SEI e MBI. Mais especificamente, a Figura 26(a) apresenta a solução do modelo SEI (linhas em preto) para $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$ e $\beta = 0,07$. São ainda apresentadas as soluções do MBI para $\gamma = 1/0,0027$, $\rho = 1/0,0083$, $\mu = 70$, $\delta = 0,00026$, $\beta = 0,007$ e $\Delta t = 0,1$, obtidas sobre 10 diferentes realizações. Em ambos os modelos foram considerados $N = 10.000$, $S(0) = 9200$, $E(0) = 200$ e $I(0) = 600$. Pode-se observar ainda que, para todas as realizações efetuadas, as soluções do MBI se aproximam da solução do modelo SEI. A proximidade entre as soluções de tais modelos ocorre independentemente do conjunto de parâmetros utilizado, conforme mostra a Figura 26(b). Nesse caso, foram considerados para o modelo SEI os valores $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,13$ e, para o MBI, os valores $\gamma = 1/0,0027$, $\rho = 1/0,0083$, $\mu = 70$, $\delta = 0,00026$, $\beta = 0,13$ e $\Delta t = 0,1$. Em ambos os modelos foram considerados $N = 10.000$, $S(0) = 9200$, $E(0) = 200$ e $I(0) = 600$.

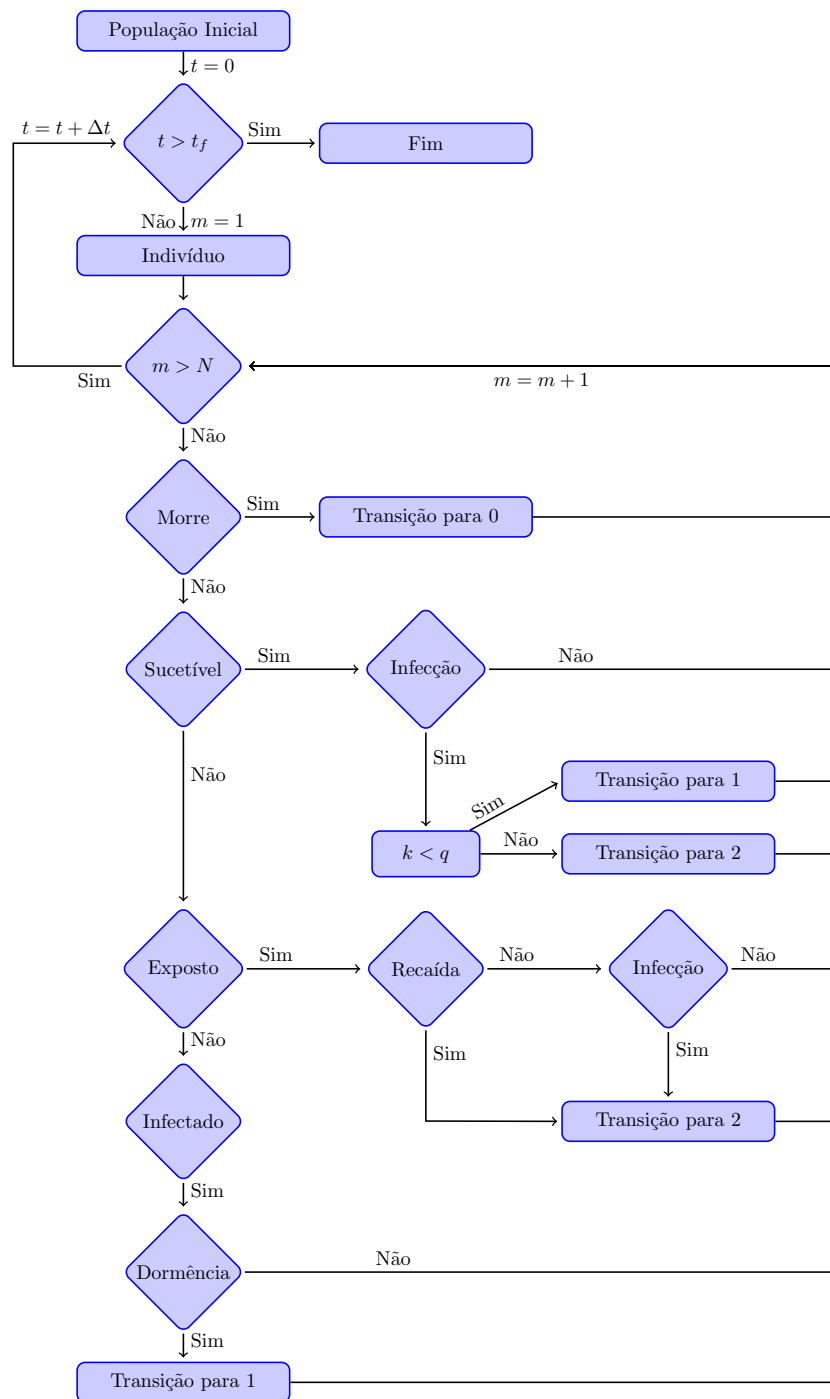


Figura 25 - Fluxograma do MBI para o modelo SEI. Os estados 0, 1 e 2 correspondem aos indivíduos suscetíveis, expostos e infectados respectivamente.

Observou-se, ainda, que as médias amostrais dos indivíduos infectados da Figura 26(a), $\bar{x}_{mbi} = 263,8$ e $\bar{x}_{sei} = 262,2$ e da Figura 26(b), $\bar{x}_{mbi} = 635,6$ e $\bar{x}_{sei} = 633,7$, não possuem diferenças significativas. A fim de quantificar com maior precisão tais diferenças, calculou-se o intervalo de confiança para a comparação das suas médias populacionais, ϕ_{mbi} e ϕ_{sei} , ao nível de significância de 5%, $(IC_{\phi_{mbi}-\phi_{sei}})$. Supondo-se que a hipótese nula seja a igualdade das médias dos dois modelos e, conseqüentemente, a alternativa como a diferença entre elas, através do resultado do teste t, determinou-se o valor- p associado a esta estatística. A Tabela 4 exhibe os resultados obtidos da comparação das soluções produzidas pelos modelos SEI e MBI. Observa-se não há diferenças estatisticamente significativas entre as médias populacionais dos dois modelos, o que implica na não rejeição da hipótese nula, o que pode ser comprovado pelo valor- p , sendo valor- $p \in [0, 1]$. Da tabela observa-se que esta estatística aponta mais evidências a favor da hipótese de nulidade, corroborando com o resultado do IC. Portanto, à 95% de confiança $\phi_{mbi} = \phi_{sei}$.

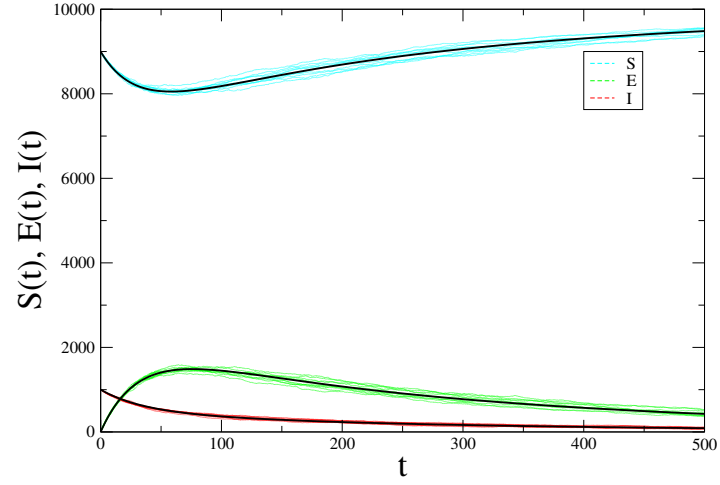
Tabela 4: Resultados obtidos pelo teste t-Student da comparação das soluções produzidas pelos modelos SEI e MBI.

	$IC(\phi_{mbi} - \phi_{sei}; 0, 95)$	valor- p
Figura 26(a)	$(-8,6312; 18,8987)$	0,4647
Figura 26(b)	$(-15,4966; 19,9541)$	0,8053

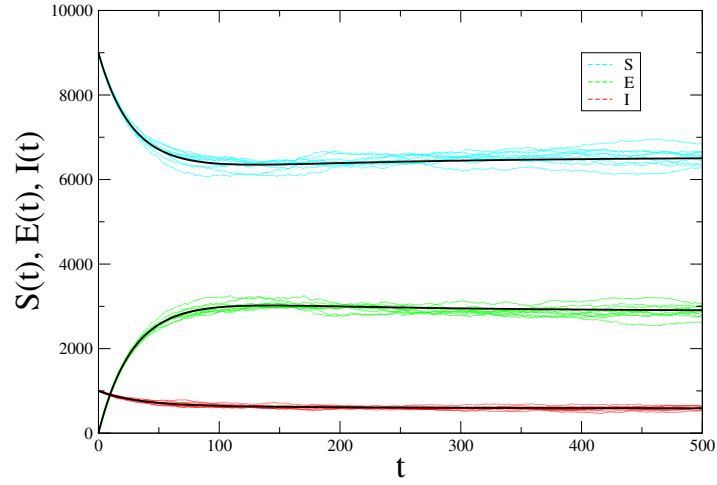
6.3 O SEI e redes complexas

Analogamente ao estudo desenvolvido no Capítulo 5, utilizou-se o MBI em conjunto com diversas redes de contato, a saber, redes aleatória, mundo pequeno, livre de escala, modular e hierárquica, na análise do efeito da topologia das mesmas nas soluções produzidas pelo MBI e, como tais, se desviam da solução produzida pelo modelo SEI.

A Figura 27 mostra o fluxograma de um MBI com redes complexas para o modelo SEI. Analogamente ao fluxograma da Figura 14, a população inicial



(a)



(b)

Figura 26 - Número de indivíduos suscetíveis, expostos e infectados em função do tempo para os modelos SEI e MBI. Em (a) foram utilizados os parâmetros $N = 10.000$, $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,07$, $S(0) = 9000$, $E(0) = 0$, $I(0) = 1000$ e $\mathcal{R}_0 \approx 0,8$. Em (b) foram utilizados os parâmetros $N = 10.000$, $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,13$, $S(0) = 9000$, $E(0) = 0$, $I(0) = 1000$ e $\mathcal{R}_0 \approx 1,48$. As linhas pretas representam a solução do modelo SEI e as demais representam o conjunto de dez soluções do MBI.

é determinada de modo aleatório. Em cada instante de tempo, cada indivíduo é considerado e verifica-se, por meio de distribuições probabilísticas, qual transição ocorrerá. A infecção não ocorre de forma aleatória, ou seja, caso um indivíduo esteja infectado, o mesmo só poderá infectar indivíduos conectados a ele. Após os N indivíduos serem avaliados, o tempo de simulação é incrementado em Δt . O algoritmo termina quando o tempo de simulação atinge o valor final t_f . Desta forma, obteve-se a solução produzida pelo MBI para uma dada rede complexa e, em seguida, estimou-se os parâmetros μ , β , γ , ρ e δ (Eq. (45)) que melhor adequavam a solução do modelo SEI à solução do MBI correspondente.

Para cada tipo de rede em estudo, foram obtidas as curvas médias de indivíduos infectados e suscetíveis, com base em 100 diferentes realizações do MBI correspondente. Os parâmetros utilizados em cada realização foram $N = 2000$, $S_0 = 1760$, $E_0 = 60$, $I_0 = 180$, $q = 0,9$, $\mu_{mbi} = 70$, $\beta_{mbi} = 0,13$, $\rho_{mbi} = 1/0,0083$, $\gamma_{mbi} = 1/0,0027$ e $\delta_{mbi} = 0,0026$. Com base no método de estimação de parâmetros descrito no Capítulo 5, foram estimados os valores μ , β , ρ , γ e δ (Tabela 5), de tal forma que a solução produzida pelo modelo SEI se tornasse a mais próxima possível à da solução do MBI correspondente.

A Figura 28 ilustra as curvas médias de indivíduos suscetíveis e infectados, solução do MBI sobre cada uma das redes em estudo, e a solução do modelo SEI correspondente. Conforme esperado, a solução do MBI com uma rede aleatória e do modelo SEI correspondente coincidem. Contudo, para as demais redes, há uma diferença entre as soluções dos modelos MBI e SEI. Isso se deve ao fato das topologias dessas redes se desviarem da topologia de uma rede puramente aleatória.

As Figuras 29, 30 e 31 ilustram as soluções do modelo MBI, mais especificamente, os números de indivíduos suscetíveis, expostos e infectados para os cinco modelos de redes complexas em estudo. Os parâmetros utilizados foram $N = 2000$, $S_0 = 1990$, $E_0 = 0$ e $I_0 = 10$, $q = 0,9$, $\mu_{mbi} = 70$, $\beta_{mbi} = 0,13$, $\rho_{mbi} = 1/0,0083$, $\gamma_{mbi} = 1/0,0027$ e $\delta_{mbi} = 0,0026$. Analogamente aos resultados apresentados no Capítulo 5, observa-se, de modo geral, que redes complexas com diferentes topologias

resultam em curvas de indivíduos suscetíveis, expostos e infectados com diferentes comportamentos. Desta forma, a evolução de uma dada doença é altamente sensível à topologia de rede utilizada, em particular na tuberculose.

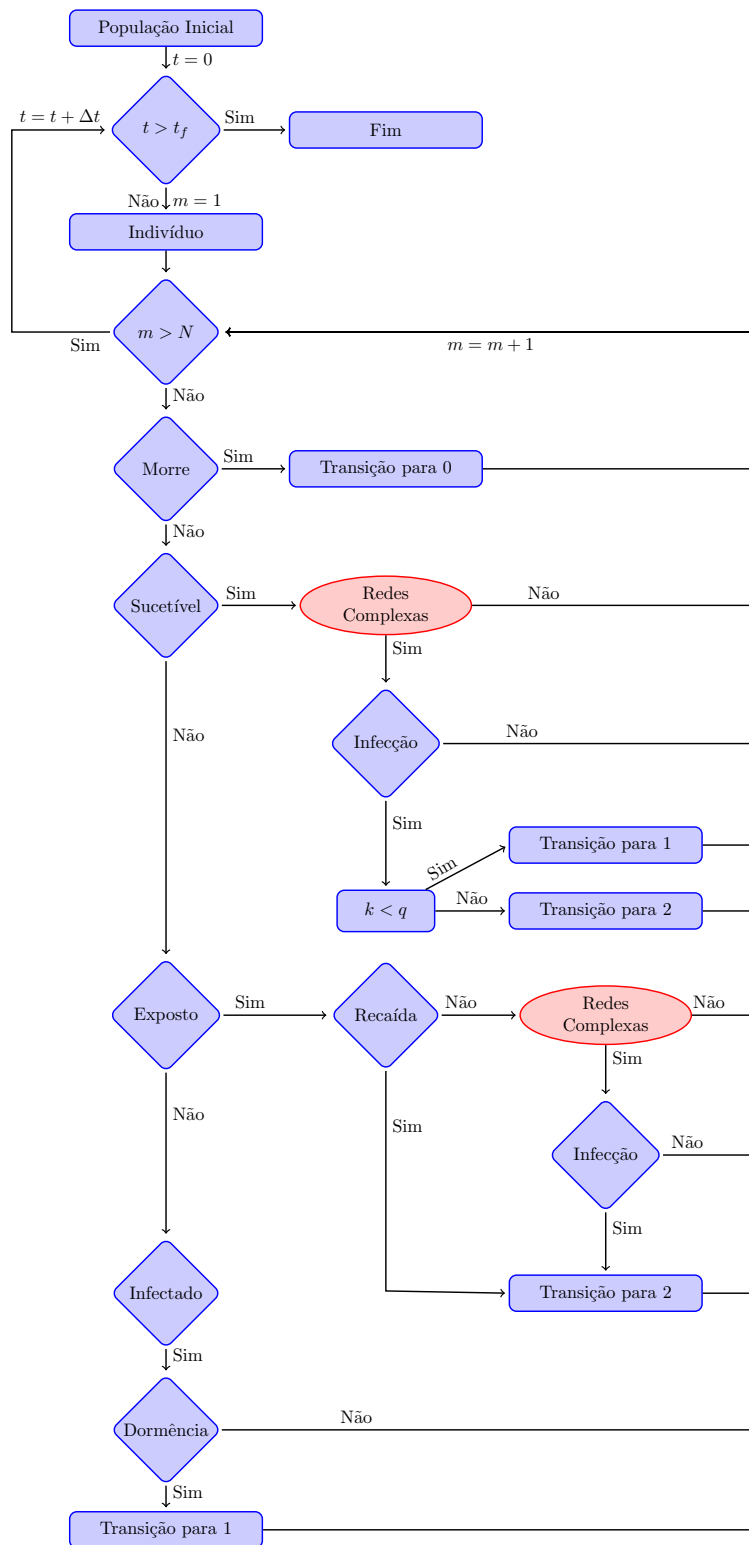


Figura 27 - Fluxograma do MBI com redes complexas para o SEI. Os estados 0, 1 e 2 correspondem os indivíduos suscetíveis, expostos e infectados, respectivamente.

Tabela 5: Parâmetros para geração das redes aleatória (ER), mundo pequeno (WS), livre de escala (BA), modular (MOD) e hierárquica (HIE) e suas propriedades topológicas.

	p	k	n_0	m	r	h	b	$\hat{\mu}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\gamma}$	$\hat{\delta}$	$\langle k \rangle$	$\langle L \rangle$	$\langle C \rangle$	Δ	$\mathcal{M}$
ER	0,07	-	-	-	-	-	-	0,0142	0,0958	0,0066	0,0023	5×10^{-5}	139,9	1,93	0,07	667,83	0,07
WS	0,01	70	-	-	-	-	-	0,0158	0,0937	0,0056	0,0019	0,128	140	2,628	0,723	75,68	0,717
BA	-	-	70	-	-	-	-	0,0136	0,1008	0,0093	0,0035	$1,3 \times 10^{-4}$	139,9	1,93	0,183	702,96	0,065
MOD	0,07	-	-	10	0,95	-	-	0,0155	0,0931	0,0056	0,0017	0,159	140,1	2,239	0,605	100,32	0,849
HIE	-	-	-	-	-	2	137931	0,0463	0,3533	0,0821	0,0203	4×10^{-8}	139,9	1,94	0,393	757,07	0,192

p probabilidade de conexão (ER e MOD), e probabilidade de reconexão (WS)

k número de vizinhos à direita

n_0 número inicial de vértices da rede

m número de módulos

r razão de ligações dentro e fora do módulo

h fator hierárquico

b número de ligações adicionais na rede

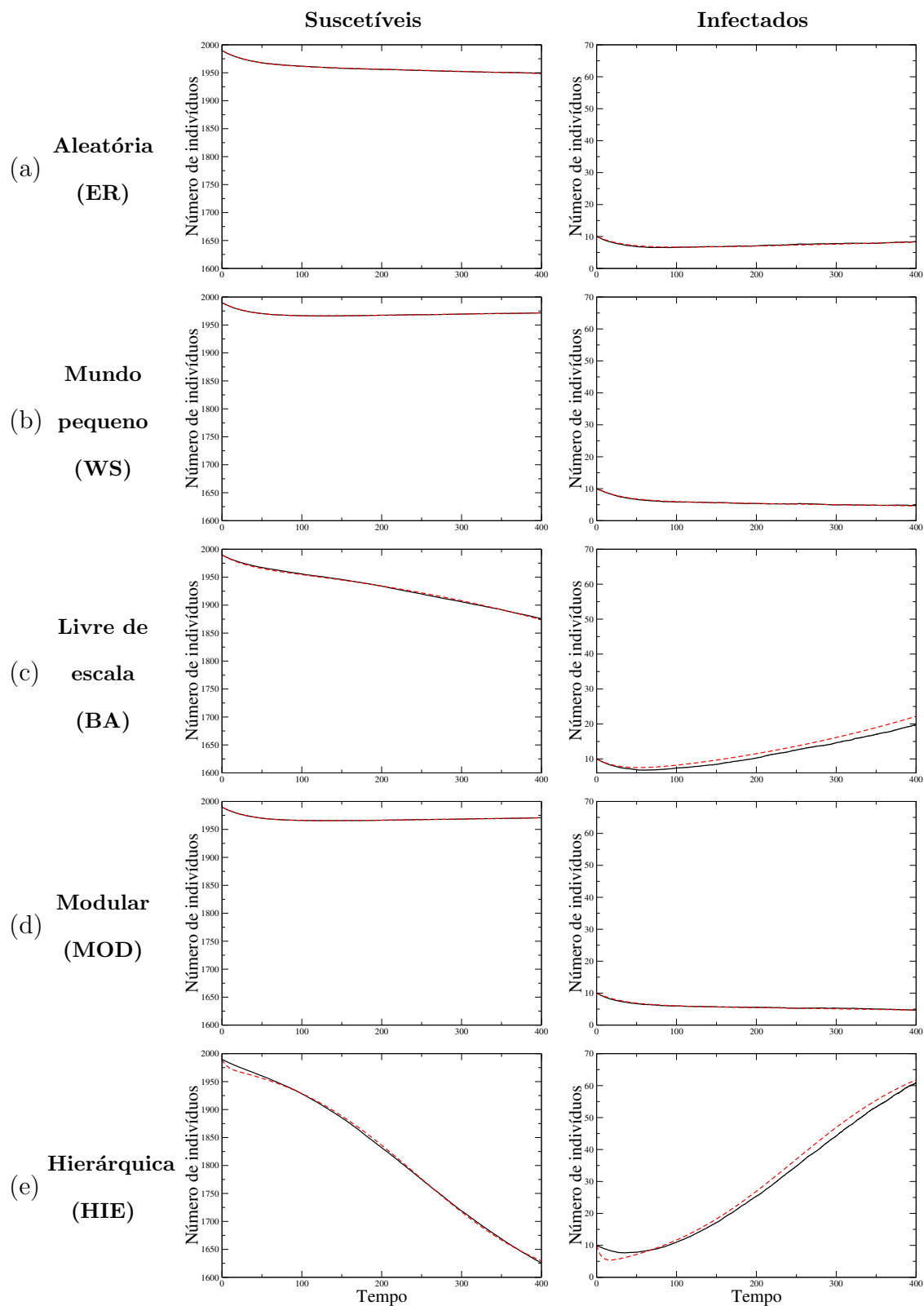


Figura 28 - Comparação do modelo MBI com redes complexas (linha contínua) e o modelo SEI (linha tracejada).

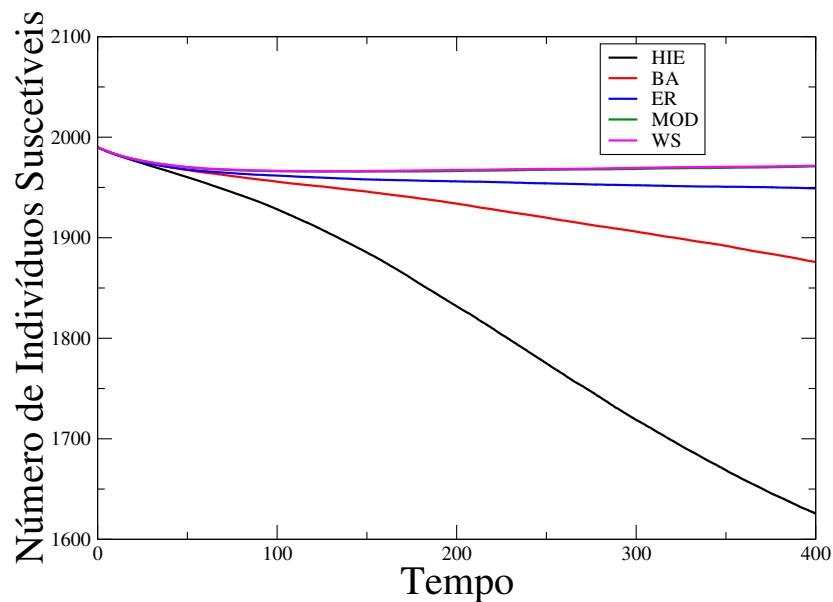


Figura 29 - Comparação das curvas de indivíduos suscetíveis, onde cada curva é a média aritmética de 100 simulações do MBI com os modelos de redes complexas em estudo.

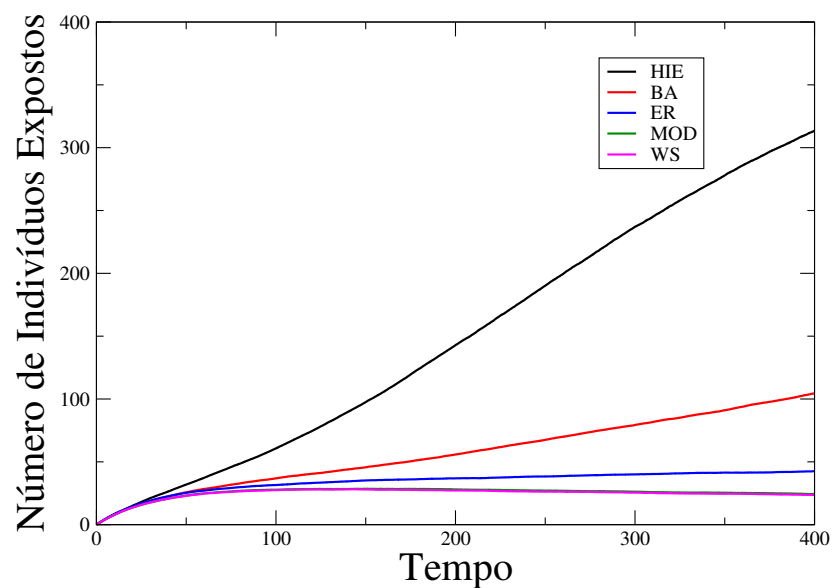


Figura 30 - Comparação das curvas de indivíduos expostos, onde cada curva é a média aritmética de 100 simulações do MBI com os modelos de redes complexas em estudo.

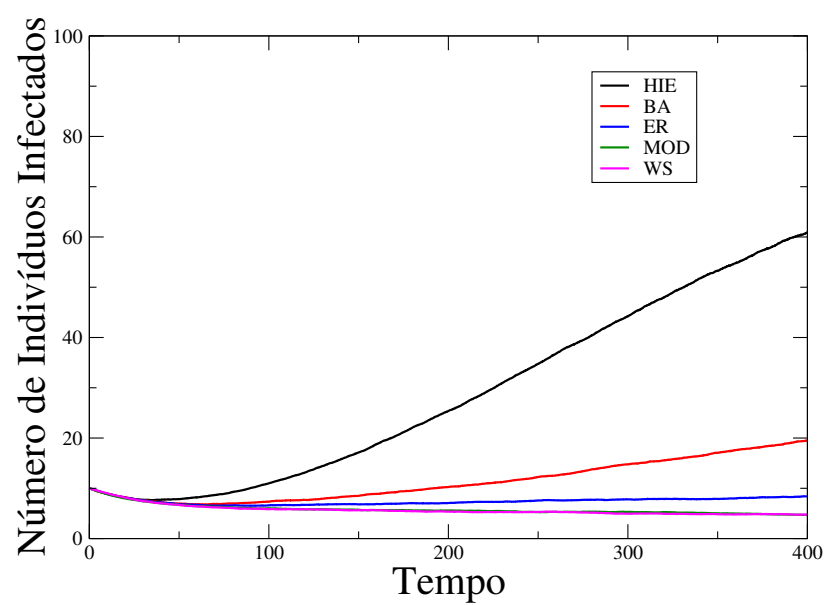


Figura 31 - Comparação das curvas de indivíduos infectados, onde cada curva é a média aritmética de 100 simulações do MBI com os modelos de redes complexas em estudo.

7 CONCLUSÕES

Nesse trabalho, foi implementado computacionalmente o MBI com base nos modelos SIR e SEI, descritos nos Capítulos 4 e 6. Como o modelo SIR e o MBI (ou analogamente, o modelo SEI e o MBI) podem utilizar as mesmas premissas, foi então realizada uma comparação entre as soluções produzidas por tais modelos, e observou-se a existência de uma equivalência entre os mesmos. Contudo, na formulação original do SIR e do MBI (ou, analogamente, do SEI e do MBI), as relações entre os indivíduos são representadas por grafos completos, e como a topologia de uma rede de contatos real não pode ser descrita por tais grafos, utilizou-se o MBI em conjunto com diversos modelos realísticos de redes complexas para a modelagem da propagação de uma doença infecciosa. Tais modelos compreendem redes aleatória, mundo pequeno, livre de escala, modular e hierárquica. Desta forma, foi feita uma análise do efeito da topologia de tais redes nas soluções produzidas pelo MBI e, como tais, se desviavam das soluções produzidas pelos modelos SIR e SEI.

Conclui-se que a evolução de uma dada doença, em particular a tuberculose, é altamente sensível à topologia de rede utilizada. Mais especificamente, observou-se que quanto maior o valor do comprimento do salto médio, mais rápida será a propagação da doença, e conseqüentemente, maior será o número de indivíduos infectados.

A pesquisa desenvolvida nesse trabalho mostra que o uso do MBI em conjunto com a teoria de redes complexas pode ser uma ferramenta simples e eficaz na modelagem de uma doença infecciosa, em particular, da tuberculose. Futuramente, deseja-se validar o modelo proposto verificando se as soluções numéricas produzidas pelo mesmo condizem com a dinâmica real da doença em estudo. Isso será

feito através da comparação dos dados produzidos por tal modelo com dados reais, disponíveis gratuitamente na literatura (WHO, 2016).

REFERÊNCIAS BIBLIOGRÁFICAS

ALBERT, R.; BARABÁSI, A. L. Statistical Mechanics of Complex Networks. **Reviews of Modern Physics**, v.74, 2002.

ALMEIDA, A. C. L. Modelos Matemáticos de Propagação de Epidemias Baseados em Redes Sociais e Detecção de Clusters de Doenças. Belo Horizonte, 2011. 113p. Tese (Doutorado) - Universidade Federal de Minas Gerais.

ALVARENGA, L. R. Modelagem de epidemias através de modelos baseados em indivíduos. Belo Horizonte, 2008. 106p. Tese (Doutorado) - Universidade Federal de Minas Gerais.

ANDERSON, R. M.; MAY, R. M. **Infectious Diseases of Humans: Dynamics and Control**. New York: Oxford University Press, 1991. 757p.

BARABÁSI, A. L. **Linked - A Nova Ciência das Redes**. São Paulo: Leopardo Editora, 2009. 256p.

BARABÁSI, A. L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v.289, 1999.

BATAGELJ, V.; MRVAR, A. PAJEK: program for large network analysis. <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>, 1998. [Online; acesso em 15 de outubro de 2017].

BHUNU, C. P.; GARIRA, W.; MUKANDAVIRE, Z.; MAGOMBEDZE, G. Modeling the effects of pre-exposure and post-exposure vaccines in tuberculosis control. **Theoret Biolo**, v.254, 2008.

BURDEN, R. L.; FAIRES, J. D. **Numerical Analysis**. Boston: Brooks Cole, 2010. 888p.

CAMARGO, A. L. Mapeamento de Séries Financeiras em Redes Complexas. Santo André, 2016. 82p. Dissertação (Mestrado) - Universidade Federal do ABC.

CAMPANHARO, A. S. L. O. Dualidade entre a análise de séries temporais e de redes complexas. São José dos Campos, 2011. 102p. Tese (Doutorado) - Instituto Nacional de Pesquisas Espaciais - INPE.

CAMPANHARO, A. S. L. O.; RAMOS, F. M. Hurst exponent estimation of self-affine time series using quantile graphs. **Physica A**, v.444, 2016.

CASTILLO-CHAVEZ, C.; FENG, Z. Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. **Math Biosci**, v.151, 1998.

CHANNING, L.; SINANOVIC, E. Modelling the cost-effectiveness of a new infant vaccine to prevent tuberculosis disease in children in South Africa. **Cost Effect Resource Allocation**, v.12, 2014.

CHAPRA, S. C.; CANALE, R. P. **Métodos Numéricos Para Engenharia**. São Paulo: McGraw-Hill, 2008. 832p.

Code::Blocks, 2008. The IDE with all the features you need, having a consistent look, feel and operation across platforms. <http://www.codeblocks.org/>, 2008. [Online; acesso em 15 de outubro de 2017].

COHEN, T.; COLIJN, C.; MURRAY, M. Modeling the effects of strain diversity and mechanisms of strain competition on the potential performance of new tuberculosis vaccines. **Proc Natl Acad Sci U S A**, v.105, 2008.

COSTA, L. F.; RODRIGUES, F. A.; TRAVIESO, G.; VILLAS, P. R. Characterization of complex networks: A survey of measurements. **Advances in Physics**, v.56, 2007.

COX, N. J.; TAMBLYN, S. E.; TAM, T. Influenza pandemic planning. **Vaccine** **21**, v.16, 2003.

DITKOWSKY, J. B.; SCHWARTZMAN, K. Potential cost-effectiveness of a new infant tuberculosis vaccine in South Africa-implications for clinical trials: a decision analysis. **PloS ONE**, v.9, 2014.

DODDS, P. S.; WATTS, D. J.; SABEL, C. F. Information exchange and the robustness of organizational networks. **PNAS**, v.100, 2003.

DOS SANTOS BENEDITO, A.; PIO DOS SANTOS, F. L. . In: ROJAS, I.; JOYA, G.; CATALA, A. (Ed.). **A Novel Technique to Estimate Biological Parameters in an Epidemiology Problem**. Springer, Cham, 2017. p.112–122.

DRAPER, N. R.; SMITH, H. **Applied Regression Analysis**. New York: John Wiley & Sons, 1998. 697p.

ERDŐS, P.; RÉNYI, A. On random graphs. **Publicationes Mathematicae**, v.6, 1959.

FAGIOLO, G. Clustering in complex directed networks. **Physical Review E**, v.76, 2007.

FLOYD, R. W. Algorithm 97: Shortest Path. **Commun. ACM**, v.5, 1962.

FOURNET, J.; BARRAT, A. Epidemic risk from friendship network data: an equivalence with a non-uniform sampling of contact networks. **Nature**, v.6, 2016.

GOMES, P. D. Um modelo matemático para a epidemia de tuberculose. Niterói, 2004. 70p. Dissertação (Mestrado) - Universidade Federal Fluminense.

GUIMERÁ, R.; SALES-PARDO, M.; AMARAL, L. Modularity from fluctuations in random graphs and complex networks. **Physical Review E**, v.70, 2004.

HARRIS, R. C.; SUMNER, T.; KNIGHT, G. M.; WHITE, R. G. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. **Hum Vaccin Immunother**, v.12, 2016.

HETHCOTE, H. W. The mathematics of infectious diseases. **SIAM review** **42**, v.4, 2000.

KEELING, M. J.; GRENFELL, B. T. Individual-based Perspectives on R_0 . **Journal of Theoretical Biology**, v.203, 2000.

KERMACK, W. O.; MCKENDRICK, A. G. A Contribution to the Mathematical Theory of Epidemics. **Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character**, v.115, 1927.

LILJEROS, F.; EDLING, C. R.; AMARAL, L. A. N.; STANLEY, H. E.; ABERG, Y. The web of human sexual contacts. **Nature**, v.411, 2001.

MARCOLINO, R. S. Modelos matemáticos epidemiológicos da tuberculose em Codó-MA. Campinas, 2016. 107p. Dissertação (Mestrado) - UNICAMP.

MARQUARDT, D. An Algorithm for Least-squares Estimation of Nonlinear Parameters. **SIAM Journal Applied Mathematics**, v.11, 1963.

MILGRAN, S. The small world problem. **Psychology Today**, v.1, 1967.

MORENO, V.; ESPINOZA, B.; PAREDES, M.; BICHARA, D.; A., M.; CASTILLOCHAVEZ, C. The role of mobility and health disparities on the transmission dynamics of Tuberculosis. **Theoretical Biology and Medical Modeling**, v.143, 2017.

NEPOMUCENO, E. G. Dinâmica, Modelagem e Controle de Epidemias. Belo Horizonte, 2005. 167p. Tese (Doutorado) - Universidade Federal de Minas Gerais.

NEWMAN, M. E. J. The spread of epidemic disease on networks. **Physical Review Letters**, v.66, 2002.

NEWMAN, M. E. J. Structure and function of complex networks. **SIAM Review**, v.45, 2003.

NEWMAN, M. E. J. **Networks: An Introduction**. New York: Oxford University Press, 2010. 720p.

NOCEDAL, J.; WRIGHT, S. T. **Numerical Optimization**. New York: Springer Verlag, 2006. 664p.

PAPOULIS, A. **Probability, random variables, and stochastic processes**. New York: McGraw-Hill, 1991. 661p.

RAVASZ, E.; BARABÁSI, A. L. Hierarchical organization in complex networks. **Physical Review E**, v.67, 2003.

REVELLE, C. S.; LYNN, W. R.; FELDMANN, F. Mathematical models for the economic allocation of tuberculosis control activities in developing nations. **Am Rev Respirat Dis**, v.96, 1967.

RODRIGUES, F. A. Caracterização, classificação e análise de redes complexas. São Carlos, 2007. 175p. Tese (Doutorado) - Universidade de São Paulo, Instituto de Física de São Carlos.

SOLE, R. V.; GAMARRA, J. G. P.; GINOVART, M.; LOPEZ, D. Controlling chaos in ecology: from deterministic to individual-based models. **Bulletin of Mathematical Biology**, v.61, 1999.

TSENG, C. L.; OXLADE, O.; MENZIES, D.; ASPLER, A.; SCHWARTZMAN, K. Cost-effectiveness of novel vaccines for tuberculosis control: a decision analysis study. **BMC Public Health**, v.11, 2011.

TURNER, P. Grace. <http://plasma-gate.weizmann.ac.il/Grace/>, 1995. [Online; acesso em 15 de outubro de 2017].

WATTS, D. J.; STROGATZ, S. H. Collective Dynamics of “Small World” Networks. **Letters to Nature**, v.393, 1998.

WHO. **Global Tuberculosis Report 2016**. World Health Organization, 2016. Global Tuberculosis Control.

APÊNDICES

Apêndice A

Análise de estabilidade

Uma importante medida na dinâmica de epidemias é o número básico de reprodução, representado por $\mathcal{R}_0$. De acordo com Anderson & May (1991), o $\mathcal{R}_0$ é definido como o número médio de infecções secundárias produzidas por um indivíduo infectado introduzido em uma população totalmente suscetível (Hethcote, 2000).

Esse conceito é central em qualquer discussão sobre dinâmica populacional. O $\mathcal{R}_0$ é um número que determina se a população estará livre da doença ou se a doença permanecerá. Desta forma, se $\mathcal{R}_0 > 1$ um indivíduo infectado consegue infectar em média mais de um indivíduo, tendendo assim a um equilíbrio endêmico (Hethcote, 2000). Caso contrário, ou seja, $\mathcal{R}_0 < 1$, a população estará livre da doença e, desta forma, um indivíduo infectado não consegue propagar a doença.

A Figura 32 mostra, esquematicamente, o funcionamento da dinâmica do $\mathcal{R}_0$. Cada círculo representa um indivíduo infectado, as flechas indicam quando um indivíduo infectado infecta um indivíduo suscetível. Neste caso, cada indivíduo infectado produz em média duas infecções secundárias, já que $\mathcal{R}_0 = 2$.

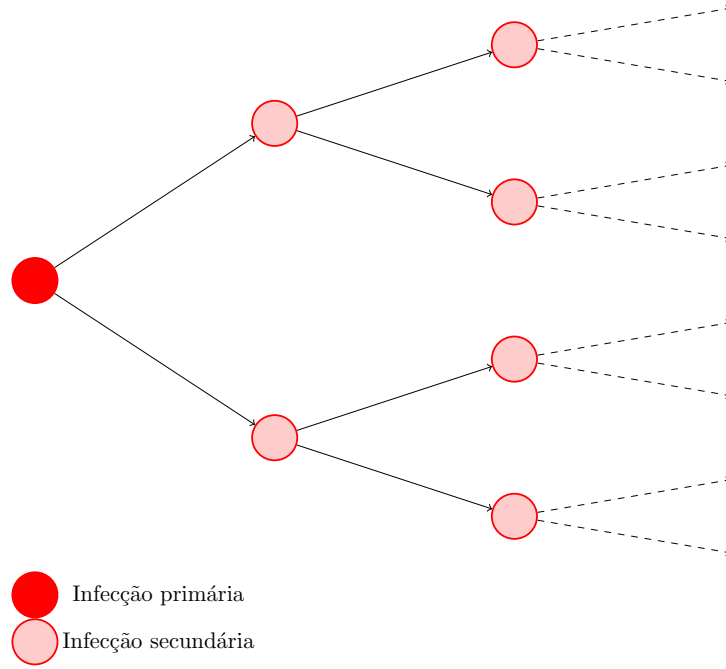


Figura 32 - O número básico de reprodução $\mathcal{R}_0 = 2$: um indivíduo infectado introduzido em uma população de indivíduos suscetíveis.

Modelo SIR

Em modelos epidêmicos compartimentais, o cálculo do $\mathcal{R}_0$ pode ser obtido quando $\frac{dI}{dt} > 0$. No sistema dado pela Equação (41) a taxa básica de reprodução é definida por

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \mu}. \quad (48)$$

Além disso, pode-se calcular os pontos de estabilidade, quando igualamos o sistema dinâmico à zero, ou seja

$$\frac{dS}{dt} = 0, \quad \frac{dI}{dt} = 0, \quad \frac{dR}{dt} = 0. \quad (49)$$

Resolvendo o sistema, obtem-se os pontos de equilíbrio endêmico, dados por:

$$S^* = \frac{N}{\mathcal{R}_0}, \quad I^* = \frac{\mu(\mathcal{R}_0 - 1)N}{\beta}, \quad R^* = \left(1 - \frac{1}{\mathcal{R}_0} - \frac{\mu(\mathcal{R}_0 - 1)}{\beta}\right)N. \quad (50)$$

Se $\mathcal{R}_0 < 1$ ou $I_0 = 0$, as trajetórias das soluções atingirão a situação de erradicação da doença, ou seja, $S = N$, $I = 0$ e $R = 0$. Se $\mathcal{R}_0 > 1$, as trajetórias das soluções com $I_0 > 0$ atingirão um equilíbrio endêmico.

Na Figura 13(a) foram utilizados os parâmetros $N = 1.000$, $\mu = \frac{1}{70}$, $\beta = 0,5$, $\gamma = \frac{1}{36}$, $S(0) = 700$, $I(0) = 300$ e $R(0) = 0$. Nesse caso, $\mathcal{R}_0 \approx 11,89 > 1$, ou seja, a doença permanecerá na população e convergirá para um equilíbrio endêmico dado por $S^* \approx 84,13$, $E^* \approx 311,05$ e $I^* \approx 604,82$.

Na Figura 13(b) foram utilizados os parâmetros $N = 1.000$, $\mu = \frac{1}{70}$, $\beta = 0,05$, $\gamma = \frac{1}{24}$, $S(0) = 700$, $I(0) = 300$ e $R(0) = 0$. Nesse caso, $\mathcal{R}_0 \approx 0,89 < 1$, ou seja, a doença não consegue permanecer na população. Logo, a convergência do ponto de equilíbrio livre da doença $S^* = N$, $I^* = 0$ e $R^* = 0$ é visível no decorrer do tempo.

Modelo SEI

Analogamente ao modelo SIR, o valor do $\mathcal{R}_0$ para o modelo SEI Equação (45), é dado por

$$\mathcal{R}_0 = \frac{\beta(\gamma + \mu(1 - q))}{\mu(\rho + \gamma + \mu)}. \quad (51)$$

E seus pontos de equilíbrio endêmico são:

$$S^* = \frac{\mu N^2}{\beta I^* + \mu N}, \quad E^* = \frac{\rho I^* N + q \beta I^* S^*}{N(\mu + \gamma) + \delta I^*}, \quad I^* = \left(\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \right) N, \quad (52)$$

onde,

$$\begin{cases} a = -\delta\beta, \\ b = \beta\delta - \mu\delta - \beta(\rho + \mu + \gamma), \\ c = \beta(\gamma + \mu(1 - q) - \mu(\rho + \mu + \gamma)). \end{cases} \quad (53)$$

Se $\mathcal{R}_0 < 1$ ou $E_0 = 0$ e $I_0 = 0$, as trajetórias das soluções atingirão a situação de erradicação da doença, ou seja, $S = N$, $E = 0$ e $I = 0$. Se $\mathcal{R}_0 > 1$, as trajetórias das soluções com $I_0 > 0$ atingirão um equilíbrio endêmico.

Na Figura 26(a) foram utilizados os parâmetros $N = 10.000$, $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,07$, $S(0) = 9000$, $E(0) = 0$ e

$I(0) = 1000$. Nesse caso, $\mathcal{R}_0 \approx 0,8 < 1$, ou seja, a doença não consegue permanecer na população. Logo, a convergência do ponto de equilíbrio livre da doença $S^* = N$, $E^* = 0$ e $I^* = 0$ é visível no decorrer do tempo

Na Figura 26(b) foram utilizados os parâmetros $N = 10.000$, $q = 0,9$, $\gamma = 0,0027$, $\rho = 0,0083$, $\mu = 1/70$, $\delta = 0,0026$, $\beta = 0,13$, $S(0) = 9000$, $E(0) = 0$ e $I(0) = 1000$. Nesse caso, $\mathcal{R}_0 \approx 1,48 > 1$, ou seja, a doença permanecerá na população e convergirá para um equilíbrio endêmico dado por $S^* \approx 6530,6$, $E^* \approx 2885,65$ e $I^* \approx 583,8$.

Apêndice B

Método de Gauss-Newton

O método de Gauss-Newton é um método iterativo de otimização utilizado para resolver problemas de mínimos quadrados não lineares. Tal metodologia é uma modificação do método de Newton (Nocedal & Wright, 2006).

Para entender o processo iterativo, considere uma função objetivo y_i e a função $f(x_i; a_0, a_1, \dots, a_n)$ do ajuste do método dos Mínimos Quadrados com n parâmetros $(a_0, a_1, \dots, a_n)$. Pode-se dizer que:

$$y_i = f(x_i) + e_i \quad (54)$$

Considere a expansão da função f em série de Taylor de n variáveis $a_0, a_1, \dots, a_n$ de primeira ordem:

$$f(x_{i+1}) = f(x_i) + \frac{\partial f(x_i)}{\partial a_0} \Delta a_0 + \frac{\partial f(x_i)}{\partial a_1} \Delta a_1 + \dots + \frac{\partial f(x_i)}{\partial a_n} \Delta a_n, \quad (55)$$

onde $\Delta a_i = a_{i_{j+1}} - a_{i_j}$, $i = 1, 2, \dots, n$ são as variações dos parâmetros ao longo do processo iterativo e j representa a j -ésima iteração. Substituindo a Equação (55) na Equação (54), pode-se escrever a igualdade abaixo:

$$y_i - f(x_i) = \frac{\partial f(x_i)}{\partial a_0} \Delta a_0 + \frac{\partial f(x_i)}{\partial a_1} \Delta a_1 + \dots + \frac{\partial f(x_i)}{\partial a_n} \Delta a_n + e_i. \quad (56)$$

O método de Gauss-Newton também pode ser escrito na forma matricial, ou seja,

$$[R] = [J][\Delta A] + [\Xi], \quad (57)$$

de modo que o resíduo $R = y_i - f(x_i)$ é um vetor coluna, $[J]$ é a matriz Jacobiana da função f , $[\Delta A]$ é o vetor coluna das diferenças dos parâmetros a cada iteração

e. Por fim, $[\Xi]$ é o vetor coluna do erro cometido no processo iterativo (Chapra & Canale, 2008).

O objetivo de tal método é encontrar os parâmetros a_k , tal que $k = 0, 1, 2, \dots, n$, para serem utilizados no ajuste. Para tal, deve ser resolvida a equação matricial:

$$-[J]^T[R] = [J]^T[J][\Delta A], \quad (58)$$

ou ainda,

$$[\Delta A] = ([J]^T[J])^{-1}[-J]^T[R]. \quad (59)$$

A Equação (59) é resolvida a cada iteração a fim de encontrar os parâmetros do ajuste de curvas (Chapra & Canale, 2008). O critério de parada para o método varia de acordo com a precisão desejada, pode se utilizar o intervalo de confiança, o fator estatístico R^2 ou o cálculo do erro (Draper & Smith, 1998).

Método do Gradiente

O método do gradiente ou de máxima descida é um método numérico iterativo de otimização que objetiva direcionar a busca do ponto ótimo através do vetor gradiente. Este procedimento, mesmo com condições iniciais deficientes e uma convergência mais lenta, comparada ao método de Gauss-Newton, quase sempre converge para o ponto ótimo (Burden & Faires, 2010).

Primeiramente, deve-se encontrar a matriz Jacobiana da função objetivo e calcular os pontos críticos igualando as derivadas parciais ao zero, obtendo então $(x^*, y^*, \dots)$. Avalia-se os pontos críticos na matriz Hessiana e verifica-se se há um ponto de máximo, mínimo ou de sela na região.

Necessita-se de um ponto inicial $(x_0, y_0, \dots)$ para que as iterações sejam realizadas. Uma vez calculado, o vetor gradiente é avaliado neste ponto e a direção de busca do ponto ótimo é estabelecida e definida pela equação:

$$f\left(x_0 + \frac{\partial f}{\partial x}h, y_0 + \frac{\partial f}{\partial y}h, \dots\right), \quad (60)$$

em que h é o passo do método.

O passo seguinte é calculado substituindo $(x_1, y_1, \dots)$ por $(x_0 + \frac{\partial f}{\partial x} h, y_0 + \frac{\partial f}{\partial y} h, \dots)$ em f e verificando qual valor de h que satisfaz $f'(h) = 0$. Como h é um ponto crítico e o método avalia a máxima descida, então ele também é o ponto de máximo. O processo descrito anteriormente constitui a primeira iteração do método. Ao substituir h na Equação (60), tem-se o novo ponto $(x_1, y_1, \dots)$ que será usado na segunda iteração e, assim, sucessivamente.

Método de Levenberg-Marquardt

O método de Levenberg-Marquardt é um método de otimização utilizado para resolver problemas de mínimos quadrados não lineares (dos Santos Benedito & Pio dos Santos, 2017). Esse método é uma combinação dos métodos clássicos de otimização de *Gauss-Newton* e do *Gradiente* (de máxima descida) (Nocedal & Wright, 2006; Draper & Smith, 1998; Chapra & Canale, 2008). O método de *Gauss-Newton* tem forte convergência perto do ponto ótimo, porém como utiliza matrizes inversas, pode-se obter matrizes singulares, de forma a não convergir corretamente para as estimativas dos parâmetros. Por outro lado, o método do Gradiente tem baixa convergência comparado ao Gauss-Newton, mas possui uma notável estabilidade e dispensa inversão de matrizes.

O método de Levenberg-Marquardt insere na Equação (59) um parâmetro λ , positivo, que ao ser multiplicado pela matriz identidade, acelera a convergência do método e/ou retira a singularidade da matriz no método de Gauss-Newton (Nocedal & Wright, 2006). Assim a Equação (59) pode ser reescrita na forma:

$$[\Delta A] = -([J^T][J] + \lambda[I])^{-1}[J^T][R] \quad (61)$$

Observe que há um sinal negativo no segundo termo da equação que faz menção à máxima descida $(-\nabla f)$ e a adição do fator $\lambda[I]$ que influencia na direção e no tamanho do passo, λ promove convergência quadrática do método de Levenberg-Marquardt e o mesmo deve ser atualizado a cada iteração

Apêndice C

Validação do MBI para o modelo SIR

Nesta seção, é feita a validação do MBI para o modelo SIR. Define-se os seguintes eventos ocorridos no intervalo de tempo $(k\Delta_t, (k+1)\Delta_t)$:

- $\Delta(k)$: o evento de um indivíduo infectado (I) se recuperar, passando para a classe dos recuperados (R);
- $\Omega(k)$: o evento de um indivíduo morrer e, conseqüentemente, nascer um outro indivíduo na classe dos suscetíveis (S);
- $\Gamma(k)$: o evento de se sortear um indivíduo suscetível;
- $\Phi(k)$: o evento de um indivíduo ter um contato com um indivíduo infectado;
- $\Lambda(k)$: o evento de um indivíduo suscetível passar para a classe dos infectados, devido ao contato com um indivíduo infectado.

A probabilidade de um indivíduo se recuperar é dada por:

$$P(\Delta(k)) = \gamma\Delta_t. \quad (62)$$

A probabilidade de um indivíduo morrer e nascer um outro no seu lugar é dada por:

$$P(\Omega(k)) = \mu\Delta_t. \quad (63)$$

A probabilidade de um indivíduo ser infectado é a probabilidade de ocorrer simultaneamente os eventos Γ e Φ , ou seja:

$$P(\Lambda(k)) = P(\Gamma(k) \cap \Phi(k)). \quad (64)$$

Por serem eventos independentes, a Equação (64) pode ser reescrita como:

$$P(\Lambda(k)) = P(\Gamma(k))P(\Phi(k)). \quad (65)$$

A probabilidade de sorteio de um indivíduo suscetível é dada por:

$$P(\Gamma(k)) = \frac{S(k)}{N}. \quad (66)$$

A probabilidade de acontecer um contato com um indivíduo infectado é proporcional ao número de infectados presentes naquele momento, sendo dada por:

$$P(\Phi(k)) = \frac{\beta\Delta_t}{N-1}I(k). \quad (67)$$

Das equações (65), (66) e (67) temos que:

$$P(\Lambda(k)) = \frac{\beta\Delta_t}{N-1} \frac{S(k)I(k)}{N}. \quad (68)$$

Calcula-se agora, os números médios de mudanças de classes ocorridos. O número médio de indivíduos, por unidade de tempo, da classe I que se recuperam da infecção e passam para a classe R é calculado a partir da Equação (62) e é dado por:

$$\gamma\Delta_t I(k). \quad (69)$$

Os números médios de indivíduos, por unidade de tempo, das classes S , I e R que morrem e são substituídos por outros da classe S são calculados a partir da Equação (63), sendo dados respectivamente por

$$\mu\Delta_t S(k) \quad (70)$$

$$\mu\Delta_t I(k) \quad (71)$$

$$\mu\Delta_t R(k) \quad (72)$$

Se um experimento é realizado X vezes, X suficientemente grande, e verifica-se a ocorrência de certo evento E vezes, então, segundo (Papoulis, 1991), a probabilidade P_e deste evento ocorrer se aproxima de:

$$P_e = \frac{E}{X}. \quad (73)$$

No MBI, em cada intervalo de tempo, executa-se o algoritmo descrito na Seção 3, uma vez para cada indivíduo, ou seja, N vezes. A probabilidade da infecção ocorrer é dada por $P(\Lambda(k))$ na Equação (68). Fazendo $X = N$ e $P_e = P(\Lambda(k))$ na Equação (73), conclui-se que o número médio de indivíduos, por unidade de tempo, que passam da classe S para a classe I é dado por:

$$P(\Lambda(k)) = N \frac{\beta \Delta_t}{N-1} \frac{S(k)I(k)}{N}. \quad (74)$$

Para valores de N razoavelmente grandes pode-se considerar $N/(N-1) \approx 1$, e reescreve-se o número médio de novas infecções como:

$$P(\Lambda(k)) = \frac{\beta \Delta_t S(k)I(k)}{N}. \quad (75)$$

As Equações (69), (70), (71), (72) e (75) permitem descrever a dinâmica do MBI:

- suscetíveis: recebe o acréscimo de todos os indivíduos renovados ($\mu \Delta_t N = \mu \Delta_t S + \mu \Delta_t I + \mu \Delta_t R$), onde $S(k) + I(k) + R(k) = N$. Terá, ainda, a perda dos seus próprios indivíduos renovados ($\mu \Delta_t S$) e dos indivíduos que foram infectados ($\beta \Delta_t I(k)S(k)/N$);
- infectados: terá a perda dos indivíduos recuperados ($\gamma \Delta_t I$) e renovados ($\mu \Delta_t I$) e o acréscimo das novas infecções ($\beta \Delta_t I(k)S(k)/N$);
- recuperados: terá a perda dos indivíduos renovados ($\mu \Delta_t R$) e o acréscimo dos indivíduos que se recuperaram da infecção ($\gamma \Delta_t I$).

Matematicamente, pode-se então escrever a dinâmica associada ao comportamento médio do MBI:

$$\begin{cases} S(k+1) = S(k) + \mu \Delta_t N - \mu \Delta_t S(k) - \frac{\beta \Delta_t I(k)S(k)}{N}, \\ I(k+1) = I(k) - \mu \Delta_t I(k) - \gamma \Delta_t I(k) + \frac{\beta \Delta_t I(k)S(k)}{N} \\ R(k+1) = R(k) + \gamma \Delta_t I(k) - \mu \Delta_t R(k). \end{cases} \quad (76)$$

Lembrando que k representa o instante de tempo $k\Delta_t$ e que

$$\frac{df}{dt} = \lim_{\Delta_t \rightarrow 0} \frac{f(t + \Delta_t) - f(t)}{\Delta_t} \quad (77)$$

nota-se que a Equação (76) coincide com a equação de diferenças da Equação (41) utilizada na obtenção de soluções do modelo SIR por integração numérica. Assim, o modelo MBI converge, em média, para o modelo SIR.

Apêndice D

Validação do MBI para o modelo SEI

Nesta seção, é feita a validação do MBI para o modelo SEI. Define-se os seguintes eventos ocorridos no intervalo de tempo $(k\Delta_t, (k+1)\Delta_t)$:

- $\Delta(k)$: o evento de um indivíduo infectado (I) entrar no estágio de latência, passando para a classe dos expostos (E);
- $\Omega(k)$: o evento de um indivíduo morrer e, conseqüentemente, nascer um outro indivíduo na classe dos suscetíveis (S);
- $\Pi(k)$: o evento de um indivíduo exposto (E) ter uma recaída, passando para a classe dos infectados (I);
- $\Gamma(k)$: o evento de se sortear um indivíduo suscetível;
- $\Theta(k)$: o evento de se sortear um indivíduo exposto;
- $\Phi(k)$: o evento de um indivíduo ter um contato com um indivíduo infectado;
- $\Lambda(k)$: o evento de um indivíduo suscetível passar para a classe dos infectados, devido ao contato com um indivíduo infectado.
- $\Psi(k)$: o evento de um indivíduo exposto passar para a classe dos infectados, devido ao contato com um indivíduo infectado.

A probabilidade de um indivíduo infectado entrar no estágio de latência é dada por:

$$P(\Delta(k)) = \rho\Delta_t. \quad (78)$$

A probabilidade de ocorrer a renovação de um indivíduo é dada por:

$$P(\Omega(k)) = \mu\Delta_t. \quad (79)$$

A probabilidade de um indivíduo exposto ter uma recaída é dada por:

$$P(\Pi(k)) = \gamma\Delta_t. \quad (80)$$

A probabilidade de um indivíduo suscetível ser infectado é a probabilidade de ocorrer simultaneamente os eventos Γ e Φ , ou seja:

$$P(\Lambda(k)) = P(\Gamma(k) \cap \Phi(k)). \quad (81)$$

Por serem eventos independentes, a Equação (81) pode ser reescrita como:

$$P(\Lambda(k)) = P(\Gamma(k))P(\Phi(k)). \quad (82)$$

A probabilidade de sorteio de um indivíduo suscetível é dada por:

$$P(\Gamma(k)) = \frac{S(k)}{N}. \quad (83)$$

A probabilidade de acontecer um contato com um indivíduo infectado é proporcional ao número de infectados presentes naquele momento, sendo dada por:

$$P(\Phi(k)) = \frac{\beta\Delta_t}{N-1}I(k). \quad (84)$$

Das equações (82), (83) e (84) tem-se que:

$$P(\Lambda(k)) = \frac{\beta\Delta_t}{N-1} \frac{S(k)I(k)}{N}. \quad (85)$$

Analogamente, a probabilidade de um indivíduo exposto ser infectado é a probabilidade de ocorrer simultaneamente os eventos Θ e Φ , ou seja:

$$P(\Psi(k)) = P(\Theta(k) \cap \Phi(k)). \quad (86)$$

Por serem eventos independentes, a Equação (86) pode ser reescrita como:

$$P(\Psi(k)) = P(\Theta(k))P(\Phi(k)). \quad (87)$$

A probabilidade de sorteio de um indivíduo exposto é dada por:

$$P(\Theta(k)) = \frac{E(k)}{N}. \quad (88)$$

A probabilidade de um indivíduo exposto ser infectado é proporcional ao número de infectados presentes naquele momento, sendo dada por:

$$P(\Phi(k)) = \frac{\delta\Delta_t}{N-1} I(k). \quad (89)$$

Das equações (87), (88) e (89) tem-se que:

$$P(\Psi(k)) = \frac{\delta\Delta_t}{N-1} \frac{E(k)I(k)}{N}. \quad (90)$$

Calcula-se agora, os números médios de mudanças de classes ocorridos. O número médio de indivíduos infectados que entram no estágio de latência, por unidade de tempo, é calculado a partir da Equação (78) e é dado por:

$$\rho\Delta_t I(k). \quad (91)$$

Os números médios de indivíduos, por unidade de tempo, das classes S , E e I que morrem e são substituídos por outros da classe S são calculados a partir da Equação (79), sendo dados, respectivamente, por

$$\mu\Delta_t S(k) \quad (92)$$

$$\mu\Delta_t E(k) \quad (93)$$

$$\mu\Delta_t I(k) \quad (94)$$

O número médio de indivíduos expostos que tem recaídas, por unidade de tempo, é calculado a partir da Equação (80) e é dado por:

$$\gamma\Delta_t E(k). \quad (95)$$

No MBI, em cada intervalo de tempo, executa-se o algoritmo descrito uma vez para cada indivíduo, ou seja, N vezes. A probabilidade de um indivíduo suscetível ser infectado é dada por $P(\Lambda(k))$ na Equação (85) e a probabilidade de

um indivíduo exposto ser infectado é dada por $P(\Psi(k))$ na Equação (90). Fazendo $X = N$, $P_{e_1} = P(\Lambda(k))$ e $P_{e_2} = P(\Psi(k))$ na Equação (73), conclui-se que o número médio de indivíduos, por unidade de tempo, que passam da classe S para a classe I e que passam da classe E para a classe I é dado por:

$$P(\Lambda(k)) = N \frac{\beta \Delta_t}{N-1} \frac{S(k)I(k)}{N}, \quad (96)$$

$$P(\Psi(k)) = N \frac{\delta \Delta_t}{N-1} \frac{E(k)I(k)}{N}. \quad (97)$$

Para valores de N razoavelmente grandes, pode-se considerar $N/(N-1) \approx 1$, e reescreve-se o número médio de novas infecções como:

$$P(\Lambda(k)) = \frac{\beta \Delta_t S(k)I(k)}{N}, \quad (98)$$

$$P(\Psi(k)) = \frac{\delta \Delta_t E(k)I(k)}{N}. \quad (99)$$

As Equações (91), (92), (93), (94), (95), (98) e (99) permitem descrever a dinâmica do MBI:

- suscetíveis: recebe o acréscimo de todos os indivíduos renovados ($\mu \Delta_t N = \mu \Delta_t S + \mu \Delta_t E + \mu \Delta_t I$), onde $S(k) + E(k) + I(k) = N$. Terá, ainda, a perda dos seus próprios indivíduos renovados ($\mu \Delta_t S$) e dos indivíduos que foram infectados ($\beta \Delta_t I(k)S(k)/N$);
- expostos: terá a perda dos indivíduos renovados ($\mu \Delta_t E$) e dos indivíduos infectados (infecção natural e exógena) ($\gamma \Delta_t E(k) + \delta \Delta_t E(k)I(k)/N$) e o acréscimo dos indivíduos que não desenvolveram a tuberculose pulmonar ($q\beta \Delta_t I(k)S(k)/N$) e os indivíduos que entraram no estágio de latência ($\rho \Delta_t I$);
- infectados: terá a perda dos indivíduos que entraram no estágio de latência ($\rho \Delta_t I$) e renovados ($\mu \Delta_t I$) e o acréscimo das novas infecções (infecção natural e exógena) ($\gamma \Delta_t E(k) + \delta \Delta_t E(k)I(k)/N$) e dos indivíduos que desenvolveram a tuberculose pulmonar ($(1-q)\beta \Delta_t I(k)S(k)/N$).

Matematicamente, pode-se então escrever a dinâmica associada ao comportamento médio do MBI como:

$$\begin{cases} S(k+1) = S(k) + \mu\Delta_t N - \mu\Delta_t S(k) - \frac{\beta\Delta_t I(k)S(k)}{N}, \\ E(k+1) = E(k) - \left(\mu + \gamma + \frac{\delta I(k)}{N}\right) E(k)\Delta_t + q \frac{\beta\Delta_t I(k)S(k)}{N}, \\ I(k+1) = I(k) - (\mu + \rho)I(k)\Delta_t + \left(\gamma + \frac{\delta I(k)}{N}\right) E(k)\Delta_t + (1-q) \frac{\beta\Delta_t I(k)S(k)}{N}. \end{cases} \quad (100)$$

Lembrando que k representa o instante de tempo $k\Delta_t$ e que

$$\frac{df}{dt} = \lim_{\Delta_t \rightarrow 0} \frac{f(t + \Delta_t) - f(t)}{\Delta_t} \quad (101)$$

nota-se que a Equação (100) coincide com a equação de diferenças da Equação (45) utilizada na obtenção de soluções do modelo SEI por integração numérica. Assim, o modelo MBI converge em média para o modelo SEI.